NASA Technical Paper 1788

Goddard Laboratory for Atmospheric Sciences Collected Reprints 1978-1979, Volume 2

MARCH 1981

(NASA-TP-1788-Vol-2) GODDARD LABORATORY FOR ATMOSPHERIC SCIENCES, COLLECTED REPRINTS 1978 - 1979, VOLUME 2 (NASA) 521 p
HC A22/MF A01 CSCL 04B

N81-21693
THRU
N81-21701
Unclas
H1/47 20684

NASA

NASA Technical Paper 1788

Goddard Laboratory for Atmospheric Sciences Collected Reprints 1978-1979 Volume 2

William C. Skillman, *Editor*
Earl R. Kreins, *Editor*
Goddard Space Flight Center
Greenbelt, Maryland

National Aeronautics and
Space Administration

Scientific and Technical
Information Branch

1981

The numbered papers in Volumes 1 and 2 of NASA Technical Paper 1788 have been accessioned as indicated below. Only the papers which have not been previously accessioned were processed as daughter items:

1. 79A48227
2. 79A32181
3. 79N20658
4. 80N16678
5. 80A39689
6. 79N20656
7. 79A35844
8. 79A33845
9. 79A35845
10. 79A29187
11. 79A29187
12. 79N20688
13.
14. 80N17638
15. 79A47689
16. 79N20633
17. 79N20660
18.
19. 80A10927
20. 79A44815
21. 78A47780
22. 79N10648
23. 78A26049
24. 79N20961
25.
26. 79A25356
27. 80A15279
28.
29. 80A26444
30. 80A26444
31. 80A10935
32. 79A50521
33. 78A23054
34. 79A23924
35. 79N13189
36.
37. 79A31871
38. 79A37698
39. 80A22480
40. 80A35540
41. 79N20645
42. 80A16184
43. 80A14265
44. 79N20654
45. 79N20650
46. 78A23607
47. 79N20644
48. 79N20574
49. 79N20639
50. 79A38006
51.
52. 79N20651
53. 79A38007
54. 78A36438
55. 81A13262
56. 79A45462
57. 79A34413
58. 80A21374
59. 79A31575
60. 79N20649
61. 80A15755
62. 78A47196
63. 79A22202
64. 78A35348
65. 78N26677
66. 78N21690
67. 79N13476
68. 79N16479
69. 79N30611
70. 79N19519
71. 80A34280
72. 78N20575
73. 79A48230
74. 80N12536
75. 79N16473
76. 78N33284
77. 80N19591
78. 80N12534
79. 78A35349
80. 79N10502
81. 79N23561
82.
83.
84. 78N20576
85. 79N28649
86. 80A20247
87. 80N19588
88.
89. 79N33725
90.
91. 79A22206
92. 80A15556
93. 79A53581
94.
95. 78N20577
96. 79N23478
97.
98. 78N16410
99.
100. 79A20134
101. 77N19564
102. 77N19563
103. 79A53565
104. 79N20672
105. 80N16403
106. 81A11971
107. 81A11970
108. 79N27635
109. 79N20689
110. 80A20194
111. 79A17829
112. 78A43280
113. 79A47678
114. 80N16402
115. 79N16494
116. 79N33531
117. 79N20674
118. 79N20676
119. 78A50493
120. 79A43934
121. 79A16717
122. 79A23668
123. 80A13586
124. 79A47192
125. 80A12813
126.

CONTENTS

VOLUME 1
I. SPECIAL PROJECTS OFFICE

II. MODELING AND SIMULATION FACILITY

CONTENTS (continued)

CONTENTS (continued)

III. UPPER ATMOSPHERE BRANCH

CONTENTS (continued)

IV. TROPOSPHERE BRANCH

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

V. HYDROSPHERIC SCIENCES BRANCH

Paper 62

(Reprinted from Nature, Vol. 274, No. 5671, pp. 539–543, August 10, 1978)

© Macmillan Journals Ltd., 1978

Ice sheet topography by satellite altimetry

R. L. Brooks
FGG Washington Analytical Services Center, Pocomoke City, Maryland 21851

H. Ray Stanley
NASA Wallops Flight Center, Wallops Island, Vancouver, Canada 23337

W. J. Campbell
U.S. Geological Survey, Tacoma, Washington 98416

H. Jay Zwally
NASA Goddard Space Flight Center, Greenbelt, Maryland 20771

R. O. Ramseier
Department of Environment, Ottawa, Canada K1A 0H3

The surface elevation of the southern Greenland ice sheet and surface features of the ice flow are obtained from the radar altimeter on the CEOS 3 satellite. The achieved accuracy in surface elevation is ~2 m. As changes in surface elevation are indicative of changes in ice volume, the mass balance of the present ice sheets could be determined by repetitive mapping of the surface elevation and the surface could be monitored to detect surging or significant changes in ice flow.

CHANGES in the volume of ice contained in the polar ice sheets on Greenland and Antarctica are intimately related to worldwide sea level and global climate. Because the sea level equivalent of the Greenland ice sheet is about 6 m and the Antarctic ice sheet about 60 m, ice volume changes of less than 1% are significant. The currently rising sea level (5 cm increase since 1940[1]) might be caused in part by changes in the volume of the polar ice sheets. Although major changes in polar ice volume have occurred during the past 10,000 years, it is not known whether the present ice sheets are growing or shrinking, nor if their ice flow is stable or subject to surges[2]. There is also speculation that man's impact on the climate through fossil fuel combustion[3] may induce ice-sheet melting.

Surveying the surface topography of the Antarctic ice sheet by satellite radio altimeter was proposed by Robin[4] before the advent of radar altimeters designed for accurate measurements of sea surface topography. A direct determination of the mass balance of the ice sheets and monitoring them for surges by repetitive mapping with satellite radar altimetry at intervals of 5–10 yr was suggested by Zwally[5]. A change of only 1 m in the 2,000 m average elevation of the ice sheets would indicate a change in ice volume that is equivalent to 3 or 4 cm of sea level. It is not feasible to determine the mass balance of the major parts of the ice sheets to such significant accuracy by conventional surveying techniques or direct measurement of mass input (snow accumulation) against mass output (melting and iceberg discharge). The increased ice output in an ice sheet surge would certainly cause detectable surface elevation changes in the surge basin. Measurements of the surface topography and its time dependence could be used, along with ice dynamics models, to assess the existing potential for ice sheet surging and to determine the nature and extent of an actual surge if one occurs.

Extensive mapping of both the surface and bedrock topography of Greenland and Antarctica is being conducted by aircraft radio echo sounding[6,7]. Recently, balloon-borne radar altimeters have also been used to map surface topography in Antarctica[8]. The accuracy achieved by these techniques, ± 30 m for aircraft and ± 60 m for balloons, is limited mainly by the uncertainty of the determination of the platform reference level using pressure altimetry. At least one order of magnitude increase in the accuracy is needed to determine significant ice volume changes and details of the surface topography.

The geodynamics experimental ocean satellite (GEOS 3), currently in orbit, is the first of a series of oceanographic satellites equipped with a radar altimeter. The results presented here from analysis of GEOS 3 data collected over southern Greenland show that the radar altimeter determines the elevation of the ice sheet surface to an accuracy of about ± 2 m and delineates surface features of the ice flow. Detailed knowledge of the surface topography is important not only for assessing volumetric changes of the ice sheets, but as a primary input for study of glacier dynamics. In addition to the improved vertical accuracy provided by satellite altimetry, it can cover the ice sheets in a much shorter time.

Sensing system

GEOS 3 was launched into an orbit with a mean altitude of 844.5 km and an inclination of 114° 52′ on 9 April 1975. This orbit covers all water, ice and terrain between 65° 08′ N and S latitudes.

The altimeter height (h) is determined from the measured time between the transmission and return of 13.9 GHz radar pulses. The height measured is to the mean surface within the altimeter 'footprint', or effective backscattering area, which is determined primarily by the altimeter characteristics. Because the footprint size and location on the observed surface can affect the height measurement and the interpretation of the results, the footprint definition and the method of determining height from the return pulse must be carefully reviewed. The footprint size is limited by the width (τ) of the transmitted pulse. As the antenna is designed to illuminate an area larger than the footprint, the footprint is pulse-limited and not beam-limited.

The ideal pulse-limited footprint is defined as the maximum circular area, on a reflecting plane normal to a radial line from the altimeter, from which backscattering can be simultaneously received. It is derived as follows. The leading spherical wave front of a pulse transmitted at time $t = 0$ first intersects the surface at time $t = h/c$, where c is the velocity of light. The circle of illumination then expands over the surface until $t = \tau$, at which time the trailing wave front intersects the surface as shown in Fig. 1a. After time $t = \tau$, point N at the centre is no longer illuminated. Backscattered signal at the altimeter is received from point N from time $2_1 = 2h/c$ until $2t = \tau$. The

552

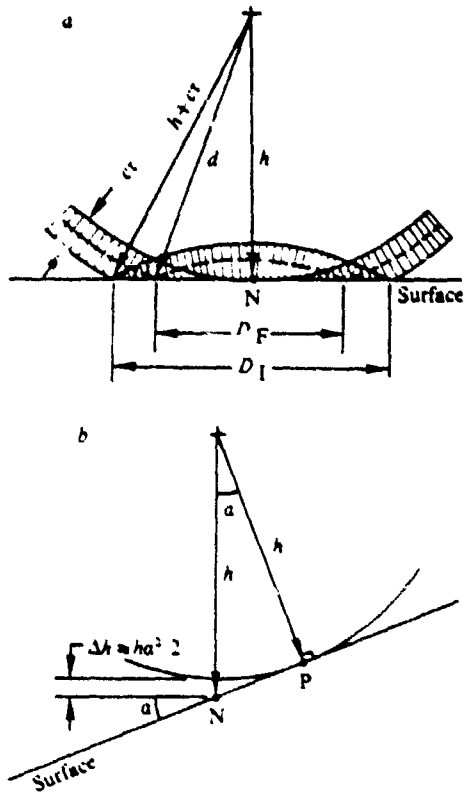


Fig. 1 a, Radar pulse wave fronts intersecting a reflecting surface. The diameter, D_F, of the footprint (effective backscattering area) and the diameter, D_I, of the illuminated area are illustrated *b*, Over a surface with slope α, the measured height, h, is to the nearest reflecting point (P). The desired height, h, above satellite nadir (N) is about $h\alpha^2/2$ greater than the measured h

distance d (see Fig. 1*a*), to the furthest point from which backscatter can also be received at time $2t + \tau$, is obtained from $2d = 2t + \tau$. Thus, $d = h + c\tau/2$ and from geometry the diameter of the ideal footprint is $D_F = 2(ch\tau)^{1/2}$, neglecting a second-order term. Effective pulse widths of 12.5 ns are achieved by pulse compression techniques in the short-pulse mode of GEOS 3 (ref. 9), and the corresponding footprint diameter is 3.6 km. The diameter D_I of the maximum circular area simultaneously illuminated, however, is a factor of $\sqrt{2}$ larger than the footprint.

The return signal increases from $2t$ to $2t + \tau$, and then remains nearly constant until the backscatter from the larger angles is attenuated by the beam limitations of the antenna, or is limited by the specular character of the surface. The altimeter split-gate tracker[10] is designed to measure the time (T) from transmission to the centroid of the return signal rise, so that $h = c/2(T - \tau/2)$. The centroid is chosen primarily to minimise the effect of surface waves on the determination of the mean surface height, because surface waves tend to distort the signal rise symmetrically about the centroid. In fact, the significant wave height of ocean waves is derived from the slope of the return signal rise.

The height determined by the altimeter is not necessarily to the satellite nadir point (N), but is to the nearest surface within the field-of-view that reflects sufficient signal to the antenna. The GEOS 3 altimeter antenna has a 3db field-of-view of 38 km (2.6° beam), which is about 10 times the footprint diameter. The antenna axis is also constrained to within 1° of nadir by the GEOS 3 stabilisation system. From a geometric standpoint, for a flat surface which has a mean slope (α) over the entire field-of-view, the nearest reflecting point (P) is upslope from nadir at a horizontal distance of $h \sin\alpha \cos\alpha \approx h\alpha$ and vertical distance of $h \sin^2\alpha \approx h\alpha^2$ (see Fig. 1*b*) (ref. 4). However, the geometrical error Δh in the measured height (h'), compared to the desired height (h) above nadir, is a factor of two smaller and is equal to $h(1 - \cos\alpha) \approx h\alpha^2/2$. Also, the measured height is the true height at a horizontal distance $h(1 - \cos\alpha)\tan\alpha \approx h\alpha^2/2$ from nadir. In principle, for uniformly sloping surfaces the indicated surface can be adjusted by this geometrical factor to obtain the true surface. For a surface slope of 1/300 (see Fig. 3, for example), the true surface is located 1.4 km horizontally inward or 4.7 m vertically below the indicated surface. Where the slope does not extend across the entire field-of-view, the geometrical displacement is $< \Delta h$.

Ice surface elevations above sea level are computed by subtracting the altimeter height measurements from the calculated satellite altitudes above a reference ellipsoid of the Earth and also subtracting the geoid height (local sea level height above the ellipsoid). Geoid heights from the Marsh–Vincent model are used.[11] The precise position of the GEOS 3 satellite (latitude, longitude, and altitude above the reference ellipsoid) is determined from orbital solutions incorporating measurements from ground-based tracking systems. The accuracy of the resultant satellite position in three dimensions depends on the type of tracking data available, and it varies from ±5 cm for laser tracking to a few metres for radar or doppler tracking.

Although an accuracy of a few metres is satisfactory for the horizontal position of the satellite, the altitude over the ice sheet is further refined by reference to the adjacent sea surface. For all the data shown here the elevation is adjusted to zero over the deep water east of Greenland where tidal and dynamic variations do not exceed 0.5 m. The data may be corrected later to compensate for the small tidal effects.

Height data points can be obtained from GEOS 3 at the rate of 100 s^{-1} or 10 s^{-1}, depending on whether the delays of individual pulses or the average of 10 pulses are analysed. Averaged data (10 s^{-1}) are used here so that each elevation data point represents an average along 0.67 km of the satellite ground track.

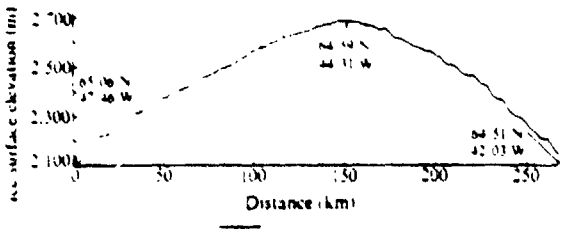


Fig. 2 Greenland ice sheet surface profile measured by the radar altimeter along orbit 0640 of GEOS 3. Vertical exaggeration is 134:1.

Analysis and discussion

The surface elevation segment in Fig. 2 from radar altimeter measurements taken on orbit 0640 of GEOS 3 shows a typical ice sheet profile. The surface along this segment is nearly symmetrical about the crest with a mean slope of 1:270 on each side. As the ice flow toward the edges of the ice sheet can be approximately described as plastic deformation, there is a well-known relationship between surface slope and ice thickness. In plastic flow, the shear stress and consequently the ice flow are proportioned to the surface slope and the ice thickness. Therefore, small surface slopes are generally indicative of thick ice. Near the 210 km point, the slope increases to 1:150 indicating thinner ice toward the mountainous east coast.

ORIGINAL PAGE IS
OF POOR QUALITY

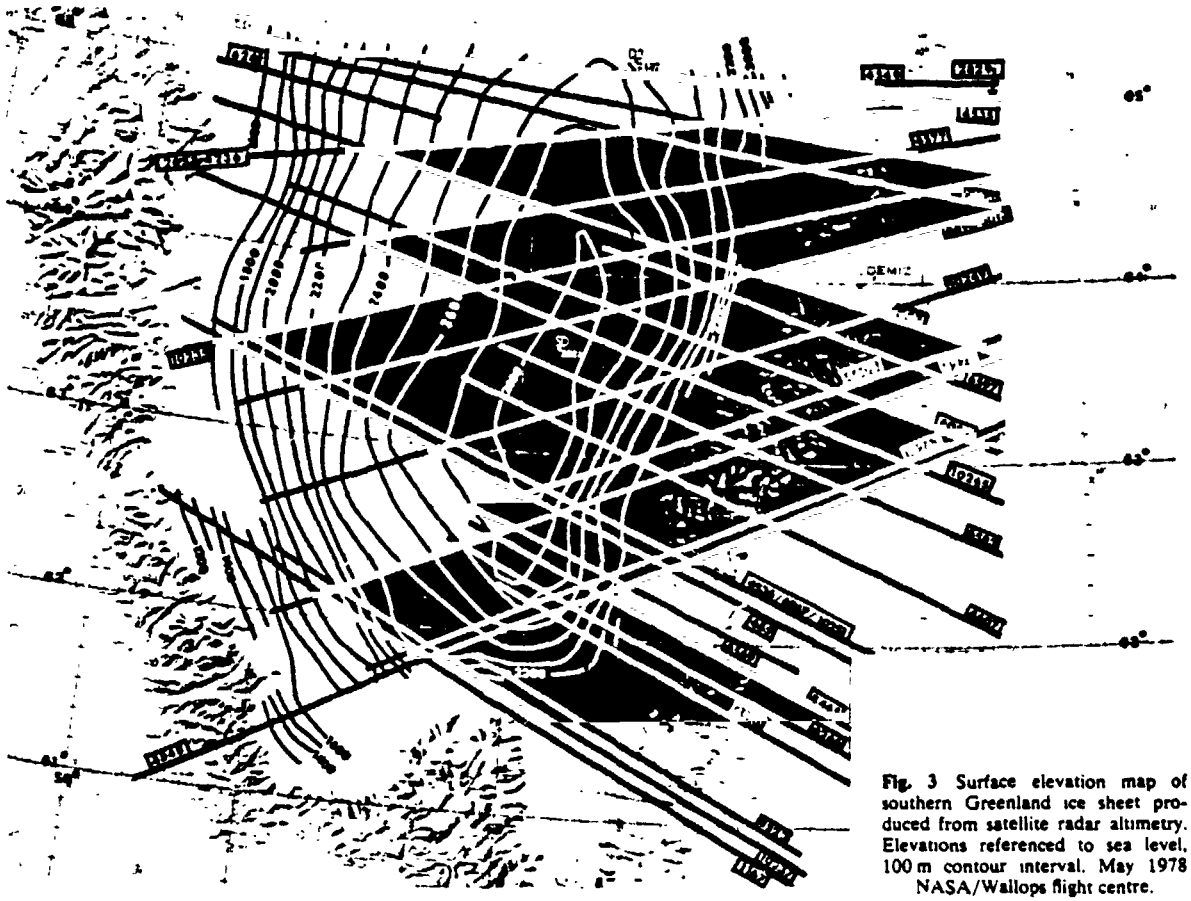


Fig. 3 Surface elevation map of southern Greenland ice sheet produced from satellite radar altimetry. Elevations referenced to sea level, 100 m contour interval. May 1978 NASA/Wallops flight centre.

A topographic map prepared from altimeter data from 34 GEOS 3 orbits between May 1975 and April 1977 is shown in Fig. 3. As the best previous topographic map of this area was made from rather sparse surface altimeter measurements having in most cases only ±30 m accuracy[12], there is little information of sufficient accuracy for comparison with the radar altimetry. However, at two locations (DYE-3 station, and South Dome) the surface elevation was established to an accuracy of several metres by Mock[13] using navigational-satellite positioning. Mock's elevations at these locations, labelled D3 and SD respectively in Fig. 3, have been referenced to sea level. Orbit 10265 data show a maximum elevation of 2,833 m at 15 km north-east of South Dome. This altimeter measurement is 9 m higher than Mock's elevation of South Dome and consistent with his observation that the surface slopes upwards to the north-east from the South Dome site. Similar agreement exists between the DYE-3 elevation and the elevations from the orbit 6 km south of the station. In addition, Mock (personal communication) has provided elevations on the eastern slope of orbit 2658 south-east of DYE-3 which agree with the satellite altimeter within 5 m. The western part of orbit 2658, however, shows elevations 100–150 m higher than previously inferred values, a difference which is equal to about 5% of the ice thickness. These large differences occur in an area where the previous control points were sparse and less accurate.

The internal consistency or precision of the radar altimeter has been examined by analysing intersecting nadir tracks along the surface. Elevation differences at intersections presented in Table 1 range from 0.07 to 3.10 m. When the time interval between intersecting orbits is long, close agreement of surface elevations is not necessarily expected because accumulation, ablation, and ice flow can cause detectable variations. For the two intersections having time intervals of only a few days, the indicated differences are <0.30 m. Because other intersects had intervals as long as 96 weeks, real elevation changes of the order of 1–2 m might account for part of those indicated differences. The mean absolute difference is 1.4 m. Although additional analysis is required, these data are consistent with a precision of ~ 2 m or better.

A typical accumulation rate over southern Greenland is 0.55 m yr^{-1} ice-equivalent[14] or about 1.5 m yr^{-1} of near-surface snow equivalent. If the ice flow is in equilibrium with the surface accumulation, then the vertical surface velocity is equal to the accumulation rate and the surface elevation will not change. Nevertheless, even if the ice flow is in equilibrium with the multi-year average accumulation rate, localised and inter-annual anomalies in accumulation rate and surface elevation will occur. For example, the effect of a localised doubling of a 0.55 m yr^{-1} accumulation rate for one year would be a temporary surface rise of 1.5 m.

Figure 4 shows an example of nearly overlapping tracks that are parallel and separated by 0.8 km across track. The mean surface elevations over distances of 10–20 km along the two tracks agree within 1 m, but local differences are as much as 5 m and are probably due to across-track surface undulations. The local across-track slopes are about the same as along-track slopes on the scale of several kilometres, indicating undulations in both directions.

The delineation of undulations or waves is an intriguing aspect of the altimeter-derived elevation profiles. Similar waves have been observed by surface levelling, for example, during the European glaciological expedition across Greenland

Table 1 Observed elevations at intersections over ice sheet

Orbit no.	Date (yr month date)	Interval (d)	Position	Surface slope	Elevation above sea level (m)	Elevation difference (m)
640	750525		64.885°N,		2,310 22	
4236	760203	254	42.688°W	1/280	2.310.36	+0.14
683	750528		64.565°N,		2,283.75	
2658	751014	139	47.335°W	1/125	2,286.71	+2.96
683	750528		64.149°N,		2,636.45	
10235	770402	675	43.367°W	1/115	2,633.35	−3.10
2673	751015		64.814°N,		2,623.89	
4236	760203	111	43.898°W	1/370	2,621.97	−1.92
4548	760225		62.105°N,		2,559.95	
4593	760228	3	44.540°W	1/960	2,560.23	+0.28
10235	770402		63.468°N,		1,916.40	
10251	770403	1	48.557°W	1/110	1,916.33	−0.07

at about 70°N (ref. 15). The generation of surface undulations by ice flow over irregularities in the bedrock topography has been discussed by several authors including Budd[16]. Budd's theoretical analysis, and analyses of measured surface undulations with respect to bedrock undulations by Beitzel[17] and Budd and Carter[18], have shown that wavelengths about 2–6 times the ice thickness, that is wavelengths of the order of 5–10 km, dominate the surface profiles. Such waves generated by flow over the bedrock are stationary, whereas kinematic surface waves[19] would move downslope and wind-drift accumulation waves[20] would move upslope. Therefore, bedrock-generated waves should be distinguishable from other surface waves by their stationary character

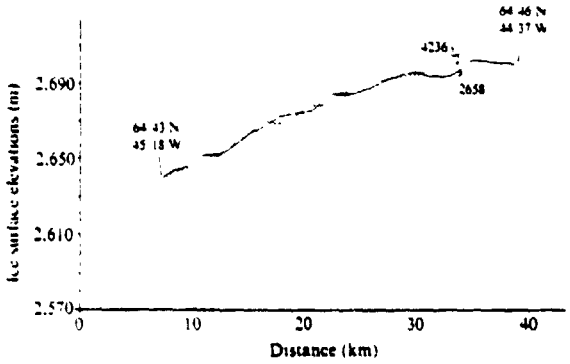


Fig. 4 Ice surface elevations on the eastern slope along parallel lines separated by 0.8 km. GEOS 3 orbits, 2658 14 October 1975; 4236 3 February 1976.

Segments of four elevation profiles on the western slope are shown in Fig. 5. Waves in the western slope profiles shown here have individual amplitudes as large as 25 m. However, the eastern slope of track 640 in Fig. 2 shows more regular wave features. A series of about eight waves having an average wavelength of 10 km and an average amplitude of 15 m is apparent. The geometrical factors that limit the amplitude of the wave features in a radar profile have been described by Robin[4]. For example, for a wavelength of 10 km the curvature of the radar front is such that it will meet the crests of the surface waves at the same time that it meets the bottom of a valley 15 m deep at the mid-point. Thus, waves with amplitudes greater than 15 m having wavelengths of about 10 km apparently have 15 m amplitude. Therefore, the amplitude of the observed waves could be larger than indicated. On another scale, the average amplitude of wave features such as sastrugi, having small spatial extents compared to the altimeter footprint, should be derivable from the return pulse-shape as ocean significant wave heights are derived.

The detailed surface topography that will result from the analysis of the complete set of altimeter profiles will determine the three-dimensional character of the observed waves, and sequential mapping will determine their time-dependent nature. Because the wave character is related to the bedrock topography, ice thickness, and ice flow parameters, much can also be learned about ice sheet dynamics from the observation of surface waves.

Conclusion

Although the GEOS 3 radar altimeter was specifically designed for ocean measurements and the operating parameters are far from optimum for ice, it can be used for mapping the surface of

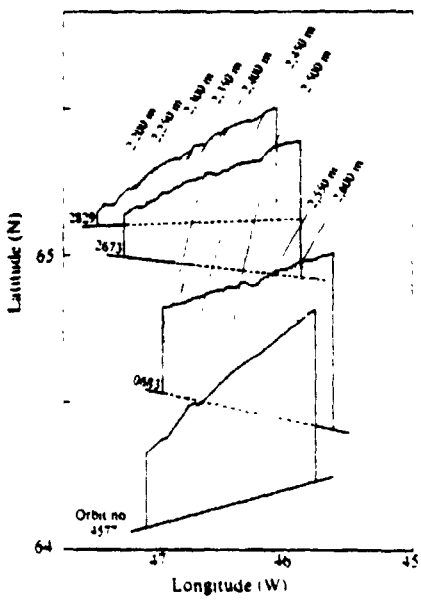


Fig. 5 Three dimensional view of ice surface on western slope of the ice sheet.

the ice sheets. The accuracy of ~2 m obtained with GEOS 3 is sufficient to establish a baseline elevation map of the southern Greenland ice sheet for determination of its mass balance or state equilibrium between mass input and ice flow. This mapping will be repeated next by the radar altimeter on Seasat A launched in June 1978. The improved precision expected with Seasat A may reduce the time interval that is needed between mappings to detect significant elevation changes. Unfortunately, the Seasat-A orbit inclination of 108° only extends the coverage to ±72° latitude which is sufficient to cover ~50% of Greenland but only the coastal margin of the East Antarctic ice sheet. For example, it does not cover the very interesting West Antarctic ice sheet[2].

By using the data from closely spaced and overlapping orbits to remove the geometrical errors on sloping surfaces, it will be possible to establish more accurately the actual surface of the sloping portions of the ice sheet. Any residual errors in the absolute elevation due to slopes will have a smaller influence on the detection of elevation changes between repetitive mappings using the same technique.

The observation of numerous surface waves of the order of 10–20 m, amplitude and 10 km, wavelength corroborates surface measurements and will provide additional insights to the dynamics of large ice masses. This also points out the necessity for accurate mapping of the entire ice sheet, because extrapolation of localised surface measurements can be strongly influenced by localised or temporary features of the surface dynamics. For example, the determination of the locations of ice domes or ice divides requires the smaller scale undulations to be averaged to determine the mean surface.

The results from the analysis of intersecting tracks and the agreement with the surface control points implies that Fig. 3 is the most accurate topographic map of southern Greenland. One notable difference from the previous maps is the greater elongation of the southern dome as shown by the extension of the 2,600 m contour near to DYE-3 and the extension of the 2,800 m contour south to 62.7°N.

An altimeter similar to the present one but specifically designed for ice sheet mapping is required for future studies. The GEOS 3 studies are providing information on altimeter operation over ice and sloping surfaces that can be used to optimise antenna beam, pulse, sensitivity, tracking circuit, and other parameters and also to evaluate designs such as a multiple-beam altimeter. An appropriate ice altimeter placed in polar orbit at intervals of 2–5 yr would provide the data required to answer important geophysical questions about the growth or shrinkage and stability of the present day ice sheets.

Received 2 February; accepted 31 May 1978.

1. Hicks, S. D. *J. geophys. Res.* **83**, 1377 (1978).
2. Weertman, J. *Nature* **260**, 284 (1976).
3. Brooker, W. S. *IEEE* **65**, 989 (1977).
4. Robin G. de Q. *Can. J. Earth Sci.* **3**, 893 (1966).
5. Zwally, H. J. *J. Glaciology* **15**, 73, 444 (1975).
6. Oswald, G. K. A. & Robin, G. de Q. *Nature* **245**, 251 (1973).
7. Gudmandsen, P. *J. Glaciol.* **15**, 73, 93 (1975).
8. Lovenso, N., Julian P. & Seemi, V. *Nature* **268**, 314 (1977).
9. Lester, C. D., Purdy, C. L. & Brooks, R. L. *NASA Tech. Mem.* X-69357 (1975).
10. Skolnik, M. I. *Introduction to Radar Systems* (McGraw-Hill, New York 1962).
11. Vincent, S. & Marsh, J. G. *NASA Tech. Mem.* X-70492 (1973).
12. Mock, S. J. & Weeks, W. F. *Cold Regions Research and Engineering Laboratory*, Res. Rep. 170 (1965).
13. Mock, S. J. *Cold Regions Research and Engineering Laboratory*, Res. Rep. 76-41 (1976).
14. Mock, S. J. *Cold Regions Research and Engineering Laboratory*, Res. Rep. 233 (1967).
15. Militzer, H. *Meddy Grønland* **173**, 7 (1964).
16. Budd, W. F. *J. Glaciol.* **10**, 59, 177 (1971).
17. Bentley, J. E. *Int. Symp. Antarctic Glaciological Exploration*, 191 (Hefter, Cambridge, 1970).
18. Budd, W. F. & Carter, D. B. *J. Glaciol.* **10**, 59, 198 (1971).
19. Paterson, W. S. B. *The Physics of Glaciers* (Pergamon, Oxford 1969).
20. Bleak, H. P. & Budd, W. *J. Glaciol.* **5**, 37, 3 (1964).

Paper 63

VOL. 84, NO. C1 JOURNAL OF GEOPHYSICAL RESEARCH JANUARY 20, 1979

Comparison of 2.8- and 21-cm Microwave Radiometer Observations Over Soils With Emission Model Calculations

W. J. BURKE

Regis College Research Center, Weston, Massachusetts 02193

T. SCHMUGGE

NASA Goddard Space Flight Center, Greenbelt, Maryland 20771

J. F. PARIS

University of Houston at Clear Lake City, Houston, Texas 77058

An airborne experiment was conducted under NASA auspices to test the feasibility of detecting soil moisture by microwave remote sensing techniques over agricultural fields near Phoenix, Arizona at midday of April 5, 1974 and at dawn of the following day. Extensive ground data were obtained from 96 bare, sixteen hectare fields. Observations made using a scanning (2.8 cm) and a nonscanning (21 cm) radiometer were compared with the predictions of a radiative transfer emission model. It is shown that (1) the emitted intensity at both wavelengths correlates best with the near surface moisture, (2) surface roughness is found to more strongly affect the degree of polarization than the emitted intensity, (3) the slope of the intensity-moisture curves decreases in going from day to dawn, and (4) increased near surface moisture at dawn is characterized by increased polarization of emissions. The results of the experiment indicate that microwave techniques can be used to observe the history of the near surface moisture. The subsurface history must be inferred from soil physics models which use microwave results as boundary conditions.

INTRODUCTION

Recent attempts to monitor the world's food supply from space [*Hammond*, 1975] would be greatly assisted by timely information concerning the soil moisture available to growing crops. *Poe et al.* [1971] and *Schmugge et al.* [1974] have reported observing correlations between the natural microwave emissions of soils and their moisture content. To test the remote sensing capabilities of microwave technology for measuring soil moisture, the National Aeronautics and Space Administration (NASA) has initiated a set of experiments. We report here on the initial phase of the experiments, a set of flights over agricultural fields near Phoenix, Arizona, concentrating on microwave and infrared radiometric data.

Experimental and theoretical considerations suggest that the following parameters most strongly influence the radiative signatures of natural terrain: (1) the frequency and polarization of the radiation, (2) the look angle of the detectors, (3) the texture of the soil, (4) the moisture and temperature profiles of the soil, (5) the roughness characteristics of the surface, and (6) the vegetative cover. The first two parameters are under experimental control but the remainder are given by nature. In the present experiment the textures, moistures, temperatures and surface roughnesses of the soils were directly measured for comparison with microwave data. The effects of vegetative covers are disregarded for the time being, by limiting ourselves to bare field observations.

Theoretical emission models are useful for interpreting the physical information contained in the microwave observations. The principle of detailed balancing has been applied by *Peake* [1959] to calculate the emissivity of soils in thermodynamic equilibrium with their environments. *Stogryn* [1970] has pointed out that emissivity is a meaningful concept only where no thermal gradients exist in soils. He has employed a Green's function method to derive a general set of equations for microwave emissions at either polarization for any dielectric and temperature profile. In this paper we present an emission model based on the radiative transfer equation. Moisture and temperature gradients are represented by step functions across soil layers of finite thickness. The model appears (1) to provide physical insight into the significance of the data, (2) to be in reasonable agreement with radiometric observations, and (3) to require relatively little computer time.

In the following sections of this paper we first summarize the Phoenix experiment and its results. The radiative transfer model along with its predictions for 2.8 cm and 21 cm emissions are then presented. In comparing radiometric observations with the model's predictions it is shown that the emitted intensity and polarization are physically significant parameters. The inversion capabilities of the model are illustrated in terms of these quantities. Finally, we discuss the experimental results as they pertain to future uses of microwave remote sensing.

THE EXPERIMENT

The first phase of the NASA sponsored feasibility test for detecting soil moisture by microwave remote sensing techniques was conducted near Phoenix, Arizona on April 5 and 6, 1974. Contributing investigators came from the Johnson Space Center, Goddard Space Flight Center, the Agricultural Research Service of the U.S. Department of Agriculture, the Environmental Research Institute of Michigan, the University of Arkansas, the University of Kansas and Texas A & M University. *X* and *L* band (2.8 cm and 21 cm) radiometers and imaging radars were flown over four flight lines, two aligned north-south and two aligned east-west. Three passes were made over the flight lines during the early afternoon of April 5, and a single pass at dawn of April 6. At the time of the

Copyright © 1979 by the American Geophysical Union.

afternoon passes soil moisture and temperature ground truth samples were taken from 96 bare sixteen hectare (40 acre) fields. Time constraints allowed for moisture and temperature measurements from only six pairs of fields during the April 6 pass.

The ground truth effort, a detailed map of the flight lines and tabular listings of the soil moisture, temperature, texture and average brightness temperatures of each field are given by *Schmugge et al.* [1976]. Active microwave results were reported by *Cihlar et al.* [1975]. Moisture samples were taken from four widely spaced points in each field from the tops and bottoms of furrows. The sample layers were 0–1, 1–2, 2–5, 5–9 and 9–15 cm. Temperature samples were taken at one point per field at the tops and bottoms of furrows. Probes with thermistors embedded at the midpoints of the above given layers were used. The soil samples were subjected to hygroscopic and textural analysis [*Schmugge et al.*, 1974]. The dielectric coefficients of the various soil types found in the Phoenix area were measured as a function of moisture content in waveguides [*Babai*, 1974; *Newton and McClellan*, 1975]. As a measure of surface roughness a set of three photographs was taken in each field against the background of a gridded panel. Due to the relatively uniform cultivation practices of Pheonix area farmers only three kinds of surfaces were found: (1) listed fields with mean furrow heights of 27 cm and peak to peak separations of 1 m, (2) planted fields with mean furrow heights of 5 cm separated by 30 cm, and (3) one flat field. No attempt has been made to estimate the spectral characteristics of the small scale roughness.

The microwave radiometers carried aboard the NASA P-3A aircraft were a nonscanning 21 cm (L-Band) instrument and the 2.8 cm (X-Band) Passive Microwave Imaging System (PMIS). The PMIS is an imaging radiometer that scans in 44 steps along an azimuthal arc of ±35° at a constant look angle of 49.5°. The antenna is a dual polarized cross-slot phased array, that is electronically stepped for scanning. The half power beam width is 2.6° × 1.6°. The L-band antenna is a flat plate dipole array whose look angle and polarization must be set manually. During the first afternoon and the dawn flights the antenna was set at a nadir look angle. The look angle was 40° for the second and third afternoon flights, measuring vertical and horizontal polarizations respectively. The half power beam width is ~15°. Since the aircraft flew at an altitude of ~800 m at a speed of 82 m/s we were able to retrieve only 1 or 2 s of uncontaminated L band data per field.

Also aboard the aircraft were two infrared radiometers in the 10–12 μm range (one scanning, the other nonscanning). These radiometers were used to determine the surface temperature of the soil. From a remote sensing point of view, it is necessary to have some way to estimate the soil temperature. By comparing surface temperatures with temperatures of the first layer of soil, observed within 20 min of an afternoon overpass, we calculated a relationship

$$T_{1R} = 1.34T_1 - 3.6 \text{ (in °C)}$$

with a correlation coefficient of 0.96. T_{1R} is the surface temperature observed with the infrared radiometer and T_1 is the measured temperature of the 0–1 cm layer.

A Simple Model for Microwave Emissions From Stratified Soils

In this section we develop a rather simple radiative transfer model for the emission of microwave radiation from bare soils. The predictions of the model are checked against PMIS observations, and an inversion technique is suggested. By comparing the predictions of the model for simulated data we are able to estimate the relative contributions from various depths of soil for X and L band emissions. Surprisingly, the model predicts a high correlation between X and L band radiative emissions.

Our radiative transfer model makes the following simplifying assumptions:

1. The radiation is incoherent.
2. There is neither attenuation nor emission between the surface and the antenna.
3. The sky brightness is isotropic and has a value of 5°K at X and L bands.
4. Moisture and temperature are functions of depth only.
5. Dielectric and thermal properties of the soil are constant across layers of finite thickness.
6. The surface of the soil is smooth.

A cross section of a stratified soil is shown on Figure 1. Layers have thicknesses ΔZ_j, which are not necessarily the same for all layers. The jth layer is bounded or the top by the jth surface and by the jth + 1 surface on the bottom. Within this layer the dispersion relation for electromagnetic wave propagation is $k_j^2 = (\omega/c)^2 \mu_j \xi_j$. The frequency is ω in radians per second, c is the velocity of light, μ is the magnetic permeability (assumed equal to one) and $\xi_j = \epsilon_{Rj} + i\epsilon_{Ij}$ is the complex dielectric coefficient. If we write $k_j = (\omega/c)(\beta_j + i\alpha_j)$, then the dispersion relation gives

$$\begin{aligned} (\beta_{xj}^2 + \beta_{yj}^2 + \beta_{zj}^2) - (\alpha_{xj}^2 + \alpha_{yj}^2 + \alpha_{zj}^2) &= \epsilon_{Rj} \\ 2\boldsymbol{\beta}_j \cdot \boldsymbol{\alpha}_j &= \epsilon_{Ij} \end{aligned} \tag{1}$$

Snell's law shows that the component of the wave vector parallel to the surface is a conserved quantity. Thus

$$\beta_{xj}^2 + \beta_{yj}^2 = \sin^2 \theta_0$$

and

$$\alpha_{xj} = \alpha_{yj} = 0$$

θ_0 is the angle of the ray emerging from the soil with respect to the surface normal. It ranges in value between 0° and 90°. Equation (1) can be solved to give

$$\beta_{zj} = \left\{ \frac{1}{2} (\epsilon_{Rj} - \sin^2 \theta_0) \left[1 + \left(1 + \frac{\epsilon_{Ij}^2}{(\epsilon_{Rj} - \sin^2 \theta_0)^2} \right)^{1/2} \right] \right\}^{1/2} \tag{2}$$

$$\alpha_{zj} = \epsilon_{Ij} / 2\beta_{zj}$$

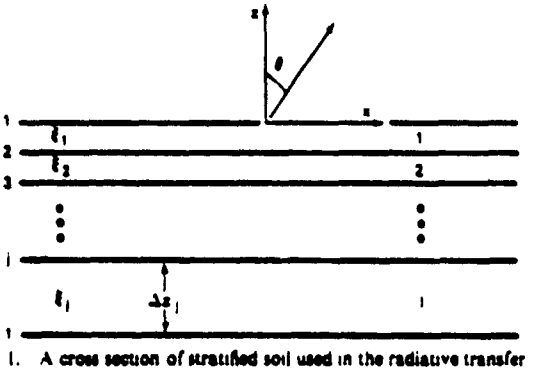


Fig. 1. A cross section of stratified soil used in the radiative transfer model.

At the boundary between the j and $j-1$ layers radiation is partially reflected and transmitted. The fractions of the incident electric field with horizontal and vertical polarizations reflected back into the jth layer are given by the Fresnel coefficients.

$$\rho_H = \frac{k_{zj} - k_{zj-1}}{k_{zj} + k_{zj-1}}$$
$$\rho_V = \frac{\xi_{j-1}k_{zj} - \xi_j k_{zj-1}}{\xi_{j-1}k_{zj} + \xi_j k_{zj-1}} \tag{3}$$

where $k_{zj} = \beta_{zj} + i\alpha_{zj}$. Thus β_z, α_z, ρ_H and ρ_V all depend on the complex dielectric coefficient and on the angle θ_0 that the ray emerges from the soil.

An attempt to construct a radiative transfer equation that describes radiation emitted from a stratified soil is now made. Within the first layer the radiative transfer equation is

$$\frac{dI_\omega}{dz} = -\gamma_1 I_\omega + \gamma_1 J_\omega \tag{4}$$

γ is usually written as a product of the density and monochromatic mass absorption coefficient. By writing the Poynting theorem in an appropriate form it can be shown that $\gamma_1 = 2\omega\alpha_{z1}(\theta_0)/c$.

I_ω is the intensity of radiation at frequency ω. J_ω is the Planck emission function. In the microwave frequency range Planck's emission law reduces to the Rayleigh-Jeans equation, where J_ω is proportional to the temperature of the medium T. Adopting a similar scaling rule for I_ω and effective temperature T_p can be defined which is directly proportional to I_ω. The subscript ω is suppressed and T_p refers to the intensity in a narrow range near ω in the p polarization state. Since J_ω is isotropic and independent of polarization no designation is necessary. The radiative transfer equation in the first layer may be written

$$\frac{dT_p}{d(\gamma_1 z)} = -T_p + T_1 \tag{5}$$

This equation can be integrated from a point just below the surface to a point just above the interface between the first and second layers. Because the dielectric properties are assumed to be constant across the layer

$$T_p(1^-) = T_1(1 - e^{-\gamma_1\Delta z_1}) + T_p(2^-)e^{-\gamma_1\Delta z_1} \tag{6}$$

The argument $N^\pm$ implies that the measurement is made above (plus) or below (minus) the Nth interface. The first term on the right hand side of (6) accounts for radiation emitted within the first layer and comes directly to the surface. The second term describes upwelling radiation at the bottom of the first layer. This in turn has two components: first, radiation emitted in the first layer and reflected at the interface between the first and second layers; and second, radiation transmitted from lower layers.

$$T_p(2^-) = R_{p2}T_1(1 - e^{-\gamma_1\Delta z_1}) + T_p(2^-)(1 - R_{p2}) \tag{7}$$

R_{p2} is the absolute value squared of the Fresnel coefficient for the p polarization (equation (3)). The radiation field just above the surface is the value just below multiplied by the transmittance $(1 - R_{p1})$

$$\begin{aligned} T_p(1^+) &= (1 - R_{p1})T_p(1^-) \\ &= (1 - R_{p1})[T_1(1 - e^{-\gamma_1\Delta z_1})(1 + R_{p2}e^{-\gamma_1\Delta z_1}) \\ &\quad + (1 - R_{p2})T_p(2^-)e^{-\gamma_1\Delta z_1}] \end{aligned} \tag{8}$$

The radiative transfer equation can be integrated again to calculate $T_p(2^-)$. Repeating the procedure for N layers gives

$$T_p(1^+, \theta_0) = \sum_{j=1}^{N} T_j(1 - e^{-\gamma_j\theta_0\Delta z_j})(1 + R_{p,j+1}(\theta_0)e^{-\gamma_j(\theta_0)\Delta z_j}) \cdot \prod_{j=1}^{i} [1 - R_{pj}(\theta_0)] \exp - \left(\sum_{j=1}^{i} \gamma_{j-1}(\theta_0)\Delta z_{j-1}\right) \tag{9}$$

The brightness temperature measured by the antenna is the sum of the reflected sky brightness and radiation emitted from the soil

$$T_{B,p}(\theta_0) = T_{\omega s}R_{p,s}(\theta_0) + T_p(1^+, \theta_0) \tag{10}$$

A computer program was written that used as its inputs the moisture and temperature profiles measured in the Phoenix fields. Dielectric coefficients were calculated using the soil type characteristic of each field. With this information (10) could be calculated and results of the model compared with PMIS and MFMR observations.

To illustrate the capabilities of the model consider the ground observations (Table 1a) taken simultaneously with the aircraft overflight of field 260A. The field had listed furrows aligned perpendicular to the flight line. Irrigation by flooding the bottom of furrows occurred one week before the flight. Predicted brightness temperatures at vertical and horizontal polarizations for look angles between 0° and 90° are listed in Table 1*b*. The average value of PMIS observations for this field, $T_v = 281.1°$ and $T_H = 265.9°$, are near the emission predictions from the tops of furrows with a look angle of 30°. Evidently, emissions from the bottoms of furrows were shadowed. With the exception of the one flat field the observed brightness temperatures of all other fields were found to coincide with emission values predicted for look angles between 35° and 25°. This shift in the effective look angle is due to surface roughness [*Ulaby et al.*, 1970].

For a model to be useful in remote sensing it must be capable of inversion. To effect such an inversion the brightness temperatures of a set of pseudo-fields have been calculated. The temperature profile $T_1 = 303°$, $T_2 = 301°$, $T_3 = 299°$, $T_4 = 297°$ and $T_5 = 295°$ was similar to that observed in Phoenix fields on April 5. The moisture profile was PM(1) = M, PM(2) = M + 2, PM(3) = M + 7, PM(4) = M + 14 and PM(5) = M + 16. M was given values between 1 and 25%. However, the moisture content of any given layer was not allowed to exceed 30%. We refer to this as the 'normal' profile. Other profiles were studied for comparisons.

Calculated brightness temperatures are plotted in Figure 2

TABLE 1*a*. Average Moisture and Temperature Profiles Observed at Tops and Bottoms of Furrows, Field 260A

Layer	Moisture	Temperature, °K
	Top	
1	4.4	305.5
2	10.8	303.7
3	19.7	301.4
4	21.6	297.3
5	23.7	290.3
	Bottom	
1	12.8	304.4
2	17.5	302.0
3	21.1	299.2
4	23.6	295.4
5	23.0	289.3

TABLE 1b. Brightness Temperatures Predicted by the Model for Tops and Bottoms of Furrows, Field 260A

	Top		Bottom	
θ, deg	T_v	T_h	T_v	T_h
0	271.9	271.9	225.2	225.2
10	272.4	271.2	225.7	224.9
20	275.4	268.2	230.2	220.5
30	280.4	262.6	238.0	212.6
40	287.3	253.2	249.5	200.5
50	295.6	238.3	265.0	181.2
60	302.6	214.7	283.9	159.3
70	299.3	177.5	300.8	127.1
80	253.5	119.3	285.3	84.7
90	30.0	30.0	30.0	30.0

as functions of the moisture content of the top centimeter PM(1) for look angles of 0°, 30° and 50° at the X band frequency. For PM(1) > 10%, at nadir look, the average slope is −3.5°K/percent moisture. This is consistent with a slope of −3.3°K/percent moisture reported by *Schmugge et al.* [1974]. Figure 2 can also be used to attempt an inversion. Unless the presence of a flat field was indicated by a large difference between T_v and T_H, it was assumed that due to surface roughness effects PMIS observations came from an effective look angle of 30°. Estimates of PM(1) from the top of furrows gained by this method diverged by no more than 1% from ground observations.

Other predictive features of the model can be seen by considering the first two Stokes parameters; $P = 1/2(T_v + T_H)$ (intensity) and $Q = T_v - T_H$ (polarization). Figures 3*a* and 3*b* are plots of predicted intensity and polarization, respectively, as functions of the moisture content of the surface layer of soil at specified look angles. The 0° intensity curve is indistinguishable from that of 30°. For look angles ≤50° the intensity depends only weakly on look angle. The polarization depends on both look angle and PM(1). The dependence of P and Q on soil temperature was tested by using temperature profiles similar to those found at dawn. We found that a 20° shift in soil temperature changed the calculated intensity by 15° and the calculated polarization by only 2°.

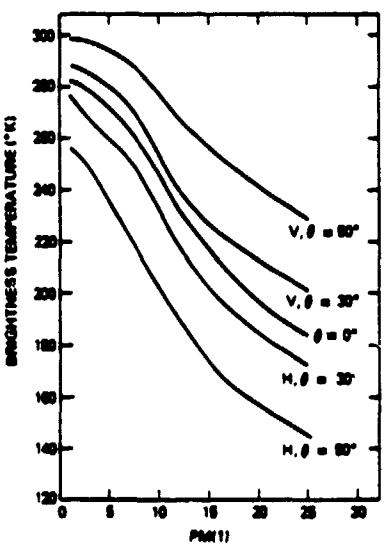


Fig. 2. Model calculations for vertical and horizontal polarization brightness temperatures at X band frequency and look angles of 0°, 30° and 50° as a function of the percent moisture in the top layer of soil.

The brightness temperature for L-band (21) cm radiation (T_L) at the nadir look angle as a function of PM(1) is given in Figure 4. In the case of the 'normal profile' T_L decreased in the moisture ranges 0 < PM(1) < 5 and 12 < PM(1) < 30% but remained nearly constant in the 5 < PM(1) < 12% range. Due to the paucity of fields with surface layer moisture between 5 and 12% at the time of the 1974 experiment, we could not determine whether the plateau in T_L is a natural feature or an artifact of the model. The plateau is found in the calculations independent of the size of the layers. If the plateau were an artifact it would result from a very poor representation of the complex dielectric coefficient of moist soil. This seems to be an unlikely source of serious error since dielectric coefficients were measured in waveguides using [illegible] soils [*Newton and McClellan*, 1975].

Sensitivity to moisture gradients was tested by comparing predictions for profiles with very steep gradients with those of the 'normal' profile. In Figure 4 we have also plotted results with PM(2) = 20% and more moisture at greater depths. We find that the case PM(1) = 1% and PM(2) = 20%, a moist soil with a very dry crust, was indistinguishable from a relatively dry soil with PM(1) = 4% and PM(2) = 6%. Because of results of this type it may be impossible to make an unambiguous determination of soil moisture with one brightness temperature observation.

In Figures 5*a* and 5*b* we have plotted the percent contribution of the various layers to the total brightness temperature under 'normal profile' conditions. Most of the escaping radiation at X-band is generated in the top 2 cm of soil. The contributions to the emitted radiation at 21 cm come from much deeper in the soil. Changing the gradients of the moisture changes the relative contributions of the different layers, but the sum of the terms is about constant for a given PM(1). *Cihlar and Ulaby* [1975] have suggested that the predicted correlations of T_v and T_L with PM(1) results from the fact that the change in dielectric coefficient at the air-soil interface dominates over subsurface gradients in determining what radiation can get through the surface.

To summarize, we have developed a simple radiative transfer model for microwave emissions from soils. By considering the Stokes parameters $P = 1/2(T_v + T_H)$ and $Q = T_v - T_H$ we have found that the model predicts the following:

1. Both P and Q are sensitive to the moisture content of soil. Q is more sensitive in the range 0–15% and P is more sensitive for moistures >15%.
2. Q is sensitive to the look angle of the antenna but P is not. If the major effect of roughness is to change the effective look angle of the detector (which is true in the geometric optics limit) then P should be independent of roughness.
3. Changes in the soil temperature are reflected by a similar shift in P and by smaller shifts in Q.
4. Both P and Q should correlate with PM(1) but not with the gradient in soil moisture. This is true despite the fact that the emitted radiation may be generated well below the surface.

Discussion

The purpose of this section is to consider the radiometric data in terms of the P and Q parameters, using the radiative transfer model as a qualitative guide. We treat the afternoon and dawn X and L band intensities in their relation to (1)

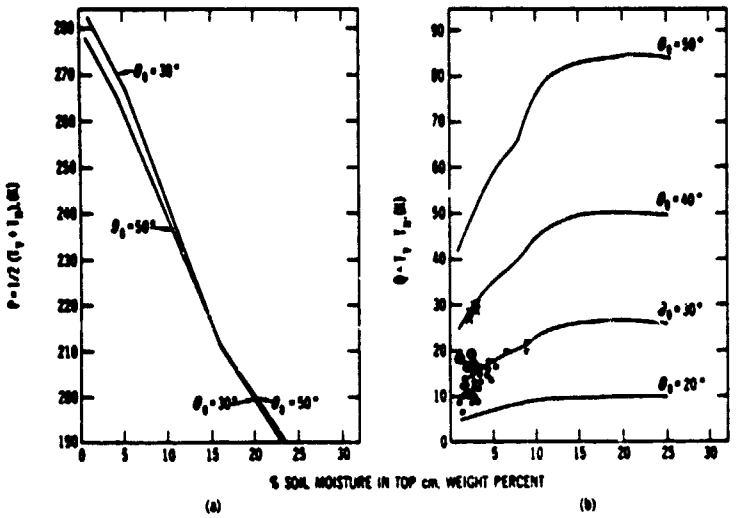


Fig. 3. Calculated values of (*a*) P and (*b*) Q at X band frequency for various look angles as a function of the moisture present in the top layer of soil. A scatter plot of observed values of Q at a look angle of 49° from the afternoon flight is superimposed.

moisture profiles, (2) temperature profiles, and (3) each other. The observed X-band polarizations are compared with the radiative transfer model predictions. From a comparison of polarizations observed during the afternoon and dawn flights we adduce evidence for the nighttime percolation of soil moisture to the surface.

Before considering the microwave signatures per se, let us pause to reflect on the way that we wish to represent the parameter we have been calling soil moisture. In the introduction we noted the soil texture as a parameter affecting the microwave radiation. Textures are expressed as the percent weight that is sand, silt and clay. Sand particles are defined by the size range 2–0.05 mm, silt particles by the range 0.05–0.002 mm, and clay particles by sizes less than 0.002 mm. Textural information is needed for determining the moisture holding characteristics of the soil. In particular, texture information is used to estimate the field capacity (FC) of the soil, i.e., the amount of water remaining in the soil 2–3 days after an irrigation. The relationship used to derive field capacity,

$$FC = 25.1 - 0.21Sa + 0.22Cl \tag{11}$$

where Sa and Cl represent the percent sand and clay in the soil sample, was obtained by performing a multiple regression on data from 100 soils where texture and moisture characteristics were known. A new parameter, the percent field capacity can be constructed

$$F_j = (PM_j/FC) \times 100$$

where PM_j is the moisture content of the jth layer. Using this parameter reduces scatter in the data due to different soil textures.

Figure 6 is a plot of the P values observed for X and L bands (crosses and circles) on the first afternoon flight of April 5, as a function of F_1. Linear regression analyses of these data sets show that for X and L bands

$$P_x = 287.1 - 1.05F_1 \tag{12}$$

and

$$P_L = 300.7 - 1.23F_1 \tag{13}$$

with correlation coefficients $R_x = -0.88$ and $R_L = -0.89$. A similar regression performed on the 21 cm data with P_L as a function of the average percent field capacity of the top two layers gives

$$P_L = 297.0 - 0.53\,(F_2) \tag{14}$$

with a correlation coefficient of −0.73. A comparison of P_x and P_L values from the same fields yields the relationship

$$P_L = 1.04P_x + 0.64 \tag{15}$$

with a correlation coefficient 0.90. The decrease in correlation coefficient with depth combined with the relatively high correlation between P_L and P_x tends to confirm the model predic-

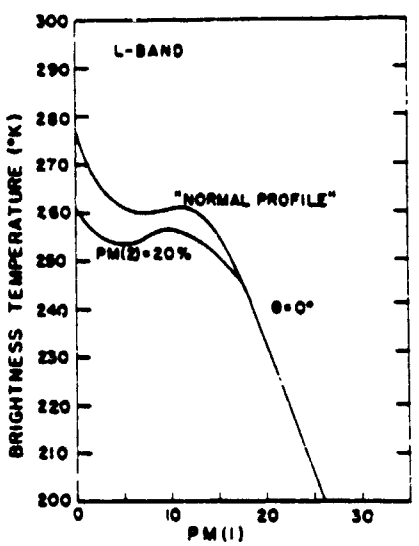


Fig. 4. Brightness temperature at 21 cm for 'normal profile' at nadir look. The model predictions for profiles with PM(2) = 20% and PM(I) > 20 for I > 2 are also given.

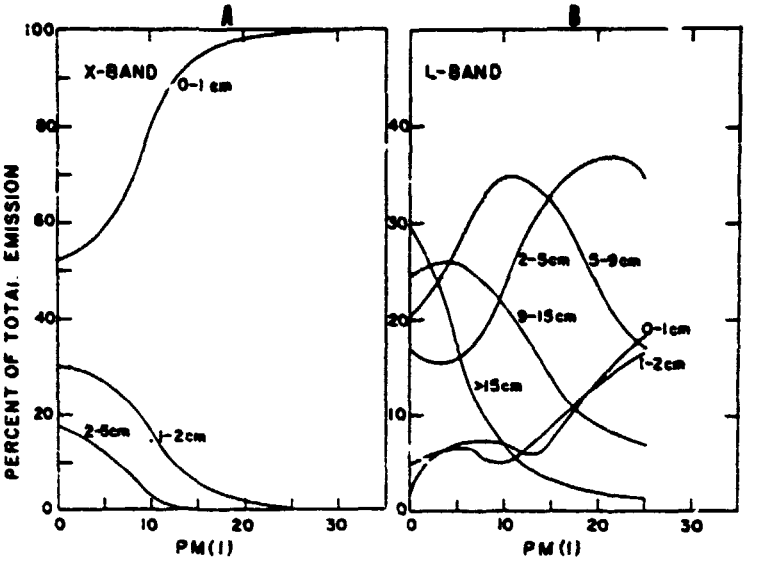


Fig. 5. Percent contributions of the various layers of soil to the values of P calculated from the radiative transfer model for (*a*) X and (*b*) L bands as functions of the percent moisture in the top centimeter of soil.

tion that surface modulations tend to dominate over subsurface effects.

The morning ground truth data set was limited to six fields. Linear regression analyses of X and L band intensities give

$$P_x = 256.4 - 0.27F_1 \tag{16}$$

and

$$P_L = 286.4 - 0.40F_1 \tag{17}$$

with correlation coefficients $R_x = -0.94$ and $R_L = -0.93$. A comparison of these relationships with (12) and (13) shows that the X band temperatures are changed more dramatically in magnitude than the L band temperatures. The reason for this lies in the fact that the major contributions to the L band intensity come from deeper in the soil than those of the X band. Since the diurnal temperature variation of deeper layers is less, the effect is not surprising. Our interpretation of this effect is corroborated by comparing dawn values of P_x and P_L with IR surface temperature [*Schmugge et al.*, 1976]. Values of P_x were always less than T_{IR}. However for many dry fields T_{IR} was less than P_L. An apparent 'emissivity' greater than one indicates that most of the L band radiation from these dry fields originated deep in the soil (Figure 5*b*) where measured physical temperatures were significantly higher than the surface IR temperature.

A second point is the difference in slopes. This is a moisture-temperature effect. Due to the relatively high heat capacity of water and the cooling effect of evaporation, the physical temperature of moist soils varies less over a day than dry soils. For example the diurnal range of a very dry field (<2% at the surface) was observed to be 34°C while for a moderately moist field (>10% at the surface) the range was 23°C. The effect of this difference on microwave signatures is to rotate the P/F_1 curves to lower slopes as we pass from afternoon to dawn observations.

A linear regression was also performed on the full set of X and L band data from the dawn flight (whether or not we had ground data on the field). The relationship was calculated to be

$$P_L = 1.32P_x - 54.9 \tag{18}$$

with a correlation coefficient of −0.89. The increased slope in comparison with the afternoon observations (equation (15)) tends to confirm the hypothesis that L band radiation comes from deeper layers of the soil. The high correlation coefficient between the X and L band results supports the conclusion of *Cihlar and Ulaby* [1975] that the dielectric properties of the surface layer have the dominant effect in determining the emission from the surface. This result needs to be tested with data over a wider range of soil moisture.

The observed values of $(T_v - T_H)$ for 2.8 cm radiation are superimposed on Figure 3*b* as a scatter plot. We note that all but a few points fall between the 10° and 20° curves. The major exceptions were four flat fields designated by crosses

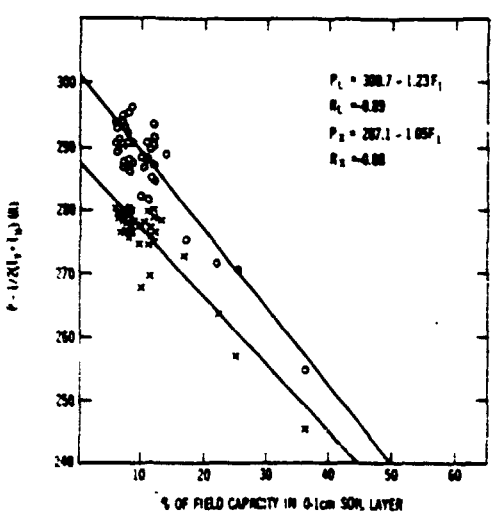


Fig. 6. Scatter plot of observed X (crosses) and L (circles) band P values during the afternoon of April 5, 1974, as a function of the percent field capacity of the top layer of soil.

and four fields that were planted and relatively smooth designated by dotted circles. The effect of the small scale roughness appears as a smaller shift of effective look angle from the specular. In general, the effects of roughness are more pronounced in listed than in planted fields. Considering only dry fields whose percent field capacity in the top layer was less than 10 we find the polarization observed in listed fields to be 12.2 ± 2.0° while on planted fields it was 14.2 ± 2.6°. Had the fields been perfectly smooth we would expect Q values between 40° and 50°. We also note that for moisture less than 15%, Q unambiguously increased with moisture. However, on the basis of polarization alone, it would be difficult to distinguish between a dry flat field and a wet rough one. Joined with the intensity information, the distinction is unmistakable.

Despite the small number of fields sampled at the time of the dawn flight, evidence for a general percolation of moisture to the surface can be found in the X band observations. The radiative transfer model predicts that Q is sensitive to moisture and to look angle as mediated by surface roughness but not to soil temperature. By comparing microwave polarization observations from fields on the first afternoon flight with those at dawn, we hold the look angle (roughness) parameter constant. If the moisture were constant and the temperature decreased by 20°, then the value of Q should decrease by about 2°. If, on the other hand, the moisture increased, the Q should also increase. Figure 7 is a scatter plot of observed morning values of Q as a function of the afternoon values. The line $Q_{dawn} = Q_{day} - 2°$ is drawn for reference. We find that all but 6 of the 96 fields fall above this line. One of the fields, number 60, was wet during the day and should be less moist due to free evaporation. The other fields were observed at the end of the flight and probably represent a sunrise effect on the surface.

We submit that this general shift in the observed polarization values constitutes the microwave signature of the percolation effect described by *Jackson* [1973]. This result is also consistent with the ground measurements. A linear regression analysis of the percent moisture in the first layer during the morning as a function of the moisture in that layer during the previous afternoon gives

$$PM(1)_{AM} = 1.63PM(1)_{PM} + 0.31 \qquad (19)$$

with a correlation coefficient of 0.94.

In some respects the results of the experiment are disappointing. It was hoped that by using a multifrequency microwave system, the moisture content of the soil might be inferred down to 100 cm from one set of observations [*Poe et al.*, 1971]. Such is not the case. Both radiative transfer theory and our observations suggest that we can only know the moisture content of the near surface with any confidence. But not all is lost. We have seen that different bits of information about the thermal and moisture properties of the soil can be gained by making observations at different times.

The microwave signatures of soil moisture, if observed periodically, can be used to construct the history of the near surface moisture. *Ciblat and Ulaby* [1975] have argued that the moisture history of the subsurface must be inferred using water balance models based on the principles of soil physics. Regular microwave observations of the surface soil moisture will provide the input boundary conditions for such models. This should afford an improvement over the point rainfall data currently used.

Summary and Conclusions

The primary goal of the April 5 and 6, 1974, flights over the Phoenix site was to study the feasibility of quantifying the effects of the nonuniform vertical distribution of moisture, surface roughness, and soil type. Despite the disappointing range of moisture in the surface layers of the studied fields, the mission was productive. Our understanding of the microwave signatures of agricultural scenes has been enhanced considerably. Some new items of understanding include the following:

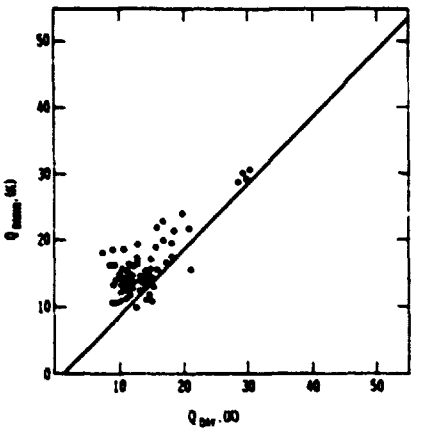


Fig. 7. Scatter plot of observed Q values at X band during the morning of April 6, as a function of their values the previous afternoon.

1. By simultaneously observing T_v and T_H we can form the parameters $P = 1/2[T_v + T_H]$ and $Q = T_v - T_H$. It has been shown that moisture and surface roughness effects can be separated by a comparison of these parameters.

2. By comparing afternoon and dawn observations of P for the X- and L-bands, we found that the slope of the P versus F plot depends on the time that the observation is made, with the afternoon observations yielding the greater slope. The improved correlation obtained when soil moisture is expressed as a percent of field capacity (F) indicates that the effects of variations of soil type can be accounted for.

3. By comparing the values of Q as observed during the afternoon and dawn flights, we find that the microwave signature of the surface rewetting effect described by *Jackson* [1973] is an increase in the polarization. As yet, the data set is too small to calibrate the change of Q with the amount of water that has moved to the surface.

4. By comparing the response of P_s and P_L to the changes between afternoon and dawn, we confirmed the prediction of a radiative transfer theory that the 21-cm radiation reaching the antenna comes from much deeper in the soil than the 2.8-cm radiation. However, the high correlation (~0.9) between P_s and P_L observed in both the afternoon and predawn data confirms the somewhat paradoxical model prediction that the surface layer dominates over subsurface gradient effects at both wavelengths.

This final conclusion is a major disappointment. It was hoped that gradient information could be obtained by comparing a single set of 2.8- and 21-cm observations. Perhaps a study based on a wider variety of moisture profiles will suggest a practical way around the impasse.

Results have indicated that the effect of soil type is well understood and that emission from the soil is dominated by the dielectric properties of the surface layer, which for this case was 1 or 2 cm thick. In addition, these results have indicated a possible new approach, that is, the use of Stokes parameters, in our efforts to quantify the effects of surface roughness. To

verify the usefulness of these parameters for separating the effects of roughness and soil moisture, data over a wider range of moisture conditions is required. To accomplish this, the experiment was essentially repeated in March 1975 with several significant changes. By moving the flight in mid-March, we increased the probability of observing the needed wider variety of moistures. The altitude of the off nadir runs was decreased to 400 m from 800 m to reduce the footprint of the 21-cm radiometer from 280 m to 140 m. Additional nonscanning radiometers were added with wavelengths of 1.67, 1.36, and 0.8 cm. The data from this experiment are still being reduced for analysis.

Acknowledgments. One of the authors (W.J.B.) wishes to thank the National Research Council for supporting his research. We also thank F. T. Ulaby and J. Ciblar of the University of Kansas for helpful discussions on the applicability of the incoherent model. Special recognition should go to Bruce Blanchard (Texas A & M University), Rudy Trabanino (Johnson Space Center), and many students at Arizona State University who designed and executed the ground truth program.

References

Babai, P., Measurement of complex dielectric constant of soil types at *X* band, *Rep. RSC-95*, Remote Sensing Center. Tex. A & M Univ., College Station, 1974.

Cihlar, J., and F. T. Ulaby, Microwave remote sensing of soil water content, *RSL Tech. Rep. 264-6*, Univ. of Kans., Lawrence, 1975.

Cihlar, J., F. T. Ulaby, and R. Mueller, Soil moisture from radar imagery, *RSL Tech. Rep. 264-4*, Univ. of Kans., Lawrence, 1975.

Hammond, A. L., Crop forecasting from space: Toward a global food watch, *Science, 188*, 434, 1975.

Jackson, R. D., Diurnal changes in soil water content during drying, in *Field Soil Water Regime*, pp. 37-51, Soil Society of America, Madison, Wis., 1973.

Newton, R. W., and W. R. McClellan, Permittivity measurements of soils at *L*-band, *Rep. RSC-58*, Remote Sensing Center, Tex. A & M Univ., College Station, 1975.

Peake, W., Interactions of electromagnetic waves with some natural surfaces, *IEEE Trans. Antennas Propagat., AP-7*, 324, 1959.

Poe, G., A. Stogryn, and A. T. Edgerton, Determination of soil moisture content using microwave radiometry, *Final Tech. Rep. 1684 FR-1*, DOC contract 0-35239, Microwave Div., Aerojet-General Corp., El Monte, Calif., 1971.

Schmugge, T., P. Gloersen, T. Wilheit, and F. Geiger, Remote sensing of soil moisture with microwave radiometers, *J. Geophys. Res., 79*, 317, 1974.

Schmugge, T., B. Blanchard, W. J. Burke, J. F. Paris, and J. Wang, Report on the April, 1974 soil moisture flights, *NASA Tech. Note TN D-8199*, Goddard Space Flight Center, Greenbelt, Md., 1976.

Stogryn, A., The brightness temperature of a vertically structured medium, *Radio Sci., 5*, 1397, 1970.

Ulaby, F. T., A. K. Fung, and S. T. Wu, The apparent temperature and emissivity of material surfaces at microwave frequencies, *Tech. Rep. 133-112*, Center for Res., Univ. of Kans., Lawrence, 1970.

(Received December 1, 1977;
revised May 15, 1978;
accepted July 25, 1978.)

Paper 64

MICROWAVE REMOTE SENSING OF SEA ICE IN THE AIDJEX MAIN EXPERIMENT

W. J. CAMPBELL, J. WAYENBERG, and J. B. RAMSEYER
Ice Dynamics Project, USGS, Tacoma, Wash., U.S.A.

R. O. RAMSEIER, M. R. VANT, R. WEAVER, A. REDMOND, and L. ARSENAULT
Environment Canada, 580 Booth Street, Ottawa, Canada

P. GLOERSEN, H. J. ZWALLY, T. T. WILHEIT, T. C. CHANG, and D. HALL
Goddard Space Flight Center, NASA, Greenbelt, Md., U.S.A.

L. GRAY
Dept. of Energy, Mines and Resources, 2414 Sheffield Drive, Ottawa, Canada

D. C. MEEKS
Aerojet Electrosystems Co., 1100 W. Hollyvale St, Azusa, Calif., U.S.A.

and

M. L. BRYAN, F. T. BARATH, C. ELACHI, F. LEBERL, and T. FARR
Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, Calif., U.S.A.

(Received 25 August, 1977)

Abstract. During the AIDJEX Main Experiment, April 1975 through May 1976, a comprehensive microwave sensing program was performed on the sea ice of the Beaufort Sea. Surface and aircraft measurements were obtained during all seasons using a wide variety of active and passive microwave sensors. The surface program obtained passive microwave measurements of various ice types using four antennas mounted on a tracked vehicle. In three test regions, each with an area of approximately 1.5×10^6 m^2, detailed ice crystallographic, dielectric properties, and brightness temperatures of first-year, multiyear, and first-year/multiyear mixtures were measured. A NASA aircraft obtained passive microwave measurements of the entire area of the AIDJEX manned station array (triangle) during each of 18 flights. This verified the earlier reported ability to distinguish first-year and multiyear ice types and concentration and gave new information on ways to observe ice mixtures and thin ice types. The active microwave measurements from aircraft included those from an X- and L-band radar and from a scatterometer. The former is used to study a wide variety of ice features and to estimate deformations, while both are equally usable to observe ice types. With the present data, only the scatterometer can be used to distinguish positively multiyear from first-year and various types of thin ice. This is best done using coupled active and passive microwave sensing.

1. Description of Experiment

The Arctic Ice Dynamics Joint Experiment (AIDJEX) was a 5-year experiment designed to achieve a better understanding of the behavior of Arctic sea ice. Three pilot experiments in the Beaufort Sea during the springs of 1970, 1971, and 1972 preceded the Main Experiment which ran from March 1975 through April 1976. The design of all the field experiments stressed the need of acquiring remotely sensed ice data in order to accomplish the scientific goals. The remote sensing programs for the early experiments demonstrated new approaches for ice observation; by the time of the main experiment, therefore, a comprehensive remote

Boundary-Layer Meteorology 13 (1978) 309–337. *All Rights Reserved*
Copyright © 1978 by D. Reidel Publishing Company, Dordrecht, Holland

sensing plan based on the use of aircraft and satellite data had been developed (Weeks and Campbell, 1975).

One of the exciting aspects of the pilot studies was demonstrating the possibility of attaining an all-time, all-weather capability of ice observation by microwave techniques (Wilheit *et al.*, 1972; Gloersen *et al.*, 1973; Campbell *et al.*, 1974). Thus, the remote-sensing program of the AIDJEX Main Experiment put great stress on both active and passive microwave sensing based on aircraft and surface observations of all ice types during all seasons. Here we wish to present an overview of this program.

The major part of the aircraft observational program was performed by the NASA Convair 990 Airborne Laboratory, which flew 18 missions during 4 extended periods in the spring, summer, and fall of 1975 and the spring of 1976, referred to as phases one to four, respectively. Important microwave data were also acquired by two Canadian aircraft of the Department of National Defence (DND): an Argus and C-47 remote-sensing aircraft. Table I lists the sensors and various remote-sensing platforms.

The difficult task of coordinating these flights and guiding them accurately over a large and constantly shifting and deforming sector of the Beaufort Sea ice was helped considerably by two-way communication between the main ice camp (Big Bear) and the air field from which most of the flights originated (Eielsen AFB near Fairbanks) via satellite link, by operating radio beacons at the three ice camps at the apexes of the triangle, and by the use of inertial navigating systems in the aircraft. The positions of the AIDJEX triangle during the four phases of the remote-sensing program are shown in Figure 1. A great deal of the actual accomplishment of these difficult and numerous missions, every one of which was successful, was due to the flying skills of the aircraft crews.

The remote-sensing programs of the pilot experiments clearly showed the need for a variety of surface measurements in order to interpret and analyze the aircraft data accurately; therefore this aspect of the main experiment was emphasized. A Canadian flextrack vehicle carrying three boom-mounted microwave radiometers and an infrared radiometer operated in the area of the main camp during the two spring seasons (Table I). Three test zones near the main camp were selected for detailed study. One was exclusively composed of first-year (FY) ice, one was primarily composed of multiyear (MY) ice, and one was a mixture of primarily MY and some second-year ice. Each zone had an area of approximately 1.5×10^4 m^2. Within these zones the surface brightness temperatures (T_b) were mapped under a variety of weather and temperature conditions. In addition, detailed dielectric structural and textural measurements were performed. Two 10-km test lines were also selected which were overflown repeatedly at low altitudes by the aircraft. These lines transected a wide variety of ice features including a large lead adjacent to the main camp in which many thin ice forms were observed.

The surface measurements were carried out during relatively stable periods (spring 1975 and 1976, fall 1975) and coincided with the various aircraft

TABLE I

Remote sensing platforms and sensors

Sensor/sensor platform	CV 990, NASA	C-47, CCRS	Argus, DND	Flextrack
Radiometer	Aerojet	Aerojet		Aerojet
frequency	19 GHz	37 GHz		4.99, 13.4, 37 GHz
polarization	1.8 cm	0.81 cm		6.0, 2.23, 0.81 cm
incident angle	nadir, ± 55° starboard and port scanning	45° forward		H, V all three, nadir to 55° forward
resolution	500 × 500 m	15 × 15 m		1.5 × 1.5, 0.7 × 0.7, 0.4 × 0.4 m
Imaging Radar	JPL		Motorola AN/ADS 94D	
type	synthetic aperature		real aperature	
frequency	1.215 GHz		9.2 GHz	
wavelength	24.5 cm		3.25 cm	
polarization	HH		HH	
incident angle	0–55° starboard		45–88° starboard and port	
resolution				
range	25 m		30 m	
azimuthal	25 m		40–200 m	
swath	14 km		25 km	
Scatterometer		Ryan 720		
frequency		13.3 GHz		
wavelength		2.25 cm		
polarization		HH, HV, VV, VH		
incident angle		160° fore and aft		
resolution		15 × 15 m		
Infrared		Deadalus		PRT-5
wavelength		8–14 μm		8–14 μm
incident angle		155° starboard and port		nadir to 55° forward
Mapping Camera	RC-10	RC-10		Handheld Nikon
Altitude Data Collected	12 000 m	300 m	900 m	0 m (surface)

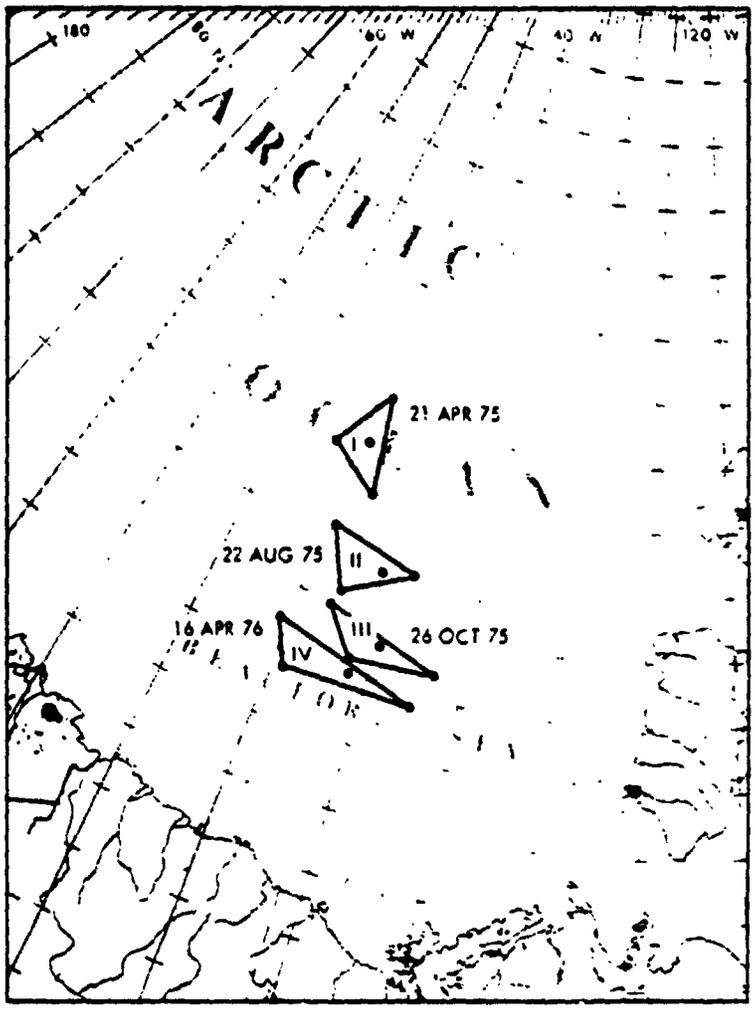


Fig. 1. Map of AIDJEX study site showing locations of AIDJEX station triangle during the four phases of the remote sensing experiment.

overflights, except during the summer of 1975 when no detailed surface measurement program took place.

2. *In Situ* Observations

All the *in situ* measurements were made at Big Bear, the original main AIDJEX camp. The main objective was to observe changes in the physical properties and surface-brightness temperatures caused by the summer melt. Three large test sites were established and surveyed during April 1975. The plan called for resurvey of

the sites in the fall. Unfortunately, Big Bear had to be evacuated during October 1975 and no radiometric measurements were made. The abandoned camp was revisited and resurveyed during April 1976; it was found that there had been a loss of about 50% of the multiyear floe test areas of the original Big Bear camp and the entire runway area which in the spring of 1975 had consisted entirely of first-year sea ice.

The passive microwave measurements were made by three microwave radiometers from Aerojet Electro Systems mounted on an adjustable platform above a flextrack vehicle (see Table I). The data, including infrared surface temperatures, were recorded on magnetic tape. The physical properties were measured using standard field techniques (Meeks *et al.*, 1974a, b). A more detailed description and analysis of the AIDJEX remote sensing program conducted at Big Bear will be published later after all data have been reduced.

2.1. Physical properties

Salinity, temperature, density, and surface roughness of the ice are among the most important properties determining the signatures of passive and active microwave sensors. In this discussion there are two classes of ice we would like to distinguish – first-year (FY) and multiyear (MY) ice. During phase I, the FY ice in the Big Bear area (Figure 1) reached a maximum thickness of 1.75 m towards the end of June 1975. The salinity curve shown in Figure 2, labeled FY, was obtained during April 1975 from our Venus site (airstrip). This a typical thick FY ice profile with a thickness of 1.61 m. Near the top, the salinity is very high, as it is at the ice-water interface. Large fluctuations near the top are often associated with rafting, i.e., overriding of a piece of ice on another. From a depth of 0.9 m to the skeleton layer, which is about 0.5-m thick at the bottom of the ice layer, the salinity fluctuates between 4 and 5%.

During the summer, melt ponds form on the surface of the ice, collecting water from melting snow and eventually from melting ice on the surface. The melt water percolates through the surface layers of ice, entraining at the same time the brine. This causes a complete desalination of the ice surface layer, as evidenced by the curve shown in Figure 2 labeled FY + 1. The bulk of the ice brine is lost to the ocean through a network of brine channels. The curve FY + 1 was plotted from samples obtained during October 1975 near the same location as the sample FY. Its thickness was 1.28 m, which means that a net thickness loss of 0.47 m occurred. The loss is 0.20 m due to surface melt and 0.27 m due to melt at the ice-water interface. The bulk salinities below the surface layers vary between 3–4%. Changes in the salinity profiles between the late fall and April of the following year are minimal except for those of the new ice which commences growth at the fall ice-water boundary and exhibits salinities of the same level as FY ice.

Similar changes causing further desalination are observed in the upper part of MY ice as evidenced by the curves labeled MY and MY + 1 in Figure 2. The most significant change takes place in MY ice at a depth of approximately 1.15 m. The

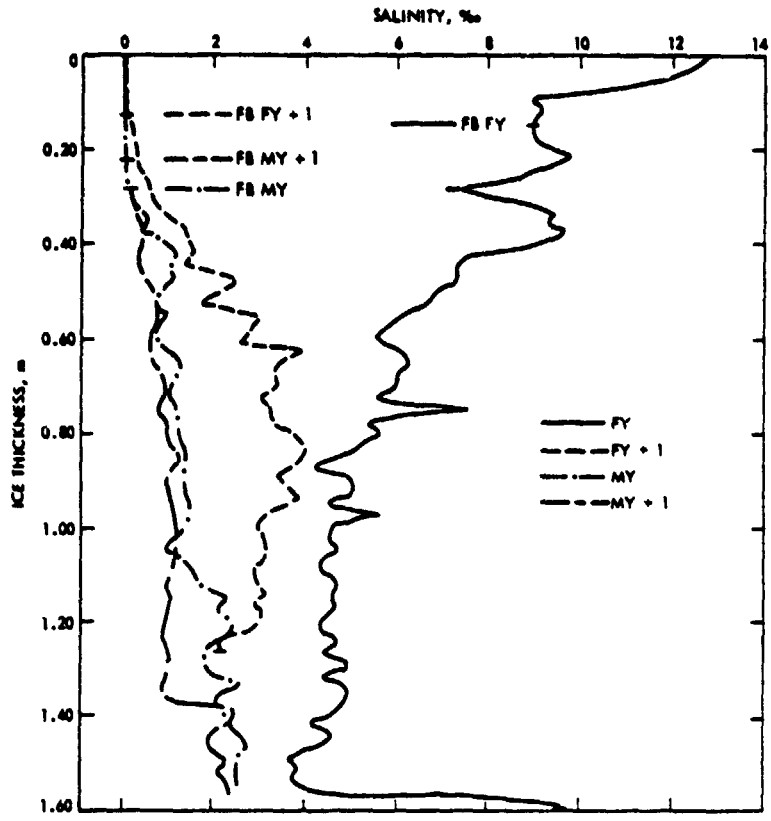


Fig. 2. Salinity as a function of depth below surface of ice observed in (a) first-year ice (FY); (b) first-year plus 1 year (FY + 1); (c) multiyear ice (MY); and (d) multiyear plus 1 year (MY + 1). Only for curves FY and FY + 1 has the total thickness been shown, whereas the other curves would extend to greater depths. FB designates the respected freeboard for each curve.

bulk increase in salinity from 1–1.5% to 2–3% now occurs at the 1.4-m depth for the MY + 1 curve.

The density for FY as shown in Figure 3 is high, usually fluctuating around 0.90 to 0.92 g cm^{-3}, with the exception of the surface layer, which often is slightly below this value. Because of the irregular surface and the high brine content of the FY ice, it is very difficult to measure the density of the surface layer to any reasonable accuracy. In most cases, however, it is above 0.8 g cm^{-3} (Ramseier *et al.*, 1975). For ice older than FY, the near surface density is much lower than for FY ice. It is not uncommon that the surface density exhibits a linear extrapolation of the trend shown near the surface in Figure 3 by curves MY and MY + 1. From the examination of the densities of many samples, it can be seen that the fluctuations in density for ice older than FY ice are always much larger than those for FY ice. The freeboard (FB), as in the cases illustrated, occurs at a depth in the ice where the density becomes greater than 0.8 g cm^{-3}

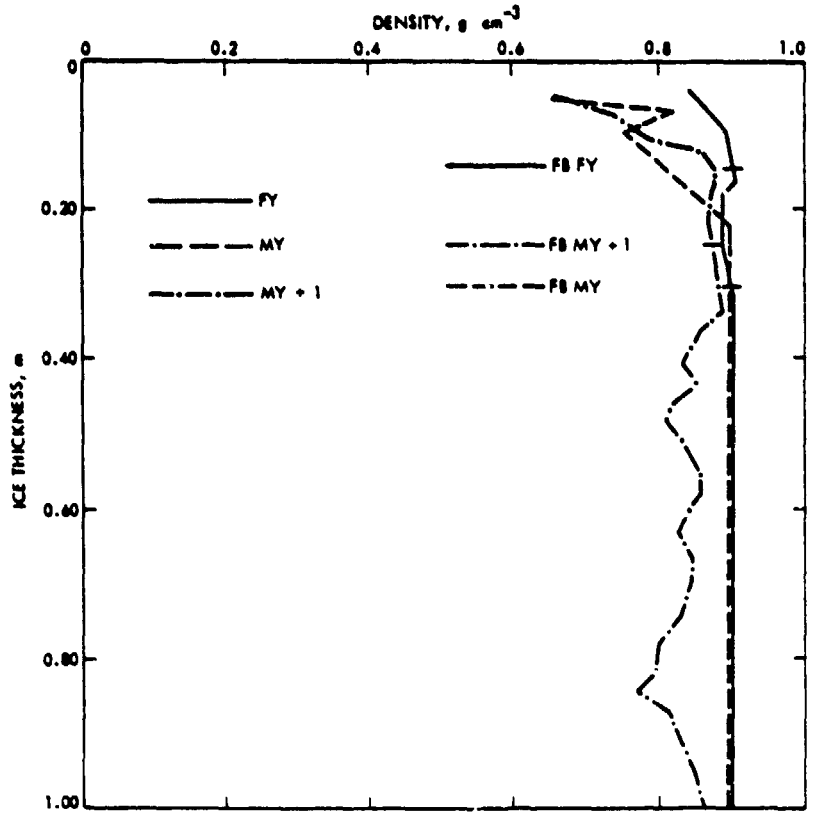


Fig. 3. Density as a function of depth below surface of ice observed in FY and MY ice. The MY + 1 was measured in the fall when its freeboard tends to be above the MY ice.

During the summer, when melt occurs, the ice is at or very near the melting point for long periods of time. In addition to the desalination mentioned earlier, the surface layers undergo complete recrystallization. Figure 4 represents a thin section of FY ice 0.10 m wide. These sections are usually prepared on a microtome to a thickness between 0.02 and 0.05 cm, and then photographed between crossed polaroid filters. This technique reveals the crystal structures of ice samples as evidenced by these Figures 4a and 4b, the latter one representing MY ice, i.e., FY ice recrystallized MY ice going from Figure 4a to 4b.

Very often FY ice has a surface layer consisting of frazil ice, i.e., the grains are equiaxed in one plane but very thin in the direction of depth. The crystallographic orientation is random and the grain size varies between 0.1 and 0.2 cm. This layer is usually responsible for the high salinity content near the surface and the somewhat lower density. The thickness of the layer can vary substantially between 0.02 and <0.36 m (Ramseier *et al.*, 1975). This layer is followed by columnar ice, which

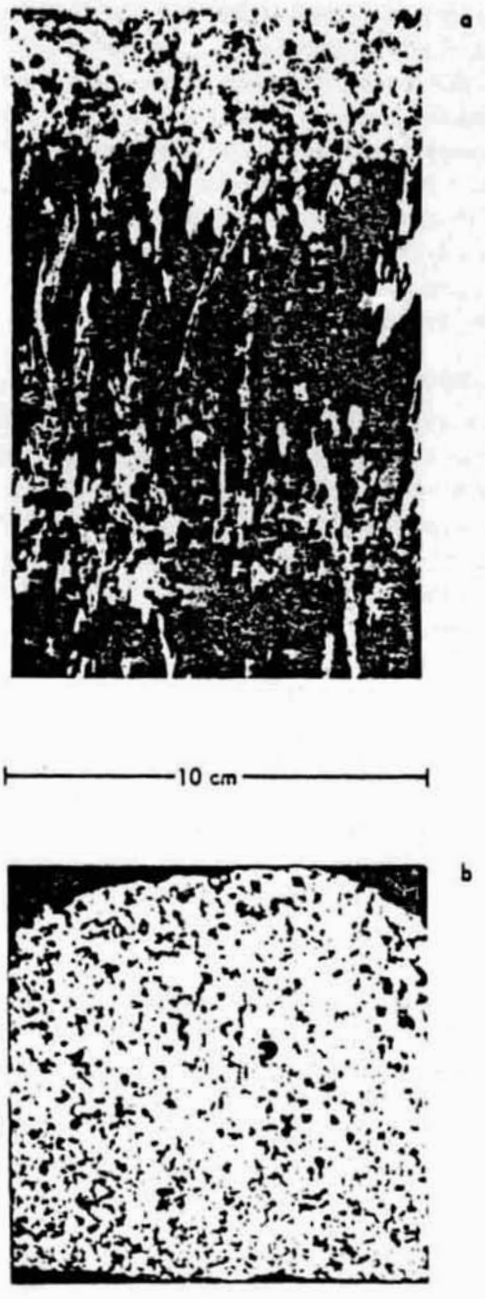


Fig. 4. Thin section of (a) FY ice and (b) MY ice. MY ice is derived from FY ice through summer melt and complete recrystallization. This is generally true for the part above the FB. Further below, the columnar structure is usually maintained but at a lower salinity. The width of the thin section is 0.10 m.

ORIGINAL PAGE IS
OF POOR QUALITY

exhibits a high density, elongated grains, and a preferred crystallographic orientation. In contrast, Figure 4b shows the surface part of a MY ice thin section. It is quite evident that, due to recrystallization, there is no resemblance between FY and MY ice. The grains are equiaxed, varying in size between 0.1–0.3 cm, and have a random crystallographic orientation. The density is low, as discussed earlier, and the ice is completely desalinated above the freeboard (Meeks *et al.*, 1974a).

It is important to appreciate these variations in material properties of FY and MY sea ice, especially the changes which take place near the surface, because it is these variations which determine the dielectric properties of the ice and produce the signatures one observes from passive and active microwave sensors.

2.2. Dielectric properties

The key to any meaningful understanding of microwave signatures of ice requires a knowledge of the dielectric behavior of sea ice, i.e., the values for the real (ε') and imaginary (ε'') part of the dielectric constant as a function of frequency, temperature, salinity, density, and orientation of the ice. For brine, ε' and ε'' are much larger than for pure ice, so, despite the smaller quantity of brine present, its influence will dominate the dielectric behavior of the mixture. It is apparent that the placement, shape, and orientation of the brine inclusions in the ice sheet will have a profound influence on the dielectric properties.

Measurements of the *in situ* properties of ice and the recent work by Vant (1976) on the dielectric properties of predominantly FY ice and some MY ice have enabled localization of the effect of some of the variables mentioned earlier.

To determine the dielectric properties, a novel type of wideband (100 MHz to 7.5 GHz) 'coaxial-cage' transmission line was designed based on an earlier study by Vant *et al.* (1974) It measured FY ice samples over a wide range of salinities (5.1 to 10.5 ‰) and temperatures (−5 °C to −4 °C) during April 1975 at Big Bear. A somewhat less comprehensive set of measurements was made on MY sea ice samples.

Figure 5 summarizes the dielectric absorption losses obtained as a function of frequency during the AIDJEX study. Results by other investigators have been included for comparison. The solid line represents the prediction for the absorption losses based on the theoretical model developed by Vant (1976). The basic physical properties of ice required in the model are the salinity S, the temperature T, and the density ρ. In addition, some derived information such as the ellipsoid axial ratio a/b and the angle of orientation θ of the ellipsoids is required. The brine inclusions are represented by the ellipsoids.

2.3. Passive microwave measurements

In general, the surface microwave measurements were performed as described by Meeks *et al.* (1974b). The test areas on the MY ice floes Neptune (N) and Apollo (A) measured 150 × 200 m and were located adjacent to Big Bear. Each short side of the rectangular area was divided into 10-m intervals, and the 16 tracks, each

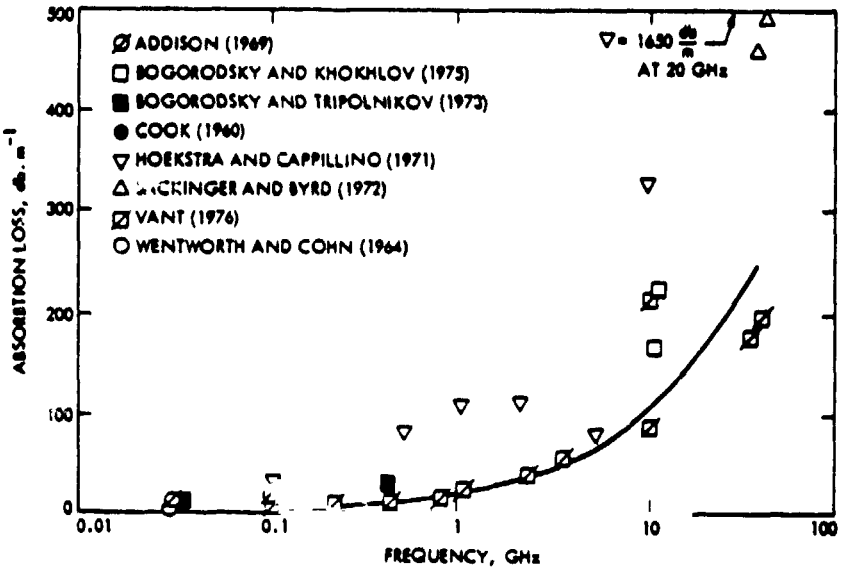


Fig. 5. Summary of loss data (in db m^{-1}) versus frequency. Data taken at temperatures between −7 °C and −12 °C, and salinities between 5‰ and 10‰. Solid curve is based on theory. The parameters used were: $T = -10$ °C, $S =$ ‰, $a/b = 20$, $\theta = 45°$, and $\rho = 0.91$ g cm^{-3}

200 m long, were then surveyed with the radiometers and PRT-5 mounted on the flextrack (see Table I).

Figure 6 shows the April 1975 brightness temperatures (T_b) plotted in the form of histograms. The measurements were made at 15, 30, and 45° viewing angles. Figure 6 shows only the 45° angle data for the wavelength of 2.2 and 0.81 cm at both vertical (V) and horizontal (H) polarizations. Also shown in the histograms are the calculated average brightness temperatures for the given variables. It is interesting to note that the N area has an average T_b which is slightly below that for A. Even though the difference is small, area N did have a section which was rather hummocky, resulting in lower T_b, as evidenced by tailing off of the histogram at the 2.2-cm wavelength for both the horizontal and vertical polarizations as compared to the equivalent A measurements. The histogram also shows that A had more ice of higher T_b than N. By comparing the shape of salinity curves such as in Figure 2, there is a strong indication that the associated brightness temperature with the FY + 1 ice is in the range of 210–220 K for H polarization and 210–250 K for V polarization. Even though, on the whole, there was not much second-year ice present, A had about double the amount of N. The rest of the ice in both areas consisted of MY ice. The arrows in each histogram represents the Juno snow-free ice T_b as plotted from Table II. Table II summarizes the average T_b in terms of frequency, polarization, look-angle (45° or nadir (N) for three test sites with Juno representing FY ice). A number of important deductions can be made.

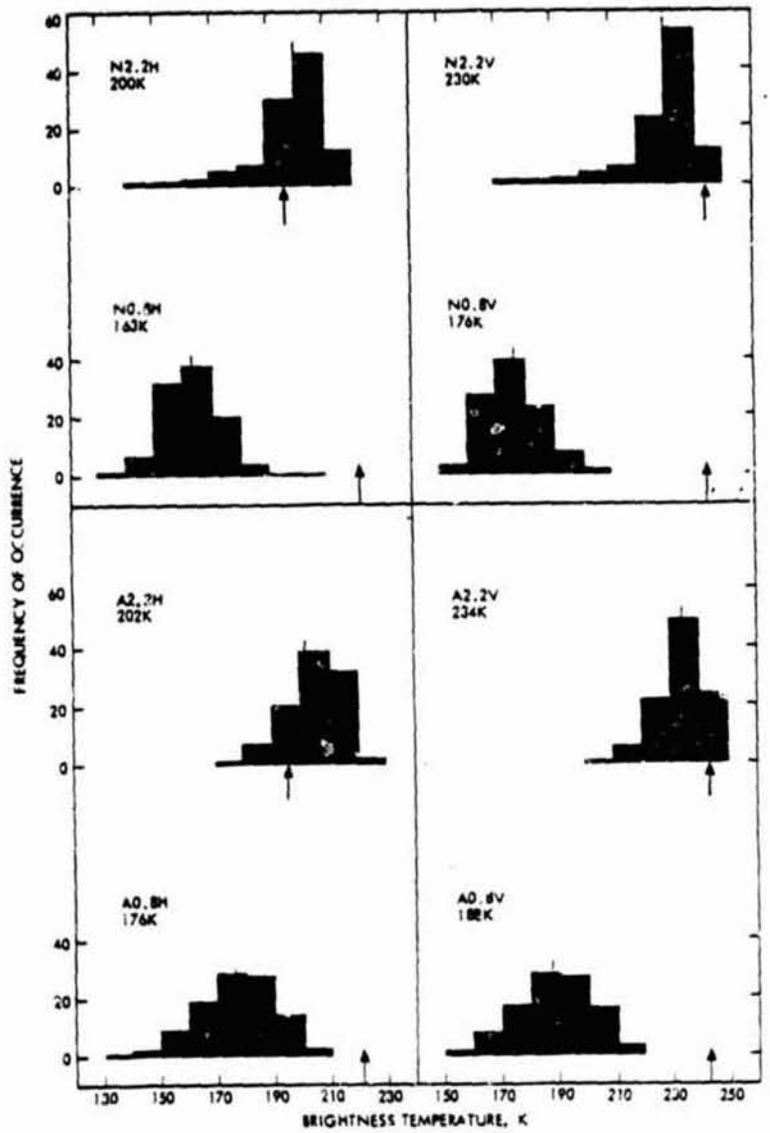


Fig. 6. Histogram of 2.2 cm and 0.81 cm radiation emanating from test areas Neptune (N) and Apollo (A) with horizontal (H) and vertical (V) polarization at a viewing angle of 45°. The large number appearing below the test area identification represents the average brightness temperature of each area at the stated wavelength and polarization. The arrows designate the T_b of snow-free FY ice

— By comparing the polarizations at both wavelengths, it is clear that the longer wavelength gives a larger ΔT_b than the shorter wavelength.

— At the shorter wavelength, the ΔT_b is greater for MY than for FY ice.

— It is quite evident that FY ice can clearly be distinguished from MY ice even at the shorter wavelength. This applies also to a nadir-looking instrument.

TABLE II

Average brightness temperatures for multiyear and first year areas

Wavelength λ cm	Neptune (MY) T_b K	Apollo (MY) T_b K	Juno (FY) T_b K	Juno snowfree T_b K
2.2 H (45)	200	202	220	196
2.2 V (45)	230	234	250	243
2.2 NADIR	222	228	237	225
0.8 H (45)	163	176	241	222
0.8 V (45)	176	188	247	243
0.8 NADIR	186	200	243	236

— There is also an indication that going from nadir to a larger viewing angle improves the discrimination between FY and MY ice.

The last two observations confirm those made by Gloersen *et al.* (1973) from data collected by the CV 990 aircraft.

The last column in Table II represents preliminary results on the effect of snow on brightness temperature. The following deductions can be made.

— In all cases the T_b is lower with the snow removed from the surface.

— The effect of snow seems to be more pronounced at the longer wavelength.

— In both cases the ΔT_b between snow-free and snow-covered ice is greater for the horizontal polarization than the vertical one. At the shorter wavelength, the ΔT_b for vertical polarization is very small and could be considered insignificant.

— The ΔT_b for FY ice at both wavelengths is greater for snow-free FY ice than for FY ice covered with snow. In fact, at the shorter wavelength the difference is significant with snow-free FY ice as compared to snow-covered FY ice.

In conclusion it can be said that the longer wavelength gives good discrimination between H and V polarizations for FY ice, whereas the shorter does not, except in the case of snow-free FY ice. The shorter wavelength gives excellent discrimination between FY and MY ice as compared to the longer wavelength where the discrimination is not so pronounced. The effect of snow cover is very pronounced on FY ice, which may have some important considerations in the interpretation of Electronically Scanning Microwave Radiometer (ESMR) imagery. Indications at present are that this is also the case for MY ice. However, this will have to await further analysis of the data.

3. Airborne Passive Microwave Measurements

3.1. General

The Electronically Scanning Microwave Radiometer (ESMR) detects radiation at a wavelength of 1.55 cm, with a brightness temperature sensitivity and stability of

about 3 K and a beamwidth of 2.8° (Table I). It was used previously, in 1971 and 1972, for producing mosaic images of earlier AIDJEX test areas (Gloersen *et al.*, 1973; Campbell *et al.*, 1975a, b, 1976). As in the past, parallel aircraft tracks were flown during the Main AIDJEX Experiment but with a 50% overlap of the ±50° scans of the ESMR to accommodate the Synthetic Aperture Radar (SAR) which views only to one side of the aircraft.

The ESMR images of the test area, Figures 7 through 10, were produced automatically by a computer using latitude and longitude data from the aircraft's inertial navigation system. The data shown in these figures are averages of the two values obtained for each map pixel as a result of the overlap. A standard map projection was used with the center of the projection located approximately in the center of the image. Because the test areas were small, the map projection is very nearly rectilinear.

3.2. Morphological variations

On comparing the microwave and photographic mosaic images for 22 August and 18 August, respectively (Figures 8 and 18), it is evident that all the medium-to-large floes that can be distinguished in the photograph are also discernible on the microwave image. These same floes are also readily seen on 27 August (Figure 9) and, except for the large aggregation of multiyear ice in the center of the test area, are discernible on 26 October (Figure 10), but not as easily on 21 April (Figure 7).

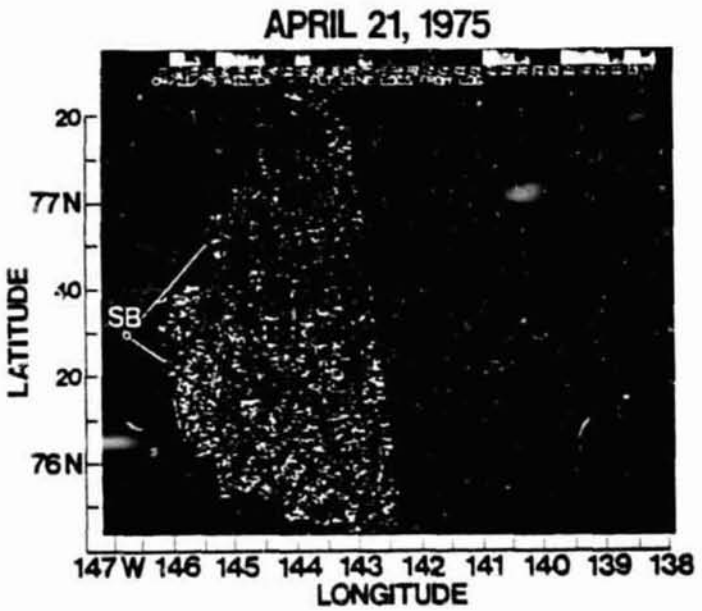


Fig. 7. Electronically Scanning Microwave Radiometer (ESMR) image at a wavelength of 1.55 cm of the AIDJEX area on 21 April, 1975.

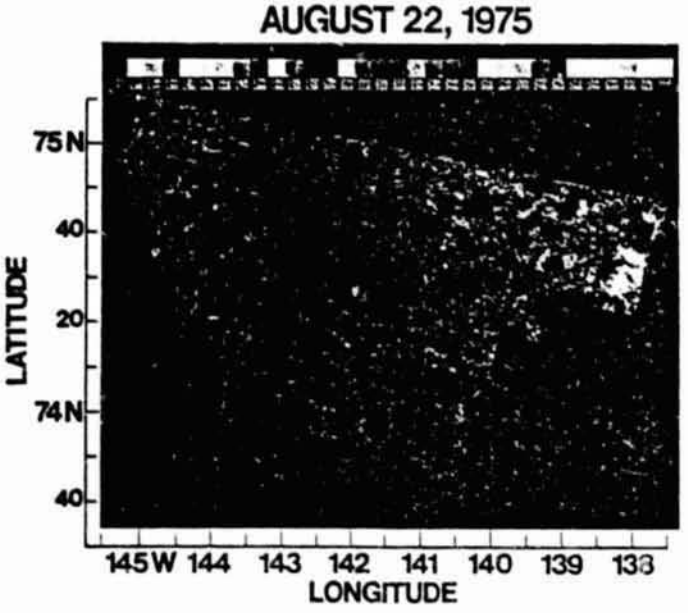
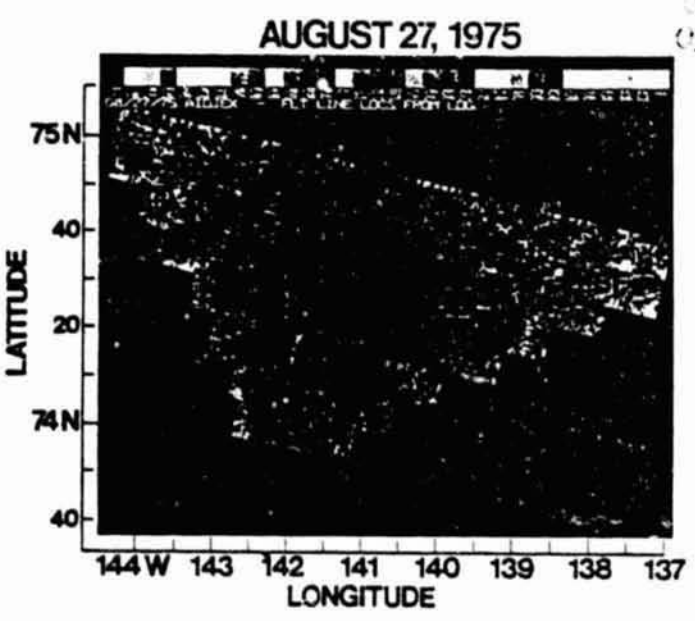


Fig. 8. Electronically Scanning Microwave Radiometer (ESMR) image at a wavelength of 1.55 cm of the AIDJEX area on 22 August, 1975.



ORIGINAL PAGE IS
OF POOR QUALITY

Fig. 9. Electronically Scanning Microwave Radiometer (ESMR) image at a wavelength of 1.55 cm of the AIDJEX area on 27 August, 1975.

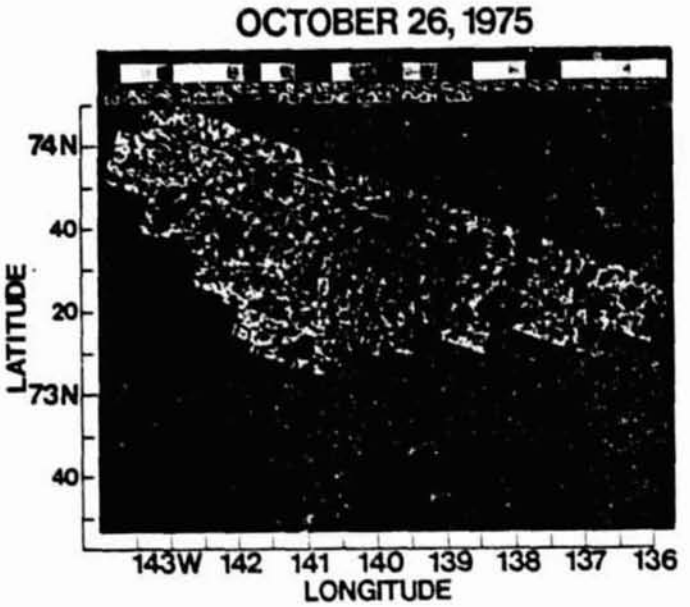


Fig. 10. Electronically Scanning Microwave Radiometer (ESMR) image at a wavelength of 1.55 cm of the AIDJEX area on 26th October, 1975.

On 21 April, the floes on which the AIDJEX camps Blue Fox (77.28 N, 143.66 W) and Caribou (75.73 N, 142.73 W) were located can be distinguished and located in the other three microwave images. This is also true of the larger floes (*A*). The floes on which the other camps were located on 21 April could not be identified on succeeding images. In any event, it is clear that considerable dynamic activity took place within the AIDJEX triangle in the period from 21 April to 22 August, and that the deformation was not as intense during the following periods. Also clearly evident during these studies is the lowering of the microwave emissivity of these individual floes in October compared to April, indicating an increase in the multi-year fraction in the floes. In fact, since these floes survived the summer melt, they are by definition multiyear floes in October. The April image indicates that the floes at that time consist of small multiyear pieces cemented together by first-year and new ice. Thus we have observed for the first time the metamorphosis of specific sea ice floes from a mixture of ice types to pure multiyear ice over the summer as evidenced by the change in the microwave signatures.

The floes exhibit a distinctly different microwave signature in August. The combined process of melting and percolating causes a net increase in the vertically integrated moisture content in the freeboard volume of each floe. This increases the imaginary part of the index of refraction (n'') of the ice; therefore its microwave emissivity (Chang *et al.*, 1976; Chang and Gloersen, 1975) increases, but only for n'' less than one. As n'' increases further, the emissivity drops, since n'' then contributes significantly to the reflectivity of the surface. This phenomenon has also

been observed experimentally in snowpacks (Edgerton *et al.*, 1971). Thus, variation in the emissivity of the floes in August images is interpreted as variation in the liquid water content in the freeboard volume of the floes. Finally, in these images, the change in lead orientation from predominantly east–west on 22 August to north–south on 27 August is noteworthy.

Histograms (not shown) of the brightness temperatures (T_b) represented in the four microwave images (Figures 7–10) are characterized by a multiple peak distribution. In April (Figure 7) only two peaks appear, with relative populations of 0.82 centered at 207 K and 0.18 centered at 229 K. In August (Figures 8 and 9), both data sets show a triple peak distribution, with relative populations of 0.02, 0.34, and 0.55 centered at 121, 228, and 243 K, respectively, for 22 August, and 0.01, 0.35, and 0.54 centered on 129, 221, and 235 K for 27 August. In addition, there is a broad distribution of data points between the first two peaks on both days. On 26 October, again only two peaks appear, with populations of 0.90 and 0.10 centered on 201 and 217 K.

The April and October distributions can be attributed to the presence of only two significant components in the field-of-view: first-year (higher T_b) and multiyear (lower T_b) sea ice. This interpretation is strengthened by the 20 K separation of the peaks which has been observed previously (Gloersen *et al.*, 1973). It is noteworthy that there was an increase in the multiyear fraction from 0.82 to 0.90 from spring to fall.

In August, under melt conditions, the distinction between multiyear and first-year sea ice has been lost, both from radiometric and morphological points-of-view (Gloersen *et al.*, 1978). The peak at 243 K (and 235) is attributable to sea ice with a moist surface and high emissivity. The peak at 121 K (and 129) is attributable to open water. The broad distribution between the first two peaks is attributed to open-water/ice mixtures of decreasing concentrations as brightness temperature increases. We are uncertain as to the interpretation of the second peak, but believe it to correspond to smaller floes with higher moisture content in the surface layer due either to higher temperatures (thinner floes) or surfaces occasionally awash. This hypothesis appears to be supported by the contrasting brightness temperatures of the large and small floes in Figures 8 and 9. The populations in the broad distribution between the first two peaks were weighted with a linearly decreasing factor (Gloersen *et al.*, 1975) to obtain the amount of unresolved open water represented by those populations; the sum of these weighted factors was added to the populations of the first peak (open water) to obtain percent open water of 8.5% and 9.6% on 22 and 27 August, respectively.

4. Airborne Active Microwave Measurements

4.1. General

A Synthetic Aperture Radar (SAR) operating at L-band (Table I) collected the data presented as mosaics of the AIDJEX study site (Figures 11 to 15 inclusive).

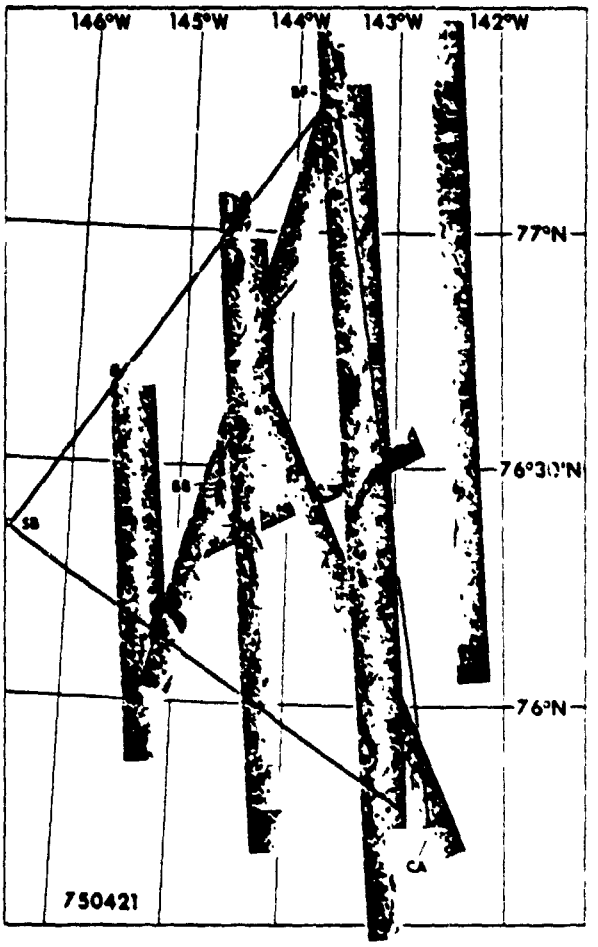


Fig. 11. L-band, HH polarized mosaic of the AIDJEX study site for 21 April, 1975. Data collected using Jet Propulsion Laboratory Synthetic Aperature Radar.

These mosaics were constructed with the available navigation data supplied by the aircraft inertial-navigation system (LTN-51). Due to the drift of this instrument and the navigation errors, which are typically sinusoidal with periods of 84 min and amplitudes of about 0.5 to 1 km h^{-1}, location on these mosaics is only roughly correct, within about 3–5 km on ~0.04° lat. and 0.12° long. (Leberl *et al.*, 1976a; Van Roessel *et al.*, 1974). In addition, because the radar operates in the slant-range mode, there is a significant geometric distortion (compression) in the area nearest the flight line (near range). This may be rectified by digital processing techniques (Thompson *et al.*, 1972). Such geometric corrections have, however, not been applied to the data presented here.

The change in the average brightness temperature across the radar image results from the decrease of the average backscatter cross-section and an increase in the

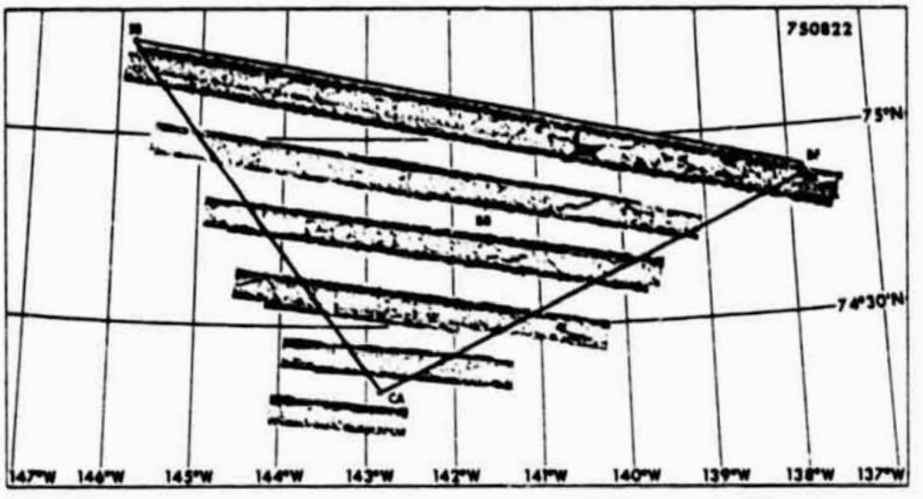


Fig. 12. L-band, HH polarized mosaic of the AIDJEX study site for 22 August, 1975. Data collected using JPL Synthetic Aperature Radar.

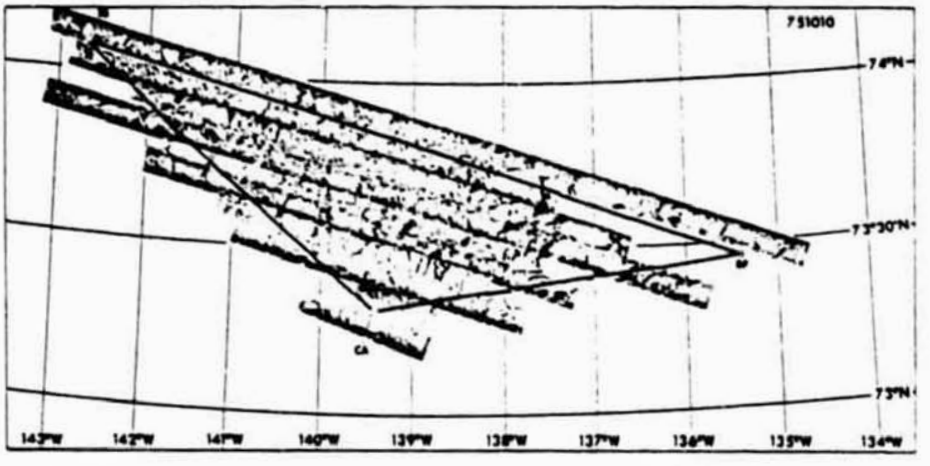


Fig 13. L-band, HH polarized mosaic of the AIDJEX study site for 10 October, 1975. Data collected using JPL Synthetic Aperature Radar.

radar-to-surface distance (i.e., slant range) as the incident angle increases. Thus, on the average, the imagery appears brighter in the near range. It should be mentioned that during the August 1975 flights, a controllable calibrated gain system was added to the radar. This system compensates as much as possible for the decrease in brightness across the image.

The synthetic-aperture radar repeatedly samples coherent electromagnetic waves to synthesize a large aperture. The coherence of the scatter leads to a speckled appearance of the radar imagery similar to the effect observed when a

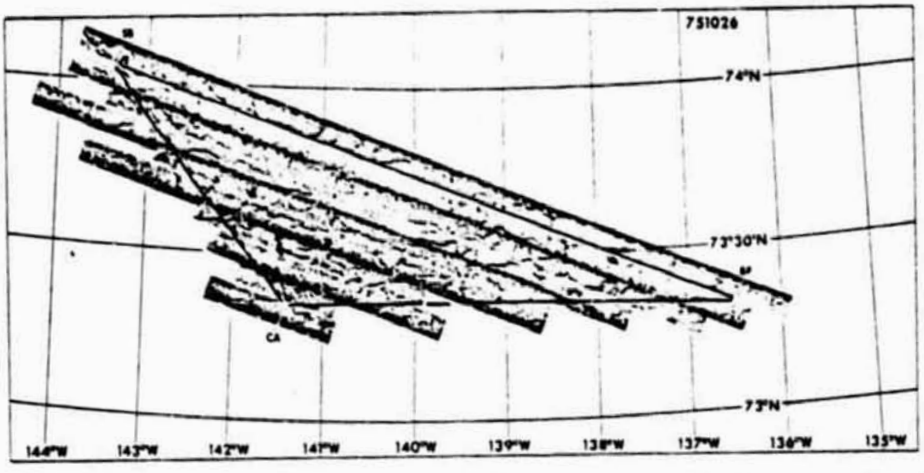


Fig. 14. L-band, HH polarized mosaic of the AIDJEX study site for 26 October, 1975. Data collected using JPL Sythetic Aperature Radar.

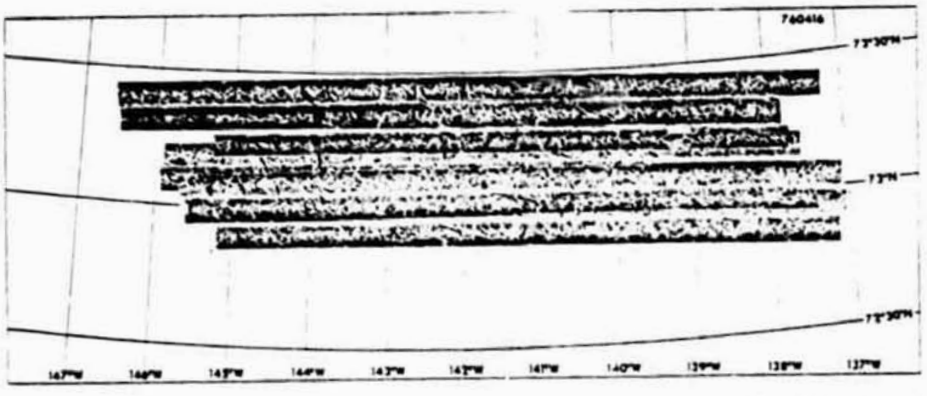


Fig. 15. L-band, HH polarized mosaic of the AIDJEX study site for 16 April, 1976. Data collected using JPL Synthetic Aperature Radar.

rough surface is illuminated with a laser light. For detailed analysis of the principle of synthetic-aperture radar, the reader is referred to the literature (Goodman, 1968; Rihacaek, 1969; Harger, 1970).

A Motorola AN/APS94D real-aperture side-looking radar (Table I) collected the data in Figure 16. The system was flown aboard an Argus aircraft to image swaths of 25, 60, or 100 km width on either or both sides of the aircraft. All the images contain a blind strip down the center of the track, which is outside the field-of-view of the antenna and is about twice as wide as the altitude of the aircraft. Cross-track resolution is constant at about 30 m and along-track resolution is approximately 8 m times the range in kilometers.

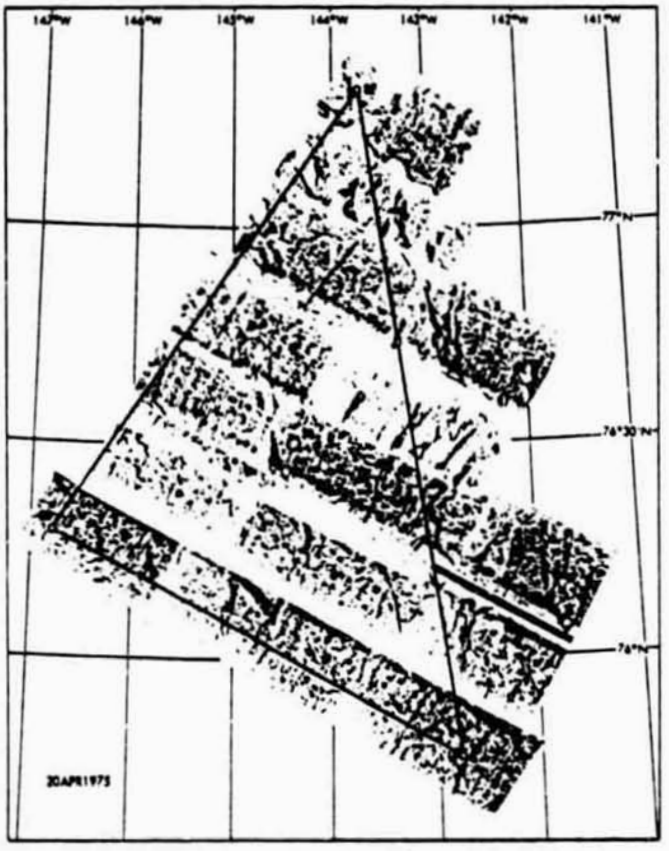


Fig. 16. Mosaic of radar images of the AIDJEX study site for 21 April, 1975, obtained from X-band HH polarized real aperture radar.

The azimuth, or along-track, resolution is determined by the along-track beamwidth of the antenna. The longest possible antenna is carried under the aircraft to obtain the narrowest possible along-track beamwidth. As in conventional radars, the range resolution, or cross-track resolution, is determined by the modulation (bandwidth) applied to the transmitted energy (usually narrow pulses). Although it is relatively easy to achieve high cross-track resolution, the transmitter modulation is usually chosen so that this resolution approximates the more limited along-track resolution, which degrades linearly with increasing range from the transmitter. For the mosaic obtained on 21 April, 1975, the swath was set at 25 km to achieve optimum information content in the imagery. The aircraft flew at an altitude of about 900 m. In contrast to the SAR system, the imagery is processed onboard the aircraft and no geometrical correction is necessary. As with the SAR, the average brightness varies across the image. The wave-like pattern visible in certain parts of the imagery is due to the motion of the aircraft.

The Ryan scatterometer, owned by the Canadian Centre for Remote Sensing (CCRS), measured quantitatively the back-scattering of microwave radiation over a range of incidence angles from approximately 5° to 60° from nadir. The system operates at 13.3 GHz (2.25-cm wavelength in free space) and is similar to the NASA 13.3-GHz scatterometer which has been in other studies of sea ice. Either horizontally or vertically polarized continuous-wave, CW, signals can be transmitted with the Ryan instrument. Both the like- and cross-polarized returns are received simultaneously. This is in contrast to the NASA scatterometer which transmits and receives only vertically polarized radiation. The slotted-wavelength-array antennas radiate a fan beam in the fore-aft direction with an azimuthal beamwidth of 120° and a range beamwidth of 3°. The two transmitting and receving antennas have approximately the same gain patterns. Further quantitative details of the data reduction are given by Royer (1975) and Bradley (1971).

4.2. Morphological observations of sea ice using L-band synthetic aperture radar

Figure 11 reveals several features. Dark lines, such as that at approximately 76°15′ N, 144°30′ W, are generally indicative of open water or recently refrozen leads. Because both of these surface types are generally smooth, the radar imagery cannot distinguish between them. However, as the ice thickens in a given polynya or lead, rafting and ridging are noted as bright linears in the darker (smoother) feature. Thus, older leads, such as the one at 76°10′ N, 145°00′ W, are rather easily identified. In these images, thin ice can be distinguished from open water by reference to the passive microwave data which, as presented in Figure 7, show numerous large leads and polynyas as having brightness temperatures of approximately 230–240 K, indicating that they are indeed ice-covered.

These radar data are especially sensitive to surface roughness, as indicated by the easy identification of the small variations in roughness in some of the refrozen leads. Even in the April 1975 data (Figure 11) which were collected prior to the inclusion of the STC (Sensitivity Time Control – an automatic gain system), bright patterns indicating greater roughness produced by ridging, rafting, and areas of highly broken and disturbed ice are noted.

The identification of various ages of ice (e.g., FY and MY ice) using only the single wavelength and single polarization of SAR data is especially difficult – and if one is restricted to using only the signature of the ice itself (i.e., the strength of the backscatter return), the identification may be impossible. This difficulty is compounded by the use of SAR data which have been placed on photographic paper. In transforming the data to such a format, a considerable amount of information has been lost as a result of the low dynamics range (approximately 20 db) of the photographic paper when compared to the higher dynamic range (approximately 60 db) of the SAR system. This larger dynamic range is only available either through the use of digitized SAR data or by studying the data while they are still in the optical processor. Such was not the case for our study.

Although it is not possible to identify the MY fraction within any given area from this image (Figure 11), it may be possible to distinguish MY floes by consideration of both the shape and texture of the floe in addition to its signature. Hence, rounded areas having dark-grey tones and having bright circular borders are most likely MY floes, and reference to the ESMR data presented earlier confirms this basic interpretation.

Figure 12, an L-HH SAR mosaic for 22 August, 1975, illustrates again several basic notions concerning the interpretation of SAR data of sea ice. Again, we note that we cannot definitely identify areas of open water from those of thin ice. However, when one considers the season, it seems reasonable to consider the large polynya at approximately 75°45′ N, 140°30′ W as being open water, or at best, to have a thin covering of frazil ice. This former interpretation of open water is confirmed by comparison with the passive microwave data presented in Figure 8. Again we note that the large floes which are discernable in the SAR data are not easily identified within the ESMR imagery. However, the identification of MY and FY ice may, from the radiometric point of view and for this season, be a rather academic point to pursue. This is because of the influence of the moisture in the snow on the radiometric signal. Radar suffers from a similar fate, but again the use of shape and texture in addition to the strength of the return helps to distinguish the major ice types.

Figures 13 and 14 present two SAR mosaics collected during October. These data, collected only two weeks apart, are most instructive in illustrating the usefulness of SAR imagery for identifying short-term changes in the ice surface morphology. Immediately one notes the change in the orientation of the dark features (which represent open leads or recently refrozen leads) from an approximate north–south orientation (Figure 13, 10 October, 1975) to having a general east–west orientation (Figure 14, 26 October, 1975). Also, as in August, large floes are readily distinguishable on both SAR images, and individual floes (as the two located at approximately 73°40′ N, 142°00′ W on Figure 14) are noted to have moved approximately 1° west and 0.5° north in two weeks (compared to Figure 13). In addition, several of the MY floes seen earlier on the August images for both the ESMR and the SAR (Figures 8 and 12) are identified, indicating that no significant exchange of ice into or out of the AIDJEX study triangle has taken place. In the October images, we observe that it is considerably easier to make basic ice-type identifications (using shape, texture, and tone in addition to the radar signature) than it is during the August seasons because of the changes which have taken place on the ice surface as a result of the summer processes and because there has been a considerable reduction in the amount of moisture on the ice surface (liquid water?).

Finally, by April of the following year (Figure 15, 16 April, 1976) the morphology of the ice has been altered considerably. It is very difficult to estimate the fraction of FY ice in this image. The ice consists primarily of large MY floes and thin, fractured new ice coupled with many areas of brash and extremely rough ice surfaces resulting from winter pressures.

It can be seen from both the available imagery and this brief review that SAR data can be used to distinguish detailed changes in the morphology of sea ice. This results partly from the high resolution of the SAR and partly from the nature of the interactions of the radar signal and the ice surface. In earlier sections of the paper, passive microwave remote-sensing systems have been shown to be quite useful for distinguishing ice types. Consequently, we believe that these two systems are complementary, and for optimal results in the interpretation of remotely sensed data, they should be studied in concert, as already shown by Campbell *et al.* (1975a). A more detailed analysis of image interpretation of SAR L-band data is given in Bryan (1976).

Finally, it should be noted that the SAR data, when used as a set of sequential data, are useful for detecting the drift of the ice within the aforementioned navigational accuracies. The accuracy with which the sea-ice drift can be measured from airborne or satellite images can be evaluated from images of a non-moving well-mapped surface. This approach was demonstrated by Hibler *et al.* (1975), for Landsat images. Thus, initial studies using 20 ground control points on the 1:250 000 topographic map of Alaska to determine navigational accuracies and the algorithm originally developed for radar mapping of the Moon (Leberl, 1976) yielded an apparent drift, and thus supplied an estimate of the error measuring absolute drift. Detailed studies of this type are discussed in Leberl *et al.* (1976b). The set of SAR data from 22–27 August, 1975 indicate a westward drift of 6.5 km day^{-1} and a northward drift of 2.9 km day^{-1} from sections within the AIDJEX study site. Absolute errors for this drift measurement were estimated to be approximately 200 m day^{-1} - or about 5%.

4.3. Side-looking airborne-radar (SLAR) observations

Due to greater backscatter at the shorter wavelength of the AN/APS94D SLAR compared with the JPL SLAR, more surface details are visible in Figure 16. The image was taken during the same 24-h period as Figure 11. Big Bear can be identified as the white spot within the circle identifying this area. The dark linear features represent refrozen leads. All leads are covered with varying thicknesses of grey ice. Big Bear was located on a small MY ice floe having an average thickness of 3.5 m. The light lines within this floe represent ridges having a height of approximately 2 m and a width of 3 m. The large lead which runs from northwest to southeast consisted of grey ice with a thickness of 0.7 m. In the original image, just north of and adjacent to the camp, the dark linear feature running east–west represents the runway. From surface measurements it is known that it consists of 1.6 m thick FY ice. One striking aspect of this mosaic is the backscattering coefficient σ. For example, the radar returns, as represented by the grey scale, for the MY ice on which Big Bear is located and the large lead to the west of the camp, are very much alike. If it were not for the geometry of the ice, one could not distinguish FY from MY ice in this image. Comparison of the SLAR mosaic (Figure 17) with the ESMR mosaic (Figure 7) shows that a one-to-one correlation exists for

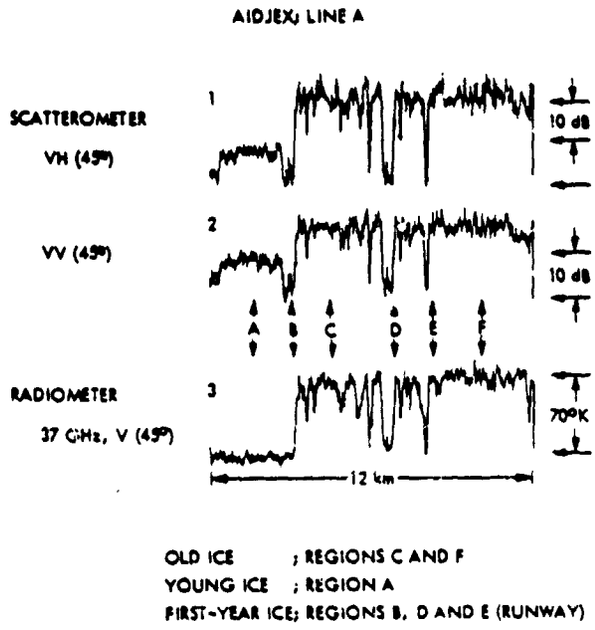


Fig. 17. Scatterometer (2.25 cm) and passive microwave radiometer (0.81 cm) observations along line A at Big Bear, AIDJEX, April, 1975. Trace 1 is cross-polarized (VH) back-scattering observed at 45° fore of nadir, and trace 2 is like-polarized (VV). Trace 3 is variation in passive microwave radiometer brightness temperature, also at 45° fore of nadir.

the major lead systems; however, the SLAR also provides additional fine-structure information. Where the ESMR image indicates a concentration of MY ice, the SLAR image indicates very rough areas with no open leads or thin ice present. This is one example of the complementary nature of the imagery obtained from the active and passive systems.

4.4. Scatterometer observations

Preliminary analog and digital processing of the scatterometer data taken over the AIDJEX site has shown that there is significant contrast in the microwave power backscattered from MY and FY ice, particularly for the cross-polarized components $\sigma\, VH(\theta)$. Figure 17 includes three analog traces of results obtained from flight-line A at the AIDJEX site. The top two traces illustrate the relative variation in the backscatter at 45° fore of nadir, while the lower trace illustrates, qualitatively, the variation in microwave brightness temperature as measured simultaneously by the Aerojet 0.81-cm radiometer. The ground area observed by the microwave radiometer corresponds approximately to the area measured by the scatterometer in traces one and two (see Table I). The plot of the horizontally polarized radiometer has been drawn to emphasize the correspondence between the results obtained by the passive radiometer operating at 0.81 cm and the active scatterometer operating at 2.25 cm. Regions *A* and *B* in the results illustrated in Figure

Fig. 18. Photomosaic of the AIDJEX area on 18 August, 1975.

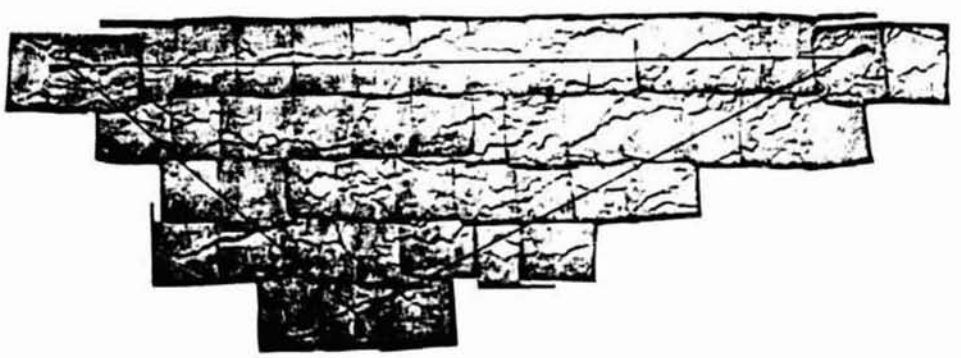


Fig. 19. Photomosaic of the AIDJEX area on 26 August, 1975.

17 correspond in time to observations made over two regions of ice in a refrozen polynya. Region *A* represents an area of grey ice which was thinner than Region *B* and which shows significantly higher, well detailed thermal infrared emission as seen in the line-scanner imagery. The snow-covered FY ice in Region *B* was thicker than at *A* and the infrared imagery showed negligible contrast between the FY ice at *B* and the MY ice beginning after Region *B*. The radiometer appears to be insensitive to the different thicknesses of sea ice in the refrozen polynya. It is interesting to note that the lower frequency scatterometer indicates a higher backscatter at the 45° to nadir over Region *B* than over the thinner grey ice in the middle of the lead.

In the regions of predominantly MY ice (*C* and *F* in Figure 17), the backscatter increases significantly in relation to that observed in the regions of relatively thick

FY ice, at *B*, *D*, or *E*. The difference in backscatter at 45° between *C* and *F*, and *B*, *D*. and *E*, is approximately 20 db for the cross-polarized σ° in trace one, and to approximately 13 db for the like-polarized results σ°_{VV} in trace two. The microwave brightness temperature is approximately 70 K lower than that observed for FY ice for the regions of predominantly ice, as seen in trace three. The first-year ice features, labeled *E* in Figure 17, corresponds to the intersection of the flight line with the FY ice used as the runway of the Big Bear site. The feature labeled *D* corresponds to snow-covered FY ice in a region of predominantly multiyear ice. Identification of this region as being FY ice would have been very difficult, if not impossible, on the basis of the visible or thermal infrared imagery alone.

Further analysis of these data, particularly the SAR and SLAR signatures, is being pursued actively. The recent dielectric results for FY ice obtained by Vant (1976) will allow improved modelling of the ice and perhaps lead eventually to automatic processing of active microwave data for ice reconnaissance. As shown in Table I, the active systems employed varied greatly in frequency, inclination angle, and data processing. It would be premature at this time to make detailed comparisons with other published results such as the one on radar observations by Anderson (1966) and Johnson and Farmer (1971) and on the scatterometer by Parashar *et al.* (1974).

5. Conclusion

We have succeeded for the first time in remotely sensing a chosen large ($\sim 10^4$ km^2) area of Arctic Basin sea ice during all seasons of one year as it drifted, deformed, and metamorphosed. This was done with a wide variety of sensors, each performing at the highest state-of-the art in its class. The ground-truth experiment entailed the most extensive measurements of sea-ice dielectric properties ever made. These and the detailed surface microwave observations were essential for the interpretation of the numerous aircraft data.

We feel that the data ensemble presented in this paper makes a convincing case that in the remote sensing of sea ice, combined passive and active microwave techniques offer the most promise. Bearing in mind that various aspects of sea-ice research need imagery at various resolutions, from meters to tens of kilometers, and that a means of observing sea ice through the clouds and in the dark is essential, microwave remote sensing offers a whole new vista of exciting possibilities. Passive microwave observations allow us to identify and delineate the distribution of ice types and mixtures and the measurement of ice concentration. Active techniques allow the structural aspects of sea ice to be observed at resolutions finer than by any other sensors. Both will sense ice anywhere under any atmospheric or oceanographic conditions. The forthcoming active and passive microwave sensors on Seasat-A and Nimbus-G will give us a greatly expanded view of the behavior of sea ice on Spaceship Earth. We will go forward not only with a clearer understanding but also with a sense of excitement.

Acknowledgement

This complex remote sensing experiment could never have been performed without the support of a number of agencies in the U.S. and Canada. NASA's generous contribution of the CV 990 with its excellent crew under the direction of Earl V. Peterson was invaluable. The very active logistical and financial support given by the Polar Continental Shelf Project and the AIDJEX Project made our dream of an experiment a reality. Additional excellent remote sensing aircraft support was given by the Canadian Airborne Sensing Unit (C-47) and Maritime Proving and Evaluation Unit (Argus). The difficult reduction of the surface microwave data was done by B. Hache of Environment Canada.

Dedication

We dedicate this work to our beloved friend William Nordberg. He pioneered in microwave ice observations, and his brilliance and enthusiasm inspired all of us.

References

Addison, J. R.: 1969, 'Electrical Properties of Saline Ice', *J. Appl. Phys.* **40**, 3105–3114.

Anderson, V. H.: 1966, 'High Altitude Side-Looking Radar Images of Sea Ice in the Arctic', Proceedings of the Fourth Symposium on Remote Sensing of Environment, Environmental Research Institute, Ann Arbor, Michigan, pp. 845–857.

Bogorodsky, V. V. and Tripol'nikov, V. P.: 1973, 'Electromagnetic Characteristics of Sea Ice in the Range of 30–400 MHz', *Dokl. Akad. Nauk S.S.S R.* **213**, 577–579.

Bogorodsky, V. V. and Khokhlov, G. P.: 1975, 'Electrical Properties of Ice in the Ice Edge Zone of the Bering Sea at 10 GHz Frequency', Proceedings of the USSR/USA Bering Sea Experiment, NASA Goddard, pp. 219–233.

Bradley, G. A.: 1971, 'Remote Sensing of Ocean Winds Using a Radar Scatterometer', University of Kansas, CRES Technical Report No. 177-22, 201 pp.

Bryan, M. L.: 1976, 'Interpretation Key for SAR (L-Band) Imagery of Sea Ice', Proceedings of the American Society of Photogrammetry, 105 N. Virginia Avenue, Falls Church, Va. 22046, Fall Convention, Seattle, Wash., pp. 406–435.

Campbell, W. J., Gloersen, P., Nordberg, W., and Wilheit, T. T.: 1974, 'Dynamics and Morphology of Beaufort Sea Ice Determined from Satellites, Aircraft, and Drifting Stations', COSPAR Approaches to Earth Survey Problems through Use of Space Techniques, P. Bock (ed.), Akademie Verlag Berlin, pp. 311–327.

Campbell, W. J., Ramseier, R. O., Weeks, W. F., and Gloersen, P.: 1975a, 'An Integrated Approach to the Remote Sensing of Floating Ice', Proceedings of the Third Canadian Symposium on Remote Sensing, Edmonton, pp. 39–72.

Campbell, W. J., Weeks, W. F., Ramseier, R. O., and Gloersen, P.: 1975b, 'Geophysical Studies of Floating Ice by Remote Sensing', *J. Glaciol.* **15**, 305–328.

Campbell, W. J., Gloersen, P., Ramseier, R. O., Webster, W. J., and Wilheit, T. T.: 1976, 'Beaufort Sea Ice Zones as Delineated by Microwave Imagery', *J. Geophys. Res.* **81**, 1103–1110.

Chang, T. C. and Gloersen, P.: 1975, 'Microwave Emission from Dry and Wet Snow', Proceedings of the Workshop on Operational Applications of Satellite Snowcover Observations, A. Rango (ed.), NASA SP-391.

Chang, T. C., Gloersen, P., Schmugge, T., Wilheit, T. T., and Zwally, H. J., 1976, 'Microwave Emission from Snow and Glacier Ice', *J. Glaciol.* **17**, Paper F-1.

Cook, J. C.: 1960, 'RF Electrical Properties of Salty Ice and Frozen Earth', *J. Geophys. Res.* **65**, 1767–1771.

Edgerton, A. T., Stogryn, A., and Poe, G.: 1971, 'Microwave Radiometric Investigations of Snow Pack'; Final Report No. 1285R-4 for USGS Contract No. 1A-08-001-11828, Aerojet General Corporation Microwave Division, El Monte, Calif.

Gloersen, P., Nordberg, W., Schmugge, T., Wilheit, T. T., and Campbell, W. J.: 1973, 'Microwave Signatures of First Year and Multiyear Sea Ice', *J. Geophys. Res.* **78**, 3564–3572.

Gloersen, P., Chang, T. C., Wilheit, T. T., and Campbell, W. J.: 1974, 'Polar Sea Ice Observations by Means of Microwave Radiometry', *Advanced Concepts and Techniques in the Study of Snow and Ice Resources*, H. S. Santeford (ed.), NAS Washington, D.C. (Lib. Cong. 74-10120; ISBN 0–309–02235–5), pp. 541–550.

Gloersen, P. and Salomonson, V. V.: 1975, 'Satellites – New Global Observing Techniques for Ice and Snow', *J. Glaciol.* **15**, 373–389.

Gloersen, P., Ramseier, R. O., Campbell, W. J., Chang, T. C., and Wilheit, T. T.: 1975, 'Variation of Ice Morphology of Selected Mesoscale Test Areas during the Bering Sea Experiment', Proceedings of the Symposium of the Joint U.S./U.S.S.R. Bering Sea Experiment, Gidrometeoizdat Leningrad, pp. 196–218.

Gloersen, P., Zwally, H. J., Chang, A. T. C., Hall, D. K. Campbell, W. J., and Ramseier, R. O.: 1978, 'Time-Dependence of Sea Ice Concentration and Multilayer Ice Fraction in the Arctic Basin', this volume, pp. 339–359.

Goodman, J. W.: 1968, *Introduction to Fourier Optics*, McGraw-Hill, San Francisco.

Harger, R. O.: 1970, *Synthetic Aperature Radar Systems*, Academic Press, New York.

Hibler, W. I. III, Tucker, W. B., and Weeks, W. F.: 1975, 'Techniques for Studying Sea Ice Drift and Deformation at Sites Far from Land using Landsat Imagery', Proceedings of the 10th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, Michigan.

Hoekstra, P. and Capillino, P.: 1971, 'Dielectric Properties of Sea and Sodium Chloride Ice at UHF and Microwave Frequencies', *J. Geophys. Res.* **76**, 4922–4931.

Johnson, J. D. and Farmer, L. D.: 1971, 'Use of Side-Looking Airborne Radar for Sea Ice Identification', *J. Geophys. Res.* **76**, 2138–2155.

Leberl, F.: 1976, Mapping of Lunar Surface from Side-Looking Orbital Radar Images', *The Moon* **15**, 329–342.

Leberl, F., Jensen, H., and Kirphan, J.: 1976a, 'Side-Looking radar Mosaicing Experiment', *Photogrammetric Engineering and Remote Sensing* **42**, 1035–1042.

Leberl, F., Farr, T., Bryan, M. L., and Elachi, C.: 1976b, 'Study of Arctic Sea Ice Drift from L-Band Synthetic Aperature Radar', Proceedings of the American Society of Photogrammetry, 42nd Annual Convention, Washington, D C., pp. 597–611.

Meeks, D. C., Ramseier, R. O., and Campbell, W. J.: 1974a, 'A Study of Microwave Emission Properties of Sea Ice – AIDJEX 1972', Proceedings of the 9th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, Mich., pp 307–322.

Meeks, D C., Poe, G. A., and Ramseier, R. O.: 1974b, 'A Study of the Microwave Properties of Sea Ice – AIDJEX 1972', University of Washington, AIDJEX Program Office, Seattle, Wash., Final Report No 1786FR-1.

Parashar, S. K., Biggs, A. W., Fung, A. K., and Moore, R. K., 1974, 'Investigation of Radar Discrimination of Sea Ice', Proceedings of the 9th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, Mich., pp. 323–332.

Ramseier, R. O., Gloersen, P., Campbell, W J., and Chang, T. C.: 1975, 'Mesoscale Description for the Principal Bering Sea Ice Experiment', Proceedings of the USSR/USA Bering Sea Experiment, NASA Goddard, pp. 234–269.

Rihaczek, A. W.: 1969, *Principles of High Resolution Radar*, McGraw-Hill, New York.

Royer, G. M., 1975, 'Computer Analysis of CW Scatterometer Data with Particular Reference to Ryan Model 720 Systems', Department of Communications, Ottawa, CRC Report No. 1265, 65 pp.

Sackinger, W and Byrd, R., 1972, 'The Dielectric Properties of Sea Ice in the Range from 26–40 GHz', IAEE Report No. 7203, University of Alaska, Fairbanks, Alaska, 88 pp.

Thompson, T W., Bishop, R. S., and Brown, W E.: 1972, 'Progress Report on 25 cm Radar Observations of the 1971 AIDJEX Studies', AIDJEX Bulletin No. 12, pp 1–15

Van Roessel, J. and Degodoy, R. C.: 1974, 'SLAR Mosaics from Project Radam', *Photogrammetric Engineering* **40**, 583–595.

Vant, M. R.: 1976, 'A Combined Empirical and Theoretical Study of the Dielectric Properties of Sea Ice over the Frequency Range 100 MHz to 40 GHz', Environment Canada, Ocean and Aquatic Sciences/Carleton University, Ottawa, Technical Report, Ph.D. Thesis, Carleton University, Ottawa, 437 pp.

Vant, M. R., Gray, R. B., Ramseier, R. O., and Makios, V.: 1974, 'Dielectric Properties of Fresh and Sea Ice at 10 and 35 GHz', *J. Appl. Phys.* **45**, 4712–4717

Weeks, W. F. and Campbell, W. J.: 1975, 'Remote Sensing Plan for the AIDJEX Main Experiment', *AIDJEX Bull.* **29**, 21–48.

Wentworth, F. L. and Cohn, M.: 1964, 'Electrical Properties of Sea Ice at 0.1 to 30 Mc/s', *Radio Sci. J.* **68D**, 681–691.

Wilheit, T. T., Nordberg, W., Blinn, J., Campbell, W. J., and Edgerton, A.: 1972, 'Aircraft Measurements of Microwave Emissions from Arctic Sea Ice', *Remote Sensing of Environment* **2**, 129–139.

Paper 65

NASA TP-1251 June 1978

Estimation of Snow Temperature and Mean Crystal Radius from Remote Multispectral Passive Microwave Measurements

A. T. C. Chang

Abstract

Recent work by A. T. C. Chang et al. has demonstrated that variation in crystal size and physical temperature of snowfield observations from space give large variations in the microwave brightness temperature. Since the brightness temperature is a function of wavelength, the microwave brightness temperature can be used to extract the snow temperature and mean crystal radius profiles. The Scanning Multichannel Microwave Radiometer (SMMR), to be launched on board the Nimbus-G and Seasat-A spacecraft, will make observations in wavelengths of 0.8, 1.4, 1.7, 2.8, and 4.6 cm.

A statistical retrieval method has been developed in this paper to determine the snowfield temperature profile and mean crystal size by using the scanning multifrequency microwave radiometer on board a spacecraft. The estimated errors for retrieval are approximately 1.5 K for temperature and 0.001 cm for crystal radius in the presence of 1 K rms noise for each SMMR channel.

NASA TP-1212 April 1978

Microwave Emission From Polar Firn

A. T. C. Chang and B. J. Choudhury

Abstract

The microwave emission from a half-space medium characterized by coordinate dependent scattering and absorbing centers has been calculated by numerically solving the radiative transfer equation by the method of invariant imbedding. Rayleigh scattering phase functions and scattering induced polarization of the radiation have been included in the calculation. For uniform temperature distribution within the medium, the results of A. T. C. Chang and B. J. Choudhury calculations agree well with those by A. W. England. Using the scattering and extinction data of polar firn as compiled by H. J. Zwally, the brightness temperature has been calculated for the 1.55 cm wavelength. For the North Polar region, the results of our calculations are in agreement with the Nimbus-5 scanning radiometer experiment. This study is the first quantitative comparison of the results of numerical calculation using the actual measured information of crystal size with the observed data.

Paper 67

NASA Technical Memorandum 79662 October 1978

Microwave Brightness of Polar Firn As Measured by Nimbus 5 and 6 ESMR

A. T. C. Chang

B. J. Choudhury

P. Gloersen

Abstract

The microwave emission from a half-space medium characterized by coordinate dependent scattering and absorbing centers has been calculated by numerically solving the radiative transfer equation by the method of invariant imbedding. A Mie scattering phase function and surface polarization have been included in the calculation. Also included are the physical temperature profile and the temperature variation of the index of refraction for ice. Using published values of grain size and temperature profile data of polar firn, the brightness temperature has been calculated for the 1.55 cm and 0.8 cm wavelengths. For selected regions in Greenland and Antarctica, the results of our calculations are in reasonable agreement with the observed Nimbus-5 and Nimbus-6 ESMR data.

Paper 68

NASA Technical Memorandum 79671 November 1978

STUDIES OF SNOWPACK PROPERTIES BY PASSIVE MICROWAVE RADIOMETRY

A. T. C. Chang, D. K. Hall, J. L. Foster,
A. Rango and T. J. Schmugge

ABSTRACT

Research involving the microwave characteristics of snow was undertaken in order to expand the information content currently available from remote sensing, namely the measurement of snowcovered area. Microwave radiation emitted from beneath the snow surface can be sensed and thus permits information on internal snowpack properties to be inferred. The intensity of radiation received is a function of the average temperature and emissivity of the snow layers and is commonly referred to as the brightness temperature (T_B). The T_B varies with snow grain and crystal sizes, liquid water content and snowpack temperature.

The T_B of the 0.8 cm wavelength channel was found to decrease moreso with increasing snow depth than the 1.4 cm channel. More scattering of the shorter wavelength radiation occurs thus resulting in a lower T_B for shorter wavelengths in a dry snowpack. The longer 21.0 cm wavelength was used to assess the condition of the underlying ground. Ultimately it may be possible to estimate snow volume over large areas using calibrated brightness temperatures and consequently improve snowmelt runoff predictions.

Paper 69

NASA Technical Memorandum 80267 May 1979

A COMPARATIVE STUDY OF MICROWAVE RADIOMETER OBSERVATIONS OVER SNOWFIELDS WITH RADIATIVE TRANSFER MODEL CALCULATIONS

A. T. C. Chang and J. C. Shiue

Laboratory for Atmospheric Sciences (GLAS)
NASA, Goddard Space Flight Center

ABSTRACT

The increase demand for water throughout the world imposes a challenging problem for water resources managers. Since the water stored in snowpacks represents an important source of the water supply, a system which monitors and inventories the water equivalent of snow should clearly improve the water resources estimate.

Recent results indicate that spaceborne microwave radiometry has the potential for inferring the snow depth and water equivalent information over large snow-covered areas. In order to assess this potential for determining the water equivalent of a snowpack, it is necessary to understand the microwave emission and scatter behavior of the snow at various wavelengths. The emitted microwave radiation is dependent on the physical temperature, crystal size, and density. The basic relationship between the properties of the snowpack and the emitted radiation can be derived by using the radiative transfer approach.

Truck-mounted microwave instrumentation was used to study the microwave characteristics of the snowpack in the Colorado Rocky Mountain region in the vicinity of Fraser, Colorado during the winter of 1978. The spectral signatures of 5.0, 10.7, 18, and 37 GHz radiometers with dual polarization were used. These data compared favorably with calculated results based on recent microscopic scattering models.

Paper 70

NASA Technical Paper 1408

PRELIMINARY RESULTS OF PASSIVE MICROWAVE SNOW EXPERIMENT DURING FEBRUARY AND MARCH 1978

A. T. C. Chang and J. C. Shiue
Goddard Space Flight Center
Greenbelt, Maryland

and

H. Boyne, D. Ellerbruch, G. Counas,
R. Wittmann, and R. Jones
National Bureau of Standards
Boulder, Colorado

1979

ABSTRACT

The purpose of this experiment was to determine if remote microwave sensing of snowpack data could be used to predict runoff, thereby allowing more efficient management of the water supply. A four-frequency microwave radiometer system was attached to a truck-mounted aerial lift and was used to gather data on snowpacks at three different sites in the Colorado Rocky Mountains. Ground truth data measurements (density, temperature, grain size, hardness, and free-liquid water content) were taken at each site corresponding to each microwave scan. Although the detailed analysis of these data is not yet complete, understanding of microwave sensing has been enhanced considerably.

Radio Science, Volume 14, Number 5, pages 793–802, September–October 1979

Remote sensing of atmospheric water vapor, liquid water, and wind speed at the ocean surface by passive microwave techniques from the Nimbus 5 satellite

A. T. C. Chang and T. T. Wilheit

NASA / Goddard Space Flight Center, Greenbelt, Maryland 20771

(Received October 18, 1978.)

The microwave brightness temperature measurements for Nimbus 5 electrically scanned microwave radiometer (ESMR) and Nimbus-E microwave spectrometer (NEMS) are used to retrieve the atmospheric water vapor, liquid water, and wind speed by a quasi-statistical retrieval technique. It is shown that the brightness temperature can be utilized to yield these parameters under various weather conditions. Observations at 19.35, 22.235, and 31.4 GHz were input to the regression equations. The retrieved values of these parameters for portions of two Nimbus 5 orbits are presented. Then comparison between the retrieved parameters and the available observations on the total water vapor content and the surface wind speed are made. The estimated errors for retrieval are approximately ±0.15 g/cm^2 for water vapor content, ±6.5 mg/cm^2 for liquid water content, and ±6.6 m/s for surface wind speed.

1. INTRODUCTION

In this paper we will discuss techniques for measuring the water vapor and liquid water content of marine atmospheres and the wind speed at the ocean's surface. Using microwave remote sensing, the importance of water vapor and wind speed measurements are well known and obvious. However, the liquid water content of the atmosphere has not been routinely measured, and thus its relative importance is not firmly established. It is, nevertheless, available as a by-product of the remote sensing approach discussed in this paper and may well prove useful.

This remote sensing approach involves the measurement of emitted microwave radiation at several wavelengths and sorting out the various contributions to the measured brightness temperature. In a recent paper by *Wilheit and Fowler* [1977], a linear approximation method is used to analyze the wind speed and cloud liquid water content from a set of microwave measurements collected by an aircraft platform. Their data set was taken under a rather narrow range of subarctic conditions which permitted use of a simpler interpretation technique than is used here. In order to describe the techniques involved, we will first review the relevant aspects of microwave radiative transfer. We will then develop a statistical technique for retrieval of geophysical parameters from the measurements. Using this statistical technique, a data set involving observations over a typhoon in the Pacific Ocean will then be analyzed.

The specific sensors applied here are the Nimbus-E microwave spectrometer (NEMS) and the electrically scanned microwave radiometer (ESMR), both of which are in orbit aboard the Nimbus 5 satellite. Detailed descriptions of these instruments can be found in *Wilheit* [1972] and *Staelin et al.* [1972]. For present purposes it is sufficient to note that the ESMR is a scanning instrument which measures the upwelling microwave radiation at 19.35 GHz while scanning 50° to either side of the spacecraft and has a spatial resolution of 25 km at nadir. The NEMS instrument views only at the nadir and obtains observations at 22.235 and 31.4 GHz and at three frequencies in the 50- to 70-GHz molecular oxygen band (not of interest in this paper). It has a spatial resolution of 200 km.

2. MICROWAVE RADIATIVE TRANSFER IN THE ATMOSPHERE

In the microwave portion of the electromagnetic spectrum and at temperatures typical of the earth and its atmosphere the Rayleigh-Jeans approximation for black body radiation works quite well. This enables radiative transfer for an isotropic, nonscattering, nonreflecting medium to be expressed as follows:

This paper is not subject to U.S. copyright. Published in 1979 by the American Geophysical Union.

$$dT_B/dx = -\alpha(T_B - T_0) \quad (1)$$

where T_B represents the intensity of radiation at a given wavelength and in the x direction expressed as an equivalent black body temperature, called brightness temperature. The absorption coefficient is α, and T_0 is the thermodynamic temperature of the medium. If this is integrated for a uniform slab of thickness a along the direction of propagation x, the result becomes

$$T_{B_{out}} = T_{B_{in}} e^{-\alpha a} + (1 - e^{-\alpha a})T_0 \quad (2)$$

where $T_{B_{in}}$ is the brightness temperature incident on the slab and $T_{B_{out}}$ represents the emerging brightness temperature.

If the expression αa is quite small or if $T_0 - T_B$ is small, the slab has a very small effect on the intensity of the radiation; whereas, if the slab is opaque ($\alpha a \to \infty$), the radiation emerging out of the slab is characterized by the thermodynamic temperature of the slab independent of the intensity of the incident radiation. In this study the atmosphere is divided into 100 layers, each 0.2 km thick. Each layer is characterized by the temperature, pressure, and concentration of constituents of the center of the slab. In a manner analogous to (1) the brightness temperature T_B radiating from a surface is given by

$$T_B = (1 - e)T_{B_{inc}} + eT_0 \quad (3)$$

$T_{B_{inc}}$ is the brightness temperature incident on the surface, e is the emissivity of the surface, and T_0 is the thermodynamic temperature of the surface. The incident energy must also be averaged appropriately over all angles which can be reflected into the viewing direction. Since most natural surfaces are quite rough when they are considered on the scale of microwave wavelengths (0.1–30 cm), we will use the Lambertian approximation [*Born and Wolf*, 1975] for this angular distribution:

$$T_{B_{inc}} = 2\int_0^{\pi/2} T_B(\theta)\cos\theta\sin\theta\, d\theta \quad (4)$$

where $T_B(\theta)$ is the incident brightness as a function of the incidence angle θ. For computational simplicity, since the weighting function ($\sin\theta\cos\theta$) within the integrand has a maximum at $\theta = 45°$, we will approximate $T_{B_{inc}}$ by

$$T_{B_{inc}} \cong T_B \qquad \theta = 45° \quad (5)$$

Thus to determine this brightness temperature which would be observed in a particular situation, we must be able to characterize the emissivity of the underlying surface and the absorption coefficient of the atmosphere as a function of their physical and chemical parameters.

The emissivity of terrestrial surfaces in the microwave region of the electromagnetic spectrum varies from 0.95 for most solid surfaces to about 0.4 for the ocean surface. It is this low emissivity and relative homogeneity of the calm ocean surface which provides the possibility of sounding the water vapor and liquid water content of the atmosphere.

The emissivity of a smooth ocean surface can be calculated from the dielectric constant through the Fresnel relations [*Jackson*, 1962]. The dielectric constant data used here are from *Lane and Saxton* [1952]. At a given frequency ν the real part of the dielectric constant ε' may be expressed as

$$\varepsilon' = \varepsilon_\infty + (\varepsilon_s - \varepsilon_\infty)/(1 + \omega^2\tau^2) \quad (6)$$

and the imaginary part ε'' as

$$\varepsilon'' = \omega\tau(\varepsilon_s - \varepsilon_\infty)/(1 + \omega^2\tau^2) + \sigma/\omega \quad (7)$$

where $\omega = 2\pi\nu$. The ε_∞, ε_s, σ, and τ values at a number of temperatures and NaCl concentration are given by *Saxton and Lane* [1952]. Conductivity value as a function of temperature is taken from the *International Critical Tables* [1928]. These can be summarized by the analytical expressions

$$\varepsilon_\infty = 4.9 \quad (8)$$

$$\varepsilon_s = 1\!\!\!\!\diagup 00.0 - 81.0N + 38.0N^2 - (3.75 - 2.1N + N^2)\frac{T}{10.0} \quad (9)$$

$$\tau = \frac{0.00199}{T}e^{2140/T} + \left[\frac{0.00972}{T}e^{2060/T} - \frac{0.00324}{T}e^{1968/T} - \frac{0.00597}{T}e^{2140/T}\right]N + \left[\frac{0.00648}{T}e^{1968/T} - \frac{0.00972}{T}e^{2060/T} + \frac{0.00398}{T}e^{2140/T}\right]N^2 \quad (10)$$

$$\sigma = 92.13N - 8.73N^2 + 3.12(T - 273)N - 0.37(T - 273)N^2 \quad (11)$$

T is the temperature in Kelvin, and N the NaCl concentration in gram normal. The form used for

the temperature dependence of τ was suggested on theoretical grounds by Saxton and appears necessary in light of the large temperature dependence of this particular parameter. The simple polynomial fit appears adequate for the remainder of these expressions. We will approximate sea water as a 0.6N NaCl solution.

When the sea surface is roughened by wind, the situation becomes more complicated. Calculations by *Stogryn* [1967] indicate that there is little effect due to roughness for viewing at nadir. In horizontal polarization from nadir the roughness increases the emissivity, and the effect increases with viewing angle out to at least 60°. In vertical polarization the roughness decreases the emissivity out to about 55° incidence angle, above which the emissivity increases with roughness and viewing angle. The observations of *Hollinger* [1971] only partially confirm the theoretical expectations, but both agree that there is little or no effect due to roughness for viewing directly at nadir, which is the case considered here.

In addition to roughening the sea surface the wind also creates foam which can also increase the emissivity of the surface. *Nordberg et al.* [1971] found that for a frequency of 19.35 GHz the foam coverage increases the emissivity at nadir by 4.3×10^{-3} for each meter per second the wind speed exceeded 7 m/s, with no effect at wind speeds below 7 m/s. Measurements have been made by *Van Melle et al.* [1973] at 2.69 GHz on an artificially generated foam; they found an effect comparable with that of *Nordberg et al.* [1971]. More recently, *Webster et al.* [1976] have measured the net effect of wind on the surface emissivity at several frequencies from 1.4 to 37 GHz at view angles of 0° (nadir) and 38°. They found little frequency dependence above 10 GHz and a somewhat weaker effect than was found by Nordberg et al. On this basis we will in this paper assume that for the frequencies of present interest (19.35, 22.235, and 31.4 GHz) each meter per second that the wind speed exceeds 7 m/s increases the surface emissivity by 3.2×10^{-3}. The determination of the effect of wind speed on surface emissivity as a function of frequencies, polarization, and view angle remains an important research problem.

In the earth's atmosphere at frequencies between 1 and 100 GHz, only three constituents need to be considered: water vapor, liquid water, and molecular oxygen. All other constituents normally found in the atmosphere have either a sufficiently low abundance or weak microwave absorption that they are ignorable. We will therefore consider only these three constituents. We will restrict this treatment also to nonraining situations to simplify the treatment of liquid water in atmosphere.

Water vapor has transitions at 22.235 and 183.3 GHz in addition to many transitions at higher frequencies. By using the treatment given by *Gaut* [1968] the absorption coefficient can be expressed as

$$\alpha = \alpha_{RES} + \alpha_{NON} \tag{12}$$

where α_{RES} refers to the contribution due to the 22.235-GHz absorption line and can be written as

$$\alpha_{RES} = \frac{343\nu^2 \Delta\nu\rho}{T^{5/2}} e^{-644/T} \left[\frac{1}{(\nu - 22.235)^2 + \Delta\nu^2} + \frac{1}{(\nu + 22.235)^2 + \Delta\nu^2} \right] \tag{13}$$

where

$$\Delta\nu = \frac{0.12\,6P(1 + 0.011(\rho T/P))}{T}$$

is the line width of the 22.235-GHz absorption line,

- ρ water vapor density, in grams per cubic meter;
- P pressure, in torrs;
- T temperature, in Kelvins;
- ν propagating frequency, in GHz.

the contribution from all other lines, α_{NON}, can be written as

$$\alpha_{NON} = 2.55 \times 10^{-3} \frac{\nu^2 \rho \Delta\nu}{T^{3/2}} \tag{14}$$

For liquid water, as already noted, only nonprecipitating clouds are considered in this study. Since in such clouds the typical particle size is much smaller than the wavelength of interest, the Rayleigh approximation [*Gunn and East*, 1954] should be quite adequate to represent the scattering by cloud particles less than 100 micron. The absorption coefficient can be written as

$$\gamma_{cloud}(\mathrm{km}^{-1}) = \frac{0.188\, m\varepsilon''}{[(\varepsilon' + 2)^2 + (\varepsilon'')^2]\,\lambda} \tag{15}$$

where m is the cloud liquid water content of atmosphere in grams per cubic meter.

The dielectric data (ε' and ε'') used are those

of *Lane and Saxton* [1952] and discussed in the previous paragraph for the special case of pure water. The combination of the factor of ν in this expression of ϵ' and ϵ'' results in a frequency dependence of approximately ν^2 for the absorption coefficient of clouds. Since clouds are often supercooled below the 265 K lower temperature limit of the Lane and Saxton dielectric data on which an expression for ϵ' and ϵ'' are based, it is important that we have chosen an expression for the relaxation time τ which has some theoretical foundation.

The absorption by the complex of molecular oxygen resonances near 0.5 cm can be approximated for frequencies below 45 GHz by treating them as a line with a resonant term and a nonresonant term. The absorption coefficient can be written as

$$\alpha = \frac{0.3P}{T^2}\left\{\frac{\Delta\nu}{(\nu-\nu_0)^2+\Delta\nu^2}+\frac{\Delta\nu}{\nu^2+\Delta\nu^2}\right\} \qquad (16)$$

where $\nu_0 = 60$ GHz and $\Delta\nu$ is the pressure broadening in GHz:

$$\Delta\nu = \begin{cases} 7.91\times10^{-3}P\left(\dfrac{300}{T}\right)^{0.85} & P > 250 \text{ torr} \\ 1.55\times10^{-3}P\left(\dfrac{300}{T}\right)^{0.85}\left(\dfrac{231+0.49(19-P)}{231}\right) & 250 \geq P \geq 19 \text{ torr} \\ 1.55\times10^{-3}P\left(\dfrac{300}{T}\right)^{0.85} & P < 19 \text{ torr} \end{cases} \qquad (17)$$

Calculations based on this approximation agree well with those based on the full *Meeks and Lilly* [1963] expression as modified by *Lenoir* [1968] for frequencies less than 45 GHz.

3 QUASI-STATISTICAL RETRIEVAL TECHNIQUE

As can be seen from the previous section, the brightness temperature which would be observed upwelling from the atmosphere depends on many geophysical variables However, several of the variables (temperature, humidity, and liquid water content) are, in themselves, continuous functions of height. Therefore this problem has an infinite number of degrees of freedom. In the most general sense, a retrieval of geophysical variables from a finite set of observations could never be performed. We must resort to approximations and constrained solutions if we are to derive useful information.

One technique for retrieving geophysical parameters is the statistical technique which was applied, quite successfully, by *Waters et al.* [1975] to the problem of approximating the atmospheric temperature profile from microwave brightness temperature measurements in the molecular oxygen band ($\lambda \simeq 5$ mm). In this technique they begin with a statistical data base that consisted of a large set of actual temperature soundings. For each of these soundings they calculate the expected upwelling brightness temperature at all frequencies at which measurements are to be made. The temperature at various heights are then regressed against the brightness temperature to determine the most probable (if the statistics are Gaussian) temperature from each level.

If there is little information about a given level in the brightness temperature, then the retrieved temperature will tend toward the a priori average of the statistical data set, and the residual of the regression will be little smaller than the variance of the a priori data set. It is to be noted that information about a given level can occur in a given measurement both directly through the equation of radiative transfer and also indirectly through correlations inherent to the atmosphere itself.

This technique may be faulted in that it contains any bias of the a priori data set. When the brightness temperatures are strongly related in a direct sense to the geophysical parameters being retrieved, this would be of little consequence, but where the coupling is through statistical correlations among the geophysical observables or when the retrieval offers only marginal improvement over the a priori statistics, then the retrieved values will tend strongly toward the commonplace situations. This may well give average statistics which suggest a very good retrieval technique but which give poor results in the extraordinary situation which is often of the most interest. We therefore have modified this

technique somewhat. We generate an artificial data set subtending the entire expected range in all the geophysical variables at issue and force any correlations among them to be zero or at least much smaller than is expected in nature. As a byproduct, this relieves us of the necessity of providing a true statistical data base for parameters which have been measured sparsely or not at all.

In the present case we wish to interpret measurements made at frequencies of 19.35, 22.235, and 31.4 GHz. On the basis of the microwave properties of the atmosphere, we would expect the brightness temperature to be affected strongly by the wind speed at the sea surface and the net content of water vapor and liquid water in the atmosphere and rather weakly by sea surface temperature, cloud temperature, and details of the vertical distribution of water vapor. Thus variables such as details of the temperature structure of the atmosphere would be expected to be quite significant. Thus for our statistical data base we select values from handbook atmospheres [*Air Force Cambridge Research Laboratory*, 1965], each with its associated water vapor content as specified as follows:

Handbook Atmosphere	Water Vapor Content, g/cm^2
U.S. Standard	1.60
tropical	3.88
subtropical summer	4.23
subtropical winter	2.12
mid-latitude summer	2.93
mid-latitude winter	0.87
subarctic summer	2.07
subarctic winter	0.42
Arctic winter	0.20

We specify the values of sea surface temperature and surface wind speed as follows:

Sea Surface Temperature Values, K	Surface Wind Speed Values, m/s
273.0	0
283.0	10
293.0	20
303.0	30

and in Table 1 we give the cloud models used. The statistical data base then consists of all possible combinations of these four parameters. This would provide zero correlation among the four parameters. However, we also specify that wherever the cloud liquid water content is not zero, the relative humidity must be 100% at that altitude. This gives a weak correlation between cloud liquid water content and net water vapor and provides some additional variability in the vertical distribution of water vapor. Thus for the present case at each of three frequencies we calculate brightness temperature for every possible combination, within our set, of sea surface wind, sea surface temperature, atmosphere model, and cloud model. There are a total of 1296 data sets generated which contain the following information: sea surface wind, net water vapor content, net liquid water content, T_B (19.35 GHz), T_B (22.235 GHz), and T_B (31.4 GHz). Conventional linear predications can then be generated for each of the first three of these based on the last three, while adding appropriate instrument noise, typically 1K rms, to the brightness temperature from a random number generator. While the linear approach works reasonably well, the dependence of the brightness temperature on the atmospheric water components is not linear. Rather, the optical depth is more nearly linearly related to the water components. For the moment, we treat the atmosphere as a single layer with an optical depth $\tau = \tau_{vapor} + \tau_{liquid}$ and a thermodynamic temperature t_1, while the surface reflectance is R and has a temperature t_2. We will also make the approximation that $t_1 \approx t_2$, and we will ignore the cosmic background. Then the upwelling brightness T_B is given by

TABLE 1. Cloud models

Model	Cloud Top, km	Cloud Bottom, km	Liquid Water Density, g/m^3
1	2	1	0.01
2	2	1	0.2
3	8	7	0.01
4	8	7	0.2
5	6	1	0.01
6	6	1	0.2
7	8	6	0.01
8	8	6	0.2
9	no cloud		

TABLE 2. Regression residuals

	A priori Mean	A priori Standard Deviation	Estimated Error
Water vapor content, g/cm^2	2.35	1.64	0.15
Liquid water content, g/cm^2	0.019	0.030	0.0065
Wind speed, m/s	15.	11.2	6.6

$$T_B = (1 - Re^{-(1+\sqrt{2})\tau})t_1 \quad (18)$$

The $\sqrt{2}$ occurs because of the approximation to a Lambertian surface discussed earlier. The logarithmic form of (18) can be written as

$$\ln(t_1 - T_B) = \ln(Rt_1) - (1 + \sqrt{2})\tau_{vapor} + (1 + \sqrt{2})\tau_{liquid} \quad (19)$$

We then note that the function $\ln(t_1 - T_B)$ is linearly related to each of the components of atmo-

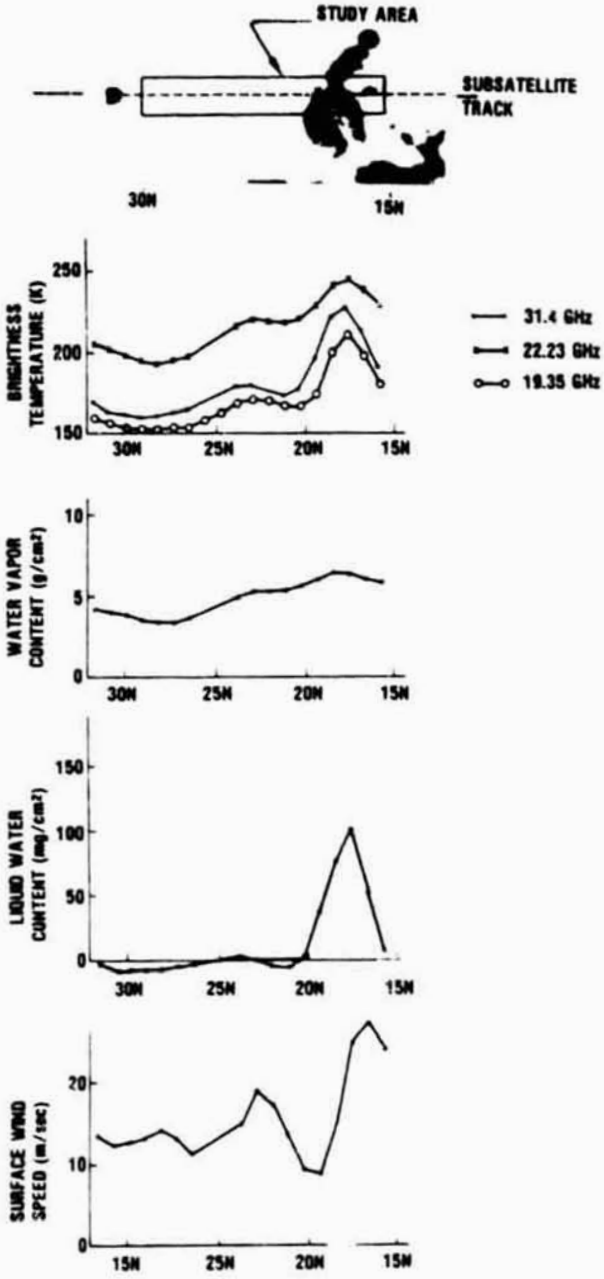


Fig. 1. The measured brightness temperatures and the ESMR image of Nimbus 5 orbit 2891, July 14, 1973. The estimates of the water vapor content, the cloud liquid water content, and the surface wind speed along the subsatellite track are presented.

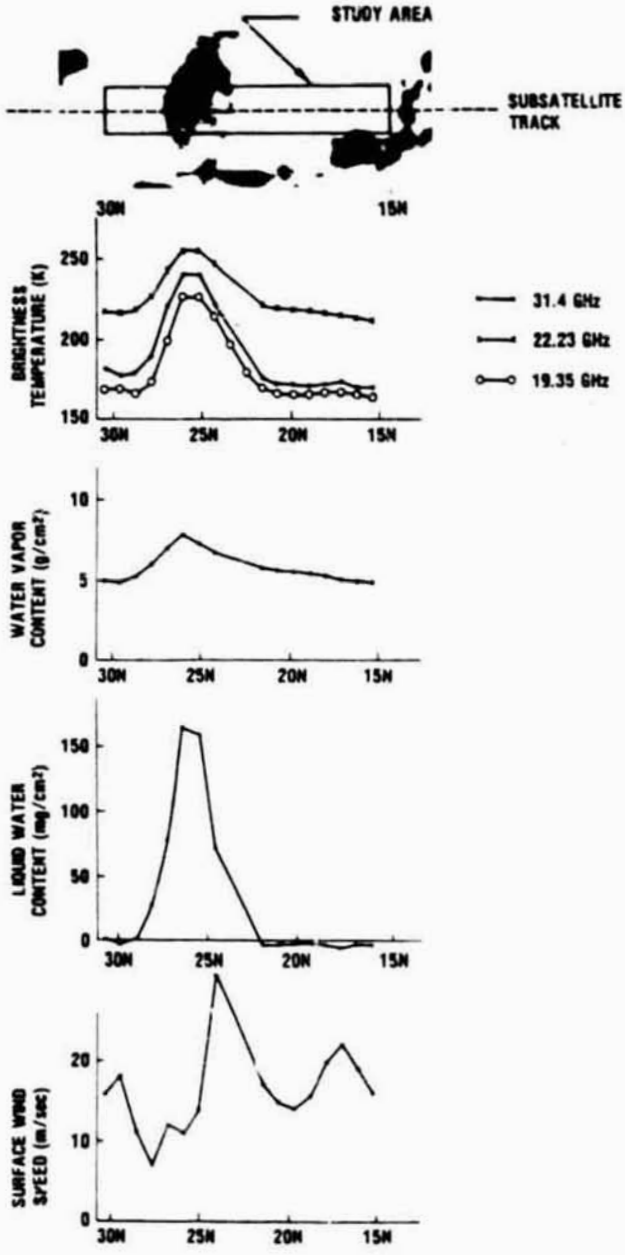


Fig. 2. The measured brightness temperatures and the ESMR image of Nimbus 5 orbit 2918, July 16, 1973. The estimates of the water vapor content, the cloud liquid water content, and the surface wind speed along the subsatellite track are presented.

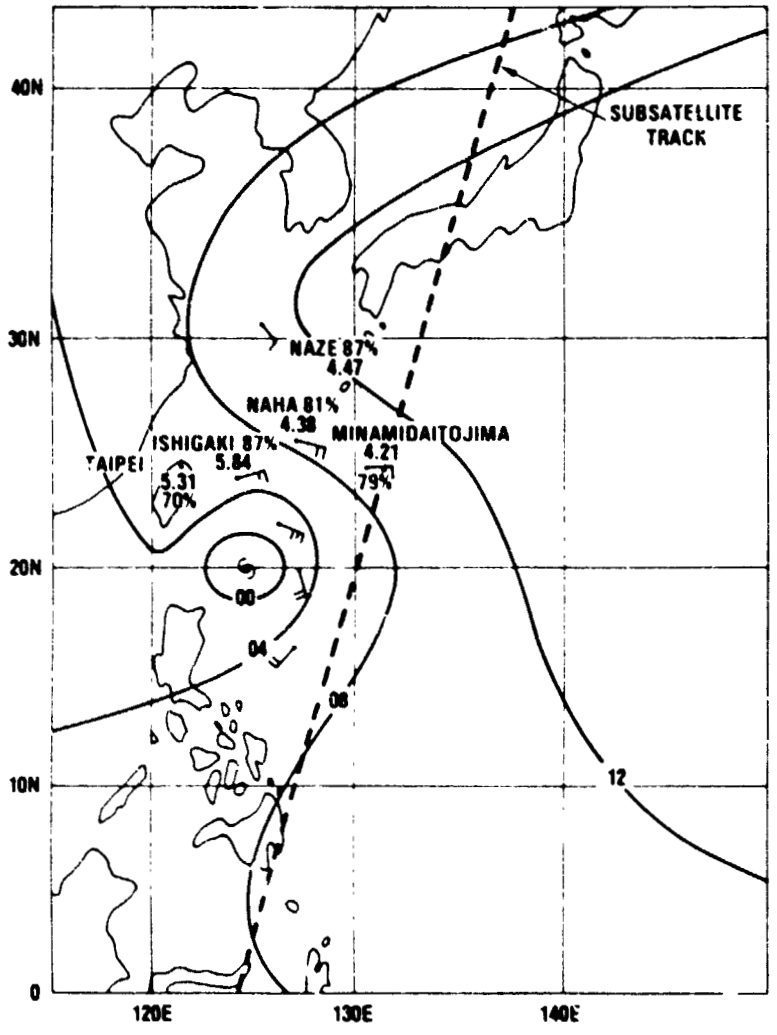


Fig 3 The surface weather chart corresponding to the studied area for 1200 UT, July 14, 1973 The percentage of relative humidity and water vapor content (in grams per cubic centimeter) for several weather stations are shown for reference

spheric water. Since the 22.235- and 31.4-GHz channels are primarily sensitive to water vapor and liquid water, we will use these logarithmic functions rather than the brightness temperature for these two frequencies and use the brightness temperature itself at 19.35 GHz. This discussion is not at all rigorous, but it serves to describe the gross behavior of the relationships, and the functions so derived will mitigate the nonlinear behavior of the problem. Since we are dealing with brightness temperature much colder than the thermodynamic temperature of the surface or atmosphere, we are not particularly sensitive to the value chosen for t_1, thus for present purposes we will set $t_1 = 280$K. We then solve for the three geophysical parameters, wind speed W, cloud liquid water content L, and water vapor V by conventional linear regression so that the parameters are solved for in the form

$$\begin{bmatrix} W \\ L \\ V \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} T_{B_{19.35}} \\ \ln(280 - T_{B_{22}}) \\ \ln(280 - T_{B_{31}}) \end{bmatrix} \tag{20}$$

where the values of a and b are given as follows:

a_1	−1008
a_2	1.831

a_3	37.92
b_{11}	2.330
b_{12}	66.81
b_{13}	76.68
b_{21}	−0.0024
b_{22}	−0.0146
b_{23}	−0.2941
b_{31}	−0.0479
b_{32}	−8.699
b_{33}	2.421

In Table 2 we also give the residual of each of the regressions which serves as an estimate of the accuracy with which the retrieval can be made and the a priori standard deviation of each of the variables which specifies the variability in the data base.

4. RESULTS

To demonstrate the retrieval technique, two orbits of Nimbus 5 ESMR and NEMS data were chosen. In order to utilize both ESMR and NEMS data for the retrieval calculations, these data must be matched up geographically. Since the instantaneous fields of view (IFOV's) for the NEMS radiometers are quite different from those of ESMR, it is necessary to sum up all the 25 km × 25 km ESMR IFOV's which are contained within the NEMS'

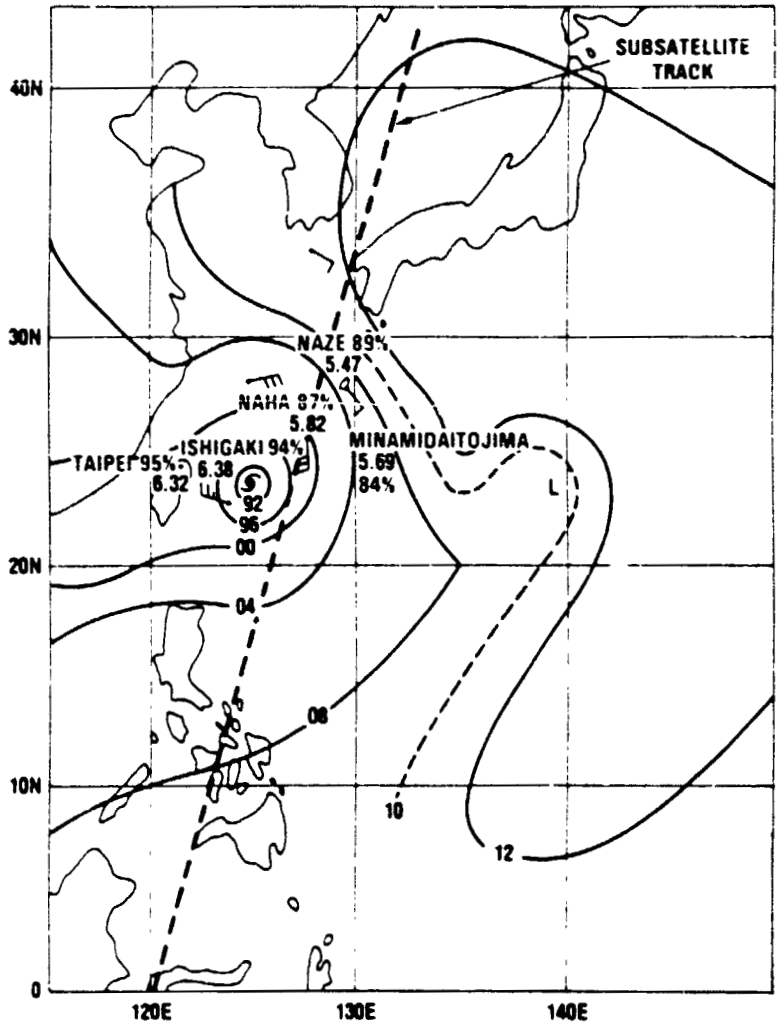


Fig. 4. The surface weather chart corresponding to the studied area for 1200 UT, July 16, 1973. The percentage of relative humidity and water vapor content (in grams per cubic centimeter) for several weather stations are shown for reference

IFOV that is approximately 200 km × 200 km. The brightness temperatures measured at 0.96, 1.35, and 1.55 cm radiometers for portions of orbits 2891 and 2918 are shown in Figures 1 and 2 plotted as a function of latitude. These matched data sets were then input into (20) to yield the estimates of the surface wind speed, cloud liquid water content, and water vapor content in the atmosphere. The results of these computations are also shown in Figures 1 and 2 along with the corresponding ESMR images. By showing rain and land areas as dark shades, a spatial perspective on the data is provided. The specific data sets used here are from Nimbus 5 orbits 2891 and 2918, which passed near the eye of Typhoon Billie at about 1500 UT on July 14 and 16, 1973, respectively. Figures 3 and 4 give the corresponding surface analyses (each 3 hours earlier than the satellite pass) along with relevant observations of wind speed and atmospheric water vapor content. Examination of the water vapor retrievals shows first that this varies smoothly with little structure. More specifically, on July 14 it can be seen that the retrieved water vapor of about 4.5 cm at 30°N agrees almost exactly with the observations. Further, the slight decrease in water vapor along the track to about 25°N is also supported by the observations. Beyond that, there are no supporting water vapor observations for this orbit. On July 16, similar agreement is observed between water vapor observations and retrievals. There are no direct observations of cloud liquid water content, but the areas with liquid water retrievals below 1 mg/cm correspond to clear areas as indicated by thermal infrared images from the same satellite. Also, the fact that the bulk of the retrieved data are very near to 0 mg/cm suggests that there is no consequential bias in this retrieval. Examination of the wind speed observations shows most observations in the 15–30 kt (7.5–15 m/s) range away from the storm center consistent with the retrievals and no observations in the high wind portions of the typhoon.

5 CONCLUSION

We have developed a quasi-statistical technique suitable for retrieval of geophysical parameters from microwave radiometric observations. We have derived specific equations for retrieving the wind speed at the ocean's surface and the atmospheric contents of both liquid and water vapor from simultaneous observations at 19.35, 22.235, and 31.4 GHz and applied these equations to data taken by the Nimbus 5 spacecraft, and by comparing the results with available observations we have shown that the results are promising.

The estimates of liquid water content in this study could be biased upward in the presence of precipitation because the present regression analysis is performed on the nonraining data ensemble. Further study in this area is required to resolve this problem.

REFERENCES

Air Force Cambridge Research Laboratory (1965), *Handbook of Geophysics and Space Environments*, McGraw-Hill, New York.

Born, M., and E. Wolf (1975), *Principles of Optics*, 182 pp., Pergamon, New York.

Gaut, N. E. (1968), Studies of atmospheric water vapor by means of passive microwave techniques, *Tech. Rep. 467*, Res. Lab. of Elec., Mass. Inst. of Technol., Cambridge.

Gunn, L. L. S., and T. W. R. East (1954), The microwave properties of precipitation particles, *Quart. J. Meteorol. Soc.*, *80*, 522–545.

Hollinger, J. P. (1971), Passive microwave measurements of sea surface roughness, *IEEE Trans. Geosci. Electron.*, *9*, 165–169.

International Critical Tables of Numerical Data, Physics, Chemistry and Technology (1928), vol. 4, pp. 231–239, McGraw-Hill, New York.

Jackson, J. D. (1962), *Classical Electrodynamics*, 219 pp., John Wiley, New York.

Lane, J. A., and J. A. Saxton (1952), Electrical properties of sea water, *Wireless Eng.*, *29*, 269–275.

Lenoir, W. B. (1968), Microwave spectrum of molecular oxygen in the mesosphere, *J. Geophys. Res.*, *73*, 361–376.

Meeks, M. L., and A. E. Lilly (1963), The microwave spectrum of oxygen in the earth's atmosphere, *J. Geophys. Res.*, *68*, 1683–1703.

Nordberg, W., J. Conaway, D. B. Ross, and T. Wilheit (1971), Measurements of microwave emission from a foam-covered, wind-driven sea, *J. Atmos. Sci.*, *28*, 429–435.

Saxon, J. A., and J. A. Lane (1952), Dielectric dispersion in pure polar liquids at very high radio-frequencies, I, Measurements on water, methyl and ethyl alcohols, *Proc. Roy. Soc. London*, *213*, 400–408.

Staelin, D. H., F. T. Barath, J. C. Blinn III, and E. J. Johnson (1972), The Nimbus E microwave spectrometer (NEMS) experiment, in *Nimbus-5 User's Guide*, pp. 141–157, NASA/Goddard Space Flight Center, Greenbelt, Md.

Stogryn, A. (1967), The apparent temperature of the sea at microwave frequencies, *IEEE Trans. Antennas Propagat.*, *15*, 278–286.

Van Melle, M. J., H. H. Wang, and W. F. Hall (1973), Microwave radiometric observations of simulated sea surface conditions, *J. Geophys. Res.*, *78*, 969–976.

Waters, J. W., K. F. Kunzi, R. L. Pettyjohn, R. K. L. Poon, and D. H. Staelin (1975), Remote sensing of atmospheric temperature profiles with Nimbus-5 microwave spectrometer, *J. Atmos. Sci.*, *32*, 1953–1969.

Webster, W. J., Jr., T. T. Wilheit, D. B. Ross, and P. Gloersen (1976), Spectral characteristics of the microwave emission from a wind-driven foam-covered sea, *J. Geophys. Res.*, *81*, 3095–3099.

Wilheit, T. T. (1972), The electrically scanning microwave radiometer (ESMR) experiment, in *Nimbus-5 User's Guide*, pp. 59–105, NASA/Goddard Space Flight Center, Greenbelt, Md.

Wilheit, T. T., and M. G. Fowler (1977), Microwave radiometeric determination of wind speed at the surface of the ocean during BESEX, *IEEE Trans. Antennas Propagat.*, *25*, 111–120, *IEEE Trans. Oceanic Eng.*, *2*, 111–120. (joint issue).

Cold Regions Science and Technology, 1 (1979) 121–128
© Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

THE SOLAR REFLECTANCE OF A SNOW FIELD

B.J. Choudhury
Computer Sciences Corporation, 8728 Colesville Rd., Silver Spring, MD 20910 (U.S.A.)

and A.T.C. Chang
NASA, Goddard Space Flight Center, Greenbelt MD 20771 (U.S.A.)

(Received February 12, 1979, accepted in revised form May 26, 1979)

ABSTRACT

The radiative transfer equation has been solved using a modified Schuster–Schwartzschild approximation to obtain an expression for the solar reflectance of a snow field. The parameters in the reflectance formula are the single scattering albedo and the fraction of energy scattered in the backward direction. The single scattering albedo is calculated from the crystal size using a geometrical optics formula and the fraction of energy scattered in the backward direction is calculated from the Mie scattering theory. Numerical results for reflectance are obtained for visible and near infrared radiation for different snow conditions. Good agreement has been found with the whole spectral range. The calculation also shows the observed effect of aging on the snow reflectance.

INTRODUCTION

Interest in developing techniques for remote measurement of snow parameters (e.g. temperature profile, density, grain size, water equivalent and free water content) have significantly increased in recent years. One of the more promising techniques to monitor these parameters is that of microwave radiometry. A recent study by Chang et al. (1976) has used a theoretical scattering model to relate the observed microwave brightness temperature to the physical temperature and snow grain size for snow on glacier ice. By utilizing this remotely determined snow grain size and density information, the solar spectral reflectance from a snow field can be inferred. The knowledge of spectral reflectance is important because it determines the amount of solar energy that will be reflected by a snow field. The present study is an attempt to calculate the solar reflectance of a snow cover in terms of parameters which depend upon the grain size and density.

For incident solar radiation, the energy reflected at the surface often constitutes only a small fraction (about 5%) of the total energy that is reflected by a snow field. For visible and near infrared radiation, the scattering by individual snow crystals is such that the major portion of the radiation gets scattered in the forward direction. The small fraction of energy that in scattered in the backward direction, however, gets enhanced considerably by the cumulative effect of many scatterers. Thus the major portion of the incident solar radiation which gets reflected by a snow cover is due to the volumetric effect of many scatterers.

The appropriate equation which needs to be solved for calculating the cumulative effect of many scatterers is the radiative transfer equation. An analytical solution of the radiative transfer equation when the individual scatterings are in the near forward direction is rather difficult, but a solution can be obtained by numerical methods (Choudhury 1977). In this paper we discuss an analytical solution with modified Schuster–Schwartzschild approximation (Sobolev 1963, Sagan and Pollack 1967). The ex-

pression for spectral reflectance obtained from this analytical solution of the radiative transfer equation can be evaluated in terms of snow parameters. The theory presented here is intended to show the influence of snow parameters on the solar reflectance of snow.

There have been several theoretical attempts to relate the spectral of a model snowcover with its physical properties (e.g., Dunkle and Bevans 1956, Giddings and La Chapelle 1961, Bergen 1970, 1971, 1975, Bohren and Barkstrom 1974). These theories either contain parameters which cannot be obtained directly from the measured physical properties or are valid for a restricted range within the visible spectrum. For an excellent review on snow see Mellor (1977).

In the following the general expressions and equations needed to study the spectral reflectance of a snow cover are given. It also includes the background information regarding the snow parameters and the solar intensity distribution which are expected to influence the spectral reflectance. The analytic solution of the radiative transfer equation is discussed and an expression for the spectral reflectance obtained. In addition comparisons are made for the obtained numerical results and the observations.

GENERAL FORMULATION

The solar energy incident on the snow surface can be divided into two components: the collimated intensity, I_c falling at an angle θ_0 with respect to the normal and the nearly isotropic diffuse intensity I_d which arises due to the atmospheric scattering of the radiation. The flux of the incident radiation is given by

$$F_i = I_c \mu_0 + I_d \pi \tag{1}$$

where $\mu_0 = \cos \theta_0$.

Since most of the incident radiation is contained in the collimated component, the flux of the radiation reflected by the snow-air interface is approximately given by

$$F_R = r(\theta_0)(I_c \mu_0 + I_d \pi) \tag{2}$$

where $r(\theta_0)$ is the Fresnel reflectivity of the surface. The effect of randomly oriented faceted grains is not considered.

The refracted portion of the incident radiation will progress through the medium (snow) getting weaker due to scattering and absorption. It is convenient to separate the refracted portion of the collimated radiation into two components: the unscattered radiation and the radiation which arises due to scattering. This separation is done because the unscattered radiation remains angularly concentrated along the direction of refraction but the angular distribution of the scattered radiation depends upon the scattering phase function. If multiple reflection at the boundaries of the medium is negligible then the unscattered intensity at depth x is given by:

$$I_u(\tau) = (1 - r(\theta_0)) n^2 I_c \, \delta(\phi - \phi_0) \, \delta(\mu - \mu_0) \exp(-\tau/\mu_0') \tag{3}$$

where n is the index of refraction, μ_0' is the angle of refraction:

$$\mu_0' = \left[1 - \frac{1 - \mu_0^2}{n^2}\right]^{1/2}$$

and τ is the optical depth.

$$\tau = \int_0^X \gamma_e(x') \, dx' \tag{4}$$

where $\gamma_e(x)$ is the extinction coefficient at depth x. The delta functions in eqn. (3) represent the constraint due to the collimated nature of the radiation.

The flux of the unscattered radiation is given by

$$F_u(\tau) = (1 - r(\theta_0)) I_c \mu_0 \exp(-\tau/\mu_0') \tag{5}$$

The intensity which arises due to scattering of the incident radiation is to be obtained by solving the radiative transfer equation (Chandrasekhar 1960):

$$\mu \frac{dI_s(\tau,\mu)}{d\tau} = -I_s(\tau,\mu) + \frac{\omega}{2} \int_{-1}^{1} p(\mu,\mu') I_s(\tau,\mu') \, d\mu' + \frac{\omega}{4\pi} n^2 I_c (1 - r(\theta_0)) p(\mu, \mu_0') \exp(-\tau'/\mu_0') \tag{6}$$

where μ is the angle of radiation, ω is the single scattering albedo and $p(\mu,\mu')$ is the scattering phase function.

The boundary conditions for the radiative transfer equation are as follows:

(1) If the snow-air interface is assumed to be a geometrically smooth surface then the appropriate boundary condition is

$$I_+(\tau = 0,\mu) = r(\mu)\, I_-(\tau = 0,\mu) + (1 - r(\mu))\, n^2\, I_d \quad (7)$$

where I_+ is the radiation going downward into the snow and I_- is the radiation going upward towards the atmosphere. The first term on the right hand side represents the intensity reflected by the snow-air interface and the second term accounts for the refracted component of the diffuse intensity.

(2) The second boundary condition is associated with the snow-soil interface. The general boundary condition will be of the form:

$$I_-(\tau_0,\mu) = r_1(\mu)\, I_+(\tau_0,\mu) \quad (8)$$

where $r_1(\mu)$ is the reflectivity of the snow-soil interface and τ_0 is the total optical thickness of snow.

Once the solution of eq. (6) is known subject to the above boundary conditions, one can calculate the flux of the radiation emerging from snow as:

$$F_e = 2\pi \int_0^1 I_-(0,\mu)\,(1 - r(\mu))(\mu d\mu/n^2) \quad (9)$$

The albedo of snow is to be calculated using eq. (1), (2) and (9) as:

$$A = (F_R + F_e)/F_i \quad (10)$$

The volumetric effect of scattering which is mainly responsible for large values of albedo is associated with the emergent flux F_e appearing in eq. (10) and is defined through the eq. (9). The calculation of the emergent flux requires the solution of the radiative transfer equation (eq. (6)).

APPROXIMATE EXPRESSION FOR SNOW REFLECTANCE

To obtain a closed form analytic solution of the radiative transfer equation (6), the following assumptions are made:

(1) The medium (snow) is a homogeneous plane parallel layer, consisting of statistically equal size ice spheres.

(2) The Schuster–Schwartzschild or two-stream approximation for the intensity is valid.

(3) The angular asymmetry in the scattering process can be described by introducing a parameter which weights the scattered intensities.

It should be noted that with the first approximation we are neglecting the presence of all internal inhomogenities like enlarged ice glands and ice lenses which sometimes form during the process of metamorphism. The other two approximations are associated directly with the method of solution. It has been shown that for an optically thick medium, the solution of the radiative equation under the above set of approximations is reasonably accurate (Sagan and Pollack 1967).

The Schuster–Schwartzschild or two-stream approximation is based upon partitioning the scattered radiation into mean intensities along the forward and the backward hemispheres. In single scattering, if β is the fraction of energy scattered in the backward hemipshere with respect to the direction of incidence, then we can replace the radiative transfer equation (6) by the following equations:

$$dI_+/d\tau = -I_+ + \omega(1-\beta)\, I_+ + \omega\beta I_- + \omega(1-\beta)\, I' \exp(-\tau) \quad (11)$$

and

$$-dI_-/d\tau = -I_- + \omega(1-\beta)\, I_- + \omega\beta I_+ + \omega\beta I' \exp(-\tau) \quad (12)$$

where

$$I' = [1 - r(0)]\,(I_c + I_d) \quad (a)$$

and I_+, I_- are respectively the mean intensities along the forward and backward directions with respect to the direction of incident radiation, and $r(0)$ is the Fresnel reflectivity for normal incidence. It should be noted that these equations represent the radiative

transfer in a one-dimensional medium (Sobolev 1963).

The radiative transfer equation (11) and (12) contains two parameters, namely, (1) ω, the single scattering albedo and (2) β, the fraction of energy scattered in the backward direction. We will discuss the method of determining these parameters before solving this radiative transfer equation.

For ice particles in the millimeter range and the radiation in the visible or near infrared region, the size of the particle is significantly larger than the wavelength. Under this condition, Mie scattering theory shows the extinction cross-section undergoes resonance oscillations for small changes in the wavelength and size of particle. This resonance effect, however, can be blurred by assuming that there is a statistical variation in the particle size about its mean radius. With this mean size value, one can use the geometrical optics approximation to calculate the single scattering albedo using the equation (Irvine and Pollack 1968, Sagan and Pollack 1967):

$$\omega = \tfrac{1}{2} + \tfrac{1}{2}[\exp(-2k_\lambda r)] \tag{13}$$

where r is the radius of the particle and k_λ the absorption per unit length of ice at wavelength, λ. In our calculation, we have used k_λ values tabulated by Irvine and Pollack (1968). Following the above discussion, instead of using eq. (13), it should be noted that one can also use Mie scattering theory to calculate the single scattering albedo.

Consistent with eq. (13) the extinction coefficient required in the calculation of optical depth (eq. (4)) is given by

$$\gamma_e = \frac{3}{2r}\left(\frac{\rho_s}{\rho_i}\right) \tag{14}$$

where ρ_s and ρ_i are respectively the density of snow and the density of ice.

The fraction of energy scattered in the backward direction, β, in a single scattering has been studied by many investigators (e.g. Bartky 1968, Sagan and Pollack 1967). This parameter can be related to the scattering phase function $p(\cos\theta)$ as:

$$2\beta = 1 - \int p(\cos\theta)\cos\theta \,\frac{d\Omega}{4\pi} \tag{15}$$

where θ is the scattering angle and $d\Omega$ is an element of solid angle in the co-ordinate system where θ is the polar angle. The normalization condition for the phase function is

$$\int p(\cos\theta)\,\frac{d\Omega}{4\pi} = 1 \tag{B}$$

When the phase function is expanded in terms of Legendre polynomials, the parameter β can then be expressed in terms of the coefficient of the first order Legendre polynomial.

For a given particle size one can use Mie scattering theory to calculate the phase function and then use eqn. (15) to calculate the value of β. Calculations performed by Irvine and Pollack (1968) show that when the size of the particle is significantly larger than the wavelength of the radiation, the value of β is about 0.075, i.e., in single scattering only 7% of the incident energy is scattered in the backward direction. In our calculation, we have used this value of β (0.075) because for visible and near infrared radiations, the scattering by individual ice crystals is indeed in the near-forward direction.

For a thick snow layer, the general solution of eqns. (11) and (12) can be obtained by standard methods (see Sobolev 1963, p. 33) as

$$I_+(\tau) = \tfrac{1}{2}\left(1 + \frac{1-\omega}{k}\right) C e^{-k\tau} - I' e^{-\tau} \tag{16}$$

and

$$I_-(\tau) = \tfrac{1}{2}\left(1 - \frac{1-\omega}{k}\right) C e^{-k\tau} \tag{17}$$

where C is the constant of integration and

$$k^2 = (1-\omega)(1-\omega(1-2\beta)) \tag{C}$$

To determine the constant of integration we now impose the boundary condition (7) to obtain

$$C = \frac{2I'}{\left(\left(1 + \frac{1-\omega}{k}\right) - r(0)\left(1 - \frac{1-\omega}{k}\right)\right)} \tag{18}$$

With the knowledge of complete solution, the reflectance for a one-dimensional system can be calculated as:

$$A = r(0) + I_-(\tau = 0)(1 - r(0))\,[1/(I_c + I_d)] \tag{19}$$

The explicit expression for reflectance is:

$$A = r(0) + \frac{(1 - r(0))^2 \alpha}{1 - r(0)\alpha} \tag{20}$$

where

$$\alpha = (k - 1 + \omega)/(k + 1 - \omega) \tag{D}$$

(Note that for $\omega = 1$ one finds $\alpha = 1$)

The Fresnel reflectivity for normal incidence, $r(0)$, can be calculated from the reflective index of ice, n using the formula:

$$r(0) = \left| \frac{n - 1}{n + 1} \right|^2 \tag{21}$$

Since the spectral dependence of n in the visible and the near infrared region is extremely weak (Irvine and Pollack 1968), we have used a constant 1.30 for n in calculating the $r(0)$.

The expression for reflectance is in close resemblance with that calculated by Dunkle and Bevans (1956) without explicitly making reference to a one-dimensional system. The definition of the parameters appearing in eqn. (20) are, however, quite different. We thus note that although there is a formal resemblance between the expression for reflectance obtained in this paper and that obtained by Dunkle and Bevans (1956), the method of obtaining this expression and the method of calculating the parameters of this expression are quite different in the two cases.

The expression for reflectance obtained in this paper is based upon simplifying approximations. We have discussed these approximations and attempted to provide their justification. The ultimate justification of the model is to be found by comparing with the experimental observations. In the next section, we provide the numerical results and compare them with the observations.

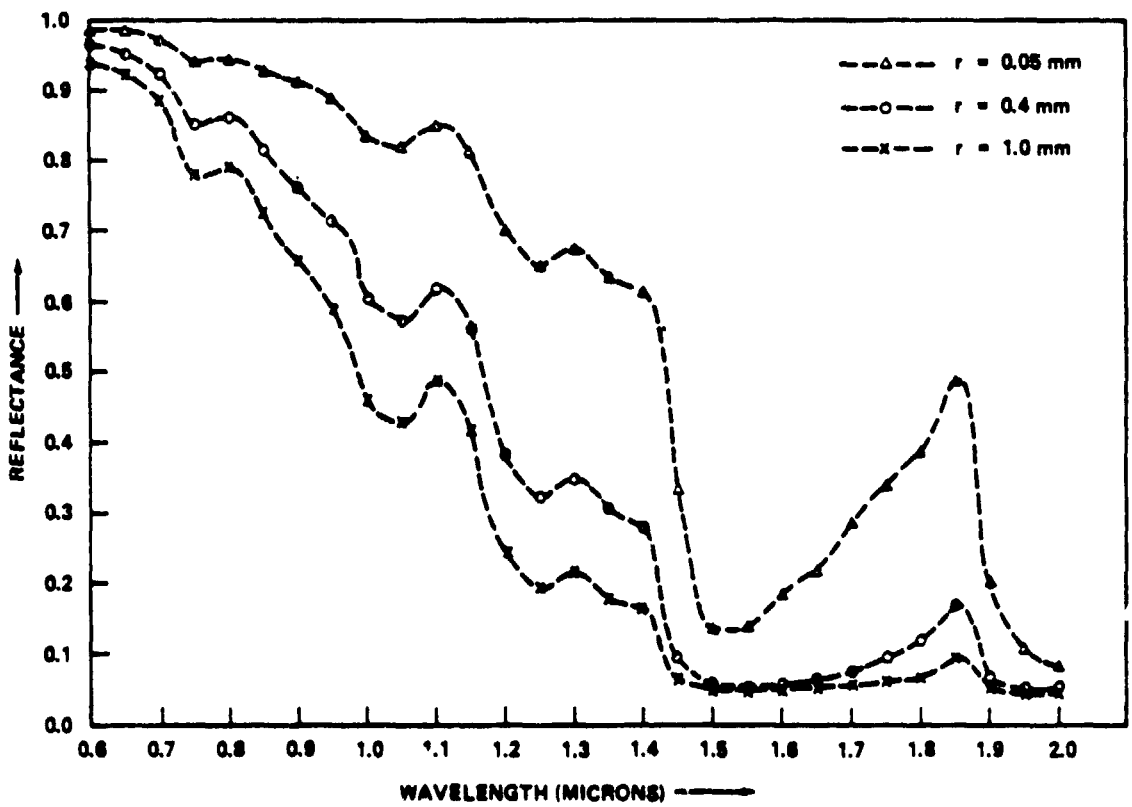


Fig. 1. Illustrations of the effect of different snow crystals on snow reflectance.

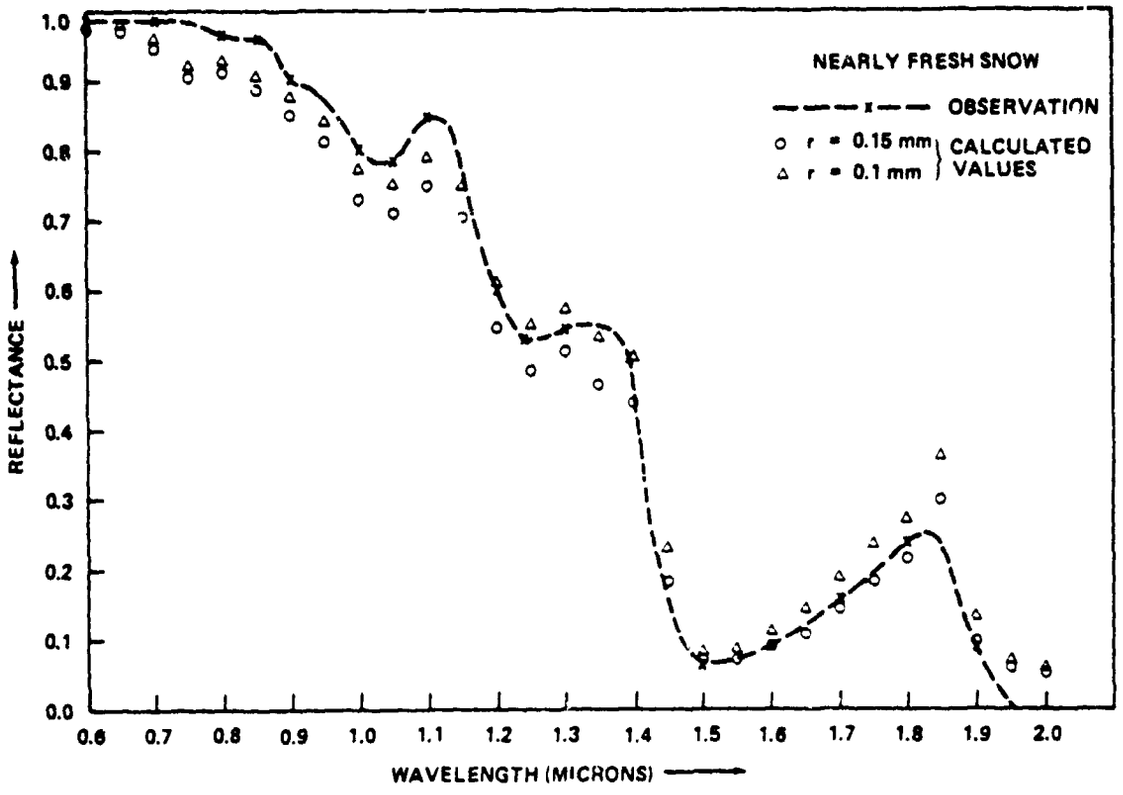


Fig. 2. Comparison of calculated and observed reflectance of a nearly fresh snow.

RESULTS AND DISCUSSION

By examining eqn. (20) the single scattering albedo ω is the only varying parameter in calculating the snow reflectance for a given wavelength. Since ω is directly related to the snow crystal radius, it is possible to characterize the snow condition by specifying the radius of the ice crystals.

The ice crystals of freshly fallen snow are usually of complex form and contain sharp corners. The shape of these crystals changes with time, depending upon the vapor content and the prevalent temperature. The process of equi-temperature metamorphism leads to the production of fairly uniform and well rounded grains. At the initial stage of this metamorphism the mean radius of the ice crystals is about 0.2 mm. The radius then continues to increase as the process of metamorphism advances. At a fairly advanced stage of this metamorphism, the radius increases to about 1.0 mm. The effect of other metamorphism, such as the temperature-gradient metamorphism or the melt-freeze metamorphism, is generally to produce non-spherical and non-uniform ice crystals. Also, the crystals are larger than with equi-temperature metamorphism. Thus, although it is not unique, one should be able to use the radius of ice crystal to characterize the stage of metamorphism.

Fig. 1 illustrates the changes in the calculated snow reflectance due to the difference in the crystal radius. The chosen radii are the typical snow crystal sizes for different stages of the equi-temperature metamorphism. The overall shape of the curve does not seem to depend upon the radius of the ice crystals but the relative magnitudes of various parts of the curve vary with the radius. The spectral reflectance in the red and near-infrared regions shows maximum sensitivity to the crystal sizes.

In Fig. 2 we show the calculated snow reflectance compared with the experimental values for a

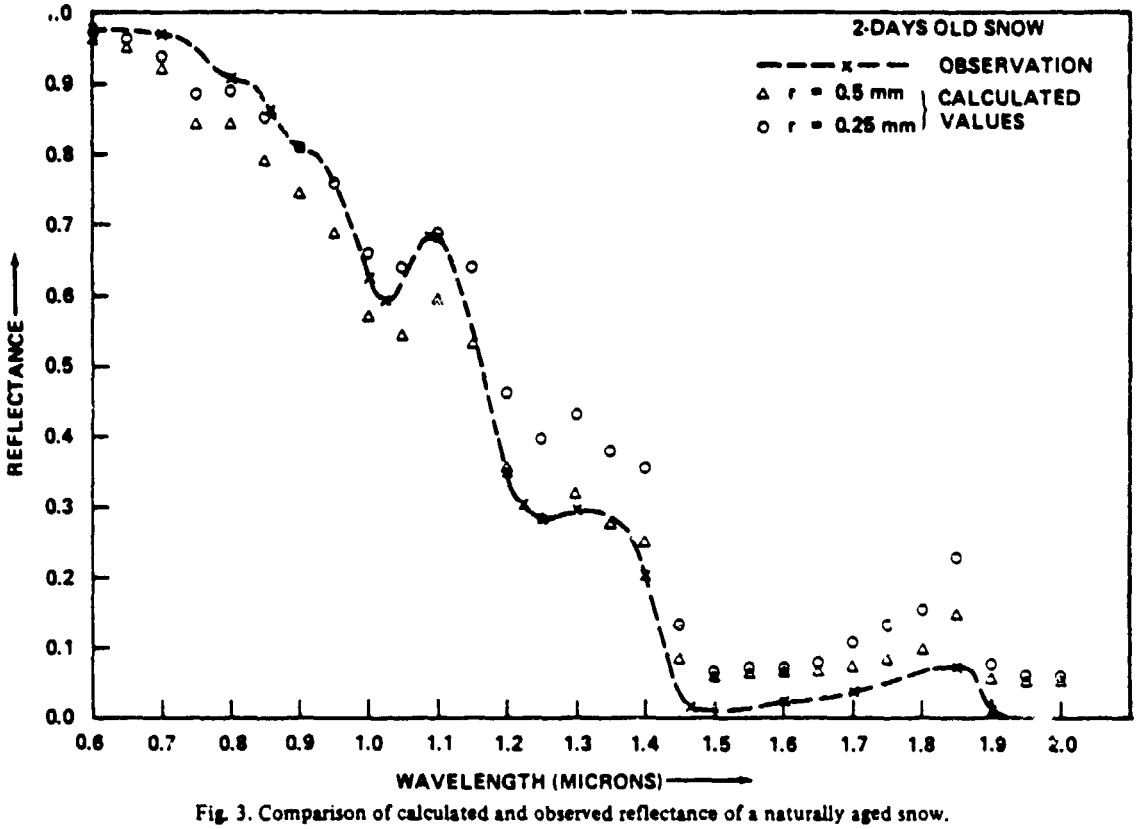


Fig. 3. Comparison of calculated and observed reflectance of a naturally aged snow.

nearly fresh snow (O'Brien and Munis 1975). The results of the calculation are in good agreement with the observations. All prominent spectral structures appearing in the observation are well duplicated in the calculations. The quantitative agreement is also good but should be treated with caution because the experimental results are relative to a standard (white barium sulfate) reflector. It is however reassuring that the radius values used in the calculation are in the range expected for a freshly fallen snow.

In Fig. 3 we compare the calculated and the observed reflectance of a naturally aged two days old snow (O'Brien and Munis 1975). There is a good qualitative agreement between them. The radii for the snow crystals which give agreement with the observation are larger than those for the freshly fallen snow (Fig. 2). The natural aging process of snow, during the observations, was such that the ambient air temperature was hovering above and below freezing point. It is unlikely that the snow was undergoing the equi-temperature metamorphism. As a result, the ice crystals were probably non-uniform within the snow. Furthermore, if the effect of the solar irradiance is interpreted as simply to advance the process of metamorphism of the snow layer in contact with the air, then the radius of the ice crystals in the top layer will be larger than those deep within. Since the absorption coefficient of ice increases with the wavelength, the effective thickness of the snow layer that contributes to the reflectance decreases with the wavelength. Therefore, the reflectance for longer wavelengths should have more influence from the top layer of snow and the reflectance for shorter wavelengths should correspond more to the deeper snow layer. This is a probable cause for the reflectance of the smaller crystals giving better agreement at shorter wavelengths while the reflectance of the larger crystals give better agreement at longer wavelengths. Although

comparisons are made with the observations of O'Brien and Munis (1975), similar observations have also been made by Vickers (1973) and Valovcin (1978).

Based on calculated results, two potential applications in remote snow reflectance monitoring are possible: (1) the use of multispectral bands to infer the radiation heat flux over a snow covered area; and (2) the use of visible and near-infrared channels to detect melting snow.

The large contrast in snow and soil reflectance makes it possible to monitor quantitatively the amount of solar radiation absorbed by the earth's surface. The degree of heat insulation by snow cover depends on the snow reflectance which is shown to be directly related to the snow crystal size in this paper. Therefore a spaceborne multispectral scanning instrument can provide a method to monitor the heat flux over a snow cover area under the National Weather and Climate Program.

The second potential application, that of detecting melting snow, is based on the calculation that the snow reflectance decreases rapidly with the increase in snow crystal size. The existence of melt water on the snow surface will tend to increase the crystal size, and also in the near infrared region the absorption per unit length of water is higher than that of ice. As a result, the reflectance of melting snow is expected to be lower than the dry snow in the near infrared region.

The theory presented here does show the qualitative and quantitative features of the observed reflectance. However the one-dimensional approach cannot distinguish the effects of the direct and the diffuse solar radiations. A more complexed analysis is needed to assess the effect of directional irradiation. In this regard the effect of surface roughness should also be considered.

REFERENCES

Bartky, C.D. (1968), The reflectance of homogeneous, plane-parallel clouds of dust and smoke, J. Quant. Spectrosc. Radiat. Transfer, 8: 51–68.

Bergen, J.D. (1970), A possible relation between grain size, density, and light attenuation in natural snow cover, J. Glaciol. 9: 154–156.

Bergen, J.D. (1971), The relation of snow transparency to density and air permeability in a natural snow cover, J. Geophys. Res., 76: 7385–7388.

Bergen, J.D. (1975), A possible relation of albedo to the density and grain size of natural snow cover, Water Resour. Res., 11: 745–746.

Bohren, C.F. and Barkstrom, B.R. (1974), Theory of the optical properties of snow, J. Geophys. Res., 76: 4527–4535.

Chandrasekhar, S. (1960), Radiative Transfer, Dover, New York.

Chang, T.C., Gloersen, P., Schmugge, T., Wilheit, T.T. and Zwally, H.J., (1976), Microwave emission from snow and glacier ice, J. Glaciol., 16: 23–39.

Choudhury, B.J. (1977), The method of invariant imbedding and radiative transfer in an innomogeneous medium, Computer Sciences Corporation, Technical Memo CSC/TM-77/6122

Dunkle, R.V and Bevans, J.R. (1956), An approximate analysis of the solar reflectance and transmittance of a snow cover, J. Meteorol., 13: 212–216.

Giddings, J.C. and LaChapelle, E. (1961), Diffusion theory applied to radiant energy distribution and albedo of snow, J. Geophys. Res., 66: 181–189.

Irvine, W.M. and Pollack, J.B. (1968), Infrared optical properties of water and ice spheres, Icarus, 8: 324–360.

Mellor, M. (1977) Engineering properties of snow. J. Glaciol., 19: 15–66.

O'Brien, H.W. and Munis, R.H. (1975) Red and near-infrared spectral reflectance of snow, in: A. Rango (Ed.), Operational Applications of Satellite Snowcover Observation, National Aeronautics and Space Administration. NASA-SP-391.

Sagan, C. and Pollack, J.B. (1967), Anisotropic nonconservative scattering and the clouds of Venus, J. Geophys. Res., 72: 469–477.

Sobolev, V.V., (1963), A Treatise on Radiative Transfer, D. Van Nostrand, Princeton, New Jersey.

Valovcin, F.R. (1978), Spectral radiance of snow and clouds in the near infrared spectral region, Air Force Geophysics Laboratory, Hanscom AFB, Mass. 01731, Technical Report, AFGL-TR-78-0289.

Vickers, A.F. (1973), Unpublisherd, see Mellor (1977).

Paper 73

Reprinted from IEEE Transactions on Geoscience Electronics, Vol. GE-17, No. 3, July 1979
0018-9413/79/0700-0063$00.75 © 1979 IEEE

Two-Stream Theory of Reflectance of Snow

B. J. CHOUDHURY, MEMBER, IEEE, AND ALFRED T. C. CHANG

Abstract—Spectral reflectance of snow under diffuse illumination is studied using the two-stream approximation of the radiative transfer equation. The scattering and absorption parameters of the radiative transfer equation—the single scattering albedo, the optical depth, and the integrated phase function are obtained from the grain size and density of snow. Analytical expressions for the intensity within the snowpack, the reflectance, and the asymptotic flux extinction coefficient, are given. Good agreement is shown between the theory and available experimental data on visible and near-infrared reflectance, and the asymptotic flux extinction coefficient. The theory may also be used to explain the observed effect of aging on the snow reflectance.

I. Introduction

INTEREST in developing techniques for remote measurement of snow parameters (e.g., density, grain size, water content) have increased significantly in recent years. One of the more promising techniques to monitor these parameters is that of microwave radiometry. A recent study by Chang and Choudhury [7] has used a radiative-transfer model to explain the observed brightness temperature of polar firn from the measured physical temperature, density, and grain size. Analysis of spectral reflectance is an alternative, and complementary to microwave radiometry. Apart from remote-sensing purposes, the study of spectral reflectance is relevant in climatology because it determines the amount of solar radiation absorbed by the earth's surface.

There have been several theoretical attempts to relate the spectral reflectance of a model snowcover with its physical properties (e.g., [11], [12], [2]-[5]). These theories either contain parameters which cannot be obtained directly from the measured physical properties or are valid for a restricted range within the visible spectrum.

Manuscript received November 19, 1973; revised April 14, 1979.
B. J. Choudhury is with the Computer Sciences Corporation, Silver Springs, MD 20910.
A. T. C. Chang is with the Laboratory for Atmospheric Sciences (GLAS), NASA/Goddard Space Flight Center, Greenbelt, MD 20771.

For visible and near-infrared radiation, the scattering by individual ice grains is such that a major portion of the radiation gets scattered in the forward direction [5]. The high reflectivity of snow is due to the multiple-scattering effect within snow. Calculation of spectral reflectance necessarily involves a study of radiative transfer in snow. For this purpose, the radiative-transfer equation is useful if snow is represented by isolated, randomly distributed ice grains. The physical picture of snow, thus stipulated, provides a simplistic model studied in aforementioned references. The snow parameters enter into the radiative transfer equation through the single scattering albedo, the extinction coefficient, and the phase function.

The physical properties of a natural snowcover are usually not known accurately and are, at best, approximately homogeneous. Although an accurate solution of the radiative-transfer equation will be useful for standard reference and comparison purposes, to date no such calculation has been performed. Based on an approximate solution of the radiative-transfer equation, a simplistic theory of spectral reflectance is given in this paper. All parameters in the theory can be calculated directly from the measured physical properties of snow, and the theory is applicable to visible and near-infrared radiation. The exact points of departure between the analysis presented below and the previous reported results vary. Instead of discussing these differences, a comparison will be made to show the relationship among the results. In addition, comparisons will be made between the theory and the observations.

II. Approximate Expression for Reflectance

To obtain a closed-form analytic expression for reflectance, the following assumptions will be made.

1) The snowcover is homogeneous.

2) Radiative transfer within the snowcover can be studied by dividing the radiation field into two streams of intensities.

3) Geometric optics calculations can be used to obtain the scattering and absorption parameters of individual ice grains.

With the first assumption the presence of internal inhomogeneities is neglected. This assumption is particularly crucial for a natural snowpack because of metamorphism. This is also a common assumption in all previous models.

The two-stream representation of the radiation field is a widely used method, and shown to be a reasonably accurate approximation of the radiative-transfer equation (e.g., [9], [10], [24], [25].) Coakley and Chylek [9], in particular, have performed a critical study of the validity of this approximation for different choices of the radiative-transfer parameters. This result depends not only upon the values chosen but also upon the combination of the parameters. Although multiple-scattering problems can be solved analytically with this approximation, the use of this approximation is a weakness of this study. A special form of this approximation was used previously in the calculation of snow reflectance [11], [4].

In general form, the two-stream approximation corresponds to replacing the radiative-transfer equation by the following coupled differential equations for mean intensities ($I_+(\tau)$ and $I_-(\tau)$) in the forward and the backward hemispheres:

$$\delta(dI_+)/(d\tau) = -I_+ + \omega(1-\beta)I_+ + \omega\beta I_-$$

$$-\delta(dI_-)/(d\tau) = -I_- + \omega(1-\beta)I_- + \omega\beta I_+$$

where ω is the single scattering albedo, and τ is the optical depth within the snow, given by

$$\tau = \gamma_e z$$

where γ_e is the extinction coefficient and z is the depth measured from the snow surface within the snowcover. The coefficients δ and β are parameters, the choice of which distinguishes alternate forms of the two-stream approximation. Some of the choices of δ and β are as follows.

1) Generalized Schuster-Schwartzschild [25]

$$\delta_1 = \tfrac{1}{2}$$

$$\beta_1 = \frac{1}{2}\int_0^1 d\mu \int_0^1 d\mu' p(\mu, -\mu').$$

2) Modified Schuster-Schwartzschild [17], [19]

$$\delta_2 = 1/\sqrt{3}$$

$$\beta_2 = \frac{1}{2}\left[1 - \frac{1}{2}\int_0^{\pi} p(\cos\theta)\cos\theta \sin\theta\, d\theta\right]$$

where $p(\mu, \mu')$ and $p(\cos\theta)$ have the usual meaning of the phase function [6] for an axially symmetric (laminar) medium.

It is expected that different set of δ and β will give different numerical results. The merit of a choice for a particular application should be decided by comparison with the exact solution of the radiative-transfer equation. Unfortunately, as stated in the introduction, such a comparison cannot be made for the present application. In the following, the method of determining the parameters and the solution of the equations will be discussed.

The phase function needed for the calculation of β can be obtained from experimental observations or from a scattering theory calculation. Since there are no available experimental observations, one needs to perform a theoretical calculation by specifying the shape, size, and the refractive index of ice grains. By discussing the difficulty and the limitations associated with the choice of shape, a recent study on snow [5] considered spherical nonabsorbing particles and used geometric optics to calculate the phase function. Geometric optics calculations for transparent spheres [5], [13] give the value of the parameter β_2 as 0.065. Using the approximate relationship between the parameters β_1 and β_2 [25] the value of the parameter β_1 is obtained as 0.12 ± 0.05. The uncertainty in β_2 is due to the stated accuracy of the approximate relationship. Because of this uncertainty, only the modified two-stream approximation will be compared with observations.

The single scattering albedo ω, can be obtained from the same calculation method which gives the phase function. As with the phase function, it is convenient to consider that snow consists of spherical particles. A simple formula which reproduces the single scattering albedo with considerable accuracy as compared with the Mie theory results is [15]

$$\omega = (1/2) + (1/2)\exp(-1.67\, k_\lambda r)$$

where k_λ is the absorption coefficient of ice for wavelength λ [14], [15] and r is the radius of the sphere. For nonspherical particles, one can use this formula by modifying the numerical coefficient in the exponential [19]. The spectral dependence of the single scattering albedo is through the absorption coefficient.

The extinction coefficient of snow γ_e is the product of the extinction cross section and the number density of ice grains [8]. By taking the extinction cross section as $2\pi r^2$, one obtains

$$\gamma_e = (3/2r)(\rho/\rho_i)$$

where ρ and ρ_i are, respectively, the density of snow and ice. All parameters in the radiative-transfer equation have now been related with the physical properties of the snowcover.

The question of appropriate boundary condition at the snow-air interface has been discussed by Barkstrom [1]. It was concluded that for visible and near-infrared radiations, this interface should not be treated as if it were a dielectric discontinuity—similar to the case of air-ocean or air-ice interface. Based on this study if the surface of the snowpack is not ice glazed (Fig. 1), then for diffuse incident radiation I_0, the boundary conditions for the solution of the equations are

$$I_+(0) = I_0$$

$$I_-(\tau_0) = R\, I_+(\tau_0)$$

where R is the reflectivity of the underlying surface and τ_0 is the total optical thickness of the snowcover:

$$\tau_0 = \gamma_e h$$

where h is depth of the snowcover.

Following standard procedures [20] the solution can be written as

$$I_+(\tau) = \frac{C_1}{2}\left[1 + \frac{1-\omega}{\delta a}\right]e^{-a\tau} + \frac{C_2}{2}\left[1 - \frac{1-\omega}{\delta a}\right]e^{a\tau}$$

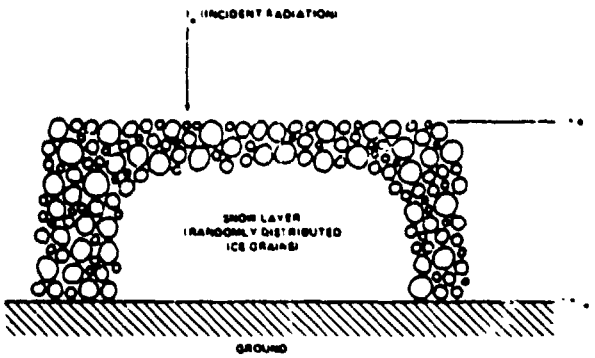


Fig. 1. Geometry of the model problem.

$$I_-(\tau) = \frac{C_1}{2}\left[1 - \frac{1-\omega}{\delta a}\right] e^{-a\tau} + \frac{C_2}{2}\left[1 + \frac{1-\omega}{\delta a}\right] e^{a\tau}$$

where

$$a^2 = \frac{(1-\omega)(1-\omega+2\omega\beta)}{\delta^2}$$

$$C_1 = \frac{2I_0}{\left[1 + \frac{1-\omega}{\delta a}\right]\left[1 + \frac{\alpha(\alpha-R)}{R\alpha-1} e^{-2a\tau_0}\right]}$$

$$C_2 = \frac{\left(\frac{\alpha-R}{R\alpha-1}\right) 2I_0 e^{-2a\tau_0}}{\left[1 + \frac{1-\omega}{\delta a}\right]\left[1 + \frac{\alpha(\alpha-R)}{R\alpha-1} e^{-2a\tau_0}\right]}$$

$$\alpha = 1 - \frac{2(1-\omega)^{1/2}}{(1-\omega+2\omega\beta)^{1/2} + (1-\omega)^{1/2}}.$$

From this solution, the reflectance $A(\tau_0)$ can be calculated as

$$A(\tau_0) = (I_-(0))/I_0$$

$$= \frac{(R\alpha-1)\alpha + (\alpha-R)e^{-2a\tau_0}}{(R\alpha-1) + \alpha(\alpha-R)e^{-2a\tau_0}}.$$

For a deep snowpack ($\tau_0 \to \infty$), the reflectance is given by

$$A(\infty) = \alpha.$$

The asymptotic flux extinction coefficient is obtained as

$$F = -\gamma_e \left[\left(\frac{dI_*(\tau)}{d\tau}\right) \Big/ (I_*(\tau))\right]_{\tau \to \infty}$$

$$= (3\rho)/(2\sqrt{5}\rho_i)\,[(1-\omega)(1-\omega+2\omega\beta)]^{1/2}.$$

III. Comparison with Previous Calculations and Observations

The reflectance of a deep snowcover and the asymptotic flux extinction coefficient can be calculated from the grain size and the density of the snowcover. Besides the assumptions made in obtaining the analytical results, the model of spherical particles was used to calculate the optical properties of individual ice grains. Although the backscattered fraction β is shown not very sensitive to the particle geometry [25], other optical parameters (i.e., the single scattering albedo and the extinction coefficient) do depend upon the geometry. The ice grains of a snowcover are not spherical. The shape is generally oblong and of variable dimensions. Although average optical properties of the grains, due to their random orientation with respect to the mean intensity, may correspond to a sphere, reservation should be exercised regarding the applicability of the theory to a real snowpack. Thus the comparison between the theory and the experimental observation presented below should not be regarded as conclusive. Note that all results are obtained for a deep snowpack ($\tau_0 \to \infty$).

Bergen [2], [4] related semiempirically the snowpack parameters to the reflectance and to the asymptotic flux extinction coefficient derived by Dunkle and Bevans [11]. Using the approximate solution of the radiative-transfer equation derived by van de Hulst [22], [23] and taking the model of spherical ice grains [5], Bohren and Barkstrom related the snow parameters to the reflectance and to the asymptotic flux extinction coefficient for visible radiation. The theoretical approach presented here differs from that of Bohren and Barkstrom, not only in the approximation method for the solution of the radiative-transfer equation but also in the equations relating the physical properties of snow to the parameters of the radiative-transfer equation. Since Bohren and Barkstrom have compared their theory with Bergen's results, the following comparison will be restricted to Bohren and Barkstrom's theory. The status of the experimental observations of the reflectance and the flux extinction coefficient is disappointing either because of quantitative disagreement [16], [21] or due to incomplete specification of snow parameters [18]. Since Bohren and Barkstrom have compared with Lilljequist's observations [16], this comparison will be discussed first with the present theory. Fig. 2 and Table I show this comparison for the asymptotic flux extinction, and the reflectance of a deep snowcover. Although present theory appears to be little inferior to Bohren and Barkstrom's theory, it is difficult to vindicate either theory because the observations have not yet been duplicated. The discrepancy between the models is particularly noticeable in the asymptotic flux extinction coefficient. Whereas Bohren and Barkstrom predict lower values, the present theory predicts higher ones than the observations.

Fig. 3 shows the calculated snow reflectance compared with the experimental values for a nearly fresh snow [18]. The snow density and mean crystal radius used in this calculation are 0.17 g/cm^3 and 0.1 mm, respectively. The results of the calculation are in good agreement with the observations. All prominent spectral structures appearing in the observation are well duplicated in the calculations. The quantitative agreement is also good but should be treated with caution because the experimental results are relative to a standard (white barium sulfate) reflector. It is, however, reassuring that the radius values used in the calculation are in the range expected for a fresh-fallen snow.

The ice crystals of fresh-fallen snow usually are of complex form and contain sharp corners. The shape of these crystals changes with time depending upon the vapor content and the prevalent temperature. The process of equi-temperature metamorphism leads to the production of fairly uniform and well-

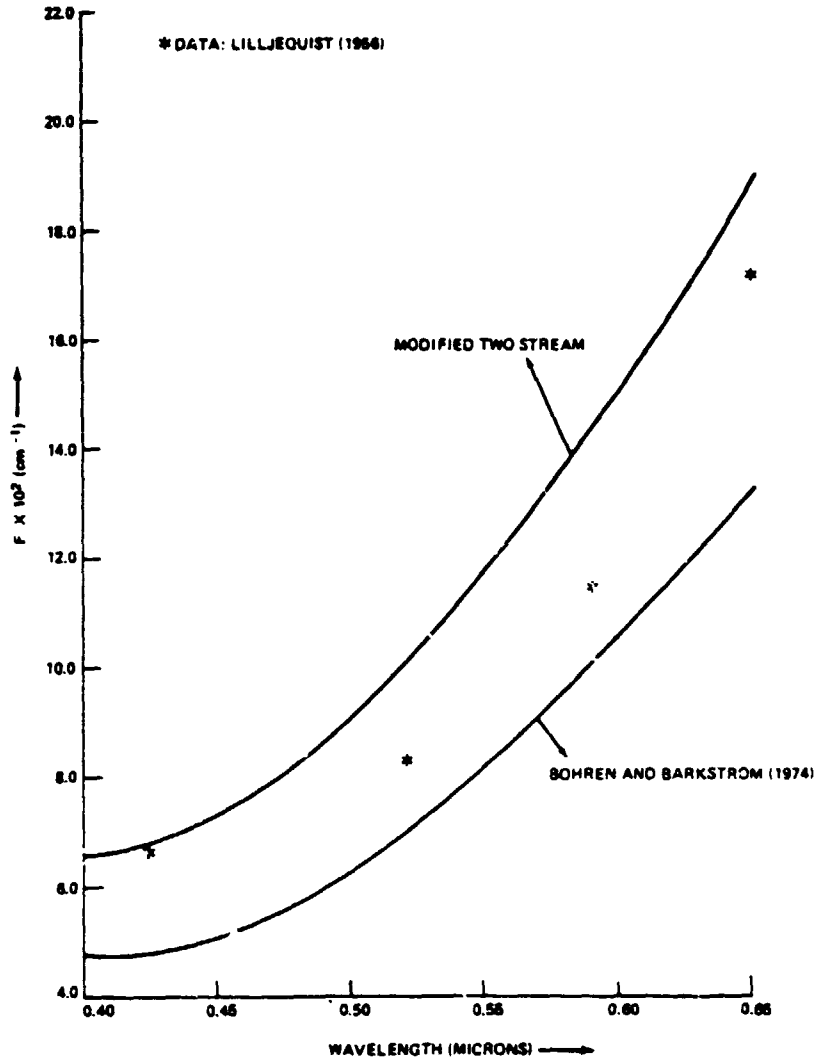


Fig. 2. Comparison of observed and calculated values of the asymptotic flux extinction coefficient.

TABLE I
COMPARISON OF OBSERVED AND CALCULATED VALUES OF THE REFLECTANCE SNOW PARAMETERS
($\rho = 0.43$ g/cm^3, $r = 0.15$ mm)

Wavelength (μ)	Modified two-stream	Observation*	Bohren and Barkstrom
0.420	0.987	0.960	0.978
0.520	0.980	0.972	0.967
0.595	0.974	0.960	0.954
0.550	0.964	0.921	0.940

*Data Lilljequist (1956)

rounded grains. At the initial stage of this metamorphism the mean radius of the ice crystals is about 0.2 mm. The radius then continues to increase as the process of metamorphism advances. At a fairly advanced stage of this metamorphism, the radius increases to about 1.0 mm. The effect of other metamorphisms such as the temperature-gradient or the melt-freeze

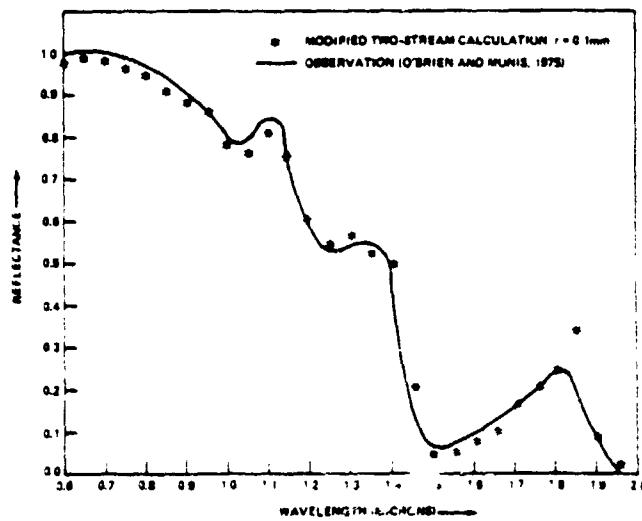


Fig. 3. Comparison of observed and calculated visible and near-infrared reflectance of nearly fresh snowpack.

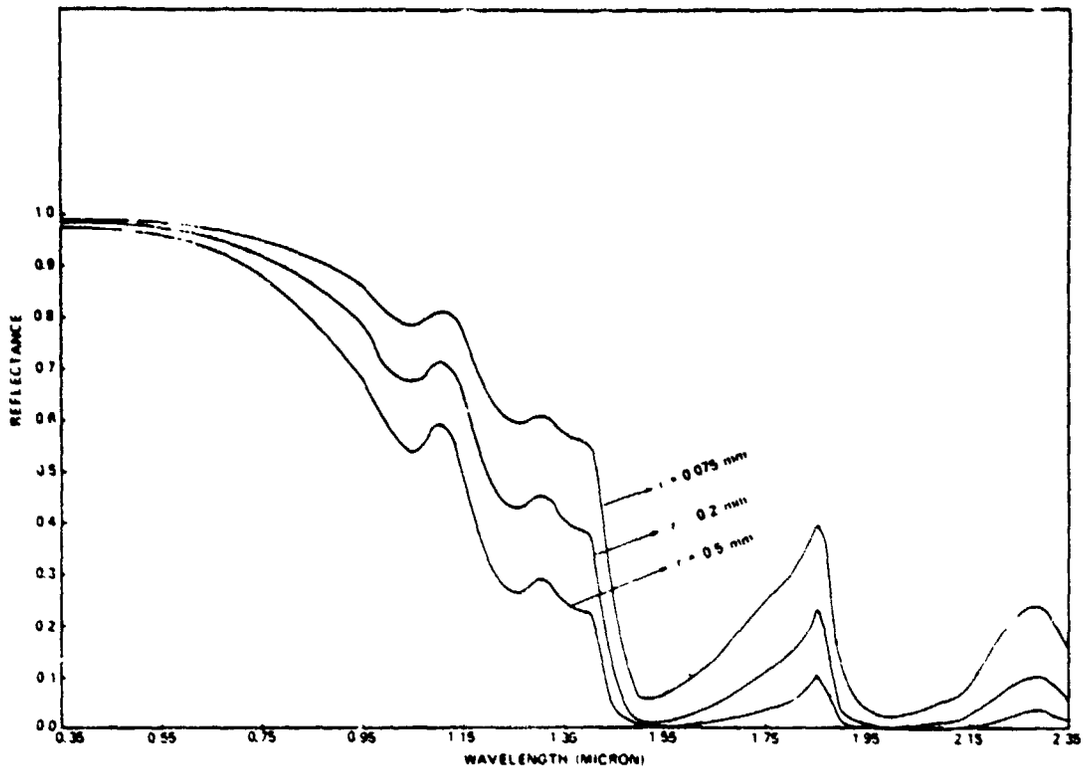


Fig. 4. Illustration of the effect of different snow crystals on snow reflectance.

metamorphism is generally to produce nonspherical and nonuniform ice crystals; also the crystals are larger than equi-temperature metamorphism. Thus although it is not unique, one should be able to use the radius of ice crystal to characterize the stage of metamorphism. Fig. 4 illustrates the changes in the calculated snow reflectance due to the difference in the crystal radius (r = 0.075, 0.2, and 0.5 mm). The chosen radii are the typical snow crystal sizes for different stages of the equi-temperature metamorphism. The overall shape of the curve does not seem to depend upon the radius of the ice crystals but the relative magnitudes of various parts of the curve vary with the radius. The spectral reflectance in the red and near-infrared regions shows maximum sensitivity to the crystal sizes.

IV. Conclusion

Based on the two-stream approximation of the radiative-transfer equations, analytic results for the reflectance and the asymptotic flux extinction coefficient of a homogeneous snowpack are derived. Using geometric optics calculations for spherical particles, the reflectance and the extinction coefficients are related to directly measurable snow parameters (density and the grain size). Good agreement was shown with the observed spectral dependence of the reflectance and the extinction coefficients using the measured (expected) snow parameters. Further observations are highly desirable to test the accuracy, and hence to refine the calculation. The applicability of the radiative-transfer equation itself should be clarified because the ice grains are generally in contact with each other.

References

[1] B. R. Barkstrom, "Some effects of multiple scattering on the distribution of solar radiation in snow and ice," *J. Glaciology*, vol. 11, pp. 357–368, 1972.
[2] J. D. Bergen, "A possible relation between grain size, density, and light attenuation in natural snow cover," *J. Glaciology*, vol. 9, pp. 154–157, 1970.
[3] —, "The relation of snow transparency to density and air permeability in a natural snow cover," *J. Geophys. Res.*, vol. 76, pp. 7385–7388, 1971.
[4] —, "A possible relation of albedo to the density and grain size of natural snow cover," *Water Resources Res.*, vol. 11, pp. 745–746, 1975.
[5] C. F. Bohren and B. R. Barkstrom, "Theory of the optical properties of snow," *J. Geophys. Res.*, vol. 79, pp. 4527–4535, 1974.
[6] S. Chandrasekhar, *Radiative Transfer*. New York: Dover, 1960.
[7] A. T. C. Chang and B. J. Choudhury, "Microwave emission from polar firn," Nat. Aeronautics and Space Administration, Tech. Paper 1212, 1978.
[8] T. C. Chang, P. Glousen, T. Schmugge, T. T. Wilheit, and H. J. Zwally, "Microwave emission from snow and glacier ice," *J. Glaciology*, vol. 16, pp 23–39, 1976.
[9] J. A. Coakley, Jr., and P. Chylek, "The two-stream approximation in radiative transfer:—Including the angle of the incident radiation," *J. Atmospheric Sci.*, vol. 32, pp. 409–418, 1975.
[10] G. A. Domoto and W. C. Wang, "Radiative transfer in homogeneous non-gray gases with non-isotropic particle scattering," *J. Heat Transfer*, vol. 96, pp. 385–390, 1974.
[11] R. V. Dunkle, and J. T. Bevans, "An approximate analysis of the solar reflectance and transmittance of a snow cover," *J. Meteorology*, vol. 13, pp. 212–216, 1956.
[12] J. C. Giddings and E. LaChapelle, "Diffusion theory applied to radiant energy distribution and albedo of snow," *J. Geophys. Res.*, vol. 66, pp. 181–189, 1961.
[13] J. E. Hansen and L. D. Travis, "Light scattering in planetary atmospheres," *Space Sciences Rev.*, vol. 16, pp. 527–610, 1974.
[14] P. V. Hobbs, *Ice Physics*. Oxford: Clarendon, 1974.
[15] W. M. Irvine and J. B. Pollack, "Infrared optical properties of water and ice spheres," *Icarus*, vol. 8, pp. 324–360, 1968.

[16] G. H. Liljequist, "Energy exchange of an antarctic snow-field, shortwave radiation (Maudheim 70°03'S 10°56'W)" Norwegian-British-Swedish Antarctic Expedition, 1949-52, Science Results 2 Part 1A, Norsk Polarinstitute, Oslo, Norway, 1956.
[17] D. R. Lyzenga, "The modified two-stream approximation of Sagan and Pollack," *Icarus*, Note, vol. 19, pp. 240-243, 1973.
[18] H. W O'Brien and R. H. Munis, "Red and near-infrared spectral reflectance of snow, in *Operational Application of Satellite Snowcover Observation*, A. Rango, Ed., Nat. Aeronautics and Space Administration, NASA SP-391, 1975.
[19] C. Sagan and J. B. Pollack, "Anisotropic nonconservative scattering and the clouds of Venus," *J. Geophysical Res.*, vol. 72, pp. 469-477, 1967.
[20] V. V. Sobolev, "*A Treatise on Radiative Transfer*. Princeton, NJ: Van Nostrand, 1963.
[21] C. W. Thomas, "On the transfer of visible radiation through sea ice and snow," *J. Glaciology*, vol. 4, pp. 481-484, 1963.
[22] H. C. van de Hulst, "The spectrum of the anisotropic transfer equation," *Astronomy Astrophys.*, vol. 9, pp. 366-373, 1970.
[23] —, "High-order scattering in diffuse reflection from a semi-infinite atmosphere," *Astronomy Astrophys.*, vol. 9, pp. 374-379, 1970.
[24] A. Venkatram and R. Viskanta, "Radiative effects of elevated pollutant layers," *J. Appl. Meteorology*, vol. 16, pp. 1256-1272, 1977.
[25] W. J. Wiscombe and G. W. Grams, "The backscattered fraction in two-stream approximations," *J. Atmospheric Sci.*, vol. 33, pp. 2440-2451, 1976.

Paper 74

NASA Technical Memorandum 80576 October 1979

ON THE ANGULAR VARIATION OF SOLAR REFLECTANCE OF SNOW

B. J. Choudhury
and
A. T. C. Chang

ABSTRACT

Spectral and integrated solar reflectance of non-homogeneous snowpacks are derived assuming surface reflection of direct radiation and subsurface multiple scattering. For surface reflection, a bidirectional reflectance distribution function derived for an isotropic Gaussian faceted surface is considered, and for subsurface multiple scattering, an approximate solution of the radiative transfer euqation is studied. Solar radiation incident on the snowpack is decomposed into direct and atmospherically scattered radiation. Spectral attenuation coefficients of ozone, carbon dioxide, water vapor, aerosol and molecular scattering are included in the calculation of incident solar radiation. Illustrative numerical results are given for a case of North American winter atmospheric condition. The calculated dependence of spectrally integrated directional reflectance (or albedo) on solar elevation is in qualitative agreement with available observations.

Paper 75

NASA Technical Paper 1404

PRELIMINARY RESULTS OF SAR SOIL MOISTURE EXPERIMENT, NOVEMBER 1975

B. J. Choudhury
Computer Sciences Corporation
Silver Spring, Maryland
and
A. T. C. Chang, V. V. Salomonson, T. J. Schmugge,
and J. R. Wang
Laboratory for Atmospheric Sciences
Goddard Space Flight Center
Greenbelt, Maryland

1979

ABSTRACT

This report presents preliminary results of the soil-moisture remote sensing experiment of November 1975, using a synthetic aperture radar (SAR) system. The experiment was performed using the Environmental Research Institute of Michigan's (ERIM) dual-frequency and dual-polarization side-looking SAR system on board a C-46 aircraft. The operating frequencies were 1.304 GHz (23 cm, L-band) and 9.375 GHz (3.2 cm, X-band). For each frequency, horizontally polarized pulses were transmitted and both horizontally and vertically polarized return signals were recorded on the signal film simultaneously. The test sites were located in St. Charles, Missouri; Centralia, Missouri; and Lafayette, Indiana. Each test site was a 4.83-km by 8.05-km (3-mile by 5-mile) rectangular strip of terrain. Concurrent with SAR overflight, ground soil samples of 0-to-5-cm and 0-to-15-cm layers were collected for soil moisture estimation. The surface features were also noted. Hard-copy image films and the digital data produced via optical processing of the signal films are analyzed in this report to study the relationship of radar backscatter to the moisture content and the surface roughness. Many difficulties associated with processing and analysis of the SAR imagery are noted. In particular, major uncertainty in the quantitative analysis appeared due to the difficulty of quality reproduction of digital data from the signal films.

Paper 76

VOL. 84, NO. C9 JOURNAL OF GEOPHYSICAL RESEARCH SEPTEMBER 20, 1979

Effect of Surface Roughness on the Microwave Emission From Soils

B. J. CHOUDHURY

Computer Sciences Corporation, Silver Spring, Maryland 20910

T. J. SCHMUGGE AND A. CHANG

Goddard Space Flight Center, Greenbelt, Maryland 20771

R. W. NEWTON

Texas A&M University, College Station, Texas 77843

The effect of surface roughness on the brightness temperature of a moist terrain has been studied through the modification of Fresnel reflection coefficient and using the radiative transfer equation. The modification involves introduction of a single parameter to characterize the roughness. It is shown that this parameter depends on both the surface height variance and the horizontal scale of the roughness. Model calculations are in good quantitative agreement with the observed dependence of the brightness temperature on the moisture content in the surface layer. Data from truck mounted and airborne radiometers are presented for comparison. The results indicate that the roughness effects are great for wet soils where the difference between smooth and rough surfaces can be as great as 50K.

INTRODUCTION

There have been several recent papers presenting theoretical models for the microwave emission from soils [e.g., *Njoku and Kong*, 1977; *Wilheit*, 1978; *Burke et al.*, 1979, *England*, 1976]. These models considered the emission from the soil for a range of moisture and temperature profiles and studied the effect of variations of these subsurface properties on the emission from the surface. The effects of surface features such as roughness were not included. However, when the results of one such set of the calculations were compared with observations by airborne radiometers [*Schmugge et al.*, 1976] there were rather large differences (~30K) between the calculated and observed brightness temperature (T_p). These differences were attributed to surface roughness. The purpose of this paper is to show that surface roughness effects can account for these differences.

The scattering of electromagnetic waves from rough surfaces has been studied by many investigators [see *Barrick*, 1970; *Wu and Fung*, 1972; *Sung and Eberhardt*, 1978]. These studies show that for a detailed quantitative calculation of the scattering by a rough surface, the knowledge of the statistical surface parameters are important. The roughness structure of an agricultural terrain depends upon the cultivation practice of that area. A typical surface may consist of furrows, clods and irregular, small amplitude undulations of different spatial dimensions. To study the effects of surface roughness on the observed dependence of the brightness temperature on the soil moisture, a simplistic model has been developed. The surface roughness effect has been incorporated into the calculation by modifying the Fresnel reflectivity. This modification is based upon the theory developed by *Ament* [1953] for a conducting surface. The emphasis in this paper is to show qualitatively the effect of surface roughness on the microwave brightness temperature. The present model is simplistic because it has only one parameter to characterize the surface, namely, the standard deviation of surface height. We realize that this description may not be an exact representation of actual soil surfaces. At present the knowledge of the statistical characteristics of soil surfaces is not sufficient to formulate a model which is not only general enough but also numerically tractable to provide a quantitative description. It is therefore not the intention of this paper to claim that this model will provide a rigorous, quantitative description of the different aspects of the microwave emission from natural terrains but it will provide a first step for including the effects of roughness in the modeling of the emission from these surfaces.

In the calculation of brightness temperature, we have used measured soil moisture and soil temperature profiles. The calculated values of the brightness temperature are in good quantitative agreement with the values observed by truck mounted and airborne radiometers. This agreement has been demonstrated for two different wavelengths. Details for the theory and the results of the calculation are given in the following sections.

THEORY

Radiative Transfer

To describe the microwave emission from the soil, we will consider the radiative transfer equation [*Chandrasekhar*, 1960]:

$$\frac{dI}{dZ} = -K_e(Z)I + S \tag{1}$$

where I is the intensity propagating in the direction Z, K_e is the total extinction per unit length, and S is the source term describing the contribution to the intensity due to scattering and due to the continuum thermal emission of the soil. In principle one should study this equation in conjunction with the equation describing the heating of the soil. It is this latter equation which will provide information about the thermal part of the source term. In this paper we will decouple these

Copyright © 1979 by the American Geophysical Union

two equations in the sense that we will consider a given temperature distribution.

To solve the radiative transfer equation we will consider a semi-infinite medium with depth dependent temperature and moisture distributions. Since soil is a highly absorbing material (i.e., large imaginary part of dielectric constant), to a good approximation, the brightness temperature T_B or the temperature equivalent of the intensity emerging from the soil will be determined by its internal temperature distribution $T(z)$. By integrating (1) with the source term as the temperature distribution of the soil one can write

$$T_B = [1 - r(0)] \int_{-\infty}^{0} T(z) K_a(Z) \exp\left(-\int_{z}^{0} K_a(Z')\, dZ'\right) dZ \quad (2)$$

where $r(0)$ is the soil surface reflectivity at normal incidence and $K_a(Z)$ is the absorption per unit length. These can be determined from the dielectric constants in the soil [*Born and Wolf*, 1975]. *Wilheit* [1978] has developed a model in which the integral is evaluated by a sum over many homogeneous layers:

$$T_B = \sum f_i T_i \quad (3)$$

where f_i is the fraction of the radiation incident on the air-soil interface that would be absorbed in the ith layer and T_i is the temperature of this layer. The values of f_i are determined by applying the electromagnetic boundary conditions to determine the energy fluxes entering and leaving each layer. The computations indicate that the radiation from the soil is characterized by two sampling depths: reflective and thermal. The reflectivity is characterized by changes in the real part of the index of refraction over a sampling depth: $\delta_r \approx 0.1\lambda$, where λ is the free space wavelength. The thermal sampling depth is determined by the losses deeper in the medium, and is given by

$$\delta_T = \sum x_i f_i / \sum f_i \quad (4)$$

where x_i is the depth of the ith layer. For a uniform dielectric this reduced to

$$\delta_T = \frac{\lambda}{4\pi \operatorname{Im}(n)} \quad (5)$$

For a low-loss dry soil, δ_T will be an order of magnitude larger than δ_r, while for a wet soil, it will be only slightly larger. A similar theoretical treatment has been developed by *Tsang et al.* [1975]. This formalism is simpler and has yielded brightness temperature values which are within 1 or 2 K of their results for the same moisture and temperature profiles.

Roughness Effects

It has been shown in *Tolstoy and Clay* [1956] that if the scattering surface is a satistically rough surface such that there is no correlation between the amplitudes of the waves scattered by two points on the surface, then the scattered intensity can be obtained by the absolute square of the average scattered amplitude. It has further been shown that if E_{sm} represents the scattered amplitude by a perfectly smooth and perfectly reflecting surface, then the average amplitude that will be specularly scattered at an angle θ by a rough surface is given by

$$\langle E_{sr} \rangle = R_0(\theta) E_{sm} \int_{-\infty}^{\infty} W(z) \exp(2ik_z z)\, dz \quad (6)$$

where $W(z)$ is the height distribution of the surface and R_0 is the reflection coefficient of a smooth surface. A typical rough surface corresponds to identifying the spectrum with a Gaussian distribution of zero mean and variance σ^2:

$$W(z) = \frac{1}{\sigma(2\pi)^{1/2}} \exp[-z^2/2\sigma^2] \quad (7)$$

For this spectrum, the average amplitude is given by

$$\langle E_{sr} \rangle = R_0(\theta) E_{sm} \exp[-2\sigma^2 k_z^2] \quad (8)$$

Since

$$k_z^2 = \left(\frac{2\pi}{\lambda}\right)^2 \cos^2\theta \quad (9)$$

The scattered intensity obtained by squaring (8) is

$$I_s(\theta) = I_s^0 |R_0(\theta)|^2 \exp(-h\cos^2\theta) \quad (10)$$

where the roughness parameter h is given by

$$h = 4\sigma^2 \left(\frac{2\pi}{\lambda}\right)^2 \quad (11)$$

From (10) one can stipulate that the gross effect of the surface roughness on the scattered intensity can be incorporated by modifying the smooth surface reflectivity $r_{sp}(\theta) = |R_0(\theta)|^2$ as

$$r_p(\theta) = r_{sp}(\theta) \exp(-h\cos^2\theta) \quad (12)$$

where the subscript p designates the polarization. The surface emissivity is obtained from (12) by

$$e_p(\theta) = 1 - r_p(\theta) \quad (13)$$

To verify this result, measurements were made by *Waite et al.* [1973] of the reflectivity for soils with different surface roughness conditions. They found that σ is not a sufficient indicator of the roughness for this model [*Hancock*, 1976]. Table 1 is a summary of their results at a look angle of 30°. The effective σ was determined by fitting the observed frequency dependence of the reflectivity in each band to that expected from (12) for a given σ. The effective σ was always greater than the measured σ. They also found that the autocorrelation length of the roughness was also important. This latter quantity is essentially an indicator of the horizontal scale of roughness.

The value of $r_{sp}(\theta)$ can be determined from the Fresnel equation for the case of a uniform dielectric, or from the layered models mentioned earlier for situation with non-uniform dielectrics. In either case it is necessary to know the variation of dielectric constant for soil with its moisture content. This is presented in Figure 1 for a clay loam soil at the wavelengths to be considered in this paper, 21 [*Lundien*, 1971] and 1.55 [*Wang et al.*, 1978] cm. It can be seen in Figure 1 that the addition of water has very little effect on the dielectric properties of the soil at low-moisture contents (<10%). Presumably this

TABLE 1. Comparison of σ's From Laboratory Measurement

Surface	Effective σ, mm		Measured σ, mm	Autocorrelation Length, mm
	1-2 GHz	4.5-7.4 GHz		
1	5.0	5.0	2.7	89
2	6.0	7.5	2.2	12
3	16.0	8.0	3.6	18

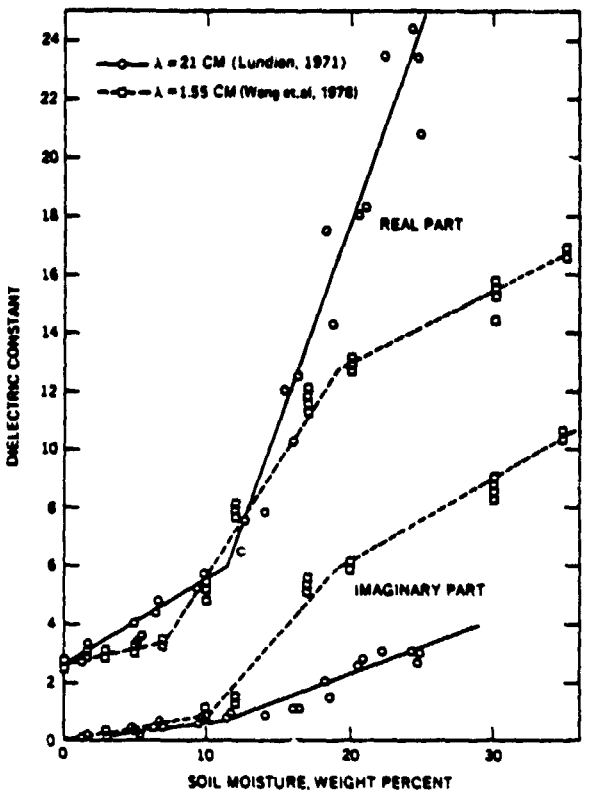


Fig. 1. Laboratory measurements of dielectric constant for a soil as a function of its moisture content. These measurements are for two clay loam soils with similar textures.

is due to the strong interaction of the water molecules with the soil particles which reduces the polarizability of the water in a thin layer around each particle. As the water content increases the water is less tightly bound and causes a greater increase in the dielectric constant for soils. The dielectric constant variation can be represented by the straight lines which are linear regression fits in these regions for the two wavelengths.

The effect of surface roughness on the emissivity of the soil as a function of soil moisture at the 21 cm wavelength is presented in Figure 2 for several values of h. These values were calculated assuming uniform moisture and temperature profiles using the Fresnel coefficient for the smooth surface case. The curves show a behavior similar to that of the dielectric constant presented in Figure 1, i.e., slow decrease of emissivity for soil moistures below about 10% and a much sharper decrease above this value. In general the effect of surface roughness is to increase the emissivity with increase being larger for the wet soil case.

The contrast in T_B between wet and dry soil is presented in Table 2 assuming a soil temperature of 300K. At $\theta = 0°$ the increase in emissivity due to roughness is given by

$$\Delta e = r_{sp} [1 - \exp(-h)] \tag{14}$$

When r_{sp} is small, i.e., for dry soils, Δe will be small, e.g., the range is 0.04 for dry soils which from Table 2 corresponds to 12K range for T_B due to surface roughness. For wet soil r_{sp} is larger and the surface roughness effect will be much larger. At the 25% moisture level, the increase in emissivity $\Delta e = 0.28$ corresponding to a 81K T_B increase.

The net effect then is to decrease to the dynamic range of the T_B change with soil moisture changes, e.g., a decrease of ΔT_B from 114K for smooth surfaces to 63K for a rough surface with $h = 0.6$. From this discussion it is seen that it is necessary to have some knowledge of the surface roughness to make an unambiguous soil moisture estimate from a T_B observation. By comparing these calculations with radiometric observations for realistic situations should yield a range of values for the parameter h.

Results and Discussion

Experimental Details

The experimental results to be discussed in this paper were obtained from a portable tower (cherry picker) platform and from an aircraft platform. The tower measurements were done at Texas A&M University from a 25 m height using 21 and 2.8 cm wavelength radiometers [*Newton*, 1977]. The radiometer measurements were supported by observation of the soil moisture and temperature at several depths down to 15 cm. The surface roughness profiles were also observed so that the values of σ^2 can be estimated.

Field Measurements

The results from the field measurements are presented in Figure 3 for fields with surfaces having three different levels of roughness: a rough plowed field, a medium rough field that had been disced, and a field that had been dragged smooth. The calculated values were obtained using the moisture and temperature profiles that were observed at the time of the

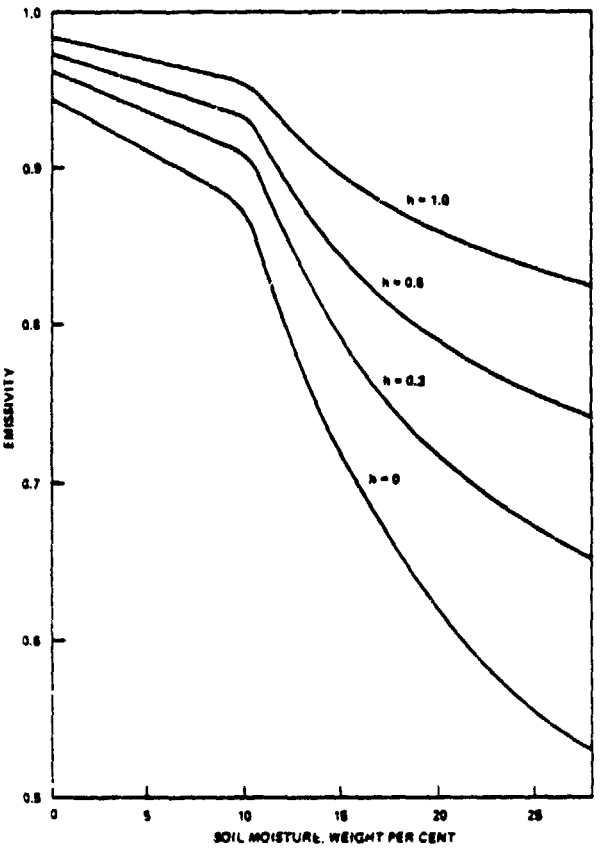


Fig. 2. Calculated values of the emissivity for a [illegible] using several values of roughness parameter h. The calculation use the dielectric constants presented in Figure 1 for the 21 cm wavelength.

TABLE 2. Calculated T_B's for a Soil Temperature of 300K

	r_{op}	$h = 0$	$h = 0.3$	$h = 0.6$	$h = 1.0$
Dry, $SM = 0\%$	0.06	282	288	294	294
Wet, $SM = 25\%$	0.44	168	201	228	249
$\Delta T_B = [T_B$ (dry) $- T_B$(wet)]		114	87	63	45

measurements. The values of h were selected to yield good agreement with the observed points for each roughness level. The observed T_B's for the smooth field are in good agreement with those listed in Table 2 for $h = 0$. In all three cases, agreement between the observed and calculated values is good at the wet and dry ends, but there are differences of 10–15K at the middle moisture levels i.e. about 15%. Recalling Figure 1, the middle moisture level is the region where T_B changes most rapidly with soil moisture, this also tends to be the region where there is the greatest uncertainty in soil moisture determination.

An additional factor contributing to the scatter at these middle moisture levels is the fact that the calculations were performed with a coherent model. Therefore there is the possibility of resonances occuring when a sharp dielectric gradient is present. We believe this caused the higher values of T_B (>285K) for the three points in Figure 3c at about 10% moisture. In the extreme case the dielectric constant changed from a value of 6 at a depth of 1.6 cm to 19.5 at 3.5 cm. A quarter wavelength in the soil for this case is given by

$$\frac{\lambda_m}{4} = \frac{\lambda_0}{4(\epsilon^{1/2})} \approx \frac{21}{4(6^{1/2})} = 2.1 \text{ cm} \qquad (15)$$

where λ_m is the wavelength in the medium. As a result constructive interference occurs causing an increase in the transmission through the surface (i.e., emissivity) for this case. To verify this hypothesis the calculations were performed as a function of wavelength from 30 to 15 cm. The maximum, 295K, occurred at 21 cm wavelength and T_B fell off to 285K at 27 and 15.5 cm wavelengths. The sharp dielectric gradient in this case was caused by the transition from values on the low moisture portion (10%) of the dielectric curve in Figure 1 at a point (20%) on the steep portion of the curve in a relatively short distance. Therefore in applying any coherent model one has to be aware of the possiblity of these resonances occurring so that false interpretations from the model can be avoided.

In Table 3 the values of $\sigma_{measured}$ for the fields are listed along with those calculated (11) using the observed values of h.

The parameter h increases with increasing roughness but does not do so as rapidly as expected from the measured σ's. For example, the ratio of the σ^2 for the rough and medium rough cases is 2.7 while the ratio of the h's is 1.7.

The effective σ's for these cases are less than the measured values. This is opposite from the situations presented in Table 1 for the laboratory measurements. This difference probably results from the different horizontal scales of the roughness in the two situations.

Aircraft Measurements

The aircraft results were obtained during flights with NASA aircraft over the Phoenix, Arizona, area and the Imperial Valley of California during March 1972 and February 1973 [*Schmugge et al.*, 1976] and flights over only the Phoenix area during March 1975 [*Schmugge*, 1976]. The aircraft altitude for these flights were 600 m in 1972 and 1973 and 300 m in 1975. On board the aircraft were microwave radiometers covering those wavelength range of 0.8–21 cm. In this paper only the results at the 21 and 1.55 cm wavelengths will be presented. The 21 cm radiometer was nadir viewing with a 15° (~1/4 radian) beamwidth, therefore its spatial resolution was approximately 1/4 the aircraft altitude. The 1.55 cm radiometer is a scanning radiometer which has an angular beam width of 2.8° (~1/20 radian). This sensor was only used on the 1972 and 1973 missions.

The aircraft flew along flight lines centered on the agricultural fields which were at least 16 hectares (40 acres) in area.

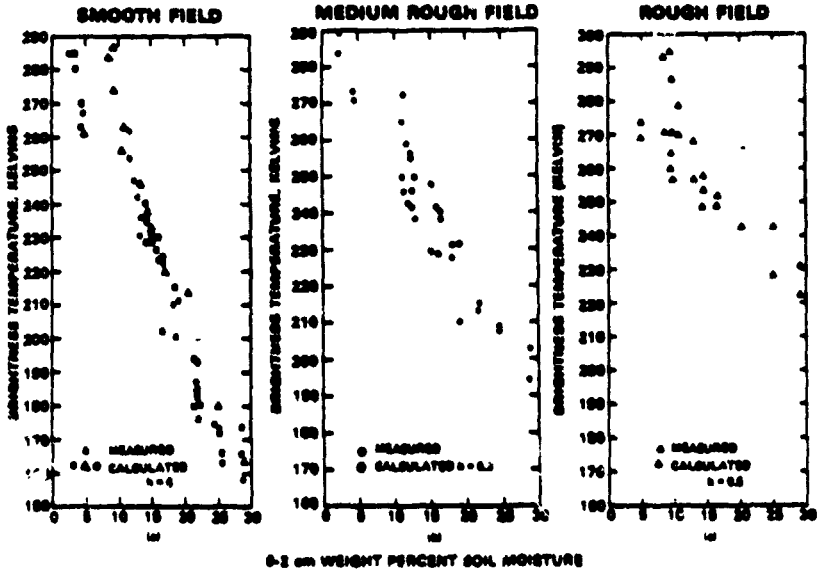


Fig. 3. Comparison of calculated and measured values of T_B at $\lambda = 21$ cm for fields with three levels of surface roughness. Measurements were made from the truck mounted at Texas A&M University [*Newton*, 1977]. In Figure 3a the calculations were done for the profiles observed in all three fields.

TABLE 3. Comparison of σ's From Field Measurements

Field	h	Effective, cm	Measured, cm
Smooth	0	0	0.9
Medium rough	0.3	0.9	2.6
Rough	0.5	1.2	4.3

These fields generally had uniform surface and moisture conditions over their total area. All the radiometer data obtained over each field were used to obtain the average brightness temperature (T_b) for the field. The soil moisture measurements were made at four locations and for several depths in each field. The values presented here are the averages for each field. For the 1975 flights soil temperature profiles were also measured. Soil textures determination were also made for the sampled fields.

Because of the range of soil texture, from sandy loams to clays, that are present at both of these sites it is necessary to account for the different water bolding capacities of these soils. This was done by normalizing the measured soil moistures to the field capacity levels (FC) for each soil. The amount of water in a soil at FC is that which remains two or three days after having been saturated and after free drainage has practically ceased. This level is determined to a large extent by the soil texture i.e. particle size composition. The value of FC for each field was estimated on the basis of the soil textures that were measured for that field [*Schmugge*, 1978].

The surface roughness characteristics were those resulting from the agricultural practices of the two areas. The dominant method of irrigation is the flooded furrow. The furrow separation was about one meter and the furrow height was about 20 cm. Superimposed on these corrugations were clods, which were generally less than 5 cm.

Plots of T_b versus the soil moisture in various layers for the 1972 and 1973 flights are presented in Figure 4. We note that the range of T_b is not as great as that expected for a smooth surface.

Calculations using the layered model [*Wilhelt*, 1978] were performed using moisture and temperature profiles measured by the personnel at the U.S. Water Conservation Laboratory at Phoenix [*Jackson*, 1973] The soil moisture and temperature profiles were observed at frequent intervals after a heavy irrigation on March 2, 1971. These data from the same area at the same time of the year were assumed to be reasonable estimates of the situations occurring during the 1972 aircraft overflights. It should be noted that the moisture and temperature profiles had been obtained from a smooth field, while the microwave radiometer resuls were obtained from rough-surfaced fields. The moisture profiles were rather uniform when wet, but dried rapidly at the surface after 5 or 6 days producing sharp moisture and dielectric constant gradients just below the surface. The calculations were performed for the early afternoon (1:30 P.M.) profiles of each day. The corresponding temperature profiles are probably quite representative of the actual situation for the 1972 flights. For the 1973 flights, occurring early in February, the temperatures were somewhat cooler. The surface temperatures as observed by the aircraft IR sensor were found to be about 15K lower than that observed during March. The February temperature profiles were then obtained from the observed March temperature profiles by adjusting the gradient to fit the observed surface temperature for February data and assuming two profiles to be equal at about 50 cm.

The dielectric constants used in the calculation were those

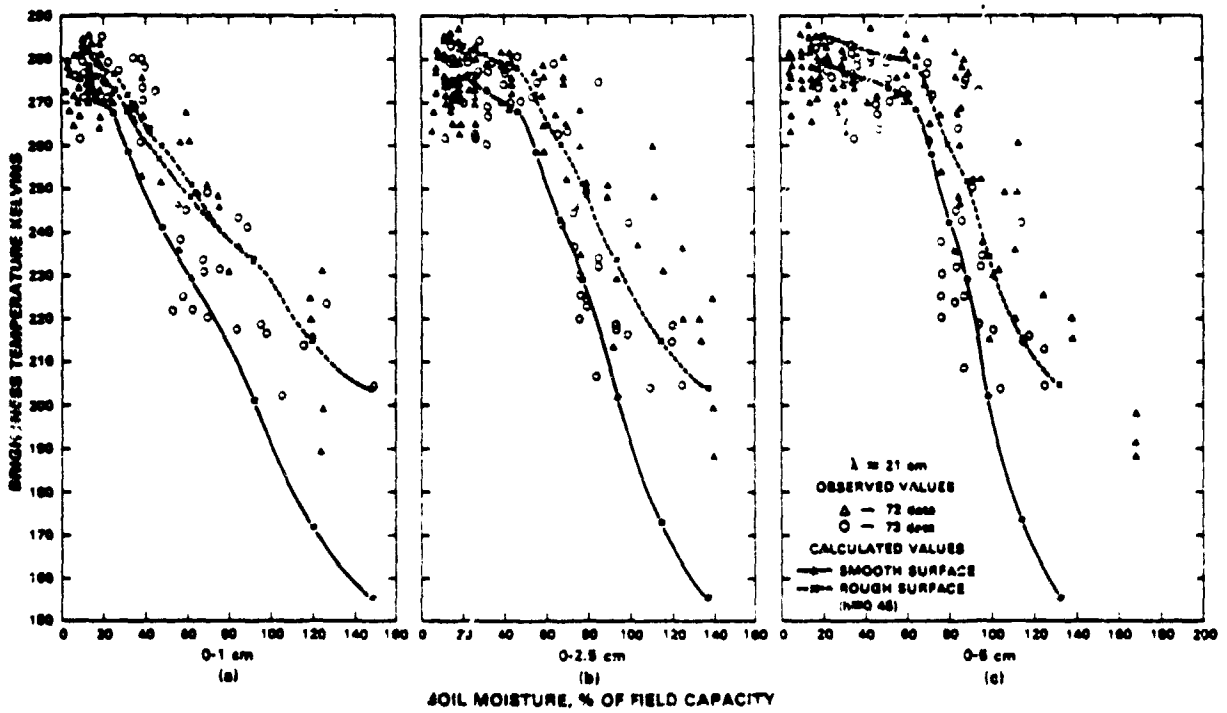


Fig. 4. Aircraft observations of T_b at λ = 21 cm during 1972 and 1973 flights over Phoenix compared with soil moisture in three depths of the soil. The two dashed curves in Figure 4a indicate the difference produced by using the 1972 and 1973 temperature profiles.

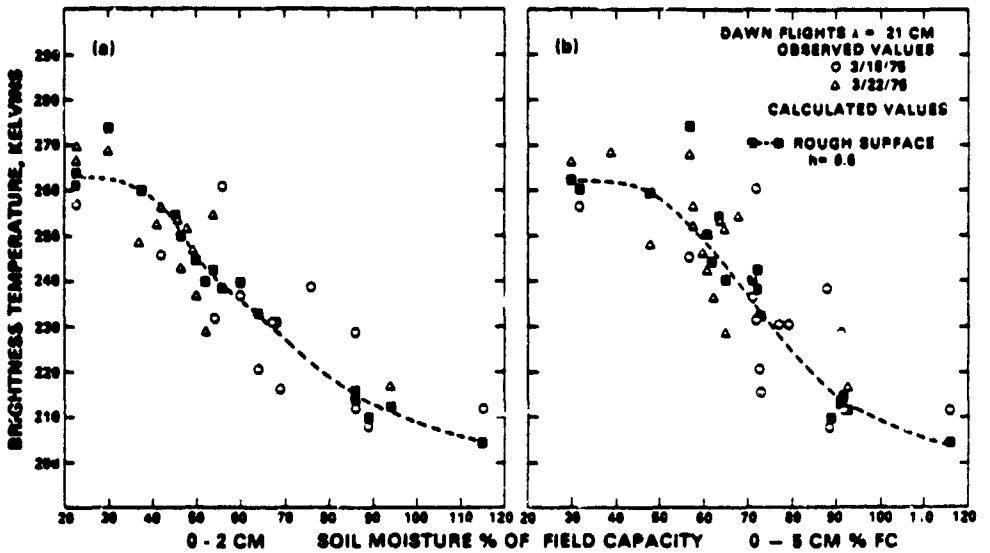


Fig. 5. Aircraft observations of T_B at $\lambda = 21$ cm during 1975 dawn (6–7 A.M.) flights of Phoenix compared with soil moisture in two depths of the soil. The curves are visual best fits to the calculated values.

presented in Figure 1. These values are for soils having texture similar to the Avondale clay loam soil at the Water Conservation Laboratory.

The solid curves in Figure 4 are the calculated values assuming a smooth surface ($h = 0$). It is clear that the aircraft T_B's do not get as low as those calculated for the smooth surface. The form of the calculated curve however does agree with the observations, i.e., little variation out to about 50% of FC and then the rapid decrease in T_B.

The dashed curves in Figure 4*a*–4*c* are for $h = 0.45$, and it is seen that the range of calculated T_B's is in good agreement the observed range and that the roughness factor has its greatest effect at the higher level of moisture as predicted by (14).

In observing the variation of T_B with soil moisture in the three layers, we note the linear decrease T_B with the soil moisture in the 0–1 cm layer, but for the 0–2.5 and 0–5 cm layers there is a region at low soil moisture in which there is little variation of T_B. Above this level there is a sharp decrease in T_B. This behavior is similar to that presented in Figure 2 and that the location of the break point for the 2.5 cm curve, at approximately 50% of FC, is in good agreement with the location presented in Figure 2, assuming a FC of 20–25% for a clay loam soil. Because of this agreement, we will assume that the radiometer is responding to the moisture variations in the top 2 or 2.5 cm.

In Figures 5 and 6 the results from the 1975 flights are presented. Figure 5 presents the results from the pre-dawn flights and Figure 6 from the midday flights. Essentially, the same dependence of T_B on soil moisture is observed for these flights as for the 1973 flights. The calculated values in this case used the moisture and temperature profiles that were measured in each field at the time of the flights. The curves are visual best

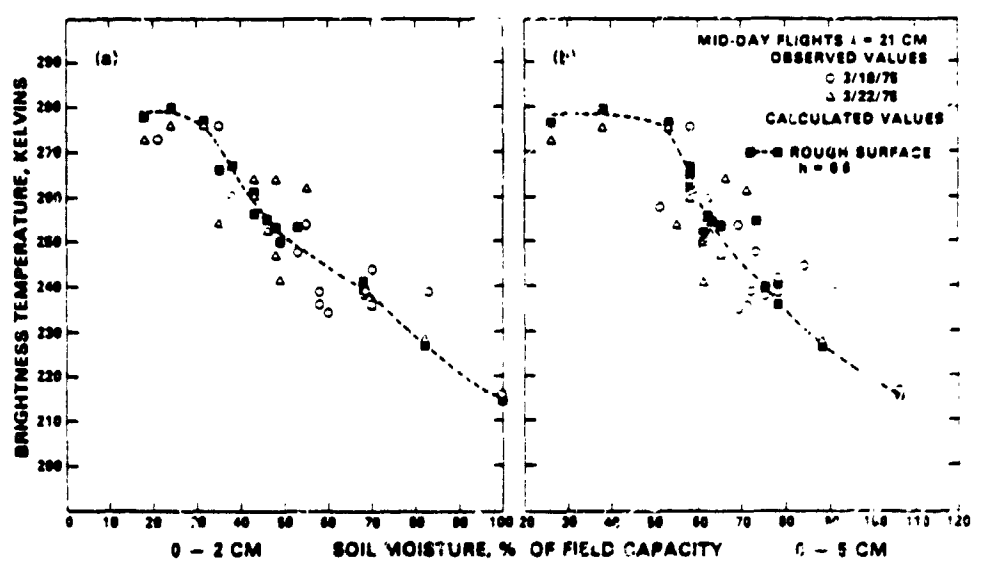


Fig. 6. Aircraft observations of T_B at $\lambda = 21$ cm during 1975 midday (1–4 P.M.) flights over Phoenix compared with soil moisture in two depths of the soil. The curves are visual best fits to the calculated values.

fits to the calculated points. Note that there is a 10–15K difference in the observed values of T_B for dry fields between the A.M. and P.M. flights due to soil temperature differences and that there is a similar difference in the calculated values.

The scatter in the calculated points about the curves is due to two causes: first, the variations in the soil moisture profiles having the same average 0–2 cm or 0–5 cm average moisture levels; and second, the variations in soil temperature. In Figure 5 the point that is the greatest distance from the curve, the one with at T_B of 275K, may be due to the resonance effect noted earlier. The large difference between the 0–2 and 0–5 cm moisture levels, almost a factor of 2, indicates the presence of a sharp moisture gradient. The observed value of T_B for this field is only 5K less than the calculated value. There is considerably more scatter in the observed values of T_B. This is probably due to uncertainty of our surface measurements. Analysis of intensively sampled field [*Bell et al.*, 1979] indicates that for a sample size of 4 the expected limit of accuracy L is 4%. Thus the true mean for the field should be within ±4% of the observed value, or approximately ±20% of FC. Most of the points in Figures 5 and 6 lie within this distance from the curve.

The best agreement is obtained in each case for $h = 0.6$. This was done by fitting the data at both the wet and dry ends. This value of h is slightly larger than the result for the 1972–1973 data. The reason for this difference is unknown, and because of the scatter in the data, may indicate the uncertainty of our estimates for the value of h.

The values of h obtained for the aircraft data generally fell in between those obtained for the medium rough and rough cases for the field measurement results (Figure 3). This indicates that for the Phoenix region none of the observed fields approached the smooth category and an assumption that agricultural fields are in the medium rough to rough (i.e., $h \geq 0.5$) may be reasonable for furture calculations at this wavelength.

The data at 1.55 cm are presented in Figure 7. They are also described by an h of 0.6; the fact that h does not scale with wavelength is indicative of the shortcomings of the model. In this case, it appears that different portions of the roughness spectrum will contribute at the different wavelengths. It is also clear that certain of our assumption concerning the roughness are violated at this wavelength.

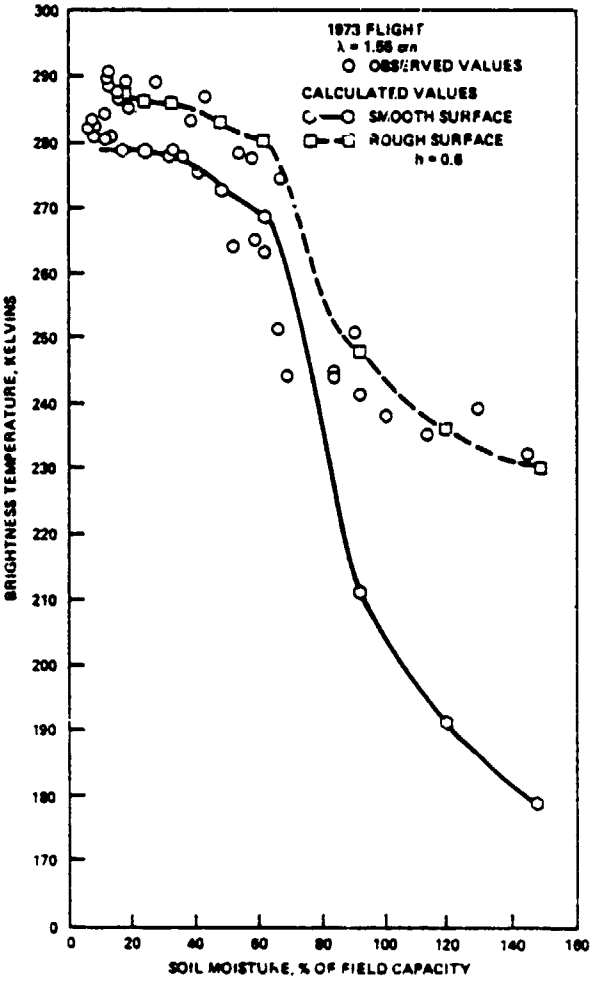


Fig. 7. Aircraft observations of T_B at $\lambda = 1.55$ cm during 1973 flight over Phoenix compared with the soil moisture in the surface cm of the soil.

Conclusions

A one parameter model for estimating the effect of surface roughness on the microwave emission from soils has been developed and compared with radiometer measurements from both tower and aircraft platforms. By a suitable choice of the parameter h, the model, when combined with a radiative transfer model for the soil, yields good agreement with the observed brightness temperatures. An effective range for the parameter was found to be from 0 for a smooth surface to 0.6 for a rough plowed surface. From the derivation of the model the parameter h is expected to be proportional to the variance of the surface height. However when compared with the measured variance for the tower measurements this dependence was not verified. A similar result was found in laboratory measurements of surface reflectivity [*Hancock*, 1976]. These latter measurements indicate that the horizontal scale of surface roughness, i.e. surface slopes, is also important in determining the magnitude of the parameter h. Because of this factor it does not appear possible to extrapolate the value of h from the measurements presented here at the 21 cm wavelength to other wavelengths and additional measurements at other wavelengths will be required to determine the dependence of h on wavelength.

At the 21 cm wavelength the value of 0.5 for h appears to be representative of the conditions observed in the aircraft data acquired over the agricultural area around Phoenix. Therefore with this model for roughness, radiative transfer model calculations should yield accurate estimates of the values of T_B for a wide range of moisture and temperature conditions.

Further work will be required to determine the wavelength dependence of h and to determine if the model can accurately predict the polarization differences expected for off-nadir observations.

Acknowledgement. The 1975 data were acquired by the members of Joint Soil Moisture Experiment which involved investigators from Texas A&M University and University of Arkansas in addition to NASA. Their cooperation is making these measurements is appreciated.

References

Ament, W. S., Towards a theory of reflection by a rough surface, *Proc. IRE*, *41*, 142–146, 1953.

Barrick, D. E., Rough surfaces, in *Radar Cross Section Handbook*, vol. 2, edited by G. T. Ruck, chap. 9, Plenum, New York, 1970.

Bell, K. R., B J Blanchard, T J Schmugge, and M W. Witczak, Analysis of surface moisture variations within large field sites, *NASA Tech. Memo. 80264*, 1979

Born, M., and E. Wolf, *Principles of Optics*, chap XIII, Pergamon, New York, 1975.

Burke, W J., T. J. Schmugge, and J F. Paris, Comparison of 2 8 and 21-cm microwave radiometer observations over soils with emission model calculations, *J Geophy Res., 84*, 287–294, 1979

Chandrasekhar, S., *Radiative Transfer*, Dover, New York, 1960.

England, A. W., Relative influence upon microwave emissivity and fine-scale stratigraphy, internal scattering and dielectric properties, *Pure Appl Geophys., 114*, 287–289, 1976

Hancock, G. D., Broad spectrum microwave measurements of rough discontinuous surfaces for the determination of moisture content, report, Dep of Elec. Eng., Univ of Arkansas, Fayetteville, 1976

Jackson, R. D., Diurnal soil-water content changes during drying, Field Solar Water Regime, SSSA *Spec Publ. 5*, pp. 37–56, Soil Sci. Soc of Amer., Madison, Wis., 1973

Lundien, J R., Terrain analysis by electromagnetic means, *Tech Rep 3-693, Rep 5*, U S Army Waterways Exp Sta., Vicksburg, Miss., 1971

Newton, R W., Microwave remote sensing and its application to soil moisture detection, *Tech. Rep RCS-81*, Remote Sensing Center, Tex. A&M Univ., College Station, 1977 (Available from University Microfilms, order no 77-20-398)

Njoku, E. G., and J A. Kong, Theory for passive microwave remote sensing of near-surface soil moisture, *J Geophys. Res., 82*, 3108–3117, 1977

Schmugge, T J., Preliminary results from the March 1975 soil moisture flight, *NASA X-913-76-216*, Goddard Space Flight Center, Greenbelt, Md., 1976

Schmugge, T., Effect of soil texture on the microwave emission of soils, in press, 1978

Schmugge, T. J., T T Wilheit, W Webster, and P Glorsen, Remote sensing of soil moisture with microwave radiometers II, *NASA Tech. Note TN D-8321*, 1976

Sung, C. C., and W D Eberhardt, Explanation of the experimental results of light backscattered from a very rough surface, *J Opt. Soc Amer., 68*, 323–328, 1978

Tolstoy, I., and C S. Clay, *Ocean Acoustics*, McGraw-Hill, New York, 1966

Tsang, L., E. Njoku, and J A. Kong, Microwave thermal emission from a stratified medium with non-uniform temperature distribution, *J. Appl. Phys., 46*, 5127–5133, 1975.

Waite, W. P., K. R. Cook, and B B Bryan, Board spectrum microwave system for remotely measuring soil moisture content, *Publ. 18*, Water Resour. Res Center, Univ of Arkansas, Fayetteville, 1973

Wang, J., T Schmugge, and D Williams, Dielectric constants of soil microwave frequencies, II, *NASA Tech Pap TP-1238*, 1978

Wilheit, T., Radiative transfer in a plane stratified dielectric, *IEEE Trans. Geosci. Elec., GE-16*, 138–143, 1978

Wu S. T., and A. K Fung, A noncoherent model for microwave emissions and backscattering from sea surface, *J Geophys. Res.*, 77, 5917–5929, 1972.

(Received October 16, 1978.
revised May 4, 1979;
accepted May 10, 1979)

Paper 77

NASA Technical Memorandum 80582 October 1979

MULTISENSOR ANALYSIS OF HYDROLOGIC FEATURES IN THE WIND RIVER RANGE, WYOMING WITH EMPHASIS ON THE SEASAT SAR

James L. Foster

Dorothy K. Hall

ABSTRACT

Synthetic Aperture Radar (SAR) imagery of the Wind River Range area in Wyoming is compared to visible and near-infrared imagery of the same area. Data from the Seasat L-Band SAR and an aircraft X-Band SAR are compared to Landsat Return Beam Vidicon (RBV) visible data and near-infrared aerial photography and topographic maps of the same area. Visible and near-infrared data provide more information than the SAR data when conditions are optimum. However, the SAR penetrates clouds and snow, and data can be acquired day or night. Drainage density detail is good on SAR imagery because individual streams show up well due to riparian vegetation causing higher radar reflections which result from the "rough" surface which vegetation creates. In the winter image, the X-Band radar data show high returns resulting from cracks on the lake ice surfaces. High returns are also evident in the L-Band SAR imagery of the lakes due to ripples on the lake surfaces induced by wind. It is concluded that utilization of multispectral data (visible, near-infrared and microwave (radar)) would optimize analysis of hydrologic features.

Paper 78

NASA Technical Memorandum 80564 September 1979

INTERIM CALIBRATION REPORT FOR THE SMMR SIMULATOR

P. Gloersen and D. Cavalieri

ABSTRACT

The calibration data obtained during the FALL 1978 NIMBUS-G underflight mission with the Scanning Multichannel Microwave Radiometer (SMMR) simulator on board the NASA CV-990 aircraft have been analyzed and an interim calibration algorithm developed. Data selected for this analysis consisted of in-flight sky, first-year sea ice, and open water observations, as well as ground-based observations of fixed targets while varying the temperatures of selected instrument components. For most of the SMMR channels, a good fit to the selected data set was obtained with the algorithm.

Paper 79

TIME-DEPENDENCE OF SEA-ICE CONCENTRATION AND MULTIYEAR ICE FRACTION IN THE ARCTIC BASIN

P GLOERSEN, H J ZWALLY, A T C CHANG, and D K HALL
NASA, Goddard Space Flight Center, Greenbelt, Md., U.S.A.

W. J CAMPBELL
United States Geological Survey, Tacoma, Wash., U.S.A.

and

R. O RAMSEIER
Canadian Department of Fisheries and Environment, Ottawa, Canada

(Received 30 September, 1977)

Abstract. The time variation of the sea-ice concentration and multiyear ice fraction within the pack ice in the Arctic Basin is examined, using microwave images of sea ice recently acquired by the Nimbus-5 spacecraft and the NASA CV-990 airborne laboratory. The images used for these studies were constructed from data acquired from the Electrically Scanned Microwave Radiometer (ESMR) which records radiation from earth and its atmosphere at a wavelength of 1.55 cm. Data are analyzed for four seasons during 1973–1975 to illustrate some basic differences in the properties of the sea ice during those times. Spacecraft data are compared with corresponding NASA CV-990 airborne laboratory data obtained over wide areas in the Arctic Basin during the Main Arctic Ice Dynamics Joint Experiment (1975) to illustrate the applicability of passive-microwave remote sensing for monitoring the time dependence of sea-ice concentration (divergence). These observations indicate significant variations in the sea-ice concentration in the spring, late fall and early winter. In addition, deep in the interior of the Arctic polar sea-ice pack, heretofore unobserved large areas, several hundred kilometers in extent, of sea-ice concentrations as low as 50% are indicated.

1. Introduction

The delineation of the extent, structure, concentration, and motion of sea ice in polar regions by microwave radiometry has been discussed in earlier papers (Wilheit *et al.*, 1972; Gloersen and Salomonson, 1975; Gloersen *et al.*, 1973, 1974a, b, 1975a, b; Campbell *et al.*, 1974, 1975, 1976, 1978; Ramseier *et al.*, 1974, 1975). In this paper, emphasis is placed on examining the time variation of the sea-ice concentration and multiyear ice fraction within the pack ice in the Arctic Basin, using microwave images of sea ice recently acquired by the Nimbus-5 spacecraft and the NASA CV-990 airborne laboratory.

The principal instrument used for these studies is the Electrically Scanned Microwave Radiometer (ESMR) which records radiation from Earth and its atmosphere at a wavelength of 1.55 cm. Such an instrument is carried on both the CV-990 aircraft (ESMR-A/C) and the Nimbus-5 satellite (ESMR-5), the experimental techniques have been described in detail previously. Since the Rayleigh-Jeans approximation applies in the wavelength regime of the ESMR and the range of sensible temperatures encountered, the received power is linearly dependent on

Boundary-Layer Meteorology 13 (1978) 339–359 *All Rights Reserved*
Copyright © 1978 by D. Reidel Publishing Company, Dordrecht, Holland

the brightness temperature of the target. Thus, radiances are described in terms of brightness temperatures rather than received power.

Data are analysed for four seasons during 1973–1975 to illustrate some basic differences in the properties of the sea ice during those times. In the Spring and Fall of 1975, the NASA CV-990 airborne laboratory made three poleward transects on 16 April, 19 April, and 22 October, 1975 in the area north of 70° N between 37° W and 167° W. Corresponding satellite data from these times are also presented and discussed. In addition, spacecraft data for August 1975 are compared with corresponding CV-990 data obtained during the Main Arctic Ice Dynamics Joint Experiment (Main AIDJEX) to illustrate the applicability of passive microwave remote sensing for monitoring the time-dependence of sea ice concentration (divergence).

2. Interpretation of the Microwave Signatures of Sea Ice

The radiative-transfer equation applicable to passive microwave observations at a given wavelength within the sea-ice canopy is as follows:

$$T_B = \varepsilon T_s e^{-\tau} + T_1 + (1-\varepsilon) T_2 e^{-\tau} + (1-\varepsilon) T_{sp} e^{-2\tau}, \tag{1}$$

where

$$T_1 = \int_0^{\tau} T(z)\, e^{-\tau + \tau'(z)}\, \mathrm{d}\tau'(z)$$

and

$$T_2 = k \int_0^{\tau} T(z)\, e^{-\tau'(z)}\, \mathrm{d}\tau'(z)$$

in which T_B is the observed brightness temperature; ε is the emissivity of the surface, which depends on wavelength as well as on various surface variables; T_S is the sensible temperature of the surface; τ is the total atmospheric opacity which is also wavelength dependent; $T(z)$ is the sensible temperature in the atmosphere at a height z; $\tau'(z)$ is the atmospheric opacity from the surface to a height z; and T_{sp} is the average temperature of free space, also wavelength dependent, but a small contribution at all the wavelengths considered here.

In Equation (1), the first term on the right represents the contribution to the observed brightness temperature of the surface as seen through the atmosphere, the second is the upwelling atmospheric radiation, the third is the downwelling atmospheric radiation as observed after reflection at the surface and transmission back through the atmosphere, and the fourth is radiation from space reflected back from the surface and attenuated by two passes through the atmosphere.

The constant k in the expression for T_2 under Equation (1) is included to account for the case of diffuse reflection from rough surfaces in an approximate manner, because strictly speaking the integral should be carried out over the entire hemisphere rather than along the line-of-sight direction, z. The value of k will depend

both on the distribution assumed for the absorption, τ, in the sky and the type of diffuse reflection assumed at the surface.

We assume that a suitable approximation for T_1 and T_2 in the polar regions is given by

$$T_1 = k^{-1}T_2 = \tau T_A , \tag{2}$$

where T_A is a weighted average of $T(z)$ in the lower troposphere. Using the index i to denote different wavelengths and polarizations, equation (1) then becomes

$$T_{Bi} = \varepsilon_i T_s + \tau_i[(1+k)T_A - \varepsilon_i(kT_A + T_s)] + (1-\varepsilon_i)T_{sp} , \tag{3}$$

where terms containing $\tau_i T_{sp}$ have been neglected.

Under the assumption that the Rayleigh extinction coefficients (Gaut *et al.*, 1970) apply in the polar regions, the opacities, τ_i, vary inversely as the square of the wavelength λ^{-2} and Equation (3) may be further approximated as:

$$T_{Bi} = \varepsilon_i(T_s - T_{sp}) + \lambda_i^{-2}A + T_{sp} \tag{4}$$

in which A is assumed to be independent of λ, to first order, since variations of the product $\tau_i\varepsilon_i$ are of second order; also T_A and T_s are independent of wavelength and k is assumed to be independent of wavelength.

The sea-ice surface of the Main AIDJEX area was found to consist of four major categories with significantly distinct microwave characteristics; (1) open water; (2) first-year sea ice with snow cover (FY); (3) thin (but thicker than one microwave wavelength) first-year ice without snow cover (FT); and (4) multiyear sea ice (MY). Taking these into account, using $T_{sp} = 3$ K, and noting that the open sea temperature is 271 K, Equation (4) becomes:

$$T_{Bi} - 3 = (271-3)\varepsilon_{0i}C \tag{5a}$$

$$+ \varepsilon_{FYi}(1-C)(T_{ice} - 3) \tag{5b}$$

$$- (\varepsilon_{FYi} - \varepsilon_{MYi})F(1-C)(T_{ice} - 3) \tag{5c}$$

$$- (\varepsilon_{FYi} - \varepsilon_{FTi})X(1-C)(T_{ice} - 3) \tag{5d}$$

$$+ \varepsilon_{FTi}X(1-C)(t_{FT} - 3) \tag{5e}$$

$$+ \lambda_i^{-2}A . \tag{5f}$$

The term (a) on the right-hand side of Equation (5) is the contribution from the open water which is assumed to be at the freezing point, 271 K, ε_{0i} is the emissivity of calm open water and C is the percent open water. Terms (b) through (e) are the contributions of the various forms of sea ice to the radiation. They consist of the emissivities (ε_{xxi}) of first-year (FY), thin first-year without snow cover (FT), and multiyear (MY) sea ice; the fraction of the area covered with sea ice, $(1-C)$; the fraction of the ice-covered area that is multiyear, F; the fraction of the ice-covered area that is thin first-year without snow cover, X; the sensible temperature of the

FY and MY, T_{ice}; and the difference between the sensible temperature of the FT and T_{ice}, t_{FT}. The terms (5b)–(5e) have been arranged so that under certain circumstances some of them will drop out, e.g., when $\varepsilon_{MYi} \simeq \varepsilon_{FTi}$ as is the case for vertical polarization at all of the wavelengths discussed (see data presented in Campbell *et al.*, 1977).

Equation (5) contains the following ten compound variables: C, $T_{ice}-3$, $C(T_{ice}-3)$, $F(T_{ice}-3)$, $X(T_{ice}-3)$, $X(t_{FT}-3)$, $FC(T_{ice}-3)$, $XC(T_{ice}-3)$, $XC(t_{FT}-3)$, and A, consisting of various products of the six variables, C, X, F, T_{ice}, t_{FT}, and A. In principle, these ten compound variables can be determined analytically from ten different microwave brightness temperatures by solving ten simultaneous linear equations of the form of Equation (5) if the coefficients are known, e.g., with the ten channels of T_B's to be obtained with the Scanning Multichannel Microwave Radiometer (SMMR) on board the Nimbus G and Seasat A spacecraft, both due to be launched in 1978. In practice, however, such solutions often result in large coefficients and alternating signs which give rise to noisy solutions when taking into account radiometer noise. This has been demonstrated in the present situation by using a set of best-judgment emissivities. It is unlikely that improved solutions would result from more accurate values of the emissivities. A sequential approach to the solution is currently under investigation in which fewer radiometer channels are used to retrieve a given variable, determination of some variables requires the prior retrieval of others, and the coefficients are obtained from regression analysis of an experimental parameter/brightness-temperature data set. The initial procedure used for this sequential retrieval consists of determining T_{ice} and C from the 4.6-cm horizontal and vertical channels under the following assumptions: (1) term (3) of Equation (5) is zero; (2) X is sufficiently small so that terms (5d) and (5e) can be neglected compared to term (5a); and (3) term (5f) can be dropped since there is negligible atmospheric interference at the 4.6-cm wavelength. Next, using these values for T_{ice} and C, F is determined from the difference between the Equations (5) written for the 1.6- and 2.8-cm wavelength horizontal channels (terms (5b), (5d), and (5e) drop out and term (5f) is neglected). Then, T_{ice}, C, and F are inserted in the Equations (5) written for the difference between the two 0.8-cm channels to permit solving for X (term (5f) drops out). Finally, A is determined from the differences between the Equations (5) for the 0.8- and 1.4-cm vertical polarization channels, using the values determined for T_{ice}, C, F, and X and is used as a measure of atmospheric interference.

The spatial resolution of such determinations will in general be about 150 km with the SMMR, but better spatial resolutions can be obtained by making additional assumptions as to the variation of the variables within the 150-km cell. If, for instance, it is assumed that T_{ice}, F, and A are essentially constant within that cell (F is zero around Antarctica), then X and C may be calculated using the 0.8-cm channels at a spatial resolution of about 30 km.

To summarize the state of knowledge of sea-ice signatures at this point, the emissivity is known to be determined in general by both the surface and volume

properties of the ice. FY sea ice appears to be optically thick with the emissivity being given by:

$$\varepsilon_{FYI} = 1 - r_{FYI}, \tag{6}$$

where r_{FYI} is the surface reflectivity obtainable from the Fresnel reflection coefficients (Baker, 1970).

Although sea ice and snow cover include air pockets, a reasonable fit to the data is obtained by assuming that the index of refraction, n, of solid ice, approximately 1.78, should be used in the Fresnel relations at all wavelengths. At the SMMR incidence angle of 50°, these yield values of 0.16 and 0.025 for $r_\parallel$ (parallel to the surface) and $r_\perp$ (orthogonal with $r_\parallel$ and the line of sight), respectively, corresponding to emissivity values of 0.84 and 0.97 for FY sea ice. Surface roughness modifies this property, however, and appears to be important at the 0.81-cm SMMR wavelength according to field measurements during the Main AIDJEX which indicated much smaller polarization effects at that wavelength for FY, FT, and MY sea-ice types.

The empirical knowledge of microwave properties of MY ice at these wavelengths is limited to the observation that they are the same as FY at wavelengths of 4.6 cm and longer, about 5 K lower at nadir than FY at 2.8 cm, about 20 K lower at nadir than FY at 1.55 cm, and about 50 K lower at nadir than FY at 0.81 cm, all at a surface temperature of 250 K (Wilheit *et al.*, 1972; Gloersen *et al.*, 1973, 1974a). On this basis, and using 0.92 for the nadir emissivity of FY at all wavelengths but 0.81 cm, the following equations apply for the MY:

$$\begin{aligned}
\varepsilon_{2.8N} &= \frac{0.92 \times 250\ \text{K} - 5\ \text{K}}{250\ \text{K}} = 0.90 \\
\varepsilon_{1.5N} &= \frac{0.92 \times 250\ \text{K} - 20\ \text{K}}{250\ \text{K}} = 0.84 \\
\varepsilon_{0.8N} &= \frac{0.95 \times 250\ \text{K} - 50\ \text{K}}{250\ \text{K}} = 0.75\,.
\end{aligned} \tag{7}$$

In order to estimate the MY emissivities at an incidence angle of 50°, the surface contribution at nadir $(1-r)$ must be removed, the remaining volume effect multiplied by 1.5 (the increase in slant range), and the surface contributions at 50°, $(1-r_\parallel)$ and $(1-r_\perp)$, must be added to the values shown in Equation (7). The emissivities so deduced are listed in Table I.

The bracketed values in this table represent measurements at the surface during the Main AIDJEX and are subject to significant variation according to the degree of surface roughness.

Also shown in Table I, for completeness, are values for the emissivities of FT sea ice, inferred from measurements taken at the AIDJEX site (Campbell *et al.*, 1978). Again, the bracketed values are based on actual radiance and sensible surface-temperature measurements; the other values are extrapolated from the values

TABLE I

λ, Polarization	ε_{FYi}	ε_{MYi}	ε_{FTi}
4.6 cm, (Nadir)	0.92	0.92	—
4.6 cm, parallel	0.84	0.84	0.78
4.6 cm, perpendicular	0.97	0.97	0.96
2.8 cm, (Nadir)	0.92	0.90	
2.8 cm, parallel	0.84	0.81	[0.78]
2.8 cm, perpendicular	0.97	0.94	[0.96]
1.6 cm, (Nadir)	0.92	0.84	
1.6 cm, parallel	0.84	0.73	0.78
1.6 cm, perpendicular	0.97	0.86	0.96
0.8 cm, (Nadir)	[0.95]	[0.75]	
0.8 cm, parallel	[0.95]	[0.64]	[0.87]
0.8 cm, perpendicular	[0.97]	[0.69]	[0.96]

inferred from the measurements at the 2.8-cm wavelength. Evidently, surface roughness has a role in determining the emissivity at the 0.8 cm wavelength also for FT sea ice.

In this paper, we are constrained to single-channel observations at the Nimbus-5 ESMR wavelength of 1.55 cm. In order to separate out the various contributions to T_B, it is necessary to use *a priori* values of T_{ice} and to depend on a time sequence of observations and *a priori* knowledge to interpret T_B in terms of FY, MY, and FT sea-ice fractions. To this end, the aircraft observations were indispensable in properly interpreting the images obtained from ESMR-5 data.

3. Microwave Characteristics of Sea Ice in the Spring

ESMR-5 brightness temperatures obtained on board Nimbus-5 since its launch in December 1972 have been displayed as images by a computerized polar map projection in which the T_B's are represented by a color scale (Gloersen *et al.*, 1974a). The spring distribution of ice types, such as MY, FY, and MY + FY shown in the Beaufort Sea portion of these images, is remarkably similar from one year to the next, and to that of earlier years as measured by the aircraft ESMR (Gloersen *et al.*, 1973; Campbell *et al.*, 1976).

In this paper we present for the first time simultaneous satellite and aircraft data extending over a significant portion of the Arctic Basin. During the period of observation, 16 to 19 April, the aircraft transects, Figure 1, indicate there was an insignificant amount of open water along the transects. This permits a simplified approach to interpreting the maps in Figure 1, i.e., we may assume that C (percent open water) and X (snow-free thin-ice fraction) are nearly zero, and Equation (5) becomes:

$$T_{Bi} = \varepsilon_{FYi}(1-F)T_{ice} + \varepsilon_{MYi}FT_{ice}\,. \tag{8}$$

Fig. 1. ESMR-5 image of the north polar region on 11 April, 1975. The relationship between the brightness temperatures and the colors in the image are indicated on the scale to the right, along with an indication of the approximate multiyear sea-ice fraction, F. The transects on the image indicate the flight paths of the NASA CV-990 airborne laboratory on 16 and 19 April.

Because T_{ice} was measured along the CV-990 transects which were mostly cloud-free using an infrared radiometer operating in a band centered on 11 μm, T_B is a known linear function of F (multiyear ice fraction). Plots of T_B vs distance along the transects are shown in Figure 2. Also shown on Figures 1 and 2 is a scale from which the reader may infer approximate values of F along the transects. While we have direct measurements of C and T_{ice} only in the area of the aircraft transects, F may be inferred over much wider areas because T_{ice} is relatively constant over the ice pack and because numerous aircraft and satellite observations indicated essentially no open water in this area at this time.

The T_B profiles shown in Figure 2 were obtained from 10-s averages of the center 10 ESMR-A/C beam positions. As in a previous study (Campbell *et al.*, 1976) various ice zones can be delineated by noting the variations of the T_B profiles. For example, one sees in Figure 2 the transitions: (1) from the FY and MY pack into the predominantly MY polar pack; (2) from the FY and MY pack into the shear zone; (3) from the shear zone into the near-shore ice containing both FY and MY; and (4) between the near-shore ice and shore-fast ice.

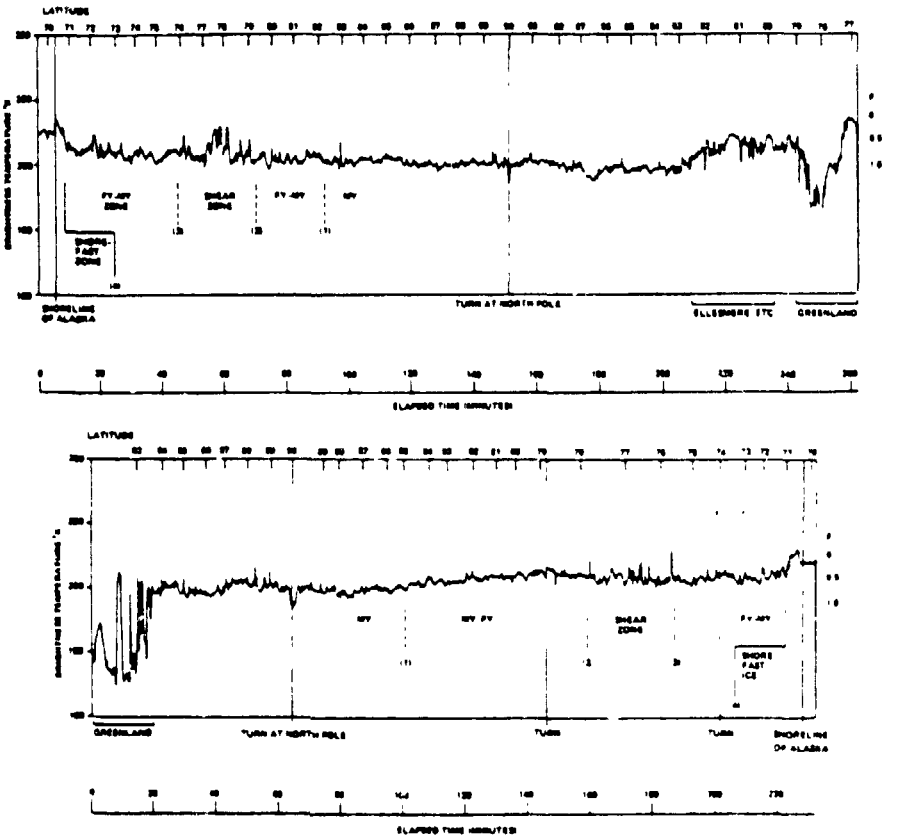


Fig. 2. ESMR-A/C brightness temperatures obtained on board the NASA CV-990 airborne laboratory on 16 and 19 April, 1975 along the transects indicated in Figure 1. The zonal characteristics (Campbell *et al.*, 1976) of the Arctic Basin sea ice are delineated.

4. Microwave Characteristics of Sea Ice in Late Summer

The radiometric distinction of FY and MY sea ice is difficult in the melting season because the freeboard layers of both have high moisture content and random crystal orientation (Campbell *et al.*, 1976) giving rise to high values of the imaginary part of the index of refraction. When such conditions are obtained uniformly, the determination of C is much simpler than under frozen surface conditions, however, because $\varepsilon_{FY_i} \simeq \varepsilon_{MY_i}$ and $T_{ice} \simeq T_{FT} \simeq 271$ K and Equation (5) becomes:

$$T_{B_i} = 268[\varepsilon_{FY_i} + (\varepsilon_{0_i} - \varepsilon_{FY_i})C + (\varepsilon_{FT_i} - \varepsilon_{FY_i})X - (\varepsilon_{FT_i} - \varepsilon_{FY_i})XC] + \lambda_i^{-2}A + 3 . \quad (9)$$

With aircraft observations, the spatial resolution is sometimes high enough to permit an additional major simplification: the FY and FT ice types are separable for some resolution cells; for these, C and X are restricted to values of 0 or 1, and Equation (9) becomes the basis for distinguishing between the two types.

The situation is probably more complex than this, however, because various phenomena occur at the ice surface during the summer. The ice surface can change rapidly from one containing extremely small amounts of free water to one in which all ice crystal surfaces are wet. According to microwave measurements of an alpine snowpack by Edgerton *et al.* (1971), and snowpack model computations (Chang and Gloersen, 1975), the emissivity of snow increases rapidly with the onset of melting and then decreases slowly as the free water content increases.

For three days prior to the first AIDJEX Phase II ESMR-A/C mapping mission on 18 August, the air temperature was usually below freezing and the ice surface was frozen as was the surface of the melt ponds, which had 2-3 cm of refrozen slush with slush underneath. The maximum T_B was observed to drop steadily by about 20 K ± 7 K, larger than the calibration uncertainty of the ESMR-A/C (Campbell *et al.*, 1978). Fortunately, detailed observations of the ice surface within the AIDJEX test area during the summer were made by Hanson (in press).

In Figure 3 for the period 11 August to 4 September we show the state of the surface in and around the AIDJEX main camp (as observed by Hanson), the 2-m air temperatures at all AIDJEX camps, the time of each CV-990 flight, and the observing periods for the ESMR-5 images (Figures 4 and 6). According to the surface observations, the surface-layer moisture steadily increased during the period of the three CV-990 flights, in agreement with the observed decreasing trend of the T_B's observed with the ESMR-A/C. It is noteworthy that the moisture content was increasing steadily even though the 2-m air temperature was fluctuating about the freezing point. There are no measurements to indicate how cool the T_B's became due to progressive ice-surface melting after the CV-990 flight. Of course, there is a limit to how much free water can remain in the surface layer, because the percolation through the layer and into the melt ponds occurs rapidly. However, because the last CV-990 flight occurred immediately following the largest seasonal melt increment (Figure 3), we assume the 20 K decrease in T_B due to increasing surface-layer moisture was the maximum. This change in T_B is significantly less than that observed with ESMR-5 for parts of the Arctic basin during this period.

It appears from the above that the minimum T_B of the moist sea ice is about 230 K. This is also the signature of dry MY just below the melt point (Table I). Therefore, any T_B appreciably lower than 230 K can be due neither to moist sea ice nor dry MY. Appreciably lower T_B's can be caused by either large areas of open water within the pack, due to ice divergence, or by extensive areas of open melt ponds. Of these two possibilities, we believe that the former mechanism is responsible for the large areas of low T_B observed in the summer ice pack shown in the ESMR-5 images (Figures 4 and 6).

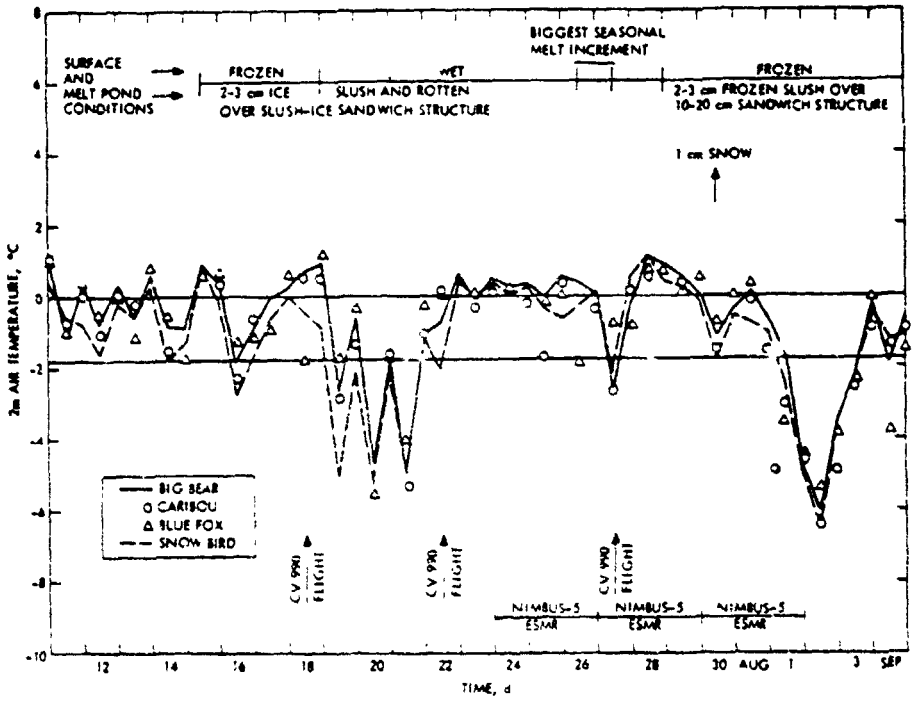


Fig. 3. Physical conditions at the AIDJEX campsites during the period 10 August to 4 September, 1975. In terms of microwave properties of the ice, the comments along the top of the figure on the physical condition of the ice are of greater importance than the air temperatures indicated by the data points and curves.

The large areas of low T_B appear as green ($T_B \sim 192$ K) zones deep within the Arctic pack, shown in the ESMR images for 24, 27, and 30 August, 1975 (Figures 4 and 6). Also shown on Figure 4 is a transect from the shoreline through and beyond the AIDJEX test area along which T_B data from the ESMR have been plotted vs distance in Figure 5 for the three different time periods during August shown in Figures 4 and 6. These data show that very large spatial and temporal variations in T_B occur. This is the first time we have succeeded in observing such pronounced large-scale variations in T_B at time scales as short as several days. This is also the first time T_B changes of such magnitude within the Arctic Basin ice pack have been observed. These changes exceed those observed for the ice surface itself with the ESMR-A/C data during that time.

The very large low T_B areas north of the Main AIDJEX array, several hundred kilometers in extent, suggest that the percent of open water is as high as 50%. It is interesting to note that the divergence in the Main AIDJEX manned station array, which at times was located at the southern extremity of this highly divergent area, was observed with ESMR-A/C (Campbell *et al.*, 1978) to have the same temporal

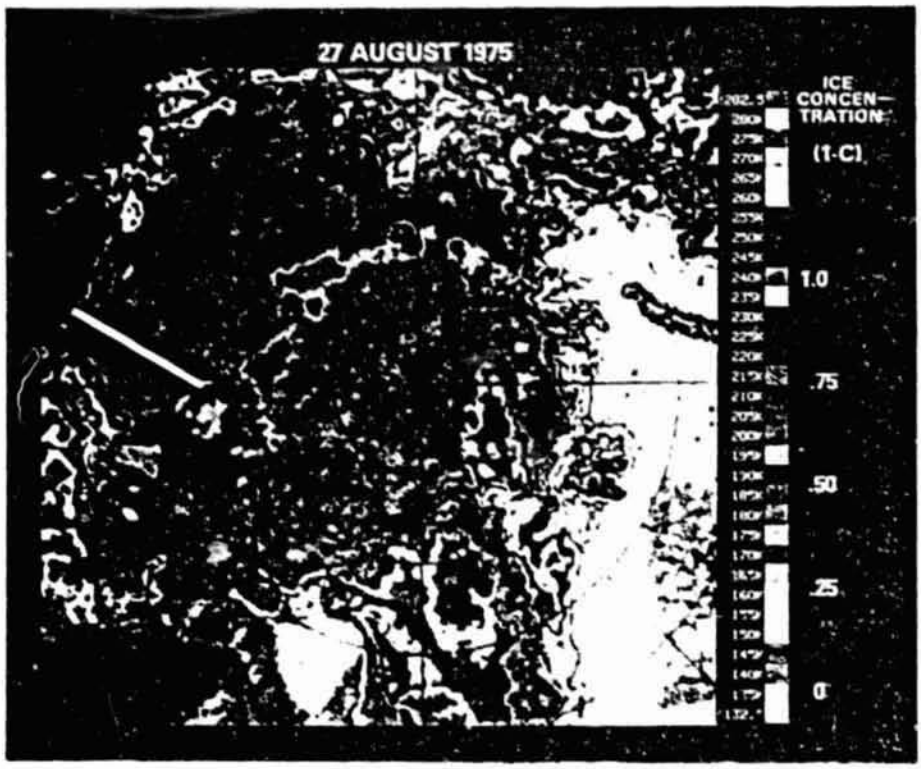


Fig. 4. ESMR-5 image of the north polar region on 27 August, 1975. An approximate sea-ice concentration scale is shown alongside the brightness-temperature/color scale to the right of the figure. A transect is shown on the image along which the time variation of the ESMR-5 brightness temperature was studied (see Figure 5).

variation as that observed with ESMR-5 (Figure 5). Using ESMR-A/C images of the main AIDJEX test area it is found that the ice within the AIDJEX triangle was consolidated on 18 August, diverging on 22 August, and more divergent on 27 August, on which occasion the percent open water in the northeast corner of the triangle is estimated from the aircraft data to be about 20% compared to 30% as deduced from ESMR-5 data (Figure 5). The discrepancy is attributed to the drop in sea-ice microwave emissivity in this period due to the increasing amount of free water in the surface layer and open-melt ponds. It is noteworthy (in Figure 5) that the divergences in the AIDJEX triangle and the region around 77° N, 135° W were out of phase and the maximum divergence in the northern region was greater.

Returning now to the other phenomenon which may significantly lower the T_B of sea ice, i.e., the localized thawing of melt ponds, evidence is presented that these effects are of secondary importance. It was observed that in the AIDJEX area the percentage of the ice surface covered with melt ponds was ca. 30%. The T_B of an

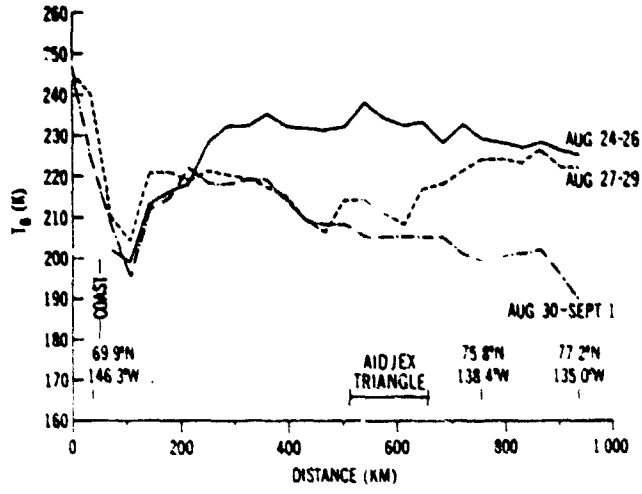


Fig. 5 Illustration of the rapid fluctuations in brightness temperature with time in late August, 1975 along the transect indicated in Figure 4. These fluctuations are attributed to rapid variations in the sea-ice concentration.

ice-free melt pond is 130 K, which is the same as for open sea water. Therefore, if the melt ponds were ice-free in an area that was 30% ponded, the decrease in T_B compared to the frozen state would be 32 K. When the melt ponds are completely ice-free, the surface of the ice floes they rest upon are moist. Based on the average of the values shown earlier, the average T_B of sea ice is ca. 240 K. Therefore the combined effect of ice-free melt ponds and moist surface is a T_B of ca. 210 K. From the ESMR-5 images shown in Figures 4 and 6, it can be seen that the T_B's within the low T_B zones discussed above range from 175 to 190 K. Therefore, although preferential thawing of the melt pond surfaces within the low T_B zones could account for part of the decrease in T_B, these data show that the major T_B decrease must be due to open water. Past observations show that the thawing and freezing of melt ponds generally occurs over large areas because the mechanisms controlling these are associated with weather systems and the existence of extensive cloud layers. These phenomena generally have much larger dimensions than those of the low T_B areas shown in Figures 4 and 6. We have observed from a NOAA-4 image that on 30 August the entire ice pack north of the AIDJEX area to the Pole was covered by a thick stratus cloud layer. Therefore, if the melt ponds were ice-free due to back-radiation from the clouds, which frequently happens in the summer, then they should be ice-free over very large areas and not just in the smaller areas of low T_B. Also, it is very difficult to see how limited preferential pond melting can account for the rapid motion of these low T_B areas observed in the sequential ESMR-5 imagery. Therefore we conclude that the low T_B's observed within the pack north of the AIDJEX camp on 24, 27, and 30 August 1975 were due to large

Fig. 6. ESMR-5 image of the Arctic Basin for 24 and 30 August, 1975, illustrating the rapidly changing regions of low brightness temperature deep within the Arctic pack. The location of the AIDJEX triangle is indicated on the image for 30 August. These large regions of low ice concentration (dark green in appearance on the image) are attributed to regions of low sea-ice concentration.

ORIGINAL PAGE IS
OF POOR QUALITY

amounts of open water. Because the surface observations show that the melt ponds in the AIDJEX area at this time were not ice-free, the percentage of open water in these green areas is ca. 50–60%.

Discovery of such large amounts of open water within the Arctic Basin is significant. Although aircraft observations (Wittman and Schule, 1966) suggested that ice concentrations as low as 90% exist at times within the polar ice pack, we believe that this is the first reported case of such large areas of ice concentration as low as 50% occurring within the pack in the Arctic Basin.

5. Microwave Characteristics of Sea Ice in Autumn

Autumn gives rise to microwave images similar to those in Spring in the sense that changes in T_B are relatively slow, at least during a two-week time period centered on 20 October, 1975, Figure 7. Again, a reasonable deduction is that such slow changes indicate values of C generally near zero as occurred during the transect by the CV-990 on 22 October, 1975 to 81° N, Figure 7. A comparison of Figures 1

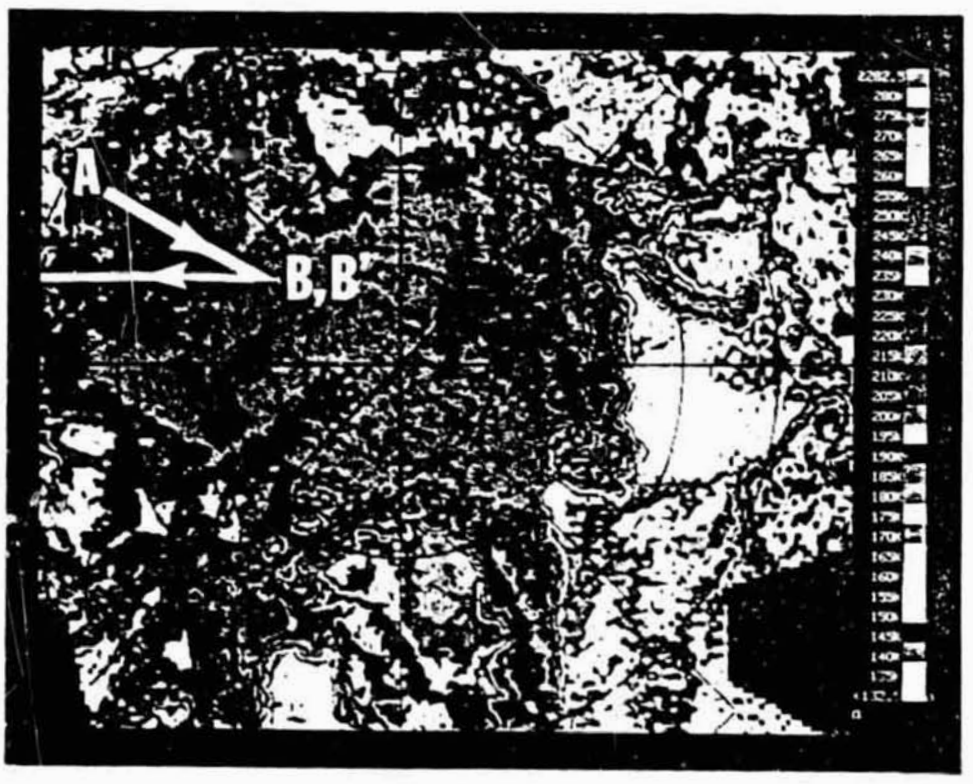


Fig. 7. ESMR-5 image of the north polar region for 20 October, 1975, indicating the CV-990 flight path for 22 October (see Figure 8).

and 7 indicates a distinctly different multiyear distribution for Spring and Fall. It is likely that this change resulted from both production of new ice throughout the Arctic Basin and transport of multiyear ice in the Beaufort Gyre and the Transpolar Drift Stream.

Measurements of T_B by the ESMR on board the CV-990 on the 22 October transect are shown in Figure 8. Here the T_B's of the center 11 beam positions (angular swath of ±15°) have been averaged together for 10 s, to produce a very low noise signal. Thus, the oscillations in Figure 8 are all due to the variation in the sea-ice properties. They appear to divide into three categories: (1) rapid oscillations due to individual or aggregated multiyear floes or large refrozen polynyas (new ice); (2) slower oscillations with a period of roughly 2° to 4° of latitude; and (3) an almost linear decrease in T_B towards the North Pole due to an increase in F. The large refrozen polynyas observed between 78° and 80° N, Figure 8, are noteworthy in that they indicate earlier low values of ice concentration deep within the pack.

The Nimbus-5 ESMR T_B's along the CV-990 tracks shown in Figures 7 and 8 have been plotted in Figure 9 for comparison with the CV-990 data. The T_B's are three-day average values over the periods 20–23 October and 26–29 October, 1975. The type (1) oscillations mentioned above are absent due mostly to spatial

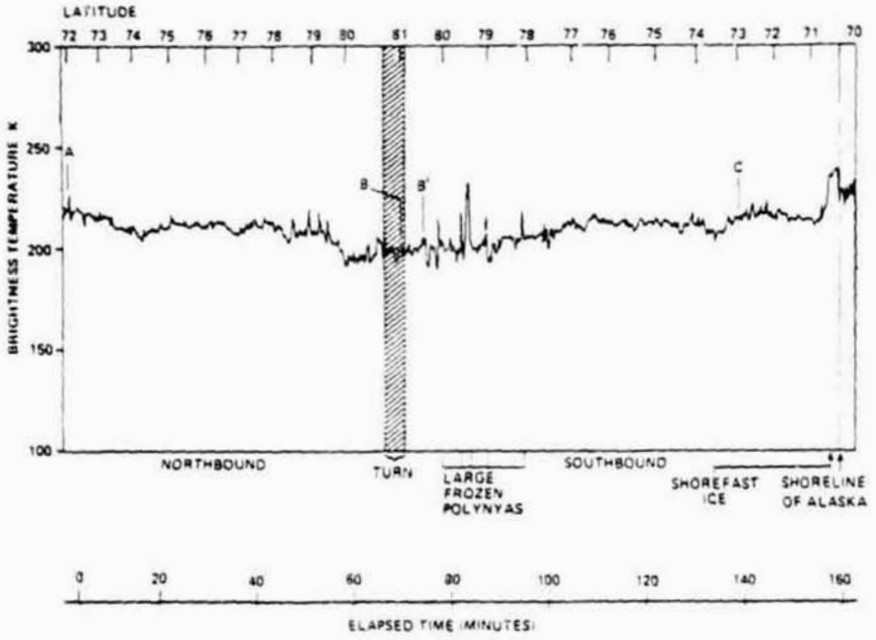


Fig. 8. ESMR-AC brightness temperatures along the transect indicated on Figure 7 illustrating zonal structure similar to that shown in Figure 7 for 22 October, 1975. Noteworthy are the large refrozen polynyas at high latitudes along the transect $B'-C$, indicating the earlier presence of large areas of open water.

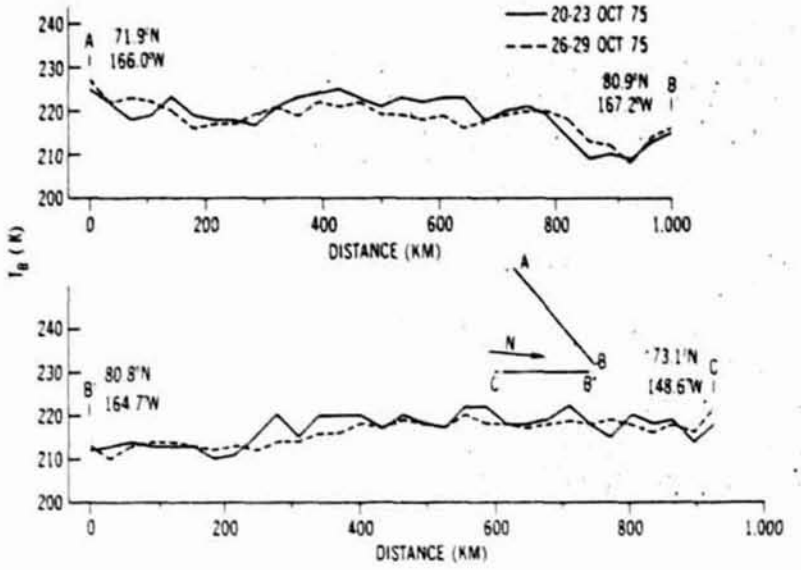


Fig. 9. ESMR-5 brightness temperatures along the CV-990 transect shown in Figure 8 for two different dates late in October, 1975, illustrating the time variation of the brightness temperature during this season. The diagram in the lower right portion of the figure illustrates the flight path of the CV-990 (see Figure 7). The letters at the end of these tracks correspond to the letters on Figures 7 and 8. The spatial variations of the brightness temperatures closely parallel those observed with the ESMR-A/C (Figure 8).

averaging. There is, however, a remarkable similarity of the type (2) and (3) variations in Figures 8 and 9 for the coincident times. On comparing the T_B's of the two time intervals in Figure 9, it may be reasonably concluded that the type (2) oscillations are attributable mostly to the variation of C, because F cannot reasonably be expected to vary so much over such a short time.

6. Microwave Characteristics of Sea Ice in Winter

The ESMR-5 polar images for the winter of 1975 were not available in the color format shown for the other seasons, but the series for the preceding winter shows evidence of considerable dynamic activity throughout the Arctic Basin. To best illustrate this, we selected the line indicated in the overall view of the Basin shown, Figure 10, for 7 January, 1974. The T_B's along this line are plotted vs distance in Figure 11a for three 3-day periods in December, 1973. The dynamic nature of the ice cover during this period suggests the presence of open water, i.e., both C and F vary with location. To illustrate this in a different way, three regions, A, B, and C, defined in Figure 11a, were selected for further analysis over a longer time, extending from the beginning of December 1973 to the beginning of February

Fig. 10. ESMR-5 brightness temperature for 7 January, 1974. The line off the north shore of Alaska indicates the location of detailed analysis of the temporal variation of the brightness temperature (see Figure 11).

1974, Figure 11b. The sea-ice concentration in the shoreline zone, *A*, evidently varies by as much as 0.3; in *B* and *C* the variation is 0.1 or less.

To display the extent of the changes, a series of eight ESMR-5 images starting on 8 December, 1973 and ending on 22 January, 1974 are shown in Figures 12 and 13. Particular attention is directed towards the changes in *C* taking place along the North American shoreline during this period. On December 8, the ice canopy contains considerable open water along the Alaskan shoreline west of Barrow; to the east, the pack is fairly well consolidated. On December 11, *C* has decreased east of Barrow while it has increased in the Mackenzie Bay/Barter Island area. On December 20, *C* has generally decreased along the entire shoreline. The subsequent images in this series display further oscillations in the value of *C* along the shoreline. Examination of similar images (not shown) after 22 January indicates stabilization of the pack throughout the Arctic Basin until the following April, as

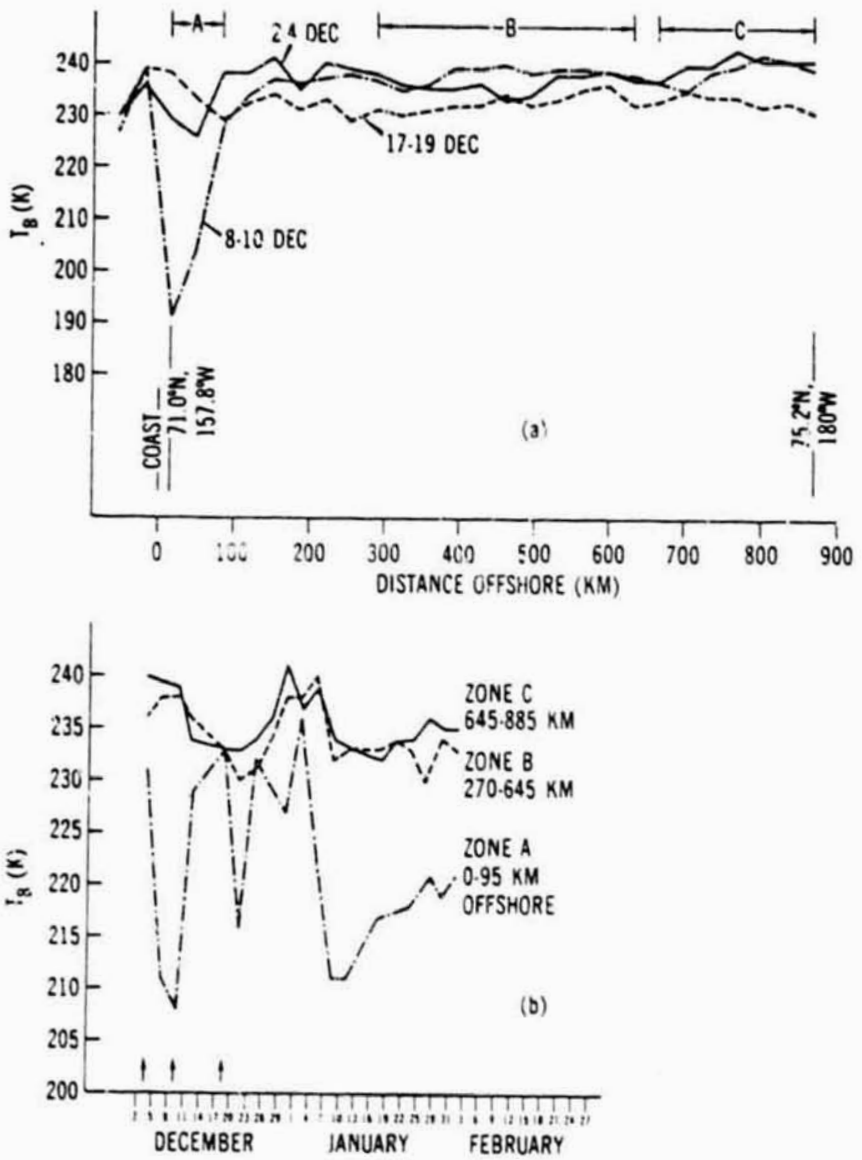


Fig. 11. (a) Temporal variation of the ESMR-5 brightness temperature along the line indicated in Figure 10 for three different dates in the first half of December (upper curves). Regions *A*, *B*, and *C* have been delineated as regions of high and moderate variations of brightness temperature during this season. (b) (lower curves) The temporal variations of the average brightness temperatures in these three regions are illustrated for the month of December, 1973, January, 1974, and February, 1974.

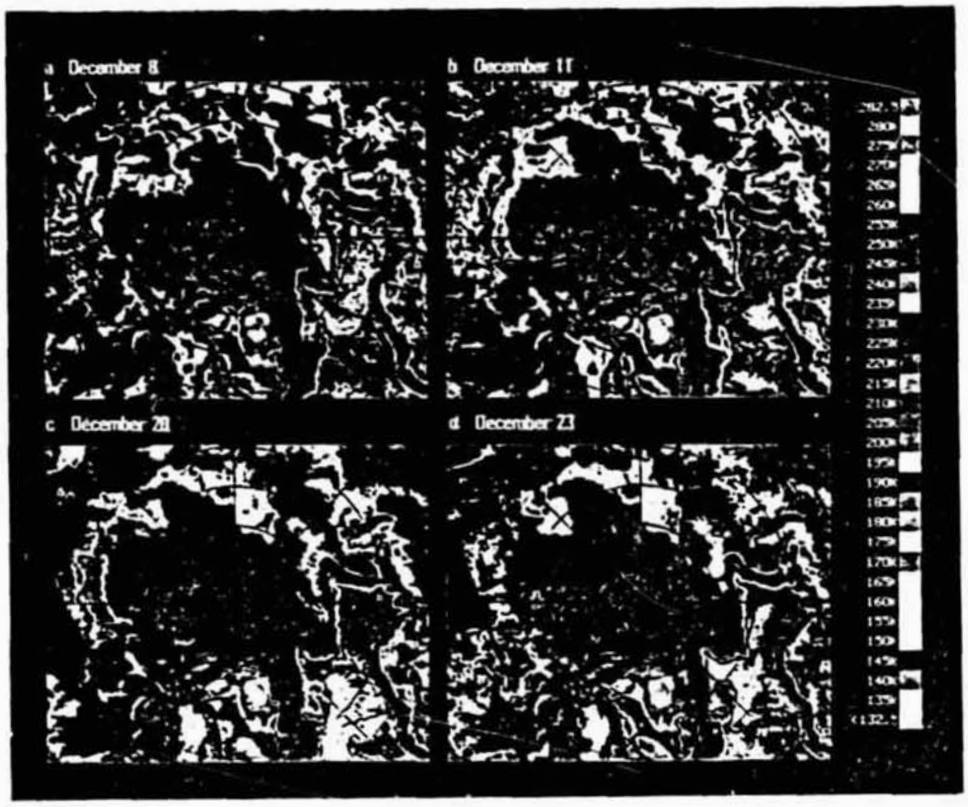


Fig. 12. ESMR-5 images of the Arctic Basin for four dates in December 1973 illustrating the brightness temperature variations in that period. Evidently the largest changes in brightness temperature occur in the coastal regions.

evidenced by the persistence of the T_B patterns, with an overall lowering of the T_B's due to the seasonal trend in the surface temperature.

Another remarkable feature in these images is the distinctly different distribution of multiyear sea ice in the basin, as compared with that shown for the winter of 1972–73, during which F was appreciably higher near the Laptev and East Siberian Seas. As might be expected, there is an apparent change in the distribution of F depending on annual variability of the climate. Systematic studies of this phenomenon may provide a valuable indicator of climatic change.

7. Concluding Remarks

We have observed variability of sea-ice concentration ($1\text{-}C$) and first-year/multiyear sea-ice fraction (F) using an imaging microwave radiometer operating at a wavelength of 1.55 cm on board the Nimbus-5 satellite. Future observations using

ORIGINAL PAGE IS
OF POOR QUALITY

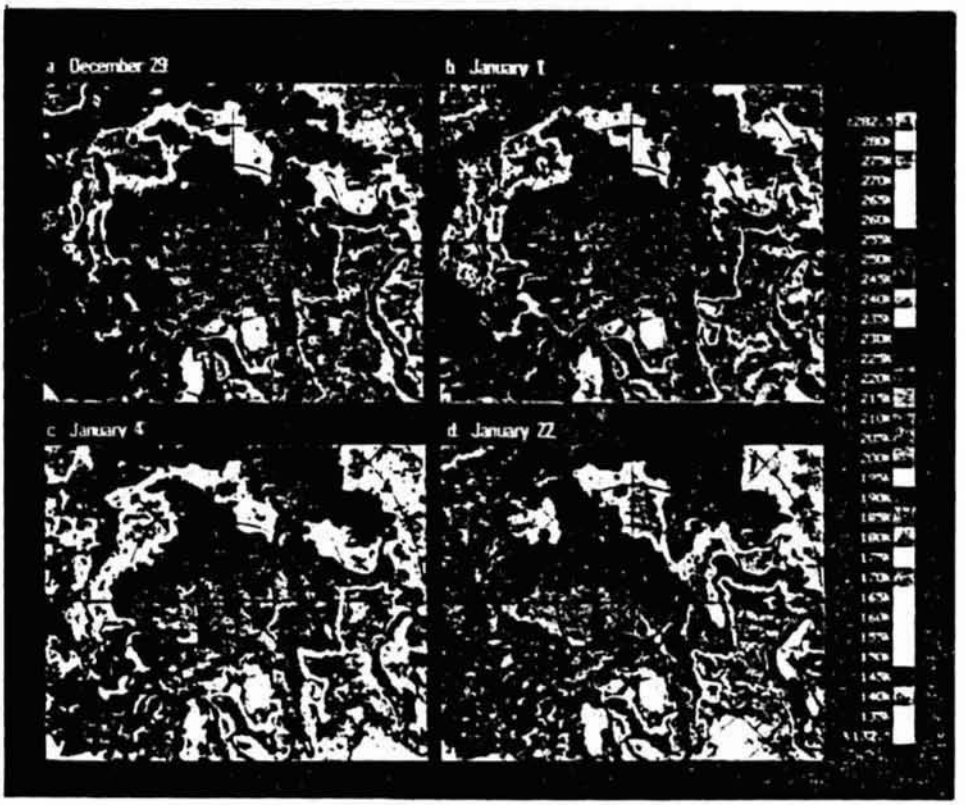


Fig. 13. Same as Figure 12 but one month later.

the two channels of the ESMR-6 in addition to the one of ESMR-5, and eventually the 10-channel SMMR due to be launched on the Nimbus-G and Seasat-A satellites in 1978, are expected to provide valuable information on the Arctic Sea-Ice canopy for monitoring Arctic weather patterns, and for input to existing ice dynamics models or to provide a basis for revision of these models, and to provide a framework for empirical studies related to safety and environmental protection in connection with resource extraction in the northern continental shelf area.

Acknowledgement

We dedicate this paper to our dear friend and colleague, William Nordberg, who passed away during the IUCRM Colloquium on Radio Oceanography at which this paper was presented. He pioneered in observations of microwave radiation from sea ice which led to the efforts described here.

References

Baker, J. G.: 1960, in D. H. Menzel (ed.), *Fundamental Formulas of Physics* **2**, 366, Dover Publications, New York.

Campbell, W. J., Gloersen, P., Nordberg, W., and Wilheit, T. T.: 1974, 'Dynamics and Morphology of Beaufort Sea Ice Determined from Satellites, Aircraft, and Drifting Stations', in P. Bock (ed.), *Cospar Approaches to Earth Survey Problems through Use of Space Techniques*, Akademie-Verlag-Berlin, pp. 311–327.

Campbell, W. J., Weeks, W. F., Ramseier, R. O., and Gloersen, P.: 1975, 'Geophysical Studies of Floating Ice by Remote Sensing', *J. Glaciology* **15**, 305–328.

Campbell, W. J., Gloersen, P., Webster, W. J., Wilheit, T. T., and Ramseier, R. O.: 1976, 'Beaufort Sea Ice Zones as Delineated by Microwave Imagery', *J. Geophys. Res.* **81**, 1103–1110.

Campbell, W. J., Wayenberg, J., Ramseyer, J. B., Ramseyer, R. O., Vant, M. R., Weaver, R., Redmond, A., Arsenault, L., Gloersen, P., Zwally, H. J., Wilheit, T. T., Chung, T. C., Hall, D., Gray, L., Meeks, D. C., Bryan, M. L., Barath, F. T., Elachi, C., Leberl, F., and Farr, T.: 1978, 'Microwave Remote Sensing of Sea Ice in the AIDJEX Main Experiment', this volume, pp. 309–337.

Chang, T. C. and Gloersen, P.: 1975, 'Microwave Emission from Dry and Wet Snow', in A. Rango (ed.), *Operational Applications of Satellite Snowcover Observations*, NASA SP-361, Washington D.C., pp. 399–407.

Chang, A. T. C., Gloersen, P., Schmugge, T., Wilheit, T. T., and Zwally, H. J.: 1976, 'Microwave Emission from Snow and Glacier Ice', *J. Glaciology* **16**, 23–39.

Edgerton, A. T., Stogryn, A., and Poe, G.: 1971, 'Microwave Radiometric Investigations of Snow Pack', Final Report No. 1285R-4 for U.S.G.S. contract no. 1A-08-001-11828, Aerojet-General Corp., Microwave Division, El Monte, California.

Gaut, N. E. and Reifenstein, E. C. III: 1970, 'Interaction of Microwave Energy with the Atmosphere', Am. Inst. of Aeronautics and Astronautics Paper No. 70–197.

Gloersen, P., Nordberg, W., Schmugge, T. J., Wilheit, T. T., and Campbell, W. J.: 1973, 'Microwave Signatures of First-Year and Multiyear Sea Ice', *J. Geophys. Res.* **78**, 3564–3572.

Gloersen, P., Wilheit, T. T., Chang, T. C., Nordberg, W., and Campbell, W. J.: 1974a, 'Microwave Maps of the Polar Ice of the Earth', *Bull. Am. Meteorol. Soc.* **55**, 1442–1448.

Gloersen, P., Chang, T. C., Wilheit, T. T., and Campbell, W. J.: 1974b, 'Polar Sea Ice Observations by means of Microwave Radiometry', in H. S. Santeford (ed.), *Proc. of the Symp. on Advanced Concepts and Techniques on the Study of Snow and Ice Resources*, pp. 541–550.

Gloersen, P. and Salomonson, V. V.: 1975, 'Satellites – New Global Observing Techniques for Ice and Snow', *J. Glaciology* **15**, 373–389.

Gloersen, P., Ramseier, R., Campbell, W. J., Chang, T. C., and Wilheit, T. T.: 1975a, 'Variation of Ice Morphology of Selected Mesoscale Test Areas During the Bering Sea Experiment', *Proc. of the Symp. of the Joint U.S./U.S.S.R. Bering Sea Experiment*, Gidrometeoizdat, Leningrad, pp. 196–218.

Gloersen, P., Ramseier, R., Campbell, W. J., Kuhn, P. M., and Webster, Jr., W. J.: 1975b, 'Ice Thickness Distribution as Inferred from Infrared and Microwave Remote Sensing During the Bering Sea Experiment', *Proc. of the Symp. of the Joint U.S./U.S.S.R. Bering Sea Experiment*, Gidrometeoizdat, Leningrad, pp. 282–293.

Meyers, W.: 1976, AIDJEX Project, private communication.

Ramseier, R. O., Gloersen, P., and Campbell, W. J.: 1974, 'Variation of the Microwave Emissivity of Sea Ice in the Beaufort and Bering Seas', in E. Schanda (ed.), *Proc. of the URSI Commission II-Specialists Meeting on Microwave Scattering and Emission from the Earth*, Berne, Switzerland, pp. 87–93.

Ramseier, R. O., Gloersen, P., Campbell, W. J., and Chang, T. C.: 1975, 'Mesoscale Description for the Principal Bering Sea Ice Experiment', *Proc. of the Symp. of the Joint U.S.-U.S.S.R. Bering Sea Experiment*, Gidrometeoizdat, Leningrad, pp. 254–270.

Wilheit, T., Nordberg, W., Blinn, J., Campbell, W., and Edgerton, A.: 1972, 'Aircraft Measurement of Microwave Emissions from Arctic Sea Ice', *Remote Sensing of Environment* **2**, 129–139.

Wilheit, T. T.: 1972, 'The Electrically Scanned Microwave Radiometer (ESMR) Experiment', *Nimbus-5 User's Guide*, NASA, Goddard Space Flight Center, pp. 59–105.

Wittman, W. J. and Schule, Jr., J. S.: 1966, 'Comments on the Mass Budget of Arctic Pack Ice', in J. O. Fleischer (ed.), *Proc. of the Symp. on the Arctic Heat Budget and Atmospheric Circulation*, in J. O. Fleischer (ed.), Rand Memo RM-5233-NSF.

Paper 80

NASA Technical Memorandum 79642, September 1978.

MINERAL PRECIPITATION IN NORTH SLOPE AUFEIS FIELDS

DOROTHY K. HALL
GODDARD LABORATORY FOR ATMOSPHERIC SCIENCES
NASA/GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND 20771

ABSTRACT

Powdered calcium carbonate ($CaCO_3$) patches averaging 4 cm in thickness were found on aufeis fields in the Canning and Shaviovik Rivers in northeastern Alaska. The presence of this material on aufeis (overflow river ice) suggests that the groundwater which feeds the aufeis is coming from depth and has flowed through calcareous bedrock. Aufeis forms when ground or river water is forced upwards through cracks in the river ice. Calcium carbonate in solution in the groundwater is excluded as the water freezes during ice growth. Calcium carbonate slush accumulates on top of the ice as the aufeis ablates in the melt season. Four patches of $CaCO_3$ covering approximately 0.1% of the total area of the Canning River aufeis were observed during the July 1978 field study. It is estimated that approximately 540 cu m of $CaCO_3$ precipitate were present in the Canning River aufeis in July of 1978. If similar percentages of $CaCO_3$ precipitate were present on other the other aufeis fields on the eastern North Slope, approximately 18,000 cu m of $CaCO_3$ would be present during a given year in the major North Slope aufeis fields. Most of this precipitate is deposited into the Arctic Ocean via river flow.

Paper 81

NASA Technical Memorandum TM 79720 March 1979

GEOMORPHIC PROCESSES ON THE NORTH SLOPE OF ALASKA

Dorothy K. Hall
Hydrospheric Sciences Branch
NASA/Goddard Space Flight Center

ABSTRACT

Three physiographic provinces comprise the North Slope of Alaska: the Arctic Mountains, the Arctic Foothills and the Arctic Coastal Plain Provinces. The features and processes in the Arctic Coastal Plain, a zone of continuous permafrost, are stressed in this paper. The evidence for and mechanisms of the geomorphic cycle are discussed starting with frost cracks. Frost cracks may form polygonal ground which leads to low-centered ice wedge polygons in areas having ice-rich permafrost. As the low-centered ice wedge polygons enlarge due to thermal erosion they may evolve into thaw lakes which are largely oriented in a northwest-southeast direction on the Arctic Coastal Plain. Eventual drainage of a deep lake may result in a closed-system pingo. Evidence of the various stages of the geomorphic cycle is ubiquitous on the Alaskan Arctic Coastal Plain and indicates the ice content of the permafrost in some areas.

N81-21694

Paper 82

Reprinted from *U.S. Army Cold Regions Research and Engineering Special Report 78-10*, August 1978.

THE 1977 TUNDRA FIRE AT KOKOLIK RIVER, ALASKA

D. Hall, J. Brown and L. Johnson

INTRODUCTION

Widespread fires occurred on the Seward Peninsula during the summer of 1977. During this period there was also one large natural fire in northern Alaska. This fire occurred due east of Pt. Lay and several kilometers southwest of the Kokolik River, Alaska (69°30'N, 161°50'W; Fig. 1). No tundra fires have previously been reported from this area. However, fires on the North Slope during the years 1969 to 1971 were reported, four of which were reportedly caused by lightning (Wein 1977, Barney and Comiskey 1973). According to the Bureau of Land Management (BLM) smoke jumper records, this was the northern-most fire ever manned by BLM in Alaska.

The fire, presumably caused by lightning, occurred on the boundary between the coastal plain and the northern foothills (McCulloch 1967). It apparently started just west of the National Petroleum Reserve — Alaska (NPRA) boundary and spread eastward into the NPRA (Fig. 1).

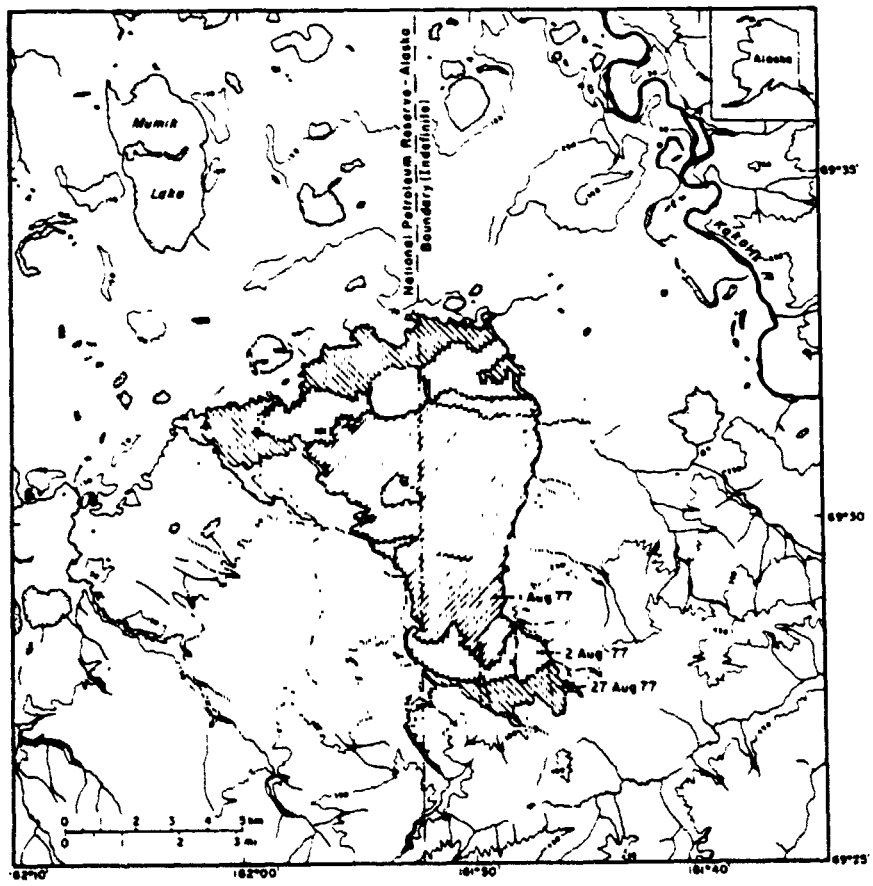


Figure 1. Location map of Kokolik River tundra fire with 1, 2, and 27 August boundaries of the fire. Boundaries of the fire were drawn from 1:63,360 Landsat photographic enlargements prepared by T. Marlar, CRREL.

Climatic conditions in northern and western Alaska during the summer of 1977 were apparently ideal for tundra fires. Cape Lisburne, operated by the U.S. Air Force, is 185 km to the southwest on the coast and is the closest available reporting station to the 1977 Kokolik River fire.

July was extremely dry (5.3-mm rainfall) and average temperatures were above normal. Maximum air temperatures in late July were in the mid-teens and low 20s (°C). In general, the North Slope experienced an extremely dry July with no precipitation reported from Umiat or Barrow for the period 9 to 24 July 1977.

The fire was discovered by the BLM on 26 July, at which time its extent was estimated at 0.4 km². But high winds prevented the firefighters from landing on 26 July. These winds, which continued through 31 July, undoubtedly contributed to the intense nature of the fire. By noon on 27 July, when the first BLM firefighters arrived, the extent was estimated at 4.0 km². BLM personnel fought the fire from 27 July to 31 July, at which time it was abandoned in order to deal with fires elsewhere.

Apparently light rains nearly extinguished the fire on 28 July, and after it started up again, fog slowed its progress on 30 July. Nevertheless, Landsat imagery indicates the fire was still burning on 2 August. The BLM fire report indicated that on 7 August the fire had not gained any "acreage" and was declared to be "controlled." It was declared out by BLM on 12 August.

We have studied the fire through the combined use of Landsat imagery, a reconnaissance field trip, and field reports from BLM. This short report presents information on the spread of the fire, the meteorological observations, the natural containment of the fire, and the ground investigations following the fire.

SATELLITE OBSERVATIONS

Landsat multispectral scanner subsystem (MSS) imagery has previously been employed to measure the progression and areal extent of wildfires (Haugen et al. 1972). Landsat provides repetitive coverage with a spatial resolution of 80 m. The MSS onboard the Landsat-1 and -2 satellites provides imagery in four spectral bands; band 4 (0.5-0.6 μm) and band 5 (0.6-0.7 μm) in the visible range; and band 6 (0.7-0.8 μm) and band 7 (0.8-1.1 μm) in the near-infrared range. Only imagery from bands 5 and 7 was used in this study. Analysis of the 1977 Landsat imagery provided considerable information on the areal extent and movement of the Kokolik River fire, as well as the rapidly changing wind conditions which affected the spread of the fire.

Figure 2 shows the area of the Kokolik River fire before the fire on 21 July 1977. Note the two small lakes southest of Mumik Lake. These lakes became surrounded by fire as seen on later Landsat scenes.

By 1 August the fire was burning quite intensively (Fig. 3). The smoke shows up clearly on band 5 imagery and provides a good indicator of wind direction. On 2 August a drastic shift in wind direction from the previous day occurred, as seen in Figure 4. National Weather Service Synoptic Maps indicate that there was a shift in the wind direction from northeast at 5 knots on 1 August to southwest at 10 knots on 2 August (NWS 1977). This information correlates well with the information derived from Landsat. In fact, the Landsat imagery is probably a better indicator of wind direction than the NWS maps for which the data have been extrapolated over large areas. Considerable haziness is visible on the 2 August image (Fig. 4). This is probably due to smoke from fires which were burning at the time on the Seward Peninsula (southwest of the Kokolik River fire).

Figures 5 and 6 are band 7 images of the 1 and 2 August scenes. Band 7 is more useful than band 5 for delineation of the areal extent of the fire. The extent was determined to be 26 km² on 1 August and 34 km² on 2 August. Digitized Landsat imagery was used to obtain these areal measurements on the Atmospheric and Oceanographic Image Processing System (AOIPS) at Goddard Space Flight Center.

Black and white positive transparencies of Landsat scenes were scanned and projected onto a television screen. The burned area was then delineated by a cursor.* Subsequently, a pixel (picture element) count of the burned area was taken. The number of Landsat pixels was determined and then the total extent in square kilometers was calculated. Digital analysis of the 1 and 2 August images revealed an 8-km² increase in extent of burn over the 24-hour period. Measurements using a 21 August image (Fig. 7) revealed an ultimate extent of 44 km² for the burn. No Landsat images could be obtained between 2 and 21 August but the fire was considered to be out by 12 August.

*The cursor on a computer terminal delineates the outline of the burned area.

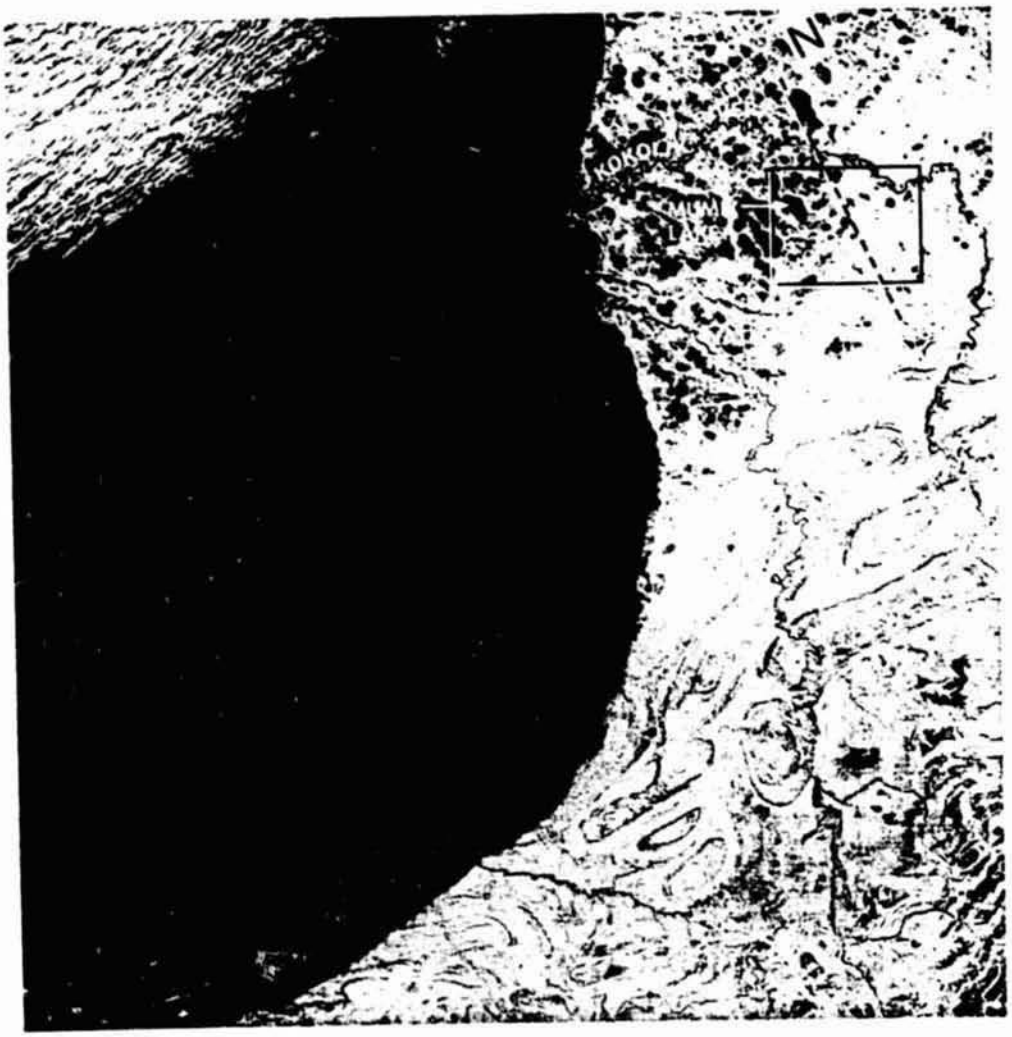


Figure 2. Landsat image of Kokolik River area, before fire, 16 July 1977, scene 2906-21410, band 7.

ORIGINAL PAGE IS
OF POOR QUALITY

Figure 3. Landsat image of Kokolik River fire in progress showing smoke plumes, 1 August 1977, scene 2922-21284, band 5.

Figure 4. Landsat image of Kokolik River fire in progress showing smoke plumes, 2 August 1977, scene 2923-21342, band 5.

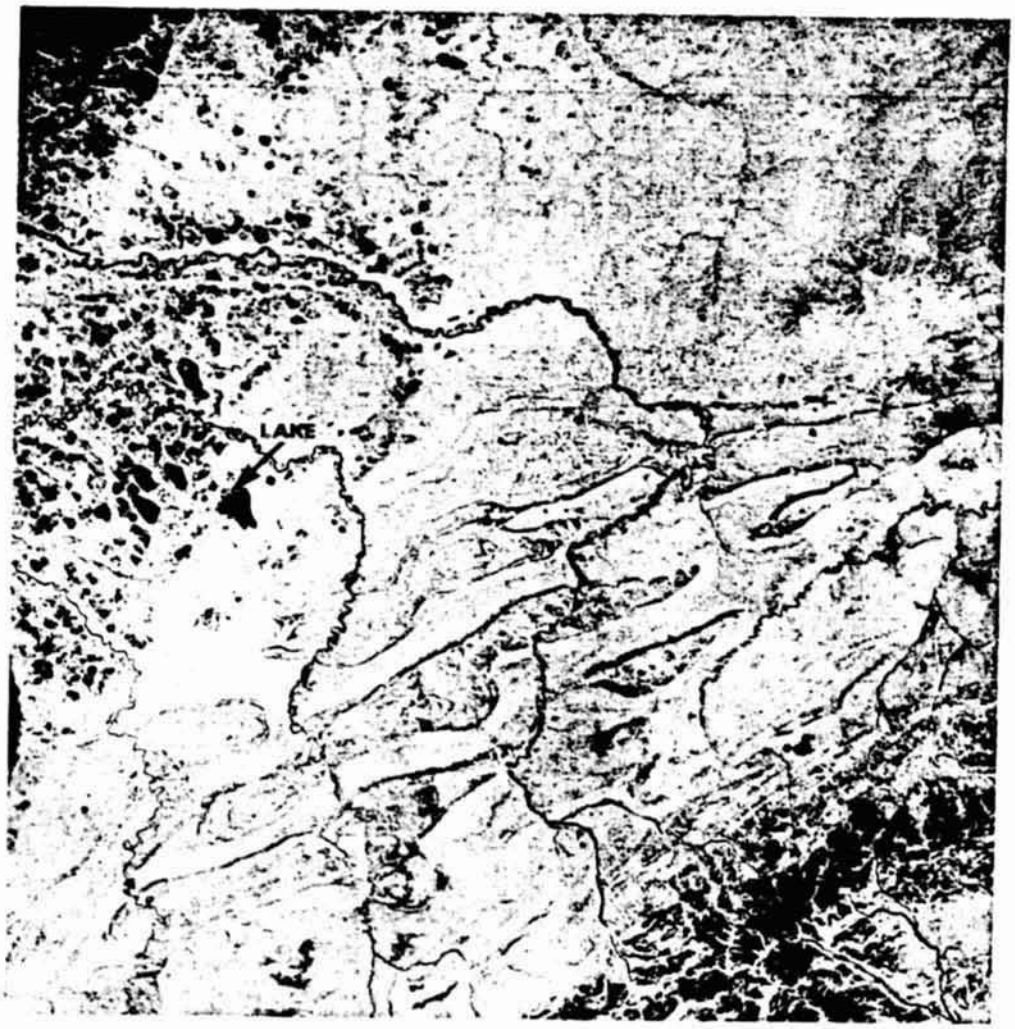


Figure 5. Landsat image of Kokolik River fire extent, 1 August 1977, scene 2922-21284, band 7.

Figure 6. Landsat image of Kokolik River fire extent, 2 August 1977, scene 2923-21342, band 7.

IGINAL PAGE IS
POOR QUALITY

Figure 7. Landsat image showing maximum extent of Kokolik River fire, 21 August 1977, scene 2942-21390, band 7.

ORIGINAL PAGE IS
OF POOR QUALITY

Figure 8 shows 1:250,000-scale enlargements of the Kokolik River burned area on 1 and 27 August 1977. By superimposing the burned areas from the Landsat imagery on the 1:250,000 Pt. Lay and Utukok River U.S. Geologic Survey topographic maps, it became apparent that the stream bordering the fire on the east formed a natural fire barrier. The fire then spread northward and southward (Fig. 1) being arrested again by streams to the north. The small streams that ultimately contained the fire are below the 80-m resolution of the Landsat imagery. However, recently available NASA aerial photography, obtained on 1 August 1977 (see cover photograph), confirms the fact that the streams contained the fire, as interpreted on the Landsat imagery.

GROUND OBSERVATIONS

Craig Johnson of the Forest Soil Laboratory, University of Alaska, and Lawrence Johnson, CRREL, spent 30 August on site after landing there by a BLM-chartered helicopter. Low-altitude observations from the helicopter indicated that that fire burned a range of vegetation and relief types which included low polygonized and upland tussock tundras. The fire was extremely hot and, although exposed mineral soils in the upland tundra were not a high percentage of the total area, the extent of bare soils was several times greater in the burned areas than in comparable unburned areas. Some of the bare areas could have been exposed previously as frost scars or boils.

A transect within the upland tundra was established from approximately 180 m within the burn, to the edge of the burn, across a small creek and up onto the unburned tundra. The exact ground location of the transect was difficult to ascertain because of poor visibility from the air, but was most likely located on the eastern edge of the burn (Fig. 1). Three 1-m²

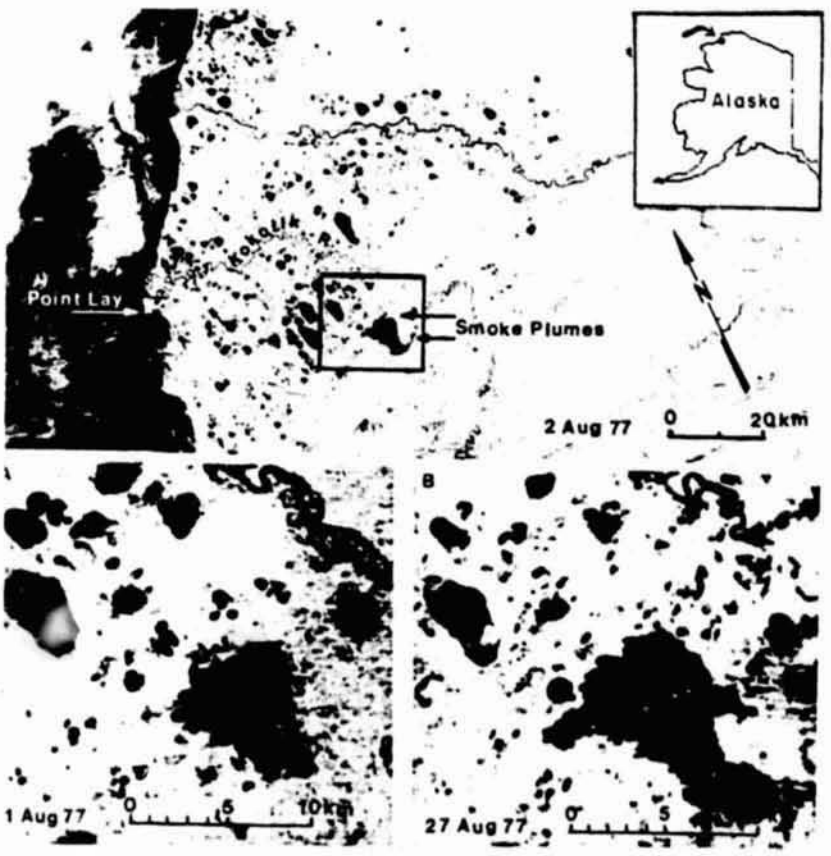


Figure 8. Landsat band 7 scene (2923-21342) of Kokolik River fire, showing extent and direction of smoke on 2 August 1977. Insert shows a 1:250,000 scale enlargement of the fire area from (A) 1 August (2922-21284 and (B) 27 August (5861-20500) scenes.

quadrats were established in each of the following burned areas: 1) severely burned tussock tundra, 2) moderately burned wet tundra, 3) lightly burned shrub tundra at the edge of the fire, and 4) unburned tussock tundra. Percentages of cover of live vascular plants, mosses, lichens, and charred surface were observed within each quadrat. Soil samples were obtained for subsequent nutrient analysis. Depth of thaw was measured at five points within the quadrat and at 25-m intervals along the 180-m transect.

By comparison with the unburned vegetation, which had large, well-developed tussocks, it was estimated that 80-90% of the biomass of the tussocks was consumed in the fire. All areas of raised relief, which characteristically consisted or tussock tundra, were severely burned and averaged less than 10% live cover. This may have been due to greater available fuel (litter and standing dead vegetation), drier soils associated with higher terrain, or both.

In contrast, irregularly distributed areas of low relief, such as swales, were moderately burned and averaged 10% live cover. Lightly burned areas averaged 50% live cover. It appeared that all aboveground shrub parts in both severely and moderately burned areas had been consumed by the fire, and if there were any living below ground parts, they had not sprouted by the end of August. In contrast, 80% of the tussocks already showed regrowth from shoots (Fig. 9). There was also regrowth in meadows containing *Carex* and *Eriophorum angustifolium* which occupy the low-lying, moderately burned areas. Similar late summer regrowth in tussocks was observed in the Elliott Highway tussock fire in 1969 (Wein and Bliss 1973, Brown and Vietor 1969).

The burned area appeared wetter on the surface than the unburned area because of a lack of moisture-absorbing organic matter and possible release of moisture from the deeper thawed zone. Depth of thaw between tussocks averaged 35.4 cm in the burned areas as compared with 26.6 cm in the unburned areas. Mackay (1977) reports upon the continued seasonal increase in thaw at the Inuvik fire following an average initial increase of 24 cm after that fire. Kriuchkov (1968) [translation in Brown et al., (1969)] discusses the initial increase in thaw followed by a decrease in subsequent years for tundra fires in Siberia.

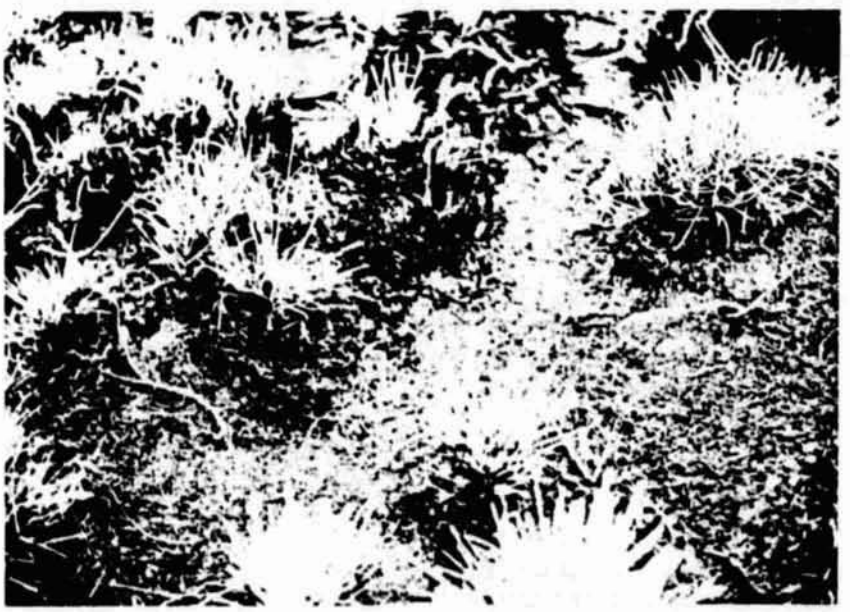


Figure 9. Closeup of severely burned tussocks. Individual tussocks show regrowth. (Photograph by L. Johnson, CRREL, on 30 August 1977.)

CONCLUSIONS

The summer of 1977 Kokolik fire totaled 44 km² of tundra vegetation according to measurements using Landsat imagery. Based on the experience gained from analysis of this fire using ground observations, Landsat imagery, and topographic maps, it appears that natural drainages form effective fire breaks on the subdued relief of the arctic coastal plain and northern foothills. Furthermore, this study reconfirmed that intensity of the fire is related to vegetation type and moisture content of the organic rich soils.

FUTURE INVESTIGATIONS

The Kokolik River fire site provides an excellent site for continued investigations of fire effects on tundra and permafrost terrains. The following studies are desirable:

(1) Comparison of the nutrient runoff and change in primary productivity between the burned and unburned areas. The influence of both increased nutrient and sediment load on the lakes contained within the fire area (Fig.1) could also be followed.

(2) Monitoring depth of thaw changes and thermokarst development under burned areas of different intensity and moisture conditions. The hypothesis proposed by L. Viereck (personal communication) that intensity of frost action may be directly related to fire frequency on the Seward Peninsula could be tested at this site.

(3) Monitoring revegetation and successional changes on the burned areas. The increased flowering of *Eriophorum vaginatum* which normally occurs after a fire could provide a harvestable seed source for revegetation efforts elsewhere. Digital analyses of the available August 1977 Landsat imagery can be used to classify burn intensity as a function of reflectance differences.

(4) Acquisition of ground truth data, particularly on vegetation succession, which can be correlated with detailed digital analyses of future Landsat imagery. Landsat data obtained during the spring and summer of 1978 should reveal valuable information on initial vegetation response in the burned areas.

LITERATURE CITED

Barney, R.J. and A.L. Comiskey (1973) Wildfires and thunderstorms on Alaska's North Slope. U.S. Department of Agriculture, Forest Service Research Note, PNW-212. Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, 8 p.

Brown, J., W. Rickard and D. Vietor (1969) The effect of disturbance on permafrost terrain. CRREL Special Report 138. AD 699327.

Haugen, R.K., H.L. McKim, L.W. Gatto and D.M. Anderson (1972) Cold regions environmental analysis based on ERTS-1 imagery. *Proceedings of the Eighth International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan*, p. 1511-1521.

Kriuchkov, V.V. (1968) Soils of the Far North should be preserved. *Priroda*, no. 12, p. 72-74. (Translation in Brown, J. et al., 1969, CRREL Special Report 138). AD 699327.

Mackay, J. Ross (1977) Changes in the active layer from 1968 to 1976 as a result of the Inuvik fire. Report of Activities, Part B. Geological Survey, Canada, Paper 77-1B, p. 273-275.

McCulloch, D.S. (1967) Quaternary geology of the Alaskan shore of Chukchi Sea. In *The Bering Land Bridge* (D.M. Hopkins, editor). Stanford, California: Stanford University Press, p. 91-120.

National Weather Service (1977) National Weather Service 6-hourly synoptic maps, Ashville, North Carolina.

Wein, R.W. (1977) Frequency and characteristics of Arctic tundra fires. *Arctic*, vol. 29, p. 213-222.

Wein, R.W. and L.C. Bliss (1973) Changes in Arctic *Eriophorum* tussock communities following fire. *Ecology*, vol. 54, p. 845-852.

Wein, R.W., T.W. Sylvester, and M.G. Weber (1975) Vegetation recovery in Arctic tundra and forest-tundra after fire. Arctic Land Use Research 47-75-62, Department of Indian Affairs and Northern Development, Ottawa, Canada, 115 p.

N81-21695 D2

Paper 83

Reprinted from *Arctic, Journal of the Arctic Institute of North America*, Vol. 31, No. 1, March 1978.

Short Papers

The 1977 Tundra Fire in the Kokolik River Area of Alaska

INTRODUCTION

Widespread fires occurred on the Seward Peninsula, Alaska, during the summer of 1977. During this period there was also one large natural fire in the northern part of the state. Presumably caused by lightning, it occurred due east of Point Lay and several kilometres southwest of the Kokolik River (69°30'N, 161°50'W) on the boundary between the coastal plain and the northern foothills[1] (Fig. 1). According to the smoke-jumper record kept by the U.S. Bureau of Land Management (BLM) this was the farthest north a fire had ever been fought by the personnel of BLM in Alaska. No tundra fires had previously been reported from this area, although during the years 1969 to 1971 fires occurred on the North Slope of which four were stated to have been caused by lightning[2,3].

Climatic conditions in northern and western Alaska during the summer of 1977, as observed at the Cape Lisburne station of the U.S Air Force, situated 185 km to the southwest of the area of the Kokolik River fire (the meteorological observation centre closest to it), were apparently ideal for tun-

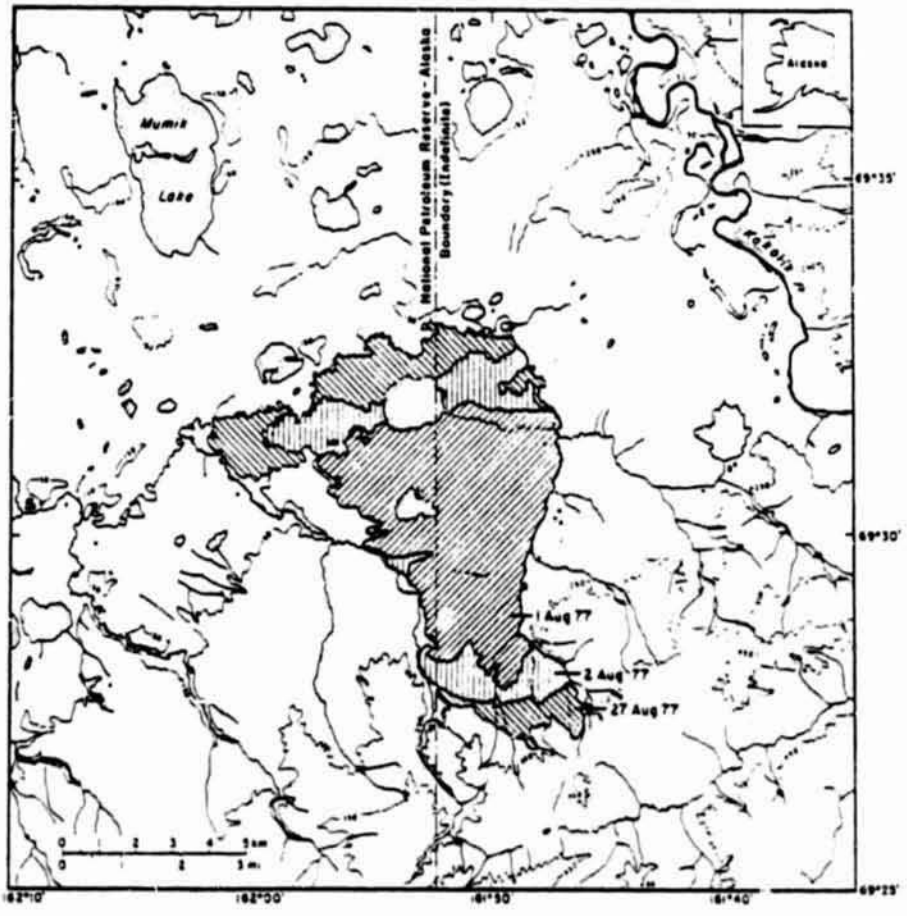


FIG. 1. Location map of Kokolik River tundra fire, with its extent on 1, 2 and 27 August 1977 indicated by shadings. Boundaries of the fire were drawn from 1:63,360 Landsat photographic enlargements prepared by Tom Marlar, U.S. Army Cold Regions Research and Engineering Laboratory.

dra fires. In July it was extremely dry (5.3 mm rainfall); and average air temperatures were above normal, maxima in the latter part of the month being in the mid-teens and low twenties (°C). The North Slope was in general extremely dry, no precipitation being reported from Umiat or Barrow for the period 9-24 July 1977.

The Kokolik River fire apparently started just west of the National Petroleum Reserve — Alaska (NPRA) boundary and then spread eastward into the NPRA itself (Fig. 1). It was discovered by personnel of the BLM on 26 July, at which time its extent was estimated at 0.4 km². High winds, which prevented the fire fighters from landing on 26 July, prevailed until 31 July and undoubtedly contributed to the intense nature of the fire. By noon on 27 July, when the first BLM fighters arrived, the extent was estimated at 4.0 km². The BLM personnel fought the fire from 27 July to 31 July, when they abandoned it in order to attend to outbreaks elsewhere. Apparently, light rains nearly extinguished the fire on 28 July, and after it had started up again fog slowed its progress once more on 30 July. Nevertheless, according to the Landsat imagery the fire was still burning on 2 August. It was stated in the BLM fire report that on 7 August the fire had not gained any "acreage" and was "controlled". It was declared out by the BLM on 12 August.

The information on the spread of the fire,

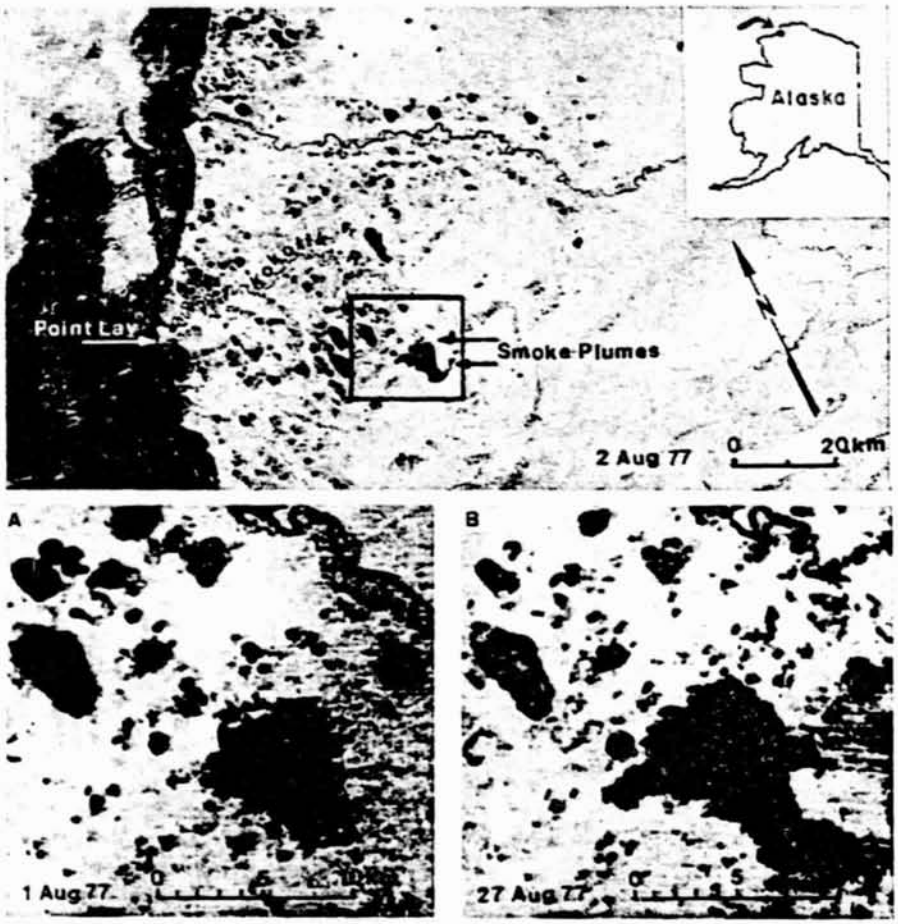


FIG. 2. TOP: Landsat band 7 scene (scene I.D. no. 2923-21342) of Kokolik River fire, showing extent and direction of smoke on 2 August 1977. BOTTOM: 1:250,000 enlargements of scenes of the fire area on (a) 1 August (scene I.D. no. 2922-21284) and (b) 27 August (scene I.D. no. 5861-20500).

as presented in this short paper, was gathered from Landsat imagery, meteorological observations, facts concerning the natural containment of the fire, and investigations of the affected area following the fire.

SATELLITE OBSERVATIONS

Landsat multispectral scanner subsystem (MSS) imagery has previously been employed to measure the progression and areal extent of wildfires[4]. Landsat provides repetitive coverage with a spatial resolution of 80 metres. The MSS on board the Landsat-1 and Landsat-2 provides imagery in four spectral bands: band 4 (0.5-0.6 μm) and band 5 (0.6-0.7 μm) in the visible range; and band 6 (0.7-0.8 μm) and band 7 (0.8-1.1 μm) in the near-infrared range. Only imagery from bands 5 and 7 were used in this study. Analysis of the 1977 Landsat imagery provided considerable information on the spread of the Kokolik River fire as well as the rapidly changing wind conditions which influenced it.

Examination of the band 5 (visible) image for 1 August clearly revealed two distinct smoke plumes emanating from the fire on its eastern boundary and a wind blowing from the east. The band 7 (near-infrared) image did not show well-defined smoke plumes and therefore provided a poor indication of wind direction; it did however provide the best available view of the 26 km^2 areal extent of the fire. The measurements of the areal extent of the fire are based on digital analysis of 1:1,000,000-scale Landsat imagery. Photographic enlargements of the burned areas were laid directly on top of a 1:63,360 topographic map of the U.S. Geological Survey (see Fig. 1).

The band 5 image for 2 August (Fig. 2) indicated a shift in wind direction, with smoke blowing to the east. National Weather Service synoptic maps indicate that there was a change in wind, from 5 knots (9 km per hour) northeasterly on 1 August to 10 knots southwesterly on 2 August[5]. These data correlate well with the results of the Landsat observations.

Considerable haziness existed on the band 5 image of 2 August, probably due to the smoke from fires that were burning on the Seward Peninsula.

A comparison of the band 7 images of 1 August and 2 August reveals an increase from 26 km^2 in the extent of the fire over the 24-hour period to 34 km^2. The images of the fire of 21 August and of 27 August both indicate that a total of 44 km^2 had burned — an increase of about 10 km^2 since 2 August (Fig. 2).

By superimposing enlargements of the Landsat imagery of the burned areas on the topographic map of scale 1:63,360, it is apparent that the stream bordering the fire on the east had formed a natural fire barrier, and that the fire then spread northward and southward (Fig. 1). Again, the stream channel to the north seemed to have stopped the spread of the fire. These small channels were not apparent on the Landsat image.

GROUND OBSERVATIONS

Craig Johnson of the Forest Soil Laboratory, University of Alaska, together with one of the present authors (Johnson) spent 30 August on site after landing there by helicopter. Low-altitude observations from the helicopter indicated that the fire burned a range of vegetation and relief types which included low-polygonized and upland-tussock tundras. The fire was extremely hot and, although exposed mineral soils in the upland tundra did not constitute a high percentage of the total area, the extent of bare soils was several times greater in the burned areas than in comparable unburned areas. Some of the bare areas could have previously been exposed as frost scars or boils.

A transect within the upland tundra was established from approximately 180 metres within the burn to the edge of the burn, across a small creek and up onto the unburned tundra. The exact ground location of the transect was difficult to ascertain, due to poor visibility from the air, but it is most likely located on the eastern edge of the burn (Fig. 1). Three one-metre-square quadrats were established in each of the following burned areas: severely burned tussock tundra; moderately burned wet tundra; lightly burned shrub tundra at the edge of the fire; and unburned tussock tundra. The percentage of area covered by live vascular plants, mosses, lichens and charred surface was observed within each quadrat. Soil samples were obtained for subsequent nutrient analysis. Depth of thaw was measured at five points within the quadrat and at 25-m intervals along the 180-m transect.

On the basis of a comparison with the unburned vegetation, which had large, well-developed tussocks, it was estimated that 80-90% of the biomass of the tussocks was consumed in the fire. All areas of raised relief, which characteristically consisted of tussock tundra, were severely burned, averaging less than 10% live-plant cover. This may have been due to either more fuel (litter and standing dead vegetation) being available, drier soils being associated with higher terrain, or both conditions. In contrast, irregu-

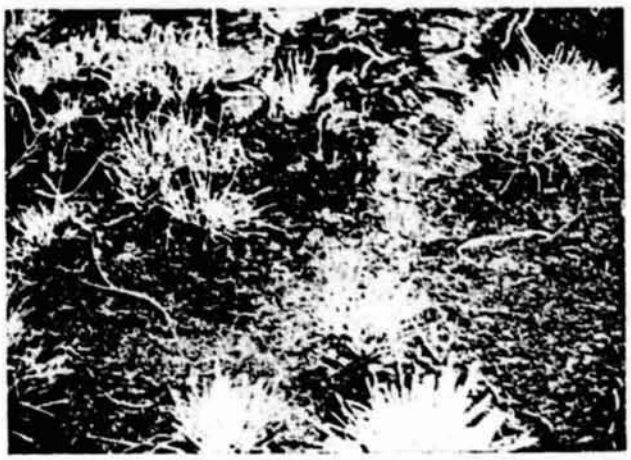


FIG. 3. Close-up view of severely burned tussocks. Individual tussocks show regrowth. Photograph by L. Johnson, 30 August 1977.

larly distributed areas of low relief, such as swales, were moderately burned, averaging 30% live plant cover, while in lightly burned areas live vascular-plant cover averaged 40%. It appeared that all shrub parts present above ground, in both severely and moderately burned areas, had been consumed by the fire, and if any below-ground shrub parts were still living, they had not sprouted by the end of August. In contrast, 80% of the tussocks already showed regrowth, as was evidenced by their shoots (Fig. 3); and there was also regrowth in wetter areas containing *Carex* spp. and *Eriophorum angustifolium*. Such regrowth of tussocks in the late summer was also observed in the tussock fire at Elliott Highway in 1969 [6,7].

The burned area appeared wetter on the surface than the unburned area, due to a lack of moisture-absorbing organic matter and the possible release of moisture from the deeper-thawed zone. Depth of thaw between tussocks averaged 35.4 cm in the burned areas, as compared to 26.6 cm in the unburned ones. Mackay[8] reports that after the Inuvik fire of 1968 there was a continued seasonal increase in thaw following an average initial increase of 24 cm. Kryuchkov[9] discusses the fact that an initial increase in thaw was followed by a subsequent decrease in it, after tundra fires in Siberia.

POSSIBLE FUTURE INVESTIGATIONS

The site of the Kokolik River fire is excellent for continued investigations, similar to those undertaken at Inuvik[10], of the effects of fire on tundra and permafrost terrains. It is desirable to:

(a) compare the nutrient runoff and change in primary productivity between the burned and unburned areas. The influence of anticipated increases in nutrient and sediment loads on the lakes contained within the fire area (Fig. 1) could also be investigated.

(b) monitor depth-of-thaw changes and thermokarst development under areas of different intensity of burn and different moisture conditions. The hypothesis proposed by L. Viereck, U.S. Forest Service (personnel communication), that intensity of frost action may be directly related to fire frequency on the Seward Peninsula could be tested at this site.

(c) monitor revegetation and successional changes on the burned areas. The increased flowering of *Eriophorum vaginatum*, which normally occurs after a fire, could provide a harvestable seed source for revegetation elsewhere. Digital analyses of the available Landsat imagery for August 1977 can be used to classify burn intensity as a function of reflectance differences.

(d) acquire ground-truth data, particularly on vegetation succession, which can be correlated with those obtained from detailed digital analyses of future Landsat imagery. Data obtained from the latter during the spring and summer of 1978 should reveal valuable information on the initial response of vegetation in the burned areas.

CONCLUSIONS

The Kokolik River fire of 1977 affected an area of 44 km^2, according to measurements made using Landsat imagery. On the evidence obtained from an analytical study of this fire by ground observations, Landsat imagery and topographic maps, it appears that natural drainages form effective fire breaks on the subdued relief of the Arctic coastal plain and northern foothills. Furthermore, this study served to reconfirm the fact

that the intensity of fire is related to vegetation type and moisture content for organically-rich soils.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support rendered them by the U.S. Army Cold Regions Research and Engineering Laboratory, through its Cold Regions Environmental Factors project, as well as from the results of the environmental investigations conducted by it, with U.S. Geological Survey support, in the National Petroleum Reserve — Alaska. They also gratefully acknowledge the assistance given them by Mr. Kenneth Lehtonen of the National Aeronautics and Space Administration in the digital analyses of the Landsat imagery, and also other staff members of NASA. Finally they wish to thank Dr. Leslie Viereck of the U.S. Forest Service, Fairbanks, Alaska, for a critical review of, and valuable contribution to, the manuscript, and also Mr. John Santora of the U.S. Bureau of Land Management, Fairbanks, who made a helicopter available for surveys of the fire and provided other information on it.

Dorothy K. Hall
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771, U.S.A.

Jerry Brown
U.S. Army Cold Regions Research and Engineering Laboratory
Hanover, New Hampshire 03755, U.S.A.

Larry Johnson
U.S. Army Cold Regions Research and Engineering Laboratory
Fairbanks, Alaska 99701, U.S.A.

REFERENCES

[1]McCulloch, D.S. 1967. Quaternary geology of the Alaskan shore of the Chukchi Sea. In: Hopkins, D.M. (ed.), *The Bering Land Bridge*. Stanford, California: Stanford University Press. pp. 91-120.

[2]Wein, R.W. 1976. Frequency and characteristics of Arctic tundra fires. *Arctic*, 29(4): 213-22.

[3]Barney, R.J. and Comiskey, A.L. 1973. Wildfires and thunderstorms on Alaska's North Slope. *U.S., Department of Agriculture, Forest Service, Research Note* PNW-212.

[4]Haugen, R.K., McKim, H.L., Gatte, L.W. and Anderson, D.M. 1972. Cold regions environmental analysis based on ERTS-1 imagery. *Proceedings, Eighth International Symposium on Remote Sensing of Environment, 1972*. Ann Arbor, Michigan: Environmental Research Institute of Michigan. Vol. 2, pp. 1511-21.

[5]N.W.S. 1977. National Weather Service 6-hourly synoptic maps. Ashville, North Carolina.

[6]Wein, R.W. and Bliss, L.C. 1973. Changes in Arctic *Eriophorum* tussock communities following fire. *Ecology*, 54(4):845-52.

[7]Brown, J., Rickard, W. and Vietor, D. 1969. The effect of disturbance on permafrost terrain. *U.S. Army Cold Regions Research and Engineering Laboratory, Special Report* 138.

[8]Mackay, J.R. 1977. Changes in the active layer from 1968 to 1976 as a result of the Inuvik fire. *Canada, Geological Survey, Paper* 77-1B. Part B. pp. 273-75.

[9]Kryuchkov, V.V. 1968. Soils of the Far North should be preserved. *Priroda*, no. 12: 72-74.

[10]Wein, R.W., Sylvester, T.W. and Weber, M.G. 1975. Vegetation recovery in Arctic tundra and forest-tundra after fire. *Arctic Land Use Research* 47-75-62.

Paper 84

PASSIVE MICROWAVE STUDIES OF SNOWPACK PROPERTIES 1/

640-78

By

D. K. Hall 2/, A. Chang 2/, J. L. Foster 3/

A. Rango 2/, and T. Schmugge 2/

Introduction

The use of remotely-acquired microwave data, in conjunction with essential ground measurements will most likely lead to improved information extraction regarding snowpack properties beyond that available by conventional techniques. Landsat visible and near-infrared satellite data have recently come into near operational use for performing snowcovered area measurements (Rango, 1975: 1978). However, Landsat data acquisition is hampered by cloudcover, sometimes at critical times when a snowpack is ripe. Furthermore, information on water equivalent, free water content, and other snowpack properties germane to accurate runoff predictions is not currently obtainable using Landsat data alone because only surface and very near-surface reflectances are detected.

Microwave data have the potential to provide information regarding internal snowpack properties, even through cloudcover because of the penetrating character of the radiation. However, operational use of remotely-collected microwave data for snowpack analysis is not imminent because of complexities involved in the data analysis. Snowpack and soil properties are highly variable and their effects on microwave emission are still being explored. Nevertheless much work is being done to develop both active microwave (Hoekstra and Spangle, 1972; Ellerbruch, et al., 1977) and passive microwave techniques (Schmugge et al., 1973; Schmugge, 1973; Linlor et al., 1974; and Chang et al., 1976) for analysis of snowpack properties. Passive microwave data obtained during recent flights by NASA aircraft over Colorado will be discussed in this paper.

1/ Presented at the Western Snow Conference, April 18-20, 1978, Otter Rock, Oregon

2/ NASA/Goddard Space Flight Center, Greenbelt, Maryland.

3/ Systems and Applied Sciences Corporation, Riverdale, Maryland.

"Reprinted Western Snow Conference 1978"

Paper 85

NASA Technical Memorandum 80310 June 1979

PASSIVE MICROWAVE APPLICATIONS TO SNOWPACK MONITORING USING SATELLITE DATA

Dorothy K. Hall

James L. Foster

Alfred T. C. Chang

Albert Rango

ABSTRACT

Nimbus-5 Electrically Scanned Microwave Radiometer (ESMR) data were analyzed for the fall of 1975 and winter and summer of 1976 over the Arctic Coastal Plain of Alaska to determine the applicability of those data to snowpack monitoring. It was found that when the snow depth remained constant at 12.7 cm, the brightness temperatures (T_B) varied with air temperature. During April and May the production of ice lenses and layers within the snow, and possibly wet ground beneath the snow contribute to the T_B variations also. Comparison of March T_B values of three areas with the same (12.7 cm) snow depth showed that air temperature is the predominant factor controlling the T_B differences among the three areas, but underlying surface conditions and individual snowpack characteristics are also significant factors.

Paper 86

NASA Technical Memorandum 79613, July 1978.

*Presented at the American Society of Photogrammetry Fall Technical Meeting, October 15-20, 1978, Albuquerque, NM.

PASSIVE MICROWAVE STUDIES OF FROZEN LAKES*

Dorothy K. Hall
James L. Foster
Albert Rango
Alfred T. C. Chang
Hydrospheric Sciences Branch
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

Lakes of various sizes, depths and ice thicknesses in Alaska, Utah and Colorado were overflown with passive microwave sensors providing observations at several wavelengths. A layer model is used to calculate the microwave brightness temperature, T_B (a function of the emissivity and physical temperatures of the object), of snowcovered ice overlain with water. Calculated T_B's are comparable to measured T_B's. At short wavelengths, e.g., 0.8 cm, T_B data provide information on the near surface properties of ice covered lakes where the long wavelength, 21.0 cm, observations sense the entire thickness of ice including underlying water. Additionally, T_B is found to increase with ice thickness. 1.55 cm observations on Chandalar Lake in Alaska show a T_B increase of 38°K with a ~124 cm increase in ice thickness.

Paper 87

NASA Technical Memorandum 80602 December 1979

LANDSAT DIGITAL ANALYSIS OF THE INITIAL RECOVERY OF THE KOKOLIK RIVER TUNDRA FIRE AREA, ALASKA

D. K. Hall*

J. P. Ormsby*

L. Johnson**

J. Brown†

ABSTRACT

During late July and early August 1977, a wildfire burned a 48 square kilometer area in the tundra of northwestern Alaska near the Kokolik River. The environmental effects of the fire were studied in the field and from aircraft and Landsat data. Three categories of burn severity were delineated using an August 1977 Landsat scene acquired shortly after the fire stopped. Measurable reflectance increases occurred in all three categories of burn severity by the following year as determined from a Landsat image acquired in August 1978. Regrowth of vegetation in the one year period following the fire was measured using Landsat digital data and compared with field measurements from selected portions within the burned area. Live ground cover increased 50 percent in the severely burned portions of the burned area and 50 - 75 percent in the lightly burned portions as determined from field measurements. Landsat-derived measurements showed an increase of 62.5 percent in reflectance for the severely burned areas, and 53 percent for the light burned areas, which is attributed to regrowth of vegetation. The most severely burned portion of the burned area decreased by 9.6 square kilometers in area in approximately one year according to measurements made using Landsat data. Within the lightly burned portion, 5.9 square kilometers had completely recovered based on spectral response. Prefire terrain and vegetation conditions were also found to influence burn severity. Field measurements showed that high relief areas generally burned more completely than lower lying areas. Satellite data before and after the fire confirmed this for much of the burned area.

*NASA/Goddard Laboratory for Atmospheric Sciences, Code 913, Greenbelt, MD 20771

**U.S. Army Cold Regions Research and Engineering Laboratory, Fairbanks, AK 99701

†U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH 03755

N81-21696

Paper 88

Reprinted from *Toxic Dinoflagellate Blooms*, Taylor, D. C. and Seliger, H. H., Ed., Elsevier/North Holland, NY, pp. 303-308, 1979.

PROSPECTS FOR MEASURING PHYTOPLANKTON BLOOM EXTENT AND PATCHINESS USING REMOTELY SENSED OCEAN COLOR IMAGES: AN EXAMPLE

JAMES L. MUELLER
Laboratory for Atmospheric Sciences, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771

ABSTRACT

An ocean color image measured from a high altitude aircraft is presented and discussed. The image shows large scale (25 to 60 km) turbidity patches off the west coast of Florida during a major Gymnodinium breve bloom in October 1977. That at least one of the larges patches was G. breve is suggested by a single in situ measurement made the previous day.

INTRODUCTION

The location, size, and shape of bloom patches are widely recognized to be important, but poorly known, parameters in marine phytoplankton ecology. For toxic dinoflagellates in p ticular, these parameters relate to questions concerning site and oceanographic conditions associated with the initiation of major blooms. More generally, spatial extent and distribution data are needed to refine estimates of overall standing stock during blooms, and to identify coherence between phytoplankton patch structure and patterned physical phenomena such as river plumes, current deflections by bottom topography, and coastal upwelling fronts.

Bloom extent and large-scale patchiness have been difficult to determine in the past. Coastal regions affected by dinoflagellate blooms, for instance, typically extend tens of kilometers offshore and hundreds of kilometers alongshore. Even when ships are augmented by helicopters, in situ sampling is too slow and expensive to adequately define phytoplankton distributions over areas this large. Reconnaissance by light aircraft, coupled with aerial photography, can provide useful information on patchiness, but only over limited areas.

The Coastal Zone Color Scanner (CZCS) recently placed in earth orbit on Nimbus-7 offers an exciting new tool for mapping phytoplankton bloom patches over individual scenes approximately 700 km x 700 km in extent. The CZCS is a six channel visible and infrared scanning radiometer having a spatial resolution of approximately one km. A detailed description of the CZCS, including data formats, is given by Hovis[1].

A growing body of literature[2-7] documents significant correlations between ocean color and chlorophyll concentration. Further research is required to derive a reliable chlorophyll algorithm using CZCS data, and to determine its precision. Nevertheless, success is assured for at least usefully mapping large-scale patchiness structure in phytoplankton blooms with the CZCS; relative turbidity estimates are sufficient for this particular problem. In order to illustrate this potential use of CZCS data, I will present and discuss an ocean color image of phytoplankton patches measured from an airplane with a prototype instrument. This image was measured off the west coast of Florida during a known Gymnodinium breve bloom.

OBSERVATIONS AND RESULTS

While transiting to fly Ocean Color Scanner (OCS) imagery over the Dry Tortugas in support of CZCS Nimbus Experiment Team research, the crew of a NASA-chartered Lear-Jet observed large patches of brown phytoplankton near the Florida coast. The next day, 14 October 1977, they flew the OCS to measure the two image segments shown here in Figure 1. The jagged edges of the image are due to aircraft roll, for which each scan line is corrected using data from the airplane's inertial navigation system. The geographical locations of the two image segments are outlined as rectangles in Figure 2. In order to depict the bottom topography affecting radiance distributions in Figure 1, the approximate isobaths for 2, 4, 6, and 10 m are also sketched in Figure 2.

The OCS is a 10 channel scanning radiometer built as a prototype for the CZCS. The ground resolution (pixel size) of the OCS is approximately 50 m. The channels are each approximately 25 nm wide with center wavelengths of: 425, 462, 502, 548, 592, 636, 672, 710, 756, and 795 nm.

Figure 1* is a false color representation of radiances measured by the OCS at 592 nm. Over the water, lowest radiances are black (<37), and radiances increase through blue (57), pale green (77), olive drab (97), golden brown (117), brown (137) and reddish brown (157). The numbers in parenthesis are digital counts of OCS system output and are related to radiance by a constant gain factor. The three highest radiance levels indicated on the color bar, gold (177), yellow (197) and white (217), do not occur over water in Figure 1. There is a slight color mismatch between the image and the color bar, so that in the image, golden brown (117) can easily be mistaken for yellow (197), and pale green (77) for white (217).

Fig. 1.* Ocean color radiance patterns measured at 592 nm with an airborne ocean color scanner off the west coast of Florida during a Gymnodinium breve bloom (red tide event). Refer to the main text for interpretation of patchiness and for an explanation of the color coding used to contour radiance levels (Courtesy of NASA Lewis Research Center.).

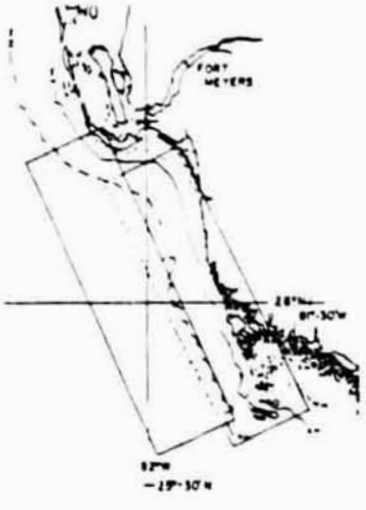


Fig. 2. Chart of the Florida coas showing the geographic locations of the two OCS image segments of Fig. 1 as overlapping rectangles. Isobaths for 2 m (.____.), 4 m (....), 6 m (_____), and 10 m (_____ ___ ___) are sketched (based on NOAA marine navigational charts) to aid in identifying bottom reflection patterns in Fig. 1.

Upwelled radiance patterns in Figure 1 are attributable to two competing phenomena. At any pixel, radiance N upwelled beneath the sea surface is given by the sum $N = N_b + N_w$ where the two radiance components are:

1) Daylight reflected from the sea bottom, N_b, which varies approximately as:

$$N_b \approx H A e^{-2kd} \qquad (1)$$

where H is incident irradiance, A is bottom albedo, K is the vertical diffuse attenuation coefficient of sea water, and d is water depth.

2) Daylight backscattered by sea water and suspended particles, N_w, which varies approximately as:

$$N_w \approx H \frac{r}{2k} (1 - e^{-2kd}) \qquad (2)$$

where r is the volume reflectance of sea water.

Approximations (1) and (2) show that the relative contributions of N_w and N_b vary with r, k, and d. Reflectance r and diffuse attenuation k both increase with increasing turbidity, although not generally in a fixed relationship to each other. Therefore, for fixed d, the ratio N_w/N_b will increase with increasing turbidity. And for horizontally constant r and k, N_w/N_b will increase with increasing depth d.

In the absence of phytoplankton blooms, or of turbidity plumes from terrestrial runoff, sea water is typically clear over the inner Florida shelf. Middleton and Barker[8], using OCS data for clear water conditions, showed that bottom depths near Tampa Bay, Florida, are highly correlated with upwelled radiance at 550 nm from 0 to 12 m depth (squared correlation coefficient = 0.94). To interpret Figure 1, I will assume on the basis of this correlation, that under turbid conditions N_b dominates in depths less than 6 m, N_w dominates in depths greater than 12 m, and that both N_w and N_b are important in depths between 6 and 12 m.

Referring to Figures 1 and 2 together, the general bottom slope, except for the Cape Romano Shoals at 25°-45'N, 81°-40'W, implies a monotonic decrease in N_b with distance offshore. The high radiance patterns (golden brown to reddish brown) at 25°-45'N, 81°-40'W clearly delineate the complicated structure of the Cape Romano Shoals. Further north along the shore, the reddish brown to olive drab banding is consistent with the trend of bottom slope inshore of the 6 m isobath. Further offshore, however, are three large, anomalously bright (pale green to olive drab) areas which are not consistent with N_b patterns implied by bathymetric gradients. These areas, centered approximately at 26°-10'N, 81°-55'W, at 26°-20'N, 81°-15'W, and at 25°-55'N, 82°-00'W,

are therefore identified as patches of high turbidity. The largest patch, extending from 25°-38'N, 81°-50'W to 25°-10'N, 82°-05' W, is labelled "Brown Phytoplankton" on the basis of visual observations by the flight crew.

At the time of this flight, the Florida Department of Natural Resources (FDNR) laboratory in St. Petersburg, Florida was following the progress of a major Gymnodinium breve bloom. The bloom moved southward down the west coast of Florida, passed through the Florida Straits, and subsequently reappeared on the east coast. On 13 October 1977, FDNR personnel measured a G. breve concentration of 1.5×10^6 cells/liter in position 26°-20'N, 82°-12'W [9]. This position falls in the middle of the northernmost turbidity patch of Figure 1, which was measured on 14 October 1977. This coincidence strongly implies that the turbid area at 26°-20'N, 82°-15'W in Figure 1 was a G. breve bloom patch approximately 25 km in diameter. It is intriguing to speculate that the larger (25 km x 60 km) patch to the south was also likely to have been G. breve which was not detected in FDNR's sampling program. Certainly, had Figure 1 been available to FDNR in near real-time, they could have modified and extended their sampling to confirm and quantify the apparent G. breve extent and patchy structure.

SUMMARY

Figure 1* is a remotely measured ocean color image showing what is probably the patchy structure of a G. breve bloom extending over a 60 km x 100 km area off the west coast of Florida. This conclusion is based on visual inspection of bathymetry to infer bottom reflection trends, and on a single ground truth measurement of G. breve concentration obtained independently (and fortuitously) by FDNR. The image shows coherent bloom patches which extend scales up to 60 km in length.

The scientific value of this OCS data would, of course, be greater with more extensive ground truth data and additional analysis to quantify, and improve discrimination between, turbidity and bottom reflection effects. Nevertheless, Figure 1 conclusively demonstrates the strong potential utility which remote ocean color sensors hold for future studies and monitoring of phytoplankton bloom extent and patchiness.

ACKNOWLEDGEMENTS

This work was supported by the NASA Environmental Quality Program (RTOP 176-30-41). I am indebted to Mr. Jack Salzman of NASA Lewis Research Center, Cleveland, Ohio for the OCS data and color image

presented here in Figure 1.

REFERENCES

1. Bovis, W.A. (1978) in: The Nimbus-7 User's Guide. Landsat/Nimbus Project; NASA Goddard Space Flight Center, Greenbelt, Md., pp. 19-31.
2. Clark, G.L., Ewing, G.C. and Lorenzen, C.J. (1970) Science 167, 1119-1121.
3. Arveson, J.C., Weaver, E.C., and Millard, J.P. (1971) American Institute of Aeronautics and Astronautics Paper No. 71-1097.
4. Mueller, J.L. (1973) Ph.D. Thesis Oregon State University, Corvallis. 239 p.
5. Mueller, J.L. (1976) Applied Optics 15, 394-402.
6. Morel, A. and Prieur, L. (1977) Limnol. Oceanogr. 22, 709-722.
7. Neville, R.A. and Gower, J.F.R. (1977) J. Geophys. Res. 82, 3487-3493.
8. Middleton, E.M. and Barker, J. (1976) NASA Goddard Space Flight Center Report X-923-76-111. Greenbelt, Md. 6 p.
9. Roberts, b. (1978) Personal communication. Florida Dept. of Natural Resources, St. Petersburg, Florida.

● Publisher's note: Color photograph may be received from author upon request.

Paper 89

NASA Technical Memorandum 80296 July 1979

RESULTS OF A STATISTICAL APPROACH TO RAINFALL ESTIMATION USING NIMBUS 5 6.7μm AND 11.5μm THIR DATA

James P. Ormsby

ABSTRACT

Nimbus 5 6.7μm and 11.5μm Temperature Humidity Infrared Radiometer (THIR) data were used in a simple multiple regression scheme to test the feasibility of using these data to estimate hourly rainfall. Throughout the test area (85°W to 105°W and 45°N to 30°N) subareas (8°X6°) were chosen from which point to point and areal statistics were obtained.

Four subsets of data were used. The first consisted of only those surface stations indicating precipitation whose latitude and longitude coincided with the THIR grid points. A second used surface stations 0.1 degree from the THIR grid points. The third was a combination of subsets one and two. A reciprocal distance weighting scheme was used to derive precipitation values in data sparse areas. A fourth subset was made using these data combined with the data from subsets one and two.

Point estimates resulted in negative correlations between estimated and grid derived "surface" precipitation. One degree areal estimates showed a slight improvement with a correlation coefficient of ~ 0.11. Single regression areal estimates resulted in correlations of ~ 0.11 and 0.20 for the 6.7μm and 11.5μm data respectively.

These poor results were attributed to problems which are inherent in the satellite data (location errors, short temporal span of data, wavelength of sensors, etc) and the lack of sufficient surface data to better verify the satellite estimate.

N81-21697

Paper 90

ARCTIC SEA ICE DECAY SIMULATED FOR A CO_2-INDUCED TEMPERATURE RISE

CLAIRE L. PARKINSON
NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, U.S.A.

and

WILLIAM W. KELLOGG*
National Center for Atmospheric Research, Boulder, Colorado 80307, U.S.A.

Abstract. A large scale numerical time-dependent model of sea ice that takes into account the heat fluxes in and out of the ice, the seasonal occurrence of snow, and ice motions has been used in an experiment to determine the response of the Arctic Ocean ice pack to a warming of the atmosphere. The degree of warming specified is that expected for a doubling of atmospheric carbon dioxide with its associated greenhouse effect, a condition that could occur before the middle of the next century. The results of three 5-year simulations with a warmer atmosphere and varied boundary conditions were (1) that in the face of a 5 K surface atmospheric temperature increase the ice pack disappeared completely in August and September but reformed in the central Arctic Ocean in mid fall; (2) that the simulations were moderately dependent on assumptions concerning cloud cover; and (3) that even when atmospheric temperature increases of 6–9 K were combined with an order-of-magnitude increase in the upward heat flux from the ocean, the ice still reappeared in winter. It should be noted that a year-round ice-free Arctic Ocean has apparently not existed for a million years or more.

1. Introduction

During the past several years considerable attention has been given to the probable effects on climate of the increase in atmospheric carbon dioxide (CO_2) due to the burning of fossil fuels. Since carbon dioxide absorbs infrared radiation, some of which would otherwise escape to space, an expected effect of adding CO_2 to the lower atmosphere is to raise the average surface temperature of the earth. Although there is still much uncertainty, a variety of climate models and model comparisons are converging on a temperature response approximating a 1.5–3 K increase in global mean surface temperature for a CO_2 doubling (e.g., Manabe and Wetherald, 1967: 2.3 K; Manabe and

The calculations for this work were carried out while both authors were at the National Center for Atmospheric Research (NCAR), which is sponsored by the National Science Foundation.

*Currently on leave, working for the World Meteorological Organization in Geneva, Switzerland, on the World Climate Programme.

Climatic Change 2 (1979) 149–162. 0165–0009/79/0022–0149$01.40.
Copyright © 1979 by D. Reidel Publishing Company

Wetherald, 1975: 2.9 K; Schneider, 1975: 1.5–3.0 K; Augustsson and Ramanathan, 1977: 2.0–3.2 K); and the response in the polar regions is expected to be several times larger than that in the equatorial belt, according to model studies by Manabe and Wetherald (1975), Sellers (1974), and Ramanathan *et al.* (1979). These latter results are corroborated by noting that, in the event of natural changes, the real atmosphere also displays a larger response in the polar regions (SMIC, 1971; van Loon and Williams, 1976, 1977; Borzenkova *et al.*, 1976). It is believed that the doubling of atmospheric carbon dioxide could be achieved by the middle of the next century, based on the assumption that the rate of fossil fuel consumption will continue to increase (Broecker, 1975; Keeling and Bacastow, 1977; Siegenthaler and Oeschger, 1978; Kellogg, 1979).

Although the magnitude of the effect is uncertain, any temperature rise in the polar regions would unquestionably influence the floating pack ice of the Arctic Ocean. An expected consequence of a temperature increase would be an increased downward heat flux at the surface and a decrease in both the areal extent of ice and ice thickness. This would produce at least one strong feedback effect, serving to further increase atmospheric temperatures. Since ice reflects a far greater percentage of incoming solar radiation than does open water, a lessened areal extent of floating pack ice would increase the amount of solar radiation absorbed at the air-ocean interface. Thus the polar surface temperature increase due to CO_2 could be expected to be reinforced by this positive feedback created through the reduction of Arctic sea ice.

The temperature response from a global doubling of atmospheric carbon dioxide would likely be strongest in the surface layer of the polar regions for several additional reasons. Bolstering the positive feedback mentioned above, the stable stratification in the polar troposphere suppresses vertical mixing and hence limits the amount of turbulence transferring the surface heat upward (Manabe and Wetherald, 1975). As a result of such regional peculiarities, the temperature response in the surface layer of the polar regions could well be 3–4 times that of the global average of 1.5–3 K for a doubling of CO_2, and a 10 K value is estimated by the model of Manabe and Wetherald (1975) for the average annual increase at 80 °N. An improved model by the same authors (1979) suggests a 7–8 K response at 80 °N for a CO_2 doubling.

The specific question of what would happen to the Arctic Ocean ice pack in the face of such a warming trend has been investigated by Budyko (1966, 1974) with a fairly simple heat-balance model, and he concluded that a 4 K rise in summertime temperature would be sufficient to cause the ice pack to disappear entirely. He conjectured, furthermore, that once the ice pack had disappeared it would not reform in winter. In the present study the same question is asked and the results from a more complete model of the pack ice are presented.

Numerical modeling can clearly serve a useful purpose in evaluating hypothetical scenarios of the future. Although mankind might inadvertently perform the 'real world experiment' by indeed doubling the atmospheric carbon dioxide content, thereby allowing direct observation of its consequences in the polar regions, it is less risky and much quicker to perform the experiment numerically. In doing so, however, one must remain aware that the response of a model is not necessarily the response which the real world

would have, especially in view of the large number of interactive processes essentially ignored in the simulations (Kellogg, 1975). Nevertheless, model results can provide an estimate of possible effects and can add an additional component to our understanding of past and future climatic changes.

Thus, for instance, we accept as a starting point the results of Manabe and Wetherald (1975, 1979) estimating a significant temperature increase in the polar regions, recognizing though that their specific numerical results, like the specific results to be presented here, must be accepted with some caution. The highly idealized land-sea geography of Manabe and Wetherald and the lack of realistic ice-climate feedback mechanisms certainly affect their results; how significantly and in which direction remain unknown. Their calculations provide at least a first-guess estimate, with 10 K and 7–8 K polar surface temperature increases. Utilizing their qualitative results but proceeding conservatively, in the initial experiment below we insert into an Arctic sea ice model a surface air temperature increase of 5 K.

The numerical model used in the current study is described briefly in the following section and in more detail by Parkinson and Washington (1979). It is a thermodynamic-dynamic model with the thermodynamics based largely on the one-dimensional calculations of Maykut and Untersteiner (1969, 1971) as simplified by Semtner (1976) and with the dynamics based on the five major stresses acting on the ice pack. The object of the model is to simulate the large-scale geographical growth and decay of sea ice over a yearly cycle. It accepts both atmospheric and oceanic variables as given and does not adjust these according to simulated sea ice conditions. Thus in this study we examine simply the first order effects of an atmospheric temperature increase on the sea ice and not the resultant effects of the change of ice on the atmosphere and ocean. In the future, when ice, atmosphere, and ocean models become fully coupled, it will be possible to perform such experiments in a totally interactive mode, something not yet possible (Kellogg, 1977, 1979).

2. Sea Ice Model

The sea ice model used to perform the numerical experiments is described in detail in Parkinson and Washington (1979). Here we summarize some of its more important features. The calculations are carried out with an 8-hour timestep on a rectangular grid superimposed over a polar stereographic projection (Figure 1). The grid covers the majority of the northern hemisphere's ice-covered oceans and has a horizontal resolution of roughly 190 km in both x and y directions. A corresponding grid is used for the southern hemisphere (Parkinson and Washington, 1979), but this is not utilized in the current study.

Vertically, at each ocean grid point there are potentially 4 layers – ocean, ice, snow, and atmosphere. There is also a lead parameterization, so that each grid square has a variable percentage of its area assumed to be ice-free (Figure 2).

The principal object of the model is to simulate a reasonable yearly cycle of the thickness and extent of the ice. This naturally requires calculations of ice and snow

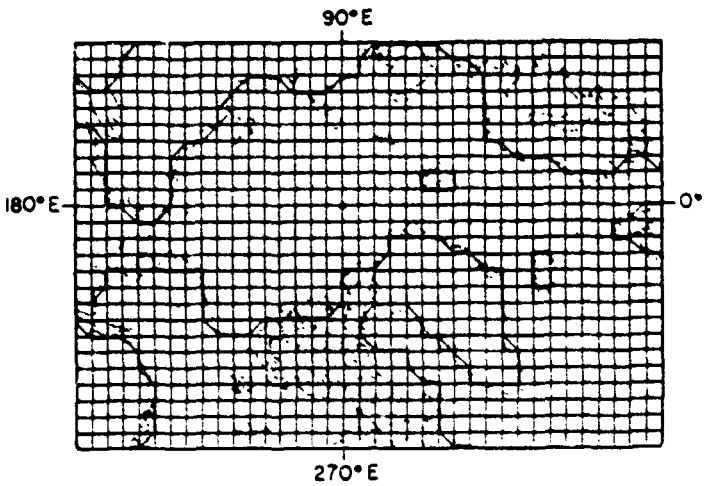


Fig. 1. Horizontal grid. Dotted lines trace the continental boundaries from standard polar stereographic projections. Model resolution of those boundaries is indicated in solid lines, as is the grid itself.

accretion and/or ablation, calculations which are here based on energy balances at the various interfaces. At the upper snow surface the modeled fluxes are sensible and latent heat, incoming longwave radiation, incoming shortwave radiation, and emitted longwave radiation from the surface to the atmosphere. These fluxes are balanced with the only heat flux modeled through the snow – the conductive flux – and the balance is used to calculate the change in surface temperature. Should the calculated temperature be above freezing, then the temperature is set exactly at freezing and the excess thermal energy is used to melt a portion of the surface snow.

At the ice-snow interface the energy balance simply requires the conductive heat flux through the snow to equal the conductive flux through the ice. Finally, at the bottom of the ice, the downward conductive flux and a constant oceanic heat flux from water to ice are combined to determine the appropriate bottom ablation or accretion. This latter

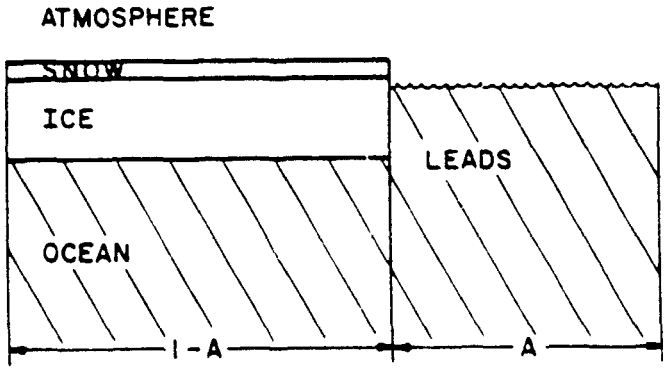


Fig. 2. Schematic diagram of the major vertical divisions within a grid square

calculation proceeds by assuming that the temperature at the bottom of the ice remains continuously at the freezing point. The amount of ablation/accretion is calculated so that the latent heat absorbed/released balances the energy excess/deficit from the other two fluxes at the interface.

The insertion of leads within the ice pack centers on a determination of the net energy input into or out of the lead. In the event of a positive net input, the energy is used to increase water temperature and laterally melt the ice. In the event of a negative net input, the deficit is balanced by cooling the water, and, should the temperature fall below freezing, then by laterally accreting new ice onto the existing ice.

The movement of the ice is determined in two steps, calculated for each timestep sequentially after the above thermodynamic calculations. The first step computes ice velocities through requiring a steady state balance between four forces – the wind stress from above the ice, the water stress from below the ice, the Coriolis force due to the rotation of the earth, and the force due to the dynamic topography or the tilt of the sea surface. The second step alters a portion of the steady state velocities by proportionately reducing all velocity vectors entering grid squares with greater than 99.5% ice coverage. In this manner the model simulates the last of the five major forces acting on the ice – internal ice resistance – and it specifies a minimum of 0.5% leads for all grid squares. Furthermore, at coastal regions, velocity components toward the coast from adjacent grid squares are set to zero, thereby preventing the disappearance of the ice from the active region of the grid. This is a brute-force method of accounting for internal ice resistance and fails to allow floe interactions to produce momentum increases as well as decreases. Hence the dynamical effects are somewhat underestimated, perhaps accounting for the finding by Parkinson and Washington (1979) that the ice dynamics had only a secondary effect on the extent and thickness of the ice. For more elaborate treatments of ice dynamics the reader is referred to Hibler (1979) and to Coon *et al.* (1974).

3. Model Experiments with Increased Temperatures

Among the input data to the basic model are mean monthly atmospheric temperatures at each grid point. These temperature data are required for calculating the fluxes of sensible heat and incoming longwave radiation for the energy balance at the upper earth surface, whether the surface cover is snow, ice, or water. The data are obtained from the long-term mean monthly distributions presented by Crutcher and Meserve (1970). Similarly, the same source is used for mean monthly dew point temperatures and for mean monthly geostrophic winds, both of which remain unchanged for the following set of experiments.

Use of the Crutcher and Meserve data fields has resulted in a fairly realistic yearly cycle of sea ice growth and decay (Figure 3). The yearly cycle has reached an approximate equilibrium state by year 5 of the simulation, with the simulated ice varying in extent from a minimum in September to a maximum in March. In September, the ice covers only a portion of the Arctic and has receded from most coastlines, with ice thicknesses reaching 3.0 m in the center of the pack. At the March maximum, ice extent

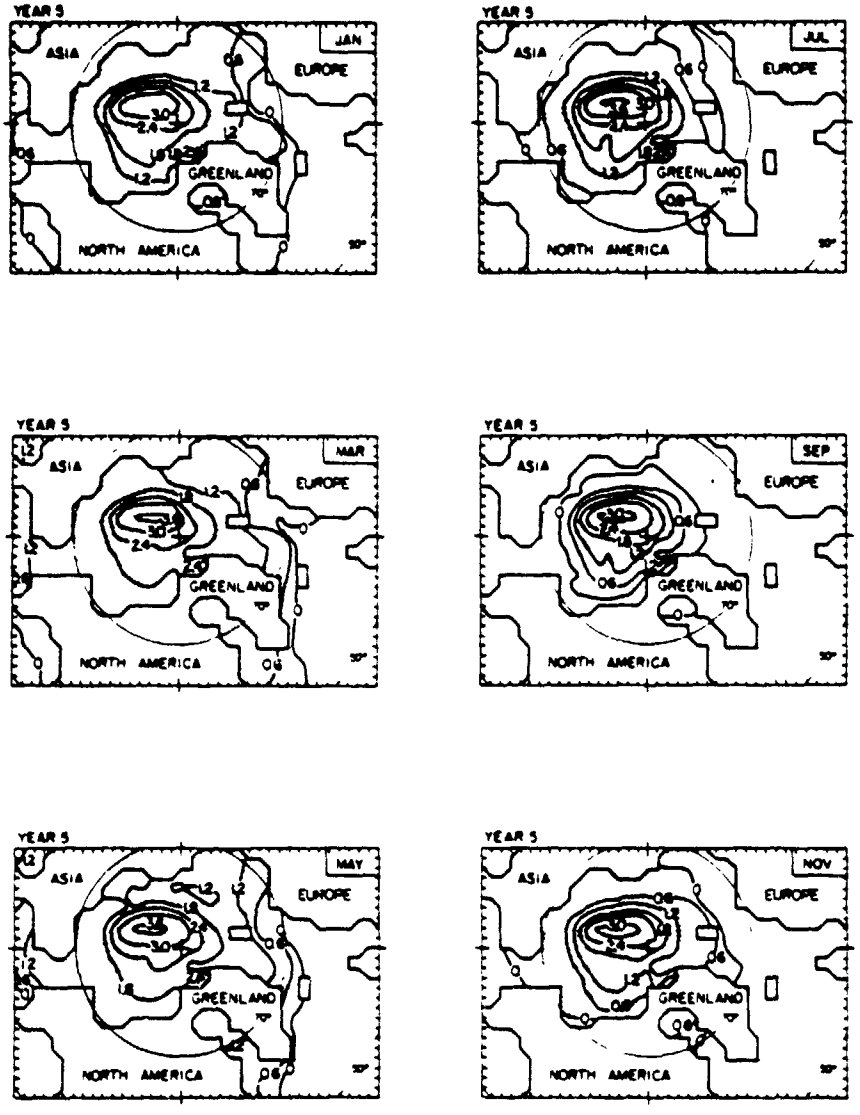


Fig. 3. Simulated yearly cycle of sea ice in the Standard Case, based on current climate. Contours show ice thickness in meters, while shading indicates ice compactness above 90%.

has greatly increased, reaching Iceland, well into the Bering Sea, and beyond the southern coast of Greenland. Central Arctic thicknesses in winter reach 3.6 m, and ice concentrations exceed 97% for the entire Arctic basin. The model results regarding ice extent and concentration can be compared to modern satellite observations, e.g. those presented in Gloersen *et al.* (1978).

The initial model run with current climatic conditions will hereafter be referred to as the Standard Case. It is the same Standard Case as appears in Parkinson and Washington (1979), where a fuller discussion can be found regarding the results.

3.1. Experiment 1

The first experiment with altered atmospheric conditions simply increases all air temperatures by 5 K, half the annual average temperature increase predicted by Manabe and Wetherald (1975) for 80 °N under a doubled CO_2 concentration. Operationally, at each grid point the 12 specified mean monthly temperatures are now 5 K greater than the Crutcher and Meserve (1970) values used above.

As expected, the effect on the ice of an increase in atmospheric temperatures is a decrease in both thickness and extent (Figure 4). Performing the experiment, however, quantifies the qualitative expectations. By year 3 of the model simulation the thickest January ice is just under 1 m and the thickest July ice is on the order of 0.3 m. The ice totally disappears in August of the third year, though it did not do so in either of the two previous years. It reappears in October and, interestingly, by January of year 4 has become slightly thicker than in January of year 3. The July results also show a thickening of the ice between years 3 and 4, as reflected in the further outward extension of the 0.3 m thickness contour. Ice concentrations too have increased somewhat between years 3 and 4, after having decreased strongly from year 1 and apparently slightly 'overshooting'. The simulation appears to be near an equilibrium then, as the results for years 4 and 5 are nearly identical. (In contrast to some interesting one-dimensional thermodynamic results of Semtner (1976), there is no multi-year cycle produced in the current study.) In each of years 3–5 there is an ice-free Arctic during August and September, with the ice cover reestablishing itself for the remainder of the year.

Clearly the 5 K increase in atmospheric temperatures has caused a significant reduction in simulated ice thicknesses and extents (Figure 3 versus Figure 4). Although the calculated ice boundaries and thicknesses are not predictions, this simplified experiment strongly suggests that an increase in atmospheric temperatures of 5 K drastically reduces the Arctic ice cover.

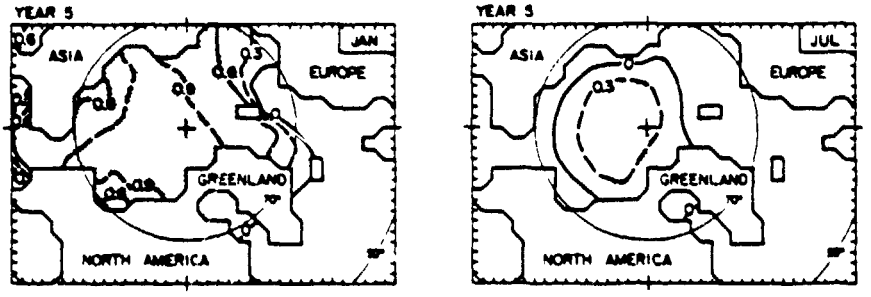


Fig. 4 January and July ice thicknesses simulated in Experiment 1. Contours show thickness in meters, while shading indicates ice compactness above 90%.

3.2. *Experiment 2*

Of course, in the real world one would expect that a temperature increase from increased carbon dioxide would be accompanied by other atmospheric, oceanic, and cryospheric alterations. However, the real world has proven so complex that hypotheses regarding the various interactive effects which would accompany a temperature increase must all remain mere educated guesses, though some are more firmly based than others. (Someday, it is hoped, fully interactive atmosphere-ocean-ice models will be developed, but none exists today.) It is for that reason that in Experiment 1 we altered only the surface air temperature, thereby eliminating ambiguity and isolating the precise cause of the simulated sea ice reductions. To attempt more realism, however, we have tried a sequence of additional experiments where atmospheric temperature increases are accompanied by additional changes in the input parameters. We will describe two such related experiments, labelled Experiments 2 and 3.

In Experiment 2 the temperature increase is greater than the 5 K increase in Experiment 1, but it remains conservative compared to Manabe and Wetherald's predicted 10 K increase at 80 °N. Furthermore, the increase varies seasonally, following the contention of Budyko (1974) that the atmospheric temperature response from an increased CO_2 concentration should be larger in winter than in summer. This seasonal contrast seems fairly consistent with the calculations of Ramanathan *et al.* (1979) when the expected contrast in low cloudiness between winter and summer is taken into account. (Since the Manabe and Wetherald model is seasonally averaged, it cannot provide information on the relative temperature response in winter versus summer.) In Experiment 2, we insert a seasonal dependence into the atmospheric temperature increases, in accordance with the speculated larger temperature response in winter. The specified increases over those of Crutcher and Meserve range from 6 K in July and August to 9 K in January and February (Table I).

As the increased temperature suggests a larger amount of available thermal energy at the surface, and as Experiment 1 has already demonstrated that the simulated ice cover is dramatically reduced under a 5 K temperature increase, it is assumed that the temperature increase will be accompanied by greater evaporation from the ocean surface and hence an increased cloud amount in all seasons. Cloud cover enters the model calculations both in determining incoming solar radiation, where the cloud reduces the magnitude of the flux reaching the surface during the daylight hours, and in the calculation of incoming longwave radiation, where the cloud increases the flux due to the radiative properties of

TABLE I: Specified increase in atmospheric temperatures over the mean monthly space-dependent values used in the Standard Case. Units are degrees Kelvin.

	J	F	M	A	M	J	J	A	S	O	N	D
Experiment 1	5	5	5	5	5	5	5	5	5	5	5	5
Experiment 2	9	9	8	8	7	7	6	6	7	7	8	8

TABLE II: Specified percent cloud cover.

	J	F	M	A	M	J	J	A	S	O	N	D
Standard Case and Experiment 1	50	50	50	55	70	75	75	80	80	70	60	50
Experiment 2	60	60	60	65	75	80	85	85	85	80	70	60

the cloud droplets. In the Standard Case the ice model employs spatially-independent cloud percentages averaged from the curves of Huschke (1969) for four Arctic regions. These values range from a 50% cloud cover in winter to an 80% cloud cover in August and September (Table II). In Experiment 2 we increase the prescribed cloud amounts to account for the presumed increased evaporation from the greater areal extent of open water. The cloud percentages used in Experiment 2 are listed in Table II.

Results from Experiment 2 are presented in Figure 5. The ice disappears in August, September, and October, while the maximum thickness occurs in April with 1.7-m-thick ice in portions of the Central Arctic. In fact, even as late as mid June there remains ice of thickness exceeding 1.3 m, leaving the predominant thickness decrease to occur between mid June and mid July, during which period maximum thicknesses drop from 1.3 m to just over 0.3 m (Figure 5). In Experiment 1, also, maximum thicknesses of 1.7 m occur in April and the bulk of the thickness decrease takes place between mid June and mid July, with maximum thicknesses decreasing from 1.6 m to 0.4 m. By comparing both Experiments 1 and 2 with the Standard Case it can be seen that the specified changes of temperature and cloud cover dramatically reduce the extent of the ice in summer but in winter reduce the extent far less significantly than the thickness. Climatologically it is presumably the extent, not the thickness, which has the greatest feedback effects.

3.3. Experiment 3

The final change we have included is an increase in the ocean heat flux. It is this ocean heat flux from the water to the undersurface of the ice which combines with the

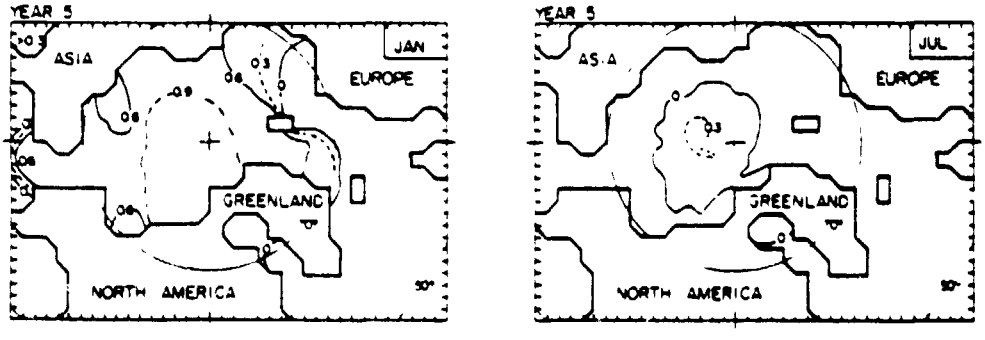


Fig. 5. January and July ice thicknesses simulated in Experiment 2.

conductive flux through the ice to determine the amount of bottom ablation or accretion. The Standard Case uses an ocean heat flux of 2 W m^{-2} after the work of Maykut and Untersteiner (1971).

We hypothesize an increased ocean heat flux in the event of a major recession of the ice pack for the following reasons: the current flux is small largely because of the salinity stratification in the upper 50 m of the water, which produces a stable upper layer (Aagaard and Coachman, 1975). In the event of a significant increase in the areal extent of open water in the Arctic, there would be mechanical mixing due to wave action and quite possibly the halocline would be reduced or eliminated. Thus in Experiment 3 we have increased the upward heat flux to account for the likelihood of changed oceanic conditions. The other elements of the model are the same as in Experiment 2.

One purpose of this additional experiment was to see if in this case the model would respond to a total disappearance of ice in mid summer by not allowing ice to reform in winter. Although the oceanic heat flux was increased by over an order of magnitude, from 2 to 25 W m^{-2}, winter ice did still form and sustain itself throughout the season. The higher flux value, incidentally, is comparable to the present estimated oceanic heat flux in the Southern Ocean (Parkinson and Washington, 1979), where ice forms abundantly in winter, so our result was not entirely unexpected. However, the increased flux naturally tends to increase bottom ablation, resulting in lowered ice thicknesses and lessened ice extent.

Comparing the results of Experiment 3 to those of Experiment 2, the difference in ice extents is most apparent in the rapidity of the retreat of the ice with the oncoming of summer. The May ice extents are decidedly less in Experiment 3, especially in the area of the Bering Straits, and the June extents are much more so. By mid July the ice has disappeared in Experiment 3, whereas it remains until August in Experiment 2. In both cases ice reappears by mid November. Ice thicknesses in winter tend to be lower by about 30–50% in the final experiment, as a result of the additional oceanic heat flux.

It was mentioned above that Budyko (1966, 1974) hypothesized that once the ice pack had disappeared from the Arctic Ocean it would not reappear in winter, and the SMIC Report (1971) (p.160) considers that this is at least a possibility. The present model allows ice to form as soon as a negative surface energy balance reduces the water

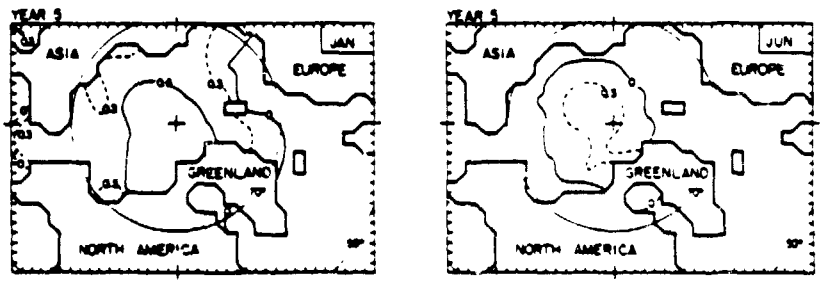


Fig. 6. January and June ice thicknesses simulated in Experiment 3. There is no ice simulated in July.

temperature to 271.2 K. This occurs by mid November in Experiment 3 in spite of the increased air temperatures and the greater tendency for ice ablation due to the large upward transport of heat from the ocean. The simulated return of the ice depends strongly on the atmospheric temperature assumptions, but since ice forms and thickens in the model in response to energy balances and not strictly temperatures, this dependence is more complicated than a simple one-to-one correspondence. Both the wintertime equatorward expansion of the ice extent and the summertime poleward decay tend to lag the movement of the atmospheric freezing line. This is true both in the model and in the real world.

Of course there are many physical factors involved in the *initial* formation of a layer of ice in mid ocean, factors not presently being modeled, including the effects of waves, the stability of the upper ocean layers as the freezing point is approached, the salinity distribution, and the critical size of a 'freezing nucleus' of pack ice required to permit new ice to form at its edges. Furthermore, once the retreat of the ice begins, in reality there would be both positive and negative feedbacks interacting among ice, ocean and atmosphere. For instance, the replacement of an ice cover by open water immediately decreases surface albedo and presumably thereby further increases surface air temperatures. However, the wider expanse of open water also suggests greater evaporation and snowfall, thereby feeding back negatively with increased albedos over the ice and surrounding land areas. Some future model may take all these various factors into account, thus permitting much firmer conclusions than can be drawn here.

4. Discussion

The present study adds to the accumulating body of literature on the potential changes in the polar regions due to a further increase of atmospheric carbon dioxide and a consequent global warming. Mercer (1978) ominously forewarns of a West Antarctic deglaciation induced by the temperature increase and indicates that deglaciation could well begin within 50 years and proceed rapidly, eventually causing a rise in global sea levels of about 5 m, thus drowning many coastal cities. Our results are less dramatic, as the melted Arctic sea ice would not cause a change in sea level. However, the melting of the Arctic ice would replace the highly reflecting ice pack with a dark ocean surface, thus causing a substantial increase in solar energy absorption during the spring, summer, and fall. As suggested earlier, this would feed back to the atmosphere a further temperature increase, thereby magnifying the initial perturbation. Another consequence of the open Arctic would be an increase in evaporation from the surface and hence an increase in precipitation over both the ocean itself and the adjacent land areas.

While its occurrence is highly unlikely in the next century, there have been rapid cooling episodes in the past (Flohn, 1974, 1975) due to natural causes, and if such a cooling were to take place, it should produce effects that are the reverse of those for a warming. Hence we should then obtain an increase in the extent of Arctic sea ice, a resulting increase in surface albedos and decrease in absorbed solar radiation, and

consequently a further temperature reduction and perhaps even a northern hemisphere glaciation or 'ice age'.

At this moment, ice age theories are numerous and somewhat contradictory, but many do include the effect of the Arctic ice cover in a prominent position. For instance, Ewing and Donn (1956) hypothesize that an ice-free Arctic would lead to an ice age due to the increased evaporation and snowfall. However, evidence tends to discount the existence of ice-free Arctic conditions at any time during the past million years or more (Hunkins *et al.*, 1971), a period during which glacial and interglacial periods have alternated repeatedly. Most theorists assume that during major glaciations the sea ice was more extensive than it is now. Among the most recent theories in which an extended sea ice cover could contribute significantly to precipitating ice age conditions is that of Beaty (1978). In Beaty's synthetic model, the primary requirement for initiation of glaciation is the existence of continents in high latitudes, and the triggering factor for a glacial episode is an increased surface albedo.

The essential point to be made in the context of this article is that man's continued insertion of carbon dioxide into the atmosphere could cause dramatic climatic consequences worldwide and that the polar regions would respond in special ways. Neither our numerical models nor our general understanding are sufficiently sophisticated to provide total confidence that the polar regions of the real world would respond in the manner that any model predicts. No model yet incorporates the full range of potential influences and feedbacks which the polar regions contain (e.g., Kellogg, 1975), though eventually our models should improve and our confidence in them increase. In the meantime, however imperfect, the current models do suggest some grave consequences, and mankind would do well to proceed with caution where possible manipulation of the world's climate is at stake.

Acknowledgements

The authors thank Dr Warren Washington for discussions and assistance in the running of the model. We also thank the NCAR Graphics Department and the Graphics Arts Branch of the Goddard Space Flight Center for drafting the diagrams.

References

Aagaard, K., and Coachman, L. K.: 1975, Toward an ice-free Arctic Ocean. *Trans. Am. Geophys. Union* **56**, 484–486.

Augustsson, T., and Ramanathan, V.: 1977, A radiative-convective model study of the CO_2 climate problem. *J. Atmos. Sci.* **34**, 448–451.

Beaty, C. B.: 1978, The causes of glaciation. *Am. Scientist* **66**, 452–459.

Borzenkova, I. I., Vinnikov, K. Ya., Spirina, L. P., and Stekhnovskiy, D. I.: 1976, Change in the air temperature of the northern hemisphere for the period 1881–1975. *Meteorol. Gidrol.* **7**, 27–35.

Broecker, W. S.: 1975, Climatic change: Are we on the brink of a pronounced global warming? *Science* **189**, 460–463.

Budyko, M. I.: 1966, Polar ice and climate. In *Proc. Symp. on the Arctic Heat Budget and Atmospheric Circulation*, J. O. Fletcher (Ed.), Rand Corp. Memorandum RM-5233-NSF, Santa Monica, California, 3–22.

Budyko, M. I.: 1974 *Climate and Life* (English edition edited by D. H. Miller), International Geophysical Series, vol. 18, Academic Press, New York and London, 508 pp.

Coon, M. D., Maykut, G. A., Pritchard, R. S., Rothrock, D. A., and Thorndike, A. S.: 1974, Modelling the pack ice as an elastic-plastic material. *Aidjex Bulletin* 24, 1–105.

Crutcher, H. L., and Meserve, J. M.: 1970, *Selected Level Heights, Temperatures and Dew Points for the Northern Hemisphere*. NAVAIR 50-1C-52 Revised, Naval Weather Service Command, Washington, D.C.

Ewing, M., and Donn, W. L.: 1956, A theory of ice ages. *Science* 123, 1061–1066.

Flohn, H.: 1974, Background of a geophysical model of the initiation of the next glaciation. *Quaternary Res.* 4, 385–404.

Flohn, H.: 1975, History and intransitivity of climate. Appendix 1.2 in *The Physical Basis of Climate and Climate Modelling*, GARP Publication Series, No. 16, World Meteorological Organization, Geneva, Switzerland, 106–118.

Gloersen, P., Zwally, H. J., Chang, A. T. C., Hall, D. K., Campbell, W. J., and Ramseier, R. O.: 1978, Time-dependence of sea-ice concentration and multiyear ice fraction in the Arctic Basin. *Boundary-Layer Meteorol.* 13, 339–359.

Hibler, W. D. III.: 1979, A dynamic thermodynamic sea ice model. *J. Phys. Oceanogr.*, in press.

Hunkins, K., Bé, A. W. H., Opdyke, N. D., and Mathieu, G.: 1971, The late Cenozoic history of the Arctic Ocean. In *The Late Cenozoic Glacial Ages*, ed. by K. K. Turekian, Yale University Press, New Haven, 215–237.

Huschke, R. E.: 1969, *Arctic Cloud Statistics from 'Air-Calibrated' Surface Weather Observations*. The RAND Corporation Memorandum RM-6173-PR, Santa Monica, California, 79pp.

Keeling, C. D., and Bacastow, R. B. 1977, Impact of industrial gases on climate. In *Energy and Climate*, Geophysics Research Board, National Academy of Sciences, Washington, D.C., 72–95.

Kellogg, W. W.: 1975, Climatic feedback mechanisms involving the polar regions. In *Climate of the Arctic*, G. Weller and S. A. Bowling (Eds.), Geophysical Institute, University of Alaska, 111–116.

Kellogg, W. W.: 1977, *Effects of Human Activities on Global Climate*, World Meteorological Organization Technical Note No. 156 (WMO No. 486), Geneva, Switzerland.

Kellogg, W. W.: 1979, Influences of mankind on climate. *Ann. Rev. Earth Planet. Sci.* 7, 63–92.

Manabe, S., and Wetherald, R. T.: 1967, Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.* 24, 241–259.

Manabe, S., and Wetherald, R. T.: 1975, The effects of doubling the CO_2 concentration on the climate of a general circulation model. *J. Atmos. Sci.* 32, 3–15.

Manabe, S., and Wetherald, R. T. 1979 On the horizontal distribution of climate change resulting from an increase in CO_2 content of the atmosphere. *J. Atmos. Sci.*, in press.

Maykut, G. A., and Untersteiner, N.: 1969, *Numerical Prediction of the Thermodynamic Response of Arctic Sea Ice to Environmental Changes*. The RAND Corporation Memorandum RM-5093-PR, Santa Monica, California, 173 pp.

Maykut, G. A., and Untersteiner, N.: 1971, Some results from a time-dependent thermodynamic model of sea ice. *J. Geophys. Res.* 76, 1550–1575.

Mercer, J. H.: 1978, West Antarctic ice sheet and CO_2 greenhouse effect: a threat of disaster. *Nature* 271, 321–325.

Parkinson, C. L.: 1978, *A Numerical Simulation of the Annual Cycle of Sea Ice in the Arctic and Antarctic*. NCAR Cooperative Thesis No. 46, National Center for Atmospheric Research, Boulder, Colorado, NTIS Accession No. PB279789, 191 pp.

Parkinson, C. L., and Washington, W. M.: 1979, A large-scale numerical model of sea ice. *J. Geophys. Res.* 84, 311–337.

Ramanathan, V., Lian, M. S., and Cess, R. D.: 1979, Increased atmospheric CO_2: Zonal and seasonal estimates of the effect on the radiation energy balance and surface temperature. *J. Geophys. Res.*, in press.

Schneider, S. H.: 1975, On the carbon dioxide-climate confusion. *J. Atmos. Sci.* 32, 2060–2066.

Sellers, W. D.: 1974, A reassessment of the effect of CO_2 variation on a simple global climatic model. *J. Appl. Meteorol.* 13, 831–833.

Semtner, A. J., Jr.: 1976, A model for the thermodynamic growth of sea ice in numerical investigations of climate. *J. Phys. Oceanogr.* 6, 379–389.

Siegenthaler, U., and Oeschger, H.: 1978, Predicting future atmospheric carbon dioxide levels. *Science* 199, 388–395.

SMIC: 1971, *Inadvertent Climate Modification: Report of the Study of Man's Impact on Climate*. MIT Press, Cambridge, Massachusetts.

van Loon, H., and Williams, J.: 1976, The connection between trends of mean temperature and circulation at the surface: Part I, Winter. *Mon. Weather Rev.* 104, 365–380.

van Loon, H., and Williams, J.: 1977, The connection between trends of mean temperature and circulation at the surface: Part IV, Comparison of the surface changes in the northern hemisphere with the upper air and with the Antarctic in winter. *Mon. Weather Rev.* 105, 638–647.

(Received 3 November, 1978; in revised form 18 April, 1979)

Paper 91

VOL. 84, NO. C1 JOURNAL OF GEOPHYSICAL RESEARCH JANUARY 20, 1979

A Large-Scale Numerical Model of Sea Ice

CLAIRE L. PARKINSON[1] AND WARREN M. WASHINGTON

National Center for Atmospheric Research, Boulder, Colorado 80307

Work at the National Center for Atmospheric Research has resulted in the construction of a large-scale sea ice model capable of coupling with atmospheric and oceanic models of comparable resolution. The sea ice model itself simulates the yearly cycle of ice in both the northern and the southern hemispheres. Horizontally, the resolution is approximately 200 km, while vertically the model includes four layers, ice, snow, ocean, and atmosphere. Both thermodynamic and dynamic processes are incorporated, the thermodynamics being based on energy balances at the various interfaces and the dynamics being based on the following five stresses: wind stress, water stress, Coriolis force, internal ice resistance, and the stress from the tilt of the sea surface. Although the ice within a given grid square is of uniform thickness, each square also has a variable percentage of its area assumed ice free. The model results produce a reasonable yearly cycle of sea ice thickness and extent in both the Arctic and the Antarctic. The arctic ice grows from a minimum in September, when the edge has retreated from most coastlines, to a maximum in March, when the ice has reached well into the Bering Sea, has blocked the north coast of Iceland, and has moved southward of the southernmost tip of Greenland. Maximum arctic thicknesses are close to 4 m. In the Antarctic the ice expands from a minimum in March to a maximum in late August, remaining close to the continent in the former month and extending northward of 60°S in the latter month. Maximum thicknesses are about 1.4 m. The distribution of modeled ice concentrations correctly reveals a more compact ice cover in the northern hemisphere than in the southern hemisphere. Modeled ice velocities obtain both the Beaufort Sea gyre and the Transpolar Drift Stream in the arctic summer as well as the Transpolar and East Greenland Drift streams in the winter. In the Antarctic, simulated velocities reveal predominantly westerly motion north of 58°S, with smaller-scale cyclonic motions closer to the continent.

1. INTRODUCTION

Large-scale numerical models of the atmosphere and of the oceans have both reached the stage where a coupling of the two is among the major steps now needed to improve the ability of the models to reproduce atmospheric and oceanic phenomena. As the models become coupled, it is imperative that the interface be given adequate consideration. Sea ice is one element of this interface which should not be neglected, for (1) it covers roughly 7% of the earth's oceans; (2) its extent shows large seasonal and interannual variations; (3) it has a strong insulation effect, restricting exchanges of heat and momentum between ocean and atmosphere; and (4) its high albedo relative to that of water significantly reduces the amount of shortwave radiative energy available near the ocean surface in the polar regions.

This article describes a large-scale numerical model of sea ice which should be capable of coupling with atmospheric and oceanic models. The horizontal resolution is roughly 200 km, and both thermodynamic and dynamic processes are incorporated. Emphasis is on reproducing the large-scale features of the annual cycle of sea ice coverage in the northern and southern hemispheres.

2. PREVIOUS ICE MODELING

Previous ice modeling can be divided into two categories: thermodynamics and dynamics. The thermodynamic calculations of the present study are patterned after those by *Untersteiner* [1964], *Maykut and Untersteiner* [1969, 1971], and *Semtner* [1976a]. Each of these earlier studies is one dimensional in space, and each specifies rather than computes atmospheric energy fluxes. None of the models is applied to the Antarctic, and none includes ice transport.

Maykut and Untersteiner [1969, 1971] construct an elaborate one-dimensional thermodynamic model of sea ice, computing the time-dependent ice thickness and vertical ice temperature profile based on diffusion equations with finite difference approximations involving a grid interval of 0.10 m. They include the effects of ice salinity, brine pockets trapped within the ice, heating from penetrating shortwave radiation, vertical variations in ice density, conductivity, and specific heat. Theirs is the most complete one-dimensional ice model devised so far; however, the model takes 38 simulation years to reach equilibrium, and this inhibits its extension to a three-dimensional framework.

Semtner [1976a] simplifies the Maykut and Untersteiner model, making it more appropriate for three-dimensional simulations. He does this largely through a reduction in the number of vertical layers, a change in the differencing scheme, the elimination of a heat source term in the diffusion equation, and the use of constants rather than variables for the specific heats and conductivities of ice and snow. The Semtner equations are formulated for n temperature levels within the ice and snow, though his article stresses the performance of the three-layer and zero-layer versions. Since he uses atmospheric and oceanic forcing identical to that of Maykut and Untersteiner, Semtner can compare the mean annual ice thicknesses from his models against the thicknesses predicted in the earlier, more complete work. Doing so for 25 separate cases, he finds an average deviation of only 0.22 m for the three-layer version and 0.24 m for the zero-layer version. It is basically the Semtner zero-layer model which has been followed in the present work for the calculations internal to the ice and snow.

Closer to the present model in spatial extent, *Pease*'s [1975] model simulates antarctic ice along longitude 155°E from 70° to 58°S. Although still not spatially three dimensional, this model can compute the advance and retreat of ice along the 155°E meridian, thus giving it a capability not available with the one-dimensional models by *Maykut and Untersteiner* [1969, 1971] or *Semtner* [1976a]. Furthermore, Pease computes rather than specifies atmospheric fluxes, and she tests the

[1] Now at Laboratory for Atmospheric Sciences, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771.

Copyright © 1979 by the American Geophysical Union.

TABLE 1. Forces Included for the Transport Calculations of Previous Models

	Wind Stress	Water Stress	Coriolis Force	Internal Ice Resistance	Dynamic Topography
Nansen [1902]	yes	yes	yes	no	no
Sverdrup [1928]	yes	no	yes	yes	no
Rossby and Montgomery [1935]	yes	yes	no	no	no
Rossby and Montgomery [1935]	yes	yes	yes	yes	no
Rossby and Montgomery [1935]	yes	no	yes	yes	no
Shuleikin [1938]	yes	yes	yes	no	no
Felzenbaum [1958]	yes	yes	yes	no	yes
Ruzin [1959]	yes	yes	yes	yes	no
Reed and Campbell [1960, 1962]	yes	yes	yes	no	no
Campbell [1964]	yes	yes	yes	yes	yes
Hunkins [1966, 1974]	yes	yes	yes	yes	yes
Rothrock [1973]	yes	yes	yes	yes	yes
Coon et al. [1976]	yes	yes	yes	yes	no
Pritchard et al. [1976]	yes	yes	yes	yes	yes
Hibler [1977]	yes	yes	yes	yes	yes

Extended from *Campbell* [1968].

effects of three parameterizations of the oceanic mixed layer. Another two-dimensional model including sea ice is that of *MacCracken and Luther* [1974]. This is a zonal atmospheric model that introduces sea ice of 0.2 m when the ocean temperature drops below freezing and calculates the one-dimensional horizontal extent of the ice, as well as its thickness, through energy balances. The ice, ocean, and land parameterizations are used to determine vertical processes at the lower boundary of the atmospheric model.

An attempt at a three-dimensional ice simulation is included in the work of *Bryan et al.* [1975] and *Manabe et al.* [1975]. They compute an ice distribution within a global ocean/atmosphere climate simulation, stressing long-term results and using mean annual insolation rather than seasonally varying forcing. Although they include transport of the ice, the ice moves strictly with the water of the upper ocean unless its thickness equals or exceeds 4 m, in which case the ice movement stops altogether. The resulting calculations lead to an unreasonable buildup of arctic ice with time, the average and maximum thicknesses steadily increasing, with values of 5.32 and 24.7 m, respectively, in year 200 of the simulation [*Manabe et al.*, 1975].

Additional discussions of ice transport include a wide range of formulations. Certainly the simplest are those relating the speed and direction of the ice strictly to the speed and direction of the water, as in the work of *Manabe et al.* [1975], or strictly to the speed and direction of the boundary layer wind, as in Zubov's rule [e.g., *Gordienko*, 1958].

Most other formulations of ice transport calculate the ice velocity by considering a balance of forces acting on the ice and by assuming a steady state solution. It appears widely agreed that the relevant forces are wind stress, water stress, Coriolis force, stress from the tilt of the sea surface (dynamic topography), and the force from interactions within and among floes (internal ice resistance). However, disagreement exists regarding the relative magnitude of these, with the result that different researchers have included different combinations in their final balance. A summary of the forces used by various authors appears in Table 1. Similarly, different researchers use different parameterizations for the forces included.

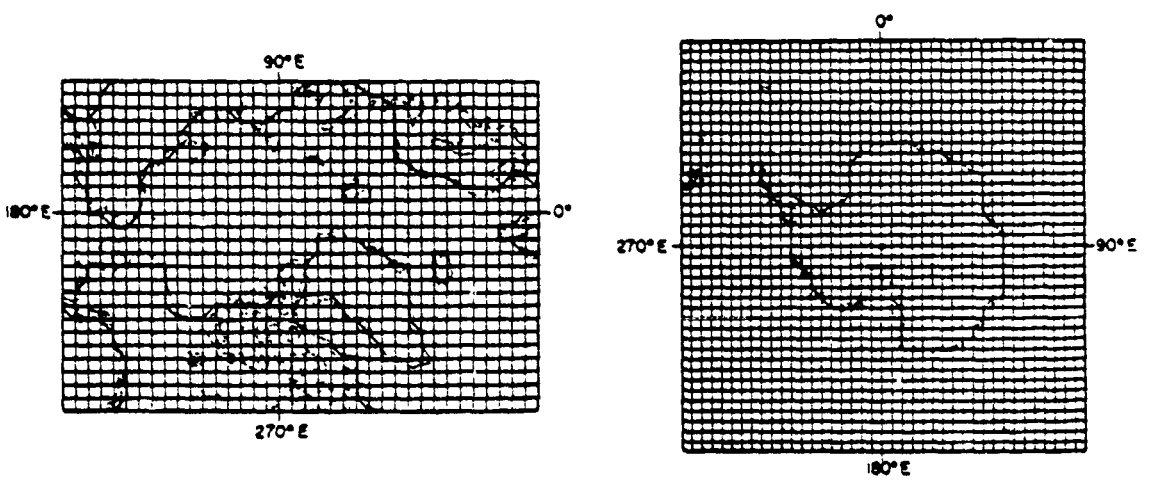


Fig. 1. Arctic and antarctic grids. Dotted lines trace the continental boundaries from standard polar stereographic projections. Model resolution of those boundaries is indicated in solid lines, as are the grids themselves. The south pole is at the center of the 41 × 41 antarctic grid, while the north pole is at position (18, 16) of the 38 × 26 arctic grid.

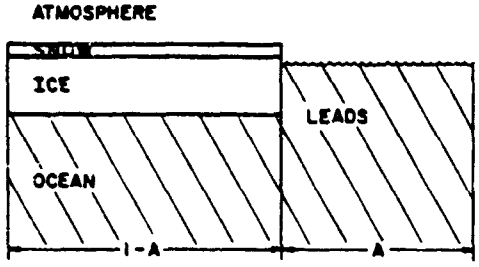


Fig. 2. Schematic diagram of the major divisions within a grid square.

One of the studies giving the most thorough discussion of the selection of parameterizations and subsequent combination of them for a mathematical solution to the ice velocity is by *Campbell* [1964]. Although his solution is not seasonally dependent, Campbell does obtain a steady state ice velocity field for the majority of the Arctic Ocean, and he does, upon increasing the eddy viscosity of ice to 3×10^6 m^2 s^{-1}, position a Pacific gyre approximately in its observed location.

However, in spite of its importance, the selection of forces and the calculation of resultant velocities are only one aspect of the transport problem. By not redistributing the ice according to the velocities obtained, the Campbell study and others in Table 1 avoid the additional difficulties posed by excessive convergence and divergence. Two studies facing these difficulties in a limited fashion are those of *Nikiforov et al.* [1970] and J. Walsh (personal communication, 1976). Both examine ice convergence as encountered by a coastal boundary, and both are concerned with ice concentrations only, not thicknesses.

In the studies by Nikiforov et al. and Walsh, thermodynamics is excluded, and ice velocities are determined strictly from maps of the observed atmospheric pressure fields. They deal with excessive ice convergence near the coast by zeroing out onshore velocity components wherever the intervening region to the coast is covered by ice of 100% compactness. Beyond the

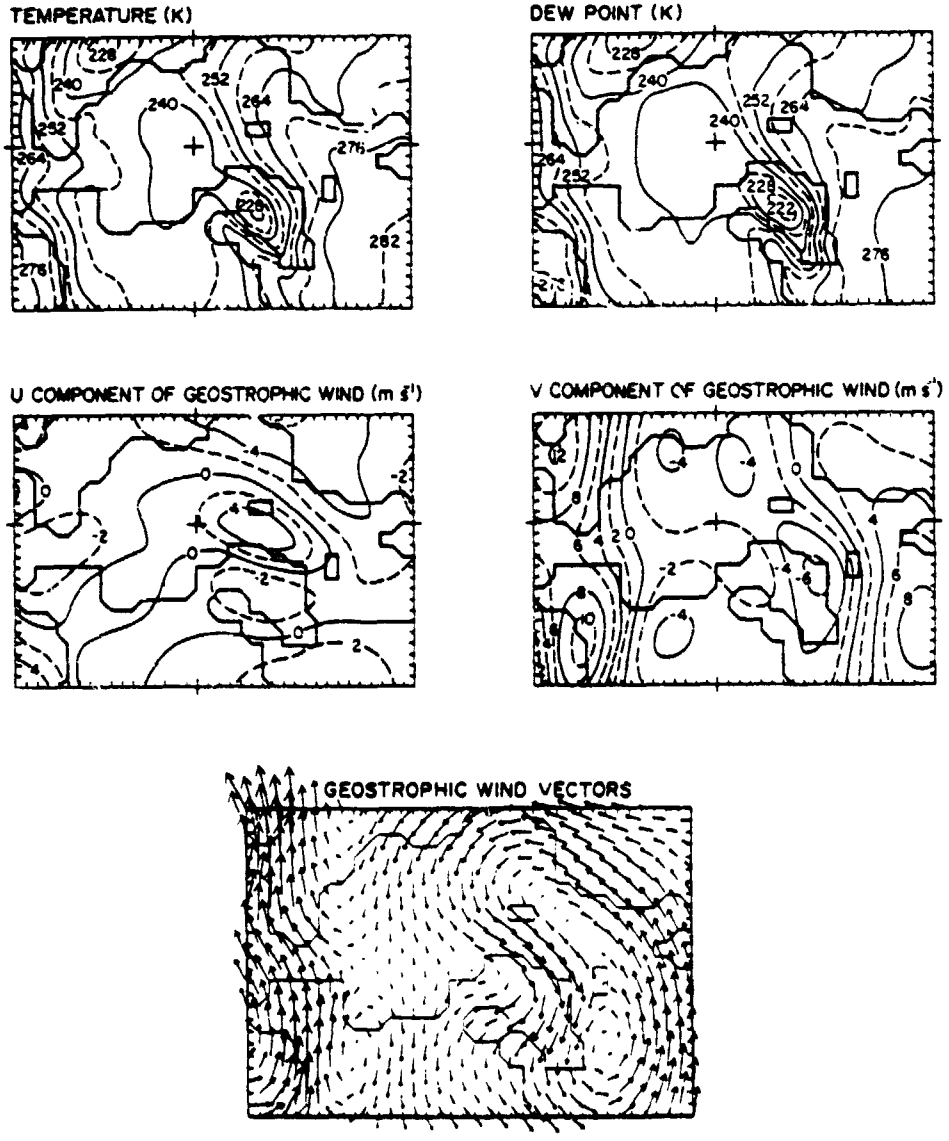


Fig. 3. January atmospheric input data for the Arctic.

coastal region, compactness n is altered through convergence/divergence and advection. The Nikiforov et al. simulation is for the region of the East Siberian, Laptev, and Chukchi seas, while the Walsh simulation is for the portion of the Arctic bounded by 60°W, 120°E, 85.5°N, and the North American and Siberian coasts.

A more detailed ice model has been developed in connection with the Arctic Ice Dynamics Joint Experiment (Aidjex) and contains three major parts: a thickness distribution model, a momentum equation for ice motions, and a stress/strain law. The thickness distribution model is based on a thickness distribution function $G(H, t)$, which is the fractional area with ice thinner than H at time t. The momentum equation includes all five major forces mentioned earlier plus an acceleration term. (The four references preceding *Hibler* [1977] in Table 1 are part of Aidjex.) The stress/strain relationships are modeled according to an elastic-plastic formulation [e.g., *Coon et al.*, 1974; *Pritchard*, 1975, 1976; *Hall et al.*, 1976].

The complete Aidjex model is considerably more detailed than the model of the present study; however, the time and space scales for which it has been tested are much smaller, and its complexity makes it impractical for use with global atmospheric or oceanic models. The application of the Aidjex work centers on roughly one twentieth of the Arctic Basin, located in the Beaufort Sea, and the time scale of interest is of the order of 1 day. By contrast, the present study creates a more simplified numerical simulation of the entire annual cycle for the majority of the world's ice-covered regions.

3. Description of the Model

a. Grid

Polar stereographic projections are used in each hemisphere, with rectangular grids superimposed. The 41 × 41 grid in the Antarctic is centered on the south pole, while the 38 × 26 grid in the Arctic positions the north pole at location (18, 16) with respect to the lower left corner. Both grids are presented in Figure 1, along with the sea-land boundaries. Since the model is specifically for floating sea ice, the antarctic ice shelves are not treated. Horizontal grid resolution ranges from 211.06 km at the poles to 181.006 km at the farthest edge of the north polar region and to 173.844 km at the farthest edge of the south polar region. Details on the grid structure can be obtained from *Parkinson* [1978, section 4.1].

Vertically, the model includes a mixed layer in the ocean, an ice layer, a snow layer, and an atmospheric boundary layer. Figure 2 schematizes the four layers plus the allowance for leads. Depending on time and location, the snow layer or both the ice and the snow layers may be nonexistent, thereby reducing the number of vertical layers from four to three or two, respectively. The remaining dimension is time, where the resolution is defined by an 8-hour time step and the length of all months is set at a uniform 90 time steps, i.e., 30 days.

b. Fields of Input Data

The thermodynamic calculations (section 3*e*) require input data fields in the form of atmospheric temperatures for sensible heat flux and incoming longwave radiation, dew points for latent heat flux, wind speeds for sensible and latent heat fluxes, and both wind velocities and dynamic topography for ice transport.

Atmospheric data. The atmospheric data were obtained from long-term mean monthly distributions as given on 5° latitude-longitude grids by *Taljaard et al.* [1969] in the southern hemisphere and by *Crutcher and Meserve* [1970] in the northern hemisphere. Sixteen-point interpolation was used to convert the atlas data to the rectangular grid of the sea ice model. In this manner, fields were obtained for surface air temperatures T_a, surface dew points T_d, and the u_{ag} and v_{ag} components of sea level geostrophic winds. The stored atmospheric data for all 12 months are accepted as accurate at time step 45 (midmonth), with values for other time steps being determined by linear interpolation.

Figures 3–6 present contour maps and plotted wind vectors from the stored data for January and July. For increased legibility the wind vectors are drawn at only half the grid points, and the vector lengths are scaled individually for each diagram. A unit length of the distance between grid points signifies 3 m s^{-1} in Figure 3, 2 m s^{-1} in Figure 4, and 5 m s^{-1} in Figures 5 and 6. This allows reasonably lengthened vectors in each diagram in spite of the weaker winds in the northern hemisphere summer.

Dynamic topography. Input fields of dynamic topography were digitized from contour maps by A. Gordon (personal communication, 1976) in the Antarctic and from a map by *Coachman and Aagaard* [1974] in the Arctic. In both hemispheres, only yearly averages were available. Furthermore, the arctic data were limited spatially to the central portions of the ocean, and thus the contours were extended subjectively. This was done in a manner maintaining the approximate patterns of a 20-m pressure field predicted by a model of the Arctic Ocean Circulation [*Semtner*, 1976*b*, Figure 7]. The Semtner diagram extends contours to the continental boundaries and includes the Norwegian-Barents Sea. The patterns correspond closely to the dynamic topography field of Coachman and Aagaard in the regions for which this latter field is presented. However, neither the North Atlantic nor the North Pacific is included, and so contours for these areas are sketched roughly from the patterns of ocean circulation in the work of *Sverdrup et al.* [1961, Chart VII]. The resulting dynamic topography fields appear in Figures 7 and 8.

c. Forcing From Above the Ice or Water

As will be seen in section 3*e*, the computations for thickness changes of ice and snow require fluxes of solar radiation, longwave radiation, sensible heat, and latent heat. The evaluation of each is discussed in turn.

Solar radiation. The flux of solar radiation $SW\downarrow$ is calculated by applying the cloudiness factor by *Laevastu* [1960] to an empirical equation by *Zillman* [1972] for global radiation under cloudless skies. The Zillman equation is

$$Q_s = \frac{S \cos^2 Z}{(\cos Z + 2.7)e \times 10^{-3} + 1.085 \cos Z + 0.10} \qquad (1)$$

S signifies the solar constant, Z the solar zenith angle, and e the vapor pressure in pascals (1 Pa = 10^{-2} mbar). A value of 1353 W m^{-2} has been taken for the solar constant [*Thekaekara and Drummond*, 1971], while the cosine of the zenith angle is calculated by the standard geometric formula

$$\cos Z = \sin \phi \sin \delta + \cos \phi \cos \delta \cos HA \qquad (2)$$

where ϕ, δ, and HA are latitude, declination, and hour angle, respectively [*Sellers*, 1965]. The approximate declination and the hour angle are determined as

$$\delta = 23.44° \times \cos [(172 - \text{day of year}) \times \pi/180]$$
$$HA = (12 \text{ hours} - \text{solar time}) \times \pi/12 \qquad (3)$$

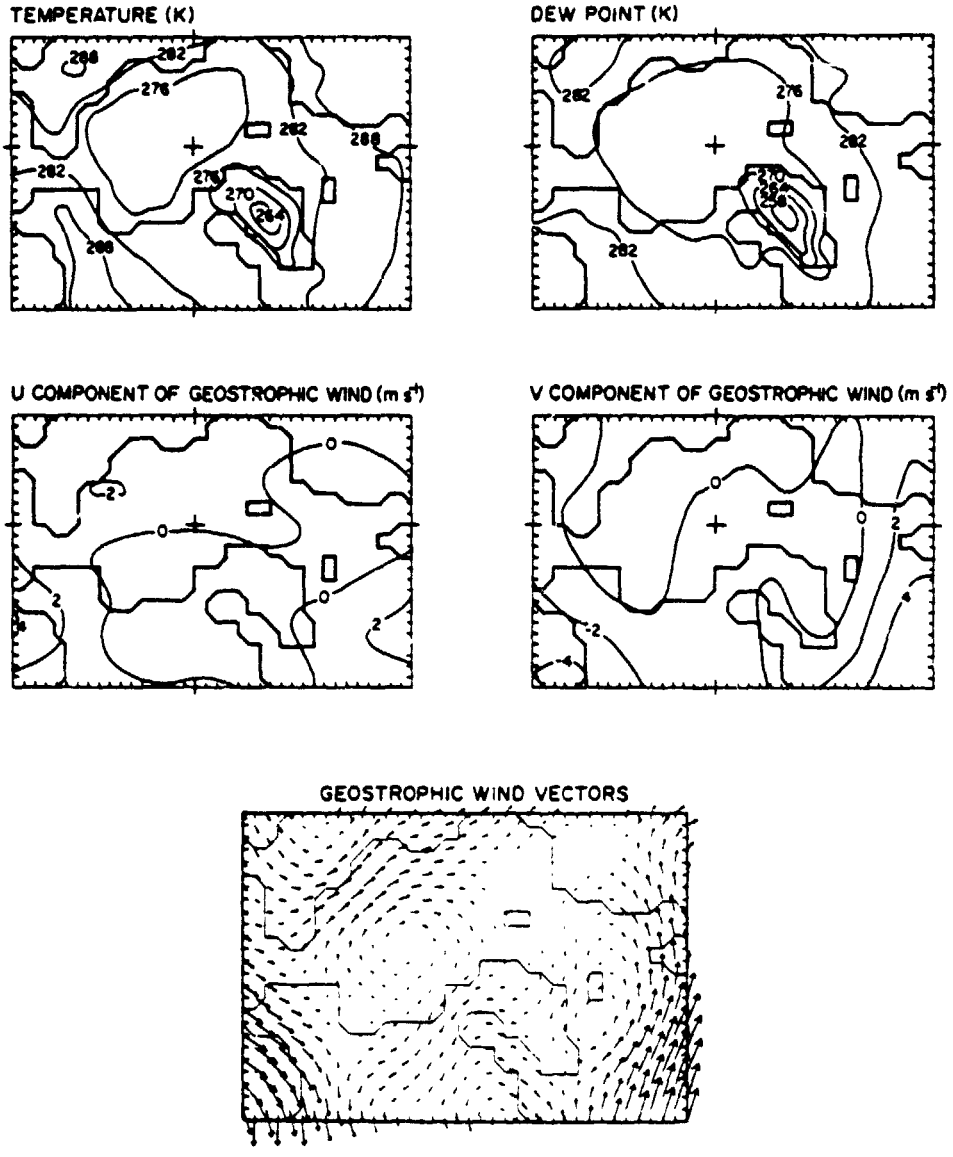


Fig. 4. July atmospheric input data for the Arctic.

while the vapor pressure e is calculated by a formulation presented in section 3c, 'latent heat.'

As was mentioned, Q_s is modified by a cloud factor to obtain total incoming shortwave radiation

$$Q = Q_s(1 - 0.6c^3) \tag{4}$$

[*Laevastu*, 1960]. Cloud cover figures for the Arctic have been averaged from the curves of *Huschke* [1969] for his four arctic regions, these curves being frequently reproduced in the subsequent literature [e.g., *Kellogg*, 1974; *Herman*, 1975; *Baker*, 1976]. As a result, in the Arctic the fractional cloud cover c is set at 0.50 for December through March, 0.55 for April, 0.70 for May, 0.75 for June and July, 0.80 for August and September, 0.70 for October, and 0.60 for November.

In the Antarctic, cloud cover varies more with latitude than with season, and so the modeled cloud values are a function of both latitude and month. The basis of the calculations is the January and July curves of *van Loon* [1972] for cloud cover versus latitude. We have approximated these two curves by fourth-order Lagrangian interpolating polynomials satisfying the functional values of van Loon at 40°, 50°, 60°, 70°, and 80°S (Figure 9) [*Parkinson*, 1978]. Cloud cover amounts for intervening months are obtained through linear interpolation.

At each grid point, $SW\downarrow$ is determined by performing the above calculations at the centers of two 1-hour and eleven 2-hour periods, multiplying by the relevant time length (1 or 2 hours), summing for the day, and then dividing by 24 hours. Because of the computer time involved, solar radiation is recalculated only every 11 time steps (3.75 days). This procedure eliminates the diurnal cycle of solar radiation and hence the diurnal thaw-refreezing of the snow and sea ice.

Incoming longwave radiation. Incoming longwave radiation is calculated each time step from *Idso and Jackson*'s [1969] formula for clear skies:

$$F\downarrow = \sigma T_a^4[1 - 0.261 \exp\{-7.77 \times 10^{-4}(273 - T_a)^2\}] \tag{5}$$

modified by a cloudiness factor of $1 + cn$. In the current calculation, the Stefan-Boltzmann constant σ is 5.67×10^{-8} W m^{-2} K^{-4}, while the surface air temperatures T_a are discussed under fields of input data (section 3*b*), cloud cover values c are presented under 'solar radiation' (section 3*c*), and n is an empirical factor set at 0.275 as averaged from the drifting station values of *Marshunova* [1966].

Sensible heat. The flux of sensible heat is calculated each time step from the standard bulk aerodynamic formula:

$$H\downarrow = \rho_a c_p C_H V_{wg}(T_a - T_{s/c}) \qquad (6)$$

Surface air temperature T_a and surface geostrophic wind speed $V_{wg} = (u_{wg}^2 + v_{wg}^2)^{1/2}$ are from section 3*b* ('atmospheric data'), while surface ice-snow-water temperature $T_{s/c}$ is calculated from a surface energy balance equation (thermodynamic calculations, section 3*e*). Following *Maykut* [1977], the transfer coefficient for sensible heat C_H and the latent heat transfer coefficient C_E are approximated at 1.75×10^{-3}. The specific

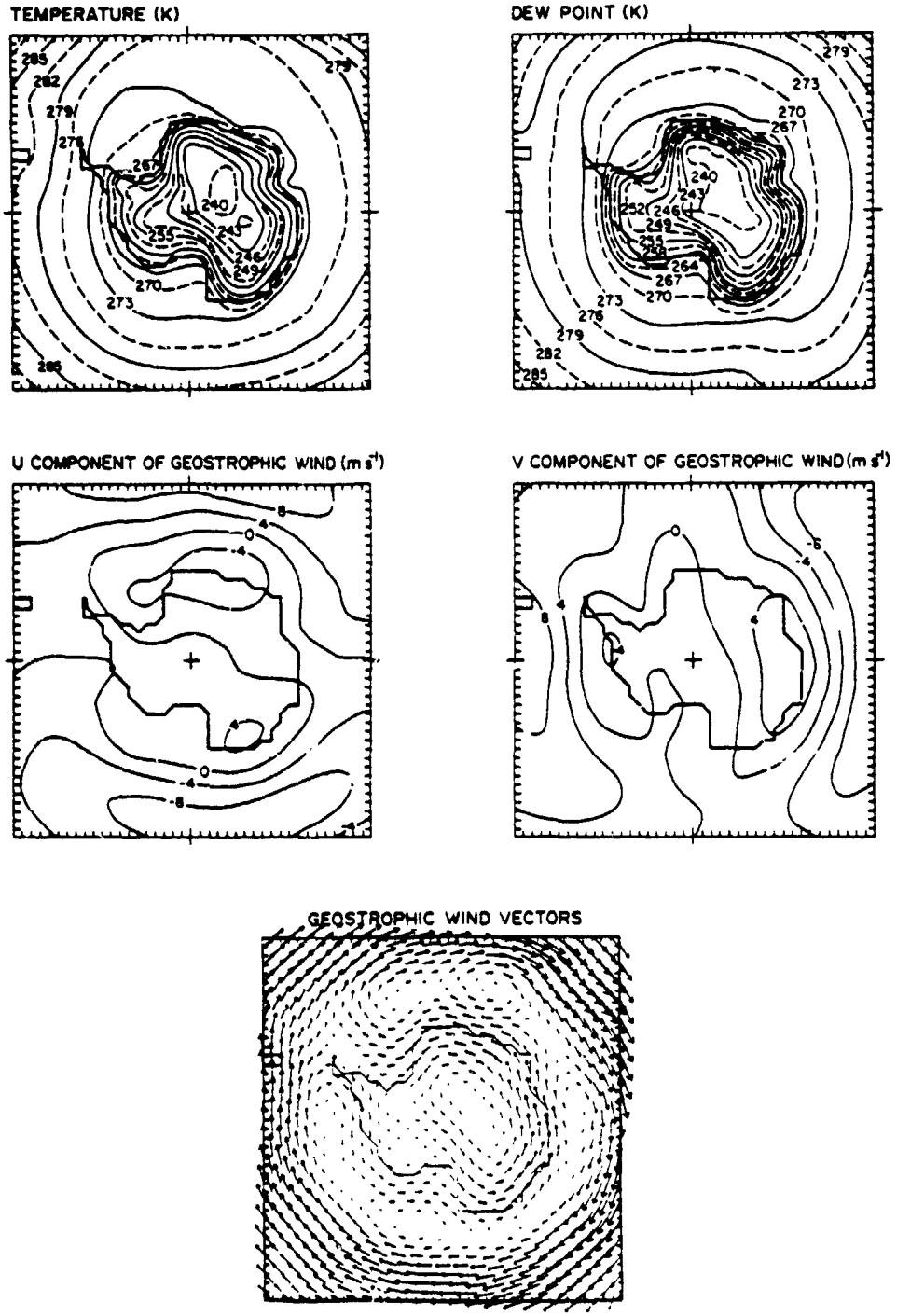


Fig. 5. January atmospheric input data for the Antarctic.

heat of air c_p is set at the dry air value of 1004 J kg^{-1} K^{-1} [*Huschke*, 1959]. Air density ρ_a is obtained from the equation of state by using a constant pressure p of 98,800 Pa in the southern hemisphere and 101,400 Pa in the northern hemisphere.

Latent heat. As in the calculations for sensible heat, calculations for the flux of latent heat proceed each time step with a bulk aerodynamic formula. Symbolically,

$$LE\downarrow = \rho_a L C_E V_{wa}(q_{10m} - q_s) \tag{7}$$

L being the latent heat of vaporization (2.5 × 10^6 J kg^{-1}) or of sublimation (2.834 × 10^6 J kg^{-1}), depending on whether an ice cover exists [*Haltiner and Martin*, 1957], and the q's being specific humidities at 10 m and the surface. The formulae for specific humidities are

$$q_{10m} = \frac{\epsilon e}{p - (1 - \epsilon)e} \tag{8}$$

and

$$q_s = \frac{\epsilon e_s}{p - (1 - \epsilon)e_s} \tag{9}$$

where $\epsilon = 0.622$ is the ratio of the molecular weight of water vapor to that of dry air [*Hess*, 1959; *Oliger et al.*, 1970]. Vapor pressure e and saturation vapor pressure e_s are determined from an empirical formula by *Murray* [1967]:

$$e_s = 611 \times 10^{a(T_{sfc} - 273.16)/(T_{sfc} - b)} \tag{10}$$

where $(a, b) = (9.5, 7.66)$ if an ice cover exists and (7.5, 35.86) if no ice cover exists. For e, (10) is used with T_{sfc} replaced by the surface dew point temperature T_d. All temperatures are in kelvins and pressures in pascals.

d. *Forcing From Below the Ice*

The upper layer of the ocean tends to be well mixed in temperature and salinity. The present model, ignoring the details of differing density stratifications, assumes the depth of this mixed layer to be 30 m and calculates changes in water temperature based on this depth (section 3*f*). In the standard case the temperature and other mixed layer properties are not used to determine the upward heat flux, but instead the energy flux $F\uparrow$ to the ice from the water beneath is taken to be a constant 2 W m^{-2} in the Arctic and a constant 25 W m^{-2} in the Antarctic. Recognizing that this involves a considerable simplification, we attempt to explain the basic reasons for it in the following paragraphs.

In studies by *Pease* [1975], *Bryan et al.* [1975], and *Welander* [1977] the flux from the mixed layer is directly proportional to the temperature difference between the water and the ice. This follows from a scale analysis of the first law of thermodynamics and was attempted in the present model also. However, the numerical value of the proportionality factor varies considerably among the three former studies, and as *Pease* [1975] indicates, it depends on an arbitrary eddy diffusivity of water.

N. Untersteiner (personal communication, 1976) and G. Maykut (personal communication, 1977) object to calculating the flux as being proportional to the temperature difference between water and ice, contending that the temperature directly under the ice remains practically at freezing until the ice disappears. Indeed, this is found to be true in the calculated temperatures of the present model, resulting in a 0 calculated oceanic heat flux regardless of the proportionality factor.

Furthermore, although only temperatures have been mentioned so far, an accurate calculation of the energy flux from ocean to ice would require incorporation of ocean salinity, a variable mixed layer depth, and interaction with the underlying ocean layers. Salt content of the upper ocean is significantly affected by the freezing of water and the melting of ice and in turn affects the vertical density structure of the ocean. As seawater freezes, the discharge of salt to the ocean beneath increases the density of the mixed layer and hence increases the chance of an unstable stratification and resulting convection with underlying water. The convection alters the depth, the temperature, and the salinity of the mixed layer, the amount of alteration depending on the initial temperature and salinity profiles versus depth. Unfortunately, these profiles vary widely with location and season, as is well quantified in antarctic waters by data from several *Eltanin* cruises [*Jacobs et al.*, 1974] and well recognized in arctic waters, in particular due to the inflow of the Atlantic layer through the Greenland-European sector.

Thus a parameterization including the salinity influence would require three-dimensional fields of temperature and salinity. Although computer space prevented insertion of such fields for the present model, presumably upon coupling with an ocean model these fields will be available. This would allow a more realistic parameterization of ice-ocean interactions, though the problem of uncertain eddy diffusivities would remain. In the meantime, it is felt that a constant ocean heat flux is the best choice for the current standard case. The task then becomes to determine a proper constant.

In their one-dimensional model, *Maykut and Untersteiner* [1971] use 2 W m^{-2} for the oceanic heat flux in the central Arctic, having tested values of 0, 1, 2, 4, 6, and 8 W m^{-2}. The choice of 2 W m^{-2} for their standard case was predominantly based on its yielding the most satisfactory results. With $F\uparrow$ = 8 W m^{-2} the ice vanished before equilibrium was reached.

In the Antarctic, as in the Arctic, accurate ocean flux values are simply not known. However, typical fluxes from water to ice are believed to be much greater in the Antarctic than in the Arctic, and A. Gordon (personal communication, 1976) suggested values of up to 25 W m^{-2}. The selection of 25 W m^{-2} for the current model followed the testing of three values [*Parkinson*, 1978] and the judgment that the results looked reasonable with $F\uparrow$ = 25 W m^{-2}. By contrast, the selection of 2 W m^{-2} for the arctic simulation strictly followed the usage of *Maykut and Untersteiner* [1971] and was not preceded by testing alternatives.

e. *Thermodynamic Calculations*

Calculations of changes in thickness of the ice and snow layers are based on energy balances at the various interfaces. In grid squares with no ice, energy balances determine instead the change in ocean temperature.

No ice. The no-ice situation is depicted in Figure 10*a*, where the net energy flux into the mixed oceanic layer is

$$Q_{net} = H\downarrow + LE\downarrow + \epsilon_w LW\downarrow + (1 - \alpha_w)SW\downarrow + F_w\uparrow - \epsilon_w \sigma T_w^4 \tag{11}$$

The evaluations of $H\downarrow$, $LE\downarrow$, $LW\downarrow$ and $SW\downarrow$ were presented earlier, and in view of the large uncertainty and spatial variance in both the sign and the magnitude of the flux $F_w\uparrow$ to the mixed layer from the deeper ocean, this flux is taken as 0. All terms in the expression for Q_{net} are for the current time step, except the water temperature T_w, which is for the previous time step. The shortwave albedo α_w of water is set at 10%, and the longwave emissivity ϵ_w is set at 97%. The albedo figure is

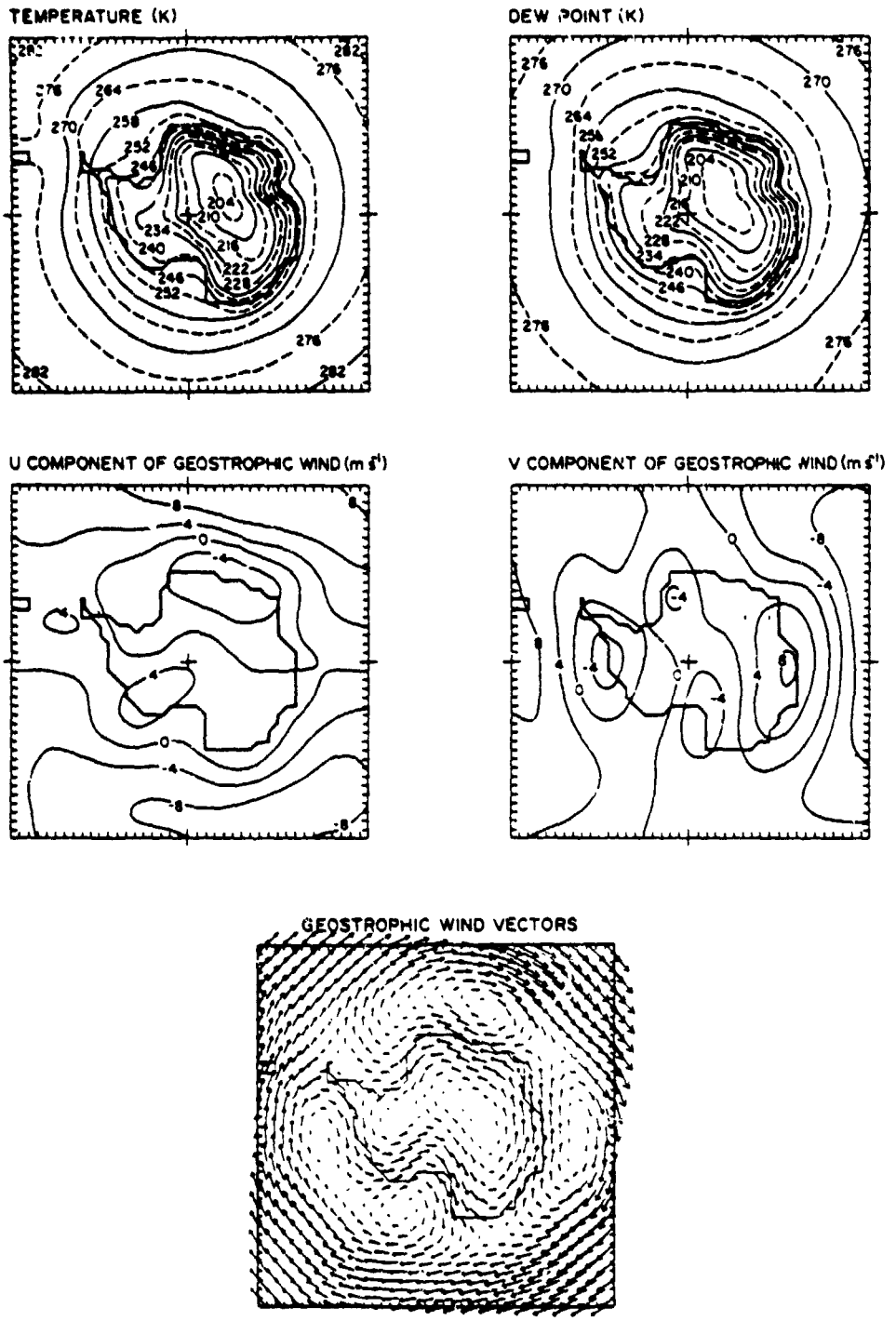


Fig 6. July atmospheric input data for the Antarctic.

the same as that used by *Badgley* [1966] for arctic leads, by *Donn and Shaw* [1966] for an open polar ocean, and by *Langleben* [1972] for along the arctic coastline.

The entire net inflow Q_{net} goes toward raising the temperature of the water, so that the time rate of change of internal energy (per unit horizontal area) becomes

$$dI/dt = Q_{net} \tag{12}$$

Internal energy I is the product of the 30-m mixed layer depth d_{mix}, the volumetric heat capacity of water c_w = 4.19 MJ m^{-3} K^{-1}, and the water temperature. Applying a finite difference approximation to (12), we obtain a resulting temperature change of

$$\Delta T_w = \frac{\Delta t \times Q_{net}}{d_{mix} \times c_w} \tag{13}$$

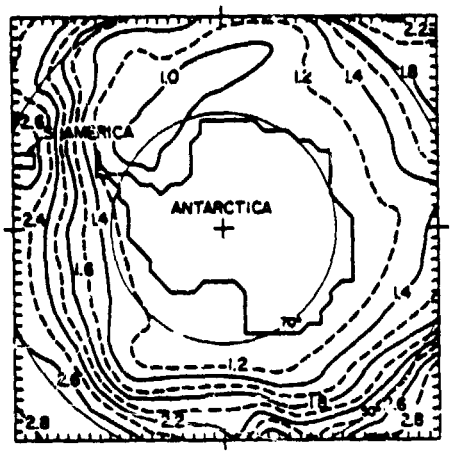


Fig. 7. Contours of dynamic topography used in the Southern Ocean. Contour interval is 0.2 dynamic m.

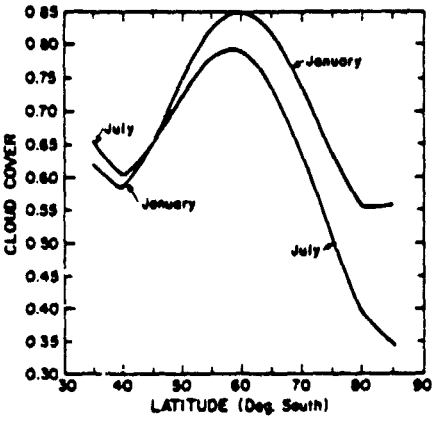


Fig. 9. Cloud cover percentages used for the southern hemisphere in January and July, presented as a function of latitude.

where Δt is the 3-hour time step. As heat is spread uniformly throughout the mixed layer, the new water temperature equals

$$T_{w,i} = T_{w,i-1} + \Delta T_w \tag{14}$$

where subscripts i and $i-1$ refer to the time step. Note that if Q_{net} is negative in (13), then the temperature is reduced, as desired. Should T_w fall below 271.2 K, then a portion of the water is assumed to freeze. The thickness of the newly frozen ice is set at $h_n = 0.01$ m, and the extent is calculated so as to release the amount of heat necessary to maintain the water temperature at freezing. The fraction of the grid square covered by ice then becomes

$$F = \frac{(T_B - T_w) \times d_{mix} \times c_w}{Q_I \times h_n} \tag{15}$$

where Q_I is the heat of fusion of ice, set at 302 MJ m^{-3}. The new ice is modeled to have a surface temperature of 271.2 K. An experiment has been run in which the thickness of newly formed ice is 0.20 m instead of 0.01 m and in which the ice thickness cannot be between 0 and 0.20 m. The results show the model to be fairly insensitive to the choice of 0.01 versus 0.20 m.

Ice with no snow cover. The fluxes involved in the case of ice with no snow cover are diagramed in Figure 10*b*. The only flux through the ice specifically modeled is the conductive one, though in reality a fraction I_0 of the net incident shortwave radiation penetrates the uncovered upper surface. *Maykut and Untersteiner* [1971] assume that $I_0 = 17\%$. This penetrating radiation normally causes the brine volume to increase and the cooling near the upper surface to be delayed. The present model follows the zero-layer model of *Semtner* [1976*a*] in parameterizing this by using 60% of I_0 as heating in the surface energy budget and ignoring the remaining 40%. In essence, the ignored 40% is treated as reflected radiation, i.e., as an effective increase in the shortwave albedo.

The conductive flux G_I through the ice equals $k_I(T_B - T_{sfc})/h_I$, as in the work of *Bryan et al.* [1975], *Pease* [1975], and *Semtner* [1976*a*]. The thickness of the ice h_I is taken from the previous time step, while the thermal conductivity of the ice k_I is set at a constant 2.04 W m^{-1} K^{-1} and the temperature at the bottom of the ice T_B is assumed to be the freezing point of seawater. Following *Maykut and Untersteiner* [1971] and *Bryan et al.* [1975], this freezing point is 271.2 K. The remaining value in the expression for conductive flux, surface temperature T_{sfc}, will now be calculated from the surface energy balance, the basic assumption being that the temperature at the surface adjusts itself in a manner maintaining that balance.

From Figure 10*b* the surface energy balance is

$$H\downarrow + LE\downarrow + \epsilon_I LW\downarrow + (1 - 0.4I_0)(1 - \alpha_I)SW\downarrow - \epsilon_I \sigma T_{sfc}^4 + (k_I/h_I)(T_B - T_{sfc}) = 0 \tag{16}$$

The longwave emissivity of the ice ϵ_I is set at 0.97, and the shortwave ice albedo α_I is set at 0.50. As in the work of *Semtner* [1976*a*], (16) is linearized by setting $T_{sfc} = T_p + \Delta T$ (where T_p is the temperature at the surface during the previous time step) and by approximating the power $T_{sfc}^4 = (T_p + \Delta T)^4$ by the two largest terms, i.e.,

$$T_{sfc}^4 \approx T_p^4 + 4T_p^3 \Delta T \tag{17}$$

Thus, (16) can be solved for

$$\Delta T = \left[H\downarrow + LE\downarrow + \epsilon_I LW\downarrow + (1 - 0.4I_0)(1 - \alpha_I)SW\downarrow - \epsilon_I \sigma T_p^4 + \frac{k_I}{h_I}(T_B - T_p)\right]\left[4\epsilon_I \sigma T_p^3 + \frac{k_I}{h_I}\right]^{-1} \tag{18}$$

If the resulting $T_{sfc} = T_p + \Delta T$ exceeds the 273.05-K melting point of ice, then T_{sfc} is set at 273.05 K exactly, and the excess energy is used to melt a portion of the ice rather than to raise its temperature beyond 273.05 K. The amount melted is determined by recalculating, with the new surface temperature of 273.05 K, the summed fluxes to the upper ice surface. As this

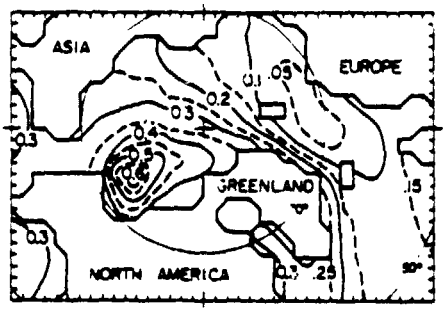


Fig. 8. Contours of dynamic topography used in the Arctic. Contour interval is 0.05 dynamic m.

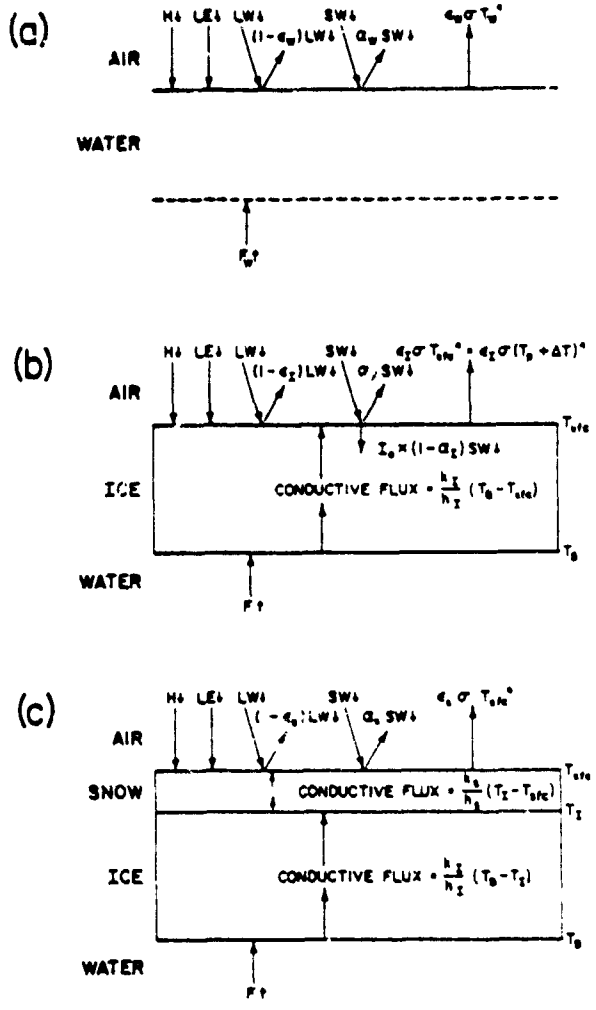


Fig. 10. Schematic diagrams of the energy fluxes involved. (*a*) Case of no ice. (*b*) Case of ice with no snow (*c*) Case of snow-covered ice.

sum represents the energy available for melting, the calculated melt equals

$$-\Delta h_I = \frac{\Delta t}{Q_I}\left[H\downarrow + LE\downarrow + \epsilon_I LW\downarrow + (1 - 0.4I_0)(1 - \alpha_I)SW\downarrow - \epsilon_I\sigma T_{s/c}^4 + \frac{k_I}{h_I}(T_B - T_{s/c})\right] \quad (19)$$

On the other hand, if the calculated $T_{s/c} = T_p + \Delta T$ yields a $T_{s/c} \leq 273.05$ K, then no ablation occurs, and instead snow is allowed to fall at a specified rate. The prescribed rate of snow accumulation in the antarctic region is 3 mm/month from the start of March through the end of November and 0 mm/month for the remaining three months. In the Arctic the rates follow those of *Maykut and Untersteiner* [1969] and are presented, along with the antarctic values, in Table 2.

With ablation and accretion at the top of the ice now determined, consideration turns to the bottom of the ice. Using the same reasoning as that for the upper surface, we find that the melt at the bottom equals

$$-\Delta h_I = \frac{\Delta t}{Q_I}\left[F_W^{\uparrow} - \frac{k_I}{h_I}(T_B - T_{s/c})\right] \quad (20)$$

This equation allows the possibility of a negative 'melt,' translating simply to ice accretion rather than ablation.

Snow-covered ice. The fluxes relevant to snow-covered ice are diagramed in Figure 10c. Because of the high extinction coefficient of snow, no penetration of shortwave radiation is allowed. Thus the I_0 parameterization needed in section 3e ('ice with no snow cover') is avoided here, and the only fluxes through the ice and snow layers are the conductive ones. With T_B, T_I, and $T_{s/c}$ equaling the temperatures at the bottom of the ice, at the snow-ice interface, and at the surface of the snow, the conductive fluxes become

$$G_I = k_I(T_B - T_I)/h_I \quad (21)$$

and

$$G_s = k_s(T_I - T_{s/c})/h_s \quad (22)$$

Snow conductivity k_s equals 0.31 W m^{-1} K^{-1}, while the shortwave albedo α_s of snow is set at 0.75 and the longwave emissivity ϵ_s is set at 0.99. Values for h_I and h_s, ice and snow thicknesses, come from the previous time step.

With these preliminaries the energy balance at the snow surface becomes

$$H\downarrow + LE\downarrow + \epsilon_s LW\downarrow + (1 - \alpha_s)SW\downarrow - \epsilon_s\sigma T_{s/c}^4 + \frac{k_s}{h_s}(T_I - T_{s/c}) = 0 \quad (23)$$

and that at the snow-ice interface becomes

$$\frac{k_s}{h_s}(T_I - T_{s/c}) = \frac{k_I}{h_I}(T_B - T_I) \quad (24)$$

To solve these equations, $T_{s/c}$ is replaced by $T_p + \Delta T$, and $(T_p + \Delta T)^4$ is again approximated by $T_p^4 + 4T_p^3\Delta T$. This leads to two equations linear in the two unknowns T_I and ΔT. Solving, we have

$$\Delta T = \left[H\downarrow + LE\downarrow + \epsilon_s LW\downarrow + (1 - \alpha_s)SW\downarrow - \epsilon_s\sigma T_p^4 + \frac{k_s k_I(T_B - T_p)}{k_s h_I + h_s k_I}\right]\left[4\epsilon_s\sigma T_p^3 + \frac{k_I k_s}{k_s h_I + k_I h_s}\right]^{-1} \quad (25)$$

and

$$T_I = \frac{k_s h_I(T_p + \Delta T) + h_s k_I T_B}{k_s h_I + h_s k_I} \quad (26)$$

TABLE 2. Monthly Values of Prescribed Snowfall Rates in Millimeters per Month

	Arctic Snowfall	Antarctic Snowfall
January	8.33	0
February	8.33	0
March	8.33	3
April	8.33	3
May	50	3
June	0	3
July	0	3
August 1–19	0	3
August 20–30	128.57	
September	128.57	3
October	128.57	3
November	8.33	3
December	8.33	0
Total	400 mm/yr	27 mm/yr

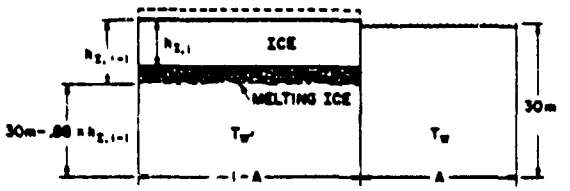


Fig. 11. Variables involved in calculating temperature changes due to vertical melt.

Should the calculated ΔT result in a new snow surface temperature $T_{s/c} = T_p + \Delta T$ greater than the melting point for snow, 273.15 K, then the excess energy goes toward melting the snow rather than increasing the temperature, so that $T_{s/c}$ will be set back at 273.15 K and T_l will be calculated from that. The depth of melted snow equals

$$-\Delta h_s = \frac{\Delta t}{Q_s}\left[H\downarrow + LE\downarrow + \epsilon_s LW\downarrow + (1-\alpha_s)SW\downarrow - \epsilon_s\sigma T_{s/c}^4 + \frac{k_s}{h_s}(T_l - T_{s/c})\right] \quad (27)$$

for which the heat of fusion of snow Q_s has been set at 110 MJ m^{-3}.

In the event of a calculated snow surface temperature $T_{s/c}$ below 273.15 K, no snow melts, but additional snow is accumulated at the rate specified for the given month (Table 2). No allowance is made for conversion of snow into ice, for instance through the melting/refreezing process. At the bottom of the ice, thickness changes are calculated analogously to those in the no-snow case, i.e.,

$$\Delta h_l = \frac{\Delta t}{Q_l} \times \left[\frac{k_l}{h_l}(T_B - T_l) - F\uparrow\right] \quad (28)$$

this being accretion if positive and ablation if negative.

f. *Lead Parameterization and Calculation of Water Temperatures*

On-site measurements have indicated that heat and moisture fluxes at the freezing point can be 2 orders of magnitude greater over seawater than over sea ice [*Ackley and Keliher*, 1976; *Zwally et al.* 1976; *Muench and Ahlnas*, 1976]. In fact, where ice-free water covers over 1% of the total surface area, the large-scale turbulent heat exchanges can be dominated by the leads [*Maykut*, 1976]. Consequently, the present model includes the parameter A registering at each grid point the percentage of the horizontal grid area covered by open water or leads. Changes in A occur as ice either melts laterally or forms at the lead surface, the magnitude of the change depending on the net energy flux into the lead itself and on the temperature of the water.

Each grid square has two water temperatures in the mixed layer: T_w, the temperature of the water in the leads, and $T_{w'}$, the temperature of the water under the ice. During any given time step each of these temperatures is altered in a maximum of three steps. For T_w the first step results from the vertical melting of the ice (section 3*f*, 'temperature adjustments for vertical melting'), the second from a lateral freezing of water and hence an increase in the area of ice coverage (section 3*f*, 'temperature and lead area changes for negative Q_o'), and the third from a partial mixing with the water in the leads (section 3*f*, 'lateral mixing of water'). Alterations in T_w arise from the net vertical energy input into the lead (section 3*f*, 'temperature and lead area changes for positive Q_o' and 'temperature and lead area changes for negative Q_o'), from lateral melting and the consequent increase of the lead area (section 3*f*, 'temperature and lead area changes for positive Q_o'), and from mixing with the water under the ice (section 3*f*, 'lateral mixing of water'). The two sets of three-step calculations are intertwined with themselves and with the change in lead area and hence are presented here in the order of calculation.

Temperature adjustments for vertical melting. If the ice thickness calculations (section 3*e*, 'ice with no snow cover' and 'snow-covered ice') result in a melting of the ice, the meltwater is assumed to mix into the water beneath and thereby to readjust the mixed layer temperature (Figure 11). The temperature of the melted water is taken as a constant T_f = 271.2 K, and the ice density ρ_l is taken as a constant 900 kg m^{-3}. The mixed layer density in the Antarctic is set at 1027 kg m^{-3}, and that in the Arctic is set at 1025 kg m^{-3} [see *Tolstikov*, 1966; *Smith and English*, 1974]. Thus the ice density/water density ratio in both hemispheres is 0.88, and so by Archimedes' principle, 88% of the ice sits below the air-water interface. Consequently, upon adding the meltwater to the mixed layer the new volume-weighted temperature $T_{w'}$ becomes

$$T_{w',i} = [T_{w',i-1}(d_{mix} - 0.88h_{l,i-1}) + T_f \times 0.88(h_{l,i-1} - h_{l,i})]\,[d_{mix} - 0.88 \times h_{l,i}]^{-1} \quad (29)$$

The subscripts i or $i-1$ appended to a variable normally refer to the number of the time step. The exception comes when a variable W is recalculated more than once during a given time step and both W_i and W_{i-1} occur within the same calculation. In that case, W_i refers to the latest calculated value of W(pernaps the one currently being calculated, as for $T_{w'}$ in (29)), and W_{i-1} refers to the latest previous value. A nonsubscripted variable automatically indicates its latest value.

Vertical energy input to the lead. Reasonably, the energy flux to the lead area is the same as to the open water area (Figure 10*a*). Thus the energy per unit horizontal area that vertically enters the lead during a given time step is

$$Q_o = \Delta t \times [(1-\alpha_w)SW\downarrow + H\downarrow + LE\downarrow + \epsilon_w LW\downarrow - \epsilon_w\sigma T_w^4 + F_w\uparrow] \quad (30)$$

Where no ice exists within a grid square (section 3*e*, 'no ice') the total energy input, positive or negative, contributes to temperature change in the mixed layer. Here, however, positive energy input will be partially used to melt the ice laterally. Briefly, if Q_o is positive then the water is warmed and the lead area increased, whereas if Q_o is negative, then the water is cooled, and should the temperature sink to freezing, some water freezes to ice, thus reducing the lead area. To present the actual amounts of temperature and lead changes, we set A_{i-1}

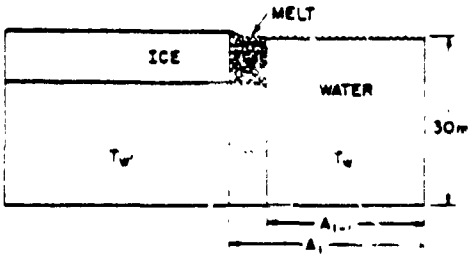


Fig. 12. Variables involved in calculating temperature changes due to lateral melt.

equal to the horizontal area of the grid square and A_{lead} equal to the horizontal area within the square covered by open water or leads. Since A is simply the fraction of leads, we have

$$A_{lead} = A_{sq} \times A \tag{31}$$

Temperature and lead area changes for positive Q_0. In the event of a positive flux, $A_{lead} \times Q_0$ is the total energy available to warm the water and laterally melt the ice. Of this amount, $A \times A_{lead} \times Q_0$ is allotted for warming, thus yielding a temperature increase of

$$\Delta T_w = \frac{A \times Q_0}{d_{mix} \times c_w} \tag{32}$$

and the remaining $(1 - A) \times A_{lead} \times Q_0$ is allotted for lateral melt, thus yielding an increase in horizontal lead area of

$$\Delta A_{lead} = \frac{(1 - A) \times A_{lead} \times Q_0}{Q_I \times h_I + Q_s \times h_s} \tag{33}$$

Although a crude parameterization, this does conform to the intuitive feeling that the larger the lead area, the greater the energy used directly in the lead rather than in melting the ice along the lead. Also should there be no ice then this parameterization does correctly use all the energy for warming the water. The only other parameterization known to the author is Semtner's, where all the energy is allotted to lateral melt regardless of the lead percentage [*Washington et al.*, 1976]. The Semtner parameterization is unrealistic when A approaches 100%.

The water temperature must be adjusted following (33). The melted ice (Figure 12) yields a volume-weighted water temperature in the newly formed lead area of

$$T = \frac{T_w(d_{mix} - 0.88h_I) + T_{ice} \times 0.88h_I}{d_{mix}} \tag{34}$$

and the water of the newly formed lead is assumed to mix immediately with that of the old lead. Thus the adjusted lead temperature becomes

$$T_{w,i} = \frac{T_{w,i-1} \times A_{i-1} + T(A_i - A_{i-1})}{A_i} \tag{35}$$

Furthermore, if the resulting A_{lead} from (33) exceeds A_{sq} (i.e., $A > 100\%$), then all ice has melted, and the energy which could have melted the additional ice instead further heats the water. In this case, A_i is set at precisely 100% for (35), and the resulting open water temperature T_w is further modified by the additional heat input from energy remaining after the total ice melt. Setting A' equal to the ratio of the calculated lead area to the grid square area, we have the adjusted water temperature

$$T_{w,i} = T_{w,i-1} + \frac{(A' - 100\%) \times (Q_I \times h_I + Q_s \times h_s)}{d_{mix} \times c_w} \tag{36}$$

Temperature and lead area changes for negative Q_0. If Q_0 is negative, water must be cooled to offset the energy deficit, the necessary cooling being

$$\Delta T_w = \frac{Q_0}{d_{mix} \times c_w} \tag{37}$$

This creates difficulties only if the calculated water temperature lies below the freezing point, 271.2 K, in which case the temperature is reduced only to freezing and the remaining energy deficit is balanced by the freezing of water at the lead surface. The freezing is done laterally onto the existing ice, maintaining the previously calculated ice thickness for the particular grid square. Should this result in an ice coverage exceeding 98% in the southern hemisphere or 99.5% in the northern hemisphere, then the ice coverage is restricted to 98 or 99.5%, and the remaining energy deficit is offset by a cooling of the water under the ice. (Hence we insist on at least a minimum lead percentage A_{min} of 2% in any grid square in the southern hemisphere and 0.5% in any grid square in the northern hemisphere. Though the model is not sensitive to the choice of 2% versus 0.5%, we have retained a larger value in the southern hemisphere than in the northern hemisphere simply because of the known larger amounts of open water throughout the Southern Ocean.) Finally, if the cooling decreases the water temperature below 271.2 K, then water is frozen to the underside of the ice. An energy balance is maintained throughout the calculations (see *Parkinson* [1978] for details). Also, when the lead percentage is decreased, the temperature under the ice is adjusted by a volume weighting.

Lateral mixing of water. The above temperature determinations are followed by a simple adjustment to account for lateral mixing:

$$\Delta T_{w'} = \tfrac{1}{4}A(T_w - T_{w'}) \tag{38}$$

$$\Delta T_w = \tfrac{1}{4}(1 - A)\left(1 - \frac{0.88h_I}{d_{mix}}\right)(T_{w'} - T_w) \tag{39}$$

The choice of the $\frac{1}{4}$ coefficient is arbitrary, though it clearly depends on the length of the time step. Through (38) and (39), temperature contrasts are reduced, but the total heat content of the water remains the same, the total increase/decrease of heat energy under the ice ($c_w \times (1 - A) \times A_{sq} \times (d_{mix} - 0.88h_I) \times \Delta T_{w'}$) balancing the heat decrease/increase in the leads ($c_w \times A \times A_{sq} \times d_{mix} \times \Delta T_w$).

g. *Transport Calculations*

The transport calculations proceed in two steps. Initially, a steady state velocity of the ice is calculated by balancing four major stresses: wind stress from above the ice τ_a, water stress from below the ice τ_w, Coriolis force D, and the stress from the tilt of the sea surface or dynamic topography G. This is followed by incorporating a fifth stress, the internal ice resistance produced by the interactions among the ice floes, while computationally attempting to move the ice according to the velocity vectors resulting from the four-stress steady state. The actual movement and the ice resistance force are discussed after the first four stresses and their balance have been presented.

Ice velocities. Using Newton's second law of motion, ice velocities $\mathbf{V}_I$ should be calculable from

$$\frac{d\mathbf{V}_I}{dt} = \sum_j \mathbf{F}_j \tag{40}$$

where the $\mathbf{F}_j$ are all forces (per unit mass) acting on the ice. As was stated above, we assume a steady state velocity, hence setting the left side of (40) to 0, and we restrict our $\mathbf{F}_j$ to four major stresses, reducing (40) to

$$0 = \tau_a + \tau_w + \mathbf{D} + \mathbf{G} \tag{41}$$

Arguments for a steady state versus a nonsteady state calculation have been given by *Campbell* [1964] and *Rothrock* [1973]. Rothrock shows through scale analysis that the acceleration term is generally 3 orders of magnitude smaller than the wind and water stresses.

Of the stresses considered here, only the Coriolis force has a precise formulation:

$$\mathbf{D} = \rho_i h_i f \mathbf{V}_i \times \mathbf{k} \tag{42}$$

[*Campbell*, 1964]. All terms have been defined (sections 3*f* ('temperature adjustments for vertical melting'), 3*e* ('snow-covered ice'), and 3*g* ('ice velocities')) except f, the Coriolis parameter, and $\mathbf{k}$, the unit vertical vector.

The stress from the dynamic topography involves a finite difference approximation to the gradient of sea surface height. With the dynamic topography fields $DT(I, J)$ plotted in Figures 7 and 8 the stress formulation becomes

$$\mathbf{G} = \{-10\rho_i h_i[(DT(I+1, J) - DT(I-1, J))\mathbf{i} + (DT(I, J+1) - DT(I, J-1))\mathbf{j}]\}(2 \times H)^{-1} \tag{43}$$

[*Campbell*, 1964]. The horizontal distance H between grid points is approximated at a constant 2×10^6 m.

The remaining two stresses are less precise. Wind stress is assumed to act in the direction of the wind which, following *Pritchard et al.* [1976], is approximated at 20° to the left of the surface geostrophic wind in the northern hemisphere and 20° to the right in the southern hemisphere. With β as the resulting angle in the x-y plane, wind stress is parameterized as

$$\tau_a = \alpha C_D \rho_a V_{ag}^2(\cos\beta\,\mathbf{i} + \sin\beta\,\mathbf{j}) \tag{44}$$

(*Hunkins* [1966]: $\tau_a = \rho_a C_D V_w^2$). Equivalently, the formulation is

$$\tau_a = \alpha C_D \rho_a V_{ag} \mathbf{V}_{ag} B \tag{45}$$

where B is the following rotation matrix:

$$B = \begin{bmatrix} \cos 20° & \sin 20° \\ -\sin 20° & \cos 20° \end{bmatrix} \tag{46a}$$

in the northern hemisphere and

$$B = \begin{bmatrix} \cos 20° & -\sin 20° \\ \sin 20° & \cos 20° \end{bmatrix} \tag{46b}$$

in the southern hemisphere. This latter formulation avoids the use of β. The drag coefficient C_D is set at 0.002 5, a figure reflecting the roughening effect of ridging [*Banke and Smith*, 1975], and $\alpha = 3$ is an adjustment factor to account for the use of monthly mean winds rather than instantaneous values [*Semtner* 1976*b*].

Following *Hunkins* [1966], water stress has been parameterized as

$$\tau_w = \rho_w(k_w|f|)^{1/2}(U_R\mathbf{i} + V_R\mathbf{j}) \tag{47}$$

with both the eddy viscosity $k_w = 24 \times 10^{-4}$ m^2 s^{-1} and the density of seawater ρ_w (section 3*f*, 'temperature adjustments for vertical melting') taken as constants. The vector $U_R\mathbf{i} + V_R\mathbf{j}$ is the vector difference between the ocean geostrophic velocity $\mathbf{V}_{og}$ and the ice velocity $\mathbf{V}_i$, i.e.,

$$U_R\mathbf{i} + V_R\mathbf{j} = (u_{og} - u_i)\mathbf{i} + (v_{og} - v_i)\mathbf{j} \tag{48}$$

The ocean geostrophic flow is obtained by balancing the dynamic topography and the Coriolis forces ($\mathbf{D} + \mathbf{G} = 0$), while the ice velocity is the unknown to be calculated by balancing the four stresses. The geostrophic balance yields an ocean geostrophic velocity

$$\mathbf{V}_{og} = \{-10[(DT(I, J+1) - DT(I, J-1))\mathbf{i} - (DT(I+1, J) - DT(I-1, J))\mathbf{j}]\}(2 \times f \times H)^{-1} \tag{49}$$

Recent work by Aidjex has suggested that perhaps a quadratic relationship between $U_R\mathbf{i} + V_R\mathbf{j}$ and τ_w would be more appropriate than the linear relation predicted by the classical Ekman approach and used in (47). The Aidjex formulation additionally includes a turning angle in the ocean boundary layer [*McPhee*, 1977].

Having defined the four stresses, we obtain the ice velocity from (41) by solving the two component equations

$$0 = \rho_i h_i f v_i - 10\rho_i h_i\left(\frac{1}{2 \times H}\right)[DT(I+1, J) - DT(I-1, J)] + \alpha C_D \rho_a V_{ag}^2 \cos\beta + \rho_w(k_w|f|)^{1/2}U_R \tag{50}$$

$$0 = -\rho_i h_i f u_i - 10\rho_i h_i\left(\frac{1}{2 \times H}\right)[DT(I, J+1) - DT(I, J-1)] + \alpha C_D \rho_a V_{ag}^2 \sin\beta + \rho_w(k_w|f|)^{1/2}V_R \tag{51}$$

To solve these, u_i and v_i are replaced by their equivalents $u_{og} - U_R$ and $v_{og} - V_R$, and (49) is inserted. In this manner, (50) and (51) simplify to a pair of simultaneous linear equations in U_R and V_R. Solving, the resulting velocity components are

$$U_R = \frac{-\rho_w(k_w|f|)^{1/2}\cos\beta - \rho_i h_i f \sin\beta}{\rho_w^2 k_w|f| + \rho_i^2 h_i^2 f^2}(\alpha C_D \rho_a V_{ag}^2) \tag{52}$$

$$V_R = \frac{-\rho_w(k_w|f|)^{1/2}\sin\beta + \rho_i h_i \cos\beta}{\rho_w^2 k_w|f| + \rho_i^2 h_i^2 f^2}(\alpha C_D \rho_a V_{ag}^2) \tag{53}$$

from which the velocity of the ice $\mathbf{V}_i = \mathbf{V}_{og} - \mathbf{V}_R$ is immediately calculated.

Ice movement. Once the ice velocities at each grid square have been found, the ice within each square is translated without rotation or distortion to the endpoint of its velocity vector; i.e., the midpoint of the ice is translated from (I, J) to $[I + (\Delta t \times u_i)/H, J + (\Delta t \times v_i)/H]$. It is at this point that attention must be given to the effect of continental boundaries and to the resistance forces among the ice floes. The modeled ice responds to continental boundaries by disallowing movement onto the shore; should either the u_i or the v_i component lead immediately into a land grid point, then that component is set to 0. Thus the boundaries have the same effect as that in the work of *Nikiforov et al.* [1970] and J. Walsh (personal communication, 1976), velocity components normal to and approaching the coast from an adjacent grid point being zeroed out and velocity components tangential to the coast remaining unchanged.

Internal ice resistance produces an effect similar to, though not as stringent as, that of continental boundaries. After the redistribution of ice according to the velocities calculated from the four-stress balance, all incoming velocities to any grid square with a resulting ice concentration $C_{\%}$ greater than 100% $- A_{min}$ are reduced. The reduction is done proportionately, after first insisting that the ice attempting to remain in the grid square remain. More specifically, we set $R_{\%}$ equal to the percentage of the grid square covered by ice that remains in the square, i.e.,

$$R_{\%} = (1 - A_{(I,J)})\left(1 - \frac{\Delta t}{H}\,u_{(I,J)}\right)\left(1 - \frac{\Delta t}{H}\,v_{(I,J)}\right) \tag{54}$$

We then reduce all incoming velocity components to this square from adjacent squares through multiplication by

$$\frac{100\% - A_{min} - R_{\%}}{C_{\%} - R_{\%}}$$

TABLE 3 Initial Values in the Standard Case

	Arctic	Antarctic
Ice thickness, m	3.5	1.5
Snow thickness, m	0.3	0.0
Percent area with open water, %	10	10
Temperature of open water, K	271.45	271.45
Temperature of water under ice, K	271.45	271.45
Temperature of upper snow surface, K	271.20	271.20
Temperature of snow-ice interface, K	271.20	271.20

Since the ice remaining in a square increases with a reduction in its velocity, this method generally requires several iterations through the 38 × 26 or 41 × 41 grid. The iterative procedure continues until no grid square has an ice concentration exceeding 100% − A_{min}.

After thus determining a field of ice velocities which does not create excessive convergence in any grid square, we complete the calculations for the time step by redistributing the ice according to the resulting velocities. If the ice velocity at point (I, J) is (u_I, v_I), then the ice in the grid square centered at (I, J) is translated, without rotation or distortion, until its center is positioned at (X, Y), where

$$X = I + (\Delta t \times u_I/H)$$
$$Y = J + (\Delta t \times v_I/H) \tag{55}$$

Hence both the volume and the areal coverage of ice which had existed in the square centered at (I, J) are distributed to the following four squares in the following percentages:

$$\begin{aligned}
&(1 - X + X_T)(1 - Y + Y_T)(100\%) \text{ to } (X_T, Y_T)\\
&(1 - X + X_T)(Y - Y_T)(100\%) \text{ to } (X_T, Y_T + 1)\\
&(X - X_T)(1 - Y + Y_T)(100\%) \text{ to } (X_T + 1, Y_T)\\
&(X - X_T)(Y - Y_T)(100\%) \text{ to } (X_T + 1, Y_T + 1)
\end{aligned} \tag{56}$$

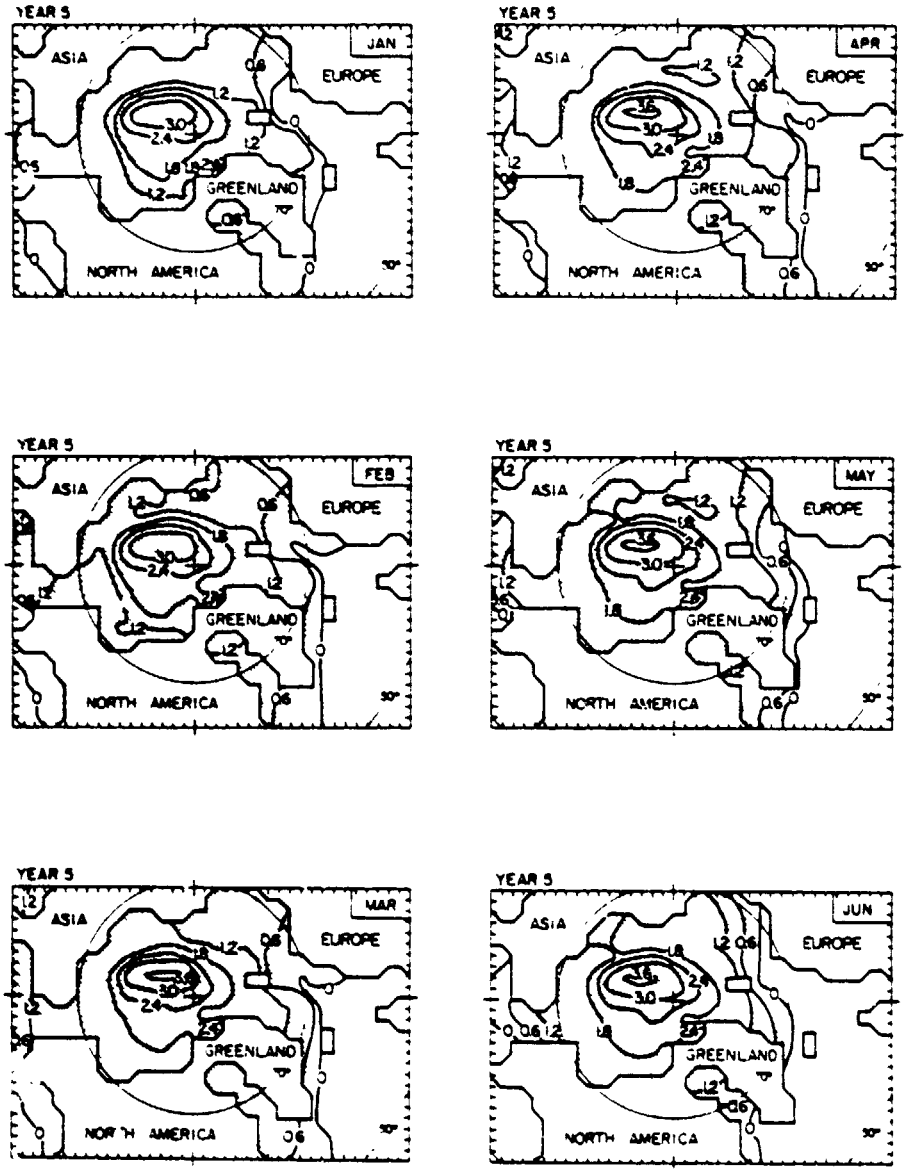


Fig. 13 Simulated half-yearly cycle of sea ice in the Arctic, January–June. Contours show ice thicknesses in meters. White shading indicates ice compactness above 90%.

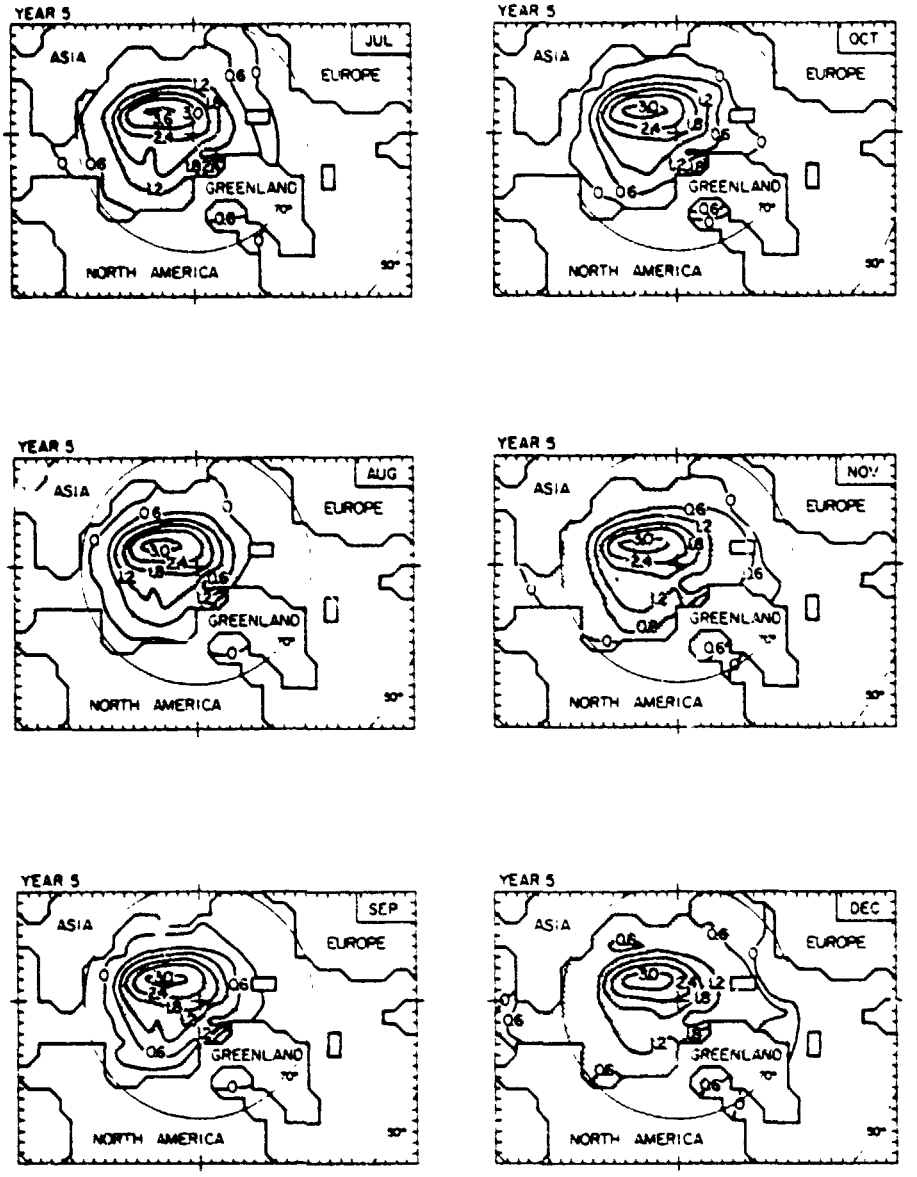


Fig 14 Simulated half-yearly cycle of sea ice in the Arctic, July–December Contours show ice thickness in meters, while shading indicates ice compactness above 90%

Subscript T signifies truncation of a number to its integral portion. The new ice thickness in an individual square is then obtained by summing the ice volumes remaining in and entering the square and dividing by the total area of ice coverage over the square. In this manner, each square retains only one ice thickness, and yet thickness can be advected from one square to another.

In the actual calculations the regions with nonzero internal ice resistance have tended to be quite small in the summer months but very significant in the winter months. This is illustrated later in Figures 19 and 20.

4. Model Results and Comparisons With Observations

In this section the model results are described and then compared with various estimated or observed ice distributions. Table 3 lists the initial conditions for each hemisphere Computations begin with a presumed time of 0800 UT, January 1, year 1, and run through 4 years of simulation in the Antarctic and 5 years in the Arctic This is sufficient time in each hemisphere for the results to approach an equilibrium yearly cycle [*Parkinson* 1978, section 5 4].

The resultant annual cycles obtained in the standard case are presented in Figures 13–16 The simulated arctic ice varies in extent from a minimum in September to a maximum in March, though the February and April extents are very close to those of March. In September the ice covers only a portion of the Arctic and has receded from most coastlines Thickness reaches 3 0 m in the center of the pack. At the March maximum, extent has greatly increased, reaching Iceland and beyond the southern coast of Greenland, while the central arctic thicknesses have reached 3 6

In the simulated Antarctic, minimum ice extent occurs in March and maximum extent in late August In March the ice reaches a latitude of 64°S off the coast of East Antarctica, the

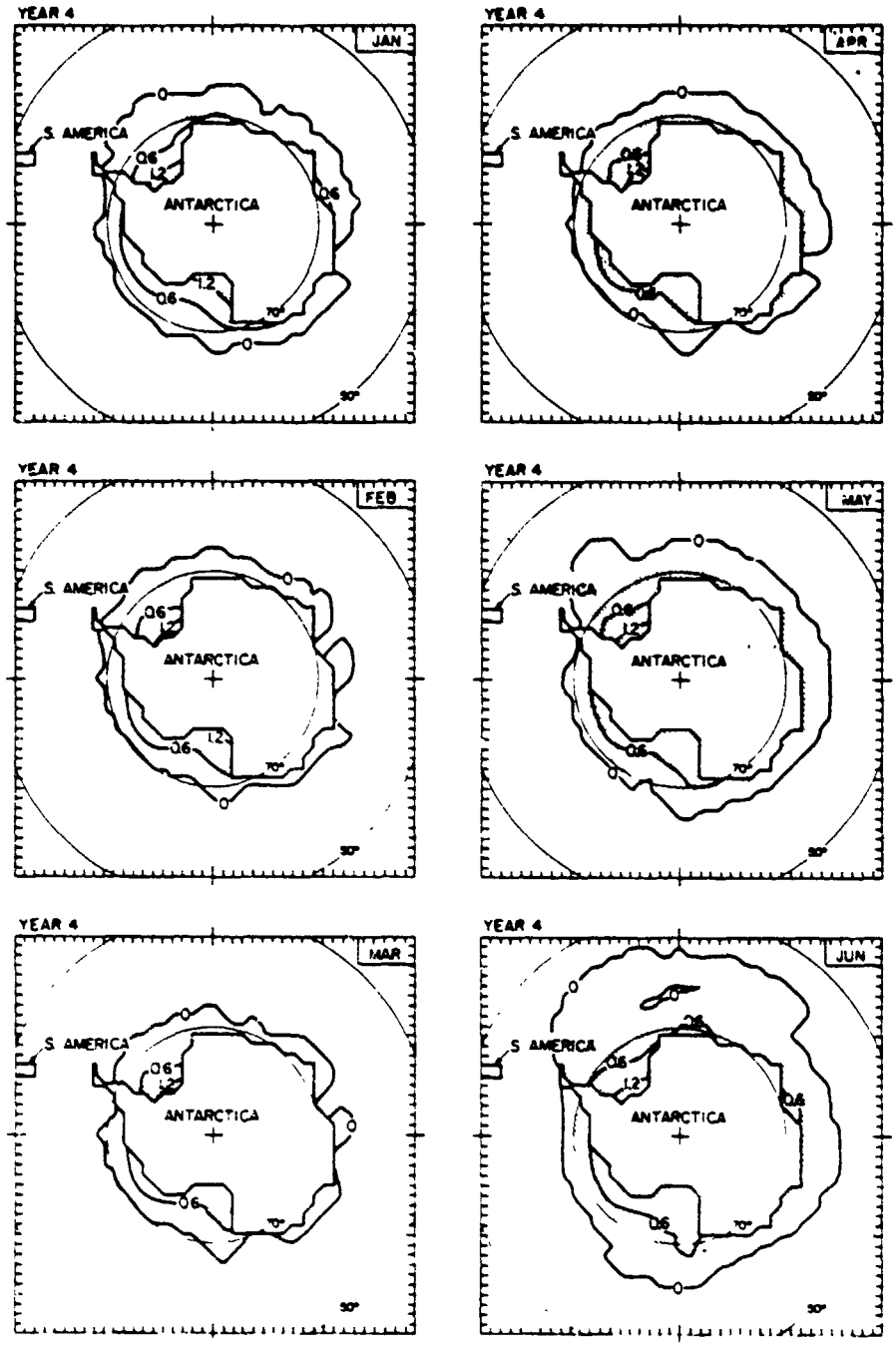


Fig. 15 Simulated half-yearly cycle of sea ice in the Antarctic, January–June. Contours show ice thickness in meters, while shading indicates ice compactness above 90%

maximum extension from the continent being in the Weddell Sea region to the east of the Antarctic Peninsula. Thickness varies from about 1.2 m in the Ross and Weddell seas to 0 m at the ice edge. In August the ice extends to the grid boundary over the approximate longitude range 20°W–10°E, reaching latitudes of 50°–51°S. Thickness varies from about 1.4 m at its coastal maximum to 0 m at the ice edge.

In the following sections the general pattern of arctic and antarctic results sketched above is examined more closely by comparison with atlas and satellite observations. The comparisons are divided into three sections, ice extent and distribution, ice thickness and lead areas, and ice drift, the fullest discussion being given for the extent and distribution. First, however, we briefly indicate the contrast between the extents and those that would be simulated by one of the very simplest of ice models, i.e., one placing the ice edge at the 271.2-K air temperature isotherm.

The model-simulated ice extent is clearly not a simple reproduction of the location of the freezing isoline for atmospheric temperatures (Figures 17 and 18). In the Arctic the fall freez-

ing line (September–November) definitely expands southward faster than the ice. By January the ice and the freezing contour show comparable extents, with the modeled ice extent beyond the freezing line off southern Alaska but within the freezing line in the Atlantic region. By the end of winter the modeled ice has expanded beyond the freezing line in all regions, and as the ice melts with the coming of summer, the ice line retreats much more slowly than the freezing line. This should be expected, owing to the inertia of the ice. By June the freezing contour encompasses only the central portion of the Arctic plus the region of Greenland. By July the mean air temperatures throughout the Arctic are above freezing, though in the model simulation, as in the real world, considerable ice remains. The ice continues melting into September, by which time the atmospheric freezing contour has expanded outward such that the ice and freezing lines again show a comparable extent. The September simulated ice extent is slightly beyond the freezing contour near Spitsbergen but not as far equatorward as the freezing line north of North America and along the West Greenland coast.

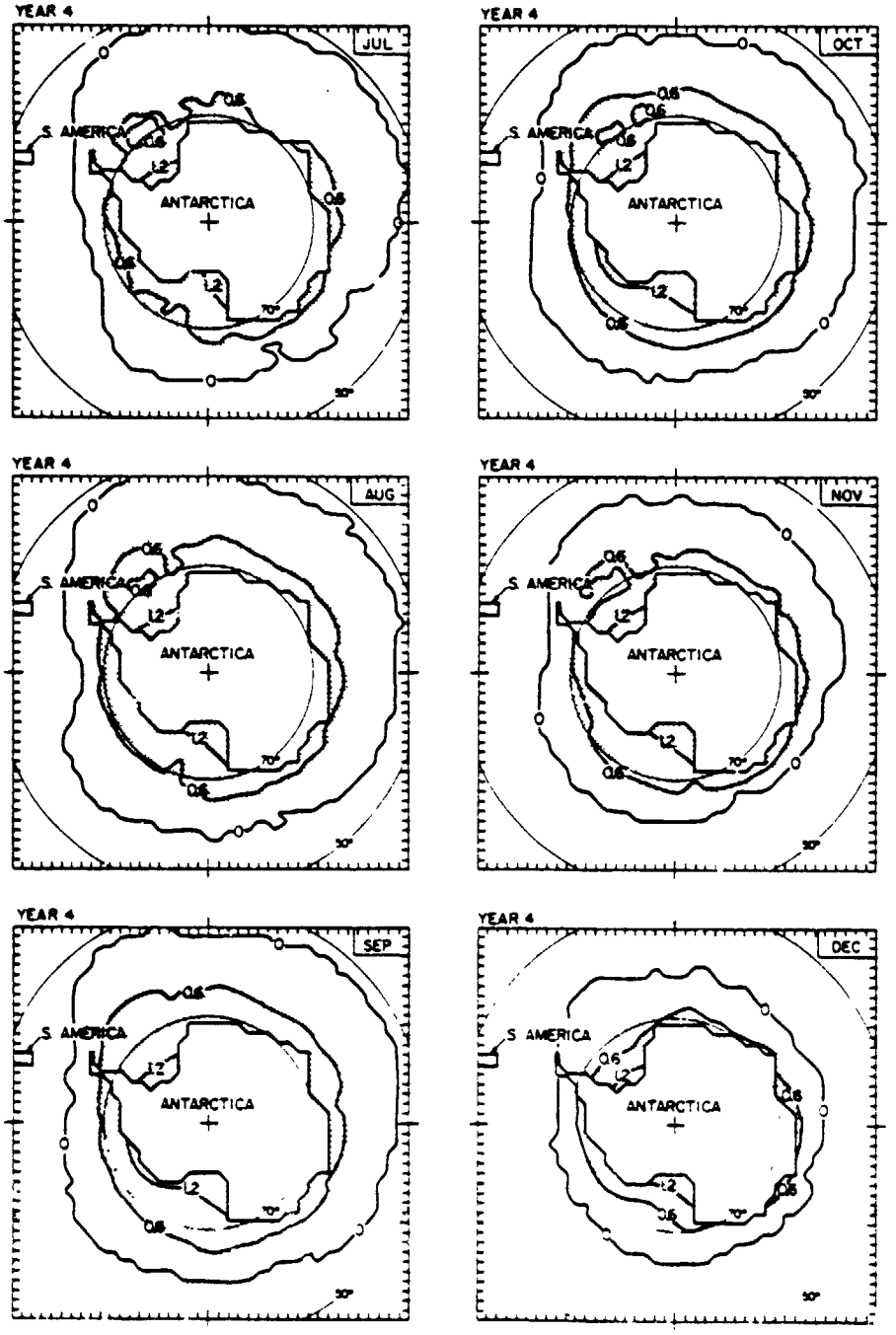


Fig. 16. Simulated half-yearly cycle of sea ice in the Antarctic, July–December. Contours show ice thickness in meters, while shading indicates ice compactness above 90%.

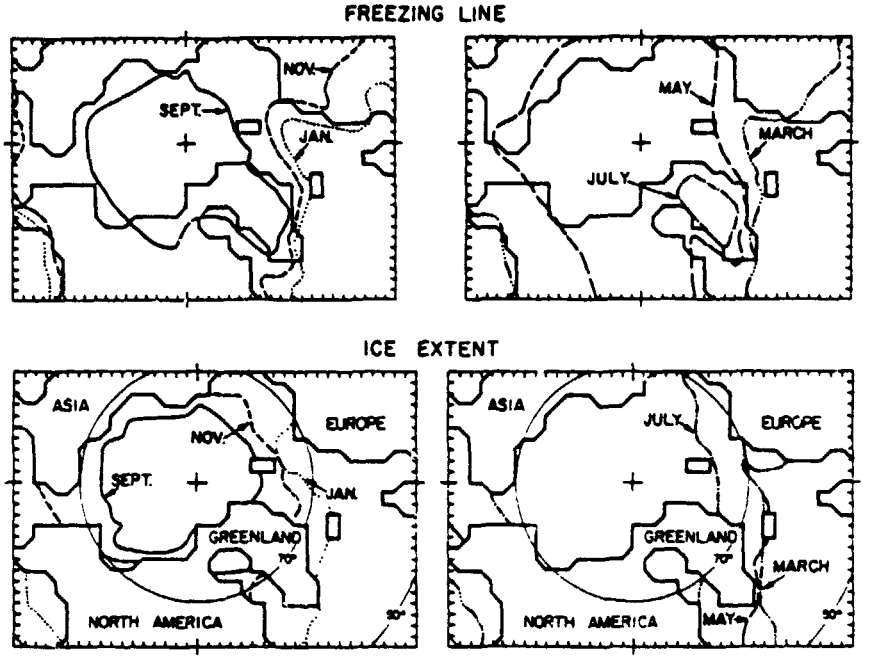


Fig. 17. Bimonthly simulated ice extents and observed mean atmospheric freezing (271.2 K) contour in the Arctic.

In the southern hemisphere also, the ice extent lags poleward of the freezing line in fall and retreats more slowly than the freezing line in spring (Figure 18). The two contours show similar extents in February, with the ice slightly equatorward of the freezing line at 180° longitude, though poleward of the freezing line off much of East Antarctica. The ice continues melting into March, while the freezing line is pushing outward. The ice extent remains poleward of freezing through June, then in July it tends to be poleward of freezing for the hemisphere from 90°E to 180° to 90°W and equatorward for the remaining hemisphere. Both contours reach their maximum extents in August The two curves are similar in August and September, though with the freezing line showing less longitudinal variation than the ice. During spring warming, the ice melt lags the poleward retreat of the freezing line, eventually catching up in most areas sometime in the peak summer months of January and February.

Ice Extent and Distribution

Actual ice distributions vary considerably from year to year, thus eliminating the possibility of a perfect yearly cycle against which to compare simulated results. This section begins therefore with an indication of the observed variability.

Pease [1975] summarizes significant differences in antarctic ice extent as depicted by three standard atlases: (1) the Soviet *Atlas of Antarctica* [*Tolstikov*, 1966], (2) the U.S. Navy *Oceanographic Atlas of the Polar Seas* [*Daniel*, 1957], and (3) the British Admiralty's *Ice Chart of the Southern Hemisphere* (1943) The U.S. atlas shows a larger extent of ice than the Soviet atlas for eastern longitudes and a smaller extent for western longitudes, this being true in every month. The British charts suggest more ice in summer than either the Soviet or U.S. sources. Furthermore, plots by *Gloersen and Salomonson* [1975] of ice extents from satellite microwave radiometry reveal a much more irregular distribution for the specific summer of 1972–1973 than is presented by any of the three atlases.

Prior to satellite sensing, apparent differences between sources on the antarctic ice boundary were often attributed to the sparsity and imprecision of observations rather than to real year-to-year contrasts. Though we remain plagued by sparse observations, satellite images have shown decisively the existence of large interannual differences, some of which are now mentioned.

Budd [1975] presents maps of satellite-observed October and December extents of antarctic sea ice for 1967 and 1968. In spite of the consecutiveness of the 2 years, clear contrasts exist in the pattern of ice distribution. In October 1968 the ice extent is more symmetric than it is in 1967 and reaches lower latitudes at almost all longitudes except those in the region north of the Weddell Sea. The ice is almost nonexistent off the east coast of the continent in December of 1967 and yet extends farther than the 1968 ice in the region of the 0° meridian. *Ackley and Keliher* [1976] have also compared antarctic ice extents at similar times in two consecutive years, finding the August 1–2 extent greater in 1973 than in 1974 in the Weddell Sea sector but less in the Ross Sea sector. Their data derive from electronically scanning microwave radiometer (ESMR) observations of the Nimbus V satellite.

Fletcher [1969] has conjectured that typical year-to-year variations in antarctic ice extent may be as large as typical variations between seasons Others emphasize changing distributions from one year to another rather than changes in overall extent [e.g., *Budd*, 1975]. As Budd mentions, large anomalies in individual regions may persist over entire seasons.

Interannual variation exists in the Arctic as well, though perhaps not to as striking an extent. Among the studies including plots of the ice boundary for various years are *Winchester and Bates* [1958] for September off northern Alaska, *Blindheim and Ljoen* [1972] for August in the vicinity of Spitsbergen, and *Haupt and Kant* [1976] for April in the entire Atlantic sector. Of the eight years presented by Haupt and Kant the ice mar-

gins in April reached the north coast of Iceland in only two years, 1968 and 1969, while they never reached the coast of Norway and always reached the southern tip of Greenland.

The lack of constancy from one year to another obviously eliminates the possibility of an absolute standard against which to judge numerical results. However, a general comparison can be made between model results and various estimates of actual ice extents, and when this is done, the overall correspondence is good.

More specifically, the simulated September and March maps in Figures 13 and 14 have ice boundaries at locations similar to those presented by *Wittmann and Schule* [1966] for the mean minimum and maximum ice extents. The major differences in September are that in the Wittmann and Schule case the ice extends farther along the east coast of Greenland and does not exist in the Baffin Bay region. The major differences in March are that Wittmann and Schule show a lesser extent in the Atlantic sector by an average of about 2° of latitude and do not show ice off the coast of southern Alaska.

Off northern Alaska the September ice edge in Figure 14 is well within the range presented by the 1953–1955 plots of *Winchester and Bates* [1958] and the 1968 plot of J. Walsh (personal communication, 1976), coming closest probably to 1968, though actually suggesting a greater extent than that year and a lesser extent than either 1953 or 1955. The April ice edge in the Atlantic (Figure 13) corresponds well with the *Haupt and Kant* [1976] curves for April extents from 1966 to 1973, though it indicates a slightly more southward extent especially off the southwest coast of Greenland. Also, the simulated results have the edge extending to Iceland which, from the Haupt and Kant maps, would be expected only 25% of the time. By contrast, in September the simulated ice has melted much more off the east coast of Greenland than *Rothrock* [1973] or *Wittmann and Schule* [1966] suggest. The correspondence between the simulated ice extent and the Rothrock curve is close for the rest of the Arctic, though the ice touches more of the coast of Siberia in the Rothrock case. For almost every area of the Arctic a reference could be found suggesting a greater extent than the simulation, and an alternate reference could be found suggesting a lesser extent.

In summary, the simulated arctic distribution and extent seem within the observed variability for most sections of the Arctic. The two areas of poor correspondence are to the south of Alaska, where too much ice is simulated, and immediately along the Greenland coast. The results show an abbreviated ice extension down the east coast of Greenland but too much ice off the southwest coast. These are both regions where the oceans are probably exerting a far greater influence in the real world than that allowed in the model. To the east of Greenland the predominant surface flow is from north to south, carrying cold arctic water and ice southward along the coast, while to the southwest of Greenland warm water flows from the Atlantic, moving northward along the coast in the West Greenland Current.

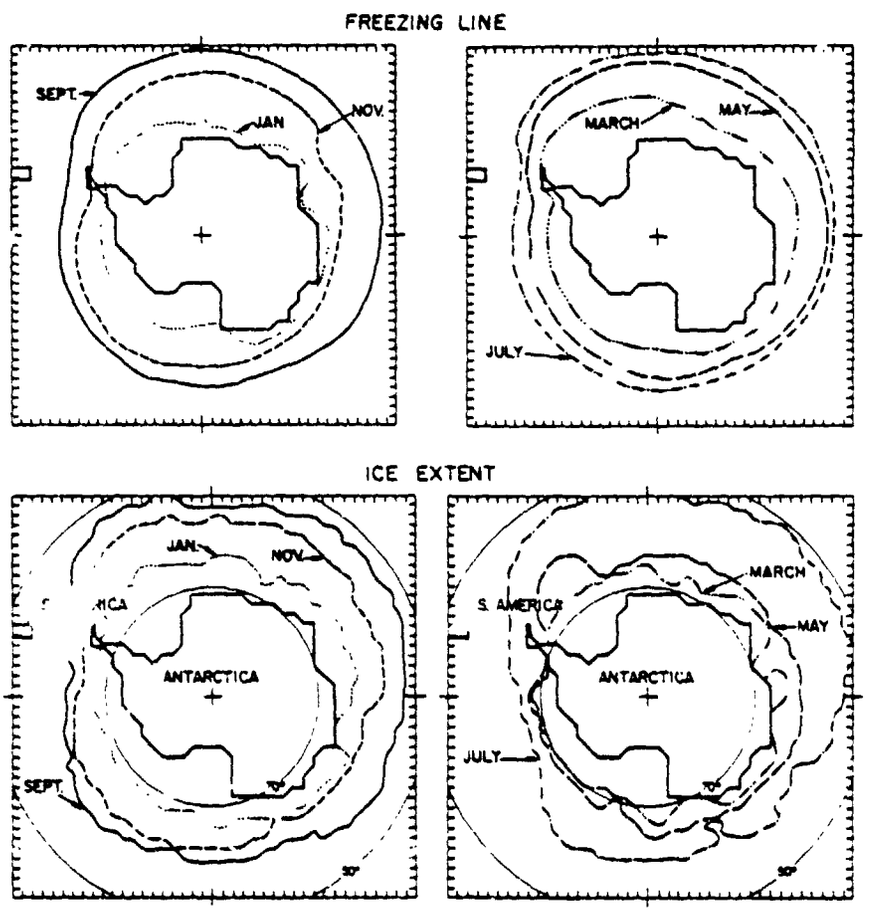


Fig. 18. Bimonthly simulated ice extents and observed mean atmospheric freezing (271.2 K) contour in the Antarctic.

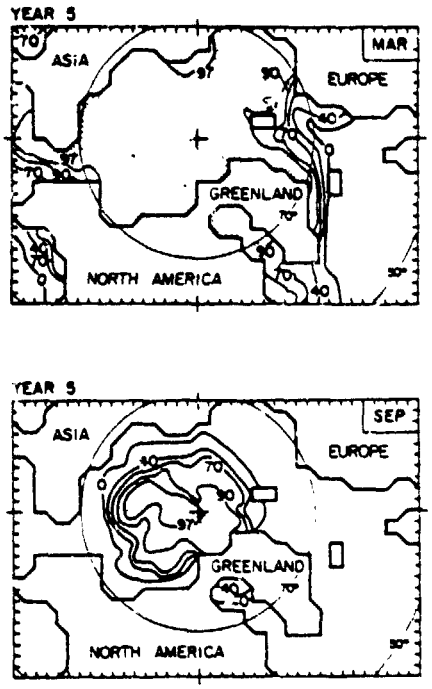


Fig. 19. Simulated contours of ice concentration (percent) in the Arctic, March and September. Shading indicates those regions where internal ice resistance was nonzero.

As is the case with the Arctic, the large-scale antarctic results also fall within the observed range. In each month the extents of Figures 15 and 16 correspond within about 2° of latitude to those of the U.S. Navy *Oceanographic Atlas of the Polar Seas* [*Daniel*, 1957]. The simulated extents tend to be somewhat less than the atlas extents from February to April, especially in the Antarctic Peninsula and Weddell Sea regions, and somewhat greater from July to December. In May the simulation shows more ice off East Antarctica and less ice along the peninsula than the *Oceanographic Atlas*.

The simulated ice limits differ from those of the Soviet *Atlas of Antarctica* [*Tolstikov*, 1966] most strongly off East Antarctica in the longitude range 40°–140°E. Within this region the simulated extents tend to be greater than the atlas extents, the strongest differences occurring in August and September, when the simulated ice reaches roughly 6° farther northward than the atlas charts. By contrast, in November the atlas and simulated extents are roughly identical in the eastern longitudes.

Finally, compared with the satellite-observed results of *Budd* [1975], the simulated December extent is noticeably greater at most eastern hemisphere longitudes than the observed extent in either 1967 or 1968, while the simulated October results seem quite realistic. It should be noted that the October 1968 observed distribution [*Budd*, 1975], along with the simulated October distribution, does not show the bulge north of the Weddell Sea suggested by several other sources [e.g., *Daniel*, 1957; *Neumann and Pierson*, 1966].

As for the times of minimum and maximum antarctic ice extent, the simulated minimum in March appears to be timed properly in terms of the typical real world cycle, while the maximum in late August appears to be slightly early. *Tolstikov* [1966], *Fletcher* [1969], and *Pease* [1975] all indicate March as the month of minimum extent, and Pease and Tolstikov indicate September as the month of maximum extent. Fletcher suggests that the maximum may occur in either September or October. On the other hand, the ESMR satellite data examined by *Ackley and Keliher* [1976] place the 1974 maximum extent near the end of August, the same as the simulated results. The fact that the ESMR data for 1973 have the maximum extent at the end of September [*Ackley and Keliher*, 1976] once again illustrates the real world variability.

b. *Ice Thickness and Lead Areas*

In addition to distribution and extent, ice thicknesses and concentrations are needed for a full picture of the ice coverage in either hemisphere. The simulated results in both aspects seem reasonable in the large-scale view, though perhaps the arctic thicknesses decrease too gradually from the thickest values in the central Arctic to the values at the ice edge and perhaps the antarctic thicknesses are too low, if only by 0–0.5 m.

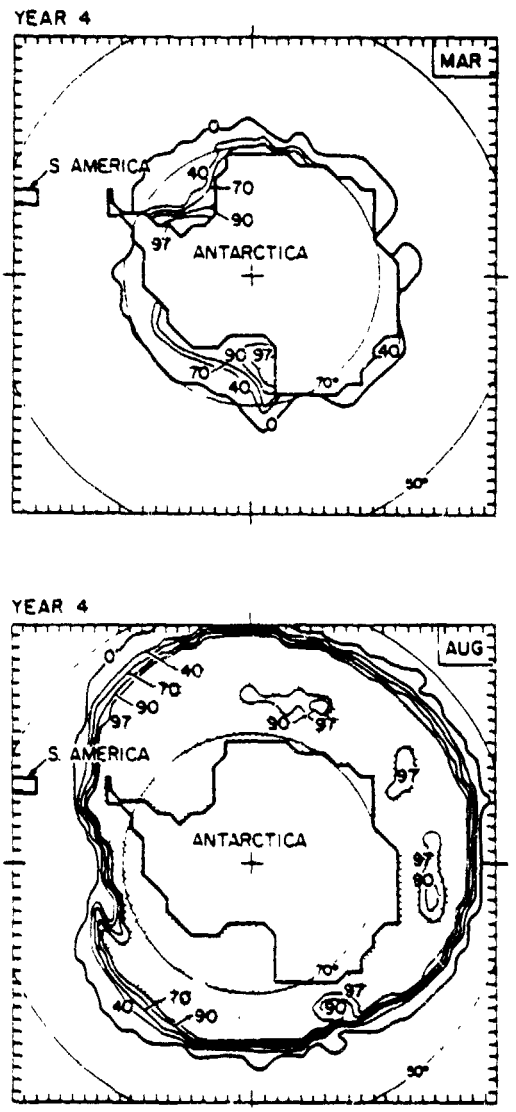


Fig. 20. Simulated contours of ice concentration (percent) in the Antarctic, March and August. Shading indicates those regions where internal ice resistance was nonzero.

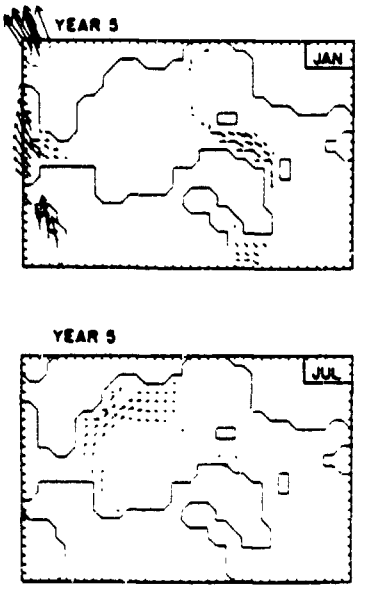


Fig. 21. Simulated arctic ice drift for January and July.

With reference to Figures 13 and 14, central arctic ice thickness tends to be about 3.0–3.6 m. Although these numbers are not entirely comparable either with figures obtained from a strictly thermodynamic model or with estimates of averaged arctic thicknesses, including ridged and otherwise deformed ice, we do mention for rough comparative purposes that the 3.0 to 3.6-m range exceeds the 2.88-m average thickness simulated by the one-dimensional model of *Maykut and Untersteiner* [1969] but is exceeded by the 4-m mean thickness suggested

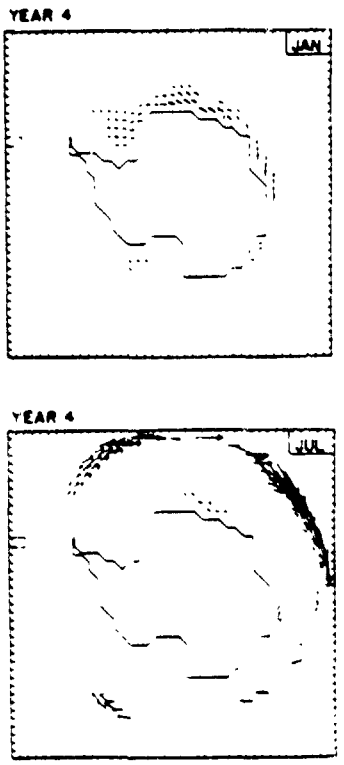


Fig. 22. Simulated antarctic ice drift for January and July

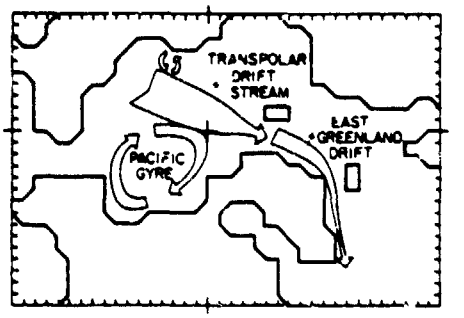


Fig. 23. Schematic diagram of the major observed features of arctic ice drift

by *Fletcher* [1965]. Spatially, the first-order tendency is for the simulated ice thicknesses to decrease outward from the peak values in the central Arctic (Figures 13 and 14), the major exception to this being in the region directly north of Greenland and Ellesmere Island, where thicknesses again increase as the coast is approached.

As for the yearly cycle of central ice thicknesses, the simulated values of 3.0 m at the end of summer and 3.6 m at the end of winter are comparable to the seasonal differences suggested by others. *Neumann and Pierson* [1966] suggest that the polar cap averages 2.0–2.5 m toward the end of summer and 3.0–3.5 m toward the end of winter, a somewhat larger range than the simulated results, while *Maykut and Untersteiner* [1969] obtain in their model a somewhat smaller range than that simulated here, theirs being 2.71–3.14 m.

Simulated antarctic thicknesses (Figures 15 and 16) decrease fairly monotonically from coastal values of 0–1.5 m out to 0 m at the ice edge. Since estimates of antarctic sea ice suggest that most of the ice is no greater than 1.0–1.5 m thick [*Baker*, 1976], the simulated thicknesses are realistic or at least no more than about 0.5 m too thin. One method of thickening the simulated ice is to decrease the oceanic heat flux and hence the bottom ablation. Such thickening, however, occurs at the expense of increasing ice extent as well. Smaller-scale features of antarctic ice thickness, such as the observed tendency for thicker ice at the northern reaches of the Weddell Sea than at the Weddell Sea coast, are not reproduced. Proper simulation of such features may require detailed modeling of the ocean circulation.

In addition to the monthly 90% ice concentration contours in Figures 13–16, Figures 19 and 20 present more detailed contour maps of concentration for the months of maximum and minimum extent. In view of the lack of precise data the simulated concentrations seem reasonable. *Wittmann and Schule* [1966] claim that at least 5% of the central Arctic consists of open water during all seasons. The modeled results show less than 5% throughout the central Arctic in winter, corresponding poorly with the Wittmann and Schule estimates. However, *Maykut* [1976] claims a lead percentage of under 1% in winter in the central Arctic, and *Weeks* [1976] claims a winter ice concentration of 99%, ranging only from 98 to 100%, and a summer concentration of 92%, ranging from 30 to 100%. Similarly, Koerner claims that only 0.6% of the area is ice free in winter, and Swithinbank claims a March total of 5% for all open water and ice of thickness less than 0.3 m, only a small fraction of this 5% being open water [*Swithinbank*, 1972]. The ice of Figures 13 and 14 is less compact than the estimates of Weeks or Koerner and more compact than the

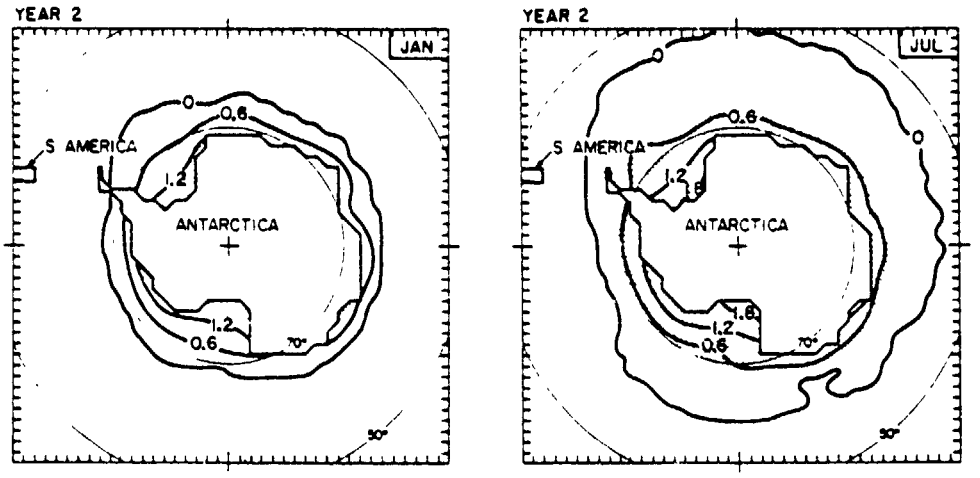


Fig. 24 Simulated January and July ice thickness in the Antarctic after eliminating ice dynamics. Contours show thickness in meters, while shading indicates ice compactness above 90%

estimates of Wittmann and Schule. As was pointed out by one of the reviewers, it should be mentioned that neither submarine sonar profiles nor satellite imagery can yet distinguish clearly between open water and thin ice. Thus the estimates of Wittmann and Schule, which were obtained from submarine data, should be interpreted as the combined percentages of open water and thin ice, thereby accounting for the higher values.

In the Antarctic the patterns of simulated ice concentrations correspond well with those of *Neumann and Pierson* [1966] and additionally reveal a broad region of lowered concentrations in the area most noted for large polynyas, to the north of Queen Maud Land at about the Greenwich meridian [*Daniel*, 1957; *Zwally and Gloersen*, 1977]. *Ackley and Keliher* [1976] report that satellite data for the antarctic winters of 1973 and 1974 reveal far more open water than most previous estimates and, incidentally, more than the simulated results, but it must be recalled that satellite imagery is not able to distinguish the thin ice from the open water. The overall ice concentrations reported by Ackley and Keliher between the continent and the ice edge average 70–80% in the Weddell Sea sector, 65–75% in the Ross Sea sector, 55–65% in the East Antarctic sector, and 50–65% in the Amundsen-Bellingshausen Sea sector. In addition to the difficulties in distinguishing thin ice from water, Ackley and Keliher warn that their method further overestimates the lead percentages since both fresh snow and clouds with high water content tend to lower brightness temperatures down toward their seawater values. In a further analysis, *Zwally and Gloersen* [1977] claim that on an average, the ice concentrations derived by Ackley and Keliher are probably about 10–15% too low.

c *Ice Drift*

Simulated January and July ice velocities appear in Figures 21 and 22. These are more difficult to compare against observations than the extents or thicknesses, since considerably less observational data exist, especially in the southern hemisphere. However, major qualitative aspects of the large-scale drift in the Arctic are fairly well known, and of benefit to future studies, procedures are now being developed for esti-

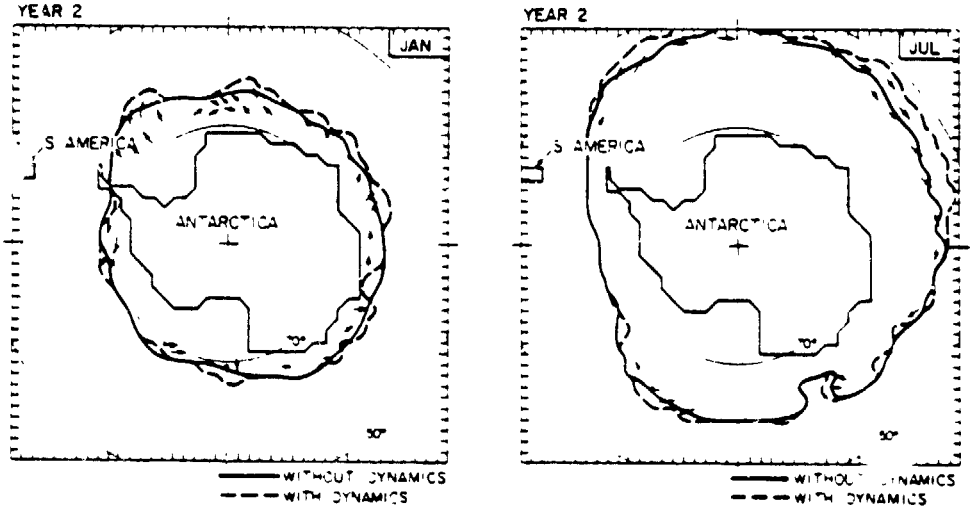


Fig. 25 January and July ice extents in the Antarctic with and without ice dynamics simulated. Arrows indicate the directions of the simulated ice velocities

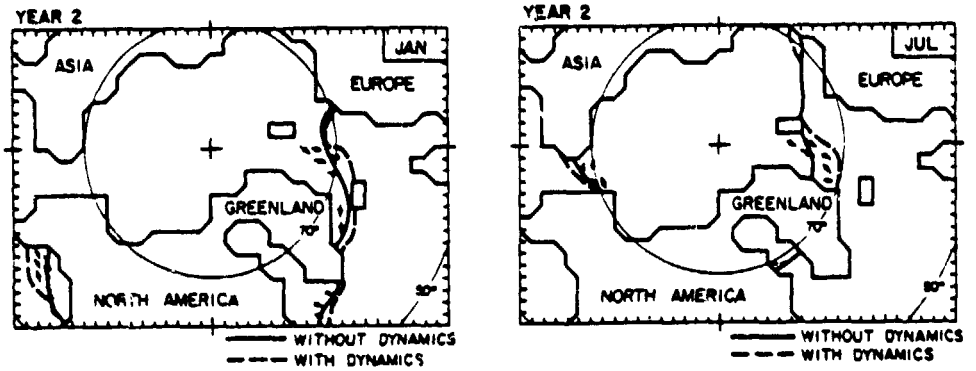


Fig. 26 January and July ice extents in the Arctic with and without ice dynamics simulated. Arrows indicate the directions of the simulated ice velocities.

mating drift from Landsat satellite photographs [*Hibler et al.*, 1976].

Beginning with the northern hemisphere, schematic plots of the observed patterns of arctic ice drift almost always emphasize two main features: (1) the anticyclonic Pacific gyre in the Beaufort Sea and Central Arctic Basin and (2) the Transpolar Drift Stream crossing the pole from the Laptev and East Siberian seas and moving out toward the Atlantic, eventually merging with the East Greenland Drift. Smaller features also often mentioned are a stagnant region north of Greenland and Ellesmere Island and a small cyclonic gyre to the east of Severnaya Zemlya [*Rothrock*, 1973]. These various observed features are schematized in Figure 23.

Both the Pacific gyre and the Transpolar Drift are visible in the simulated results. The gyre appears prominently in the July drift chart, its center occurring at roughly 78°N, 210°E. This location for the gyre center is quite close to the 79°N, 200°–205°E positioning determined from data of a Soviet expedition of 1961 [*Bushuyev et al.*, 1970]. However, the simulated January results do not indicate a gyre pattern, and indeed, the internal ice resistance has reduced the January ice velocities to near zero throughout most of the central Arctic. Although this velocity reduction is probably excessive, two supportive points can be noted regarding these January results. First, the wind field specified for the present model, a field taken from mean monthly observational data (section 3*b*, 'atmospheric data'), also does not contain a gyre in the Beaufort Sea in the winter months (Figure 3). Second, there is little doubt that winter ice velocities are slowed owing to compaction. This second point is illustrated by the reduction of wind coefficients from summer to winter [*Nikiforov et al.*, 1970; *McPhee*, 1977] and also by observed instances when highly concentrated winter ice has remained motionless [e.g., *Pritchard*, 1976]; however, these instances of stationary ice may not be typical.

The Transpolar Drift Stream appears in the simulated results for both January and July (Figure 21), with the continuing southward motion in the East Greenland Drift also prominent in the January results. As for the smaller-scale features of arctic drift, there is indeed a relatively stagnant region to the north of Greenland and Ellesmere Island in both January and July plots, and there is a slight cyclonic curvature in the region to the east of Severnaya Zemlya in January, though the island itself is not modeled.

The simulated velocity fields for the Antarctic show westerly motion wherever the ice reaches northward of about 58°S (Figure 22). Ice velocities tend to be greatest in this region of strong westerly atmospheric flow, the weaker velocities near the coast being more variable in direction but predominantly easterly and having cyclonic curvature. The cyclonic flow in the Weddell Sea is well corroborated by the documented drift of a large iceberg from 1967 to 1976 [*McClain*, 1976], while the overall flow patterns of Figure 22 correspond closely with the large-scale ocean circulation in the Southern Ocean [*Tolstikov*, 1969; *Zwally et al.*, 1976].

5. Experiments on the Importance of Ice Transport and Internal Ice Resistance

When the model is run without allowing a horizontal transport of the ice, the results show only slight contrasts with the standard case. Since the basic trends become apparent early in the simulation, these experiments are run for only 2 years rather than the longer time periods of the standard cases.

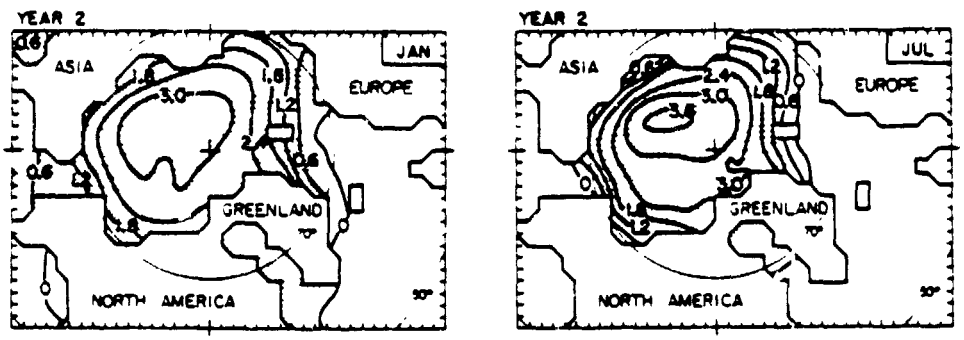


Fig. 27 Simulated January and July ice thickness in the Arctic after eliminating ice dynamics. Contours show thickness in meters, while shading indicates ice compactness above 90%

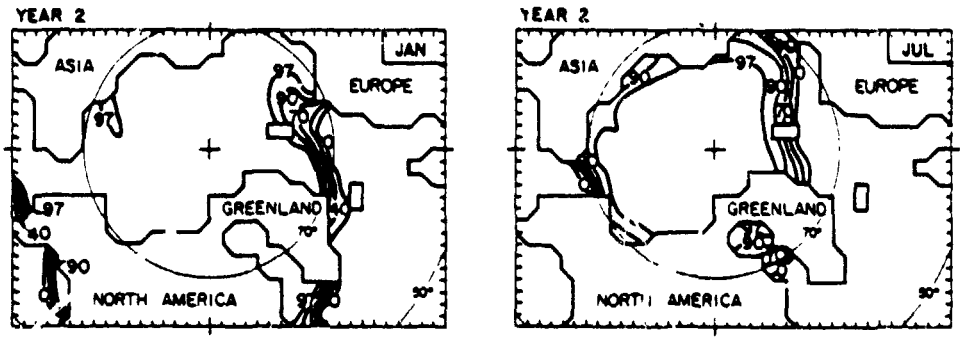


Fig. 28. Simulated January and July ice compactness (percent) in the Arctic after eliminating ice dynamics.

In the Antarctic (Figure 24) the extents are generally not as great as they are when ice transport allows a slight equatorward flow. Figure 25 overlays the extents in year 2 of the no-transport case with those in year 2 of the standard case and plots the directions of the ice velocity near the ice edge. In regions where the meridional component of flow is southerly, the extent is greatest when transport is included, while in regions of negligible flow the extents are identical, and in the few regions with a northerly flow component the extents are greatest without transport (Figure 25). Naturally, the contrast between extents with or without transport would be much greater if the predominant ice flow at the equatorward edge of the pack had been meridional rather than zonal.

In the North Atlantic the simulated meridional flow component at the ice edge tends to be somewhat stronger than in the Antarctic. As this flow is equatorward, the ice extent is thereby more noticeably reduced when transport is eliminated (Figure 26). Furthermore, the region of ice of thickness greater than 3 m is also slightly reduced without transport (Figure 27 versus year 2 of the standard case), presumably reflecting the small outward flow component in the Pacific gyre. At the same time the ice is more compact without transport (Figure 28), again reflecting the divergent velocity component. Local regions of convergent flow will naturally show less compaction without transport; however, on a large scale in the Arctic the transport increases the extent of the ice and decreases the compactness.

In view of the large amount of computer time required for the transport calculations and the relatively small effect they make on simulated ice distributions, it would appear appropriate in some uses to model thermodynamics alone. This could be true, for instance, in large-scale atmospheric general circulation models where the ice simulation is desired primarily for its influence in increasing shortwave albedo and decreasing heat exchanges between ocean and atmosphere.

In a further experiment on ice transport we calculate ice velocities simply by balancing wind stress, water stress, Coriolis force, and dynamic topography, without then proceeding to reduce these velocities due to internal ice resistance as described in section 3g ('ice movement'). Instead, we allow the thicknesses to increase according to the volume of ice converging in the grid square. This results in unrealistic ice distributions, such that in regions of ice convergence, thicknesses become excessive, while sometimes in adjacent grid squares the ice has disappeared altogether. From this we conclude that if ice transport is modeled, some parameterization must allow for the effect of internal ice resistance. The regions of nonzero internal ice resistance for 2 months in each hemisphere in the standard case are shaded in Figures 19 and 20.

6. Discussion

This study has described the construction of a numerical model of the growth and decay of sea ice and has compared the results of the model against observations. The basic large-scale features of the simulations correspond nicely with observations in both hemispheres. However, the model retains many simplifications which limit its accuracy, and in this concluding section we discuss some of these limitations and indicate prospects for future development.

Certain restrictions are obvious and are built into the model equations. Major among these is the insistence on uniform ice thickness (with an allowance for leads) over the 200 × 200 km grid squares. Another example is the reduction of such variables as the drag coefficient, the coefficients of sensible and latent heat exchange, the conductivities of ice and snow, and the values of the oceanic heat flux to constant values.

Other restrictions derive from the nature of the input data, for instance, the use of mean monthly atmospheric forcing. As was emphasized in section 4*a*, significant year-to-year variations in ice extent exist in both the Arctic and the Antarctic. By employing a constant yearly cycle of atmospheric conditions the present numerical model is unable to simulate such real world interannual variability.

Equally important, the use of mean monthly averages smoothes out the temperature and wind extremes, both of which can be important in creating anomalous local conditions which can persist and expand in influence over time. For instance, large polynyas often open during intense storms [*Vowinckel*, 1966; *Knapp*, 1972], and a concentrated ice pack will sometimes remain motionless until the wind stress becomes untypically large [*Pritchard*, 1976]. The use of mean winds prevents a proper simulation of such occurrences. Similarly, mean temperatures eliminate the short but sometimes significant melting periods accompanying anomalously warm conditions. This ties in with the additional restriction that the diurnal cycle is not included.

Further complicating matters in the case of the wind input, since opposing wind directions partially cancel, the use of averaged wind vectors does not produce averaged wind speeds. Thus the wind speeds specified in the model are particularly unrealistic in regions without dominant wind directions within the individual months.

In addition to the atmospheric simplifications, oceanic simplifications also prevent precise correspondence between simulated results and real world ice configurations. For instance, ocean salinity is not modeled, and the spatial variations of salinity are not included, although, among other consequences, the salinity affects the freezing temperature and the

density of the water. Relatedly, the use of a constant mixed layer depth ignores the influence of vertical convection, which is tied closely to the mutual interrelatedness of the freezing and melting of the ice on the one hand and the properties of the mixed oceanic layer on the other. Clearly, the failure to include salinity and a variable mixed layer depth restricts the ability of the model to simulate in precise detail the freeze-melt cycle.

Many of the current restrictions of the model can be eliminated upon coupling with atmospheric and oceanic models. An atmospheric model will allow calculation of winds, air temperatures, and dew points on a shorter time scale than monthly averages, and a full oceanic model will allow a variable mixed layer, interaction with the deeper ocean and adjacent river and oceanic flows, calculation of the changing salinity field, and determination of the underlying oceanic circulation. The model should then be able to simulate year-to-year variations and should allow analyses of various interactions among air, sea, and ice. Any contribution that such a coupled model could make to understanding the atmosphere-hydrosphere-cryosphere climatic system would be valuable.

Acknowledgments. This work was carried out at the National Center for Atmospheric Research (NCAR) under a National Science Foundation grant (NSF grant ATM76-08492) through Ohio State University. The authors thank J. Rayner for helping to obtain funding from NSF and for offering comments at various stages of the project. A. Semtner, Jr., is thanked for the use of his thermodynamic model, H. van Loon for many discussions and suggestions on data sources, R. Jenne for providing access to magnetic tapes of atmospheric data, and A. Gordon for maps of the dynamic topography of the Southern Ocean. We feel that this paper has been improved by several valuable comments from the reviewers G. Maykut and M. MacCracken. We also appreciated comments and suggestions from J. Walsh, L. Morrison, E. Thompson, G. Demko, A. J. Arnfield, T. Mayer, N. Untersteiner, T. Seliga, E. Leavitt, C. Bull, and E. Taaffe, and we benefited from our association with the Institute of Polar Studies at Ohio State University. A. Modahl typed the manuscript and provided much appreciated editorial assistance, while G. Meehl and the NCAR Graphics Department drafted the figures. NCAR is sponsored by the National Science Foundation.

References

Ackley, S. F., and T. E. Keliher, Antarctic sea ice dynamics and its possible climatic effects, *Aidjex Bull., 33*, 53–76, 1976.

Badgley, F. I., Heat budget at the surface of the Arctic Ocean, in *Proceedings of the Symposium on the Arctic Heat Budget and Atmospheric Circulation*, edited by J. O. Fletcher, pp. 267–277, Rand Corporation, Santa Monica, Calif., 1966.

Baker, D. J., Jr. (Ed.), *The Polar Sub-Programme of the Global Atmospheric Research Programme*, 97 pp., Department of Oceanography, University of Washington, Seattle, 1976.

Banke, E. G., and S. D. Smith, Measurement of form drag on ice ridges, *Aidjex Bull., 28*, 21–27, 1975.

Blindheim, J., and R. Ljoen, On the hydrographic conditions in the West Spitsbergen Current in relation to ice distribution during the years 1956–1963, in *Sea Ice: Proceedings of an International Conference, Reykjavik, Iceland, May 10–13, 1971*, edited by I. Karlsson, pp. 33–41, National Research Council of Iceland, Reykjavik, 1972.

Bryan, K., S. Manabe, and R. C. Pacanowski, A global ocean-atmosphere climate model, II, The oceanic circulation, *J. Phys. Oceanogr., 5*, 30–46, 1975.

Budd, W. F., Antarctic sea-ice variations from satellite sensing in relation to climate, *J. Glaciol., 15*, 417–427, 1975.

Bushuyev, A. V., N. A. Volkov, Z. M. Gudkovich, and V. S. Loshchilov, Results of expedition investigations of the drift and dynamics of the ice cover of the Arctic Basin during the spring of 1961, *Aidjex Bull., 3*, 1–21, 1970.

Campbell, W. J., *On the Steady-State Flow of Sea Ice*, 167 pp., Department of Atmospheric Sciences, University of Washington, Seattle, 1964.

Campbell, W. J., Sea-ice dynamics, in *Arctic Drifting Stations: A Report on Activities Supported by the Office of Naval Research*, pp. 189–196, Arctic Institute of North America, Washington, D. C., 1968.

Coachman, L. K., and K. Aagaard, Physical oceanography of arctic and sub-arctic seas, in *Arctic Geology and Oceanography*, edited by M. Nelson and Y. Herman, pp. 1–72, Springer, New York, 1974.

Coon, M. D., G. A. Maykut, R. S. Pritchard, D. A. Rothrock, and A. S. Thorndike, Modeling the pack ice as an elastic plastic material, *Aidjex Bull., 24*, 1–105, 1974.

Coon, M. D., R. Colony, R. S. Pritchard, and D. A. Rothrock, Calculations to test a pack ice model, in *Numerical Methods in Geomechanics, II*, edited by C. S. Desai, pp. 1210–1227, American Society of Civil Engineers, New York, 1976.

Crutcher, H. L., and J. M. Meserve, Selected level heights, temperatures and dew points for the northern hemisphere, *NAVAIR Rep. 50-1C-52*, revised, U.S. Nav. Weather Serv. Command, Washington, D. C., 1970.

Daniel, H. C., *Oceanographic Atlas of the Polar Seas, Part I, Antarctic, Publ. 705*, 70 pp., U.S. Navy Hydrogr. Office, Washington, D. C., 1957.

Donn, W. L., and D. M. Shaw, The heat budgets of an ice-free and an ice-covered Arctic Ocean, *J. Geophys. Res., 71*, 1087–1093, 1966.

Felzenbaum, A. I., The theory of the steady drift of ice and the calculation of the long period mean drift in the central part of the Arctic Basin, *Probl. Severn., 2*, 5–14, 1958.

Fletcher, J. O., The heat budget of the Arctic Basin and its relation to climate, *Rep. R-444-PR*, 179 pp., Rand Corp., Santa Monica, Calif., 1965.

Fletcher, J. O., Ice extent in the Southern Ocean and its relation to world climate, *Memo RM-5777-NSF*, 108 pp., Rand Corp., Santa Monica, Calif., March 1969.

Gloersen, P., and W. J. Campbell, Satellites—New global observing techniques for ice and snow, *J. Glaciol., 15*, 373–389, 1975.

Gordienko, P., Arctic ice drift, in *Proceedings of the Conference on Arctic Sea Ice, Publ. 598*, pp. 210–222, National Academy of Sciences/National Research Council, Washington, D. C., 1958.

Hall, R. T., G. A. Maykut, and D. A. Rothrock, Sea ice modeling: Its testing with Landsat and potential use in FGGE, in *Proceedings of the Symposium on Meteorological Observations From Space: Their Contribution to the First Garp Global Experiment*, pp. 104–106, Committee on Space Research, International Council of Scientific Unions, Paris, 1976.

Haltiner, G. J., and F. L. Martin, *Dynamical and Physical Meteorology*, 470 pp., McGraw-Hill, New York, 1957.

Haupt, I., and V. Kant, Satellite ice surveillance studies in the Arctic in relationship to the general circulation, in *Proceedings of the Symposium on Meteorological Observations From Space: Their Contribution to the First Garp Global Experiment*, pp. 179–187, Committee on Space Research, International Council of Scientific Unions, Paris, 1976.

Herman, G., *Radiative-Diffusive Models of the Arctic Boundary Layer*, 170 pp., Department of Meteorology, Massachusetts Institute of Technology, Cambridge, 1975.

Hess, S. L., *Introduction to Theoretical Meteorology*, 362 pp., Holt, Rinehart, and Winston, New York, 1959.

Hibler, W. D., III, Modeling pack ice as a viscous-plastic continuum: Some preliminary results, in *A Symposium on Sea Ice Processes and Models, September 5–9, 1977*, vol. II, pp. 46–55, University of Washington, Seattle, 1977.

Hibler, W. D., III, W. B. Tucker, and W. F. Weeks, Techniques for using Landsat imagery without ground references to study sea ice drift and deformation, *Aidjex Bull., 31*, 115–135, 1976.

Hunkins, K., Ekman drift currents in the Arctic Ocean, *Deep Sea Res., 13*, 607–620, 1966.

Hunkins, K., The oceanic boundary layer and ice-water stress during Aidjex 1972, *Aidjex Bull., 26*, 109–127, 1974.

Huschke, R. E., *Glossary of Meteorology*, 638 pp., American Meteorological Society, Boston, Mass., 1959.

Huschke, R. E., Arctic cloud statistics from "air-calibrated" surface weather observations, *Memo. RM-6173-PR*, 79 pp., Rand Corp., Santa Monica, Calif., 1969.

Idso, S. B., and R. D. Jackson, Thermal radiation from the atmosphere, *J. Geophys. Res., 74*, 5397–5403, 1969.

Jacobs, S. S., C. B. Bauer, P. M. Bruchhausen, A. L. Gordon, T. S. Root, and F. L. Rosselot, *Climax Reports, Cruises 47–50, 1971–55, 1972*, 503 pp., Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York, 1974.

Kellogg, W. W., Meteorological observations in support of weather forecasting, in *U.S. Contribution to the Polar Experiment (Polex)*,

Part I: Polex-Garp (North), pp. 26–49, National Academy of Sciences, Washington, D. C., 1974.

Knapp, W. W., Satellite observations of large polynyas in polar waters, in *Sea Ice: Proceedings of an International Conference, Reykjavik, Iceland, May 10–13, 1971*, edited by T. Karlsson, pp. 201–212, National Research Council of Iceland, Reykjavik, 1972.

Laevastu, T., Factors affecting the temperature of the surface layer of the sea, *Comment. Phys. Math., 25*, 1, 1960.

Langleben, M. P., The decay of an annual cover of sea ice, *J. Glaciol., 11*, 337–344, 1972.

MacCracken, M. C., and F. M. Luther, Climate studies using a zonal atmospheric model, in *Proceedings of the International Conference on Structure, Composition and General Circulation of the Upper and Lower Atmospheres and Possible Anthropogenic Perturbations, Melbourne, Australia, January 14–25, 1974*, vol. II, edited by W. L. Godson, pp. 1107–1128, University of Melbourne, Melbourne, Australia, 1974.

Manabe, S., K. Bryan, and M. J. Spelman, A global ocean-atmosphere climate model, I. The atmospheric circulation, *J. Phys. Oceanogr.*, 5, 3–29, 1975.

Marshunova, M. S., Principal characteristics of the radiation balance of the underlying surface and of the atmosphere in the Arctic, in *Soviet Data on the Arctic Heat Budget and Its Climatic Influence*, edited by J. O. Fletcher, B. Keller and S. M. Olenicoff, pp. 51–131, Rand Corporation, Santa Monica, Calif., 1966.

Maykut, G. A., Energy exchange over young sea ice in the central Arctic, *Aidjex Bull., 31*, 45–74, 1976.

Maykut, G. A., Estimates of the regional heat and mass balance of the ice cover, in *A Symposium on Sea Ice Processes and Models, September 6–9, 1977*, vol. 1, pp. 65–74, University of Washington, Seattle, 1977.

Maykut, G. A., and N. Untersteiner, Numerical prediction of the thermodynamic response of arctic sea ice to environmental changes, *Memo. RM-6092-PR*, 173 pp., Rand Corp., Santa Monica, Calif., 1969.

Maykut, G. A., and N. Untersteiner, Some results from a time-dependent thermodynamic model of sea ice, *J. Geophys. Res., 76*, 1550–1575, 1971.

McClain, E. P., Monitoring earth surface characteristics important to weather and climate with earth satellites, in *Proceedings of the Symposium on Meteorological Observations From Space: Their Contribution to the First Garp Global Experiment*, pp. 129–135, Committee on Space Research, International Council of Scientific Unions, Paris, 1976.

McPhee, M. G., An analysis of pack ice drift in summer, in *A Symposium on Sea Ice Processes and Models, September 6–9, 1977*, vol. 1, pp. 240–250, University of Washington, Seattle, 1977.

Muench, R. D., and K. Ahlnas, Ice movement and distribution in the Bering Sea from March to June 1974, *J. Geophys. Res., 81*, 4467–4476, 1976.

Murray, F. W., On the computation of saturation vapor pressure, *J. Appl. Meteorol., 6*, 203–204, 1967.

Nansen, F., The oceanography of the north polar basin, in *Norwegian North Pole Expedition, 1893–1896, Scientific Results*, vol. 3, pp. 357–386, Longmans, Green, Toronto, Ont., 1902.

Neumann, G., and W. J. Pierson, Jr., *Principles of Physical Oceanography*, 545 pp., Prentice-Hall, Englewood Cliffs, N. J., 1966.

Nikiforov, E. G., Z. M. Gudkovich, Yu. N. Yefimov, and M. A. Romanov, Principles of a method for calculating the ice redistribution under the influence of wind during the navigation period in arctic seas, *Aidjex Bull., 3*, 40–64, 1970.

Oliger, J. F., R. E. Weilick, A. Kasahara, and W. M. Washington, Description of NCAR global circulation model, *NCAR-TN-56+STR*, 94 pp., Nat. Center for Atmos. Res., Boulder, Colo., 1970.

Parkinson, C. L., A numerical simulation of the annual cycle of sea ice in the Arctic and Antarctic, Coop. thesis 46, 191 pp., Nat. Center for Atmos. Res., Boulder, Colo., 1978 (Also *Accession PB279789* Nat. Tech. Inform. Serv., Springfield, Va.)

Pease, C. H., A model for the seasonal ablation and accretion of antarctic sea ice, *Aidjex Bull., 29*, 151–172, 1975.

Pritchard, R. S., An elastic-plastic constitutive law for sea ice, *J. Appl. Mech., 42*, 379–384, 1975.

Pritchard, R. S., An estimate of the strength of arctic pack ice, *Aidjex Bull., 34*, 94–113, 1976.

Pritchard, R. S., M. D. Coon, and M. G. McPhee, Simulation of sea ice dynamics during Aidjex, *Aidjex Bull., 34*, 73–93, 1976.

Reed, R. J., and W. J. Campbell, *Theory and Observations of the Drift of Ice Station Alpha*, 255 pp., Department of Meteorology, University of Washington, Seattle, 1960.

Reed, R. J., and W. J. Campbell, The equilibrium drift of ice station Alpha, *J. Geophys. Res., 67*, 281–297, 1962.

Rossby, C. G., and R. B. Montgomery, The layer of frictional influence in wind and ocean currents, *Pap. Phys. Oceanogr. Meteorol., 3*(3), 101, 1935.

Rothrock, D. A., The steady drift of an incompressible arctic ice cover, *Aidjex Bull., 21*, 49–77, 1973.

Ruzin, M. I., The wind drift of ice in a heterogeneous pressure field, *Arkt. Antarkt. Nauch. Issled. Inst. Tr., 226*, 123–135, 1959.

Sellers, W. D., *Physical Climatology*, 272 pp., University of Chicago Press, Chicago, Ill., 1965.

Semtner, A. J., Jr., A model for the thermodynamic growth of sea ice in numerical investigations of climate, *J. Phys. Oceanogr., 5*, 379–389, 1976*a*.

Semtner, A. J., Jr., Numerical simulation of Arctic Ocean Circulation, *J. Phys. Oceanogr., 6*, 409–425, 1976*b*.

Shuleikin, V. V., The drift of ice-fields, *C. R. Acad. Sci. USSR, 19*, 589–594, 1938.

Smith, J. D., and T. J. English, Seasonal variations in density and depth of the mixed layer in the Arctic Ocean (abstract), *Eos Trans. AGU, 55*, 76, 1974.

Sverdrup, H. U., The wind-drift of the ice on the North Siberian Shelf, in *Norwegian North Polar Expedition With the 'Maud,' 1918–1925 Scientific Results*, vol. 4, no. 1, 46 pp., Griegs Boktrykkeri, Bergen, Norway, 1928.

Sverdrup, H. U., M. W. Johnson, and R. H. Fleming, *The Oceans, Their Physics, Chemistry and General Biology*, 1087 pp., Prentice-Hall, Englewood Cliffs, N. J., 1961.

Swithinbank, C., Arctic pack ice from below, in *Sea Ice: Proceedings of an International Conference, Reykjavik, Iceland, May 10–13, 1971*, pp. 246–254, edited by T. Karlsson, National Research Council of Iceland, Reykjavik, 1972.

Taljaard, J. J., H. van Loon, H. L. Crutcher, and R. L. Jenne, Climate of the upper air, I, Southern hemisphere, vol. 1, Temperatures, Dew Points, and Heights at Selected Pressure Levels, *NAVAIR Rep 50-1C-55*, 135 pp., U.S. Navy Weather Serv. Command, Washington, D. C., 1969.

Thekaekara, M. P., and A. J. Drummond, Standard values for the solar constant and its spectral components, *Nature Phys. Sci., 229*, 6–9, 1971.

Tolstikov, Ye. I. (Ed.), *Atlas of Antarctica*, vol. I, USSR Academy of Sciences, Moscow, 1966.

Tolstikov, Ye. I. (Ed.), *Atlas of Antarctica*, vol. II, USSR Academy of Sciences, Moscow, 1969.

Untersteiner, N., Calculations of temperature regime and heat budget of sea ice in the central Arctic, *J. Geophys. Res., 69*, 4755–4766, 1964.

van Loon, H., Cloudiness and precipitation in the southern hemisphere, in *Meteorological Monographs*, vol. 13, *Meteorology of the Southern Hemisphere*, edited by C. W. Newton, pp. 101–111, American Meteorological Society, Boston, Mass., 1972.

Vowinckel, E., Some estimates about the size and variation of the antarctic pack ice belt, *Rep. 19*, 20 pp., Inst. of Polar Stud., Ohio State Univ., Columbus, 1966.

Washington, W. M., A. J. Semtner, Jr., C. Parkinson, and L. Morrison, On the development of a seasonal change sea-ice model, *J. Phys. Oceanogr., 5*, 679–685, 1976.

Weeks, W. F., Sea ice conditions in the Arctic, *Aidjex Bull., 34*, 173–205, 1976.

Welander, P., Thermal oscillations in a fluid heated from below and cooled to freezing from above, *Dyn. Atmos. Oceans, 1*, 215–223, 1977.

Winchester, J. W., and C. C. Bates, Meteorological conditions and the associated sea ice distribution in the Chukchi Sea during the summer of 1955, *Polar Atmosphere Symposium, Part I, Meteorology Section*, edited by R. E. Sutcliffe, pp. 323–334, Pergamon, New York, 1958.

Wittmann, W. I., and J. J. Schule, Jr., Comments on the mass budget of arctic pack ice, in *Proceedings of the Symposium on the Arctic Heat Budget and Atmospheric Circulation*, edited by J. O. Fletcher, pp. 215–246, Rand Corporation, Santa Monica, Calif., 1966.

Zillman, J. W., A study of some aspects of the radiation and heat budgets of the southern hemisphere oceans, *Meteorol. Stud., 25*, 562

pp., Bur. of Meteorol., Dep. of the Interior, Canberra, Australia, 1972.

Zwally, H. J., and P. Gloersen, Passive microwave images of the polar regions and research applications, *Polar Rec., 18*, 431–450, 1977.

Zwally, H. J., T. T. Wilheit, P. Gloersen, and J. L. Mueller, Characteristics of antarctic sea ice as determined by satellite-borne microwave imagers, in *Proceedings of the Symposium on Meteorological Observations From Space: Their Contribution to the First Garp Global Experiment*, pp. 94–97, Committee on Space Research, International Council of Scientific Unions, Paris, 1976.

(Received February 6, 1978;
revised June 9, 1978,
accepted June 15, 1978.)

Paper 92

Reprinted from MONTHLY WEATHER REVIEW, Vol. 107, No. 11, November 1979
American Meteorological Society
Printed in U. S. A.

Assessment of the Urban Heat Island Effect Through the Use of Satellite Data

JOHN C. PRICE

Laboratory for Atmospheric Sciences (GLAS), NASA/Goddard Space Flight Center, Greenbelt, MD 20771

16 February 1979 and 22 June 1979

ABSTRACT

A recent NASA satellite is obtaining high spatial resolution thermal infrared data at times of day appropriate for the study of the urban heat island effect. Quantitative estimates of the extent and intensity of urban surface heating are obtained by analysis of digital data acquired over the New York City–New England area. In many large cities satellite sensed temperatures are 10–15°C warmer than in surrounding rural areas. A thorough interpretation of the elevated urban surface temperature will require studies of 1) the relationship between remotely sensed surface temperatures and air temperatures, and 2) compensation for observed very localized heating due to industry and/or power plants.

1. Introduction

Interest in urban meteorology has grown rapidly in recent years, prompted by concern for the effects of air pollution on human health, federal requirements for air quality, uncertainty as to the environmental effects of energy consumption, and a general desire to improve the quality of life in urban areas. The gradual dispersion of population into suburban areas and the tendency toward movement of manufacturing and industry into sparsely populated areas makes the topic of significance to a growing fraction of the country.

For meteorological purposes the heat island effect is characterized by both direct (*in situ*) and indirect (circulation pattern) variations when compared to conditions in the surrounding countryside. The effect is directly measurable as a relative increase in air temperature within and above the city. This tendency is not universal—in some cases the air above a city may be cooler than that in surrounding rural areas. In most large cities the increase in air temperature is accompanied by elevated concentrations of CO_2, trace gases and atmospheric particulates. For very large cities these changes are observable to heights of hundreds of meters (Bornstein, 1968; McCormick and Baulch, 1962).

An indirect and more subtle meteorological effect is manifest through altered flow patterns and vertical convection in urban areas. Presumably, this modification is the result of air temperature changes associated with urbanization but the change in surface roughness produced by homes and buildings may also have an effect through modification of vertical transport mechanisms at low level. Changes in flow patterns pose possible implications for local cloudiness, for dispersion of the urban air mass and for rainfall, both in the city proper and in its surroundings (Duckworth and Sandberg, 1975; Huff and Vogel, 1978).

A major limiting factor in the study of the heat island effect has been the lack of observational data which clearly are relevant to the problem. Particularly limiting are the inability to describe quantitatively the areal extent and distribution of changes of air temperature, and the general paucity of data obtained at varying locations and under varying atmospheric conditions, including the annual cycle.

The Heat Capacity Mapping Mission [HCMM (Goddard Space Flight Center, 1978)], a small Applications Explorer satellite launched 26 April 1978, has obtained data suitable for heat island studies. The satellite acquires high-resolution thermal infrared (10.5–12.5 μm) data from an orbit especially selected to infer surface temperatures near the time of day of the diurnal maximum. Although surface temperatures estimated from a radiation measurement are not the same as air temperatures, they are potentially as useful because they provide an indication of the surface heating effect

FIG. 1. Satellite overpass of the New York–New England area showing surface heating effect (light areas) in urban areas.

at the base of the atmosphere. The satellite capability is discussed in the succeeding sections.

2. Data characteristics

From an altitude of 620 km the HCMM satellite acquires visible and infrared data at a spatial resolution of 500 m. Although data transmission to ground stations is in analog form, a high-performance design has demonstrated the capability to deliver temperature data with a noise equivalent temperature of 0.4 K. (Bohse *et al.*, 1979.) The remote measurements represent an equivalent blackbody temperature as modified by absorption and reemission during passage through the intervening atmosphere, i.e., a brightness temperature given by (radiance/Stefan Boltmann constant)$^{1/4}$. Under most conditions the variation of this atmospheric correction may be neglected over small areas such as that of a city and its surroundings. The development of a relationship between surface temperatures, as obtained by remote observation, and a local air temperature (e.g., at 2 m height) is a complex problem. The relatively poor spatial resolution of early meteorological satellite data (Rao, 1972) has tended to hinder development of such a relationship. More recent studies (Carlson *et al.*, 1977; Matson, *et al.*, 1978) have utilized 1 km spatial resolution satellite data acquired during mid morning and early evening when surface temperature contrasts are relatively weak. High-resolution HCMM data should prove useful for placement of air monitoring stations in cities and for spatial interpolation between such sites.

The satellite passes over northern midlatitudes in the early afternoon and at approximately 0200 LT. A portion of an overpass at 1330 LT on 6 June 1978 is displayed in Fig. 1. Standard image products for the mission are displayed with the convention cold–black, hot–white. This is contrary to the convention used for meteorological satellite imagery and results in cold clouds being displayed as dark, or black. In Fig. 1 dark patches in the lower left and speckles in the upper left center correspond to clouds.

Man's influence on the earth's surface temperature is strikingly evident in the picture. The bright areas represent regions of significant surface heating. New York City and the entire metropolitan area stand out clearly against the darker (cooler) waters of the Atlantic Ocean. Central Park is faintly visible as a dark strip running down the

ORIGINAL PAGE IS
OF POOR QUALITY

center of Manhattan Island. The large cities along the Connecticut shoreline and the Hudson and Connecticut Rivers are easily recognized. Smaller cities and even roads are also apparent, as well as many features relating to surface geology and hydrological conditions.

Weather at the time of the observation was clear and brisk following passages of a weak cold front on the preceding afternoon (5 June). Through New England and eastern New York state air temperature at 1800 GMT (1300 EST) ranged from 20 to 25°C, with winds generally from the northwest at 2–5 m s^{-1} except along the coast. The air temperatures agree generally with satellite derived brightness temperatures for the rural areas in the scene. However, a detailed comparison is not feasible at this time because of the spatial variability of satellite derived temperatures the need for a correction for atmospheric moisture and, principally, the uncertainty as to the relationship between air temperature and surface temperature.

3. Magnitude of the surface heat effect

Through analysis of the digital data represented by Fig. 1 it is possible to establish the extent and magnitude of surface heating in a number of cities compared with their surroundings. The New York metropolitan area is so large that it is difficult to establish a precise background level for comparison with values in the city center. In other cases the determination of urban–suburban–country thresholds is reasonably objective, requiring comparison with a map at a scale appropriate to the imagery, e.g., 1:1 000 000. The temperature contrast of a number of cities and towns with their surrounding rural areas is given in Table 1. The contrast is based on brightness blackbody temperature as observed by the satellite radiometer.

As one would expect, the values for temperature contrast are in rough agreement with the corresponding population figures. However, the temperature excesses in large cities are much higher than values typical of air temperatures (which is usually of the order of 2–4°C). In many cases the differences exceed 10°C. Most numerical studies (Myrup, 1969) have treated a city as a rough surface which interacts with the atmosphere in a manner similar to vegetation, rough soil, etc. The high values of surface temperature contrast suggest that a more complex description may be necessary. The highest values of temperature contrast (up to 15°C) in Table 1 are not necessarily representative of the heat island effect as specified by air temperature. Within some of the larger cities a few (1–5) isolated points indicate temperatures 2–3°C above all other values. These elevated temperatures are presumed to be associated with power plants or industrial activity. More detailed study, specifically ground truth verification of accurately mapped HCMM data, will be required in order to verify this hypothesis. It is apparent, however, that the relative importance of industry as compared to urban population need not be the same for air temperature as for remotely sensed surface temperature.

Even after discounting the effect of industry the excess of urban surface temperatures over these surrounding areas is still much greater than that of air temperatures. This effect may be due to trapping of energy within the "urban canyon" (Nunez and Oke, 1977). The vertical structure in city centers provides a potential for absorption of radiation and for heat storage at a level beneath that of strong vertical

TABLE 1. Population (1970) and observed peak temperature contrast of cities and surrounding areas.

City	Population (×1000)	Metropolitan area population (×1000)	Temperature contrast (°C)
New York, NY		7895	17.0
Hartford, CT	158	817	15.0
Schenectady, N.Y.	78		15.0
Providence, RI	179	720	13.2
Binghamton, NY	64		12.4
Bridgeport, CT	157		12.1
Worcester, MA	177	637	11.5
New Haven, CT	138	745	11.2
Stanford, CT	109		11.2
Fitchburg, MA	43		11.2
Syracuse, NY	197	637	10.9
Waterbury, CT	108		10.9
Albany, NY	116		10.3
Troy, NY	62		10.3
Pittsfield, MA	57		9.7
Rome, NY	50		8.9
Utica, NY	91		8.0
Barre, VT	10		8.0
Montpelier, VT	8.6		6.8

TABLE 2. Area and excess radiated power from urban areas (9°C threshold).

City	Area (km^2)	Radiated power (kW)
New York City area	547	40 000
Providence	48.2	3190
Hartford	26.2	1770
Schenectady	5.3	368
Bridgeport	14.2	954
Syracuse	6.5	417
Binghamton	6.0	394
New Haven	5.8	372
Worcester	5.1	327
Albany	3.*	227
Stanford	3.2	202
Fitchburg	2.1	131
Waterbury	1.2	74
Troy	1.4	86
Pittsfield	0.*	42
Rome	0.5	27

mixing in the atmosphere. This hypothesis can be tested by comparison with appropriately taken air temperature measurements. The urban canyon effect would tend to produce air temperature maxima in regions of highest vertical relief, by inhibiting exchange of radiant energy and thermal energy with the atmosphere.

A second but less likely explanation emphasizes rooftop heating by the sun at a height above the level of standard air temperature measurements. This mechanism also implies high radiation temperatures, as observed by the satellite, but lower air temperature values near ground level. However, one would expect a strong mixing interaction to extract energy efficiently from building tops and transfer it to the atmosphere below.

For investigation of urban heating, the peak temperature is less significant than a summation (area integral) of the excess power radiated as a result of the surface temperature elevation. In Table 2 the area and integrated power are listed, where an arbitrary threshold of 9°C excess has been used to specify the urban area. (At 300 K 1 K temperature increase results in an increase in radiated power of 6 W m^{-2}). It is interesting to note that Torrance and Shum (1976) estimate the typical energy consumption rate in large cities to be in the range 20–100 W m^{-2}, which is of the order of the excess energy radiated to space in city centers due to the elevation of urban surface temperature over that of the surrounding countryside. Of course, the complete energy budget is much more complex, as most heat generated by automobiles, space heating and industry enters the environment as atmospheric warming near ground level. However, by modifying the temperature structure of the lower atmosphere this exhaust heat tends to reduce the vertical temperature gradient and hence increase the surface temperature in urban areas.

Although the mechanism for the satellite-observed large elevation of urban surface temperature is unknown, the appearance in the imagery of very small towns in New England points to the great sensitivity of remote sensing for monitoring surface temperature anomalies associated with human habitation.

5. Conclusion

The availability of high spatial resolution data in the thermal infrared opens a new avenue to the study of the urban heat island effect. Although many details must be explored for a full understanding of the significance of the satellite data, imagery such as Fig. 1 illustrates the great potential utility of the data. The more recent launch of a NOAA satellite, TIROS N, having characteristics very similar to the HCMM radiometer with 1 km spatial resolution, 1430 LT overpass at midlatitudes promises a continuing capability for study of the heat island effect.

Acknowledgment. I thank Ms. Tina Lien for assistance with the image processing.

REFERENCES

Bohse, J. R., M. Bewtra and W. L. Barnes, 1979: Heat capacity mapping radiometer data processing algorithm, calibration and flight performance evaluation. Goddard Space Flight Center, Greenbelt, M.D., 162 pp.

Bornstein, R. D., 1968: Observations of the urban heat island effect in New York City. *J. Appl. Meteor.*, **7**, 575–582.

Carlson, T. N., J. N. Augustine and F. E. Boland, 1977: Potential application of satellite temperature measurements in the analysis of land use over urban areas. *Bull. Amer. Meteor. Soc.*, **58**, 1301–1303.

Duckworth, F. S., and J. S. Sandberg, 1954: The effect of cities upon horizontal and vertical temperature gradients. *Bull. Amer. Meteor. Soc.*, **35**, 198–207.

Heat Capacity Mapping Mission User's Guide, 1978: Goddard Space Flight Center, Greenbelt, MD, 120 pp.

Huff, F. A., and J. L. Vogel, 1978: Urban, topographic and diurnal effects on rainfall in the St. Louis region. *J. Appl. Meteor.*, **17**, 565–577.

Matson, M., E. P. McClain, D. F. McGinnis, Jr. and J. A. Pritchard, 1978: Satellite detection of urban heat islands. *Mon. Wea. Rev.*, **106**, 1725–1734.

McCormick, R. A., and D. M. Baulch, 1962: The variation with height of the dust loading over a city, as determined from atmospheric turbidity. *J. Air Pollut. Control*, **12**, 492–496.

Myrup, L. O., 1969: A numerical model of the urban heat island. *J. Appl. Meteor.*, **8**, 908–918.

Nunez, M., and T. R. Oke, 1977: The energy balance of an urban canyon. *J. Appl. Meteor.*, **16**, 11–19.

Rao, P. K., 1972: Remote sensing of urban heat islands from an environmental satellite. *Bull. Amer. Meteor. Soc.*, **53**, 647–648.

Torrance, K. E., and J. S. W. Shum, 1976: Time varying energy consumption as a factor in urban climate. *Atmos. Environ.*, **10**, 329–337.

Paper 93

Reprinted from *Reviews of Geophysics and Space Physics*, Vol. 17, No. 6, pp. 1262-1264, 1282-1288, published in Sept. 1979 by the American Geophysical Union.

RESEARCH ON SNOW AND ICE

REMOTE SENSING OF SNOW AND ICE

A. Rango

Goddard Space Flight Center, Greenbelt, Md.

Remote sensing has become a valuable tool in snow and ice studies because of its unique capability for acquiring measurements of glaciological conditions over large areas [Rango, 1977]. Two major objectives are to develop techniques for improved monitoring of existing conditions and to incorporate the new data into various forecasting or management systems.

Since 1974, various investigators have gained experience with and developed techniques for snowcover interpretation from visible and infrared data from satellites such as Landsat and NOAA. Numerous techniques are available for analyzing the data ranging from simple photointerpretation to automated digital methods [Schneider, et al., 1976; Meier and Evans, 1975; Rango and Itten, 1976; Barnes and Smallwood, 1975; Dallam and Foster, 1975; Katibah, 1975; Luther, et al., 1975; and Algazi and Suk, 1975]. In addition, the use of Skylab data for snow-mapping has been investigated [Barnes, et al., 1975; 1977] with most significant results pertaining to the use of the 1.55μm - 1.75μm band for discriminating clouds from snow [Bartolucci, et al., 1975; Barnes and Bowley, 1977; and Valovcin, 1976].

Radiative transfer modeling has been used to calculate the solar reflectance of snow and to estimate the effect of snow aging on reflectance [Choudhury and Chang, 1978a, 1978b]. O'Brien and Munis [1975a] have examined the reflectance of snow at discrete visible and near infrared wavelengths with emphasis on effects due to aging and melting

As a result of the extensive experience with satellite snowcover data several applications have developed. NOAA has established techniques for using satellite imagery to detect, measure, and map mean monthly snowcover over the Northern Hemisphere [Wiesnet and Matson, 1975; and Matson, 1977]. Regression analysis using nine years of data yielded several equations with correlation coefficients significant enough to have possible applications for 30, 60, and 90-day forecasting of seasonal hemispheric, and continental snowcover [Wiesnet and Matson, 1976].

Rango, et al., [1977b] used meteorological satellite snow extent data to derive a regression relationship between early April snowcovered area and April-June seasonal yield on the Indus River in Pakistan. In these large data-sparse regions the satellite snowcover data period of record actually exceeds the conventional data base.

Thompson [1975] in Wyoming found that the snowcovered area on a particular date was better related to the accumulated runoff/total seasonal runoff ratio than to just the seasonal runoff in a statistically significant expression. A long-term data base was obtained by compositing aircraft and Landsat snowcover data with resulting analysis indicating that snow extent was useful in reducing seasonal runoff forecast error when incorporated into procedures to update water supply forecasts in California on a 15-day basis as the melt season progressed [Rango, et al., 1977a]. Several runoff models including the Streamflow Synthesis and Reservoir Regulation (SSARR) model [Speers, et al., 1979] have options permitting the use of snowcover input data and variable elevation zones for calculating snowmelt. In addition, several hydrologic models, although not originally requiring snowcover input, have been modified to accept satellite snow extent data for the generation of daily discharge values [Leaf, 1975; and Hannaford, 1977].

Based on promising results in snow mapping, seven federal and three state agencies have con-

ducted a program to test the usefulness of the satellite data in operational snowmelt-runoff forecasting [Rango, 1978]. Both empirical and modeling approaches have been evaluated in connection with the satellite snowcover data in four regions of the western U.S., namely, Arizona [Schumann, 1975; Warskow, et al., 1975; and Kirdar, et al., 1977], California [Brown and Hannaford, 1975], Colorado [Washicheck and Mikesell, 1975], and the northwest states [Limpert, 1975]. Preliminary results were documented [Rango, 1975c] and a final workshop will conclude the project in April 1979.

Investigations into the use of remote sensing techniques for the measurement of more fundamental snow properties are also being carried out. The use of visible and near infrared wavelengths for inferring snow properties such as depth [McGinnis, et al., 1975a], water equivalent [Sharp and Thomas, 1975; Merry, et al., 1977], and density [McMillan and Smith, 1975] has been tested with only limited success because of the restriction to sensing of surface characteristics at these wave lengths. Development of gamma ray techniques for measurement of snow water equivalent has also continued [Bissell, 1975a; Fritzsche and Feimster, 1975].

Microwave monitoring of snowpack properties has received considerable attention because this portion of the electromagnetic spectrum has the capability for penetrating snow allowing for inference of internal characteristics. Satellite snowcovered area measurements have been made using Nimbus 5 [Kunzi and Staelin, 1975; Kunzi, et al., 1976] and Nimbus 6 [Rango, et al., 1979] radiometers. Several other passive microwave studies have covered the modeling of microwave emission from snow [Chang and Gloersen, 1975; Chang, et al., 1976; Zwally, 1977], comparison of model calculations and satellite observed brightness temperatures from polar firn [Chang and Choudhury, 1978; and Chang, et al., 1978], development of a method to determine snowfield temperature profile and mean crystal size by using multifrequency microwave radiometer measurements [Chang, 1978], and correlation of microwave emission to water equivalent, depth, and free water content [R.T. Hall, et al., 1978 and Shiue, et al., 1978]. For dry snow conditions on the high plains significant relationships between snow depth or water equivalent and microwave brightness temperature were developed [Rango, et al., 1979]. Associated active microwave and snow studies have included modeling of the electromagnetic reflection from snow [Linlor and Jiracek, 1975; and Linlor, 1976] and analysis of data from experimental ground-based measurement programs [Angelakos, 1977; Ellerbruch, et al., 1977; Linlor, et al., 1977; Angelakos, 1978; Ulaby, 1976; Ulaby and Stiles, 1977; and Ulaby, et al., 1978].

Landsat images have proved to be very useful for collecting certain basic data from glaciers, for example, long term surface velocities are readily determined by comparison of displacement on images taken at different times [Krimmel and Meier, 1975]. Surging glaciers are easily identified and their associated short term high flow rates have been measured in various locations [Krimmel, et al., 1976; Post, et al., 1976; and Meier, 1976]. Large glaciers and icecaps have also been monitored [Williams, 1976]. The location of the snowline on a glacier can easily be mapped with Landsat and when observed at the end of the melt season can be related to the annual net mass balance [Braslau and Bussom, 1979]. Under direction of the U.S. Geological Survey, Landsat is being used currently to compile a worldwide glacier atlas.

Formation and dissipation of river ice on the Ottawa River was monitored daily using visible imagery from NOAA satellites. The break-up of 14 ice-covered reaches was observed during the melt period in April 1976 [McGinnis and Schneider, 1978a]. Radar monitoring of river ice provides an all weather capability during cloudy periods and was tested on the St. Lawrence River. A contour map showing the accumulation pattern of frazil and brash ice was obtained [Dean, 1977].

Visible and near infrared observations over the Great Lakes have been used with reasonable success for mapping ice cover and type [Sydor, 1976; and McMillan and Forsyth, 1976]. However, for operational purposes cloud cover is a significant problem, and as a result, Side-Looking Airborne Radar (SLAR) has been used successfully for ice monitoring purposes [Schertler, et al., 1975]. Additional studies were conducted [Bryan and Larson, 1975] and properties of frozen northern lakes [Weeks, et al., 1977].

Radar studies on Alaskan lakes indicate that discrimination between lakes frozen completely to the bottom versus lakes with fresh water beneath the ice is possible, thus providing additional information on lake depth [Sellman, et al., 1975; and Elachi, et al., 1976]. The measurement of lake ice thickness was accomplished using a short-pulse radar system which can be ground-based or airborne [Cooper, et al., 1976]. Monitoring of lake ice using passive microwaves has also been investigated [Bryan and Hall, 1976], and it appears that ice thickness variations can be distinguished [D.K. Hall and Bryan, 1977; and D.K. Hall, et al., 1978].

The use of remote sensing in sea ice studies is directed toward answering fundamental questions regarding amount of ice covered ocean, ice movement, and ice formation and ablation. Data generated by remote sensing is used in determining the influence of sea ice on atmospheric and oceanic processes. It seems fortunate that the current increase in scientific interest about sea ice coincides with a time of rapid evolution of both remote sensing platforms and sensors. Campbell, et al., [1975] present an overview of mesoscale and macroscale studies of floating ice in three sensor categories: visual, passive microwave, and active microwave.

Using visible satellite imagery the primary advances have been in tracking ice floe movement using sequential imagery [Campbell, 1976a; Hibler, et al., 1975; Shapiro and Burns, 1975; and Campbell, 1977], ice lead and polynya dynamics [Campbell, 1976b], seasonal sea ice metamorphosis [Campbell, 1976c], and dynamics of ice shear zones [Campbell, 1976d]. A statistical method for discriminating sea ice from clouds with 90 percent or greater accuracy has been developed [Gerson and Rosenfeld, 1975]. Visible

imagery has been used to compile statistics on ice conditions for applications such as off-shore oil and gas exploration [Barnes, et al., 1977]. In addition visible and thermal imagery have been used for estimations of sea ice thickness, although the presence of snowcover may cause a limitation to relative amounts only [Poulin, 1975; R.T. Hall, 1975; Kuhn, et al., 1975; LeSchack, 1975; and Rothrock, 1975].

The advantage of microwave observations of sea ice rests in the capabilities to penetrate clouds and to make observations during the polar night. In the Beaufort Sea five ice zones were discriminated using aircraft multispectral passive microwave observations [Campbell, et al., 1976]. Brightness temperature levels and relative fluctuations are used to distinguish between shorefast sea ice, shear zone, mixed first-year and multi-year sea ice, mixed first-year ice and medium to large multi-year floes, and the polar ice zone. The radiometric signatures are most pronounced at 0.8 and 1.5 cm wavelength [Campbell, et al., 1976]. The time variation of sea ice concentration and multi-year ice fraction within pack ice in the Arctic Basin was examined using the 1.55 cm radiometer on the Nimbus 5 spacecraft [Gloersen, et al., 1978] with significant variations between seasons being observed. Previously unobserved areas, several hundred kilometers in extent, of sea ice concentration as low as 50 percent were discovered deep in the interior of the Arctic polar sea ice pack. Sea ice observations by Nimbus 3 in the polar regions are reviewed by Zwally and Gloersen [1977].

Active microwave studies of sea ice have shown the capability of synthetic aperature radar for displaying the orientation of leads in the ice and the percentage of open water in the entire sea ice scene [Bryan, et al., 1977]. Experiments with radar scatterometers indicate that various types of sea ice categories can be distinguished with about an 87 percent correct identification accuracy at the 13.3 GH_z frequency [Parashar, et al., 1977]. The profiling of sea ice thickness has also been attempted using an impulse radar [Morey, 1975]. A large variety of active and passive microwave measurements of sea ice were made as part of the Main Arctic Ice Dynamics Joint Experiment 1975-1976 [Campbell, et al., 1978].

Abrams, G., and A.T. Edgerton, Snow parameters from Nimbus-6 electrically scanned microwave radiometer, Aerojet Electrosystems Co., Azusa, Calif, 195 pp. NTIS: N78-21572, 1977.

Adams, W.P., Areal differentiation of snow cover in east central Ontario, Water Resources Research, 12(6), 1226-1234, 1976.

Ahlnas, K., and G. Wendler, Arctic sea-ice conditions in early spring viewed by satellite, Arctic and Alpine Research 9(1), 61072, 1977.

Alexander, L., L. Eichen, F.D. Haselden, R.F. Pascucci, and D.M. Ross-Brown, Applications of remote-sensing technology, Spaceworld M-5-149, 15-39, 1976.

Algazi, Ralph V. and Minsoo Suk, An all digital approach to snow areal mapping and snow modeling, in Operational applications of satellite snowcover observations, NASA, Wash., D.C., SP-391/18, 249-257, 1975.

Allison, L.J., R. Wexler, C.R. Laughlin, W.R. Bandeen, Remote sensing of the atmosphere from environmental satellites, NASA, Goddard Space Flight Center, Greenbelt, Md., 124 pp., 1977.

Amstutz, David Elam, Stereophotogrammetric reconnaissance of ocean wave/sea ice interaction, Ph.D. dissertation, Oregon State Univ, 172 pp., 1977.

Andrews, J.T., P.T. Davis and C. Wright, Little ice age permanent snow cover in the eastern Canadian Arctic: extent mapped from LANDSAT-1 satellite imagery, Inst. of Arctic and Alpine Res., Univ. of CO., Geografiska Annaler, Ser. A, Physical Geography, Stockholm, 58(1/2), 71-81, 1976.

Angelakos, D.J., Microwave scattering properties of snow fields, Univ. California, Berkeley, 6 pp. NTIS: N78-19577, 1977.

Angelakos, D.J., Research of microwave scattering properties of snow fields, Univ. California, Berkeley, 32 pp. NTIS: N78-29303, 1978.

Apel, John R., SEASAT: a spacecraft views the marine environment with microwave sensors, in T. Nejat Veziroglu (ed.), Remote Sensing Energy-Related Studies, Wash., D.C. Hemisphere Publ. Corp., 47-60, 1975.

Apinis, J.J., and W.H. Peake, Passive microwave mapping of ice thickness, Final Report, Ph.D. thesis, Ohio State University, Columbus, 156, 1976.

Aul, J.S., and P.F. Ffolliott, Use of areal snow cover measurements from ERTS-1 imagery in snowmelt-runoff relationships in Arizona, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/8, 1975.

Barnes, James C., and Michael D. Smallwood, Synopsis of current satellite snow mapping techniques, with emphasis on the application of near-infrared data, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/15, 199-123, 1975.

Barnes, J.C., M.D. Smallwood, and J.F. Cogan, Study to develop improved spacecraft snow survey methods using Skylab/EREP data, ERT Document No. 412-F, Final Report to NASA, Contract No. NAS 9-12305, 92 pp., 1975.

Barnes, James C., and Clinton J. Bowley, Study of near-infrared snow reflectance using Skylab S192 multispectral scanner data, Environmental Res. and Technology, Inc., Concord, Mass., National Environmental Satellite Service, Wash. D.C., 54 pp. NTIS: PB-267 504, 1977.

Barnes, J.C., C.J. Bowley, J.T. Parr, and M.D. Smallwood, Snow mapping experiment, in Skylab Explores the Earth, NASA SP-380, pp. 191-224, 1977.

Barnes, J.C., C.J. Bowley, M.D. Smallwood, and J.H. Willand, Use of satellite data to evaluate surface ice conditions for off-shore oil and gas exploration, International Conference on Port and Ocean Engineering under Arctic Conditions, 4th, Proceedings, St. Johns, Newfoundland, pp. 1019-1034, 1978.

Barnes, P.W. and E. Reimnitz, Flooding of sea ice by the rivers of northern Alaska, U.S. Geological Survey, Prof. Paper 929, 356-359, 1976.

Bartolucci, L.A., R.M. Hoffer, and S.G. Luther, Snowcover mapping by machine processing of SKYLAB and LANDSAT MSS data, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/21, Wash., D.C., 295-311, 1975.

Bentley, Charles R., Advances in geophysical exploration of ice sheets and glaciers, J. of Glaciol., 15(73), 113-135, 1975.

Biggs, A.W., Volume scattering from ice and water in inhomogeneous terrain, in A.N. Ince (ed.), Electromagnetic Wave Propagation Involving Irregular Surfaces and Inhomogeneous Media, Papers and Discussions presented at the Electromagnetic Wave Propagation Panel Symposium held in The Hague, Netherlands, 25-29 March 1974, AGARD Conference Proceedings 144, 16-1 - 16-13, 1975.

Biggs, A.W., Remote sensing of surface properties, in A.N. Ince (ed.), Electromagnetic Wave Propagation Involving Irregular Surfaces and Inhomogeneous Media, Papers and Discussions Presented at the Electromagnetic Wave Propagation Panel Symposium held in The Hague, Netherlands, 25-29 March 1974, AGARD Conference Proceedings 144, 13-1 - 13-12, 1975.

Bissell, Vernon C., Application of bayesian decision theory to airborne gamma snow measurement, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/28, Wash., D.C., 409-420, 1975a.

Blue, M.D., Permittivity of water at millimeter wavelengths, Ga. Institute Technology, 39 pp. NTIS: N76-30911, 1976.

Braslau, D, and D.E. Bussom, Landsat sensing of glaciers with application to mass-balance and runoff, in Proc. Modeling Snow Cover Runoff, USA Cold Regions Res. and Eng. Lab., Hanover, N.H., 6 pp., 1979.

Breaker, Lawrence C. and Michael C. Mcmillan, Sierra Nevada snow melt from SMS-2, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/14, 187-197, Wash., D.C., 1975.

Brown, A.J., and J.F. Hannaford, Interpretation of snowcover from satellite imagery for use in water supply forecasts in the Sierra Nevada, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/4, Wash., D.C., 39-51, 1975.

Brown, C.E., and E.A. Halprin, Millimeter-wave backscatter from wet and snow covered terrain, Proceedings of the IEEE 1977 National Aerospace and Electronics Conference, NAECON 1977.

Bryan, M. Leonard, and R.W. Larson, Study of fresh-water lake ice using multiplexed imaging radar, J. of Glaciology 14(71), 445-457, 1975.

Bryan, M.L., and D.K. Hall, A comparative study of active and passive microwave imagery over the north slope of Alaska, Proceedings of the Association of American Geographics, Vol. 8, pp. 164-168, 1976.

Bryan, M.L., W.D. Stromberg, and T.G. Farr, Computer processing os SAR L-Band imagery, Photogramm. Eng. and Remote Sensing 43(10), 1283-1294, 1977.

Burson, Z.G., and A.E. Fritzsche, Water equivalent of snow data from airborne gamma radiation surveys: international field year for the Great Lakes, Gov. Rep. Announce., NTIS 1973 EGG-1183602P, Springfield, Va., 1975.

Campbell, W.J., Skylab floating ice experiment: U.S. Geological Survey, Tacoma, Wash., 67 pp. NTIS: E75-10161, 1975.

Campbell, W.J., W.F. Weeks, R.O. Ramseier, and P. Gloersen, Geophysical studies of floating ice by remote sensing, J. of Glaciology 15(73), 305-328, 1975.

Campbell, W.J., P. Gloersen, W.J. Webster, T.T. Wilheit and R.O. Ramseier, Beaufort Sea ice zones as delineated by microwave imagery, J. of Geophys. Res., 81(6), 1103-1110, 1976.

Campbell, W.J., Tracking ice floes by sequential ERTS imagery, U.S. Geological Survey, Prof. Pap. 929, 337-339, 1976a.

Campbell, W.J., Ice lead and polynya dynamics, U.S. Geological Survey, Prof. Pap. 929, 340-342, 1976b.

Campbell, W.J., Seasonal metamorphosis of sea ice, U.S. Geological Survey, Prof. Pap. 929, 343-345, 1976c.

Campbell, W.J., Dynamics of Arctic ice-shear zones, U.S. Geological Survey, Prof. Pap. 929, 346-349, 1976d.

Campbell W.J., Morphology of Beaufort Sea ice, U.S. Geological Survey, Prof. Pap. 929, 350-355, 1976e.

Campbell, W.J., Visual observations of floating ice from Skylab, NASA-SP-380, 353-379, Wash., D.C., 1977.

Campbell, W.J., J. Wayenberg, J.B. Ramseyer, R.O. Ramseier, M.R. Vant, R. Weaver, A. Redmond, L. Arsenault, P. Gloersen, H.J. Zwally, T.T. Wilheit, T.C. Chang, D. Hall, L. Gray, D.C. Meeks, M.L. Bryan, F.T. Barath, C. Elachi, F. Leberl, and T. Farr, Microwave remote sensing of sea ice in the AIDJEX man experiment, Boundary-Layer Meteorol(Netherlands), 13(1-4), 309-337, 1978.

Chang, A.T.C., Estimation of snow temperature and mean crystal radius from remote multispectral passive microwave measurements, NASA, Greenbelt, Md., NTIS N78-26677/2ST, 1978.

Chang, A.T.C., and B.J. Choudhury, 1978 microwave emission from polar firn, NASA Tech. Pap. 1212, Wash., D.C., 1978.

Chang, A.T.C., B.J. Choudhury, and P. Gloersen, Microwave brightness of polar firn as measured by Nimbus 5 and 6 ESMR, NASA Tech. Memo. 79662, Goddard Space Flight Center, Greenbelt, Md., 1978, 15 pp.

Chang, T.C., Microwave emission from snow and glacier ice, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Md., 1975, 31 pp. NTIS: N75-21775.

Chang, T.C. and P. Gloersen, Microwave emission from dry and wet snow, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/27, 399-407, 1975.

Chang, T.C., P. Gloersen, T. Schmugge, T.T. Wilheit and H.J. Zwally, Microwave emission from snow and glacier ice, J. of Glaciology 16(74), 23-29, 1976.

Choudhury, B.J. and A.T.C. Chang, The solar reflectance of a snow field, NASA Tech. Memo. 78085, 1978a.

Choudhury, B.J. and A.T.C. Chang, Two-stream theory of speetral reflectance of snow, NASA Technical Memorandum 79639, Goddard Space Flight Center, Greenbelt, MD, 16 pp., 1978b.

Cooper, D.W., R.A. Mueller, and R.J. Schertler, Measurement of lake ice thickness with a short-pulse radar system, NASA Technical Note TN D-8189, Lewis Research Center, Cleveland, 23 pp., 1976.

Dallam, William C. and James Foster, Digital snow mapping technique using LANDSAT data and general electric image 100 system, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/19, 259-260, Wash., D.C., 1975.

Dean, Arnold M., Jr., Remote sensing of accumulated frazil and brash ice in the St. Lawrence River, CRREL Report 77-8, Hanover, NH, 1977.

Denhartog, Stephen L., Aerial photointerpretation of a small ice jam: special report, CRREL, 21, 1977.

Dunbar, Moira, and W.F. Weeks, Interpretation of young ice forms in the Gulf of St. Lawrence using side-looking airborne radar and infrared imagery, CRREL, Res. Rep. 337, 1975, 49 pp.

Dunne, J.A., The SEASAT-A project' an overview, Marine Technol. Soc., IEEE Oceans 10A/1-5, Wash., D.C., 1976.

Dunne, J.A., The experimental oceanographic satellite SEASAT-A, Inter-Union Comm. on Radio Meteorology, Boundary-Layer Meteorology (Netherlands) 13(1-4), 393-404, 1978.

Elachi, C., and W.E. Brown, Jr., Imaging and sounding of ice fields with airborne coherent radars, J. of Geophysical Research 80(8), 1113-1119, 1975.

Elachi, C., M.L. Bryan and W.F. Weeks, Imaging radar observations of frozen Arctic lakes, Remote Sensing of Environment 5(3), 169-175, 1976.

Ellerbruch, D.A., W.E. Little, H.S. Boyne, and D.D. Bachman, Microwave characteristics of snow, Proceedings of the 45th Annual Western Snow Conference, Albuquerque, N.M., pp. 68-74, 1977.

Fritzsche, Allen E., and Eugene L. Feimster, Snow water equivalent surveys of the Souris River basin, EG&G, Inc., Las Vegas, Nv., 50 pp. NTIS: PB-250 633, 1975.

Gerson, Donald, Computer estimation of the

presence of sea ice in satellite pictures: Maryland Univ., College Park, Md., 95 pp. NTIS: AD-A012 960, 1975.
Gerson, D.J. and A. Rosefeld, Automatic sea ice detection in satellite pictures, Remote Sensing of Environment 4(3), 187-198, 1975.
Gloersen, P., Beaufort Sea ice zones by means of microwave imagery, U.S. Goddard Space Flight Center, Greenbelt, Md., X-910-75-80, 1975.
Gloersen, P., and V.V. Salomonson, Satellites, new global observing techniques for ice and snow, J. of Glaciology 15(73), 373-387, 1975.
Gloersen, P., H.J. Zwally, A.P.C. Chang, D.K. Hall, W.J. Campbell, and R.O. Ramseier, Time-dependence of sea-ice concentration and multi-year ice fraction in Arctic Basin, Boundary-Layer Meteorology, 13, 339-359, 1978.
Goldfinger, A.D., Digital simulation of a synthetic aperture radar, Johns Hopkins Univ., Silver Spring, Md., NTIS AD-A011 831, 1975.
Hagman, Brenda Blanton, An analysis of Great Lakes ice cover from satellite imagery, NOAA, Ann Arbor, Mich., NOAA-TM-ERL-GLERL-9, NOAA-76101302, 1976.
Hagman, Brenda Blanton, On the use of microwave radiation for Great Lakes ice surveillance, NOAA, Ann Arbor, Mich., NOAA-TM-ERL-GERL-13, NOAA-77072206, 1976.
Hall, D.K., and M.L. Bryan, Multispectral remote observations of hydrologic features on the north slope of Alaska, Proceedings of the American Society of Photogrammetry, Little Rock, Arkansas, pp. 393-[illegible] 4, 1977.
Hall, D.K., A. Chang, J.L. Foster, A. Rango, and T. Schmugge, Passive microwave studies of snowpack properties, Proceedings of the 46th Annual Western Snow Conference, Otter Rock, Oregon, 33-39, 1978.
Hall, D.K., J.L. Foster, A. Rango, and A.T.C. Chang, Passive microwave studies of frozen lakes, Proceedings of the American Society of Photogrammetry, Fall Technical Meeting, Albuquerque, N.M., pp. 195-208, 1978.
Hall, R.T., Spatial variability of ice thickness distribution as determined from LANDSAT-A, Proceedings of the Tenth International Symposium on Remote Sensing of Environment, Ann Arbor, Mich., 1, 611-19, 1975.
Hannaford, J.F., Investigation application of satellite imagery to hydrologic modeling snowmelt runoff in the southern Sierra Nevada, Phase 1 Final Report NAS 5-22957, Goddard Space Flight Center, Greenbelt, Maryland, 48 pp., 1977.
Helms, Ward J. and Robert H. Willard, Polar communications via geostationary satellites, AIDJEX Bulletin 28, 173-187, 1975.
Hibler, W.D., III, W.B. Tucker, III and W.F. Weeks, Techniques for studying sea ice drift and deformation at sites far from land using LANDSAT imagery, Proceedings of the Tenth International Symposium on Remote Sensing of Environment, Ann Arbor, Mich., 1, 595-609, 1975.
Hibler, W.D., III, Techniques for using LANDSAT imagery without land references to study sea ice drift and deformation, AIDJEX Bulletin 31, 115-135, 1976.
Hubertz, J.M., Paul E. LaViolette, Surface currents off the east coast of Greenland as deduced from satellite photographs of ice floes, Geophys. Res. Letters 2(9), 400-402, 1975.
Hundemann, Audrey S., Remote sensing applied to hydrology (a bibliography with abstracts), NTIS/PS-77/0677, 1977.
Jayaweera, K.O.L.F., Use of enhanced infrared satellite imagery for sea ice and oceanographic studies, Ocean Eng. 3(5), 293-8, 1976.
Jean, B.R., G.J. Reisor, M.T. Shay, J.A. Permenter, Radar studies of Arctic ice and development of a real-time Arctic ice type identification system, Texas A&M Univ., College Station, Report No. RSC-3005-5, 1975.
Jean, B.R., Radar studies of Arctic ice and development of a real-time Arctic ice type identification system: final report, Texas A&M Univ., College Station, Tex., Remote Sensing Center, 32, 1976.
Katibah, Edwin F., Operational use of LANDSAT imagery for the estimation of snow areal extent, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/10, Wash., D.C., 129-142, 1975.
Kick, W., Application of geodesy, photogrammetry, history and geography to the study of long-term mass balances of central Asiatic glaciers, in Proceedings...Moscow...IAHS-AISH 104, 150-160, 1975.
Kirdar, E., H.H. Schumann, and W.L. Warskow, The application of aerial and satellite snow-mapping techniques for multi-purpose reservoir system operations in Arizona, Proceedings of the 45th Annual Western Snow Conference, Albuquerque, N.M., pp. 95-101, 1977.
Knepper, Daniel H. Jr., David H Sauchyn, Rebecca M. Summer, Nicholas R. French, The application of LANDSAT data to delimitation of avalanche hazards in Montane, Colorado, Colorado Univ., Boulder, Colo., Inst. of Arctic and Alpine Research, 105 pp. NTIS: E77-10211, 1977.
Kovacs, A., H.L. McKim, C.J. Merry, Islands of grounded ice, Arctic 28(3), 213-216, 1975.
Krimmel, R.M., and M.F. Meier, Glacier applications of ERTS images, J. of Glaciology 15(73), 391-401, 1975.
Krimmel, R.M. and M.F. Meier, Measuring snow-covered area to predict reservoir inflow, U.S. Geological Survey, Prof. Pap. 929, 173-175, 1976.
Krimmel, R.M., A.S. Post, and M.F. Meier, Surging and nonsurging glaciers in the Pamir mountains, USSR, U.S. Geological Survey, Prof. Pap. 929, 178-179, 1976.
Kuhn, P.M., L.P. Stearns, and R.O. Ramseier, Airborne infrared imagery of Arctic Sea ice thickness, National Oceanic and Atmospheric Administration, Boulder, CO., 23 pp. NTIS: PB-246 426, 1975.
Kunzi, K.F. and D. H. Staelin, Measurements of snow cover over land with the Nimbus-5 microwave spectrometer, Proceedings of the tenth International Symposium on Remote Sensing of Environment, Ann Arbor, Center for Remote Sensing Information and Analysis, Environmental Research Institute of Michigan 2, 1245-1253, 1975.
Kunzi, K.F., A.D. Fisher and D.H. Staelin, J.W. Waters, Snow and ice surfaces measured by the Nimbus 5 microwave spectrometer, J. of Geophysical Research, 81(27), 4965-4980, 1976.

Larson, L.W., An application of the aerial gamma monitoring techniques for measuring snow cover water equivalents on the Great Plains, Symposium on Snow Management on the Great Plains, Bismarck, ND, 1975.

LaViolette, P.E. and J.M. Hubertz, Surface circulation patters off the east coast of Greenland as deduced from satellite photographs of ice floes, Geophys. Res. Letters, 2(9), 400-402, 1975.

LeSchack, Leonard A., Potential use of satellite IR for ice thickness mapping, Development and Resources Transportation Co., Silver Spring, Md., for National Environment Satellite Service, NOAA, Dept. of Commerce, Wash., D.C. Contract 3-35384, 1975.

Limpert, Fred A., Operational application of satellite snow cover observations, northwest United States, NASA, Wash., D.C., SP-391, 71-85, 1975.

Lindenlaub, John and James Russell, An introduction to quantative remote sensing, The Laboratory Univ., West Lafayette, Ind., Infor. Note 110474, 1975.

Linlor, William I., and George R. Jiracek, Electromagnetic reflection from multilayered snow models, J. of Glaciology, 14(71), 501-515, 1975.

Linlor, W.I., Multilayered models for electromagnetic reflection amplitudes, NASA Technical Report Tk R-438, Ames Research Center, Moffett Field, CA., 51pp., 1976.

Linlor, W.I., D.J. Angelakos, F.D. Clapp, J.L. Smith, Coherent microwave backscatter of natural snowpacks, Univ. California, Berkeley, NTIS: N78-19578, 18 pp., 1977.

Luther, S.G., L.A. Bartolucci, and R.M. Hoffer, Snow cover monitoring by machine processing of multitemporal LANDSAT MSS data, NASA, SP-391, 279-294, 1975.

MacDonald, William R., Glaciology in Antarctica, U.S. Geological Survey, Wash., D.C., Prof. Pap. 929, 194-195, 1976.

Matson, Michael, Winter snow-cover maps of North America and Eurasia from satellite records, 1966-1976: National Environmental Satellite Service, Wash., D.C. 36 pp. NTIS: PB-267 393, 1977.

Mayer, Walter G., Scale model ultrasonic study of Arctic ice: final technical report, Georgetown Univ., Wash., D.C., GUUS-01751, 1975.

Mayer, Walter G., Peter H. Huang, Leslie E. Pitts, and Thomas J. Plona, Sonic reflection from an ice plate in sea water: technical report no. 7, 1 Nov 76-30 Apr 77, Georgetown Univ., Wash., D.C. GUUS-05777, 1977.

McGinnis, D.F., Jr., Progress report on estimating snow depth using VHRR data from NOAA environmental satellites, NASA, Wash., D.C. SP-391, 313-324, 1975.

McGinnis, David F., Jr., John A. Pritchard, and Donald R. Wiesnet, Determination of snow depth and snow extent from NOAA 2 satellite very high resolution radiometer data, Water Resources Research 11(6), 897-902, 1975.

McGinnis, David F., Jr., John A. Pritchard, Donald R. Wiesnet, Snow depth and snow extent using VHRR data from the NOAA-2 satellite, NOAA, Wash., D.C., Tech Memo. NESS 63, 1975b

McGinnis, D.F., and S.R. Schneider, Monitoring river ice break-up from space, Photogrammetric Engineering and Remote Sensing, Vol. 44, No 1, pp. 57-68, 1978a.

McGinnis, D.F. and S.R. Schneider, Satellite detection of an extremely light snowfall in Arizona, Monthly Weather Review, Vol. 106, No. 9, 1380-1383, 1978b.

McMillan, Michael, and James L. Smith, Remote sensing of snowpack density using short-wave radiation, NASA, SP-391, 361-373, 1975.

McMillan, Michael C., and David Forsyth, Satellite images of Lake Erie ice: January - March 1975, NOAA Tech. Memo. NESS-80, NOAA-76080324, 1976.

Meier, Mark F. and W.E. Evans, Comparison of different methods for estimating snowcover in forested, mountainous basins using LANDSAT (ERTS) images, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/17, 215-234, 1975.

Meier, Mark F., Application of remote-sensing techniques to the study of seasonal snow cover, J. of Glaciology 15(73), 251-265, 1975.

Meier, M.F., Monitoring the motion of surging glaciers in the Mount McKinley Massif, Alaska, U.S. Geological Survey, Prof. Paper 929, 185-187, 1976.

Meisner, Douglas, Satellite remote sensing of snow cover in the Adirondack Mountains, State Univ. New York, Syracuse, NTIS. PB-267063, 100 pp., 1977.

Meisner, D.E., T.M. Lillesand, and A.R Eschner, Satellite remote sensing of snowcover in the Adirondock mountains, Proceedings of the American Society of Photogrammetry, Washington, D.C., pp. 159-180, 1977.

Merry, C.J., H.L. McKim, S. Cooper, S.G. Ungar, Preliminary analysis of water equivalent/snow characteristics using Landsat digital processing techniques, Proceedings of the 1977 Eastern Snow Conference, Belleville, Ontario, Canada, 1977.

Merry, Carolyn J., and Harlan L. Mckim, Computer processing of LANDSAT digital data and sensor interface development for use in New England reservoir management, special report, CREEL, 68, 1978.

Meier, Mark F. and W.E. Evans, Comparison of different methods for estimating snowcover in forested, mountainous basins using LANDSAT (ERTS) images, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/17, 215-234, 1975.

Meier, Mark F., Application of remote-sensing techniques to the study of seasonal snow cover, J. of Glaciology 15(73), 251-265, 1975.

Meier, M.F., Monitoring the motion of surging glaciers in the Mount McKinley Massif, Alaska, U.S. Geological Survey, Prof. Paper 929, 185-187, 1976.

Meisner, D.E., T.M Lillesand, and A.R. Eschner, Satellite remote sensing of snowcover in the Adirondock mountains, Proceedings of the American Society of Photogrammetry, Washington, D.C., pp. 159-180, 1977.

Mitchell, P.A., Aerial ice reconnaissance and satellite ice information microfilm file, Naval Oceanographic Office Ref Pub . 17(76), 30, 1976.

Mitchell, Peter A., Aerial ice reconnaissance

and satellite ice information, Naval Oceanographic Office NP-17(76)-Suppl-1, 1977.
Moore, R.K., J.P. Claassen, R.L. Erickson, R.K. Fong, M.J. Komen, Radar systems for a Polar mission, Vol. 1; Final Report, Kansas Univ., Lawrence, NASA-CR-156640, RSL-TR-291-2-V-1, 1977.
Moore, R.K., J.P. Claassen, R.L. Erickson, R.K.T. Fong, and B.C. Hanson, Radar systems for a Polar mission, Volume 3, Appendices A-D, S,T; Final Report, Kansas Univ., Lawrence, NASA-CR-156641, RSL-TR-291-2-V-3, 1976.
Morey, Rexford M., Airborne sea ice thickness profiling using an impulse radar, Geophysical Survey Systems, Inc., Burlington, Mass., USCG-D-178-75, CGR/DC-28/75, 1975.
Morrison, R.B., Enhancement of topographic features by snow cover, U.S. Geological Survey, Prof. Pap. 929, 72-75, 1976.
Muller, F., H. Blatter and G. Kappenberger, Temperature measurement of ice and water surfaces in the north water area using an airborne radiation thermometer, J. of Glaciology 15(73), 241-250, 1975.
Nye, J.F., A test of the ice thickness redistribution equations by measurements on ERTS pictures, AIDJEX Bull. No. 28, 141-149, 1975.
O'Brien, Harold W. and Richard H Muris, Red and near-infrared spectral reflectance of snow, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/24, 345-360, 1975a.
O'Brien, Harold W., and Richard K. Munis, Red and near-infrared spectral reflectance of snow, CRREL, Res. Rep. 332, 24 pp., 1975b.
O'Brien, Harold W., Observations of the ultraviolet spectral reflectance of snow, CRREL, Rep. 77-27, 24 pp., 1977.
Onstott, R.G., G.J Dome, R.A. Hand, James Hague, and H. Pape, Backscatter properties of sea ice with radar, Arctic operations description and preliminary data summary, Kansas Univ., Center for Research Inc., Lawrence Remote Sensing Lab., RSL-TM-331-1, 1977.
Ostheider, M., Evaluation of NOAA-2 VHRR imagery for Arctic sea ice studies, in Proceedings of the Tenth International Symposium on Remote Sensing of Environment, 1975, Ann Arbor, Center for Remote Sensing Information and Analysis, Environmental Research Institute of Michigan, 1, 621-631, 1975.
Outcalt, Sam I., Analysis of the near-surface energy transfer environment from thermal infrared imagery, J. of Glaciology, 15(73), 267-275, 1975.
Parashar, S.K., State of the art - radar measurement of sea ice, Kansas Univ., Center for Research Inc., Lawrence, Remote Sensing Lab, RSL Tech. Note, 291-1, 1975a.
Parashar, S.K., R.M. Haralick, R.K. Moore and A.W. Biggs, Radar scatterometer discrimination of sea-ice types, Transactions on Geoscience Electronics, GE-15(2), 83-87, 1977.
Post, A.S., Environmental geology of the central Gulf of Alaska coast, U.S. Geological Survey Prof. Pap. 929, 117-119, 1976.
Post, A., M.F. Meier, and L.R. Mayo, Measuring the motion of the Lowell and Tweedsmuir surging glaciers of British Columbia, Canada, U.S. Geological Survey, Prof. Paper 929, 180-184, 1976.
Poulin, Ambrose O., Significance of surface temperature in the thermal infrared sensing of sea and lake ice, J. of Glaciology 15(73), 277-283, 1975.
Poulin, Ambrose O., The potential of thermal infrared imagery for supplemental map information in snow-covered areas, Army Engineer Topographic Labs, Fort Belvoir, Va., 43 pp. NTIS: AD-A028 384, 1976.
Rango, Albert, Applications of remote sensing to watershed management, Proceedings of the ASCE Irrigation and Drainage Division Symp on Watershed Management, Logan Utah, 700-714,1975a.
Rango, A., Operational applications of satellite snowcover observations project, Proceedings of the Tenth International Symposium on Remote Sensing of Environment, Ann Arbor, Center for Remote Sensing Information and Analysis, Environmental Research Institute of Michigan, 2, 1367-1376, 1975b.
Rango, Albert, Operational applications of satellite snowcover observations, Proceedings of a workshop, South Lake Tahoe, California, NASA SP-391, 430 pp., 1975c.
Rango, Albert, An overview of the applications systems verification test on snowcover mapping, in Operational Applications of Satellite Snowcover Observations, NASA, Wash., D.C. SP-391/1, 1-12, 1975d.
Rango, A., V.V. Salomonson, and J.L. Foster, Employment of satellite snowcover observations for improving seasonal runoff estimates, in Operational Applications of Satellite Snowcover Observations. NASA SP-391/12, 157-174, 1975.
Rango, A., and K.I. Itten, Satellite potentials in snowcover monitoring and runoff predicton, Nordic Hydrology, 7(4), 209-230, 1976.
Rango, Albert, Remote sensing: snow monitoring tool for today and tomorrow, Proceedings of the 45th Annual Western Snow Conference, Albuquerque, 75-81, 1977.
Rango, A., J.F. Hannaford, R.L. Hall, M. Rosenzweig, and A.J. Brown, The use of snowcovered area in runoff forecasts, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Md., 33 pp. NTIS: N78-16410, 1977a.
Rango, A., V.V. Salomonson, and J.L. Foster, Seasonal streamflow estimation in the Himalayan region employing meteorological satellite snow cover observations, Water Resources Res., 13/1, 109-112, 1977b.
Rango, A., Pilot tests of satellite snowcover/runoff forecasting systems, Proceedings of the 46th Annual Western Snow Conference, Otter Rock, Oregon, 7-14, 1978.
Rango, A., A.T.C. Chang, and J.L. Foster, The utilization of spaceborne microwave radiometers for monitoring snowpack properties, Nordic Hydrology, Vol. 10, No. 1, 1979.
Reimitz, E., P.W. Barnes, Influence of sea ice on sedimentary processes off northern Alaska, U.S. Geological Survey, Professional Paper 929, 360-361, 1976.
Rothrock, D.A., Testing the redistribution of sea ice thickness from ERTS photographs, AIDJEX Bull., Univ. of Wash., Seattle, 29, 1-19, 1975.
Sabatini, Romeo R., Dennis L. Hlavka, and Ronald Arcese, Applications of the Nimbus 5 ESMR to rainfall detection over the oceans and to sea-ice detection, Earth Satellite Corp., Wash., D.C., 80 pp. NTIS: AD-A013 245, 1975.

Salomonson, V.V. and A. Rango, Summary of the operational application of satellite snowcover observations working session, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/29, 421-426, 1975.

Schell, J.A., B.R. Jean and W.C. Hulse, Radar studies of Arctic ice and development of a real-time Arctic ice type identification system, Texas A&M Univ., College Station, Tex., 69 pp. NTIS: AD-A040 275, 1976.

Schertler, R.J., R.A. Mueller, R.J. Jirberg, D.W. Cooper, T. Chase, J.E. Heighway, A D. Holmes, R.T. Gedney, and H. Mark, Great Lakes all-weather ice information system, Proceedings of the Tenth International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan, pp. 1377-1404, 1975.

Schneider, Stanley R., The operational program of satellite snowcover observations at NOAA/NESS, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/7, 87-101, 1975.

Schneider, S.R., D.R. Wiesnet, and M.C. McMillan, River basin snow mapping at the National Environmental Satellite Service, NOAA TM NESS 83, 1976.

Schumann, Herbert H., Operational applications of satellite snowcover observations and LANDSAT data collection systems operations in central Arizona, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/2, 13-28, 1975.

Sellman, P.V., W.F. Weeks, and W.J. Campbell, Use of side-looking airborne radar to determine lake depth on the Alaskan north slope, CRREL Special Report 230, 1975.

Seifert, R.D., R.F. Carlson, and D.L. Kane, Operational applications of NOAA-VHRR imagery in Alaska, in Operational Applications of Satellite Snowcover Observations, NASA, Wash., D.C., 143-155, 1975.

Sharp, James M. and Randall W. Thomas, A comparison of operational and LANDSAT-aided snow water content estimation systems, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/23, 325-344, 1975.

Shapiro, Lewis H. and John J. Burns, Satellite observations of sea ice movement in the Bering Strait region, in Alaska Science Conference, 24th, Univ of Alaska, 1973, Proceedings Alaska Univ., Fairbanks, Geophysical Institute, 379-386, 1975.

Shiue, J.C., A.T.C. Chang, H. Boyne, and D. Ellerbruch, Remote sensing of snowpack with microwave radiometers for hydrological applications, Proceedings of the Twelfth International Symposium on Remote Sensing of Environment, Manila, pp. 877-886, 1978.

Speers, D., D. Kuehl, and V. Schermerhorn, Development of the operational snow band SSARR model, in Proc. Modeling Snow Cover Runoff, USA Cold Regions Res. and Eng. Lab., Hanover, N.H., 10pp., 1979

Staelin, D.H., P.W. Rosenkranz, F.T. Barath, E.J. Johnston, and J.W. Waters, Microwave spectroscopic imagery of the earth, Science, Wash., D.C. 197(4307), 991-993, 1977.

Steinhoff, Harold W., and Albert H. Barnes, Determination of snow depth and water equivalent by remote sensing, Office of Water Research and Technology, Wash., D.C., 76,W77-05104, OWRT-A-019 -Colo., 1976.

Stringer, William J., LANDSAT survey of near-shore ice conditions along the Arctic Coast of Alaska, quarterly progress rep. 3, NASA Wash., D.C., CR-148739, 1975.

Stringer, William J., and Stephen A. Barrett, LANDSAT survey of near-shore ice conditions along the Arctic Coast of Alaska, LANDSAT follow-on investigation number 21300, Alaska Univ., College, Ak., 212 pp. NTIS: E78-10136, 1978

Super, A.D. and A.D. Osmer, Remote sensing as it applies to the International Ice Patrol, Proceedings of the tenth International Symposium on Remote Sensing of Environment, 1975, Ann Arbor, Center for Remote Sensing Information and Analysis, Environmental Res. Institute of Michigan, 2, 1231-1234, 1975.

Sydor, Michael, Western Lake Superior ice, U.S. Geological Survey, Wash., D.C., Prof. Pap. 929, 169-172, 1976.

Thomas, I.L., A.J. Lewis, and N.P. Ching, Snowfield assessment from LANDSAT (South Island, New Zealand), Photogramm. Eng. Remote Sensing, Falls Church, Va., 44/4, 493-502, 1978.

Thompson, A.G., Utilization of LANDSAT monitoring capabilities for snowcover depletion analysis, in Operational Applications of Satellite Snowcover Observations, NASA SP-391/9, 113-127, 1975.

Ulaby, F.T., Snow backscatter in the 1-8 Ghz region, NASA-CR-115160, RSL-TR-177-61, 1976.

Ulaby, F.T., and W.H. Stiles, Backscatter and emission of snow, presented at the NASA Microwave Remote Sensing Symposium, Houston, Texas, 1977

Ulaby, F.T., A.K. Fung, and W.H. Stiles, Backscatter and emission of snow: Literature review and recommendations for future investigations, University of Kansas, Center for Research, RSL Technical Report 369-1, 139pp., 1978

Valovcin, Francis R., Snow/cloud discrimination, U.S. Air Force, Geophysics Lab., Hanscom AFB, Ma., 1976, 17 pp. NTIS: AD-A032 385.

Vant, M.R., A combined empirical and theoretical study of the dielectric properties of sea ice over the frequency range 100 MHz to 40 GHz, in Dissertation Abstracts International, B, 37(12), Pt. 1, 6263B-64-B, 1977

Warskow, W.L., T.T. Wilson, Jr., and K. Kirdar, Application of hydrometeorological data obtained by remote sensing techniques for multipurpose reservoir operations, NASA, Wash., D.C. SP-391, 29-37, 1975.

Washicheck, Jack N. and Tony Mikesell, Operational applications of satellite snowcover observations in Rio Grande drainage of Colorado, in Operational Applications of Satellite Snowcover Observations, NASA, Wash., D.C. SP-391-5, 53-69, 1975.

Weeks, W.F. and W.J. Campbell, Remote Sensing plan for the AIDJEX main experiment, AIDJEX Bulletin No. 29, 21-48, 1975

Weeks, W.F., P. Sellmann, and W.J. Campbell, Interesting features of radar imagery of ice-covered north slope lakes, J. of Glaciology, 18(78), 129-136, 1977

Wiesnet, Donald R. and Michael Matson, Monthly winter snowline variation in the northern hemisphere from satellite records, 1966-75, NOAA/NESS Tech. Memo. 74, Wash., D.C., 1975

Wiesnet, D.R., Remote sensing and its application to hydrology, J.C. Rodds, (ed.), Facets of Hydrology, London, New York, etc., John Wiley and Sons, 37-59, 1976.

Wiesnet, D.R., and M. Matson, A possible forecasting technique for winter snow cover in the northern hemisphere and Eurasia, Monthly Weather Review, Vol. 104, No. 7, pp 828-835, 1976.

Wiesnet, D.R., and D.F. McGinnis, Mapping snow extent in the Sierra Nevada of California, U.S. Geological Survey, Prof. Paper 929, pp. 176-177, 1976.

Williams, R.S., Vatnajokull Icecap, Iceland, U.S. Geological Survey, Prof. Paper 929, 188-193, 1976.

Williams, Richard S., Jr. and William D. Carter, (eds.), ERTS-1, A new window in our planet, Geological Survey Prof. Pap. 929, 1976.

Woolever, G.R., L.A. Kidd, J.P Welsh, J.A. McIntosh and L.D. Farmer, Utilization of remote sensing techniques for U.S. Coast Guard missions, Proceedings of the Tenth International Symposium on Remote Sensing of Environment, 1975, Ann Arbor, 1, 3-16, 1975.

Yates, H.W., and W.R. Bandeen, Meteorological applications of remote sensing from satellites, Proc. IEEE 63(1), 148-163, 1975.

Zwally, H. Jay and P. Gloersen, Passive microwave images of the polar regions and research applications, Polar Record 18(116), 431-450, 1977.

Zwally, H. Jay, Microwave emissivity and accumulation rate of polar firn, J. of Glaciology, 18(79), 195-215, 1977.

Zwally, H.J., T.T. Wilheit, P. Gloersen, and J.L. Mueller, Characteristics of antarctic sea ice as determined by satellite-borne microwave imagers, COSPAR 19th Plenary Meeting, Phila., Pa., Paper C.2a.3, 1976.

N81-21698

Paper 94

Reprinted from the Specialty Conference on
IRRIGATION AND DRAINAGE IN THE NINETEEN-EIGHTIES
ASCE/Albuquerque, New Mexico/July 17-20, 1979

SEASONAL AND DAILY SNOWMELT RUNOFF ESTIMATES UTILIZING SATELLITE DATA

By Albert Rango

ABSTRACT

For effective water management both seasonal (volumetric) and short-term runoff forecasts are required. Satellite-derived snowcovered area data have been found to be useful information for aiding in these forecasts. Regression approaches incorporating snowcovered area data have been found to be successful in reducing error in seasonal forecasts on certain watersheds. Methods using snowcovered area to update seasonal forecasts as snowmelt progresses are also being used in quasi-operational situations. The input of snowcovered area to snowmelt models for short term predictions has been attempted in two ways; namely, the modification of existing hydrologic models and/or the use of models that were specifically designed to use snowcovered area. A daily snowmelt runoff model has been used with Landsat data to simulate discharge on remote basins in the Wind River Mountains of Wyoming. Daily predicted and actual flows compare closely, and, summarized over the entire snowmelt season (April 1 - September 30), the average difference is only three percent. The model and snowcovered area data are currently being tested on additional watersheds to determine the method's transferability.

Paper 95

PILOT TESTS OF SATELLITE SNOWCOVER/RUNOFF FORECASTING SYSTEMS 1/

636-78

By

Albert Rango 2/

Background

Since 1972 when Landsat-1 and NOAA-2 were launched, speculation that high resolution satellite data could be used to measure snowcovered area and, subsequently, to assist in runoff prediction has been optimistically proposed (Salomonson and Rango, 1974). The use of photographic methods for interpreting snowcovered area has been favored over digital methods, and a handbook describing the various snowcover photointerpretive methods has been compiled (Barnes and Bowley, 1974). Various snowcover analysis techniques, successful in different areas, have been reported at a workshop at Lake Tahoe (Rango, 1975).

Although methods for extracting snowcover information from satellite data are now well established, the means by which snowcovered area data can be used for runoff prediction are still being developed. Using long-term meteorological satellite data, Rango, Salomonson, and Foster (1977) demonstrated that snowcovered area at a particular point early in the snowmelt season is highly correlated with seasonal streamflow. Various investigators are now exploring means by which remotely-sensed snowcovered area data can most effectively be used in runoff prediction. In order to assure a thorough test directly applicable to water resources users, NASA entered into a cooperative, quasi-operational, four-year test program with various agencies in four major Western U.S. snow zones. This project, entitled "Operational Applications of Satellite Snowcover Observations" (OASSO), is part of NASA's Applications Systems Verification and Transfer (ASVT) program designed to demonstrate and test the usefulness of developing technology in remote sensing.

1/ Presented at the Western Snow Conference, April 18-20, 1978, Otter Rock, Oregon.

2/ Senior Research Hydrologist, NASA/Goddard Space Flight Center, Greenbelt, Maryland.

"Reprinted Western Snow Conference 1978"

Paper 96

NASA Technical Memorandum 79713 February 1979.

REMOTE SENSING OF SNOW AND ICE: A REVIEW OF RESEARCH IN THE UNITED STATES 1975-1978*

Albert Rango
Laboratory for Atmospheric Sciences (GLAS)

ABSTRACT

Research work in the United States from 1975-1978 in the field of remote sensing of snow and ice is reviewed. Topics covered include snowcover mapping, snowmelt runoff forecasting, demonstration projects, snow water equivalent and free water content determination, glaciers, river and lake ice, and sea ice. A bibliography of 200 references is included.

*Contribution to the Report of the U.S. National Committee to the International Union of Geodesy and Geophysics (IUGG) for the 1975-1978 IUGG Quadrennial Report.

Paper 97

N81-21699

Nordic Hydrology, 10, 1979, 25-40

No part may be reproduced by any process without complete reference

The Utilization of Spaceborne Microwave Radiometers for Monitoring Snowpack Properties

A. Rango, A. T. C. Chang, and J. L. Foster
Goddard Space Flight Center, Greenbelt, Maryland, U.S.A.

Snow accumulation and depletion at specific locations can be monitored from space by observing related variations in microwave brightness temperatures. Using vertically and horizontally polarized brightness temperatures from the Nimbus 6 Electrically Scanning Microwave Radiometer, a discriminant function can be used to separate snow from no snow areas and map snowcovered area on a continental basis. For dry snow conditions on the Canadian high plains significant relationships between snow depth or water equivalent and microwave brightness temperature were developed which could permit remote determination of these snow properties after acquisition of a wider range of data. The presence of melt water in the snowpack causes a marked increase in brightness temperature which can be used to predict snowpack priming and timing of runoff. As the resolutions of satellite microwave sensors improve the application of these results to snow hydrology problems should increase.

Introduction

Runoff from melting snow provides greater than 65% of the total streamflow across most of the mountainous western United States. The characterization of the snowpack through various types of physical measurements allows for the prediction of snowmelt runoff and better multipurpose management of the scarce water supply. In the upper midwest United States and the Canadian high plains, knowledge of snowpack characteristics is extremely valuable for flood forecasting purposes in the early spring.

In order to adequately monitor the snow resource for runoff prediction reasons, the measurement of various parameters is critical, namely, snowcovered area (SCA), depth, density, snow water equivalent (S_n), and free water content. In addition to the snow properties, the condition of the underlying soil is especially significant for estimating the amount of snowmelt water that will reach the stream channel as runoff. Historically, most snow measurement in the western United States has dealt with depth, density, and S_n observations at pre-located snow courses on or near the beginning of each month from approximately February through May. A shifting emphasis in forecast needs from seasonal streamflow to short term runoff estimates in recent years has seen a movement to continuous automatic monitoring of S_n at many locations. The telemetering of these data in real time through conventional "line of sight" radio, meteor burst, and satellite transmissions has increased the use of snowmelt models for runoff forecasting.

Conventional S_n measurements are representative of a given point in a watershed, and seldom are there more than a few point measurements in a specific basin. The regression analysis approach to seasonal streamflow forecasting somewhat compensates for this deficiency by treating the S_n measurements as indicies. When the same data are used in models, however, the point measurements are assumed to be representative of large sub basin areas or the entire watershed. Such assumptions can be false and, as a result, the watershed models can be calibrated using erroneous information. This consequently may lead to model outputs that are significantly in error in specific situations.

Remote sensing may provide data more compatible to model operation because of the capability to make measurements over the entire watershed area. A true indication of snow character over the entire watershed could then be more logically obtained from these measurements. SCA measurement is a particularly good example because, during snowmelt, the area covered by snow indicates the portion of the watershed that can potentially contribute melt water for runoff. If energy balance or degree day techniques are being used to forecast snowmelt, the SCA data define the zones of the watershed where the respective equations should be applied. A model developed by Martinec (1975) requires SCA as the major snow input parameter, and simulated hydrographs using this model have closely matched actual hydrographs.

Since 1973 SCA has been successfully mapped from space using Landsat and NOAA satellites by various investigators (Rango 1975). A cooperative demonstration project is currently being conducted in which several western United States water management agencies are using the satellite-derived SCA in a quasi-operational mode for runoff predictions (Rango 1978). At present, the most desirable SCA technique involves the combined use of the high resolution, low frequency-of-coverage Landsat data and the low resolution, high frequency-of-coverage NOAA data. Once a sufficient SCA data base is acquired and appropriate techniques developed and documented for incorporating SCA into runoff

prediction, operational use of SCA by water management agencies should result.

A remote sensing technique to estimate snow depth or S_n over an entire basin would add considerably to the already existing SCA capability. Alone such a technique would provide much more detailed knowledge of the snowpack in both mountainous and flatland basins for runoff prediction purposes. Combined with SCA, a large area estimate of total snow volume would be possible. Such an estimate would specify the maximum potential of snow water production and also locate that water in relation to the various elevation zones or sub basins of the watershed.

In order to measure snow depth or S_n a sensor that measures sub surface, as opposed to solely surface, features is required. Radiometers sensitive to microwave energy emitted by the earth's surface and other natural objects such as snow can be used. Depending on the actual wavelength of the sensor, the observed radiation can come from the underlying soil, snow, and atmosphere. The objective when using passive microwave radiometers is to measure the emitted microwave radiation, expressed as brightness temperature, and infer the characteristics of the snowpack. The use of radiometers at several wavelengths may be necessary to separate the emission of the snow from the soil.

The microwave energy emitted from the earth's surface at 0.81 cm, for example, depends upon the earth's physical temperature and emissivity (the product of which equals the microwave brigtness temperature). The emissivity may vary due to changes in the dielectric constant of the emitting surface; the emissivity for water is typically 0.4 and that for the earth's surface 0.85 or greater. The emitted microwave energy results in low brightness temperatures of 120 to 170 K for data recorded over water and high brightness temperatures of 215 to 300 K over land. The dielectric constant for snow is usually lower than that of dry soil and the scattering characteristics of snow particles (Chang et al. 1976) will reduce the brightness temperature further, thus providing the needed contrast in the brightness temperature range for snow field monitoring.

Radiation from dry snow is strongly influenced by crystal or grain size because the radiation emitted by the snowpack is scattered on its way to the surface by the snow grains. Since most of the radiation emitted at short wavelengths (about 1 cm) comes from depths on the order of meters (Zwally and Gloersen 1977), the depth of snow (and thus the number of potential scatterers) also affects the brightness temperature. Larger grain sizes cause more scattering resulting in a lower emissivity. In an area, however, that produces snow grains of similar size from year to year because of similar winter climatic conditions, the microwave emissivity should provide an indication of the relative snow accumulation in that location. The onset of melting which produces liquid water in the snowpack radically alters the snow emissivity. The water coats the snow grains and causes a significant increase in internal absorption of the microwave radiation, and, as a result, causes an increase in the snow emissivity.

Passive microwave data from space have been available since December 1972 when Nimbus 5 was launched with the Electrically Scanning Microwave Radiometer (ESMR) onboard sensing at the 1.55 cm wavelength. The ESMR on Nimbus 5 scans through an angle ±50° in a plane perpendicular to the orbital path of the spacecraft and is designed to measure only the horizontally polarized component of the microwave radiation. Further data became available in June 1975 with the launch of Nimbus 6 with an ESMR instrument capable of receiving dual-polarized microwave radiation from the earth at 0.81 cm wavelength. The Nimbus 6 instrument scans the earth every five and one-third seconds from 35° to the right to 35° to the left about the forward direction with a constant incidence angle of approximately 50° at the earth's surface. The spatial resolution is about 20 km cross-track and 35 km down-track with the noise equivalent temperature for the radiometer about 2 K. Frequency of coverage from the Nimbus 5 and 6 ESMR instruments varies depending on latitude, but on the average is approximately once every 24 hours for Nimbus 5 and once every 48 hours for Nimbus 6. Of course the major drawback of these sensors for snow hydrology application is the poor resolution of about 25 km. A secondary drawback is the limited spectral coverage provided. The data have been examined, however, for pertinence to future sensors that will in all probability be more applicable because of projected increases in spatial resolution and number of spectral intervals. Improvement in spectral coverage is already available. A five channel, dual polarized Scanning Multichannel Microwave Radiometer (SMMR) was launched on Nimbus 7 in 1978. Observations are made at 4.55, 2.81, 1.67, 1.43, and 0.81 cm; however, resolution at the shortest wavelength is still not less than 20 km, and at the long wavelengths it is much larger (Gloersen and Barath 1977). But it is expected that within 5-10 years the utilization of large antennas in space will permit resolutions in the short wavelengths to be reduced to about 1 km. The use of larger antennas and development of improved state-of-the-art radiometers should permit the noise equivalent temperature to remain the same or even improve some as the resolution is improved. With the possibility of such forthcoming advances, valuable information for the application of microwave data to snow hydrology should be obtained from the analysis of existing spaceborne ESMR data.

Related Research

The theory of passive microwave emission from snow and ice has been the subject of several studies (Chang and Gloersen 1975; Chang et al. 1976; Zwally, 1977; Chang and Choudhury 1978; and Chang 1978), and the relation of the microwave theory to actual satellite microwave observations in the polar regions was examined by Zwally and Gloersen (1977). In dealing with applications to

snow hydrology, Meier (1972) indicates that monitoring of passive microwave emission has a great potential for measuring areas covered by snow. The snowline on Mt. Ranier defined using the 270 K brightness temperature on a microwave image compared favorably with the snowline mapped from aerial photography. Künzi and Staelin (1975) have used the microwave spectrometer on Nimbus 5 to measure snowcovered area. Subsequent comparison with snowcover data for the northern hemisphere obtained by NOAA visible sensors has indicated generally good agreement, with the Nimbus 5 maps usually showing a smaller snowcovered area. Measurement of natural and artificial snowpacks in Oregon at several wavelengths (Meier and Edgerton 1971) revealed relationships that show a pronounced decrease in brightness temperature with increasing snow mass (water equivalent) for dry snow. In this same study a rapid increase in brightness temperature of about 70 K was observed as the snow changed from a dry to a wet condition with only about one percent liquid water. An experiment in Colorado with a four frequency, dual polarized radiometer system mounted on a truck platform produced data showing a similar inverse relationship between brightness temperature and S_n (Shiue et al. 1978). Further confirmation of this qualitative relationship is indicated by results from an aircraft measurement program in Colorado using multichannel microwave radiometers over varying snow conditions (Hall et al. 1978). The aircraft data have also shown about a 60 K rise in brightness temperature when liquid water is present in the snowpack. Detailed ground-based multichannel radiometer measurements made in Switzerland (Schanda et al. 1978) have provided a means for classifying dry snow, ice surfaces, and wet snow with a separation of wet snow into two moisture grades a possibility. All these multichannel microwave measurements have direct pertinence to eventual interpretation of the SMMR data.

Investigation Procedure

The general objective of this study was to evaluate the use of spaceborne microwave data for monitoring snowpack characteristics useful in hydrology and climatology. Specifically the investigation was designed to compare snow accumulation and depletion data at several National Weather Service (NWS) stations with variations in ESMR 6 brightness temperatures over a six-month period to see if comparable trends were evident. Further, the feasibility of developing a statistical technique to objectively separate snowcovered from snowfree areas was explored and applied to snow mapping of the North American continent. In order to determine if a relationship between ESMR 5 or 6 data and snow depth (or water equivalent) existed, regression analyses were carried out using point data over a large area and then for a smaller homogeneous region. Data acquisition

during melting snow conditions was also attempted so that the effect of free water in the snowpack on any relationships could be assessed.

The primary remote sensing data source in this study was Nimbus 6 ESMR data. Digital brightness temperatures of both horizontal and vertical polarization were available and used in all phases of the study. In addition, Nimbus 5 ESMR data in a similar format were used in the microwave/snow depth correlation studies. Conventional daily snow depth data were obtained from the NWS publication "Climatological Data" for selected stations in the northern United States. In addition snow depth and water equivalent measurements on an irregular basis were available for Alberta and Saskatchewan observing stations from the Atmospheric Environment Service, Environment Canada. For comparison purposes northern hemisphere snow and ice boundary maps published weekly by NOAA's National Environmental Satellite Service (NESS) and derived from observations made by NOAA-4 and synchronous meteorological satellites were also obtained.

The accumulation and depletion of snow as compared to brightness temperature variations was studied for the following NOAA NWS reporting stations: Lafayette, Indiana; Urbana, Illinois; and Bismarck and Williston, North Dakota. A discriminant analysis interpretation of snowcovered area using ESMR data was carried out over the 48 conterminuous United States and large portions of Canada, specifically, an area bounded in latitude by 25°N to 65°N and in longitude by 60°W to 130°W. Snow depth and brightness temperature correlations were run for data from the four NWS stations previously mentioned, but a detailed study was also conducted over a region of relatively homogeneous vegetation and topography on the Canadian high plains. This area includes the southern portions of Alberta and Saskatchewan and is bounded in latitude by 49°N to 53°30'N and in longitude by 103°W to 114°W.

Results

Snow Accumulation and Depletion

Comparison of snow accumulation and microwave changes from 1 November 1975 to 30 April 1976 at Lafayette, Indiana and Urbana, Illinois indicated a general but weak inverse variation between snow depth and ESMR-6 brightness temperature. Little snow accumulation was evident during this winter season (no more than 12.5 cm at any time) and the snowcover did not last for more than 10 days at a time. Under these shallow snow conditions, the brightness temperature variation due to the snow layer could not be easily separated from the effects of soil moisture, surface roughness, and vegetation cover.

The snow conditions for the 1975-1976 winter in North Dakota differed consi-

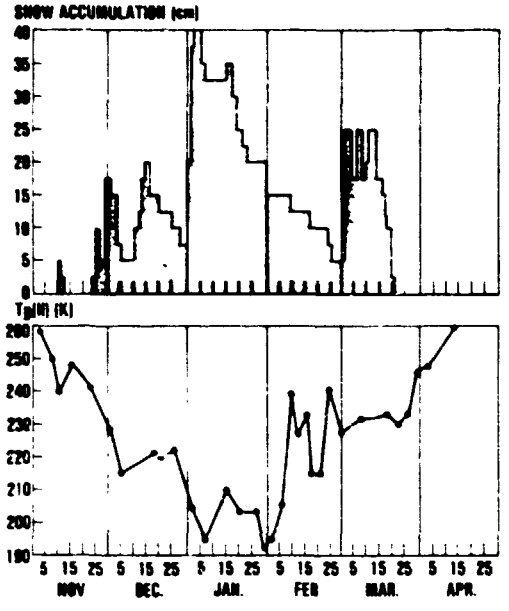


Fig. 1. Snow accumulation and Nimbus 6 ESMR horizontally polarized brightness temperature data (1975-1976) for Williston, North Dakota, U.S.A.

derably from the Lafayette and Urbana sites. From 25 November 1975 to 20 March 1976, the Williston, North Dakota site was completely covered with snow while Bismarck, North Dakota was snowcovered three-quarters of the time. Figure 1 shows the brightness temperature and snow depth variation for Williston.

The measured brightness temperatures (horizontal polarization) for the Williston area were between 190 and 210 K and the snow depths were between 20 cm and 40 cm in the month of January, 1976. During the same period, the brightness temperatures were between 208 and 220 K with 7.5 cm to 17.5 cm of snow depth in the Bismarck area. This low brightness temperature is probably due to the effect of volume scattering of the snow grains (Chang et al. 1976). By comparing the snow accumulation of the two sites in North Dakota, the amount of decrease in brightness temperature seems directly proportional to the snow depth. The modulations of the brightness temperature by snow accumulation give rise to the eventual possibility of inferring snow depth from remote measurements.

Snow Area Determination

In order to assign a snow or no-snow classification for an area, a statistically oriented discriminant analysis was prepared from data on snow accumulation and ESMR-6 brightness temperatures. The purpose of the discriminant analysis was

to find the dividing function which best separated the snow and no-snow condition. The position of the point along the discriminant function was used to assign each individual data point to its most probable category. The main assumptions considered in the discriminant analysis were that the two groups had homogeneous variance-covariance matrices, and the variables were normally distributed. The first assumption ensures that the two groups are spectrally different from each other, and the second assumption is necessary if significance tests are to be utilized. The variables used in the discriminant analysis are the horizontally polarized brightness temperature, T_H, and the vertically polarized brightness temperature, T_V. The discriminant function derived according to Mather (1976), can be expressed as:

$$R = 0.9626\,T_V - 0.9890\,T_H \qquad (1)$$

where R is the projection on the discriminant axis. The averaged R for the snow class is 17.17, while the average R for the no-snow class is 11.21. Therefore, any individual data point with calculated R value closer to the average of snow class than the average of no-snow class will be assigned to the snow class and vice versa. When this technique is used to assign the snow, no-snow classes for the area of interest, the probability of misclassification is about 0.2 per cent for a normally distributed data set.

Because ESMR brightness temperature differences can result from factors other than just snow and no-snow categories, some precautions should be taken in the discriminant analysis. A combination of water and land mass within an ESMR's field of view could give a brightness temperature similar to that of a snowfield. The signature of a relatively wet ground could also resemble the snow signature. Under these conditions, more than one discriminant function will be required to separate the different categories. In this study, a geographical and climatological approach was used to analyze and resolve these discrepancies. For example, the coast of Florida has usually been misclassified as a snowcovered area by the discriminant function. This obvious misclassification is due to the mixed signatures of water and land mass within the ESMR footprint. By knowing the geographical location, the misclassified pixels can be reassigned to a no-snow class without using other discriminant functions.

By using the derived discriminant function it was possible to separate snow and no-snow areas over the North American continent using the Nimbus-6 ESMR brightness temperatures. The calibrated brightness temperatures for each time period (usually seven days) are projected on a 1° latitude by 1° longitude grid map before applying the discriminant function. The map sequence started in mid November 1975 and continued to mid April 1976. In order to validate the snow coverage maps derived from the ESMR data, a comparison was made with the northern hemisphere averaged snow and ice boundary maps, published weekly by NOAA/NESS. The NOAA/NESS boundary maps were derived from observa-

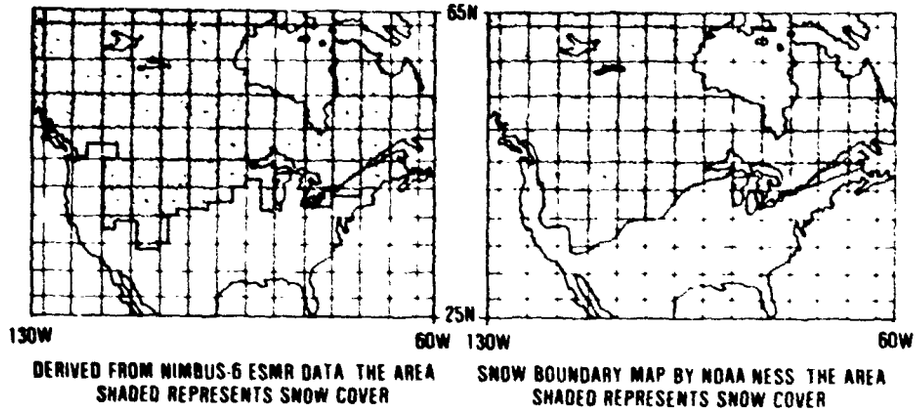


Fig. 2. Snow coverage maps of North America for the period of 15-21 December 1975.

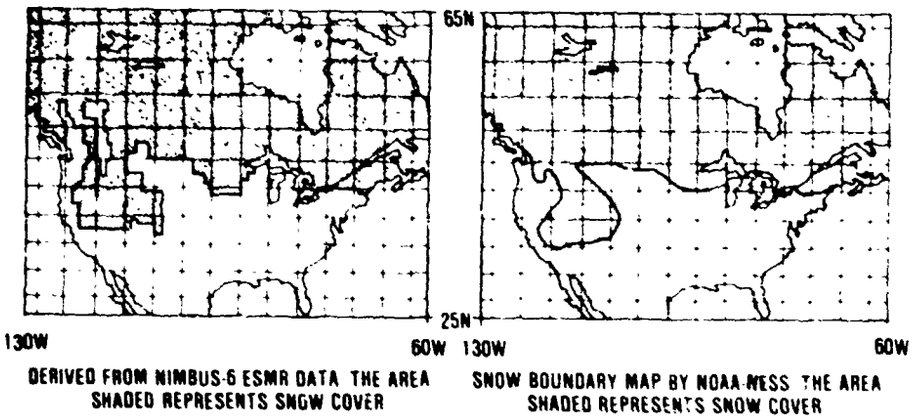


Fig. 3. Snow coverage maps of North America for the period of 15-21 March 1976.

tions made by NOAA-4 and synchronous meteorological satellites. Figures 2 and 3 are example comparisons of the snow coverage maps for the periods of 15-21 December 1975 and 15-21 March 1976, respectively. The snowcovered area data from both maps were generally comparable with the ESMR boundary maps showing a slightly less overall snow areal extent. Specific differences in locations of snowcovered areas between the two snow boundary maps is probably due to the 1° x 1° grids used to plot the poor resolution ESMR data and the lack of a geometrically registered ESMR data product. The large grid units result in blocky snowline boundaries compared to the smooth snowlines in the NOAA maps. The ESMR registration problem can result in further discrepancies between the two

snowlines. In addition, the fact that the microwave emission from the soil can penetrate very thin snowcover without appreciable effect on the brightness temperature can result in a no-snow classification whereas the NOAA visible sensors will record a snow classification. Another potential problem in generating snow coverage maps is the effect of the surface slope. Since the amount of radiation emitted and reflected from a smooth surface is a function of incidence angle, the average slope within the instrument's instantaneous field of view will affect the measured brightness temperature and the difference between the vertical and horizontal brightness temperatures. The surface slope effect tends to provide disagreement for snow coverage maps in mountainous areas.

Snow Depth Relations

Snow depth and ESMR-6 brightness temperature data for the four previously mentioned stations in Indiana, Illinois, and North Dakota were examined first. A multiple linear regression was performed with the vertically (T_V) and horizontally (T_H) polarized brightness temperatures as the predictor variables and snow depth in cm (Y) as the criterion variable. For approximately 500 data points the following regression relationship significant at the 0.01 level resulted:

$$Y = \quad \quad (1)$$

where

the coefficient of determination, $R^2 = 0.31$
the standard deviation of the snow depth values, $S_D = 7.9$ cm
the standard error of the estimated snow depth values, $S_E = 6.6$ cm

Although this relationship is significant, the relatively low R^2 value and the fact that the S_E of 6.6 cm is not much of an improvement over the S_D of 7.9 cm limits the applicability for estimating snow depth. Considerable ground cover and climatological variations between the study areas may explain the marginal regression relationship. Mixed cover types including dormant agricultural land, pasture, and trees have a varying effect on the brightness temperature of the snow which likely masks the brightness temperature response to varying snow depth. The generally colder environment of North Dakota as compared to the warmer Indiana-Illinois area, where melt-freeze conditions persist during the winter, will result in different snow emission conditions. In Indiana-Illinois melting and then freezing conditions produce significant crusts and ice layers, and free water will exist in the snowpack more of the time. In North Dakota the colder temperatures will result in a drier and more homogeneous snowpack with generally small grains or crystals serving as efficient scatterers of the emitted microwave radiation. As a result the North Dakota snowpack would produce microwave brightness temperatures considerably lower than Illinois-Indiana and with more consistent varia-

tions in response to changes in snow depth.

In order to try to minimize these variations between study areas and perhaps improve the snow depth and brightness temperature relationship, data from a more homogeneous area, the Canadian high plains, were examined. Fig. 4 shows the area that was studied in southern Alberta and Saskatchewan. Four general vegetation types are prevalent over this relatively flat area. In order to make the cover conditions as uniform as possible snow depth and brightness temperature data analysis was restricted to the areas of short and high grass prairies. Snow depth values for each of the numbered stations were used to draw isonivals which were then averaged over the 1° × 1° grid blocks for comparison with the averaged ESMR brightness temperatures. Both vertical and horizontal brightness temperatures were used in simple linear regression analyses. Fig. 5 illustrates the snow depth versus brightness temperature data for Nimbus 5 and the resulting significant (at the 0.002 level) regression line and statistics. The Nimbus 5 data are from the nighttime pass on 14 March 1976 and the snow depth data are from 15 March 1976. Air temperatures prior to 15 March were well below 0°C with little chance of significant melting, and, as a result, dry snow conditions were assumed. Fig. 6 presents a comparable plot for snow depth and Nimbus 6 vertically polarized brightness temperature from the daytime pass on 15 March 1976. Analysis of the horizontally polarized data produced essentially similar results with marginally lower and higher R^2 and S_E values, respectively. The snow depth data in Figs. 5

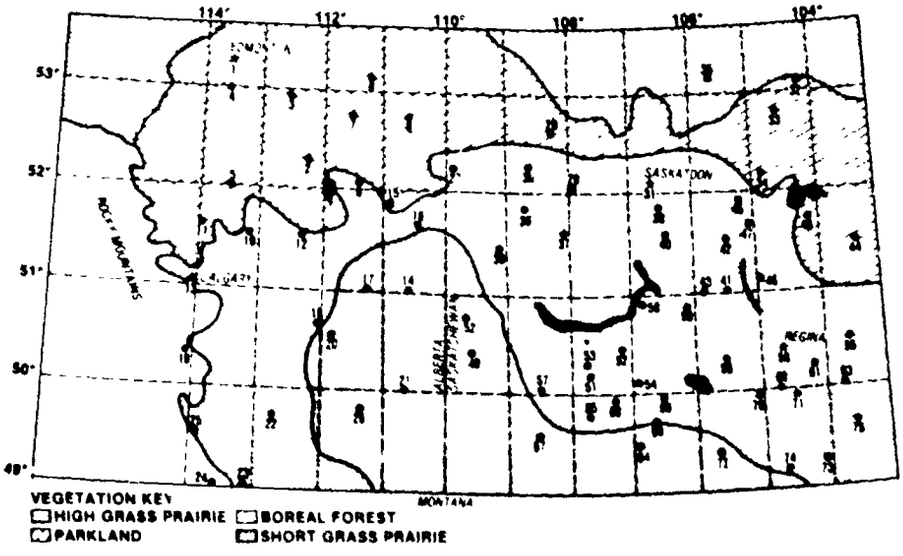


Fig. 4. The Canadian high plains snow/microwave study area.

and 6 are from the same ground measurement so that the figures may be compared. It does appear that in simple regression analysis that the Nimbus 6 data produce better relationships than the Nimbus 5 data. This most likely results from the fact that the emission from the relatively thin snowcover at 1.55 cm contains a more significant contribution from the variable underlying soil layer than at 0.81 cm. Multiple regression approaches using brightness temperatures from Nimbus 5 and 6 as predictor variables provided no marked improvement in the relationship. It appears that in this case a simple linear regression using the Nimbus 6 vertically or horizontally polarized brightness temperature as the predictor variable and snow depth as the criterion variable would be sufficient. S_n data were also available and were used in place of snow depth as the criterion variable. The relationships were similar to those described previously for snow depth, but in general the goodness-of-fit was somewhat less. Fig. 7 shows the data and the resulting regression statistics for the Nimbus 6 vertically polarized brightness temperatures for comparison. The pattern of data points is very similar to Fig. 6 because the snow depth and S_n values are closely related.

The acquisition of concurrent snow survey and satellite brightness temperature data has been difficult, especially for dry snow conditions. Both satellites, because of their research nature, experience significant intervals when data are not available. The ground snow data are not taken continuously and many times are not available during appropriate satellite passes. Furthermore, when both types of data are available, weather conditions causing snowmelt result in significant changes in the relationships previously presented. Free water in the snow causes all brightness temperatures to be markedly increased, in this case as much as 35 K on the Canadian high plains. As a result when temperatures are above freezing for any extended period prior to an otherwise good data set, it has to be eliminated from analysis for at least the dry snow conditions. On the other hand, such an occurrence allows the inference of the effect of free water on the microwave emission. The effect of free water is so strong that the regression relationships can be reversed so that deeper snowpacks (with presumably more free water) can have a greater brightness temperature than that of a shallow wet snowpack or bare soil. This was observed during melting in the Canadian high plains study site.

Examination of snow depth and brightness temperature relationships in other regions such as a site in the Montana-North Dakota area has been accomplished on a preliminary basis. Significant relationships between snow depth and brightness temperature exist as on the Canadian high plains, but there may be difficulties in extrapolating relationships from one area to another. The most probable cause of dissimilarities in the relationships between areas are differences in underlying soil types and moisture conditions and differing climatic influences on the snowpack. Better correspondence between areas may result after more data have been obtained.

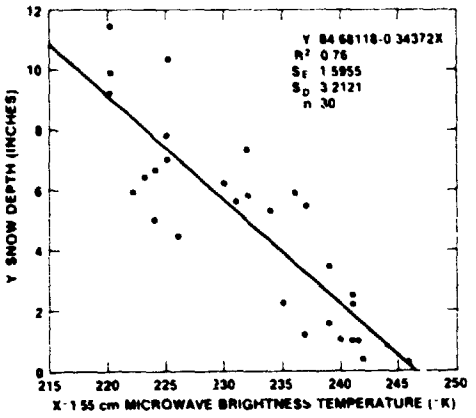


Fig. 5. Nimbus 5 microwave brightness temperature versus snow depth on the Canadian high plains. Nimbus 5 data from nighttime pass 14 March 1976 summarized by one degree latitude-longitude grid; snow depth data from 15 March 1976 summarized over same grid; data included from short and high grass prairie areas only.

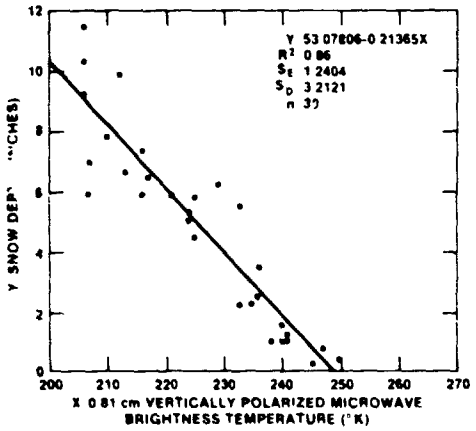


Fig. 6. Nimbus 6 vertically polarized microwave brightness temperature versus snow depth on the Canadian high plains. Nimbus 6 data from daytime pass 15 March 1976 summarized by one degree latitude-longitude grid; snow depth data from 15 March 1976 summarized over same grid; data included from short and high grass prairie areas only.

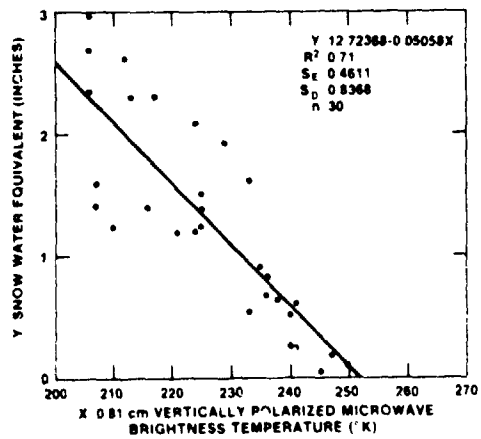


Fig. 7. Nimbus 6 vertically polarized microwave brightness temperature versus snow water equivalent on the Canadian high plains. Nimbus 6 data from daytime pass 15 March 1976 summarized by one degree latitude-longitude grid; water equivalent data from 15 March 1976 summarized over same grid; data included from short and high grass prairie areas only.

Applications

Presently and for the next few years passive microwave resolution from space will be no better than 20-30 km. This will change in 5-10 years as large antenna structures in space become feasible, and resolution should improve to as little as 1 km at the short wavelengths. This projected improvement has a direct bearing on future applications of the data.

The capability for mapping the snowcovered area on a continental basis is now provided by visible sensor satellites such as NOAA. Clouds, however, can be a significant problem in the derivation of these maps. ESMR-type data could be used to supplement the NOAA maps by providing an "all-weather" capability during cloudy intervals. As the resolution improves the microwave sensor may pro ide a better alternative for compiling the operational product.

Currently the accumulation and depletion of snow at established measurement points is best done on the ground. The microwave capability for doing this, however, is best suited for unmeasured and perhaps remote locations. The additional advantage of the microwave approach is that detection of the time of melt and initiation of snowpack priming is also possible. This information, although extremely important for hydrograph timing, is not measured at conventional observation stations because of the difficulty associated with direct measurement. In most cases temperature records are used to infer melting conditions in the snowpack. Thus, the microwave sensing provides a possible new and additional source of snow information with the greatest application most likely for areas lacking temperature measurements.

As had been hypothesized by previous modeling, ground-based, and aircraft experiments, there appears to be a significant relationship between snow depth and brightness temperature which could be used for snow surveys over large areas. Because of poor sensor resolution currently, the use of such relationships is restricted to large, homogeneous regions such as the high plains. Even this could be valuable for runoff prediction purposes during rapid spring melt in these areas once enough representative data sets are collected. Again as resolution improves applications should be found in other zones such as intermountain valleys and large, mountain open areas. Such capabilities would then become more directly applicable to seasonal and short-term runoff forecasts.

The realization of these future potentials requires refinement of the various techniques through acquisition and analysis of additional data. The multispectral, dual polarization microwave capability that exists on the Nimbus 7 satellite in the SMMR instrument appears to be the ideal data source at the moment and should permit better discrimination of snowpack properties. Modifications of the SMMR instrument in 5-10 years will be tested with larger antennas, thus significantly improving the resolution which will further promote utilization of the passive microwave data.

Conclusions

1. Variations in snow accumulation and depletion at specific locations are reflected by related variations in passive microwave brightness temperatures observed from the Nimbus satellites. Qualitative monitoring of snowpack build-up and disappearance during the winter appears feasible.
2. Using both vertically and horizontally polarized brightness temperatures from the Nimbus 6 ESMR instrument, a statistically-oriented discriminant function can be used to separate snow from no-snow areas and map snowcovered area on a continental basis. This "all weather" microwave approach is especially valuable when mapping using visible wavelength sensors is prevented by the presence of clouds. The microwave-based snowcover maps compare favorably with northern hemisphere snow and ice boundary maps published weekly by NOAA.
3. In a relatively homogeneous, flat area, such as the Canadian high plains, a significant regression relationship between snow depth (criterion variable) and microwave brightness temperature (predictor variable) can be developed. The estimation of snow depth under dry snow conditions is thus a possibility in the study area using microwave data; after further data sets are acquired these estimates could be expanded over a greater range of snow depth. Similar specific relationships would have to be derived for use in other snow study areas.
4. The presence of melt water in the snowpack radically changes the microwave emission characteristics resulting in as much as 35 K increases in brightness temperature over dry snow conditions. Such changes allow the monitoring and easy detection of a melting snowpack with direct pertinence to the estimation of the timing of snowmelt runoff.

References

Chang, T. C., and Gloersen, P. (1975) Microwave emission from dry and wet snow, Proceedings of the Workshop on Operational Application of Satellite Snowcover Observations, NASA SP-391, Washington, D. C., pp. 399-407.

Chang, T. C., Gloersen, P., Schmugge, T., Wilheit, T. T., and Zwally, H. J. (1976) Microwave emission from snow and glacier ice, *Journal of Glaciology*, Vol. 16, No. 74, pp. 23-39.

Chang, A. T. C. (1978) Estimation of snow temperature and mean crystal radius from remote multispectral passive microwave measurements, NASA Technical Paper 1251, Washington, D. C., 11 pp.

Chang, A. T. C. and Choudhury, B. J. (1978) Microwave emission from polar firn, NASA Technical Paper 1212, Washington, D. C., 20 pp.

Gloersen, P., and Barath, F. T. (1977) A scanning multichannel microwave radiometer for Nimbus-G and Seasat-A, *IEEE Journal of Oceanic Engineering*, Vol. OE-2, No. 2, pp. 172-178

Hall, D. K., Chang., A., Foster, J. L., Rango, A., and Schmugge, T. (1978) Passive microwave studies of snowpack properties, Proceedings of the 46th Annual Western Snow Conference, Otter Rock, Oregon, pp. 33-39.

Künzi, K. F., and Staelin, D. H. (1975) Measurements of snow cover over land with the Nimbus-5 microwave spectrometer, Proceedings of the Tenth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, pp. 1245-1253.

Martinec, J. (1975) Snowmelt-runoff model for stream flow forecasts, *Nordic Hydrology*, Vol. 6., pp. 145-154.

Mather, P. M. (1976) *Computational methods of multivariate analysis in physical geography*. John Wiley and Sons, New York, 532 pp.

Meier, M. F., and Edgerton, A. T. (1971) Microwave emission from snow – a progress report, Proceedings of the Seventh International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, pp. 1155-1163.

Meier, M. F. (1972) Measurement of snow cover using passive microwave radiation, The Role of Snow and Ice in Hydrology, Proceedings of the Banff Symposium, IAHS-AISH Publication No. 107, Banff, pp. 739-750.

Rango, A. (ed.) (1975) Operational applications of satellite snowcover observations, National Aeronautics and Space Administration, NASA Special Publication SP-391, Washington, D. C., 430 pp.

Rango, A. (1978) Pilot tests of satellite snowcover/runoff forecasting systems, Proceedings of the 46th Annual Western Snow Conference, Otter Rock, Oregon, pp. 7-14.

Schanda, E., Hofer, R., Wyssen, D., Musy, A., Meylan, D., Morzier, C., and Good, W. (1978) Soil moisture determination and snow classification with microwave radiometry, Proceedings of the Twelfth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, pp. 1779-1789.

Shiue, J. C., Chang, A. T. C., Boyne, H., and Ellerbruch, D. (1978) Remote sensing of snowpack with microwave radiometers for hydrological applications, Proceedings of the Twelfth International Symposium of Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, pp. 877-886.

Zwally, H. J. (1977) Microwave emissivity and accumulation rate of polar firn, *Journal of Glaciology*, Vol. 18, No. 79, pp. 195-215.

Zwally, H. J. and Gloersen, P. (1977) Passive microwave images of the polar regions and research applications, *Polar Record*, Vol. 18, No. 116, pp. 431-450.

Received: 2 November, 1978

Address:
Code 913,
Goddard Space Flight Center,
Greenbelt, Maryland 20771,
U S.A.

Paper 98

14328 JANUARY 1979 HY1

JOURNAL OF THE
HYDRAULICS DIVISION

SNOW-COVERED AREA UTILIZATION IN RUNOFF FORECASTS

By Albert Rango,[1] Jack F. Hannaford,[2] M. ASCE, Roderick L. Hall,[3] M. ASCE, Michael Rosenzweig,[4] and A. Jean Brown,[5] M. ASCE

INTRODUCTION

Since 1972 several earth resources and environmental satellites, such as Landsat and National Oceanic and Atmospheric Administration (NOAA) have been launched that have direct application to snow-cover mapping. The characteristics of these satellites and their potentials for snow-cover monitoring and subsequent runoff prediction have been treated previously (4). Although the utilization of snow-covered area (SCA) as an additional parameter in seasonal runoff predictions seems logical and has been shown to be useful (3), the duration of satellite data is too short for conclusive testing of SCA in conventional approaches. In order to expeditiously estimate the potential value of satellite SCA data in runoff predictions, simplified linear multiple regression analyses of longer-term aircraft visual observations of SCA for two watersheds in the southern Sierra Nevada in California were conducted.

Study Area Description.—The Kings and Kern Rivers are adjacent watersheds (Fig. 1) that discharge into the Central Valley near Fresno and Bakersfield, Calif., respectively. Each basin ranges in elevation from below 1,000 ft (300 m) in the foothill areas to over 14,000 ft (4,300 m) along the Sierra Nevada crest that is the eastern boundary for both watersheds. The Kings River has an east-west orientation with high subbasin divides and subbasin drainage in deep canyons. The Kern River, on the other hand, has a north-south orientation with the Sierra crest along the eastern drainage boundary and the similarly

Note.—Discussion open until June 1, 1979. To extend the closing date one month, a written request must be filed with the Editor of Technical Publications, ASCE. This paper is part of the copyrighted Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol. 105, No. HY1, January, 1979. Manuscript was submitted for review for possible publication on January 31, 1978.

[1]Research Hydro., Goddard Space Flight Center, Greenbelt, Md.
[2]Principal, Sierra Hydrotech, Placerville, Calif.
[3]Principal, Sierra Hydrotech, Placerville, Calif.
[4]Acting Chf., Electric Power Analysis Div., Energy Information Administration, Dept. of Energy, Washington, D.C., formerly, Operations Research Analyst, Goddard Space Flight Center, Greenbelt, Md.
[5]Chf., Snow Surveys Branch, California Dept. of Water Resources, Sacramento, Calif.

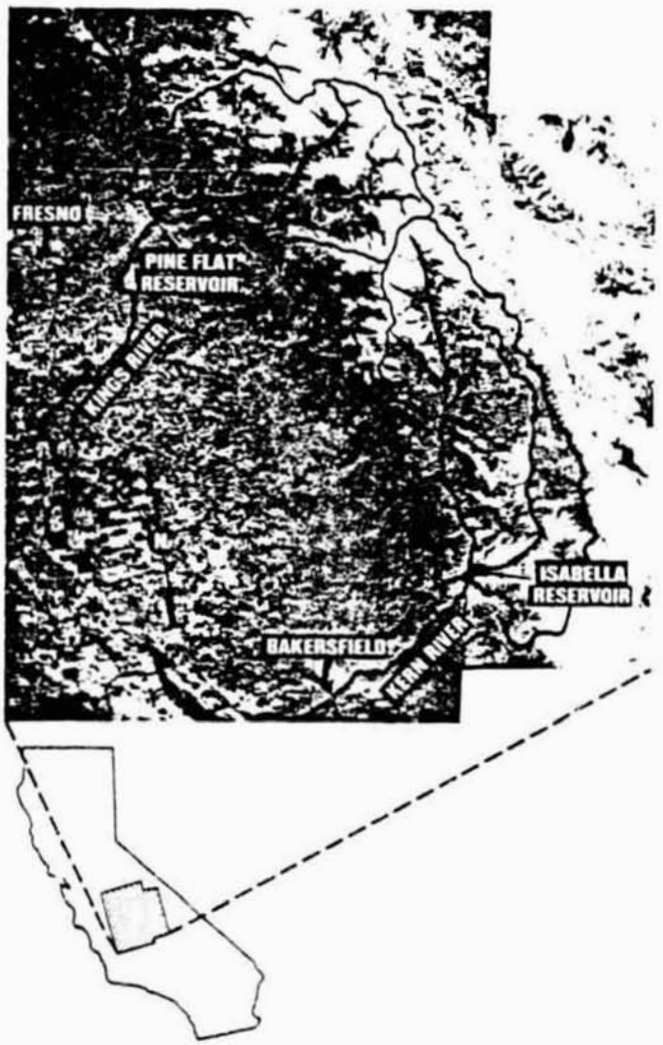


FIG. 1.—Kings and Kern River Watersheds in California as Seen in 0.6-μm to 0.7-μm Channel of Landsat on April 30 and May 1, 1976

ORIGINAL PAGE IS
OF POOR QUALITY

high Great Western Divide along the western boundary of the basin. The Kern River is characterized by plateau areas with broad meadows and timbered slopes, although the North Fork heads in steep rocky areas near the Kings-Kern divide and flows in a deep canyon through most of its length to Lake Isabella. Area-elevation graphs in Fig. 2 show the relatively uniform distribution of area with elevation on the Kings River as contrasted with the concentration of area

TABLE 1.—Seasonal Streamflow for Kings and Kern River Watersheds

Season (1)	APRIL-JULY RUNOFF			
	Kings		Kern	
	In acre-feet (cubic meters) (2)	Percentage of average (3)	In acre-feet (cubic meters) (4)	Percentage of average (5)
1976	301,800 (372,400,000)	26	103,900 (128,200,000)	25
1969	2,631,800 (3.248 billion)	227	1,349,500 (1.665 billion)	321
Average	1,157,000 (1.428 billion)		420,000 (518,000,000)	

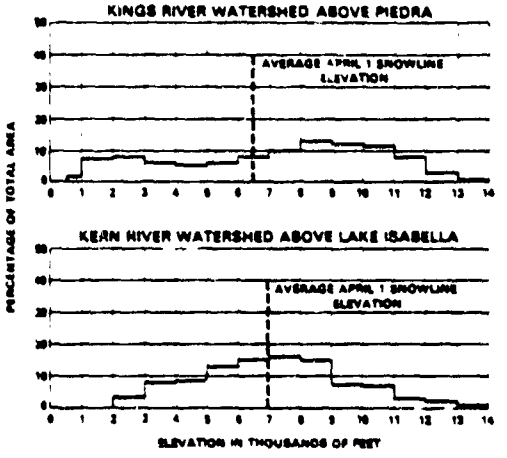


FIG. 2.—Area-Elevation Statistics on Kings and Kern River Watersheds and Location of Average April 1 Snow-Line Elevation (1 ft = 0.305 m)

between 6,000 ft (1,830 m) and 9,000 ft (2,750 m) on the Kern River. The average elevation of the April 1 snowline as taken from California Department of Water Resources (CDWR) records is also shown in Fig. 2 for both watersheds.

The 1,545-sq mile (4,002-km^2) Kings River has an average annual runoff of 1,567,600 acre-ft (1.934 billion m^3) that represents 19 in. (480 mm) of runoff, 74% of which occurs during the April-July snowmelt period. Snowpack accumulation increases with elevation to about 9,500 ft (2,900 m) and is fairly consistent

at about 30 in. (760 mm) of water above that elevation, although local topography may affect accumulation to some extent. Average annual precipitation at the 9,000-ft (2,750-m) elevation is about 35 in. (890 mm). Precipitation measurements made along the frontal slope at the western side of the basin appear to be representative of or at least proportional to precipitation at the higher elevations, although some minor variations may occur.

The 2,074-sq mile (5,372-km^2) Kern River watershed (above Lake Isabella) has an average annual runoff of 626,600 acre-ft (773,200,000 m^3) that represents 5.7 in. (145 mm) of runoff, about 67% of which occurs during the April–July snowmelt. Precipitation varies both with elevation and location in the basin. At 9,000 ft (2,750 m), average annual precipitation along the Great Western Divide exceeds 35 in. (890 mm), while at the same elevation along the Sierra crest precipitation may be as low as 16 in. (410 mm). Precipitation, snowpack accumulation, and snow cover appear much more variable over the Kern basin than over the Kings basin.

Precipitation and resulting runoff are extremely variable from season to season in the southern Sierra, emphasizing the importance and need for an adequate water-supply forecasting program. Table 1 illustrates the wide range of runoff experienced within the recent past.

Historical Water-Supply Forecast.—The CDWR makes water supply forecasts of April–July (snowmelt period) runoff for all major snowmelt streams in California, including the Kings and Kern Rivers. The California Cooperative Snow Survey Program was initiated in 1929 and the first forecasts using snow survey data were issued in 1930. Forecasts are issued by CDWR as *Bulletin 120*, "Water Conditions in California," in four reports stating conditions as of February 1, March 1, April 1, and May 1. The Kings and Kern Rivers, as well as other selected Sierra streams, have weekly updates of water supply forecasts from February 1–July 1. Present methods for updating as the snowmelt season progresses are limited by the quality and type of real-time data available during the melt period. Forecast error tends to be concentrated in the remaining runoff volume. Current forecast procedures are based upon about 45 yr of data to reflect the extreme variability that has been noted from season to season in these basins.

Data Sources.—One of the primary reasons that the Kings and Kern Rivers in the southern Sierra were selected as watersheds for investigation of SCA as a water-supply forecast parameter was availability of historic data on snow-covered area. Since 1952, the U.S. Army Corps of Engineers (USACE) has collected and assembled information on SCA from the Kings, Kern, and several other watersheds. SCA has been mapped from a low flying light aircraft by an observer using topographic maps with suitable landmark identification. Aircraft observations generally started before May 1 and continued periodically until the SCA of the Kings River was depleted to less than 100 sq miles (250 km^2). Most years had only three or four observations, but heavy snow years sometimes had as many as eight observations.

Data on SCA from aircraft observations and satellite imagery were plotted versus time to provide estimates of SCA on specific dates for use in analysis. Only during 1973 were there adequate data from both aircraft and satellite for comparison. During 1974 only one aircraft observation was made and the USACE program was subsequently discontinued. Fig. 3 is a plot of snow-covered area

for the Kings River for 1973 showing both aircraft and Landsat data. Unfortunately, aircraft observations for 1973 were not made by the same personnel who had compiled the earlier data. After reviewing the data with past and present USACE personnel, minor adjustments (using highway and weather station snow depth information plus snow survey water equivalent data) were made in the historic aircraft snow-cover observations to make them more comparable to the satellite observations. There is still a consistently greater snow-covered area observed by Landsat than observed by the aerial surveys. This difference was noted previously (1) and attributed to the fact that aerial surveys excluded lower elevation transient snow cover from their measurements.

The conventional water equivalent (also referred to as water content) data applicable to the Kings and Kern Rivers were obtained by cooperators in the California Snow Survey Program and sent to CDWR. Other pertinent hydrome-

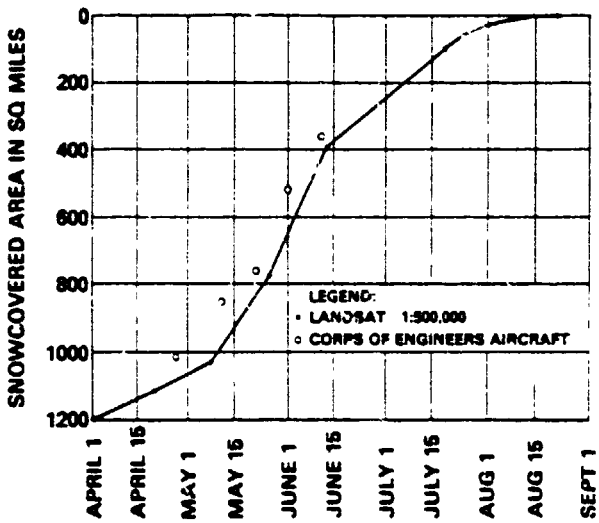


FIG. 3.—Areal Extent of Snow Cover from Satellite and Aircraft Observations for Kings River Watersheds, in 1973 (1 sq mile = 2.59 km^2)

teorological information such as precipitation and runoff records were obtained by CDWR on an operational basis for water-supply forecasting and other purposes. The data are developed by CDWR into basin indices for application to regression equations or multigraphical solutions for predicting April-July runoff. Forecast procedures after April 1 are currently based on the April 1 forecast updated using observed precipitation and limited telemetered automatic snow sensor data. The indices used for development of the April 1 forecast procedures are as follows:

1. Snowpack Index—This index is based upon the observed water equivalent at approximately 20 snow courses in each basin as of April 1. On some basins, including the Kings and Kern, two separate indices are developed for the high and low elevation snow zones, respectively. This index is expressed in terms

of percentage of average, as are most indices in the forecast procedures, and represents the relative quantity of water stored as snowpack on the date of forecast. Adjustment may be made for precipitation occurring between the actual date of measurement and April 1.

2. October–March Precipitation Index—This index, developed from approximately six lower elevation mountain stations, provides an indication of basinwide seasonal wetness.

3. April–June Precipitation Index—This index is a measure of precipitation occurring during the snowmelt season at about six stations and permits a level of updating as the season progresses after April 1. Observed precipitation data are used to replace average precipitation figures as the snowmelt season progresses.

4. October–March Runoff Index—This index relates both to basin wetness and volume of water not stored in the basin as a result of early season runoff.

5. Previous Year Runoff Index—This index is expressed as a volume for the previous April–July and may be related to the carryover from the previous runoff season.

For use in this study, these indices were developed according to the CDWR procedures with data supplied by CDWR. Analysis was performed for the period of record represented by the combined aircraft and satellite observations of snow-cover data, which was 25 yr on the Kings River and 23 yr on the Kern.

Analyses

Independent analyses were undertaken almost simultaneously by National Aeronautics and Space Administration (NASA) and Sierra Hydrotech—CDWR teams utilizing similar basic data to demonstrate the potential effect of SCA in water supply forecasting on the Kings and Kern Rivers. Although the objectives of the two investigators were somewhat different, similar results were obtained in both investigations. The NASA study (Investigation 1) was intended to demonstrate that SCA on a given date is applicable to forecasting seasonal runoff. The Sierra Hydrotech-CDWR study (Investigation 2) was intended to go one step further and to develop and demonstrate a procedure for updating water supply forecasts during the period of snowmelt utilizing SCA as a parameter. Both investigations were exploratory in nature and not intended to represent the most advanced techniques in statistical methods or water-supply forecasting.

Investigation 1

Approach.—An evaluation of conventionally based and SCA-based seasonal runoff predictions on May 1 was made. In this approach only the low altitude estimate of SCA was used in analysis. Although aircraft observations began in 1952 and ended in 1973, observations were not readily available for each watershed in every intervening year. As a result, at the time Investigation 1 was conducted, only 20 yr and 18 yr of aircraft SCA data were initially available for the Kings and Kern River watersheds, respectively. Conventional data were developed into forecast indices only for the years with existing SCA data. The existing forecast procedure used by CDWR was employed as the model for developing the "modified" (reduced data base) conventional regression equation.

in deriving a regression model using SCA for predicting seasonal flow, standard step-wise techniques were first utilized to determine the order of entry of the predictors. Several alternative orders and combinations were then considered to investigate potential reductions in the number of variables required while achieving an acceptable significance level (≤ 0.05). The "modified" conventional model was run against all the statistically acceptable SCA models in a prediction mode, and the various runoff forecast values were compared to the actual runoff figures.

On the May 1 forecast date all data were available except the April-June precipitation index. To simulate a real forecasting situation on May 1, the actual April precipitation was combined with the expected (average) May and June precipitation to obtain the best estimate of the April-June precipitation index.

Both the "modified" conventional and the set of "snow-cover" models were exercised to determine which would provide the better forecast for each watershed. Since the number of available data points was limited and several variables were being considered, a series of regressions was used to make the forecasts. This technique consisted of deleting the forecast year from the data base, deriving the regression equation coefficients from the remaining data, and then making a forecast for that deleted year. The absolute value of the difference between the forecast and the actual runoff represented the error of the forecast. The forecast and forecast error were computed for each year. The average and the standard deviation of the errors were calculated and tabulated for each watershed and the best "snow-cover" model was selected based on minimizing these values.

Results.—The regression model used by the CDWR on the Kings River is of the following form:

$$Y_1 = a_1X_1 + b_1X_2 + c_1X_3 + d_1X_4 + k_1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

in which Y_1 = Kings River April-July runoff, in acre-feet; X_1 = April 1 snowpack index; X_2 = October-March precipitation index; X_3 = previous year April-July runoff; and X_4 = April-June precipitation index. In this "modified" conventional equation the regression coefficients, a_1, b_1, c_1, and d_1, and the regression constant, k_1, are slightly different than their counterparts in the CDWR model because of the reduced data base resulting from the testing of SCA.

The regression model used by the CDWR on the Kern River is of the form:

$$Y_2 = a_2X_1' + b_2X_2 + c_2X_4 + d_2X_1'' + e_2X_5 + k_2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

in which Y_2 = Kern River April-July runoff, in acre-feet; X_1' = April 1 high elevation snowpack index; X_1'' = April 1 low elevation snowpack index; and X_5 = May 1 snowpack index. The regression coefficients and constant in the "modified" conventional equation are again slightly different than those in the current CDWR equation.

On the Kings River the resulting "snow-cover" model for all years of record had the following form:

$$Y_1 = a_3X_4 + b_3X_2X_6 + k_3 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

in which X_6 = May 1 SCA, as a percentage of basin area; $a_3 = 1.18869$; $b_3 = 0.17573$; and $k_3 = 45.58954$. On the Kern River the best alternative model had the following form:

$$Y_3 = a_4X_2X_4 + b_4X_1' + c_4X_3 + k_4 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

in which $a_4 = 0.04332$; $b_4 = 2.54$; $c_4 = 2.02$; and $k_4 = -135.022$. Both final Kings and Kern "snow-cover" models resulted from the step-wise regression analysis.

The statistics for the models on the Kings and Kern Rivers are shown in Table 2. These models were then compared on each watershed by evaluating the difference between actual and forecast runoff (which was assumed to be the forecast error) for each year in the data base. The average yearly forecast errors and the standard deviations of the errors are tabulated in Tables 3(*a*) and 3(*b*) along with the change in forecast parameter resulting from the incorporation of SCA. Although slight increases in forecast error occur when SCA

TABLE 2.—Comparison Statistics for Regression Equations

Statistics (1)	Kings River "modified" conventional model (2)	Kings River snow-covered area model (3)	Kern River "modified" conventional model (4)	Kern River snow-covered area model (5)
Degrees-of-freedom	4, 15	2, 17	5, 12	3, 14
F-test value	161.3*	218.4*	156.0*	355.3*
R^2 value	97.1	95.9	97.9	98.4
Standard error of estimate	120.9	145.7	35.6	30.5
Standard deviation of the seasonal yield	712.5	712.5	243.2	243.2

*Significant at the 0.005 level.

TABLE 3.—Comparison of "Modified" Conventional and Snow-Cover Model Average Forecast Errors, in acre-feet

Error (1)	Conventional model (2)	Snow-cover model (3)	Change, as a percentage (4)
(*a*) Kings River (n = 20)			
Average forecast error	114,900	120,900	+5
Standard deviation of forecast errors	106,400	107,700	+1
(*b*) Kern River (n = 18)			
Average forecast error	40,110	28,670	−29
Standard deviation of forecast errors	25,280	23,310	−8

Note: 1 acre-ft = 1,233 m^3.

is included in the prediction procedures on the Kings River, major reductions in forecast error using SCA are realized on the Kern River.

Investigation 2

Approach.—Investigation of the application of SCA as a parameter in CDWR water-supply forecasting has been limited to the April-July snowmelt period because most watersheds are 100% snow covered before April 1. On the average, only about 10%–15% of the annual precipitation falls after April 1. Prior to April 1, most of the total error in water-supply forecasts is attributed to the uncertainty of the amount of precipitation occurring after the date of forecast. As the snowmelt season progresses, however, procedural error contributes an increasing portion of total forecast error, justifying additional analytical work to improve techniques for correcting or updating the forecast during the snowmelt period.

This approach was predicted on the operational requirements for accurate updating of water-supply forecasts throughout the period of snowmelt runoff. Forecasts prepared by CDWR have historically been for the April–July snowmelt period. Updating has been primarily on the basis of precipitation observed subsequent to the April 1 forecast. Any procedural error in the April 1 forecast would be forced into the forecast of remaining runoff during the melt season. A forecast made on June 1 might contain the same procedural error as the forecast made on April 1, even though half of the snowmelt runoff for the season may have already occurred. The desirability of providing a forecast technique that would reduce the magnitude of procedural error as the season progresses is obvious.

Only a limited amount of data is available from these high mountain watersheds during the period of snowmelt. Precipitation from manned stations and some telemetered stations is available on a daily basis. Snowpack water equivalent measurements on a few snow courses are made about May 1, and some continuous snow sensor records are available, but data are limited. Additionally, the melt process during April introduces uncertainty into the meaning of observed water equivalent at specific locations. May 1 measurements have been used with some success in the Kern basin to reflect precipitation and melt occurring during April.

Observed runoff and depletion of SCA as the melt season progresses provide additional parameters on a near real-time basis to reflect the progress of melt in the watershed. This investigation developed and demonstrated techniques for updating conventional CDWR forecast procedures during the progress of melt. Forecast procedures were developed for April 1, May 1, May 15, June 1, and June 15 for the Kings and Kern River basins. The use of Landsat SCA data for 1973–1976 and previous aircraft observations available for Investigation 2 provided 25 yr of record on the Kings and 23 yr of record on the Kern for analysis. Procedure stability was an important factor to assure a logical sequence of operational forecasts during the progress of the season.

Basic data utilized in the conventional CDWR procedure were used to prepare the April 1 procedure. Two procedures were then developed for May 1 and each subsequent date, one with and one without SCA to observe the effect of SCA in improving forecast reliability. In both procedures, runoff between April 1 and date of forecast was used as an additional parameter. Since existing

CDWR procedures have techniques for handling precipitation during the snowmelt season, precipitation subsequent to date of forecast was assumed known and does not contribute to "procedural error" described in the analysis.

The general form of the forecast procedure equation is

$$Y_3 = a_3X' + b_3X'' + c_3X_2 + d_3X_3 + e_3X_4 + f_3X_7 + g_3X_8 + h_3X'X_7 + k_3 \quad . \quad . \quad . \quad (5)$$

in which Y_3 = basin runoff, in acre-feet, from date of forecast through July 31; X' = high snow index; X'' = low snow index; X_7 = October–March runoff; X_8 = runoff April 1 through date of forecast; and X_9 = SCA, in square miles. Regression coefficients are represented by a_3–h_3 and k_3 represents the regression constant. The conventional April 1 procedures use X', X'', X_2, X_3, X_4, and X_7. Procedures for other times use X_8 or X_8 and X_9, depending upon whether SCA is to be included or not. The SCA times April 1 snowpack index (adjusted for precipitation between April 1 and date of forecast) was used as an index of the volume of water available for snowmelt runoff during the melt period. Constraints on time and period of record did not permit investigation of more complex nonlinear analysis techniques, and exploration of hydrologic models used in water supply forecasting was not justified at this time.

Employing techniques presently utilized by CDWR, "forecasts" were made for each year of record and compared to observed runoff. Because of the limited data set, independent test data were not available and forecasts were made using data employed in derivation of the regressions. Although not statistically acceptable, the intention here was only to see if the SCA was a predictor worth pursuing for runoff prediction techniques. If the answer to this question is positive, more rigorous techniques would be used to incorporate SCA into operational procedures. Standard errors and other pertinent statistical measures were calculated for each date of forecast and results with and without SCA as a parameter were then compared, recognizing the limitations of these simple regression techniques.

Results.—Fig. 4 shows the variation in standard error, expressed as a percentage of April–July runoff, for forecast updates, depicting the effective reduction in forecast error as snowpack is depleted. Updating procedures with SCA are shown as a dashed line while updating procedures utilizing SCA are shown as a solid line. Fig. 5 shows the same variation in standard error, expressed as a percentage of remaining snowmelt runoff for forecast updates. The dashed and solid lines represent standard error of procedures without SCA and with SCA, respectively. The dotted lines represent error remaining if the procedure were used according to standard CDWR practice at the current time. In interpretation of Figs. 4 and 5, it should be noted that although procedural error, in acre-feet, remains constant throughout the period, it will increase in terms of percentage of remaining runoff as the melt season progresses.

On the Kings River (Fig. 4), standard error increases slightly between April 1 and May 1, probably as a result of additional forecast parameters used on May 1 that increase degrees-of-freedom lost. After May 1, standard error declines appreciably until on June 15 it is approx 70% of the error on April 1. This reduction in error is expressed in terms of percentage of remaining runoff in Fig. 5. The improvement over the existing procedure is apparent. The addition of SCA as a parameter, however, seems to offer little or no significant improvement in procedural error during the melt season.

On the Kern River (Fig. 4), standard error for the procedure without SCA follows approximately the same pattern as on the Kings. If SCA is included, however, substantial reduction in standard error is apparent as the season progresses. By including SCA as a parameter, May 1 error is reduced approx 45% and May 15 error about 40%, representing a corresponding decrease in the volumetric error of remaining runoff. The late season values of standard error on the Kern and the Kings are now relatively close. It is suggested that the use of SCA as a forecast parameter during the snowpack depletion period has allowed forecast accuracy on the two watersheds to be brought more into line with each other than possible with conventional parameters alone. The reduction in terms of percentage of remaining runoff is depicted in Fig. 5. Further inspection of changes of regression coefficients from date to date suggests that the Kern River equations are relatively stable—more so than those on the Kings River. Even though the precise numerical value of decrease in procedural error to

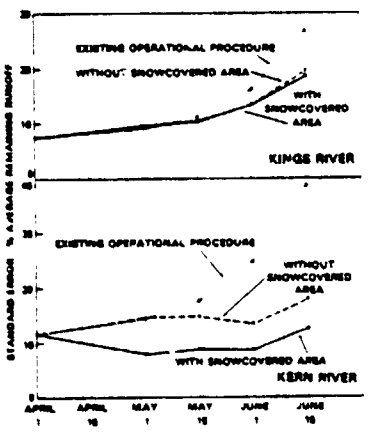


FIG. 4.—Standard Error of Forecast Procedure (with and without Snow-Covered Area) Versus Date during Snowmelt

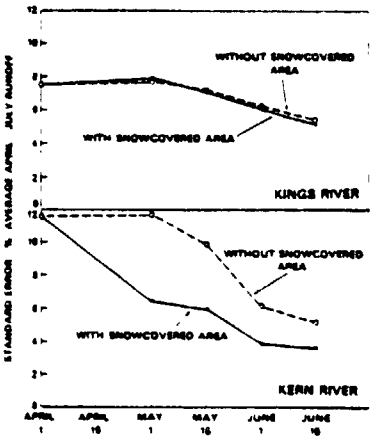


FIG. 5.—Standard Error of Various Forecast Procedures Versus Date during Snowmelt

be obtained by using these methods can not be generalized for all watersheds, it is apparent that SCA provides information pertinent to updating forecasts that is not readily available from other sources investigated here.

Examination of Results

Use of SCA as a parameter in forecasting snowmelt runoff may result in significant improvement of forecasting procedures under certain circumstances. The SCA in Investigation 1 reduced the average May 1 forecast error by 29% and the standard deviation of forecast error by 8% on the Kern River, but appeared to have no substantial or significant effect on the Kings River. Similarly, under Investigation 2, there appeared to be considerable improvement for each update on the Kern River using SCA, but no significant changes on the adjacent

Kings River. It may be hypothesized that watershed characteristics, as well as availability of data representative of the watershed, may be related to the response of forecast procedures to SCA. The following is a review of factors that may influence the effectiveness of SCA as a parameter in water supply forecasting.

The conventional April 1 forecast procedure for the Kings River is relatively more accurate (when expressed in terms of percentage of April–July runoff) than is that for the Kern River. April 1 procedural standard error represents about 7.5% of average April–July runoff on the Kings River and about 11.5% on the Kern, assuming that precipitation after April 1 is known. The higher degree of accuracy for the Kings River procedure may result partially from greater unit runoff and data that are more representative of conditions within the watershed. In any event, the higher initial degree of accuracy on the Kings River may make it considerably more difficult to obtain a marked improvement as a result of SCA or other update parameters as the season progresses.

The relatively inconsistent relationship between precipitation, snowpack accumulation, elevation, and location within the Kern River watershed described previously may be one of the more important reasons why SCA represents an effective parameter in updating Kern forecasts. The Kings River has a much more uniform area-elevation distribution than the Kern River (Fig. 2). The relatively large area between 6,000 ft–9,000 ft (1,830 m–2,750 m) on the Kern River is subject to extreme variability in snowpack accumulation and depletion, perhaps enhancing the value of SCA as a prediction parameter. It might be visualized that the Kings River consists of a number of smaller basins somewhat similar in character and can be predicted well with a forecast procedure representing basins of that character. The Kern River, on the other hand, consists of a number of smaller basins of diverse character. It might be possible to break the Kern area into a number of subbasins and forecast each subbasin independently. The SCA may provide an attractive intermediate solution to water-supply forecasts in areas with inhomogeneous characteristics and limited hydrologic data.

Most watersheds in the central and southern Sierra appear to be quite homogeneous from a hydrologic standpoint, more so than perhaps most other western United States watersheds. Northern Sierra, eastern Sierra, and other watersheds in California, however, appear much more diverse than the Kings River and those watersheds immediately to the north of the Kings, suggesting that SCA might prove to be an effective parameter for water-supply forecasting in California. Watersheds with a substantial degree of area within a limited elevation range, an erratic precipitation and snow accumulation pattern not strongly related to elevation, and poor coverage with precipitation data or stations that do not give a reliable index to the water producing areas of the basin may show the greatest response to use of SCA as a parameter in volumetric forecasting.

Even though the Kings River did not appear to respond significantly to use of SCA in water-supply forecasting in this preliminary investigation, one should not discount possible applications on streams similar to the Kings River. The SCA on the Kings River has been found useful in hydrologic modeling of daily snowmelt and runoff (2). Hydrologic modeling procedures are used in some operational forecasting, and it is hoped that near real-time satellite imagery

may prove to be useful for these types of predictions.

Procedures for updating the remaining volume of snowmelt runoff using SCA will be used operationally on the Kings and Kern Rivers during 1978. In order to assure widespread use of SCA-derived operational forecasts, however, it will be necessary to receive SCA information on a regular near real-time basis (≤72 hr). The possibility of cloud cover during a Landsat overflight may at times result in an 18-day or greater interval between observations. Some type of alternative observational capability, such as NOAA or aircraft SCA estimates, during such periods may be required.

Conclusions

1. Long-term SCA data from aircraft and satellite observations have been shown to be useful in reducing seasonal runoff forecast error on the Kern River watershed when incorporated into water-supply forecast procedures. Both one-time and regular updates of forecasts were improved using SCA. Similar analysis on the Kings River indicated that SCA-produced forecasts were generally as good as conventional forecasts but no significant improvement was noted.
2. Comparison of the Kings and Kern River watersheds indicates that certain watershed conditions may enhance the usefulness of remotely sensed SCA data. The SCA will most likely reduce forecast procedural error on watersheds with: (a) A substantial part of the area within a limited elevation range; (b) an erratic precipitation or snowpack accumulation pattern, or both, not strongly related to elevation; and (c) poor coverage by precipitation stations or snow courses restricting adequate indexing of water-supply conditions.
3. Assuming that operational acquisition and delivery problems associated with space information will be resolved, accumulated satellite date should provide a means for enhancing operational seasonal streamflow forecasts for areas that depend on snowmelt-derived water supplies. In many cases, satellite-acquired SCA data can provide for much more objective, uniform, and controlled information than that possible from aircraft platforms.

Acknowledgments

This study was partially funded by NASA under contract No. NAS5-22957. Valuable data for the study were kindly furnished by the U.S. Army Corps of Engineers, Sacramento District, Sacramento, Calif., and the California Department of Water Resources, Snow Surveys Branch, Sacramento, Calif. Susan Howard, formerly at Goddard Space Flight Center (GSFC), Greenbelt, Md., and Betsy O'Brien of GSFC provided important data analysis support for Investigation 1.

Appendix I.—References

1. Barnes, J. C., and Bowley, C. J., "Handbook of Techniques for Satellite Snow Mapping," *Final Report NAS5-21803*, National Aeronautics and Space Administration Goddard Space Flight Center, Greenbelt, Md., 1974.
2. Hannaford, J. F., "Application of Satellite Imagery to Hydrologic Modeling Snowmelt Runoff in the Southern Sierra Nevada," *Final Report NAS5-22957*, National Aeronautics and Space Administration Goddard Space Flight Center, Greenbelt, Md., 1977.

3. Leaf, C. F., "Areal Snow Cover and Disposition of Snowmelt Runoff in Central Colorado," *Research Paper RM-66*, United States Department of Agriculture Forest Service, Fort Collins, Colo., 1971.
4. Rango, A., and Itten, K. I., "Satellite Potentials in Snowcover Monitoring and Runoff Prediction," *Nordic Hydrology*, Vol. 7, 1976, pp. 209–230.

APPENDIX II.—NOTATION

The following symbols are used in this paper:

a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5, c_1, c_2, c_4, c_5, d_1, d_2, d_3, e_2, e_3, f_3, g_3, h_3 = regression coefficients;
k_1, k_2, k_3, k_4, k_5 = regression constants;
SCA = snow-covered area;
X_1 = April 1 snowpack index;
X_1' = April 1 high elevation snowpack index;
X' = high elevation snowpack index;
X_1'' = April 1 low elevation snowpack index;
X'' = low elevation snowpack index;
X_2 = October–March precipitation index;
X_3 = previous year April–July runoff;
X_4 = April–June precipitation index;
X_5 = May 1 snowpack index;
X_6 = May 1 SCA, as a percentage of basin area;
X_7 = October–March runoff;
X_8 = runoff April 1 through date of forecast;
X_9 = SCA, in square miles;
Y_1 = Kings River April–July runoff, in acre-feet;
Y_2 = Kern River April–July runoff, in acre-feet; and
Y_3 = basin runoff, in acre-feet, from date of forecast through July 31.

14326 SNOW-COVERED AREA USE IN RUNOFF FORECASTS

KEY WORDS: Aerial photography; Aircraft; Forecasting; Remote sensing; Runoff; Satellites (artificial); Snow; Water resources

ABSTRACT. Long-term data on snow-covered area from aircraft and satellite observations have been investigated for application to water-supply forecasting in California's southern Sierra Nevada Mountains. These observations have proven useful in reducing seasonal runoff forecast error on the Kern River watershed when incorporated into procedures to update water-supply forecasts as the melt season progresses. Similar use of snow-covered area on the Kings River watershed produced results that were about equivalent to methods based solely on conventional data. Snow-covered area will be most effective in reducing forecast procedural error on watersheds with: (1) A substantial amount of area within a limited elevation range; (2) an erratic precipitation or snowpack accumulation pattern, or both, not strongly related to elevation; and (3) poor coverage by precipitation stations or snow courses restricting adequate indexing of water-supply conditions. When satellite data acquisition and delivery problems are resolved, the derived snow-cover information should provide a means for enhancing operational streamflow forecasts.

REFERENCE: Rango, Albert, and Hannaford, Jack F., "Snow-Covered Area Utilization in Runoff Forecasts," *Journal of the Hydraulics Division*, ASCE, Vol. 105, No. HY1, Proc. Paper 14326, January, 1979, pp. 53-66

Paper 99

N81-21700

Nordic Hydrology 10, 1979, 225-238

No part may be reproduced by any process without complete reference

Application of a Snowmelt-Runoff Model Using Landsat Data

A. Rango and J. Martinec
Goddard Space Flight Center, Greenbelt, Maryland, U.S.A.,
and
Federal Institute for Snow and Avalanche Research
Weissfluhjoch/Davos, Switzerland

The snowmelt-runoff model developed by Martinec (1975) has been used to simulate daily streamflow on the 228 km^2 Dinwoody Creek basin in Wyoming, U.S.A. using snowcover extent from Landsat and conventionally measured temperature and precipitation. For the six-month snowmelt seasons of 1976 and 1974 the simulated seasonal runoff volumes were within 5 and 1%, respectively, of the measured runoff. Also the daily fluctuations of discharge were simulated to a high degree by the model. Thus far the limiting basin size for applying the model has not been reached, and improvements can be expected if the hydrometeorological data can be obtained from a station inside the basin. Landsat provides an efficient way to obtain the critical snowcover input parameter required by the model.

Introduction

A snowmelt-runoff model has been developed on the basis of experimental measurements in two small mountain watersheds of central Europe (Martinec 1975). In the present study, the model was applied to simulate the snowmelt-runoff in a basin significantly larger than those on which it was developed. The changing areal extent of the seasonal snowcover is an essential variable in this procedure. Progress in satellite remote sensing (Rango and Itten 1976) has enabled aircraft photographs to be replaced by a more efficient monitoring of

snowcover by Landsat. Another aim of this study was to test the use of the snowmelt-runoff model under conditions of a normal hydrometeorological network, i.e., a low density of measurement points usually at low elevation stations, in contrast to a well-equipped experimental or representative basin. The model seems particularly amenable to these kinds of data because it is relatively simple and requires as input data only the extent of snowcover, temperature (degree days), and precipitation.

Related Research

Earth resources and environmental satellites have provided a means for the timely, efficient and accurate monitoring of watershed snowcovered area. Rango and Itten (1976) describe several techniques available for analysing the data ranging from simple photointerpretation to automated digital methods. The most appropriate method should be selected based on required accuracy, turn-around time, existing remote sensing expertise, and facilities available. Rango and Itten (1976) indicate that snowcovered area measurements from Landsat are as accurate as aircraft surveys, less expensive and hazardous, and can be used to cover much more extensive areas.

The satellite-derived snow extent data have shown significant promise for use in seasonal streamflow estimation. Rango, Salomonson, and Foster (1977) used meteorological satellite snow extent data to derive a regression relationship between early April snowcovered area and April-June seasonal yield on the Indus River in Pakistan. In large data sparse areas such as these, the snowcovered area data may be the only available hydrologic information. Thompson (1975) in Wyoming found that the snowcovered area on a particular date was better related to the runoff accumulated up to the date of the Landsat pass expressed as a percentage of the seasonal runoff than to just the seasonal runoff. Such relationships are not only useful for volumetric flow forecasts but also for short term estimates of time distribution of seasonal runoff. Long-term snowcovered area data from aircraft and, subsequently, from Landsat have been found to be useful in reducing seasonal runoff forecast error when incorporated into procedures to update water supply forecasts as the melt season progresses (Rango, et al. 1979). This study found that snowcovered area will be most effective in reducing forecast error on watersheds with substantial area in a limited elevation range, erratic precipitation or snowpack accumulation patterns, and poor coverage by precipitation stations or snow courses.

Regarding the use of snowmelt runoff models for daily flow simulation Viessman (1968) states that the areal extent of the snowpack must be known as this relates to the potential snowmelt contribution from a basin and to initial ground conditions which materially affect the disposition of melt water or rainfall. Models such as by Martinec (1975) or the Streamflow Synthesis and Reservoir Regulation (SSARR) model (U.S. Army Corps of Engineers 1975) recognize this

and call for direct input of snowcovered area. The effects of elevation in relation to snowcovered area have been considered by various investigators. Martinec (1970) provides the option of treating snowcover in 500 m elevation zones, whereas the SSARR model has been modified by Speers, Kuehl, and Schermerhorn (1979) for snowmelt runoff simulation utilizing up to 20 elevation bands. The Swedish HBV model also has the capability of handling snowcover in 10 different elevation zones in a basin, assuming uniform snow conditions within each band (Bergström 1979). Several hydrologic models, although not originally requiring snowcover input, have been modified to accept satellite snow extent data for the generation of daily discharge values (Leaf 1975; Hannaford 1977).

Characteristics of the Test Site

Previous applications of the snowmelt-runoff model have occurred on a variety of small experimental watersheds in Europe. They are shown in Fig. 1 in comparison with the watershed used in this study, the Dinwoody Creek basin in west central Wyoming, U.S.A. Dinwoody Creek is in the Rocky Mountains, whereas Modrý Důl is in the Krkonoše Mountains, Lago Mar in the Pyrenees, Lainbachtal in the Bavarian Alps, and the Dischma basin in the Swiss Alps.

Dinwoody Creek is located in the portion of the Rocky Mountains referred to as the Wind River Mountains where the range of elevation is from about 2,000 m to 5,000 m. Two major rivers rise out of the Wind River Range, namely, the

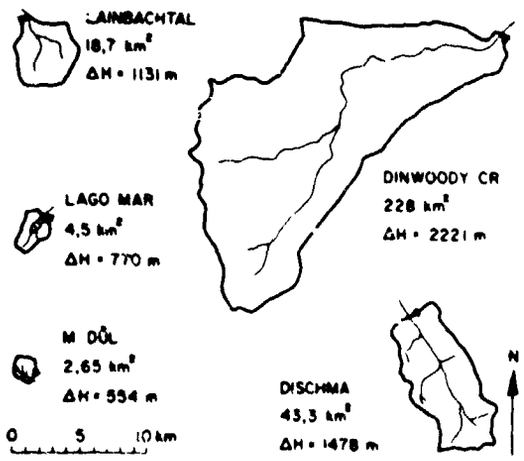


Fig. 1. Area and elevation range (ΔH) of basins in which the snowmelt-runoff model (Martinec 1975) has been applied.

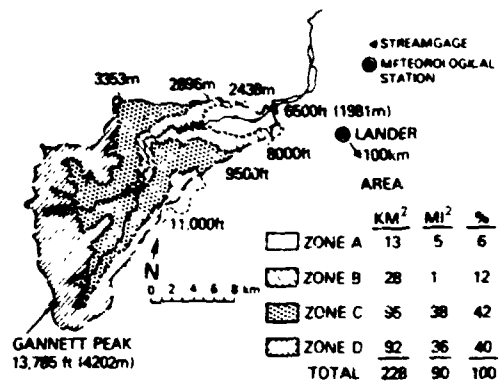


Fig. 2. Elevation zones and areas of the Dinwoody Creek basin.

Green and Wind Rivers, and flow diversions for irrigation are numerous; only in the extreme headwaters or on small tributary streams do relatively unimpaired records exist. Dinwoody Creek is a tributary to the Wind River and flows from southwest to northeast. As shown in Fig. 2 the streamgage is at 1,981 m, and the highest point is 4,202 m.

The Dinwoody Creek basin is typical of many in the western United States in that although it has a high water yield from melting snow no conventional hydrometeorological measurements are made inside the basin. Hourly temperature and precipitation are measured by the National Weather Service in the valley at Lander, Wyoming at an elevation of 1,696 m and 100 km from Dinwoody Creek, and these data were used for input to the snowmelt-runoff model. The U.S. Soil Conservation Service measures snow depth and density at 3,100 m inside the basin on the 1st of each month from February to May. Streamflow is measured year-round at the streamgage by the U.S. Geological Survey. The areal extent of snowcover used in the study was extracted from all available 0.6-0.7 μm band

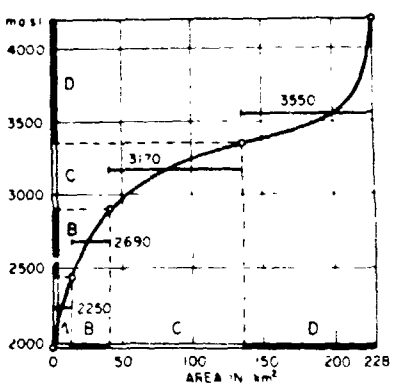


Fig. 3. Area-elevation curve of the Dinwoody Creek basin with the hypsometric mean altitude of each zone shown.

Fig. 4. Landsat image of the Dinwoody Creek basin on 14 May 1976.

images from Landsat and evaluated separately for the elevation zones shown in Fig. 2. Fig. 3 shows the area-elevation curve for Dinwoody Creek with the hypsometric mean elevation of each zone indicated. Fig. 4 is a Landsat image of the Dinwoody Creek basin that was used for snow mapping as viewed through a mask of the watershed boundary.

Application of the Model in the Dinwoody Creek Basin

The deterministic approach used in the development of the model facilitates its application in new conditions. For the Dinwoody Creek basin, it was merely found necessary to increase the number of elevation bands to four as shown in Fig. 2 and to take into account a longer time lag in comparison with the Dischma basin. The model equation (Martinec 1975) was thus rearranged as follows:

$$Q_{n+1} = c_n\Big\{[a_{An}(T_n + \Delta T_{An})\,S_{An} + P_{An}]\,\frac{A_A \cdot 10^{-2}}{86400} + [a_{Bn}(T_n + \Delta T_{Bn})\,S_{Bn} + P_{Bn}]\,\frac{A_B \cdot 10^{-2}}{86400} + [a_{Cn}(T_n + \Delta T_{Cn})\,S_{Cn} + P_{Cn}]\,\frac{A_C \cdot 10^{-2}}{86400} + [a_{Dn}(T_n + \Delta T_{Dn})\,S_{Dn} + P_{Dn}]\,\frac{A_D \cdot 10^{-2}}{86400}\Big\}$$

$$(1 - k_{n+1}) + Q_n\,k_{n+1} \qquad (1)$$

where

Q is the average daily discharge [m^3s^{-1}]
c_n is the runoff coefficient
a_n is the degree-day factor [$cm \cdot ℃^{-1} \cdot d^{-1}$]
T_n is the measured number of degree-days
ΔT_n is the correction by the temperature lapse rate [℃·d]
S_n is the snow coverage (100% = 1.0)
P_n is the precipitation contributing to runoff [cm]
A is the area [m^2]
k_n is the recession coefficient
n is an index referring to the sequence of days
A, B, C, D as subscripts refer to the four elevation zones
$\frac{10^{-2}}{86400}$ converts cm·m^2 per day to m^3s^{-1}

Values determined for certain model parameters are pertinent throughout the entire snowmelt season. The watershed area of Dinwoody Creek as obtained from topographic maps was 228 km^2 with 13 km^2 for zone A, 28 km^2 for zone B, 95 km^2 for zone C, and 92 km^2 for zone D. The recession coefficient is determined from the equation

$$k = 0.884\,Q^{-0.0677} \qquad (2)$$

where Q is discharge in m^3sec^{-1}.

The equation was derived by plotting Q_n against Q_{n+1} for recession flow cases for 1973-1976. From the envelope of points defined it is possible to determine k for any desired Q. By plotting k versus Q on log paper, a straight line can be defined and its equation determined (Fig. 5). It is evident that k increases with decreasing Q. The result for Dinwoody Creek confirms the experience of k generally being greater in larger basins. From analysis of the hydrographs, the time lag for Dinwoody Creek is approximately 18 hours, i.e., snowmelt runoff

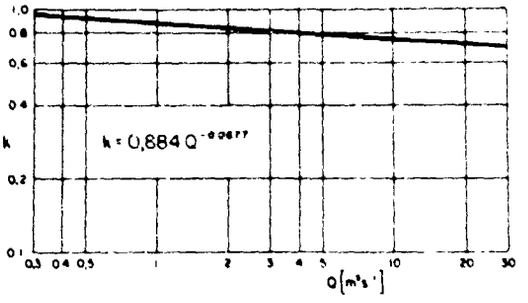


Fig. 5. Relation between the recession coefficient k and the discharge Q for Dinwoody Creek.

from the basin starts rising at about midnight so that temperatures from 0600 to 0600 hours correspond to discharge from 2400 to 2400 hours.

Other model parameters logically change throughout the snowmelt season and can be adjusted every 15 days, if necessary. The degree day factor, a was shown by Martinec (1960) to be related to the relative snow density, ρ, by the equation $a = 1.1\rho$. Thus the general seasonal increase in ρ could be used as an index for the increase in a. The Dinwoody Creek basin a was gradually increased from 0.35 on 1 April to 0.60 on 30 September in zone D. In view of the length of the snowmelt season, seasonal variations of other parameters have to be evaluated. Based on sparse information on climate (Barry and Chorley 1970), the temperature lapse rate was estimated to be higher in the Wind River Mountains than in the Alps and to vary from 0.85°C per 100 m in April to 0.95°C per 100 m in July to 0.80°C per 100 m in September. A lack of direct measurements prohibits the actual determination of the lapse rate. Regional differences between the meteorological station at Lander airport and the mountainous basin were accounted for by subtracting up to 2°C from the Lander data. The runoff coefficient was also assumed to vary during the season; it was estimated in the range from 0.85 in April to 0.75 in July to 0.90 in September.

Three variables need to be currently assessed for model calculations of daily snowmelt runoff, namely, snowcovered area, temperature (degree days), and precipitation. Landsat 0.6-0.7 μm images are used at a scale of 1:1,000,000 with a zoom transfer scope which allows registration and interpretation of the snowcover on Dinwoody Creek at 1:250,000 scale. The snowline is traced across the entire watershed and then the snowcovered area is planimetered manually in each of the four elevation zones. The points for each Landsat pass are used to construct a temporal snowcover depletion curve for each zone. Once the depletion curves for the April-September period are drawn, the daily snowcover values are read off and substituted into Eq. (1).

Air temperature expressed in degree days is used in this simple model as an index of snowmelt. For the Dinwoody Creek basin the Lander airport temperature

data were used. The number of degree days for each 24 hour period is determined by summing the hourly temperatures and dividing by 24 and using 0°C as the base temperature. Temperatures below the freezing point are regarded as 0°C. Maximum and minimum temperatures could also be used in this degree day determination. The degree day figures refer to the 24 hour periods starting at 0600 hours. These temperature data are extrapolated to the hypsometric mean elevation of the respective elevation zone by the previously discussed temperature lapse rate. The resulting degree days are used for calculating snowmelt. Extrapolation errors would be minimized if the temperature was measured in the basin and near the mean elevation.

Daily precipitation amounts at Lander were employed to satisfy the model input requirements. Lacking an acceptable method for extrapolating the precipitation data both horizontally and vertically, the Lander data were used as zonal inputs as recorded. Again, measurement of precipitation in the basin would greatly aid in the application of the model for snowmelt runoff simulation. For the final snowcovered area determinations, the sequence of precipitation events at Lander was used for identification of late season transient snowfall temporarily causing an increase in snowcovered area but not contributing to the snowmelt hydrograph.

Once all the necessary input data were prepared on a daily basis, Eq. (1) was used to calculate snowmelt depths by zones and to transform these values to runoff by the previously mentioned recession techniques. In the Dinwoody Creek basin snowmelt starts about 1 April and continues well into September. The model was used to simulate daily streamflow from 1 April-30 September for 1976 and 1974. Specific parameter values decided on before running the model were used throughout. An optimization of these values with the aim to improve the agreement of the computed and measured runoff did not appear necessary. Model runs both with no updates through the entire period and with runoff updates every two months were performed. This updating provided certain improvements in simulations, but because possible application on ungaged watersheds is of considerable interest, only simulations with no updating will be discussed. Simulated flow was compared with discharge measured at the U.S. Geological Survey streamgage both on a seasonal and daily basis. Total volume differences for the six month period were compared on a percentage basis. To facilitate comparison of the daily discharge amounts a nondimensional «goodness of fit» function was used as proposed by Nash and Sutcliffe (1970) in the following equation:

$$R^2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(q_i-\bar{q})^2 - \frac{1}{n}\sum_{i=1}^{n}(q_i-q_i')^2}{\frac{1}{n}\sum_{i=1}^{n}(q_i-\bar{q})^2} \qquad (3)$$

where

R^2 is a measure of model efficiency
q_i = observed discharge
q_i^* = simulated discharge
$\bar{q}$ = mean of observed discharge
n = number of discharge values

The measure of model efficiency that they propose (R^2) is analogous to the coefficient of determination and is a direct measure of the proportion of the variance of the recorded flows explained by the model (Kite 1975).

Results

The changing areal extent of the seasonal snowcover monitored by Landsat is important information for the snowmelt-runoff model. Fig. 6 shows the zonal depletion curves of the snow coverage in 1976. In plotting these curves it is advisable to bear in mind their characteristic shape (Leaf 1967) and to disregard short term deviations which may be caused by occasional snow storms in the summer. This new snow is taken into account as precipitation contributing to runoff on the first melting day after the snowstorm.

With parameters and variables determined on each day as desribed in the previous section, a day-to-day simulation of the runoff was carried out. A temporal comparison with the discharge measured by the U.S. Geological Survey at the outlet of the basin is illustrated in Figs. 7 and 8. Since the simulation was

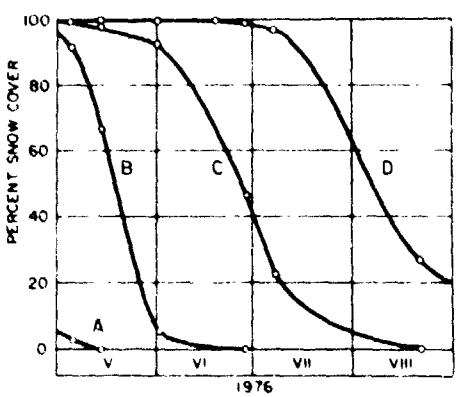


Fig. 6. Depletion curves of snow coverage in the Dinwoody Creek basin in elevation zones A, B, C, and D for May – August 1976.

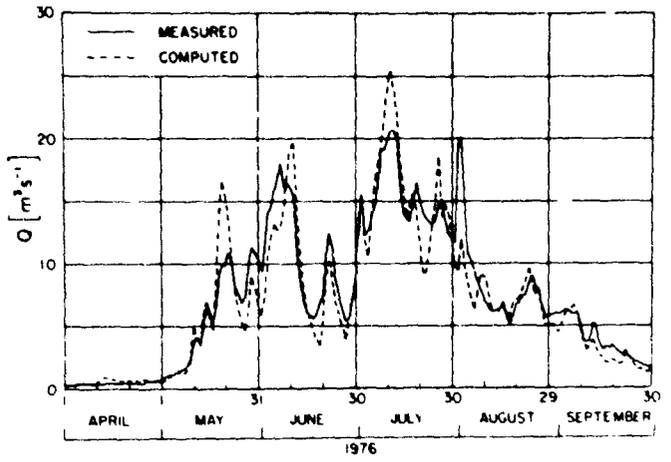


Fig. 7. Simulated and actual streamflow for Dinwoody Creek for April - September 1976.

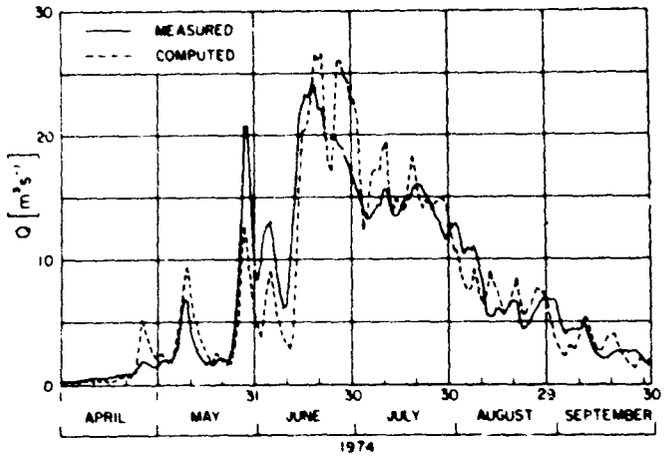


Fig. 8. Simulated and actual streamflow for Dinwoody Creek for April - September 1974.

continued for 6 months without any updating by the actual discharge, these examples indicate possibilities of the runoff simulation in ungaged basins by using Landsat data and temperature measurements.

Comparison on a seasonal basis for 1976 indicates a 5% difference between the 6 month volumes for the computed and measured flows. For 1974 this volumetric difference is only about 1%. Such differences are quite reasonable when compared to conventional simulation procedures. The daily differences in simulated and actual discharge were evaluated using R^2 Eq. (3). For 1976 R^2 was

0.86 and for 1974 it was 0.83. This indicates that about 85% of the variation in the actual daily runoff values is explained by this modeling approach.

The agreement between the simulated and measured runoff could be improved if temperature and precipitation were measured inside the basin. Although there is a procedure for extrapolating the Lander temperature data to the Dinwoody Creek basin it does not take into account the inherent climatic differences between the valley and mountain locations. The uncertainties of temperature extrapolations can be reduced by installing inexpensive automatic meteorological stations (Strangeways and McCulloch 1965) in the areas of interest. The precipitation values assumed for Dinwoody are probably in even greater error because of the same climatic differences. As an example, Johanson (1971) showed that as the precipitation gage density increased from a single gage for a 2,500 mi^2 (6,475 km^2) watershed to one per 250 mi^2 (648 km^2) the calibration error between simulated and recorded flows improved from 19% to almost zero.

An important application of runoff models is, of course, for use in discharge forecasting. Fig. 9 shows different runoff patterns in the two years which have been simulated with a reasonable accuracy. By replacing a simulation based on data from a past season with an operational forecast, the operation of a reservoir for water power generation or for water supply could be improved.

To enhance this prospect, the processing of the Landsat data should be accelerated in order to update the depletion curves of the snow coverage within several days after each satellite overflight. The behavior of the depletion curves in relation to the initial snow accumulation and to subsequent temperatures should be studied enabling extrapolations to be made. The range of short-term discharge forecasts depends on possibilities of temperature forecasts. Statistical temperature

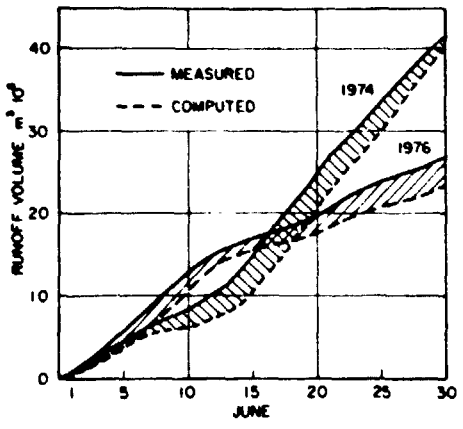


Fig. 9. Cumulative curves of computed and measured runoff volumes for June 1976 and 1974 for Dinwoody Creek.

data can also be used to predict the runoff to be expected with a given probability. In these cases the model does not require information on the water equivalent of the snowcover at the start of the melting season. This value is naturally important for seasonal forecasts of runoff volume, but it can hardly be measured on an areal basis in a large mountain basin. It remains to be seen whether water equivalent can be measured in the future by remote sensing or at least approximately estimated by frequent snowcover monitoring.

Conclusions

The snowmelt runoff model developed by Martinec (1975) was designed to operate under conditions where snowmelt is the major contributor to runoff and in mountain basins with a great elevation range. Originally the model was developed in small basins, but it appears to be applicable in watersheds of several hundred km^2, based on results from the Dinwoody Creek basin in Wyoming. The limiting upper size apparently has not yet been reached. Improved simulation accuracy would most likely result from the location of a hydrometeorological station in the basin at or near the hypsometric mean elevation. In fact, it seems that the quality of the hydrometeorological data has had a more significant effect on simulation accuracy than the increase in basin size from 43 km^2 to 228 km^2. Conversely, snowmelt runoff computations in the Dinwoody Creek basin were facilitated by the relatively low summer precipitation in the area.

Landsat provides the means for obtaining the critical snowcover input parameter required by the snowmelt model on the Dinwoody Creek basin. The satellite platform is the most efficient way to obtain snow extent on a basin of this size. It must be considered, however, that certain locations such as the Swiss Alps or the northwestern United States have a high frequency of cloudiness which severely hampers the effectiveness of Landsat. In such conditions geosynchronous satellites or the use of cloud-penetrating sensors operating in the microwave region must be considered.

The determination of model parameters and variables on a rational basis such as in this model with little or no optimization facilitates its application in new basins. Although only streamflow simulation was attempted in this study, the use of this model for discharge forecasts has great potential. A means of extrapolating the snowcover depletion curves based on more commonly observed parameters, such as temperature and snow water equivalent, would be required. Coupling this with either forecasts of temperature or statistical temperature data would permit use of the model as a water management tool. The fact that the model can provide reasonable flow simulations for a 6 month period without any updating by actual discharge measurements further indicates the possible application to ungaged watersheds.

Acknowledgments

The authors would like to thank James L. Foster, Goddard Space Flight Center, Ralph Peterson, General Electric Company, Beltsville, Maryland, U.S.A., and H.J. Etter, G. Klausegger, J.V. Niederhäusern, and E. Wengi, Federal Institute for Snow and Avalanche Research for performing essential data analysis for this project. Necessary data were supplied by the U.S. Geological Survey, U.S. Soil Conservation Service, and the National Oceanic and Atmospheric Administration. Special gratitude is extended to the Acting Director of Applications, Goddard Space Flight Center, Dr. L. Meredith, and to the Director of the Federal Institute for Snow and Avalanche Research, Prof. M. de Quervain for their continuing support of this international cooperation.

References

Barry, R G., and Chorley, R. J (1970) *Atmosphere, Weather, and Climate*. Holt, Rinehart, and Winston, Inc , New York, 320 pp.

Bergstrom, S (1979) Spring flood forecasting by conceptual models in Sweden, Proceedings of the Workshop on Modeling of Snow Cover Runoff, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, pp 397-405.

Hannaford, J.F. (1977) Investigation application of satellite imagery to hydrologic modeling snowmelt runoff in the southern Sierra Nevada, Phase 1 Final Report, NAS 5-22957, Goddard Space Flight Center, Greenbelt, Maryland, 48 pp.

Johanson, R.C. (1971) Precipitation network requirements for streamflow estimation, Technical Report 147, Department of Civil Engineering, Stanford University, Stanford, California.

Kite, G.W. (1975) Performance of two deterministic hydrological models, Proceedings of the Symposium on Application of Mathematical Models in Hydrology and Water Resources Systems, IAHS - AISH Publication No. 115, Bratislava, pp. 136-142.

Leaf, C F (1967) Aerial extent of snow cover in relation to streamflow in central Colorado, International Hydrology Symposium Proceedings, Fort Collins, Colorado, pp 157-164.

Leaf, C F. (1975) Applications of satellite snow cover in computerized short-term streamflow forecasting, in Operational Applications of Satellite Snowcover Observations, NASA SP-391, Washington, D.C., pp. 175-186.

Martinec, J. (1960) The degree-day factor for snowmelt-runoff forecasting, IAHS Publication No. 51, Surface Waters, pp. 468-477.

Martinec, J (1970) Study of snowmelt runoff process in two representative watersheds with different elevation range, IAHS Publication No. 96, Symposium of Wellington, pp. 29-39

Martinec, J. (1975) Snowmelt-runoff model for stream flow forecasts, *Nordic Hydrology*, 6, No. 3, pp. 145-154.

Nash, J. E., and Sutcliffe, J. V. (1970) River flow forecasting through conceptual models; Part I – A discussion of principles, *Journal of Hydrology, 10,* No. 3, pp. 282-290.

Rango, A. and Itten, K.I. (1976) Satellite potentials in snowcover monitoring and runoff prediction, *Nordic Hydrology, 7,* pp. 209-230.

Rango, A., Salomonson, V. V., and Foster, J. L. (1977) Seasonal streamflow estimation in the Himalayan region employing meteorological satellite snow cover observations, *Water Resources Research, 14,* No. 2, pp. 359-373.

Rango, A., Hannaford, J.F., Hall, R. L., Rosenzweig, M., and Brown, A. J. (1979) Snow-covered area utilization in runoff forecasts, *Journal of the Hydraulics Divison,* ASCE, Vol. 105, No. HY1, pp. 53-66.

Speers, D., Kuehl, D., and Schermerhorn, V. (1979) Development of the operational snow band SSARR model, Proceedings of the Workshop on Modeling of Snow Cover Runoff, U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, pp. 369-378.

Strangeways, C., and McCulloch, J. S. G. (1965) A low priced automatic hydrometeorological station, *Bulletin IAHS,* Vol. X, No. 4, pp. 57-62.

Thompson, A. G. (1975) Utilization of Landsat monitoring capabilities for snowcover depletion analysis, in Operational Applications of Satellite Snowcover Observations, NASA SP-391, Washington, D.C., pp. 113-127.

U.S. Army Corps of Engineers (1975) Program description and user manual for SSARR model – streamflow synthesis and reservoir regulation, Program 724-K5-G0010, North Pacific Division, Portland, Oregon, 188 pp.

Viessman, W. (1968) The synthesis of snowmelt hydrographs, in Snow Hydrology, Proceedings of the National Workshop Seminar, Canadian National Committee, International Hydrological Decade, Frederickton, New Brunswick, pp. 66-79.

Received. 26 April, 1979

Address:
A. Rango,
Code 913,
Goddard Space Flight Center,
Greenbelt, Maryland 20771,
U.S.A.

J. Martinec,
Federal Institute for Snow and Avalanche Research,
Weissfluhjoch/Davos,
Switzerland.

Paper 100

Reprinted from JOURNAL OF APPLIED METEOROLOGY, Vol. 17, No. 10, October 1978
American Meteorological Society
Printed in U.S.A.

Remote Sensing of Surface Soil Moisture[1]

T. SCHMUGGE

NASA/Goddard Space Flight Center, Greenbelt, MD 20771

(Manuscript received 28 December 1977, in final form 13 May 1978)

ABSTRACT

The unique thermal and dielectric properties of water afford two possibilities for remotely sensing the moisture content in the surface layer of the soil. Observations of the diurnal range of surface temperature, the microwave brightness temperature (emissivity) and radar backscatter of the soil have shown correlations of up to 0.9 with the moisture in the surface layer (~5 cm thick). The microwave techniques appear to maintain their sensitivity to moisture variations in the presence of a crop canopy. Observations of microwave brightness temperature from satellite platforms have qualitatively confirmed this sensitivity for a wide range of conditions.

1. Introduction

The unique thermal and dielectric properties of water afford two possibilities for remotely sensing the moisture content in the surface layer of the soil. The large heat capacity and thermal conductivity of water enable moist soils to have a large thermal inertia. This thermal inertia can be remotely sensed by observing the diurnal range of surface temperature.

The dielectric constant for water is an order of magnitude larger than that of dry soils at microwave wavelengths (30 cm$>\lambda>$1 cm). As a result the surface emissivity and reflectivity for the soils at these wavelengths are strong functions of its moisture content. The changes in emissivity can be observed by passive microwave techniques (radiometry) and the changes in reflectivity can be observed by active microwave techniques (radar).

Both of these approaches, thermal and microwave, have been demonstrated in extensive field and aircraft measurements. Correlations of about 0.9 have been obtained between soil moisture in the surface layer (~5 cm thick) and microwave brightness temperatures or diurnal range of surface temperature. The microwave techniques maintain their sensitivity to soil moisture variation in the presence of a crop canopy. Qualitative observations of the passive microwave sensitivity have been made from satellite platforms at wavelengths of 21 and 1.55 cm. Thus, it appears to be possible to monitor the moisture status of the surface soil using these techniques.

Since NASA is planning or proposing spacecraft tests of these approaches this paper will provide an opportunity to present the results on which these proposals are based and to discuss the relative advantages of each method.

2. Thermal methods

The amplitude of the diurnal range of surface temperature for the soil is a function of both internal and external factors. The internal factors are thermal conductivity (K) and heat capacity (C), where $P=(KC)^{\frac{1}{2}}$ defines what is known as "thermal inertia." The external factors are primarily meteorological—solar radiation, air temperature, relative humidity, cloudiness, wind, etc. The combined effect of these external factors is that of the driving function for the diurnal variation of surface temperature. Thermal inertia, then, is an indication of the soil's resistance to this driving force. Since both the heat capacity and thermal conductivity of a soil increase with an increase of soil moisture, the resulting diurnal range of surface temperature will decrease.

The basic phenomena are illustrated in Fig. 1, which presents surface temperatures as measured with a thermocouple for a field versus time, before and after irrigation. These data were obtained at the U. S. Water Conservation Laboratory in Phoenix (Idso *et al.*, 1975).

The solid line in Fig. 1 is the plot of surface temperature before irrigation, and the filled circles reflect the data on the day following irrigation. There is a dramatic difference in the maximum temperature achieved on these two days. On succeeding days the maximum temperature increases as the field dries out.

The summary of results from many such experiments is shown in Fig. 2 where the amplitude of the diurnal range is plotted as a function of the soil moisture as measured at the surface and at 0–1 cm, 0–2 cm and

[1] Presented at the Second Conference on Hydrometeorology, 25–27 October 1977, Toronto, Ontario, Canada.

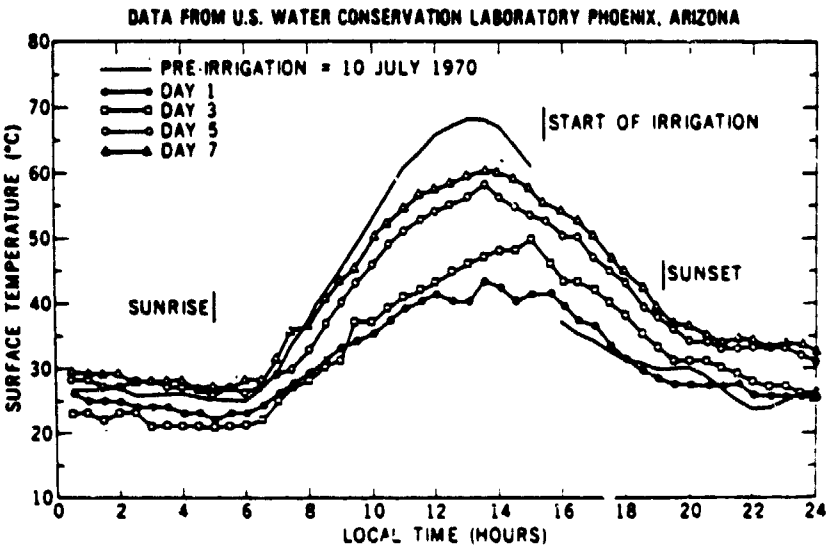


FIG. 1. Diurnal surface temperature variation.

0–4 cm layers. There is a good correlation with the soil moisture in the 0–2 cm and 0–4 cm layers of the soil, and this response is related to the thermal inertia of the soil. Initially, when the surface is moist, the temperatures are *more or less* controlled by evaporation. Once the surface layer dries below a certain level, the temperature will be determined by the thermal inertia of the soil. These results indicate that for this particular soil, the diurnal range of surface temperature is a good measure of its moisture content.

When these measurements are repeated for different soils, there are differences which depend on the soil type. However, there are certain characteristics that are independent of the soil type, and these relate to the evaporation of the water from the soil. Soil physicists have characterized the drying of a soil in three stages:

- The wet stage, where the evaporation is solely determined by the meteorological condition.
- An intermediate or drying stage where it starts out being in the wet stage early in the day, but because there is not a sufficient amount of water in the soil to meet the evaporative demand, the evaporation rate falls off.
- The dry stage, where evaporation is solely determined by the molecular transfer properties of water within the soil.

There is a striking change in both the albedo and the evaporation rate as the soil dries during the transition from the wet stage to drying stage.

Temperature measurements were repeated for different soil types. The soils ranged from sandy or light soils to heavy clay soils. It is clear that for a given diurnal temperature difference, there can be a wide range of moisture content for these soils (Idso *et al.*, 1975).

However, the ΔT values observed as the soils dried through the transitions between the stages mentioned above were approximately the same for all of the soil types studied. Thus it has been concluded (Idso *et al.*, 1975) that while the relation between ΔT and moisture content depends on soil type, the relation between ΔT and pressure potential (the tension with which water is held by soil particles) is independent of soil type. This is the basis for expressing moisture values as a percent of field capacity (FC), where field capacity is taken to be the moisture content at the $-\frac{1}{3}$ bar pressure potential.

It should be emphasized that these experiments were all made in a field, using thermocouples, and were not remotely sensed. In March 1975, an experiment was performed in which remotely sensed thermal infrared temperatures from an aircraft platform were compared with the *in situ* thermocouple measurements over a 5-day period. There was good agreement between the thermocouple measurements and the remotely sensed radiation measurements made from the aircraft (Reginato *et al.*, 1976; Schmugge *et al.*, 1978), indicating that the conclusions based on the thermocouple measurements would also be valid for radiation temperature observation.

In Fig. 3 the results from both the field experiments (from Fig. 2) and the aircraft experiments are presented. The field results are expressed as a percent of field capacity so they can be compared with the aircraft results obtained over a wide range of soil textures. The good agreement between the field and aircraft results indicate that the results based on the field measurements can be extrapolated to the remote sensing technique also.

This technique is not applicable to fields with a vegetative canopy. However, the difference between canopy temperature and ambient air temperature has

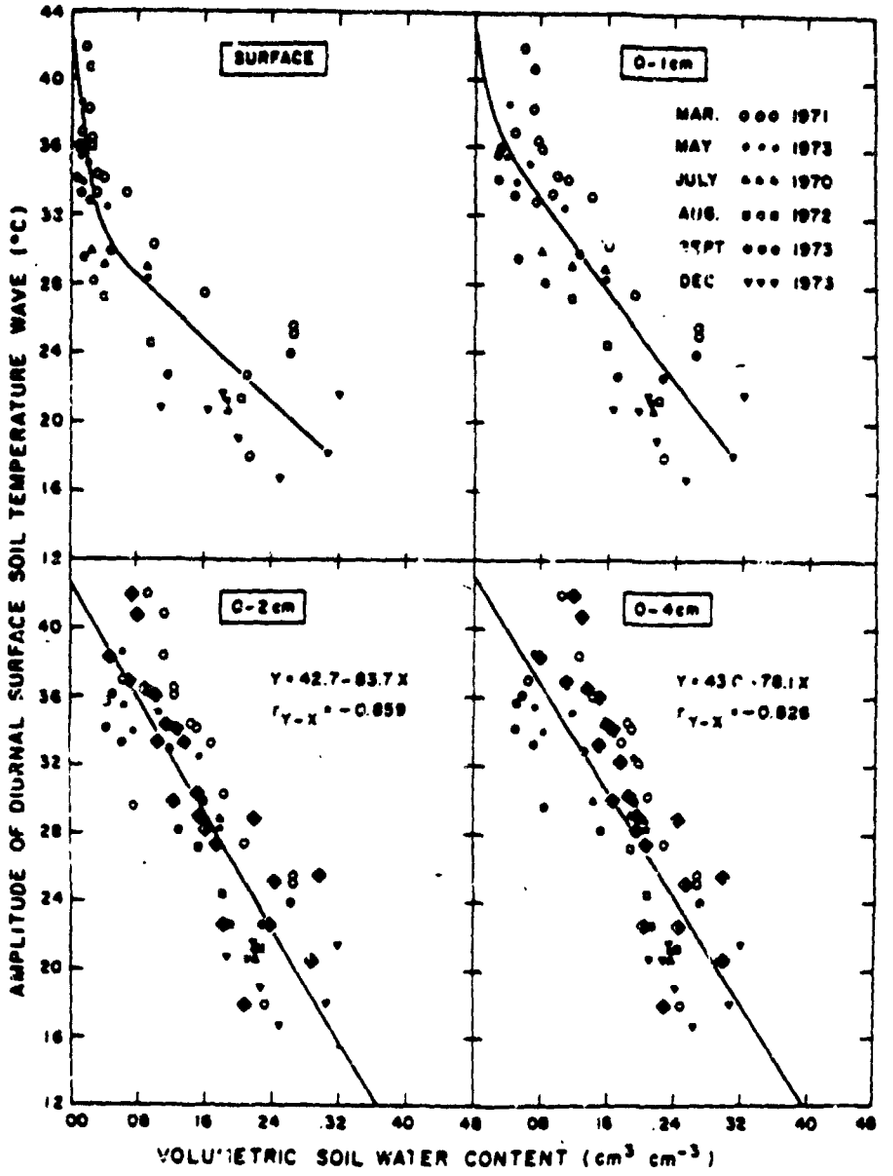


FIG. 2. Summary of results for the diurnal temperature variation versus soil moisture. From Idso *et al.* (1975).

been shown to be an indicator of crop water use (Jackson *et al.*, 1977), thus extending the usefulness of the thermal IR approach.

This approach will be studied further by additional high altitude aircraft flights and by the Heat Capacity Mapping Radiometer launched on the first Applications Explorer Satellite in April 1978. This sensor has two channels (10–12 μm and 0.5–1.1 μm), the latter for measuring surface albedo. The spatial resolution will be 0.6 km. The satellite will be in a 600 km sun-synchronous orbit with a 1400 LST equator crossing to observe the maximum surface temperature. The minimum will be observed either 12 h before or after to provide the diurnal range. This coverage will be repeated every eight days.

3. Microwave methods

a. Soil dielectric properties

As noted in the Introduction the dielectric properties of a soil are strongly dependent on its moisture content because of the large dielectric constant for water, approximately 80 as compared with 3 or 4 for dry soils. This dependence is shown in Fig. 4 which presents the

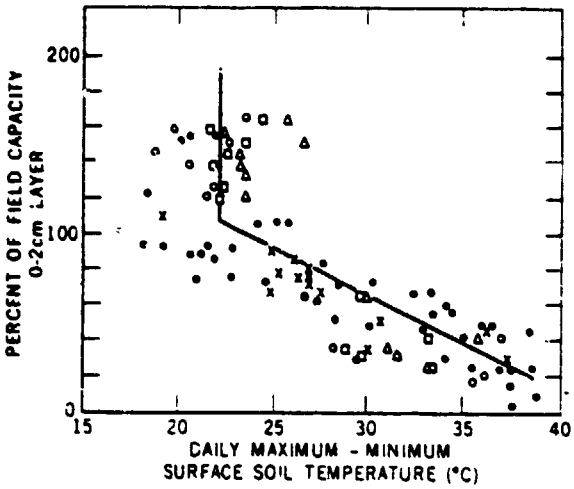


FIG. 3. Plot of ΔT versus soil moisture in the 0–2 cm layer. The symbols represent the different types of temperature measurement: (●), (□), surface thermocouple; (○), hand-held radiometer; (△), aircraft data over test plot; (×), aircraft data over the general agricultural fields. [(●), (□), (○), (△) from Reginato *et al.* (1976); (×) from Schmugge *et al.*, (1978)]

results of laboratory measurements at wavelengths of 21 and 1.55 cm. The wavelength dependence is due to the difference in the dielectric properties of water at the two wavelengths.

At low levels there is a slow increase with soil moisture but above a certain point there is a sharp increase in the slope of the curve which is due to the behavior of the water in the soil. When water is first added to a soil it is tightly bound to the soil particles and in this state the water molecules are not free to become aligned and the dielectric properties of this water are similar to those of ice. As the layer of water around the soil particle becomes larger, the binding to the particle decreases and the water molecules behave as they do in the liquid, hence the greater slope at the higher soil moisture values. The transition depends on the soil texture, i.e., particle size distribution being lower for a sand and large for a clay. This effect has been demonstrated in laboratory measurements of the dielectric constant (Lundien, 1971; Newton, 1976).

Recall that the dielectric constants of the medium describe propagation characteristics for an electromagnetic wave in the medium. Therefore, they determine the emissive and reflective properties for a smooth surface.

b. Passive microwave response to soil moisture

A microwave radiometer measures the thermal emission from the surface and at these wavelengths the intensity of the observed emission is essentially proportional to the product of the temperature and emissivity of the surface (Rayleigh-Jeans approximation). This product is commonly referred to as brightness temperature. All our results will be expressed as brightness temperatures (T_B). The value of T_B observed by a radiometer at a height h above the ground is

$$T_B = \tau(rT_{sky} - (1-r)T_{surf}) + T_{atm}, \qquad (1)$$

where r is the surface reflectivity and τ the atmospheric transmission. The first term is the reflected sky brightness temperature which depends on wavelength and atmospheric conditions; the second term is the emission from the surface ($1-r=e$, where e is the emissivity); and the third term is the contribution from the atmosphere between the surface and the receiver. At the longer wavelengths, i.e., these best suited for soil moisture sensing, the atmospheric effects are minimal and will be neglected in this discussion.

The range of dielectric constant presented in Fig. 4 produces a change in emissivity from greater than 0.9 for a dry soil to less than 0.6 for a wet soil, assuming an isotropic soil with a smooth surface. This change in emissivity for a soil has been observed by truck-mounted radiometers in field experiments (Poe *et al.*, 1971; Newton, 1976), and by radiometers in aircraft (Schmugge *et al.*, 1974) and satellites (Eagleman and Lin, 1976). In no case were emissivities as low as 0.6 observed for real surfaces. It is believed that this is primarily due to the effects of surface roughness which generally has the effect of increasing the surface emissivity.

As can be seen in Fig. 4 there is a greater range of dielectric constant for soils at the 21 cm wavelengths. This fact combined with a larger soil moisture sampling

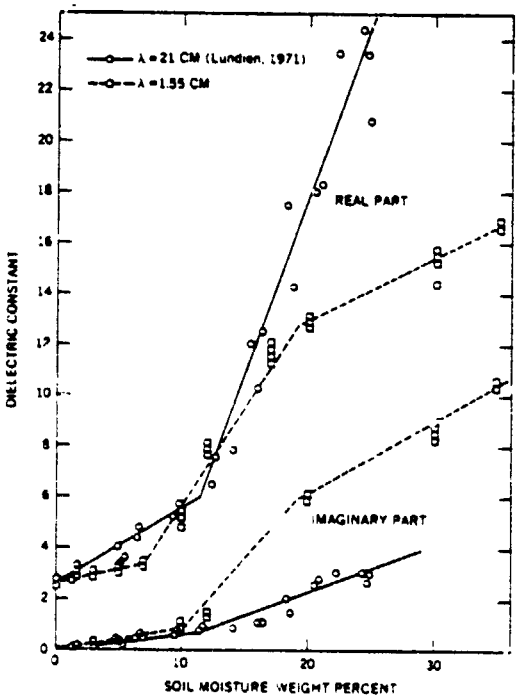


FIG. 4. Dependence of the soil's dielectric constant on its moisture content.

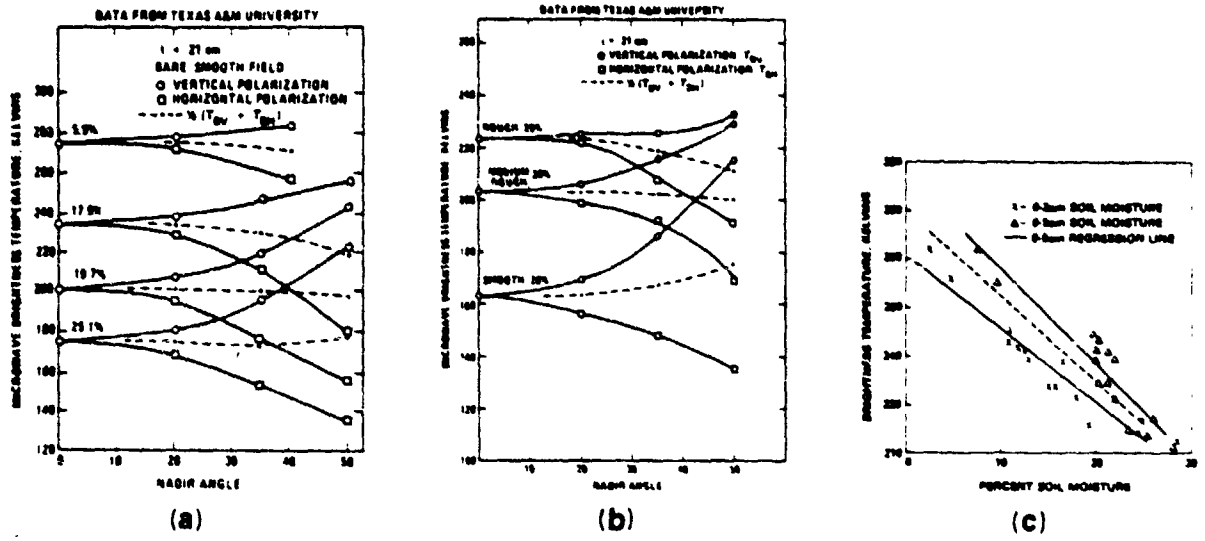


FIG. 5. Results from field measurements performed at Texas A&M University (a) T_B versus angle for different moisture levels, (b) T_B versus angle for different surface roughness at about the same moisture level, (c) T_B versus soil moisture in different layers for the medium rough field (Newton, 1976).

depth and better ability to penetrate a vegetative canopy make the longer wavelength sensors better suited for soil moisture sensing.

In Fig. 5, the field measurements of Newton (1976) are plotted versus angle of observation for various moisture contents and for three levels of surface roughness. The horizontal polarization is that for which the electric field of the wave is parallel to the surface and the vertical polarization is perpendicular to it.

These results indicate the effect of moisture content on the observed values of T_B and the effect of surface roughness which is to increase the effective emissivity at all angles and to decrease the difference in T_B for the two polarizations at the larger angles.

For the smooth field there is a 100 K change in T_B in going from wet to dry soils and it is clear that this range is reduced by surface roughness. The effect of the roughness is to decrease the reflectivity of the surface and thus to increase its emissivity. For a dry field the reflectivity is already small (<0.1) so that the resulting increase in emissivity is small. As seen in Fig. 5b surface roughness has a significant effect for wet fields where the reflectivity is larger (~ 0.4). Thus the range of T_B for the rough field is reduced to about 60 K. The smooth and rough fields represent the extremes of surface conditions that are likely to be encountered, e.g., the rough surface was on a field with a heavy clay soil (clay fraction $>60\%$) that had been

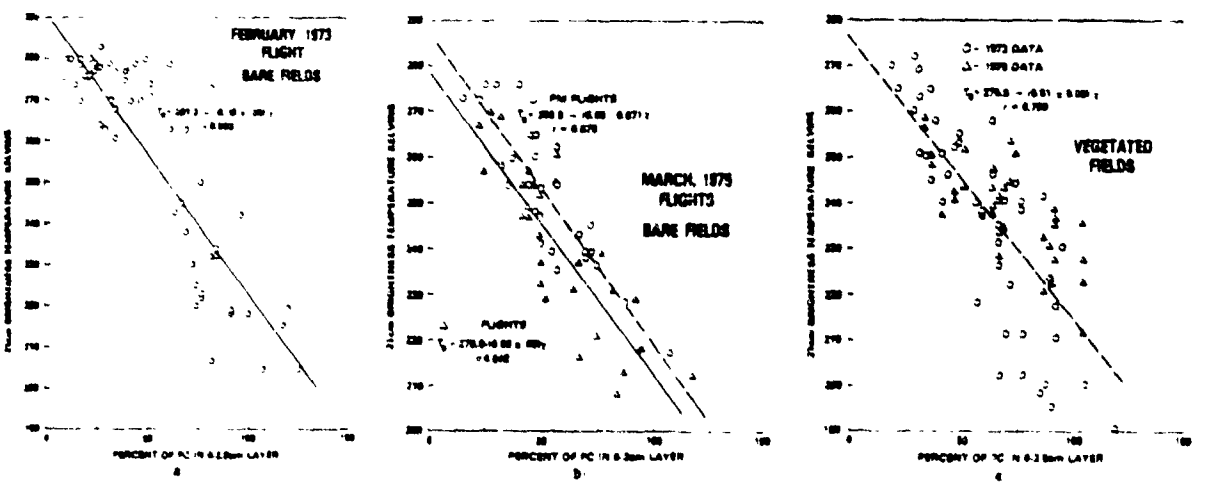


FIG. 6. Aircraft observations of T_B over agricultural fields around Phoenix (a) bare field results from 1973 flight, (b) bare field results from 1975 flights; (c) vegetated field results from both years.

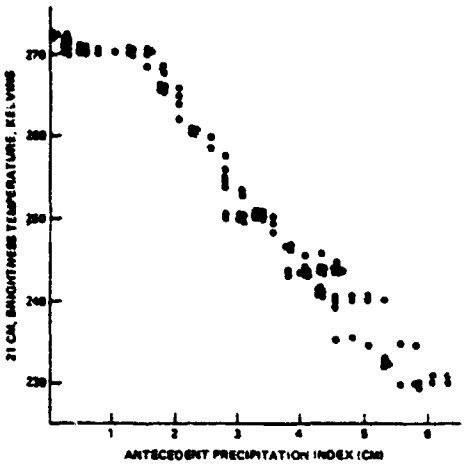


FIG. 7. Skylab observations of T_B at 21 cm compared with antecedent precipitation over Texas and Oklahoma (McFarland, 1976).

deep plowed which produced large clods. Therefore the medium rough field, with a T_B range of 80 K, is probably more representative of the average surface roughness condition that will be encountered. Another important observation from Fig. 5 is that the average of the vertical and horizontal T_B's is essentially independent of angle out to 40°. This indicates that the sensitivity of this quantity, $\frac{1}{2}(T_{BV}+T_{BH})$, to soil moisture will be independent of angle. This factor will be useful if the radiometer is to be scanned to provide an image.

When the brightness temperatures for the medium rough field are plotted versus soil moisture in the 0–2 cm layer there is an approximate linear decrease of T_B (Fig. 5c). As the thickness of the layer increases both the slope and intercept of the linear regression result also increase. This is because the moisture for the high T_B cases increases, while it remains essentially the same for the low T_B or wet cases. This type of behavior was also seen in the results obtained from aircraft platforms and has led us to conclude that the soil moisture sampling depth is in the 2–5 cm range for the 21 cm wavelength. This is in agreement with the predictions of theoretical results for radiative transfer in soils (Wilheit, 1978; Burke *et al.*, 1978).

The results from aircraft experiments are summarized in Fig. 6 where the results from flights in February 1973 (Schmugge *et al.*, 1976) and March 1975 (Schmugge, 1976) over Phoenix, Arizona, are presented. The T_B values are plotted versus soil moisture expressed as a percent of field capacity as was done for the thermal inertia case (Fig. 3) to normalize the effect of soil texture differences. The agreement of the slopes for the three regressions indicates that the results are repeatable. The differences in the intercepts are due to the differences in soil temperature. This is particularly evident in the difference between the 1975 morning and afternoon results. Also, note that the range of T_B (80 K) is in good agreement with the medium rough field of Fig. 5.

The effect of a vegetative canopy will be that of an absorbing layer that depends on the amount of the vegetation and the wavelength of observations. In Fig. 6c the results for vegetated fields from the two years are presented. The vegetation was either alfalfa or wheat with the wheat being 20–30 cm high in 1973 and 50–60 cm high for the 1975 data. The slope of the curve is in good agreement with those for the bare fields. The intercept is lower due to the cooler soil temperatures. Thus the sensitivity to soil moisture is maintained through the moderate vegetative canopies considered here. This result is supported by the field measurements of Newton (1976) who found the sensitivity maintained through 125 cm of closely planted sorghum.

As has been reported by McFarland (1976) and by Eagleman and Lin (1976), the sensitivity of the 21 cm radiometer to soil moisture has already been demonstrated from space during the Skylab mission. McFarland showed a definite relationship between the Skylab 21 cm brightness temperatures and the Antecedent Precipitation Index (API). Fig. 7 presents these results for a pass starting over the Texas and Oklahoma panhandles and proceeding to the southeast toward the Gulf of Mexico. Each point plotted is the observed brightness temperature and the API calculated from all the rain gages within the 110 km footprint. Since there is considerable overlap for the radiometer footprints presented here, this plot should be considered as a comparison of the running average of T_B with API. As such it shows the sensitivity of spaceborne radiom-

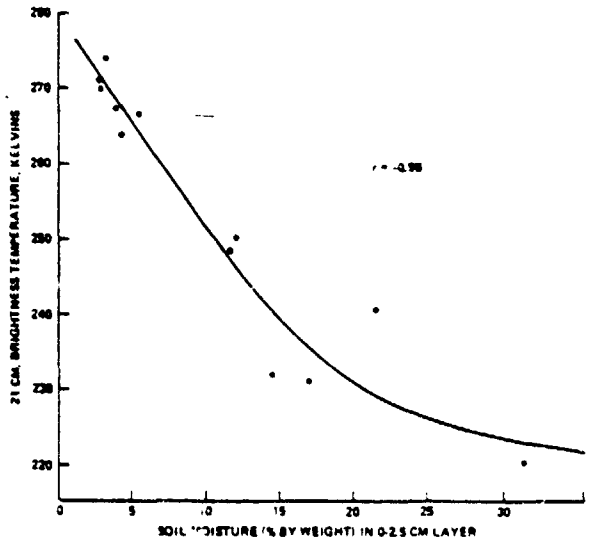


FIG. 8. Skylab observation of T_B at 21 cm compared with soil moisture estimates from five passes over the southern Great Plains (Eagleman and Lin, 1976)

eter to the soil moisture variations caused by the rainfall.

Eagleman and Lin (1976) carried the analysis of the Skylab data a step further, and compared the brightness temperature with estimates of the soil moisture over the radiometer footprint. The soil moisture estimates were based on a combination of actual ground measurements and calculations of the soil moisture using a climatic water balance model. A summary of their results is presented in Fig. 8 for 12 footprints obtained during five different Skylab passes over the states of Texas, Oklahoma and Kansas. The correlation coefficient for these 12 data points is 0.96, which is very good considering the difficulty of obtaining soil moisture information over a footprint of such a size and considering the fact that the brightness temperature was averaged over the wide range of cultural conditions that occurred over the area.

These results from space supported by the more detailed aircraft and ground measurements presented earlier give strong support to the possibility of using microwave radiometers for soil moisture sensing. Therefore, to pursue this technique further, NASA is giving strong consideration to flying a 21 cm radiometer on a future mission to monitor soil moisture variations globally. A candidate system would have a 10 m×10 m antenna which provides 20–40 km spatial resolution from a 800 km orbit. The proposed launch date for this mission is the mid 1980's.

c. *Active microwave response to soil moisture*

The backscattering from an extended target, such as a soil medium, is characterized in terms of the target's scattering coefficient σ°. Thus, σ° represents the link between the target properties and the scatterometer responses. For a given set of sensor parameters (wavelength, polarization and incidence angle relative to nadir), σ° of bare soil is a function of the soil surface roughness and dielectric properties which depends on the moisture content. The variations of σ° with soil moisture, surface roughness, incidence angle and observation frequency have been studied extensively in ground-based experiments conducted by scientists at the University of Kansas (Batlivala and Ulaby, 1977) using a truck mounted 1-18 GHz active microwave system.

To understand the effects of look angle and surface roughness consider the plots of σ° versus angle presented in Fig. 9 for five fields with essentially the same moisture content but with considerably different surface roughness. At the longest wavelength (1.1 GHz, Fig. 9a), σ° for the smoother fields is very sensitive to incidence angle near nadir, while for the rough field σ° is almost independent of angle. At an angle of about 5° the effects of roughness are minimized. As the wavelength decreases (Figs. 9b and 9c) all the fields appear rougher, especially the smooth field, and as a result the intersection point of the five curves moves out to larger angles. At 4.25 GHz the intersection occurs at 10°, and it was this combination of angle and frequency that yielded the best sensitivity to soil moisture independent of roughness.

These experiments were performed in both 1974 and 1975, the first on a field with high clay content (62%), the second with a lower clay content. Although both experiments provided the same specifications of the radar parameters for soil moisture sensing, i.e., frequency around 4.75 GHz and a 7–17° nadir angle,

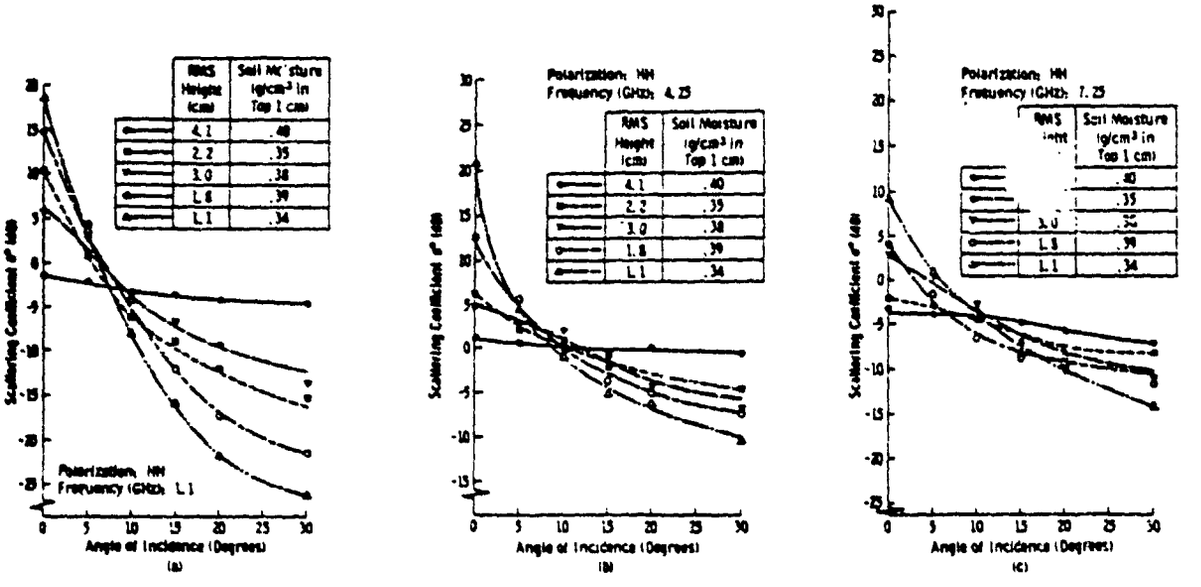


FIG. 9. Angular response of scattering coefficient for the five fields in high levels of moisture content (a) L-band (1.1 GHz), (b) C-band (4.25 GHz), (c) X-band (7.25 GHz). 1975 soil moisture experiment (Batlivala and Ulaby, 1977).

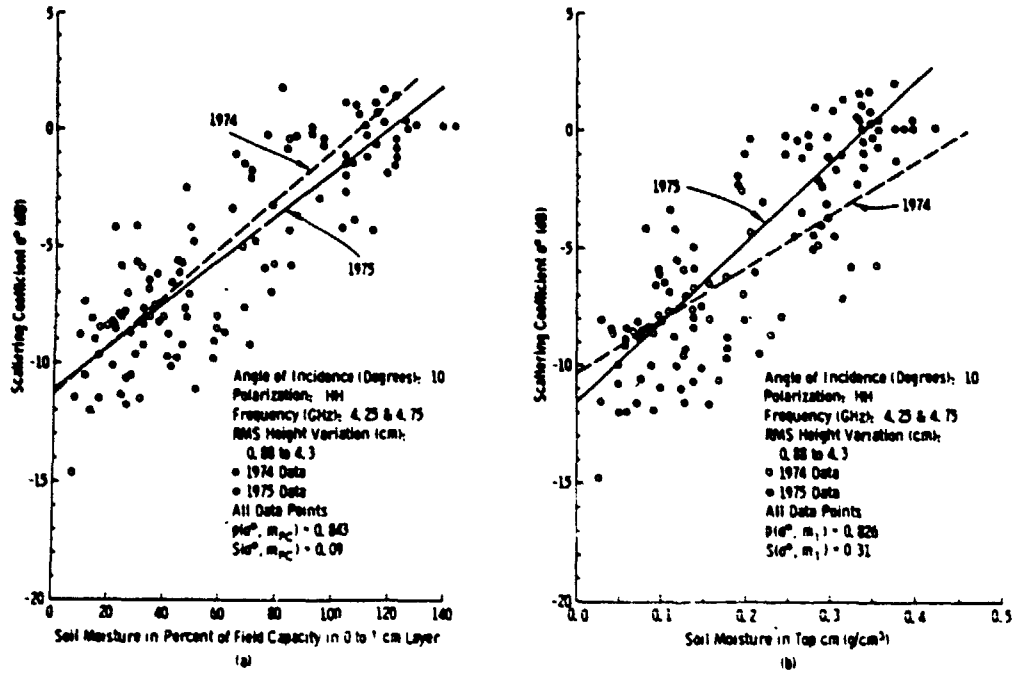


FIG. 10. Backscattering coefficient plotted as a function of soil moisture given (a) in percent of field capacity of the top 1 cm and (b) volumetrically in the top 1 cm. 1974 and 1975 bare soil experiment data are combined (Batlivala and Ulaby, 1977).

the observed sensitivity of σ° to soil moisture was different for the two experiments (Fig. 10b). When the soil moisture content is expressed as a percent of field capacity to account for textural differences, the sensitivities became almost identical (Fig. 10a) with a correlation of 0.84. This dependence on the percent of field capacity is similar to that observed with the thermal inertia and passive microwave techniques.

There have been some recent experiments studying the active microwave approach from aircraft and the results should be available in the near future. In 1978 there will be additional experiments performed with a scatterometer operating near the optimum frequency and should demonstrate the capabilities of this approach.

3. Discussion

At the present time none of the three methods presented here has the clear advantage for being the preferred method of remote sensing of soil moisture. The thermal IR approach has the advantage of providing useful thermal data that may be an indicator of crop status and is capable of providing soil moisture data at high spatial resolutions. However, the usefulness of this approach is lost in the presence of cloud cover. The ability of the microwave sensors to penetrate non-raining clouds makes them very attractive for use as soil moisture sensors. The passive microwave technique has been demonstrated by both aircraft and spacecraft instruments, but the spatial resolution is limited by the size of the antenna which can be flown. For example, at a wavelength of 21 cm, a 10 m×10 m antenna is required to yield 20 km resolution from a satellite altitude of 800 km. It is possible to make use of the coherent nature of the signal in active microwave systems (Synthetic Aperture Radar, SAR) to obtain better spatial resolutions (Moore, 1975). However, the capabilities of such systems for soil moisture sensing remain to be demonstrated from either aircraft or spacecraft platforms. Also, the strong effects of incidence angle and surface roughness makes the unambiguous determination of soil moisture difficult with this type of sensor.

While it is clear that no one system will satisfy all requirements that may be desirable for a soil moisture sensing system (i.e., frequent, high-resolution coverage on a global basis), a microwave radiometer with the characteristics mentioned above would provide wide area coverage with 10–20 km resolution every two or three days. This system could be supplemented with either the thermal IR or radar high-resolution data on a sampling basis.

REFERENCES

Batlivala, P. P., and F. T. Ulaby, 1977. Estimation of soil moisture with radar remote sensing. *ERIM Proc. 11th Int. Symp. Remote Sensing of Environment*, Vol. 2, Ann Arbor, Environ. Res. Inst. of Michigan, 1557–1566.

Burke, W. J., T. Schmugge and J. F. Paris, 1978: Comparison of 2.8 and 21 cm microwave radiometer observations over soils with emission model calculations. *J. Geophys. Res.* (in press).

Eagleman, J. R., and W. C. Lin, 1976: Remote sensing of soil moisture by a 21 cm passive radiometer. *J. Geophys. Res.*, 81, 3660–3666.

Idso, S. B., T. J. Schmugge, R. D. Jackson and R. J. Reginato, 1975: The utility of surface temperature measurements for the remote sensing of soil water status. *J. Geophys. Res.*, 80, 3044–3049.

Jackson, R. D., R. J. Reginato and S. B. Idso. 1977: Wheat canopy temperature: A practical tool for evaluating water requirements. *Water Resour. Res.*, 13, 651–656.

Lundien, J. R., 1971: Terrain analysis by electromagnetic means. Tech. Rep. No. 3–693, Rep. 5, U. S. Army Waterways Experiment Station, Vicksburg, MS, 85 pp.

McFarland, M. J., 1976: The correlation of Skylab L-band brightness temperatures with antecedent precipitation. *Preprints Conf. on Hydrometeorology*, Fort Worth, Amer. Meteor. Soc. 60–65.

Moore, R. K., 1975: Microwave remote sensors. *Manual of Remote Sensing*, Vol. 1, Amer. Soc. Photogrammetry (see Chap. 9).

Newton, R. W., 1976. Microwave sensing and its application to soil moisture detection. Tech. Rep. RSC-81 of the Remote Sensing Center at Texas A&M University, 500 pp. [University Microfilms No. 77-20-398].

Poe, G. A., A. Stogryn and A. T. Edgerton, 1971: Determination of soil moisture content using microwave radiometry. Final Rep. No. 1684 FR-1, DOC Contract 0-35239, Aerojet-General Corp., El Monte, CA, 169 pp.

Reginato, R. J., S. B. Idso, J. F. Vedder, R. D. Jackson, M. B. Blanchard and R. Goettelman, 1976: Soil water content and evaporation determined by thermal parameters obtained from ground-based and remote measurements. *J. Geophys. Res.*, 81, 1617–1620.

Schmugge, T. J., 1976: Preliminary results from the March, 1975 soil flight. NASA Tech. Memo TMX-71197, pp. 219.

——, T. Wilheit, W. Webster, Jr., and P. Gloersen, 1976: Remote sensing of soil moisture with microwave radiometers-II. NASA Tech. Note TN D-8321, [NTIS No. N76-32625].

——, B. Blanchard, A. Anderson and J. Wang, 1978: Soil moisture sensing with aircraft observations of the diurnal range of surface temperature. *Water Resour. Bull.*, 14, 169–178.

——, P. Gloersen, T. Wilheit and F. Geiger, 1974: Remote sensing of soil moisture with microwave radiometers. *J. Geophys. Res.*, 79, 317–323.

Wilheit, T. T., 1978: Radiative transfer in a plane stratified dielectric. *IEEE Trans. Geosc. Electron.*, GE-16 138–143.

Paper 101

WATER RESOURCES BULLETIN
VOL. 14, NO. 1 AMERICAN WATER RESOURCES ASSOCIATION FEBRUARY 1978

SOIL MOISTURE SENSING WITH AIRCRAFT OBSERVATIONS OF THE DIURNAL RANGE OF SURFACE TEMPERATURE[1]

T. Schmugge, B. Blanchard, A. Anderson, and J. Wang[2]

ABSTRACT: Aircraft observations of the surface temperature were made by measurements of the thermal emission in the 8-14 μm band over agricultural fields around Phoenix, Arizona. The diurnal range of these surface temperature measurements were well correlated with the ground measurement of soil moisture in the 0-2 cm layer. The surface temperature observations for vegetated fields were found to be within 1 or 2°C of the ambient air temperature indicating no moisture stress. These results indicate that for clear atmospheric conditions remotely sensed surface temperatures can be a reliable indicator of soil moisture conditions and crop status.
(KEY TERMS: remote sensing; soil moisture; crop status; thermal inertia.)

In a recent paper Reginato, *et al.*, (1976) discussed the possibility of using remotely sensed surface temperature data for estimating soil moisture. Their observations were for their laboratory field at the Water Conservation Laboratory in Phoenix. Data acquired by aircraft radiometers operating in the 8-14 μm band for agricultural fields around Phoenix support their conclusions and extend them to a wider range of conditions.

Temperature observations for fields with a vegetative canopy are also presented. These temperatures were generally within 1 or 2 K of the air temperatures reported for the area which is to be expected for non-moisture stressed plants (Idso and Ehrler, 1976).

BACKGROUND

The amplitude of the diurnal range of surface temperature for the soil is a function of both internal and external factors. The internal factors are thermal conductivity (K), density (ρ) and heat capacity (C), the combination $P = (K\rho C)^{1/2}$ defines what is known as "thermal inertia." The external factors are primarily meteorological: solar radiation, air temperature, relative humidity, cloudiness, wind, etc. The combined effect of these external factors is that of the driving function for the diurnal variation of surface temperature. Thermal inertia then is an indication of the soil's resistance to this driving force.

[1] Paper No. 77083 of the *Water Resources Bulletin*. Discussions are open until October 1, 1978.

[2] Respectively, Atmospheric and Hydrospheric Applications Division, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771; Remote Sensing Center, Texas A&M University, College Station, Texas 77843; NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771; and Lockheed Electronics Company, Inc., Houston, Texas 77058

Since both the heat capacity and thermal conductivity of a soil increase with an increase of soil moisture, the resulting diurnal range of surface temperature will decrease. Typical values of P in units of cal/cm^2 sec$^{1/2}$°C will range from 0.02 for dry soils to about 0.1 for wet soils.

The basic phenomena are illustrated in Figure 1, which presents surface temperatures as measured with a thermocouple for a field versus time, before and after irrigation. These data were obtained by Dr. Ray Jackson and his colleagues at the U.S. Water Conservation Laboratory in Phoenix and have recently been published (Idso, *et al.*, 1975).

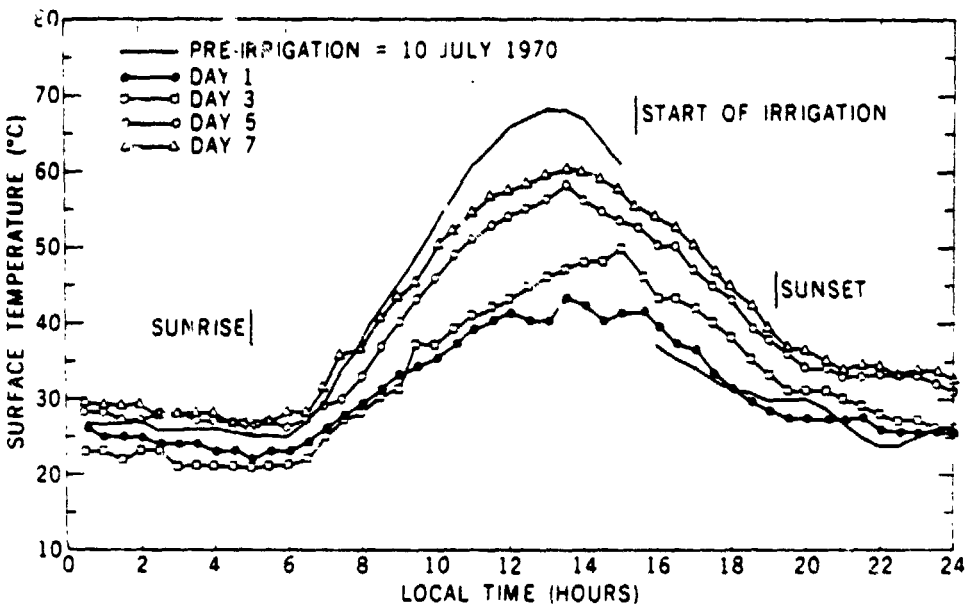


Figure 1. Diurnal Surface Temperature Variation as Measured by a Thermocouple. (Data from U.S. Water Conservation Laboratory in Phoenix, Arizona)

The solid line in Figure 1 is the plot of surface temperature before irrigation, and the solid circles reflect the data on the day following irrigation. There is a dramatic difference in the maximum temperature achieved on these two days. On succeeding days the maximum temperature increases as the field dries out.

The summary of results from many such experiments is shown in Figure 2 where the amplitude of the diurnal range is plotted as a function of the soil moisture measured at the surface and in 0-1, 0-2, and 0-4 cm layers of the soil, and this response is related to the thermal inertia of the soil. Initially, when the surface is moist, the temperatures are more or less controlled by evaporation. Once the surface layer dries below a certain level, the temperature will be determined by the thermal inertia of the soil. These results indicate that for this particular soil, the diurnal range of surface temperature is a good measure of its moisture element.

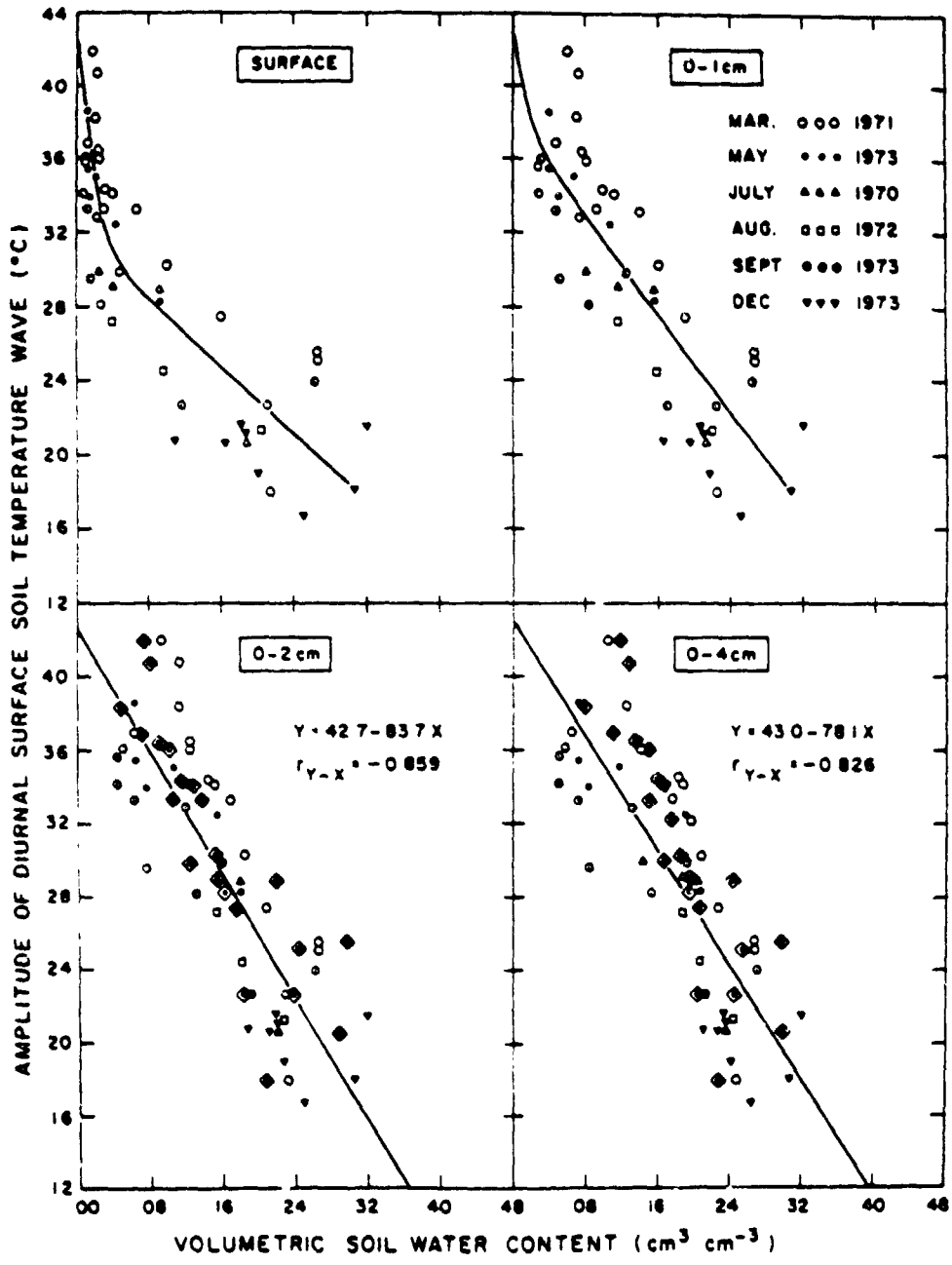


Figure 2. Summary of Results for the Diurnal Temperature Variation Versus Soil Moisture (from Idso, *et al*., 1975).

When these measurements are repeated for different soils ranging in texture from sandy loam to clay there are differences which depend on the soil type, as shown in Figure 3. However, when Idso, *et al.*, (1975) compared ΔT with the soil water pressure potential (Figure 4) the results were largely independent of soil type. This indicates that it may be possible to remotely sense the state of the water independent of soil type.

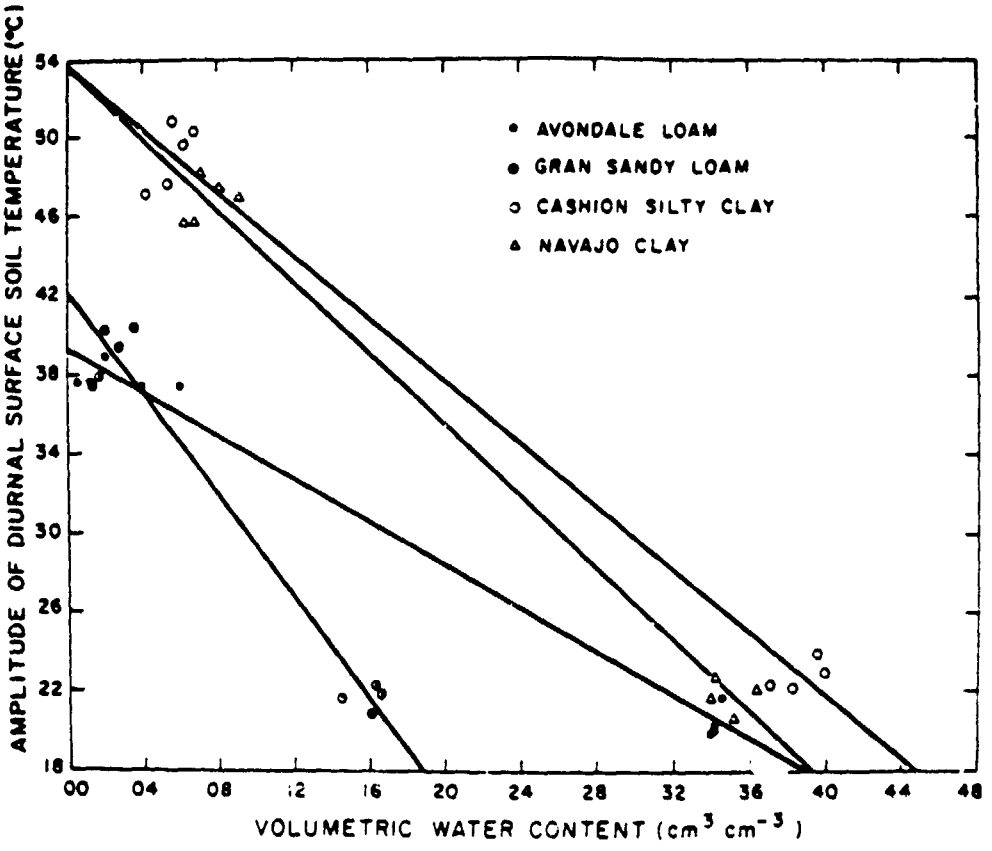


Figure 3 The Amplitude of the Diurnal Surface Soil Temperature Difference Versus the Mean Daylight Volumetric Soil Water Content of the 0-2 cm Depth Increment for Four Different Soils (Idso, *et al.*, 1975).

Two recent papers (Idso and Ehrler, 1976, and Jackson, *et al.*, 1977) have reported that observations of the canopy temperature of a crop were indicative of the moisture status of the crop. Idso and Ehrler found that a positive value for $T_C - T_A$, where T_C is the canopy temperature and T_A is the air temperature, is an indicator of moisture stress in the plant. In the second paper (Jackson, *et al.*, 1977) this concept was extended to the estimation of water use by wheat. They found that the summation of $T_C - T_A$ over time yielded a factor termed the stress degree day which was well correlated with measured water use in six experimental plots.

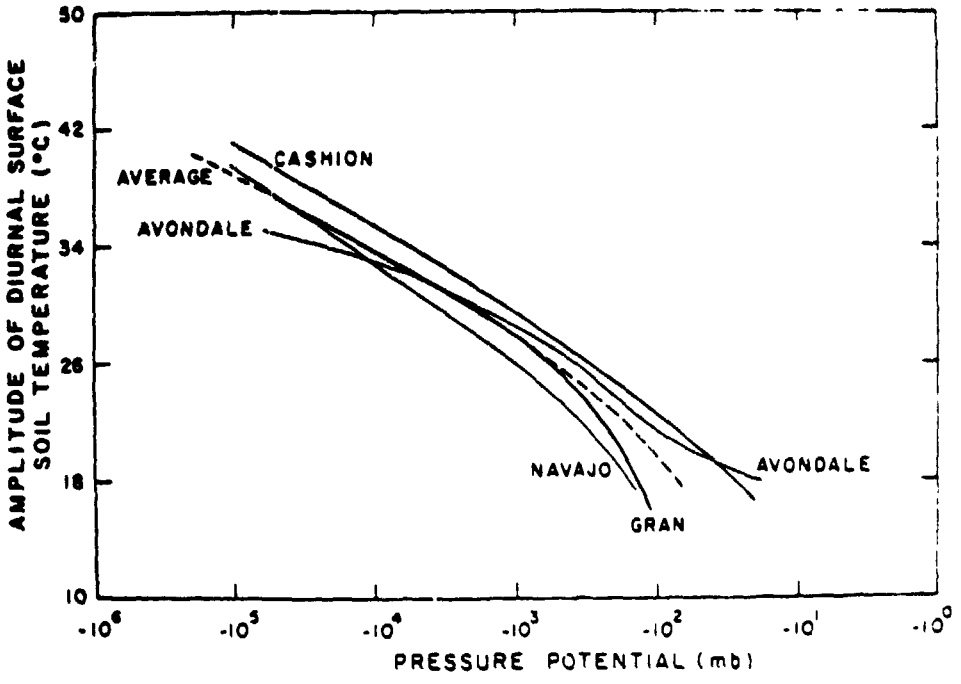


Figure 4. The Amplitude of the Diurnal Surface Soil Temperature Difference Versus the Mean Daylight Soil Water Pressure Potential of the 0-2 cm Depth Increment for Four Different Soils (Idso, *et al.*, 1975).

EXPERIMENTAL DETAILS

The aircraft measurements were made by a Barnes PRT-5 radiometer installed on the NASA P-3A aircraft. The flights were in April 1974 and March 1975. The radiometer operating in 8-14 μm band had a 2° field of view and a precision of 1°C. The surface measurements of the soil moisture were made at four points in a 400 x 400 m (40 acre) field. At each point soil samples were taken for the following depths: 0-1, 1-2, 2-5, 5-9, and 9-15 cm. For furrowed fields the five samples were taken from both the top and bottom of furrows. The values of soil moisture presented here are the averages of the four points. The details of the surface measurements are described elsewhere (Blanchard, 1975). To correlate the surface measurements with the aircraft observations air photos were used to locate the aircraft as a function of time.

Surface temperature observations were made during pre-dawn and early afternoon flights on March 18 and 22, 1975. The area was cloudy for the afternoon flight on March 22, thus rendering the surface temperature measurements on this day unusable for remote sensing purposes. The observations on March 18 agree with those reported in Idso, *et al.*, (1975) for both the diurnal range of soil temperature and the soil-air temperature difference.

The 1974 flights were on the afternoon of April 5 and the early morning of April 6. A complete set of soil moisture measurements were made for afternoon flight only. For this flight the range of the PRT-5 was set for a maximum temperature 40°C: surface

temperatures above this value were extrapolated using the gray levels of an imager (Texas Instrument RS-14). The details of this process are given in the report of that mission (Schmugge, *et al.*, 1976).

As noted in Figure 3 ΔT for given soil moisture depends on soil type, but that ΔT is related to soil moisture expressed in terms of pressure potential independent of soil type. Since we do not have data on the moisture characteristics of the soils studied we have chosen to relate the measured soil moisture to field capacity for the soil estimated from soil texture information for each field. This was done using regression relationships developed from data of over 100 soils for which the moisture characteristics and texture information were available (Schmugge, *et al.*, 1976). The expression used for estimating the field capacity (FC) was.

$$FC = 25.1 - 0.21 \times \text{sand} + 0.22 \times \text{clay}$$

where sand and clay are their respective fraction expressed in percent.

RESULTS

Three passes were flown over each line for both the morning and afternoon flights on March 18, 1975. In the morning the flights were from 6:00 to 7:31 MST, sunrise being at 6:08. The minimum temperatures were observed during the second pass between 6:26 and 6:52, i.e., before solar heating had an effect on the surface. The afternoon flights were between 12:10 and 1:20 MST with the maximum temperatures observed during the third pass 1:00 and 1:20. It is possible that higher surface temperatures may have been attained later in the afternoon (i.e., around 2:00) but it is expected that the difference would be approximately 1 or 2°C.

The data for the difference between the third pass in afternoon and the second pass in the morning are presented in Figures 5a and 5b. The soil moisture values used here are the averages of the data taken during the predawn and afternoon flights. Jackson, *et al.*, (1976) have shown that the 24-hour average soil moisture is well represented by the averages of measurements made at 0500 and 1400 hours. These were approximately the times of the 1975 measurements, and thus our moisture values should accurately reflect the average moisture content for the day. On the basis of detailed sampling done for one field the uncertainty of the 0-2 cm moisture sample using only four points per field is estimated to be 15% (Schmugge, *et al.*, 1976). The estimates of field capacity from soil texture also introduces uncertainty at about the same level resulting in total uncertainty of about 25%.

In Figure 5a the ΔT values are plotted versus the raw soil moisture, the correlation coefficient (r) is 0.82. The value of r increased to 0.89 when soil moisture is expressed as a percent of field capacity, in spite of the increased uncertainty introduced by our estimate of field capacity.

The April 1974 data are presented in Figure 6. The correlation for these data alone is 0.70; when the 1975 data are added the correlation is improved to 0.78. The soil moisture measurements were only for the afternoon, and would be expected to be drier than the 24-hour average. Scatter in the data for low soil moistures is due partially to 1 to 2°C uncertainty of the measurements above 40°C, and also the uncertainty of the soil estimates. The range of ΔT's observed in both of these data sets agree with those of Reginato *(et al.*, 1976), i.e., about 40°C for dry fields and 20°C for the wet fields.

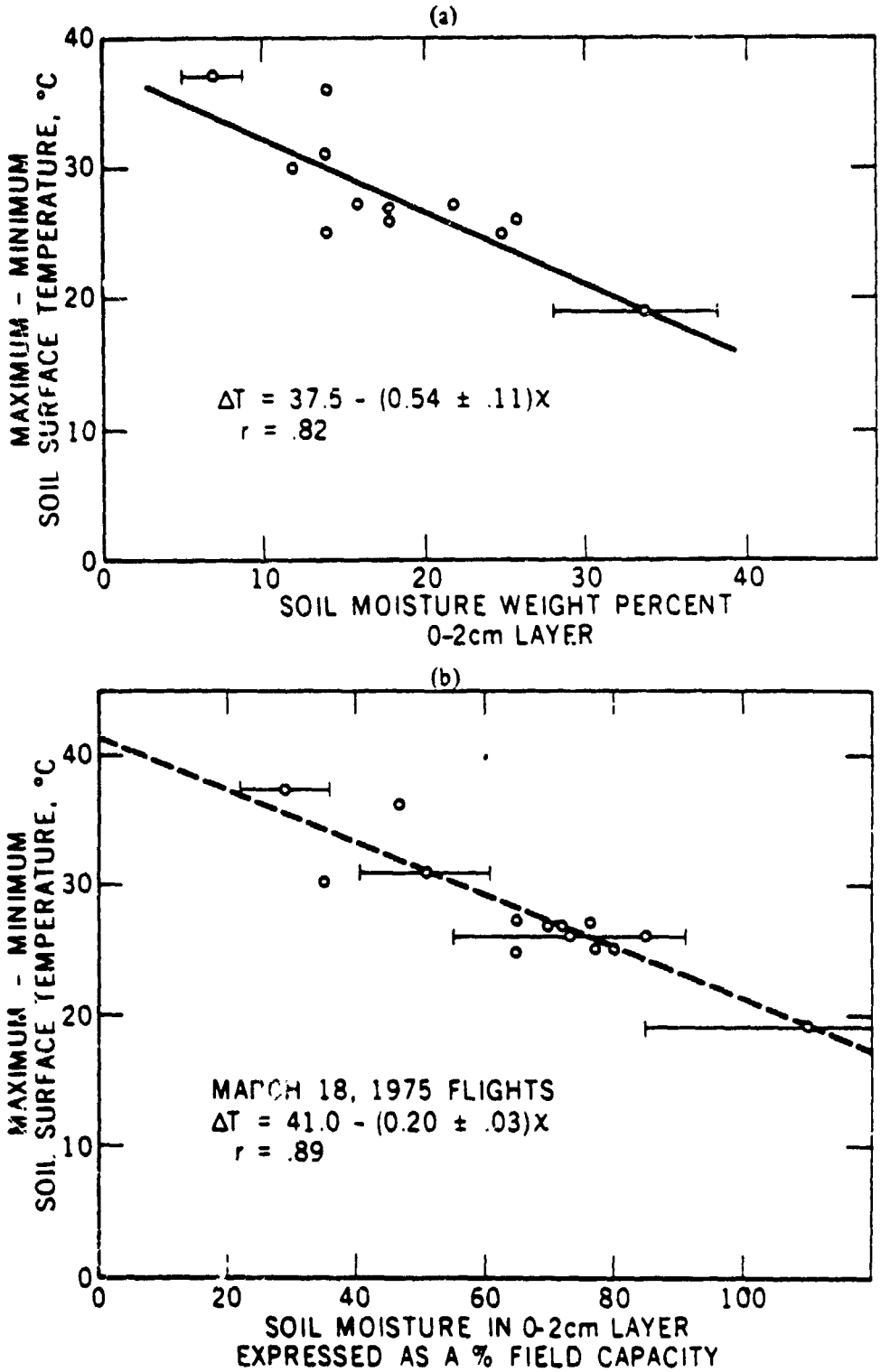


Figure 5. Aircraft Observations of ΔT from the March 18, 1975, Flight Plotted Versus (a) Soil Moisture in the 0-2 cm Layer Expressed in Weight Percent and (b) Soil Moisture in the 0-2 cm Layer Expressed as a Percent of Field Capacity

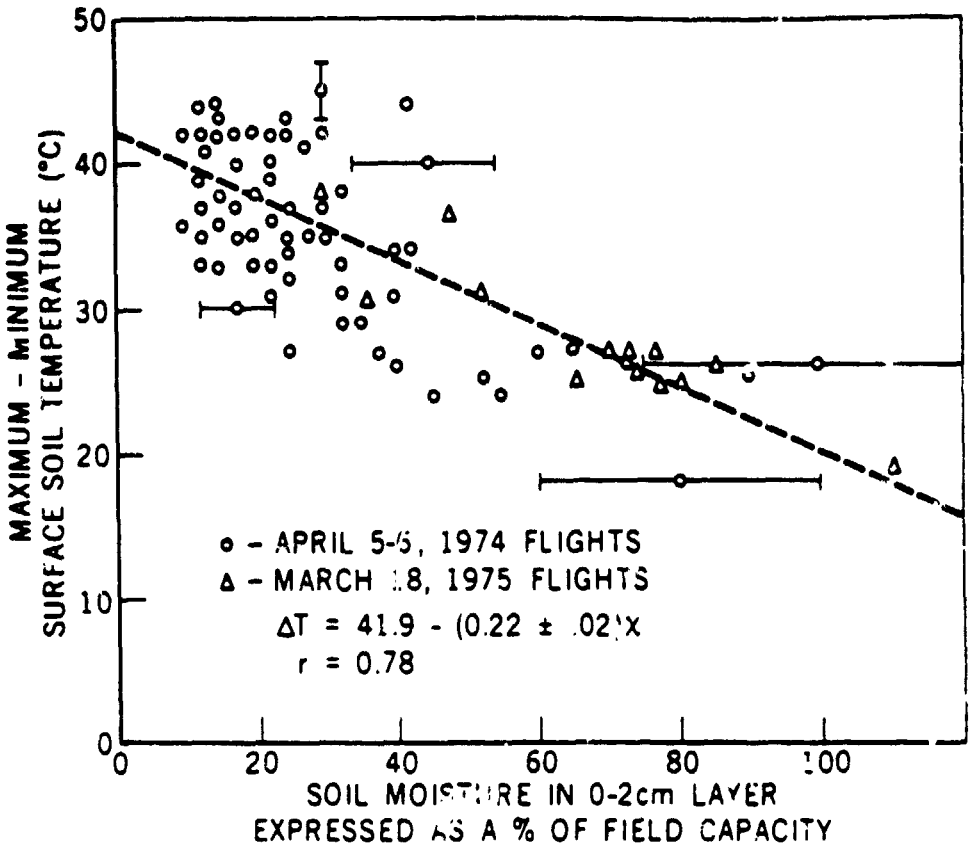


Figure 6. The Combined Data of the 1974 and 1975 Flights Over Phoenix.

When our 1975 aircraft results are compared with the total measurement set of Reginato, *et al.*, (1976) the agreement is very good. In Figure 1 we have plotted our data on Figure 2a of their paper. To compare the soil moisture values between the two experiments a value of 0.176 was taken as field capacity of the Avondale loam soil (Jackson, private communication). Our data from Figure 5 are plotted on their graph using the value 0.176 as 100% of field capacity. It is evident that the values of ΔT measured from aircraft data have the same variation with soil moisture as their combined data set acquired by several different means.

The aircraft observations of canopy temperature for several vegetated fields are presented in Table 1. The 2:00 p.m. temperature at the Phoenix airport was 21.7°C. The airport is located approximately 20 km east of the southern end of the flight line. The canopy temperatures were approximately within a degree or so of air temperature. This would appear to indicate a slight amount of moisture stress, which probably isn't the case since the 0-15 cm average soil moisture is above 50% of field capacity in each case. However, the two wheat fields with the highest temperature had the shortest plants, implying that some of the bare ground may be contributing to the observed emission.

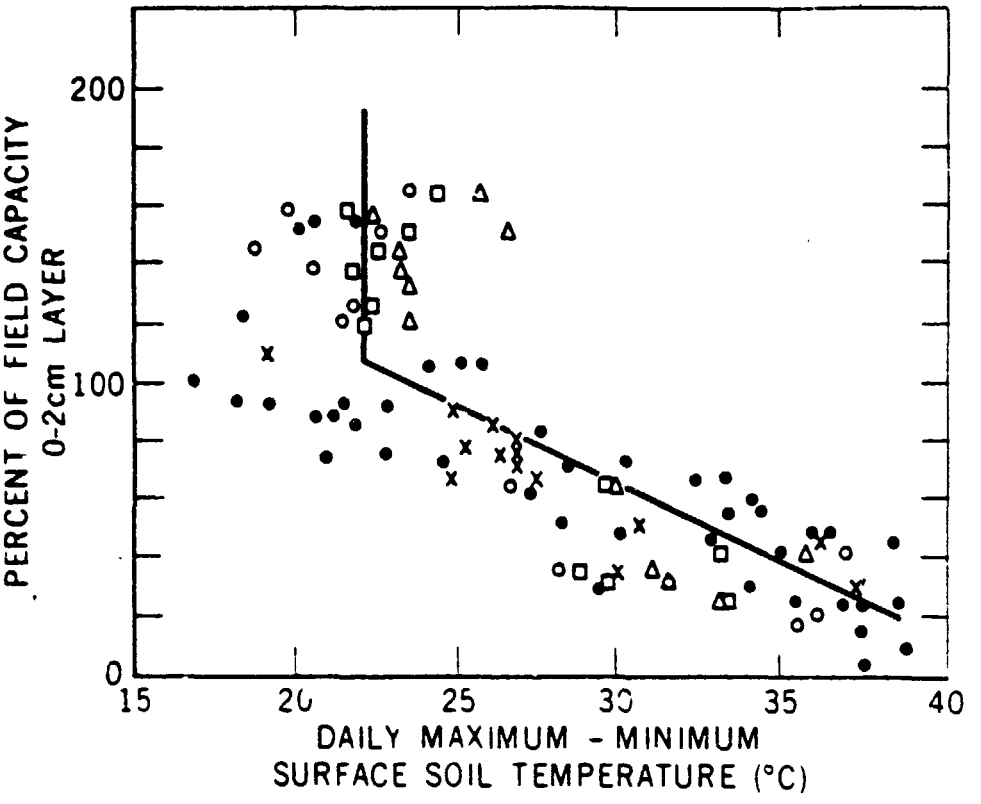


Figure 7. Plot of ΔT Data from this Paper Combined with that from Reginato, *et al.*, (1976) Versus Soil Moisture in the 0-2 cm Layer. (The symbols represent the different types of temperature measurement: ●, ○ surface thermocouple, □ hand held PRT-5, Δ aircraft data over their test plot, and × data from this paper.)

TABLE 1. Surface Temperature for Cropped Fields.

Field No.	Temp. °C	Crop Cover Type and Height	0-15 cm Soil Moisture	% of FC
97	23.1	alfalfa, 50 cm	22.0	70.5
106	24.1	wheat, 20 to 30 cm	29.0	94.5
111	21.3	wheat, 60 to 70 cm	27.3	87.5
112	22.7	wheat, 20 to 30 cm	28.8	96.0
113	21.5	wheat, 30 to 40 cm	27.5	85.0
114	22.6	wheat, 30 to 40 cm	22.6	69.5
115	24.1	wheat, 15 to 25 cm	24.8	89.0

DISCUSSION

These results have demonstrated the feasibility of using aircraft observations of the diurnal range of surface temperature for soil moisture sensing over a wider range of field conditions than those presented in the paper of Reginato, *et al.*, (1976). The comparison of remotely sensed canopy temperature with air temperatures as reported by the local weather station is a means of qualitatively assessing the state of moisture stress in the plant. This approach may be applied to determine the status of pasture grasses. Thus the condition of rangeland, which is distributed throughout the agricultural areas of the world, may be used as an index of the local soil moisture conditions. Indices of this nature would enhance opportunities for estimating crop yields and for improving current flood prediction technology.

There are many limitations on the applicability of this technique, e.g., clouds, atmospheric water vapor, and variations in surface slope, but there is the potential that frequent high resolution coverage would yield much useful data. This potential will be studied from spacecraft altitudes when NASA launches the Heat Capacity Mapping Mission satellite in the spring of 1978. The purpose of the mission is to measure the diurnal range of surface temperatures over large areas with an infrared radiometer operating in the 10-12 μm band and having a spatial resolution of 500 meters. These diurnal measurements will be repeated every eight days for most areas; daytime only coverage will be repeated every three days. Thus this sensor should acquire sufficient data to adequately test the applicability of this technique over wide areas.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Drs. H. Bouwer and R. Jackson of the U. S. Water Conservation Laboratory at Phoenix for allowing us to use their facility as a base of operation for the experiment and for allowing us the use of some of their unpublished data.

LITERATURE CITED

Blanchard, B. J., 1975. Ground Truth Report, 1975 Phoenix Microwave Experiment. Technical Report RSC-70. Remote Sensing Center, Texas A&M University.

Idso, S. B., T. J. Schmugge, R. D. Jackson, and R. J. Reginato, 1975. The Utility of Surface Temperature Measurements for the Remote Sensing of Soil Water Status. Journal of Geophysical Research, 80(21):3044-3049.

Idso, S. B. and W. L. Ehrler, 1976. Estimating Soil Moisture in the Root Zone of Crops: A Technique Adaptable to Remote Sensing. Geophysical Research Letters, 3(1):23-25.

Jackson, R. D., R. J. Reginato, and S. B. Idso, 1976. Timing of Ground Truth Acquisition During Remote Assessment of Soil-Water Content. Remote Sensing of Environment, 4:249-255.

Jackson R. D., R. J. Reginato, and S. B. Idso, 1977. Wheat Canopy Temperature: A Practical Tool for Evaluating Water Requirements. Water Resources Research 13(3):651-656.

Reginato, R. J., S. B. Idso, J. F. Vedder, R. D. Jackson, M. B. Blanchard, and R. Goettelmay, 1976. Soil Water Content and Evaporation Determined by Thermal Parameters Obtained from Ground-Based and Remote Measurements. Journal of Geophysical Research, 81(9):1617-1620.

Schmugge, T. J., B. J. Blanchard, W. J. Burke, J. F. Paris, and J. R. Wang, 1976. Results of Soil Moisture Flights During April 1974. NASA Technical Note TN D-8199.

WATER RESOURCES BULLETIN
VOL. 14, NO. 2 AMERICAN WATER RESOURCES ASSOCIATION APRIL 1978

SELECTING RECONNAISSANCE STRATEGIES FOR FLOODPLAIN SURVEYS[1]

Scott C. Sollers, Albert Rango, and Donald L. Henninger[2]

ABSTRACT: Multispectral aircraft and satellite data over the West Branch of the Susquehanna River were analyzed to evaluate potential contributions of remote sensing to floodplain surveys. Multispectral digital classifications of land cover features indicative of floodplain areas were used by interpreters to locate various floodprone area boundaries. The boundaries thus obtained were found to be more striking and continuous in the Landsat data than in the low altitude aircraft data. The digital approach permitted satellite results to be displayed at 1:24,000 scale and aircraft results at even larger scales. Results indicate that remote sensing techniques can delineate floodprone areas more easily in agricultural and limited development areas than in areas covered by a heavy forest canopy. At this time it appears that the remote sensing data would be best used as a form of preliminary planning information or as an internal check on previous or ongoing floodplain studies. In addition, the remote sensing techniques can assist in effectively monitoring floodplain activities after a community enters into the National Flood Insurance Program.
(KEY TERMS remote sensing; floodplain surveys; multispectral digital classification; planning information.)

INTRODUCTION

The concept of using the capabilities of remote sensing to assist in the management of floodplains has an understandable appeal because of the inherent complexity of conventional survey methods and the need to monitor conditions over extensive floodplain areas. Before 1972, aircraft remote sensor capabilities had been considered and generated some interest in the floodplain management community. This interest in remote sensing was further spurred in 1972 when the Earth Resources Technology Satellite, now referred to as Landsat, was launched. This satellite provided the previously unavailable capabilities of constant altitude and stability and reduced sun angle variability, while recording multispectral variations on a repetitive basis.

Coupled with this idea of somehow delineating a narrow floodplain boundary from afar (from a few kilometers altitude with aircraft to over 900 kilometers with Landsat)

[1]Paper No. 77084 of the *Water Resources Bulletin*. Discussions are open until December 1, 1978.
[2]Respectively, formerly with the U.S. Army Corps of Engineers, San Francisco District, San Francisco, California 94105, presently with Hornblower, Weeks, Noyes & Trask Inc., Public Finance Division, San Francisco, California 94111; Code 913, Goddard Space Flight Center, Greenbelt, Maryland 20771, and formerly with the Center for Remote Sensing of Earth Resources, Pennsylvania State University, University Park, Pennsylvania 16802, presently with Lockheed Electronics Company, Houston Texas 77058.

was the existing knowledge that historical, geomorphological, and botanical indicators associated with the boundary between floodprone and nonfloodprone areas could conceivably be recognized using remote sensing techniques. Further justification for investigating this potential grew out of the aftermath of the 1972 Hurricane Agnes floods in the eastern United States where several billions of dollars in damage occurred in floodplain areas. Emphasis on improving the National Flood Insurance Program resulted, and the Flood Disaster Prvention Act of 1973 was instrumental in making an increased number of communities eligible for flood insurance protection. With this increased participation the need for and backlog of floodplain surveys rose dramatically.

In 1968, when the flood insurance program was started, the U.S. Army Corps of Engineers (USACE) estimated that 5,000 communities were identified as floodprone. By May 1973 the Federal Insurance Administration (FIA) had increased that estimate to 10,000 communities, 13,600 by December 1973, and 16,300 by June 1974. A year later in June of 1975 the total identified floodprone communities had risen to 21,411 with little potential for further increase (Comptroller General of the United States, 1976). The result of this rapid increase in communities requiring surveys has been a staggering increase in associated costs. In fiscal year 1977 FIA will spend close to $75 million on the mapping program with $45 million apportioned for private contractors and $30 million for cooperating Federal agencies. The USACE will receive $15 million from FIA and will spend about $3 million additional for their own studies (G. Phippen, personal communication, 1976). It is expected that the entire mapping program, when completed, will exceed $1 billion. The costs for a flood insurance study currently average from $750 to $3,000 per stream kilometer (Comptroller General of the United States, 1976) depending on length of stream, sources of available data, basin configuration, and scheduling of tasks. Table 1 presents the average USACE costs per task involved in preparing a flood hazard map. The main objective of utilizing remote sensing in the mapping of floodplains would be either to reduce overall costs by assisting in the preparations of conventional products or to provide superior products at increased efficiency at similar costs.

TABLE 1. Average Costs Associated with Specific Tasks in Mapping Flood Hazard Areas.*

Task	Cost
Reconnaissance of Site	$300/km
Aerial Surveys	$500/km
Hydrology	$190/km
Hydraulics	$250/km
Profile Concurrence	$1,000/study
Land Surveys	$200–225/cross section
Coordination	$1,500/study
Travel	$2,000/study
Map Preparation	$5,000/study
Report	$2,500/study
Total cost to survey and prepare report for 32 km reach is $72,000 or $2,250/km. This assumes 16 km of detailed survey work with six cross sections per km.	

*Based on average costs to complete each task as calculated for previous studies by USACE, January 1976.

The purpose of this paper is to evaluate the current capabilities of remote sensing for floodplain management and to report on a specific case study regarding the feasibility of utilizing remotely sensed multispectral data to delineate floodplains. Because the 80 m resolution of Landsat is not entirely optimum for floodplain applications, higher resolution aircraft data were investigated in parallel as a cooperative effort by the USACE, the National Aeronautics and Space Administration (NASA), and the Pennsylvania State University.

RELATED RESEARCH

In-Situ Programs

Historical techniques involve observations of high water marks and flood damage related to specific floods. Several investigators have mapped flood lines based on trash accumulation, scarred trees, and sediment deposition (Leopold and Skibitzke, 1967; Sigafoos, 1964; and Lee, Parker, and Yanggen, 1972). Several limitations compromise the effectiveness of this technique. First, the flood which produced the evidence may have obliterated similar indications of earlier and less severe floods; and second, rare floods and their high-water marks may not have been observed on the stream which one desires to map.

Several investigators have relied on geomorphological features to indicate the location and frequency of flooding (Burgess, 1967; Reckendorf, 1963; and Wolman, 1971). The predominant feature used is the terrace but alluvial fans, natural levees, bars, oxbows, abandoned channels, marshes, deltas, and swales are also of significant value. These indicators, however, are of little more than local value and may not even be present from one watershed to the next.

Field investigators (Parsons and Herriman, 1970; McClelland, 1950; Coleman, 1963; and Woodyer, 1966) have shown that soils are configured horizontally and vertically to reflect flooding patterns. These patterns have been mapped by agencies such as the Soil Conservation Service (SCS) and have been used to delineate boundaries by Yanggen, Beatty, and Brovold (1966), McCormack (1971), Cain and Beatty (1968), and Viaene (1969). The difficulty in employing this approach is simply that many areas have not been mapped and that original field work is expensive and time-consuming.

Finally, vegetation has been shown to exhibit patterns related to flood conditions (Everitt, 1968; Wistendall, 1958; Helfley, 1937, Sollers, 1974; and Sigafoos, 1961, 1964). Various species possess tolerance to standing water or poorly-drained soils and are typically associated with the floodplain. Others require well-drained soils and are usually found in terrace locations. Broad size-groupings also are related to distance from the floodplain. The problem associated with using vegetation indicators is that, beyond fairly general relationships, distinct boundary delineation is difficult due to the capacity of vegetation species to tolerate and often flourish under less than optimum conditions. The result is a heterogeneous mixture that becomes difficult to interpret at close inspection. The best perspective for this approach is a distant one.

Floodplain Delineation Using Remote Sensing

Remotely sensed information from aircraft and satellites has been used to perform floodplain mapping by two complementary methods. The dynamic method images floods

as they actually occur or soon after the high waters have receded. It takes advantage of the fact that visible evidence of inundation in the near infrared region of the spectrum remains for up to two or more weeks after the flood. This evidence includes significantly reduced near infrared reflectivity in the flooded areas caused by the presence of increased surface-layer soil moisture, moisture stressed vegetation, and isolated pockets of standing water. Satellite data from Landsat 1 and 2 provide the most pertinent spacecraft information for flood observations because of the relatively high resolution, cartographic fidelity, and the near infrared sensors onboard. Flood mapping from Landsat photographic data has been reported by Hallberg, Hoyer, and Rango (1973), Deutsch and Ruggles (1974), and Rango and Salomonson (1974). Williamson (1974) has employed digital Landsat data for similar flood mapping. The compilation of a flood map from an actual event constitutes a floodprone map for that section of stream for a particular flood frequency. This dynamic map can be continually improved as additional floods are observed on the stream in question.

The second method, referred to as the static approach, utilizes the fact that many floodplains have been recognized with remote sensing because of permanent or long-term features formed by historical floods. These natural and artificial indicators have been enumerated by Burgess (1967). Floodprone areas also tend to have multispectral signatures that are distinctly different that those of surrounding nonfloodprone areas. Harker (1974) performed a multispectral analysis of digitized aircraft photography in Texas that indicated a reasonable correlation between floodprone area boundaries based on computer processed multispectral digital data and those produced by conventional techniques. A combination of high altitude aircraft and satellite data were used to produce floodprone area maps to meet Arizona state legislature requirements set forth in House Bill 2010 (Clark and Altenstadter, 1974). Rango and Anderson (1974) used Landsat exclusively to provide small-scale floodprone area maps in the Mississippi River Basin that compared favorably with existing surveys.

Flood and floodprone area observations from Landsat are indeed promising, but only on a regional basis. Most satellite photographic flood and floodplain mapping has been done at scales no larger than 1:250,000. Digital Landsat maps of floods and floodplains have been produced at 1:24,000 and 1:62,500 scales, but they do not meet national map accuracy standards. For most legal requirements, it is necessary to produce maps at even larger scales.

CENTRAL PENNSYLVANIA STUDY

Test Sites

Capitalizing on previous work, the USACE elected to pursue the potential of mapping floodplains based on natural indicators. The study area was a portion of the West Branch of the Susquehanna River in north central Pennsylvania as shown in Figure 1. The test site was selected because it exhibited a variety of land cover types including residential, light industry, agriculture, and forest. Physiography is characterized by both steep and gently sloping valleys and floodplains of varying widths. The section of the West Branch of the Susquehanna River from point A to point B (Figure 1) is in the Ridge and Valley Province, where the valley is broad with a moderately wide floodplain predominantly used for agriculture, and will subsequently be referred to as the "agricultural and developed"

area. The section of the river from point B to point C (Figure 1) is in the Allegheny Plateau Province, where the large forested valley is steep with a narrow floodplain. This portion of the area will subsequently be referred to as the "forested" area.

The study area has a humid continental climate (U.S. Department of Agriculture, 1966) with warm summers and long cold winters. The average annual temperature is 10.7°C, with January and July mean temperatures of −14°C and 23.1°C, respectively. The average annual precipitation is about 102 cm, which includes an average total seasonal snowfall depth of 94 cm.

Data Sources

Ground Data – The Flood Plain Information report (U.S. Army Corps of Engineers, 1973) prepared for the West Branch of the Susquehanna River was used to obtain the floodplain limits, established on the basis of engineering parameters, for comparison with those limits established using remotely sensed data. Maps of the West Branch of the Susquehanna River at 1:24,000 scale, showing the extent of the 100-year return period flood as well as the extent of flood waters during Hurricane Agnes in 1972, were provided by the USACE.

Among other sources of information available for the study area were the USGS 7.5-minute quadrangle topographic sheets. The Soil Survey of Clinton County, Pennsylvania (U.S. Department of Agriculture, 1966), as well as the SCS worksheets for the Lycoming County Soil Survey (in progress) provided soils information. Various geologic maps and reports were also consulted (Stose and Ljungstedt, 1932; Flint, 1947; MacClintock and Apfel, 1944; and Peltier, 1949).

An extensive field analysis of the entire test region was conducted in July 1973. River banks and terraces were inspected to determine vegetation species type and composition, bare soil texture, and drainage of the floodplain to facilitate the calculation of spectral signatures.

Aircraft Data – The NC130B aircraft of NASA flew the test area at altitudes of approximately 1525 meters (5000 feet) and 4575 meters (15,000 feet) in April and June 1973. Color positive and color infrared photography was taken, along with data from 14 channels of the Bendix 24-channel multispectral scanner.

A multispectral scanner (MSS) is an optical-mechanical scanning device used to detect levels of electromagnetic energy emanating from the earth's surface in many discrete wavelength intervals (channels). Through the use of a rotating mirror, the area beneath the aircraft is scanned in successive contiguous lines in a direction perpendicular to the flight of the aircraft. The energy received from the earth's surface is reflected by the mirror through a series of lenses and prisms which refract the energy into components of selected wavelengths onto an array of detectors. Each detector then produces an electrical output signal proportional to the energy received. These signals can then be used to modulate a light source to expose photographic film, or they can be recorded on magnetic tapes for later analysis.

At the 1525 meter altitude, the size of a ground resolution element (or pixel) of the 24-channel MSS is approximately 3 meters on a side, depending on minor aircraft altitude and velocity variations. At the 4575 meter altitude, the size of a ground resolution element or pixel is approximately 9 meters on a side. NASA provided the computer compatible tapes, containing the uncalibrated digitized MSS data, as well as imagery of selected channels.

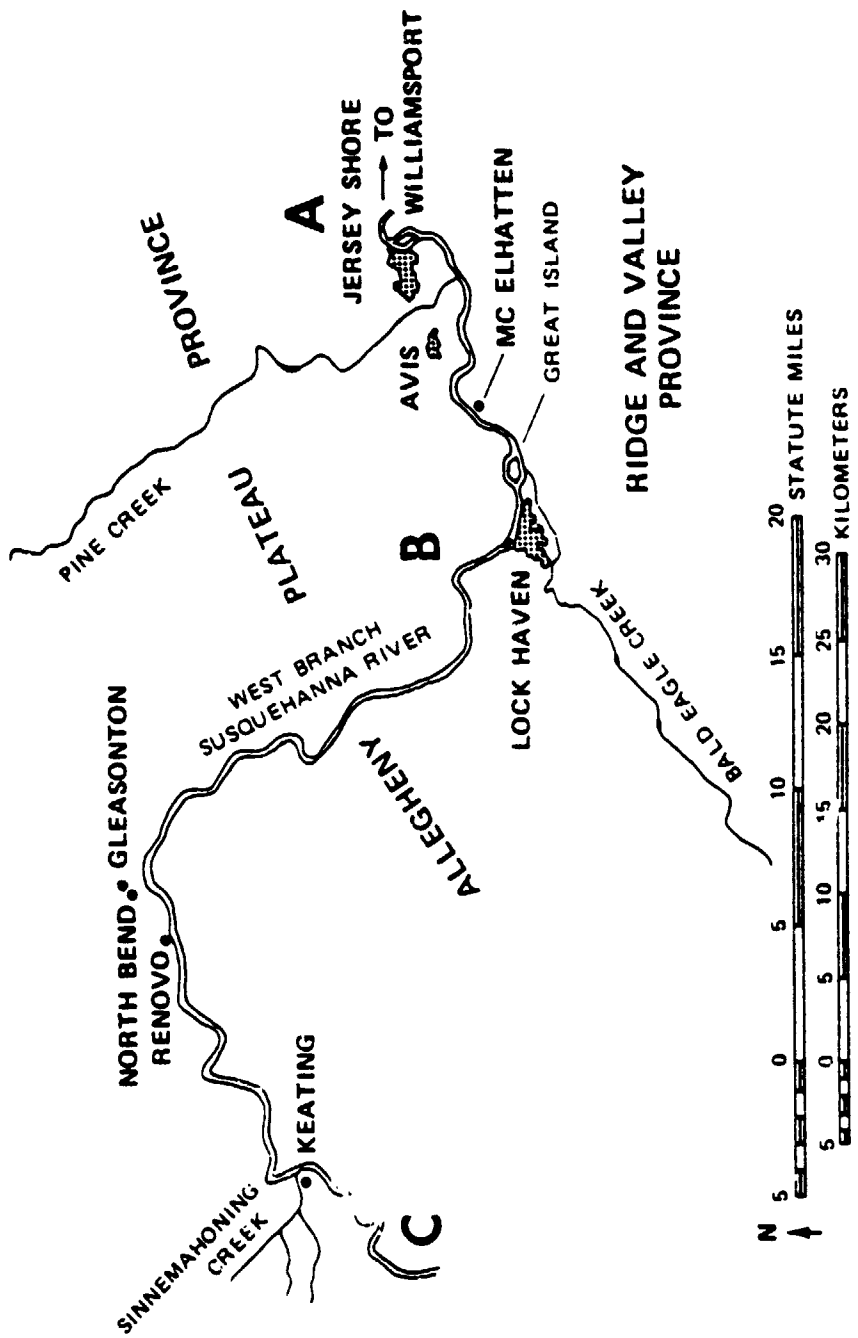


Figure 1 Central Pennsylvania Study Site Along the West Branch of the Susquehanna River.

Satellite Data – Most observations by Landsat have been taken with the MSS in the visible and near infrared wavelengths. The four MSS channels (discrete wavelength intervals) cover the 0.5–0.6, 0.6–0.7, 0.7–0.8, and 0.8–1.1 μm portions of the electromagnetic spectrum. The resolution of the Landsat MSS is approximately 80 m, although smaller features may sometimes be detected because of favorable geometric and contrast characteristics of a given object on the earth's surface.

In the agricultural and developed portion of the study area (from point A to point B, Figure 1), data from the 16 May and 25 October 1973 scenes (identification numbers 1397-15245 and 1459-15221, respectively) were selected for analysis. These scenes were selected to obtain the maximum area of exposed bare soil. Data from the 6 September 1972 scene (identification number 1045-15240) were selected for analysis of the forested portion of the study area (from point B to point C, Figure 1). This scene was selected for maximum expression of tree foliage. All Landsat MSS data were supplied by NASA in the form of computer compatible tapes as well as imagery of selected channels.

Approach

The Penn State Office of Remote Sensing of Earth Resources (ORSER) system for analyzing multispectral scanner data is based on multivariate statistical techniques. Each observation, identifiable by scan line and element number, consists of a vector composed of multispectral scanner response values with as many components as there are channels. The programs used in this study are all operational and are documented at the user level (Borden, *et al.*, 1975).

The first step is to select the particular targets and areas of interest and the computer tapes corresponding to these areas. A subset of data is then produced for the specific area of interest. The following step is to produce a brightness map employing all available channels which can be used for verifying general location and identifying specific targets. No previous knowledge of target spectral signatures is required for the brightness map.

Subsequently, a program is employed to identify areas of local spectral uniformity based on variation between spectral signatures of near neighbors as "the measure of similarity." The output shows the pattern of uniformity and contrasts from which the user can designate coordinates of training areas to be used for developing multispectral signatures for use in supervised classifying routines. Multivariate statistics of the signatures are then calculated for the training areas. Using these statistics, supervised classification and mapping can be done for the entire study area. The output is a digital character map with each category of classification represented by a unique symbol assigned by the user. Unsupervised classification or clustering options may have to be employed in combination with supervised classification to effectively classify small area or linear features, such as streams.

The ORSER system has the capability to perform a geometric correction on a character map to rectify simple distortions resulting from sensor, satellite, and earth effects. These corrected classification maps can be overlaid on other maps of the same scale, such as 1:24,000 scale topographic maps. This scaling feature facilitates the comparison of MSS (aircraft or Landsat) classification results with available ground truth.

Once classification of land cover has been accomplished, a number of interpreters attempted to use the classified results to determine which classes were indicative of flood-prone areas. In the agricultural and developed region, bare floodplain soils were used as

the key feature for drawing the floodprone area boundary line. In the forested area, different vegetation classifications were related to the floodprone areas. Attempts to draw in the floodprone area boundaries were made using classified data only and then with the addition of ancillary data such as topographic maps.

Results

Floodplain Classification Using Digital Aircraft Data – Several sites were selected to test the applicability of digital aircraft data to floodplain mapping. The two major categories of interest were agricultural areas, especially those with bare soils, and forested areas. Initial individual test site selection was based on various physical characteristics of the areas, such as vegetation, topography, and (in the case of forested test sites) freedom from alterations due to the activities of man (such as housing developments) within the recent past.

The results from computer analysis of the digital MSS data in agricultural areas distinguished between floodplain and nonfloodplain areas in small isolated portions of the test sites. In addition, the computer classification within one test site separated an area of moderately well-drained soil from the surrounding well-drained soils. Comparison with SCS data for this area indicated the moderately well-drained soil to be less extensive than shown on the computer output, but field inspection supported the results of the digital classification. In general, the results from computer analysis of the digital MSS data in agricultural areas were not sufficiently conclusive to delineate a continuous floodplain line. The presence of extensive bare soils in the agricultural and developed area made the April data more useful than June data in detecting soil differences.

Results using the MSS data from the 1925 meter June flight indicate that classification differences in the forested area can be related not only to species differences, but also to differences in crown closure or canopy density within a forested area. These density differences may be a result of species composition, site quality, or a cultural practice. Local topography and the aspect of the test sites also contributed to classification differences unrelated to the natural vegetation. A comparison of the classification results of each of the available data sources (i.e., the April and June flights at both 1525 and 4575 meter altitude) showed the MSS data collected during June to be more effective than the April data owing to greater vegetative cover existing in June. Classification differences obtained using the April data were additionally ambiguous because of extensive shadow patterns which were prevalent over much of the study area at the time of this particular flight. The June MSS data at an altitude of 1525 meters were the only data which, when used for classification, were useful in delineating an area related to the floodplain.

Floodplain Classification Using Digital Landsat Data – Preliminary analysis of the 16 May and 25 October 1973 Landsat scenes indicated both to be potentially good data sets for the purposes of differentiating floodplain bare soils from nonfloodplain bare soils. Previous experience has shown that by merging tapes from two different seasons it is often possible to improve the classification of certain targets, such as hardwoods and conifers (using data merged from summer and early winter scenes). Therefore, the two scenes were merged and treated as an 8-channel data set on this basis. Merged tapes can be used in any of the ORSER programs, using the same analytical procedures employed for a 4-channel subset tape.

The results of computer analysis of eight channels of Landsat MSS data merged from two scenes were sufficiently conclusive to delineate a continuous floodplain boundary in the agricultural and developed area, which was then quantitatively compared to the USACE 100-year return period floodplain boundary. The basis for the distinction between the boundaries of the floodplain and nonfloodplain areas was spectral differences in the bare soils of the two areas, which could be differentiated using the available computer routines.

Four individuals independently interpreted the resulting classification map. Each interpreter first delineated a floodplain solely on the basis of the classification map and then was allowed to consult the corresponding USGS 7.5-minute quadrangle topographic maps to refine the interpretation. Each interpreter's floodplain delineation was compared to the USACE 100-year return period floodplain boundary in two ways: (1) on the basis of total area measured to be floodprone, and (2) by calculations of the correlation coefficient between the pairs of measured distances from the center line of the river to the two floodplain lines on both sides of the river. For this second comparison a total of 100 pairs of measurements were made.

Results of the two methods of comparison for each of the four interpreters are shown in Table 2. The area of the USACE 100-year return period floodplain plus river is 4200 ha (10,371 acres). Using only the computer classification map, the results range from a 13.7% (574 ha) underestimation of the USACE floodplain to a 4.4% (184 ha) overestimation. Additional comparisons made using information on the USGS 7.5-minute quadrangle sheets improved the underestimation percentages to 7.7% (324 ha); however, the overestimation was increased to 8.8% (369 ha).

TABLE 2. Floodplain Boundary Comparisons for the Agricultural and Developed Portion of the Study Area.

	Area Comparisons (percent)[1]		Correlation Coefficients[2]	
Interpreter	Without Aid of Topographic Maps	With Aid of Topographic Maps	Without Aid of Topographic Maps	With Aid of Topographic Maps
1	−8.6	−0.4	0.85	0.92
2	+4.4	+8.8	0.86	0.92
3	−2.2	−0.8	0.86	0.90
4	−13.7	−7.7	0.83	0.87

[1] The total area of the floodplain and river within the 100-year return period floodplain determined by the U.S. Army Corps of Engineers was 4200 hectares (10,371 acres). A positive value here indicates an overestimation and negative value an underestimation.

[2] Based on 100 observations with all correlation coefficients highly significant.

Correlation coefficients representing the "nearness of fit" of the computer classification map and the USACE 100-year floodplain aided by information gleaned from USGS maps were calculated for each interpreter. These coefficients ranged from 0.87 to 0.92 (Table 2), indicating a rather close association of the two floodplain delineations. A scattergram illustrating the relationship between the two floodplain delineations, using

one interpreter's results, is shown in Figure 2. In this case, the horizontal differences between the two floodplain delineations ranged from 0 to 74 meters with a mean-difference of 13 meters.

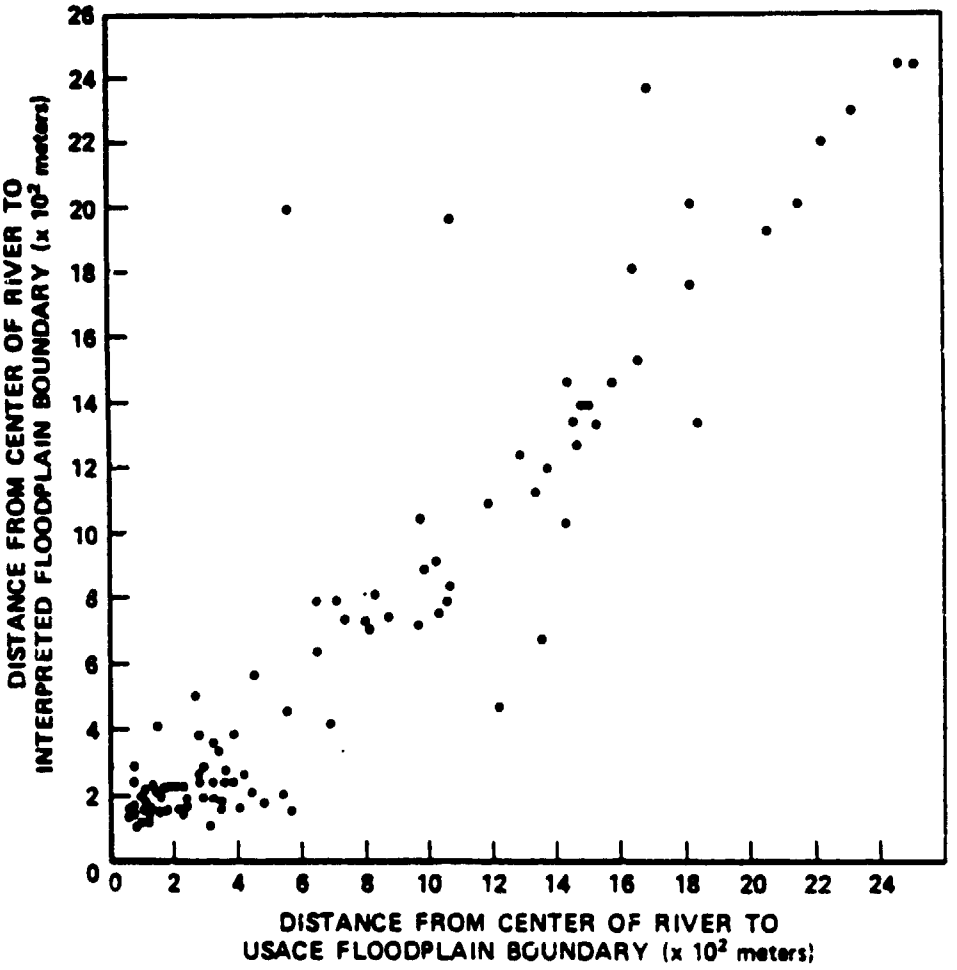


Figure 2. Scattergram Comparing the USACE and Landsat Floodplain Boundaries for the Agricultural and Developed Area (Interpreter 1).

A substantial lack of correlation between the USACE floodplain limit and the floodplain delineation based on the computer classification map existed in two situations. In the first case, small isolated areas in the floodplain having vegetated or developed land cover were not recognized as floodplain in the computer classification. This was not totally unexpected since the classification map was developed primarily on the basis of differentiating floodplain from nonfloodplain bare soils. Interpretative skills played an important role in overcoming the lack of information in these relatively small areas.

The second situation in which a discrepancy appeared between the two lines occurred where the computer classification identified "floodplain" areas based on the presence of bare soils while the USACE considered them outside of the floodplain. The elevation of these areas is only slightly higher than the adjacent floodplain area and could be remnants of old river terraces. Field inspection revealed a strong possibility that inundation by flooding waters could occur since the meander of the river channel would allow flood waters to flow in the direction of the areas in question. Based on the results of this field inspection, the three cross-sectional observations in this area were not included in the calculation of the range and mean but were included in the calculation of the correlation coefficients.

As opposed to the 8-channel data set used in the agricultural and developed area analysis, the delineation of the floodprone area in the forested region of the study area was based on classification results from a set of spectral signatures developed from only three channels of Landsat-1 data on a single date using the unsupervised classification approach. The resultant nine signatures categorized as floodprone do not correlate with individual floodplain features, such as floodplain bare soil or floodplain vegetation, but represent a variety of physical features associated with the floodplain in this area.

The separation of open and developed areas from forested areas accounts for a major portion of the floodplain classification. Generally, these open areas are wide and flat terraces adjacent to the river. In several instances upland open areas were incorrectly classified as floodplain. These areas can be readily identified as nonfloodplain because of their position relative to areas classified as river and because their shape suggests a feature other than a floodplain. This discrepancy was easily accounted for during interpretation of the classified data.

An attempt at improving the delineation of a floodprone area boundary in forested areas on the classification map was made by three interpreters in the same fashion as performed in the agricultural areas. In all cases the interpreted floodplain area overestimated the USACE floodplain area ranging from 751 ha (38%) to 1427 ha (72%). This overestimation, however, may indicate that the interpreted line identifies a boundary that represents a higher flooding frequency than the 100-year flood. The correlation coefficients between the USACE floodprone boundaries and the interpreted boundaries (0.29–0.35) in the forested sites were much lower than for the agricultural areas. There is a strong correlation, however, between each of the interpreted boundaries (0.82–0.90), indicating a high degree of repeatability. On several sites where the interpreted line grossly overestimated the USACE 100-year flood line, inspection of aerial photos revealed close proximity of mountain streams that would affect and enlarge the previously mapped floodplain. In general, however, floodplain delineation in the agricultural and developed areas was much easier than similar delineation in forested areas.

Discussion

Aircraft – A continuous floodplain line could not be delineated on the basis of computer analysis of the aircraft collected MSS data. However, the computer analysis did indicate a break between floodplain and nonfloodplain within small areas which correlated with one or more floodplain limits derived from USACE maps, soils data, and USGS sources.

The inability to consistently map a floodplain boundary using aircraft data based on natural indicators was due to several factors regarding the data collection medium and the

study area. The study area has a very complex topography and many land cover types. The slopes range from nearly level to quite steep, with greatly varying aspects. The land cover types include urban and residential areas, small agricultural fields, and heterogeneous forest stands. Research in more uniform study sites has yielded considerably more successful results. The impact of variable terrain and land use on pattern recognition is profound.

No attempt was made at a sensitivity analysis of spectral band selection to discern only those natural features that are strongly associated with flood frequency. The decision was made to proceed with a multispectral analysis since this approach was the strength of the ORSER software. There may well be, however, features in nature that are more easily identified by selective band processing and this approach should be conducted in any follow-on investigation, especially if similar terrain is selected.

The method and type of aircraft digital data collection system were also factors affecting the potential for successful findings in this portion of the investigation. The data volume alone yielded enormous bits of information that had to be arrayed, reformatted, screened, catalogued, and analyzed. In this regard the satellite technique is preferable in that significantly less data are required to cover the same size area. The aircraft platform, being subject to variable atmospheric buttressing, will occasionally render irregular land parcel data. This phenomenon has a serious impact on the transferability of signatures derived in one location to another. In addition, the platform, because of the narrow field of view afforded, needs to be directed precisely over the target to obtain radiometrically accurate results. In many cases during this study, continuation of analysis along a floodplain segment was aborted due to lack of coverage. This problem would have been obviated by using shorter flight lines to accommodate stream meander patterns.

The supplemental, high quality, aerial photography collected for this study was extremely useful as a source of ground truth to which computer classification results of the digital MSS data could be easily compared. It was also used to identify areas which required more intensive on-site investigation.

Satellite -- As opposed to the discontinuous, aircraft-derived floodplain boundary, satellite data afforded a continuous boundary that could be statistically compared to a boundary based on engineering parameters. The comparison revealed strong agreement in the agricultural and developed sites and indicated a marked overestimation in forested sites. The discrepancy between the two areas may be due in part to the fact that two scenes of four channels each were merged and used as an eight-channel set in the agricultural and developed study area, whereas data from a single scene consisting of only three good quality channels were used in the forested study area. Thus greater differentiation potential existed for the agricultural and developed study area. It is also possible that the remote sensing boundary identified in the forested area was an indicator of the limit of a flood with a return period greater than 100 years.

It appears that Landsat digital MSS data is superior to low altitude aircraft-collected digital MSS data for floodplain mapping. However, this should not be taken as a recommendation of an optimum altitude for data collection, but only as a comparison of data from the two altitude ranges available. Based on past experience with high altitude aircraft imagery, it is possible that MSS data from such a platform would be directly amenable to floodplain identification and mapping.

SUMMARY AND CONCLUSIONS

The central Pennsylvania study differed from most of the previously mentioned flood-prone area studies in that the approach primarily employed multispectral classification techniques using digital MSS aircraft and satellite data. This approach not only makes maximum use of the resolution capabilities of the sensor systems but also contributes to objective interpretation of the floodprone areas. The analysis of these floodprone areas is not fully automatic, however, but requires the attention of an expert in remote sensing who is familiar with floodplain areas to allow for the final interpretation and location of the floodprone boundary. The Pennsylvania State ORSER digital approach permitted Landsat results to be displayed at a convenient 1:24,000 scale and aircraft data at even larger scales.

In this case study it appeared that the high resolution of the low altitude aircraft survey detracted from the identification of the floodprone area boundary. The small pixel size of aircraft data resulted in an overabundance of detail which camouflaged the detection of the subtle floodplain boundaries in many areas. In contrast, the lower resolution of the Landsat MSS data seemed more effective in delimiting floodplain boundaries in complex areas because the larger pixel size integrated over a number of specific features to come up with a single radiance value. When compared to other radiance values, the satellite-derived floodprone area boundaries tend to stand out more predominantly than they do in the low altitude aircraft data.

Based on the results from remote sensing floodplain studies, including the central Pennsylvania study, several comments on the general suitability of using remote sensing data in floodplain management can be made. Digital multispectral scanner data and automatic digital analysis combined with a certain amount of user interpretation is to be preferred over conventional photointerpretation. It appears that remote sensing analysis can delineate floodprone areas best in agricultural and limited development areas as opposed to areas covered by a heavy forest cover. Visible and near infrared channels are recommended for analysis.

The digital data analysis can produce flood hazard boundary maps at a usable scale for rural areas, namely, 1:24,000. Similar maps at scales useful to urban areas must await improved resolution MSS data from space or from high altitude aircraft data. Even if the maps were at the appropriate scales, it is unlikely that floodplain management agencies would immediately adopt this new procedure. Rather, it is more likely that the remote sensing technique would be used as a form of preliminary planning information or as an internal check on previous or ongoing floodplain studies. Remote sensing can be used as another form of local knowledge and could be important in identifying areas where major discrepancies in the conventional map may exist and further surveys are merited.

Once a survey has been completed, whether it be conventional or remote sensing based, the remote sensing data can provide detailed land use analysis in floodprone areas that can serve as a base for assessment of potential flood damage. The enforcement phase of the National Flood Insurance Program has received little attention to date because of severe FIA manpower limitations (Comptroller General of the United States, 1976). The Comptroller General report recommends in addition to improving community involvement in the enforcement of floodplain management regulations, the FIA "provide a means of systematically spot checking community compliance with program requirements." It is quite conceivable that remote sensing can assist in effectively monitoring floodplain activities. Even at Landsat resolution, major and minor land use changes in

C - E

floodplain areas can be detected. Such an enforcement system would permit a check on reporting of the local communities in the program.

The continued acquisition of remote sensing data over the United States will serve to record actual flooding events on an increasing number of streams. Such data will increase the availability of actually observed flooded area maps and update flood hazard boundary maps where they already exist. Continued remote sensing research in floodplain management should concentrate on attaining the optimum resolution with multispectral sensors, whether from high altitude aircraft or space platform. Finally, as the enforcement phase of the National Flood Insurance Program receives increasing emphasis, the capabilities of remote sensing should receive serious consideration as an integral part of the program.

ACKNOWLEDGMENTS

This study was conducted as a joint venture under the auspices of the U.S. Army Corps of Engineers (Contract No. DACW 73-74-C-0036) and the National Aeronautics and Space Administration (Contract No. S-50960-G). The case study was performed by the Pennsylvania State University, Office for Remote Sensing of Earth Resources under the direction of Drs. G. W. Peterson and G. J. McMurtry. Special thanks is given to M. L. Stauffer and T. E. Walton, III, for conducting essential portions of the case study. We would also like to thank Dr. V. V. Salomonson for his helpful review of this paper in draft form.

LITERATURE CITED

Borden, F. Y., D. N. Applegate, B. J. Turner, H. M. Lachowski, and J. R. Hootsy, 1975. ERTS and Aircraft Multispectral Scanner Digital Data Users Manual. ORSER-SSEL Technical Report 10-75, Office for Remote Sensing of Earth Resources, The Pennsylvania State University, University Park, PA., 218 pp.

Burgess, L. C. N., 1967. Airphoto Interpretation as an Aid in Flood Susceptibility Determination. International Conference on Water for Peace, Washington, D.C., 4:867-881.

Cain, M. and T. Beatty, 1968. The Use of Soil Maps in the Delineation of Flood Plains. Water Resources Research, 4(1):173-182.

Clark, R. B. and J. Altenstadter, 1974. Application of Remote Sensing Techniques in Land Use Planning: Floodplain Delineation. Office of Arid Lands Studies Bulletin 6, University of Arizona, Tucson, 19 pp.

Coleman, C. S., 1963. How Fairfax County Tackles Soil and Land Use Problems. Agr. Engineering, 44:614-615.

Comptroller General of the United States, 1976. Formidable Administrative Problems Challenge Achieving National Flood Insurance Program Objectives. Report to the Congress, PED-76-94, U.S. General Accounting Office, 81 pp.

Deutsch, M. and F. Ruggles, 1974. Optical Data Processing and Projected Applications of the ERTS-1 Imagery Covering the 1973 Mississippi River Valley Floods. *Water Resources Bulletin*, 10(5):1023-1039.

Everitt, B. L., 1968. Use of the Cottonwood in an Investigation of the Recent History of a Floodplain. Amer. J. Science, 266:417-439.

Flint, R. F., 1947. Glacial Geology and the Pleistocene Epoch. John Wiley and Sons, 589 pp.

Hallberg, G. R., B. E. Hoyer, and A. Rango, 1973. Application of ERTS-1 Imagery to Flood Inundation Mapping. Proceedings of the Symposium on Significant Results Obtained from the Earth Resources Technology Satellite-1, NASA SP-327, 1:745-753.

Harker, G. R., 1974. The Delineation of Floodplains Using Automatically Processed Multispectral Data. Tech. Report RSC-60, Remote Sensing Center, Texas A&M University, College Station, TX, 227 pp.

Helfley, H. M., 1937. Ecological Studies on the Canadian River Flood Plain in Cleveland County, Oklahoma. Ecol. Monographs, 7:347-398.

Lee, G. B., D. E. Parker, and D. A. Yanggen, 1972. Development of New Techniques for Delineation of Flood Plain Hazard Zones, Part I: By Means of Detailed Soil Surveys. Partial Technical Completion Report to the Office of Water Resources Research, University of Wisconsin, Water Resources Center, Madison, Wisconsin, 77 pp.

Leopold, L. B. and H. E. Skibitzke, 1967. Observations on Unmeasured Rivers. Geogr. Annals, 49:247-255.

MacClintock, P. and E. T. Apfel, 1944. Correlation of the Drifts of the Salamance Re-entrant. Geological Society of America Bulletin, New York, 55(10):1143-1164.

McClelland, J. E., E. M. White, and F. F. Riecke, 1950. Causes of Differences in Soil Series of Missouri River Bottomlands in Monora County, Iowa. Acad. Sci. Proceedings, 57:253-258.

McCormack, D. E., 1971. Use of Soil Surveys in the Identification of Flood Plains. Ohio J Science, 71:370-375.

Parsons, R. B. and B. C. Herriman, 1970. Haploxerolls and Agrixerolls Developed in Recent Alluvium, Southern Williamette Valley, Oregon. Soil Sci., 109:299-309.

Peltier, L. C., 1949. Pleistocene Terraces of the Susquehanna River, Pennsylvania. Bulletin G23, Pennsylvania Topographic and Geologic Survey, 158 pp.

Rango, A. and A. T. Anderson, 1974. Flood Hazard Studies in the Mississippi River Basin Using Remote Sensing. *Water Resources Bulletin*, 10(5):1060-1081.

Rango, A. and V. V. Salomonson, 1974. Regional Flood Mapping from Space. Water Resources Research, 10(3):473-484.

Reckendorf, F. F., 1973. Techniques for Identifying Flood Plains in Oregon. Ph.D. Dissertation, Oregon State University, Corvallis, Oregon, 344 pp.

Sigafoos, R. S., 1961. Vegetation in Relation to Flood Frequency Near Washington, D.C. U. S. Geological Survey Prof. Paper 424-C, Washington, D.C., pp. 248-250.

Sigafoos, R. S., 1964. Botanical Evidence of Floods and Flood-Plain Deposition. U.S. Geological Survey Prof. Paper 485-A, Washington, D.C., 36 pp.

Sollers, S.C., 1974. Substrate Conditions, Community Structure, and Succession in a Portion of the Floodplain of Wissahickon Creek. Bartonia, 42:24-42.

Stose, G. W. and O. A. Ljungstedt, 1932. Geological Map of South-Central Pennsylvania. Pennsylvania Topographic and Geologic Survey.

U.S. Army Corps of Engineers, 1973. Flood Plain Information, West Branch of the Susquehanna River and Bald Eagle Creek (Lock Haven area), Clinton County, Pennsylvania. Prepared by the Baltimore District, Corps of Engineers, Baltimore, Md., 47 pp.

U.S. Department of Agriculture, 1966. Soil Survey of Clinton County, Pennsylvania. U.S. Soil Conservation Service, 141 pp.

Vlaene, R. M., 1969. Evaluation of Floodplain Delineation Based on Soil Maps in the Turtle Creek and Kickapoo River Watersheds, Wisconsin. Unpublished M.S. Thesis, University of Wisconsin, Madison, Wisconsin, 97 pp.

Williamson, A. N., 1974. Mississippi River Flood Maps from ERTS-1 Digital Data. *Water Resources Bulletin*, 10(5):1050-1059.

Wistendahl, W. A., 1958. The Flood Plain of the Raritan River, New Jersey. Ecol. Monographs, 28: 129-153.

Wolman, M. G., 1971. Evaluating Alternative Techniques of Floodplain Mapping. Water Resources Research, 7(6):1383-1392.

Woodyer, K. D., 1966. Bankfull Frequency in Rivers. J. Hydrology, 6:114-142.

Yanggen, D. A., M. T. Beatty, and A. J. Brovold, 1966. Use of Detailed Soil Surveys for Zoning. J. Soil and Water Conservation, 21:123-126.

Paper 103

Reprinted from JOURNAL OF APPLIED METEOROLOGY, Vol. 18, No. 9, September 1979
American Meteorological Society
Printed in U. S. A.

Relation Between West Coastal Rainfall and Nimbus 6 SCAMS Liquid Water Data over the Northeastern Pacific Ocean

W. VIEZEE AND H. SHIGEISHI

SRI International, Menlo Park, California 94025

A. T. C. CHANG

NASA, Goddard Space Flight Center, Greenbelt, Maryland 20771

(Manuscript received 7 March 1979, in final form 18 May 1979)

ABSTRACT

We describe a research study in which we explored the application to rainfall prediction of cloud liquid water data obtained from the SCAMS experiment of Nimbus 6. The study area is the Pacific Northwest coast of the United States, where rainfall is produced by extratropical storms that approach from across the Pacific Ocean.

SCAMS data related to cloud liquid water over the ocean, and coastal rainfall data, are analyzed for 20 different storm systems in the northeastern Pacific Ocean; these produced significant rainfall from Washington to central California during the period October 1975–March 1976. Results show that the distribution of storm-cloud water analyzed from the SCAMS data over the ocean foreshadows the distribution of coastal rainfall accumulated from the storm at a later time.

We conclude that passive microwave sensor measurements of cloud water over the ocean, when used in conjunction with numerical and other objective guidance, can be used to enhance the accuracy of predictions of coastal rainfall distribution.

Limitations in the SCAMS measurements and in the data analysis and interpretation are noted.

1. Introduction

In the Pacific Northwest west of the Cascades, rapid increase in mountain snowpack or in runoff can be predicted reliably only if precursor meteorological conditions can be diagnosed over the northeastern Pacific Ocean, where little or no conventional weather data exist.

At present, the regional Limited-Area Fine-Mesh Model (LFM) of the National Meteorological Center (NMC) predicts precipitation for the West Coast. However, as pointed out by Fawcett (1977), the model is less accurate for the western United States than for other geographic areas, due to the absence of observations over the eastern Pacific and the failure to correctly model atmospheric convection and effects of complicated terrain on the atmosphere. The output of the numerical prediction models is currently complemented by objective precipitation forecasts from the Model Output Statistics (MOS) technique and the map-type Probabilities of Precipitation (PoP's) method (Klein and Glahn, 1974; Rasch and MacDonald, 1975); also, cloud-image data from satellites are increasingly used in forecast preparation. It remains difficult, however, to identify those storm systems over the ocean that produce surges of heavy precipitation over land, especially between October and April, which is the time of greatest rainfall.

Passive microwave sensors carried on research satellites provide measurements inside extensive storm-cloud systems that the visible and infrared radiometers cannot obtain. For example, the Electrically Scanning Microwave Radiometer (ESMR) and Scanning Microwave Spectrometer (SCAMS) instruments on the Nimbus 5 and 6 satellites have provided data directly related to precipitable water, cloud liquid water and rainfall over the ocean (Staelin *et al.*, 1976; Wilheit *et al.*, 1977; Viezee *et al.*, 1978). This information should be exploited in research related to the precipitation and the hydrology of the Pacific West Coast states. Knowledge about cloud water conditions over the ocean beyond the range of weather radars may help to anticipate the extent of significant coastal precipitation.

This study emphasizes the analysis and interpretation of data related to total nonprecipitating cloud liquid water obtained from the Nimbus 6 SCAMS. The objective is to see if satellite microwave data related to storm-cloud water over the northeastern Pacific Ocean can enhance precipitation prediction for the Pacific West Coast.

2. Method of data analysis

The Nimbus 6 SCAMS was launched into a noon-midnight sun-synchronous orbit in June 1975. The sensor was designed to map the earth by scanning

0021-8952/79/091151-07$05.75
© 1979 American Meteorological Society

TABLE 1. Storm cases in the northeastern Pacific Ocean selected for analysis.

Case No.	Dates	Case No.	Dates
1	7–9 October 1975	11	23–25 December 1975
2	15–17 October 1975	12	6–8 January 1976
3	24–26 October 1975	13	12–14 January 1976
4	27–29 October 1975	14	25–27 January 1976
5	3–5 November 1975	15	11–13 February 1976
6	12–14 November 1975	16	19–21 February 1976
7	17–19 November 1975	17	23–25 February 1976
8	3–5 December 1975	18	26–28 February 1976
9	10–12 December 1975	19	17–19 March 1976
10	19–21 December 1975	20	20–22 March 1976

laterally through nadir, yielding full coverage about every 12 h. Integrated liquid water and water vapor abundances are obtained over the ocean from the 22.2 and 31.6 channel radiances. The 22.2 GHz channel is approximately twice as sensitive to water vapor as the 31.6 GHz channel, and the 31.6 channel is approximately twice as sensitive to the liquid water content of the atmosphere, so two measurements enable both water vapor and liquid water estimates to be made. No valid estimates can be made over land. Details of the statistical retrieval technique are described by Staelin *et al.* (1975).

We have analyzed SCAMS data to obtain the distribution and amounts of liquid water in 20 cyclonic storm situations.[1] These occurred in the northeastern Pacific Ocean and resulted in rainfall along the west coast of the United States during the winter season October 1975–March 1976.

Table 1 lists the 20 3-day periods for which storm cases were selected. In each case, the storm system moved eastward toward the West Coast.

Nimbus 6 SCANS digital data tapes were used in the analyses. Two time periods at intervals of 10–13 h (ascending and descending node) were available for each storm day.

Daily precipitation records were examined in conjunction with sequences of SMS-2 cloud images to determine the time period during which each selected storm system affected rainfall along the West Coast. For this time period (usually three successive days), rainfall observations at 27 coastal stations extending from Quillayute, Washington (47.95°N) to Santa Maria, California (34.90°N) were analyzed in terms of variation in latitude or distribution of coastal precipitation.

The computer program used in analyzing the Nimbus 6 SCAMS digital data and the coastal rainfall data has three major subroutines. One subroutine prints the satellite scan spot values for a given time period on a Mercator map extending from 30°–54°N latitude and 120°–168°W longitude. The second subroutine averages these data on a 1° mesh gridpoint array (25 rows, 27 columns). For each time period, the third subroutine characterizes the storm (on the basis of the SCAMS data) in terms of the distribution of liquid water as a function of latitude from north to south across the storm area, and computes the degree of linear correlation between the observed distribution of coastal precipitation and the antecedent conditions of the storm's liquid water distribution over the ocean.

To illustrate the above method of analysis, Fig. 1 shows the conditions of SCAMS-derived cloud water associated with a major storm system over the ocean identified by SMS-2 data at 0230 PST 7 October 1975, and about 24 h later, at 0100 PST 8 October 1975. These conditions preceded the coastal rainfall shown in Fig. 2, which shows the observed distribution of 24 h coastal rainfall for 8–10 October and the 72 h cumulative total for the three days. In Fig. 1, the SCAMS liquid water contents derived from SCAMS data (middle frames) are analyzed on the 1° mesh, grid-point array. Each printed number represents the vertically integrated (columnar) liquid water, averaged over the grid square, expressed in units of 10^{-2} mm (mg cm^{-2}). Uncertainties in these data have been discussed by Rosenkranz and Staelin (1972) and by Staelin *et al.* (1976). An estimated rms accuracy of 0.1 mm (10 mg cm^{-2}) in the individual data points can be considered. Maximum values range from 100–138 mg cm^{-2} in the northern part of the cloud system on 7 October, and from 100–129 mg cm^{-2} in the southern part on 8 October. The characteristic distributions of total liquid water (1° latitude strips across the grid-point area), expressed in units of 10^8 tons, are shown on the right side. The data of Figs. 1 and 2 show that the distribution of cloud water analyzed from SCAMS has a maximum ($>0.9\times10^8$ tons) at about 0100 PST 8 October between latitude 45 and 37°N; coastal rainfall has a maximum on 9 October and also in the three-day cumulative total, between these same latitudes. Thus, the distribution of the cloud-water content obtained from the Nimbus 6 SCAMS over the ocean is positively correlated with the distribution of coastal rainfall observed at a later time. Analyses such as illustrated in Figs. 1 and 2 were made for all 20 storm cases.

3. Results

For each of the 20 cases analyzed, the distribution of 72 h cumulative rainfall along the West Coast from Astoria, Oregon, to Santa Maria, California (46–34°N), was correlated with the distribution of storm-cloud water over the ocean analyzed from the SCAMS data at times ranging from 36–96 h earlier. The particular antecedent time at which the positive correlation was maximum was identified for each case. Table 2 shows

[1] Similar analyses were made using SCAMS data related to total precipitable water but results were not significant. A different application of these data to precipitation prediction is discussed by Vlasse *et al.* (1978).

ORIGINAL PAGE IS
OF POOR QUALITY

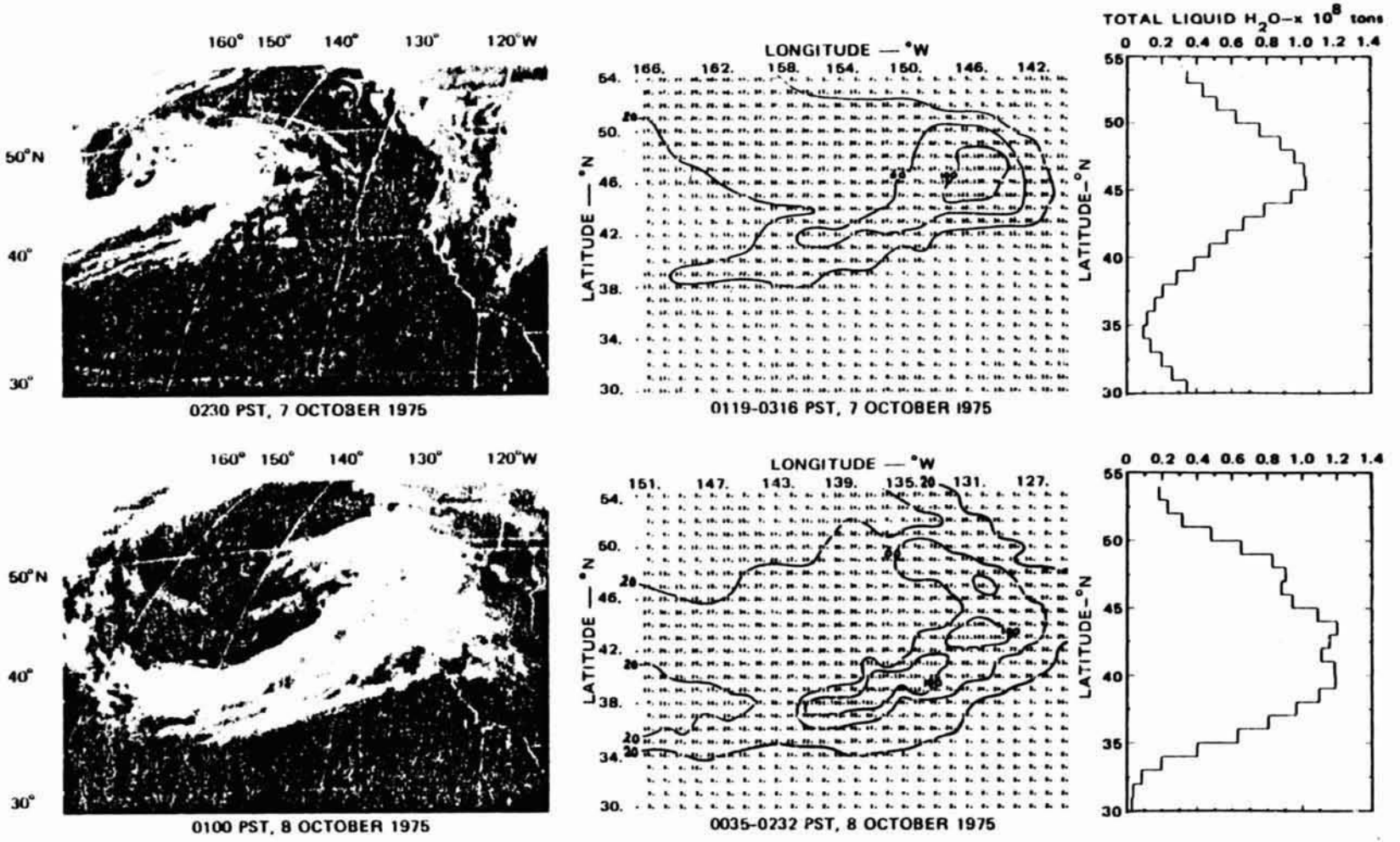


FIG. 1. SMS-2 infrared cloud image (left frame) and analysis of Nimbus 6 SCAMS cloud liquid water content (middle frame, units of mg cm^{-2}) for storm system at 0230 PST 7 October 1975, and approximately 24 h later at 0100 PST, 8 October. Distributions of cloud liquid water across the storm area (right frames) are used to obtain information on coastal rainfall distribution.

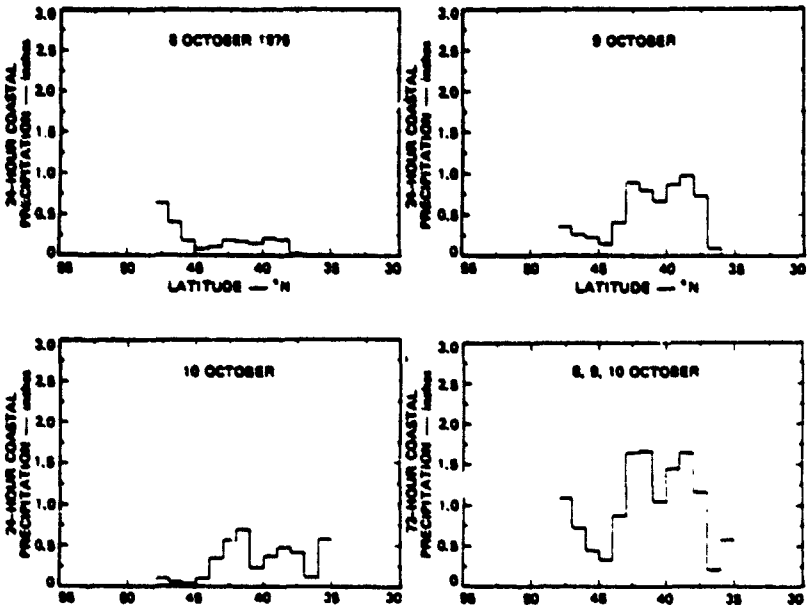


FIG. 2. Observed coastal rainfall distribution from Washington (48°N) to central California (34°N) during 3-day period in which storm system of Fig. 1 affected weather conditions.

how these maximum correlations are distributed with respect to the corresponding antecedent time period for all 20 cases analyzed and, in parentheses, for those 14 cases that had maximum correlation coefficients $\geq +0.70$. In the primary (20-case) data sample, the lowest and highest correlations were +0.39 (Case 4) and +0.99 (Case 12), respectively. It is seen that the 72 h coastal rainfall distribution is most frequently correlated with the storm-cloud liquid water distribution over the ocean analyzed from SCAMS data that are available 60–72 h earlier. The 72 h cumulative rainfall was correlated with the antecedent cloud-water distributions, because it represents the total precipitation from the storm system. Also, high positive correlations occurred most frequently when the two stations in the state of Washington (Quillayute and Hoquiam), where rainfall is significantly affected by orography, were eliminated from the data samples.

Table 2 shows that for the storm cases examined in October, November, January and March, the distribution of the 72 h cumulative coastal rainfall is highly correlated with the distribution of storm liquid-water content over the ocean at a time period 60–72 h earlier, exclusively. The December and February cases include results with the variable antecedent time periods. Storm development and movement were not accounted for in our analyses but must be considered when interpreting the correlation coefficients. For example, Fig. 3 presents the individual data sets and linear correlations (r) of the three-day cumulative coastal

TABLE 2. Frequency (number of cases) of maximum positive correlation between 72 h rainfall along the West Coast (46–34°N) and antecedent SCAMS cloud-liquid water over the eastern Pacific Ocean.

Antecedent time of maximum correlation (h)	Frequency of maximum positive correlation						
	October 1975	November 1975	December 1975	January 1976	February 1976	March 1976	Σ
96							
84			1 (1)		1 (1)		2 (2)
72	4 (2)*		1 (1)	1 (1)		2 (2)	8 (6)
60		3 (3)	1	2 (1)	1 (1)		7 (5)
48			1 (1)				1 (1)
36					2		2
Totals	4 (2)	3 (3)	4 (3)	3 (2)	4 (2)	2 (2)	20 (14)

* Numbers in parentheses refer to cases with maximum correlations $\geq +0.70$.

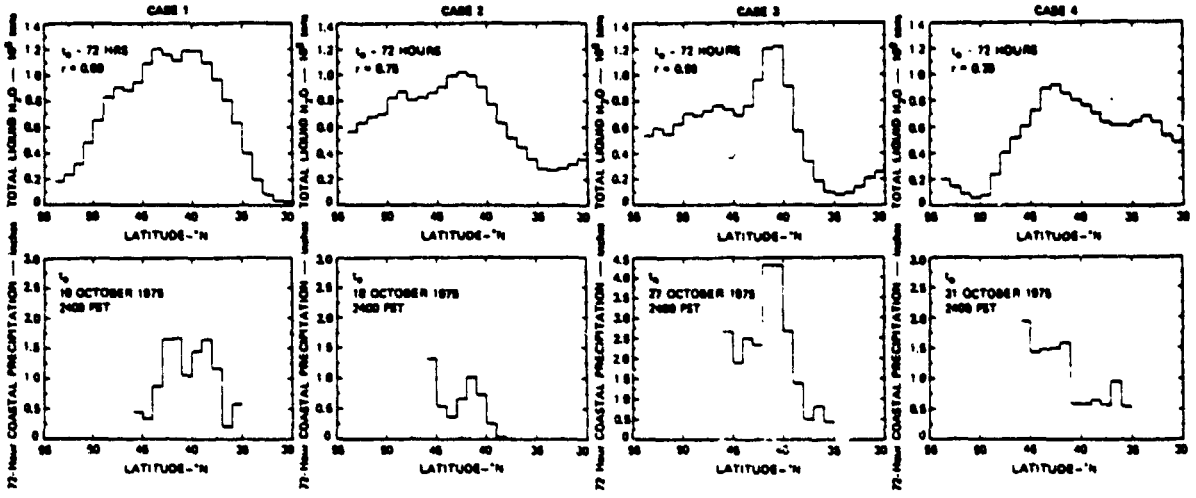


FIG. 3. Data sets and linear correlations (r) for the 3-day cumulative coastal rainfall distribution observed at t_0 (lower frame) and the antecedent distribution of SCAMS liquid water over the ocean at $t_0 - 72$ h (upper frame) for cases 1–4 (October 1975).

rainfall distribution observed at t_0 (lower frames) and the antecedent distribution of SCAMS liquid water over the ocean at $t_0 - 72$ h (upper frames) for cases 1–4 (October 1975). For cases 2 and 3, the storm's direction of movement was eastward and, consequently, the correlation coefficients are relatively high. For cases 1 and 4, however, storm movement was southward and northward, respectively. Translation of the antecedent water distributions to account for storm movement during the next 72 h would have increased the coefficients to +0.80 or larger.

For the 14 storm cases associated with correlation coefficients $\geq +0.70$, Fig. 4 shows the relationship, expressed by linear regression, between total storm water (W) over the ocean and total 72 h rainfall (P) subsequently recorded along the coast from 46°–34°N. The standard deviation of P (± 14.67 cm) about the linear fit is too large to give much practical significance to the regression equation. This indicates that the point measurements of liquid water obtained from SCAMS do not represent the rainfall for the entire storm. It is encouraging, however, that in spite of the large field of view of SCAMS (145–200 km) and the simple algorithm used to derive liquid water from SCAMS measurements, a positive correlation between remotely sensed total storm water over the ocean and recorded total rainfall along the coast is evident. Such a quantitative relationship would be very useful in operational weather prediction and hydrology and should be further explored with cloud-water data such as those available from the Scanning Multichannel Microwave Radiometer (SMMR) of Nimbus 7.

4. Discussion

Many of the storms examined were associated with a cloud system oriented from southwest to northeast. In these cases, the northern part of the cloud system crosses the coast first, after which rainfall from this part no longer contributes to our objective analysis. Our finding that the distribution of total storm rainfall is best correlated with the storm-cloud liquid water distribution over the ocean 60–72 h earlier results because, after this time period, the storm-cloud system is no longer entirely off the coast. Consequently, its total liquid water content shows decreasing positive correlation with the total storm rainfall. To relate storm liquid water over the ocean at time periods <36 h before landfall to the distribution of coastal rainfall

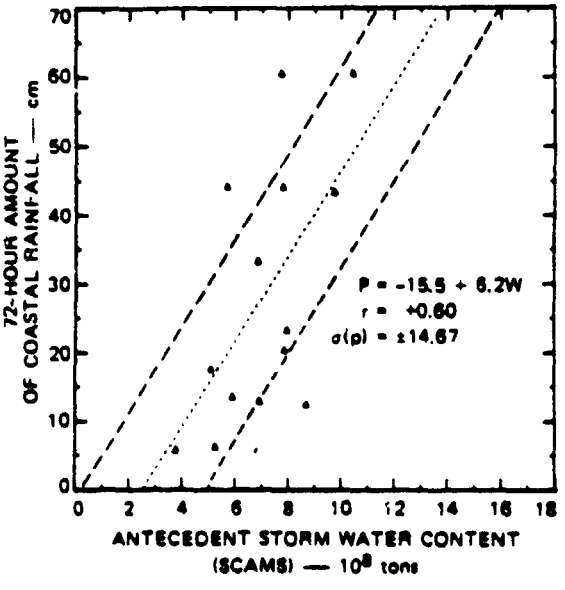


FIG. 4. Total 72 h coastal rainfall (P) from 46°–34°N versus antecedent SCAMS-derived total storm-water content (W) over the northeastern Pacific Ocean.

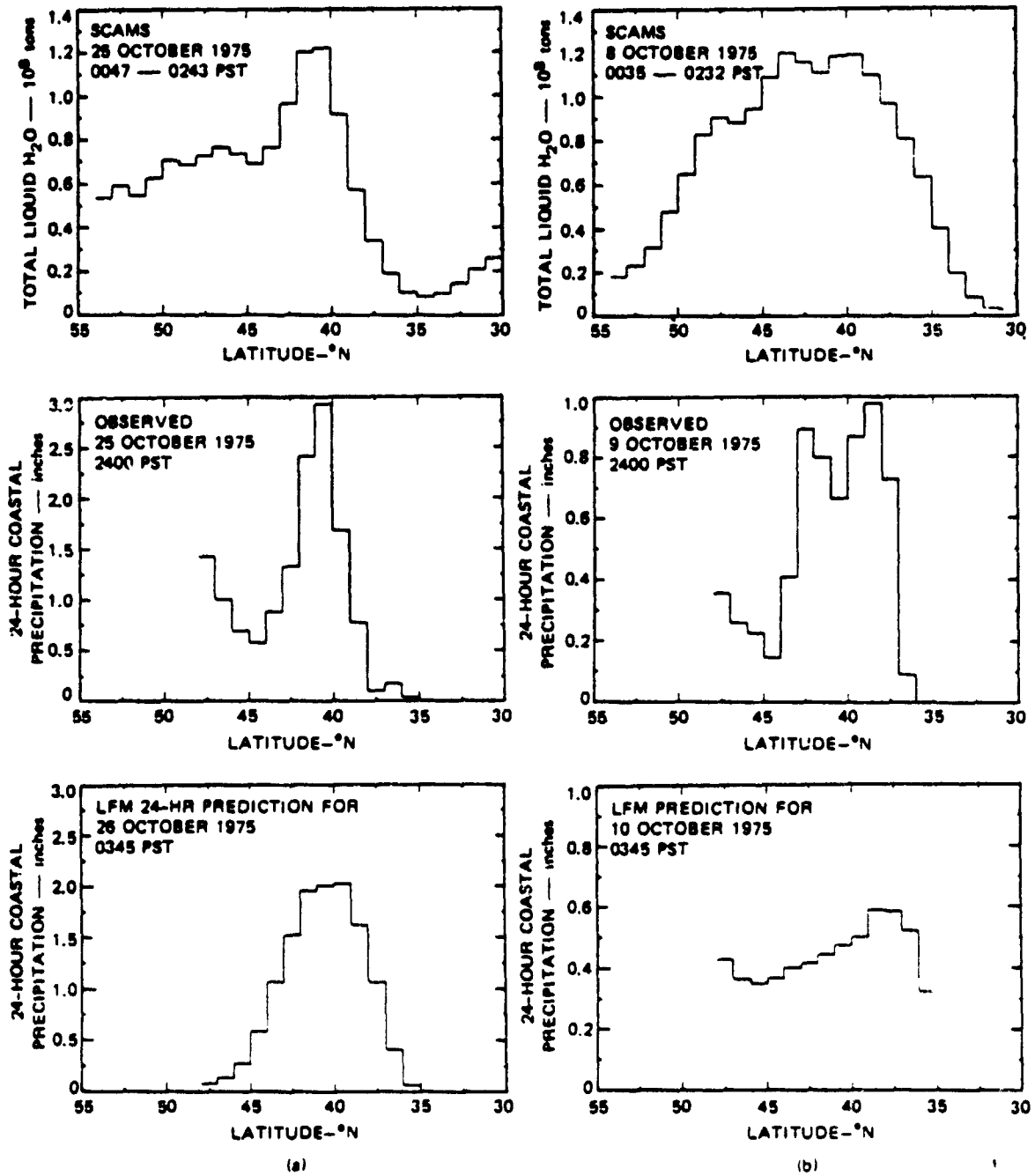


FIG. 5. Comparison between LFM prediction of coastal rainfall distribution (lower frames), observed rainfall distribution (middle frames), and antecedent SCAMS liquid water distribution over the ocean (upper frames) for two different cases. Data comparisons show how cloud liquid water data from passive microwave sensors over the ocean can be used to (a) reinforce and (b) amend numerical predictions of coastal rainfall distribution

requires additional data interpretation and analysis. The results of two case studies are presented below.

In interpreting the study results, we must remember that cloud water in the form of ice crystals does not interact significantly with microwave radiation and is not accounted for in the SCAMS data. Furthermore, the algorithm by which SCAMS radiance data are converted to liquid water is based on the assumption that the transmittance of water clouds is proportional to the total (integrated) liquid-water mass regardless of distribution of drop size. This Rayleigh approximation applies to clouds with water-droplet radii $\lesssim 50$ μm. Thus, the SCAMS liquid-water values are derived by neglecting scattering and increased absorption due to

precipitation. Grody (1976) has pointed out that the SCAMS algorithm used to derive liquid water results in an overestimate for cases where precipitation is present. Since precipitation was always present in the storm-cloud systems examined, our study emphasizes (and its results reflect) the application of *relative* variations (distributions) in storm-water content rather than of absolute values.

As noted, storm development and movement were not accounted for in our analyses but must be considered when interpreting the correlation coefficients. The analyses are straightforward when storm systems move from west to east, and our grid area (25° latitude×26° longitude) encloses only one system at a time, as was the case in October, November, January and March. For December and February, however, two frontal-wave systems moving from southwest to northeast were frequently present within the area of analysis. In such cases, the computer program should be expanded to enable the separation of liquid-water distributions of the two storms and to account for direction of movement.

5. Conclusions

Our study indicates that passive microwave sensor measurements of storm-cloud water over the ocean, when used in conjunction with numerical and other objective guidance, can enhance the accuracy of predictions of coastal rainfall distribution. This application is further examined by comparing our data with LFM forecasts of the amount of coastal precipitation, at times corresponding to those of some of our case studies. LFM forecase periods of up to 36 h are available. Forecasts for longer time periods would have been preferable for comparison with the results of our case studies, which showed high positive correlation between 72 h coastal rainfall distributions and the 60–72 h antecedent SCAMS liquid water distributions over the ocean. From the available data, the LFM 24 h precipitation predictions were selected for further analysis in the following manner:

- The case studies for which LFM data are available were reexamined for high positive correlation between the distributions of observed 24 h coastal rainfall and of SCAMS data over the ocean at antecedent time periods less than those considered previously.
- For these "short-term" periods, our computer program was modified so that corresponding data sets of observed rainfall distribution over land and antecedent total liquid water distribution over the ocean could be correlated on a smaller space and time scale. Also, the coastal distributions of LMF-predicted 24 h precipitation closest in validation time to the time of observed precipitation (2400 PST) were included in the analysis.

Two examples are illustrated in Fig. 5. Fig. 5a shows the 24 h LFM prediction of 24 h cumulative coastal rainfall distribution for 0345 PST 26 October 1975 (lower frame), the rainfall distribution observed 3.5 h earlier at 2400 PST 25 October (middle frame), and the storm-water distribution analyzed from SCAMS over the ocean ~24 h earlier at 0047–0243 PST 25 October (top frame). The LFM-predicted rainfall distribution shows general agreement with the observed distribution and with the SCAMS data analysis. In this particular case, if SCAMS data had been available, their analysis would have reinforced the accuracy of the LFM prediction. On the other hand, Fig. 5b shows a case for which the LFM prediction could have been amended on the basis of the SCAMS data analysis, since the latter is in better agreement with the observed rainfall distribution than is the LFM prediction.

Application of satellite remote measurements of cloud liquid water to rainfall prediction, as discussed above, requires the availability of near-real-time satellite data. It remains to be determined whether the potential for enhanced accuracy in predicting coastal rainfall distribution justifies the cost of making such data available.

Acknowledgments. We gratefully acknowledge the contribution of Daniel E. Wolf, Research Mathematician, SRI International, who processed the SCAMS data tapes and assisted in the development of the computer program. The samples of LFM precipitation forecasts used in the research study were supplied by the Techniques Development Laboratory of NOAA. SMS-2 visible and infrared cloud images were obtained from the regional repository of SMS imagery at the Department of Meteorology, San Jose State University.

The work reported here was performed at SRI International under Contract NAS5-24450 with NASA Goddard Space Flight Center. The computer program was written with funding support from SRI International's research and development program.

REFERENCES

Fawcett, E. G., 1977: Current capabilities in prediction at the National Weather Service's National Meteorological Center. *Bull. Amer. Meteor. Soc.*, 58, 143–149.

Grody, N. C., 1976: Remote sensing of atmospheric water content from satellites using microwave radiometry. *IEEE Trans. Antennas Propag.*, 24, 155–162.

Klein, W. H., and H. R. Glahn, 1974: Forecasting local weather by means of model output statistics. *Bull. Amer. Meteor. Soc.*, 55, 1217–1227.

Rasch, G. E., and A. E. MacDonald, 1975: Map type precipitation probabilities for the western region. NOAA Tech. Memo. WR-96, 138 pp.

Rosenkranz, P. W., and D. H. Staelin, 1972: Microwave emissivity of ocean foam and its effect on nadiral radiometric measurements. *J. Geophys. Res.*, 77, 6528–6538.

Staelin, D. H., K. F. Kunzi, R. L. Pettyjohn, R. K. L. Poon and R. W. Wilcox, 1976: Remote sensing of atmospheric water vapor and liquid water with the Nimbus 5 microwave spectrometer. *J. Appl. Meteor.*, 15, 1204–1214.

Vieese, W., P. A. Davis and D. E. Wolf, 1978: Technique to use satellite microwave spectrometer data in moisture budget studies of cyclones. *Mon. Wea. Rev.*, 106, 1627–1633.

Wilheit, T. T., A. T. C. Chang, M. S. V. Rao, E. B. Rodgers and J. S. Theon, 1977: A satellite technique for quantitatively mapping rainfall rates over the oceans. *J. Appl. Meteor.*, 16, 551–560.

NASA Conference Publication 2076 – Fourth NASA Weather and Climate Program Science Review
January 24-25, 1979, NASA/GSFC, Greenbelt, MD.

Paper No. 39

APPLICATION OF NIMBUS-6 MICROWAVE DATA TO PROBLEMS IN PRECIPITATION PREDICTION FOR THE PACIFIC WEST COAST

W. Viezee and H. Shigeishi, *SRI International, Menlo Park, California,* and A. T. C. Chang, *Laboratory for Atmospheric Science, Goddard Space Flight Center, Greenbelt, Maryland*

1. INTRODUCTION

Reliable predictions of the rapid augmentation of mountain snowpack or of runoff in the Pacific Northwest westward of the Cascades depend largely on the extent to which pertinent percursor meteorological conditions can be diagnosed over the northeastern Pacific Ocean, where little or no conventional weather data exist. Although interpretations from satellite cloud-image data are increasingly used in weather forecast preparation, it is difficult to identify in advance those storm systems that produce surges of heavy precipitation over land, especially between October and April.

At present, the regional Limited-Area Fine-Mesh Model (LFM) of the National Meteorological Center (NMC) predicts precipitation for the West Coast. However, as pointed out by Fawcett (1977), the absence of observations over the eastern Pacific and the failure to model correctly atmospheric convection and complicated terrain effects on the atmosphere make the model less accurate over the western United States than for other geographic areas. Weather forecasters, therefore, must still depend on personal insight and experience in conjunction with numerical guidance when forecasting precipitation. The output of the numerical prediction models is currently complemented by objective precipitation forecasts from the Model Output Statistics (MOS) technique and the map-type PoPs (Klein and Glahn, 1974; Rasch and MacDonald, 1975); however, precipitation in moderate and heavy amounts remains one of the most difficult weather variables to predict.

Passive microwave sensors carried on satellite platforms provide measurements inside extensive storm cloud systems that the visible and infrared radiometers cannot obtain. For example, the electrically scanning microwave radiometer (ESMR) and scanning microwave spectrometer (SCAMS) instruments on the NIMBUS 5 and 6 satellites have provided data directly related to precipitable water, cloud liquid water, and rainfall over the ocean. This information should be exploited in research related to the precipitation prediction problems and the hydrology of the Pacific West Coast States, since conditions antecedent to significant precipitation and those with which numerical prediction models must be initialized develop over the ocean.

This paper reports on the preliminary results of a research study that emphasizes the analysis and interpretation of data related to total precipitable water and nonprecipitating cloud liquid water obtained from NIMBUS–6 SCAMS.

Sixteen cyclonic storm situations in the northeastern Pacific Ocean that resulted in significant rainfall along the west coast of the United States during the winter season October 1975 through February 1976 are analyzed in terms of their distributions and amounts of total water vapor and liquid water, as obtained from SCAMS data. The water-substance analyses for each storm case are related to the distribution and amount of coastal precipitation observed during the subsequent time period when the storm system crosses the coastline. Concomitant precipitation predictions from the LFM are incorporated into the study also. The overall objective of the research is to explore techniques by which satellite microwave data over the ocean can be used to improve precipitation prediction for the Pacific West Coast states.

Paper 105

NASA Technical Memorandum 80597 December 1979

THE DIELECTRIC PROPERTIES OF SOIL–WATER MIXTURES AT MICROWAVE FREQUENCIES

J. R. Wang
Laboratory for Atmospheric Sciences

ABSTRACT

Recent measurements on the dielectric constants of soil–water mixtures have shown the existence of two frequency regions in which the dielectric behavior of these mixtures was quite different. At the frequencies of 1.4 GHz – 5 GHz there were strong evidences that the variation of the dielectric constant ϵ with water content W depended on soil type. While the real part of ϵ, ϵ', for sandy soils rose rapidly with the increase in W, ϵ' for the high-clay content soils rose only slowly with W. As a consequence, ϵ' was generally higher for the sandy soils than for the high–clay content soil at a given W. On the other hand, most of the measurements at frequencies < 1 GHz indicated the increase of ϵ with W independent of soil types. At a given W, ϵ' (sandy soil) $\cong \epsilon'$ (high-clay content soil) within the precision of the measurements.

These observational features could be satisfactorily interpreted in terms of a simple dielectric relaxation model suggested by Schwarz (1962), with an appropriate choice of the mean relaxation frequency f_m and the range of the activation energy β. It was found that smaller f_m and larger β were required for the high–clay content soils than the sandy soils in order to be consistent with the measured data. These requirements on f_m and β were strongly supported by the observed dependence of the soil surface area per unit weight on the soil types.

Paper 106

NASA Technical Memorandum 80311 June 1979

PASSIVE MICROWAVE REMOTE SENSING OF SOIL MOISTURE: THE EFFECT OF TILLED ROW STRUCTURE

J. R. Wang, R. W. Newton,[2] and J. W. Rouse[3]

June 1979

ABSTRACT

The tilled row structure is known to be one of the important factors affecting the observations of the microwave emission from a natural surface. Measurements of this effect were carried out with both L- and X-band radiometers mounted on a mobile truck on a bare 40 m x 45 m row tilled field. The soil moisture content during the measurements ranged from ~ 10% to ~ 30% by dry weight. The results of these measurements showed that the variations of the antenna temperatures with incident angle θ changed with the azimuthal angle α measured from the row direction. In particular, at $\theta = 0°$ and $\alpha \neq 45°$, the observed horizontally and vertically polarized antenna temperatures, $T_{BH}(\theta, \alpha)$ and $T_{BV}(\theta, \alpha)$, were not equal. In general, $T_{BH}(0°, \alpha) > T_{BV}(0°, \alpha)$ when $0° < \alpha < 45°$ and $T_{BH}(0°, \alpha) < T_{BV}(0°, \alpha)$ when $45° < \alpha < 90°$. The difference between $T_{BH}(0°, \alpha)$ and $T_{BV}(0°, \alpha)$ was observed to decrease with α approaching 45° and/or with soil moisture content.

A numerical calculation based on a composite surface roughness – a small scale RMS height variations superimposed on a large periodic row structure – was made and found to predict the observed features within the model's limit of accuracy. It was concluded that the difference between $T_{BV}(0°, \alpha)$ and $T_{BH}(0°, \alpha)$ was due to the change in the local angle of field emission within the antenna field of view caused by the large-scale row structure.

[1]Laboratory for Atmospheric Sciences, NASA Goddard Space Flight Center

[2]Remote Sensing Center, Texas A&M University, College Station, Texas

[3]Department of Electrical Engineering, University of Missouri, Columbia, Missouri

Paper 107

NASA Technical Memorandum 79659 October 1978

AN EMPIRICAL MODEL FOR THE COMPLEX DIELECTRIC PERMITTIVITY OF SOILS AS A FUNCTION OF WATER CONTENT

J. R. Wang and T. J. Schmugge

ABSTRACT

The recent measurements on the dielectric properties of soils have shown that the variation of dielectric constant with moisture content depends on soil types. The observed dielectric constant increases only slowly with moisture content up to a transition point. Beyond the transition it increases rapidly with moisture content. The moisture value at transition region was found to be higher for high clay content soils than for sandy soils. Many mixing formulas reported in the literature were compared with, and were found incompatible with, the measured dielectric variations of soil-water mixtures. A simple empirical model was proposed to describe the dielectric behavior of the soil-water mixtures. This model employs the mixing of either the dielectric constants or the refrac- indices of ice, water, rock and air, and treats the transition moisture value as an adjustable parameter. The calculated mixture dielectric constants from the model were found to be in reasonable agreement with the measured results over the entire moisture range of 0-0.5 cm. /cm^3. The transition moistures derived from the model range from 0.09 to 0.30 and are strongly correlated with the wilting points of the soils estimated from their textures. This relationship between transition moisture and wilting point provides a means of estimating soil dielectric properties on the basis of texture information.

Paper 108

NASA Technical Memorandum 80277 May 1979

AN ALGORITHM FOR RETRIEVAL OF OCEAN SURFACE AND ATMOSPHERIC PARAMETERS FROM THE OBSERVATIONS OF THE SCANNING MULTICHANNEL MICROWAVE RADIOMETER (SMMR)

T. T. Wilheit and A. T. C. Chang

NASA/Goddard Space Flight Center

ABSTRACT

The Scanning Multichannel Microwave Radiometer (SMMR) is a 5-frequency (6.6, 10.7, 18, 21, and 37 GHz) dual polarized microwave radiometer which was launched in two separate satellites, Nimbus-7 and Seasat in 1978. A formalism is developed which can be used to interpret the data in terms of sea surface temperature, sea surface wind speed, and the atmospheric overburden of water vapor and liquid water. It is shown with reasonable instrumental performance assumptions, these parameters can be derived to useful accuracies. Although the algorithms are not derived for use in rain; it is shown that, at least, token rain rates can be tolerated without invalidating the retrieved geophysical parameters.

NASA Conference Publication 2076 – Fourth NASA Weather and Climate Program Science Review, January 24-25, 1979, NASA/GSFC, Greenbelt, MD.

Paper No. 56

ANTARCTIC SEA ICE VARIATIONS 1973-75

H. J. Zwally, C. Parkinson, F. Carsey, P. Gloersen, *Laboratory for Atmospheric Sciences (GLAS), Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, MD 20771*
W. J. Campbell, *U.S. Geological Survey, Tacoma, WA 98416*
R. O. Ramseier, *Department of Environment, Ottawa, Canada K1A OH3*

6

ABSTRACT

Variations in the extent and concentration of sea ice cover on the Southern Ocean are described for the three-year period 1973-75 using information derived from the Nimbus-5 passive microwave imager (ESMR).

VI. CLIMATE AND RADIATIONS BRANCH

PRECEDING PAGE BLANK NOT FILMED

Paper 110

Reprinted from *Bulletin of the American Meteorological Society*, 60, pp. 1414-1426, 1979.

A Hydrological Analysis of East Australian Floods Using Nimbus-5 Electrically Scanning Radiometer Data

Lewis J. Allison and Thomas J. Schmugge

Goddard Space Flight Center

Greenbelt, Md. 20771

and

Gavin Byrne

Commonwealth Scientific and Industrial Research Organization (CSIRO) Canberra, A.C.T., Australia

Abstract

The chronological development and diminution of six floods in eastern Australia during January, February, and March 1974 were mapped for the first time by the Nimbus Electrically Scanning Microwave Radiometer (ESMR). Day and nighttime ESMR (19.35 GHz) coverage was analyzed for the low gradient, flooded Darling River system in New South Wales. Apparent movement of surface water as indicated by low brightness temperatures (<250 K, day and <240 K, night) was easily followed around the curved runoff basin along the northern shoreline of the flooded Darling River during this 3-month period. This pattern was in good agreement with flood crest data at selected river height gage stations, even under cloudy conditions.

1. Introduction

The Nimbus-5, a research and development satellite launched on 12 December 1972, carried the Electrically Scanning Microwave Radiometer (ESMR), which recorded microwave brightness temperatures (T_B) from which it has been possible to estimate meteorological and hydrological parameters on a global scale.

The microwave brightness temperature (T_B) of the earth's surface is essentially the product of the surface temperature and surface emissivity, with additional attenuation and emissions due to atmospheric water vapor and liquid water droplets (Allison, 1975; Zwally and Gloersen, 1977). Ice crystal clouds (cirrus) have little effect on microwave radiation and are essentially transparent at the 1.55 cm ESMR wavelength. Molecular oxygen absorption can also be neglected at this wavelength (Wilheit *et al.*, 1977). Emissivity is the ratio of the radiation from a surface at a measurable temperature for a given wavelength interval to that from a black-body source at the identical temperature. When the emissivity is 1, the surface radiates as a black body. When it is less than 1, the surface behaves as a gray body. The emissivity of the surface depends on, among other things, the medium's dielectric properties, and for a smooth surface can range from 0.4 for water, which has a large dielectric constant, to about 0.7 for wet soils, and to >0.9 for dry soils and dense vegetation. Microwave emission from surfaces is also affected by roughness, topographic slope, stratigraphy (layering of different materials), density of rocks, and vegetation cover (Moore *et al.*, 1975; Ulaby, 1977; Allison, 1977). For example, a smooth bare dry soil surface whose physical temperature was 290 K and emissivity was 0.90–0.95 would indicate a T_B of 261–275 K. This is to be contrasted to the low T_B over the calm ocean, which is approximately 120 K, a result of low water emissivities of ~0.4 at 19.35 GHz. The state of the sea surface (wave height, roughness, and foam) would increase the observed brightness temperature for winds above 7 m/s at the rate of 1 K/(m s^{-1}) (Nordberg *et al.*, 1971).

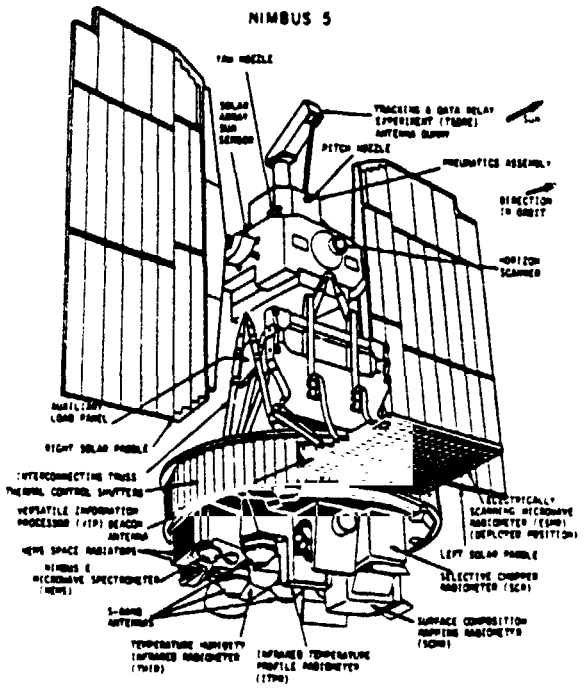


FIG. 1. The Nimbus-5 spacecraft and associated experiments.

Nimbus-5 ESMR data have been used to detect rainfall over the ocean by Allison *et al.* (1974). Rao

0003-0007/79/121414-14507.50

© 1979 American Meteorological Society

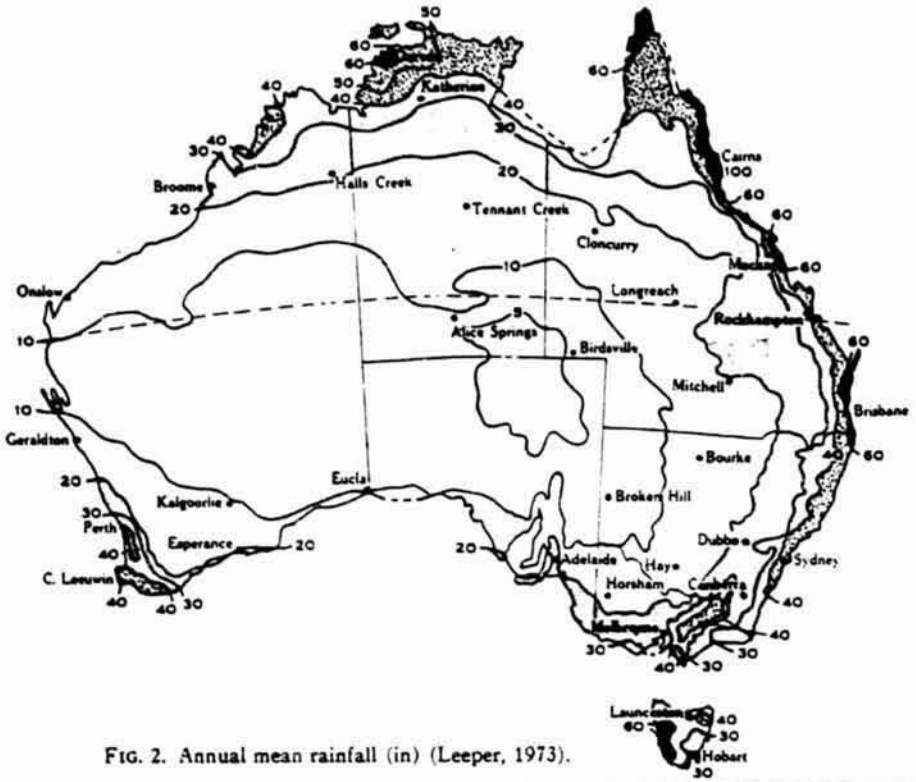


FIG. 2. Annual mean rainfall (in) (Leeper, 1973).

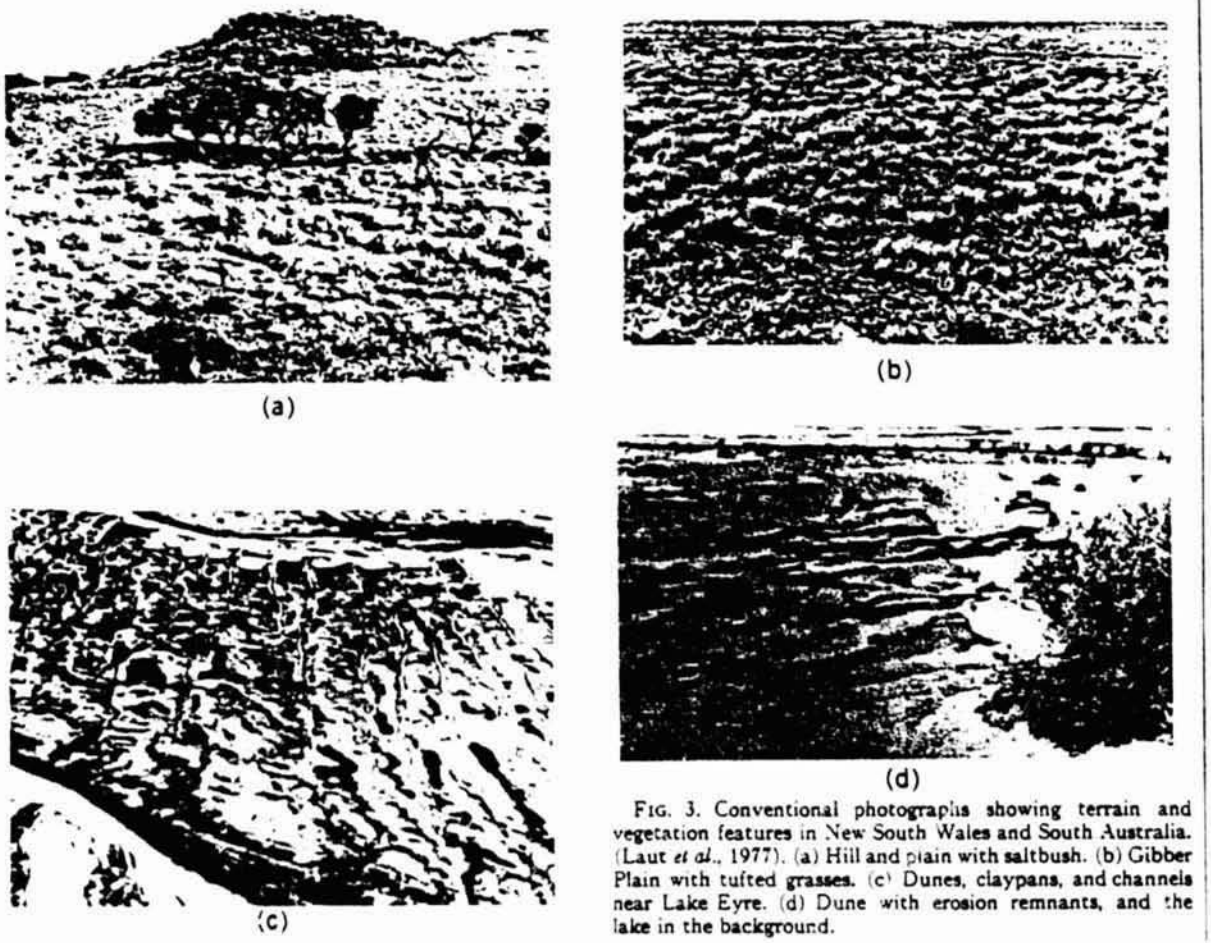


FIG. 3. Conventional photographs showing terrain and vegetation features in New South Wales and South Australia. (Laut *et al.*, 1977). (a) Hill and plain with saltbush. (b) Gibber Plain with tufted grasses. (c) Dunes, claypans, and channels near Lake Eyre. (d) Dune with erosion remnants, and the lake in the background.

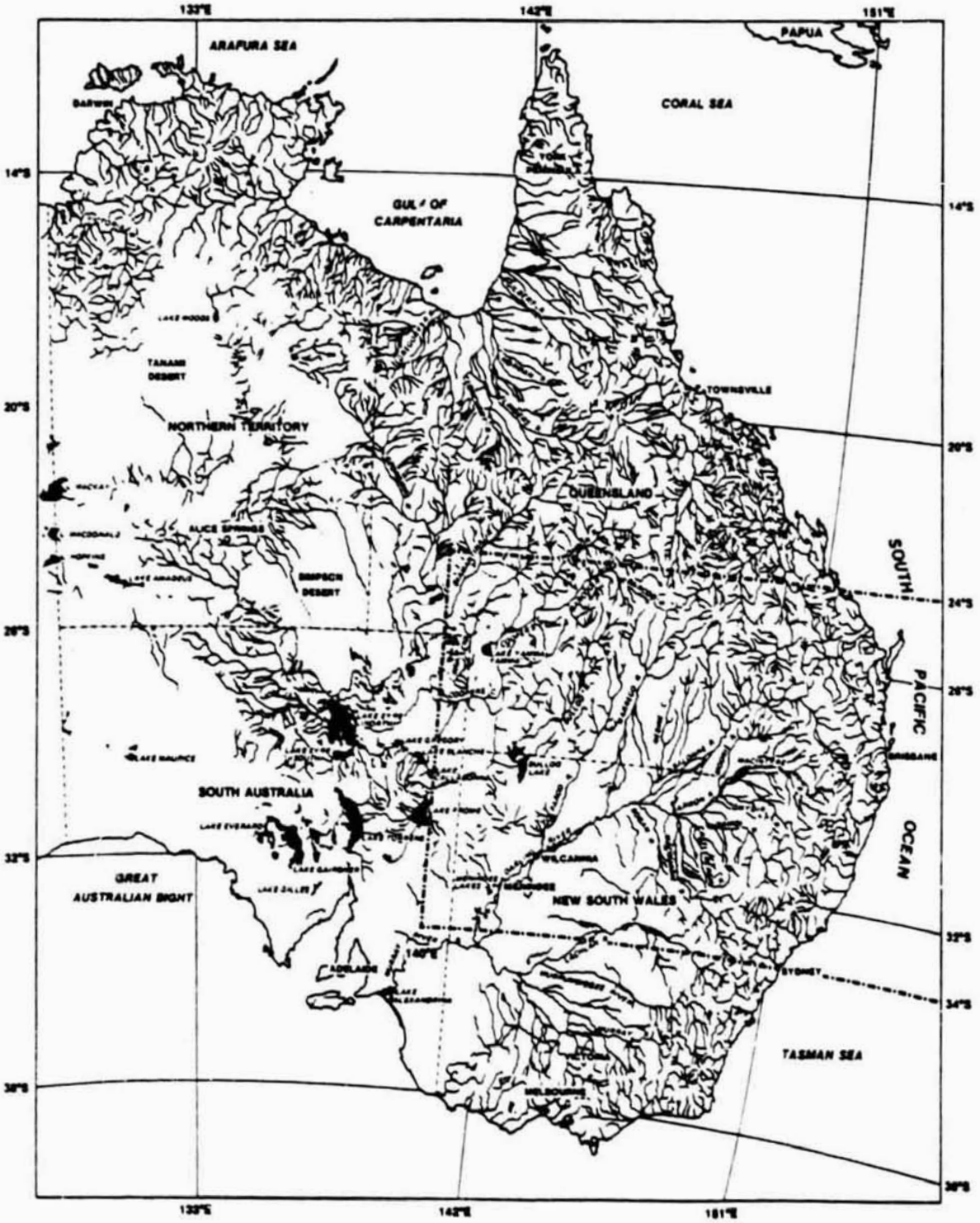


FIG. 4. Major river systems with minor tributaries and lake locations, eastern half of Australia. Dot-dash box outlines the study area analyzed in this paper. (Department of National Development, 1956.)

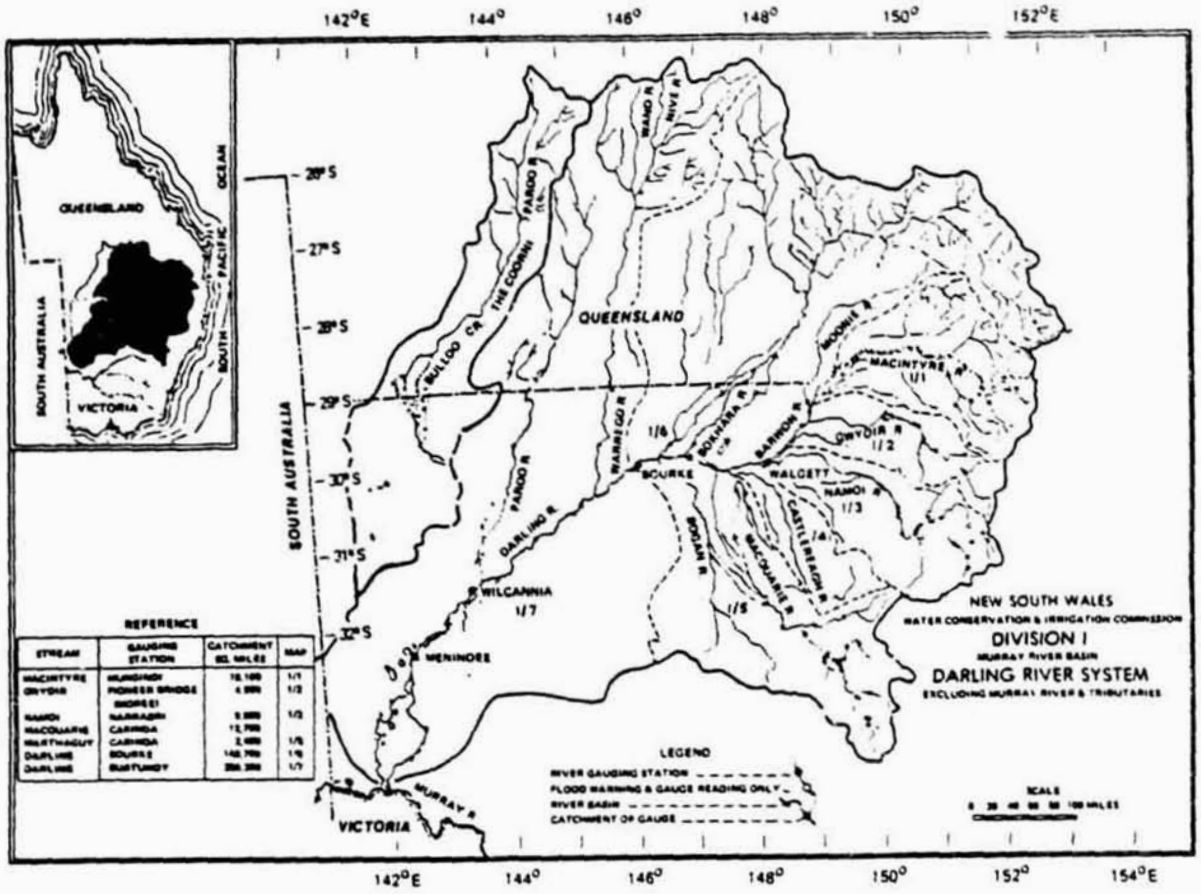


FIG. 5. A detailed chart of the Darling River system, New South Wales, Australia.

et al. (1976), Kidder and Von der Haar (1977), Wilheit *et al.* (1977), Adler and Rodgers (1977), Austin and Geotis (1978), and Stout and Martin (1979). This is possible because the oceans provide a relatively constant, low T_B (≤ 150 K) background for observing the atmosphere. Over land the background T_B is higher, typically >250 K, and more variable; as a result attempts to measure rainfall over land with the Nimbus-5 ESMR have failed (Meneely, 1975). However, there has been some limited success in delineating rain over land using theoretical analysis and actual case studies of the dual polarized Nimbus-6 ESMR (37 GHz) (Rodgers *et al.*, 1978, Savage and Weinman, 1975). The effect of rainfall over the ocean is to increase the T_B observed from space up to 250 or 260 K for high rain rates. The rate of increase in T_B depends on the wavelength. At the 1.55 cm Nimbus-5 ESMR wavelength, this maximum is reached for a rain rate of approximately 10 mm/h (Wilheit *et al.*, 1977). This relationship has been used to determine rainfall rates over the oceans, thus providing information for regions of the earth for which there were very few conventional data available. These data have been combined to produce an atlas of the monthly averages of rainfall over the oceans (Rao and Theon, 1977; Rao *et al.*, 1976).

The extent of a disastrous regional flood in the Mississippi Valley at varying recurrence intervals was mapped by use of Landsat (Rango and Anderson, 1974; Rango and Salomonson, 1974; Otterman *et al.*, 1976), NOAA VHRR (Wiesnet *et al.*, 1974), and Nimbus-5 ESMR data (Schmugge *et al.*, 1974). However, a complete time sequence of a major flood had not yet been demonstrated by satellite analysis. This paper will present Nimbus-5 microwave data that were used to detect and map the movement of flooded areas in eastern Australia from January to March 1974.

2. Electrically-scanning microwave radiometer experiment

The Nimbus-5 spacecraft (Fig. 1) was flown in a near-polar (81° retrograde) sunsynchronous circular orbit of about 1100 km. Each orbit crossed the equator northward at approximate local noon and southward at midnight with a 27° longitude separation and a

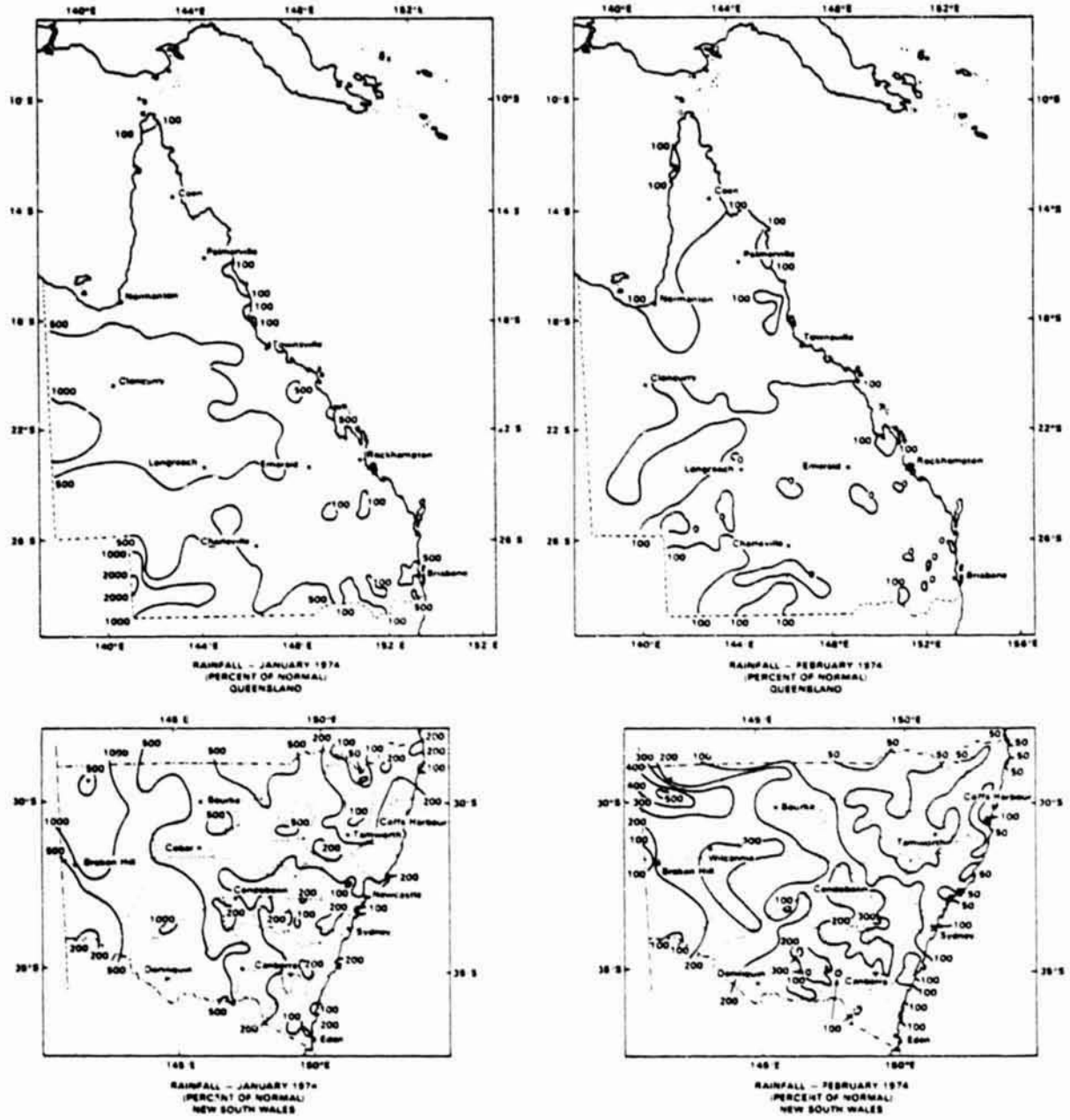
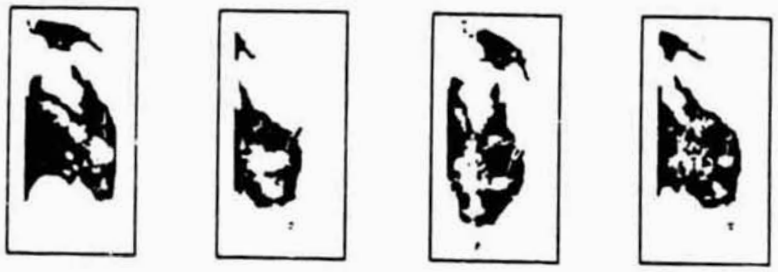


FIG. 6. Rainfall (in percent normal) for January and February 1974, Queensland and New South Wales.

period of 107 min. The Nimbus-5 ESMR measures the emitted radiation from the earth's surface and intervening atmosphere in the 250 MHz band centered at 19.35 GHz (1.55 cm), with a noise equivalent temperature difference ($NE\Delta T$) of approximately 2 K. It scans transverse to the flight path every 4 s from 50° to the left through nadir to 50° to the right in 78 steps with some overlap. The scanning process is controlled by an onboard computer and the radiometer scans a region whose width is approximately 30° of longitude at 30°N. The half-power beamwidth of the antenna is 1.4° at nadir and the 2.2° at 50° edges with resultant footprints of 22 km and 45 by 160 km, respectively. Since the ESMR senses horizontally polarized radiation, a scan-angle dependent correction was applied to the data to make each observation more closely equivalent to a linearly polarized nadir observation (Wilheit, 1973). The instrument and its basic physics are discussed in more detail by Wilheit (1972), Wilheit *et al.* (1973), and Allison *et al.* (1974).

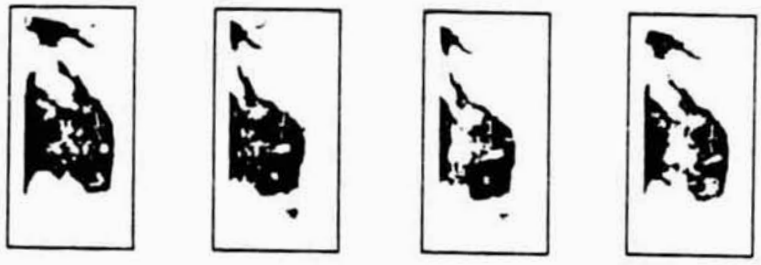


9 JAN. 1974 (NIGHT) RAIN (a)

10 JAN. 1974 (NIGHT) RAIN (b)

16 JAN. 1974 (DAY) RAIN (c)

19 JAN. 1974 (NIGHT) (d)

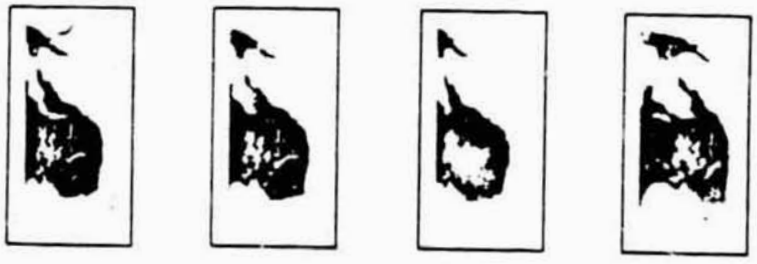


21 JAN. 1974 (NIGHT) (e)

24 JAN. 1974 (NIGHT) (f)

31 JAN. 1974 (NIGHT) (g)

2 FEB. 1974 (NIGHT) (h)

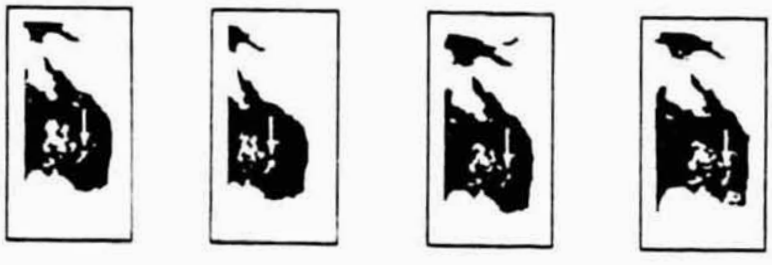


7 FEB. 1974 (NIGHT) (i)

14 FEB. 1974 (NIGHT) (j)

19 FEB. 1974 (NIGHT) RAIN (k)

23 FEB. 1974 (NIGHT) (l)



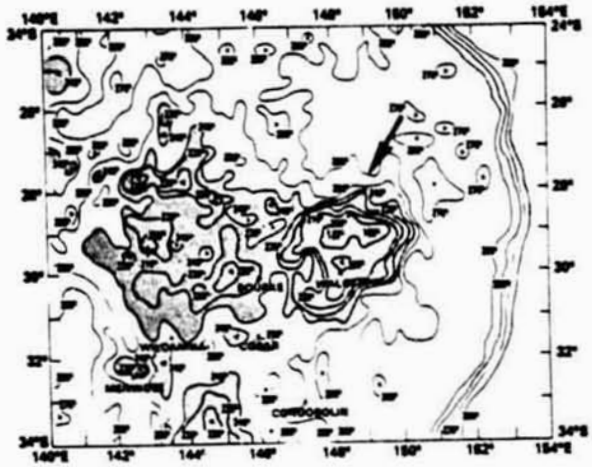
28 FEB. 1974 (NIGHT) (m)

5 MARCH 1974 (NIGHT) (n)

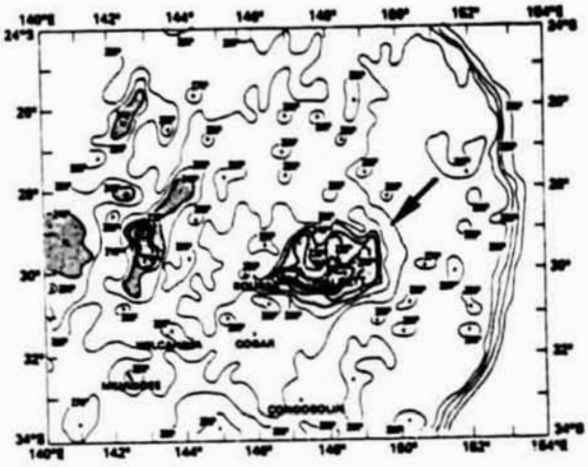
14 MARCH 1974 (NIGHT) (o)

21 MARCH 1974 (NIGHT) (p)

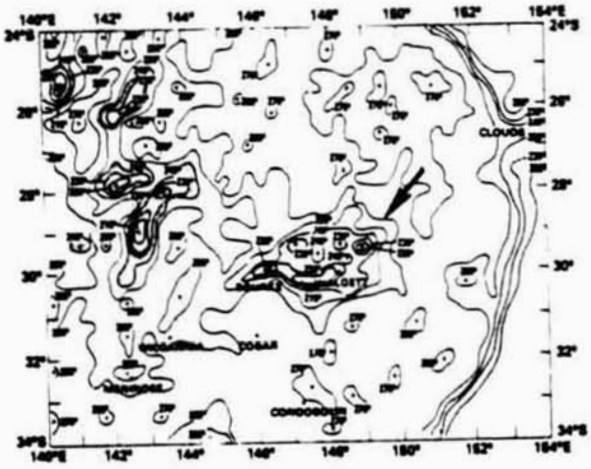
FIG. 7. Nimbus-5 ESMR (19.35 GHz) photofacsimile pictures of the East Australian floods from 9 January 1974 to 21 March 1974. The arrows indicate the wet ground in the Darling River system during this period.



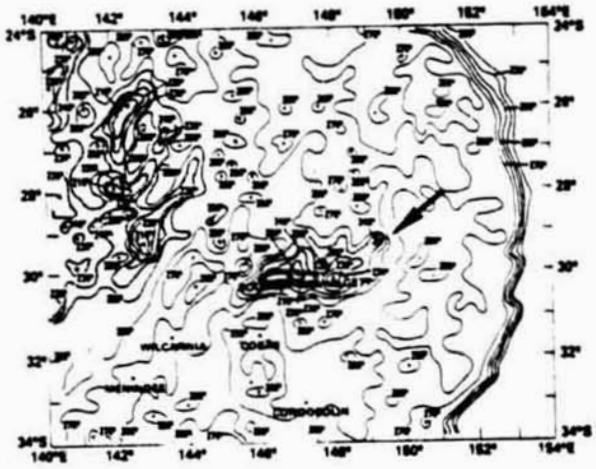
(a) 10 JANUARY 1974 ORBIT 5306 (NIGHT-RAIN)



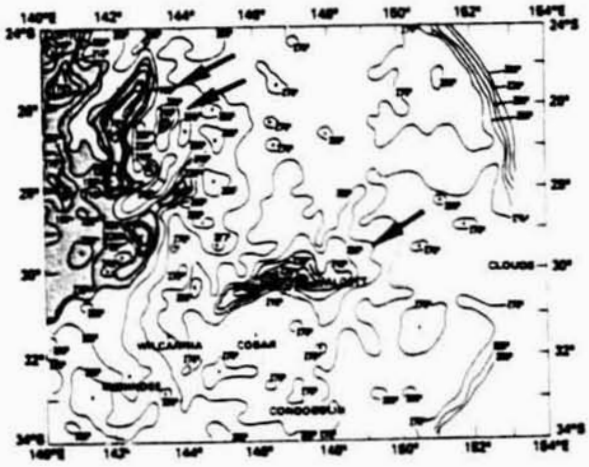
(b) 22 JANUARY 1974 ORBIT 5467 (NIGHT)



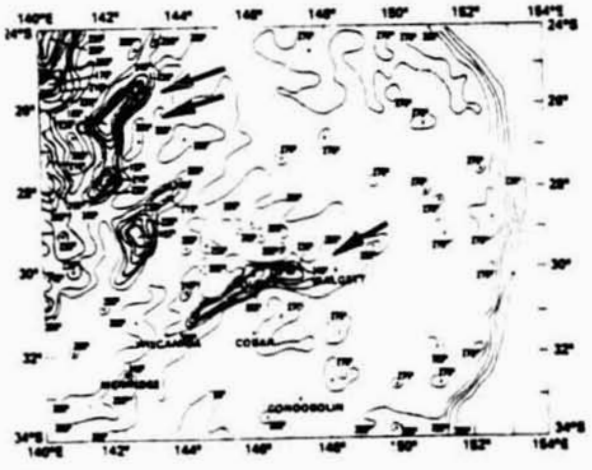
(c) 24 JANUARY 1974 ORBIT 5494 (NIGHT)



(d) 30 JANUARY 1974 ORBIT 5568 (DAY)



(e) 2 FEBRUARY 1974 ORBIT 5615 (NIGHT)



(f) 7 FEBRUARY 1974 ORBIT 5682 (NIGHT)

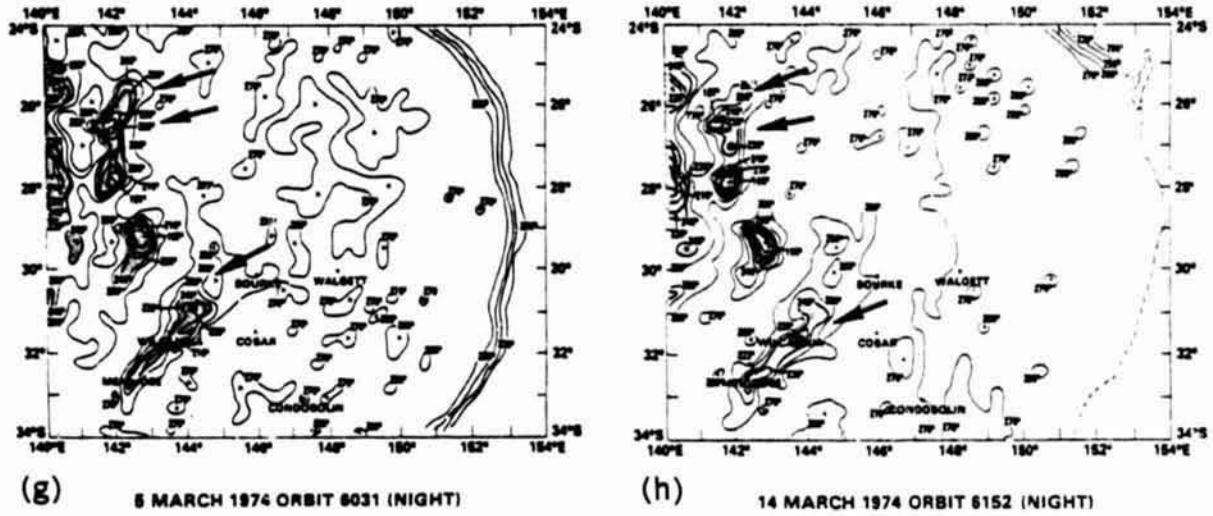


FIG. 8. Nimbus-5 ESMR (19.35G Hz) grid-print maps of brightness temperatures (T_B, K). The single arrow indicates the flooded Darling River system. The double arrow indicates the Lake Yamma Yamma and flooded Coopers Creek area, 10 January–14 March 1974.

3. The 1974 East Australian floods

a. Background information

Figure 2 shows the annual mean rainfall in inches for Australia. The wetter areas are indicated by grey tones (>20 in) and are located along the coastal strips to the north, northeast, east, southeast, and southwest. During the summer months (December–February) the rainy intertropical zone of convergence (ITC) is indicated by a low pressure monsoonal area that forms over the north equatorial coast. Extensive extremely dry desert regions (<10 in (250 mm)/yr) are shown covering large areas of central and western Australia. Evaporation is high and on the average is equivalent to 87% of the rainfall as compared to 60% for Europe and North America. Thus only a small proportion of rainfall becomes surface runoff except during heavy rain periods. During the winter months, June–August, the north coast is dry as the ITC moves northward into the Northern Hemisphere while the southern coastal strips receive rain from the west-to-east traveling lows in the "Roaring Forties" (Leeper, 1973). An average of 10–20 in of rain per year falls in semiarid Queensland and New South Wales (with the exception of coastal strips) and permits sheep and cattle grazing in vast treeless grasslands and low open hilly woodland, examples of which are shown in Fig. 3a and b (Laut *et al.*, 1977). A large number of deep artesian wells in the Great Artesian Basin (the eastern quarter of Australia in the Central Basin) supplement the surface water supply for the sheep and cattle industries. Orographic maps of Australia indicate that elevations ranging from 300 to 1000 m exist in the eastern highlands, while the Central Basin is generally below 150 m. This basin has two major drainage systems: the Lake Eyre interior drainage and the Murray-Darling River. These two gently sloped systems have <1 in (2.5 cm) runoff annually. The typical sand dunes, erosion channels, and claypans of the Lake Eyre region in south central Australia are shown in Fig. 3c and d.

FIG. 9. Landsat 1 (MSS-7, frame no. 1563-23530) recorded on 6 February 1974. Lake Yamma Yamma (appendage lower left) is shown near darker flooded Coopers Creek lowland (center right).

The lakes in the central internal drainage basin in south Australia are dry most of the time and only fill during periods of heavy rainfall. In 1974 it was possible to monitor their flooding with the Nimbus-5 ESMR. The particular lakes that will be discussed in

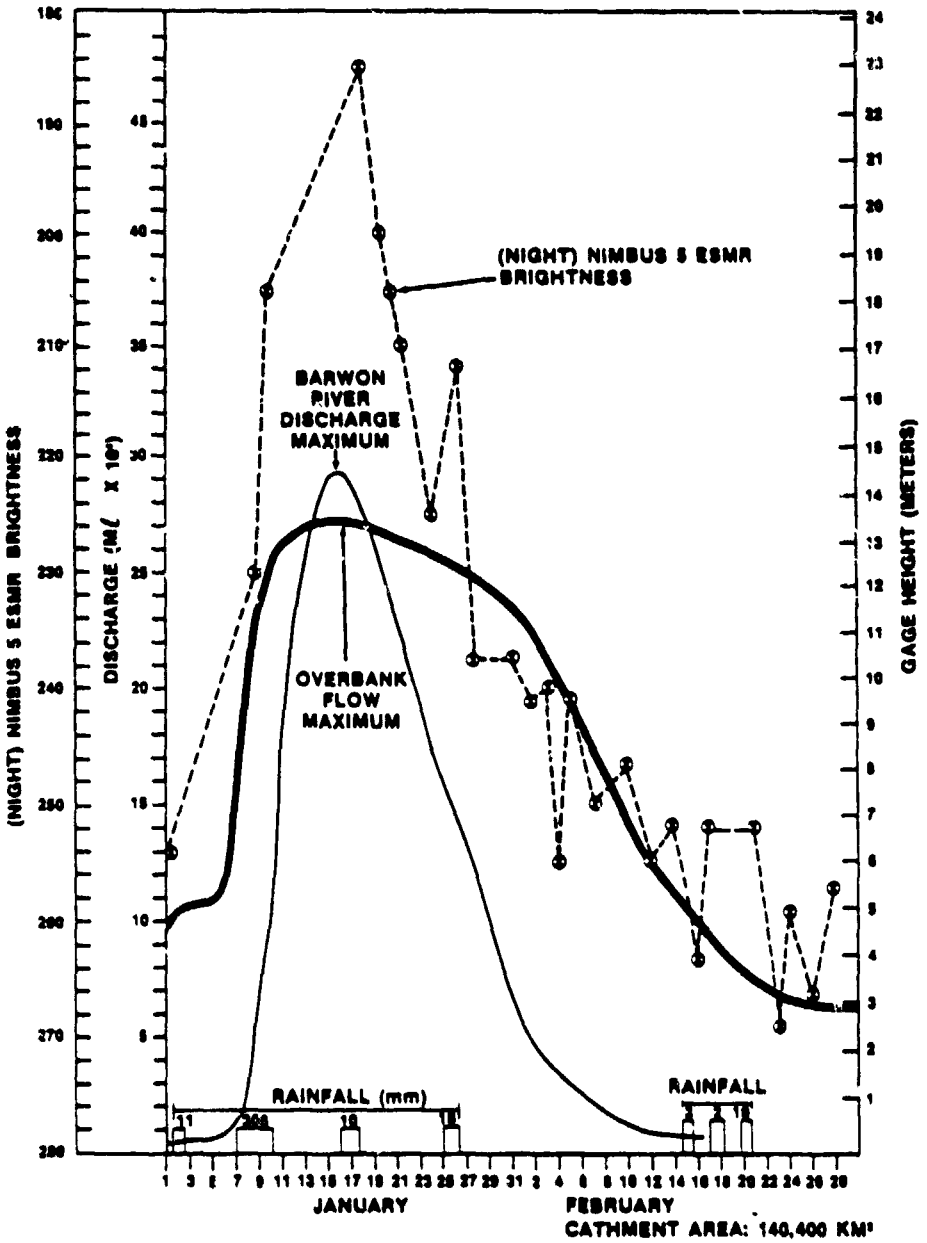


FIG. 10. Barwon River at Walgett, New South Wales. 1974 flood hydrograph and superimposed Nimbus-5 night ESMR brightness temperatures (T_B, K) and daily rainfall (mm).

some detail are: Bulloo Lake, which is the terminus of the Bulloo River in southwestern Queensland, and Lake Yamma Yamma with the neighboring lowlands of Cooper's Creek (Fig. 4).

The area of flooding that will be discussed is along the Darling River in New South Wales. Early in the period the flooding was concentrated around the town of Walgett in the northeast section of New South Wales (Fig. 5). The flooded area will appear to move downriver, ending up at Menindee in March 1974.

The January–March 1974 period was an anomalous season in that heavy rainfall and severe flooding occurred in the northern coastal plains of Queensland (south of the Gulf of Carpentaria, the Coopers Creek area of southwest Queensland, the Lakes Eyre, Frome, Gairdner, and Torrens area in South Australia, and the Darling River basin in New South Wales and south Queensland. Figure 6 shows the rainfall in both Queensland and New South Wales for January and February 1974, which ranged from 50 to 2000% of the normal 25–400 mm per month (Bureau of Meteorology, 1974). The runoff due to this heavy rainfall flowed slowly southward through normally dry channels into a network of rivers that feed into Lake Eyre and the Darling River system (Warner, 1977; Vickery, 1974). A similar flood occurred in Australia during the

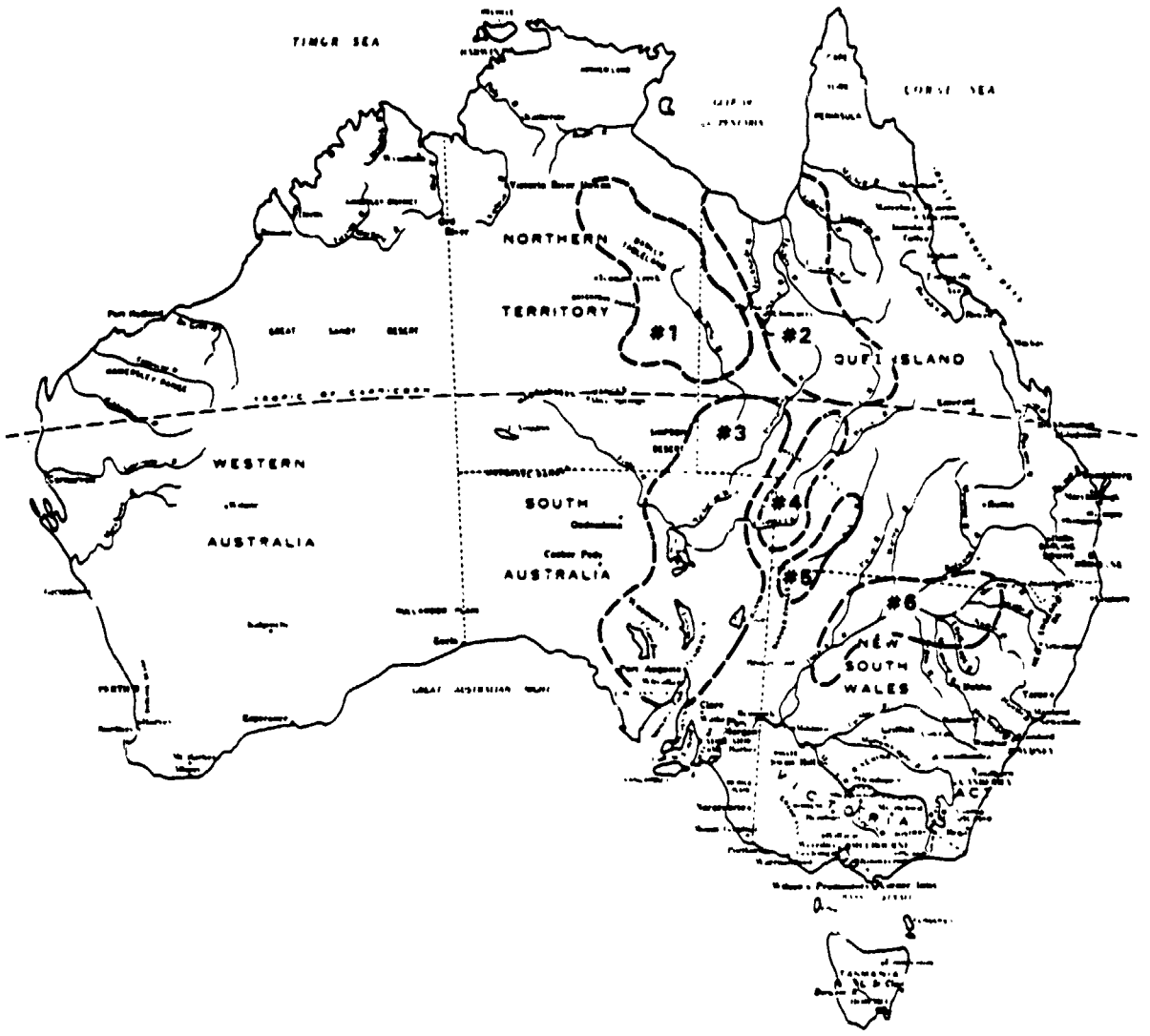


FIG. 11 Six regions in Eastern Australia that were covered by Nimbus-5 ESMR brightness temperatures ≤ 240 K (night) and ≤ 250 K (day).

1949–50 period in which large areas of the Lake Eyre drainage division were inundated and rivers flooded up to 30 km in width in certain areas. Flooding in the lower reaches of this inland water system can occur 3 months after the rainfall event in the headwaters of these drainage divisions. (Australian Water Resources Council, 1976; Gibbs and Maher, 1967; Pittock, 1975; Foley, 1954.)

In the last 100 years, the worst previous flood years in Australia occurred in 1852–3, 1863, 1870, 1875, 1887, 1889–94, 1909–11, 1916–17, 1921, 1925, 1931, 1934, and 1949–50 (Foley, 1954).

b. Chronology of the 1974 flood

At the beginning of January 1974 a normal summer monsoonal trough extended from the Gulf of Carpentaria southeastward through Queensland and northern New South Wales. By 9–10 January 1974, this surface trough became well defined and ensuing heavy rains (>300 mm) fell in Queensland and New South Wales. In addition Tropical Storm Wanda moved westward inland over southeast Queensland and increased the anomalous rainfall (>200 mm) from 24 to 26 January 1974. (Bureau of Meteorology, 1974.) The Nimbus-5 ESMR photofacsimile pictures (Fig. 7a, b), starting 9 January 1974, show the NW-SE-oriented trough with white indicating "cold" brightness areas of rain and or wet ground over a dark "warm" ground background. Ocean areas are white since they consist of lower T_B values due to low surface emissivities (0.4). Heavy rains (300 mm from 7 to 10 January at Walgett, fell on the upland portions of the Darling River watershed producing a large, almost circular, region of lowered T_B values, seen in Fig. 7c–e. As the region dried out, the total area of lowered T_B decreased and began to move down to the south and take the

shape of the lower portion of the drainage basin (Fig. 7f–j). Additional rain (35–62 mm) on 19–20 February increased the size of the low T_B region (Fig. 7k). The region continued to move southward and decrease in size until 21 March when the sequence stops (Fig. 7l–p). To the west of the Darling River basin, the left side of Fig. 7, are the interior lakes of South Australia. These lakes show a large change in size from 9 January to 23 February as the runoff from the uplands drains into these interior basins.

A more quantitative presentation of these features is given in Fig. 8, which contains contour maps of brightness temperature values for the Darling River basin. Regions with $T_B < 240$ K for night orbits and $T_B < 250$ K for day orbits are shaded or ziptoned to indicate regions with substantial flooding. This is based on the assumption that areas with persistent low brightness values are those having a substantial amount of standing water within the sensor field of view. Results presented over Illinois in 1973 (Schmugge *et al.*, 1977) and in Australia in 1976 (Barton, 1978) indicate that wet soils alone would not produce the persistent low T_B values. Barton (1978), in an analysis of ESMR data over Australia, found that there was no response to wet ground if there was vegetative cover, e.g., in his case at Deniliquin with native grasses. In the other analysis, Schmugge *et al.* (1977) found that if wet bare soil was exposed, the surface layer would dry rapidly, producing a high T_B value. It was concluded from these studies that bare soil regions with persistent low T_B values were covered with a large amount of standing water in a warm season period. The particular values of T_B chosen to delineate areas with substantial flooding were those that would indicate that 25–30% of the area within the sensor's field of view was water covered. This percentage was determined by observing that the T_B of unflooded grassland was 280 K, while the lowest T_B observed for the interior lakes was in the 170–180 K range. Thus for a daytime orbit,

$$\text{Percent of standing water} = \frac{T_B\,(\text{dry}) - T_B}{T_B\,(\text{dry}) - T_B\,(\text{wet})} = \frac{280\text{ K} - T_B}{280\text{ K} - 170\text{ K}}\,100$$

It is recognized that this is only an estimate and that regions with temporarily saturated soils could also have T_B in the 250 K range.

Figure 8 shows the T_B contour map for the time of the most intense rainfall over the area with extensive ground wetness. The low T_B values are the result of increased ground wetness rather than being due to the rain itself, since the lowest T_B expected for intense rain is approximately 250 K (Wilheit *et al.*, 1977). By 22 January (Fig. 8b) the shaded area has decreased to a 150 by 200 km section centered around the town of Walgett. The other shaded area north of the town of Wilcannia is Lake Bulloo in an internal drainage basin. Above that is Lake Yamma Yamma.

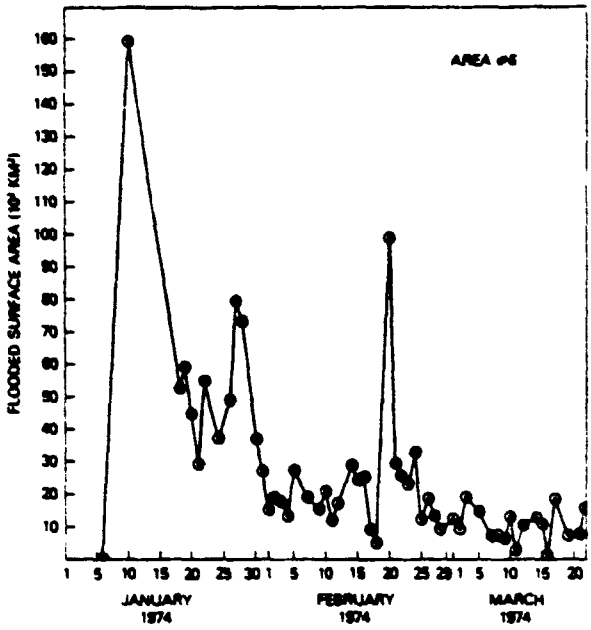


FIG. 12. Estimated total flooded surface area (km²) for area 6 (shown in Fig. 11), in Eastern Australia from January to March 1974 derived from Nimbus-5 ESMR (19.35 GHz) data (day: $T_B \leq 250$ K, night: $T_B \leq 240$ K).

and the Cooper Creek lowland is apparently just beginning to fill with water. The shaded area around Walgett continues to shrink on 24 January while the Lake Bulloo and Lake Yamma Yamma–Cooper Creek areas continue to grow and intensify. Additional rains on 25 and 26 January (18 mm at Walgett), produced little change in the 30 January map of the Darling River basin except for the movement of the shaded area to the southwest. There was a substantial increase in the shaded area for the lake regions where heavier rains had fallen. The 2 and 7 February 1974 maps show the continued shrinkage of the shaded area in the Darling River basin and the movement of the flooded area to the southwest. The leading edge moved from 70 km to approximately 30 km northwest of the town of Wilcannia during this interval, or about 8 km per day. In each case, the lowest T_B was 210 K or less, indicating at least 50–60% water within the field of view of the radiometer. The two lakes in the northwestern portion of this analysis have become well defined with T_B as low as 170 K, indicated by the double arrows in Figs. 8e to h. Figure 9 is a Landsat 1 (MSS-7) image of the same area on 6 February 1974 when waters in the Barcoo River channels overflowed into Lake Yamma Yamma and Coopers Creek lowlands, reaching a maximum width of 60 km. The flood boundaries are readily discernable because of large differences in surface reflectance between the dark flooded area and the lighter dry uplands (Robinove, 1978; Short *et al.*, 1976). From this image, the total extent of the flooded area is estimated to be 160 km by 50 km, which is in reasonable agreement with the area within the 180 K contour line in Fig. 8f

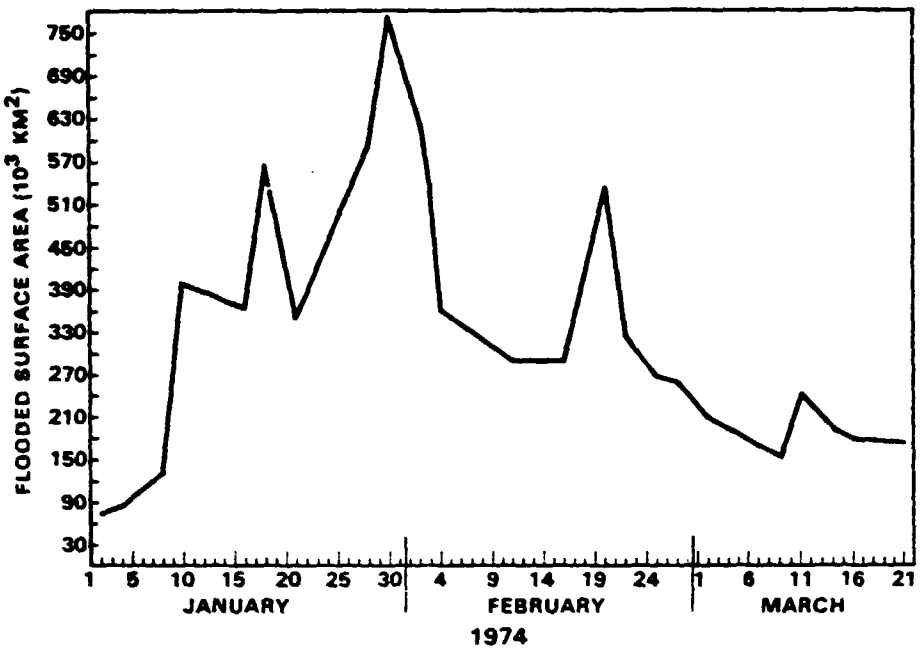


FIG. 13. Estimated total flooded surface area (km²) for the six regions shown in Fig. 11, in Eastern Australia from January to March 1974 derived from Nimbus-5 ESMR (19.35 GHz) data (day: $T_B \leq 250$ K, night: $T_B \leq 240$ K).

The 5 and 14 March maps (Fig 8g, h) indicate the continued movement of the shaded area southwestward down the Darling River, with the lowering of T_B becoming less intense, i.e., the minimum is now about 220 K. Similarly the shaded region of the lakes has also shrunk.

To study the temporal variation of T_B, a hydrograph (Fig. 10) for the Barwon River at Walgett, provided by the New South Wales Water Resources Commission, was compared with the values of T_B for a 50 km by 50 km area around the town of Walgett. Generally, the low T_B values relate well to the river gage height that is an indicator of the overbank flow. In particular, the peak of gage height occurs on 17 January, slightly prior to the minimum value of $T_B = 184$ K on 19 January 1974. This value indicates that 80–90% of this 50 by 50 km area was flooded. The subsequent rise of T_B corresponds well to the decrease in gage height.

An initial attempt was made to measure the flooded area from the Nimbus-5 ESMR grid print maps of brightness temperatures. Six regions that were covered by $T_B \leq 240$ K (night) and ≤ 250 K (day) were planimetered (Fig. 11). The daily integrated changes for area 6 along the Darling River basin is shown in Fig. 12. Starting at zero before the rains of 9–11 January, it reaches a peak on 10 January that probably includes a substantial area with bare wet soils. By 15 January, these surface soils have dried out, the area reduces to a value of 50 000–60 000 km², and the area continues to decrease with the exception of sharp spikes resulting from rains in late January and mid-February 1974. Recall that the shaded areas contained at least 20–30% standing water; therefore, the actual flooded area will be somewhat less than that indicated in Fig. 12. Fig. 13 shows the daily integrated flooded area changes for all areas 1–6. One major flood peak occurred on 30 January 1974, while minor peaks occurred on 10 and 18 January, 20 February, and 11 March 1974. Color pictures of the near-flood area maximum on 1 February and near-minimum on 11 March 1974 are shown on the cover. A more complete description of the wet season flooding in the South Carpentaria Plains (Area 2) may be found in a paper by Simpson and Doutch (1977).

4. Conclusions

The passive microwave brightness temperature observations of the Nimbus-5 ESMR were used to detect and monitor the development of flooded areas in the Darling River basin of New South Wales, Australia. This capacity was based on the observation of areas of persistent lowered values of T_B, which were attributed to the presence of flooded areas or standing water within the ESMR field of view. In the Darling River basin, the temporal variation of both size and location was monitored through the first quarter of 1974. The area with substantial flooding varied in size from approximately 60 000 km² immediately following the heavy rains in mid-January to about 10 000 km² in late March 1974. Its center moved approximately 600 km from the town of Walgett to halfway between the town of Wilcannia and Menindee, and it changed in shape from roughly

circular to oblong as it moved downstream. The motion is believed to be that of the floodwaters since there were no anomalously large rains over the downstream portion of the basin following the heavy rains early in January 1974.

These results indicate that microwave brightness temperature observations are another tool for monitoring surface water. This satellite capability will be considerably enhanced when improved spatial resolution becomes available in the future. Longer wavelength sensors will also provide greater sensitivity to soil moisture and standing water with less sensitivity to vegetation cover.

References

Adler, R. F., and E. B. Rodgers, 1977: Satellite-observed latent heat release in a tropical cyclone. *Mon. Wea. Rev.*, 105, 956–963.

Allison, L. J., 1975: A multisensor analysis of Nimbus 5 data recorded on January 22, 1973. *NASA TN D-7911*, Goddard Space Flight Center, Greenbelt, Md., p. 41.

——, 1977: Geological applications of Nimbus radiation data in the Middle East. *NASA TN D-8469*, Goddard Space Flight Center, Greenbelt, Md., p. 79.

——, E. B. Rodgers, T. T. Wilheit, and R. W. Fett, 1974: Tropical cyclone rainfall as measured by the Nimbus 5 electrically scanning microwave radiometer. *Bull. Am. Meteorol. Soc.*, 55, 1074–1089.

Austin, P. M., and S. G. Geotis, 1978: Evaluation of the quality of precipitation data from a satellite-borne radiometer. Final report, *NASA Grant NSG 5024*, Department of Meteorology, Massachusetts Institute of Technology, Cambridge, Mass., p. 32.

Australian Water Resources Council, 1976: Review of Australia's water resources, 1975. Department of Natural Resources, Australian Government Publishing Service, Canberra, Australia, p. 170.

Barton, I. J., 1978: A case study comparison of microwave radiometer measurements over bare and vegetated surfaces. *J. Geophys. Res.*, 83, 3513–3517.

Bureau of Meteorology, Department of Science, 1974: *Monthly Weather Review, New South Wales and Queensland, Australia*, January and February, p. 24.

Department of National Development, 1956. Conservation of surface water. *Atlas of Australian Resources*, Canberra, Australia, p. 12.

Foley, J. C., 1954: The climate of Australia. *Atlas of Australian Resources*, Climate Regions, Department of National Development, Canberra, Australia, p. 16.

Gibbs, W. J., and J. V. Maher, 1967: Rainfall deciles as drought indicators. *Bull. 48*, Bureau of Meteorology, Melbourne, Australia, p. 117.

Kidder, S. Q., and T. H. Vonder Haar, 1977: Seasonal oceanic precipitation frequencies from Nimbus 5 microwave data. *Geophys. Res.*, 82, 2083–2086.

Laut, P., G. Keig, M. Lazarides, E. Loffler, C. Margules, R. M. Scott, and M. E. Sullivan, 1977: Environments of South Australia. Province 8, Northern Arid, CSIRO, Canberra, Australia, p. 254.

Leeper, G. W., 1973: The Australian environment. Commonwealth Scientific and Industrial Research Organization (CSIRO), Melbourne University Press, Melbourne, Australia, p. 163.

Meneely, J. M., 1975: Application of the Nimbus 5 ESMR to rainfall detection over land surfaces. *Final report E/S 1008*, Earth Satellite Corp., Washington, D.C., p. 48.

Moore, R. K., L. J. Chastant, L. J. Porcello, J. Stevenson, and F. T. Ulaby, 1975: Microwave remote sensors. *Manual of Remote Sensing, Vol. 1*, American Society of Photogrammetry, Falls Church, Va., pp. 499–534.

Nordberg, W., J. Conaway, D. B. Ross, and T. Wilheit, 1971: Measurements of microwave emission from a foam-covered, wind-driven sea. *J. Atmos. Sci.*, 28, 429–435.

Otterman, J., P. D. Lowman, and V. V. Salomonson, 1976: Surveying earth resources by remote sensing from satellites. *Geophysical Surveys, 2*, D. Reidel Publishing Co., Dordrecht, Holland, pp. 431–467.

Pittock, A. B., 1975: Climate change and the patterns of variations in Australian rainfall. *Search*, 6, 498–504.

Rango, A., and A. T. Anderson, 1974: Flood hazard studies in the Mississippi river basin using remote sensing. *Water Resour. Bull.*, 10(5), American Geophysical Union, Washing, D.C., pp. 1060–1081.

——, and V. V. Salomonson, 1974: Regional flood mapping from space. *Water Resour. Res.* 10(3), American Geophysical Union, Washington, D.C., pp. 473–484.

Rao, M. S. V., and J. S. Theon, 1977: New features of global climatology revealed by satellite-derived oceanic rainfall maps. *Bull. Am. Meteorol. Soc.*, 58, 1285–1288.

——, W. V. Abbott, and J. S. Theon, 1976: Satellite-derived global oceanic rainfall atlas. *NASA SP-410*, National Aeronautics and Space Administration, Washington, D.C., p. 184.

Robinove, C. J., 1978. Interpretation of a landsat image of an unusual flood phenomenon in Australia. *Remote Sensing Environ.*, 7, 219–225.

Rodgers, E. B., H. Siddalingaiah, A. T. C. Chang, and T. Wilheit, 1978: A statistical technique for determining rainfall over land employing Nimbus 6 ESMR measurements. *Preprints, Third Conference on Atmospheric Radiation (Davis, Calif.)*, AMS, Boston, pp. 126–129.

Savage, R. C., and J. A. Weinman, 1975: Preliminary calculations of the upwelling radiance from rain clouds at 37.0 and 19.35 GHz. *Bull. Am. Meteorol. Soc.*, 56, 1272–1274.

Schmugge, T. J., A. Rango, L. J. Allison, and T. T. Wilheit, 1974: Hydrologic applications of Nimbus 5 ESMR data. *NASA X-910-74-51*, Goddard Space Flight Center, Greenbelt, Md., p. 21.

——, J. M. Meneely, A. Rango, and R. Neff, 1977: Satellite microwave observations of soil moisture variations. *Water Resour. Bull.*, 13(2), American Water Resources Association, pp. 265–281.

Short, N. M., P. D. Lowman, S. C. Freden, and W. A. Finch, 1976. Mission to earth, Landsat views the world. *NASA SP-360*, National Aeronautics and Space Administration, Washington, D.C., pp. 422–423.

Simpson, C. J., and H. F. Doutch, 1977: The 1974 wet season flooding in the southern Carpentaria plains, Northwest Queensland, BMR. *J. Austr. Geol. Geophys.*, 2, 43–51.

Stout, J. E., and D. W. Martin, 1979: Estimating rain with geosynchronous satellite images. (Unpublished manuscript.)

Ulaby, F. T., 1977: Microwave remote sensing of hydrologic parameters. Proceedings of the Eleventh International Symposium on Remote Sensing of Environment, Vol. 1, 25–29 April 1977, Environmental Research Institute of Michigan, Ann Arbor, Michigan, pp. 67–86.

Vickery, J., 1974: The greater flooding of the Lake Eyre Basin—1974. *Proc. R. Geogr. Soc. Aust.*, 75, 23–36.

Warner, R. F., 1977: Hydrology. *Australia, a Geography*, edited by D. N. Jeans, Chap. 3, St. Martin's Press, New York, p. 571.

Wiesnet, D. R., D. F. McGinnis, and J. A. Pritchard, 1974: Mapping of the 1973 Mississippi River floods by the NOAA-2 satellite. *Water Resour. Bull.*, 10(5), American Water Resources Association, pp. 1040–1049.

Wilheit, T. T., 1972: The electrically scanning microwave radiometer (ESMR) experiment. *Nimbus 5 Users Guide*, NASA Goddard Space Flight Center, Greenbelt, Md., pp. 55–105.

——, 1973: ESMR corrections to the user's guide. *Nimbus 5 Data Catalog, Vol. 3*, Goddard Space Flight Center, Greenbelt, Md., pp. 5–4 to 5–6.

——, J. Theon, W. Shenk, L. J. Allison, and E. Rodgers, 1973: Meteorological interpretations of the images from Nimbus 5 electrically scanning microwave radiometer. *J. Appl. Meteorol.*, 15, 168–172.

——, A. T. C. Chang, M. S. V. Rao, E. B. Rodgers, and J. S. Theon, 1977: A satellite technique for quantitatively mapping rainfall rates over the ocean. *J. Appl. Meteorol.*, 16, 551–560.

Zwally, H. J., and P. Gloersen, 1977: Passive microwave images of the polar regions and research applications. *Polar Rec.*, 18, 431–450.

Paper 111

Authorized Reprint from
Special Technical Publication 653
Copyright
American Society for Testing and Materials
1916 Race Street, Philadelphia, Pa. 19103
1978

L. J. Allison,[1] *Raymond Wexler,*[1] *C. R. Laughlin,*[1] *and W. R. Bandeen*[1]

Remote Sensing of the Atmosphere from Environmental Satellites

REFERENCE: Allison, L. J., Wexler, Raymond, Laughlin, C. R., and Bandeen, W. R., "**Remote Sensing of the Atmosphere from Environmental Satellites.**" *Air Quality Meteorology and Atmospheric Ozone. ASTM STP 653.* A. L. Morris and R. C. Barras, Eds., American Society for Testing and Materials, 1978, pp. 83-155.

ABSTRACT: The use of satellites for the remote sensing of the earth and its atmosphere is reviewed. Meteorological satellites have been in use since 1960. The polar orbiting satellites at 650 to 1500 km make 12 to 14 orbits daily with nearly full earth coverage. The geosynchronous satellites at 36 000 km remain approximately fixed in space relative to the earth and view the same area continuously.

Cameras obtaining earth images in the visible detect clouds and monitor storm systems. Infrared radiometers measure temperatures of the surface or cloud tops and also determine amounts of ozone and water vapor in the atmosphere. Radiometers sensitive to different spectral regions in absorption bands (such as carbon dioxide [CO_2] at 15 and 4.3 μm) are used to determine atmospheric vertical temperature profiles.

The earth's radiation budget is determined from satellite measurements of the reflected solar radiation and the emitted longwave terrestrial radiation.

Atmospheric winds are obtained by cloud tracking from geosynchronous satellites. A major problem is height determination. Infrared "window" measurements of the brightness temperature from clouds provide some information on cloud top heights, but cloud emisisivity and sensor resolution are still sources of error.

Microwave imagery has been used to determine rainfall intensities over the ocean, snow, and ice cover and flood mapping through clouds. It has also been possible to distinguish between first and multiyear ice. There is a potential capability for measuring soil moisture using long-wavelength (21 cm) radiometers. Landsat visual and near-infrared measurements also allow flood monitoring and observations of crop growth.

The Nimbus 4 Backscatter Ultraviolet (BUV) instrument has provided seven years of measurements from which total ozone and vertical ozone distribution can be determined.

Future missions include TIROS N satellites with improved capabilities for measuring sea surface temperatures, mapping snow and ice fields, monitoring global ozone, and continued measurements of the earth's radiation budget. SEASAT A will be devoted to oceanographic measurements.

Nimbus G sensors will observe certain atmospheric pollutants, oceanic parameters, and weather and climate variables. The Shuttle, in addition to serving as a launch

[1]Goddard Space Flight Center, Greenbelt, Md. 20771.

vehicle for free flying satellites, will make periodic calibration checks of instruments on the free flyers and provide facilities for developing, testing, and demonstrating new remote sensors.

KEY WORDS: satellite meteorology, remote sensing theory, satellite instruments, meteorology, operational satellite data, first GARP global experiment, environmental satellite programs, climate program, ozone

Imaging systems that respond to electromagnetic radiation (EMR) within the spectral response range of the human eye have had a long and productive history of development since Daguerre and Niepce reported on their first photographs in 1839. Aerial photography was utilized as early as 1920 by petroleum geologists and was used in 1944 for mineral exploration. Today, aerial and satellite color photography is used in almost all the earth sciences, including geology, soils, forestry, agriculture, hydrology, geography, and range and wildlife management. The existence of EMR adjacent to the blue-violet region of the visible spectrum was demonstrated by J. W. Ritter in 1801, by showing that this invisible EMR was even more effective than visible light in blackening silver chloride. The existence of the infrared (IR) spectral region was demonstrated as early as 1800 by Sir Frederick Herschel, and initial work in IR observations of celestial bodies was done by S. P. Langley in the late 1800s.

Although the existence of these regions of the EMR spectrum beyond the range of human vision and photographic sensitivity has been known to the scientific world for over 170 years, nearly all useful applications have taken place over only the past 40 years. The term "remote sensing" has become a commonly accepted term and is used in the present context to refer to the whole realm of possibilities afforded by the entire EMR spectrum for observations of the earth's surface and atmosphere from satellites. As sunlight filters down from the top of the atmosphere, it interacts with all particles in any thin slab of the atmosphere—atomic, molecular, and larger—and each slab scatters and absorbs radiation (as does the surface, that is, the lower boundary of the atmosphere). All photons scattered in the direction of the satellite from all intervening slabs of the atmosphere (as well as the surface) have the same energy as the incident photons. The solar energy absorbed throughout the atmosphere (and by the surface) is ultimately reemitted at different wavelengths so that atmospheric interactions throughout the EMR spectrum are important to remote sensing.

Any physical interaction between the radiation field and the atmosphere offers a potential means for probing that part of the environment involved in the interaction. It is only necessary to find interactions that are strong enough to be observable, sensitive enough to allow detection of variations, and isolated enough from other interactions to permit unambiguous interpretations.

In the following section, the basis for remote sensing will be outlined, in

the sequel, the evolution of remote sensors will be illustrated along with examples of practical applications and a prospectus of future developments. Many volumes already exist on single topics covered here, so that completeness has been curtailed in the interest of presenting a broad overview.

Basis of Remote Sensing

The prime source of energy that produces and maintains the atmospheric motions and the spatial and temporal variations of weather is the solar radiation intercepted by the earth. The rate at which this energy is received at the earth's distance from the sun is nearly 1400 w/m^2, but, on the average, approximately 30 percent of this is reflected back to space by the cloud surfaces, the clear atmosphere, the dust and ice crystals suspended in the atmosphere, and by the earth's surface. Figure 1 illustrates schematically how the flux from the sun can arrive at an Earth-viewing satellite detector. The radiation may have been scattered, reflected, absorbed, and re-emitted several times, and each time in a way that is characteristic of the composition, temperature, and pressure of the atmospheric gases and

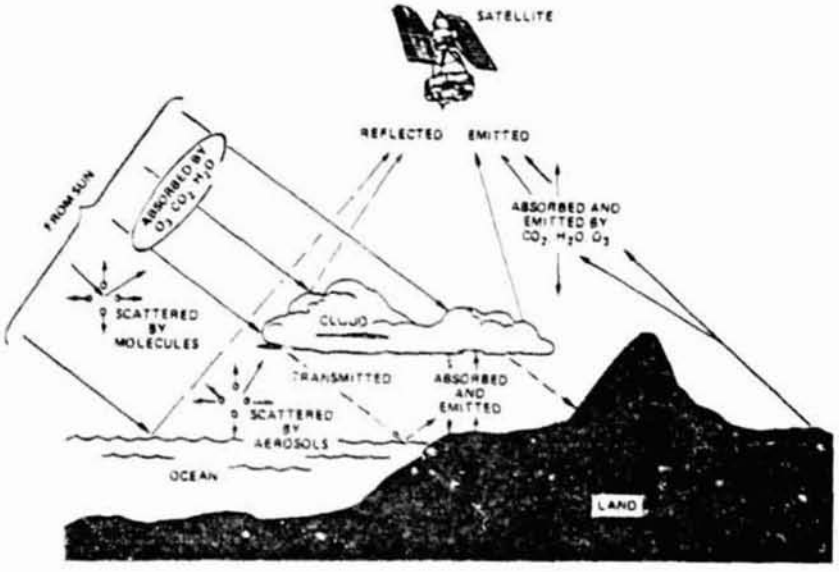


FIG. 1—*A schematic illustration of earth-sun-satellite radiative interactions.*

the physical properties of the various scatterers. A more complete description of light scattering in planetary atmospheres is given elsewhere [1].[2]

Figure 2 shows the extent of the EMR spectrum and indicates the kinds of physical processes involved at different wavelengths, along with the names associated with the various intervals. Generic types of instrumentation commonly associated with these intervals are also indicated.

Figure 3 shows the spectral characteristics of sunlight. The upper solid curve is for above the earth's atmosphere where the maximum occurs at a wavelength of about 0.47 μm. About 20 percent of the total energy is in the region below 0.47 μm; 44 percent is in the visible band between 0.4 and 0.76 μm, and 99 percent is at the so-called short wavelengths between 0.15 and 4.0 μm. The lower solid curve is typical of sunlight for cloud-free areas at the surface, so that the difference represents that which is scattered or absorbed. The atmosphere is essentially transparent in the visible and near-infrared portions of the spectrum, and these bands of least (but not insignificant) attenuation are called "windows." The windows of importance to remote sensing are shown in Table 1. Atmospheric absorption at wavelengths below 0.3 μm is so intense that little energy reaches the surface. The lower solid curve of Fig. 3 is typical of the diffuse shortwave radiation that reaches the surface after absorption by the indicated atmospheric gases with the remainder being scattered by water vapor, dust, and other atmospheric molecules.

Figure 4 is an expansion of Fig. 3 with the addition of typical thermal infrared radiation emitted from the surface. In this band, from about 4 to 20 μm, strong absorption bands occur that are caused by vibrational-rotational changes due to water vapor and carbon dioxide. These absorption bands break up both the near-infrared and thermal infrared regions into a series of windows of modest transparency interspersed with regions shuttered by absorption. Figure 5 shows the transmittance of the atmosphere within this band and illustrates the complex and pervasive nature of the EMR spectrum that must be dealt with in remote sensing [2].

At wavelengths larger than 22 μm, a whole set of rotational transitions due to water vapor out to 1 mm and beyond, effectively close the long-wavelength infrared and, 1-mm regions to remote sensing. No important molecular transitions occur for wavelengths longer than a few centimetres.

Figure 6 shows a further expansion to include the microwave region where the indicated atmospheric absorption bands are between the windows, as given in Table 1, while the rest is very transparent. The effects of clouds are relatively minor over most of the microwave region, but for measurements such as the sea surface temperature, their effects must be taken into account. Rainfall interacts with EMR in ways that are strongly dependent on wavelength and drop-size in the 1 to 10-cm region. Only the most intense

[2]The italic numbers in brackets refer to the list of references appended to this paper.

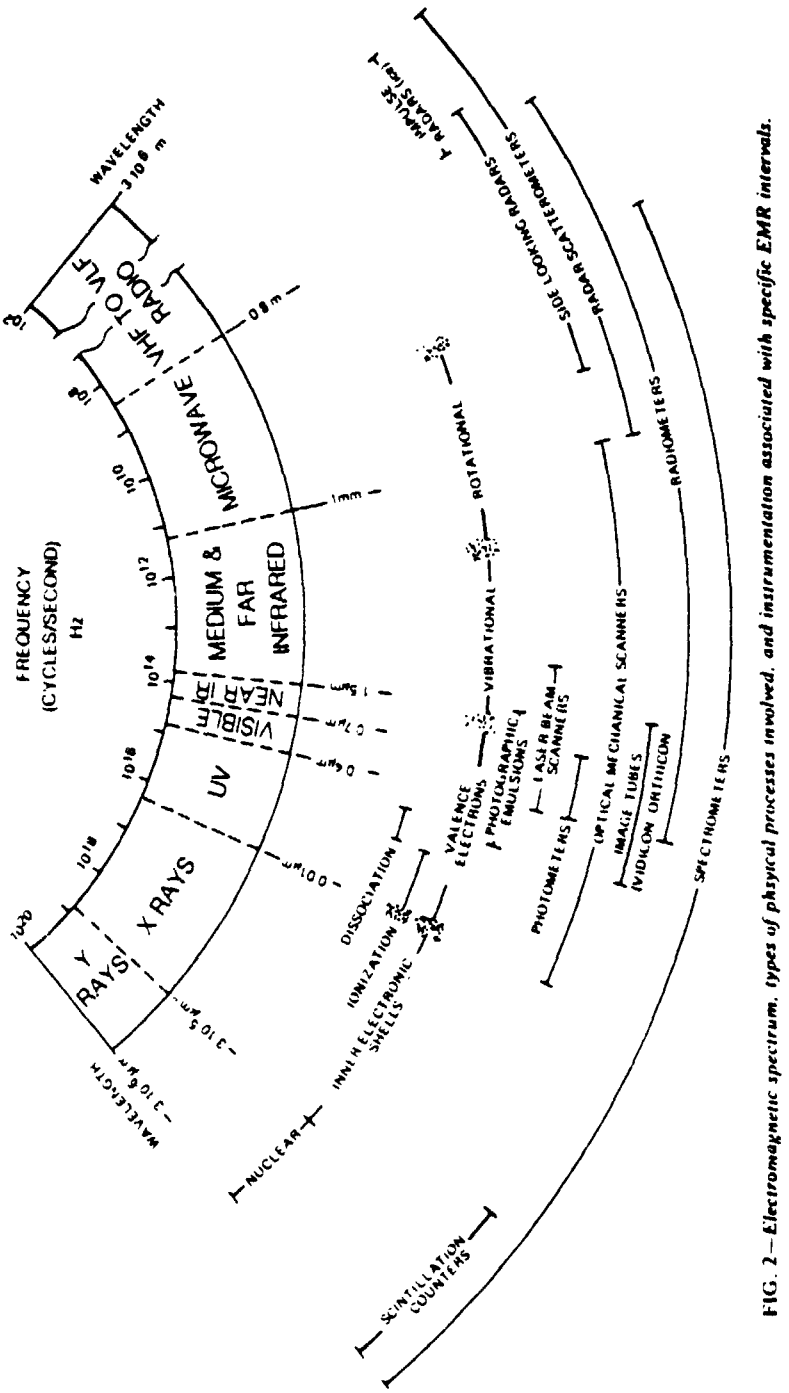


FIG. 2—*Electromagnetic spectrum, types of physical processes involved, and instrumentation associated with specific EMR intervals.*

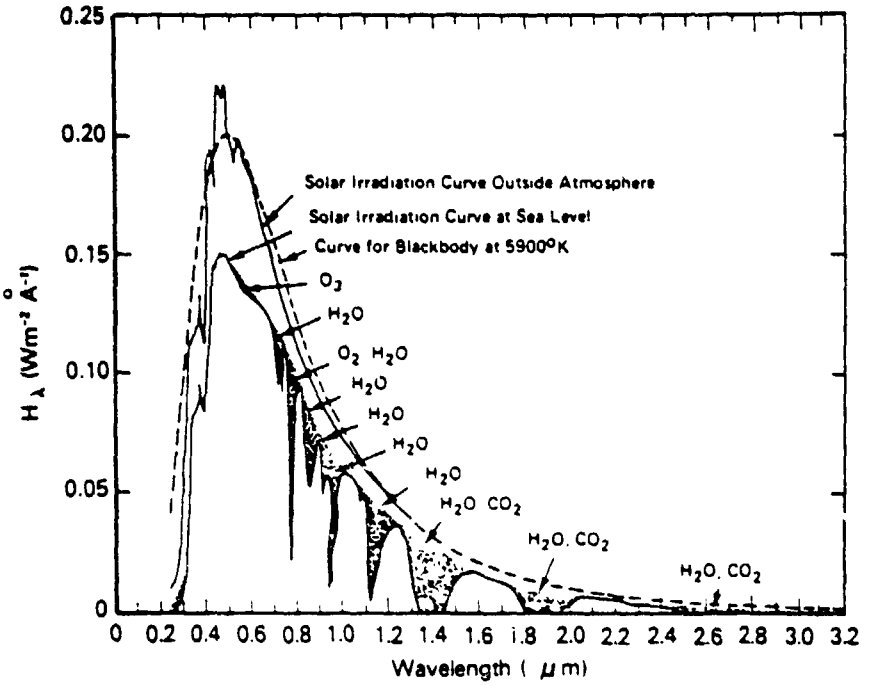


FIG 3—*Spectral irradiance (HX) of direct sunlight before and after it passes through the earth s atmosphere The stippled portion gives the atmospheric absorption the sun is at zenith* [91].

TABLE 1—*Major atmospheric windows available for spacecraft Remote Sensing* [90].

Ultraviolet and visible μm	0.30 to 0.75
	0.7[illegible] to 0.91
Near infrared, μm	1 0 to 1.12
	1.19 to 1.34
	1.55 to 1.[illegible]5
	2.05 to 2.4
Mid-infrared, μm	3.5 to 4 16
	4 5 to 5 0
Thermal infrared, μm	8 0 to 9 2
	10.2 to 12.4
	17 0 to 22.0
Microwave, mm	2.06 to 2.22
	3.0 to 3 [illegible]5
	[illegible] 5 to 11 5
	20 0 +

rainfall can be detected by microwave at wavelengths beyond 10 cm because rain echoes are proportional to the inverse fourth power of the wavelength.

Evolution of Environmental Satellite Instrumentation

The development of the first meteorological satellite was a final result of man's determination to increase his perspective of the world in which he lives. An early experiment in satellite meteorology began in 1959 with the U.S. launching of the Vanguard 2 and Explorer 6 and 7. The first two

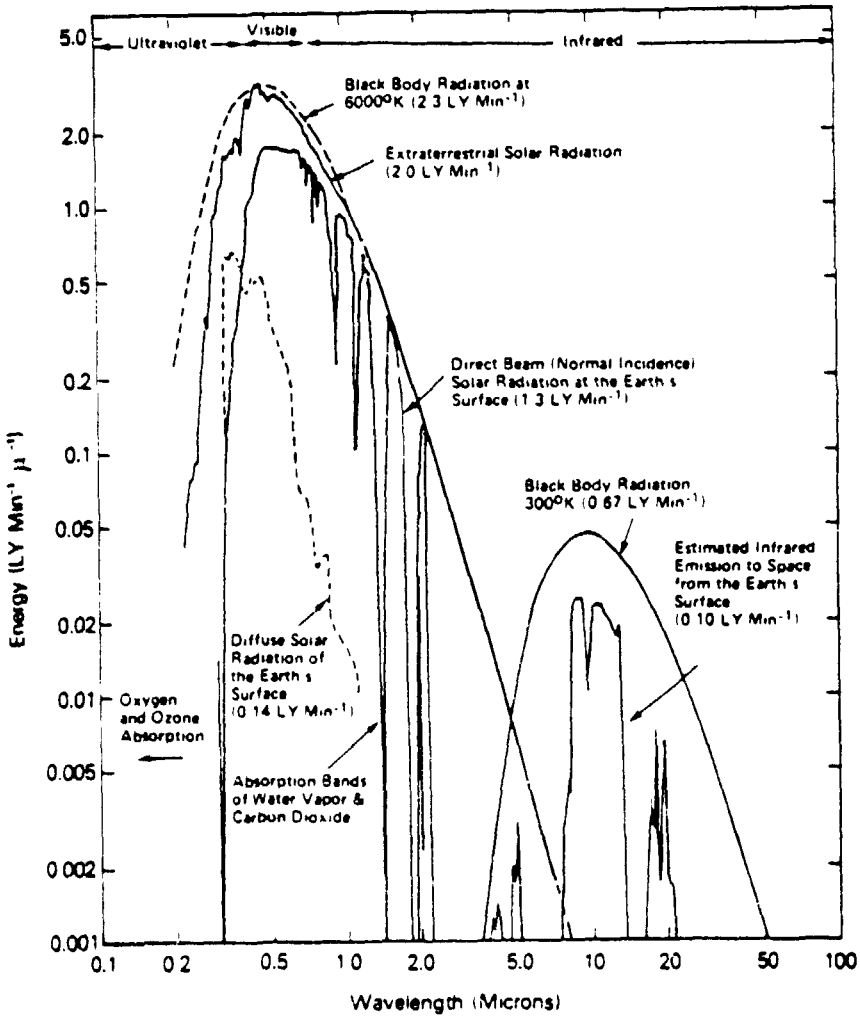


FIG 4—*Electromagnetic spectra of solar and terrestrial radiation* [92]

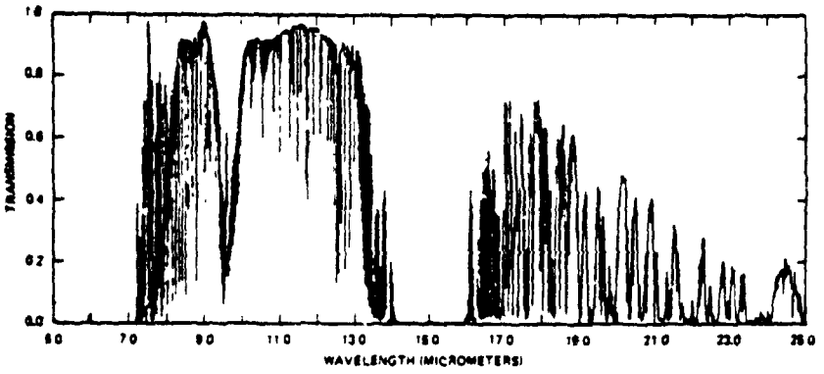


FIG. 5—*Measured transmittances through the atmosphere for the 5 to 25-μm region* [2].

spacecraft were equipped with two photocells and primitive television scanning devices, respectively, which were designed to observe large-scale cloud patterns. Explorer 7 was fitted with a low resolution, omnidirectional radiometer to measure incoming reflected solar radiation and outgoing thermal radiation emitted by the earth's atmospheric system, in order to determine the heat budget. These early flights, while not entirely successful, provided U.S. scientists with valuable in-house experience and encouraged the development of the first experimental spin-stabilized TIROS satellite [3]. Figure 7 shows an artist's conception of the wide variety of U.S meteorological satellites which were launched into polar and geostationary orbits from 1960 to 1976.

Polar-orbiting environmental satellites circle the earth at altitudes which vary from 650 to 1500 km. Near-to-global data are recorded by tape recorder from 12 to 14 daily orbits, with each orbit being 90 to 110 min in duration By command, ground stations receive the data and further process it for interested user agencies. Satellites in this category are the TIROS, ESSA, ITOS, NOAA, NIMBUS, and LANDSAT series. With successive launches, it became increasingly apparent that the polar orbiter had a distinct advantage for many users by repetitively viewing large areas with high, medium, and low resolution sensors under similar solar illumination and orbital operating conditions.

Geostationary satellites were positioned into orbits approximately 36 000 km above the equator at predetermined longitudes. Since the satellite has a period equal to the sideral rotation of the earth (23 h, 56 min), it remained fixed in space relative to the earth, and thus observed the same area continuously with visible and later infrared instruments. Satellites in this category are the ATS and SMS/GOES series. Table 2 lists in chronological order, the basic scientific-instrumental accomplishments of the U.S. environmental satellite program from 1960 to 1976.

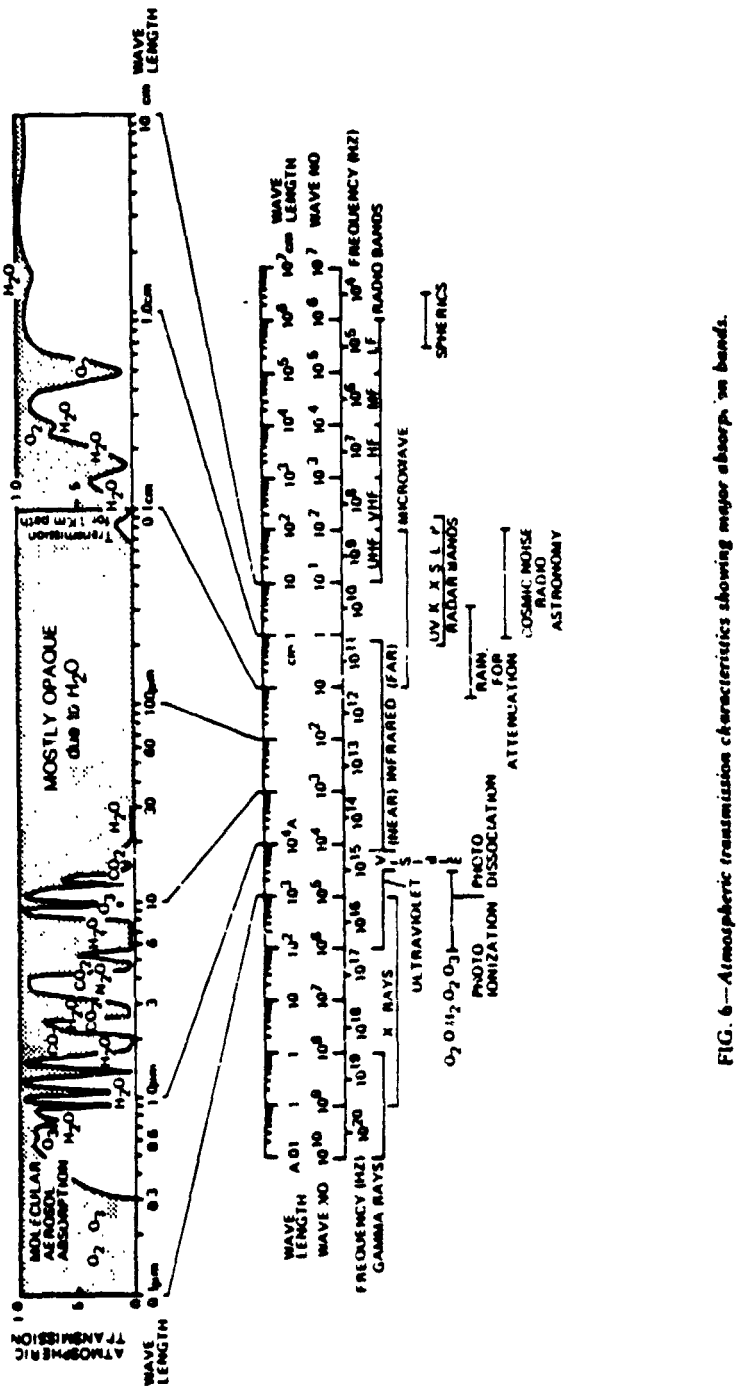


FIG. 6—*Atmospheric transmission characteristics showing major absorption bands.*

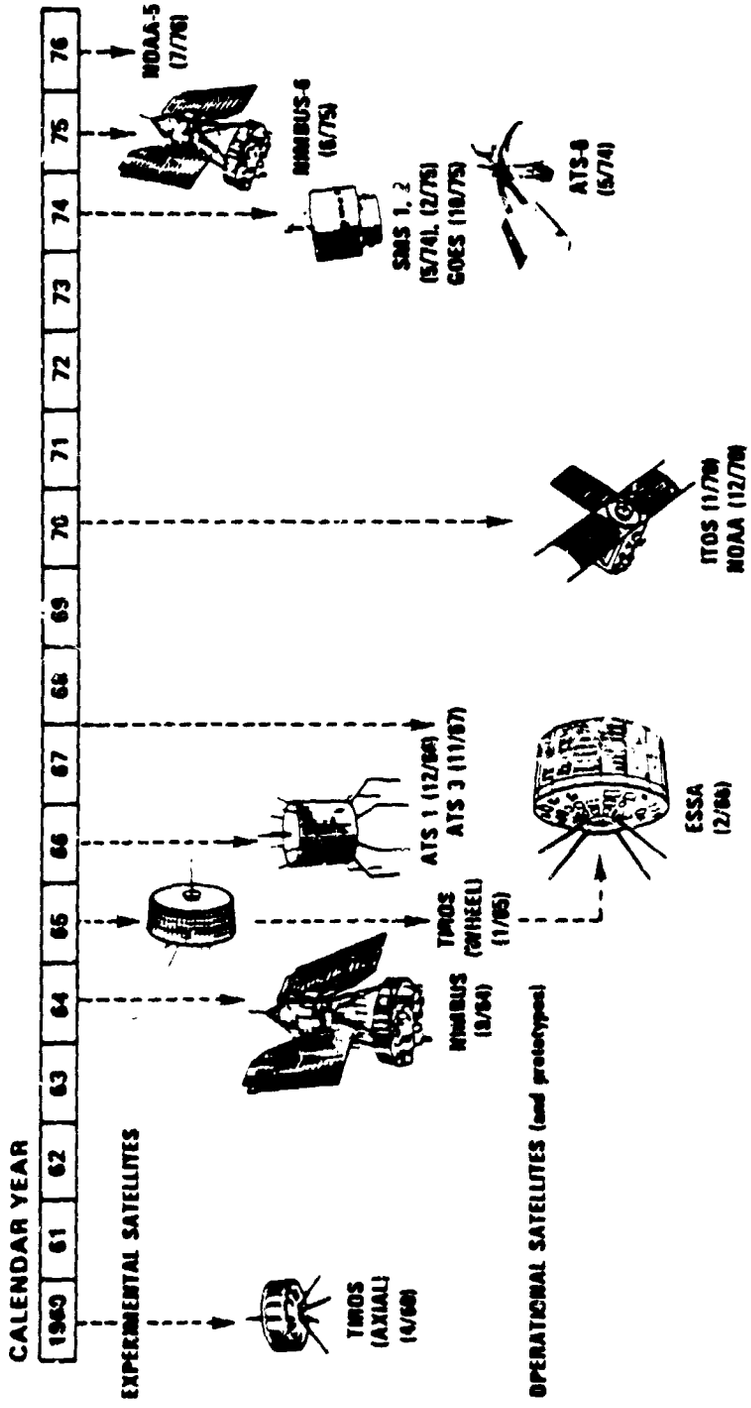


FIG. 7.—*An artist's conception of the U.S. meteorological satellite program from 1960 to 1976.*

In the subsequent sections, the development of three major instrument classifications, that is, cameras, scanners, and atmospheric sounders will be discussed briefly.

Cameras

The evolution of meteorological satellite cameras is shown in Fig. 8. TIROS 1, 2, 3, 4, and 7 carried shuttered television cameras. The photosensitive vidicon tube face, 1.27 cm in diameter, which viewed the earth scene below was internally scanned (500 times per image) and converted to a recorded analog signal which was later transmitted to the earth upon command.

An example of early TIROS 1 television pictures is shown in Fig. 9 [4]. The Advanced Vidicon Camera System (AVCS) and the Automatic Picture Transmission system (APT) were tested on the Nimbus 1 and 2 satellites. The 3-camera AVCS was intended to provide fairly high resolution cloud cover—800 TV lines, 926 m (0.5 nautical miles [NM])—over the full earth and was later flown on ESSA, ITOS, and NOAA operational satellites. A more complete description of the degradation and other limitations of the vidicon tube is given by Schwalb and Gross [5]. The Nimbus 1 and 2 APT cameras used special long-storage vidicon tubes whose dielectric surfaces were capable of holding the earth image for the required 200 s, slow-scan readout period. The APT vidicon system (800 TV lines, 3704-m [2-NM] resolution) was designed for medium resolution cloud cover imaging for direct readout, and was flown on ESSA, ITOS, and NOAA satellites. The APT data from three successive orbits covers 3445 km (1860 NM) and is received daily by 600 to 800 ground stations, including 78 stations in 43 foreign countries [6].

The Image Dissector Camera (IDCS), an electrically scanning photocathode, was flown on ATS 1 and 3, Nimbus 3 and 4, and the ESSA-NOAA series, and had a higher dynamic range (1300 TV lines) than the AVCS system. But due to its complexity, the IDCS was replaced by the simpler Spin Scan Cloud Cover Camera on the geostationary ATS satellites [7,8].

Scanners

The evolution of meteorological satellite scanner is shown in Fig. 10. The Medium Resolution Infrared Radiometer (MRIR) was the first 5-channel scanning radiometer to be flown on the TIROS-Nimbus series. It recorded the visible and infrared radiation emitted and reflected back from the earth, clouds and oceans with a ground resolution of 64 km (35 NM). For the first time, estimates on a quasi-global scale were made of cloud heights, sea surface and ground temperatures, and radiation budget. This instrument, which suffered from a lack of inflight calibration, was the forerunner

TABLE 2—*Basic scientific-instrumental accomplishments of the United States environmental satellite program from 1960 to 1978.*

Satellite	Launch	Ceased Operation	Remarks
TIROS 1	04/01/60	06/15/60	First cloud cover imaging in the visible (Vidicon).
TIROS 2	11/23/60	02/07/61	First multispectral imaging in the visible and infrared.
TIROS 3	07/12/61	10/30/61	...
TIROS 4	02/08/62	06/12/62	...
TIROS 5	06/19/62	05/05/63	...
TIROS 6	09/18/62	10/11/63	...
TIROS 7	06/19/63	02/03/66	...
TIROS 8	12/21/63	01/22/65	First Automatic Picture Transmission (APT).
Nimbus 1	08/28/64	09/23/64	Carried Advanced Vidicon Camera System (AVCS), APT, and High Resolution Infrared Radiometer (HRIR) for night coverage.
TIROS 9	01/22/65	02/15/67	First TIROS satellite in sun-synchronous orbit.
TIROS 10	07/01/65	07/31/66	...
ESSA 1	02/03/66	05/08/67	First operational weather satellite; carried 2 wide-angle TV cameras.
ESSA 2	02/28/66	10/16/70	Carried APT cameras; APT carried on all even-numbered ESSA satellites.
Nimbus 2	05/15/66	11/15/66	First direct-readout of HRIR. Carried Medium Resolution Infrared Radiometer (MRIR).
ESSA 3	10/02/66	10/19/68	Carried first operational AVCS cameras; AVCS carried on all odd-numbered ESSA satellites.
ATS 1	12/06/66	10/16/72	Carried first Spin Scan Cloud Camera (SSCC), continuously imaging in visible.
ESSA 4	01/26/67	06/19/67	...
ESSA 5	04/20/67	02/20/70	...
ATS 3	11/05/67	10/30/75	Carried first Multicolor Spin Scan Cloud Camera (MSSCC).
ESSA 6	11/10/67	11/04/69	...
ESSA 7	08/16/68	07/19/69	...
ESSA 8	12/15/68	03/12/76	...
ESSA 9	02/26/69	11/29/73	
Nimbus 3	04/14/69	09/25/70	First vertical sounding of the atmosphere (SIRS, IRIS) data collection and location (IRLS), and solar UV monitoring (MUSE). Also carried MRIR.
ITOS 1	01/23/70	06/17/71	Second generation operational prototype (Improved TIROS Operational Satellite).
Nimbus 4	04/08/70	10/17/77	First vertical sounding to 60 km (SCR) and ozone sounding globally (BUV).
NOAA 1	12/11/70	08/19/71	First NOAA-funded, second-generation operational satellite.

TABLE 2—*Continued*

Landsat 1	07/23/72	...	First Earth resources technology satellite. MSS failed 03/04/77.
NOAA 2	10/15/72	01/30/75	First operational satellite to carry visible and IR Scanning Radiometer (SR) and Vertical Temperature Profile Radiometer (VTPR).
Nimbus 5	12/12/72	06/02/75	First passive microwave sounder (NEMS) and imager (ESMR) flown.
NOAA 3	11/06/73	08/31/76	First operational satellite to permit direct broadcast of VTPR data.
SMS 1	05/17/74	01/08/76 (Standby)	First Synchronous Meteorological Satellite (GOES prototype). Carried visible and IR imaging radiometer.
ATS 6	05/30/74	...	First three-axis-stabilized geostationary satellite. (Infrared and visible scanner failed on 08/15/74).
NOAA 4	11/15/74	...	...
Landsat 2	01/22/75	09/16/76	...
SMS 2	02/06/75	04/04/78 (Standby)	Second operational geostationary SMS/GOES prototype.
Nimbus 6	06/12/75	09/01/76	First sounder including 4.3-μm CO_2 band flown (HIRS). First measurement of Earth radiation budget at synoptic/planetary scales (ERB).
GOES 1	10/16/75	...	First Geostationary Operational Environmental Satellite. (reactivated 04/04/78).
NOAA 5	07/29/76	...	SR failed 03/15/78.
Landsat 3	03/05/78	...	..

of several scanners: the High Resolution Infrared Radiometer (HRIR) [*9*], the Temperature-Humidity Infrared Radiometer (THIR) on the Nimbus series, the operational Scanning Radiometer (SR) and the Very High Resolution Radiometer (VHRR) on the NOAA series, the Very High Resolution Radiometer (VHR) on the Defense Meteorological Satellite Program (DMSP), and the Advanced Very High Resolution Radiometer (AVHRR) on the planned TIROS-N series (1978).

A typical scan sequence of the two-channel THIR on Nimbus 5 is shown in Fig. 11. the two detectors (11 and 6.7 μm, 8 and 23-km ground resolution at nadir, respectively) simultaneously view the spacecraft housing, A at zenith (zero seconds). Seven synchronous pulses start at C, followed by six 1-V calibration steps. A space scan starts between D and G, followed by an earth scan (117 deg wide) to I, a space scan to K, a housing scan to M, and then back to zenith. The entire sequence lasts 1.23 s. The space and housing scan serve as part of the in-flight calibration check. Note the small noise ripple on the housing and space scans [*10*]. The noise equivalent Δ temperature for the 11-μm channel is approximately 1.5 K at 185 K and 0.28 K at 300 K. Corrections of 1 (dry desert areas) to 10 K (moist

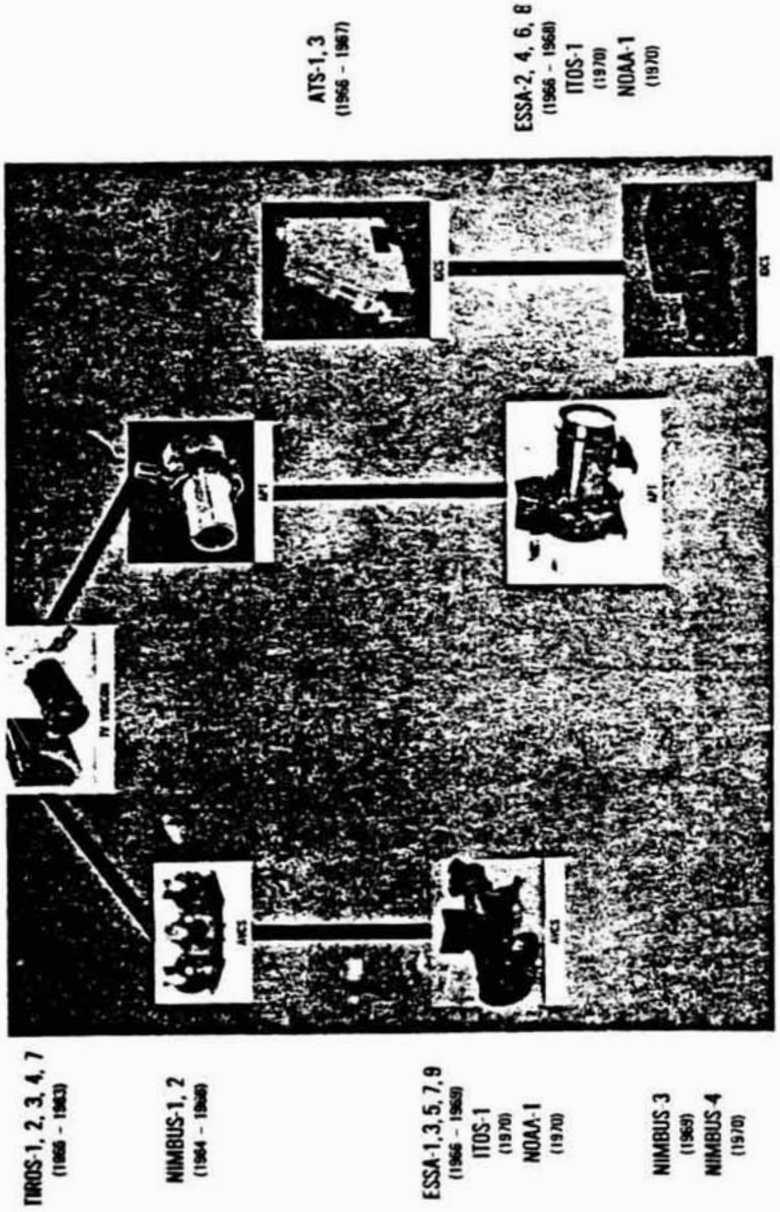


FIG 8—*Evolution of U.S. meteorological satellite camera systems from 1960 to 1970.*

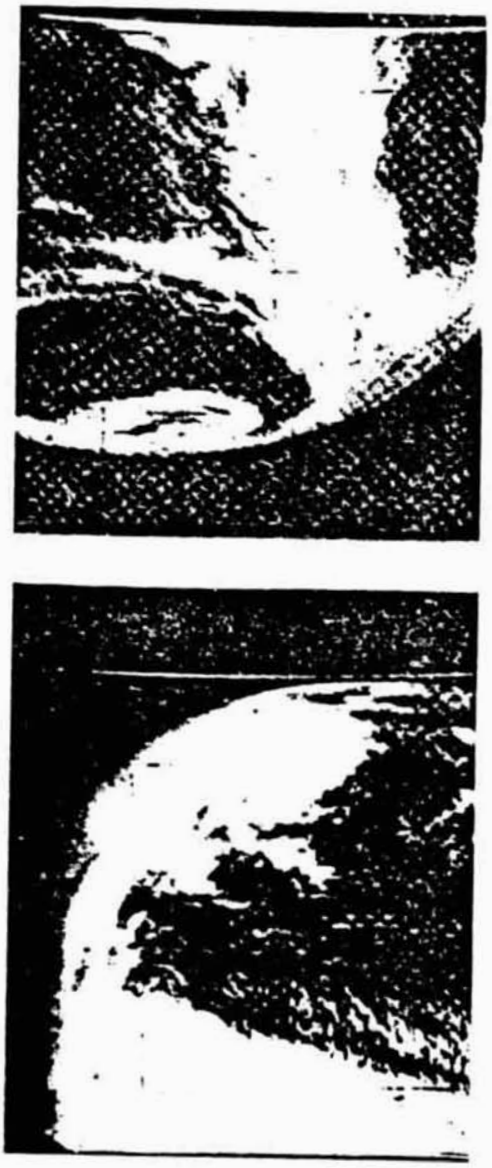


FIG. 9—*Examples of early TIROS I (1960) television pictures.*

ORIGINAL PAGE IS
OF POOR QUALITY

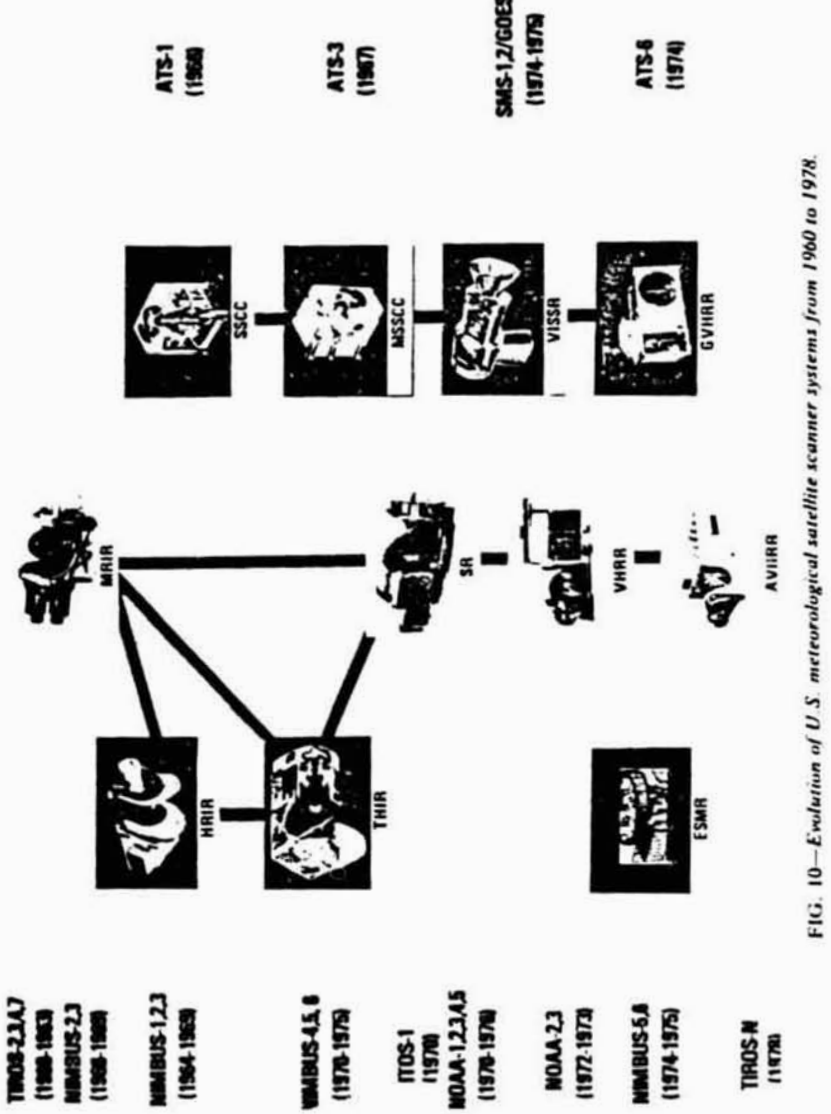
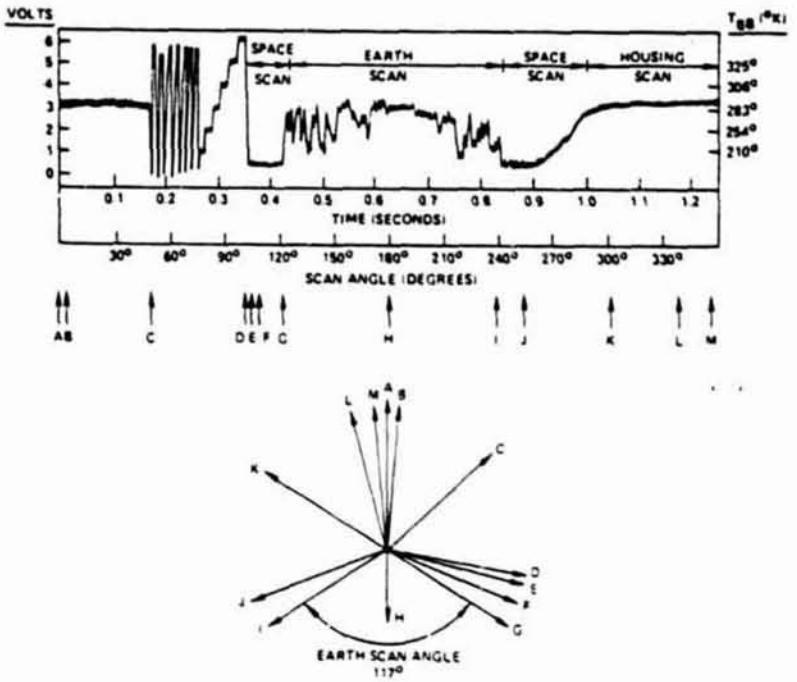


FIG. 10—*Evolution of U.S. meteorological satellite scanner systems from 1960 to 1978.*



LEGEND

REFERENCE LETTER	ANGLE (DEGREES)	TIME (MSEC)	EVENT
A	0°	0	SPACECRAFT ZENITH
B	5°	17.4	RADIOMETER IFOV JUST STARTING TO LEAVE HOUSING
C	48°	166.7	SCAN MIRROR POSITION PIP NO. 1 OCCURS AND RADIOMETER SYNC WORD AND CALIBRATION SIGNAL SEQUENCE IS STARTED. 6.7 MICRON CHANNEL GAIN RETURNED TO NORMAL
D	100°	347.5	RADIOMETER IFOV JUST STARTING TO SEE ALL OF SPACE
E	103.5°	359.4	CALIBRATE SIGNAL SEQUENCE ENDS AND RESTORE PERIOD STARTS
F	110.7°	384.4	RESTORE PERIOD ENDS
G	121.5°	422.2	EARTH SCAN PERIOD BEGINS (600 NMI ORBIT)
H	180°	625.0	SPACECRAFT NADIR
I	238.5°	828.8	EARTH SCAN PERIOD ENDS (600 NMI ORBIT)
J	250°	868.9	RADIOMETER IFOV JUST STARTING TO SEE HOUSING
K	302°	1048.5	SCAN MIRROR POSITION PIP NO. 2 OCCURS AND 6.7 MICRON CHANNEL GAIN IS ATTENUATED BY A FACTOR OF 3
L	345°	1197.9	RADIOMETER IFOV COMPLETELY FILLED BY HOUSING
M	356°	1232.6	RADIOMETER Z AXIS

FIG. 11—*Nimbus 5 THIR (11 μm) scan sequence.*

tropics), must be added to equivalent blackbody temperatures of the earth's surface and clouds for losses due to atmospheric water vapor. An example of SR imagery is shown in Fig. 12, which depicts the development of a Pacific storm as recorded by the visible channel of NOAA-2 from 12 to 14 March 1974 at approximately 2200 to 2300 GMT. Figure 13 shows a global montage of nighttime infrared data from NOAA-2 SR for 25 to 26 Aug. 1974. The locations of several eastern and western Pacific typhoons,

FIG. 12—*Development of a Pacific storm using NOAA 2 visible scanning radiometer data.* [93]

FIG. 11—A global montage of nighttime infrared data from NOAA-2 SR (10.5 to 12.5 μm) for the northern hemisphere, 8/25/74 to 8/26/74.

the Asian monsoon cloudiness and the intertropical zone of convergence over Central Africa are indicated.

The ATS-1 and -3 (Fig. 10) were equipped with a Spin Scan Cloud Cover Camera (SSCC) that provided full disk (daytime only) pictures of the earth's cloud cover on a 20-min basis. Film loops of these early pictures were used to provide winds daily at two levels over the Atlantic and Pacific Oceans. The SSCC and the Multicolor Spin Scan Cloud Cover Camera (MS SCC) on ATS-3 [*11*] were the forerunners of the operational Visible and Infrared Spin Scan Radiometer (VISSR) flown on SMS-1 (1974), SMS-2 (1975), and GOES (1975), and the Geosynchronous Very High Resolution Radiometer (GVHRR) on ATS-6 (1974). The latter two instruments permitted the first continuous day and night full disk cloud observation.

The VISSR scans from west to east, due to spacecraft rotation, with eight identical visible and two redundant infrared channels. The ground resolution at nadir of the visible (0.55 to 0.70 μm) and infrared (10.5 to 12.6 μm) channels are 0.8 and 8 km (0.43 and 4.3 NM), respectively. The SMS/GOES rotates at 100 rpm; the VISSR scan mirror views the earth for one twentieth of each 360 deg rotation The radiometer scans from north to south in 18.2 min and 1821 successive scan steps. Adding to this the time needed for ground data acquisition and processing, the final VISSR picture is available to users every 30 min. The satellite can also be programmed to scan in a limited mode for research and development studies over severe storm areas [*12.13*]. Figure 14 shows an example of a SMS-1 visible, full-disk picture with a sector of the United States and Hurricane Amy off its east coast on 30 June 1975 (1852-m resolution).

Atmospheric Sounders

The theoretical basis for the determination of the vertical profile of temperature in the atmosphere was suggested by Kaplan [*14*]. Basically, the concept utilizes measurements by a radiometer with channels sensitive to different spectral regions in and near an absorption band (CO_2, dioxide [O_2], and water [H_2O]) such that the weighting functions peak in different parts of the atmosphere. The measured radiances are then "inverted" mathematically in order to derive the temperature structure of the atmosphere. However, there is no unique solution to the problem. For this reason it is necessary to start with a "first guess" vertical temperature profile which is obtained either from a forecast or climatological profile. Various techniques, such as inverse matrices, regression methods, and iteration, are used to obtain a solution. The procedure generally involves knowledge of the transmission functions of the different channels. The transmission functions are derived by matching theoretical absorption-based models with experimental measurements of absorption and then applying the models to the atmosphere. Errors arise due to imperfect models and to extrapolation from

FIG. 14—SMS-1 visible image, 1852 m (1 NM) resolution, 30 June 1975, with sector of the United States, Caribbean, and Hurricane Amy off the U.S. eastern coast.

experimental measurements to different atmospheric pressures and temperatures. In some cases the transmission functions have been modified empirically by comparing measured radiances with computed radiances [*15*]. Given the transmission functions, the measured radiances are compared with radiances computed from the first guess profile, which are then modified by iteration until minimum differences between computed and measured radiances are obtained.

The evolution of atmospheric sounders is illustrated in Fig. 15. A Medium Resolution Infrared Radiometer (MRIR) was originally flown on Nimbus 2 (1966). It contained a 14 to 16-μm (carbon dioxide) radiometer which essentially measured stratospheric temperatures. Nimbus 3 carried a Satellite Infrared Spectrometer (SIRS) and an Infrared Interferometer Spectrometer (IRIS), both of which provided data allowing the first temperature soundings to be derived. Nimbus 4 carried improved versions of these instruments.

Another version of a sounder, the Infrared Temperature Profile Radiometer (ITPR) was flown on Nimbus 5. It had four channels in the CO_2 band between 13 and 15 μm, one channel near 20 μm for sensing atmospheric moisture and two "window" channels at 11 and 3.7 μm. The window channel measures the surface temperature minus a correction for atmospheric moisture. The purpose of the two window channels was to detect clouds. The 3.7-μm channel is much more sensitive to warmer temperatures than the 11-μm channel. Hence, if partial cloud cover is in the field of view, the 3.7-μm channel will measure a higher brightness temperature than the 11-μm channel.

A recent improved infrared sounder flown on Nimbus 6 was the High Resolution Infrared Radiation Sounder (HIRS) which contained 7 channels in the 15-μm CO_2 band, four channels in the 4.3-μm CO_2 band (also sensitive to nitrous oxide [N_2O]), two water vapor channels at 6.7 and 8.2-μm, and two window channels 3.7 and 11 μm. The spatial resolution of HIRS was about 25 km.

Nimbus 5 also carried a nadir-viewing Microwave Spectrometer (NEMS) which contained 3 channels in the 5-mm O_2 band, and two channels near 1 cm which are sensitive to water vapor and clouds over the ocean and to temperature and emissivity over land. Nimbus 6 contained a scanning version, the Scanning Microwave Spectrometer (SCAMS) which provides nearly full earth coverage every 12 h. The weighting functions in the 5-mm O_2 band peaked near 20 000 and 50 000 Nm^{-2} (200 and 500 mbar) and near the surface. The advantage of the microwave spectrometer is its ability to sound the atmospheric temperature profile through clouds. The combinations of NEMS, ITPR, SCAMS, and HIRS provide improved sounding capabilities as compared to the infrared sounders alone.

Nimbus 4 and 5 also contained the Selective Chopper Radiometer (SCR) with eight spectral channels in the 15-μm CO_2 band for measuring temperature profiles up to 50 km. An advanced version, the Pressure Modulator

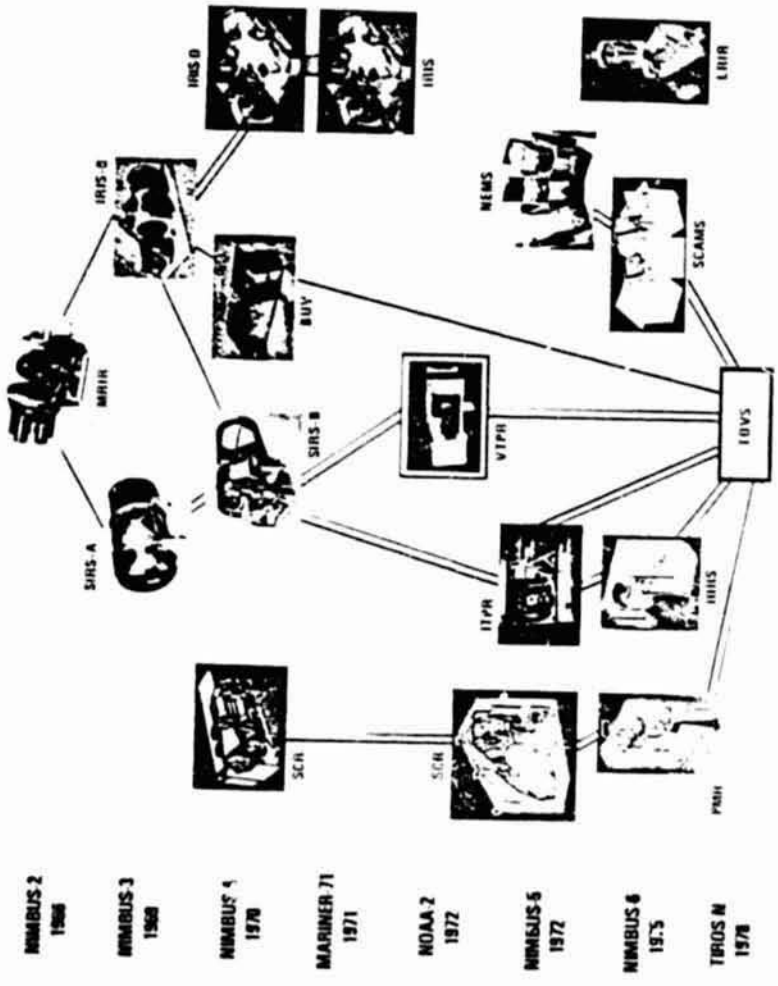


FIG. 15—*Evolution of meteorological atmospheric sounders from 1966 to 1978.*

Radiometer (PMR) was flown on Nimbus 6. SCR and PMR have provided some stratospheric temperature profiles on a research basis.

Figure 16 (*right*) shows the weighting functions of the HIRS and SCAMS, and (*left*) the average temperature difference between soundings derived from HIRS and raobs at 30 to 60 deg N. It can be seen that the combination of the 15 and 4.3-μm channels provides smaller temperature differences than the 15 or 4.3-μm channels alone. Between 80 000 and 30 000 Nm^{-2} (800 and 300 mbar) the root mean square temperature differences are between 1.5 and 2°C. Differences are greater near the surface (about 2.5°C) and near the tropopause (3.5°C). It should be pointed out that raobs are not necessarily ground truth, since two raobs obtained at the same site and time often have temperature differences of 1 to 2°C. Hence, in the range between 80 000 and 30 000 Nm^{-2} (800 to 300 mbar), the HIRS soundings may be considered as good as raobs. A very good agreement between HIRS derived temperature profile and raob was obtained on 29 June 1975 (Fig. 17). The inversion between 90 000 and 70 000 Nm^{-2} (900 and 700 mbar) was not detected by HIRS but all other features agree well. The importance of the 4.3-μm CO_2 channels for low level temperature determination is evident since without those channels the derived temperatures differed by as much as 10 K from the raob temperatures. Figure 18 shows a comparison of water vapor profiles derived from HIRS and raobs. Since there were only two water vapor channels, the shape of the profile had to be assumed for the HIRS derivation. However, the results showed good agreement and indicate that useful information can be obtained from the water vapor channels.

An operational atmospheric sounder, the Vertical Temperature Profile Radiometer (VTPR) has been flown on NOAA satellites since October 1972. It contained five 15-μm CO_2 channels, one rotational H_2O channel at 18.7 μm and a window channel at 12 μm. In order to obtain satisfactory results, it was found necessary to use radiosondes for "tuning" the solution of the regression coefficients. Then a greater density of soundings was used, which gave RMS differences of 2 to 3°C at 85 000 to 20 000 Nm^{-2} (850 to 200 mbar), respectively, from nearby radiosondes [*16*].

The Geosynchronous Operational Environmental Satellite VISSR Atmospheric Sounder (GOES VAS) to be flown on GOES D (1980), extends the SMS/GOES Visible and Infrared Spin-Scan Radiometer (VISSR) capability to include additional thermal channels for measurements of the vertical profiles of atmospheric temperature and water vapor. The VAS radiometric measurements consist of eight visible channels and two IR channels. The spatial resolution is 900 m for the band, 7 km for the 11-μm band, and 14 km in the sounding bands. The VAS can operate in the normal VISSR mode, a multispectral image scanning mode, and a dwell sounding mode. This latter mode provides the most accurate localized temperature soundings under the clear sky conditions in the vicinity of severe storms. A VAS demonstration experiment sponsored by The Na-

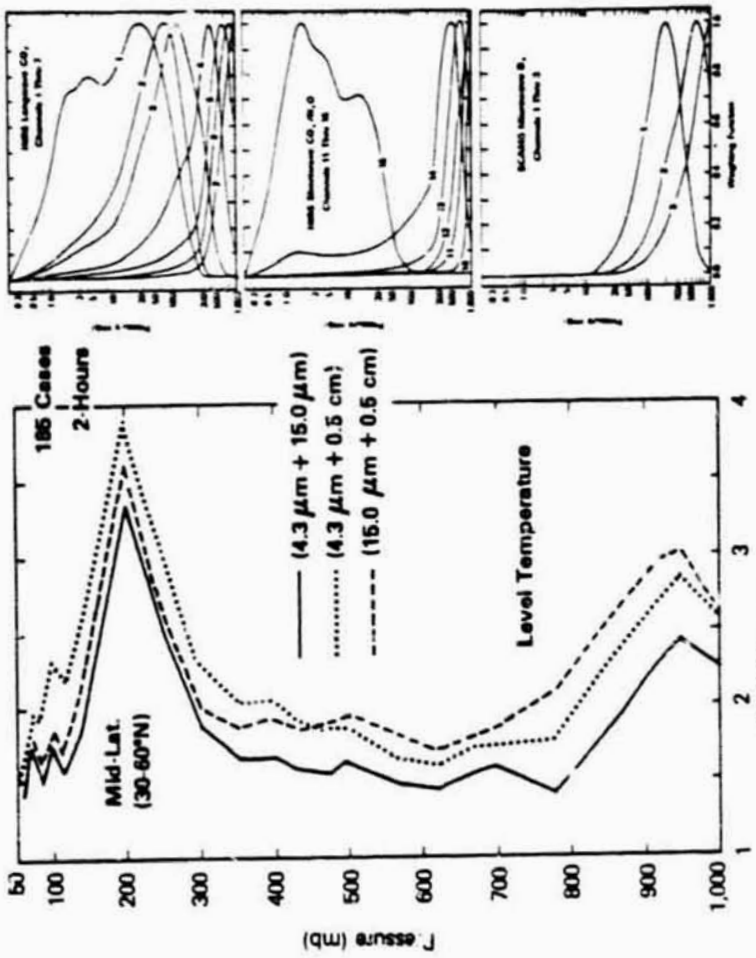


FIG. 16—(left) *Average temperature difference (RMS, °C) between Nimbus 6 HIRS-SCAMS and radiosondes taken within 2 h of satellite overflight.* (right) *Weighting functions for Nimbus 6 HIRS-SCAMS.*

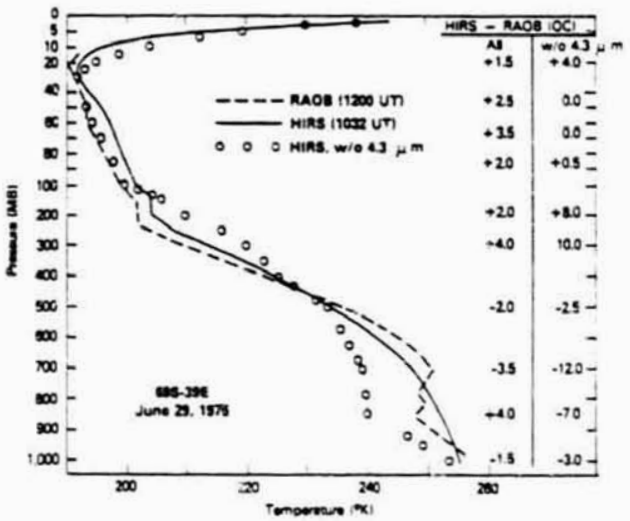


FIG. 17—*Comparison of Nimbus 6 HIRS and radiosonde temperature profile on 29 June 1975.*

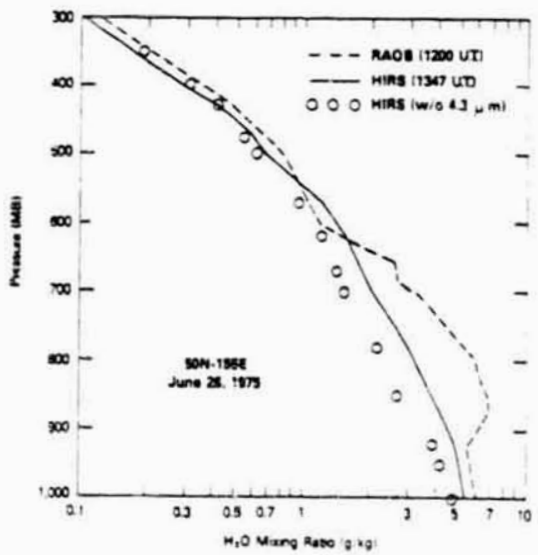


FIG. 18—*Comparison of Nimbus 6 HIRS and radiosonde water vapor profile on 26 June 1975.*

tional Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA) will examine the accuracy and utility of geosynchronous soundings at a prototype operational ground system at the University of Wisconsin and at a limited data facility at Goddard Space Flight Center (GSFC).

Some Recent Appli.ations of Environmental Satellite Data

Forecasting

The Scanning Radiometer (SR) replaced the AVCS television cameras used earlier on the ESSA, ITOS, and NOAA satellites and thus gave user agencies a vital nighttime coverage [*17*]. Figure 19 shows an overview of the present NOAA-NESS operational satellite data flow. Forecasters at 50 National Weather Service Field Offices use the satellite pictures to improve short-term forecasts and provide advisory services to the public and industry regarding aviation, agriculture, and shipping [*18*].

An interesting application of SMS/GOES infrared data is its use in a numerical model to forecast nighttime minimum ground temperatures for the citrus industry in central Florida. Figure 20 is a real-time surface temperature map of Florida (7.4 km [4-NM] resolution) on 10 Jan. 1976 which is one of a 30-min series used to track the cold air as it moved southward across the state. Based upon this movement, frost-freeze forecasts are issued 4 times daily to the citrus growers, who must decide whether or not to take protective action in their groves. An Application System Verification Test (ASVT) for this frost-freeze system will be run for three years to determine the economic value of the satellite system.[3]

Observations

A system entitled GOES-TAP provides the opportunity to receive SMS/GOES standard sector pictures every 30 min by television stations, local governments, Federal agencies, and universities. The SMS/GOES satellites are equipped with a Data Collection System (DCS) which relays environmental data sensed by surface platforms, for example, rain and river tide gages, ships, buoys, and automatic weather stations. Each spacecraft can receive data from 10 000 platforms every 6 h and relay the data to a central facility for processing.

Simultaneous pairs of visible images from SMS-1 and -2 satellites have been analyzed stereographically to yield cloud heights ranging from 1 to 18 km with a 2-sigma repeatability of 0.4 km. These measurements were made from approximately 50 deg N to 50 deg S, in a large overlap area

[3]Bartholic, J. F., private correspondence, 1976.

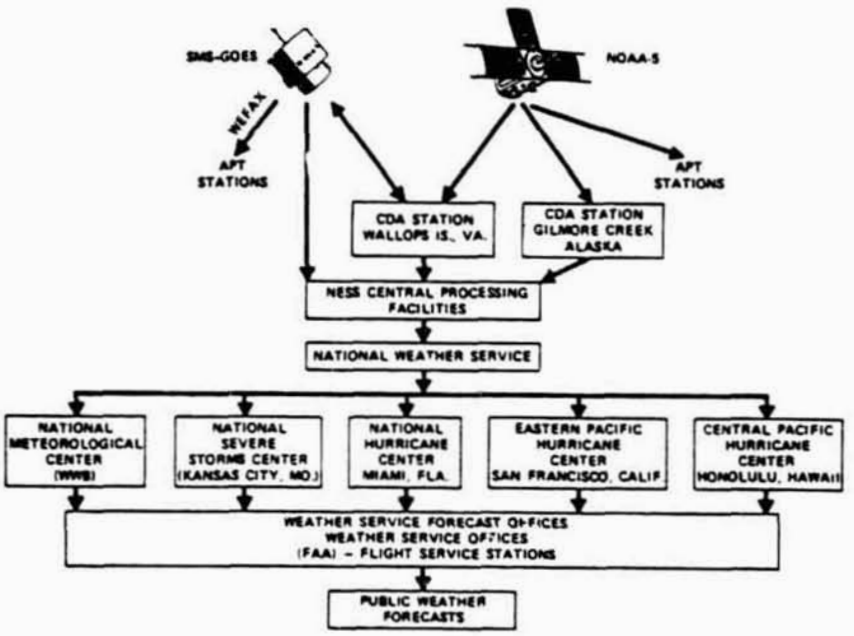


FIG. 19—*An overview of the present NOAA-NESS. Dept. of Commerce operational satellite data flow diagram, the end product of which is weather forecasts.*

80 deg of longitude wide, which is common to both satellites in the northern and southern hemisphere.

Figure 21 indicates cloud heights derived stereographically from SMS-1 and -2 on 17 Feb. 1975 over the Gulf of Mexico and Caribbean Sea [*19*].

A new high resolution ground station for receiving pictures direct from weather satellites was demonstrated recently at GSFC. The Local User Terminal (LUT), Fig. 22, is far superior to the APT system developed in the 1960s. It provides greater picture resolution from VHRR data by day, 0.9 versus 3.2 km; by night, 0.9 versus 7.2 km. The higher resolution provides meteorologists with better cloud and frontal system identification and delineates potentially severe storm clouds for early warning advisories.

After the launch of ATS-1 in 1966, it was recognized that winds could be determined by tracking clouds. Although individual clouds are affected by growth and dissipation processes, other clouds, and gravity waves, it was found that clouds generally move with the wind fields with accuracies approximating operational rawinsonde data [*20*].

In order to calculate winds from satellite imagery of clouds, careful geographical registration of succession photographs must be done. This is done by matching landmarks on the sides of the area in which cloud motions are measured. Then the displacements of clouds from successive pictures

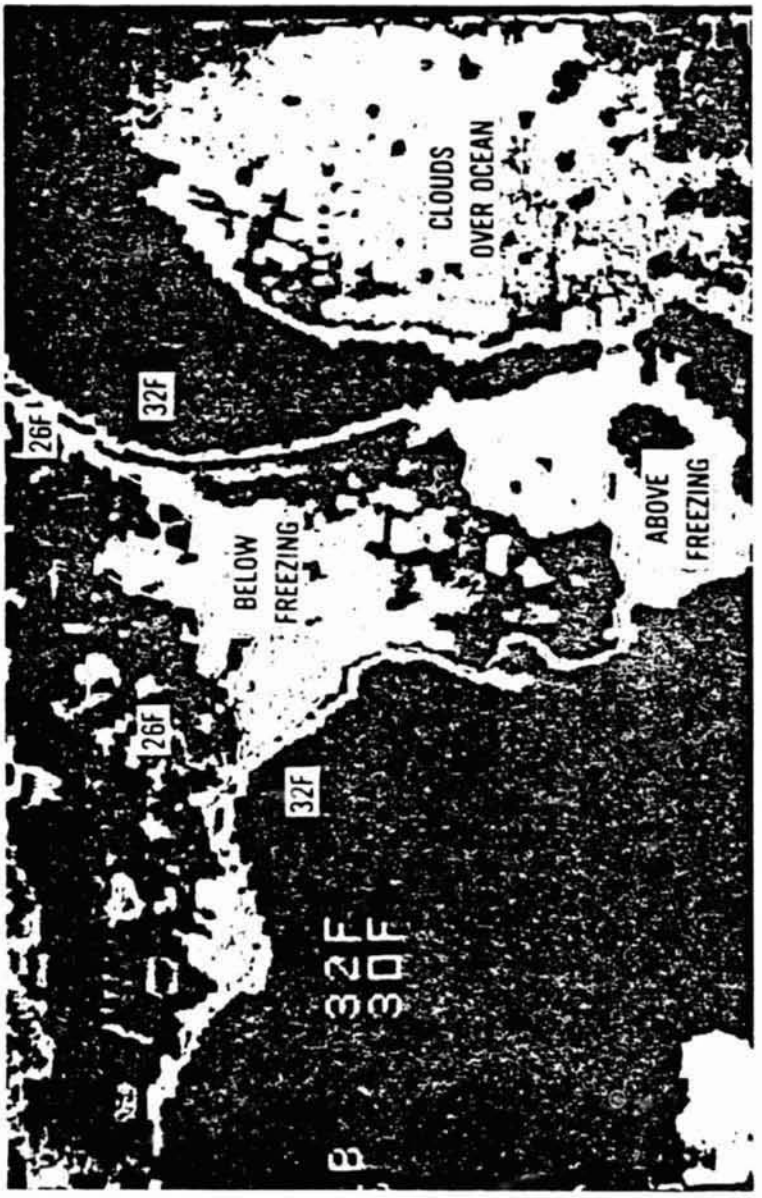


FIG. 20—*SMS-1. VISSR infrared surface and cloud (T_B) temperature map of Florida, on 10 Jan. 1976, 1852-m (4-NM) resolution.*

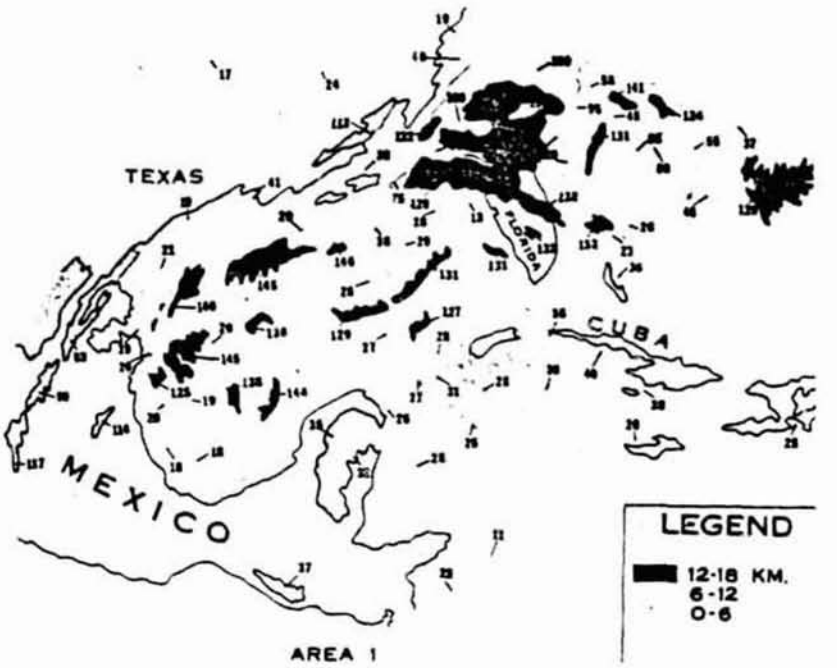


FIG. 21—*Cloud heights (km) derived stereographically from SMS-1 and -2 on 17 Feb. 1975 over the U.S., Gulf of Mexico, and Caribbean Sea.*

are measured and the winds calculated. In the SMS photographs, resolutions are 2 to 4 km (visible) and 8 km (infrared) so that often patches of clouds rather than individual clouds are tracked.

Operational determination of wind from SMS/GOES are made at 30-min intervals. The lifetimes of individual cumulus clouds are also of the order of ½h. Nevertheless, mesoscale patterns of cumulus clouds generally have lifetimes of hours, and the motion of these patterns are not greatly different from the rawinsonde measured air motion [*21*]. Measurements at shorter intervals of 15 or 7 min, or even less, have been made, and it has been found that the areal coverage has increased significantly. However, accuracy of measurement, especially for low wind speeds, is more critical at the lower time intervals.

A major problem is the assignment of heights to winds determined by cloud tracking. A skilled operator can distinguish between low, middle, and high clouds. Use of the infrared to determine cloud top temperatures may provide more accurate heights. For a thick cloud the emissivity of the top 100 m may be considered to have an emissivity of unity, so that heights can be determined by matching the measured equivalent black-body temperature with a temperature/height profile determined by radiosonde. However,

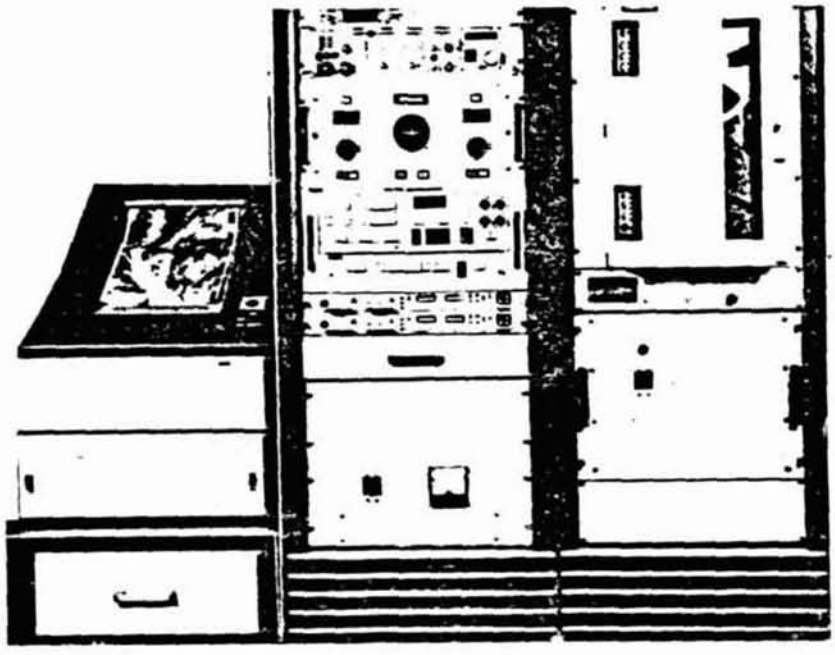


FIG. 22—*The Local User Terminal, developed by NASA for high resolution APT-type operation.*

some clouds, such as cirrus, have emissivities appreciably lower than unity. In addition, other clouds, such as cumulus, may not uniformly fill the field of view of the sensor. For these reasons, heights determined from infrared measurements may be subject to considerable error. Low level clouds over the oceans are assigned to the 90 000-Nm^{-2} (900-mbar) level which is statistically near the cloud base over oceans. Measurements by Hasler et al [*22*] indicate that low level clouds tend to move with the wind at the cloud base. The heights of middle and upper level clouds are determined by empirical methods using infrared measurements over the densest portions of the clouds, and by using estimated emissivities ranging upward from about 0.7.

Winds from cloud tracking are generally obtained by man-machine inter-active devices such as McIDAS at the University of Wisconsin, the operational MMIPS at NOAA/NESS, and AOIPS at GSFC (Fig. 23). The man is necessary in order to determine which clouds are being tracked as a single layer and carried with the wind rather than being formed by gravity waves. The man also distinguishes between low, middle, and upper clouds. The computer complex is necessary for speed and accuracy [*23*].

An example of low level winds of Tropical Storm Holly over the Atlantic obtained from 3-min GOES-1 pictures produced on the AOIPS is shown

ORIGINAL PAGE IS
OF POOR QUALITY

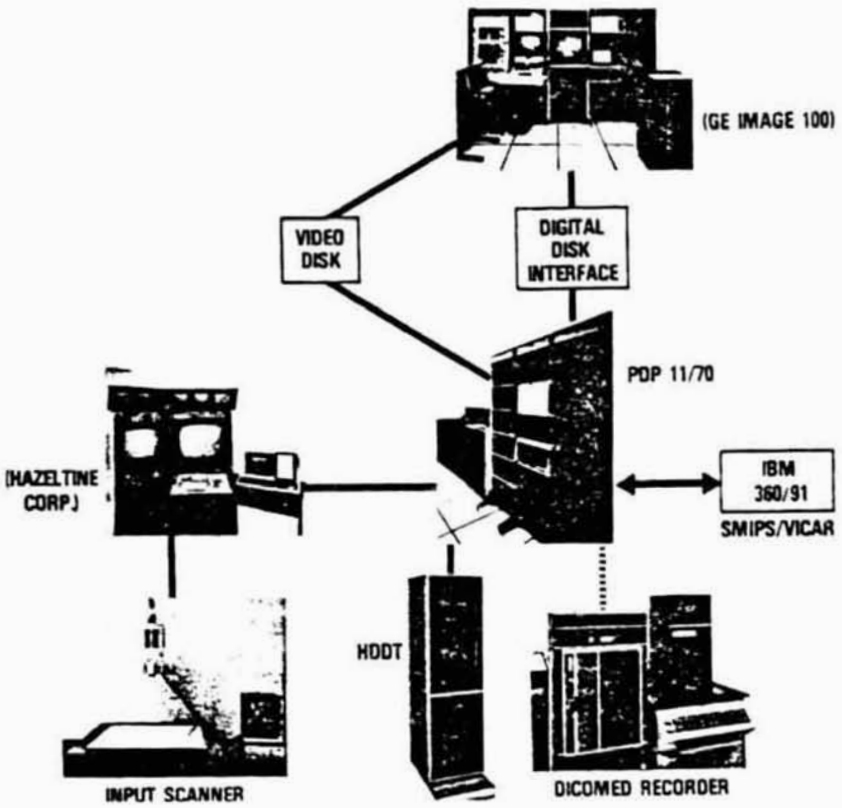


FIG. 23—*Components of the Atmospheric and Oceanographic Information Processing System (AOIPS) at Goddard Space Flight Center.*

in Fig. 24. The wind data are quite dense over the southern sector. The wind coverage for ½-h intervals would be only a small fraction of the number. The map shows an asymmetric flow pattern, with strongest wind speeds to the east and lightest speeds to the south of the storm center [*24*].

Another example of low level winds on a larger scale, Fig. 25 was produced by McIDAS during a Data Systems Test for FGGE. Prominent low level features are the anticyclonic circulations shown in the western North Atlantic and in the eastern South Pacific.

Sea surface temperature (SST) is determined presently from intake temperatures by ships of opportunity, drifting and moored buoys, and from satellites by infrared radiometry [*25*]. Ship and buoy SST measurements are accurate to about 1.5°C (RMS) when compared to oceanographic research vessel temperatures. Since 1970, NOAA meteorological satellites have carried the SR and where clouds were not in the field-of-view, horizontal

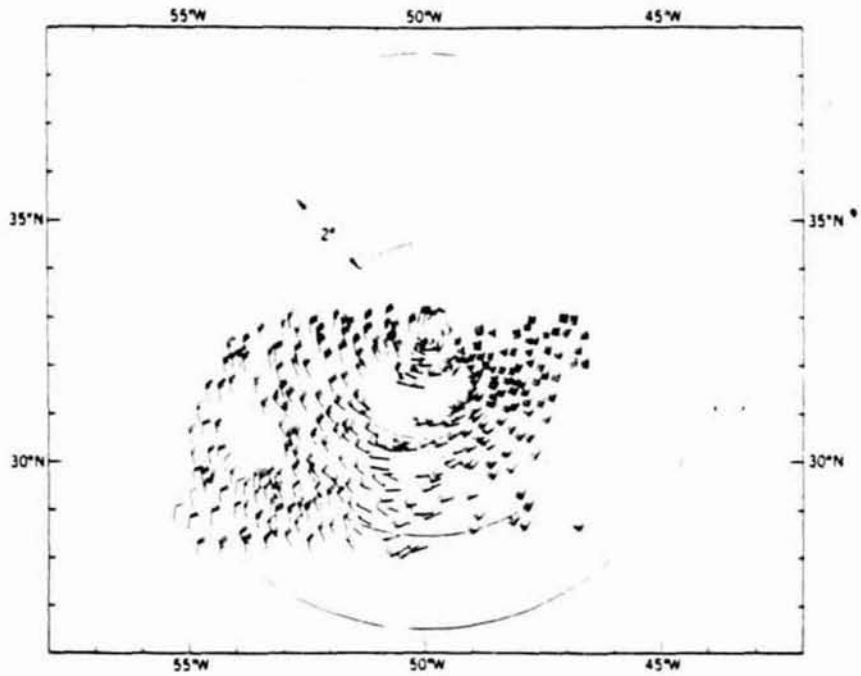


FIG. 24—*An example of low level winds derived from cloud motion on the AOIPS from 3-min visible GOES-1 pictures recorded over Tropical Storm Holly, 26 Oct. 1973, 1300 GMT.*

sea surface temperatures have been obtained. For a nominal case when the SR calibration temperature was 25°C and the ocean scene temperature was 27°C, the NEΔT is approximately 0.3°C. The NEΔT increases to 1.4°C for the same instrument temperature when the scene is very cold, that is, −88°C (cloud contaminated). A system noise can result in a NEΔT of 1.2 and 3.5°C for scene temperatures of 27 and −68°C, respectively. An example of a monthly NOAA-4 SST map of the globe is shown in Fig. 26 [*26*]. Corrections for water vapor absorption were applied to the SR-SST data from coincident VTPR soundings. An updated triweekly Gulf Stream analysis (Fig. 27) derived from the VHRR data has been used by the oil tanker industry to save on fuel shipping costs [*27-29*].

Visible, infrared, and microwave sensors are also capable of detecting ice over lakes and oceans. NOAA sensors (VHRR and VISSR) have been used to detect changes in ice cover over the Great Lakes and in the Arctic. In the Great Lakes, four scales in percent of ice cover are used to designate ice cover. Figure 28 indicates that the frozen area of the Great Lakes (black) in the winter of 1977 was larger than in 1975 [*30*].

LANDSAT visible imagery has been used to study ice floe morphology

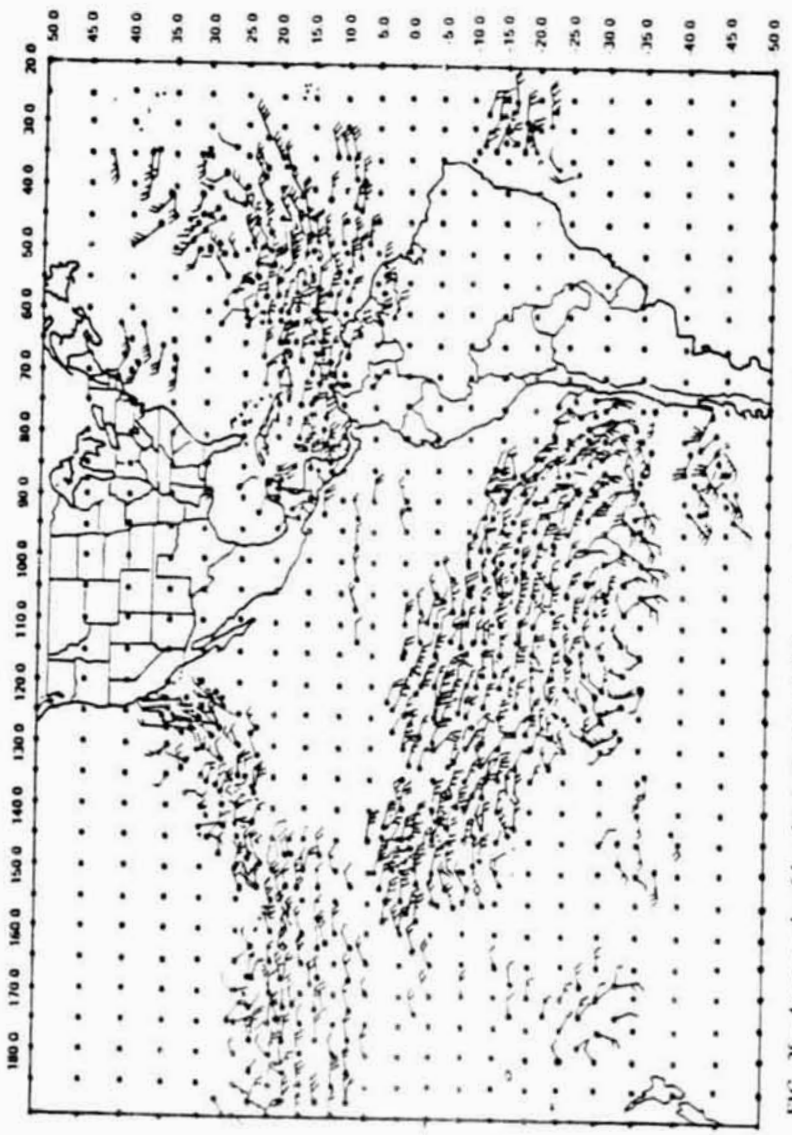


FIG. 25—*An example of low level winds derived from cloud motion on the McIDAS, at the University of Wisconsin, Madison, Wis., from visible SMS-1 and -2 pictures, 30-min data.*

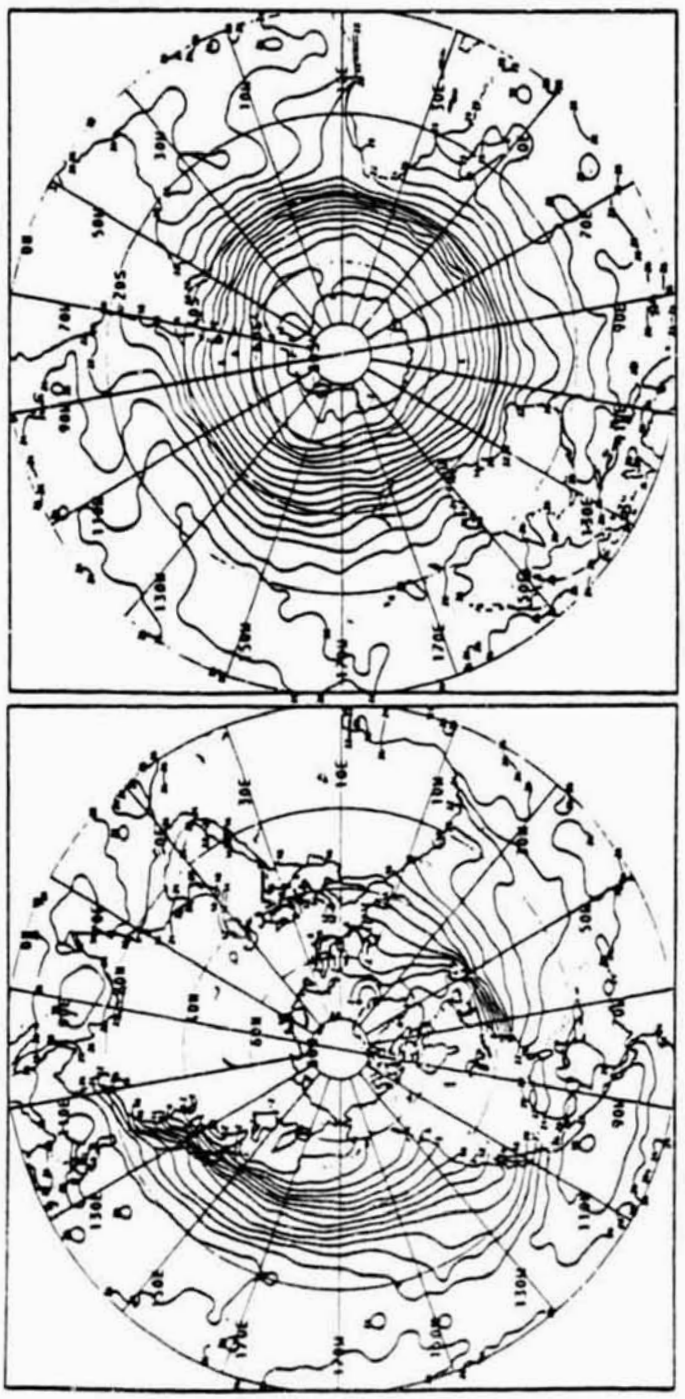


FIG. 26—*NOAA-4 global infrared SR mean monthly sea surface temperature charts for March 1975.*

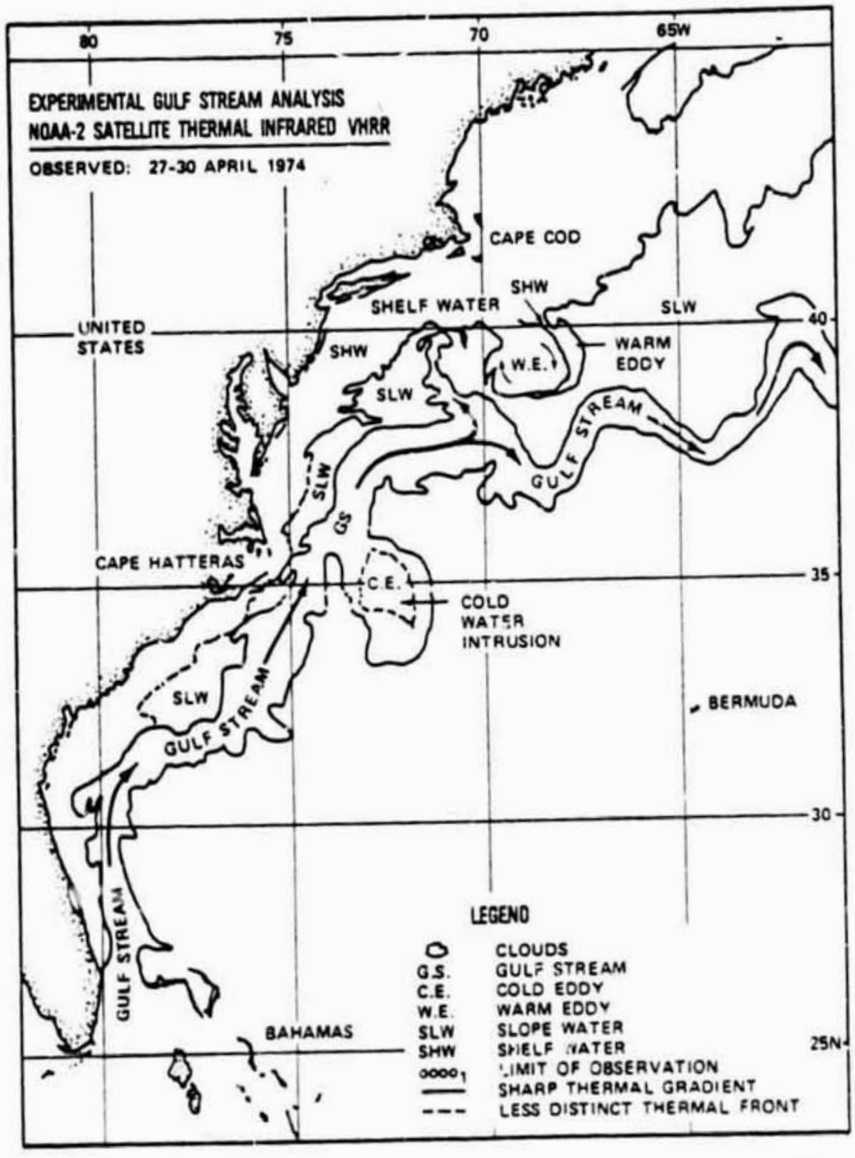


FIG. 27—*A simplified Gulf Stream infrared sea surface temperature analysis derived from NOAA-2 VHRR data from 27-30 April 1974.*

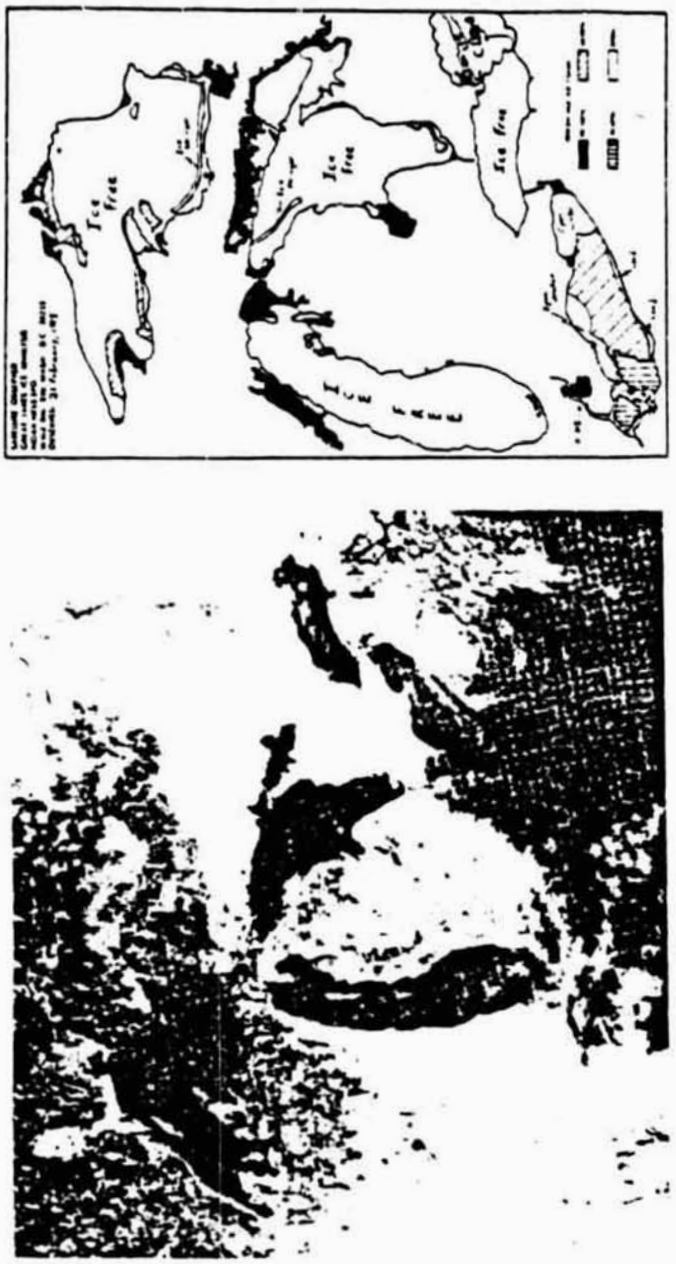


FIG. 28—*A NOAA-NESS Great Lakes ice analysis for 7 Feb. 1977 and 21 Feb. 1975 derived from VHRR and VISSR data.*

FIG. 28—*Continued.*

and dynamics in the Arctic on the scale of several days to months; lead and polynya dynamics of one to several days has also been studied. Nimbus 4 IRLS (Interrogation, Recording, and Location System) drifting buoys have permitted the mapping of gross ice drift in the Beaufort Sea [*31*].

With microwave imagery from ESMR-5 (19.35 GHz) and −6 (37 GHz), first and multiyear ice can be distinguished. The emissivity of first year ice is close to unity in the wavelength region of 0.3 to 11 cm. Multiyear ice has a lower emissivity, either because of the presence of empty brine pockets above sea level or the crystal structure of ice [*32*]. The brightness temperature (T_B) of Greenland glacier ice showed variations of 50°C, with the highest emittance corresponding roughly to the highest elevation. These variations in emissivity are related to the ice crystal size in the glacier snow cover [*33*-35].

The seasonal ice boundary in the Arctic and Antarctic can be detected easily by the sharply higher T_B of the ice as compared to the open water (Figs. 29 and 30). Figure 31 shows the percentage of ice in 5-deg latitude belts in the Antarctic taken from ESMR-5 imagery between July 1973 and October 1974. In February the ice cover is 95 percent at 75 to 80 deg S and 35 percent at 65 to 70 deg S. Maximum ice cover is during August and September, with 95 percent at 65 to 60 deg S and about 24 percent at 55 to 60 deg S. The year-to-year variation in a belt is only a few percent, as may be seen by comparing the curves for July to October of 1973 and 1974 [*36,37*].

In an analysis of Arctic ice using visible imagery from ESSA satellites from 1966 to 1974, Sanderson [*38*] showed that the total ice cover during the winter months has very little variation from year to year, although there may be wide variations regionally. This is because excess ice in some regions is counteracted by less ice in other regions. The same appears to be true for Antarctic ice [*39*], and for winter snow cover in the northern hemisphere [*40,41*].

Rainfall measurements and estimations have been made by conventional meteorological networks using ground, ship, and aircraft observations, rain gages, and weather radars. With the advent of the meteorological satellite new techniques have been developed to estimate average hourly, daily, and monthly rainfall from satellite cloud photography using cloud categories and brightness, cloud top temperatures, and monthly cloud nephanalyses [*42-44*]. The Nimbus 5 ESMR (19.35 GHz) was used to delineate rain areas and provide semiquantitative rainfall rates within oceanic tropical cyclones on a 12-h basis using a theoretical model for calibration [*45,46*]. Figure 32 shows the "dark" rain bands of Hurricane Fifi (1974) in the photofacsimile image, while the THIR image shows best the overall cloud cover and cloud height. A rainfall atlas depicting global oceanic rainfall rate based upon the Nimbus 5 ESMR brightness temperature was published recently. Figure 33 shows the first satellite-derived annual distribution of

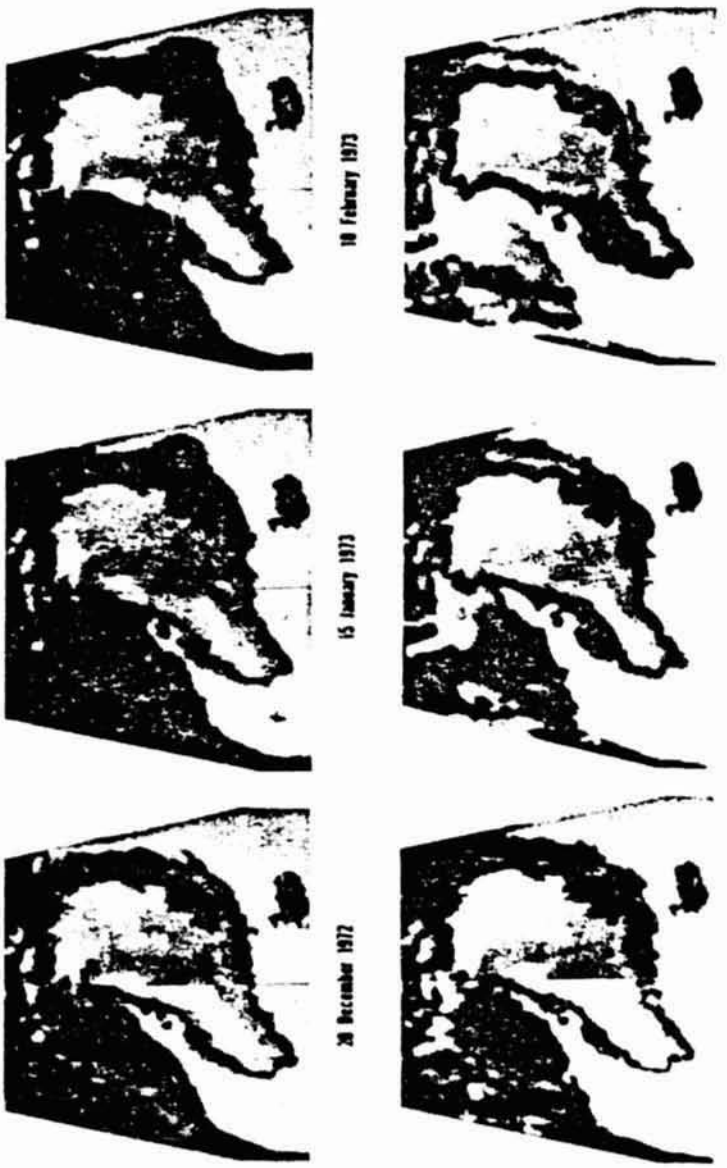


FIG. 29—*Seasonal ice boundary changes in the Greenland area from December 1972 to July 1973 as recorded by Nimbus 5 ESMR (19.35 GHz). Open water and old snow appears white (cold), while new ice and snow appear black (warm).*

FIG. 30.—Seasonal ice boundary changes in the Antarctic area from January to August 1973 as recorded by Nimbus 5 ESMR (19.35 GHz).

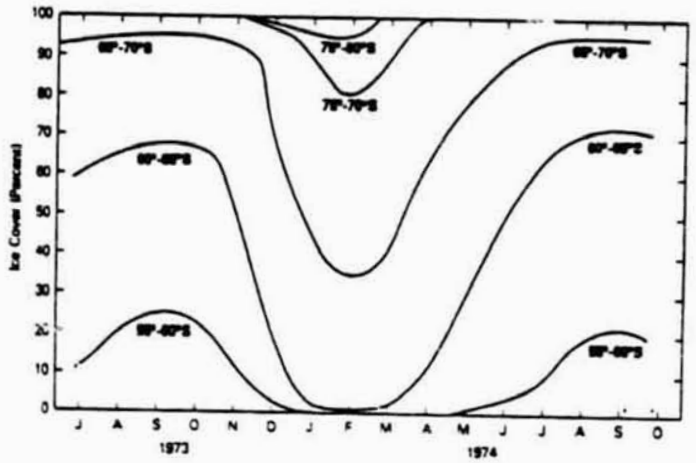


FIG. 31—*Monthly ice cover changes, in percent over the Antarctic from July 1973 to October 1974, as derived from Nimbus 5 ESMR (19.35 GHz) imagery.*

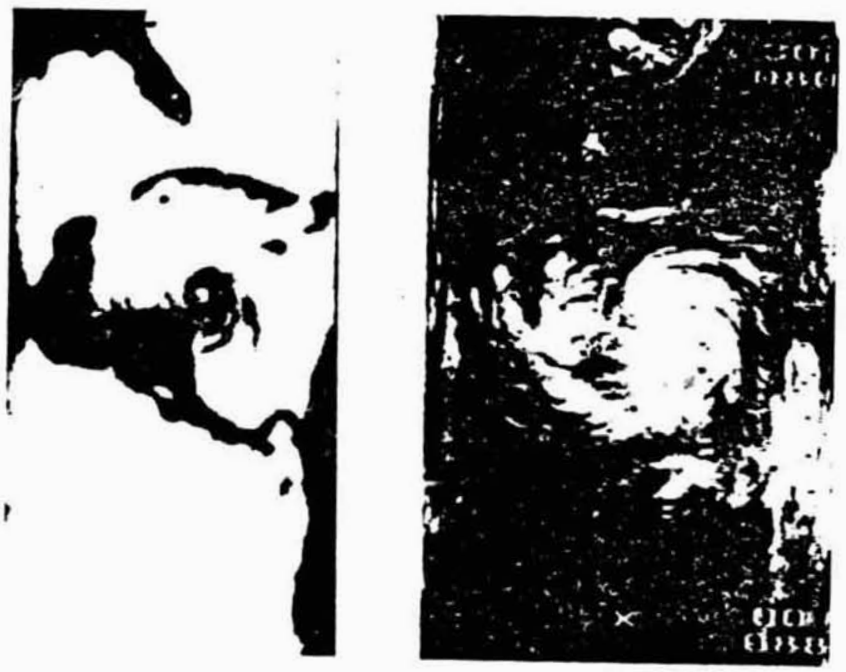


FIG. 32—*Nimbus 5 ESMR (19.35 GHz) and THIR (11 μm) facsimile pictures over Hurricane Fifi, on 18 Sept. 1974.* (left) *Electrically scanning microwave radiometer.* (right) *Temperature humidity infrared radiometer.*

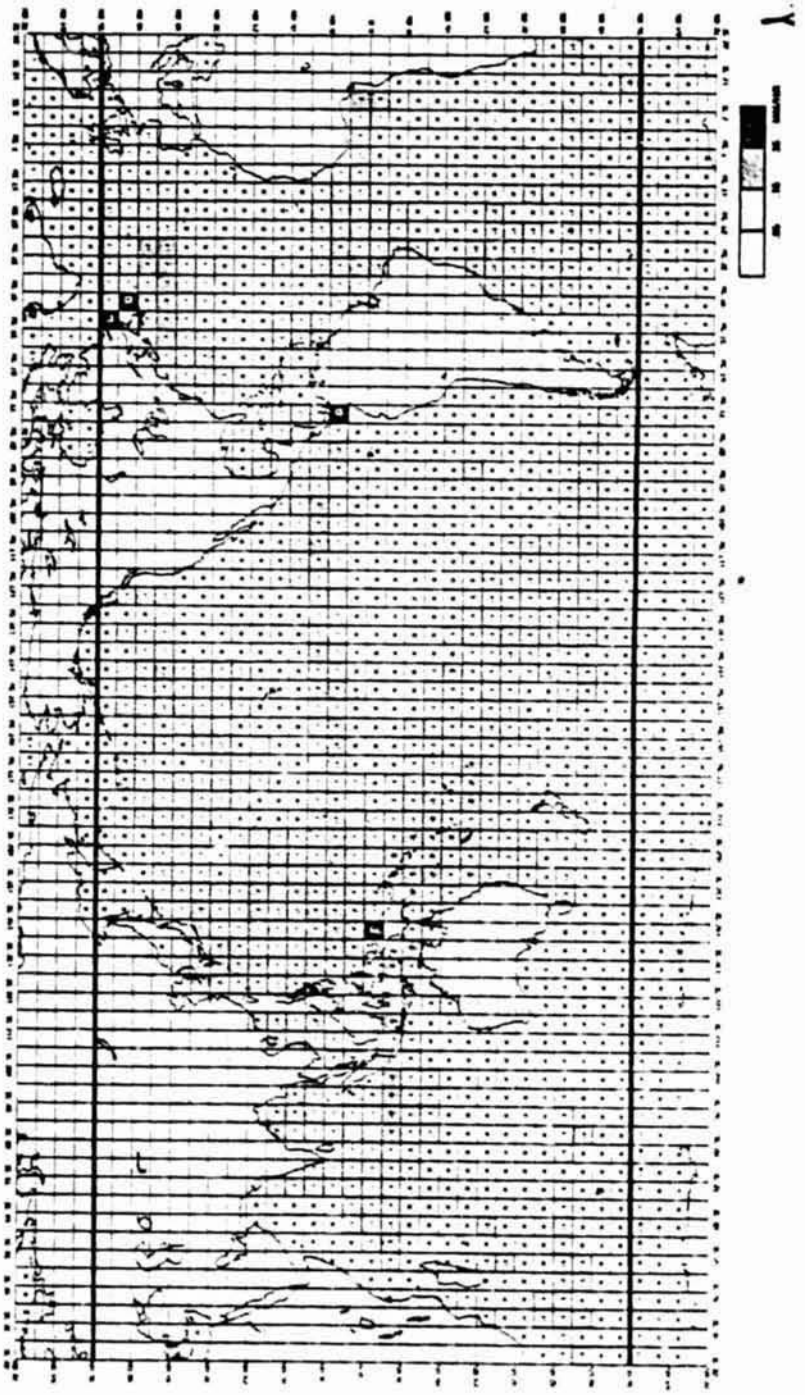


FIG. 33—*Global annual oceanic rainfall rate (averaged in mm/h) from Nimbus 5 ESMR (19.35 GHz) data from January 1974 to December 1974.*

rainfall over the oceans for January 1974 to December 1974 [47]. The weekly and monthly maps permit the monitoring of the movement of the Intertropical Zone of Convergence, the advance of the Indian Monsoon, and Pacific ocean rainfall associated with El Nino phenomenon.

Satellite microwave radiometers have an excellent potential capability to detect soil moisture. The T_B recorded by a microwave radiometer is equal to the product of the emissivity and temperature of the surface in the field of view. The emissivity of the soil is determined by the dielectric properties of the surface layer, a few tenths of a wavelength in thickness. The emissivity varies with the type of soil, percentage of moisture, and vegetative cover [48].

A photofacsimile microwave image, recorded by Nimbus 5 ESMR on 22 Jan. 1973 (Fig. 34) indicated a persistent soil moisture feature over the lower Mississippi Valley. Abnormally heavy precipitation had occurred in October, November, and December 1972, leaving the alluvial Mississippi Valley soils soaked with standing and subsurface water. Wet soils radiate as cold surfaces in the microwave imagery which shows as white in this figure. A grid print map analysis (*b*) shows two major areas of T_B values of < 220 K in the Mississippi Valley which lie within a large envelope where $T_B = 240$ K. This area overlaid the outwash aquifers in the drainage field [46].

Data from two radiometers, (21 and 2.2 cm) flown aboard Skylab were also compared with data obtained from ESMR-5 (1.55 cm) for a 300-km swath across north-central Texas in June 1973 [49]. Soil moisture in the top 2.5 cm varied from 15 to 70 percent of field soil capacity. The T_B range was 45 K for the 21-cm radiometer, 15 K for 2.2-cm radiometer, and only 5 K for 1.55-cm radiometer. The small response of the 1.55-cm instrument was ascribed to the presence of vegetation canopy which inhibits penetration for the shorter wavelengths.

Experimental flights over agricultural areas near Phoenix, Arizona, and Imperial Valley, California showed that the range in T_B between wet and dry ground for little or no vegetative cover is about 90 K for 21 cm and 60 K for 1.55 cm. An example of one flight is shown in Fig. 35 where the 21-cm radiometer indicated T_B of 206 to 210 K for 35 percent field soil capacity while dry desert (3 percent field capacity) showed a T_B of 272 K. The slope of the T_B curve was greatest beyond 25 percent field capacity [50]. The 1.55-cm radiometer was found to be incapable of detecting soil moisture variation through a plant canopy. The 21-cm radiometer showed only slight brightness differences over vegetated and bare fields; hence, it was found to be more suitable for observing a wider spectrum of soil moisture than radiometers using shorter wavelengths [51].

Landsat 1 (1972) data, (80-m resolution) was also found to be very useful as a regional tool for flood mapping during the Spring 1973 Mississippi River disaster. Particularly in the 0.8 to 1.1 μm (Band 7) spectral region of the Multispectral Scanner Subsystem (MSS), water or moist ground surface,

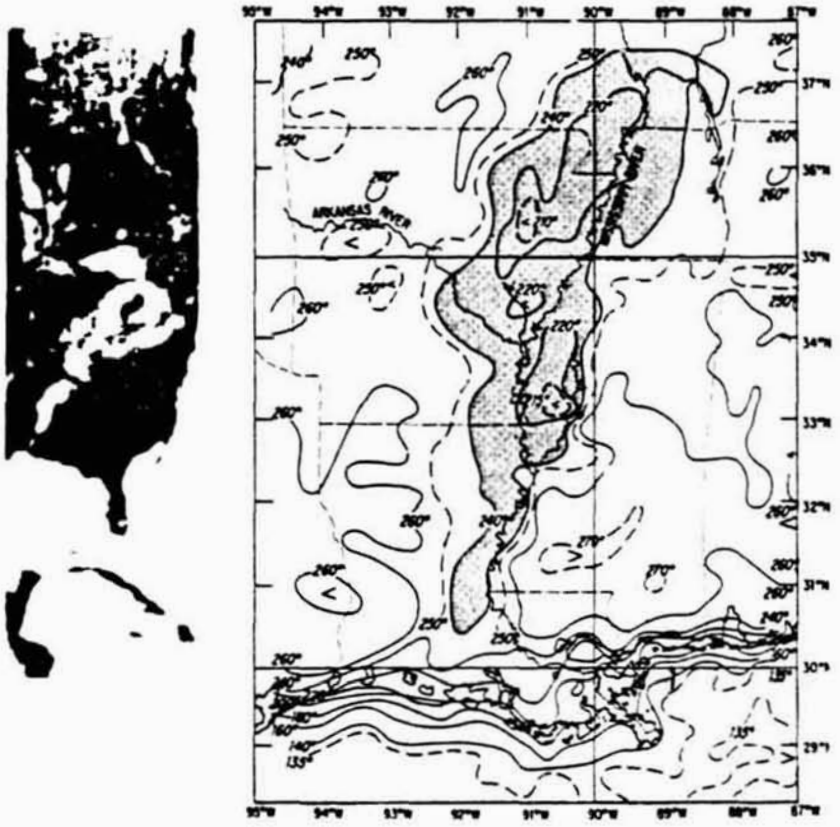


FIG. 34—*Nimbus 5 ESMR (1.55 cm) facsimile picture over the eastern U.S. on 22 Jan. 1973 and a grid print map analysis* (T_B) *in °K over the Mississippi Valley and Gulf coast, indicating high soil moisture (cold* T_B, *<240 K), in grey tone.*

or both, show up much darker than dry soil or vegetated areas [*52*]. Figure 36 shows the preflood stage of the Mississippi River in the St. Louis area. Areas C and B indicate the confluence of the Missouri and the Mississippi and the Illinois and Mississippi River, respectively, and D indicates areas of significant flooding. The stream flow in this area was the highest since 1844 and large areas were inundated [*53*]. Similarly, Band 7 was used to show the western Negev (vegetated-dark) and nonvegetated (light) Sinai desert area on the Israel-Egypt border (Fig. 37). A fence runs along this demarcation line, and the bleakness of the ground south and west of the fence is due to over-grazing by camel, sheep, and goats, denuding the existing vegetation [*54–57*].

LANDSAT C (1978) which was launched 3/5/78, carried a fifth channel, a thermal infrared band in the MSS which will have a lesser sensor spatial

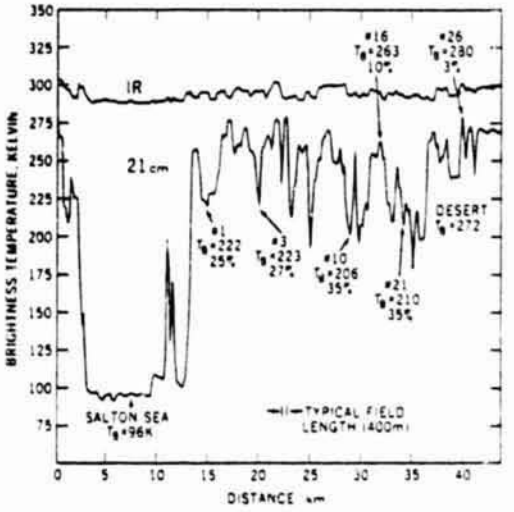


FIG. 35—*Aircraft thermal infrared (10 to 12 μm) and microwave (21 cm) brightness temperature versus flight distance at th north end of the Imperial Valley, Calif. The 21-cm brightness temperature and measured soil moisture for several vegetated fields the Salton Sea and desert land are indicated.*

resolution than LANDSAT A and B, namely, 200 m. This channel will improve the capability to classify crops, vegetation, and soils.

A Heat Capacity Mapping Mission (HCMM) is planned by GSFC to map the northern hemisphere surface temperatures and derive thermal inertial images near hours of maximum heating (1:30 p.m.) and cooling (2:30 a.m.), with 500-m resolution. It is expected that soil moisture patterns, soil and rock compositions, and geothermal sources can be distinguished better with this satellite than with LANDSAT [*58*].

A Thematic Mapper, LANDSAT D (1981), is under study at GSFC with plans for a 18-day repetitive coverage with a 30-m ground resolution for visible channels and 120 m for infrared. Four bands will most likely be centered at 0.48 μm (blue), 0.56 μm (green), 0.66 μm (red) and 0.83 μm (reflective infrared). Two additional bands, near 1.65 and 11.5 μm, will also be added. The Thematic Mapper will be especially useful for improved monitoring and prediction of the agricultural food resources on a global scale. The system feasibility has been demonstrated by LANDSAT 1 in the Large Area Crop Inventory Experiment (LACIE) [*59*].

Identification of snow cover from satellite generally has been made from visible imagery, and to some extent from infrared imagery. Microwave detection of snow has also been used in some areas. The primary snow mapping sensors on the NOAA satellites is the VHRR (Very High Resolution Radiometer) with one channel in the visible (0.6 to 0.7 μm) and another in

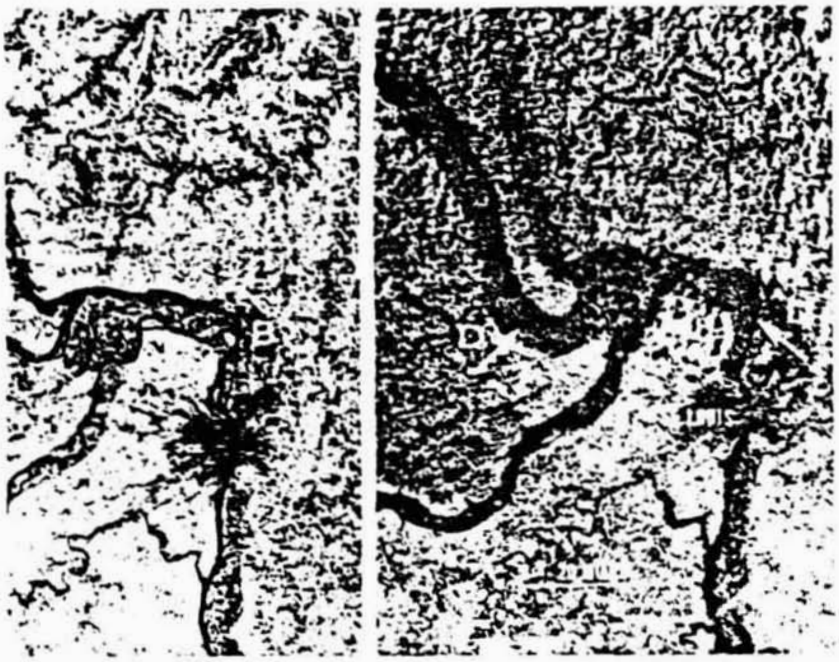


FIG. 36—*Pre-flood (2 Oct. 1972) and flood (31 March 1973) stage for the Mississippi River shown by Landsat MSS. Band 7 (0.8 to 1.1 μm). Point A indicates St. Louis, Mo., Points C and B indicate the confluence of the Missouri and the Mississippi Rivers and the Illinois and Mississippi Rivers respectively. Point D indicates areas of significant flooding.*

the window infrared (10.5 to 12.5 μm), and the Scanning Radiom ter (SR) with similar sensors but poorer resolution [*60, 61*].

Snow has about the same high albedo as many clouds. However, snow can be distinguished from clouds by pattern recognition and stability, since clouds do not retain the same shape for more than an hour. Cloud shadows on the ground can also be recognized at low sun angle.

Wiesnet and Matson [*40*] have prepared northern hemisphere average snow charts for the period 1966 to 1975, during the months September through March. These charts were prepared from photo interpretation of satellite imagery obtained from ESSA, ITOS, NOAA, and SMS-1 satellites.

An example of snow detection in the visible using LANDSAT imagery (80-m resolution) is shown in Fig. 38 for the Sierra Nevada Mountains in California. The winter of 1975 was a normal snow year as compared to 1977, a drought year.

Since LANDSAT and NOAA satellites have demonstrated ability to accurately measure snow-covered areas on various size watersheds, an Application Systems Verification Test (ASVT) Program was initiated to test the results of these studies in Arizona, California, Colorado, and

FIG. 37—*Western Negev, (vegetated-dark) and Sinai desert non-vegetated* (light) *on the Israel-Egypt border as shown by LANDSAT 1, MSS, Band 7 picture, 24 Aug. 1973 (80-m resolution).*

Oregon [*41*]. In addition, seasonal stream flow estimations can be made for flood and watershed conditions [*62*].

Similar wintertime LANDSAT imagery (Fig. 39) over Lake Michigan on 24 Nov. 1972 shows smoke plumes from seven steel mills and power-plants in the Chicago, Illinois-Gary, Indiana area. These plumes line up directly with stratoform clouds over the lake, and the location of snow on the eastern lakeshore. This picture is an observable example of weather modification due to man's industrial activities [*63,64*].

Snow detection using the window region of the infrared (10 to 12 μm) is possible because of the lower T_g associated with snow cover. However, in

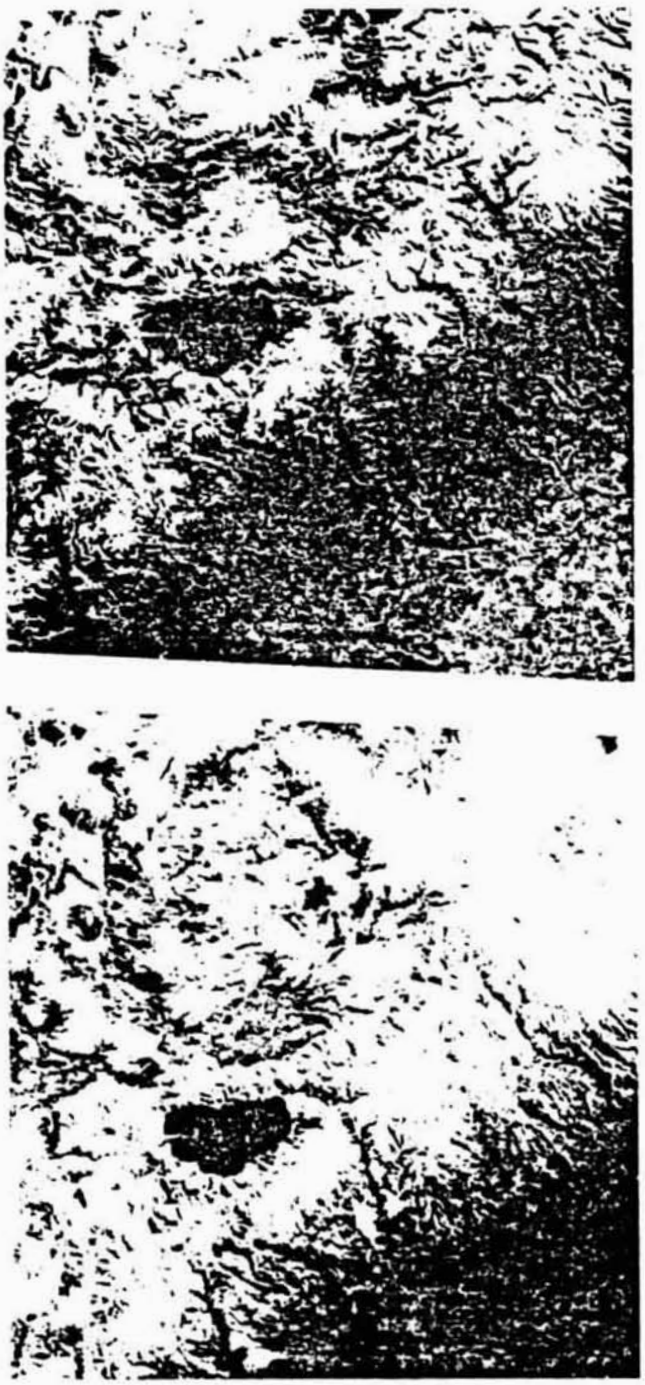


FIG. 38—*Snow cover difference in the Sierra Nevada Mts., near Lake Tahoe, Calif. (left) 25 Feb. 1975 (normal year), and (right) 14 Feb. 1977 (drought year) LANDSAT 2 MSS, Band 5, (0.6 to 0.7 μm) pictures.*

FIG. 39—*An example of weather modification due to man's industrial activities shown in a LANDSAT I. (E.R.T.S 5) picture recorded on 24 Nov. 1972 over Lake Michigan. SW to NE wind arrows and barbs point to smoke stack plume trajectory over the lake which leads to cloud sheet and snow fall development further downwind. Air temperature visibility and dew point numbers are plotted to left of weather station circle.*

mountainous terrain care must be taken when interpreting temperature differences which depend on elevation.

Microwave has the advantage of detecting snow cover through clouds. The emissivity of snow is markedly different from that of bare ground. However,

emissivity also varies with the age of the snow, depth, and the amount of prior melting and freezing. A comparison of snow cover depths and microwave imagery from ESMR-6 (37 GHz) over Minnesota in December 1973 showed that the 230 K T_B isotherm was a good delineation of the snow boundary [*51*]. Airborne microwave observations of snow indicate that it may be possible to distinguish dry and wet snow due by their different dielectric properties [*34, 49*].

The Earth Radiation Budget (ERB) experiment flown on Nimbus 6 (1975), made measurements of the reflected shortwave and emitted longwave radiation from the earth and the solar constant. Radiation budget calculations were made over the entire earth disk (3500-km radius) from 1100-km satellite altitude, with wide-angle channels (130 deg field of view), while narrow-angle channels (0.25 by 5 deg) measured regional values and the angular distribution of outgoing radiation. The ERB solar channels indicated a 1394 w/m^2 value for the solar constant while rocket measurements taken on 29 June 1976 indicate a lower value, that is, 1367 w/m^2 [*66*]. The planetary global albedo, longwave radiation fluxes, and net radiation were computed to be 30 percent, 240 w/m^2, and -4 w/m^2, respectively, for the months of July and August 1975, and thus show good agreement with prior Nimbus 3 calculations [*67*]. Albedo charts for August 1975 prepared from wide angle data (0.2 to 3.8 μm) show maxima over northern tropical South America, the bright sands of the Sahara Desert, and the Tibetan Plateau, while minima were found in the oceanic subtropical high-pressure ridges. The narrow-angle measurements showed that Greenland was a highly specular reflector of shortwave radiation, but a diffuse emitter of longwave radiation, whereas the Sahara Desert displayed the opposite characteristic. The variation in global albedo with latitude for different months is shown in Fig. 40. Gruber [*68*] (1976) used SR data (0.5 to 0.7 μm and 10.5 to 12.0 μm-channels) to compute net radiation balance profiles from June 1974 to May 1975 (Fig. 41, *left*). There was a net radiation loss in the winter hemisphere and a gain in the summer hemisphere. The net gain in the southern hemisphere summer is larger than the northern hemisphere summer. Mean annual meridional transport of energy computed by Gruber (Fig. 41, *right*) agreed well with others, except that northern hemisphere transport was appreciably greater in the southern hemisphere.

Measurements of total atmospheric ozone content were made in the last 40 years by a poorly spaced, quasi-global network of ground-based Dobson spectrometers, balloons, and sounding rockets. Many of the more recent satellite ozone measurements were made by Nimbus-3 and -4 IRIS [*69, 70*] and Nimbus-4 Backscatter Ultraviolet instrument (BUV) [*71*]. The BUV, a double (cascaded) Ebert-Fastie spectrophotometer which was launched in April 1970, is still providing daily ozone data after seven years of operation. A similar instrument, launched on Atmosphere Explorer 55 in late

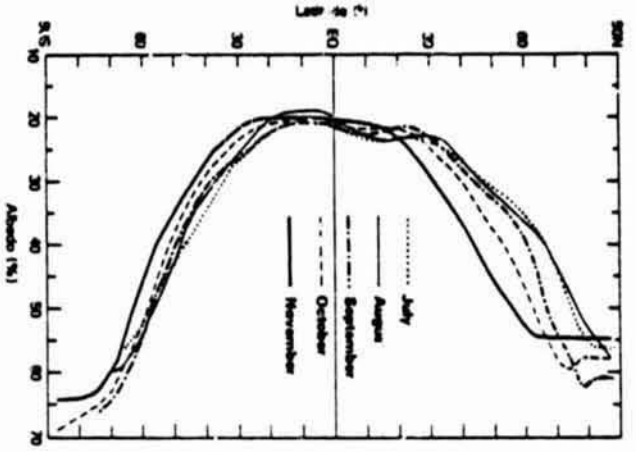


FIG. 40—*Variation in global albedo as measured by Nimbus 6 ERB from July through November 1975, using solar and wide angle channels.*

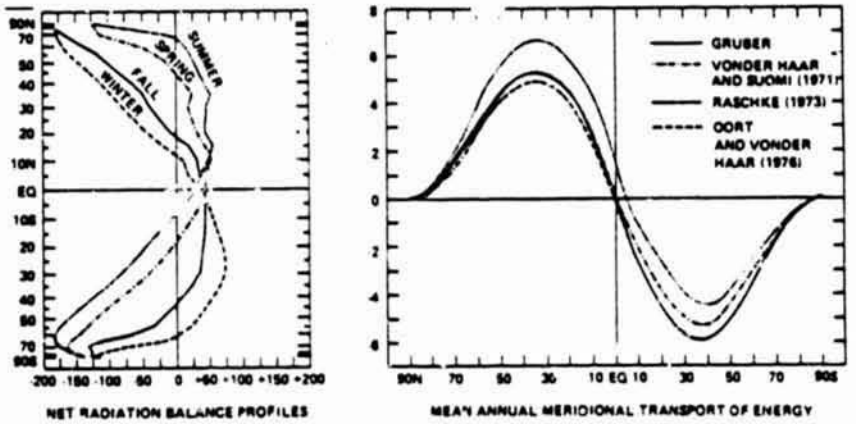


FIG. 41—(a) *Computed global net radiation balance using 2-channel NOAA-4 SR data.* (b) *Comparison of global mean annual meridional transport of energy by different satellite techniques.*

1975, is providing ozone data in an equatorial orbit over the region of the most active ozone formation. The BUV determines the total ozone content and vertical ozone distribution by measuring the backscattered ultraviolet solar energy at the satellite nadir at 12 wavelengths from 2550 to 3400 Å in the ozone absorption band (Fig. 42). The shorter wavelengths are more sensitive to ozone high in the stratosphere, and the longer wavelengths are most sensitive to ozone low in the stratosphere. To infer the ozone profile, the atmospheric albedo must be determined to within one percent.

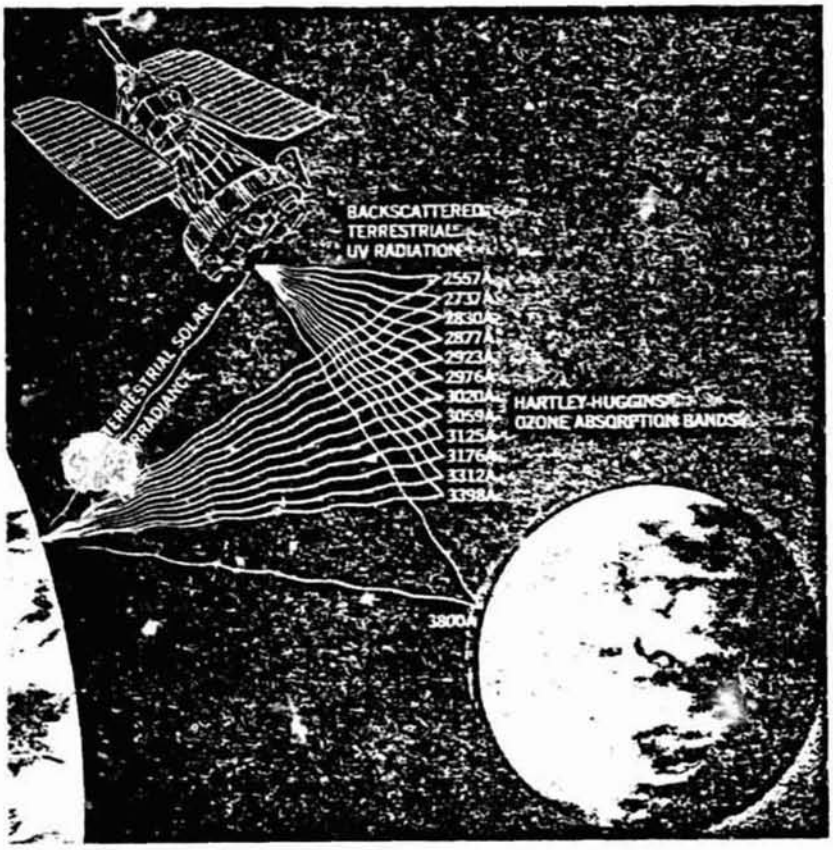


FIG. 42—*Schematic of Nimbus 4 Backscatter Ultraviolet (BUV) instrument which determines total ozone content and vertical ozone distribution by measuring the backscattered ultraviolet solar energy at the satellite nadir at 12 wavelengths from 2557 to 3398* Å *in the ozone absorption band.*

Figure 43 shows two periods (April to May 1970) of global ozone data processed from Nimbus 4 BUV data. A full-scale effort to produce an atlas of the seven years of BUV ozone data is now underway at Goddard. The Solar Backscatter Ultraviolet and Total Ozone Mapping Spectrometer (SBUV/TOMS) experiment to be flown on Nimbus G in 1978 will measure the ozone vertical profile and solar UV spectrum, and will provide a total ozone map by means of a mechanical scan across the Nimbus track. This instrument is a refinement of the BUV system with special provisions for minimizing the effects of space radiation [72].

The Limb Radiance Inversion Radiometer (LRIR) flown on Nimbus 6 (1975) also made measurements of the vertical distribution of ozone, temperature, and water vapor from 15 to 60 km on a global scale [73]. Vertical distributions were determined by inverting measured limb radiance

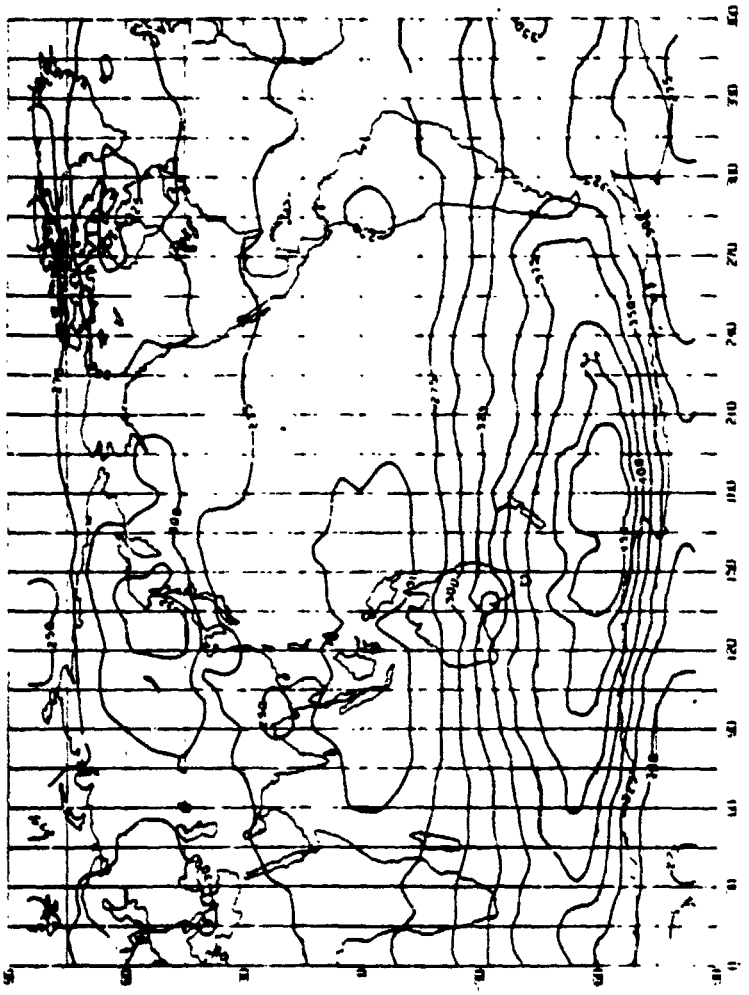


FIG. 43a—*Global total ozone content map (units in m atm-cm), derived from Nimbus-4 BUV data, for the month around the autumnal equinox (21 Sept. 1970) (from the Ozone Processing Team, NASA/GSFC).*

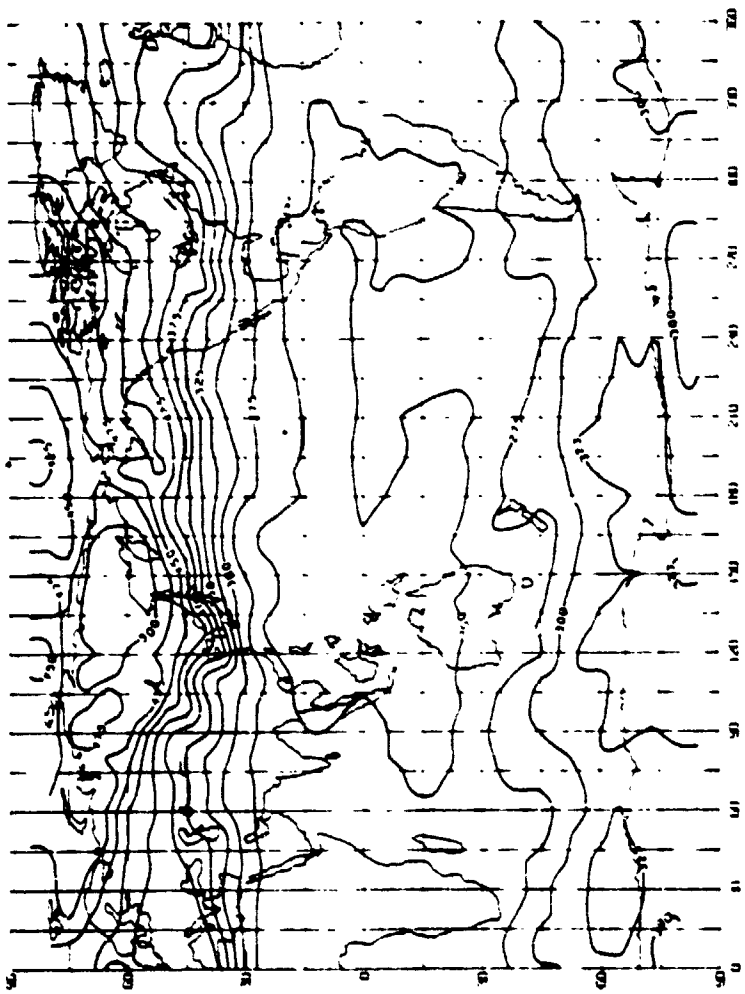


FIG. 43b—*Global total ozone content map (units in m atm-cm), derived from Nimbus-4 data, for the month around the vernal equinox (21 March 1971) (from the Ozone Processing Team, NASA/GSFC).*

profiles from the LRIR, an infrared multispectral scanning radiometer in four spectral regions: two in the 15-μm CO_2 band, one in the 9.6-μm ozone band, and one in the 23 to 2-μm rotational water vapor band.

Future Environmental Satellite Programs

The Global Atmospheric Research Program (GARP) was established in 1967 by the World Meteorological Organization (WMO) and the International Council of Scientific Unions (ICSU) in response to resolutions of the United Nations. It is an international cooperative effort to increase the accuracy of medium range weather forecasting (from 1 day to 2 weeks), to develop mathematical models of the atmosphere which will permit reliable long-range weather forecasts, to guide the design of a cost-effective global observing and forecasting system for routine use by the international community, and to investigate the mechanisms underlying climate variations which may lead eventually to some level of climate prediction. The First GARP Global Experiment (FGGE) will have 145 countries participating, and global observations will be taken from 1 Sept. 1978 to 31 Aug. 1979 with intensive observing periods during January and February and during May and June 1979 [*74*]. The system will utilize the following:

1. World weather watch surface-based stations (global).
2. 4 Polar orbiting weather satellites (global).
3. 5 Geostationary weather satellites (50 deg N to 50 deg S).
4. Special and commercial aircraft releasing dropsondes (tropics).
5. 50 ships releasing radiosondes (tropics).
6. 300 ocean drifting buoys (southern hemisphere).
7. Commercial ships equipped with weather instruments (southern hemisphere).

Figure 44 shows the full complement of weather satellites to be in orbit by 1978-1979 for this experiment, while Fig. 45 illustrates schematically the global data flow and planned utilization for the FGGE program.[4] NASA has already conducted six Data System Tests (DST) to date to address problems in data management and utilization, and to demonstrate the full potential of satellite temperature sounding data [*75*].

A Climate Research Program, now under study at GSFC will make full use of the TIROS-N, GOES-METEOSAT, and proposed CLIMSAT and ERBS-A satellites shown in Fig. 46. With this system, most of the climate-related observational requirements now anticipated can be met [*76*]. In those areas where the requirements cannot be met by direct observations

[4]Greaves, J., private correspondence, 1977.

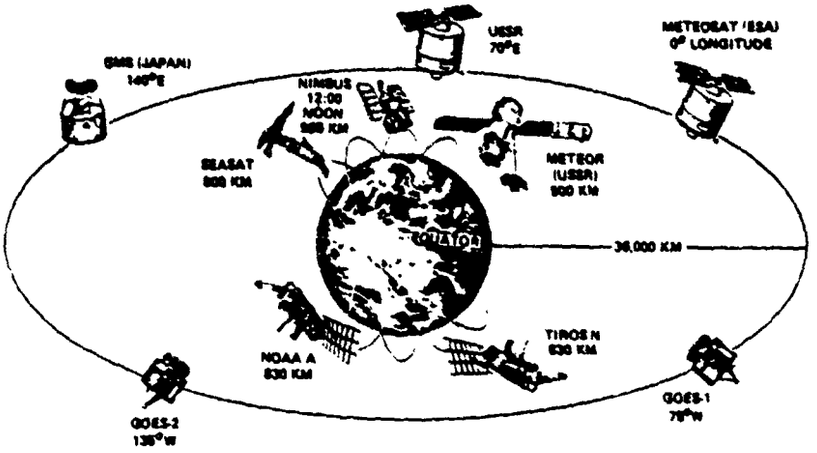


FIG. 44—*Weather satellites to be in orbit during the First GARP Global Experiment (FGGE) 1978-1979.*

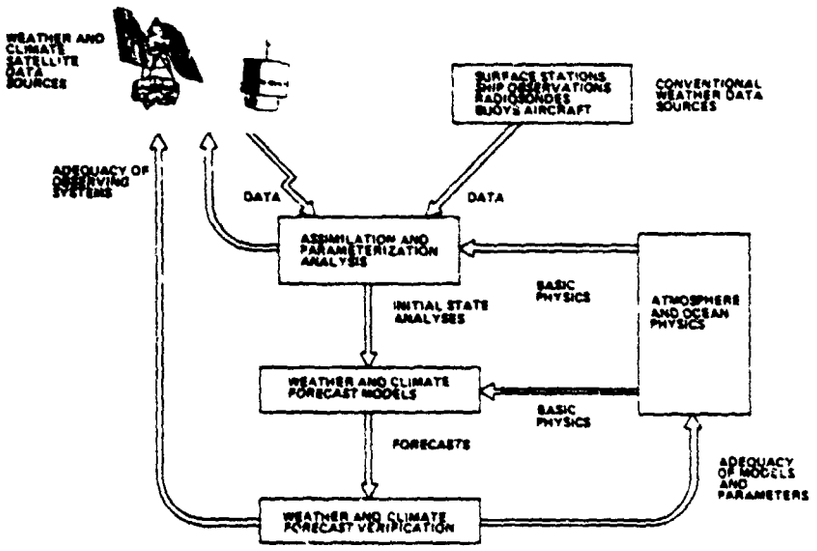


FIG. 45—*A schematic of the global data flow and planned data utilization for FGGE program.*

special intensive studies are planned so that the climate parameter can be derived and modeled later. Figure 47 illustrates the complicated climatic cause-and-effect (feedback) linkages which have an important effect on the

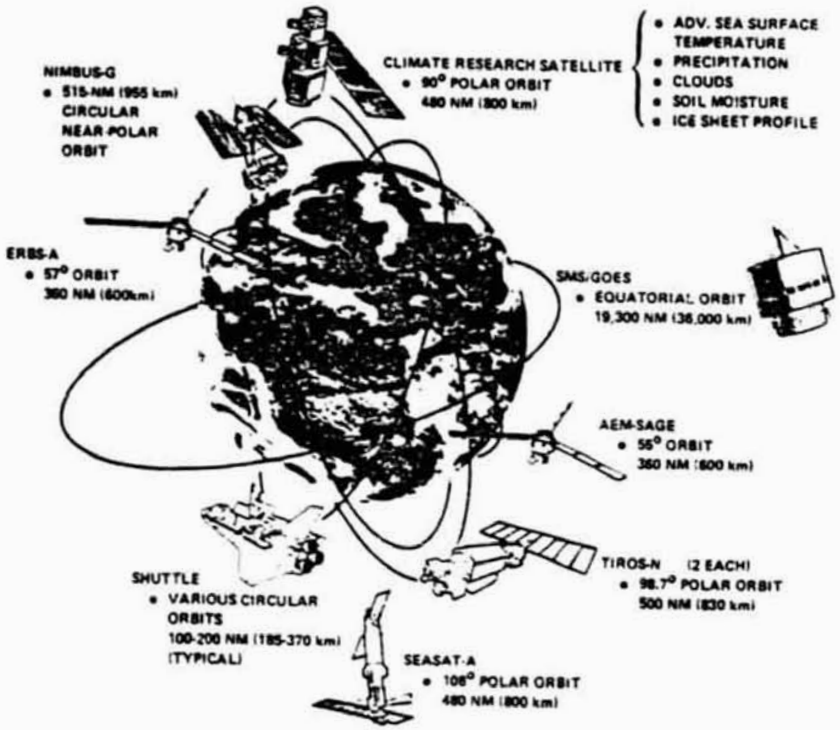


FIG. 46—*Proposed weather satellites and supporting research systems for the U.S. Climate Program.*

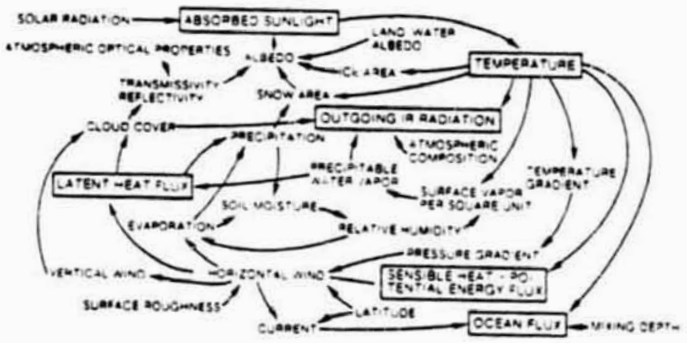


FIG. 47—*Climatic cause-and-effect (feedback) linkages* [77].

earth's climate. These and other meteorological parameters which may be partially monitored by satellites will be recorded and fed into mathematical climate models for further study [*77–80*].

The TIROS-N, a two-satellite system to be launched in 3rd quarter of 1978, was designed to provide better meteorological data than the present NOAA polar orbiting satellites. It's major improvements will be (*a*) higher spatial accuracy; (*b*) an increase in the number of tropospheric and stratospheric temperature and water vapor soundings under clear sky and cloudy conditions by the TIROS Operational Vertical Sounder (TOVS); (*c*) increased spectral radiometric information by the Advanced Very High Resolution Radiometer (AVHRR) for more accurate sea surface temperature mapping and delineation of melting snow and ice fields; and (*d*) a remote platform location and data collection capability (ARGOS) and increased proton, electron, and alpha particle spectral information by the Space Environment Monitor (SEM) for improved solar disturbance prediction [*81,82*]. Direct readout of Automatic Picture Transmission (APT) and High Resolution Picture Transmission (HRPT) data to a worldwide ground network will be provided for operational coverage. For purposes of the Climate Program, and to support NOAA operational needs, it has been suggested that two instruments be added to TIROS-N: (*a*) a solar Backscatter Ultra-Violet Instrument System (SBUV) to monitor global ozone and solar ultra-violet flux, and (*b*) an Earth Radiation Budget Satellite Instrument (ERBSI).

The SBUV measures backscattered and direct solar ultraviolet radiation in 12 channels of the UV spectrum. The instrument is nadir-viewing with a field of view covering 165 km on the earth. These data are used to derive the total ozone in the vertical column and also the vertical profile of the ozone concentration. The second instrument is the Earth Radiation Budget Instrument. This instrument system, consisting of two optical units, measures the total upwelling radiation in the 0.2 to 50-μm band and the reflected visible radiation in the 0.2 to 5.0-μm band. Three spatial resolutions are used in parallel, a wide field-of-view (1000 km) sensor, and a narrow field-of-view (87 km) scanning unit. These data are used in combination to derive global, zonal, and regional values of the components of the earth's radiation budget. For calibration purposes, a solar constant measurement channel is also included. Interest in including both the ERBS and SBUV instruments on TIROS-N has been expressed by NOAA. Table 3 gives the TIROS-N payload characteristics.

The orbital parameters and instrument payloads for an earth radiation budget and stratospheric aerosol satellite (ERBS-A) and CLIMSAT A and B are still under consideration for final selection.

The oceans, which cover about two-thirds of the earth's surface, have important influences on our weather and climate and are an important resource needing further exploration. The SEASAT-A program will demon-

TABLE 3—TIROS-N instrumentation.

Instrument	Channels	Resolution, km	Swath Width, km	Instrument Observation	Climate Parameters
ERBSI (proposed)	2 to 50 μm	1000	3000	visible and IR wide field and scanner	earth radiation budget
	0.2 to 5 μm	87			
SBUV (proposed)	12 channels 2550 to 3440 Å	165	nadir only	ultraviolet backscatter	ozone profile, total ozone
AVHRR/3	0.58 to .68 μm	1	3000	visible and IR imaging	sea surface temperature, snow cover, albedo, clouds, vegetation cover, ice sheets
	0.725 to 1.0				
	1.53 to 1.73				
	3.55 to 3.93				
	10.3 to 11.3				
	11.5 to 12.5				
TOVS—3 instruments					
HIRS	20 channels	30	1000	Multispectral IR radiation	temperature profiles, humidity profiles
SSU	3 channels 668 cm^{-1}	147	1450	Selective absorption of CO_2	stratosphere temperature profiles
MSU	4 channels 5.5 mm	100	1700	Passive microwave radiance	temperature profiles
ARGOS-Data collection and position location for surface sensors					

strate synoptic monitorings of the motions and temperatures of the oceans and provide better ocean ship routing and warnings of severe wind-wave, rain, sea current, and ice conditions which should benefit the oceanographic community.

SEASAT-A (Fig. 48) was launched in June 1978 and carried instruments fully dedicated to oceanic prediction [*83*]. The sensor applications are as follows:

1. *Altimeter*—A nadir-looking instrument that measures the displacement between the satellite and the ocean surface to a processed accuracy of 10 cm every 18 km and a RMS roughness of that surface to 1 m.
2. *Radar*—A 100-km, swath-width image of the ocean surface with a 25-m spatial resolution viewed every 18 days.
3. *Scatterometer*—Low to intermediate surface wind velocity determined over a swath-width of 1200 km, providing global coverage ($\pm$ 75 deg latitude) every 36 h on a 100-km grid basis.
4. *Microwave Radiometer*—Ice boundaries and leads, atmospheric water vapor, sea surface temperature, and intermediate to high wind speeds are provided over a swath-width of 1000 km every 36 h. Spatial resolution of ice features is 25 km, and 125 km for sea surface temperature.
5. *Visible and IR Radiometer*—Provide a 7-km spatial resolution imagery for feature identification for the microwave data.

An example of the ocean data distribution plan for SEASAT-A is shown in Fig. 49.

Nimbus G, to be launched in the 4th quarter of 1978, is a multidisciplinary satellite with application to pollution monitoring, oceanography, and weather and climate (Fig. 50). The payload consists of eight instruments [*84*].

1. *Scanning Multichannel Microwave Radiometer* (*SMMR*)—Measures radiances in five wavelengths and ten channels to extract information on sea surface roughness and winds, sea surface temperature, cloud liquid water content, precipitation (mean droplet size), soil moisture, snow cover, and sea ice [*85*].
2. *Stratospheric and Mesospheric Sounder* (*SAMS*)—Measures vertical concentrations of H_2O, N_2O, methane (CH_4), carbon monoxide (CO), and nitric oxide (NO); measures temperature of stratosphere to ~90 km and trace constituents.
3. *Solar-Backscattered Ultraviolet/Total Ozone Mapper System* (*SBUV TOMS*)—Measures direct and backscattered solar UV to extract information on variations of solar irradiance, vertical distribution of ozone and total ozone on a global basis.
4. *Earth Radiation Budget* (*ERB*)—Measures short- and longwave upwelling radiances and fluxes and direct solar irradiance to extract information on the solar constant, earth albedo, emitted longwave radiation, and the anisotropy of the outgoing radiation.

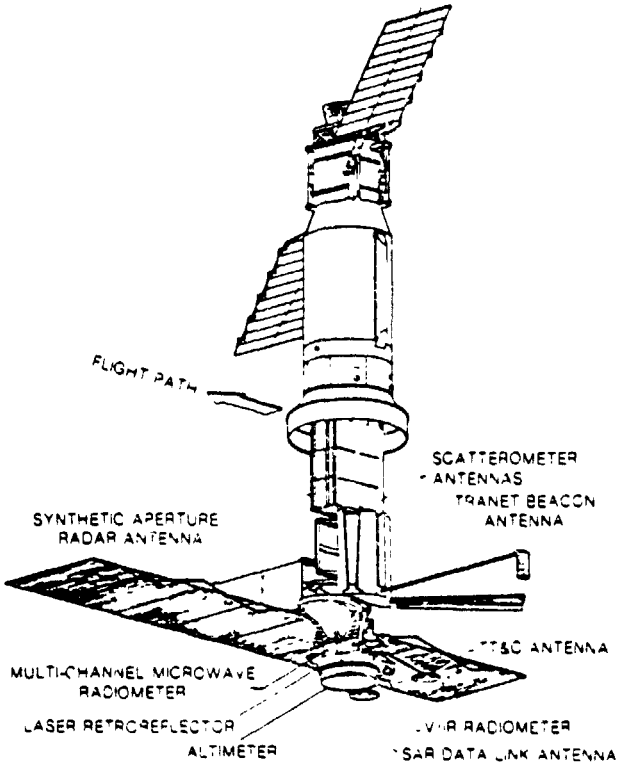


FIG. 48—*SEASAT-A with associated experiments.*

5. *Coastal Zone Color Scanner (CZCS)*—Measures chlorophyll concentration, sediment distribution, *gelbstoff* (yellow substance) concentration as a salinity indicator, and temperature of coastal waters and open ocean.

6. *Stratospheric Aerosol Measurements II Experiment (SAM II)*—Measures the concentration and optical properties of stratospheric aerosols as a function of altitude, latitude, and longitude. Tropospheric aerosols can be mapped also if no clouds are present in the IFOV.

7. *Temperature-Humidity Infrared Radiometer Experiment (THIR)*—Measures the infrared radiation from the earth in two spectral bands (11 and 6.7 μm) both day and night to provide pictures of cloud cover, three-dimensional maps of cloud cover, temperature maps of clouds, land and ocean surfaces, and atmospheric moisture.

8. *Limb Infrared Monitoring of the Stratosphere Experiment (LIMS)*—Makes a global survey of selected gases from the upper troposphere to the lower mesosphere. Inversion techniques are used to derive gas concentrations and temperature profiles.

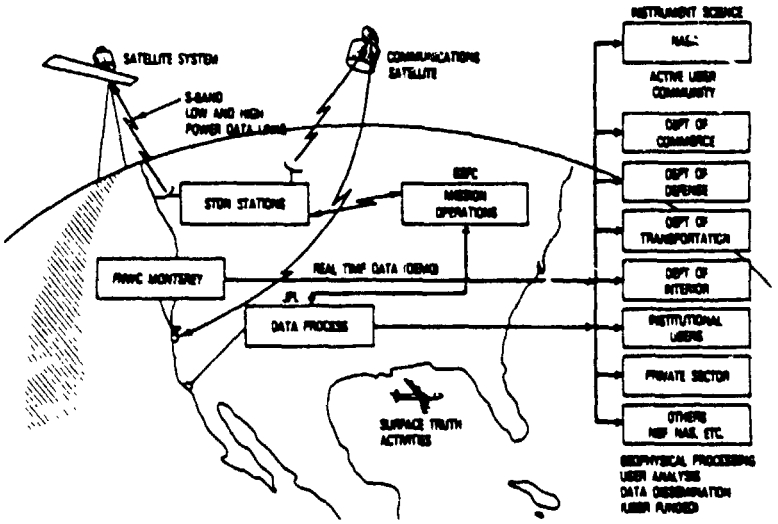


FIG. 49—*A schematic of SEASAT-A ocean data distribution plan.*

The Nimbus G launch will coincide with the FGGE and substantially enhance the results of that experiment.

The main objective of the proposed STORMSAT mission is to observe meso- and synoptic scale meteorological parameters from geostationary altitude for use in early detection and prediction of severe local storms [*86,87*]. Vertical temperature and moisture sounding under cloudy and clear sky conditions, and high resolution imagery from the Advanced Atmospheric Sounding and Imaging Radiometer (AASIR) and Microwave Atmospheric Sounding Radiometer (MASR), can best be obtained from a three-axis stabilized satellite. This system offers an order of magnitude sensing improvement over a spinning GOES VAS system. Preliminary design studies of the AASIR and the spacecraft and ground system design for STORMSAT has been completed [*88*] (Fig. 51). Figure 52 shows the chronological development of geostationary satellites to the present STORMSAT concept [*89*].

From early in the 1980s, and at least throughout the following decade, the Shuttle will be NASA's transportation system for access to space. Beyond that, the joint venture with European countries in developing the Spacelab will provide a functioning manned laboratory facility in space. These two programs will play vital roles in the development of future remote sensors. The four functional categories that these programs will provide are as follows: (*a*) calibration of instruments on operational satellites, (*b*) direct monitoring of slowly changing earth/atmosphere parameters, (*c*) develop-

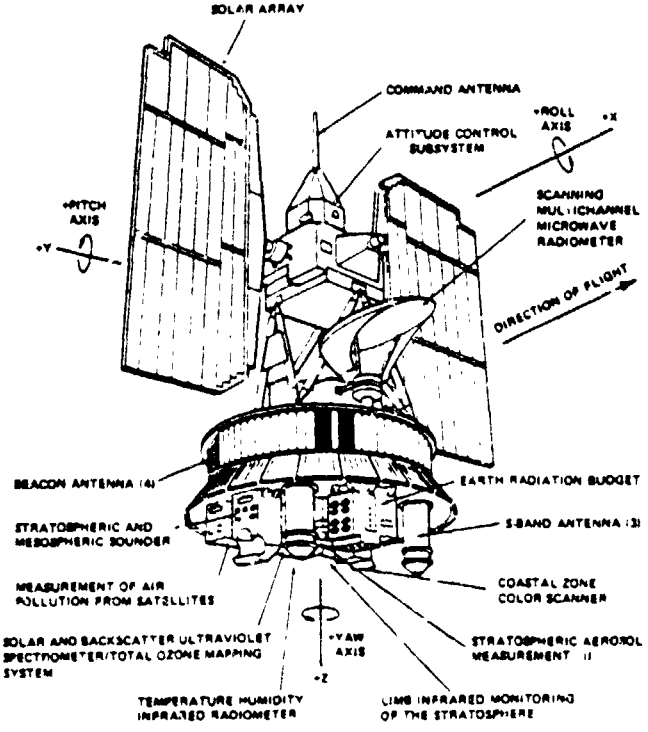


FIG 50—*Nimbus-G with associated experiments*

ment and demonstration of new remote sensors, and (*d*) special experiments impossible or impractical by other means.

The difficulties inherent to calibration and correction for long-term drift of satellite instruments, in the past, has presented manifold problems. First, logistic problems associated with deployment of enough ground systems and interrelating their independent errors imposed severe limitations. A second factor is that ground instruments do not generally measure the same quantities as satellite instruments. For example, a satellite radiometer measures a T_B integrated over the entire scene in view, while ground truth data are restricted usually to point measurements. Thus, at least calibrations of satellite instruments over their several years of useful lifetimes and the ability to interrelate succeeding satellites, is a crucial requirement at least for NASA's emerging Climate Program. The Shuttle can fulfill this requirement by periodic flight of instruments, calibrated against standards before and after flight, that can provide comparative measurements at points of orbit conjunction with the satellites when identical scenes are in view.

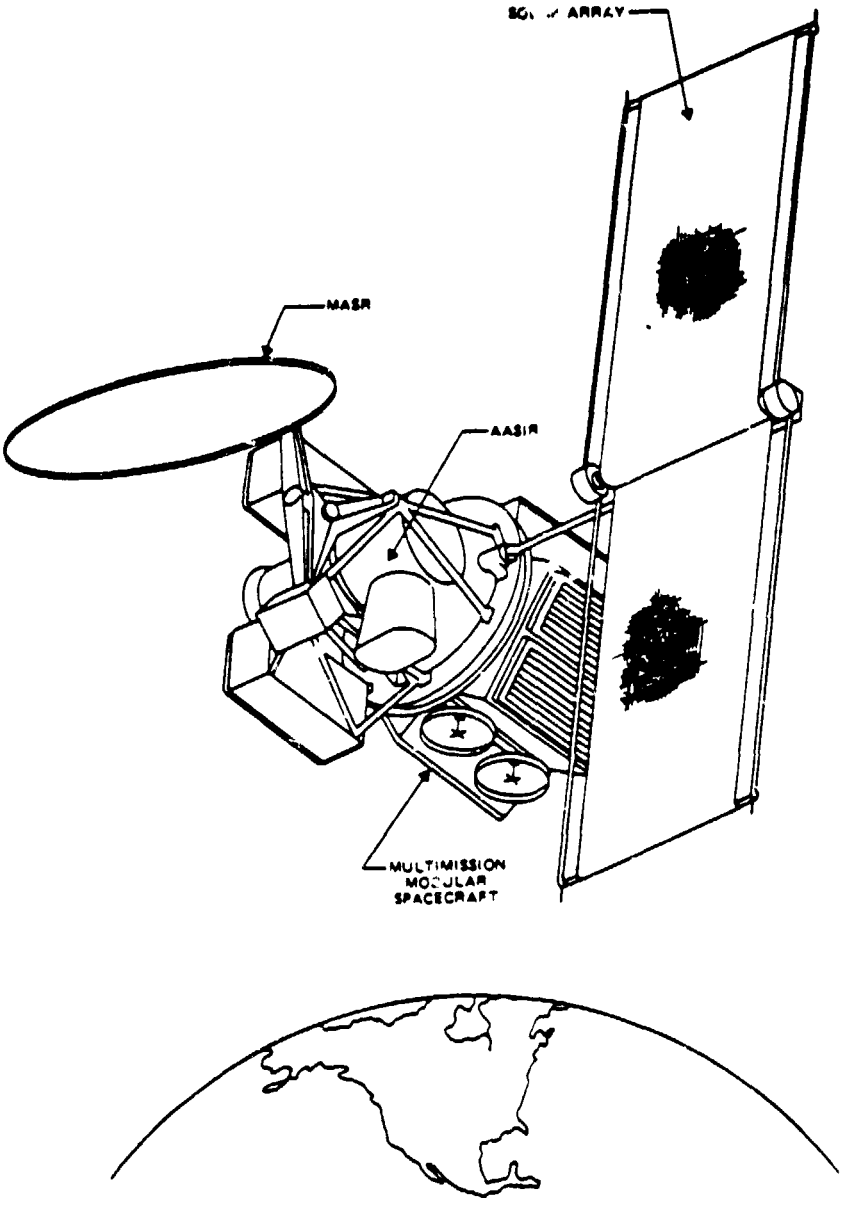


FIG. 51—*Conceptual STORMSAT spacecraft with associated experiments*

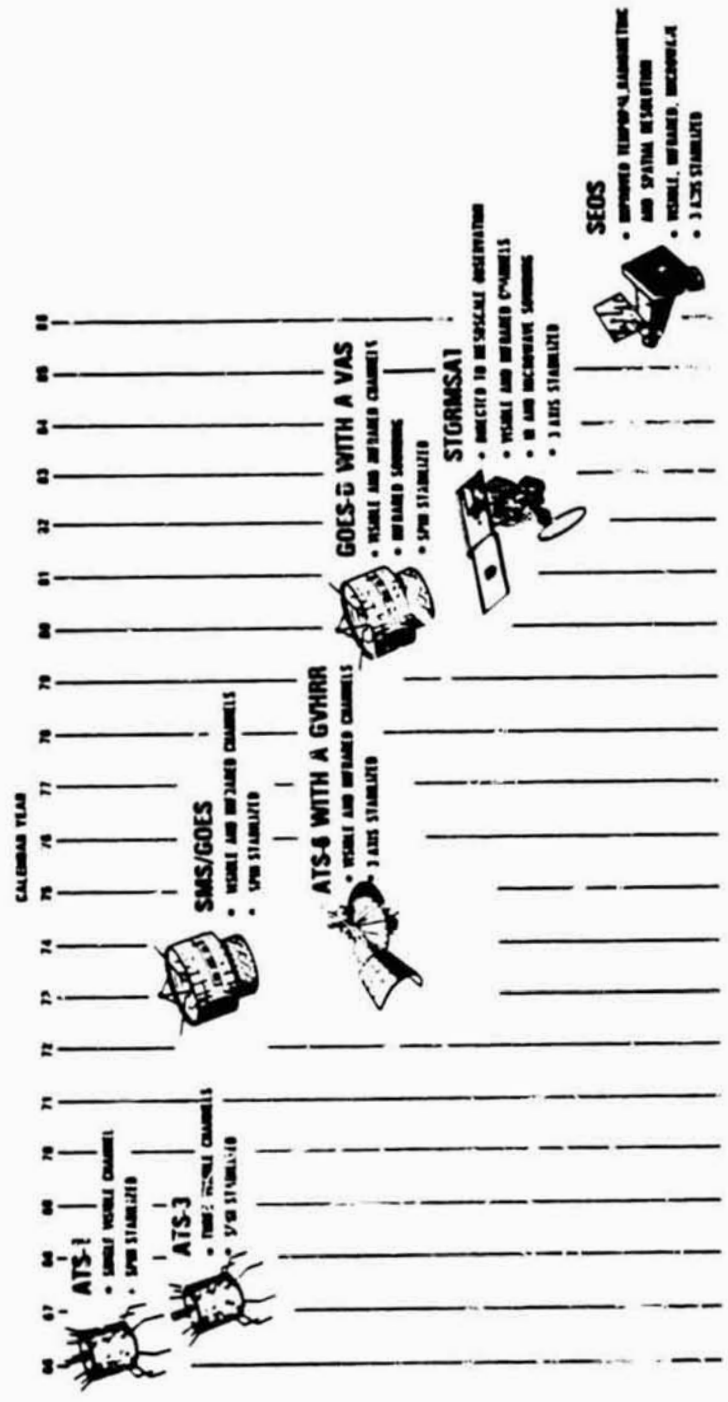


FIG. 52—*Chronological development of geostationary meteorological satellites leading to the STORMSAT and later SEOS concept.*

Monitoring of certain slowly varying parameters can be accomplished directly by regular flights in conjunction with Shuttle-Spacelab missions. The requirements for measuring the solar spectral irradiance can be met adequately by two flights per year spaced at six-month intervals (Fig. 53).

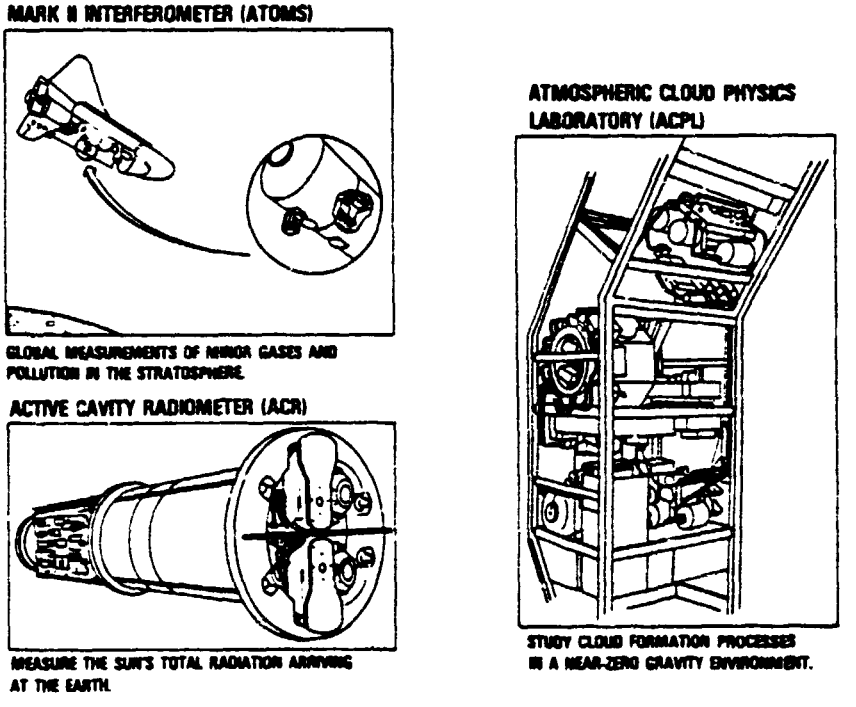


FIG. 53—*Instrument development proposed for Shuttle-Spacelab use.*

Experience with the Nimbus-6 ERB instrument substantiates the desirability of continuing total solar flux measurements at six-month intervals. Such measurements will provide an independent and highly calibrated set of data by themselves, but will also serve to interrelate the total solar flux measurements of the ERBSI proposed for TIROS-N. Other parameters that can be directly monitored by the Shuttle-Spacelab include a number of the well mixed atmospheric constituents, ozone, and tropospheric aerosols. For example, annual measurements are probably adequate for CO_2, CFMs, N_2O, NO_x, and CH_4.

The Spacelab will also provide the needed facilities to develop, test, and demonstrate new remote sensors. These include:

1. *Sea Surface Temperature*—Improvements that appear possible include increased signal-to-noise ratios and additional wavelengths to reduce errors due to atmospheric water vapor and clouds.

2. *Vertical Temperature Profile*—Accuracy of passive sounding below 20-km altitude can be improved with measurements at 4 μm and with bandwidths of 2 wave numbers. Active Lidar promises even higher accuracy along with an order of magnitude-improved vertical resolution.

3. *Winds*—Lidar techniques may be able to measure wind speed and direction in cloud-free regions.

4. *Stratospheric Aerosols*—Stellar sources provide a large number of occultations not limited to the ecliptic plane, so that reasonably good coverage of stratospheric aerosols is possible.

5. *Tropospheric Aerosols*—At present, no technique exists for measuring tropospheric aerosols, but a pulsed lidar system is believed to hold promise.

As plans are formulated for the development of these instruments (as well as any others that may yet be identified) to the point where space flight development and testing can be visualized, the advantages that Spacelab can provide will be capitalized upon to the maximum extent possible.

There are two particularly important special experiments in the area of remote sensing that are especially suited to the Spacelab capabilities. These are extended cloud physics and radiation studies and precipitation over land and water. In the case of the extended cloud and radiation studies a group of instruments, including two separate radiometers and a lidar system, are required. In the case of the precipitation studies, a microwave radiometer, a visible band radiometer, an IR radiometer, and a 10-cm radar are required. A number of short duration flights are adequate to meet the objectives of both of these studies and Spacelab offers a unique capability for their accomplishment.

Conclusion

The meteorological satellites have used instrumentation sensitive to different portions of the electromagnetic spectrum to observe and measure atmospheric and surface properties. In the ultraviolet, solar variability has been measured and total atmospheric ozone determined. In the visible, imagery of cloud systems has (*a*) provide better location of storms in data-sparse areas, (*b*) observed ice and snow boundaries, and (*c*) monitored floods. In the infrared, sea surface and cloud top temperatures have been measured and water vapor flow patterns determined. By using infrared CO_2 absorption bands, centered at 4.3 and 15 μm, radiometers have derived atmospheric vertical temperature profiles. Microwave radiometer measurements in the 5-mm, O_2 absorption bands have improved these atmospheric soundings and allowed penetration through clouds. Microwave measurements near 1.55 and 0.8 cm have determined rainfall intensities over the ocean, snow and ice boundaries through clouds, and have provided some indications of soil moisture.

Cloud tracking from geosynchronous satellites have allowed wind determinations at different levels from which important properties, such as divergence and vorticity, have been derived. Geosynchronous satellites have also monitored the development and motions of severe storms such as hurricanes, or severe thunderstorms with hail or tornados. These observations and measurements have not only increased our knowledge of atmospheric processes but have also been used operationally in short-term forecasting. For practical use, the satellite provides observations denser in time and space scales than any national network of surface observing stations.

Future approved satellite missions include improved resolutions for some of the foregoing measurements, but also include new experiments such as pollution detection and oceanographic monitoring. Proposed satellite missions include STORMSAT, a geosynchronous satellite, for better monitoring of severe storms with temperature and moisture sounding capabilities.

A proposed Climate Program would provide observations from a system of satellite sensors for numerical modeling of climate, leading to a better understanding of climate processes and, hopefully, to an improved capability to predict climate change.

Acknowledgment

The authors wish to thank Professor G. C. Gill, University of Michigan for his encouragement in writing this paper.

References

[1] Hansen, J. E. and Travis, L. D., *Space Science Reviews* Vol. 16, 1974, pp. 527-610.

[2] Kunde, V. G., Hanel, R. A., and Herath, L. W., "High Spectral Resolution Ground-Based Observations of Venus in the 450-1250 cm^{-1} Region," NASA Report X-622-76-249, National Aeronautics and Space Administration, Washington, D.C., 1976.

[3] Stoldt, N. W. and Havanac, P. J., "Compendium of Meteorological Satellites and Instrumentation," NSSDC Report 73-02, National Space Science Data Center, NASA, Goddard Space Flight Center, Greenbelt, Md., 1973.

[4] Allison, L. J. and Neil, E. A., "Final Report on the TIROS 1 Meteorological Satellite System," NASA Technical Report R-131, Goddard Space Flight Center, Greenbelt, Md., 1962.

[5] Schwalb, A. and Gross, J., "Vidicon Data Limitations," ESSA Technical Memo and NESC Technical Memo 17, U.S. Department of Commerce, ESSA, NESC, Washington, D.C., 1969.

[6] Conlan, E. F., "Operational Products from ITOS Scanning Radiometer Data," NOAA Technical Memo NESS 52, U.S. Dept. of Commerce, Washington, D.C., 1973.

[7] Ostrow, H. and Weinstein, O. in *Proceedings*, SPIE 13th Annual Technical Symposium, Society of Photographic Optical Instrumentation Engineers, Washington, D.C., 1968, pp. 251-262.

[8] Minzner, R. A. and Oberholtzer, J. D., "Space Applications Instrumentation Systems," NASA Technical Memo X-2066, National Aeronautics and Space Administration, Washington, D.C., 1972.

[9] Press, H. and Huston, W. B., *Astronautics and Aeronautics*, Vol. 3, March 1968, pp. 56-65.
[10] McCulloch, A. W., in *Nimbus 5 User's Guide*, ERTS-Nimbus Project, Goddard Space Flight Center, Greenbelt, Md., 1972, pp. 11-47.
[11] Suomi, V. E., and Von der Haar, T. H., *Journal of Spacecraft and Rockets*, Vol. 6, No. 3, 1969, pp. 342-344.
[12] Corbell, R. P., Callhan, C. J., and Kotsch, W. J., *The GOES/SMS User's Guide*, NOAA-NESS, NASA, Washington, D.C. 1976.
[13] Bristor, C. L., "Central Processing and Analysis of Geostationary Satellite Data," NOAA Technical Memo NESS 64, 1975.
[14] Kaplan, L. D., *Journal of the American Optical Society*, Vol. 49, 1959, pp. 1004-1007.
[15] Smith, W. L., Woolf, H. M., Abel, P. G., Hayden, C. M., Chalfunt, M., and Grody, N., "Nimbus 5 Sounder Data Processing System," NOAA Technical Memo NESS 57, 1974.
[16] Hayden, C., "The Use of the Radiosonde in Deriving Temperature Soundings from the Nimbus and NOAA Satellite Data," NOAA Technical Memo NESS 76, 1976.
[17] Fortuna, J. J. and Hambrick, L. N., "The Operation of the NOAA Polar Satellite System," NOAA Technical Memo NESS 60, U.S. Dept. of Commerce, NOAA, NESS, Washington, D.C., 1974.
[18] Hoppe, E. R. and Ruiz, A. L., 1974, *Catalog of Operational Satellite Products*, NOAA Technical Memo NESS 53, U.S. Dept. of Commerce, NOAA, NESS, Washington, D.C., 1974.
[19] Minzner, R. A., Shenk, W. E., Steranka, J., and Teagle, R. D., "Cloud Heights Determined Stereographically From Imagery Recorded Simultaneously by Two Synchronous Meteorological Satellites, SMS-1 and SMS-2," *Proceedings*, American Geophysical Union, Spring 1976, Washington, D.C.
[20] Bauer, K. E., in *Studies of the Atmosphere Using Aerospace Probes*, U. of Wisconsin, Madison, Wis., pp. 45-78.
[21] Hubert, L. E. and Whitney, L. F., "Comparability of Low Level Cloud Vectors and Rawins for Synoptic Analysis," NOAA Technical Memo NESS 70, 1974.
[22] Hasler, A. F., Shenk, W., and Skillman, W., "Wind Estimates from Cloud Motions," *Journal of Applied Meteorology*, Vol. 15, 1976, pp. 10-15
[23] Billingsley, J. B., in *Proceedings*, 7th Conference on Aerospace and Aeronautical Meteorology and Symposium on Remote Sensing from Satellites, 16-19 Nov. 1976, Melbourne, Fla., pp. 268-275.
[24] Gentry, R. C., Rodgers, E. B., Shenk, W. E., and Oliver, V., in *Proceedings*, 7th Conference on Aerospace and Aeronautical Meteorology and Symposium on Remote Sensing from Satellites, 16-19 Nov. 1976, Melbourne Fla., pp. 115-118.
[25] Szekielda, K. H., *Oceanography and Marine Biology Annual Review*, Vol. 14, 1976, pp. 99-166.
[26] Brower, R. L., Gohrband, H. S., Pichel, W. G., Signore, T. L., and Walton, C. C., "Satellite Derived Sea-Surface Temperatures from NOAA Spacecraft," NOAA Technical Memo NESS 78, U.S. Dept. of Commerce, NOAA, NESS, Washington, D.C., 1976.
[27] Stumpf, H. G., *Mariners Weather Log*, Vol. 18, 1974, pp. 149-152.
[28] Gulfstream Staff, *Gulfstream*, Vol. 11, No. 9, 1977, pp. 6-7
[29] Cogan, J. L. and Willand, J. H., "Mapping of Sea Surface Temperature by the DMSP Satellite," ERT Document No. 1065, Final Report EPRF 7475 (ERT), Earth Resources Technology, Inc., Concord, Mass., 1975
[30] Barazotto, R. M., *Mariners Weather Log*, Vol. 21, No. 2, 1977, pp. 73-77.
[31] Campbell, W. J., Gloersen, P., Nordberg, W., and Wilheit, T., "Dynamics and Morphology of Beaufort Sea Ice," NASA Report GSFC X-650-73-391, Goddard Space Flight Center, Greenbelt, Md., 1973.
[32] Gloersen, P., Nordberg, W., Schmugge, T., Wilheit, T., and Campbell, W., "Microwave Signatures of First-Year and Multiyear Sea Ice," NASA Report X-652-72-312, Goddard Space Flight Center, Greenbelt, Md., 1972.
[33] Chang, T. C., Gloersen, P., Schmugge, T., Wilheit, T. T., and Zwally, H. J., *Journal of Glaciology*, Vol. 16, No. 74, 1976, pp. 23-39
[34] Zwally, H. J., *Journal of Glaciology*, Vol. 18, No. 79, 1977, pp. 195-215.
[35] Zwally, H. J. and Gloersen, P., *Polar Record*, Vol. 18, No. 116, 1977, pp. 431-450

[36] Curran, R. and Wexler, R., "Separation of Cloud Contribution to the Monthly Zonal Albedo," unpublished manuscript, 1977.
[37] Zwally, H. J., Wilheit, T. T., Gloersen, P., and Mueller, J. L., in *Proceedings*, Symposium on Meteorological Observations from Space, Their Contributions to the First GARP Experiment, National Center for Atmospheric Research, Boulder, Colo., 1976, pp. 94-97.
[38] Sanderson, R. M., *Meteorological Magazine*, Vol. 104, No. 1240, 1975, pp. 313-323.
[39] Ackley, S. F. and Keliher, T. E., *AIDJEX Bulletin*, No. 33, 1976, pp. 53-76.
[40] Wiesnet, D. and Matson, M., "Monthly Winter Snow Line Variation in the Northern Hemisphere from Satellite Records," NOAA Technical Memo NESS 74, 1975.
[41] Rango, A., *Proceedings*, 10th International Symposium on Remote Sensing of the Environment, 6-10 Oct. 1975, Environmental Research Institute of Michigan, University of Michigan, Ann Arbor, Mich., pp. 1367-1376.
[42] Follansbee, W. A., "Estimation of Daily Precipitation over China and the USSR Using Satellite Imagery," NOAA Technical Memo NESS 81, U.S. Dept. of Commerce, NOAA, NESS, Washington, D.C., 1976.
[43] Martin, D. W. and Scherer, W. D., *Bulletin of the American Meteorological Society* Vol. 54, 1973, pp. 661-674.
[44] Griffith, C. G. and Woodley, W. L., "Rainfall Estimates from Geosynchronous Satellite Imagery," *Proceedings*, 11th Space Congress, Coco Beach, Fla., 17-19 Aug. 1974.
[45] Allison, L. J., Rodgers, E. B., Wilheit, T. T., and Fett, R. W., *Bulletin of the American Meteorological Society*, Vol. 55, No. 9, 1974, pp. 1074-1089.
[46] Allison, L. J., Rodgers, E. B., Wilheit, T. T., and Wexler, R., "A Multisensor Analysis of Nimbus 5 Data Recorded on January 22, 1973," NASA Technical Note D-7911, Goddard Space Flight Center, Greenbelt, Md., 1975.
[47] Rao, M. S. V., Abbott, W. V., and Theon, J. S., "Satellite-Derived Global Oceanic Rainfall Atlas (1973-1974)," NASA Special Publication 410, National Aeronautics and Space Administration, Washington, D.C., 1976.
[48] Yates, H. W. and Bandeen, W. R., in *Proceedings*, Institute of Electrical and Electronics Engineers, Vol. 63, No. 1, 1975, pp. 148-163.
[49] Schmugge, T., Gloersen, P., Wilheit, T., and Geiger, F., *Journal of Geophysics Research*, Vol. 79, 1974, p. 317.
[50] Schmugge, T., Wilheit, T., Webster, W., Jr., and Gloersen, P., "Remote Sensing of Soil Moisture with Microwave Radiometers-II" NASA Technical Note D-8321, National Aeronautics and Space Administration, Washington, D.C., 1976.
[51] Meneely, J. M., "Application of the Electrically Scanning Microwave Radiometer (ESMR) to Classification of the Moisture Condition of the Ground," Final Report, Earth Satellite Corporation, Washington, D.C., 1977.
[52] Rango, A. and Anderson, A. T., in *Proceedings*, 3rd Earth Resources Technology Satellite-1 Symposium, Vol. 1, Section B, NASA, Goddard Space Flight Center, Greenbelt, Md., 1973, pp. 1127-1166.
[53] Short, N., Lowman, P. D., Jr., and Freden, S. C., "Mission to Earth: LANDSAT Views the World," NASA Special Publication 360, National Aeronautics and Space Administration, Washington, D.C., 1976.
[54] Otterman, J., Ohring, G., and Ginzburg, A., *Remote Sensing of the Environment*, Vol. 3, 1974, pp. 133-148.
[55] Otterman, J. and Fraser, R. S., *Remote Sensing of the Environment*, Vol. 5, 1976, pp. 247-266.
[56] Allison, L. J., "Geological Applications of Nimbus Radiation Data in the Middle East," NASA Technical Note D8469, Goddard Space Flight Center, Greenbelt, Md., 1977.
[57] Otterman, J., Lowman, P. D., and Salomonson, V. V., *Geophysical Surveys*, Vol. 2, 1976, pp. 431-467.
[58] Kahle, A. B., Gillespie, A. R., and Goetz, A. F. H., *Geophysical Research Letters*, Vol. 3, No. 1, 1976, pp. 26-28.
[59] MacDonald, R. B., in *Proceedings*, 2nd Annual W. T. Pecora Memorial Symposium, Sioux Falls, S. Dak., 25-29 Oct. 1976, p. 20.
[60] McGinnis, D. F., Pritchard, J. A., and Wiesnet, D. R., *Water Resources Research*, Vol. 11, No. 6, 1975, pp. 897-902.

[61] Barnes, J. C. and Bowley, C. J., *Handbook of Techniques for Satellite Snow Mapping*, ERT Document 0407-A, Earth Resources Technology, Inc., Concord, Mass., 1974.
[62] Rango, A. and Salomonson, V. V., *Water Resources Research*, Vol. 13, No. 1, 1977, pp. 109–112.
[63] Mercanti, E. P., *Astronautics and Aeronautics*, Vol. 12, No. 5, May 1974, pp. 28–39.
[64] Horan, J., *IEEE Spectrum*, Vol. 13, No. 3, Institute of Electrical and Electronic Engineers, March 1976, pp. 59–62.
[65] Schmugge, T., Wilheit, T. T., Gloersen, P., Meier, M. F., Frank, D., and Dirmhern, I., in *Advanced Concepts and Techniques in the Study of Snow and Ice Resources*, H.S. Santeford and J. L. Smith, Eds., National Academy of Sciences, Washington, D.C., 1974, pp. 551–562.
[66] Duncan, C. H., Harrison, R. G., Hickey, J. R., Kendall, J. M., Sr., Thekaekara, M. P., and Willson, R. C., *Applied Optics*, Vol. 16, No. 10, 1977, pp. 2690–2697.
[67] Smith, W. L., Hickey, J., Howell, H. B., Jacobowitz, H., Hilleary, D. T., and Drummond, A. J., *Applied Optics*, Vol. 16, No. 2, 1977, pp. 306–318.
[68] Gruber, A. in *Proceedings*, Symposium on Radiation in the Atmosphere (IAMAP), 19–28 Aug. 1976, Garmisch-Partenkirchen, Germany.
[69] Prabhakara, C., Conrath, B., and Hanel, R. A., *Journal of Atmospheric Sciences*, Vol. 27, 1970, pp. 689–697.
[70] Prabhakara, C., Salomonson, V. V., Conrath, B. J., Steranka, J., and Allison, L. J., *Journal of the Marine Biological Association of India*, Vol. 14, No. 1, 1973, pp. 42–54.
[71] Krueger, A. J., Heath, D. F., and Mateer, C. L., *Pure Applied Geophysics*, No. 106–108, 1973, pp. 1254–1263.
[72] Heath, D. F., Krueger, A. J., Roeder, H. A., and Henderson, B. D., *Optical Engineering*, Vol. 14, No. 4, 1975, pp. 323–331.
[73] Gille, J. C., House, F. B., Craig, R. A., Thomas, J. R., in *Nimbus 6 User's Guide*, NASA, Goddard Space Flight Center, Greenbelt, Md., 1975, pp. 141–162.
[74] Global Atmospheric Research Program (GARP), "The First GARP Global Experiment," GARP Publications Series No. 11, WMO-ICSU Joint Organizing Committee, 1973.
[75] "World Weather Watch, Global Observing System-Satellite Sub-System" WMO No. 114, Secretariat of the World Meteorological Organization, Geneva, Switzerland, 1975.
[76] Global Atmospheric Research Program (GARP), "The Physical Basis of Climate and Climate Modeling," GARP Publications Series No. 16, WMO-ICSU Joint Organizing Committee, 1975.
[77] Kellogg, W. W. and Schneider, S. H., *Science*, Vol. 186, No. 4170, 1974, pp. 1163–1172.
[78] Schneider, S. H. and Dickinson, R. E., *Reviews of Geophysics and Space Physics*, Vol. 12, No. 3, 1974, pp. 447–493.
[79] "Weather and Climate," Report of the Panel of Weather and Climate to the Space Applications Board of the Assembly of Engineering, National Research Council, National Academy of Sciences, 1975.
[80] Laurmann, J. A., *Bulletin of American Meteorological Society*, Vol. 56, No. 10, 1975, pp. 1084–1088.
[81] Schneider, J. R., "Guide for Designing RF Ground Receiving Stations for TIROS-N," NOAA Technical Report NESS 75, U.S. Dept. of Commerce, NOAA, NESS, Washington, D.C., 1976.
[82] Hussey, W. J., "The TIROS-N Polar Orbiting Environmental Satellite System," National Environmental Satellite Service, National Oceanic and Atmospheric Administration, Washington, D.C., 1977, p. 32.
[83] NOAA Staff Members, "NOAA Program Development Plan for SEASAT-A Research and Applications," U.S. Dept. of Commerce, NOAA, NESS, Washington, D.C., 1977.
[84] Staff Members, Goddard Space Flight Center, "Nimbus G, Nimbus Observation Processing System (NOPS) Design Study Report," Goddard Space Flight Center, Greenbelt, Md., Jan. 1976.
[85] Gloersen, P. and Barath, F. T., *IEEE Journal of Oceanic Engineering*, Vol. OE-2, No. 2, 1977, pp. 172–178.
[86] Fujita, T., *Weatherwise*, Vol. 26, No. 2, 1973, pp. 56–83.
[87] Fujita, T., Forbes, G. S., Umenhofer, T. A., *Weatherwise*, Vol. 29, No. 3, 1976, pp. 117–145.

[88] "Stormsat Ground System Concept Study," CSC/SD-76/6088, Contract NAS 5-11999 Computer Sciences Corp., Silver Springs, Md., 1976.
[89] "The Federal Plan for Meteorological Services and Supporting Research," NOAA, U.S. Dept. of Commerce, 1976, Washington, D.C.
[90] Lintz, J., Jr. and Simonett, D., Eds., *Remote Sensing of Environment*, Addison-Wesley Publishing Co., Reading, Mass., 1976.
[91] Valley, S. L., Ed., *Handbook of Geophysics and Space Environments*, Air Force Cambridge Research Laboratory, Bedford, Mass., 1965.
[92] Sellers, W. D., *Physical Climatology*, The University of Chicago Press/Chicago and London, 1975.
[93] Johnson, J. D., Parmenter, F. C., and Anderson, R., "Environmental Satellites," U.S. Dept. of Commerce, Washington, D. C., NOAA, NESS, 1976.

Paper 112

Reprinted from JOURNAL OF APPLIED METEOROLOGY, Vol. 17, No. 6, June 1978
American Meteorological Society
Printed in U.S.A.

A Fourier Approach to Cloud Motion Estimation

ALBERT ARKING

Goddard Space Flight Center, Greenbelt, Md. 20771

ROBERT C. LO[1] AND AZRIEL ROSENFELD

Computer Science Center, University of Maryland, College Park 20742

(Manuscript received 17 August 1977, in final form 13 January 1978)

ABSTRACT

A Fourier phase difference technique for cloud motion estimation from pairs of pictures is described. This technique makes use of the phase of the cross-spectral density; it allows motion estimates to be made for individual spatial frequencies, which are related to cloud pattern dimensions. Results obtained using this technique are presented and are compared with the results of a Fourier-domain cross-correlation scheme, using both artificial and real cloud data. It is concluded that the phase difference technique is relatively sensitive to the presence of mixtures of motions, changes in cloud shape and edge effects. Under these circumstances, the cross-correlation scheme yields a more reliable estimate of cloud motion.

1. Introduction

Sequences of high-resolution photographs from satellites in geosynchronous orbit have been utilized by many investigators to measure cloud motion and derive winds. The techniques for extracting such information range from a manual analysis of closed "movie loops" (Fujita *et al.*, 1969; Hubert and Whitney, 1971) to completely automated methods in which the images and/or data are computer processed but selection of cloud targets is controlled manually from an interactive display console (Serebreny *et al.*, 1969; Smith and Phillips, 1972).

The main advantage of manual and semi-automatic methods is in target selection, making it possible to incorporate a wide variety of criteria, not necessarily defined beforehand, for choosing cloud elements representative of atmospheric motion. These criteria are sometimes related to the size of the cloud element, the idea being that small elements move with the wind, while massive elements may disguise the actual motion of the atmosphere due to cloud formation and deformation processes.

Early studies by Leese and Epstein (1963) have shown that one can identify the predominant dimensions of cloud patterns by a spectral analysis of satellite photographs. Subsequently, Leese *et al.* (1971) used the Fourier transform as an efficient computational tool for obtaining the cross correlation between successive cloud cover pictures, from which cloud displacements are derived. The fact that displacement information appears separately for each spectral component in the Fourier transform of the cross-covariance function was pointed out by Weinstein (1972). The combined advantages of the Fourier transform—as an efficient computational tool for measuring displacement and as a means for extracting cloud pattern dimensional information—have motivated the present study of the Fourier transform properties of real and simulated sequences of satellite cloud cover pictures and their cross correlations. Specifically, we describe an attempt to derive winds from cloud motion, while discriminating with respect to cloud size.

The mathematical aspects of the Fourier transform method are outlined in Section 2, the results of a simulation study are presented in Section 3, a preliminary experiment with real pictures is described in Section 4, and the outlook is discussed and conclusions drawn in Section 5.

2. Properties of the Fourier transform

A function defined in the space domain, say $f(x,y)$, where x and y are coordinates in a two-dimensional Cartesian space, can be transformed into a function in the frequency domain by using the Fourier transform $\mathbf{F}$, i.e.,

$$\mathbf{F}[f(x,y)] = \int_{-\infty}^{\infty}\int f(x,y)e^{-i2\pi(ux+vy)}dxdy = F(u,v), \quad (1)$$

[1] Present affiliation: Computer Sciences Corp., Silver Spring, Md.

0021-8952/78/0735-0744$05.00
© 1978 American Meteorological Society

GRAY LEVELS OF FIRST IMAGE

8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
8	8	8	8	8	8	8	8	9	9	9	8	8	8	8	8
8	8	8	8	8	8	9	9	9	9	9	8	8	8	8	8
8	8	8	8	8	8	9	9	10	10	9	9	9	8	8	8
8	8	8	8	8	9	9	10	10	10	10	9	9	8	8	8
8	8	8	8	9	9	9	10	10	10	10	9	9	8	8	8
8	8	8	8	9	9	10	10	10	10	10	9	9	8	8	8
8	8	8	8	9	9	10	10	10	10	10	9	9	8	8	8
8	8	8	8	16	17	9	9	17	10	9	9	9	8	8	8
8	8	8	8	16	18	9	17	21	17	9	9	8	8	8	8
8	8	8	8	8	8	8	17	17	17	9	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8

GRAY LEVELS OF SECOND IMAGE

8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
8	8	8	8	8	8	8	8	8	9	9	9	8	8	8	
8	8	8	8	8	8	8	8	9	9	9	9	9	9	8	8
8	8	8	8	8	8	8	19	17	10	10	17	9	9	8	
8	8	8	8	8	8	9	17	20	10	18	22	17	9	8	
8	8	8	8	8	9	9	9	10	10	18	18	17	9	8	
8	8	8	8	8	9	9	10	10	10	10	10	9	9	8	
8	8	8	8	8	8	9	9	10	10	10	10	10	9	9	8
8	8	8	8	8	8	8	9	9	9	9	10	9	9	9	8
8	8	8	8	8	8	8	8	9	9	9	9	9	9	8	8
8	8	8	8	8	8	8	8	8	9	9	9	9	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	9	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8

FIG. 1a. Gray levels of a pair of images simulating cloud motion. A large mass of relatively dim cloud is undergoing a displacement (in pixels) of $X=2$ and $Y=1$, on which are superimposed smaller but brighter clouds undergoing a displacement of $X=4$, $Y=6$.

POWER SPECTRUM OF FIRST IMAGE

8.73	0.51	0.15	0.13	0.19	0.13	0.05	0.03	0.01	0.03	0.05	0.13	0.19	0.13	0.15	0.51
0.47	0.28	0.10	0.08	0.18	0.12	0.06	0.04	0.02	0.04	0.05	0.13	0.19	0.16	0.09	0.38
0.28	0.29	0.16	0.05	0.14	0.12	0.03	0.06	0.06	0.03	0.05	0.11	0.17	0.18	0.07	0.16
0.26	0.26	0.15	0.02	0.12	0.10	0.01	0.07	0.07	0.05	0.06	0.11	0.15	0.15	0.08	0.13
0.18	0.22	0.15	0.01	0.09	0.08	0.03	0.04	0.06	0.05	0.07	0.10	0.11	0.14	0.10	0.06
0.10	0.16	0.14	0.05	0.05	0.07	0.03	0.05	0.06	0.03	0.04	0.06	0.07	0.09	0.08	0.01
0.05	0.10	0.10	0.05	0.06	0.05	0.01	0.03	0.05	0.04	0.03	0.04	0.03	0.05	0.07	0.03
0.00	0.04	0.06	0.06	0.04	0.02	0.01	0.02	0.04	0.03	0.02	0.03	0.02	0.02	0.04	0.03
0.02	0.01	0.01	0.02	0.02	0.01	0.00	0.01	0.02	0.01	0.00	0.01	0.02	0.02	0.01	0.01
0.00	0.03	0.04	0.02	0.02	0.03	0.02	0.03	0.04	0.02	0.01	0.02	0.04	0.06	0.06	0.04
0.05	0.03	0.07	0.05	0.03	0.04	0.03	0.04	0.05	0.03	0.01	0.05	0.06	0.05	0.10	0.10
0.10	0.01	0.08	0.09	0.07	0.06	0.04	0.03	0.06	0.05	0.03	0.07	0.05	0.05	0.14	0.16
0.18	0.06	0.10	0.14	0.11	0.10	0.07	0.05	0.06	0.04	0.03	0.08	0.09	0.01	0.15	0.22
0.26	0.13	0.08	0.15	0.15	0.11	0.06	0.05	0.07	0.07	0.01	0.10	0.12	0.02	0.15	0.26
0.28	0.16	0.07	0.18	0.17	0.11	0.05	0.03	0.06	0.06	0.03	0.12	0.14	0.05	0.16	0.29
0.47	0.38	0.09	0.16	0.19	0.13	0.05	0.04	0.05	0.04	0.06	0.12	0.18	0.08	0.10	0.28

POWER SPECTRUM OF SECOND IMAGE

8.73	0.54	0.16	0.15	0.19	0.13	0.07	0.03	0.01	0.03	0.07	0.13	0.19	0.15	0.16	0.54
0.56	0.42	0.14	0.11	0.17	0.13	0.05	0.04	0.02	0.02	0.07	0.13	0.19	0.16	0.10	0.40
0.33	0.29	0.15	0.05	0.15	0.13	0.03	0.05	0.04	0.04	0.06	0.11	0.17	0.16	0.03	0.21
0.25	0.27	0.16	0.02	0.13	0.12	0.03	0.05	0.07	0.05	0.04	0.08	0.15	0.16	0.06	0.13
0.17	0.21	0.14	0.02	0.08	0.08	0.03	0.04	0.08	0.06	0.05	0.07	0.11	0.14	0.10	0.05
0.12	0.17	0.13	0.05	0.05	0.07	0.03	0.04	0.08	0.06	0.04	0.06	0.07	0.09	0.07	0.01
0.06	0.10	0.10	0.05	0.05	0.06	0.02	0.04	0.06	0.04	0.02	0.02	0.03	0.06	0.05	0.01
0.01	0.02	0.04	0.03	0.02	0.03	0.01	0.02	0.04	0.04	0.02	0.02	0.01	0.03	0.02	0.03
0.02	0.01	0.02	0.02	0.02	0.01	0.01	0.03	0.05	0.03	0.01	0.01	0.02	0.02	0.02	0.01
0.01	0.03	0.04	0.03	0.01	0.02	0.02	0.04	0.04	0.02	0.01	0.03	0.02	0.03	0.04	0.02
0.06	0.01	0.05	0.06	0.03	0.02	0.02	0.04	0.06	0.04	0.02	0.06	0.05	0.05	0.10	0.10
0.12	0.01	0.07	0.09	0.07	0.06	0.04	0.06	0.08	0.04	0.03	0.07	0.05	0.05	0.13	0.17
0.17	0.05	0.10	0.14	0.11	0.07	0.05	0.06	0.08	0.04	0.03	0.08	0.08	0.02	0.14	0.21
0.25	0.13	0.06	0.16	0.15	0.08	0.04	0.05	0.07	0.05	0.03	0.12	0.13	0.02	0.16	0.27
0.33	0.21	0.03	0.16	0.17	0.11	0.06	0.04	0.04	0.05	0.03	0.13	0.15	0.05	0.15	0.29
0.56	0.40	0.10	0.16	0.19	0.13	0.07	0.02	0.02	0.04	0.05	0.13	0.17	0.11	0.14	0.42

FIG. 1b. Power spectra of the images in 1a. The (0,0) component is in the upper left corner, with wavenumber increasing downward and to the right.

CROSS-COVARIANCE FUNCTION

76.4	76.6	76.7	76.7	76.6	76.4	76.2	76.0	75.8	75.7	75.6	75.7	75.7	75.9	76.0	76.2
76.6	76.8	77.0	77.0	76.9	76.6	76.4	76.1	75.9	75.8	75.7	75.7	75.8	75.9	76.0	76.3
76.7	76.9	77.1	77.2	77.1	76.8	76.6	76.3	76.0	75.9	75.7	75.7	75.8	75.9	76.1	76.4
76.8	77.0	77.2	77.3	77.2	77.0	76.6	76.4	76.1	75.9	75.7	75.6	75.8	75.9	76.2	76.5
76.7	77.0	77.2	77.5	77.4	77.1	76.8	76.9	76.6	76.1	75.7	75.6	75.7	75.9	76.1	76.5
76.8	77.1	77.2	78.1	78.7	77.9	77.2	77.4	77.2	76.3	75.7	75.6	75.7	75.9	76.0	76.3
77.3	77.4	77.5	78.6	79.9	78.4	77.1	77.0	76.8	76.0	75.7	75.6	75.7	75.8	75.9	76.3
77.4	77.5	77.2	77.6	78.2	77.4	76.3	76.2	76.0	75.7	75.7	75.6	75.6	75.7	75.8	76.5
76.4	76.6	76.3	76.3	76.3	76.2	75.9	75.8	75.7	75.7	75.6	75.6	75.6	75.6	75.7	76.1
75.7	75.8	75.9	75.9	75.8	75.8	75.7	75.7	75.7	75.6	75.6	75.6	75.6	75.6	75.6	75.7
75.7	75.7	75.8	75.8	75.7	75.7	75.6	75.6	75.6	75.6	75.6	75.6	75.6	75.6	75.6	75.7
75.7	75.7	75.7	75.7	75.7	75.7	75.7	75.6	75.6	75.6	75.6	75.6	75.6	75.7	75.7	75.7
75.8	75.8	75.8	75.8	75.8	75.7	75.7	75.7	75.6	75.6	75.6	75.6	75.6	75.7	75.7	75.7
75.8	75.9	75.9	76.0	75.9	75.8	75.7	75.7	75.7	75.6	75.6	75.6	75.7	75.7	75.7	75.8
76.0	76.1	76.1	76.1	76.1	76.0	75.8	75.7	75.7	75.6	75.6	75.6	75.7	75.7	75.8	75.9
76.2	76.3	76.4	76.4	76.3	76.2	76.0	75.9	75.7	75.6	75.6	75.7	75.7	75.8	75.9	76.1

FIG. 1c. Cross covariance of the images in 1a. Zero displacement corresponds to the upper left corner; X displacement increases to the right and Y displacement increases downward.

where u and v are frequency domain variables. Conversely, a function in the frequency domain can be transformed into the space domain by using the inverse Fourier transform

$$\mathbf{F}^{-1}[F(u,v)] \equiv \int_{-\infty}^{\infty}\int F(u,v)e^{2\pi i(ux+vy)}du\,dv \equiv f(x,y). \quad (2)$$

For application to discrete data, the Fourier transform process is redefined as

$$F(u,v) = \frac{1}{M.N}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp\left[-2\pi i\left(\frac{ux}{M}+\frac{vy}{N}\right)\right], \quad (3)$$

$$f(x,y) = \frac{1}{M.N}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)\exp\left[2\pi i\left(\frac{ux}{M}+\frac{vy}{N}\right)\right], \quad (4)$$

where M and N are the sample sizes in the x and y directions, respectively.

One indication of the relative displacement between successive pictures is the location of the peak of the cross covariance of the two images (Leese *et al.*, 1971). Letting $f(x,y)$ and $g(x,y)$ represent the brightness distribution of the two images, the cross-covariance is defined as

$$C_{fg}(x,y) = \int_{-\infty}^{\infty}\int f(\xi,\eta)g(\xi-x,\eta-y)d\xi d\eta. \quad (5)$$

The Fourier transform provides a fast way to calculate the cross-covariance function. From Eqs. (1) and (2), it follows that

$$C_{fg}(x,y) = \mathbf{F}^{-1}[F^*(u,v)G(u,v)], \quad (6)$$

where $F^*(u,v)$ is the complex conjugate of $F(u,v)$ while $G(u,v)$ is the Fourier transform of $g(x,y)$. The Fourier transform of C_{fg}, namely, F^*G is called the cross spectral density of f and g. The positions of the

PHASE ANGLES OF CROSS SPECTRAL DENSITY

0.0	-3.1	-6.2	4.3	0.0	-3.7	7.0	4.2	0.0	-4.2	-7.0	3.7	0.0	-4.3	6.2	3.1
-4.1	7.9	2.5	-1.8	-5.9	6.1	2.3	-2.0	-4.9	7.2	3.2	-2.0	-5.9	6.0	2.3	-0.8
4.5	0.2	-4.2	8.0	3.9	-0.1	-2.5	7.4	4.1	1.0	-3.5	8.0	3.9	0.2	-5.2	-7.7
-2.0	-6.0	6.1	-3.5	-2.3	-5.6	6.6	1.8	-2.0	-6.0	5.2	1.6	-2.0	-6.0	5.9	1.7
7.8	3.8	0.1	-4.8	-7.6	4.3	0.0	-4.6	8.0	4.8	-0.6	-4.8	8.0	3.8	0.0	-4.1
1.5	-2.4	-6.3	5.7	1.7	-1.9	-3.9	5.5	2.3	-1.1	-5.3	5.9	2.0	-2.1	-5.9	0.0
-4.5	8.0	4.2	-0.2	-3.0	-6.5	5.8	0.1	-4.0	8.0	3.1	-0.7	-3.1	7.8	4.2	1.3
-7.7	1.5	-1.9	-4.9	-6.4	3.5	-2.5	-6.1	6.0	2.1	-3.1	7.5	2.0	0.7	-2.1	-6.0
0.0	-4.5	-7.3	4.2	0.0	-4.4	2.7	2.9	0.0	-2.9	-7.7	4.4	0.0	-4.2	-7.3	4.5
7.7	6.0	2.1	-0.7	-2.0	-7.5	3.1	-2.1	-6.0	6.1	2.5	-3.5	6.4	4.9	1.9	-1.5
4.5	-1.3	-4.2	-7.8	3.1	0.7	-3.1	8.0	4.0	-0.1	-5.8	6.5	3.0	0.2	-4.2	8.0
-1.5	0.0	5.9	2.1	-2.0	-5.9	5.3	1.1	-2.3	-5.5	3.9	1.9	-1.7	-5.7	6.3	2.4
-7.8	4.1	0.0	-3.8	-8.0	4.8	0.6	-4.8	-8.0	4.6	0.0	-4.3	7.6	4.8	-0.1	-3.8
2.0	-1.7	-5.9	6.0	2.0	-1.6	-5.2	6.0	2.0	-1.8	-6.6	5.6	2.3	3.5	-6.1	6.0
-4.5	7.7	5.2	-0.2	-3.9	8.0	3.5	-1.0	-4.1	-7.4	2.5	0.1	-3.9	8.0	4.2	-0.2
4.1	0.8	-2.3	-6.0	5.9	2.0	-3.2	-7.2	4.9	2.0	-2.3	-6.1	5.9	1.8	-2.5	-7.9

FIG. 1d. Phase angles of the cross-spectral density of the two images. The (0,0) component is in the upper left corner, with wavenumber increasing downward and to the right. (Note that a factor of $2\pi/16$ is suppressed.)

X DISPLACEMENT ESTIMATES

3.1	3.2	5.4	4.3	3.7	5.2	2.8	4.2	4.2	2.8	5.2	3.7	4.3	5.4	3.2	3.1
4.0	5.4	4.4	4.1	4.0	3.7	4.3	2.9	3.9	4.0	5.3	3.9	4.1	3.6	3.1	3.3
4.3	4.4	3.8	4.1	3.9	2.5	6.1	3.3	3.1	4.5	4.5	4.1	3.8	5.4	2.5	3.8
4.0	3.9	-6.4	-1.1	3.2	3.8	4.8	3.9	3.9	4.8	3.6	3.6	4.0	4.1	4.2	3.7
4.0	3.7	4.9	2.7	4.1	4.3	4.6	3.4	3.2	5.4	4.2	3.2	4.2	3.8	4.1	4.2
3.9	4.0	3.9	4.0	3.7	1.9	6.6	3.2	3.4	4.3	4.7	3.9	4.1	3.8	-5.9	-1.6
3.5	3.8	4.3	2.8	3.5	3.7	5.7	4.0	4.0	4.9	3.9	2.3	5.1	3.6	3.0	5.7
6.8	3.5	3.0	1.5	6.1	6.0	3.6	3.9	3.9	5.2	5.4	5.5	1.3	2.9	3.9	1 7
4.5	4.2	3.0	4.2	4.4	-7.1	-0.3	2.9	2.9	-0.3	-7.1	4.4	4.2	3.0	4.2	4.5
1.7	3.9	2.9	1.3	5.5	5.4	5.2	3.9	3.9	3.6	6.0	6.1	1.5	3.0	3.5	6.8
5.7	3.0	3.6	5.1	2.3	3.9	4.9	4.0	4.0	5.7	3.7	3.5	2.8	4.3	3.8	3.5
-1.6	-5.9	3.8	4.1	3.9	4.7	4.3	3.4	3.2	6.6	1.9	3.7	4.0	3.9	4.0	3.9
4.2	4.1	3.8	4.2	3.2	4.2	5.4	3.2	3.4	4.6	4.3	4.1	2.7	4.9	3.7	4.0
3.7	4.2	4.1	4.0	3.6	3.6	4.8	3.9	3.9	4.8	3.8	3.2	-1.1	-6.4	3.9	4.0
3.8	2.5	5.4	3.8	4.1	4.5	4.5	3.1	3.3	6.1	2.5	3.9	4.1	3.8	4.4	4.3
3.3	3.1	3.6	4.1	3.9	5.3	4.0	3.9	2.9	4.3	3.7	4.0	4.1	4.4	5.4	4.0

Y DISPLACEMENT ESTIMATES

4.1	5.0	7.2	6.2	5.9	6.2	4.7	6.2	4.9	4.5	5.8	5.8	5.9	5.7	3.9	3.9
7.4	7.7	6.7	6.2	6.2	6.1	4.8	6.6	7.0	6.2	6.7	6.0	6.2	5.8	7.5	6.9
6.5	6.2	5.7	-4.5	6.2	5.5	6.9	5.5	6.1	6.9	7.3	6.4	5.9	6.2	4.9	6.6
6.2	6.2	6.0	1.4	5.2	6.1	6.6	6.4	6.0	5.2	5.8	6.4	6.0	6.1	5.9	5.8
6.2	6.1	6.4	5.4	6.7	6.3	3.9	5.9	5.7	5.9	4.7	5.3	6.0	5.9	5.9	-4.0
6.0	5.6	5.5	5.9	4.7	4.6	6.3	5.5	6.3	6.9	7.5	6.6	5.1	6.1	5.9	-1.3
3.2	6.5	6.1	4.7	3.4	6.0	-7.7	6.1	6.0	5.9	6.2	7.8	-5.1	7.1	6.3	7.2
-7.7	6.1	6.8	6.9	-6.4	7.9	-5.2	7.0	6.0	5.1	-0.4	3.1	2.0	5.0	5.2	5.5
-7.7	5.5	5.2	5.0	2.0	3.1	-0.4	5.1	6.0	7.0	-5.2	7.9	-6.4	6.9	6.8	6.1
3.2	7.2	6.3	7.1	-5.1	7.8	6.2	5.9	6.0	6.1	-7.7	6.0	3.4	4.7	6.1	6.5
6.0	-1.3	5.9	6.1	5.1	6.6	7.5	6.9	6.3	5.5	6.3	4.6	4.7	5.9	5.5	5.6
6.2	-4.0	5.9	5.9	6.0	5.3	4.7	5.9	5.7	5.9	3.9	6.3	6.7	5.4	6.4	6.1
6.2	5.8	5.9	6.1	6.0	6.4	5.8	5.2	6.0	6.4	6.6	6.1	5.2	1.4	6.0	6.2
6.5	6.6	4.9	6.2	5.9	6.4	7.3	6.9	6.1	5.5	6.9	5.5	6.2	-4.5	5.7	6.2
7.4	6.9	7.5	5.8	6.2	6.0	6.7	6.2	7.0	6.6	4.8	6.1	6.2	6.2	6.7	7.7
4.1	3.9	3.9	5.7	5.9	5.8	5.8	4.5	4.9	6.2	4.7	6.2	5.9	6.2	7.2	5.0

FIG. 1e. X and Y displacement estimates derived from the cross-spectral density.

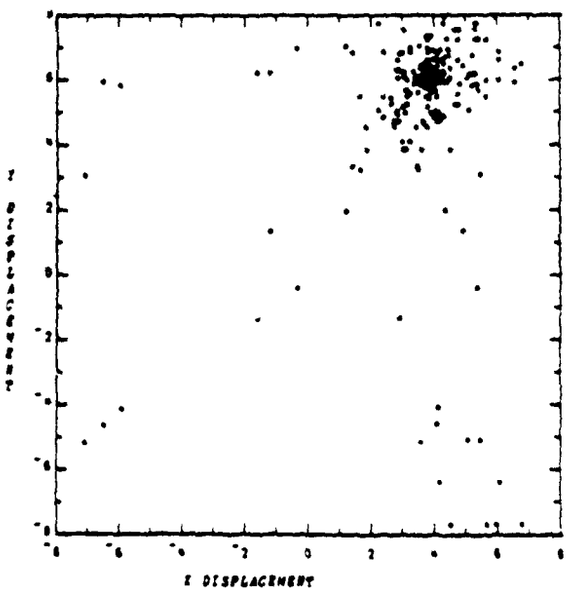


FIG. 1f. Scatter plot of the estimates in 1e.

maxima of the cross-covariance function determine the relative displacement between the two pictures. The ratio of the computer times required for the Fourier transform method and the conventional method of calculating the cross covariance is approximately $(N/2)^N$ (Brigham and Morrow, 1967), where N is the linear size of the array.

If the discrete version of the Fourier transform is used to compute the cross covariance, as in (6), the resulting function is *cyclic*, which is equivalent to assuming that the functions in the space domain are periodic. In other words, the Fourier transform method assumes that when movement of objects takes place, what moves out of the picture at one boundary must enter at the opposite boundary. Further examination and discussion of this assumption is presented in Section 3.

The Fourier transform possesses still another property which makes it possible to take a different approach to the estimation of motion. Assume that uniform translation is the only difference between the two functions, i.e., $g(x,y)=f(x-a, y-b)$. Then from

the definition of the Fourier transform in Eq. (1) it follows (Briggs, 1968) that

$$G(u,v)=F(u,v)e^{-2\pi i(ua+vb)}. \quad (7)$$

The Fourier transform at a specific frequency (u,v) can be represented by

$$F(u,v)=|F(u,v)|e^{i\varphi(u,v)}, \quad (8)$$

where $|F|$ is the amplitude and φ is the phase angle. Applying the same representation to $G(u,v)$, we have

$$G(u,v)=|F(u,v)|\exp\left[i\varphi(u,v)-2\pi i\left(\frac{ua}{M}+\frac{vb}{N}\right)\right]. \quad (9)$$

Thus,

$$F^*(u,v)G(u,v)=|F(u,v)|^2\exp\left[-2\pi i\left(\frac{ua}{M}+\frac{vb}{N}\right)\right]. \quad (10)$$

The term $F^*(u,v)G(u,v)$ is complex with amplitude $|F|^2$ and phase angle $-2\pi(ua+vb)$, for all u's and v's. The differences between the phase angles of two neighboring frequencies have the properties

$$\varphi(u,v)-\varphi(u+1,v)=\frac{2\pi}{M}a, \quad (11)$$

$$\varphi(u,v)-\varphi(u,v+1)=\frac{2\pi}{N}b. \quad (12)$$

We thus find that differences in phase angle between neighboring terms in the F^*G matrix yield x and y components of displacement. (When referring to phase angles in the examples to follow the factors $2\pi/M$ and $2\pi/N$ will be suppressed.)

If the case where the two images differ only in that one is a uniform translation of the other, the phase differences are the same between all neighboring frequencies. However, when the displacement between the two images is not uniform—and specifically when displacement is different for features with different characteristic dimensions—then the phase angle difference will be a function of frequency. In the following sections, we will examine to what extent the phase angle differences continue to yield estimates of the displacement and how these estimates are related to the size of features in the image.

3. Simulation studies

A number of simulation studies were conducted to investigate the comparative effectiveness of cross-correlation and phase difference techniques for estimating motion (Lo and Parikh, 1973; Lo *et al.*, 1974). In particular, cases were studied that involved 1) distortions such as rotation, scale change, brightness change and noise, in addition to simple translation; 2) mixtures of motion; and 3) edge effects—objects moving off one boundary without coming back in at the opposite boundary, so that the displacements are non-cyclic.

It was found that if these conditions were not too strongly present, both techniques yielded reasonable estimates of the motion (or, in mixture cases, estimates of the predominant motion). However, as the distortion, mixture or edge factors become stronger, performanc began to deteriorate. This situation is illustrated in the simulations shown in Figs.

In Fig. 1 a large mass of relatively dim is moving at a comparatively slow speed (displacement components $X=2$ and $Y=1$) relative to smaller clouds

GRAY LEVELS OF FIRST IMAGE

0	0	20	0	0	0	0	20	0	0	0	0	0	0	0	0	
0	20	20	20	0	0	20	20	20	0	0	0	0	0	0	0	
20	20	20	20	20	20	20	20	20	20	0	0	0	0	0	0	
0	20	20	20	0	0	20	20	20	0	0	0	0	0	0	0	
0	0	20	0	0	0	0	20	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	
0	0	20	20	20	20	20	0	0	0	0	0	0	0	0	0	
0	20	20	20	20	20	20	20	0	0	0	0	0	0	0	0	
0	20	20	20	20	20	20	20	0	0	0	0	0	0	0	0	
20	20	20	20	20	20	20	20	20	0	0	0	0	0	0	0	
0	20	20	20	20	20	20	20	0	0	0	0	0	0	0	0	
0	20	20	20	20	20	20	20	0	0	0	0	0	0	0	0	
0	0	20	20	20	20	20	0	0	0	0	0	0	0	0	0	
0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	

GRAY LEVELS OF SECOND IMAGE

0	0	0	0	0	0	0	0	20	0	0	0	0	20	0	0	
0	0	0	0	0	0	0	20	20	20	0	0	20	20	20	0	
0	0	0	0	0	0	20	20	20	20	20	20	20	20	20	20	
0	0	0	0	0	0	0	20	20	20	0	0	20	20	20	0	
0	0	0	0	0	0	0	0	20	0	0	0	0	20	0	0	
0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	
0	0	0	0	20	20	20	20	20	0	0	0	0	0	0	0	
0	0	0	20	20	20	20	20	20	20	0	0	0	0	0	0	
0	0	0	20	20	20	20	20	20	20	0	0	0	0	0	0	
0	0	20	20	20	20	20	20	20	20	20	0	0	0	0	0	
0	0	0	20	20	20	20	20	20	20	20	0	0	0	0	0	
0	0	0	20	20	20	20	20	20	20	0	0	0	0	0	0	
0	0	0	0	20	20	20	20	20	0	0	0	0	0	0	0	
0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

FIG. 2a. Analogous to Fig. 1, but for images simulating the motion of two sets of clouds with the same brightness but different displacements. The large mass is displaced $X=2$, $Y=2$, and the two smaller masses are displaced $X=6$, $Y=0$.

POWER SPECTRUM OF FIRST IMAGE

5.86	3.81	0.33	1.23	0.52	0.26	0.38	0.22	0.08	0.22	0.38	0.26	0.52	1.23	0.33	3.81
1.28	1.26	0.76	0.64	0.27	0.15	0.16	0.06	0.12	0.08	0.13	0.14	0.10	0.32	0.82	1.01
2.06	0.94	0.62	0.71	0.09	0.11	0.14	0.22	0.30	0.33	0.22	0.07	0.12	0.91	0.72	1.00
1.12	0.71	0.51	0.11	0.38	0.16	0.36	0.27	0.02	0.11	0.24	0.26	0.15	0.06	0.38	0.81
0.52	0.34	0.07	0.25	0.28	0.19	0.28	0.36	0.39	0.04	0.42	0.06	0.67	0.46	0.09	0.28
0.36	0.19	0.14	0.54	0.45	0.32	0.53	0.07	0.39	0.12	0.22	0.20	0.33	0.38	0.11	0.17
0.26	0.14	0.14	0.40	0.34	0.26	0.41	0.31	0.14	0.19	0.38	0.32	0.29	0.35	0.12	0.14
0.30	0.01	0.22	0.14	0.32	0.10	0.20	0.44	0.05	0.47	0.27	0.10	0.27	0.19	0.23	0.02
0.08	0.17	0.32	0.19	0.32	0.39	0.32	0.55	0.08	0.55	0.32	0.39	0.32	0.19	0.39	0.17
0.30	0.02	0.23	0.19	0.27	0.10	0.27	0.47	0.05	0.44	0.20	0.10	0.32	0.14	0.22	0.01
0.26	0.14	0.12	0.35	0.28	0.32	0.38	0.19	0.14	0.31	0.41	0.26	0.34	0.40	0.14	0.14
0.36	0.17	0.11	0.38	0.33	0.20	0.22	0.12	0.39	0.07	0.53	0.32	0.45	0.54	0.14	0.19
0.52	0.28	0.09	0.36	0.67	0.06	0.42	0.04	0.39	0.36	0.28	0.19	0.28	0.25	0.07	0.34
1.12	0.81	0.38	0.06	0.13	0.26	0.24	0.11	0.02	0.17	0.36	0.16	0.38	0.11	0.51	0.71
2.06	1.00	0.72	0.91	0.12	0.07	0.22	0.33	0.30	0.12	0.14	0.11	0.09	0.71	0.62	0.94
1.28	1.01	0.82	0.32	0.10	0.14	0.13	0.08	0.12	0.06	0.16	0.15	0.27	0.64	0.76	1.26

POWER SPECTRUM OF SECOND IMAGE

5.86	2.78	1.30	1.15	0.52	0.20	0.30	0.28	0.08	0.28	0.30	0.20	0.52	1.15	1.30	2.78
1.28	1.98	0.17	0.82	0.10	0.16	0.18	0.08	0.12	0.09	0.16	0.17	0.27	0.46	0.35	2.58
2.06	0.94	0.15	0.58	0.12	0.11	0.08	0.31	0.30	0.33	0.19	0.13	0.09	0.91	0.41	0.49
1.12	0.87	0.61	0.12	0.13	0.22	0.38	0.20	0.02	0.16	0.27	0.13	0.38	0.05	0.51	0.10
0.52	0.06	0.07	0.15	0.67	0.09	0.28	0.28	0.39	0.23	0.42	0.18	0.28	0.50	0.09	0.43
0.36	0.10	0.11	0.56	0.33	0.17	0.49	0.09	0.39	0.10	0.09	0.26	0.45	0.31	0.07	0.20
0.26	0.17	0.08	0.40	0.28	0.35	0.40	0.31	0.14	0.29	0.37	0.32	0.34	0.32	0.03	0.14
0.30	0.06	0.28	0.16	0.27	0.11	0.27	0.42	0.05	0.43	0.32	0.10	0.32	0.23	0.29	0.04
0.08	0.12	0.28	0.19	0.32	0.40	0.42	0.50	0.08	0.50	0.42	0.40	0.32	0.19	0.28	0.12
0.30	0.04	0.29	0.23	0.32	0.10	0.32	0.43	0.05	0.42	0.27	0.11	0.27	0.16	0.28	0.06
0.26	0.14	0.03	0.32	0.34	0.32	0.37	0.29	0.14	0.31	0.40	0.35	0.28	0.40	0.08	0.17
0.36	0.20	0.07	0.31	0.45	0.26	0.09	0.10	0.39	0.09	0.49	0.17	0.33	0.56	0.11	0.10
0.52	0.43	0.09	0.50	0.28	0.18	0.42	0.23	0.39	0.28	0.28	0.09	0.67	0.15	0.07	0.06
1.12	0.10	0.51	0.05	0.38	0.13	0.27	0.16	0.02	0.20	0.38	0.22	0.13	0.12	0.61	0.87
2.06	0.49	0.41	0.91	0.09	0.13	0.19	0.33	0.30	0.31	0.08	0.11	0.12	0.58	0.15	0.94
1.28	2.58	0.35	0.46	0.27	0.17	0.16	0.09	0.12	0.08	0.18	0.16	0.10	0.82	0.17	1.98

FIG. 2b. Power spectra of the images in 2a. The (0,0) component is in the upper left corner, with wavenumber increasing downward and to the right.

of higher brightness ($X=4$, $Y=6$). Fig. 1a shows the two pictures. Fig. 1b shows their power spectra and Fig. 1c shows their cross-covariance, which has a peak corresponding to the displacement (4,6) of the higher contrast clouds. Fig. 1d shows the phase angles of the Fourier transform product, and Fig. 1e shows the resulting X and Y displacement estimates from the phase difference method, obtained by taking hori-

CROSS-COVARIANCE FUNCTION

54.7	70.3	67.2	64.1	60.9	60.9	64.1	42.2	26.6	17.2	15.6	23.4	20.3	25.0	29.7	43.8
62.5	71.9	75.0	71.9	65.6	62.5	53.1	43.8	25.0	15.6	12.5	15.6	21.9	25.0	34.4	43.8
57.8	70.3	82.8	67.2	60.9	53.1	45.3	31.3	20.3	7.8	9.4	10.9	17.2	23.4	29.7	45.3
53.1	62.5	65.6	59.4	53.1	43.8	34.4	25.0	9.4	4.7	3.1	6.3	12.5	18.8	29.1	39.1
45.3	53.1	54.7	56.3	48.4	40.6	29.7	17.2	10.9	6.3	3.1	6.3	7.8	17.2	26.6	37.5
42.2	50.0	56.3	53.1	50.0	43.8	34.4	25.0	14.1	7.8	7.8	6.3	12.5	18.8	26.6	35.9
43.8	54.7	59.4	59.4	54.7	50.0	43.8	32.8	21.9	15.6	14.1	14.1	17.2	20.3	28.1	34.4
46.9	56.3	60.9	62.5	59.4	56.3	48.4	40.6	31.3	23.4	20.3	21.9	21.9	25.0	26.6	35.9
43.8	56.3	62.5	57.8	54.7	54.7	53.1	46.9	35.9	26.6	23.4	26.6	28.1	26.6	25.0	32.8
39.1	54.7	59.4	53.1	50.0	50.0	56.3	51.6	35.9	28.1	26.6	28.1	31.3	28.1	26.6	28.1
35.9	46.9	54.7	48.4	46.9	48.4	51.6	46.9	35.9	26.6	23.4	26.6	28.1	26.6	23.4	26.6
34.4	40.6	45.3	46.9	46.9	43.8	42.2	37.5	31.3	23.4	20.3	21.9	21.9	21.9	20.3	23.4
28.1	37.5	40.6	40.6	39.1	37.5	34.4	26.6	20.3	15.6	18.1	15.6	15.6	14.1	15.6	21.9
26.6	34.4	37.5	34.4	31.3	34.4	28.1	21.9	7.8	7.8	10.9	9.4	9.4	9.4	14.1	20.3
32.8	37.5	39.1	39.1	39.1	37.5	32.8	20.3	14.1	7.8	9.4	9.4	7.8	12.5	17.2	26.6
46.9	50.0	53.1	50.0	53.1	53.1	43.8	37.5	21.9	14.1	12.5	12.5	10.9	15.8	25.0	35.9

FIG. 2c. Cross-covariance of the images in 2a. Zero displacement corresponds to the upper left corner; X displacement increases to the right and Y displacement increases downward.

PHASE ANGLES OF CROSS SPECTRAL DENSITY

0.0	-2.8	-6.2	-3.1	8.0	6.0	5.6	3.6	0.0	-3.6	-5.6	-6.0	8.0	3.1	6.2	2.8
-5.0	-3.9	3.2	1.2	7.2	3.9	1.1	3.9	-2.0	-5.3	-5.0	-7.8	7.2	-0.3	-2.2	-1.3
-1.4	-6.0	6.6	-6.2	5.3	2.0	-6.4	-0.6	-4.0	-6.0	5.3	6.3	5.3	2.0	-1.8	4.6
-1.6	-7.1	5.6	2.2	-0.7	0.1	-3.7	7.6	-6.0	-5.2	3.5	3.3	-0.7	-1.5	-2.0	-0.1
2.4	6.3	4.0	7.5	3.4	-0.1	-4.0	7.9	8.0	-1.1	4.0	0.2	3.4	-0.1	-4.0	-7.7
3.2	3.0	2.5	-1.3	-7.0	-3.4	-5.2	-7.5	6.0	5.0	0.0	-0.1	-7.0	4.7	-5.9	7.6
5.7	2.4	1.6	-2.0	-7.3	-6.2	-7.8	6.0	4.0	-4.4	-0.2	-2.0	-7.3	3.0	-7.3	6.0
1.1	6.3	-2.2	-0.9	-7.0	7.9	-2.4	5.2	2.0	-6.4	2.7	-4.0	-7.0	0.1	6.3	5.3
0.0	-2.9	-3.8	-2.3	8.0	6.0	-2.1	6.3	0.0	-6.3	2.1	-6.0	8.0	2.3	3.8	2.9
-1.1	-5.3	-6.3	-0.1	7.0	4.0	-2.7	6.4	-2.0	-6.2	2.4	-7.9	7.0	0.9	2.2	-6.3
-5.7	-6.0	7.3	-3.0	7.3	2.0	0.2	4.4	-4.0	-6.0	7.8	6.2	7.3	2.0	-1.6	-2.4
-3.2	-7.6	5.9	-4.7	7.0	0.1	0.0	-5.0	-6.0	7.5	5.2	3.4	7.0	1.3	-2.5	-3.0
-2.4	7.7	4.0	0.1	-3.4	-0.2	-4.0	1.1	8.0	-7.9	4.0	0.1	-3.4	-7.5	-4.0	-6.3
1.6	0.1	2.0	1.5	0.7	-3.3	-3.5	5.2	6.0	-7.6	3.7	-0.1	0.7	-2.2	-5.6	7.1
1.4	-4.6	1.8	-2.0	-5.3	-6.3	-5.3	6.0	4.0	0.6	6.4	-2.0	-5.3	6.2	-6.6	6.0
5.0	1.3	2.2	0.3	-7.2	7.8	5.0	5.3	2.0	-3.9	-1.1	-3.9	-7.2	-1.2	-3.2	3.9

FIG. 2d. Phase angles of the cross spectral density of the two images. The (0,0) component is in the upper left corner, with wavenumber increasing downward and to the right. (Note that a factor of $2\pi/16$ is suppressed.)

zontal and vertical differences in Fig. 1d. Fig. 1f is a scatter plot of these estimates, which cluster around the displacement (4,6). [Due to the cyclic assumptions in the Fourier transform method, the displacements are determined modulo the image dimension (16) and are shown in the range -8 to $+8$.]

Fig. 2 is analogous to Fig. 1 for a pair of pictures (shown in Fig. 2a) containing two sets of clouds with

X DISPLACEMENT ESTIMATES

2.8	3.3	-3.0	4.9	2.0	0.5	2.0	3.6	3.6	2.0	0.3	2.0	4.9	-3.0	3.3	2.8
-1.2	-7.1	2.0	-6.0	3.3	2.8	-2.8	5.9	3.3	-0.3	2.9	0.9	7.6	1.9	-1.0	3.8
4.6	3.4	-3.2	4.4	3.3	-7.6	-5.8	3.4	2.0	4.7	-0.9	0.9	3.3	3.8	-6.4	5.9
5.5	3.3	3.4	2.8	-0.8	3.8	4.7	-2.4	-0.8	7.3	0.2	4.0	0.9	0.4	-1.8	1.5
-3.9	2.3	-3.5	4.1	3.6	3.9	4.1	-0.1	-6.9	-5.1	3.8	-3.2	2.5	3.9	3.7	5.9
0.2	0.5	3.9	5.7	-3.6	1.8	2.4	2.5	1.0	5.0	0.1	6.9	4.3	-5.5	2.5	4.3
3.3	0.8	3.6	5.3	-1.1	1.6	2.2	2.0	-7.6	-4.2	1.8	5.3	5.7	-5.7	2.7	0.3
-5.2	-7.4	-1.3	6.0	1.1	-5.7	7.5	4.2	-7.6	7.0	6.6	3.0	-7.0	-5.2	1.0	4.2
2.9	0.9	-1.6	5.7	2.0	-7.9	7.6	6.3	6.3	7.6	-7.9	2.0	5.7	-1.6	0.9	2.9
4.2	1.0	-6.2	-7.0	3.0	6.6	7.0	-7.6	-4.2	7.5	-5.7	1.1	6.0	-1.3	-7.4	-5.2
0.3	2.7	-5.7	5.7	5.3	1.8	-4.2	-7.6	2.0	2.2	1.6	-1.1	5.3	3.6	0.8	3.3
4.3	2.5	-5.5	4.3	6.9	0.1	5.0	1.0	2.5	2.4	1.8	-3.6	5.7	3.9	0.5	0.2
5.9	3.7	3.9	3.5	-3.2	3.8	-5.1	-6.9	-0.1	4.1	3.9	3.6	4.1	-3.5	2.3	-3.9
1.5	-1.8	0.4	0.9	4.0	0.2	7.3	-0.8	-2.4	5.7	3.8	-0.8	2.8	3.4	3.3	5.5
5.9	-6.4	3.8	3.3	0.9	-0.9	4.7	2.0	3.4	-5.8	-7.6	3.3	4.4	-3.2	3.4	4.6
3.8	-1.0	1.9	7.6	0.9	2.9	-0.3	3.3	5.9	-2.8	2.8	3.3	-6.0	2.0	-7.1	-1.2

Y DISPLACEMENT ESTIMATES

5.0	1.0	6.6	-4.4	0.8	2.1	4.5	-0.4	2.0	1.7	-0.6	1.8	0.8	3.5	-7.6	4.1
-3.7	2.1	-3.4	7.5	1.9	1.9	7.5	4.6	2.0	0.7	5.7	1.9	1.9	-2.3	-0.4	-5.8
0.2	1.1	1.0	7.6	6.0	1.9	-2.8	7.8	2.0	-0.8	1.8	2.9	6.0	3.5	0.1	4.7
-4.0	2.7	1.6	-5.3	-4.1	0.3	0.3	-0.3	2.0	-4.1	-0.5	3.1	-4.1	-1.4	2.0	7.6
-0.9	3.2	1.5	-7.2	-5.6	3.3	1.2	-0.5	2.0	-6.1	4.0	0.4	-5.6	-4.7	1.9	0.7
-2.4	0.7	1.0	0.7	0.3	2.7	2.6	2.5	2.0	-6.7	0.2	1.9	0.3	1.6	1.4	1.6
4.6	-3.9	3.8	-1.1	-0.3	1.9	-5.4	-0.2	2.0	2.0	-2.9	2.0	-0.3	3.0	2.4	0.7
1.1	-6.8	1.6	1.4	1.0	2.0	-0.3	-0.1	2.0	-0.1	0.6	2.0	1.0	-2.2	2.5	2.3
1.1	2.3	2.5	-2.2	1.0	2.0	0.6	-0.1	2.0	-0.1	-0.3	2.0	1.0	1.4	1.6	-6.8
4.6	0.7	2.4	3.0	-0.3	2.0	-2.9	2.0	2.0	-0.2	-5.4	1.9	-0.3	-1.1	3.8	-3.9
-2.4	1.6	1.4	1.6	0.3	1.9	0.2	-6.7	2.0	2.5	2.6	2.7	0.3	0.7	1.0	0.7
-0.9	0.7	1.9	-4.7	-5.6	0.4	4.0	-6.1	2.0	-0.5	1.2	3.3	-5.6	-7.2	1.5	3.2
-4.0	7.6	2.0	-1.4	-4.1	3.1	-0.5	-4.1	2.0	-0.3	0.3	0.3	-4.1	-5.3	1.6	2.7
0.2	4.7	0.1	3.5	6.0	2.9	1.8	-0.8	2.0	7.8	-2.8	1.9	6.0	7.6	1.0	1.1
-3.7	-5.8	-0.4	-2.3	1.9	1.9	5.7	0.7	2.0	4.6	7.5	1.9	1.9	7.5	-3.4	2.1
5.0	4.1	-7.6	3.5	0.8	1.8	-0.6	1.7	2.0	-0.4	4.5	2.1	0.8	-4.4	6.6	1.0

FIG. 2e. X and Y displacement estimates derived from the cross spectral density.

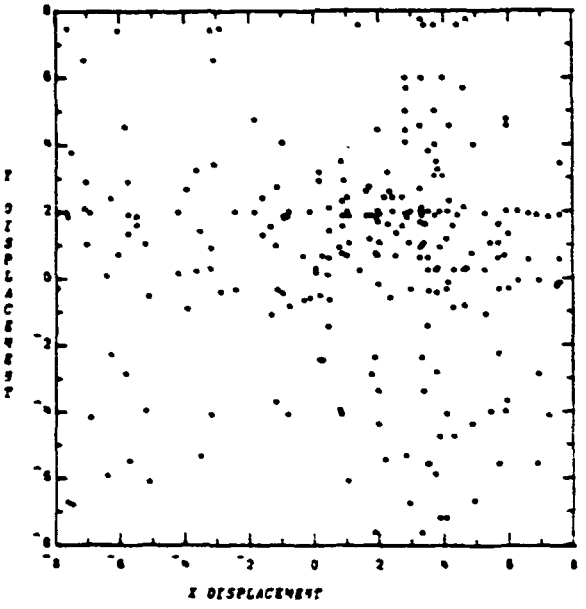


FIG. 2f. Scatter plot of the estimates in 2e.

the same brightness but different velocities, represented by the displacements (6,0) and (2,2). Here the cross covariance (Fig. 2c) corresponds to the displacement (2,2) of the larger cloud mass; but the phase difference estimates (Fig. 2e) show little tendency to cluster (Fig. 2f).

It would seem from these examples that the phase difference estimation approach is more sensitive to the presence of mixtures of motions than the cross-covariance approach. Similar results can be obtained by studying simulations involving edge effects and other distortions. The details can be found in the two references cited at the beginning of the section.

4. Experiments with real data

Several experiments with pairs of real cloud windows were also carried out, as described by Lo *et al.* (1974). The windows used were selected from ATS-1 geosynchronous satellite images, taken at intervals of 47 min. Each window is 64 by 64 pixels, and has been scaled to the gray level range 0 to 63.

The two windows shown in Fig. 3a contain relatively small clouds. In order to better evaluate the synoptic environment in which these cloud patterns are embedded, two larger windows (256 by 256 pixels) containing the smaller windows at their centers are shown in Fig. 3b. It is evident that the clouds in this scene tend to dissipate and reform relatively fast, so that considerable distortion has occurred between the frames. Nevertheless, many of the cloud patterns in the first window are easily recognizable in the second. Based on such patterns, a hand estimate of the relative displacement was made using computer printouts of the pictures so that the displacement could be accurately measured. The estimated displacement, averaged over several cloud elements, was $X=-5$, $Y=-3$.

The phase difference method applied to the pair of windows in Fig. 3a gave rise to a displacement estimate of $(-12, -4)$. This estimate was obtained as follows: X and Y displacement estimates were obtained for each spatial frequency component. The means and standard deviations of these estimates were computed, and estimates deviating from the mean by more than one standard deviation were discarded. The mean was then recomputed for the surviving estimates, and was used as the final estimate.

The cross-covariance method, on the other hand, yielded a displacement estimate of $(-6, -3)$, as determined by the position of the cross-covariance peak. In this example, therefore, the cross-covariance method yields better results (assuming the hand estimate to be most reliable).

An attempt was made to improve the results obtained from the phase difference method by applying it to selected bands of spatial frequencies, rather than to all frequencies. The following results were obtained:

Band	Frequency range	X	Y
a	$0 \leqslant (u^2+v^2)^{\frac{1}{2}} \leqslant 5$	−8	2
b	$5 \leqslant (u^2+v^2)^{\frac{1}{2}} \leqslant 10$	−13	16
c	$10 \leqslant (u^2+v^2)^{\frac{1}{2}} \leqslant 64$	−12	−4

Note that the last result is the same as that obtained when all frequencies were used, and that those results still do not agree very well with the hand estimate. The cross-covariance estimates were also reevaluated using only selected spatial frequencies (i.e., the cross spectral density F^*G was computed, the frequencies outside the selected band were suppressed, and the inverse Fourier transform of the result was computed to yield a "filtered" cross covariance). These results were as follows:

Frequency Band	X	Y
a	−6	−3.5
b	−5	−3
c	No clearcut peak	

Results from bands a and b are quite close to the result $(-6, -3)$ obtained when no filtering was used, and also agree well with the hand estimate $(-5, -3)$.

5. Discussion and conclusions

The results of the experiments reported here indicate that both the cross-covariance and phase difference methods can be good estimators of motion when the objects being tracked do not change their shape, size and orientation to more than a limited degree. These techniques are less effective when a mixture of motions

(a)

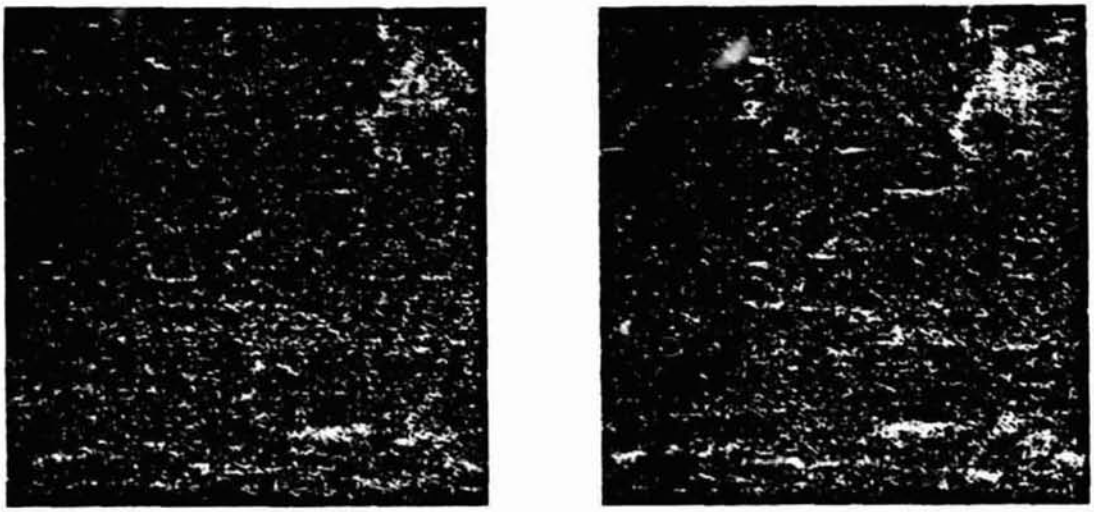


FIG. 3. A pair of images selected as an example from the ATS-1 satellite. (a) The 64×64 pixel portion of each image from which cloud motion is to be determined. (b) The surrounding area (256 by 256 pixels) in which the images in 3a are centrally imbedded.

exists, unless one of the motions is strongly dominant. These properties indicate that the Fourier transform phase difference estimation methods should be reliable in problems such as landmark matching, where the features do not change appreciably as viewed in satellite photographs. In the atmosphere, where the clouds grow, dissipate, rotate and move relative to each other, the estimation of these motions using Fourier transform methods is not reliable.

The main obstacle to accurate cloud motion estimation is believed to be the presence of mixtures of motions. One possible way to solve this problem is to use spatial frequency filtering to separate the motions of different types of cloud features, since cloud size is closely related on the one hand to spatial frequency and on the other to cloud motion. The effort to separate clouds by their sizes through frequency filtering is, however, thwarted by the fact that for real pictures each component in the frequency domain is a complicated function of the features in the space domain. Cloud patterns are not perfect sinusoidal functions. More than one frequency component is always needed to represent a cloud element, and conversely a high-frequency component may receive contributions from small cloud elements, and also from sharp corners of a large cloud elements. To further complicate the matter, the small cloud element or sharp corner may not exist in the second picture with which the first picture is being correlated. It is not found possible to effectively separate cloud types by spatial frequency filtering.

Another possible approach to separating cloud types for motion estimation purposes is to segment the cloud cover windows using thresholding or pattern classification techniques (Lo and Mohr, 1974; Lo, 1975). The brightness, the equivalent blackbody temperature in various regions of the infrared spectrum and the size of clouds are all closely related to the altitude of the clouds. Since clouds at certain altitudes in a limited region tend to move at the same velocity, thresholding techniques could be designed to separate clouds according to their brightness and or spectral equivalent black body temperatures. Picture analysis techniques which determine cloud size should also be investigated for the separation of clouds.

It can be concluded from this study that a tech-

nique which separates clouds according to their motions must be designed and utilized to pre-process a picture *before* applying Fourier transform motion estimation techniques such as the cross-covariance and phase difference methods. In the absence of such separation, the cross-covariance method yields a more reliable estimate of motion than the phase difference method.

Acknowledgments. This research was supported by NASA under Grant NGR-21-002-378. The authors are indebted to JoAnn Parikh and Jeffrey Mohr for the programming effort involved in this study, and the help of Mrs. Shelly Rowe in preparing this paper.

REFERENCES

Briggs, B. H., 1968: On the analysis of moving patterns in geophysics—I. Correlation analysis. *J. Atmos. Terr. Phys.*, **30**, 1777–1788.

Brigham, E. O., and R. E. Morrow, 1967: The fast Fourier transform. *IEEE Spectrum*, **4**, No. 12, 63–70.

Fujita, T., K. Watanabe and T. Izawa, 1969: Formation and structure of equatorial anticyclones caused by large-scale cross-equatorial flows determined by ATS-I photographs. *J. Appl. Meteor.*, **8**, 649–667.

Hubert, L. F., and L. F. Whitney, Jr., 1971: Wind estimation from geostationary satellite pictures. *Mon. Wea. Rev.*, **99**, 665–672.

Leese, J. A., and E. S. Epstein, 1963: Application of two-dimensional spectral analysis to the quantification of satellite cloud photographs. *J. Appl. Meteor.*, **2**, 629–644.

——, C. S. Novak and B. B. Clark, 1971: An automated technique for obtaining cloud motion from geosynchronous satellite data using cross correlation. *J. Appl. Meteor.*, **10**, 118–132.

Lo, R. C., 1975: The application of a thresholding technique in cloud motion estimation from satellite observations. Tech. Rep. 357, Computer Science Center, University of Maryland.

——, and J. Mohr, 1974: Applications of enhancement and thresholding techniques to Fourier transform cloud motion estimation. Tech. Rep 326, Computer Science Center, University of Maryland.

——, and J. A. Parikh, 1973: A study of the application of Fourier transforms to cloud movement estimation from satellite photographs. Tech. Rep. 242, Computer Science Center, University of Maryland.

——, —— and J. A. Parikh, 1974: Applications of Fourier transform methods of cloud movement estimation to simulated and satellite photographs. Tech. Rep. 292, Computer Science Center, University of Maryland.

Serebreny, S. M., R. G. Hadfield, R. M. Trudeau and E. J. Wiegman, 1969: Comparison of cloud motion vectors and rawinsonde data. Final Rep., Contract E-193-68 (NESC, ESSA), Stanford Research Institute, 53 pp.

Smith, E. A., and D. R. Phillips, 1972: Automated cloud tracking using precisely aligned digital ATS pictures. *IEEE Trans. Comput.*, **21**, 715–729.

Weinstein, F. S., 1972: An advantageous method of performing cross-correlational analysis. *Proc. IEEE*, **60**, 449–450.

Paper 113

Reprinted from JOURNAL OF THE ATMOSPHERIC SCIENCES, Vol. 36, No. 7, July 1979
American Meteorological Society
Printed in U. S. A.

A Stability Theorem for Energy-Balance Climate Models

ROBERT F. CAHALAN[1] AND GERALD R. NORTH[2]

Physics Department, University of Missouri, St. Louis 63121

(Manuscript received 29 August 1978, in final form 15 March 1979)

ABSTRACT

This paper treats the stability of steady-state solutions of some simple, latitude-dependent, energy-balance climate models. For north-south symmetric solutions of models with an ice-cap-type albedo feedback, and for the sum of horizontal transport and infrared radiation given by a linear operator, it is possible to prove a "slope-stability" theorem; i.e., if the local slope of the steady-state iceline latitude versus solar constant curve is positive (negative) the steady-state solution is stable (unstable). Certain rather weak restrictions on the albedo function and on the heat transport are required for the proof, and their physical basis is discussed in the text.

1. Introduction

The parallel study of climate models within a hierarchy of models of increasing complexity is now a well-established strategy in climate theory [for a discussion see the review of Schneider and Dickinson (1974)]. One hopes that some features will be common to all models from the simplest zero-dimensional globally averaged models to the most complex three-dimensional general circulation models. The one-dimensional Budyko-Sellers models have proven to be useful in exploring such properties. For example, the catastrophic transition to an ice-covered planet if the solar constant is lowered by a few percent was first discovered independently by Budyko (1968, 1969) and Sellers (1969) and appears to be common to a large variety of more complicated models (e.g., Temkin and Snell, 1976) ranging up to general circulation models (Manabe and Wetherald, private communication).

The purpose of this paper is to sketch the proof of a stability theorem for a large class of energy-balance climate models which include the Budyko-Sellers models. The linear stability analysis of solutions of several individual models in the class has been given previously by a number of investigators (Schneider and Gal-Chen, 1973; Held and Suarez, 1974; North, 1975a,b; Ghil, 1976; Su and Hsieh, 1976; Fredericksen, 1976; Drazin and Griffel, 1977; North, 1977). In addition, approaches to the full nonlinear stability problem have been suggested by Ghil (1976) and North *et al.* (1979). Our theorem clarifies the central result of these studies, and applies to a larger class of mean annual models than those represented above.

We shall not dwell on the well-known assumptions and limitations inherent in the Budyko-Sellers approach, since these have been adequately discussed in the recent literature. We shall start by motivating our study with a simple example in this section. In Section 2 we proceed to obtain formal solutions to the class of models being considered. In Section 3 we derive the so-called slope-stability theorem for that class.

We first consider a very simple model, variants of which have been discussed recently by several authors (Sellers, 1974; Crawfoord and Källen, 1978; Fraedrich, 1978). The model is a zero-dimensional energy-balance model (globally averaged) which may be defined by

$$C\frac{d}{dt}T_0 + I(T_0) = Q\tilde{a}(T_0), \tag{1.1}$$

where C is the heat capacity per unit area, T_0 the globally (and annually) averaged temperature, $I(T_0)$ the infrared radiation rate, Q the solar constant divided by 4, and $\tilde{a}(T_0)$ the globally (and annually) averaged co-albedo, which is presumed to be a function of T_0 because of the ice-cap albedo feedback. The steady-state solutions to (1.1) are easily obtained by setting d/dt to zero and solving the resulting algebraic relation for T_0 as a function of Q. For a given Q there are typically several roots, as illustrated by the solution curve in Fig. 1, obtained from a model with a linear infrared law and a cubic co-albedo.

[1] Present affiliation, Advanced Study Program, The National Center for Atmospheric Research, Boulder, CO 80307.

[2] Present affiliation: Laboratory for Atmospheric Sciences, Goddard Space Flight Center, Greenbelt, MD 20771.

0022-4928/79/071178-11$06.75
© 1979 American Meteorological Society

In order to examine the linear stability, let $T_0 = T_0^s + \delta(t)$, where T_0^s is a solution to the steady-state problem. Then to the first order in $\delta(t)$,

$$C\dot{\delta}(t) + (I' - Q\tilde{a}')\delta(t) = 0, \qquad (1.2)$$

where a prime denotes the derivative, evaluated at $T_0 = T_0^s$. The stability is determined by the sign of the expression in parentheses. This sign can be expressed in terms of the slope of the solution curve $Q(T_0)$, as follows. Differentiating the steady-state equation $I = Q\tilde{a}$ and substituting, Eq. (1.2) becomes

$$C\dot{\delta}(t) + \frac{dQ}{dT_0}\tilde{a}\delta(t) = 0 \qquad (1.3)$$

This last equation embodies the "slope-stability" theorem

$$\left.\begin{aligned} \frac{dQ}{dT_0} > 0 &\leftrightarrow \text{stability} \\ \frac{dQ}{dT_0} < 0 &\leftrightarrow \text{instability} \end{aligned}\right\} \qquad (1.4)$$

The theorem is easily interpreted since branches with negative slope have an apparent negative heat capacity and are therefore unphysical. Budyko (1972) advanced a heuristic argument along these lines as a stability proof, but it is not obvious that it applies to systems with spatial extension. One such example is a model of a star like the sun which has uniform temperature, is held together by its own gravity, and a heat balance is maintained by nuclear reactions in the interior and blackbody radiation at the surface. Increase of the heating rate leads to an increased radius and a *cooler* star (Nauenberg and Weisskopf, 1978). We presume such a star is stable. Counterexamples like this suggest that when possible we should construct rigorous proofs for stability. Moreover, rigorous mathematical results should help physical insight advance into still uncertain areas of climate theory.

This paper concentrates on one-dimensional energy-balance models which retain the sine of the latitude, x, as the single spatial variable. Hemispheric symmetry is assumed, so that only values of x from 0 to 1 need to be considered. The horizontal transport and infrared radiation laws are assumed to be represented by operators which are linear in the temperature field, and whose sum has an inverse with certain physically reasonable properties. Rather than assuming an explicit temperature dependence for the co-albedo as in the global model (1.1), an implicit temperature dependence is introduced through a parameter x_s, the sine of the latitude of the ice cap edge; this edge is assumed to be attached to a given isotherm. We shall consider only temperature fields having a unique x_s, and further assume that the temperature decreases northward across x_s.

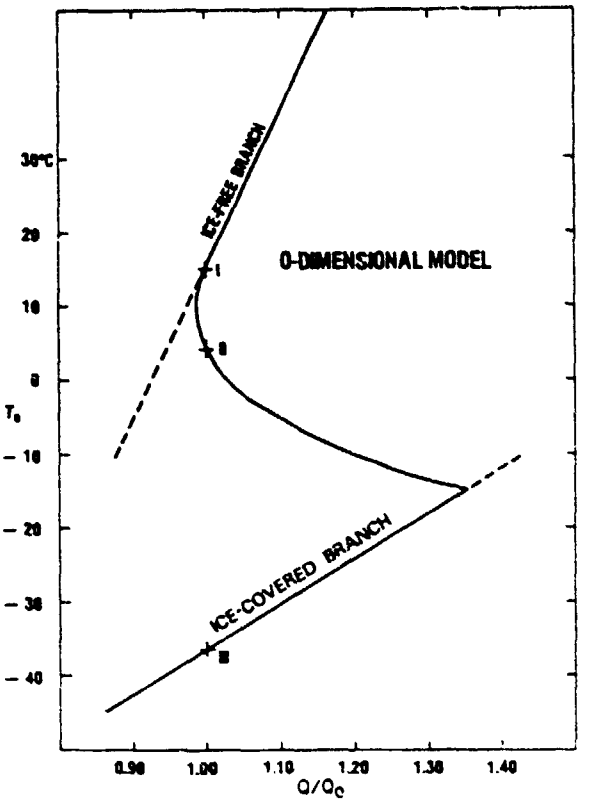


FIG. 1. Global average temperature T_0 as a function of the solar constant Q (in units of the present value Q_0), obtained from a global model having a globally and annually averaged co-albedo depending on T_0 due to ice cap-albedo feedback, as discussed in the text. The crosses indicate three solutions corresponding to the present value of the solar constant.

Just as the multiple solutions of the global model for a given solar constant are specified by the value of T_0, the different solution branches of the one-dimensional models are completely specified by the value of x_s. Fig. 2 shows an example of an exact solution to a specific model of this type (North, 1975a). A given point on the curve corresponds to a unique temperature field. We shall prove a stability theorem related to the sign of the local slope of this graph. Drazin and Griffel (1977) have found that under certain circumstances there can exist north-south unsymmetrical solutions. In such cases more than x_s and Q would be required to specify a given solution and our theorem does not hold.

The outline of the paper is as follows: Section 2 employs the formalism of linear operators and their corresponding Green's functions to construct steady-state solutions to the model equations, and to derive a transcendental equation whose roots determine the stability of these solutions. Section 3 establishes the slope-stability theorem by showing that the sign of the minimum root, which determines whether perturbations will grow in time or decay

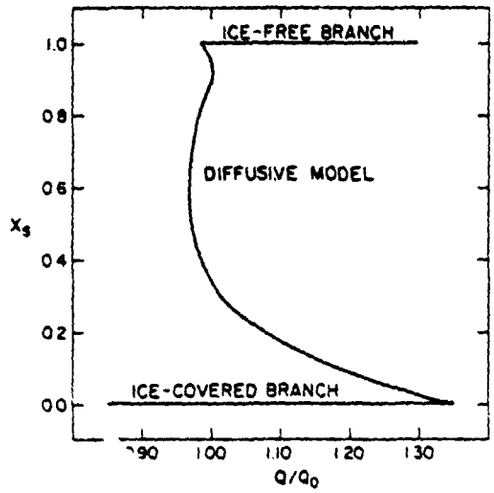


FIG. 2. A graph of the sine of the latitude of the iceline, x_s, as a function of the solar constant Q (in units of the present value Q_0) obtained from a zonal model having diffusive heat transport, as discussed by North (1975a). The co-albedo depends on the temperature field implicitly via x_s, which is attached to an isotherm. Multiple solutions correspond to a given value of the solar constant but different values of x_s.

back to equilibrium, is identical to the sign of the slope of the solution curve $x_s(Q)$ at the point which corresponds to the particular temperature field being perturbed. Also discussed in Section 3 are the assumptions for this theorem, its relation to the infinite instability of small ice caps found in some cases (indicated by the cusp near $x_s = 1$ in Fig. 2), and its relation to the formulation in terms of T_0 rather than x_s. Section 4 gives a brief discussion of results and concludes the paper. Two appendices are devoted to special cases: Appendix A treats the Green's functions of the Budyko and diffusive models, Appendix B the stability for a step-function albedo.

2. Steady-state solutions and stability eigenvalues

The class of models considered may be defined in terms of their corresponding energy balance equation

$$L[T_E](x) + f(x) = QS(x)a(x,x_s). \tag{2.1}$$

Here $T_E(x)$ is the equilibrium (sea level) temperature field, $f(x)$ a given positive function associated with the outgoing radiation rule, Q the solar constant divided by 4, $S(x)$ the mean annual normalized solar distribution reaching the top of the atmosphere, $a(x,x_s)$ the co-albedo which is a function of the sine of the latitude, x, and the sine of the latitude of the ice cap edge, x_s. L is a linear operator representing horizontal transport and the part of the infrared rule which is linear and homogeneous in $T(x)$.

We may illustrate the various terms in (2.1) with some specific examples. In the diffusive transport model of North (1975a,b)

$$L_D[T](x) = -\frac{d}{dx}D(1-x^2)\frac{d}{dx}T(x) + BT(x) \tag{2.2}$$

and $f(x) = A$ = constant. The first term here takes account of the heat transported by transient eddies, and $A + BT$ is Budyko's infrared law. A mean circulation term $v(x)\cdot\nabla T$ as in Sellers (1969) might also be included, as well as possible latitude-dependence in D, A and B. In the Budyko model (Budyko, 1969; Chylek and Coakley, 1975)

$$L[T](x) = \gamma\int_0^1 dy[\delta(x-y) - 1]T(y) + BT(x)$$
$$= \gamma(T(x) - T_0) + BT(x), \tag{2.3}$$

and again $f(x) = A$.

The co-albedo in the solar input term on the right-hand side (RHS) of (2.1) has often been assumed to be discontinuous at $x = x_s$, or as a function of temperature it is taken as discontinuous at a given iceline value T_s. These are equivalent formulations. However, there is no observational evidence for such a sharp transition. Even neglecting the effect of clouds and seasonal snow cover, we may expect that the zonal average of an iceline having large longitudinal variations such as Earth's would introduce considerable smoothing. Previous proofs of the stability theorem have been restricted to the step function co-albedo discussed in Appendix B. Unless otherwise stated we shall regard the source terms in (2.1) as smooth functions of x. As examples of smooth co-albedo functions one might generalize the step function of Appendix B to include a linear transition of width Δx_s around x_s, or simply replace it with a smooth function such as $\tanh[(x - x_s)/\Delta x_s]$. The presence of other feedback mechanisms, such as variable cloudiness, might also require some explicit temperature-dependence in the co-albedo which cannot be directly formulated in terms of latitude, and we shall exclude such processes here.

In addition to (2.1) it is necessary to specify the iceline condition

$$T_E(x_s) = T_s. \tag{2.4}$$

where T_s has a fixed value (Budyko, 1969) usually taken to be -10°C. This condition introduces nonlinearity into the model. Finally, we specify the boundary conditions such that no net heat flows across the equator (north–south symmetry), $x = 0$, or into the pole, $x = 1$. In what follows it is necessary to assume that a Green's function G_0 for L exists satisfying the boundary conditions and

$$L[G_0](x,y) = \delta(x-y). \tag{2.5}$$

where $\delta(x - y)$ is the Dirac function. The existence of G_0 limits the class of operators L, but this is not a severe restriction since any physical, linear system such as the ones we consider should have a unique response to a localized heat source.

As an example for the case of diffusive transport, $L = L_D$ given by Eq. (2.2), we may construct $G_0(x,y)$ explicitly

$$G_0^D(x,y) = \sum_{n=0}^{\infty} \frac{f_n(x)f_n(y)}{L_n}, \tag{2.6}$$

where $f_n(x)$, $n = 0, 1, 2, \ldots$, are the orthonormal eigenfunctions of L_D, and $L_n < 0$ are the corresponding discrete eigenvalues. [Ghil (1976) has shown that standard Sturm-Liouville results apply here.] For D constant, the eigenfunctions are just the even-indexed, normalized Legendre polynomials, $(2l + 1)^{1/2}P_l(x)$, and $L_l = Dl(l + 1) + B$ with $l = 2n$. In this latter case a closed form can be found for G_0^D in terms of hypergeometric functions. The spectral form (2.6) is discussed in Appendix A along with its counterpart for the Budyko model.

Returning to the general case, the equilibrium temperature field is given by the nonlinear integral equation

$$T_E(x) = \int_0^1 dy G_0(x,y)[QS(y)a(y,x_s) - f(y)], \tag{2.7}$$

where $T_E(x)$ enters the integrand through x_s.

Evaluating (2.7) at $x = x_s$ yields

$$T_s = \int_0^1 dy G_0(x_s,y)[QS(y)a(y,x_s) - f(y)]. \tag{2.8}$$

For a given x_s, Eq. (2.8) states that Q is determined; in fact, we may solve for it and obtain

$$Q = Q(x_s) = \frac{T_s + \int_0^1 dy G_0(x_s,y)f(y)}{\int_0^1 dy G_0(x_s,y)S(y)a(y,x_s)}. \tag{2.9}$$

From Eq. (2.9) we have the desired relationship between x_s and Q, so that we may plot a graph like that in Fig. 2. Given values x_s and Q, we may compute the unique temperature field $T_E(x)$ corresponding to them from (2.7). Hence, the steady-state problem is formally solved.

For a fixed value of the solar constant, we consider now a small perturbation of the temperature field about the steady-state solution, and apply the standard linear stability technique. The variation in temperature will also produce a variation in ice cap size through the ice-albedo feedback mechanism (provided $0 < x_s < 1$; special cases $x_s = 0, 1$ will be treated separately). Thus in addition to

$$T(x,t) = T_E(x) + \delta T(x,t), \tag{2.10}$$

we must also have

$$x_s(t) = x_s + \delta x_s(t), \tag{2.11}$$

where the ice cap variation δx_s may be determined in terms of δT by expanding the iceline condition to first order in the small quantities. To first order in δx_s and δT we have the temperature at the perturbed iceline

$$\begin{aligned} T(x_s + \delta x_s,t) &= T(x_s,t) + \left(\frac{\partial T}{\partial x}\right)_{x_s} \delta x_s \\ &= T_E(x_s) + \delta T(x_s,t) + T_E' \delta x_s, \end{aligned}$$

where T_E' is $(dT_E/dx)_{x=x_s}$. Due to the iceline conditions $T(x_s + \delta x_s, t) = T_s = T_E(x_s)$, the lowest order terms cancel, and we have

$$\delta x_s = \delta T(x_s,t)/(-T_E'). \tag{2.12}$$

Ice cap models for which x_s may be defined must have $T_E(x)$ decrease as we cross to the north of the iceline so that $T_E' < 0$. Thus Eq. (2.12) says that a positive variation in temperature away from equilibrium causes the ice cap to shrink, and the amount of shrinkage is inversely proportional to the temperature drop across the equilibrium iceline. The dependence on the temperature drop becomes clear if we picture the neighborhood of the iceline having a linear falloff in T_E and constant δT. Thus in order for x_s to follow an isotherm, we need $\delta T/\delta x_s = |\text{slope}|$. Note that in writing Eq. (2.12) we have implicitly assumed that T_E' exists, which may not hold when the source terms have a discontinuity and L contains integral operators as in the Budyko case. No discontinuities in temperature arise in Sturm-Liouville type problems, but in other cases it may be necessary to smooth the source terms.

The allowed temperature variations are determined by the requirement that the time-dependent energy balance equation must be satisfied to first order in the small quantities. Expanding

$$\begin{aligned} C\frac{\partial T}{\partial t}(x,t) + L[T](x,t) \\ = QS(x)a(x, x_s + \delta x_s) - f(x) \end{aligned}$$

to first order in δx_s and δT and eliminating T_E by employing Eq. (2.1) leads to

$$\begin{aligned} C\frac{\partial}{\partial t}\delta T(x,t) + L[\delta T](x,t) \\ = QS(x)a_2(x,x_s)\delta x_s, \end{aligned} \tag{2.13}$$

where a_2 indicates $\partial a(x,x_s)/\partial x_s$. If we substitute Eq. (2.12) for the iceline shift, we have a linear equation in $\delta T(x,t)$ which has solutions of the form

$$\delta T(x,t) = \delta T(x)e^{-\lambda t/C}. \tag{2.14}$$

Substituting (2.14) into Eq. (2.13) and cancelling exponentials, we obtain for $\delta T(x)$ the equation

$$(L - \lambda)[\delta T](x) = QS(x)a_2(x,x_s)\delta T(x_s)/(-T_E'). \quad (2.15)$$

The possible values of the parameter λ, the "stability eigenvalues" λ_n of this equation, determine the stability of the equilibrium solution $T_E(x)$ under a perturbation given by the associated eigenfunction $\delta T_n(x)$. If any of these perturbations (modes) have an eigenvalue with a negative real part, the equilibrium state is unstable. Equilibrium states having all $\mathrm{Re}\lambda_n > 0$ return to equilibrium with a decay time $\sim C/\mathrm{Re}\lambda_{\min}$.

Eq. (2.15) is an important tool in the study of linear stability and has been previously analyzed for a number of simple climate models. [See for example the appendix of Drazin and Griffel (1977) where the stability eigenvalue equation for the diffusion model with step-function albedo is given by Eq. (A6).] Our interest here, however, is not in determining the stability of a given solution to a given model. Instead we wish to relate the stability of such a solution to the sign of the slope of the steady-state iceline function at the point corresponding to that solution.

We proceed by first eliminating the eigenfunctions from (2.14) in order to obtain a scalar transcendental equation whose roots give the stability eigenvalues. In the next section we then relate these roots to properties of the steady-state solution whose stability is being tested.

We may define a Green's function G_λ which is a generalization of G_0 in Eq. (2.5) and satisfies

$$(L - \lambda)[G_\lambda](x,y) = \delta(x - y), \quad (2.16)$$

along with the boundary conditions at $x = 0, 1$. We may think of G_λ as a function of the continuous parameter λ except at the eigenvalues of L where it is not defined. For the diffusive case G_λ may be constructed explicitly:

$$G_\lambda^D(x,y) = \sum_{n=0}^{\infty} \frac{f_n(x)f_n(y)}{L_n - \lambda}, \quad (2.17)$$

where we have used the notation of Eq. (2.6) (cf. also Appendix A). From Eq. (2.17) one sees directly that G_λ^D is singular at $\lambda = L_n$, $n = 0, 1, 2, \ldots$. In fact, in these examples it is a meromorphic function of λ with simple poles at L_n.

Now the stability equation (2.15) may be rewritten in the form

$$\delta T(x) = \int_0^1 dy G_\lambda(x,y)[QS(y)a_2(y,x_s)\delta T(x_s)/(-T_E')], \quad (2.18)$$

where the boundary conditions are incorporated into G_λ. By evaluating this expression at $x = x_s$ we may cancel $\delta T(x_s)$ and obtain a formula for T_E', i.e.,

$$-T_E' = \int_0^1 dy G_\lambda(x_s,y)[QS(y)a_2(y,x_s)]. \quad (2.19)$$

On the other hand, if we differentiate (2.7) and evaluate the result at $x = x_s$, we obtain an equivalent expression for the slope of T_E at the iceline:

$$\left(\frac{-dT_E}{dx}\right)_{x_s} = -T_E' = \int_0^1 dy \partial G_0(x_s,y)/\partial x_s \times [QS(y)a(y,x_s) - f(y)]. \quad (2.20)$$

Similarly we may differentiate (2.8) with respect to x_s, noting that the first term on the RHS after differentiation is just given by (2.20). Since T_s is constant we obtain, after substitution,

$$-T_E' = \int_0^1 dy G_0(x_s,y) \frac{\partial}{\partial x_s}[QS(y)a(y,x_s)]. \quad (2.21)$$

In the differentiation on the RHS of (2.21) we must allow Q to depend on x_s. Substituting the result into Eq. (2.19) and combining the terms involving $a_2(y,x_s)$ yields

$$\frac{K_s}{Q}\frac{dQ}{dx_s} = \int_0^1 dy[G_\lambda(x_s,y) - G_0(x_s,y)]S(y)a_2(y,x_s), \quad (2.22)$$

where

$$K_s = \int_0^1 dy G_0(x_s,y)S(y)a(y,x_s). \quad (2.23)$$

If we subtract Eq. (2.5) from (2.16) and use the linearity property, we obtain

$$L[G_\lambda - G_0](x,y) = \lambda G_\lambda(x,y), \quad (2.24)$$

so that the term appearing in square brackets may be rewritten as

$$G_\lambda(x,y) - G_0(x,y) = \lambda \int_0^1 dz G_0(x,z)G_\lambda(z,y). \quad (2.25)$$

Substituting this expression into Eq. (2.22) gives the final form of the eigenvalue equation as

$$\frac{K_s}{Q}\frac{dQ}{dx_s} = \lambda F(\lambda), \quad (2.26)$$

where the roots λ_j are determined by the properties of the function $F(\lambda)$, defined by

$$F(\lambda) = \int_0^1 dy \int_0^1 dz G_0(x_s,z) G_\lambda(z,y)S(y)a_2(y,x_s). \quad (2.27)$$

$F(\lambda)$ depends on λ through Green's function G_λ defined by Eq. (2.16). For a given climatic state $T_E(x; Q)$, represented by a single point (x_s,Q) on an iceline curve such as in Fig. 2, the left-hand side of Eq. (2.26) will be fixed. Thus (2.26) is a transcendental equation for λ; the climate $T_E(x; Q)$ will be stable only when there are no roots λ of (2.26) having

a negative real part. In the next section we shall argue that for physically reasonable models, the sign of the real part of the smallest value of λ is identical to the sign of the iceline slope, dQ/dx_s.

3. Slope-stability theorem

In the previous section we derived a transcendental equation whose roots λ_j are the stability eigenvalues. If the lowest eigenvalue has a positive (negative) real part, the system is stable (unstable). In this section we shall argue that for a certain class of models the slope-stability theorem

$$\left.\begin{aligned} \frac{dQ}{dx_s} > 0 &\leftrightarrow \text{stability} \\ \frac{dQ}{dx_s} < 0 &\leftrightarrow \text{instability} \end{aligned}\right\} \qquad (3.1)$$

follows from the properties of the eigenvalue equation (2.26). First, we state the main assumptions which limit the class of models considered. Aside from the assumption of north-south symmetric solutions where the single index x_s may be used [implicit is also the assumption that $T_E(x)$ is decreasing at $x = x_s$], we impose the following conditions:

(i) The feedback is of the ice-cap type. Specifically $a_2(x,x_s) \geq 0$ for all x, x_s between 0 and 1.
(ii) The Green's function $G_0(x,y)$ is positive.
(iii) The generalized $G_\lambda(x,y)$ is positive for negative real values of λ with the asymptotic behavior $\delta(x-y)/(-\lambda)$ as $\lambda \to -\infty$.

The first assumption is obvious but does eliminate certain cloud band feedback mechanisms from consideration.

To clarify condition (ii), we first note that when an arbitrary heat source $\rho(y)$ is introduced it produces a temperature distribution $T(x) = \int dy \times G_0(x,y)\rho(y)$. If in addition $\rho(y)$ is increased by adding heat at a rate q at latitude x_0, so that $\Delta\rho(y) = q\delta(y - x_0)$, then the temperature distribution changes by $\Delta T(x) = qG_0(x,x_0)$. Thus saying that G_0 is positive is equivalent to the statement that heat added at one latitude will not lead to a decrease in temperature at any other latitude. Fig. 3 shows an example of the behavior of the Green's function corresponding to L_D, Eqs. (2.2) and (2.6) with D constant. This condition holds for L of Sturm-Liouville type, and a proof of this is sketched in Appendix A. We suspect that it holds for any physically reasonable L including some integral operators such as Budyko's (cf. Appendix A). Similar conditions have been applied to other stability problems, e.g., Joseph (1976).

The increasing localization of G_0 in Fig. 3 for decreasing values of D (or increasing B) is related to condition (iii). For the Sturm-Liouville type systems

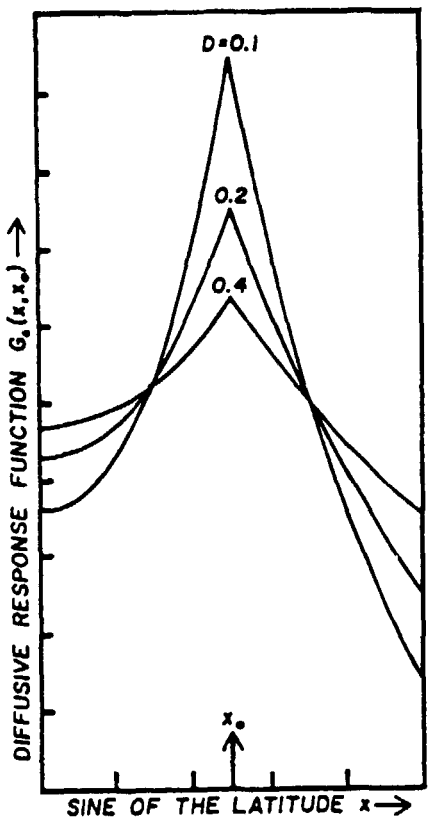


FIG. 3. The diffusive response function $G_0(x,x_0)$ given in Eq. (2.6) of the text, computed with $B = 1$ and various (constant) values of D. The area under the curve is independent of D (units = 0.2).

the expression $\delta(x-y)/(-\lambda)$ corresponds to keeping the first term in an asymptotic expansion for large negative λ and using the completeness relation $\sum_n f_n(x)f_n(y) = \delta(x-y)$. Adding this relation divided by $-\lambda$ to (2.17) leads to

$$G_\lambda^D(x,y) = -\delta(x-y)/\lambda + \sum_n f_n(x)f_n(y)\frac{L_n}{(L_n-\lambda)\lambda}; \qquad (3.2)$$

the second term on the RHS is equivalent to $L[G_\lambda^D](x,y)/\lambda$, which can also be obtained formally from rearranging (2.16). Although more singular at $x = y$ than the first term, since it contains derivatives of delta functions, it is smaller by a factor of order $1/\lambda$. This suggests that if $G_\lambda^D(x,y)$ only appears as a factor in an integrand multiplied by sufficiently well-behaved functions, (iii) will hold. It is possible to exhibit this asymptotic behavior of L explicitly in certain specific cases.

This ends the discussion of the conditions (i), (ii), and (iii) which are sufficient for our proof to be valid. The theorem may hold also for a weakened form of these conditions in certain cases. In what

follows the reader may wish to refer to Appendix B where a special case is discussed.

We now return to the discussion of the stability eigenvalues. From (ii) we clearly have

$$K_s > 0, \tag{3.3}$$

since each factor in the integrand of (2.23) is positive.

The behavior of $F(\lambda)$ may be determined from that of G_λ. According to Eq. (2.25) $G_\lambda \to G_0$ as $\lambda \to 0$, and for small positive values of λ we have

$$G_\lambda(x,y) \approx G_0(x,y) + \lambda G_0^{(2)}(x,y) > 0; \tag{3.4}$$

the "iterated kernel," given by

$$G_0^{(2)}(z,y) = \int dz' G_0(z,z') G_0(z',y), \tag{3.5}$$

represents the first term in a series expansion of G_λ in powers of λ. Eq. (2.25) is in fact an *inhomogeneous Fredholm equation of the second kind*, for G_λ. Its solution can be represented by such a power series in λ, a *Neumann series*, which converges as long as $|\lambda| < L_0$, where L_0 is the first eigenvalue of L. The coefficient of λ^n in the Neumann series is the nth iterate of the kernel G_0. Hence each term is positive for positive λ. Thus, G_λ is positive for $-\infty < \lambda < 0$ because of assumption (iii) and for $0 \leq \lambda < L_0$ by the argument above; as $\lambda \to L_0$, $G_\lambda \to +\infty$. Furthermore, if we compute the derivative with respect to λ of Eq. (2.16) the solution of the resulting equation may be written in the form

$$dG_\lambda(x,y)/d\lambda = \int dz G_\lambda(x,z) G_\lambda(z,y), \tag{3.6}$$

which again is positive in this interval so that G_λ is strictly monotonic for $\lambda < L_0$. This is sufficient information to conclude that as λ decreases through zero, the stability function $\lambda F(\lambda)$ will also decrease through zero, and will monotonically approach its asymptotic value, determined by the behavior of G_λ to be

$$\lambda F(\lambda) \xrightarrow[\lambda \to -\infty]{} -\int_0^1 dy G_0(x_s,y) S(y) a_2(y,x_s). \tag{3.7}$$

We may now establish the slope-stability theorem by considering the possible values of λ determined by the relation

$$\frac{K_s}{Q}\frac{dQ}{dx_s} = \lambda F(\lambda), \tag{2.26}$$

for a fixed point on the equilibrium iceline curve $x_s = x_s(Q)$. The graphical solution of this transcendental equation is illustrated schematically in Fig. 4. The constants L_1, L_2, etc., are the higher eigenvalues of L, but for now we concentrate on the region $\lambda < L_0$. When the iceline slope is positive, indicated by the solid horizontal line in Fig. 4, the smallest possible stability eigenvalue λ_0 is also positive, and determined by the intersection of the horizontal line with $\lambda F(\lambda)$, computed from the Neumann series for G_λ. Perturbations in this region decay to zero in a characteristic time C/λ_0. This is the situation in the positive slope branches in Fig. 2. As we approach one of the critical points indicated in Fig. 2, dQ/dx_s as well as λ_0 approach zero and we have a situation of neutral stability. Finally, in the unstable regions

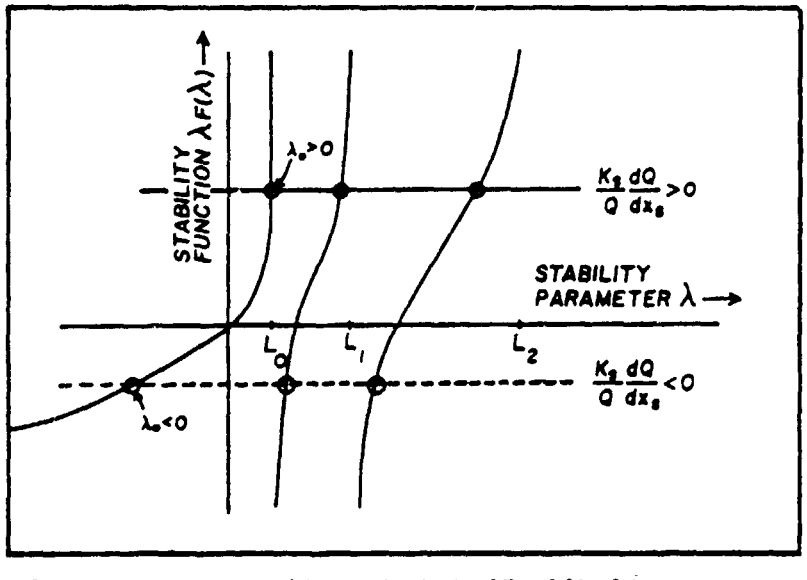


FIG. 4. A schematic graph of the right-hand-side of Eq. (2.26) of the text versus the stability parameter λ. The solid and dashed horizontal lines represent positive and negative values of the left-hand-side of Eq. (2.26), determined by the slope at a given point on an iceline curve such as in Fig. 2. Intersections marked with circles give the stability eigenvalues, and λ_0 is the minimum stability eigenvalue.

(negative slope) shown in Fig. 2, the iceline slope is negative, and we have the situation indicated by the dotted horizontal line in Fig. 4. In this case, the stability eigenvalue λ_0 is negative, so that the corresponding perturbation δT_0 increases in time according to Eq. (2.14). Eventually the linearization of $T[x_s(t),t]$ and $a[x,x_s(t)]$ breaks down, and the exact solution approaches one of the available stable solutions for the given value of Q, for example, that of an ice-covered or ice-free planet.

If, as Drazin and Griffel (1977) have suggested, the iceline curve can have a cusp in some cases, so that $dQ/dx_s \to -\infty$ as $x_s \to 1$, the dotted line in Fig. 4 might conceivably fall below the asymptotic value of $\lambda F(\lambda)$, so that a region of tremendous instability ($\lambda_0 \to -\infty$) would suddenly become stable. It remains then in our proof to show that the dashed line never falls below the asymptotic level of $\lambda F(\lambda)$ as given by (3.7).

The slope of the iceline curve for equilibrium solutions may be computed from Eq. (2.21) and has the form

$$\frac{K_s}{Q}\frac{dQ}{dx_s} = (-T_E')/Q - \int_0^1 dy G_0(x_s,y)S(y)a_2(y,x_s), \quad (3.8)$$

where the positive constant K_s is defined in Eq. (2.23). Since for the class of models (or solutions) considered the temperature decreases as one crosses the ice cap poleward, the first term in (3.8) is positive. The second term, on the other hand, is strictly negative. This latter is in fact just the asymptotic value of $\lambda F(\lambda)$ which is given in (3.7). Hence, $(K_s/Q)(dQ/dx_s)$ is always larger than the asymptotic plateau by the positive amount $(-T_E')/Q$. This concludes the proof of the slope-stability theorem.

We should note that the cusp behavior of the x_s vs Q curve pointed out by Drazin and Griffel (1977) and shown near $x_s = 1$ in Fig. 2 comes about for the case of $a(x,x_s)$ a step function. In this case $a_2(x,x_s)$ is a delta function so that the second term of (3.8) can be evaluated explicitly. The cusp comes about because the resulting expression is proportional to $G_0(x_s,x_s)$, which diverges as $x_s \to 1$. Any smoothing of the albedo at the ice cap edge eliminates this divergence. (In fact, we have found numerically that a smoothing width Δx_s of the order of 0.1 is sufficient to remove the negative slope portion of the curve in Fig. 2 near $x_s = 1$ altogether.) One also notes that the cusp is removed even in the discontinuous albedo case if $S(1) = 0$, in agreement with Drazin and Griffel.

We may now examine the special cases, $x_s = 0$ and $x_s = 1$, the ice-covered and ice-free situations. In both cases we may neglect the ice feedback since, for example, in the $x_s = 0$ case the equator is well below -10°C so that an enormous perturbation would be required to cause the ice cap to recede from the equator. Hence an infinitesimal perturbation of the steady-state solution causes only a perturbation in the temperature field and $\delta x_s = 0$. One easily deduces that in this case the stability eigenvalues are $L_n > 0$, and the solution is stable. The same argument works for $x_s = 1$.

Finally, we may relate the slope-stability theorem as stated for the $x_s(Q)$ iceline curve to the formulation in terms of the corresponding $T_0(Q)$ curve, as discussed for the globally averaged model in the introduction. We shall see that the "heat capacity" argument does not invariably follow in the present class of spatially extended systems. First note that in the special cases $x_s = 0$, 1 as discussed in the preceding paragraph, the connection is correct since dT_0/dQ is clearly positive and the solution is stable. Elsewhere we may relate the iceline and T_0 slopes by first integrating (2.7) over the hemisphere and then differentiating with respect to Q:

$$dT_0/dQ = \int_0^1 dx \int_0^1 dy G_0(x,y)S(y)a(y,x_s) + Q\frac{dx_s}{dQ}\int_0^1 dx \int_0^1 dy G_0(x,y)S(y)a_2(y,x_s).$$

Since both of the integrals above are bounded and strictly positive, and dx_s/dQ changes sign at an ordinary bifurcation by passing through infinity, we may assert that these derivatives have the same sign at least when $|dx_s/dQ|$ is large enough. However they could have opposite sign near a cusp for which $dx_s/dQ \to 0$. Hence, the theorem must be stated as (3.1) in terms of x_s rather than as (1.4).

4. Discussion

We have proven a theorem that states that the stability of a model climate solution depends on a property of the curve which determines the steady-state solutions, the latitude of the ice cap edge versus the solar constant. If the local slope of this curve is positive (negative) the corresponding steady-state solution will be stable (unstable). The theorem covers a fairly broad class of models of the ice-cap-feedback type, but certain assumptions are necessary. The most important of these relate to the Green's function for the linear part of the problem. Existence, positivity and a certain type of asymptotic behavior are sufficient for our proof to hold. These conditions are satisfied by generalized diffusive (Sturm-Liouville) models and hold for some nonlocal models such as Budyko's.

We have seen that positive values of the global "heat capacity" do not guarantee stability in these models, although negative values definitely imply instability. Instability may be simply shown by an

opportune choice of perturbation, but to demonstrate stability of a solution under an arbitrary perturbation is less simple, particularly in systems with spatial extension, which may not follow intuitions appropriate for their global or homogeneous counterparts.

We have also pointed out that the stability of small ice caps, as determined by the slope of the iceline curve near $x_s = 1$, is highly sensitive to the smoothness of the albedo across the iceline. Such dependence unfortunately suggests that the true sensitivity and stability of small ice caps cannot be estimated from such simple models. This question is clearly important, considering the possibility of global warming and the effect of a melting ice cap on the present level of the oceans. The question of whether energy and mass balance requirements are sufficient to model the problem, or whether specific dynamical mechanisms must be included, can only be answered through continued study of ice cap variations on many time scales.

Finally, one may imagine attempting to generalize our theorem to models having more degrees of freedom, with additional dynamics and feedback mechanisms. Rather than resort to pure speculation, we shall remain within the present framework of one-dimensional energy-balance models, and briefly discuss three types of extensions which may affect the stability: different iceline conditions, nonlinearities due to other feedbacks and climatic noise.

Generalizations of our conditions of isothermal icelines having north-south symmetry are of interest. Pollard (1976), for instance, has shown that if the iceline lags slightly behind the -10°C isotherm in time rather than following it instantaneously the theorem still holds. We do not yet know if some similar generalization holds for cases when more than one index might be required to specify the branch of a solution. For instance, in unsymmetrical solutions with only one ice cap, one must specify which ice cap.

Several nonlinearities merit further study:

1) As Stone (1973) has suggested, the diffusion coefficient may be proportional to some power of the temperature gradient.

2) Any dependence of atmospheric carbon dioxide, water vapor or cloudiness on the surface temperature would tend to introduce some nonlinearity in the infrared emission. The net effect has been much contested, but since analytical solutions to linear models are known, it would be of much interest to study the effect of a nonlinear perturbation.

3) Although we have generalized the discontinuous coalbedo to allow for smooth variations in latitude, a more realistic treatment of cloudiness would also allow the co-albedo to vary smoothly with temperature. Even if no analytical solution to such a nonlinear problem can be found, it may be possible to solve the linear stability problem analytically.

Climatic noise is typically modeled by the addition of stochastic forcing terms having given statistical properties. These then determine the statistics of the response, or in this case the temperature fluctuations. Such fluctuations will normally be governed by our stability theorem, but new stability questions arise and here we mention two which merit further study:

1) The probability of finite amplitude fluctuations may be sufficiently large to effect transitions between different branches of the solution curve. One must then consider the relative stability of the states involved. For example, states which have marginal stability in the deterministic response may be only metastable in the stochastic case.

2) The statistics of the forcing may itself depend on the climatic state; for example, its variance may depend on the temperature. In that case the stability may not be related in a simple way to the deterministic solution curve.

Many questions remain unanswered concerning climatic stability, and we hope that the present paper will help generate further interest in and study of such questions.

Acknowledgments. Both authors wish to thank the Advanced Study Program at the National Center for Atmospheric Research where part of this work was done. We also thank Ms. Haydee Saimún for some of the programming and graphics. We acknowledge that a proof like that in the Introduction was first suggested to us by Dr. Robert Dickinson. We also wish to thank Dr. Michael Ghil for several helpful suggestions. The research for this paper was supported in part by the Climate Dynamics Office of the National Science Foundation, and in part by the Goddard Laboratory for Atmospheric Sciences/NASA.

APPENDIX A

Response Functions in the Budyko and General Diffusion Models

In the Budyko model, the combined infrared and meridional transport terms are given by

$$L[T](x,t) = \int_0^1 dz[(\gamma + B)\delta(x - z) - \gamma]T(z,t), \quad \text{(A1)}$$

and in this case it is easily verified that the generalized response function defined in Eq. (2.16) is given by

$$G_\lambda(x,y) = \frac{\delta(x - y)}{\gamma + B - \lambda} + \frac{\gamma}{(B - \lambda)(\gamma + B - \lambda)}. \quad \text{(A2)}$$

which clearly satisfies our conditions (ii) and (iii) discussed in Section 3. Note that in the absence of ice feedback, any heat added to a given latitude belt does not affect the temperature gradient at any other latitude in this model, since G_0 is constant for $x \neq y$. Also, the delta function implies that any discontinuity in albedo yields a discontinuity in temperature, so that one must be very careful in applying the iceline condition.

Next we consider a class of models given by energy-balance equations of the form

$$C\frac{\partial T}{\partial t} - \frac{\partial}{\partial x}D(x)(1-x^2)\frac{\partial}{\partial x}T + A(x) + B(x)T - QS(x)a(x,x_s) = 0, \quad \text{(A3)}$$

along with the boundary conditions of vanishing heat flux at the equator and pole, appropriate for a prototype planet with hemispheric symmetry, so that

$$D(x)[(1-x^2)]^{1/2}\frac{dT}{dx}\bigg|_{x=0,1} = 0. \quad \text{(A4)}$$

The Green's function for the equilibrium solution of such a model satisfies

$$\left[-\frac{d}{dx}D(x)(1-x^2)\frac{d}{dx} + B(x)\right]G_0(x,y) = \delta(x-y), \quad \text{(A5)}$$

as well as the boundary conditions (A4). We wish to show that

$$G_0(x,y) \geq 0 \quad \text{if} \quad B(x), D(x) > 0. \quad \text{(A6)}$$

The proof is identical for G_λ as long as $B - \lambda > 0$.

We may easily demonstrate the positivity of G_0 at the point $x = y$ from the so-called "spectral representation." The eigenvalue equation

$$\left[-\frac{d}{dx}D(x)(1-x^2)\frac{d}{dx} + B(x)\right]f_n(x) = L_n f_n(x), \quad \text{(A7)}$$

determines the orthonormal eigenfunctions f_n as well as the corresponding eigenvalues L_n of the linear operator. [See Ghil (1976) for a comparison with standard Sturm-Liouville systems.] Multiplying Eq. (A7) by $f_n(x)$, integrating over all x, and employing the boundary conditions to integrate the first term by parts yields

$$\int_0^1 dx\left[D(x)(1-x^2)\left|\frac{df_n}{dx}\right|^2 + B(x)f_n^2\right] = L_n > 0, \quad \text{(A8)}$$

so that the spectrum is positive definite. It can be shown that the eigenfunctions f_n form a complete system; hence we may represent G_0 as

$$G_0(x,y) = \sum_n \frac{f_n(x)f_n(y)}{L_n}, \quad \text{(A9)}$$

which along with Eq. (A8) clearly implies

$$G_0(x,x) > 0. \quad \text{(A10)}$$

For the case in which B and D are constant, we obtain $f_n = (2n+1)^{1/2}P_n$ and $L_n = Dn(n+1) + B$, where the P_n are the even-indexed Legendre polynomials. In general, for reasonable $D(x)$, $B(x)$, the f_n are regular everywhere and we shall, therefore, assume in the following that $G_0(x,y)$ is everywhere finite [with the exception of the singular point in Eq. (A4), viz., $x, y = 1$].

In order to show that G_0 is positive for $x \neq y$ we return to Eq. (A5) and choose a new independent variable $z(x)$ such that

$$\frac{dx}{dz} = D(x)(1-x^2) > 0. \quad \text{(A11)}$$

The integral of (A11) is

$$z(x) = \int_0^x \frac{dx'}{D(x')(1-x'^2)}, \quad \text{(A12)}$$

so that $z(0) = 0$ and $z(1) = \infty$. Multiplying (A5) by dx/dz leads to

$$\left[-\frac{d^2}{dz^2} + C(z)\right]G_0(z,z_0) = \delta(z-z_0), \quad \text{(A13)}$$

where $C(z) = (dx/dz)B(z) > 0$ and we have employed the well-known rule that $\delta[x(z) - y] = \delta(z - z_0)/|dx/dz|$, where $x(z_0) = y$. According to Eq. (A13), the curvature of G_0 is determined by the sign of G_0. G_0 is concave upward in the upper half plane, concave downward in the lower half plane, and has a point of inflection if it vanishes. Thus, if G_0 is to remain finite, it must be asymptotically flat as $z \to \infty$, and according to the boundary conditions it is also asymptotically flat as $z \to 0$. Since G_0 is positive at $z = z_0$, and thus concave upward on either side of $z = z_0$, it cannot vanish anywhere and still become flat as $z \to 0, \infty$. Thus, it must be positive everywhere.

APPENDIX B

Step-Function Albedo

In order to make contact with previous studies, we may specialize our stability condition in Eq. (2.26) to the case in which the absorption changes discontinuously at the iceline, so that

$$a(x,x_s) = a_u\theta(x_s - x) + a_l\theta(x - x_s), \quad \text{(B1)}$$

where a_u and a_l are given constants for which $a_u > a_l$, and θ is the unit step function. Since the derivative of θ is a delta function, Eq. (2.27) reduces to

$$F(\lambda) = S(x_s)(a_u - a_l)\int_0^1 dz\,G_\lambda(x_s,z)G_\lambda(z,x_s). \quad \text{(B2)}$$

If we insert the spectral representation

$$G_\lambda(x,x_s) = \sum_n \frac{f_n(x)f_n(x_s)}{L_n - \lambda}, \qquad \text{(B3)}$$

where the f_n are the orthonormal eigenfunctions of L which we assume to be complete, and the L_n are the eigenvalues which we assume to be positive definite (as proven for general diffusive models in Appendix A), then Eq. (B2) takes the form

$$F(\lambda) = S(x_s)(a_u - a_l) \sum_n \frac{[f_n(x_s)]^2}{L_n(L_n - \lambda)}. \qquad \text{(B4)}$$

For $\lambda > 0$, $F(\lambda)$ has simple poles at $\lambda = L_n$ and changes sign as λ passes each L_n. However, for $\lambda < 0$ both $F(\lambda)$ and $dF(\lambda)/d\lambda$ are positive definite, and $\lambda F(\lambda)$ decreases monotonically to $-\infty$. Thus, for this case the graphical solution of Eq. (2.26),

$$\frac{K_s}{Q}\frac{dQ}{dx_s} = \lambda F(\lambda),$$

will have the qualitative form shown in Fig. 4, already discussed for the general case in Section 3.

REFERENCES

Budyko, M. I., 1968: On the origin of glacial epochs. *Meteor. Gidrol.*, 2, 3–8.

——, 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, 14, 611–619.

——, 1972: The future climate. *Trans. Amer. Geophys. Union*, 53, 868–874.

Chylek, P., and J. A. Coakley, 1975: Analytical analysis of a Budyko-type climate model. *J. Atmos. Sci.*, 32, 675–679.

Crawford, C., and E. Kallén, 1978: A note on the condition for existence of more than one steady-state solution in Budyko-Sellers type models. *J. Atmos. Sci.*, 35, 1123–1125.

Drazin, P. G., and D. H. Griffel, 1977: On the branching structure of diffusive climatological models. *J. Atmos. Sci.*, 34, 1696–1706.

Fraedrich, K., 1978: Structural and stochastic analysis of a zero-dimensional climate system. *Quart. J. Roy. Meteor. Soc.*, 104, 461–474.

Frederiksen, J. S., 1976: Nonlinear albedo-temperature coupling in climate models. *J. Atmos. Sci.*, 33, 2267–2272.

Ghil, M., 1976: Climate stability for a Sellers-type model. *J. Atmos. Sci.*, 33, 3–20.

Held, I. M., and M. J. Suarez, 1974: Simple albedo feedback models of the icecaps. *Tellus*, 26, 613–628.

Joseph, D. D., 1976: *Stability of Fluid Motions*, Part I. Springer-Verlag, 70–81 and Appendix D.

Nauenberg, M., and V. F. Weisskopf, 1978: Why does the sun shine? *Amer. J. Phys.*, 46, 23–32.

North, G. R., 1975a: Analytical solution to a simple climate model with diffusive heat transport. *J. Atmos. Sci.*, 32, 1301–1307.

——, 1975b: Theory of energy-balance climate models. *J. Atmos. Sci.*, 32, 2033–2043.

——, 1977: Analytical models of the earth's macroclimate. *Proceedings of the First International Conference on Mathematical Modeling*, Vol. 4, University of Missouri-Rolla, 2291–2302.

——, L. Howard, D. Pollard and B. Wielicki, 1979: Variational formulation of Budyko-Sellers climate models. *J. Atmos. Sci.*, 36, 255–259.

Pollard, D., 1976: A functional for energy-balance climate models. *1976 GFD Summer Program Lecture Notes*. Woods Hole Oceanographic Institution [NTIS No. PB-265 181].

Schneider, S. H., and T. Gal-Chen, 1973: Numerical experiments in climate stability. *J. Geophys. Res.*, 78, 6182–6192.

——, and R. E. Dickenson, 1974: Climate modeling. *Rev. Geophys. Space Phys.*, 2, 447–493.

Sellers, W. D., 1969: A climate model based on the energy balance of the earth-atmosphere system. *J. Appl. Meteor.* 8, 392–400.

——, 1974: Climate models and variations in the solar constant. *Geofis. Int.*, 14, 303–315.

Stone, P. H., 1973: The effect of large-scale eddies on climatic change. *J. Atmos. Sci.*, 30, 521–529.

Su, C. H., and D. Y. Hsieh, 1976: Stability of the Budyko climate model. *J. Atmos. Sci.* 33, 2273–2275.

Temkin, R. L., and F. M. Snell, 1976: An annually averaged hemispherical climatic model with diffuse cloudiness feedback. *J. Atmos. Sci.*, 33, 1671–1685.

Paper 114

NASA Technical Memorandum 80555 September 1979

A SPECTRAL FILTER FOR ESMR'S SIDELOBE ERRORS

Dennis Chesters
NASA/Goddard Space Flight Center
Code 915
Greenbelt, Maryland 20771

ABSTRACT

Fourier analysis is used to remove periodic errors from a series of NIMBUS-5 ESMR microwave brightness temperatures. The observations were all taken from the midnight orbits over fixed sties in the Australian grasslands during 1974. The angular dependence of the data indicates calibration errors consisting of broad sidelobes and some miscalibration as a function of beam position. Even though an angular recalibration curve cannot be derived from the available data, the systematic errors can be removed with a spectral filter. The 7-day cycle in the drift of the orbit of NIMBUS-5, coupled to the look-angle biases, produces an error pattern with peaks in its power spectrum at the weekly harmonics. About ±4°K of error is removed by simply blocking the variations near two- and three-cycles-per-week.

Paper 115

NASA Technical Memorandum 79576 June 1978

ALBEDO CLIMATOLOGY ANALYSIS AND THE DETERMINATION OF FRACTIONAL CLOUD COVER

Robert J. Curran and Raymond Wexler
NASA/Goddard Space Flight Center
Greenbelt, Maryland 20771

Myron L. Nack
Computer Sciences Corporation
Silver Spring, Maryland 20910

ABSTRACT

In a previous paper we developed a transformation of surface albedo to albedos at the top of clear, totally cloud covered, and fractionally cloud covered realistic atmospheres. In this paper we present monthly and zonally averaged surface cover climatology data, which we use to construct monthly and zonally averaged surface albedos. Our albedo transformations are then applied to the surface albedos, using solar zenith angles characteristic of the Nimbus 6 satellite local sampling times, to obtain albedos at the top of clear and totally cloud covered atmospheres. We then combine these albedos with measured albedo data to solve for the monthly and zonally averaged fractional cloud cover. The measured albedo data were obtained from the wide field of view channels of the Nimbus 6 Earth Radiation Budget (ERB) experiment, and consequently our fractional cloud cover results are representative of the local sampling times. These fractional cloud cover results are compared with recent studies. The cloud cover results not only show peaks near the intertropical convergence zone (ITCZ), but the monthly migration of the position of these peaks follows general predictions of atmospheric circulation studies.

Paper 116

NASA Technical Memorandum 80332 July 1979

AN IMPROVED SCHEME FOR THE REMOTE SENSING OF SEA SURFACE TEMPERATURE

G. Dalu*, C. Prabhakara
Laboratory for Atmospheric Sciences (GLAS)
NASA/Goddard Space Flight Center
Greenbelt, Maryland 20771

R. C. Lo, M. J. Mack
Computer Sciences Corporation
Silver Spring, Maryland 20910

ABSTRACT

A radiometer which has two channels in the 11 to 13 μm window region, has been proposed for inclusion on the forthcoming NOAA satellites. It will be useful in estimating the sea surface temperature to within an accuracy of 1°C. But this study shows that this accuracy could be improved to within 0.3°C, if an independent estimate of total precipitable water vapor is available. In order to remotely sense the total precipitable water vapor, a broad-band channel in the 18 μm water vapor band should be introduced in addition to the two channels in the 11 to 13 μm window region. With these three channels the total water vapor could be estimated over oceans, which would improve the accuracy of the sea surface temperature estimation. In addition the effect of the surface emissivity is taken into account in this scheme.

*NAS/NRC Research Associate on leave from C.N.R., Istituto di Fisica della Atmosfera, Rome, Italy.

Paper 117

NASA Conference Publication 2076 – Fourth NASA Weather and Climate Program Science Review, January 24-25, 1979, NASA/GSFC, Greenbelt, MD.

Paper No. 41

THE EFFECT OF SURFACE REFLECTION AND CLOUDS ON THE ESTIMATION OF TOTAL OZONE FROM SATELLITE MEASUREMENTS

R. S. Fraser, *Goddard Space Flight Center, Greenbelt, Maryland*
Z. Ahmad, *Systems and Applied Sciences Corporation, Riverdale, Maryland*

ABSTRACT

The total amount of ozone in a vertical column is being measured by Nimbus 4 and 7 observations of the intensity of ultraviolet sunlight scattered from the earth. The algorithm for deriving the amount of ozone from the observations uses the assumption that the surface reflects the light isotropically and the albedo is independent of wavelength. The effects of anisotropic surfaces and clouds on the estimate of total ozone are computed for models of the earth-atmosphere system.

Paper 118

Reprinted from JOURNAL OF THE ATMOSPHERIC SCIENCES, Vol. 36, No. 7, July 1979
American Meteorological Society
Printed in U. S. A

Perturbation of the Zonal Radiation Balance by a Stratospheric Aerosol Layer[1]

HARSHVARDHAN[2]

Laboratory for Atmospheric Sciences, NASA/Goddard Space Flight Center, Greenbelt, MD 20771

(Manuscript received 22 November 1978, in final form 22 March 1979)

ABSTRACT

The effect of stratospheric aerosols on the earth's monthly zonal radiation balance is investigated using a model layer consisting of 75% H_2SO_4, which is the primary constituent of the background aerosol layer. The reduction in solar energy absorbed by the earth-atmosphere system is determined through the albedo sensitivity, defined here as the change in albedo per unit mid-visible optical depth of the aerosol layer. The optically thin approximation is used in conjunction with the Henyey-Greenstein phase function for scattering to simplify computations. Satellite-derived planetary albedos are used as the frame of reference about which the change in albedo is computed. An infrared radiative transfer model is used to estimate the increased greenhouse effect attributed to the aerosol layer. The infrared heating tends to compensate for the albedo effect in altering the radiation balance. The results indicate that the dominant influence of the thin model stratospheric aerosol layer is an increased reflection of solar energy all over the globe except for the polar-winter region, but the change in the radiation balance is seen to be uniform and small equatorward of 50°. The largest perturbations occur in the spring and fall in the polar regions with the equator-to-pole difference in radiation balance increasing by 6.5–7.0 W m^{-2} for an aerosol mid-visible optical depth of 0.1.

1. Introduction

The role of stratospheric aerosols in enhancing the reflection of solar radiation has been investigated recently by numerous authors (Harshvardhan and Cess, 1976; Pollack *et al.*, 1976; Cadle and Grams, 1975; Pinnick *et al.*, 1976; Herman *et al.*, 1976). Model calculations by the above and also by Luther (1976) have shown that the dominant climatic effect of these aerosols is enhanced solar reflection with increased infrared opacity playing a smaller role. However, Herman *et al.* (1976) and Luther (1976) have also pointed out that the degree of albedo enhancement is a strong function of the albedo of the underlying surface. Broadly speaking, this underlying surface may be divided into two classes, *viz.*, cloud-free portions of the earth-atmosphere system and the cloudy portion. This classification does not imply a high and low albedo differentiation because the effective albedo of a cloud layer above a snow or ice-covered surface may be lower than that of the surface alone (Dave and Braslau, 1975). However, with the above exception, the effective albedo of the cloud-free regions is between 0.15 and 0.3, whereas that of cloudy regions is greater than 0.4.

The present study examines the albedo sensitivity of the earth-atmosphere system to a stratospheric aerosol layer under different assumptions regarding surface and cloud albedos and solar zenith angle averaging procedures. Work to date encompasses a broad range of simplifying assumptions. Harshvardhan and Cess (1976) considered an albedo of 0.3 for the unperturbed earth-atmosphere system. Pollack *et al.* (1976) assumed a ground albedo of 0.1 and 50% cloud cover with a detailed radiative transfer scheme to yield aerosol layer perturbations, while Herman *et al.* (1976) assigned ground albedos to 10° latitude belts in the Northern Hemisphere with seasonal variations and assumed a 50% cloud cover of reflectivity equal to 0.5. The marked changes in albedo sensitivity obtained by Herman *et al.*, with changing seasons at high latitudes, due to changes in the mean solar zenith angle and surface albedo, suggests that incorrect inferences may be drawn if global average surface albedos and cloud cover are assumed in computing albedo enhancement.

In this work the albedo of the cloud-free portion of the earth-atmosphere system in 10° latitude belts is deduced following the recommendation of Budyko (1974). An empirical relation for the effective cloud reflectivity is derived from calculations of Dave and Braslau (1975). With these two sets of albedos as a base, the global albedo change in the presence of a model stratospheric aerosol layer is computed for

[1] This paper is a condensation of NASA Technical Memorandum 79573, Goddard Space Flight Center, Greenbelt, Md.

[2] Present affiliation: Meteorology Program, University of Maryland, College Park 20742.

0022-4928/79/071274-12$07.00
© 1979 American Meteorological Society

climatological cloud-cover fractions. The procedure is extended to include the zonal variation of albedo sensitivity by month, based on satellite-derived planetary albedos. With the inclusion of the effect of the stratospheric aerosol layer on the infrared radiation leaving the top of the atmosphere, the net energy deficit on a zonal basis is obtained. This quantity may be used as an input to energy-balance climate models used for investigating the climatic effects of stratospheric aerosols.

2. Planetary albedo

The earth-atmosphere system albedo α_p is commonly expressed as

$$\alpha_p = \alpha_c C + \alpha_s(1 - C), \tag{1}$$

where α_s is the cloud-free or clear-sky albedo and α_c the albedo with cloud cover; C is the fractional cloud cover. The expression for α_p may be used on a global basis or for individual latitude belts (Budyko, 1974; Cess, 1976). For the remainder of the text α_p will be used to denote the albedo for each 10° latitude belt, with the hemispheric planetary albedo A computed as

$$A = 4/S \int_0^{\pi/2} \alpha_p(\phi) Q(\phi) \cos\phi d\phi, \tag{2}$$

where ϕ is the latitude, Q the latitudinal distribution of incoming solar radiation and S the solar constant. The perturbation of this planetary albedo with the introduction of a stratospheric aerosol layer enhancement (as following a major volcanic eruption) is related to the optical depth of the aerosol layer. The albedo sensitivity may be defined as $dA/d\tau$ where the optical depth τ is computed at some reference wavelength, usually the mid-visible wavelength of 0.55 μm. Since the stratospheric aerosol layer is optically thin ($\tau \ll 1$), the change in albedo is directly proportional to the change in optical depth and the albedo sensitivity may be written as $\Delta A/\Delta\tau$ (Russell and Hake, 1977). In succeeding sections, models of the planetary albedo and reflectivity of the aerosol layer are postulated and the albedo sensitivity computed.

a. Surface and clear-sky albedo

The fraction of solar radiation incident on a surface that is reflected by it is the surface albedo α_{s0}. Using the mean value of albedo for various natural surfaces, Budyko (1974) has suggested appropriate values of land albedos to be used for different climatic and seasonal regimes. The values vary from 0.8 for stable snow cover in high latitudes to 0.13 for steppe and forests. The albedo of a water surface to be used for the ocean fraction of latitude belts has also been given by Budyko based on the work of Sivkov (in Kondratyev, 1975). As there is a strong zenith-angle-dependence of the direct beam albedo, there is a marked seasonal variation at midlatitudes and high latitudes. The albedo of both land and ocean change dramatically with snow and ice cover and this seasonal variation must also be incorporated in the computation of global albedo. In the present study, data on snow and ice cover compiled by Curran *et al.* (1978) were used to generate the seasonal march of albedo at high latitudes. The monthly zonal surface albedos for land and ocean were computed after identifying the surface type and incorporating snow and ice cover. The clear-sky albedo α_s is the effective albedo of a cloud-free atmosphere over the land and ocean surface identified by the albedos α_{s0}. It is possible to make a detailed radiative transfer calculation averaged over the solar spectrum to obtain α_s; however, for our purposes, a simple parameterization that takes into account molecular and particle scattering as well as gaseous absorption in the atmosphere is sufficient. One such relation may be obtained from the model of Braslau and Dave (1973) for midlatitude summer water vapor and ozone distributions and an absorbing tropospheric aerosol layer. Choosing their Model C1, the planar reflectivity of a clear sky over a black surface may be written as

$$\alpha_s(R = 0, \theta) = 0.05 + 2.4 \times 10^{-6}\theta^2 + 3.2 \times 10^{-9}\theta^4, \tag{3}$$

where θ is the solar zenith angle (deg) and R is the reflectivity of the underlying surface. To extend this relation to reflecting surfaces, a multiple reflection model is used which yields (Coakley and Chýlek, 1975; Wiscombe, 1975)

$$\alpha_s(R,\theta) = \alpha_s(R = 0, \theta) + \frac{RT_cT_d}{1 - R\alpha_d(R = 0)}, \tag{4}$$

where T_c and T_d are the planar and diffuse transmission, respectively, and $\alpha_d(R = 0)$ is the diffuse or global albedo over a black surface. The atmospheric absorptivity used to compute T_c and T_d has been assumed to be 15% which is an average value based on Braslau and Dave's computations. Using the above parameterization, one can generate the monthly zonal clear-sky albedos from surface albedo values and mean monthly solar zenith angle and also obtain the insolation-weighted, mean clear-sky albedo of each latitude belt. This is the value of α_s that should be used in mean annual global energy balance climate models as it is a measure of the total solar radiation reflected back to space by the clear-sky portion of each latitude belt on an annual basis. The values of clear-sky albedo are compared with the minimum albedo obtained from 29 months of satellite observations by Vonder Haar and Ellis (1975) in Fig. 1. Although sampling by various satellites occurred at just one local time dur-

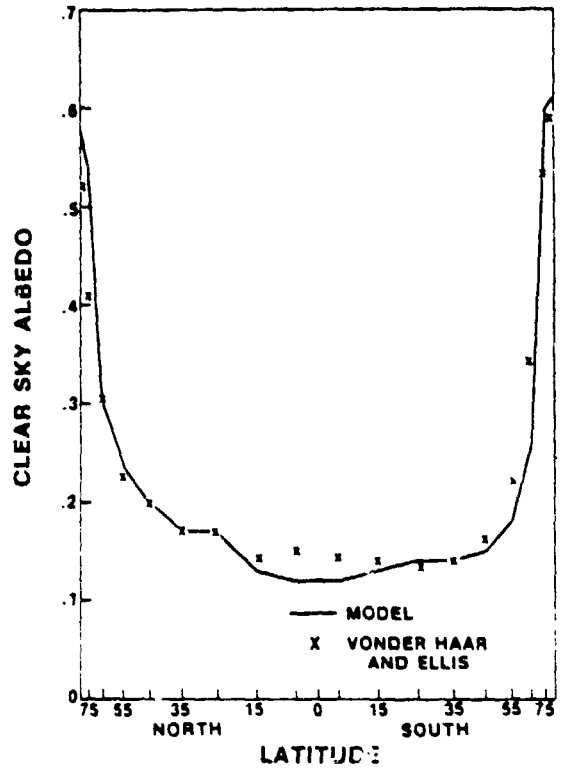


FIG. 1. Model clear-sky albedos compared with minimum albedos from Vonder Haar and Ellis (1975).

ing daylight hours, the 29-month set includes late morning, near noon and early afternoon orbits; so the results may be considered to correspond to a mean solar zenith angle.

Viewing geometry and the presence of clouds over snow and ice make observations near the poles unreliable; so only the portion of the globe between latitudes 65°N and 65°S should be considered. Substantial differences between the model and the minimum albedo exist at high latitudes in the Southern Hemisphere and near the Equator. The former is a region of extensive cloudiness [up to 80% average cloud cover (van Loon, 1972)], and it is possible there were no totally cloud-free satellite viewings recorded, resulting in the 5–10% positive discrepancy. Since one set of radiatively active constituents is used in the atmospheric model to represent the whole globe, the parameterization is adequate.

b. Cloudy-sky albedo

The spectrally averaged reflectivity of an isolated cloud layer is a function of the physical properties of the cloud and the zenith angle of the incident solar radiation. The albedo of a cloudy sky further depends on the reflectivity of the underlying surface. Since the number of variables involved in specifying a cloud completely for a radiative transfer computation is large and the nature of cloud cover and underlying surface is variable, parameterization of the reflectivity is necessary. In the present work, the cloud-sky albedo is assumed to depend on the mean solar zenith angle and reflectivity of the underlying surface. An empirical relation between the effective albedo of a cloudy sky over a nonreflecting surface, derived from the results of Dave and Braslau (1975), is extended to yield cloud-sky albedos using a multiple-reflection model as in Wiscombe (1975).

Dave and Braslau (1975) have presented results for a detailed atmospheric model with midlatitude summer water vapor and ozone distributions and two tropospheric aerosol concentrations, and also a nonabsorbing cloud layer between 3 and 4 km. For parameterization, their model with average aerosols (C1-ST) has been chosen and an excellent fit to the planar reflectivity over a black surface is obtained with

$$\alpha_c(R = 0, \theta) = 0.194 + 4.9 \times 10^{-5}\,\theta^2, \quad (5)$$

where θ and R have been defined previously. To extend this relation to reflecting surfaces, Eq. (4) is used, again with 15% absorption. The cloud model used by Dave and Braslau has an optical depth of 3.35 at 0.55 μm and a diffuse albedo of 33.2%. This model (henceforth called Model C) has been used to calculate the cloud-sky albedo α_c for each latitude belt and month of the year, and the insolation-weighted annual mean cloud-sky albedo is shown in Fig. 2 along with the values of α_c deduced by Cess (1976) using satellite measurements of the global albedo and mean annual cloud cover. The model is seen to be an underestimate at all but equatorial latitudes, indicating that the optical depth used in the model is too low. Choice of a larger optical depth and assumption that the cloud layer is nonabsorbing, yields

$$\alpha_c(R = 0, \theta) = 0.3 + 4.25 \times 10^{-5}\,\theta^2, \quad (6a)$$

henceforth called Model A. Another model albedo intermediate between that of Eq. (5) and Model A is assumed to be

$$\alpha_c(R = 0, \theta) = 0.25 + 4.625 \times 10^{-5}\,\theta^2 \quad (6b)$$

and is called Model B. The mean annual results have been presented in Fig. 2 along with the model given by (5) and Cess (1976).

The highest reflectivity is given by Model A and is seen to be an overestimate throughout, while Model B provides good correspondence at midlatitudes and high latitudes. The three cloud models have been used in Eq. (1) to give the latitudinal distribution of mean global albedo using the clear-sky albedos of Fig. 1 and cloud-cover fractions C re-

ported by Cess (1976). The results are presented in Fig. 3. Also marked are values of α_p from satellite measurements (Ellis and Vonder Haar, 1976) and from the Nimbus 6 ERB experiment (Jacobowitz *et al.*, 1979). Except for latitudes poleward of 65°, the variation is seen to be simulated quite well by Model B at midlatitude. The Northern Hemisphere planetary albedo A, defined by Eq. (2) for the three models, is 0.323 for Model A, 0.308 for Model B and 0.291 for Model C. It may be noted that the planetary albedo derived from the satellite results of Ellis and Vonder Haar is 0.306. The three models are seen to cover the range of reasonable values of the cloud-sky albedo and the planetary albedo, and as such may be used as the basis for computing the albedo sensitivity.

3. Stratospheric aerosol layer

There is a persistent tenuous layer of submicron size aerosols in the stratosphere that is composed primarily of an aqueous solution of sulfuric acid (Cadle and Grams, 1975). The number concentration of this aerosol layer increases after major volcanic events, and as the residence time of these aerosols is of the order of a year or more, this enhanced concentration persists for up to three years. Recent measurements (Hofmann *et al.*, 1975; Rosen *et al.*, 1975) during a period free of major volcanic activity have given a picture of the worldwide distribution

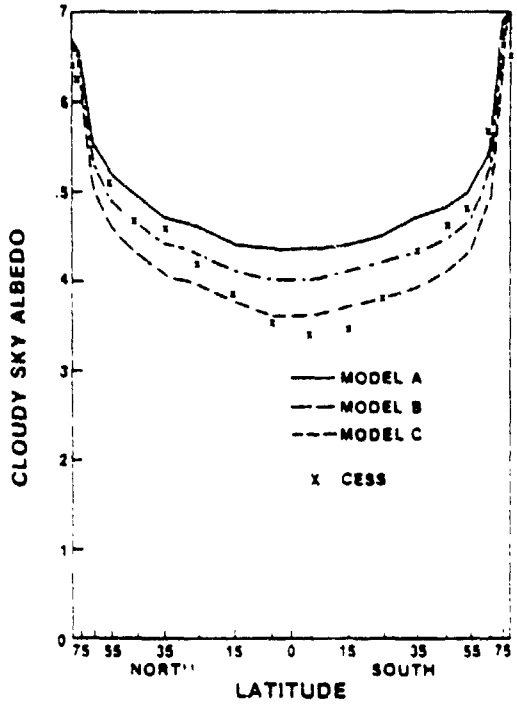


FIG. 2. Cloudy-sky albedos for three models compared with values deduced by Cess (1976).

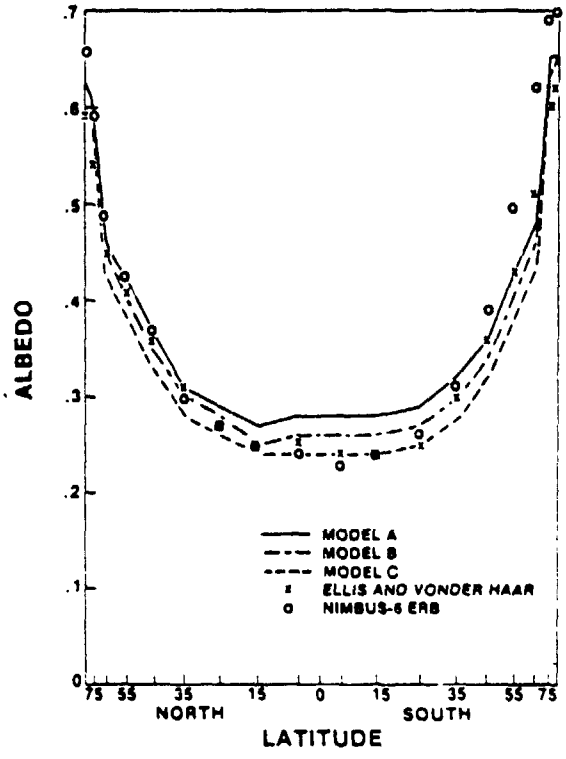


FIG. 3. Planetary albedos compared with satellite-derived results.

of the background stratospheric aerosol layer. Following a volcanic eruption in which the ejected material penetrates into the stratosphere, there is a very significant enhancement of the aerosol layer. Even after the larger particles that form part of the volcanic debris have fallen back into the troposphere, the aerosols formed from the gaseous ejection remain. Initially this enhanced concentration is limited to the vicinity of the explosion but in a matter of months may spread globally (Dyer and Hicks, 1968). The enhancement of the layer may be 10 or 20 fold in the latitude belt containing the explosion (Volz, 1970) and this can definitely be expected to have radiative effects.

One result of this severe enhancement is the additional reflection of solar energy back to space, and previous work mentioned in Section 1 shows that this indeed is the major radiative effect. However, the magnitude of this effect is crucial in determining the radiative perturbation that may be expected from an enhancement. The albedo sensitivity introduced in Section 2 is a measure of the increased solar reflectivity. A radiative model of the aerosol is needed to compute the reflectivity of the aerosol layer. Mie theory for spherical particles may be used to compute radiative parameters. The aerosol composition is assumed to be 75% H_2SO_4 [Rosen, 1971 (optical constants are tabulated in Palmer and Williams, 1975)] and a modified gamma size distribu-

tion, normalized to 1 particle cm^{-3} is adopted (Shettle and Fenn, 1976):

$$\frac{dn(r)}{dr} = 324r \exp(-18r), \tag{7}$$

with $dn(r)/dr$ the number of particles per cubic centimeter per micron within the radius interval dr, where r is in microns. At the mid-visible wavelength of 0.55 μm, the radiative parameters are $\beta_{ex} = \beta_{sc} = 1.1 \times 10^{-4}$ km^{-1}, $\tilde{\omega} = 1$ and $g = 0.73$, with β_{ex} and β_{sc} the extinction and scattering coefficients, respectively, $\tilde{\omega}$ the single scattering albedo and g the asymmetry parameter.

a. Aerosol reflectivity

As in the case of a cloud deck overlying a reflecting surface, a multiple-reflection model may be used to determine the albedo sensitivity of the stratospheric aerosol layer. For this the reflective property of the isolated aerosol layer has to be determined based on the model introduced above. Here all computations are made at a wavelength of 0.55 μm; however, radiative properties may be averaged over the solar spectrum. The reflectivity of the layer at 0.55 μm is roughly 25% greater than a solar averaged value (Harshvardhan and Cess, 1976).

Since the optical depth of the aerosol layer is small even after major volcanic events, the optically thin approximation is used to compute the reflectivity of the layer. In this limit the planar albedo a is given by (Coakley and Chýlek, 1975)

$$a(\mu_0) = \frac{\tilde{\omega}\tau}{\mu_0} \frac{1}{2} \int_0^1 d\mu\, p(\mu, -\mu_0), \tag{8}$$

where τ is optical depth, μ_0 the cosine of the angle of incidence of the sunlight, and $p(\mu,\mu')$ is the azimuth-independent part of the scattering phase function. The global albedo $\bar{a}$ is

$$\bar{a} = \tilde{\omega}\tau \int_0^1 d\mu \int_0^1 d\mu' p(\mu, -\mu'). \tag{9}$$

It may be noted that both $a(\mu_0)$ and $\bar{a}$ are directly proportional to the optical depth in this limit. The scattering phase function is computed from the model parameters and the integrations in (8) and (9) may be carried out for selected solar zenith angles following the procedure of Wiscombe and Grams (1976). The use of the Henyey-Greenstein phase function (Hansen, 1969), which is

$$p_{HG}(\zeta) = \frac{1 - g^2}{(1 + g^2 - 2g\zeta)^{3/2}}, \tag{10}$$

where g is the asymmetry parameter and ζ the cosine of the scattering angle, introduces only small errors in flux and is used here because closed form solutions can be obtained for $a(\mu)$ and $\bar{a}$. These are

$$a(\mu_0) = \frac{\tilde{\omega}\tau}{\mu_0}\left\{\tilde{\omega}_0 - \frac{\tilde{\omega}_1}{2}\mu_0 - \sum_{l=1}^{\infty} \tilde{\omega}_{2l+1} P_{2l+1}(\mu_0)\left[\frac{(-1)(-3)(-5)\ldots(-2l+1)}{(2l+2)(2l)(2l-2)\ldots 2}\right]\right\}, \tag{11}$$

$$\bar{a} = \tilde{\omega}\tau\left\{\tilde{\omega}_0 - \frac{\tilde{\omega}_1}{4} - \sum_{l=1}^{\infty} \tilde{\omega}_{2l+1}\left[\frac{(-1)(-3)(-5)\ldots(-2l+1)}{(2l+2)(2l)(2l-2)\ldots 2}\right]^2\right\}, \tag{12}$$

where $\tilde{\omega}_0 = 1$, $\tilde{\omega}_1 = 3g/4$ and $\tilde{\omega}_{2l+1} = (4l+3)g^{2l+1}$ and P_l is the Legendre polynomial of order l.

Eqs. (11) and (12) give the reflectivity of an isolated aerosol layer for direct and diffuse solar radiation, respectively. This layer is treated as a thin lid covering the earth-atmosphere system in order to compute the albedo sensitivity. An assumption made here is that the reflection from the portion of the atmosphere above the layer is negligible. The altered albedo of the underlying system which is composed of the atmosphere, clouds and land is obtained from the multiple-reflection model [Eq. (4)]

$$\alpha(\mu_0) = a(\mu_0) + \frac{\alpha'(\mu_0)[1 - a(\mu_0)](1 - \bar{a})}{1 - \bar{a}\alpha'(\mu_0)}, \tag{13}$$

where $\alpha'(\mu_0)$ is the unperturbed system albedo and $\alpha(\mu_0)$ the albedo in the presence of the aerosol layer. As there is no absorption in the model aerosol layer at 0.55 μm ($\tilde{\omega} = 1$), $T_c = 1 - a(\mu_0)$ and $T_d = 1 - \bar{a}$. Eq. (13) may be applied to each latitude belt and the resulting albedo integrated over each hemisphere using Eq. (2), to yield the albedo sensitivity. Each latitude belt is identified with a different mean solar zenith angle, so the planar reflectivity of the layer, $a(\mu_0)$, will be different. Also, land form and cloud cover are variable over the globe, so the radiative perturbation caused by the introduction of an enhanced stratospheric aerosol layer will be latitude dependent. If seasonal effects are considered instead of mean annual averages, the changing solar zenith angle, cloud cover and land reflectivity will produce a seasonal dependence as was shown by Herman *et al.* (1976).

4. Planetary albedo changes

As mentioned in Section 2, the hemispheric planetary albedo is calculated by integrating the albedo α_p of each latitude belt over the hemisphere using Eq. (2). The unperturbed zonal albedo α_p is calculated using (1) in which mean annual cloud-cover fractions may be used. To compute the perturbed

TABLE 1. Albedo of aerosol layer at 0.55 μm; $g = 0.73$.

Latitude (°N)	Cosine of mean solar zenith angle	$a(\mu_0)/\tau$
85	0.257	1.045
75	0.274	0.941
65	0.316	0.739
55	0.385	0.520
45	0.453	0.384
35	0.513	0.302
25	0.560	0.254
15	0.592	0.227
5	0.609	0.214
		$\bar{a}/\tau = 0.401$

planetary albedo, climatological cloud-cover fractions should be used as the albedos of the cloudy and cloud-free portions of the zone are significantly different and so may be expected to be perturbed to different extents with the addition of a reflecting layer. Moreover, lower latitudes cover more area and have greater insolation over the year, so planetary albedo changes will be weighted in favor of low-latitude changes. Apart from the above, the wide variation in planar reflectivity $a(\mu_0)$ of the stratospheric aerosols with latitude will be a dominant effect. Table 1 gives the planar albedo at 0.55 μm for the model aerosol layer for annual mean solar zenith angles corresponding to the different latitude belts. The effect of latitude on the albedo sensitivity may be illustrated by using (13) to compute the change in the albedo of the earth-atmosphere system for arbitrary unperturbed system albedos. This is presented in Fig. 4 for the model nonabsorbing aerosol layer. The ordinate is the albedo sensitivity $[\alpha(\mu_0) - \alpha'(\mu_0)]/\tau$ and $\alpha'(\mu)$ is the unperturbed system albedo. Results are presented for direct beam radiation at the solar zenith angles corresponding to different latitudes and for diffuse radiation. Two features may be noted from this figure. The albedo sensitivity for latitudes equatorward of 45° is less than that for diffuse radiation. These latitude belts contain over 70% of the surface area and receive more solar insolation than regions poleward of 45°. This explains some of the discrepancies in planetary albedo sensitivity obtained by using different averaging procedures as described later. The other notable feature is the almost linear relationship between the albedo sensitivity and surface reflectivity at high latitudes. This feature is present also in the case of slightly absorbing aerosols as shown in Fig. 5, in which results at two mean solar zenith angles corresponding to 85 and 5°N latitude are presented for $\tilde{\omega} = 0.99$, 0.95 and 0.9.

Here three methods of estimating planetary albedo changes are compared. A hemispheric annual average albedo sensitivity is obtained by con-

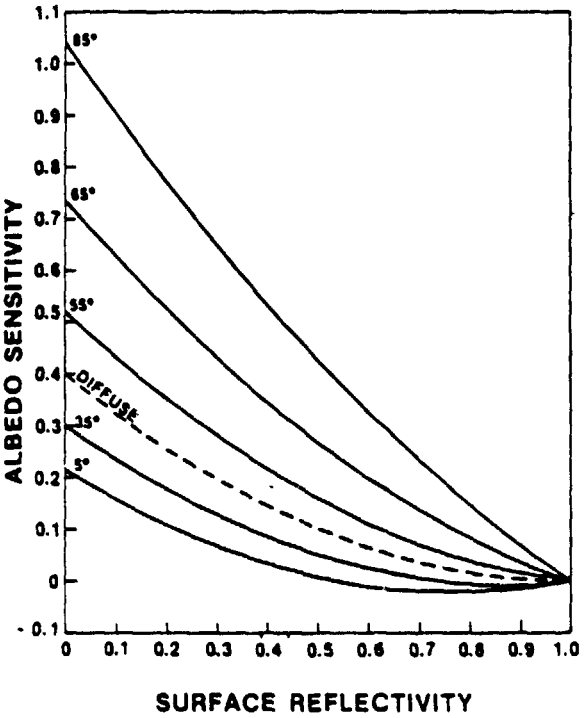


FIG. 4. Albedo sensitivity for model aerosol layer for direct-beam radiation at annual mean solar zenith angles corresponding to latitudes 85°, 65°, 55°, 35°, 5° and for diffuse radiation.

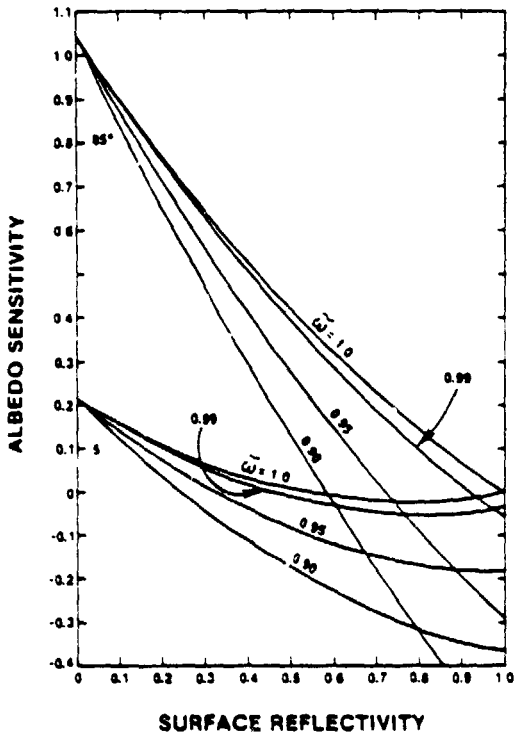


FIG. 5. As in Fig. 4 except for only 85° and 5° latitudes with $\tilde{\omega} = 1.0$, 0.99, 0.95 and 0.9.

sidering the change in the Bond albedo of the earth atmosphere system. For a particular set of albedos α_s and α_c and cloud-cover fraction C, the planetary albedo is computed from Eq. (2). Now it is assumed that a uniform optically thin layer of aerosol covers the planet. Using the multiple-reflection model, the new system albedo is

$$A_{new} = \tilde{a} + \frac{A_{old}(1 - \tilde{a})^2}{1 - \tilde{a}A_{old}} \quad (14)$$

Note that the diffuse or Bond albedo $\tilde{a}$ of the stratospheric layer has been used which for the present model is 0.4τ (Table 1), such that the albedo sensitivity is

$$\frac{A_{new} - A_{old}}{\tau} = \frac{0.4(1 - A_{old})^2}{1 - 0.4\tau A_{old}} \approx 0.4(1 - A_{old})^2 \quad (15)$$

when $\tau \ll 1$. Another approach is to compute the change in reflectivity of each latitude belt with constant cloud cover. This is a zonal annual average accomplished by applying the reflection scheme to each α_p and then integrating over the hemisphere using (2). The change in planetary albedo is then

$$A_{new} - A_{old} = 4/S \int_0^{\pi/2} (\alpha_{p,new} - \alpha_{p,old}) Q \cos\phi d\phi, \quad (16)$$

where, for each latitude belt,

$$\alpha_{p,new}(\phi) = a(\phi) + \frac{\alpha_{p,old}(1 - \tilde{a})[1 - a(\phi)]}{1 - \tilde{a}\alpha_{p,old}} \quad (17)$$

with ϕ the latitude. The zenith-angle-dependence of the planar reflectivity has been translated to a latitude-dependence. The mean annual solar zenith angle has been chosen for this purpose.

The zonal annual average computation may also be carried out separately for the cloudy and clear sky portions of the atmosphere. Here the multiple-reflection model takes the form

$$\alpha_{c,s,new}(\phi) = a(\phi) + \frac{\alpha_{c,s,old}(1 - \tilde{a})[1 - a(\phi)]}{1 - \tilde{a}\alpha_{c,s,old}}, \quad (18)$$

$$\alpha_{p,new} - \alpha_{p,old} = (\alpha_{c,new} - \alpha_{c,old})C + (\alpha_{s,new} - \alpha_{s,old})(1 - C). \quad (19)$$

The three methods of computing the albedo sensitivity have been carried out for the clear-sky albedos and the three model cloudy albedos given in Figs. 1 and 2, respectively. The hemispheric albedo sensitivity calculated using the three methods is listed in Table 2 along with the unperturbed planetary albedo for the three different cloud models and also for albedo values deduced by Cess (1976) from satellite-derived results in Ellis and Vonder Haar (1976).

The results of the albedo sensitivity computation are significant in two ways. First, there is a substantial reduction in the sensitivity when a zonal average is taken instead of a hemispheric average. The mean value of this reduction is 31% for the four cases considered. This lower value is a more appropriate measure of the mean annual planetary albedo sensitivity to a uniform stratospheric aerosol perturbation. The explanation for this reduction may be found in Table 1 and Fig. 4. The planar reflectivity of the aerosol layer is a very strong function of the zenith angle and the use of hemispheric averaging is equivalent to assuming a diffuse reflectivity which is greater than the planar reflectivity for mean solar zenith angles corresponding to latitudes equatorward of 45°. The increased reflectivity of the layer in polar regions is compensated by the effect of lower latitudes which have more area and receive more solar insolation.

The other feature of the results is the insignificant change in albedo sensitivity when cloudy and clear-sky portions of each latitude zone are considered separately. This may seem surprising because the albedo sensitivity is a strong function of the reflectivity of the underlying surface as is shown in Fig. 4. However, the relationship is monotonic and nearly linear for the higher latitudes; the same is

TABLE 2. Albedo sensitivity at 0.55 μm.

			Albedo sensitivity $\Delta A/\Delta\tau$					
	Planetary albedo		Hemispheric annual average		Zonal annual average (cloudy and clear combined)		Zonal annual average (cloudy and clear separate)	
Models	N.H.	S.H.	N.H.	S.H.	N.H.	S.H.	N.H.	S.H.
A	0.323	0.327	0.184	0.182	0.123	0.123	0.132	0.123
B	0.308	0.309	0.192	0.191	0.130	0.132	0.137	0.139
C	0.291	0.289	0.202	0.203	0.138	0.142	0.143	0.147
Cess	0.306	0.301	0.193	0.196	0.132	0.136	0.138	0.140

N.H. = Northern Hemisphere, S.H. = Southern Hemisphere.

true for lower latitudes when the underlying system albedo is less than 0.5. The effective albedo of a partly cloudy area lies between the cloudy and clear-sky albedos, the exact value depending on the cloud-cover fraction C[Eq. (1)]. Owing to the near linearity of the albedo sensitivity relationship, the value obtained by considering cloudy and clear-sky portions separately and considering an effective zonal albedo are very nearly the same. The significance of this result is that it is not essential to know the exact cloud cover fraction and cloudy and clear-sky albedos to compute the albedo sensitivity in this case. This argument holds for slightly absorbing aerosols as well, as shown in Fig. 5. Satellite measurements of effective zonal albedos are available by month in 10° latitude belts (Ellis and Vonder Haar, 1976) and present earth radiation budget studies are geared to obtaining this information. Therefore it is possible to estimate the albedo sensitivity of different latitude zones over the annual cycle with current data.

5. Radiative perturbations

a. Monthly albedo sensitivity

The procedure for computing the monthly albedo sensitivity is similar to the annual case. The unperturbed zonal albedo α_{old} is taken to be the mean monthly values quoted in Ellis and Vonder Haar (1976). By the discussion in the previous section, we need not take into account the cloudy and cloud-free portions of the zone separately. The reflectivity $a(\mu_0)$ of the aerosol layer is a function of the solar zenith angle, the variation of which must be included in the computation. To simplify matters, the 21st day of each month has been chosen to compute the aerosol layer reflectivity which has been averaged over the zenith angles made by the path of the sun on that day for the midpoint of each latitude belt. The albedo perturbation caused by the presence of a nonabsorbing stratospheric aerosol layer of incremental optical thickness $\Delta\tau$ will then be $(\alpha_{new} - \alpha_{old})$, where α_{new} is given by Eq. (17). With the assumption that the optical depth at 0.55 μm (denoted by τ_{vis}) is representative of the aerosol layer, the albedo sensitivity $\Delta\alpha/\Delta\tau$ obtained from (17) has been plotted in Fig. 6.

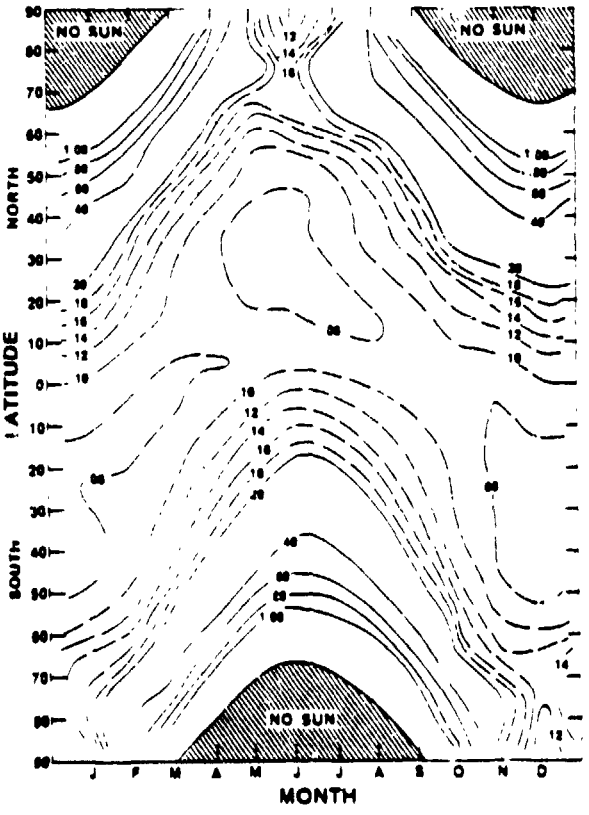


FIG. 6. Monthly zonal albedo sensitivity using model reflectivity at 0.55 μm.

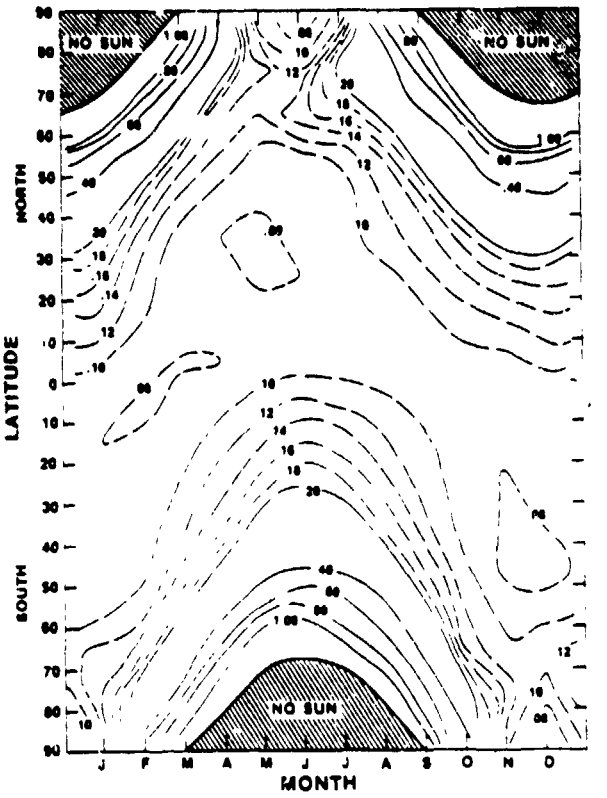


FIG. 7. As in Fig. 6 except with solar-averaged reflectivity.

The isopleths in Figs. 6 and 7 are in units of change of albedo per unit change in optical depth or, alternately, the change in albedo in percent of incident solar energy for an optical depth perturbation of $\Delta\tau = 0.01$. The dominant feature is seen to be the zenith-angle-dependence translated here in a monthly and latitudinal dependence. Over high latitudes the

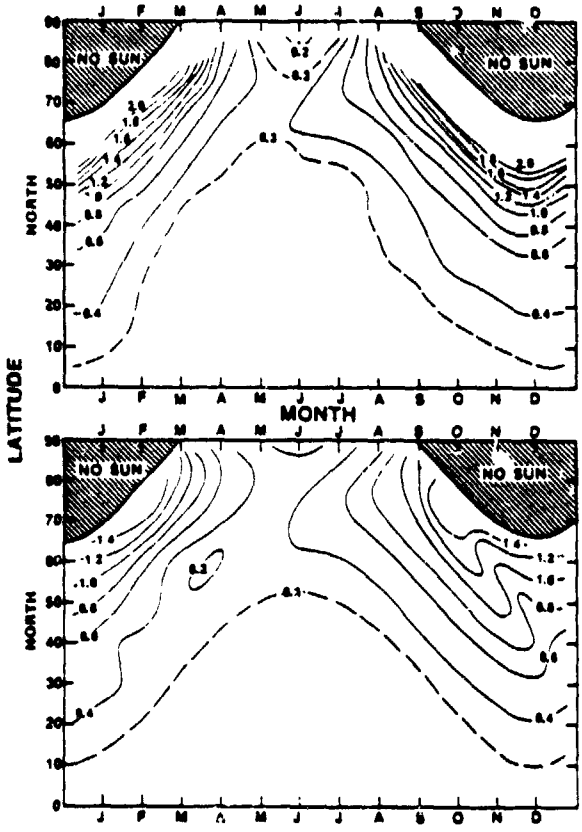


FIG. 8. Change in system albedo in percent with addition of aerosol layer of optical depth 0.03 at 0.5 μm; present model (above) and Herman *et al.* (1976) (below).

effect of surface reflectivity may also be noticed, such as the lower values of $\Delta\alpha/\Delta\tau$ in late winter and early spring when there is snow cover on the ground. This phenomenon has been discussed in Section 4 with the help of Fig. 4. Fig. 7 is a plot of the solar-averaged albedo sensitivity which is obtained by repeating the computations carried out at 0.55 μm at a number of wavelengths and taking a weighted mean. The aerosol considered here is slightly absorbing in the solar infrared and Eq. (17) may be rephrased in terms of the transmission as

$$\alpha_{new} = a + \frac{\alpha_{old} T\hat{T}}{1 - a\alpha_{old}}, \qquad (20)$$

where T is planar transmission a $\hat{T}$ is spherical.

These figures may be compared with Figs. 5 and 11, respectively, of Herman *et al.* (1976) who have plotted the change in albedo in percent for a stratospheric aerosol perturbation optical depth of 0.03 at 0.5 μm. Their results are for the Northern Hemisphere for a prescribed surface albedo variation according to three classes of surface (bare ground, open water and ice or snow). Clouds are assumed to cover 50% of the area and are simulated by results corresponding to a ground albedo of 0.5. Their results for nonabsorbing aerosols show very clearly the high-latitude regions of minimum albedo sensitivity in spring and maximum albedo sensitivity in the fall. Their results for 75% H_2SO_4 average over solar wavelengths correspond quite closely to Fig. 7 presented here except for high-latitude regions in the spring and fall. The comparison is facilitated by replotting Fig. 7 for the Northern Hemisphere for an optical depth of 0.03 at 0.5 μm. This is shown in Fig. 8 in which Fig. 11 of Herman *et al.* (1976) has been reproduced for comparison. (It should be noted that the scattering phase function used in their work is slightly different from the one used here.) One reason for the discrepancy is the low values of albedo for high-latitude regions quoted by Ellis and Vonder Haar (1976), who recognize that their results are quite uncertain in the 70–90° latitude range.

b. Radiation balance and climate modeling

The annual and monthly zonal albedo sensitivity results have important implications in energy-balance climate modeling. The sensitivity of these models is usually related to perturbations in the magnitude of the solar constant. For example, both Sellers (1969) and Budyko (1969) refer to an increase in volcanic dust loading and a decrease in the solar constant interchangeably. However, even a uniform dust layer of global extent will not lead to a uniform reduction in the solar energy received by a column of the earth–atmosphere system below the stratospheric layer. The computations presented here have shown how this reduction is distributed in latitude and with season. The basic equation of zonal energy balance on an annual mean basis is (Budyko, 1969)

$$Q(\phi)[1 - \alpha(\phi)] - I(\phi) = D(\phi), \qquad (21)$$

where $Q(\phi)$ is the solar insolation at latitude ϕ, $\alpha(\phi)$ the albedo, $I(\phi)$ the outgoing terrestrial infrared radiation and $D(\phi)$ the divergence of heat. The presence of a stratospheric aerosol layer modifies α through increased reflectivity and I through an increased greenhouse effect acting in the opposite direction. The reduction in solar energy absorbed $Q\Delta\alpha$ caused by a uniform stratospheric layer of $\tau_{vis} = 0.1$ is plotted in Fig. 9 which follows from Fig. 7 and the distribution of $Q(\phi)$ for each month.

The net loss in a column of the earth-atmosphere system, however, is $Q\Delta\alpha + \Delta I$ (note that ΔI is a gain, hence negative), which is the change in the radiation balance and it is necessary to compute the infrared greenhouse sensitivity of the aerosol layer. This is done using an atmospheric radiation model similar to the one by Harshvardhan and Cess (1976) in which an emissivity formulation has been used to calculate the infrared flux. Water vapor and carbon dioxide are considered to be the gaseous absorbers

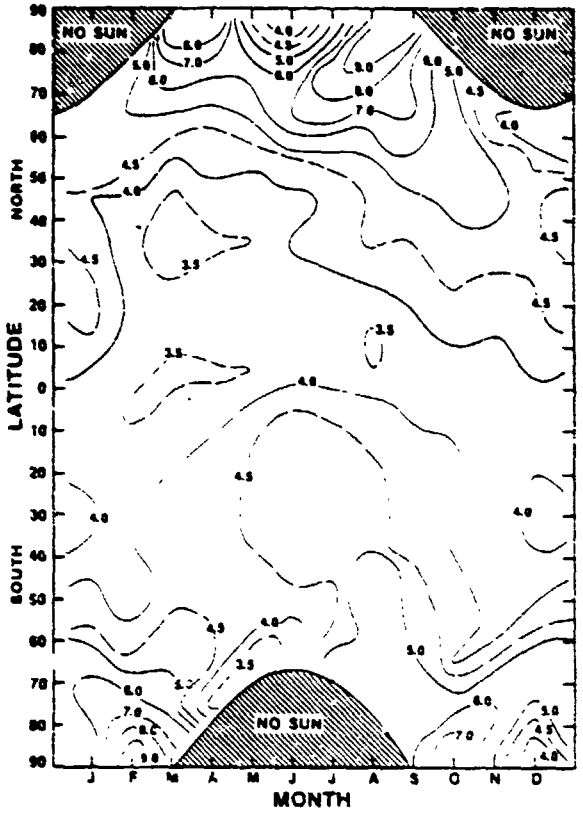


FIG. 9. Reduction in mean monthly solar energy absorbed (W m^{-2}) with addition of aerosol layer of $\tau_{vis} = 0.1$.

in the atmosphere and a 4 km thick layer of the model aerosol layer is added to determine the enhanced greenhouse effect. The computations have been carried out for the model atmospheres given by McClatchey *et al.* (1972) for clear skies and a cloudy sky with cloud top based on the tabulation of Ohring and Adler (1978) for the Northern Hemisphere. The aerosol layer was positioned to correspond to the observed level of peak concentration at different latitudes (Rosen *et al.*, 1975). Results are presented in Table 3 in which the infrared greenhouse sensitivity is expressed in (W m^{-2})/τ_{vis} and is seen to be a maximum in the tropics and a minimum for the subarctic winter model. This effect is greatest when the temperature difference between the effective radiative temperature of the atmosphere and the stratospheric aerosol layer is greatest, which is why the clear-sky cases are more sensitive than the cloud-sky cases and there is a large variation from one atmospheric model to another.

From the table it can be seen that an increase in the aerosol optical thickness of $\tau_{vis} = 0.1$, for example, would result in an increase in radiation trapped in the earth-atmosphere column under clear-sky conditions of 1.4 W m^{-2} in the tropics and 0.5 W m^{-2} in the subarctic winter. Cloudiness may be taken

TABLE 3. Infrared greenhouse sensitivity.

Model	Infrared greenhouse sensitivity (W m^{-2})/τ_{vis}	
	Clear sky	Cloudy sky
Tropical	13.9	7.7
Midlatitude summer	13.1	7.9
Midlatitude winter	8.2	5.3
Subarctic summer	10.6	6.4
Subarctic winter	5.3	4.0

into account by a weighting of the clear- and cloudy sky results. This greenhouse moderation when combined with solar depletion results in the distribution of net energy loss $Q\Delta\alpha + \Delta I$ as shown in Fig. 10 for $\tau_{vis} = 0.1$. Rather than use climatological cloudiness, the average of the clear- and cloudy-sky results has been used. Also, the greenhouse effect in spring and fall has been taken as intermediate between winter and summer. There is a secondary effect that has not been included in this analysis. Infrared absorption by the aerosol layer warms that region of the atmosphere and emission at the higher temperature moderates the infrared greenhouse sensitivity. Model calculations show this to be about 25% if the layer warms by 2.5 K for $\tau_{vis} = 0.1$.

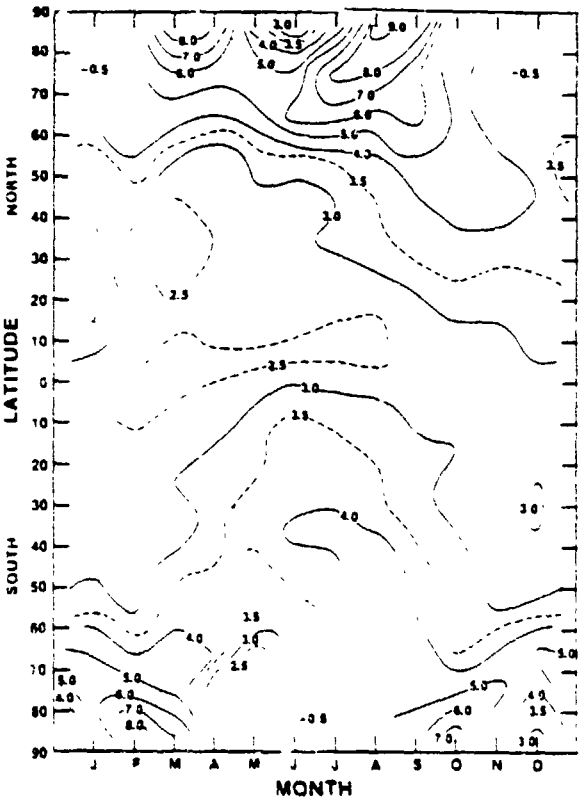


FIG. 10 Reduction in the zonal radiation balance (W m^{-2}) with addition of aerosol layer of $\tau_{vis} = 0.1$.

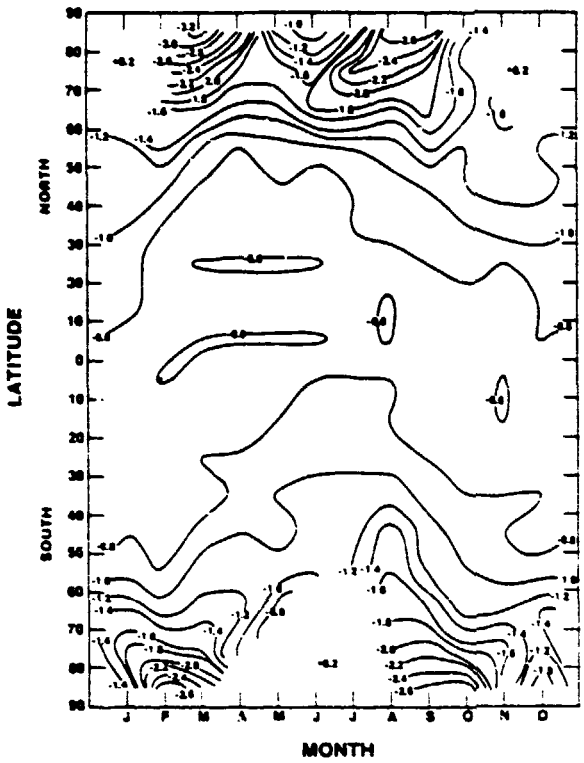


FIG. 11. Change in the effective blackbody radiative temperature of the atmosphere (K) with addition of aerosol layer of $\tau_{vis} = 0.1$.

One observation from Fig. 10 is that the presence of a stratospheric aerosol layer composed of 75% H_2SO_4 results in a loss of radiation absorbed within the earth-atmosphere system at all latitudes and seasons except the polar winter. The relationship between this net decrease in energy absorbed and mean atmospheric or surface temperature may be established through climate models. The maximum perturbation in the radiation balance occurs at high latitudes in the spring and fall, whereas polar winters show a net gain in the radiation balance due to the increased greenhouse effect. Latitudes equatorward of 50° show a nearly uniform reduction in the radiation balance.

The magnitude of the change in the net radiation is only a small percentage of the range of typical values of the radiation balance. Monthly radiation balance over all latitudes have been presented by Ellis and Vonder Haar (1976) and values range from -175 W m^{-2} in the South Polar winters to $+110$ W m^{-2} in the Southern Hemisphere tropics in summer. The uncertainty in mean monthly net radiation measurements ranges from 3 to 14 W m^{-2} such that the computed values of the change in the radiation balance caused by the aerosol layer of $\tau_{vis} = 0.1$ are less than the uncertainty in the satellite derived data. The eruption of Mt. Agung produced a perturbation of about 0.2 in τ_{vis} for about a year in the Southern Hemisphere (Volz, 1970). It is unlikely that perturbations of $\tau_{vis} > 0.1$ will occur for extended periods of time following volcanic eruptions. Furthermore, because of the relatively low sensitivity over most of the globe, even eruptions of magnitude greater than that of Agung will not appreciably alter the net radiation.

The perturbation in the radiation balance can also be expressed as a change in the effective blackbody radiative temperature of the atmosphere. This change is given by

$$\Delta T = -\frac{Q\Delta\alpha + \Delta I}{4\sigma T^3}, \tag{22}$$

where σ is the Stefan-Boltzmann constant, T the effective blackbody radiative temperature of the atmospheric column which is computed from the emitted infrared radiation tabulated in Ellis and Vonder Haar (1976), and $(Q\Delta\alpha + \Delta I)$ the decrease in the radiation balance. Fig. 11 is a plot of ΔT(K) for $\tau_{vis} = 0.1$. The major features detected in the plot of radiation balance perturbation appear enhanced in this representation. It can be seen that the major radiative effect of the stratospheric aerosol layer occurs in the spring and fall when there is a pronounced change in the equator-to-pole radiative energy gradient. As the results pertain to the entire atmospheric column, it is not possible to estimate the change in the diabatic heating at various levels of the atmosphere or estimate surface temperature changes. However, it appears from this analysis that a uniform layer of stratospheric aerosols would have only a small effect on the long term radiative regime equatorward of 50–60°, if the change in turbidity corresponded to that caused by the Agung eruption. Even this small change in the radiation balance, though probably undetectable from a space platform, could induce climatically significant changes through an alteration in the dynamics caused by the change in the equator-to-pole radiative energy gradient. Of course, this hypothesis can only be tested by a dynamic model. Also, the localized radiative effect in the vicinity of the aerosol layer would still be considerable as discussed in Harshvardhan and Cess (1976).

6. Summary

The present study has examined the role of stratospheric aerosols in altering the radiation balance of the earth–atmosphere system. It has been confirmed that for an aerosol layer of 75% H_2SO_4 the dominant influence is an increased reflection of solar energy all over the globe except for the polar winters when there is no solar insolation. An important conclusion resulting from the albedo sensitivity computations is the necessity of including solar zenith angle

effects in considering albedo enhancement. It was also pointed out that the separation of areas into cloud-covered and cloud-free is unnecessary in the calculation of the radiative effects of the reflecting layer. This enables the use of mean monthly planetary albedos obtained from satellite measurements to calculate the change in the radiation balance of each zonal belt caused by the presence of a stratospheric layer. The perturbation in the radiative balance may be used to simulate the radiative effect of a stratospheric aerosol layer in climate studies. Results indicate that the perturbation is strongest at high latitudes, particularly in the polar region, in the spring and fall.

Acknowledgments. This work was done while the author was a National Research Council Resident Research Associate at NASA/Goddard Space Flight Center, Greenbelt, Md.

REFERENCES

Braslau, N., and J. V. Dave, 1973: Effect of aerosols on the transfer of solar energy through realistic model atmospheres. Part II: Partly absorbing aerosols. *J. Appl. Meteor.*, 12, 616–619.

Budyko, M. I., 1969: The effect of solar radiation variations on the climate of the earth. *Tellus*, 21, 611–619.

——, 1974: *Climate and Life*. Academic Press, 508 pp.

Cadle, R. D., and G. W. Grams, 1975: Stratospheric aerosol particles and their optical properties. *Rev. Geophys. Space Phys.*, 13, 475–501.

Cess, R. D., 1976: Climate change: an appraisal of atmospheric feedback mechanisms employing zonal climatology. *J. Atmos. Sci.*, 33, 1831–1843.

Coakley, J. A., and P. Chýlek, 1975: The two-stream approximation in radiative transfer: including the angle of the incident radiation. *J. Atmos. Sci.*, 32, 409–418.

Curran, R. J., R. Wexler and M. L. Nack, 1978: Albedo climatology analysis and the determination of fractional cloud cover. NASA Tech. Memo. 79576, Goddard Space Flight Center, Greenbelt, MD. 45 pp.

Dave, J. V., and N. Braslau, 1975: Effect of cloudiness on the transfer of solar energy through realistic model atmospheres. *J. Appl. Meteor.*, 14, 388–395.

Dyer, A. J., and B. B. Hicks, 1968: Global spread of volcanic dust from the Bali eruption of 1963. *Quart. J. Roy. Meteor. Soc.*, 98, 545–554.

Ellis, J., and T. H. Vonder Haar, 1976: Zonal average earth radiation budget measurements from satellites for climate studies. Atmos. Sci. Pap. No. 240, Colorado State University, 50 pp.

Hansen, J., 1969: Exact and approximate solutions for multiple scattering by cloudy and hazy planetary atmospheres. *J. Atmos. Sci.*, 26, 478–487.

Harshvardhan, and R. D. Cess, 1976: Stratospheric aerosols: Effect upon atmospheric temperature and global climate. *Tellus*, 28, 1–10.

Herman, B. M., S. R. Browning and R. Rabinoff, 1976: The change in earth-atmosphere albedo and radiational equilibrium temperatures due to stratospheric pollution. *J. Appl. Meteor.*, 15, 1057–1067.

Hofmann, D. J., J. M. Rosen, T. J. Pepin and R. G. Pinnick, 1975: Stratospheric aerosol measurements I: Time variations at northern midlatitudes. *J. Atmos. Sci.*, 32, 1446–1456.

Jacobowitz, H., W. L. Smith, H. B. Howell, F. W. Nagle and J. R. Hickey, 1979: The first eighteen months of planetary radiation budget measurements from the Nimbus 6 ERB experiment. *J. Atmos. Sci.*, 36, 501–507.

Kondratyev, K. Ya., Ed., 1975: *Radiation Characteristics of the Atmosphere and the Earth's Surface*. NASA TTF-678, Amerind Publishing Co., 580 pp.

Luther, F. M., 1976: Relative influence of stratospheric aerosols on solar and longwave radiative fluxes for a tropical atmosphere. *J. Appl. Meteor.*, 15, 951–955.

McClatchey, R. A., R. W. Fenn, J. E. A. Selby, F. E. Volz and J. S. Garing, 1972: Optical properties of the atmosphere. 3rd ed. *Environ. Res. Pap.*, No. 411, AFCRL, Hanscom Air Force Base, Bedford, Mass. [NTIS AD753075].

Ohring, G., and S. Adler, 1978: Some experiments with a zonally averaged climate model. *J. Atmos. Sci.*, 35, 186–205.

Palmer, K. F., and D. Williams, 1975: Optical constants of sulfuric acid; application to the clouds of Venus? *Appl. Opt.* 14, 208–219.

Pinnick, R. G., J. M. Rosen and D. J. Hofmann, 1976: Stratospheric aerosol measurements III: optical model calculations. *J. Atmos. Sci.*, 33, 304–314.

Pollack, J. B., O. B. Toon, C. Sagan, A. Summers, B. Baldwin and W. van Camp, 1976: Volcanic explosions and climatic change: a theoretical assessment. *J. Geophys. Res.*, 81, 1071–1083.

Rosen, J. M., 1971: The boiling point of stratospheric aerosols. *J. Appl. Meteor.*, 10, 1044–1046.

——, D. J. Hofmann and J. Laby, 1975: Stratospheric aerosol measurements II: the worldwide distribution. *J. Atmos. Sci.*, 32, 1457–1462.

Russell, P. B., and R. D. Hake, Jr., 1977: The post-Fuego stratospheric aerosol: Lidar measurements, with radiative and thermal implications. *J. Atmos. Sci.*, 34, 163–177.

Sellers, W. D., 1969: A global climatic model based on the energy balance of the earth-atmosphere system. *J. Appl. Meteor.*, 8, 392–400.

Shettle, E. P., and R. W. Fenn, 1976: Models of the atmospheric aerosols and their optical properties. *AGARD Conf. Proc.*, No. 183, *Optical Propagation in the Atmosphere*, 45 pp. [NTIS ADA028615]

van Loon, H., 1972: Cloudiness and precipitation. *Meteor. Monogr.*, No. 35, Amer. Meteor. Soc., 101–111.

Volz, F. E., 1970: Atmospheric turbidity after the Agung eruption of 1963 and size distribution of the volcanic aerosol. *J. Geophys. Res.*, 75, 5185–5193.

Vonder Haar, T. H., and J. Ellis, 1975: Albedo of the cloud-free earth-atmosphere system. *Preprints Second Conf. Atmospheric Radiation*, Arlington, Amer. Meteor. Soc., 107–110.

Wiscombe, W. J., 1975: Solar radiation calculations for arctic summer stratus conditions. *Climate of the Arctic*, G. Weller and S. A. Bowling, Eds., Geophysical Institute, University of Alaska, 245–254.

——, and G. W. Grams, 1976: The backscattered fraction in two-stream approximations. *J. Atmos. Sci.*, 33, 2440–2451.

Paper 119

Reprinted with permission from *Journal of Quantitative Spectroscopy and Radiative Transfer*, 19, pp. 621-632. copyright 1978, Pergamon Press, Ltd.

EFFECT OF TROPOSPHERIC AEROSOLS UPON ATMOSPHERIC INFRARED COOLING RATES

HARSHVARDHAN† and R. D. CESS

Laboratory for Planetary Atmospheres Research, State University of New York, Stony Brook, NY 11794, U.S.A.

(*Received 7 October* 1977)

Abstract—An investigation has been made of the impact of wind-blown dust particles upon local climate of arid regions. The case of Northwest India is specifically considered, where a dense layer of dust persists for several months during the summer. In order to examine the effect of this dust layer on the infrared radiative flux and cooling rates, a method is presented for calculating the i.r. flux within a dusty atmosphere which allows the use of gaseous band models and is applicable in the limit of small single scattering albedo and pronounced forward scattering. The participating components of the atmosphere are assumed to be water vapor and spherical quartz particles only. The atmospheric window is partially filled by including the water vapor continuum bands for which empirically obtained transmission functions have been used. It is shown that radically different conclusions may be drawn on dust effects if the continuum absorption is not considered. The radiative transfer model, when applied to a dusty atmosphere, indicates that there is a moderate enhancement in the atmospheric greenhouse and a 10% increase in the mean i.r. radiative cooling rate, relative to the dust free case, within the lower troposphere. These results have been compared with previous work by other authors in the context of the possibility of dust layers inhibiting local precipitation.

1. INTRODUCTION

RECENT concern with the role of atmospheric aerosols in altering the energy budget of the earth-atmosphere system, and thus affecting global climate, has led to numerous investigations of stratospheric aerosols with long residence times [HARSHVARDHAN and CESS,[1] POLLACK *et al.*,[2] PINNICK *et al.*,[3] COAKLEY and GRAMS[4]], as well as tropospheric aerosols [RASOOL and SCHNEIDER,[5] YAMAMOTO and TANAKA,[6] RECK,[7] WANG and DOMOTO[8]] which, though typically present in higher concentrations, have residence times of the order of days or weeks. However, situations exist in which tropospheric particulate matter may remain suspended for extended periods, and through an alteration of the radiative balance affect local climate. Currently there is considerable concern with the climatic impact of dust layers, particularly over arid regions and how they might enhance desertification.

A study by BRYSON and BAERREIS[9] of the dust layer over the Rajasthan Desert in northwestern India suggests that wind-blown dust plays an important role in the radiative balance of the atmosphere and could sufficiently affect the dynamics of the region to alter precipitation over the desert. Their argument is that the dense layer of dust which persists over the entire area for several months during the summer causes a significant increase in the i.r. radiative cooling rate, which in turn increases the subsidence over the desert, thus inhibiting precipitation.

In the current study we reexamine this problem employing recent models of i.r. gaseous absorption. In order to accomplish this, it is essential to construct a radiative model for the atmosphere that incorporates wind-blown dust as an added atmospheric opacity through both absorption and scattering processes. Recent attempts to include the scattering term in the expressions for flux and cooling rates have involved the expansion of fluxes in terms of a backscattering parameter [WANG and DOMOTO[8]], the two-stream approximation [RECK[7]], and a four-stream discrete ordinate method [ACKERMAN *et al.*[10]]. An earlier model, due to SARGENT and BECKMAN,[11] used a semi-isotropic scattering model within the equation of transfer which was solved numerically over the entire frequency range. The present work employs an exponential kernel approximation within the radiative transfer equation, which yields optical depth

†Present address: NASA Goddard Space Flight Center, Greenbelt, MD 20771, U.S.A.

parameters that can very easily be applied to atmospheric band models. The scattering effects are represented in terms of gross scattering parameters such as the single scattering albedo, $\tilde{\omega}$, and the asymmetry factor, $<\cos\theta>$, that still retain the non-grayness and anisotropy of the scattering process. The detailed angular dependence of intensity is lost, but as we are interested only in the radiative flux, this is not a handicap.

The radiative model developed here has been used to estimate the effect of a desert aerosol layer on i.r. cooling rates. In order to keep the variables to a minimum, water vapor has been taken to be only gaseous absorber and spherical quartz dust particles to be the only scatters. Data obtained by PETERSON,[12] from measurements made over the Rajasthan Desert of north-western India, have been used in the study and climate modifying effects have been evaluated.

2. WATER VAPOR MODEL AND FLUX COMPUTATION

The object of any radiative transfer model is to represent the selective absorption by atmospheric constituents in a realistic manner. The emissivity formulation given by RODGERS[13] is simple yet accurate for flux computations and has been employed here. The MAYER–GOODY[14] statistical model has been used to represent the rotation band and the 6.3 μm vibration–rotation band. The continuum bands, on the other hand, are represented by empirical fits taken from the work of RAMANATHAN[15] and ROBERTS *et al.*[16]

The upward and downward fluxes in a cloudless atmosphere have been given by Rodgers in terms of T_ν and T^*_ν, defined below, and yield

$$F(P) = \int_{P_0}^{P} \pi B(\theta)\frac{\partial T^*_\nu\uparrow}{\partial P'}\,dP' - \int_{P}^{0} \pi B(\theta)\frac{\partial T_\nu\downarrow}{\partial P'}\,dP' + \pi B(\theta_0)T^*_\nu\uparrow. \tag{1}$$

At any pressure level, P, the net radiative flux is $F(P)$, where P_0 is the surface pressure and θ_0 the surface temperature. Here $T^*_\nu\uparrow$ is the modified upward slab transmissivity defined by eqn (3) below, evaluated over a path from P_0 to P, while $T^*_\nu\downarrow$ is the downward slab transmissivity defined by eqn (2) and evaluated over a path from P to the top of the atmosphere. The slab transmissivity is defined analogous to the gas emissivity [Chap. 5 of GOODY[14]] and, in the strong line limit [RODGERS,[13] GOODY[14] and CESS[17]],

$$T_\nu(\hat{a},\theta) = T_c(r_d\hat{a},\theta) = \frac{\sum_i \bar{T}_{\Delta\nu_i}(r_d\hat{a})B_i(\theta)\Delta\nu_i}{\sum_i B_i(\theta)\Delta\nu_i} \tag{2}$$

while the modified slab transmissivity is given by [Chap. 6 of GOODY[14]]

$$T^*_\nu(\hat{a},\theta) = T^*_c(r_d\hat{a},\theta) = \frac{\sum_i \bar{T}_{\Delta\nu_i}\left(\frac{dB}{d\theta}\right)_i\Delta\nu_i}{\sum_i \left(\frac{dB}{d\theta}\right)_i\Delta\nu_i}. \tag{3}$$

Here $\hat{a} = a\bar{P}$, where a is the absorber amount and $\bar{P}$ the Curtis–Godson broadening pressure [GOODY,[14] p. 238] which accounts for the inhomogeneity of the atmospheric path, r_d is the diffusivity factor taken to be 1.66, $\bar{T}_{\Delta\nu_i}$ is the mean transmission of the ith spectral range of frequency width $\Delta\nu_i$, and $B_i(\theta)$ is the Planck function at its center with θ the absolute temperature. The suffix "c" refers to column transmissivities.

Thus the calculation of the radiative flux at any level within an atmosphere with a given distribution of temperature, pressure and absorber amount reduces to that of evaluating the transmission, $\bar{T}_{\Delta\nu}$. Using Goody's model of random line positions with an exponential distribution of intensity, in the strong line limit, the transmission in the spectral regions of the rotation band may be expressed as [RODGERS[13]]

$$\bar{T}_{\Delta\nu} = \exp\left\{-\frac{2}{\Delta\nu}(\hat{a})^{1/2}\sum_{j=1}^{N}[S(j)\alpha_L(j)]^{1/2}\right\}. \tag{4}$$

where $\Delta\nu$ is the band interval having N lines of intensity $S(j)$ and Lorentz half-width $\alpha_L(j)$. The quantity $\Sigma[S(j)\alpha_L(j)]^{1/2}$ is tabulated at 20 cm^{-1} intervals from 0 to 1000 cm^{-1} on p. 184 of GOODY[14] for three temperatures. We select 260°K for use in this study. For this and other bands, the temperature dependence of the transmission has been neglected following RODGERS,[13] who states that with the degree of approximation involved in using theoretically computed emissivities to calculate the flux in a real atmosphere, nothing is gained by introducing a temperature dependence for the transmission $\bar{T}_{\Delta\nu}$. For the 6.3 μm band, line parameters obtained from the work of BENEDICT and CALFEE[18] [who have enumerated $S(j)$ and $\alpha_L(j)$ at 287.7°K] were used to compute the transmission employing eqn (4), with 20 cm^{-1} intervals from 1000 to 2000 cm^{-1}, but summing over the individual lines.

In this work, the continuum bands have also been included since they contribute to the opacity within the atmospheric window, where dust extinction is expected to have a maximum influence. An average transmission through each of the regions was calculated using analytical expressions provided by RAMANATHAN,[15] who fitted experimental data, while we have followed the suggestion of ROBERTS *et al.*[16] that only the self-broadened component of the continuum absorption in the 800–1200 cm^{-1} region be retained. Where the continuum bands overlap portions of other bands, the multiplicative property of the mean spectral transmission has been employed [GOODY[14]].

It remains to specify the vertical distribution of water vapor and temperature within the atmosphere and, for illustrative purposes, we simply choose the global model employed by CESS.[17] The vertical distribution of water-vapor partial pressure, P_w, is given by

$$P_w/P_{w0} = \exp(-z/H_w), \tag{5}$$

where $P_{w0} = 0.0128$ atm is the global average for P_w at the surface, z is altitude and H_w the water vapor scale height given by

$$H_w = \frac{H}{1 + (5385\, R\Gamma/g\theta_0)} \tag{6}$$

with H the atmospheric scale, R and g the gas constant and acceleration due to gravity, respectively, while $\Gamma = -d\theta/dz$ denotes the lapse rate and θ_0 the surface temperature. With this distribution, the scaled absorber amount, $\tilde{a}$, may be expressed in terms of the atmospheric and water-vapor scale heights and pressures [CESS[17]].

Although a standard atmosphere or a particular sounding may be used with the model, a hydrostatic atmosphere with a constant scale height, H, evaluated at 260°K, has been employed. The surface pressure, P_0, is taken to be 1 atm and surface temperature, θ_0, to be 288°K, which is a global mean. A constant lapse rate within the troposphere, $\Gamma = 6.5$°K km^{-1}, prescribes the temperature distribution.

3. ATMOSPHERIC MODEL WITH DUST EXTINCTION

The presence of a dust layer within the atmosphere makes the computation of the radiative flux more complicated since the contribution to the extinction of radiation has a scattering component as well as an absorbing component. In this case, the solution of the equation of transfer for the flux cannot be written in terms of the transmissivity as in eqn (1), and band models for gaseous absorption cannot be used. However, the technique presented here permits flux computations to be made using eqn (1) under certain restrictive conditions, which are satisfied by the dust model considered here and could probably be satisfied by other polydispersions.

The procedure consists of solving the equation of transfer in terms of the mean intensity, using an exponential kernel approximation following WANG.[19] Details of the procedure may be obtained from HARSHVARDHAN.[20] The Eddington phase function [SHETTLE and WEINMAN[21]] is employed, such that

$$p(\cos\theta) = 1 + 3(\cos\theta)\cos\theta, \tag{7}$$

where $\langle\cos\theta\rangle$ is the asymmetry factor with θ the scattering angle and $p(\cos\theta)$ satisfies the normalization condition

$$1/2\int_{-1}^{1} p(\cos u)\,\mathrm{d}(\cos\theta) = 1. \tag{8}$$

Further,

$$1/2\int_{-1}^{1} p(\cos\theta)\cos\theta\,\mathrm{d}(\cos\theta) = \langle\cos\theta\rangle. \tag{9}$$

With this choice of $p(\cos\theta)$, the solution of the equation of transfer is simplified but gross anisopropy effects, which influence the emergent flux, are retained. It is assumed that the optical properties of the atmosphere are homogeneous. This implies that the single scattering albedo, $\tilde{\omega}$, is independent of altitude. The single scattering albedo may be written as

$$\tilde{\omega} = \frac{\beta_{sc}^{d}}{\beta_{ex}^{d} + \beta_{ab}^{g}}, \tag{10}$$

where β_{sc}^{d} and β_{ex}^{d} are, respectively, the monochromatic scattering and extinction coefficients for the polydispersion and β_{ab}^{g} is the monochromatic absorption coefficient of the gas (water vapor). In a realistic atmosphere $\tilde{\omega}$ will vary with altitude, as these quantities depend upon the concentration, but here the value of the single scattering albedo at the ground is taken to be the uniform value of $\tilde{\omega}$ throughout the atmosphere. This is the "worst-case" choice for $\tilde{\omega}$ if dust extinction falls off more rapidly with altitude than gaseous absorption, as is shown later.

The atmosphere is represented schematically in Fig. 1. The equation of transfer is given by

$$\mu\frac{\mathrm{d}}{\mathrm{d}\tau}I(\tau,\mu) = I(\tau,\mu) - \frac{\tilde{\omega}}{2}\int_{-1}^{1}\mathrm{d}\mu' p(\mu,\mu')I(\mu') - (1-\tilde{\omega})B(\tau) \tag{11}$$

for axisymmetric radiation. All radiation quantities are monochromatic and the suffix ν is suppressed for clarity. Here I is the monochromatic intensity, B is the Planck function, μ is the cosine of the scattering angle θ, and τ is the optical depth. Introducing the mean intensity J and the radiative flux, $F = 4\pi\bar{F}$, substitution of the expression for the phase function in eqn (11) yields

$$\mu\frac{\mathrm{d}}{\mathrm{d}\tau}I(\tau,\mu) = I(\tau,\mu) - \tilde{\omega}J(\tau) - 3\tilde{\omega}\langle\cos\theta\rangle\mu\bar{F}(\tau) - (1-\tilde{\omega})B(\tau), \tag{12}$$

where

$$J = \frac{1}{2}\int_{-1}^{1}\mathrm{d}\mu I(\mu)$$

and

$$\bar{F} = \frac{1}{2}\int_{-1}^{1}\mathrm{d}\mu\mu I(\mu).$$

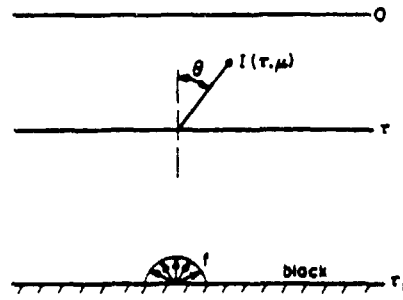


Fig. 1. Plane parallel atmosphere with diffuse surface radiation.

Equation (12) yields the expression for the rate of radiation leaving an element of unit area. Upon integration over μ,

$$\frac{d\bar{F}}{d\tau} = (1-\tilde{\omega})(J-B). \tag{13}$$

The solution of eqn (12) yields $\bar{F}(\tau)$ in terms of J and B. Following WANG,[19] if the exponential integral functions E_2 and E_3 are replaced by exponentials, a second-order differential equation in the mean intensity, J, can be obtained.

This differential equation, along with the appropriate boundary conditions, are listed below for the approximation $E_2(t) \approx a\exp(-bt)$ and $E_3(t) \approx a/b\exp(-bt)$ and are

$$\frac{d^2J}{d\tau_{eff}^2} - J = \frac{(1-a)(1-\tilde{\omega})}{1-\tilde{\omega}(1-a)}\frac{d^2B}{d\tau_{eff}^2} - B, \tag{14}$$

$$\frac{dJ(0)}{d\tau_{eff}} - v\frac{dB(0)}{d\tau_{eff}} = u[J(0) - vB(0)], \tag{15a}$$

$$\frac{dJ(\tau_{1eff})}{d\tau_{eff}} - v\frac{dB(\tau_{1eff})}{d\tau_{eff}} = -u[J(\tau_{1eff}) - vB(\tau_{1eff}) - \Phi], \tag{15b}$$

where

$$\tau_{eff} = \left(\frac{(1-\tilde{\omega})(b^2 - 3a\tilde{\omega}\langle\cos\theta\rangle)}{1-\tilde{\omega}(1-a)}\right)^{1/2}\tau, \tag{16}$$

$$u = \left\{\frac{\left(1 - \frac{3a\tilde{\omega}\langle\cos\theta\rangle}{b^2}\right)(1-\tilde{\omega}(1-a))}{(1-\tilde{\omega})}\right\}^{1/2}, \tag{17}$$

$$v = \frac{(1-\tilde{\omega})(1-a)}{1-\tilde{\omega}(1-a)}, \tag{18}$$

and

$$\Phi = \frac{4a}{1-\tilde{\omega}(1-a)}\frac{f}{4\pi}, \tag{19}$$

with f the surface emission at $\tau = \tau_1$. Furthermore, in the derivation of eqn (14), we also obtain the following expression for the radiative flux:

$$\bar{F}(\tau_{eff}) = \frac{1-\tilde{\omega}(1-a)}{bu}\left(\frac{dJ(\tau_{eff})}{d\tau_{eff}} - v\frac{dB(\tau_{eff})}{d\tau_{eff}}\right). \tag{20}$$

For monochromatic radiation, the above set of equations provides an approximate solution of the equation of transfer. Here the monochromatic solution is utilized to modify the transmission computed from atmospheric band models employed in the present work that are used to calculate the flux in a dust-free atmosphere. It is possible to obtain a solution for the band-averaged flux in a form that can be directly applied to atmospheric band models if the single scattering albedo, $\tilde{\omega}$, is small, a condition that is met in the i.r. region where gaseous absorption dominates.

With this assumption, the solution of eqn (14), subject to boundary conditions[15] and with the use of eqn (20), yields the following for the monochromatic radiative flux:

$$F\uparrow = \int_{\tau_{eff}}^{\tau_{1eff}} \pi B(t)\exp[-(t-\tau_{eff})]\,dt + f\exp[-(\tau_{1eff}-\tau_{eff})] + 0(\delta),$$

$$F\downarrow = -\int_{0}^{\tau_{eff}} \pi B(t)\exp[-(\tau_{eff}-t)]\,dt + 0(\delta), \tag{21}$$

where the perturbation quantity δ is related to the scattering properties by

$$\delta = \left\{ \frac{\left(1 - \frac{3a\tilde{\omega}\langle\cos\theta\rangle}{b^2}\right)[1 - \tilde{\omega}(1-a)]}{(1-\tilde{\omega})} \right\}^{1/2} - 1. \tag{22}$$

For pure absorption, $\tilde{\omega} = 0$, and by the above, $\delta = 0$. For the case of small $\tilde{\omega}$ and neglecting $0(\tilde{\omega}^2)$, we have

$$\delta \simeq \frac{a\tilde{\omega}}{2}\left(1 - \frac{3}{b^2}\langle\cos\theta\rangle\right). \tag{23}$$

Sample calculations, applying to the atmospheric window where the aerosol effect is of greatest importance, illustrate that the terms of $0(\delta)$ within eqn (21) are small. Moreover, their neglect yields an upper bound (or "worst case") with respect to aerosol influence upon the radiative flux, consistent with our use of eqn (10). We thus neglect the $0(\delta)$ term in eqn (21) and, since $b^2/3 \simeq 1$ (see following discussion), this effectively constitutes an approximation for strong forward scattering, in that $\delta \simeq 0$ for $\langle\cos\theta\rangle \to 1$. The effect of scattering is, however, still retained through the scaling in optical depth as given by eqn (16) which, in the absence of scattering, reduces to $\tau_{\text{eff}} = b\tau$ whereas, for small $\tilde{\omega}$,

$$\frac{\tau_{\text{eff}}}{b\tau} = 1 - \frac{a\tilde{\omega}}{2}\left(1 + \frac{3}{b^2}\langle\cos\theta\rangle\right). \tag{24}$$

Thus, while terms of $0(\delta)$ within eqn (21) vanish in the limit of pure forward scattering, this limit produces the maximum influence upon the scaled optical depth τ_{eff}.

Replacing monochromatic quantities in eqn (21) by their average values over a small frequency interval $\Delta\nu$ and noting that $\exp(-|\tau_{\text{eff}} - t|)$ is the monochromatic transmission, eqns (2) and (3) defining the transmissivity can be used to convert eqn (21) to eqn (1), which gives the radiative flux in terms of the transmissivity. The substitution of an exponential kernel automatically converts the column transmissivity to a slab transmissivity with b the diffusivity factor ($b = r_d$). In the previous section, a value of 1.66 for the diffusivity factor was adopted for gaseous absorption. For compatibility, the same value should be used so that $b = 1.66$. The ratio a/b is 0.5 since $E_3(0) = 0.5$ [SPARROW and CESS[22]] so that $a = 0.83$. With this choice, eqn (16) may then be written as

$$\tau_{\text{eff}} = 1.66\left\{ \frac{(1-\tilde{\omega})(1 - 0.9\,\tilde{\omega}\langle\cos\theta\rangle)}{(1 - 0.17\,\tilde{\omega})} \right\}^{1/2}\tau. \tag{25}$$

Here τ includes the contribution of dust extinction which is $\tau_\alpha^d = \beta_\alpha^d h$, with h the height of the dust column such that $\tau = \tau_{ab}^g + \tau_\alpha^d$. The changed optical path has a physical interpretation in that it takes into account the forward (or backward) scattering of radiation by the polydispersion. For fixed $\tilde{\omega}$, a positive value of $\langle\cos\theta\rangle$ (forward scattering) reduces the effective optical path relative to the isotropic scattering case whereas backward scattering has the opposite effect.

When eqn (25) is used in band models for the gaseous absorber, it is used in the form

$$\bar{a}_{\text{eff}}(i) = 1.66\left\{ \frac{(1-\bar{\omega}_i)(1 - 0.9\,\bar{\omega}_i\langle\cos\theta\rangle_i)}{(1 - 0.17\,\bar{\omega}_i)} \right\}^{1/2}\bar{a}_i \tag{26}$$

where the suffix i stands for the ith spectral interval, $\Delta\nu_i$, and $\bar{\omega}_i$ and $\langle\cos\theta\rangle_i$ represent quantities averaged over $\Delta\nu_i$.

The approach to the problem of radiative transfer presented here is mathematically similar to the method of WANG and DOMOTO,[8] hereinafter referred to as WD. The governing equations of the diffuse field presented as eqn (4) in WD are equivalent to the exponential kernel approximation with $a = 1$ and $b = \sqrt{3}$. Their equations, in a more generalized form, include

irradiance at the upper free surface and a non-black lower surface. Using the above values of a and b, parameters defined by eqns (16) and (17) become

$$\tau_{\text{eff}} = \sqrt{3}\tau\{(1-\tilde{\omega}\langle\cos\theta\rangle)\}^{1/2} \tag{27}$$

and

$$u = \{(1-\tilde{\omega}\langle\cos\theta\rangle)/(1-\tilde{\omega})\}^{1/2}. \tag{28}$$

WD define a backscattering parameter, β, as

$$2\beta = 1-\langle\cos\theta\rangle. \tag{29}$$

With this definition of β, the phase functions used in the two investigations are identical. The radiative flux is expanded in terms of powers of β which, for typical polydispersions, is small, around 0.15 for aerosols and hazes in the Earth's atmosphere. The zeroth order equations of WD ($\beta = 0$) correspond to totally forward scattering, or $\langle\cos\theta\rangle = 1$. In this limit, from eqn (27),

$$\tau_{\text{eff}} = \sqrt{3}\tau(1-\tilde{\omega}) \tag{30}$$

and it is this optical depth scaling that WD employed, irrespective of the value of $\langle\cos\theta\rangle$. The first and higher-order equations are corrections to the lead term to take into account backscattering. These correction terms, although multiplied by powers of β, can also be seen to be of the form $(\tilde{\omega}\beta)^n$ from eqns (16) and (17) of WD. It should be noted that WD have identified the albedo as a, and the backscatter parameter as b. From eqn (23) above, for $a = 1$ and $b = \sqrt{3}$, the correction terms in the present work are of $0(\delta) = 0(\beta\tilde{\omega})$ and, therefore, WD's expression for flux is also of the form

$$\text{lead term} + 0(\delta) + \ldots$$

Scaling τ_{eff} as in eqn (30) is equivalent to adding on the absorption component of the particulate only, and correction terms are required so that the flux may approach the correct value. Wishing to use their formulation at solar wavelengths where $\tilde{\omega}$ may not be small, WD have restricted themselves to a single perturbation about $\langle\cos\theta\rangle = 1$. In the present work, we have effectively perturbed about $\langle\cos\theta\rangle \approx 1$ and $\tilde{\omega} = 0$ simultaneously. For this reason, terms of $0(\delta)$ are not required in the double limit $\langle\cos\theta\rangle \to 1$ and $\tilde{\omega} \to 0$ as in the shortwave region.

A comment is in order here on the single scattering albedo defined by eqn (10). Whereas dust extinction is linear with amount (for a fixed particle size distribution), the band absorption of water vapor in most spectral regions varies as the square-root of the amount of gas [eqn (4)]. Therefore, the appropriate value of $\tilde{\omega}$ for each spectral range is more suitably written as

$$\tilde{\omega} = \frac{\tau_{sc}^{d}}{\tau_{ex}^{d} + \tau_{ab}^{g}} = \frac{\tilde{\omega}^{d}\tau_{ex}^{d}}{\tau_{ex}^{d} + \tau_{ab}^{g}}, \tag{31}$$

where superscript d refers to dust alone and superscript g refers to water vapor alone. Choosing the ground level to evaluate $\tilde{\omega}$, such that τ refers to the total optical depth, will place an upper bound on dust effects if τ_{ex}^{d} falls off more rapidly with altitude than τ_{ab}^{g}. This may be verified for typical tropospheric conditions where $H \approx 8$ km, $H_w \approx 2$ km and $H_d \approx 1.2$ km [ELTERMAN[23]] with H, H_W and H_D the atmospheric, water-vapor and dust scale heights respectively.

4 RESULTS AND DISCUSSION

Based on the radiative model developed in the previous sections, sample computations of the flux and cooling rates have been made for a tropospheric dust model which represents conditions found over tropical deserts. Utilizing measurements made over the Rajasthan Desert

of northwestern India, the particle size distribution adopted by PETERSON and WEINMAN[24] is

$$\left.\begin{aligned}\frac{dN}{dr} &= 13\, r^{-3.5} \quad \text{for} \quad 0.1 \le r \le 15\ \mu\text{m},\\ \frac{dN}{dr} &= 4.1 \times 10^{4} \quad \text{for} \quad 0.03 \le r \le 0.1\ \mu\text{m},\end{aligned}\right\} \tag{32}$$

where dN/dr is the number of particles per cubic centimeter per micron within the radius interval dr. The authors have calculated β_{α}^{d}, $\tilde{\omega}^{d}$ and $\langle\cos\theta\rangle$ as a function of frequency for spherical quartz particles. An exponential variation of dust concentration with height of the form

$$N(z) = N(0)\exp(-z/H_d) \tag{33}$$

has been adopted. Here $N(z)$ refers to the total concentration in number of particles per unit volume at any altitude z with H_d the dust scale height.

As long as the form of the particle size distribution is unchanged, the dust extinction coefficient, $\beta_{\alpha}^{d}(z)$, is directly proportional to $N(z)$, and so a natural parameter to use in presenting the results of the computations is the total optical depth at some reference wavelength. Although computations are being made for i.r. wavelengths, we can adopt as a reference the mean visible wavelength $0.55\,\mu$m, as there is a unique relationship between $\tau(0.55\,\mu\text{m})$ and the optical depth at any other wavelength. The optical depth at $0.55\,\mu$m is generally referred to as the visible optical depth, τ_{vis}.

Computations made for the upward, downward and net flux, using eqn (1) with transmissivities calculated from the modified transmission functions, are presented in Figs. 2 and 3 for the lowest 7 km of the troposphere with and without a dust layer of $H_d = 2$ km. The surface dust concentration is that given by eqn (32). The i.r. radiative cooling rate, defined as

$$\frac{d\theta}{dt} = -\frac{1}{\rho(z)C_p}\frac{dF(z)}{dz} \tag{34}$$

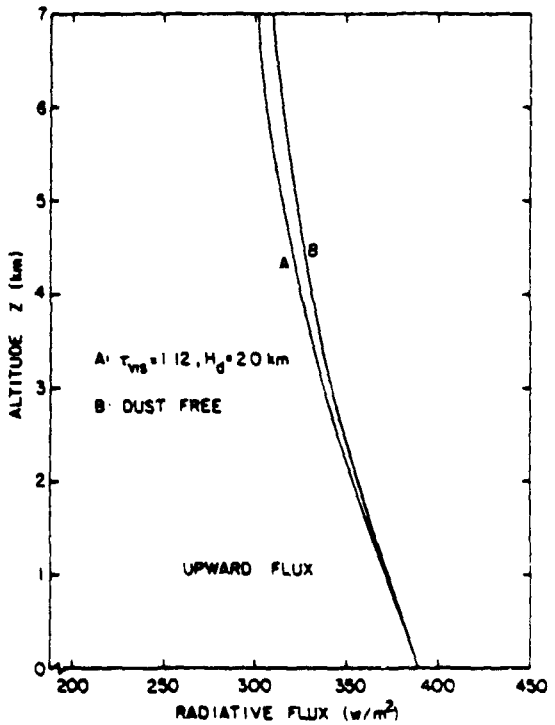


Fig. 2. Upward i.r. flux in the lower troposphere.

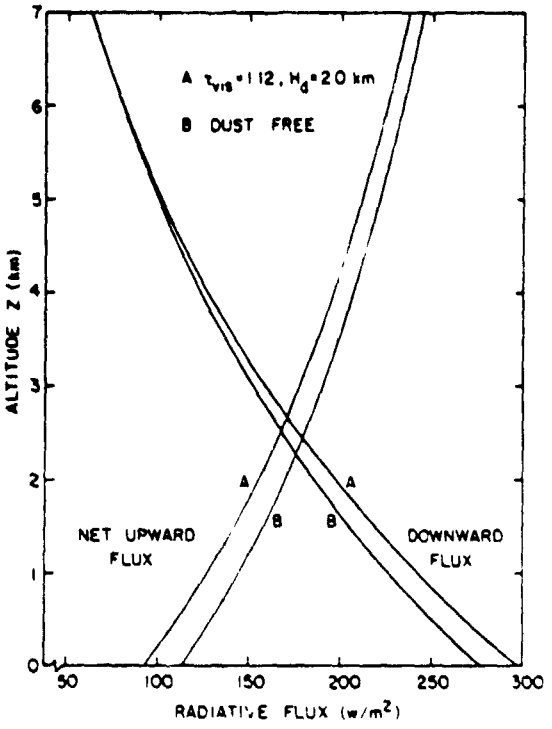


Fig. 3. Downward and net infrared flux in the lower troposphere.

for the dust free atmosphere, is compared with that for dusty atmospheres with $H_d = 0.5$, 1 and 2 km in Fig. 4 and with atmospheres for which $H_d = 2$ km but varying surface dust concentrations in Fig. 5.

The changes in radiative flux could have been anticipated. An enhanced greenhouse effect may be noted in Fig. 3, the net upward flux leaving the atmosphere being reduced by 3%. The net upward flux at the surface, however, is reduced by 10% owing to the strongly enhanced downward emission. Of possible climatic interest is the change in the cooling rate profile. Even a small amount of dust increases the cooling rate near the surface and tends to smooth out regions of minimum cooling in the troposphere. DAS,[25] in a study of mean vertical motion and non-adiabatic heat sources over India during the monsoon, had concluded that an additional source of radiative cooling was required to account for the theoretically calculated subsidence over NW India. BRYSON and BAERREIS[9] suggested that the heavy dust layer over this region could account for this discrepancy. Their calculations showed that a 50% increase in the mean cooling rate of the lower troposphere was possible.

Our results indicate a substantial increase in the cooling rate near the surface, but the mean increase in the cooling rate throughout the lower troposphere is only 10%. This wide difference in predicted dust effects is not a result of our water vapor distribution, corresponding to global mean conditions. Comparison with ascent data over NW India given in PETERSON[12] shows that the precipitable water in the lowest 400 mb of the theoretical model used here is actually about half that in the same column computed from the sounding. Thus, if we were to double our water vapor concentration, the dust contribution would be even less. Moreover, the approximations employed in our model tend at most to overestimate the dust effects. The discrepancy actually arises from the modeling of water vapor absorption, particularly in the critical region of the atmospheric window. The inclusion of an adequate representation of the water vapor continuum greatly reduces the effect of dust opacity. This may be illustrated by considering different water vapor continuum models. The work of DAS,[25] BRYSON and BAERRIES[9] and SARGENT and BECKMAN[11] used the representation given by ELSASSER and CULBERTSON,[26] which underestimates the window absorption by an order of magnitude. Figure 6 illustrates the drastic effect on the clear sky radiative cooling rate of adopting different continuum models. The

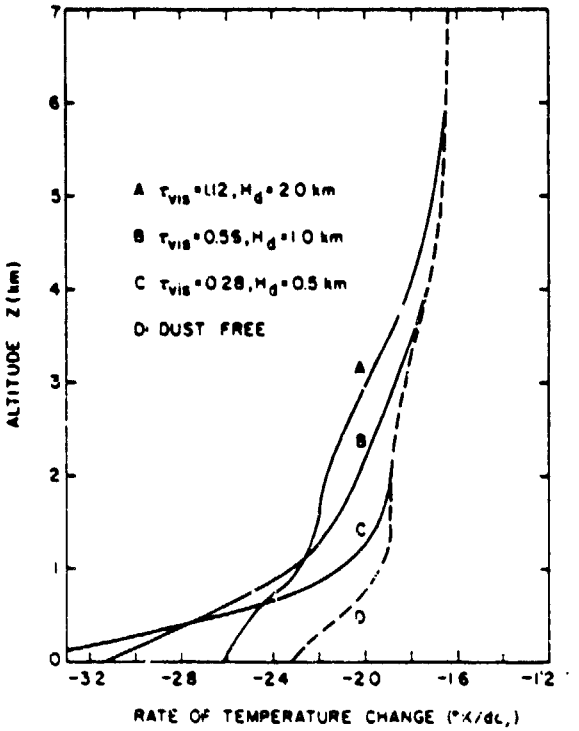


Fig. 4. Infrared cooling rate as a function of dust scale height.

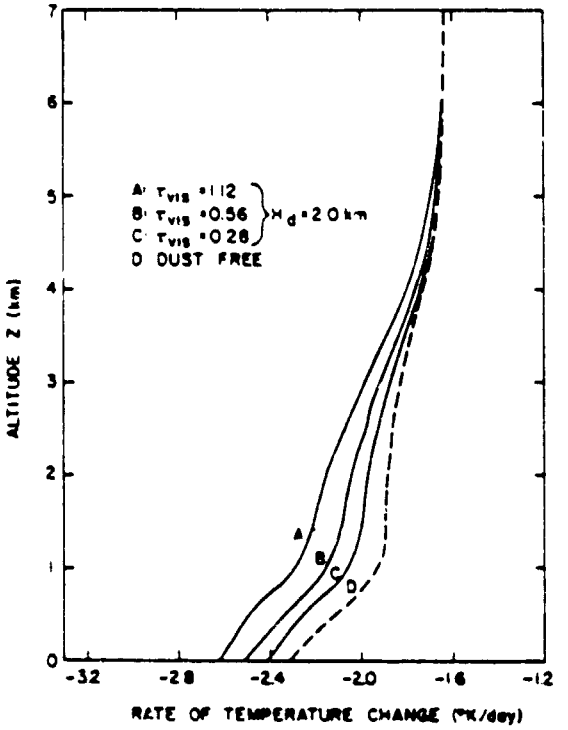


Fig. 5. Infrared cooling rate as a function of dust surface concentration.

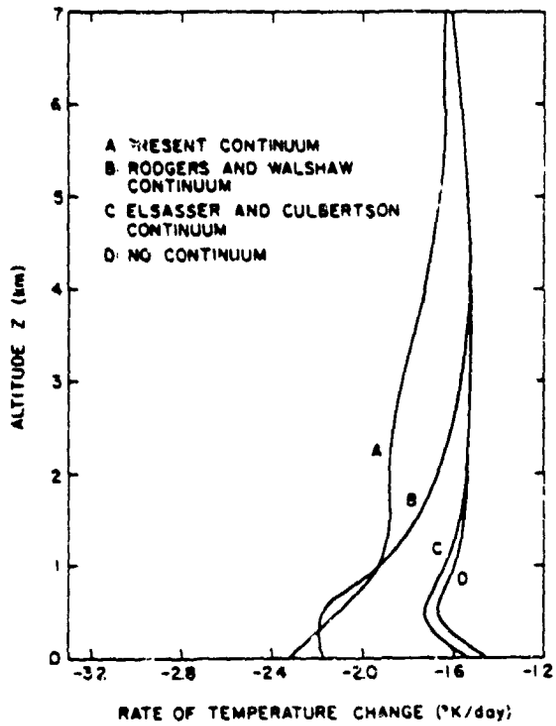


Fig. 5. Infrared cooling rate for different water vapor continuum models.

results strongly suggest that an estimate of radiative cooling, based on a more adequate water vapor model, would have accounted for almost all the cooling required for the subsidence calculated over the desert region by Das. Therefore, it is our belief that the dust layer alone cannot alter the magnitude of the local subsidence appreciably. There is a fairly significant change in the cooling rate, of about 25%, in localized regions of the troposphere, but whether this has any stabilizing effect that inhibits precipitation processes is open to question. The very slight increase in i.r. cooling rates in the presence of aerosols is in agreement with the work of GRASSL,[27] who included aerosol absorption within the atmospheric window region to compute changes in cooling rate and found little change for tropical atmospheres as well as for the U.S. Standard Atmosphere.

5 SUMMARY

In this study, the possible climatic effects of long residence time tropospheric dust layers has been investigated, particularly with respect to changes in the i.r. radiative cooling rate. To accomplish this, a radiative model was adopted for gaseous absorption and modified to include extinction by a polydispersion. This modified radiative model is in a form suitable for use in other atmospheric studies such as the effect of urban pollution on the local radiative balance, as well as numerical experiments with general circulation models in which computation speed is of high priority.

The results of the investigation show that there is a mean increase in the radiative cooling in the presence of heavy dust layers of the order of 10%, although locally the increase may be as high as 25% in particular regions of the lower troposphere. The significance of this change in altering the dynamics over arid areas and effecting precipitation cannot be determined on the basis of work to date. However, the important economic and social implications of altered precipitation patterns over arid regions would make this area of study most worthwhile.

Acknowledgement—This work was supported by the National Science Foundation through Grant ENG-7682547

REFERENCES

1. HARSHVARDHAN and R. D. CESS, *Tellus* **28**, 1 (1976).
2. J. B. POLLACK, C. SAGAN, A. SUMMERS, B. BALDWIN and W. VAN CAMP, *J. Geophys. Res.* **81**, 1071 (1976).
3. R. G. PINNICK, J. M. ROSEN and D. J. HOFFMANN, *J. Atmos. Sci.* **33**, 304 (1976).
4. J. A. COAKLEY, JR. and G. W. GRAMS, *J. Appl. Met.* **15**, 679 (1976).
5. S. I. RASOOL and S. H. SCHNEIDER, *Science* **173**, 138 (1971).
6. G. YAMAMOTO and M. TANAKA, *J. Atmos. Sci.* **29**, 1405 (1972).
7. R. A. RECK, *Atmos. Environ.* **8**, 823 (1974).
8. W. C. WANG and G. A. DOMOTO, *J. Appl. Met.* **13**, 521 (1974).
9. R. A. BRYSON and D. A. BAERREIS, *Bull. Am. Met. Soc.* **48**, 136 (1967).
10. T. P. ACKERMAN, K. LIOU and C. B. LEOVY, *J. Appl. Met.* **15**, 28 (1976).
11. S. L. SARGENT and W. A. BECKMAN, *J. Atmos. Sci.* **30**, 88 (1973).
12. J. T. PETERSON, *ONR Tech. Rept.* **38**, Dept. of Meteorology, Univ. of Wisconsin, Madison, Wis. (1968).
13. C. D. RODGERS, *Q. J. Roy. Met. Soc.* **93**, 43 (1967).
14. R. M. GOODY, *Atmospheric Radiation*. Oxford University Press, London (1964).
15. V. RAMANATHAN, *J. Atmos. Sci.* **33**, 1330 (1976).
16. R. E. ROBERTS, J. E. A. SELBY and L. M. BIBERMAN, *Appl. Opt.* **15**, 2085 (1976).
17. R. D. CESS, *JQSRT* **14**, 861 (1974).
18. W. S. BENEDICT and R. F CALFEE, *ESSA Professional Paper* 2 U.S. Dept. of Commerce, Washington, D.C. (1967).
19. L. WANG, *Ap. J.* **174**, 671 (1972).
20. HARSHVARDHAN, Ph.D. Thesis, State University of New York at Stony Brook, Stony Brook, N.Y. (1976).
21. E. P. SHETTLE and J. A. WEINMAN, *J. Atmos. Sci.* **27**, 1048 (1970).
22. E. M. SPARROW and R. D. CESS, *Radiation Heat Transfer*. McGraw-Hill, New York (1978).
23. L. ELTERMAN, *Appl. Opt.* **3**, 1139 (1964).
24. J. T. PETERSON and J. A. WEINMAN, *J. Geophys. Res.* **74**, 6974 (1969).
25. P. K. DAS, *Tellus* **14**, 212 (1962).
26. W. M. ELSASSER and M. F CULBERTSON, *Meteor Monogr.* **4**, 23 (1960).
27. H. GRASSL, *Beitr. Phys. Atmos.* **47**, 1 (1974).

Paper 120

Reprinted from JOURNAL OF THE ATMOSPHERIC SCIENCES, Vol. 36, No. 6, June 1979
American Meteorological Society
Printed in U. S. A.

Determination of the Ground Albedo and the Index of Absorption of Atmospheric Particulates by Remote Sensing. Part II: Application[1]

MICHAEL D. KING[2]

Institute of Atmospheric Physics, The University of Arizona, Tucson 85721

(Manuscript received 29 August 1978, in final form 12 February 1979)

ABSTRACT

A hemispherical radiometer has been used to obtain spectrally narrow-band measurements of the downward hemispheric diffuse and total (global) flux densities at varying solar zenith angles on 14 days over Tucson. Data are presented which illustrate the effects of temporally varying atmospheric conditions as well as clear stable conditions on the ratio of the diffuse to direct solar radiation at the earth's surface. The ground albedo and the effective imaginary term of the complex refractive index of atmospheric particulates are derived from the diffuse-direct ratio measurements on seven clear stable days at two wavelengths using the statistical procedure described by King and Herman (1979). Results indicate that the downwelling diffuse radiation field in the mid-visible region in Tucson can be adequately described by Mie scattering theory if the ground albedo is 0.279 ± 0.100 and the index of absorption is 0.0306 ± 0.0082.

1. Introduction

A hemispherical radiometer designed by Huttenhow (1976) has been used in order to obtain spectrally narrow-band measurements of the downward hemispheric diffuse and total (global) flux densities at Tucson, Arizona. By combining the global (diffuse plus direct) flux density with the diffuse flux density, the diffuse-direct ratio can be computed. Data have been collected for 14 days between 5 May and 16 June 1977 at four wavelengths spaced throughout the visible region. The diffuse-direct ratio at each wavelength was measured during the course of each day, resulting in data for a range of solar zenith angles. Measurements have been collected for not only clear and stable atmospheric conditions but also for days during which the atmosphere was unstable and cloudy.

In Part I of this series (King and Herman, 1979) a statistical method was presented whereby measurements of the diffuse-direct ratio as a function of solar zenith angle can be analyzed to assess the magnitude of both the ground albedo and the index of absorption (imaginary part of the complex refractive index) of atmospheric particulates. In this paper a description is given of the hemispherical radiometer used to make the measurements. Data are then presented of the diffuse-direct ratio as a function of solar zenith angle and wavelength in order to illustrate some of the measurement sensitivities observed. Finally, data for days during which the atmosphere was clear and stable are analyzed and the optimum values of the ground albedo and the index of absorption of atmospheric particulates are presented.

2. Description of hemispherical radiometer

The basic requirement of the radiometer is to provide accurate narrow-band measurements of the hemispheric flux densities, both diffuse and global, received at the earth's surface. Fig. 1 is a schematic illustration of the hemispherical radiometer used in the present investigation. A plastic diffuser material has been selected for the first optical element since it is capable of producing a diffuse radiation field below the element independent of viewing direction. After passing through the diffuser element, the radiant energy within certain selected wavelength intervals is isolated by transmission through narrow-band interference filters. The distance between the diffuser element and interference filter is dictated by the maximum permissible angle of incidence of radiation on the filter ($\sim 7°$). After passing through the filter, the radiant energy is detected by a photodiode which produces an output current proportional to

$$F^{-}_{meas}(\tau_t) = \int_0^{2\pi} \int_0^{\pi/2} I(\tau_t, -\theta, \phi) f(\theta) \sin\theta d\theta d\phi, \quad (1)$$

where $I(\tau_t, -\theta, \phi)$ is the intensity of light propagating in the downward direction at the level τ_t, a function

[1] The research reported in this article was supported by the National Science Foundation under Grant ATM75-15551-A01 and the Office of Naval Research under Grant N00014-75-C-0370.

[2] Present affiliation: Laboratory for Atmospheric Sciences, Goddard Space Flight Center, NASA, Greenbelt, MD 20771.

0022-4928/79/061072-12$07.00

© 1979 American Meteorological Society

of zenith angle θ and azimuth angle ϕ; τ_t is the total optical depth of the atmosphere; and $f(\theta)$ is the optical response function of the sensor. In order for $F^{-}_{\text{meas}}(\tau_t)$ to agree well with the downward hemispheric diffuse flux density $F^{-}(\tau_t)$, it is necessary for $f(\theta)$ to equal $\cos\theta$.

Since there is a dielectric boundary between the atmosphere and the diffuser, some reflection occurs at the surface due to a discontinuity in the refractive index across the boundary. This dielectric discontinuity results in Fresnel reflection at the surface of the diffuser element which varies with polar angle θ. Any physical diffuser element will have an optical response function other than $\cos\theta$ as a direct result of reflection at the surface. It is therefore necessary to design a corrector plate to surround the diffuser element such that the combination of diffuser plus corrector will cause $f(\theta)$ to be nearly equal to $\cos\theta$. Since reflection can be thought of as reducing the effective area of the sensor element, it is necessary for the effective area of the sensor normal to the given beam of radiation to vary as a function of θ in such a way as to compensate for Fresnel reflection losses.

Huttenhow (1976) investigated several designs for the geometrical shape of the diffuser element and corrector plate. He concluded that measurement errors less than 2% could be attained by using a cylindrical diffuser element surrounded by a stepped corrector plate with only two steps, as illustrated in Fig. 1. It is possible to empirically determine the radial positions and heights where baffles can be placed in order to obtain $\cos\theta$ compensation for a range of polar angles. The more steps there are in the external corrector plate the more polar angles there are where $f(\theta)=\cos\theta$.

The housing for the hemispherical radiometer was constructed out of brass with the interior painted with a flat black paint to reduce the effects of reflected light from the sides of the housing. The radiometer contained a drawer port for interchanging the interference filters. The amplifier assembly was contained within the housing of the radiometer with the output fed into an external digital voltmeter to be recorded. The radiometer was mounted on a tripod for mobility and easy leveling. The tripod assembly contained a semi-circular metal slide ring on which was mounted a small brass ball whose location was adjustable along the ring. This ball and slide ring were used for occulting the direct sun while still allowing the maximum amount of diffusely scattered light from the solar aureole to be measured. Due to this requirement, an occulting ball was selected of sufficient diameter to block only the direct sun. With the ring lowered below the level of the table, global (diffuse plus direct) flux density measurements were obtained. Since a relative measurement between the diffuse and direct

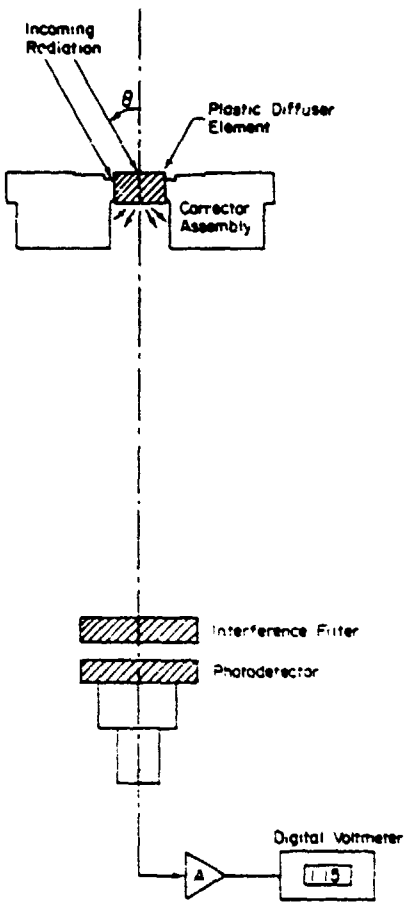


FIG. 1. Schematic illustration of the hemispherical radiometer.

flux densities was of interest, it was not necessary to calibrate the instrument in an absolute sense.

It is exceedingly critical that the top surface of the diffuser element be at the same level as the top level of the corrector assembly in order to assure that $f(\theta)$ approaches zero when $\theta=90°$. Any height difference in excess of about 0.025 mm can produce substantial errors in the response function for large polar angles.

3. Measurement sensitivity of the diffuse radiation field

In order to obtain estimates of the index of absorption of atmospheric particulates and the reflectivity of the earth's surface, measurements have been collected of the diffuse and global flux densities using the hemispherical radiometer described in the preceding section. These measurements have been made from the roof of the Civil Engineering Building (altitude 762 m MSL) on The University of Arizona campus (32°14′N, 110°57′W) because of the relatively clear and unobstructed view it affords of the entire upward hemisphere. Simultaneous to the collection of the flux density measurements, the directly trans-

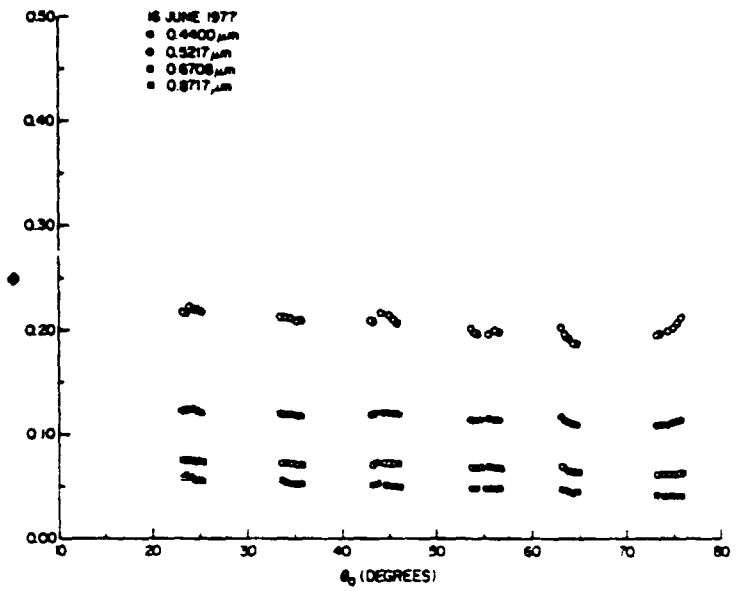


FIG. 2. Diffuse-direct ratio measurements for 16 June 1977 at four wavelengths.

mitted solar radiation was measured during the course of the day with a multi-wavelength solar radiometer described by Shaw *et al.* (1973). The diffuse-direct ratio Φ, as defined by Herman *et al.* (1975) and King and Herman (1979), is obtained from the diffuse $F^{-}_{meas}(\tau_i)$ and global $G_{meas}(\tau_i)$ flux density measurements once the solar zenith angle θ_0 is computed.

Under the assumption that $f(\theta_0)=\mu_0$, where μ_0 is the cosine of the solar zenith angle θ_0, it follows that the diffuse-direct ratio is given by

$$\Phi=\mu_0\frac{F^{-}_{meas}(\tau_i)}{G_{meas}(\tau_i)-F^{-}_{meas}(\tau_i)}. \tag{2}$$

Due to the deviation of the instrumental response function from $\cos\theta$, however, it can be shown that

$$\frac{\Phi-\hat{\Phi}}{\hat{\Phi}}=\frac{\mu_0}{f(\theta_0)}\frac{F^{-}_{meas}(\tau_i)}{F^{-}(\tau_i)}-1, \tag{3}$$

where $\hat{\Phi}$ and $F^{-}(\tau_i)$ are the measurements which would be obtained if $f(\theta)=\cos\theta$.

Measurements of the angular response function $f(\theta)$ have been obtained in the laboratory and used to simulate errors in the diffuse-direct ratio for a number of intensity distributions representative of six different atmospheric models and five solar zenith angles. Results of this investigation indicate that the Φ measurements are systematically higher than theory when $\theta_0 \lesssim 70°$ and systematically lower than theory when $\theta_0 \gtrsim 70°$ for all models. The maximum systematic errors ($\sim 2\%$) occur when $\theta_0 \approx 55°$, the angle where the largest error in the instrumental response function occurs. Over the range of solar zenith angles $35° \lesssim \theta_0 \lesssim 75°$, $\Delta\Phi/\hat{\Phi} \approx 0.7\%$ for a Mie (particulate) optical depth τ_M (0.5550 μm) = 0.05. The magnitude of this systematic error varies somewhat with the atmospheric model and decreases slightly as the Mie optical depth increases.

Fig. 2 illustrates a typical data set of Φ vs θ_0 which has been obtained for Tucson at four wavelengths (0.4400, 0.5217, 0.6708 and 0.8717 μm) on 16 June 1977. For all the filters reported here the full bandwidth at half-peak transmittance was less than or equal to 0.012 μm. From the directly transmitted solar radiation measurements the Mie optical depths $\tau_M(\lambda)$ were obtained by the method described by King and Byrne (1976). This day was cloud-free and stable with Mie optical depths ranging between 0.0500 ($\lambda=0.4400$ μm) and 0.0356 ($\lambda=0.8717$ μm). Because the wavelength dependence of $\tau_M(\lambda)$ was relatively small, a log-normal type of aerosol size distribution was obtained [see King *et al.* (1978) for details]. Examination of Fig. 2 suggests that Φ decreases slightly with θ_0 until a certain zenith angle is reached after which point Φ starts to increase. The increase of Φ at larger solar zenith angles is the most evident for the shortest wavelength, becoming less evident the longer the wavelength. The solar zenith angle dependence of the diffuse-direct ratio is in essential agreement with the theoretical computations presented by King and Herman (1979, Fig. 3), a case for which $\lambda=0.5550$ μm and $\tau_M=0.0500$.

On occasions when the occulting ball is incorrectly positioned, a diffuse flux density in excess of the proper amount is recorded. As a consequence of this error, the measured value of the diffuse-direct ratio is systematically higher than it should be [see Eq. (2)]. Although it is not always possible to identify systematic errors with absolute certainty, the more

obvious ones are evident when repeated measurements are made over a small range of solar zenith angles. Fig. 2 suggests that a few such errors did occur, mostly notably at $\lambda=0.4400$ μm and $\theta_0\approx45°$. In these situations the appropriate index of absorption κ and ground albedo A should be determined by analyzing only the more accurate data points. After elimination of the data for the times having occulting errors, the majority of the measurements yield characteristics similar to the theoretical computations illustrated in Part I (Fig. 3).

It sometimes happens that the total optical depth of the atmosphere varies with time during the course of a single day. As pointed out by King and Herman (1979), a fluctuation in τ_M is sufficient to alter the magnitude of the diffuse-direct ratio. A particularly dramatic illustration of the effect of temporal and spatial fluctuations in the atmosphere can be seen from the diffuse-direct ratio data for 16 May 1977 (see Fig. 3). On this day measurements were collected at four wavelengths during both the morning and the afternoon. As the solar zenith angle decreased throughout the morning from an initial value of 54°, the diffuse-direct ratio decreased. Had the atmosphere been horizontally homogeneous with a fixed Mie optical depth and aerosol size distribution throughout the day, the diffuse-direct ratio would have been a function only of λ and θ_0. Instead, the afternoon measurements did not repeat those of the morning but were noticeably lower in magnitude for a fixed solar zenith angle and wavelength. Although this day was cloud-free, it was visibly very turbid. Noticeable streaks were visible in the sky during the morning which appeared very similar to the stratospheric dust striations which occurred following the eruption of Volcán de Fuego in October 1974. By the afternoon all visible dust striations were gone but the late afternoon sky in the vicinity of the sun was very white suggesting a still quite appreciable dust content in the atmosphere. Data collected with a multi-wavelength solar radiometer on 16 May 1977 show very clearly that the total (and hence Mie) optical depth of the atmosphere was steadily decreasing throughout the day.

Garrison *et al.* (1978) report observations of the diffuse-direct ratio in the near ultraviolet for a similar situation in which the diffuse-direct ratio differed appreciably between morning and afternoon. They similarly attribute these fluctuations to variations in the optical depth of the atmosphere.

Since the Mie optical depth is a very important atmospheric parameter determining the magnitude of the diffuse-direct ratio, it is of interest to compare the magnitude of Φ for various days having different Mie optical depths. In order to compare measurements of the diffuse-direct ratio for several different days, estimates of Φ were obtained for $\lambda=0.5217$ μm and $\theta_0=55°$ by assuming that Φ can be approximated by a linear function of θ_0 for data collected at zenith angles near 55°. The results, including statistical error bar estimates for both Φ and τ_M, are presented in Fig. 4 for seven days between 6 May and 16 June 1977. As the Mie optical depth increases, the diffuse-direct ratio generally increases as expected from the computations illustrated by King and Herman (1979, Fig. 4) for $\lambda=0.5550$ μm. The magnitude of Φ does not increase monotonically as τ_M increases, however, simply due to daily differences in the aerosol size

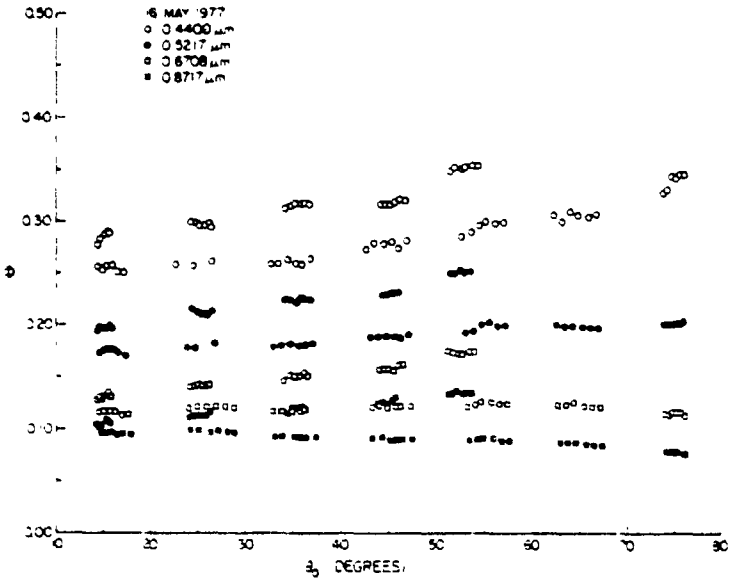


FIG. 3. Diffuse-direct ratio measurements for 16 May 1977 demonstrating the effects of temporal fluctuations in the atmosphere.

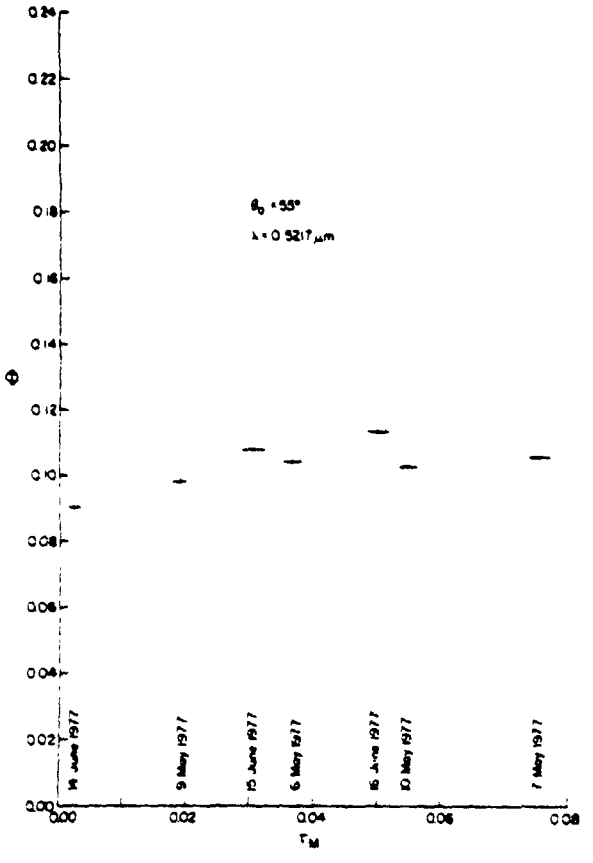


FIG. 4. The diffuse-direct ratio for seven days of varying Mie optical depths for $\lambda = 0.5217$ μm and for $\theta_0 = 55°$

distribution. The rate of increase of Φ as a function of τ_M is less than any of the cases illustrated by King and Herman (1979). Although the wavelengths are slightly different, it will be seen in the next section that the primary reason for the smaller change of Φ with increasing τ_M for the data presented here is due to an index of absorption κ larger than 0.01.

Illustrations similar to Fig. 4 have been constructed for $\theta_0 = 55°$ at each of the other three wavelengths (0.4400, 0.6708 and 0.8717 μm). Data for the two longer wavelengths are qualitatively in agreement with the results presented in Fig. 4 with Φ tending to increase with increasing τ_M. The results for $\lambda = 0.4400$ μm tend to be less predictable due to the increased statistical and systematic errors characteristic of this wavelength.

Fig. 5 illustrates Φ as a function of λ for four days between 9 May and 16 June 1977 at a fixed solar zenith angle of 55°. The values of the diffuse-direct ratio presented here were obtained in the same manner as those of Fig. 4, by making a linear least-squares fit to Φ as a function of θ_0 for data collected around $\theta_0 = 55°$. All days for which data have been collected, including the ones presented here, have diffuse-direct ratios which monotonically decrease with wavelength and exhibit slight positive curvature.

In assessing the agreement between the measured wavelength sensitivity and theoretical expectations, it becomes necessary to perform radiative transfer calculations for the specific aerosol size distribution and optical depth applicable to a particular day and wavelength. Since computations are required only of the transmitted hemispheric flux density at the earth's surface and not the intensity field in any particular direction, considerable computational time can be saved by expressing the elements of the intensity vector and phase matrix in Fourier series in $(\phi' - \phi)$, the difference between the azimuth angles of the meridian planes containing the directions of incidence and scattering. This procedure has been adopted by Dave (1970) and by Herman and Browning (1975) using Gauss-Seidel iteration. Though the Fourier analysis procedure can readily be used to calculate the complete intensity field, one of its main advantages when particulates are present in the atmosphere is in computing the hemispheric flux densities since only the constant, azimuth-independent terms are required (Herman and Browning, 1975). Theoretical computations of the diffuse-direct ratio have been performed for three wavelengths (0.5217, 0.6708 and 0.8717 μm)

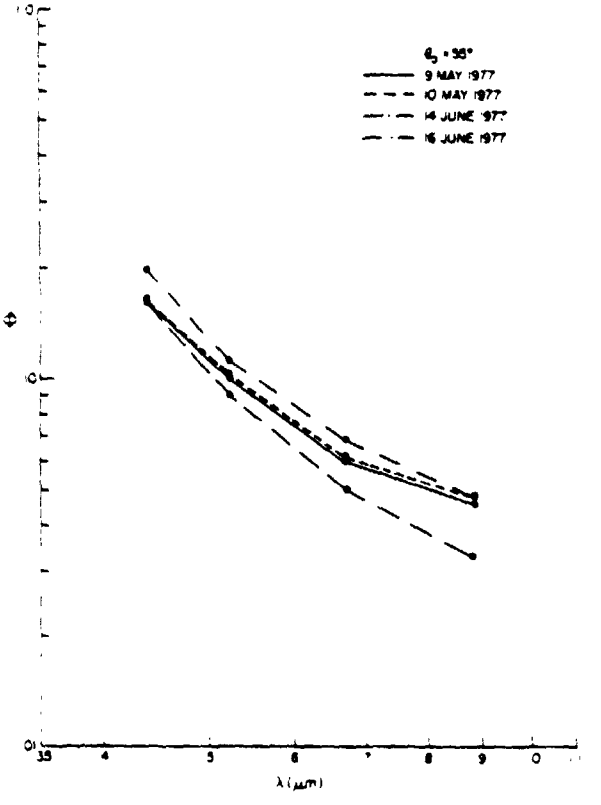


FIG. 5. The diffuse-direct ratio as a function of wavelength for $\theta_0 = 55°$ and for four different days.

on five days for which the Mie optical depths and aerosol size distributions are known. The theoretical results (assuming κ and A to be wavelength independent) are in close agreement with those of Fig. 5 in having Φ decrease monotonically in λ with slight positive curvature. Since both the theoretical and experimental sensitivities of Φ as a function of λ are in close agreement, having no anomalously high or low values at wavelengths in the ozone Chappuis absorption band (0.5217 and 0.6708 μm), it appears that the sensitivity of the diffuse-direct ratio to total ozone content is quite small as suggested by Herman *et al.* (1975) and King and Herman (1979).

4. Experimental results

The statistical method for inferring the ground albedo and the index of absorption of atmospheric particulates described by King and Herman (1979) has been applied to diffuse-direct ratio data collected on seven days during May and June 1977. After determining the spectral Mie optical depth values and columnar aerosol size distribution for a particular day, radiative transfer computations were performed for the two wavelengths 0.5217 and 0.6708 μm. The methods which have been adopted for estimating the Mie optical depth and aerosol size distribution are described by King and Byrne (1976) and King *et al.* (1978), while the method used to compute the hemispheric flux density and diffuse-direct ratio is described by Herman and Browning (1975). The real part of the complex refractive index of atmospheric particulates is assumed to be 1.45 both for the determination of the aerosol size distribution and for the application of the diffuse-direct radiation method. Fortunately, neither of these techniques is very sensitive to the real part of the particle refractive index (King *et al.*, 1978; King and Herman, 1979).

The values of the Mie optical depth, index of absorption, ground albedo and mean diffuse-direct ratio for 10 data cases are presented in Table 1. In view of the amount of scatter in the values of the ground albedo A, it does not seem justified to separate the data taken at two different wavelengths (0.5217 and 0.6708 μm) and seek a wavelength dependence. The standard deviations in the individual values of τ_M, κ and A shown in Table 1 have been derived by standard error propagation methods for each data set and thus reflect the combined effect which each individual data point has on the determination of the regression coefficients τ_M, κ and A [see King and Byrne (1976) and King and Herman (1979) for details]. The magnitude of the ground albedo presented here shows a relatively large day-to-day variability which is due, in part, to uncertainties in the aerosol size distribution and Mie optical depth. The analysis procedure which we have used has not attempted to incorporate information on the uncertainties of either the Mie optical depth or aerosol size distribution. The mean values of the imaginary index of refraction and ground albedo, weighted by the reciprocal of the variances, are found to be 0.0306 and 0.279, respectively, with corresponding standard deviations of 0.0082 and 0.100 (see Table 1). Considering the magnitude of the ground albedos and Mie optical depths obtained for these days, this amount of absorption on the part of the atmospheric particulates implies an absorption by the atmosphere on the order of 3% or less of the radiation incident at small solar zenith angles, increasing to 5% at zenith angles of 65° on 10 May 1977.

By comparing the data of Fig. 4 with the theoretical computations presented in Part I of this series (Fig. 4, applicable to λ=0.5550 μm) it is apparent that the tendency for very little increase of the diffuse-direct ratio measurements with Mie optical depth is

TABLE 1. Summary of the Mie optical depth, index of absorption, ground albedo and mean diffuse-direct ratio obtained for seven days during 1977.

Date	Wavelength (μm)	Mie optical depth	Index of absorption	Ground albedo	Mean diffuse-direct ratio
6 May	0.5217	0.0367±0.0014	0.0330±0.0048	0.325±0.004	0.1078
6 May	0.6708	0.0341±0.0014	0.0074±0.0051	0.469±0.064	0.0645
7 May	0.6708	0.0629±0.0010	0.0336±0.0033	0.289±0.013	0.0671
9 May	0.5217	0.0189±0.0012	0.0334±0.0090	0.394±0.007	0.0990
10 May	0.5217	0.0548±0.0014	0.0333±0.0031	0.123±0.006	0.1049
10 May	0.6708	0.0589±0.0012	0.0340±0.0074	0.172±0.011	0.0643
14 June	0.5217	0.0025±0.0010	0.0334±0.0091	0.370±0.012	0.0859
15 June	0.5217	0.0304±0.0019	0.0354±0.0048	0.236±0.015	0.1005
16 June	0.5217	0.0501±0.0016	0.0367±0.0037	0.199±0.011	0.1164
16 June	0.6708	0.0445±0.0013	0.0233±0.0036	0.348±0.025	0.0695
Weighted mean			0.0306	0.279	
Standard deviation			0.0082	0.100	

Note: The observed Mie optical depths and estimated aerosol size distributions for 6 May and 15 June 1977 are illustrated in King *et al.* (1978).

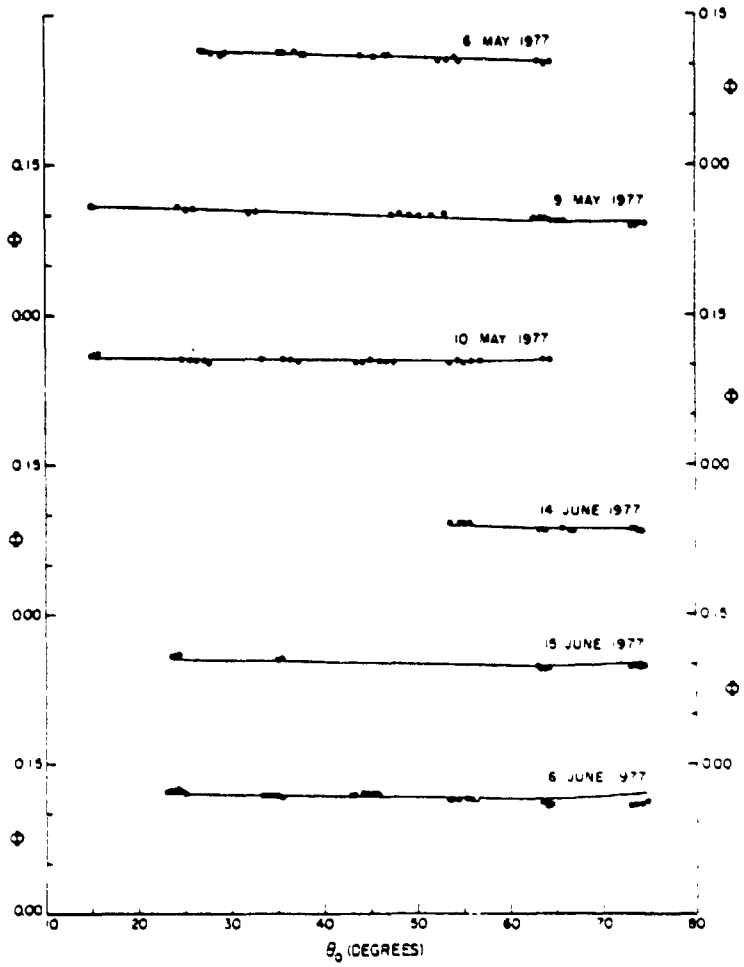


FIG. 6. Observed and computed diffuse-direct ratio versus solar zenith angle for six days at $\lambda = 0.5217$ μm. The ordinate scale to which each curve refers is repeated alternately on the left and then the right side of the figure in order to separate the individual data cases.

consistent with a relatively large value of 0.0306 for the index of absorption. In addition, the theoretical computations for 6 May and 15 June require the diffuse-direct ratio to be less on 6 May for fixed values of κ and A simply due to differences in the aerosol size distributions. The reduced magnitude of 10 May over 16 June is similarly due to a difference in aerosol size distributions and is to be expected theoretically (see Fig. 4).

On any given day and wavelength it is not always possible to determine values for κ and A due to uncertainties in either the diffuse-direct ratio, the Mie optical depth or the aerosol size distribution. Temporal and spatial fluctuations in the atmosphere can also make it not feasible to analyze a particular day or wavelength, as is apparent upon examination of Fig. 3, since it is necessary to measure Φ over a range of solar zenith angles while the atmosphere remains constant. Although 7 May appeared very clear at the time of the observations, subsequent examination of the solar radiometer data indicated an appreciable time variation of optical depth in the late morning and afternoon. The hemispherical radiometer data were fortunately collected in the early morning on this day but some time variations in optical depth were nevertheless detectable at 0.5217 μm (but not so severely at 0.6708 μm). The analysis for 7 May 1977 at $\lambda = 0.5217$ μm has not been included in Table 1 for this reason. It is desirable, particularly at the longer wavelengths, to have data for days exhibiting a relatively large Mie optical depth since the larger the optical depth the greater the sensitivity to A and κ (King and Herman, 1979). The days presented in Table 1 at $\lambda = 0.6708$ μm were the four days which had the largest Mie optical depths at this wavelength.

Due to the normally small Mie and Rayleigh optical depths at wavelengths in the near-infrared, sensitivities to κ and A are so greatly reduced that the diffuse-direct technique is normally not a viable remote sensing method for deriving these parameters. A pre-

liminary analysis of some of the data sets at 0.8717 μm indicates that the measurements may have been systematically higher than theory at this wavelength (but the general lack of sensitivity to κ and A for the Mie optical depths measured at 0.8717 μm make it difficult to say whether these errors are significant). The near-ultraviolet wavelength region, on the other hand, does seem attractive for applying the diffuse-direct technique since the Mie optical depths are normally larger at these wavelengths. No attempt has been made in the present investigation to analyze the measurements at 0.4400 μm since the relatively small measurement signal contributed to larger statistical and systematic errors (see Figs. 2 and 3) and because the Mie optical depths were particularly small during this period at 0.4400 μm. Since there is negligible ozone absorption at $\lambda=0.4400$ μm, it would be a good wavelength to investigate further in the future if measurements can be made with a sufficient degree of accuracy.

Figs. 6 and 7 present the diffuse-direct ratio data and results of the fitting procedure for all 10 cases of Table 1. The solid curves represent the regression fit to the data points using the optimum values of the coefficients κ and A. Most hemispherical radiometer measurements were made only in the afternoon but the data for 7 May (Fig. 7) were collected in the morning while the data for 6 and 9 May include both morning and afternoon observations. The atmosphere on 14 June was incredibly clean with Mie optical depths of about 0.01 at most wavelengths. Although it was clear throughout the day, large systematic occulting errors occurred at small solar zenith angles thus limiting the range of useful solar zenith angles on 14 June to $\theta_0 \gtrsim 53.56°$. By comparing Figs. 6 and 7 it is clear that theory requires Φ to increase slightly toward large solar zenith angles at $\lambda=0.5217$ μm, while not at all at $\lambda=0.6708$ μm. Had the optical depths been much larger on these days, as on the morning of 16 May (see Fig. 3), the diffuse-direct ratio would have increased monotonically in zenith angle for all wavelengths. The sample standard deviation of the data points about the regression fit is typically 0.0028 at $\lambda=0.5217$ μm and 0.0016 at $\lambda=0.6708$ μm, representing random fluctuations on the order of 2.5% at both wavelengths. These errors are larger than those attributed solely to measurement error (~0.7% as discussed in Section 3) and thus reflect the combined effect of measurement errors, atmospheric fluctuations and the accuracy of the determination of the Mie optical depth and aerosol size distribution.

The best months in Tucson for obtaining clear skies and stable atmospheric conditions for a long enough period of time to obtain good optical depth data are May, June, October and November. Due to the small solar declination angles in the fall, however, a restricted range of solar zenith angles is available ($\theta_0 \gtrsim 40°$) thus making May and June the best time to collect diffuse-direct ratio data in Tucson. In addition to the large range of solar zenith angles which are available in the spring, there is normally a secondary peak value of Mie optical depth at this time of year (with the absolute peak usually occurring in July or August). Unfortunately, the optical depths during May and June of 1977 were smaller than seasonally expected from previous years and thus the sensitivity of the diffuse-direct ratio measurements to A and κ was not as large as anticipated.

Although the measurements reported in this investigation were obtained near the center of the city of Tucson, the mean value of the ground albedo is in close agreement with the results obtained elsewhere in the southwest. Since the surface albedo determined by the diffuse-direct radiation method represents an area averaged albedo consistent with the transfer of radiation in the earth's atmosphere, the most representative values to be compared with it are other area-averaged radiation measurements such as those

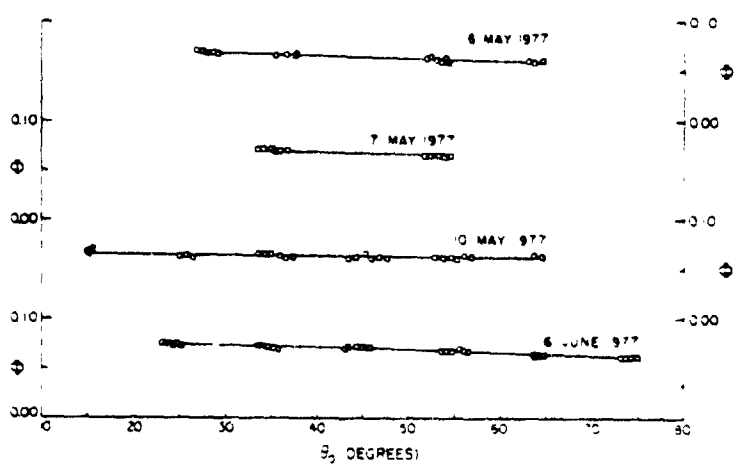


FIG. 7. Observed and computed diffuse-direct ratio versus solar zenith angle for four days at $\lambda=0.6708$ μm. The ordinate designation is the same as Fig. 6

obtained by low flying aircraft with upward and downward pointing radiometer systems. Kung *et al.* (1964) and Griggs (1968) obtained values ranging between 0.22 and 0.25 for the surface albedo in southern Arizona using measurement systems having a very wide wavelength response. DeLuisi *et al.* (1976) found values ranging between 0.20 and 0.35 for Quartzite, Arizona, where crude wavelength resolution was obtained by attaching Schott glass cutoff filters to their aircraft mounted Eppley pyranometers.

An alternative type of surface albedo determination in the southwestern United States has been obtained by Otterman and Fraser (1976) by examining the earth-atmosphere system reflectivities obtained from the Landsat satellite. By modeling the scattering properties of the atmospheric particulates they inferred values for the surface albedo of 0.264 in the wavelength range of 0.5 to 0.6 μm with little sensitivity to the particulate model. The mean value of 0.279 found in this investigation is in close agreement with that of these investigations. The variability of the values of ground albedo presented in Table 1 is, however, quite large. In order to better establish the ground albedo and its daily and spectral variability, it is necessary to collect more observations on days for which the Mie optical depths are large since the sensitivity of the diffuse-direct ratio to ground albedo increases in direct proportion to τ_M. As pointed out earlier, the Mie optical depths during May and June 1977 were smaller than normally expected for that time of year.

Since there are very few assessments of the imaginary index of refraction of naturally suspended atmospheric aerosol particles which have been obtained by examining the scattered radiation field, the number of studies which can be compared on an equivalent basis with the present one are few. DeLuisi *et al.* (1976) examined the percent absorption by the atmosphere between two different heights using as inputs previously determined values of the Mie optical depth and surface reflectivity. They reported values of κ on the order of 0.013 for Quartzite, Arizona, but with error bar estimates extending to 0.028. Kuriyan *et al.* (1979) examined the intensity and degree of polarization of the sunlit sky as a function of azimuth angle in the Los Angeles area. They found that during days having dry continental air masses the radiation data could best be ascribed to a low Mie optical depth ($\tau_M \approx 0.05$) and large index of absorption ($0.02 \lesssim \kappa \lesssim 0.05$). They felt that the relatively large values for the index of absorption (roughly consistent with those of Table 1) were due to the low humidity of the continental air mass.

Eiden (1966) compared theoretical and measured ellipticities of light scattered by a volume of atmospheric air and obtained an index of absorption κ ranging between 0.01 and 0.1 at Mainz. He also pointed out that an increased humidity decreases the effective value of κ supporting the conclusions by Kuriyan *et al.* (1979).

Since Eiden (1966) considered only three different Junge models for the aerosol size distribution, a better estimate of the size distribution should in theory lead to a more accurate determination of the imaginary index of refraction of the aerosol particles. Grams *et al.* (1974) measured the angular variation of the intensity of light scattered from a collimated beam by the atmospheric aerosol while simultaneously collecting the particles for size distribution analysis. With an accurate determination of the aerosol size distribution thus obtained, theoretical angular scattering intensity computations were performed and compared to simultaneous measurements of the same in order to determine the index of absorption of the atmospheric particulates. Results of this procedure applied to 23 data sets obtained at Big Spring, Texas, indicate that κ ranged between 0.0011 and 0.0214 with a geometric mean value of 0.0050.

Although Bergstrom (1973) and Eiden (1971) have pointed out the problems inherent in using bulk indices of refraction of collected samples of atmospheric dust in computing the radiation field of the free atmosphere, it is nevertheless of interest to compare the values obtained in the present investigation with laboratory measurements of the bulk index of absorption. Fischer (1970) has collected samples of aerosol particles with a jet impactor and, by means of an integrating sphere to collect the scattered light, inferred the index of absorption of the particulates. Assuming that the density of the aerosol particles is approximately 2.0 g cm^{-3}, Fischer's (1970) values for the imaginary index of refraction lie in the range $0.010 \lesssim \kappa \lesssim 0.045$ in the mid-visible wavelength region. More recently Fischer (1973) examined specific particulate samples from both urban and rural locations from which he found that the index of absorption is typically 0.010 in rural locations while being more typically 0.028 in urban locations.

Lin *et al.* (1973) used opal glass instead of a sphere to integrate the light scattered by collected samples of New York City particulates. They determined that the index of absorption ranged between 0.028 and 0.050 with a mean value of 0.040. Using the Kubelka-Munk theory of diffuse reflectance Lindberg and Laude (1974) inferred values for the index of absorption of samples of New Mexico dust lying between 0.007 and 0.008.

One of the limitations of the present procedure for inferring the ground albedo and the index of absorption of atmospheric particulates is the sensitivity of the diffuse-direct ratio to particles $\lesssim 0.10$ μm in radius. Since these particles are very difficult to sense from spectral Mie optical depth measurements (King *et al.*, 1978), the effects of small particles on the values of κ and A inferred by the diffuse-direct technique were considered. This was accomplished by comparing

results obtained assuming a Junge (1955) size distribution for the aerosol particles with one slope ($\nu^*=3.0$) but with two different radii ranges (0.01–5.0 μm and 0.10–5.0 μm). Results indicate that the effect of small particles is insignificant in the determination of the surface reflectivity though it may affect significantly (factor of $\sim$2.5) the value of the index of absorption inferred. The fact that the presence of small particles can have a large effect on the value of κ inferred by the diffuse-direct technique may readily be understood. Since absorbing particulates $\lesssim$0.10 μm in radius are more efficient absorbers than scatterers (i.e., $Q_{sca} \propto r^4$ and $Q_{abs} \propto r$, where Q_{sca} and Q_{abs} are the Mie efficiency factors for scattering and absorption, respectively), the neglect of these particles in computing the diffuse radiation field in the earth's atmosphere affects the theoretical values of the diffuse-direct ratio. For a fixed Mie optical depth of the atmosphere, the presence of a significant number of small absorbing particles decreases the single-scattering albedo and consequently the computed diffuse-direct ratio over those values computed if only particles larger than 0.10 μm are present. Since these computations are to be compared with measurements of Φ at a particular value of τ_M, the inclusion of a significant number of particles less than 0.10 μm in radius tends to *decrease* the inferred value of κ over the value determined if only particles larger than 0.10 μm are included. As a consequence of this effect, one should consider the values reported in the present investigation as being an upper limit of the true index of absorption. Most of the aerosol size distributions obtained for the seven days analyzed in Table 1 were log-normal in character, however, and thus showed little tendency to have very many particles smaller than 0.10 μm in radius. These types of distributions are typical of days for which the Mie optical depths are small (King *et al.*, 1978). The observed Mie optical depths and aerosol size distributions for 6 May and 15 June 1977 are illustrated by King *et al.* (1978, Figs. 4 and 5).

The effect of particulates with radii $\gtrsim$4.0 μm is of little importance since the reduced number of particles at these sizes is such that their contribution to the Mie optical depth, and hence to the diffuse radiation field, is on the order of a few tenths of one percent.

Grams *et al.* (1974) investigated the effect of including particles with radii $<$0.6 μm in computing the theoretical scattering intensities, since their primary refractive index analysis only included log-normal size distributions with particle radii $>$0.6 μm. They concluded that a Junge distribution extension to smaller particles with $\nu^*=3.0$ (i.e., $dN/d\log r \propto r^{-1}$) tended to *increase* their inferred geometric mean value of the index of absorption from 0.005 to 0.008. Although no mention was made of the lower radius limit of the Junge distribution extension, it is interesting to note that the direction of the effect of small particles is just the opposite of that of the diffuse-direct technique.

5. Conclusions

For a successful application of the diffuse-direct technique described in the present investigation, it is necessary to have a clear atmosphere devoid of any cloud cover for a long enough period of time for the solar zenith angle to undergo a large change, preferably of about 40°. It is further required that accurate normal incidence and downward hemispheric flux densities be obtained over narrow bandpass wavelength intervals. A hemispherical radiometer capable of quasi-monochromatic flux density measurements has been designed by Huttenhow (1976) and used in the present investigation. From measurements of the hemispheric diffuse and total (global) flux densities, the diffuse-direct ratio is obtained after computing the solar zenith angle at the time of the measurements.

Due to fluctuations in the magnitude of the diffuse-direct ratio associated with daily differences in Mie optical depth and aerosol size distribution (see Fig. 4), a procedure for interpreting measurements of the diffuse-direct ratio has been adopted which uses as input parameters values of the Mie optical depth and aerosol size distribution determined from alternative measurements of the directly transmitted solar radiation. In addition, the complete analysis requires an accurate radiative transfer program and a statistical optimization procedure for combining these factors in order to obtain the particulate index of absorption and ground albedo.

Although the combination of atmospheric conditions and instrumentation are difficult to satisfy in most locations, they can normally be met in Tucson during the months of April, May and June. Measurements have been collected on 14 days during May and June 1977 using the hemispherical radiometer described in Section 2. The clear and stable days have been analyzed and presented in Table 1 for two wavelengths (0.5217 and 0.6708 μm), with the diffuse-direct ratio data and corresponding regression fits being illustrated in Figs. 6 and 7. The mean value of the ground albedo presented here is in agreement with other measurements obtained both from satellite (Otterman and Fraser, 1976) and low flying aircraft (Kung *et al.*, 1964; Griggs, 1968; DeLuisi *et al.*, 1976). The relatively large variability among the individual data cases presented in Table 1 is primarily a consequence of the small Mie optical depths which occurred during May and June 1977. In order to better establish the ground albedo and its daily and spectral variability it is necessary to collect more observations on days for which the Mie optical depths are large since the sensitivity of the diffuse-direct ratio to ground albedo increases in direct proportion to τ_M (King and Herman, 1979).

The values for the imaginary part of the complex refractive index of atmospheric particulates are somewhat higher than initially anticipated but not inconsistent with the results of other investigators (Eiden, 1966; Fischer, 1970, 1973; Lin *et al.*, 1973; DeLuisi *et al.*, 1976; Kuriyan *et al.*, 1979). As a word of caution, however, it is important to note that spectral Mie optical depth measurements in the visible and near-infrared wavelength regions are relatively insensitive to the aerosol size distribution at particle radii much less than about 0.1 μm (King *et al.*, 1978). This can lead to an overestimation of the imaginary index of refraction of atmospheric particulates using the diffuse-direct technique if there are an optically significant number of small (Aitken) particles present in the atmosphere. Fortunately, the aerosol size distributions obtained for most of the days analyzed in this investigation were log-normal in character with the radius of maximum concentration lying between 0.3 and 0.9 μm. These data show little tendency to have very many particles smaller than 0.1 μm in radius. Since the sensitivities of the diffuse-direct ratio to ground albedo and index of absorption increase with increasing Mie optical depth, situations for which the concentration of small particles increases, it is necessary to obtain a good estimate for the aerosol size distribution at small radii when applying the methods described in this series of articles. On high optical depth situations in which the aerosol size distribution is nearly Junge with a large value of ν^*, it is reasonably simple to extend the minimum particulate radius for inversion purposes to 0.06 μm (King *et al.*, 1978).

Because the parameters κ and A determined by the diffuse-direct technique are those giving the best agreement between measurement and radiative transfer theory, they are of the most interest to meteorologists concerned with the impact of aerosol particles on climate. As pointed out by Yamamoto and Tanaka (1972), Wang and Domoto (1974) and Herman and Browning (1975), it is both the index of absorption and the surface reflectivity which determine whether an increase in the Mie optical depth will lead to a warming or a cooling of the earth's atmosphere. Assuming that the mean values of the surface albedo and index of absorption obtained at the 0.5217 and 0.6708 μm wavelengths apply to 0.5550 μm (the wavelength for which the most extensive set of radiative transfer computations have been performed), it is readily found that the effect of particulates over Tucson is one of heating the earth's atmosphere. This is primarily a result of the relatively large surface reflectivity of the southwestern desert region since a value for the index of absorption on the order of 0.010 is sufficient to produce a net warming over a wide range of aerosol size distributions.

Acknowledgments. This study represents part of the author's Ph.D. dissertation at the University of Arizona, prepared under the direction of Professor Benjamin Herman, to whom he would like to express his deepest gratitude for the guidance and encouragement offered during his entire graduate career. The author would further like to thank Professor S. Twomey for several helpful comments during various stages of this research and Capt. Jay D. Huttenhow for designing the hemispherical radiometer used in this investigation.

REFERENCES

Bergstrom, R. W., 1973: Comments on the estimation of aerosol absorption coefficients in the atmosphere. *Beitr. Phys. Atmos.*, 46, 198–202.

Dave, J. V., 1970: Intensity and polarization of the radiation emerging from a plane-parallel atmosphere containing monodispersed aerosols. *Appl. Opt.*, 9, 2673–2684.

DeLuisi, J. J., P. M. Furukawa, D. A. Gillette, B. G. Schuster, R. J. Charlson, W. M. Porch, R. W. Fegley, B. M. Herman, R. A. Rabinoff, J. T. Twitty and J. A. Weinman, 1976: Results of a comprehensive atmospheric aerosol-radiation experiment in the southwestern United States. Part II: Radiation flux measurements and theoretical interpretation. *J. Appl. Meteor.*, 15, 455–463.

Eiden, R., 1966: The elliptical polarization of light scattered by a volume of atmospheric air. *Appl. Opt.*, 5, 569–575.

——, 1971: Determination of the complex index of refraction of spherical aerosol particles. *Appl. Opt.*, 10, 749–754.

Fischer, K., 1970: Bestimmung der absorption von sichtbarer strahlung durch aerosolpartikeln. *Beitr. Phys. Atmos.*, 43, 244–254.

——, 1973: Mass absorption coefficient of natural aerosol particles in the 0.4–2.4 μm wavelength interval. *Beitr. Phys. Atmos.*, 46, 89–100.

Garrison, L. M., L. E. Murray, D. D. Doda and A. E. S. Green, 1978: Diffuse-direct ultraviolet ratios with a compact double monochromator. *Appl. Opt.*, 17, 827–836.

Grams, G. W., I. H. Blifford, Jr., D. A. Gillette and P. B. Russell, 1974: Complex index of refraction of airborne soil particles. *J. Appl. Meteor.*, 13, 459–471.

Griggs, M., 1968: Aircraft measurements of albedo and absorption of stratus clouds, and surface albedos. *J. Appl. Meteor.*, 7, 1012–1017.

Herman, B. M., and S. R. Browning, 1975: The effect of aerosols on the earth-atmosphere albedo. *J. Atmos. Sci.*, 32, 1430–1445.

——, —— and J. J. DeLuisi, 1975: Determination of the effective imaginary term of the complex refractive index of atmospheric dust by remote sensing: the diffuse-direct radiation method. *J. Atmos. Sci.*, 32, 918–925.

Huttenhow, J. D., 1976: Design and analysis of a spectrally narrow-band radiometer. M.S. thesis, University of Arizona. 68 pp.

Junge, C. E., 1955: The size distribution and aging of natural aerosols as determined from electrical and optical data on the atmosphere. *J. Meteor.*, 12, 13–25.

King, M. D., and D. M. Byrne, 1976: A method for inferring total ozone content from the spectral variation of total optical depth obtained with a solar radiometer. *J. Atmos. Sci.*, 33, 2242–2251.

——, and B. M. Herman, 1979: Determination of the ground albedo and the index of absorption of atmospheric particulates by remote sensing. Part I: Theory. *J. Atmos. Sci.*, 36, 163–173.

——, D. M. Byrne, B. M. Herman and J. A. Reagan, 1978: Aerosol size distributions obtained by inversion of spectral optical depth measurements. *J. Atmos. Sci.*, 35, 2153–2167.

Kung, E. C., R. A. Bryson and D. H. Lenschow, 1964: Study of a continental surface albedo on the basis of flight measurements and structure of the earth's surface cover over North America. *Mon. Wea. Rev.*, 92, 543–564.

Kuriyan, J. G., S. K. Mitra and D. M. Landau, 1979: Effective optical characteristics of atmospheric aerosol particles in the L. A. basin. Submitted to *Tellus*.

Lin, C. L, M. Baker and R. J. Charlson, 1973: Absorption coefficient of atmospheric aerosol: A method for measurement. *Appl. Opt.*, 12, 1356–1363.

Lindberg, J. D., and L. S. Laude, 1974: Measurement of the absorption coefficient of atmospheric dust. *Appl. Opt.*, 13, 1923–1927.

Otterman, J., and R. S. Fra[illegible]76: Earth-atmosphere system and surface reflectivities in arid regions from LANDSAT MSS data. *Remote Sens. Environ.*, 5, 247–266.

Shaw, G. E., J. A. Reagan and B. M. Herman, 1973: Investigations of atmospheric extinction using direct solar radiation measurements made with a multiple wavelength radiometer. *J. Appl. Meteor.*, 12, 374–380.

Wang, W.-C., and G. A. Domoto, 1974: The radiative effect of aerosols in the earth's atmosphere. *J. Appl. Meteor.*, 13, 521–534.

Yamamoto, G., and M. Tanaka, 1972: Increase of global albedo due to air pollution. *J. Atmos. Sci.*, 29, 1405–1412.

Paper 121

Reprinted from *Journal of the Atmospheric Sciences*, 35, pp. 2153-2167, 1973.

Aerosol Size Distributions Obtained by Inversion of Spectral Optical Depth Measurements

MICHAEL D. KING,[1,4] DALE M. BYRNE,[2,5] BENJAMIN M. HERMAN[1] AND JOHN A. REAGAN[3]

The University of Arizona, Tucson 85721

(Manuscript received 17 February 1978, in final form 15 August 1978)

ABSTRACT

Columnar aerosol size distributions have been inferred by numerically inverting particulate optical depth measurements as a function of wavelength. An inversion formula which explicitly includes the magnitude of the measurement variances is derived and applied to optical depth measurements obtained in Tucson with a solar radiometer. It was found that the individual size distributions of the aerosol particles (assumed spherical), at least for radii $\gtrsim 0.1$ μm, fall into one of three distinctly different categories. Approximately 50% of all distributions examined thus far can best be represented as a composite of a Junge distribution plus a distribution of relatively monodispersed larger particles centered at a radius of about 0.5 μm. Scarcely 20% of the distributions yielded Junge size distributions, while 30% yielded relatively monodispersed distributions of the log-normal or gamma distribution types. A representative selection of each of these types will be presented and discussed. The sensitivity of spectral attenuation measurements to the radii limits and refractive index assumed in the numerical inversion will also be addressed.

1. Introduction

A relationship between the size of atmospheric aerosol particles and the wavelength dependence of the extinction coefficient was first suggested by Ångström (1929). Since that time Ångström's (1929) empirical formula for the wavelength dependence of the extinction coefficient has been directly related to a parameter of a Junge size distribution when the radii extend from 0 to ∞ (van de Hulst, 1957; Junge, 1963). Curcio (1961) used the wavelength dependence of the particulate extinction coefficient in the visible and near-infrared regions to infer the aerosol size distributions existing above water in the Chesapeake Bay area. He determined that the majority of aerosol size distributions could best be represented by a two-component size distribution consisting of a Junge-type distribution plus a small component of larger particles. This type of composite distribution was the most capable of explaining the wavelength dependence of the attenuation measurements he observed.

Yamamoto and Tanaka (1969) were the first to apply a numerical inversion algorithm to spectral measurements of extinction coefficient in order to determine an aerosol size distribution. These authors applied the linear inversion techniques developed by Phillips (1962) and Twomey (1963) to the problem of numerically solving the Fredholm integral equation of the first kind which arises in this problem. Although they clearly demonstrated that these numerical procedures are quite successful for obtaining size distributions by remote sensing, other investigators more recently have still continued to estimate parameters of model size distributions from spectral attenuation measurements (e.g., Quenzel, 1970; Shaw *et al.*, 1973). Although these fitting procedures are reasonably satisfactory, they are more restrictive than inversion procedures in that they assume that the atmospheric particulates follow one of several possible analytical distributions.

Grassl (1971) presents an iterative method for numerically inverting spectral attenuation data. After demonstrating the success of this algorithm on spectral attenuation coefficients generated for three model size distributions, the size distributions obtained by inversion of two real data cases are presented. In order to accurately determine aerosol size distributions from spectral optical depth measurements obtained from direct solar observations, it is necessary to collect optical depth measurements over a sufficient number of wavelengths to obtain a good estimate of both the ozone absorption and particulate optical depths

[1] Institute of Atmospheric Physics.
[2] Optical Sciences Center.
[3] Department of Electrical Engineering.
[4] Present affiliation: Goddard Space Flight Center, Greenbelt, MD 20771.
[5] Present affiliation: Pratt & Whitney Aircraft, Government Products Division, Box 2691, West Palm Beach, FL 33402.

0022-4928/78/2153-2167$07.50
© 1978 American Meteorological Society

separately. In making these corrections, Grassl (1971) used tabulated values for a model atmosphere in lieu of alternative observations.

In the present investigation an earlier theoretical study (Herman *et al.*, 1971) of angular scattering intensities has been extended to the problem of inferring columnar aerosol size distributions by inversion of spectral optical depth measurements. An inversion formula which explicitly includes the magnitudes of the measurement variances is derived and applied to optical depth measurements obtained in Tucson with a solar radiometer. Aerosol size distribution results have been obtained for 57 days and the results of a representative selection are presented, together with a discussion of the relative frequency of occurrence of various types of distributions. Some of the practical difficulties to be considered when inverting spectral attenuation measurements will be discussed. The most important of these are the radii limits of maximum sensitivity and the particulate refractive index assumed in the inversion.

2. Method of solution

From wavelength measurements of the directly transmitted solar flux density as a function of solar zenith angle, one can obtain spectral values of the particulate (Mie) optical depth (Shaw *et al.*, 1973; King and Byrne, 1976). The spectral variation of Mie optical depth, designated $\tau_M(\lambda)$, is produced through attenuation by aerosol and is primarily determined by the aerosol size distribution. Assuming that the atmospheric particulates can be modeled by equivalent spheres of known refractive index to a sufficient degree of accuracy, the integral equation which relates optical depth to an aerosol size distribution can be written as

$$\tau_M(\lambda)=\int_0^\infty\int_0^\infty \pi r^2 Q_{ext}(r,\lambda,m)n(r,z)dzdr, \qquad (1)$$

where $n(r,z)dr$ is the height-dependent aerosol number density in the radius range r to $r+dr$; m the complex refractive index of the aerosol particles; λ the wavelength of the incident illumination; and $Q_{ext}(r,\lambda,m)$ the extinction efficiency factor from Mie theory. The effects of possible variations of the particulate refractive index with wavelength and particle radius will be discussed in Section 4. In the present formulation, the aerosol is assumed to consist of homogeneous spherical particles which are nondispersive over the wavelength range of the observations. Performing the height integration, Eq. (1) can be rewritten as

$$\tau_M(\lambda)=\int_0^\infty \pi r^2 Q_{ext}(r,\lambda,m)n_c(r)dr, \qquad (2)$$

where $n_c(r)$ is the unknown columnar aerosol size distribution, i.e., the number of particles per unit area per unit radius interval in a vertical column through the atmosphere.

To determine $n_c(r)$, the transform of (2) must be obtained. Since an expression for $n_c(r)$ cannot be written analytically as a function of the $\tau_M(\lambda)$ values, a numerical approach must be followed. Therefore, the integral in (2) is replaced by a summation over coarse intervals in r, each of which is composed of several subintervals as described by Herman *et al.* (1971) for the case of the angular distribution of scattered light of one wavelength. In order to examine the specific kernel functions which result if that procedure is applied to the present problem, we let $n_c(r)=h(r)f(r)$, where $h(r)$ is a rapidly varying function of r, while $f(r)$ is more slowly varying. With this substitution, Eq. (2) becomes

$$\tau_M(\lambda)=\int_{r_0}^{r_q} \pi r^2 Q_{ext}(r,\lambda,m)h(r)f(r)dr$$

$$=\sum_{j=1}^{q}\int_{r_j}^{r_{j+1}} \pi r^2 Q_{ext}(r,\lambda,m)h(r)f(r)dr, \qquad (3)$$

where the limits of integration have been made finite with $r_1=r_c$ and $r_{q+1}=r_b$. If $f(r)$ is assumed constant within each coarse interval, a system of linear equations results which may be written as

$$\mathbf{g}=\mathbf{A}\mathbf{f}+\boldsymbol{\epsilon}, \qquad (4)$$

where $\boldsymbol{\epsilon}$ is an unknown error vector whose elements ϵ_i represent the deviation between measurement (g_i) and theory ($\sum_j A_{ij}f_j$). This deviation arises from quadrature and measurement errors, as well as any uncertainties as to the exact form of the kernel function [in this case, $\pi r^2 Q_{ext}(r,\lambda,m)$].

Returning to (3), it follows that the elements of (4) are given by

$$g_i=\tau_M(\lambda_i), \qquad i=1, 2, \ldots, p,$$

$$A_{ij}=\int_{r_j}^{r_{j+1}} \pi r^2 Q_{ext}(r,\lambda_i,m)h(r)dr, \quad j=1, 2, \ldots, q, \qquad (5)$$

$$f_j=f(\bar{r}_j),$$

where $\bar{r}_j$ are the midpoints of the coarse intervals. Writing (5) as a quadrature results in an expression similar to that obtained by Herman *et al.* (1971) where their weighting functions $W_{k_j}(r_{k_j})=h(r_{k_j})\Delta r_{k_j}$ in the terminology used here and r_{k_j} are the midpoint radii of the subintervals.

In terms of an integral over $x=\log r$, Eq. (5) may be rewritten as

$$A_{ij}=\int_{x_j}^{x_{j+1}} K(x,\lambda_i)dx, \qquad (6)$$

where

$$K(x,\lambda)=\pi 10^{3x}Q_{ext}(10^x,\lambda,m)\,h(10^x)\ln 10. \quad (7)$$

Eqs. (6) and (7) are those obtained by Yamamoto and Tanaka (1969) if $h(r)$ takes the form of a Junge (1955) size distribution,

$$h(r)=r^{-(\nu^*+1)}, \quad (8)$$

with ν^* assumed to have a value of 3.0.

Phillips (1962) and Twomey (1965) have discussed the instability in the solution vector f which results if (4) is directly solved by minimizing $\sum_i \epsilon_i^2$. Phillips suggested that, due to ever-present error, a constraint be added that discriminates against such instability. In order to select a physical solution among the family of solutions which satisfy (4), Phillips introduced a smoothing constraint such that the sum of the squares of the second derivatives of the solution points is minimized. For a quadrature of equal division, the solution vector f is obtained by minimizing a performance function Q, defined as

$$Q=\sum_{i=1}^{p}\epsilon_i^2+\gamma\sum_{j=2}^{q-1}(f_{j-1}-2f_j+f_{j+1})^2, \quad (9)$$

where γ is some non-negative Lagrange multiplier. Minimizing Q with respect to the unknown f_k coefficients, when γ equals zero, is equivalent to making an unweighted least-squares fit to the data.

Since it is further known that some of the $\tau_M(\lambda_i)$ measurements are more precise than others, it is desirable to include that *a priori* information in the mathematical formalism. For the case in which the measurements are correlated with known covariances $\sigma^2_{\tau_M\tau_M}$, a more general form of (9) would be

$$Q=\sum_{i=1}^{p}\sum_{j=1}^{p}C_{ij}^{-1}\epsilon_i\epsilon_j+\gamma\sum_{j=2}^{q-1}(f_{j-1}-2f_j+f_{j+1})^2, \quad (10)$$

where C_{ij} is an element of the measurement covariance matrix C whose elements are given by $C_{ij}=\sigma^2_{\tau_i\tau_j}$. This follows from the Gauss-Markov theorem in the absence of a constraint (see, e.g., Liebelt, 1967), and thus the minimum value of Q represents the statistically optimum estimate of f.

Following the method suggested by Twomey (1963) whereby the performance function is differentiated with respect to each of the f_k coefficients, a set of simultaneous equations results which may be written as

$$-\sum_{i=1}^{p}\sum_{j=1}^{p}C_{ij}^{-1}A_{ik}\epsilon_j+\gamma\sum_{j=1}^{q}H_{kj}f_j=0, \qquad k=1,2,\ldots,q, \quad (11)$$

where H_{kj} is an element of the smoothing matrix H defined by Twomey (1963) as

$$\mathbf{H}=\begin{bmatrix} 1 & -2 & 1 & 0 & 0 & & & & \\ -2 & 5 & -4 & 1 & 0 & 0 & & & \\ 1 & -4 & 6 & -4 & 1 & 0 & & & \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & & \\ & & & \ddots & & & & & \\ & & & & 0 & 1 & -4 & 5 & -2 \\ & & & & & 0 & 1 & -2 & 1 \end{bmatrix}. \quad (12)$$

The series of equations in (11) can be written in matrix form as

$$-\mathbf{A}^T\mathbf{C}^{-1}\boldsymbol{\epsilon}+\gamma\mathbf{H}\mathbf{f}=0, \quad (13)$$

where $\mathbf{A}^T$ is the transpose of A. Eliminating $\boldsymbol{\epsilon}$ between (4) and (13) leads to the solution

$$\mathbf{f}=(\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}+\gamma\mathbf{H})^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{g}. \quad (14)$$

For the case in which the statistical errors in the measurements are assumed equal and uncorrelated, C reduces to $s^2\mathbf{I}$, where s represents the constant rms error and I is the identity matrix. With this assumption, Eq. (14) reduces to the form derived by Twomey (1963).

For the attenuation problem considered here, the statistical errors in the measurements are assumed uncorrelated but known to be unequal (King and Byrne, 1976). As a consequence of this, the covariance matrix becomes diagonal with elements given by $C_{ij}=\sigma^2_{\tau_M}(\lambda_i)\delta_{ij}$, where δ_{ij} is the Kronecker delta function. This gives a relative weighting to each of the measurements, placing greater emphasis on those measurements which have the smallest error bars. With C defined in this manner, Eq. (14) is equivalent to making a weighted least-squares fit to the data subject to a constraint.

Initially, a zeroth-order weighting function $h^{(0)}(r)$ is assumed in (5) from which first-order $f^{(1)}(\bar{r}_i)$ values are computed with the aid of (14). Since the solution vector $\mathbf{f}^{(1)}$ represents a modifying factor to the assumed form of $h^{(0)}(r)$, the $f^{(1)}(\bar{r}_i)$ values are then used to calculate a first-order weighting function, $h^{(1)}(r)$, which better represents the size distribution than the initially assumed weighting function. The first-order weighting function is then substituted back into (5) from which a second-order $\mathbf{f}^{(2)}$ is obtained through (14). This iterative procedure is continued until a stable result is obtained (Herman *et al.*, 1971).

Two advantages result from separating the size distribution function into two parts as described above. The most obvious advantage is that the quadrature error which results from (3), when $f(r)$ is assumed constant in each coarse interval, will be less the better $h(r)$ comes to describing the size distribution. In fact, if the weighting function represents the size distribution exactly, f will be a vector whose elements are unity (Herman *et al.*, 1971). The second advantage, though less obvious, is equally important. Since the smoothing constraint minimizes the second derivatives

of the solution points on a *linear* scale, it is a much more appropriate constraint in cases in which $f(\bar{r}_j)$ is nearly constant. Since the columnar size distribution typically varies over many orders of magnitude, a direct inversion for $n_c(r)$ would consist of minimizing curvature of a function which implicitly has large curvature. A Junge distribution, for example, is nearly a delta function for large values of ν^* on a linear scale and hence has large curvature, even though it has zero curvature on a log scale (Twomey, personal communication).

For the results presented in this paper, the initial weighting functions were assumed to have the form of a Junge size distribution given by (8). In practice, several different values of ν^* are used to calculate the zeroth-order weighting function and the final results after successive iterations are intercompared. One test of the procedure is the similarity of the results obtained when different values of ν^* are used. This point will further be addressed in the next section.

In performing the inversion described above, it is necessary to select a value for γ. Since γ enters (14) in a manner such that elements of $\gamma\mathbf{H}$ are to be added to $\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$ to produce the desired smoothing, the magnitude of $\gamma H_{kj}/(\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A})_{kj}$ is of importance, not the magnitude of γ alone. In selecting γ, therefore, $\gamma_{rel}=\gamma H_{11}/(\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A})_{11}$ is allowed to vary in the range 10^{-8} to 1 until a minimum value of γ_{rel} is reached for which all elements of the solution vector $\mathbf{f}$ are positive (i.e., negative values of the elements of $\mathbf{f}$ constitute an unphysical solution).

3. Aerosol size distribution results

The method for determining the columnar aerosol size distribution described in the preceding section has been carried out at the University of Arizona since August 1975. The Mie optical depth measurements have been determined by the method described by King and Byrne (1976). In this procedure the ozone absorption optical depths, and hence total ozone content of the atmosphere, are inferred from the spectral variation of total optical depth in the visible and near infrared wavelength regions. Using the O_3 absorption optical depths determined in this manner, the values of the Mie optical depth are obtained by subtracting the molecular scattering and estimated O_3 absorption contributions from the total optical depth. At the present time inversions have been carried out for 57 different days, of which the results of a representative selection are presented below. For consistency of presentation, all inversions were performed assuming the complex refractive index of the aerosol particles was wavelength and size independent and given by $m=1.45-0.00i$. Although this value was arbitrarily selected, there is considerable evidence that the real part of the index of refraction lies somewhere between 1.33 (pure water) and 1.54 (silicate particles) as discussed by Yamamoto and Tanaka (1969) and others. The results of a combined direct sampling and remote sensing experiment over Tucson in November 1974 are suggestive of a real index between 1.40 and 1.45 for the aerosol particles (Reagan *et al.*, 1977). This refractive index is in good agreement with the refractive index of $H_2SO_4 \cdot b\ H_2O$ droplets which comprised about 30% of the particles analyzed by the electron microscope in this experiment (Reagan *et al.*, 1977). The effect of varying the refractive index assumed in the inversion procedure was addressed by Yamamoto and Tanaka (1969) where it was shown that the inverted size distribution maintains its *shape* under various values of refractive index when $m \approx 1.0$. The validity of this result for indices of refraction between 1.45 and 1.54 will be examined in Section 4.

Optical depth measurements made at between six and eight different wavelengths ranging between 0.4400 and 1.0303 μm have been used in the size distribution determinations to be presented below. Due to both the extinction cross sections (which increase significantly with radius) and the number densities of the natural aerosol particles (which normally decrease with radius), this spectral region of the attenuation measurements limits the radius range of maximum sensitivity to the large and giant aerosol particles only ($0.1 \lesssim r \lesssim 4.0$ μm). Although this matter has been considered by Yamamoto and Tanaka (1969) for both Junge- and Woodcock-type aerosol size distributions, it is very important to realize that there is no absolute rule which determines the radii limits having the most significant contribution to the attenuation measurements. As will be shown in the next section, this radius range is dependent on both the form of the size distribution function and the values of the Mie extinction cross sections over the radius range. Since the size distribution function is not known in advance, it is apparent that occasional trial and error is required in order to determine the radius range over which the inversion can be performed.

It is convenient, for purposes of illustration, to categorize the inversion results according to both the form of the columnar aerosol size distribution and the spectral dependence of the Mie optical depths. In lieu of $n_c(r)$ or, equivalently, dN_c/dr, the size distribution results are presented in terms of $dN_c/d\log r$, representing the number of particles per unit area per unit log radius interval in a vertical column through the atmosphere. In most cases the columnar aerosol size distributions can be classified in terms of three different types of distributions (although gradations between two different types are occasionally observed making this classification somewhat arbitrary). An example of the first type is illustrated in Fig. 1 for 13 August 1975, a day for which the spectral Mie optical depth measurements very nearly follow Ångström's (1929) empirical formula given by $\tau_M(\lambda)=\beta\lambda^{-\alpha}$. The observed Mie optical depths and corresponding standard devia-

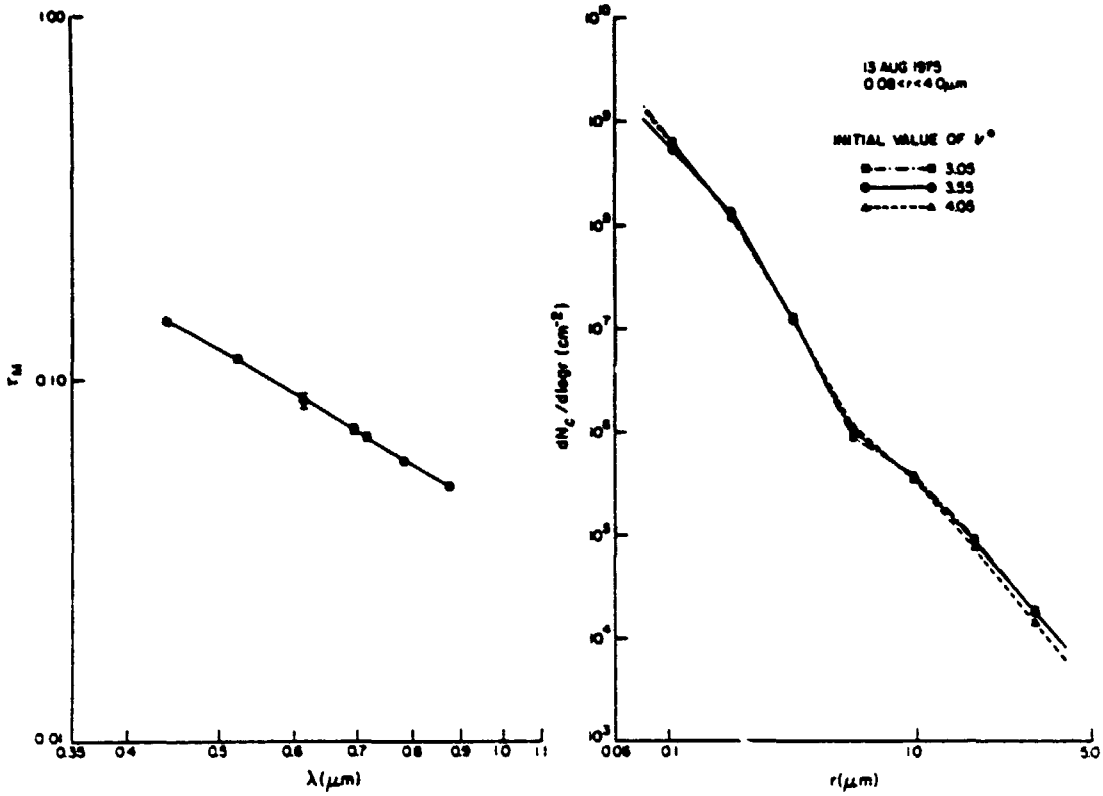


FIG. 1. Observed Mie optical depths and estimated size distributions for 13 August 1975. The three distribution curves on the right represent the results using different initial weighting functions (see text) while the curve on the left indicates the regression fit to the data using the inverted size distributions.

tions are shown in the left portion of the figure while the size distributions obtained by inverting these data for three initial ν^* cases are shown in the right portion. The solid curve in the left portion indicates how the inverted size distributions are able to reproduce the $\tau_M(\lambda)$ measurements (i.e., the direct problem $\mathbf{g}=\mathbf{Af}$).

In applying the inversion procedure described in the preceding section, several different initial Junge distribution parameters ν^* are assumed in formulating the zeroth-order weighting functions $h^{(0)}(r)$ so that the results after subsequent iterations can be intercompared. In practice, a best-fit value for the Ångström turbidity coefficient α is determined from the observed values of $\tau_M(\lambda)$ by linear least-squares methods, from which a corresponding ν^* value is determined from the well-known relationship $\nu^*=\alpha+2$. Inversions are then performed for three different values of the Junge parameter about this value (*viz.*, $\nu^*-0.5$, ν^* and $\nu^*+0.5$). For the data case illustrated in Fig. 1 $\alpha=1.55$, and thus inversions were performed using the initial values $\nu^*=3.05$, 3.55 and 4.05 as indicated in the right-hand figure. It is readily apparent that all initial values of the weighting functions $h^{(0)}(r)$ were able to be perturbed such that the results after subsequent iterations yielded solutions quite independent of the initial "guess."

The aerosol size distributions illustrated in Fig. 1 can best be described as constituting either Junge or two-slope types of size distributions. They have been observed on approximately 20% of all days examined thus far. Fig. 2 illustrates the spectral optical depth measurements and corresponding size distributions for three additional days (29 August and 29 October, 1975) where only one distribution function is shown for each day. In all instances the sensitivity to the initial weighting function $h^{(0)}(r)$ was negligible. The Mie optical depth data for cases of this type always tend to exhibit linear or slightly positive curved spectral dependences of $\log\tau_M(\lambda)$ vs $\log\lambda$ with relatively steep slopes ($\alpha\approx1.2$) and relatively high turbidities [$\tau_M(0.6120\,\mu m)\approx0.08$]. The one-to-one correspondence between the slope of the data (α) and the slope of the inverted size distribution (ν^*) is readily apparent on examination of Fig. 2.

On occasions when the Mie optical depths are small, $\tau_M(\lambda)$ tends generally to increase with wavelength (i.e., $\alpha<0.00$). On most of these occasions the spectral Mie optical depth measurements exhibit negative curvature. An example of this type is illustrated in Fig. 3 for the data of 24 October 1975. In a similar manner to the method described above, an Ångström turbidity coefficient was estimated from the data to be

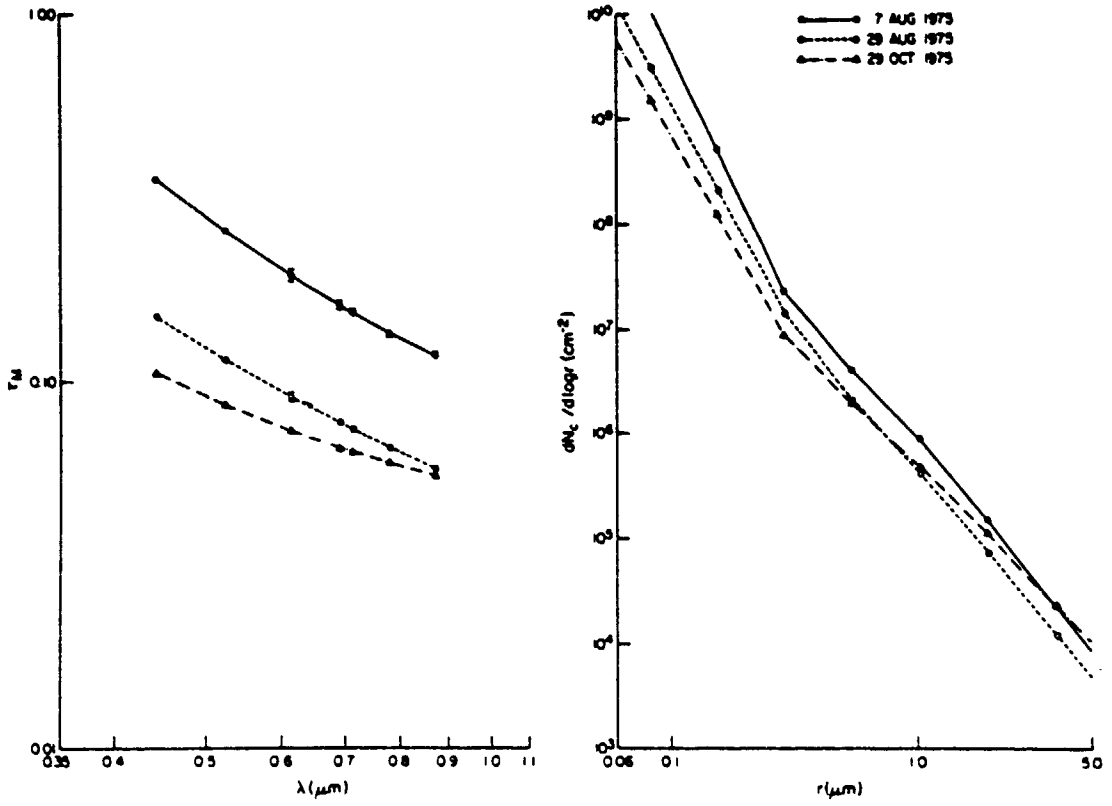


FIG. 2. Observed Mie optical depths and estimated size distributions for 7 August, 29 August and 29 October, 1975.

$\alpha = -0.27$, and thus inversions were performed using weighting functions having initial Junge parameters $\nu^* = 1.23$, 1.73 and 2.23. It is readily apparent on examination of Fig. 3 that distinctly non-Junge size distributions result from data of this type. This is not unexpected since the tendency for negative curvature on the part of the data suggests an absence of both small and large particles.

The three different size distributions illustrated in Fig. 3 indicate a relative insensitivity to the choice of the Junge parameter in the initial weighting functions except in the radius range where little retrievable information exists. The inversion result presented here was specifically selected because it clearly demonstrates that the inversion procedure is capable of perturbing the initial guess as required. Most data (but not all) do not necessitate such a drastic alteration of the initial weighting function $h^{(0)}(r)$. All days for which the Mie optical depths exhibit negative curvature while increasing with wavelength produce inverted size distributions which are relatively monodisperse, as is the case with the data of Fig. 3. These types of distributions have been observed on approximately 30% of all days examined thus far.

Fig. 4 illustrates the spectral optical depth measurements and corresponding size distributions for 12 November 1975 and 6 and 15 May 1977, days for which the $\tau_M(\lambda)$ data increase with wavelength. In all instances the inverted size distributions are relatively monodisperse in character with very little sensitivity to the initial weighting function $h^{(0)}(r)$. Data of this type are often difficult to invert due to the problems associated with determining the radius range having the major contribution to the attenuation measurements. For this reason it is frequently necessary to invert a data set several different times with slight alterations in the radius range. To this end, the agreement between distributions having different initial weighting functions affords a very convenient subjective test of stability. None of the attenuation measurements over the Chesapeake Bay (Knestrick *et al.*, 1962) which were considered by Yamamoto and Tanaka (1969) showed any spectral dependencies similar to Figs. 3 and 4. The large Junge parameter ($\nu^* = 3.0$) used in their weighting functions would probably have made it more difficult to invert any data of this type had they occurred.

Not all spectral optical depth measurements which exhibit negative curvature or negative Ångström turbidity coefficients [i.e., $\tau_M(\lambda)$ increases as a function of λ] show quite as dramatic a monodisperse character to the size distributions as the data of Figs. 3 and 4. Fig. 5 illustrates the spectral optical depth measurements and corresponding size distributions for 10 October

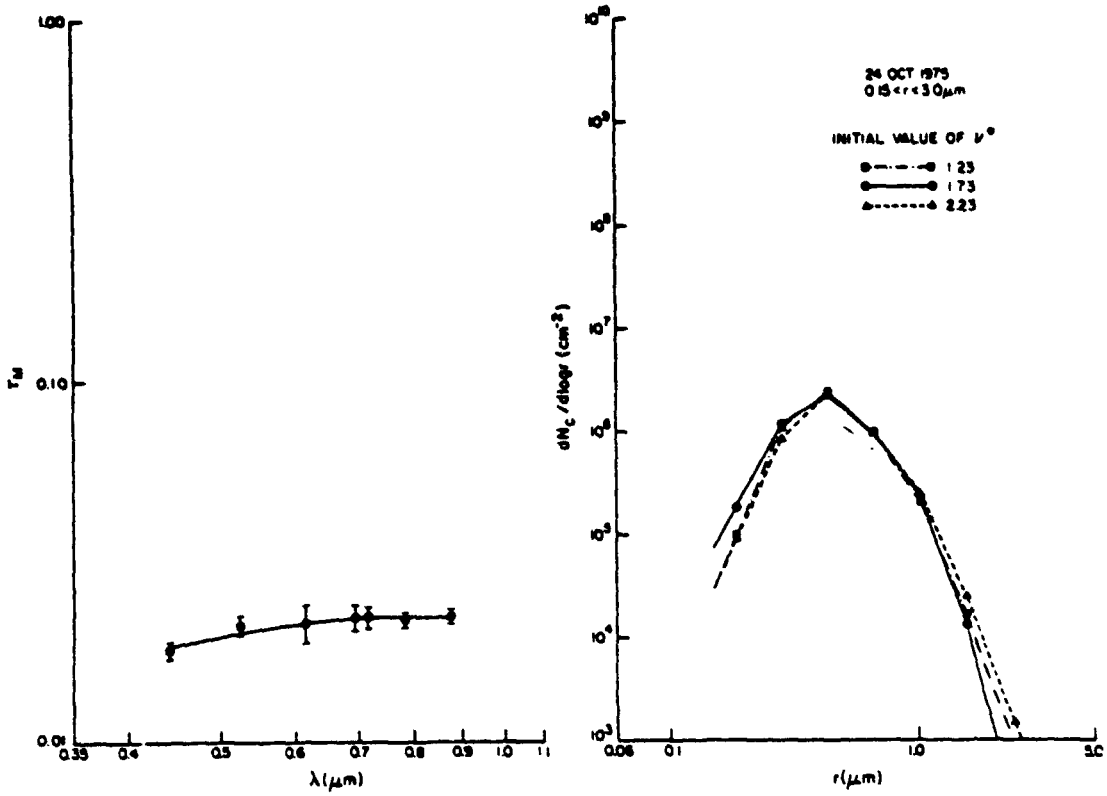


FIG. 3. As in Fig. 1 except for measurements collected on 24 October 1975.

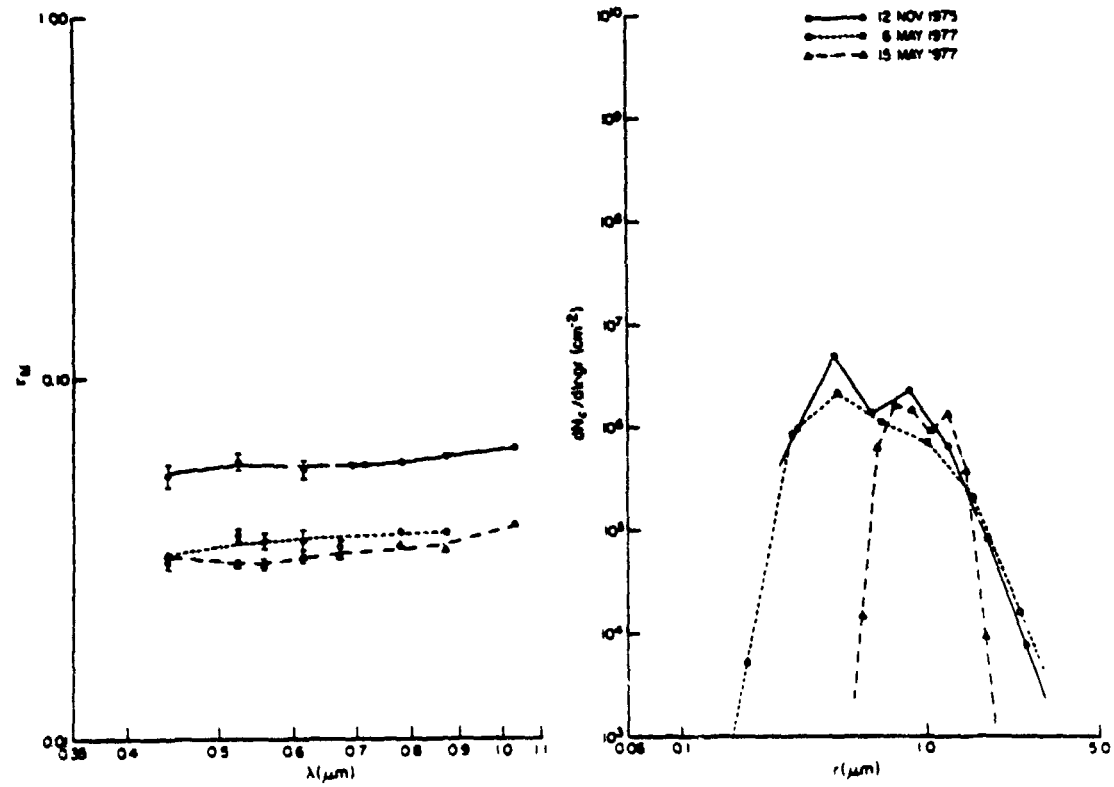


FIG. 4. As in Fig. 2 except for measurements on 12 November 1975, 6 May 1977 and 15 May 1977.

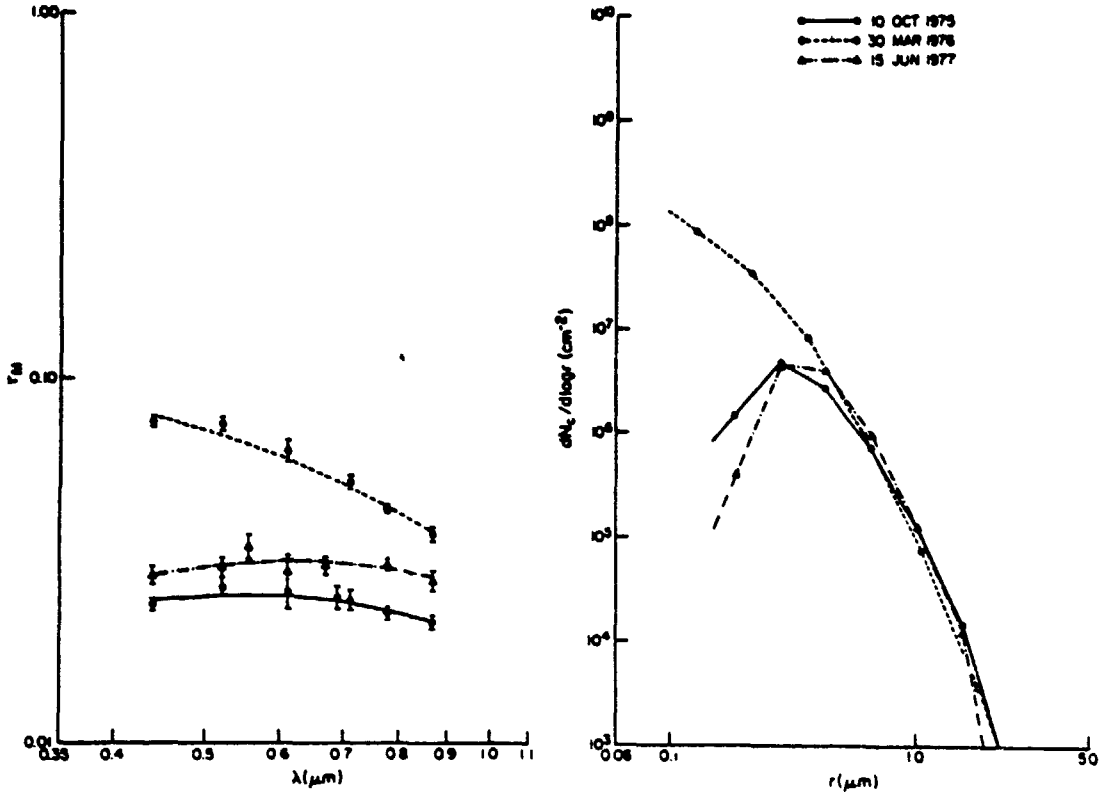


FIG. 5. As in Fig. 2 except for measurements on 10 October 1975, 30 March 1976 and 15 June 1977.

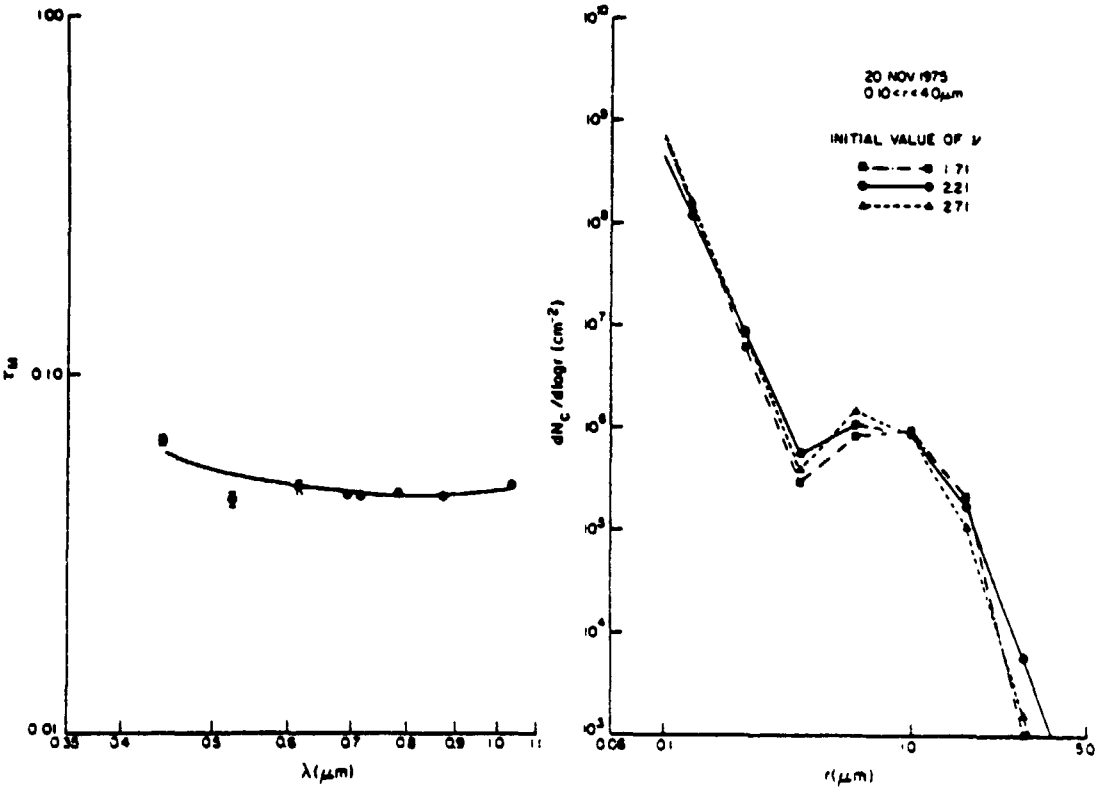


FIG. 6. As in Fig. 1 except for measurements collected on 20 November 1975.

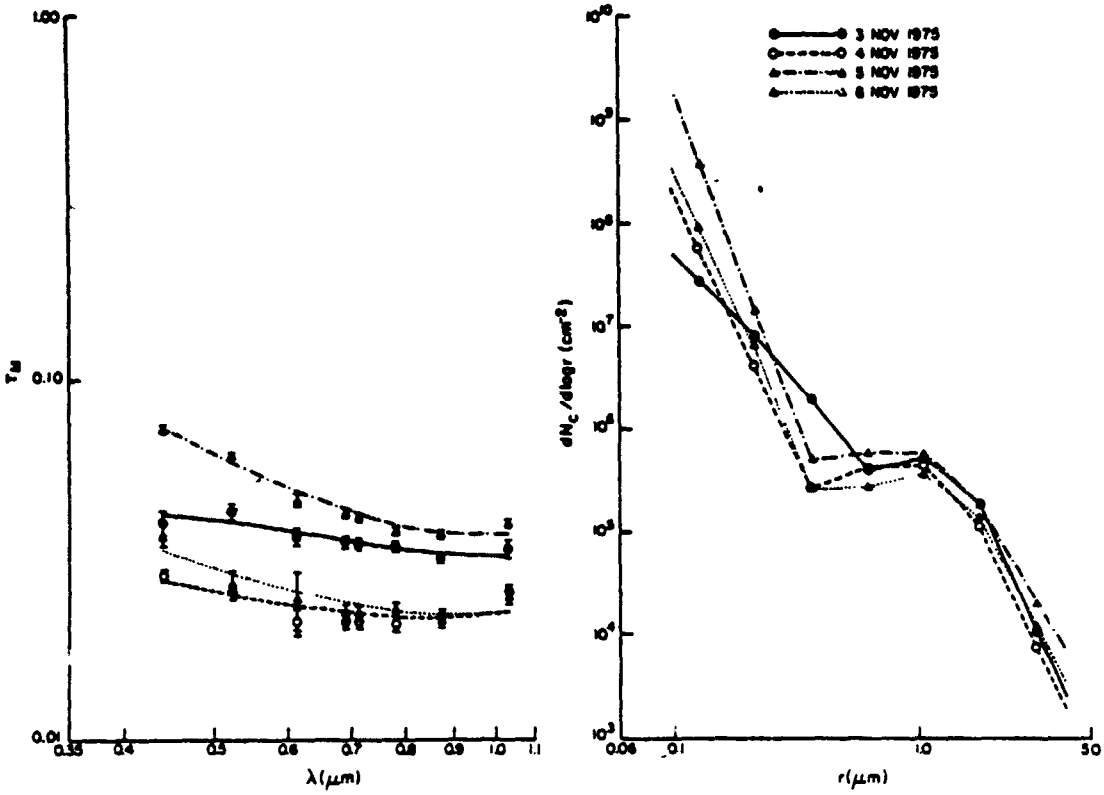


FIG. 7. As in Fig. 2 except for measurements on four consecutive days from 3 to 6 November 1975.

1975, 30 March 1976 and 15 June 1977. On these days (as well as several others) the $\tau_M(\lambda)$ measurements had characteristic negative curvature resulting in size distributions which tended to be peaked but quite broad and asymmetric. There appears to be correspondence between the slope of the $\tau_M(\lambda)$ measurements at the short wavelengths and the radius at which the distribution attains a maximum. The Mie optical depth measurements which produce this second class of distributions (type II) are typically very small [$\tau_M(0.6120\ \mu m) \approx 0.03$] with Ångström turbidity coefficients which are typically negative ($\alpha \approx -0.23$). The single data case of 30 March 1976 (see Fig. 5) deviates the most from these values, and the corresponding size distribution could equally well be categorized as a type I (Junge-type) distribution, thus making classification according to three distinct types somewhat arbitrary.

Perhaps the most interesting distribution type which has been observed thus far is one for which the spectral optical depth values are intermediate between the large turbidities of type I and the small turbidities of type II. On occasions when the Mie optical depths are intermediate in magnitude [$\tau_M(0.6120\ \mu m) \approx 0.05$], $\tau_M(\lambda)$ tends generally to decrease with wavelength ($\alpha \approx 0.5$) but with positive curvature. An example of this type (which will be referred to as type III) is illustrated in Fig. 6 for the data of 20 November 1975. On this day the Ångström turbidity coefficient was estimated to be 0.21, and thus inversions were performed using Junge distribution initial weighting functions with $\nu^* = 1.71$, 2.21 and 2.71. Fig. 6 clearly demonstrates that not only do non-Junge aerosol size distributions result but that the inversion procedure is quite capable of perturbing the initial weighting functions $h^{(0)}(r)$ as required. All days for which the Mie optical depths exhibit relatively small, but positive, Ångström turbidity coefficients with positive curvature imply aerosol size distributions which can be represented as a combination of a power law (type I) and a relatively monodisperse (type II) distribution, as is the case with the data of Fig. 6. These types of distributions have been observed on about 50% of the days examined thus far. It is consistent with the present results, however, to consider types I and II to be subsets of the more general type of distribution shown in Fig. 6 since all distributions can be represented by a combination of types I and II where the relative concentrations of these two types vary in space and time.

Fig. 7 illustrates the spectral optical depth measurements and corresponding aerosol size distributions for an episode of four consecutive days in November 1975. Note in particular that $\tau_M(\lambda)$ for the three longest wavelengths is very similar in shape for all four days as is the size distribution function for the larger particles. However, the slopes of the $\tau_M(\lambda)$ measurements

at the shorter wavelengths are directly correlated with the slopes of the Junge part of the aerosol size distributions at the smaller radii.

The resulting two-component size distributions shown in Figs. 6 and 7 are consistent with the production mechanisms of atmospheric aerosol particles. Those particles with radii $\lesssim 0.5$ μm are produced by a combination of nucleation from the gas phase and subsequent coagulation. Particles having radii $\gtrsim 1.0$ μm are principally the result of mechanical and wind stresses at the earth's surface. It is known from aerosol physics that these production mechanisms are not very effective in the 0.5–1.0 μm range, and thus it is not surprising that a separation of two regions might sometimes be observed. A more complete discussion of aerosol physics can be found in Twomey (1977). A review of the multimodal nature of size distributions and their physical characteristics is given by Whitby (1978).

Curcio (1961) assumed that the atmospheric particulates were composed of such a two-component system with the smaller particles given by a Junge distribution and the larger ones by a Woodcock or Gaussian distribution. Most of the spectral attenuation measurements he presented, as well as theoretical computations using this model, have wavelength characteristics similar to those presented above for wavelengths in the visible and near-infrared regions. Curcio similarly found that the higher attenuation data cases tended to be dominated by Junge-type aerosol size distributions in essential agreement with the results presented above.

Although Yamamoto and Tanaka (1969) inverted data with wavelength characteristics similar to those in Figs. 6 and 7, their inversions show none of the bimodal characteristics which we have observed. This may be partly due to their representing the size distributions on a scale of $n(r)$ or dN/dr, rather than $dN/d\log r$, which tends to stretch the appearance of the distributions.

As pointed out earlier, gradations between clear-cut distribution types occasionally occur, making classification into three distinct categories somewhat arbitrary. Fig. 8 illustrates the spectral optical depth measurements and corresponding size distributions for 1 and 3 November 1976 and 18 May 1977, days for which the aerosol size distributions are very similar and the optical depth data show very subtle curvature differences. In estimating the relative frequencies of abundance, the size distribution for 3 November 1976 was considered to be partly type I and partly type III, while that for 18 May 1977 was considered to be type I. The distribution on 18 May 1977, in particular, could have been placed with some justification in any of the three categories.

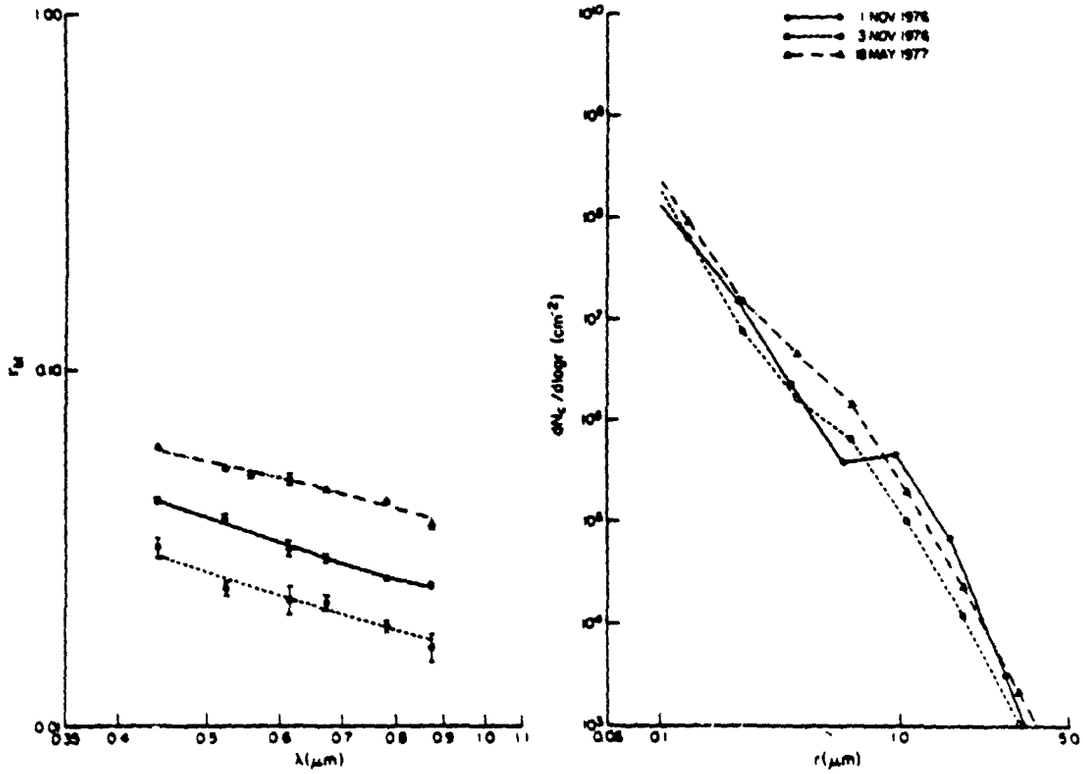


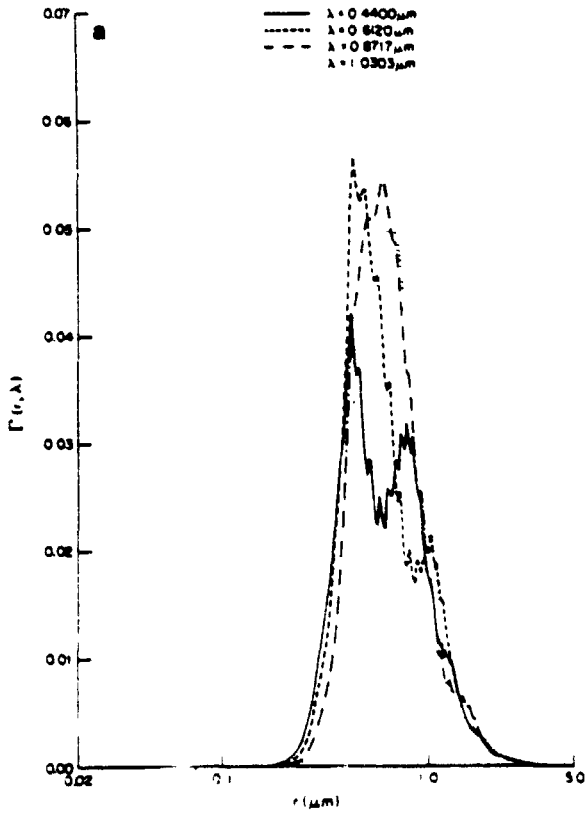
FIG. 8. As in Fig. 2 except for measurements on 1 November 1976, 3 November 1976 and 18 May 1977.

4. Practical considerations

Some of the practical difficulties associated with obtaining satisfactory size distribution determinations from inversion of attenuation coefficients as a function of wavelength or angular scattering intensities as a function of scattering angle have been discussed by various investigators (e.g., Yamamoto and Tanaka, 1969; Dave, 1971). Two aspects of the remote sensing determination of aerosol size distributions, namely, the radii limits of maximum sensitivity and the assumed index of refraction of the atmospheric aerosol particles, are worth reexamining in the context of the present investigation. The necessity of weighting the kernel by a modifying function $h(r)$, as well as the need to scale the Lagrange multiplier γ according to the magnitude of one of the elements of the matrix $\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$, has been discussed previously.

In Section 3 it was pointed out that the radius range of the atmospheric particulates which contributes the most significantly to the Mie optical depth measurements is itself a function of the aerosol size distribution to be determined. In order to see why this should be the case, consider the contribution function $\Gamma(r,\lambda)$ defined as

$$\Gamma(r,\lambda) = \pi r^2 Q_{ext}(r,\lambda,m)\frac{dN_c}{d\log r}. \qquad (15)$$



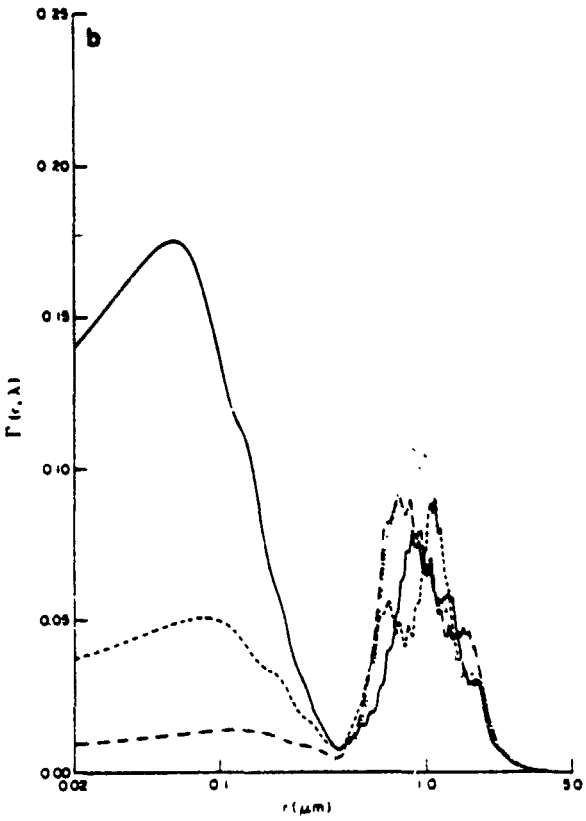


FIG. 9. Contribution function $\Gamma(r,\lambda)$ as a function of $\log r$ and λ for (a) the aerosol size distribution of 24 October 1975 (see Fig. 3) and (b) the aerosol size distribution of 20 November 1975 (see Fig. 6).

With this definition,

$$\tau_M(\lambda) = \int_{\log r_a}^{\log r_b} \Gamma(r,\lambda)\, d\log r, \qquad (16)$$

and thus the integral under a curve of $\Gamma(r,\lambda)$ vs $\log r$ represents the regression fit to the $\tau_M(\lambda)$ measurement for the spectral band which is centered at λ. The contribution function defined here is similar to the probability density function defined by Fraser (1975). However Fraser uses a unit volume particle size distribution based on dN/dr, and normalizes the extinction cross section by the resulting attenuation coefficient.

Fig. 9a illustrates the contribution function $\Gamma(r,\lambda)$ as a function of $\log r$ for the columnar aerosol size distribution of 24 October 1975 ($\nu^* = 2.23$ case of Fig. 3) and for four different wavelengths (0.4400, 0.6120, 0.8717 and 1.0303 μm) on this day. Fig. 9a indicates that the radii which contribute to the spectral Mie optical depth measurements on 24 October 1975 are restricted to lie in the range $0.15\ \mu\text{m} \lesssim r \lesssim 3.0\ \mu\text{m}$. With this type of distribution (i.e., type II where τ_M increases with wavelength), it generally follows that the radius

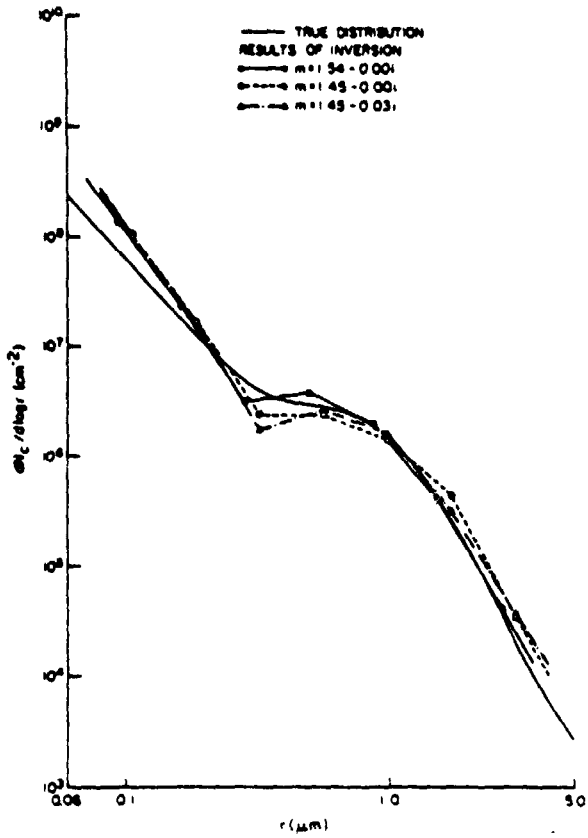


FIG. 10. Inversion solutions obtained for simulated $\tau_M(\lambda)$ data computed for an aerosol size distribution having a combination of Junge plus log-normal composites and a refractive index of $m = 1.54 - 0.00i$. The refractive indices assumed in the inversions are indicated in the figure.

range for which there is significant contribution to the optical depth measurements is more restricted than for other distribution types. Attempting to perform an inversion for radii beyond this sensitive region results in an unstable, highly oscillatory and hence unphysical solution. As a consequence, some trial and error is usually required in order to isolate the proper radius interval. Therefore, for a set of optical depth measurements over a given spectral range, it does not necessarily follow that the upper and lower limits for sensitivity are fixed as implied by Yamamoto and Tanaka (1969) but instead vary, depending on the form of the unknown size distribution.

As an example of this, a very different contribution function is obtained for the more typical composite size distributions having both Junge- and log-normal type components (type III). Fig. 9b illustrates the contribution function for the size distribution of 20 November 1975 ($\nu^* = 2.71$ case of Fig. 6). The contribution function in the radius interval $0.4\ \mu m \lesssim r \lesssim 4.0\ \mu m$ is primarily produced by the log-normal part of the aerosol size distribution and therefore has a form similar (although shifted to larger radii) to the contribution function of Fig. 9a. This demonstrates that a distribution with a greater abundance of large particles can have the benefit of extending the range of maximum sensitivity to somewhat larger radii than would otherwise be detectable were the distribution simply Junge with measurements extending only to $\lambda \approx 1.0\ \mu m$. The contribution function at the shorter radii is due primarily to the Junge part of the aerosol size distribution and therefore depends on both the effective ν^* value and the constant of the Junge distribution at these radii. Since the inverted aerosol size distributions for 20 November 1975 (and in fact most distributions of type III) were obtained by assuming a minimum radius of 0.1 μm, the contribution function below 0.1 μm represents only what would be observed were the slope of the Junge part of the size distribution extended to smaller radii. As Fig. 9b illustrates, significant contribution to the optical depth measurements could be produced below 0.1 μm at the shorter wavelengths. However, an integral over the extended radius range of $\Gamma(r,\lambda)$ would produce a wavelength dependence of $\tau_M(\lambda)$ which no longer has the gentle slope with well-developed positive curvature that is observed. This implies that the aerosol size distribution cannot continue with such a large slope (ν^*) for radii $< 0.1\ \mu m$ but must decrease at some point not far below 0.1 μm. The aerosol size distributions of type III are therefore bimodal with the radius of the first maximum being $\lesssim 0.1\ \mu m$, but with the specific radius being undetectable with spectral attenuation measurements at visible and near-infrared wavelengths.

The bimodal size distributions of type III are not normally as difficult to invert as those of type II but they do have certain idiosyncrasies of their own. If inversions are attempted using an upper radius which is too large (or occasionally too small), instabilities develop such that subsequent iterations produce more and more particles at the larger radii. When this occurs, the number of particles in the region between the Junge and log-normal parts of the size distribution are typically at a minimum. The inversion result within this intermediate region is intrinsically unstable if the number densities are small enough so that little contribution is made to the resultant optical depths. Both instabilities were found to be critically dependent on the value of the upper radius limit. A radius range of $0.1\ \mu m < r < 4.0\ \mu m$ is normally the most stable for data of this type (when inverting with $m = 1.45 - 0.00i$). When the proper radius range is isolated, all cases having different initial Junge distribution weighting functions, as well as all iterations within each case, are stable and similar as seen upon examination of Fig. 6.

The sensitivity of the inverted size distribution to index of refraction has been treated analytically by Yamamoto and Tanaka (1969). Using van de Hulst's (1957) anomalous diffraction theory, Yamamoto and Tanaka showed that the inverted aerosol size distributions maintain their *shape* under various indices of

refraction (shifting slightly in both magnitude and radius). Representing the refractive index as a complex number $m=n-\kappa i$ and extending Yamamoto and Tanaka's (1969) results to a scale of $dN_e(r)/d\log r$, it can be shown that

$$\frac{dN_e\left(\frac{0.54r}{n-1}\right)}{d\log r}=\left(\frac{n-1}{0.54}\right)^3\frac{dN_e(r)}{d\log r}. \tag{17}$$

In this expression $dN_e(r)/d\log r$ represents the size distribution obtained for $n=1.54$ and $dN_e[0.54r/(n-1)]/d\log r$ represents the size distribution obtained for a different real refractive index n where the radius is scaled as indicated. [Within the limits of applicability of the anomalous diffraction approximation, this expression is strictly applicable when $\kappa\ (n-1)$ is held fixed for both inversions.]

In order to see how well this performs for refractive indices between 1.45 and 1.54, we considered an aerosol size distribution consisting of a combination of Junge and log-normal components for radii between 0.02 and 10.0 μm. Assuming that this distribution function is applicable for $m=1.54-0.00i$, Mie optical depths were then computed at each of seven wavelengths between 0.4400 and 1.0303 μm. These values of $\tau_M(\lambda)$ were used without the addition of random error to determine particle size distributions by the method described in Section 2 for $m=1.54-0.00i$, $1.45-0.00i$ and $1.45-0.03i$. For these inversions the radii limits differed from that of the true distribution as described by the discussion on radius sensitivity. The results are presented in Fig. 10 where the solid curve represents the assumed model distribution and the broken curves are the results of computation (inversion) for the refractive indices indicated. The spectral optical depth calculations, although not explicitly illustrated, exhibit the characteristic wavelength dependence expected for this type of size distribution (i.e., little wavelength dependence but with positive curvature).

It can be seen on examination of Fig. 10 that the inversion result for the refractive index $m=1.54-0.00i$ agrees very well with the model distribution, particularly for $r\gtrsim 0.16$ μm. This is due to the fact that the refractive index assumed in the model distribution when generating the data is the same as that used in the inversion. The deviation which does occur between the true and computed distribution is due primarily to truncating the lower radius assumed in the inversion (in this case $r_a=0.07$ μm and $r_b=3.5$ μm). Extension of the lower radius much less than that used in this example, however, would lead to instabilities in the inversion.

Fig. 10 suggests that the inverted size distributions obtained when $n=1.45$ (for $r_a=0.08$ μm, $r_b=4.00$ μm) are quite similar regardless of the imaginary part of the complex refractive index (at least when $\kappa\lesssim 0.03$). It is similarly apparent that the *shapes* of the size distributions for $n=1.45$ and $n=1.54$ are similar as suggested by Yamamoto and Tanaka (1969). The reduction in magnitude and the shift to larger radii which does occur as the real part of the refractive index decreases is in essential agreement with that of (17).

Although no random error was added to the simulated measurements, the same set of data was used in performing inversions with all three refractive indices. This type of analysis has been performed for several other types of aerosol size distributions, including Junge and log-normal, and the results indicate that (17) is able to explain the major differences in the inversion for all three refractive indices considered. Additional sensitivity analysis using real data and a wider range of aerosol refractive indices is required to fully examine the problem of potential refractive index biases in the retrieval of aerosol size distribution from spectral attenuation measurements.

The effect of variations of the refractive index with wavelength and/or particle size has not been considered in the present investigation. Although these effects are not expected to be significant for most aerosol systems, one can imagine situations in which these effects can be large. Gillespie *et al.* (1978) considered the effects of an atmospheric aerosol consisting of small particles with one refractive index ($m=1.8-0.5i$) and large and giant particles with another refractive index ($m=1.5-0.0i$). They computed the volume extinction coefficient at three wavelengths between 0.55 and 1.06 μm and compared the results with those of a model with the same size distribution and a single refractive index ($m=1.50-0.005i$). Their results indicate that the difference in the extinction coefficient between the models is small with differences of 15% at 0.55 μm and only 2% at 1.06 μm. This effect would alter slightly the inverted size distribution result of the large aerosol particles ($0.1\lesssim r\lesssim 1.0$ μm) while having little effect on the size distribution of the giant particles ($r\gtrsim 1.0$ μm). For a less dramatic difference in the particulate refractive index with size range the differences in the extinction coefficient (or optical depth) computed with a mean refractive index will be even further reduced.

5. Summary and conclusions

An inversion solution to the linear system of equations $\mathbf{g}=\mathbf{A}\mathbf{f}$ has been derived which explicitly considers the magnitude of the standard deviations in a set of measurements $\mathbf{g}$. This formula [Eq. (14)] has been applied to the problem of determining the columnar aerosol size distribution from spectral measurements of the particulate (Mie) optical depth in the wavelength region $0.4400\ \mu\text{m}\lesssim\lambda\lesssim 1.0303\ \mu$m. An iterative method of solution is described whereby an estimate of the size distribution is included in the elements of the A matrix. With this procedure, the inverted solution vector f

amounts to a modifying function to the assumed form of the size distribution.

As demonstrated by the family of curves in the right portion of Figs. 1, 3 and 6, this iterative algorithm is quite capable of perturbing the initial estimates of the size distribution as required. In all instances three different Junge size distributions had been initially assumed so that the results after subsequent iterations could be intercompared. Tests with both real and generated spectral measurements of the Mie optical depth indicate that this method of solution is quite capable of inverting a wide variety of observation types. The only difficulty which normally arises is in selecting the radius range having the major contribution to the measurements.

The subject of radii limits has been discussed in some detail in Section 4 where it was shown that the radii which contribute most significantly to the magnitude of the Mie optical depth measurements vary somewhat with the type of the size distribution to be retrieved. Satisfactory size distribution determinations can normally be obtained for the radii range $0.10\ \mu m < r < 4.0\ \mu m$ when measurements are available for wavelengths throughout the visible and near-infrared regions. The greatest difficulty arises when the size distribution is a relatively monodisperse distribution contained in a narrow radius interval around 0.5 μm. Uncertainties in the complex refractive index have been found to have little effect on the spectral attenuation problem, particularly with regard to absorption, as illustrated by the family of curves in Fig. 10.

Columnar aerosol size distributions have been determined by inversion of spectral Mie optical depth measurements for 57 days in Tucson, Arizona. The optical depth measurements and corresponding aerosol size distributions are illustrated in Figs. 1–8 for a few of these days. The results generally indicate (at least for $r \gtrsim 0.1\ \mu m$) that the aerosol size distribution on a particular day can be represented either as a Junge distribution (type I), a relatively monodisperse distribution such as a log-normal or gamma distribution (type II), or as a two-component system consisting of a combination of both of these types (type III).

Type I distributions have been observed on 20% of the days examined thus far while types II and III have been observed on approximately 30 and 50% of the days, respectively. The type I distributions occur mainly in the fall and late spring in Tucson when the optical depths are the largest [$\tau_M(0.6120) \approx 0.08$] and when the spectral optical depth measurements exhibit linear or slightly positive curvature on a $\log\tau_M$ vs $\log\lambda$ scale with relatively steep slopes ($\alpha \approx 1.2$). Type II distributions, on the other hand, occur in the late fall and early summer when the optical depths are the lowest [$\tau_M(0.6120) \approx 0.03$]. Under these circumstances $\tau_M(\lambda)$ tends generally to increase with wavelength ($\alpha \approx -0.2$) while exhibiting negative curvature on a $\log\tau_M$ vs $\log\lambda$ scale. The size distributions of type III occur throughout the winter months when virtually no distributions of types I and II occur. These distributions occur when the optical depths are intermediate in magnitude [$\tau_M(0.6120) \approx 0.05$], the Ångström slope $\alpha \approx 0.5$, and the $\log\tau_M$ measurements show positive curvature as a function of $\log\lambda$ particularly at the longer wavelengths. Both the distributions and the optical depths of this type appear to be combinations of types I and II.

Curcio (1961) found that the majority of his observations of attenuation coefficient along a horizontal path in the Chesapeake Bay area could best be described as resulting from a combination of type I and type II distributions (i.e., type III). He also determined that larger optical depths resulted from Junge-type size distributions. This is in support of the results obtained in the present investigation for vertical attenuation at a different location. Rangarajan (1972) investigated the Ångström wavelength exponent for 520 observations at Poona, India. His results are in agreement with those reported here in that the lowest (near zero) α values occurred on occasions of lowest turbidity. Rangarajan's (1972) median value of 0.5 is the same as our type III distributions, those found to occur the most frequently.

Acknowledgments. The research reported in this article has been supported by the National Science Foundation under Grants DES75-15551 and ATM75-15551-A01 and the Office of Naval Research under Grant N00014-76-C-0438.

REFERENCES

Ångström, A., 1929: On the atmospheric transmission of sun radiation and on dust in the air. *Geogr. Ann.*, 11, 156–166.

Curcio, J. A., 1961: Evaluation of atmospheric aerosol particle size distribution from scattering measurements in the visible and infrared. *J. Opt. Soc. Amer.*, 51, 548–551.

Dave, J. V., 1971. Determination of size distribution of spherical polydispersions using scattered radiation data. *Appl. Opt.*, 10, 2035–2044.

Fraser, R. S., 1975: Degree of interdependence among atmospheric optical thicknesses in spectral bands between 0.36–2.4 μm. *J. Appl. Meteor.*, 14, 1187–1196.

Gillespie, J. B., S. G. Jennings and J. D. Lindberg, 1978: Use of an average complex refractive index in atmospheric propagation calculations. *Appl. Opt.*, 17, 989–991.

Grassl, H., 1971. Determination of aerosol size distributions from spectral attenuation measurements. *Appl. Opt.*, 10, 2534–2538.

Herman, B. M., S. R. Browning and J. A. Reagan, 1971. Determination of aerosol size distributions from lidar measurements. *J. Atmos. Sci.*, 28, 763–771.

Junge, C. E., 1955: The size distribution and aging of natural aerosols as determined from electrical and optical data on the atmosphere. *J. Meteor.*, 12, 13–25.

——, 1963 *Air Chemistry and Radioactivity*. Academic Press, 382 pp.

King, M. D., and D. M. Byrne, 1976: A method for inferring total ozone content from the spectral variation of total optical depth obtained with a solar radiometer. *J. Atmos. Sci.*, 33, 2242–2251.

Knestrick, G. L., T. H. Cosden and J. A. Curcio, 1962: Atmospheric scattering coefficients in the visible and infrared regions. *J. Opt. Soc. Amer.*, 52, 1010–1016.

Liebelt, P. B., 1967: *An Introduction to Estimation Theory*. Addison-Wesley, 273 pp.

Phillips, D. L., 1962. A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput. Mach.*, 9, 84–97.

Quenzel, H., 1970: Determination of size distribution of atmospheric aerosol particles from spectral solar radiation measurements. *J. Geophys. Res.*, 75, 2915–2921.

Rangarajan, S., 1972: Wavelength exponent for haze scattering in the tropics as determined by photoelectric photometers. *Tellus*, 24, 56–64.

Reagan, J. A., J. D. Spinhirne, D. M. Byrne, D. W. Thomson, R. G. dePena and Y. Mamane, 1977: Atmospheric particulate properties inferred from lidar and solar radiometer observations compared with simultaneous *in situ* aircraft measurements: A case study. *J. Appl. Meteor.*, 16, 911–928.

Shaw, G. E., J. A Reagan and B. M. Herman, 1973: Investigations of atmospheric extinction using direct solar radiation measurements made with a multiple wavelength radiometer. *J. Appl. Meteor.*, 12, 374–380.

Twomey, S., 1963: On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. *J. Assoc. Comput. Mach.*, 10, 97–101.

——, 1965: The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements. *J. Franklin Inst.*, 279, 95–109.

——, 1977: *Atmospheric Aerosols. Developments in Atmospheric Science*, Vol. 7, Elsevier, pp.

van de Hulst, H. C., 1957: *Light Scattering by Small Particles*. Wiley, 470 pp.

Whitby, K. T., 1978: The physical characteristics of sulfur aerosols. *Atmos. Environ.*, 12, 135–159.

Yamamoto, G., and M. Tanaka, 1969: Determination of aerosol size distribution from spectral attenuation measurements. *Appl. Opt.*, 8, 447–453.

Paper 122

Reprinted from JOURNAL OF THE ATMOSPHERIC SCIENCES, Vol. 36, No. 1, January 1979
American Meteorological Society
Printed in U. S. A.

Determination of the Ground Albedo and the Index of Absorption of Atmospheric Particulates by Remote Sensing. Part I: Theory[1]

MICHAEL D. KING[2] AND BENJAMIN M. HERMAN

Institute of Atmospheric Physics, The University of Arizona, Tucson 85721

(Manuscript received 31 May 1978, in final form 26 September 1978)

ABSTRACT

A statistical technique is developed for inferring the optimum values of the ground albedo and the effective imaginary term of the complex refractive index of atmospheric particulates. The procedure compares measurements of the ratio of the hemispheric diffuse to directly transmitted solar flux density at the earth's surface with radiative transfer computations of the same as suggested by Herman *et al.* (1975). A detailed study is presented which shows the extent to which the ratio of diffuse to direct solar radiation is sensitive to many of the radiative transfer parameters. Results indicate that the optical depth and size distribution of atmospheric aerosol particles are the two parameters which uniquely specify the radiation field to the point where ground albedo and index of absorption can be inferred. Varying the real part of the complex refractive index of atmospheric particulates as well as their vertical distribution is found to have a negligible effect on the diffuse-direct ratio. The statistical procedure utilizes a semi-analytic gradient search method from least-squares theory and includes a detailed error analysis.

1. Introduction

Within the past decade there has been an increasing concern about the observed decline in the mean temperature of the Northern Hemisphere which began in the 1940's. Because this decline followed a steady rise in mean temperature which began at the end of the last century, there have been countless attempts at explaining this behavior. Bryson (1968) discusses the various factors which could alter the basic climate of of the earth. These factors have included a change in sunspots, an increase in the carbon dioxide content of the atmosphere due to the consumption of fossil fuels, and a change in the properties and number of particulates in the atmosphere. All of these factors influence the complicated energy budget of the earth-atmosphere-ocean system.

Solid and liquid particles suspended in the atmosphere affect climate through two major processes. They directly affect the transfer of radiant energy in the clear atmosphere as well as affecting the optical and microphysical properties of clouds (Twomey, 1974, 1977; Ackerman and Baker, 1977). The first of these processes will be considered in some detail in this series of articles.

The first successful applications of radiative transfer models to the problem of actual particles in the atmosphere were made by Rasool and Schneider (1971) and Yamamoto and Tanaka (1972). The latter-named authors have computed the global albedo which results as particles of varying values of the imaginary index of refraction (index of absorption) are introduced into a model atmosphere. They conclude that an increasing turbidity will always lead to an enhanced global albedo (i.e., cooling) if the index of absorption is less than 0.05.

Since Yamamoto and Tanaka only considered surfaces with a ground albedo less than 0.15, Wang and Domoto (1974) and Herman and Browning (1975) extended the analysis to include surfaces with greater reflection. Both sets of authors found that heating can result if the surface reflectivity exceeds about 0.4, regardless of the index of absorption of the particulates. These studies address a very important point, that both ground albedo and index of absorption are important in assessing the effect of aerosol particles on the earth and its atmosphere.

It has been suggested in a preliminary study by Herman *et al.* (1975) that the ratio of the hemispheric diffuse to the directly transmitted solar flux density at the earth's surface be used as an indirect means of inferring the index of absorption of the atmospheric particulates. The diffuse radiation field in the earth's atmosphere at the shorter wavelengths (visible and ultraviolet) is complicated by its dependence on such things as the vertical distribution, optical depth, size distribution and index of refraction of the efficiently scattering atmospheric particulates as well as the position of the sun and the reflectivity of the earth's surface. In order to apply measurements of the diffuse radiation

[1] The research reported in this article was supported in part by the Office of Naval Research under Grants N000 14-75-C-0370 and N000 14-67-A-0209-0[illegible] by the National Science Foundation under Grants DES75-[illegible]551 and ATM75-15551-A01.

[2] Present affiliation: Goddard Space Flight Center, Greenbelt, MD 20771.

0022-4928/79/010163-11$06.75
© 1979 American Meteorological Society

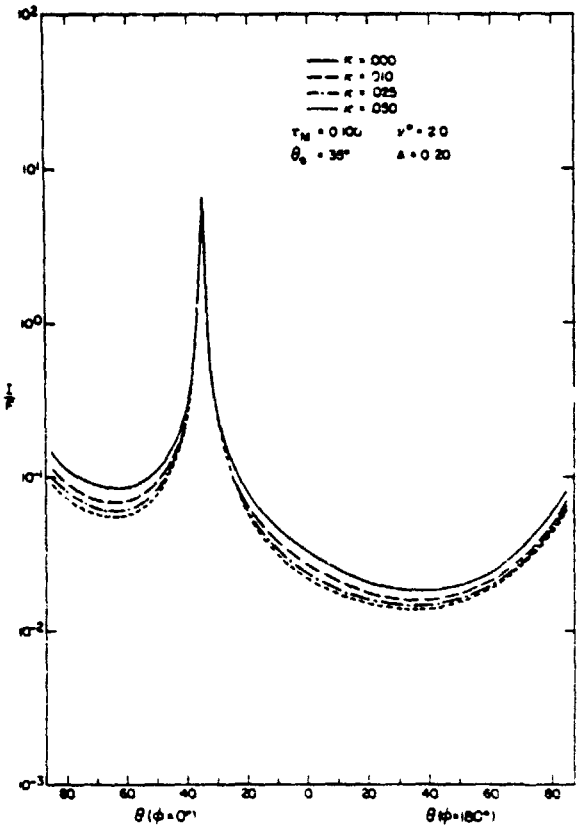


FIG. 1. The ratio of the transmitted intensity to directly transmitted solar flux density for $n=1.54$ and for four values of the imaginary part of the complex refractive index.

field to theoretical calculations of the same, as Herman *et al.* (1975) have suggested, it therefore becomes necessary to establish the relative importance and sensitivity of the radiation field to all of the radiative transfer parameters.

This paper considers the sensitivities and assumptions of primary importance for a successful application of theoretical calculations to real data in a real atmosphere. It will be demonstrated that the ratio of the hemispheric diffuse to directly transmitted solar flux density at the earth's surface can be measured as a function of solar zenith angle and, with certain *a priori* information about the Mie optical depth and aerosol size distribution, compared to theoretical calculations to assess the magnitude of both the ground albedo and the index of absorption of atmospheric particulates. An inversion procedure is developed which makes use of the laws of diffuse reflection and transmission together with a gradient search method from least-squares theory. Formulas are derived for estimating the magnitude of the errors in both ground albedo and the imaginary index of refraction of the particles.

In Part II of this series (King, 1979) measurements of the diffuse-direct ratio which have been collected in Tucson are presented and discussed. Data for days during which the atmosphere was clear and stable are analyzed and the optimum values of the ground albedo and index of absorption of atmospheric aerosol particles are presented.

2. Theoretical sensitivity of the diffuse radiation field

To assess the sensitivity of the diffuse radiation field to the many radiative transfer parameters mentioned above, computations were made for 364 model atmospheres illuminated from above by the sun at five solar zenith angles. The method used to calculate the diffuse radiation field is described by Herman and Browning (1965). This solution to the equation of radiative transfer consists of a Gauss-Seidel iteration technique and many numerical integrations over optical depth (τ), zenith angle ($\theta=\cos^{-1}\mu$) and azimuth angle (ϕ). In all calculations to be presented below, the bottom boundary of the atmosphere is assumed to reflect radiation according to Lambert's law with some albedo A. The computations are made for a plane-parallel, vertically inhomogeneous atmosphere with the dust vertically distributed according to Elterman (1968). In all calculations the wavelength of incident illumination is $\lambda=0.5550$ μm which, at the pressure level of Tucson, has a Rayleigh (molecular) optical depth $\tau_R=0.0860$. The phase matrix for the particles is computed from Mie theory [see van de Hulst (1957) for details] and is thus a function of the complex refractive index and size distribution of the particulates, the shape having been assumed spherical. Absorption by atmospheric ozone is neglected in all computations for reasons discussed below.

The aerosol size distribution for the sensitivity tests was assumed to be that proposed by Junge (1955), given as

$$n(r)=Cr^{-(\nu^*+1)}, \tag{1}$$

where C and ν^* are constants and $n(r)$ is the number density of particles per unit interval of radius r. All model calculations presented below are for Junge distributions where the radius r extends from 0.01 to 10.01 μm.

Fig. 1 illustrates the sensitivity of the transmitted intensity field in the $\theta=0°$ and 180° plane (i.e., the vertical plane containing the sun) to the imaginary term κ of the complex refractive index of the aerosol particles. The refractive index $m=n-\kappa i$ is a complex number: the real part n is the ordinary index of refraction of the material, while the imaginary part κ determines the absorption of electromagnetic radiation. All intensities are plotted relative to the directly transmitted solar flux density at the earth's surface, given by the Lambert-Beer law as

$$F=F_0e^{-\tau_t/\mu_0}. \tag{2}$$

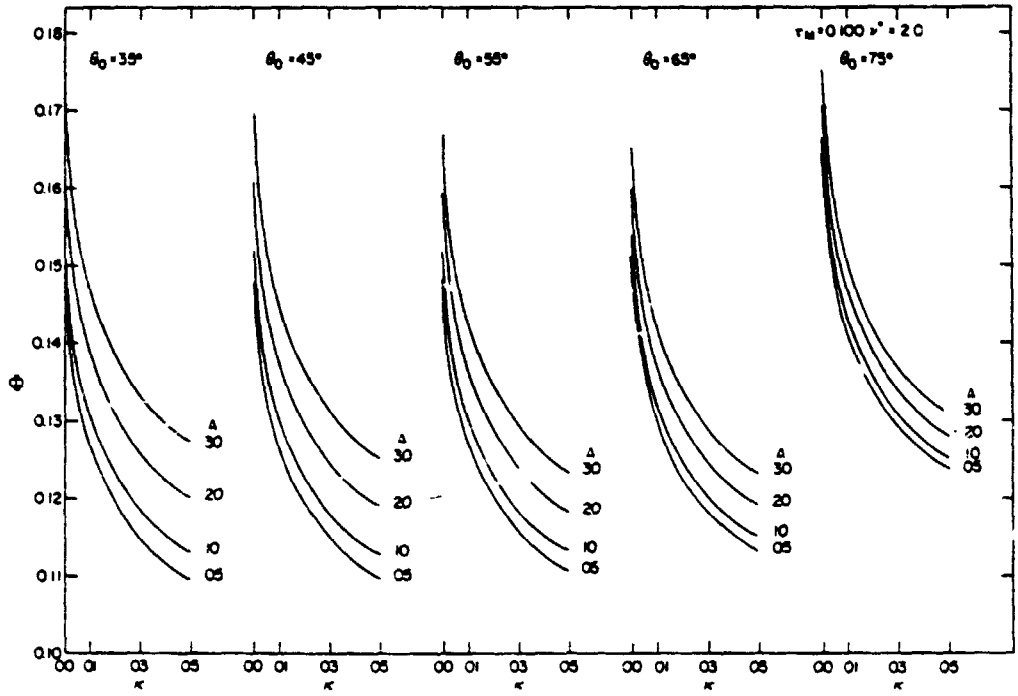


FIG. 2. The ratio of the hemispheric diffuse to directly transmitted solar flux density as a function of index of absorption, ground albedo and solar zenith angle.

In this expression F_0 is the solar flux density incident on the top of the atmosphere ($\tau=0$ level), μ_0 the cosine of the solar zenith angle θ_0, and τ_t the total optical depth of the atmosphere. From measurements of the directly transmitted solar radiation at wavelengths for which absorption due to molecular species other than ozone can be neglected, it is possible to determine the optical depth τ_M due to particulates alone, known as the Mie optical depth (King and Byrne, 1976).

Although the particular model atmosphere illustrated in Fig. 1 has $\tau_M=0.100$, $\nu^*=2.0$, $\theta_0=35°$, $A=0.20$ and $n=1.54$, the obvious sensitivities in the figure are similar for all other model atmospheres investigated, models which had values of A from 0.05 to 0.30, κ from 0.00 to 0.20, τ_M from 0.00 to 0.30 and ν^* from 2.0 to 4.0. In general, the solar aureole (the strong intensity region in the vicinity of the sun) is very insensitive to κ, being primarily produced by Fraunhofer diffraction around the particles, while the intensity at larger scattering angles shows a marked sensitivity to absorption.

Computations of the intensity field similar to Fig. 1 have been compared to measurements of the intensity by Herman *et al.* (1971) where theoretical calculations were made using measured values for τ_M and θ_0, while assumed values were used for the index of refraction, Junge size distribution parameter ν^* and ground albedo. Although no attempt was made to optimize the assumed values of ν^*, A and n, the agreement between measurement and theory was considered quite acceptable, and demonstrated for the first time that inclusion of aerosol particles into a Rayleigh atmosphere yields better agreement between observation and theory. It is possible, however, to optimize the determination of the index of refraction and ground albedo by examining the diffuse radiation field as a function of solar zenith angle. The method by which this is accomplished is presented in Section 4.

The hemispheric diffuse flux densities at the earth's surface are given by

$$F^{\pm}(\tau_t)=\int_0^{2\pi}\int_0^1 I(\tau_t,\pm\mu,\phi)\mu d\mu d\phi, \qquad (3)$$

where the $+(-)$ sign refers to intensities propagating in the upward (downward) direction at the level τ_t. If the downward hemispheric diffuse flux density is calculated using (3), it follows that $F^-(\tau_t)$ should monotonically decrease as κ increases since the intensity field has this behavior for all observation angles (μ,ϕ) as seen upon examination of Fig. 1. Fig. 2 presents a family of curves for the diffuse flux density versus imaginary index of refraction κ (hereafter referred to as the index of absorption) where again the diffuse flux density is plotted relative to the directly transmitted solar flux density. This ratio,

$$\Phi=F^-(\tau_t)/F_t \qquad (4)$$

shows the required sensitivity to κ as well as sensitivities to θ_0 and A. Yamamoto and Tanaka (1972) investi-

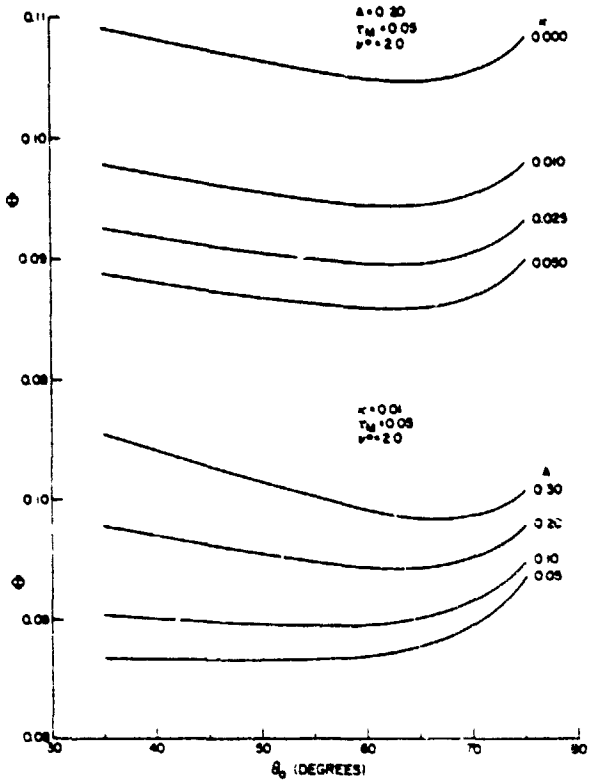


FIG. 3. The diffuse-direct ratio as a function of solar zenith angle for $\tau_M = 0.05$, $\nu^* = 2.0$, and for different values of κ and A.

gated some of the sensitivities of Φ, including its dependence on κ and λ, and came to the same conclusions for the sensitivity of Φ to absorption. However, Yamamoto and Tanaka made no attempt to investigate this sensitivity as a potential remote sensing means of inferring κ and A directly. Herman *et al.* (1975), on the other hand, show curves similar to Fig. 2 and suggest using the sensitivity of Φ, termed the diffuse-direct ratio, to infer the index of absorption of the atmospheric particulates.

In addition to the obvious instrumental advantage of making a relative (ratio) measurement in lieu of an absolute flux density measurement, there is a theoretical advantage associated with formulating the ratio of diffuse to direct solar flux density. Since ozone has a very broad absorption feature extending from about 0.4 to 0.9 μm, known as the Chappuis continuum, ozone must be considered when calculating the diffuse radiation field at the earth's surface having wavelengths in this region. Since most of the absorption by O_3 occurs above 20 km while the majority of scattering by molecules and particles is confined to heights well below 20 km, O_3 serves mainly to attenuate the direct solar beam before the radiation can interact with the scattering atmosphere. It is expected, therefore, that O_3 absorption will affect the diffuse and direct flux densities by nearly the same fraction such that the diffuse-direct ratio will be largely insensitive to absorption by O_3. Since approximately 7.5% of the total ozone content of the earth's atmosphere is located within the troposphere (Elterman, 1968), however, there is an imperfect separation with height. As a direct consequence of the presence of tropospheric O_3 one would expect to see some minor sensitivity to O_3 absorption particularly for wavelengths in the heart of the Chappuis band (centered near 0.6120 μm). This subject will further be addressed in part II of this series (King, 1979) where the experimental results are presented.

The sensitivity of the diffuse-direct ratio to solar zenith angle and ground albedo is also illustrated in Fig. 2, where it is apparent that the affect which a changing solar zenith angle has on Φ depends upon the values of κ and A. For small indices of absorption and large values of surface albedo (e.g., $\kappa = 0.01$, $A = 0.30$), Φ decreases slightly with θ_0 until about 65° after which point it starts to increase. For smaller values of the surface reflectivity, however, Φ increases with zenith angle for all $\theta_0 > 45°$, regardless of the index of absorption. Fig. 3 explicitly illustrates the combined effects of κ and A on the zenith angle dependence of the diffuse-direct ratio for a model atmosphere having $\tau_M = 0.05$ and $\nu^* = 2.0$. The family of curves for a fixed ground albedo (top portion of the figure) shows that the predominant effect of the index of absorption is to alter the *magnitude* of Φ with very little effect on its functional relationship with respect to solar zenith angle. Ground albedo, on the other hand, affects both the magnitude and *shape* of Φ as a function of θ_0. This is readily seen upon examination of Fig. 3 for the family of curves having a fixed value for the index of absorption of the atmospheric particulates (bottom portion of the figure). This suggests that a measurement of the diffuse-direct ratio versus solar zenith angle contains enough information to determine both the imaginary index of refraction of the atmospheric particulates as well as the reflectivity of the earth's surface, given a knowledge of the aerosol size distribution and Mie optical depth, provided a measurement accuracy on the order of 1% can be achieved. The instrumentation used to make the measurements was capable of this requirement (King, 1979).

Up to this point little mention has been made of the sensitivity of the diffuse-direct ratio to Mie optical depth and aerosol size distribution. This is because these parameters are not variables to be determined from the diffuse flux density measurements but are, instead, input parameters to be determined *a priori*. Since an error in either τ_M and or ν^* leads to errors in the determination of κ and A, the Mie optical depth and aerosol size distribution play a very important role in the successful application of the diffuse-direct technique to real data in a real atmosphere.

Fig. 4 shows Φ as a function of τ_M for two refractive indices and three Junge size distributions. As the Mie optical depth increases, the diffuse-direct ratio increases

for aerosol particles having a fixed refractive index and size distribution, a feature which can also be deduced from the results of Yamamoto and Tanaka (1972). It is obvious from this figure that any uncertainty in τ_M leads to an uncertainty in the determination of κ (and A), even if the aerosol size distribution is known. It is similarly apparent that any uncertainty in the aerosol size distribution will lead to incorrect assessments of the value for κ, even if the Mie optical depth is known with a high degree of accuracy. In general, neither τ_M nor $n(r)$ are known precisely, and thus they both must be estimated from alternative measurements.

The method developed for obtaining accurate Mie optical depth measurements (King and Byrne, 1976) optimizes the determination of the ozone optical depth, the major uncertainty in obtaining τ_M from measurements of the total optical depth. By making use of the spectral variation of Mie optical depth it is possible to obtain an estimate of the columnar aerosol size distribution $n_c(r)$, defined as the number of particles per unit area per unit radius interval in a vertical column through the atmosphere (King *et al.*, 1978). This work further demonstrates that the index of refraction has very little effect on the size distribution determined from spectral attenuation measurements (at least for $1.45 \leqslant n \leqslant 1.54$ and $0.00 \leqslant \kappa \leqslant 0.03$), thus allowing a determination of $n_c(r)$ essentially independent of κ (and

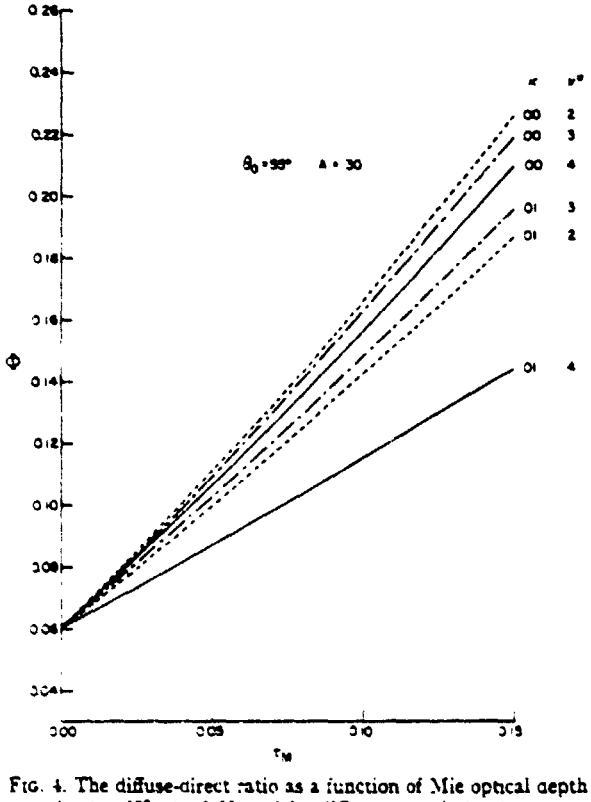


FIG. 4. The diffuse-direct ratio as a function of Mie optical depth for $\theta_0 = 55°$, $A = 0.30$, and for different κ and ν^* values.

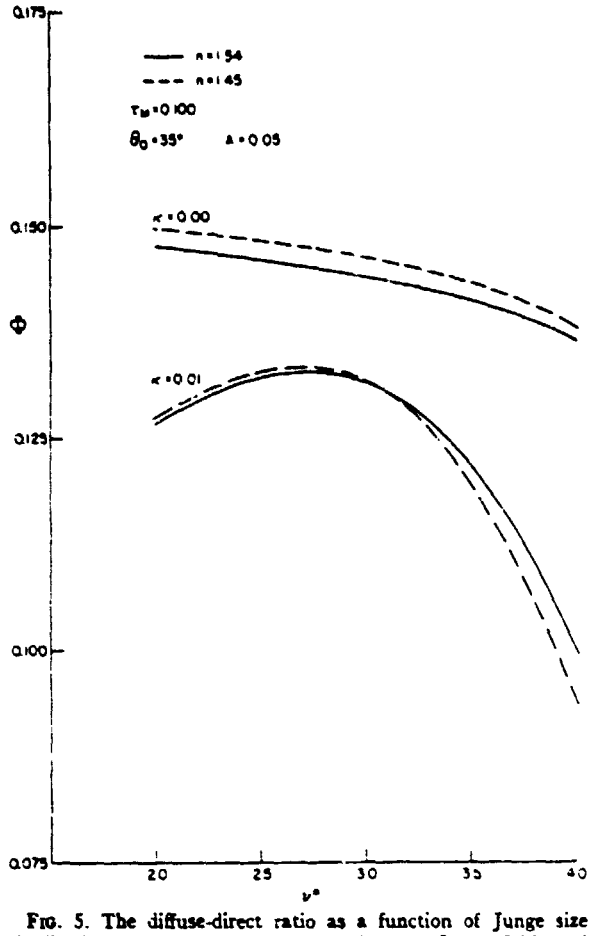


FIG. 5. The diffuse-direct ratio as a function of Junge size distribution parameter ν^* for $\tau_M = 0.10$, $\theta_0 = 35°$, $A = 0.05$, and four values of the complex refractive index.

of course A). By making use of the optical depth and size distribution of the atmospheric particulates thus obtained, it is possible to make explicit radiative transfer calculations for a particular day at a given wavelength similar to the set presented in Figs. 2 and 3. In this manner errors in the determination of κ and A are minimized since the need to interpolate in wavelength, size distribution and Mie optical depth can thus be eliminated.

The sensitivity of the diffuse-direct ratio to Junge size distribution parameter ν^* is also illustrated in Fig. 4 where it is apparent that the effect which a changing Junge parameter has on Φ depends upon the value of κ. For $\kappa = 0.00$, Φ decreases monotonically with ν^* regardless of the Mie optical depth while for larger values of the index of absorption Φ is nonmonotonic in ν^*, having the largest value for $\nu^* \approx 3.0$. Fig. 5 explicitly illustrates Φ as a function of ν^* for four values of the complex refractive index. In addition to the sensitivities of the diffuse-direct ratio to ν^* just mentioned, the same sensitivities are evident for both real indices of refraction illustrated, viz., 1.45 and 1.54.

TABLE 1. Values of $\delta(\tau_t)$ and $i(\tau_t,\mu_0)$, at $\mu_0=0.819$ and 0.259 and at $\lambda=0.5550$ μm, for a variety of atmospheric models.*

τ_M	κ	ν*	$\delta(\tau_t)$	$i(\tau_t,0.819)$	$i(\tau_t,0.259)$
0.00			0.07478	0.04061	0.03619
0.05	0.000	2	0.08788	0.07852	0.06012
0.05	0.000	3	0.09137	0.07693	0.05794
0.05	0.000	4	0.09608	0.07380	0.05604
0.05	0.025	2	0.07693	0.06658	0.05190
0.05	0.025	3	0.08261	0.06729	0.05126
0.05	0.025	4	0.07408	0.05114	0.04053
0.10	0.000	2	0.09989	0.11384	0.07901
0.10	0.000	3	0.10643	0.11065	0.07514
0.10	0.000	4	0.11535	0.10452	0.07179
0.10	0.025	2	0.07863	0.09026	0.06327
0.10	0.025	3	0.08923	0.09152	0.06239
0.10	0.025	4	0.07357	0.06035	0.04313

* All calculations are for Junge distributions where the radius extends from 0.01 to 10.01 μm.

It is evident from this figure, as well as from computations made for other values of τ_M, θ_0 and A, that the effect of the real term of the complex refractive index on the magnitude of the diffuse-direct ratio is quite negligible compared to the much greater effect of absorption. According to Gebbie *et al.* (1951), Eiden (1966) and Hänel (1968), natural haze has a real refractive index lying somewhere between 1.33 and 1.57; the former value corresponds to that of water droplets and the latter value to several salts, sulfates and silicates arising from natural and man-made sources. In desert regions like Tucson, it is reasonable to expect n to lie in the region between 1.45 and 1.54 due to the sources of the particulates and the relatively low humidity of the atmosphere.

The main significance of Fig. 5 is to illustrate the relative insensitivity of the diffuse-direct ratio to the real, as opposed to imaginary, part of the complex refractive index of the suspended atmospheric particulates. From inversion of spectral optical depth data in order to infer the columnar aerosol size distribution, King *et al.* (1978) have suggested that scarcely 20% of all days over Tucson have aerosol size distributions which can be adequately described as Junge. Most days appear to have a combination of a Junge plus a log-normal type of distribution. Although the actual size distribution to be used in analyzing the diffuse-direct ratio measurements on a particular day is the one obtained by inverting the spectral Mie optical depth measurements, Fig. 5 nevertheless indicates that the diffuse-direct technique is largely insensitive to the real part of the complex refractive index, regardless of size distribution.

Since spectral optical depth measurements are capable of sensing only the columnar aerosol size distribution, a logical concern is the sensitivity of the hemispheric diffuse flux density to the vertical distribution of the particles. By assuming that the shape of the size distribution is height independent, Herman and Browning (1975) performed computations of the reflected flux density out of the top of an atmosphere having an Elterman (1968) height profile for the particulate concentration as well as two perturbed Elterman profiles simulating pollution conditions under an inversion layer. Their results indicate that the total Mie optical depth is the principal factor determining the diffuse flux density with little effect attributed to the vertical distribution. Similar computations for the transmitted diffuse flux density at the earth's surface indicate that errors less than 0.75% arise if $\tau_M=0.10$. The relative insensitivity of the diffuse direct ratio to the height profile of the aerosol particles is a fortunate result since obtaining that additional information would require another instrument such as a monostatic lidar [see Spinhirne (1977) for details].

3. Diffuse reflection and transmission

It is sometimes convenient to express the diffusely reflected light at the top of the atmosphere and the diffusely transmitted light at the bottom of the atmosphere in terms of a reflection function $S(\tau_t,\mu,\phi;\mu_0,\phi_0)$ and a transmission function $T(\tau_t;\mu,\phi;\mu_0,\phi_0)$. In terms of these functions, the reflected and transmitted intensities from a horizontally homogeneous atmosphere illuminated from above by a parallel beam of radiation of incident flux density F_0 may be expressed in the forms

$$I(0,+\mu,\phi)=\frac{F_0}{4\pi\mu}S(\tau_t;\mu,\phi;\mu_0,\phi_0), \tag{5}$$

$$I(\tau_t,-\mu,\phi)=\frac{F_0}{4\pi\mu}T(\tau_t;\mu,\phi;\mu_0,\phi_0). \tag{6}$$

Eqs. (5) and (6) can be viewed as defining the functions S and T for the case of zero ground reflectivity. The advantage of including the factor of $1/\mu$ in the expressions for the emerging intensities given above is that the reflection and transmission functions obey the Helmholtz principle of reciprocity as discussed in detail by Chandrasekhar (1960).

In order to examine the effect which a non-zero ground reflectivity has on the downward hemispheric diffuse flux density $F^-(\tau_t)$, and hence on Φ, it is necessary to consider an extension of (6) to include not only plane-parallel illumination from above the atmosphere but also diffuse radiation from below. The inward intensity at the level τ_t, for cases when the surface is assumed to reflect radiation according to Lambert's law with some albedo A, is given by (Chandrasekhar, 1960)

$$I(\tau_t,-\mu,\phi)=\frac{F_0}{4\pi\mu}\left[T(\tau_t;\mu,\phi;\mu_0,\phi_0)+\frac{AJ}{1-AS(\tau_t)}\mu\mu_0\gamma_1(\tau_t,\mu_0)\frac{s(\tau_t,\mu)}{\mu}\right]. \tag{7}$$

In this expression,

$$s(\tau_t,\mu)=\frac{1}{4\pi}\int_0^{2\pi}\int_0^1 S(\tau_t;\mu,\phi;\mu',\phi')d\mu'd\phi',$$

$$t(\tau_t,\mu)=\frac{1}{4\pi}\int_0^{2\pi}\int_0^1 T(\tau_t;\mu,\phi;\mu',\phi')d\mu'd\phi',$$

$$\bar{s}(\tau_t)=2\int_0^1 s(\tau_t,\mu)d\mu,$$

$$\gamma_1(\tau_t,\mu)=e^{-\tau_t/\mu}+\frac{t(\tau_t,\mu)}{\mu}. \tag{8}$$

In order to determine Φ, it is necessary to evaluate the downward hemispheric diffuse flux density $F^-(\tau_t)$ for application of (4). From definitions, it follows that

$$F^-(\tau_t)=\int_0^{2\pi}\int_0^1 I(\tau_t,-\mu,\phi)\mu d\mu d\phi$$

$$=F_0\left[t(\tau_t,\mu_0)+\frac{A}{1-A\bar{s}(\tau_t)}\mu_0\gamma_1(\tau_t,\mu_0)\bar{s}(\tau_t)\right]. \tag{9}$$

Substituting (9) into (4) and recalling the definition of F, it follows that

$$\Phi=e^{\tau_t/\mu_0}\left[t(\tau_t,\mu_0)+\frac{A\bar{s}(\tau_t)}{1-A\bar{s}(\tau_t)}\mu_0\gamma_1(\tau_t,\mu_0)\right], \tag{10}$$

which, upon rearranging the order of terms, can be shown to yield

$$\Phi=\frac{t(\tau_t,\mu_0)e^{\tau_t/\mu_0}+A\mu_0\bar{s}(\tau_t)}{1-A\bar{s}(\tau_t)}. \tag{11}$$

This expression for the diffuse-direct ratio gives its dependence explicitly upon ground albedo, solar zenith angle and optical depth in terms of the reflection and transmission functions for zero ground reflectivity. The dependence of Φ on the index of refraction and size distribution of the atmospheric particulates is, of course, through the functions $t(\tau_t,\mu_0)$ and $\bar{s}(\tau_t)$. These functions may be determined for specific values of κ and $n_e(r)$ through solutions of the transfer equation for zero ground reflectivity. The diffuse-direct ratio may then be determined for any ground albedo through application of (11).

For a combination of optical depth and ground albedo such that $A\bar{s}(\tau_t)\ll 1$, it follows from (11) that Φ is very nearly a linear function of A with an intercept of $t(\tau_t,\mu_0)e^{\tau_t/\mu_0}$ and a slope of $\mu_0\bar{s}(\tau_t)$. This result will be employed in the next section so that the lengthy calculations required to determine $\bar{s}(\tau_t)$ may be avoided. The function $\bar{s}(\tau_t)$ physically represents the value of the spherical albedo, i.e., the ratio of the total radiation reflected from a spherical atmosphere to the total radiation incident from a distant source when the surface albedo $A=0.00$. For small optical depths $\bar{s}(\tau_t)$ is proportional to τ_t where the proportionality constant is a function of the single scattering phase function. Values of $\bar{s}(\tau_t)$ and $t(\tau_t,\mu_0)$, for $\mu_0=0.819$ ($\theta_0=35°$) and $\mu_0=0.259$ ($\theta_0=75°$) and for $\lambda=0.5550$ μm, are presented in Table 1 for a variety of atmospheric models. When $\kappa=0.00$, $\bar{s}(\tau_t)$ increases monotonically with τ_t to a limiting value of 1.0 when $\tau_t\to\infty$. Since the slope of Φ as a function of A is approximately equal to $\mu_0\bar{s}(\tau_t)$ in the clear atmosphere, it follows that the slope decreases with increasing zenith angle (due to μ_0) and increases with increasing Mie optical depth [due to $\bar{s}(\tau_t)$]. The first of these sensitivities can readily be seen upon examination of Figs. 2 and 3.

For the case in which polarization is included, as it has been in the radiative transfer computations presented above, the principle of reciprocity is complicated somewhat, and the reflection and transmission functions become four-vectors. Chandrasekhar (1960) derives a relationship analogous to (7) for the case of scattering according to a phase matrix. The general conclusions and functional relationships derived in this section remain essentially unaltered, however, so nothing is gained by the increased complexity of the mathematical formulation.

4. Determination of ground albedo and index of absorption

The reflection and transmission functions in a planetary atmosphere depend on the size distribution and index of refraction of the atmospheric particulates as well as the optical depth and solar zenith angle explicitly indicated (see Table 1). For the cases in which the Mie optical depth and aerosol size distribution can be determined, the functions $t(\tau_t,\mu_0)$ and $\bar{s}(\tau_t)$ depend primarily on solar zenith angle and index of absorption. From (11) it therefore follows that the functional form of the diffuse-direct ratio may be written as

$$\Phi(\theta_0,\kappa,A)=\frac{a(\theta_0,\kappa)+b(\theta_0,\kappa)A}{1-c(\kappa)A}, \tag{12}$$

where

$$a(\theta_0,\kappa)=t(\tau_t,\mu_0)e^{\tau_t/\mu_0}, \tag{13}$$

$$b(\theta_0,\kappa)=\mu_0\bar{s}(\tau_t), \tag{14}$$

$$c(\kappa)=\bar{s}(\tau_t). \tag{15}$$

Since $c(\kappa)A\ll 1$ for Mie optical depths at least as large as $\tau_M=0.30$, a point at which $c(\kappa)=\bar{s}(\tau_t)\lesssim 0.14$, the observations of Φ may be fitted to a straight line of the form

$$\Phi(\theta_0,\kappa,A)=a(\theta_0,\kappa)+b(\theta_0,\kappa)A \tag{16}$$

in order to determine the index of absorption κ and ground albedo A.

If one assumes that each measurement is made with

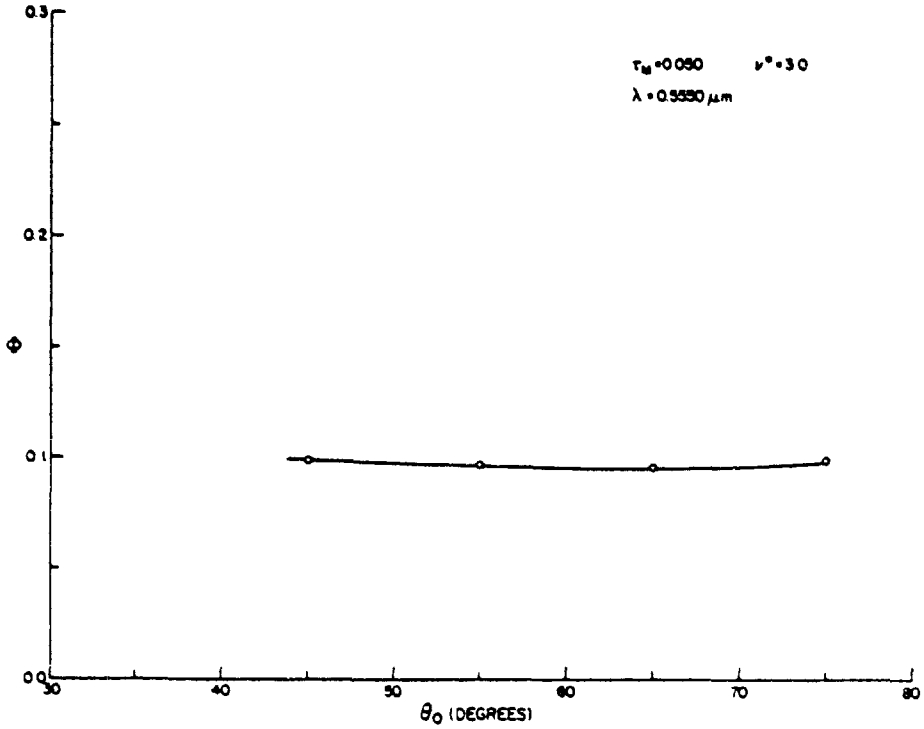


FIG. 6. The diffuse-direct ratio versus solar zenith angle for a model atmosphere having $\tau_M=0.05$ and $\nu^*=3.0$.

the same precision, Bevington (1969) shows that maximizing the probability that Φ_i observations have the functional form $\Phi(\theta_i,\kappa,A)$ is equivalent to minimizing the statistic χ^2 defined as

$$\chi^2=\sum_i[\Phi_i-\Phi(\theta_i,\kappa,A)]^2$$
$$=\sum_i[\Phi_i-a(\theta_i,\kappa)-b(\theta_i,\kappa)A]^2, \quad (17)$$

where the summation extends over all solar zenith angles θ_i for which measurements have been made and calculations performed.

Minimizing χ^2 as defined by (17) is equivalent to making an unweighted least-squares fit to the data. The minimum value of χ^2 can be determined by setting the partial derivative of χ^2 with respect to each of the coefficients (κ,A) equal to zero. This procedure results in two simultaneous equations

$$\frac{\partial\chi^2}{\partial A}=-2\sum_i\epsilon_i b(\theta_i,\kappa)=0, \quad (18)$$

$$\frac{\partial\chi^2}{\partial\kappa}=-2\sum_i\epsilon_i\left[\frac{\partial a(\theta_i,\kappa)}{\partial\kappa}+A\frac{\partial b(\theta_i,\kappa)}{\partial\kappa}\right]=0, \quad (19)$$

where

$$\epsilon_i=\Phi_i-a(\theta_i,\kappa)-b(\theta_i,\kappa)A. \quad (20)$$

Due to the complicated dependence of the functions a and b on κ, the set of Eqs. (18) and (19) is nonlinear in the unknowns κ and A. As a consequence, no analytic solution for the coefficients exists.

By combining (20) with (18), it immediately follows that

$$A=\sum_i b(\theta_i,\kappa)[\Phi_i-a(\theta_i,\kappa)]/\sum_i b^2(\theta_i,\kappa). \quad (21)$$

For any arbitrary value of κ, however, χ^2 will not necessarily be a minimum since $\partial\chi^2/\partial\kappa$ will not be identically zero as required by (19). As κ is varied, the functions $a(\theta_i,\kappa)$ and $b(\theta_i,\kappa)$ are determined from the theoretical set of computations, from which A and χ^2 can be computed using (21) and (17), respectively. The coefficient κ is continuously varied until a minimum value of χ^2 is determined, always with the knowledge that the value of A computed using (21) is that value of A giving the minimum value of χ^2 for a given κ.

In order to see how this procedure works, consider the data set of Fig. 6 which pertains to a model atmosphere having a Junge distribution of the aerosol particles with $\nu^*=3.0$. All "measurements" of the diffuse-direct ratio are for $\theta_0=45^\circ(10^\circ)75^\circ$ and $\lambda=0.5550$ μm, a wavelength for which $\tau_M=0.050$ and $\tau_R=0.086$ in this example. Since the theoretical $\Phi(\theta_0,\kappa,A)$ function can be well approximated by a linear function of A, radiative transfer computations are required for only two ground albedos at any given index of absorption and solar zenith angle. From these two computations, the coefficients $a(\theta_0,\kappa)$ and $b(\theta_0,\kappa)$ may readily be determined through application of (16). Due to the excessive

computer time required to generate a set of calculations $\Phi(\theta_0,\kappa,A)$ for any given value of κ and θ_0, it is desirable to be able to interpolate in κ for fixed values of θ_0 and A and in θ_0 for fixed values of κ and A.

Fig. 2. illustrates a data set of $\Phi(\theta_0,\kappa,A)$ for values of $\theta_0=35°(10°)75°$, $\tau_M=0.10$ and $\nu^*=2.0$. Although curves for four values of A are illustrated, it is now obvious that this is redundant since computations at two values of A are sufficient to predict all other values. It is also apparent upon examination of this figure that $\Phi(\theta_0,\kappa,A)$ is a smooth, monotonically decreasing function of κ for fixed values of θ_0 and A, a feature readily lending itself to interpolation. For this purpose a spline under tension interpolation has been adopted [see Cline (1974) for details] and thus computations of $\Phi(\theta_0,\kappa,A)$ are required for a finite number of κ values only (usually four). It is similarly necessary to interpolate a set of calculations $\Phi(\theta_0,\kappa,A)$ in θ_0 because measurements are normally made at solar zenith angles θ, other than those for which calculations are performed. Since $\Phi(\theta_0,\kappa,A)$ is a smooth function of θ_0 for fixed values of κ and A, as illustrated in Fig. 3, spline under tension interpolation has been used where computations are made for a finite number of θ_0 values.

After computing the diffuse-direct ratio at four values of κ, two values of A, and every 10° in θ_0 for a range of zenith angles surrounding the measurements (at least four values of θ_0), it is reasonably simple to compute χ^2 at any value of κ and A, making use of spline under tension interpolation in κ and θ_0 as well as (17). Fig. 7 shows such a χ^2 hypersurface in coefficient space for the theoretically generated data of Fig. 6. The solid curves indicate the points of constant χ^2. The parameters κ and A which give the best fit of the measurements to the nonlinear function $\Phi(\theta_0,\kappa,A)$ are determined by the location of the minimum value of χ^2 in this two-dimensional space. Searching this hypersurface for the parameters which minimize χ^2 is greatly facilitated through the use of (21) since this expression gives analytically the value of A which minimizes χ^2 at any value of κ. The albedo values given by (21) are shown in Fig. 7 as a dashed line which must necessarily pass through the absolute minimum of the function χ^2.

By varying κ and following the magnitude of χ^2 along the path of lowest value of χ^2, the parameter κ may be found. The variation of χ^2 as a function of κ along the ravine (dashed line) of Fig. 7 is illustrated in Fig. 8. The value of κ at which χ^2 attains a minimum is found to be 0.0100, an index of absorption for which the corresponding value of A is 0.200 in this example. It may readily be seen upon examination of Fig. 7 that these values for the parameters κ and A correspond to the absolute minimum of the χ^2 hypersurface.

In any real experiment the absolute minimum value of χ^2 will not be identically zero as in this example due to various sources of error. After having determined the optimum values for the coefficients κ and A, a best fit set of $\Phi(\theta_i,\kappa,A)$ values can be computed using (10). The solid curve in Fig. 6 is the best fit to the data where $\kappa=0.0100$ and $A=0.200$. Although the effect of random errors in the Φ_i measurements was not investigated using theoretical computations, the procedure described above has been applied to measurements obtained at The University of Arizona during May and June 1977. The results of this experiment and a discussion of the various sources and magnitudes of error are presented by King (1979).

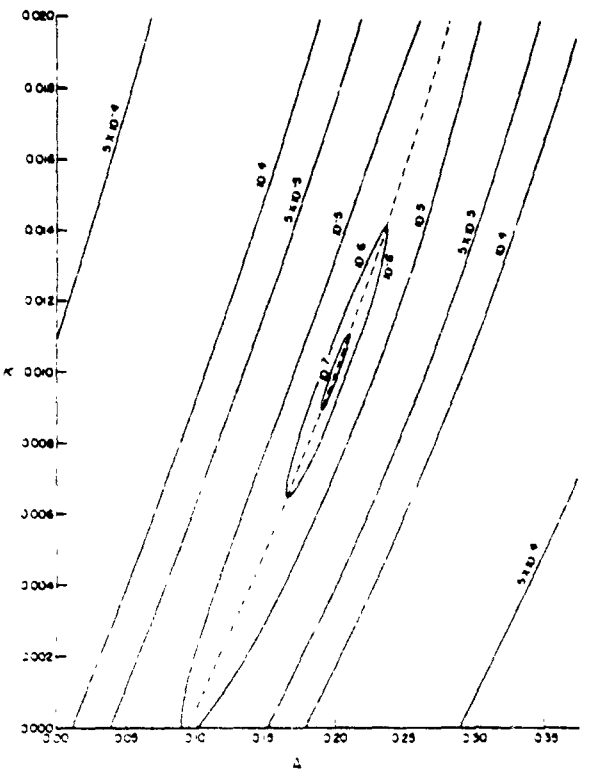


FIG. 7. χ^2 hypersurface for theoretically generated data of Fig. 6.

Computations similar to those used in the preceding example have been performed by assuming that the real part of the complex refractive index of the aerosol particles is 1.45, rather than 1.54. In this way, the theoretical set of computations $\Phi(\theta_0,\kappa,A)$ was altered while the "measurements" Φ_i remained the same. The results of the inversion algorithm for this situation indicate that χ^2 attains a minimum value when $\kappa=0.0099$ and $A=0.203$. This strongly supports the conclusion made previously that the diffuse-direct technique is sensitive primarily to the index of absorption and ground albedo, having substantially reduced sensitivity to the real part of the refractive index of atmospheric particulates.

Having determined the minimum value of χ^2 and the optimum values of the regression coefficients κ and A, Bevington (1969) shows that the uncertainties in

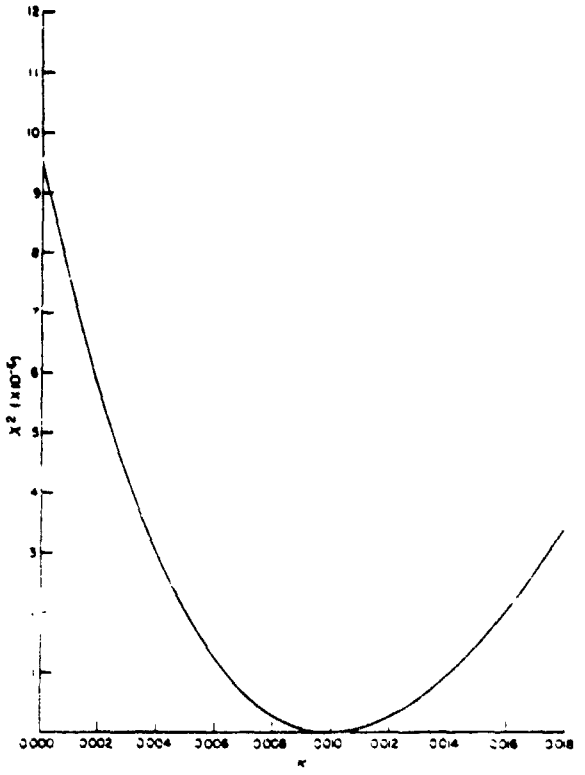


FIG. 8. Variation of χ^2 with κ along the gradient search path of Fig. 7 showing the pronounced minimum which occurs at $\kappa=0.01$.

these coefficients are related to the curvature matrix α, whose elements are given by

$$\alpha_{jk}=\frac{1}{2}\frac{\partial^2\chi^2}{\partial\xi_j\partial\xi_k}, \tag{22}$$

where ξ_j represents the coefficients κ and A. The covariances $\sigma^2_{\xi_j\xi_k}$ are then obtained from the (j,k) elements of the α^{-1} matrix through the expression

$$\sigma^2_{\xi_j\xi_k}\approx s^2[\alpha^{-1}]_{jk}, \tag{23}$$

where s^2 is the sample variance for the fit, given by $s^2=\chi^2/\nu$, and $\nu=N-2$ is the number of degrees of freedom after fitting N data points with 2 parameters. The diagonal elements of (23) thus represent the variances σ_A^2 and σ_κ^2. The elements of the curvature matrix and the method of computation for this problem are discussed in detail by King (1977).

The procedure described above for finding the coefficients κ and A which minimize the statistic χ^2 most nearly parallels the gradient search method from nonlinear least-squares theory. By incorporating the analytical expression for A given by (21), both parameters κ and A are incremented simultaneously such that the resultant direction of travel in parameter space is along the gradient of χ^2. This method of solving nonlinear least-squares problems which are nonlinear in only one coefficient has been developed by King and Byrne (1976) for obtaining the total atmospheric ozone.

5. Summary

A statistical technique has been developed for inferring optimum values of the ground albedo and imaginary index of refraction of atmospheric aerosol particles from surface measurements of the downward hemispheric diffuse and total flux densities. This method is based on a careful consideration of the sensitivities of the diffuse radiation field to the many radiative transfer parameters and represents an optimized extension of the diffuse-direct technique suggested by Herman *et al.* (1975). The statistical algorithm presented in this paper makes use of the laws of diffuse reflection and transmission for the planetary problem, i.e., the situation of an atmosphere resting on solid ground idealized as a surface which reflects light according to Lambert's law. As a consequence of this formulation, the ground albedo inferred by the diffuse-direct technique represents a weighted average of the albedo over the entire area which affects the transfer of radiation. A simple technique is given for estimating the standard deviations of the solution parameters κ (index of absorption) and A (surface albedo).

As illustrated by the family of curves in Figs. 2 and 3, the ratio of the diffuse to direct solar flux densities decreases with increasing absorption (κ), increases with surface reflectivity (A) and shows very little change with solar zenith angle (θ_0). Fig. 3 clearly indicates, however, that the shape of the diffuse-direct ratio as a function of solar zenith angle is the primary factor which determines the ground albedo while the magnitude determines the index of absorption. As a consequence of this sensitivity to surface albedo it is clearly desirable to obtain data on days during the spring and summer months in order to allow a large span of solar zenith angles to be examined. In Part II of this series (King, 1979) the results of such an experiment will be described.

A consideration of the sensitivities of the diffuse radiation field has led to the conclusion that several radiative transfer parameters such as the real part of the particle refractive index, total ozone content, vertical distribution of the atmospheric particulates and the law of reflection at the ground have very little effect on the diffuse-direct ratio at the ground. Koepke and Kriebel (1978) recently examined the effects of measured bidirectional reflectance characteristics of four vegetated surfaces on the radiation scattered from a realistic atmosphere containing molecules and particles. They concluded that the anisotropy of the reflection properties has negligible influence on the transmitted intensity field.

The principal assumptions upon which the technique is based are that the angular distribution of energy scattered by a polydispersion of randomly oriented

particles can be adequately described by Mie theory (applicable to spheres) and that the coefficients κ and A which best describe the radiation field at the earth's surface can be determined by minimizing the statistic χ^2. Although the suspended atmospheric particulates are somewhat irregular in shape, it is the purpose here to find a consistent set of parameters which are able to predict accurately the scattered radiation field in the earth's atmosphere. Since the assumption of sphericity is not totally correct, the index of absorption and ground albedo inferred by the diffuse-direct technique may not be representative of the optical properties of the individual particles or the reflectivity of the earth's surface. They nevertheless represent a consistent set of parameters which are able to predict accurately the scattered radiation field in the earth's atmosphere. Minimization of the statistic χ^2 is a result of the principle of maximum likelihood and forms the basis of least-squares curve fitting.

Acknowledgments. The authors are grateful to Dr. S. Twomey for many helpful suggestions and thought-provoking discussions which led to improvements in this experiment. Acknowledgment is made to the National Center for Atmospheric Research for the computing time used in the early stages of this research. The spline under tension interpolation algorithm made available by the Computing Facility was invaluable in this work.

REFERENCES

Ackerman, T., and M. B. Baker, 1977: Shortwave radiative effects of unactivated aerosol particles in clouds. *J. Appl. Meteor.*, 16, 63–69.

Bevington, P. R., 1969: *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, 336 pp.

Bryson, R. A., 1968: All other factors being constant...: A reconciliation of several theories of climate change. *Weatherwise*, 21, 56–61.

Chandrasekhar, S., 1960: *Radiative Transfer*. Dover, 393 pp.

Cline, A. K., 1974: Scalar- and planar-valued curve fitting using splines under tension. *Comm. Assoc. Comput. Mach.*, 17, 218–220.

Eiden, R., 1966: The elliptical polarization of light scattered by a volume of atmospheric air. *Appl. Opt.*, 5, 569-575.

Elterman, L., 1968. UV, visible, and IR attenuation for altitudes to 50 km, 1968. Rep. AFCRL-68-0153, Environ. Res. Pap. No. 285, Air Force Cambridge Research Laboratories, 49 pp.

Gebbie, H. A., W. R. Harding, C. Hilsum, A. W. Pryce and V. Roberts, 1951: Atmospheric transmission in the 1 to 14 μ region. *Proc. Roy. Soc. London*, A206, 87–107.

Hänel, G., 1968: The real part of the mean complex refractive index and the mean density of samples of atmospheric aerosol particles. *Tellus*, 20, 371–379.

Herman, B. M., and S. R. Browning, 1965: A numerical solution to the equation of radiative transfer. *J. Atmos. Sci.*, 22, 559–566.

——, and ——, 1975: The effect of aerosols on the earth-atmosphere albedo. *J. Atmos. Sci.*, 32, 1430-1445.

——, —— and R. J. Curran, 1971: The effect of atmospheric aerosols on scattered sunlight. *J. Atmos. Sci.*, 28, 419–428.

——, —— and J. J. DeLuisi, 1975: Determination of the effective imaginary term of the complex refractive index of atmospheric dust by remote sensing: The diffuse-direct radiation method. *J. Atmos. Sci.*, 32, 918–925.

Junge, C. E., 1955: The size distribution and aging of natural aerosols as determined from electrical and optical data on the atmosphere. *J. Meteor.*, 12, 13–25.

King, M. D., 1977: Determination of the complex refractive index of atmospheric aerosols by the diffuse-direct technique: A statistical procedure. Ph.D. dissertation, University of Arizona, 147 pp.

——, 1979: Determination of the ground albedo and the index of absorption of atmospheric particulates by remote sensing. Part II: Application. *J. Atmos. Sci.*, 36 (in press).

——, and D. M. Byrne, 1976: A method for inferring the total ozone content from the spectral variation of total optical depth obtained with a solar radiometer. *J. Atmos. Sci.*, 33, 2242–2251.

——, ——, B. M. Herman and J. A. Reagan, 1978: Aerosol size distributions obtained by inversion of spectral optical depth measurements. *J. Atmos. Sci.*, 35, 2153–2167.

Koepke, P., and K. T. Kriebel, 1978: Influence of measured reflection properties of vegetated surfaces on atmospheric radiance and its polarization. *Appl. Opt.*, 17, 260–264.

Rasool, S. I., and S. H. Schneider, 1971: Atmospheric carbon dioxide and aerosols: Effects of large increases on global climate. *Science*, 173, 138–141.

Spinhirne, J. D., 1977: Monitoring of tropospheric aerosol optical properties by laser radar. Ph.D. dissertation, University of Arizona, 176 pp.

Twomey, S., 1974: Pollution and the planetary albedo. *Atmos. Environ.*, 8, 1251–1256.

——, 1977. The influence of pollution on the shortwave albedo of clouds. *J. Atmos. Sci.*, 34, 1149–1152.

van de Hulst, H. C., 1957: *Light Scattering by Small Particles*. Wiley, 470 pp.

Wang, W. C., and G. A. Domoto, 1974. The radiative effect of aerosols in the earth's atmosphere. *J. Appl. Meteor.*, 13, 521–534.

Yamamoto, G., and M. Tanaka, 1972: Increase of global albedo due to air pollution. *J. Atmos. Sci.*, 29, 1405–1412.

Paper 123

Reprinted with permission from *Journal of Quantitative Spectroscopy and Radiative Transfer*, 22, pp. 483-488, copyright 1979, Pergamon Press, Ltd.

TROPOSPHERIC AEROSOLS: EFFECTS UPON THE SURFACE AND SURFACE-ATMOSPHERE RADIATION BUDGETS

L. A. KONYUKH and F. B. YUREVICH

Luikov Heat and Mass Transfer Institute, Byelorussian Academy of Sciences, Minsk, 220728, U.S.S.R.

R. D. CESS

Laboratory for Planetary Atmospheres Research, State University of New York at Stony Brook, Long Island, NY 11794, U.S.A.

and

HARSHVARDHAN

Meteorology Program, University of Maryland, College Park, MD 20742, U.S.A.

(*Received* 11 *May* 1979)

Abstract—Employing a global atmospheric radiation model which incorporates tropospheric aerosols, it is suggested that: (1) tropospheric aerosols exert a significantly different influence upon the surface radiation budget as opposed to the planetary radiation balance, and (2) that aerosol-induced effects upon the surface radiation budget are strongly dependent upon solar zenith angle. These results indicate that extreme caution should be exercised when interpreting surface radiation measurements for localized dust events in terms of aerosol-induced influences upon global climate.

1. INTRODUCTION

There has been recent concern that an increase in tropospheric aerosols, due to man's activities, could alter the radiation balance of the earth-atmosphere system, leading to global climatic change. Although our knowledge of the radiation properties of such aerosols, as well as the manner in which they interact with the climate system, is meager, most model results suggest that an increase in tropospheric aerosols would lead to global cooling (e.g. Reck,[1] Wang and Domoto,[2] Lee and Snell[3]). Bryson and Dittberner,[4] in fact, have suggested that global cooling by anthropogenic aerosols might offset the predicted global warming resulting from increased levels of carbon dioxide.

Recently, however, Idso and Brazel[5] have concluded that increasing tropospheric particulates should lead to global heating and thus complement the CO_2-climate problem. They base this conclusion upon studies of several atmospheric dust-loading events at locations within Arizona, during which dust-induced changes in both the net solar radiation and the net i.r. radiation were measured at the ground. These measurements indicated that, for initial increases in tropospheric dust concentration, there was a greater reduction in net surface i.r. radiation than in net surface solar radiation. Extrapolating this finding to a planetary radiation balance thus led Idso and Brazel to conclude that the dominant influence due to increased atmospheric particulates was an enhanced i.r. atmospheric opacity, which would lead to global warming.

But changes in net surface radiation do not necessarily reflect comparable changes in the components of the planetary radiation balance, since the surface energy balance additionally contains a convective component. In that the presence of aerosols will alter the static stability of the troposphere, then the convective component should by itself be dependent upon tropospheric aerosol concentration. A further discussion of surface vs planetary radiation budgets, with reference to the CO_2-climate problem, is given by Manabe and Wetherald.[6] An additional difficulty lies in the fact that the observed dust-loading events occurred at different times of the day and during different months; it does not appear that the solar zenith angle for the individual dust events correctly simulated an appropriate global mean value. There are, of course, numerous other caveats concerning local surface vs planetary radiation budgets, but the above two appear to be the most worrisome.

In view of the above, it would seem useful to perform a model study to illustrate: (1) possible differences between surface and planetary radiation budgets, and (2) the influence of

solar zenith angle upon the surface radiation budget. These two points comprise the objectives of the present paper.

2. RADIATION BUDGETS

An increase in tropospheric aerosols will produce two effects upon the earth's radiation balance which, in turn, could lead to an alteration of the global climate. One effect pertains to a change in solar absorption by the surface-atmosphere system, while the other concerns a change in the outgoing i.r. flux at the top of the atmosphere. To illustrate how these separate effects can influence global climate, let Q_0 denote solar radiation absorbed by the surface-atmosphere system, while F_0 is the corresponding outgoing i.r. flux. Furthermore, it will be assumed that $Q_0 = Q_0(\tau_{vis}, T_s)$ and $F_0 = F_0(\tau_{vis}, T_s)$, where T_s is the global surface temperature while τ_{vis} denotes the aerosol optical depth at the wavelength $\lambda = 0.55\ \mu$m. In the following, we will employ τ_{vis} as a measure of the tropospheric aerosol concentration; the dependence of Q_0 and F_0 upon τ_{vis} incorporates the two previously discussed influences upon the planetary radiation balance. The dependences upon T_s additionally incorporate temperature-dependent climate feedback mechanisms.

Since a planetary radiation balance requires that $Q_0 = F_0$, it readily follows that the response of global surface temperature to a change in aerosol concentration can be expressed by

$$\frac{dT_s}{d\tau_{vis}} = \frac{\partial Q_0/\partial \tau_{vis} - \partial F_0/\partial \tau_{vis}}{\partial F_0/\partial T_s - \partial Q_0/\partial T_s}. \tag{1}$$

Stability of the climate system requires that

$$\partial F_0/\partial T_s - \partial Q_0/\partial T_s > 0,$$

as can be determined, for example, from the stability analyses of Frederiksen[7] and Su and Hsieh.[8] Thus, the sign of the quantity

$$\delta_0 = \partial Q_0/\partial \tau_{vis} - \partial F_0/\partial \tau_{vis} \tag{2}$$

determines whether an increase in tropospheric aerosols will lead to global cooling ($\delta_0 < 0$) or to global heating ($\delta_0 > 0$). This approach is analogous to Schneider's[9] treatment of cloud amount as a climate change mechanism.

In their observational study of dust events, Idso and Brazel[5] have empirically suggested that a modest increase in tropospheric aerosols should lead to global heating. But they arrive at this conclusion by considering the sign of a quantity which may be quite different than δ_0. Their measurements refer to a surface radiation budget, such that they are dealing with the surface counterpart to Eq. (2)

$$\delta_s = \partial Q_s/\partial \tau_{vis} - \partial F_s/\partial \tau_{vis}, \tag{3}$$

where Q_s is the solar radiation absorbed by the surface and F_s is the net i.r. radiation at the surface. While the sign of δ_s is indicative of aerosol-induced changes in the surface radiation budget, this *does not* relate to changes in the surface energy balance, since this balance is given by

$$Q_s = F_s + H + LE, \tag{4}$$

where H and LE are, respectively, the sensible and latent heat transports from the surface. One would anticipate that these quantities would also be influenced by a change in τ_{vis}.

For the purpose of illustrating, by means of a model calculation, possible differences between δ_0 and δ_s, we employ a particularly simple tropospheric aerosol model, that due to Peterson and Weinman.[10] This assumes an aerosol composed of quartz particles having the size distribution

$$\frac{dN(r)}{dr} = \begin{cases} C, & r \le r_0; \\ C(r_0/r)^{3.5}, & r \ge r_0. \end{cases} \tag{5}$$

here $N(r)$ is the particle size density, r is the particle radius, $r_0 = 0.1\ \mu m$, and C is arbitrary. The vertical distribution of aerosols is taken to be

$$N(z) = N(0) \exp(-z/H_a), \tag{6}$$

where $N(z)$ here refers to the total concentration in number of particles per unit volume at altitude z, with H_a the aerosol scale height.

The motivation for employing this particular aerosol model is that it is the same one previously employed by Harshvardhan and Cess[11] within a global atmospheric i.r. radiation model, from which we find that

$$\frac{\partial F_0}{\partial \tau_{vis}} = \begin{cases} -2.5\ \mathrm{W\,m^{-2}}, & H_a = 1\ \mathrm{km}; \\ -8.9\ \mathrm{W\,m^{-2}}, & H_a = 2\ \mathrm{km}; \end{cases} \tag{7}$$

$$\frac{\partial F_t}{\partial \tau_{vis}} = \begin{cases} -21.0\ \mathrm{W\,m^{-2}}, & H_a = 1\ \mathrm{km}; \\ -17.5\ \mathrm{W\,m^{-2}}, & H_a = 2\ \mathrm{km}. \end{cases} \tag{8}$$

These results are for cloud-free conditions, which is consistent with the study by Idso and Brazel.[5] The range for H_a is further consistent with $H_a = 1.2$ km, as suggested by Peterman.[12]

Note in particular the substantial differences between $\partial F_0/\partial \tau_{vis}$ and $\partial F_t/\partial \tau_{vis}$. This, in order to compare model estimates of δ_0 and δ_t, results for $\partial Q_0/\partial \tau_{vis}$ and $\partial Q_t/\partial \tau_{vis}$ are required, and this is accomplished in the following section. Again restriction will be made solely to cloud-free conditions.

3. SOLAR RADIATION

In that the imaginary refractive index of quartz is of $O(10^{-7})$ for solar wavelengths, the present aerosol model absorbs essentially no solar radiation. Furthermore, when integrating over the solar spectrum, we will assume no aerosol effects for $\lambda < 0.3\ \mu m$ and $\lambda > 0.9\ \mu m$, since ozone absorbs virtually all incident solar radiation for $\lambda < 0.3\ \mu m$, whereas for $\lambda > 0.9\ \mu m$ there is significant water vapor absorption (Joseph[13]). Thus, since the model aerosol neither absorbs nor interacts with gaseous absorbers, then $\partial Q_0/\partial \tau_{vis} = \partial Q_t/\partial \tau_{vis}$, because the reductions in surface-atmosphere absorption and surface absorption, due to aerosol scattering, are the same. It further follows that

$$\partial Q_0/\partial \tau_{vis} = -(S_0/4) \int_{0.3\,\mu m}^{0.9\,\mu m} (d\alpha_\lambda/d\tau_{vis})(S_\lambda/S_0)\, d\lambda, \tag{9}$$

where $S_0 = 1360\ \mathrm{W\,m^{-2}}$ is the solar constant, S_λ is its spectral counterpart, and α_λ is the spectral albedo of the earth-atmosphere system.

The evaluation of $\partial Q_0/\partial \tau_{vis}$, and correspondingly $\partial Q_t/\partial \tau_{vis}$, thus reduces to evaluating $d\alpha_\lambda/d\tau_{vis}$. To accomplish this, α_λ may be expressed as

$$\alpha_\lambda = 1 - \frac{(1 - R_\lambda)(1 - \alpha_s)}{1 - \alpha_s R_\lambda}, \tag{10}$$

where α_s is the surface albedo while R_λ denotes the spectral reflectivity of the atmosphere. Furthermore, letting $R_{R\lambda}$ and $R_{a\lambda}$ represent, respectively, the Rayleigh and aerosol contributions to R_λ, then

$$R_\lambda = 1 - \frac{(1 - R_{R\lambda})(1 - R_{a\lambda})}{1 - R_{R\lambda} R_{a\lambda}} \tag{11}$$

Since the aerosol is nonabsorbing, the fact that the molecular and aerosol scatters have different vertical distributions produces no effect upon Eq. (11). The Rayleigh reflectivity, $R_{R\lambda}$, applicable to global mean conditions, has been evaluated employing the scattering formulation due to Wang,[14] together with the Rayleigh optical depths of Penndorf.[15]

To evaluate $R_{a\lambda}$, we employ the global mean expression applicable for the present case in

which the single scattering albedo is unity (Wiscombe and Grams[16]),

$$R_{s\lambda} = 2b_\lambda \tau_\lambda, \tag{12}$$

where τ_λ is the aerosol optical depth and b_λ is the backscattered fraction, defined as (Wiscombe and Grams[16])

$$b_\lambda = \frac{1}{2\pi}\int_0^\pi P_\lambda(\theta)\theta \sin\theta \, d\theta$$

with $P_\lambda(\theta)$ denoting the scattering phase function. Thus, from Eqs. (11) and (12) with restriction to $R_{s\lambda} \ll 1$ as is the case with Eq. (12), it follows that

$$dR_\lambda/d\tau_{vis} = 2b_\lambda(1 - R_{R\lambda})^2(\sigma_\lambda/\sigma_{vis}), \tag{13}$$

where σ_λ is the scattering coefficient. Equations (9), (10), and (13) thus describe $\partial Q_0/\partial\tau_{vis}$ and, correspondingly, $\partial Q_s/\partial\tau_{vis}$, for $R_{s\lambda} \ll 1$.

Both $\sigma_\lambda/\sigma_{vis}$ and b_λ have been evaluated employing Mie scattering theory, the numerical calculations being performed on the BESM-6 computer at the B.S.S.R. Academy of Sciences. Figure 1 illustrates the wavelength dependence of $\sigma_\lambda/\sigma_{vis}$. For these wavelengths, b_λ is essentially constant, with

$$b_\lambda = 0.27. \tag{14}$$

The present tropospheric aerosol model has been employed to assure consistency with the corresponding i.r. calculations. More detailed calculations, employing a slightly different aerosol model, have been performed by Braslau and Dave[17] for solar wavelengths only, and the present results are quite similar to theirs. To illustrate this, Fig. 3 of Wiscombe and Grams[16] suggests that, for b_λ given by Eq. (14), the appropriate mean solar zenith angle is $\approx 65°$. For this value, together with $\alpha_s = 0.1$, we obtain from the Braslau–Dave results

$$\partial Q_0/\partial\tau_{vis} = -74\ \mathrm{W\,m^{-2}}, \quad \partial Q_s/\partial\tau_{vis} = -82\ \mathrm{W\,m^{-2}}.$$

The difference between the two quantities is due to aerosol interaction with solar absorption by atmospheric gases, but this effect is small. From the present study, the comparable result

$$\partial Q_0/\partial\tau_{vis} = \partial Q_s/\partial\tau_{vis} = -75\ \mathrm{W\,m^{-2}}$$

is obtained.

4. RESULTS AND DISCUSSION

Values for both δ_0 and δ_s, as defined, respectively, by Eqs. (2) and (3), are summarized in

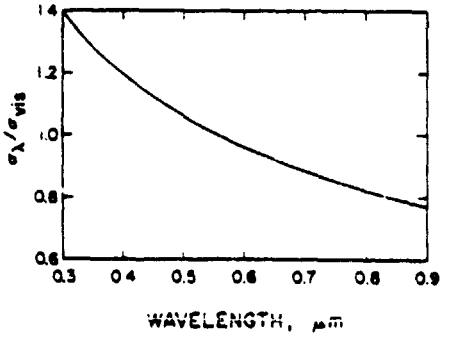


Fig. 1.

Table 1.

α_s	0.15	0.20	0.25
$\delta_0(H_a = 1$ km)	-65.1	-58.0	-51.3
$\delta_s(H_a = 1$ km)	-46.3	-39.5	-32.8
$\delta_0(H_a = 2$ km)	-58.7	-51.6	-44.9
$\delta_s(H_a = 2$ km)	-50.1	-43.0	-36.3

Table 1; the differences between δ_0 and δ_s are due, within the present model, to the differences between $\partial F_0/\partial\tau_{va}$ and $\partial F_s/\partial\tau_{va}$ [see Eqs. (7) and (8)]. As can be seen from this table, there is significant diversity between δ_0 and δ_s as well as dependence upon surface albedo α_s. A recent estimate suggests that $\alpha_s = 0.15$ on a global average (Hummel and Reck[18]), while Idso[19] indicates a range of values from $\alpha_s = 0.14$ to 0.30 as appropriate to local surfaces in areas where dust-loading events have been observed. Thus, it would appear that, on the average, the local albedos tend to exceed the global value, and from Table 1 this effect could tend to amplify the difference between a locally observed δ_s and the applicable global δ_0. We certainly do not contend that the δ_0 and δ_s values of Table 1 are by themselves indicative of real tropospheric aerosols; we employ Table 1 solely to illustrate possible relative differences between a local dust loading δ_s as opposed to the planetary quantity, δ_0.

As discussed in the Introduction, a second problem associated with observations of localized dust events, in contrast to global average conditions, concerns the fact that different dust events will, due to variations in the times of their occurrence, correspond to different local solar zenith angles. To illustrate the zenith-angle dependence of δ_s, we have employed the results of Braslau and Dave[17] to estimate $\partial Q_s/\partial\tau_{va}$ as a function of solar zenith angle, recalling the equivalence of their aerosol model with ours. These results are summarized in Table 2 for $\alpha_s = 0.20$, from which δ_s is found to undergo a sign change with decreasing solar zenith angle.

The primary point concerning Tables 1 and 2 is that localized dust events produce aerosol-induced changes in solar and i.r. radiation at the surface which could differ substantially from global planetary changes. For example, if a localized dust event corresponds to $\alpha_s = 0.20$ and a zenith angle of 30°, then for $H_a = 1$ km, Table 2 gives $\delta_s = 3.9\ \mathrm{W\,m^{-2}}$, which differs in both sign and magnitude from the corresponding planetary result of Table 1, $\delta_0 = -65.1\ \mathrm{W\,m^{-2}}$. From this, the sole conclusion of the present model study is simply that aerosol-induced changes in solar and i.r. radiation, applicable to a planetary radiation balance, cannot be realistically inferred from measurements which pertain to changes, during a localized dust-loading event, in the surface radiation budget.

Table 2.

Solar zenith angle	0°	30°	60°
$\delta_s(H_a = 1$ km)	10.1	3.9	-34.8
$\delta_s(H_a = 2$ km)	6.6	0.4	-38.3

Acknowledgements—We wish to thank Academician R. I. Soloukhin, of the B.S.S.R. Academy of Sciences, for his interest and assistance in making this cooperative research effort possible. We are also indebted to Mrs. I. I. Philipovich for her programming aid. One of us (R.D.C.) appreciates the hospitality and cooperation shown him during several visits to the B.S.S.R. Academy of Sciences. This work was jointly supported by the Byelorussian Academy of Sciences and the United States National Science Foundation, with the latter support stemming from NSF Grant No. ENG 77-09124

REFERENCES

1. R. A. Peck, *Atmos. Environ.* 8, 823 (1974).
2. W. C. Wang and G. A. Domoto, *J. Appl. Meteor.* 13, 521 (1974).
3. P. S. Lee and F. M. Snell, *J. Atmos. Sci.* 34, 847 (1977).
4. R. A. Bryson and G. J. Dittberner, *J. Atmos. Sci.* 33, 2094 (1976).

5. S. B. Idso and A. J. Brazel, *Science* 198, 731 (1977).
6. S. Manabe and R. T. Wetherald, *J. Atmos. Sci.* 24, 241 (1967).
7. J. S. Frederiksen, *J. Atmos. Sci.* 33, 2267 (1976).
8. C. H. Su and D. Y. Hsieh, *J. Atmos. Sci.* 33, 2273 (1976).
9. S. H. Schneider, *J. Atmos. Sci.* 29, 1413 (1972).
10. J. T. Peterson and J. A. Weinman, *J. Geophys. Res.* 74, 6974 (1969).
11. Harshvardhan and R. D. Cess, *JQSRT* 19, 621 (1978).
12. L. Elterman, *Appl. Opt.* 3, 1139 (1964).
13. J. H. Joseph, *Solar Energy* 13, 251 (1971).
14. L. S. Wang, *Astrophys. J.* 172, 671 (1972).
15. R. Penndorf, *J. Opt. Soc. Am.* 47, 176 (1957).
16. W. J. Wiscombe and G. W. Grams, *J. Atmos. Sci.* 33, 2440 (1976).
17. N. Braslau and J. V. Dave, *J. Appl. Meteor.* 12, 601 (1973).
18. J. R. Hummel and R. A. Reck, *J. Appl. Meteor.* in press.
19. S. B. Idso, *Atmos. Environ.* 9, 766 (1975).

Paper 124

Reprinted from APPLIED OPTICS, Vol. 18, page 2852, August 15, 1979
Copyright © 1979 by the Optical Society of America and reprinted by permission of the copyright owner.

Adjacency effects on imaging by surface reflection and atmospheric scattering: cross radiance to zenith

J. Otterman and R. S. Fraser

An analytical solution is discussed for the nadir radiance as measured from a satellite, based on a simplified single-scattering approximation in which the scattered radiation is not subject to extinction. In the solution, terms can be identified as due to a reflection from the vicinity of the object pixel, and, respectively, (1) upward scattering to zenith above the object pixel, and (2) downward scattering from the entire atmosphere to the object pixel. The first term is referred to as the cross radiance, the second as the cross irradiance. The cross radiance is proportional to the forward scattering optical thickness, as defined, and the cross irradiance to the backscattering optical thickness. The cross radiance usually constitutes the predominant effect. The effect, even at low atmospheric turbidity, can be large enough to constitute a significant fraction of the radiance registered at the satellite, thus hampering determination of spectral signature of the object pixel or identification of pixels with inherently the same spectral signature. Explicit expressions and computer solutions for the cross radiance from annular or from rectangular reflecting areas are presented. The effect depends on the height distribution and on the sharpness of the forward peak of the scattering particles.

I. Introduction

Adjacency effects are well known in photography, arising in the development processes of high-contrast scenes. Because developer consumption depends on local exposure, concentration of the developer in the middle of a highly exposed area becomes lower than at its edge. Unless vigorous mixing is resorted to, development proceeds at an unequal rate. This adjacency effect spuriously increases the contrast at the edge. Some diffraction effects on imaging by a diffraction-limited optical system can, although different in nature, be regarded as adjacency effects: part of the energy from the high reflectivity area spills over to the adjoining low reflectivity pixels, increasing their apparent brightness. What is common in these diverse phenomena is that they become more acute with an increasing contrast. In this paper, another type of adjacency effect is studied.

A previous paper[1] presented an analysis of the enhancement in the global irradiance of the surface and in the radiant emittance from the top of the atmosphere due to surface reflection and subsequent atmospheric scattering. Here the effects of this surface reflection and subsequent scattering on the radiance measured from a satellite at zenith are analyzed as adjacency effects: nadir radiance due to a reflection over one area (the source) and a subsequent scattering over an adjacent area (object pixel) is studied.

The analysis of these effects is oriented toward a specific problem area, i.e., the multispectral radiometry by computer of imagery from satellites or from high altitude aircraft. In multispectral radiometry, the surface is analyzed by assessing the reflectivity of individual pixels in the various bands. Configurations of pixels are not recognized, unlike in photointerpretation by human operators. High contrast between a small group of pixels A and a surrounding group B simplifies the discrimination of A from B. It will be seen, however, that such a high contrast hinders the identification of group A with another group A' of pixels, which have the same inherent reflectivities as A but are part of a large homogeneous field.

The satellite-borne radiometer measures the nadir radiance in a field of view that covers the object pixel on the surface. The signal component of this radiance, contributed by the object pixel, is given by

$$L_{\nu} = rG_t \exp(-Q)/\pi, \tag{1}$$

where r is the object pixel reflectivity (Lambert's law of reflection is assumed), G_t is the total (global) surface irradiance,[2] and Q is the atmospheric vertical optical

J. Otterman is with Tel Aviv University, Department of Geophysics & Planetary Sciences, Ramat Aviv, Israel, and R. S. Fraser is with NASA Goddard Space Flight Center, Greenbelt, Maryland 20771.
Received 10 May 1978.

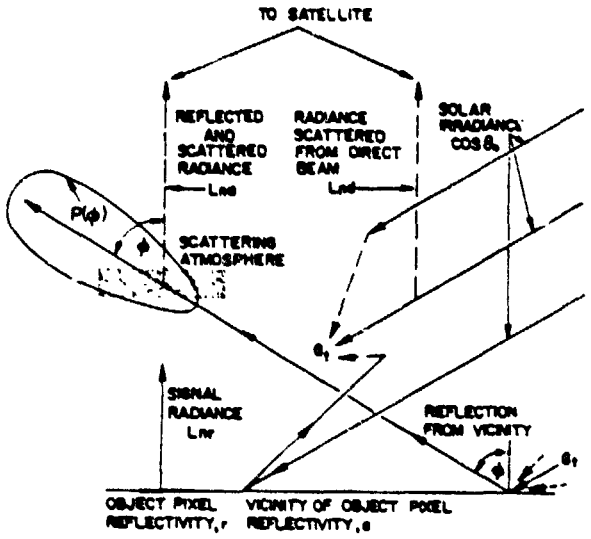


Fig. 1. The geometry of surface reflection and subsequent scattering (a) to the zenith and (b) to the surface. Three components of radiance at the zenith are schematically indicated: (i) L_{nr}, reflection from the object pixel (ii) L_{nd}, radiance scattered from the direct beam and (iii) L_{na}, radiance due to scattering from reflection in the vicinity of the object pixel.

thickness, which accounts for extinction due both to scattering and absorption. To this signal component, the scattering by the atmosphere adds components that do not depend at all on the reflectivity r [see Eq. (13)]. Inferring of r from the measured radiance is thus quite complex.

Thematic mapping by multispectral radiometry can be successfully conducted without determining the actual surface reflectivities by resorting to an approach known as the extrapolation mode. In an extrapolation mode, a known field is pointed out to the computer, and the computer seeks out and prints out all areas with similar satellite-messured spectral radiances. Our basic aim is to identify and map all the pixels with the same spectral reflectivities as that of the sample area presented to the computer. For a successful identification of the pixels with inherently the same characteristics, the measured radiance for any pixel in the entire frame should be a uniquely determinate function of r, monotonically increasing with r (with a reasonable slope[3]). The obvious condition is that the atmosphere be horizontally homogeneous in the frame. To what extent the radiance is not a uniquely determinate function of r, even for a horizontally homogeneous atmosphere, is discussed here. The analysis is limited to cases of low optical thickness.

II. Cross Radiance and Cross Irradiance

The radiance measured by a satellite-borne radiometer L_n in addition to the signal component L_{nr} [Eq. (1)] includes two other components: the radiance scattered from the direct beam L_{nd} and the radiance L_{na}, which is scattered by the atmosphere above the object pixel from the radiation field reflected by terrain in the vicinity (see Fig. 1). In deriving the expressions for L_{nd} and L_{na}, an approach of simplified single-scattering is used in which the scattered photons are not subject to extinction. Thus, a photon once scattered either leaves through the top of the atmosphere or strikes the surface.[1]

Let Q be the total vertical optical thickness (consisting of Rayleigh optical thickness R; Mie scattering optical thickness M; and W, absorption optical thickness either by gaseous components or aerosols), and let θ_0 be the solar zenith angle. The solar irradiance at the top of the atmosphere is $\mu_0 = \cos\theta_0$ of which a fraction $[1 - \exp(-Q/\mu_0)]$ is either scattered or absorbed from the direct beam; therefore, assuming homogeneity of composition with height, a fraction $[1 - \exp(-Q/\mu_0)] \times R/Q$ of the direct beam undergoes Rayleigh scattering and a fraction $[1 - \exp(-Q/\mu_0)]M/Q$ undergoes Mie scattering. $P_R(\phi)$ and $P_M(\phi)$ are Rayleigh and Mie scattering phase functions (normalized so that their integral over the 4π solid angle is one), and the angle between the beam and the direction of scattering is ϕ. The nadir radiance scattered from the direct beam is thus

$$L_{nd} = \mu_0[1 - \exp(-Q/\mu_0)][RP_R(180° - \theta_0) + MP_M(180° - \theta_0)]/Q \tag{2}$$

A more complex equation is obtained when extinction of the scattered radiation is considered. This equation is

$$L_{nd} = \frac{\mu_0[1 - \exp[-Q(1 + \sec\theta_0)]][RP_R(180° - \theta_0) + MP_M(180° - \theta_0)]}{Q(1 + \mu_0)} \tag{2a}$$

This equation can be used to define L_{nd}. Equation (2a) is definitely more accurate than Eq. (2) when the extinction is primarily by absorption, but the simple Eq. (2) can be more accurate when the extinction is primarily by scattering with a large forward peak. In the latter case, a large fraction of extinct photons are scattered through only a small angle. Such scattering can be regarded approximately as an exchange of photons among nearly parallel pencils of scattered radiation.

We compute the radiance L_{na} scattered by a vertical column of the atmosphere with a pixel-size cross section, when the source of illumination is the infinite plane surrounding the column. First to be computed is the case where reflectivity of the entire plane is uniform.

Throughout the paper, Lambert's law of reflectivity is assumed. The reflectivity is denoted a outside the object pixel. This reflectivity generally is different from the reflectivity of the object pixel r. A general reciprocity relation exists when the surface reflectivity and the atmosphere are uniform. the radiation reflected from part A of the plane and scattered by the atmosphere above part B of the plane is equal to the radiation reflected from the surface in B and scattered above A. The computation of L_{na} is then equivalent to computing the total scattering to zenith from a radiation field of reflection from a small pixel with a reflectivity a. Using the simplified single-scattering approach, we can write this radiance as

Table I. Two α and Two β Term Representations of an Aerosol Scattering Phase Function for Two Wavelengths

λ (μm)	α_1	n_1	α_2	n_2	β_1	m_1	β_2	m_2
0.55	0.644	3	0.038	80	0.309	0	0.009	12
0.90	0.628	3	0.046	60	0.316	0	0.011	10

Refractive index of 1.5–0.0i, and particle size distribution of $-dN/d\log r = 4$ (N is the number of particles and r is the particle radius) between $r = 0.3$ μm and 10 μm. The computations were carried out by Y. Kaufman using an approach of least absolute differences.

$$L_{ns} = (aG_t/\pi Q)\int_0^{\pi/2}[1-\exp(-Q/\cos\phi)]\cos\phi[RP_R(\phi) + MP_M(\phi)]2\pi\sin\phi d\phi, \quad (3)$$

when ϕ is the zenith reflection angle. The Rayleigh scattering phase function in our normalization is

$$P_R(\phi) = 3(1+\cos^2\phi)/16\pi. \quad (4)$$

The aerosol scattering phase function is approximated as follows:

$$P_M(\phi) = \left[2\sum_i \alpha_i(n_i+1)\cos^{n_i}\phi + \sum_j \beta_j(m_j+1)\cos^{m_j}\phi\right]/4\pi \quad 0 \le \phi \le (\pi/2)$$

$$= \sum_j \beta_j(m_j+1)\cos^{m_j}\phi/4\pi \quad (\pi/2) \le \phi \le \pi, \quad (5)$$

where n_i and m_j are nonnegative integers, m_j are even, and

$$\sum_i \alpha_i + \sum_j \beta_j = 1. \quad (6)$$

In this paper we do not aim at series expansions, since the required number of α_i terms for a sharply peaked phase function exceeds a thousand. Series expansion of phase functions for various size distributions and types of aerosols, in which economy in the number of α_i terms is imposed, will be the subject of another paper. For the purposes of this paper, we indicate that two α_i terms, with n_2 of the order of $30n_1$, already provide a reasonable representation to a phase function that is peaked around $\phi = 0$ but at the same time is characterized by appreciable scattering in any direction (see Table I). Introducing Eqs. (4) and (5) into Eq. (3), we obtain

$$L_{ns} = \frac{aG_t}{2\pi Q}\int_0^{\pi/2}\cos\phi[1-\exp(-Q/\cos\phi)]\left\{\left[2\sum_i \alpha_i(n_i+1)\cos^{n_i}\phi + \sum_j \beta_j(m_j+1)\cos^{m_j}\phi\right] + \frac{3}{4}(1+\cos^2\phi)R\right\}\sin\phi d\phi$$

$$= \frac{aG_t}{2\pi Q}\left\{\left[2\sum_i \alpha_i(n_i+1)C_{n_i+1}(Q) + \sum_j \beta_j(m_j+1)C_{m_j+1}(Q)\right]M + \frac{3}{4}[C_1(Q)+C_3(Q)]R\right\}. \quad (7)$$

where the functions $C_n(Q)$, related to the exponential integral, are given by

$$C_n(Q) = \int_0^{\pi/2}[1-\exp(-Q/\cos\phi)]\cos^n\phi\sin\phi d\phi. \quad (8)$$

It should be noted that, in Eq. (7), reflection from the entire surface is derived. Thus when Eq. (7) is used for $a \neq r$, a negligible area of the object pixel is tacitly assumed. The radiance L_n measured at the satellite is

$$L_n = L_{nr} + L_{nd} + L_{ns}. \quad (9)$$

The surface irradiance G_t appears both in L_{nr} and L_{ns}. The simplified single-scattering approach[1] leads to the equation (valid only at small or moderate zenith angles)

$$G_t = \mu_0\frac{\exp(-Q/\mu_0) + [1-\exp(-Q/\mu_0)]f}{1-2abC_1(Q)}, \quad (10)$$

where $C_1(Q)$ is defined by Eq. (8), and f and b, respectively, denote the probability of scattering into the forward and backward hemisphere in an encounter by a photon with the atmosphere:

$$f = [(1+\alpha)M + R]/2Q, \quad (11)$$

$$b = [(1-\alpha)M + R]/2Q, \quad (12)$$

where

$$\alpha = \sum_i \alpha_i$$

defines the anisotropy of the aerosol scattering [see Eq. (5)]. Since $M + R$ and the absorbing optical thickness add up to Q, $f + b$ will be equal to unity only if there is no absorption:

$$L_n = \left\{\frac{r\exp(-Q)}{\pi} + \frac{a}{\pi}\left\{\left[\sum_i 2\alpha_i(n_i+1)C_{n_i+1} + \sum_j \beta_j(m_j+1)C_{m_j+1}\right]\frac{M}{2Q} + \frac{3}{4}(C_1+C_3)\frac{R}{2Q}\right\}\right\} \times \mu_0\frac{\exp(-Q/\mu_0) + [1-\exp(-Q/\mu_0)]f}{1-2abC_1} + \mu_0[1-\exp(-Q/\mu_0)] \times \frac{\sum_j \beta_j(m_j+1)\mu_0^{m_j}M + \frac{3}{4}(1+\mu_0^2)R}{4\pi Q}, \quad (13)$$

where the last term is L_{nd}, and $C_n(Q)$ is written C_n. Examining this equation for a homogeneous atmosphere, i.e., when Q, R, M, and α are constant in a frame, it becomes clear that the irradiance L_n is not a uniquely determinate function of r, since it depends also on a, the reflectivity of terrain in the general vicinity of the object pixel. For the purpose of an order of magnitude assessment, a simplified version of Eq. (13) is now presented. Noting that, for $Q \ll 1$,

$$(n+1)C_{n+1}(Q) \simeq Q, \quad (14)$$

we can write Eq. (13) as follows:

$$L_n = \frac{r\exp(-Q) + a\left[(1+\alpha)\frac{M}{2} + \frac{R}{2}\right]}{\pi(1-2abQ)} \times \mu_0[\exp(-Q/\mu_0) + [1-\exp(-Q/\mu_0)]f] + L_{ns} \simeq \frac{\pi(1-Q) + aQ}{\pi(1-2abQ)}\mu_0[\exp(-Q/\mu_0) + [1-\exp(-Q/\mu_0)]f] + L_{ns} \quad (15)$$

The reflectivity a appears twice in Eq. (15). The term in the numerator, afQ, describes a reflection from outside the object pixel and a subsequent scattering to the satellite. This effect can be called the cross radiance.[4] The term in the denominator, $2abQ$, describes the reflection and subsequent scattering downward to the object pixel. This effect can be called the cross irradiance. Compared with the more accurate Eq. (13), Eq. (15) omits powers in Q that are higher than the first (in the terms with a). We can, therefore, restate Eq. (15) in a more convenient form:

$$L_n = \frac{r}{\pi}(1 - Q + \frac{a}{r}fQ + 2abQ)\mu_0 \times [\exp(-Q/\mu_0) + [1 - \exp(-Q/\mu_0)f]] + L_{nd}. \quad (16)$$

The cross radiance afQ/r and the cross irradiance $2abQ$ (both effects when reflection is integrated over the entire surface) can now be analyzed in terms of their functional dependences and orders of magnitude. Both terms are proportional to the scattering optical thickness, the first to the forward scattering optical thickness fQ and the second to the back scattering optical thickness bQ. Both terms are proportional to the reflectivity a, but the first term is divided by r, i.e., it is proportional to the a/r ratio. Since in many scenes a/r can have values of two or three in some of the spectral bands and for aerosols f is higher than b (much higher in the case of large particles), the cross-radiance effects in most situations are higher by an order of magnitude or more than the cross-irradiance effects.[5] For a/r of three, $Q = 0.2$, $f = 0.75$, the cross-radiance term is 0.45. If in the same example $a = 0.2$ and $b = 0.25$ (in a case of a purely scattering atmosphere for which $f + b = 1$), the cross-radiance term is only 0.02. The a/r ratio will generally differ from one band to another. It is quite obvious that the spurious radiance will be typically quite different in the various spectral bands, and the measured band-to-band reflectivity ratios can be sharply altered. The effect thus poses a serious handicap to thematic mapping.

III. Cross Irradiance as Response to Reflection from a Differential Area

In quantifying the problem, a was defined as the reflectivity of terrain in the vicinity of the object pixel without indicating how far this vicinity extends. The obvious question is: over what area should a be computed? The reflectivity of pixels further removed from the object pixel should, intuitively, be given decreasing weight both in the cross-radiance and the cross-irradiance effects. In what follows, the problem is solved for the predominant term of the cross radiance.

The radiance to zenith that originates by scattering out from a pencil of radiation reflected from a unit surface area (which will be referred to as a source or a source pixel) is analyzed as this pencil traverses the atmosphere in a direction ϕ (measured from zenith). The reflected radiation from a unit surface area is aG_t. The projected area into direction ϕ is $\cos\phi$, and, therefore, a radiation pencil into a unit solid angle in this direction is $aG_t \cos\phi/\pi$. Scattering to zenith out of this pencil by a thin layer in the atmosphere is now considered (see Fig. 1). Assuming the satellite to be at infinity, the radiometer field of view, which on the ground encompasses the object pixel, forms a vertical atmospheric column with a cross section equal in area to the object pixel. (The satellite might not be exactly above the object pixel, but still the direction to the satellite defines the local vertical in terms of this discussion.) Such unit cross section at a height h presents to the source a solid angle of $\cos\phi/(h^2 + y^2)$, where y is the distance from the source to the object pixel. Let ΔM denote the vertical optical thickness of the scattering layer, which presents to the reflected pencil an optical thickness $\Delta M/\cos\phi$. In the discussion, an assumption is made that both the unit source area and the area of the object pixel are infinitesimal and that the scattering volume is of infinitesimal thickness. Extinction of the reflected pencil, on the inclined path below the ΔM layer, and of the radiance to zenith, on the vertical path above the ΔM layer, is not considered now, thus allowing us to superimpose scattering of different layers. From these statements, it follows that the radiance $\Delta L_{nau}(\phi)$ scattered to zenith by such a ΔM layer is

$$\Delta L_{nau}(\phi) = aG_t \cos\phi(\Delta M)P_M(\phi)/\pi(h^2 + y^2) = aG_t \cos^3\phi(\Delta M)P_M(\phi)/\pi h^2, \quad (17)$$

where the second equality follows, since

$$\cos^2\phi = h^2/(h^2 + y^2). \quad (18)$$

The phase function $P_M(\phi)$ for Mie scattering is represented by a number of integer cosine power terms. Equation (17) is restated for only a single term of the phase function $\beta M(m + 1)\cos^m\phi/4\pi$:

$$\Delta L_{nau}(y) = \left(\frac{aG_t}{\pi}\right)\frac{(m + 1)\beta\Delta M}{4\pi h^2}\left(\frac{h^2}{h^2 + y^2}\right)^{(m+3)/2}. \quad (19)$$

The analysis for the Rayleigh scattering is along identical lines, with terms $m = 0$ and $m = 2$.

The optical density of an infinitesimal layer ΔM is given as an arbitrary vertical profile of scatterers

$$\Delta M(h) = F(h/H)M\Delta h/H. \quad (20)$$

where H denotes a characteristic height of vertical profile of the scatterers[6] (density scale height in the case of an exponentially decreasing concentration). Integrating over the vertical column of the scatterers, we obtain an expression of L_{nau}, an entire contribution to the nadir radiance over the object pixel from a reflection over a unit source at a distance y from the object pixel. In this integration, extinction of neither the reflected pencil before scattering nor of the scattered radiance is considered:

$$L_{nau} = \frac{aG_t}{\pi}\int_0^\infty \frac{(m + 1)\beta MF(h/H)}{4\pi h^2 H}\left(\frac{h^2}{h^2 + y^2}\right)^{(m+3)/2} dh \quad (21)$$

This expression, L_{nau}, divided by G_t, can be regarded as a response of the nadir radiance to an irradiance on a unit source. The total response is given by integrating $L_{nau}(y)$ over the entire ground plane, over which the reflectivity a can be variable, i.e., it will appear within the integral:

$$L_{ra} = \int L_{rau}(y)dA. \tag{22}$$

Introducing dimensionless variables of height

$$\sigma = h/H \tag{23}$$

and of horizontal distance

$$\kappa = y/H \tag{24}$$

into Eq. (22), where L_{rau} is given by Eq. (21), and α is constant, we have

$$L_{ra} = \frac{aG_t}{\pi}\int_0^\infty \frac{F(\sigma)d\sigma}{\sigma^2}\beta M\frac{m+1}{4\pi}\int_A\left[\frac{\sigma^2}{\sigma^2+\kappa^2}\right]^{(m+3)/2}\frac{dA}{H^2}. \tag{25}$$

In the above, extinction is not considered. From Fig. 1 one can note that the path of the cross radiance, i.e., the path from the source to the scattering column and then to zenith, is always longer than a vertical path length through the atmosphere. Thus, whereas $\exp(-Q)$ denotes the extinction of the reflected signal radiance L_{ra}, L_{rau} should be corrected for extinction by a term $\exp[-g(\kappa)Q]$, where $g \geq 1$.

The total optical thickness gQ along the path of the reflected radiation is essentially identical to a vertical passage through the atmosphere, when the reflection occurs in the immediate vicinity of the object pixel. Quite obviously, therefore, when $\kappa \ll 1$ (i.e., $y \ll H$), $g \simeq 1$. In Sec. IV we discuss L_{ra} for one or more geometrically thin, noninfinitesimal scattering layers, and, subsequently, we discuss scatterers distributed with height, applying superposition. In both of these sections we suggest correction for extinction by $\exp(-Q)$, which is accurate only for an area close to the object pixel $k \ll 1$.

IV. Cross Irradiance for a Concentrated Layer of Scatterers

In this section we analyze the cross radiance due to a geometrically thin layer of scatterers at a height H. $F(\sigma)$ is a Dirac delta function. Superposition of infinitesimal layers is applied, since extinction of both the reflected and the scattered radiation is neglected. Only after superposition integrals are solved do we discuss the extinction of both the reflected and the scattered radiation. The solution can be used either on a single layer or by superimposing the effects on several layers. For one layer, the concentration of scatterers is nil [$F(\sigma)$ is zero], except at $\sigma = 1$, and the integral

$$\int_0^\infty F(\sigma)f(\sigma^2,\kappa^2)d\sigma$$

becomes $f(1,\kappa^2)$. The cross radiance as a fraction of aG_t/π is denoted by l_{ra}. Equation (25) becomes

$$l_{ra} = \beta M[(m+1)/4\pi](1+\kappa^2)^{-(m+3)/2}A/H^2 \tag{26}$$

For forward scattering terms we would have $2\alpha M$ instead of βM. This equation describes the cross radiance contributed by an area A at a distance y away from the center of the object pixel (for example, an annular area $2\pi y\Delta y = A$). The term βM spells out the optical thickness associated with the $\cos^m\phi$ phase function. The term $(m+1)/4\pi$ spells out a peaking effect, due to an enhanced scattering around the zero scattering angle. The term

$$(1+\kappa^2)^{-(m+3)/2}$$

describes the falloff of the cross radiance with the distance, and it can be readily seen that the more pronounced the peaking, the sharper the falloff. Since $\kappa = y/H$, for a low height H of the scattering layer, the falloff is sharper. And finally, the effective size of a source area is measured not in square meters but in terms of A/H^2, i.e., in the dimensionless units.

For an annular area $A/H^2 = 2\pi y dy/H^2 = 2\pi\kappa d\kappa$, the contribution of a source area extending from y_1 and further, i.e., extending from κ_1 to infinity, is given by

$$l_{ra} = \beta M[(m+1)/4\pi]\int_{\kappa_1}^\infty 2\pi\kappa(1+\kappa^2)^{-(m+3)/2}d\kappa$$
$$= (\beta M/2)(1+\kappa_1^2)^{-(m+1)/2} \tag{27}$$

Again, for forward scattering only we would have 2α instead of β.

Consider scattering by a concentrated aerosol layer at $H = 2$ km, with the forward scattering peak characterized by a term with $m = 400$ and a term with $m = 8$. It can be readily computed from Eq. (27) that, for $m = 400$, half of the cross radiance effect is contributed by a very limited area within a radius of only 118 m ($\kappa_1 = 0.0588$) from the center of the object pixel. For Landsat applications this range of 118 m implies that half of the effect originates from only one ring of 8 pixels around the object pixel. On the other hand, where $m = 8$, the half-effect radius extends to 800 m ($\kappa_1 = 0.408$), i.e., half of the effect involves 10–12 rings of pixels around the object pixel. This is further analyzed in the following.

Consider the contribution from a source area in the form of a rectangle and introduce new surface Cartesian coordinates x and w, with an origin at the center of the object pixel and the directions parallel to the sides of the rectangle. The angle ϕ (see Fig. 1) is defined by the height H and the distance from the center of the object pixel to the differential source area. This distance, previously termed as y, is now $(x^2+w^2)^{1/2}$. Thus, in order to derive a solution in terms of the new coordinates, it suffices to introduce x^2+w^2 for y^2 in Eq. (26). The differential source area is now $dxdw$ or, in terms of dimensionless variables, $d\chi d\nu$. Thus, for an arbitrary rectangular source, we have

$$l_{ra} = \beta M[(m+1)/4\pi]\int_{\nu_1}^{\nu_2}d\nu\int_{\chi_1}^{\chi_2}(1+\nu^2+\chi^2)^{-(m+3)/2}d\chi \tag{28}$$

This function has been integrated for Landsat dimension pixels (60 m × 80 m). Results from each pixel in two concentric rings around the object pixel are presented in Fig. 2. (Only 9 pixels are shown, a corner of the array of 25 pixels.) For $m = 400$, the fractional contribution to the cross radiance from each pixel in the array varies from 1% to 6%. The total for the two rings

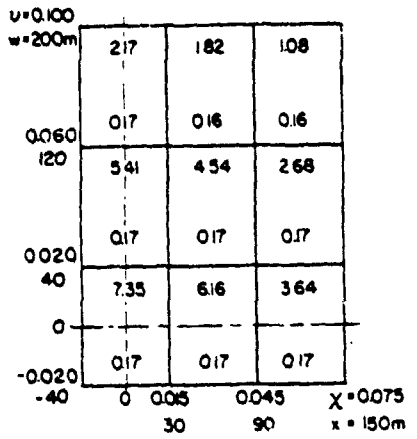


Fig. 2. Relative contribution (in percent) of two rings of Landsat pixels (24 pixels) to the cross radiance over the object pixel, for m = 400 (upper numbers), and m = 8 (lower numbers), H = 2000 m.

of 24 pixels is nearly 75%. (The 7.35% that results from the object pixel itself, acting as a source, does not contribute to spurious effects.) There is a sharp falloff from the ring of pixels nearest to the object pixel to the second ring. For m = 8, the contribution from each pixel is either 0.17% or 0.16% and practically does not vary within the array.

The case of a rectangular source can be easily extended to a semi-infinite plane with reflectivity a (with the other half of the plane nonreflecting). The limits for χ will be $\pm\infty$. A simple analytic solution exists for these limits. We have

$$L_{ns} = \beta M[(m+1)/4\pi]\int_{\nu_1}^{\infty} d\nu \int_{-\infty}^{\infty} (1+\nu^2+\chi^2)^{-(m+3)/2} d\chi$$

$$= \beta M \frac{m+1}{4\pi}\int_{\nu_1}^{\infty} (1+\nu^2)^{-(m+2)/2} \frac{2^{(m+2)/2}\left(\frac{m}{2}\right)!}{(m+1)!!} d\nu$$

$$= \beta M \frac{m!!}{2\pi(m-1)!!}\int_{\nu_1}^{\infty} \frac{d\nu}{(1+\nu^2)^{(m+2)/2}}. \qquad (29)$$

The term $\beta M/2$ represents the solution for the entire plane with a uniform reflectivity and can be regarded as the magnitude of the effect [see Eq. (27) for $\kappa_1 \to 0$]. Since

$$\frac{m!!}{\pi(m-1)!!}\int_0^{\infty} \frac{d\nu}{(1+\nu^2)^{(m+2)/2}} = 0.5, \qquad (30)$$

at $\nu_1 = 0$ the cross radiance is obviously 0.5 ($\beta M/2$). It will be noted that the slope (in terms of fractional magnitude vs the dimensionless distance) at $\nu_1 = 0$ is given by $m!!/\pi(m-1)!!$. The functions $m!!/\pi(m-1)!!$, plotted in Fig. 3, and a discontinuity in L_{ns} describe how the nadir radiance increases crossing from a black half-plane to a reflecting half-plane.

Now we consider extinction of the reflected pencil and of the scattered radiance—first for the case of a single concentrated layer, i.e., superposition at the same height of infinitely many infinitesimal layers with optical thickness dq. The probability of undergoing one and only one encounter, for photons entering the layer at an angle ϕ from below the layer (pencil of reflected radiation) and emerging after the scattering vertically upward, is given by

$$\int_0^M \exp(-q/\cos\phi)\exp(-M+q)dq/\cos\phi$$

$$= \frac{\exp(-M)[1-\exp(-M\sec\phi+M)]}{\cos\phi(\sec\phi-1)}, \qquad (31)$$

where q denotes the optical thickness at the point of the scattering, and $q = 0$ denotes the optical thickness at the bottom of the layer. We approximate this expression by $M/\cos\phi$. The superposition integrals for a single concentrated layer are then corrected for extinction by the term $\exp(-M)$, which is also the extinction term for the signal radiance. This solution is very simple, but the question is how serious is the error in the approximation. The approximation is that the term $[1-\exp[-M(\sec\phi-1)]]/M(\sec\phi-1)$ is taken as equal to unity. The error increases with increasing ϕ and M. For $\phi = 60°$, i.e., $\sec\phi = 2$ and $M = 0.1$ and 0.25, the term is, respectively, 0.95 and 0.88. Contributions to cross radiance from a ring at a distance $\kappa = \tan^{-1} 60° = 1.73$ are overestimated, at most by 5% for $M = 0.1$ and by 12% for $M = 0.25$, when we superimpose infinitesimal layers into a concentrated layer. The qualification at most needs to be inserted, since, when we incorporate in the expression such an extinction by scattering, we lose track of the extinct photons. They still might wind up at the satellite field of view above the object pixel, and the probability of it is high if f is high.

There is no such uncertainty regarding extinction by absorption, that it indeed reduces the reflected and scattered radiation. Absorption is included in the qQ optical thickness, since the absorbing optical thickness W is a component of Q. However, special consideration should be given to situations where absorbers are not distributed with the same vertical profile as the scatterers. Absorbers can specifically be above the scatterers (as in the case of ozone) or, in a ground-hugging layer, effectively below the scatterers (as in the case of water vapor).

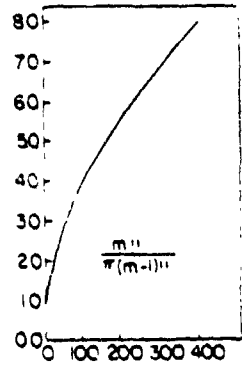


Fig. 3. The plot of the $m!!/(m-1)!!$ vs m, the slope of L_{ns} (normalized to 1 at infinity) vs the dimensionless distance at a boundary between reflecting and black half-planes, in the direction perpendicular to the boundary, a case of thin layer of scatterers.

In the first case, the extinction factor by absorption is $\exp(-W)$, i.e., the g factor is 1 and is independent of κ. Thus, the absorption affects the signal and the cross radiance in the same way, and, in relative terms, it has no effect. In the second case, the extinction factor is $\exp(-W_0/\cos\phi) = \exp[-W(h^2 + y^2)^{1/2}/h] = \exp[-W_0(1 + \kappa^2)^{1/2}]$, when the scatterers are concentrated at $\sigma = (h/H) = 1$, and W_0 denotes the absorbing optical thickness of the ground-hugging layer. Introducing this factor into Eq. (27) one obtains.

$$L_{na} = \beta M[(m+1)/4] \int_{\kappa_1}^{\infty} \times \exp[-W_0(1+\kappa^2)^{1/2}](1+\kappa^2)^{-(m+3)/2} d\kappa^2. \quad (32)$$

A source at $\phi = 60°$, i.e., at $\kappa = 1.73$, is reduced through such absorption by a factor $\exp(-2W_0)/\exp(-W_0) = \exp(-W_0)$ as compared with a source in the immediate vicinity of the object pixel $\kappa \ll 1$. Equation (32) can be integrated after a change of variables $\xi = (1 + \kappa^2)^{1/2}$ through integration by parts

$$L_{na} = \frac{\beta M}{2} \int_{\xi_1}^{\infty} \exp(-W_0\xi)\xi^{-(m+2)} d\xi = \frac{\beta M}{2} \left[\frac{(-1)^{m+2} W_0^{m+1} E_1(-W_0\xi_1)}{(m+1)!} + \frac{\exp(-W_0\xi_1)}{\xi_1^{m+1}} \times \sum_{k=0}^{m} \frac{(-1)^k W_0^k \xi_1^k}{(m+1)m \cdots (m+1-k)} \right]. \quad (32a)$$

where E_1 is the exponential integral of order one. Equation (32a) has been evaluated for three values of W_0, $m = 0$, 2, 8, and 400, and for $\kappa_1 = 0$ and $\kappa_1 = 1.73$. The integration indicates that, for $W_0 = 0.05$, $W_0 = 0.15$, and $m = 2$: (a) the total L_{na} over the entire surface is reduced by a factor of 0.976 and 0.933, respectively, as compared with a case of absorption with the same optical thickness but above the scatterers; and (b) the fraction contributed by sources extending beyond $\kappa_1 = 1.73$ are reduced to 0.116 and 0.101, respectively, of the L_{na} as compared with a fraction of $2^{-(m+1)} = 0.125$ of the L_{na} in the case of no absorption [Eq. (27)] or as compared with an absorbing layer above the scatterers. Fuller data are presented in Table II. It can be seen from these data that just multiplying L_{na} by $\exp(-W_0)$ is a completely appropriate treatment for $m = 8$ or higher, while for a very low m, more detailed analysis, based on Eq. (32), is necessary.

V. Scatterers Exponentially Decreasing with Height

Assuming that the scatterers decrease exponentially with height, i.e., $F(h/H) = \exp(-h/H)$, Eq. (25) becomes

$$I_{na} = \beta M[(m+1)/4\pi] \int_0^{\infty} \sigma^{m+1} \exp(-\sigma) d\sigma \times \int_{A_1} (\sigma^2 + \kappa^2)^{-(m+3)/2} dA/H^2. \quad (33)$$

where extinction of both the reflected and the scattered radiation is not considered, and κ, as before, denotes the dimensionless distance from the center of the object

Table II. Effects of a Ground-Hugging Absorbing Layer with an Optical Thickness W_0 on Cross Radiance for a Term in the Phase Function with an Exponent m

m	0	2	8	400
$W_0 = 0$	0.500 1.0	0.125 1.0	0.00195 1.0	0.19×10^{-120} 1.0
$W_0 = 0.05$	0.436 0.870	0.116 0.976	0.00185 0.994	0.18×10^{-120} 0.9999
$W_0 = 0.10$	0.397 0.799	0.108 0.954	0.00175 0.988	0.17×10^{-120} 0.99975
$W_0 = 0.15$	0.366 0.745	0.101 0.933	0.00165 0.982	0.17×10^{-120} 0.9996

The upper number represents the relative contribution of reflection from sources farther than $\kappa = 1.73$ (i.e., at 60° and higher from the point in the scattering layer above the object pixel). The lower number represents the reduction of the total cross radiance from the entire plane as compared with the case of an absorption W_0 above the scatterers. This lower number when multiplied by $\exp(-W_0)$ represents the reduction of the cross radiance as compared with the case of no absorption.

$$\int_0^{\infty} \left(\frac{\sigma^2}{\sigma^2 + \kappa_1^2}\right)^{\frac{m+1}{2}} e^{-\sigma} d\sigma$$

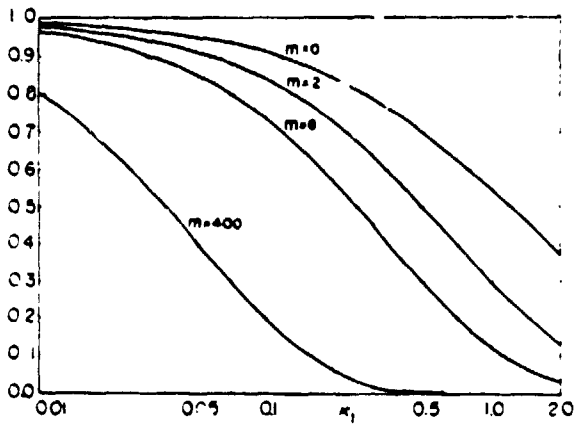


Fig. 4. Contribution to the cross radiance of the reflection from an infinite annular region further than κ_1, exponentially distributed scatterers, numerical integration of

$$\int_0^{\infty} \left(\frac{\sigma^2}{\sigma^2 + \kappa_1^2}\right)^{(m+1)/2} e^{-\sigma} d\sigma$$

pixel. When $(dA/H^2) = 2\pi\kappa d\kappa$, the second integral of Eq. (33) becomes easily integrable. We integrate the effect from a given distance κ_1 to infinity:

$$I_{na}(\kappa_1, \infty) = \beta M[(m+1)/4\pi] \int_0^{\infty} \sigma^{m+1} \exp(-\sigma) d\sigma \times \int_{\kappa_1}^{\infty} 2\pi(\sigma^2 + \kappa^2)^{-(m+3)/2} \kappa d\kappa = \frac{\beta M}{2} \int_0^{\infty} \exp(-\sigma) \left(\frac{\sigma^2}{\sigma^2 + \kappa_1^2}\right)^{(m+1)/2} d\sigma \quad (33a)$$

This integral has been evaluated for $m = 0$, 2, 8, and 400 and for $0 < \kappa_1 < 2.0$. The results are presented in Fig. 4.

It can be seen by examining Fig. 4 that there is a vast difference between the decrease with distance of the

high exponent terms and that of the low exponent terms. For $m = 8$, the contribution from outside of a circle with radius equal to one-tenth of the scale height is 74%, whereas for $m = 400$, it is only 20%. The comparable numbers for $m = 2$ and $m = 0$ are 84% and 92%, respectively. In terms of actual distances, the differences are even sharper. Large particle aerosols are likely to be encountered close to the surface, and a representative scale height can be 2 km, of which one-tenth is 200 m. Beyond this distance, the contribution of a sharp scattering peak can drop to only 20%. On the other hand, for Rayleigh scattering, density scale height is 9300 m. Thus the fractional contribution from areas farther than 930 m is still (3/4)92 + (1/4)84 = 90% (in Rayleigh scattering 3/4 is isotropic, $m = 0$, and 1/4 according to $\cos^2\phi$ relation). But in reality this fraction will be substantially smaller when extinction is considered over long paths from the source areas distant to the scattering column above the object pixel.

The case of an infinite half-plane is now given by

$$
\begin{aligned}
I_{ns} &= \beta M[(m+1)/4\pi] \int_0^\infty \sigma^{m+1} \exp(-\sigma) d\sigma \\
&\quad \times \int_{\nu_1}^\infty d\nu \int_{-\infty}^\infty \frac{d\chi}{(\sigma^2+\nu^2+\chi^2)^{(m+3)/2}} \\
&= \beta M[(m+1)/4\pi] \int_0^\infty \sigma^{m+1} \exp(-\sigma) d\sigma \\
&\quad \times \int_{\nu_1}^\infty \frac{2^{(m+2)/2}\left(\frac{m}{2}\right)! d\nu}{(m+1)!!(\sigma^2+\nu^2)^{(m+2)/2}} \\
&= \beta M \frac{m!!}{2\pi(m-1)!!} \int_0^\infty \sigma^{m+1} \exp(-\sigma) d\sigma \\
&\quad \times \int_{\nu_1}^\infty \frac{d\nu}{(\sigma^2+\nu^2)^{(m+2)/2}}. \qquad (34)
\end{aligned}
$$

The results are plotted in Fig. 5.

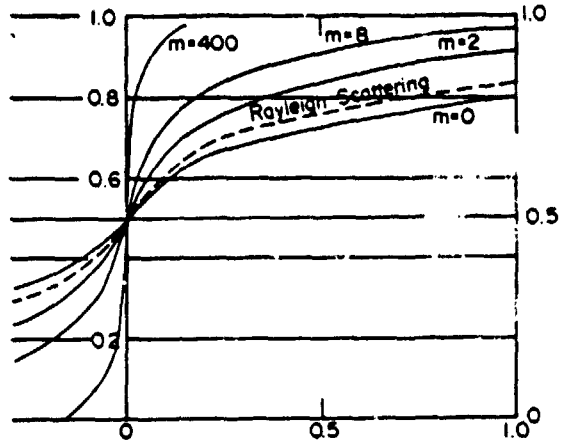


Fig. 5. The radiance L_{ns}, normalized to 1 at + infinity vs the dimensionless distance ν_1 perpendicular to a boundary between a reflecting half-plane and a black half-plane; the reflecting half-plane extends from $\nu = 0$ to $+\infty$; exponentially distributed scatterers; numerical integration of

$$\frac{m!!}{\pi(m-1)!!} \int_0^\infty e^{-\sigma}\sigma^{m+1} d\sigma \int_{\nu_1}^\infty \frac{d\nu}{(\sigma^2+\nu^2)^{(m+2)/2}}$$

VI. Assessment of the Adjacency Effects for Landsat Multispectral Scanner System (MSS) and the Thematic Mapper

This year (1979) marks the seventh anniversary of the successful and continued observation of the earth through the multispectral scanner system (MSS), which first operated on Landsat 1 and now operates on Landsat 2 and 3. Analysis of the data provided by this four-spectral-band radiometer made monitoring of renewable earth resources practical in a number of disciplines such as agriculture, forestry, hydrology, and rangeland studies.[8,9] The approach of multispectral radiometry led to the successful computerization of data analysis and the establishment of almost routine computer thematic mapping of a few crops for large fields in some regions of the U.S.A.[10] The atmospheric effects were found to be not significantly detrimental to the successful multispectral radiometry of a Landsat frame (i.e., a case of single-temporal analysis), even when a fairly turbid but homogeneous intervening atmosphere was artificially introduced.[11] Thus successful radiometry of the surface from Landsat satellites is clearly indicated for relatively large areas, for which the veil of the scattered atmospheric radiance can be assumed to be homogeneous.

This conclusion does not necessarily apply to small areas. Thanks to the rather large aperture of the MSS, the diffraction problems, even for a single pixel surrounded by much brighter terrain, are quite small.[12] But the adjacency effects as discussed here can be quite significant for an isolated pixel and even for a group of several scores of Landsat-size pixels. To demonstrate the point, Eq. (27) can be applied or, more conveniently, Fig. 2 can be reexamined. We analyze the contribution to cross radiance above an object pixel located in the center of the array of 25 pixels shown in Fig. 2. For $m = 400$, only 17.5%, while for $m = 8$ some 96% is contributed from outside the array. If we computed a representative relative spurious signal of 0.45, a spurious signal of $0.96 \times 0.45 = 0.42$ will occur when $m = 8$ in the center of a field of 25 Landsat pixels, i.e., a field 300 m × 400 m. Such a field constitutes over 100 Thematic Mapper pixels. (Thematic Mapper is a multispectral radiometer planned for the Landsat program, with ground resolution of about 30 m).

VII. Conclusions

A treatment of the subject of surface reflection and subsequent atmospheric scattering has been developed. Some specific conclusions can be drawn.

(1) The effects are amenable to analytical treatment

for low optical thickness conditions. The parameters governing these effects have been identified, which can lead to quantitative assessment of the effects.

(2) The effects can easily mask the true spectral signature of a group of pixels (of Landsat or Thematic Mapper dimensions) when surrounded by terrain of sharply different reflectivity (for example, vegetation surrounded by bare sandy soil or any surface surrounded by snow).

(3) The effects are proportional to the difference in reflectivities between the object pixel and the surrounding terrain and also (for a low optical thickness) to the optical thickness of the scatterers.

(4) An increased radiance measured above the object pixel due to a reflection from the vicinity and subsequent upward scattering is termed here cross radiance. The effect accrues only from the vertical column of the atmosphere above the object pixel and is proportional to the forward scattering optical thickness fQ.

(5) The ground irradiance increases through reflection from neighboring terrain and subsequent backscattering from the atmosphere. This backscattering accrues from the entire atmosphere. This effect, called cross irradiance, also induces a spurious signature, i.e., affects the inferred spectral reflectivities and is proportional to the backscattering optical thickness bQ. The effect of cross irradiance is typically smaller, at least by an order of magnitude, than the effect of cross radiance. However, it should be pointed out that this conclusion is based on an assumption that the forward hemisphere of the phase function does not contribute to the backscattering of the reflected radiation. For low angles of reflection, this assumption is definitely incorrect. Indeed, over a highly reflective plain, the horizon appears sometimes very bright, and the resulting cross irradiance is considerably higher than as computed by the $2ab\bar{C}_1(Q)$ term. A more detailed treatment of cross irradiance including its angular distribution is being prepared,[13] but it can be pointed out now that the effects of reflection at low angles depend quite strongly on local slopes, on the absence or presence of protrusions from the surface, such as isolated trees or houses, and on absorption in the ground-hugging layer.

(6) The distance over which a source contributes significantly to the cross-radiance effect depends inversely on the average or effective height of the scattering layer. Indeed, the effects of a source should be formulated and analyzed in terms of a dimensionless measure of distance, i.e., in terms of the ratio of the horizontal distance from the object pixel to the height of the scatterers.

(7) The distance over which the cross radiance effectively accrues depends crucially on the phase function of the scatterers. The sharper the forwardscattering peak (which is characteristic of large particles), the stronger nearby effects are. The contribution of further regions becomes negligible. For Rayleigh scattering in the absence of a strong forward peak, the sources extend over large distances. Since the effect thus accrues over a very large number of pixels, it is a more broadly averaged effect.

(8) For an absorbing atmosphere, the cross radiance relative to the signal radiance remains the same if the absorbing layer is high relative to the scatterers. When the absorbing layer is low, the nearby effects remain proportionally the same, but there is a sharper decrease for distant sources.

(9) Appropriate computer programs could be developed to eliminate most of the spurious effects on the spectral signatures. However, for accuracy and relevancy of such corrections, it is necessary to know not only the aerosol optical thickness but the height of the scatterers and their phase function. This information in general is not available, and methods to obtain it should be studied.

(10) Other spurious effects can mask the spectral signatures. The problem of inhomogeneities in the atmosphere is well recognized. A reflection from a bright mountain slope's contribution enhances the irradiance in the plain in the foothills area, which can locally be a predominant effect.

Comments by an *Applied Optics* reviewer resulted in revisions, which the authors believe have significantly improved the presentation, and suggested direction for some future work. Computations by S. Rehavi, Y. Kaufman, and M. Podolak, all of Tel Aviv University, are gratefully acknowledged.

References

1. J. Otterman, Appl. Opt. 17, 3431 (1978).
2. The surface irradia. e G_i depends on solar zenith angle, atmospheric parameters, and also on reflection from terrain surrounding the object pixel and subsequent scattering.
3. For low surface reflectivities, the slope of space measure reflectivity against the surface reflectivity is significantly smaller than unity, and there is a loss in discrimination capability
4. The phenomenon has some similarities to an effect known as the cross talk, i.e., an unwanted coupling between separate telephone channels. Here the channels are the vertical propagation paths through the atmosphere, and pixels on the surface correspond to the individual speakers. The object pixel corresponds to a speaker for whom we analyze the quality of transmission.
5. See Sec. VII where the validity of this statement is qualified.
6. If the size distribution varies with height, β is a function of h. The treatment can be generalized by introducing

$$\beta_i \Delta M(h) \to F_i(h/H_i) M \Delta h/H_i$$

7. G_i is assumed constant, i.e., variability of the cross irradiance between sources, which occurs when a is variable, is neglected.
8. E. P. Mercanti, Ed., Astronaut. Aeronaut. 36 (1973).
9. J. Otterman, P. D. Lowman, and V. V. Salomonson, Geophys. Surv. 2, 431 (1976).
10. A. Park, in *Contribution of Space Observations to Global Food Information Systems*, E. A. Godby and J. Otterman, Eds. (Pergamon, New York, 1978), pp. 105–113.
11. R. S. Fraser, O. P. Bahethi, and A. H. Al-Abbas, Remote Sensing Environ. 6, 229 (1977).
12. J. Otterman, J. Br. Interplanet. Soc. 23, 349 (1970).
13. J. Otterman, Y. Kaufman, M. Podolak, and S. Rehavi, "Adjacency Effects on Imaging by Surface Reflection and Atmospheric Scattering: Sky Radiance and Cross Irradiance," in preparation (1979).

Paper 125

Reprinted from *Monthly Weather Review*, 107, pp. 1388-1401, 1979.

Remote Sensing of Seasonal Distribution of Precipitable Water Vapor over the Oceans and the Inference of Boundary-Layer Structure

C. PRABHAKARA AND G. DALU[1]

Laboratory for Atmospheric Sciences, NASA/Goddard Space Flight Center, Greenbelt, MD 20771

R. C. LO AND N. R. NATH

Computer Sciences Corporation, Silver Spring, MD 20910

(Manuscript received 2 April 1979, in final form 5 July 1979)

ABSTRACT

From the depth of the water vapor spectral lines in the 8–9 μm window region, measured by the Nimbus 4 Infrared Interferometer Spectrometer (IRIS) with a resolution of about 3 cm^{-1}, the precipitable water vapor w over the oceans is remotely sensed. In addition the IRIS spectral data in the 11–13 μm window region have been used to derive the sea surface temperature (SST). Seasonal maps of w on the oceans deduced from the spectral data reveal the dynamical influence of the large-scale atmospheric circulation. With the help of a model for the vertical distribution of water vapor, the configuration of the atmospheric boundary layer over the oceans can be inferred from these remotely sensed w and SST. The gross seasonal mean structure of the boundary layer inferred in this fashion reveals the broad areas of trade wind inversion and the convectively active areas such as the ITCZ. The derived information is in reasonable agreement with some observed climatological patterns over the oceans.

1. Introduction

The general circulation of the atmosphere displays some important climatological features over the oceans such as the subtropical anticyclones and the intertropical convergence zones. The circulation of the subtropical anticyclones over the oceans extends to about 40°N and 40°S in the respective hemispheres. The equatorward trade winds associated with these anticyclones are instrumental in transporting water vapor to intertropical convergence zones and thereby help to maintain the directly driven Hadley circulation in the tropics (Riehl, 1954; Malkus, 1956). A salient feature of the trade wind circulation in the lowest layers of the atmosphere is the trade wind inversion which is produced as a consequence of the large-scale dynamics of the subtropical anticyclones. The base of this inversion is found to slope upward along the downstream (NE to SW in the Northern Hemisphere) in response to the joint influence of small-scale convective motions and large-scale subsidence (Malkus, 1956; Mak, 1976). Upward transport of water vapor by convective motions is inhibited by the trade wind inversion layers. From the foregoing discussion we see that the distribution of water vapor over the oceans is influenced by both the large-scale and small-scale motions and as such one may consider water vapor as a tracer of atmospheric motions.

Water vapor undergoes changes of phase in the atmosphere. Its source is essentially at the surface, while the principal sink is in the atmosphere where it condenses and precipitates out. Despite these physical and dynamical influences on water vapor, the mean vertical profile of relative humidity in the atmosphere apparently has a simple character. Based on Telegadas and London's (1954) climatology of the meridional distribution of relative humidity, Manabe and Wetherald (1967) have modeled the vertical profile of relative humidity such that it decreases from a value of 77% at the surface to ~40% at the 500 mb level.

Such a decrease in relative humidity is in good agreement with the findings of a recent study by Nieman (1977), who derived several mean relative humidity profiles on a seasonal basis for different latitude belts, from 1972 global radiosonde data. Further, these mean relative humidity profiles for different latitude zones and seasons show a linear decrease with height. This result suggests that "given sufficient time, the atmosphere tends to restore a certain climatological distribution of relative humidity responding to change of temperature" (Manabe and Wetherald, 1967).

[1] NAS/NRC Research Associate on leave from C.N.R., Istituto di Fisica della Atmosfera, Rome, Italy.

0027-0644/79/101388-14$07.50
© 1979 American Meteorological Society

Over the oceans there are two important phenomena which can produce significant departures in the relative humidity profile from its mean value. When low-level convergence is present in the atmosphere, a deep convective layer with above average relative humidity is formed. On the other hand, if stable conditions prevail, such as those associated with inversions, the relative humidity above the inversion is significantly less than the mean value. These departures in the relative humidity profile are reflected as deviations in the total precipitable water vapor from some mean value. Such deviations in precipitable water vapor will be appreciable when the relative humidity profile is perturbed from the mean conditions in the first few kilometers above the surface. The reasoning given above constitutes the basis of this study to infer the boundary-layer structure over the oceans.

The precipitable water vapor over the oceans needed for this study is remotely sensed from the spectral measurements made by the Infrared Interferometer Spectrometer (IRIS) that was flown on Nimbus 4 satellite.

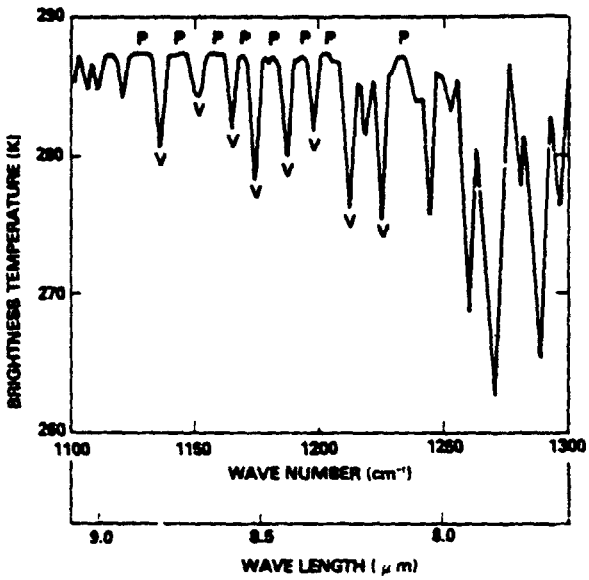


FIG. 1. Nimbus 4 IRIS brightness temperature spectrum from 1100 to 1300 cm^{-1}. The symbols P and V denote the peaks and valleys in the spectrum.

2. Spectral information in the window region

The Nimbus 4 IRIS gathered spectral measurements from 400 cm^{-1} to ~1400 cm^{-1} over the globe with a spectral resolution of 2.8 cm^{-1}, where the noise in the spectral data is about 0.5 erg cm^{-1} sr^{-1} s^{-1} (Hanel *et al.*, 1972). In a typical brightness temperature spectrum, the various water vapor bands, i.e., the rotation band around 20 μm, and the 6.3 μm vibration rotation band, as well as the window regions at 11 and 9 μm show some spectral details at 2.8 cm^{-1} resolution. In Fig. 1 a part of the brightness temperature spectrum from 1100 to 1300 cm^{-1} is shown to illustrate these spectral details.

The sharp maxima (peaks) and minima (valleys) in the figure suggest that the absorption coefficient of water vapor changes significantly in narrow spectral intervals. This implies further that radiances, corresponding to the brightness temperature, at the peaks and valleys in the spectrum, arise from significantly different altitudes in the atmosphere.

In this study we are interested in examining the information pertaining to the total precipitable water in the atmosphere, the bulk of which is contained in the first few kilometers near the surface. Hence we have considered the water vapor spectral information in the window regions of 11–13 μm and 8–9 μm. These window regions are referred to as the 11 and 9 μm regions, respectively. In the 9 μm region the continuum absorption due to self-broadening effect is weak (Bignell, 1970; Burch, 1970), while the lines are strong compared to the 11 μm region. For this reason the water vapor spectral measurements in the 9 μm window region are used to deduce the precipitable water vapor over the ocean. We will develop this scheme using radiative transfer theory.

The intensity I_ν measured at the top of the atmosphere, at wavenumber ν, in a nonscattering atmosphere in local thermodynamic equilibrium, is given by

$$I_\nu = B_\nu(p_0)\tau_\nu(p_0) + \int_{\tau_\nu(p_0)}^{1} B_\nu(p)d\tau_\nu(p), \qquad (1)$$

where B_ν is the Planck intensity, $\tau_\nu(p)$ is the transmission of the atmosphere from any given pressure p to the top of the atmosphere, and p_0 is the surface pressure. The emissivity of the surface is assumed to be unity.

When we apply this equation to the spectral data in the 9 μm water vapor window region, the first term on the right-hand side, i.e., the contribution from the surface, is essentially governed by the surface temperature and the total precipitable water content in the atmosphere. The second term accounts for the emission from the whole atmosphere, and it depends on the manner in which the water vapor and temperature are distributed. At the peaks in the 9 μm region, where absorption is weak, the contribution from the second term is much smaller than the first. Thus, the intensity at the peaks depends only weakly on the atmospheric stratification. However, in the valleys of the spectrum, where the absorption is considerably larger, the second term in the equation exceeds the first. Taking the difference ΔI between the mean intensity I_p at the peaks and the mean intensity I_v in the valleys, we obtain

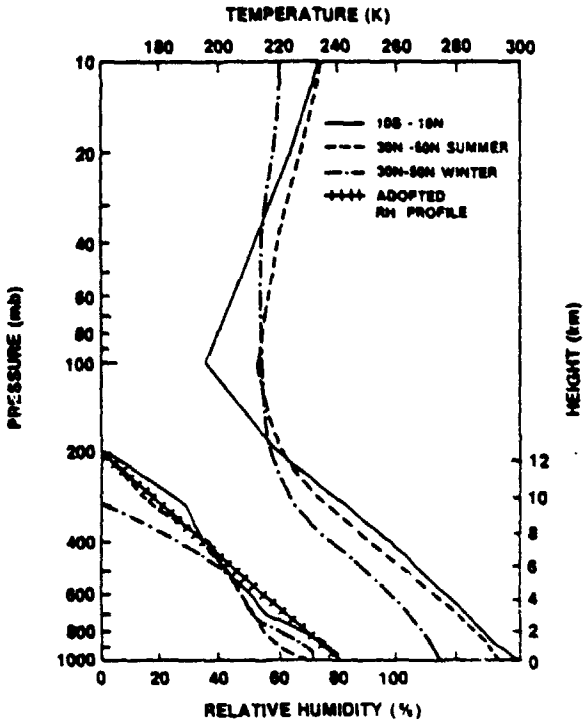


FIG. 2. Mean relative humidity and temperature profiles for tropics and midlatitudes summer and winter (from Nieman, 1977). The adopted RH profile is also shown.

$$\Delta I = I_p - I_v \approx \bar{B}(p_0)[\tau_p(p_0) - \tau_v(p_0)] + \int_{[\tau_p(p_0)-\tau_v(p_0)]}^{0} \bar{B}(p)d[\tau_p(p) - \tau_v(p)], \quad (2)$$

where $\bar{B}$ is the average Planck intensity over the peaks and valleys in the spectrum which are typically spaced within 10 cm^{-1} apart, and τ_p and τ_v are mean transmittances in the peaks and valleys of the spectrum, respectively.

Eq. (2) may be replaced with the following functional relationship

$$I = f[T_s, T(p), \mathrm{RH}(p)], \quad (3)$$

where I (called the line strength) is the brightness temperature difference corresponding to ΔI, $T(p)$ and $\mathrm{RH}(p)$ are temperature and relative humidity as functions of pressure p, and T_s is the surface temperature.

Given the surface temperature T_s, a mean temperature profile $\bar{T}(p)$ could be constructed assuming an average lapse rate. Further, if a mean relative humidity profile $\overline{RH}(p)$, based on climatological data, could be defined we can express the mean line strength $\bar{I}$ as

$$\bar{I}(T_s) = f[T_s, \bar{T}(p), \overline{RH}(p)]. \quad (4)$$

It is apparent from these assumptions that $\bar{I}$ is a function of the mean precipitable water vapor $\bar{w}(T_s)$ in a column of the atmosphere.

The temperature profiles in the atmosphere for a given T_s may differ from such mean conditions. Also the relative humidity profile could differ from the climatological mean conditions. In such a case the line strength I departs from $\bar{I}$ by an amount

$$\Delta I = \bar{I} - I = g[T_s, T'(p), \mathrm{RH}'(p)], \quad (5)$$

where T' and RH′ are deviations from the mean conditions.

The surface temperature T_s can be independently determined from the 11 μm window measurements of IRIS using a method developed by Prabhakara *et al.* (1974). From Eq. (5), therefore, it is plausible that from the quantity ΔI, we can infer one dominant mode or property of the atmosphere implicitly contained on the right-hand side of the equation. Over water bodies, as will be discussed later, there are strong correlations between $T'(p)$ and $\mathrm{RH}'(p)$. Further such correlations exist in the lowest layers of the atmosphere near the surface. Principally, this results in a significant deviation of precipitable water vapor from the mean value $\bar{w}$.

Thus if we can deduce these deviations from the IRIS spectral measurements over the global oceans, we can infer some salient features of the lowest layers of the atmosphere.

3. Determination of precipitable water vapor in different atmospheres

In order to study the manner in which I, as defined previously, changes as a function of the atmospheric conditions, we have developed a radiation computational scheme that can simulate the spectral data with a resolution of 2.8 cm^{-1} in the water vapor absorption regions from 400 cm^{-1} to 1400 cm^{-1}. The transmission function τ of the water vapor needed in these calculations is taken to be the product of three components: τ_l associated with water vapor lines, τ_p produced by the continuum due to foreign broadening, and τ_e resulting from the e-type absorption (Bignell, 1970). The transmission function of the water vapor lines τ_l is derived using a multiple-regression scheme similar to the one proposed by Smith (1969). The details of the foreign broadening and the e-type given by Kunde and Maguire (1974) are adopted.

This simulation program is used to synthesize spectral data for several different atmospheres ranging from tropics to high latitudes.

a. Average atmospheric conditions

The atmospheric temperature and relative humidity distributions corresponding to mean conditions in the latitudinal belts 0–10°, 10–30° and

TABLE 1a. Coincident data from ship radiosonde stations and Nimbus 4 IRIS.

Day	Ship	T_{IRIS} (K)	T_{SHIP} (K)	RH(p_0) (%)	l_{IRIS} (°C)	w_{IRIS} (g cm^{-2})	w_{SHIP} (g cm^{-2})
19 Apr	4 Y M	275.9	277.9	93	3.37	1.16	0.93
22 Apr	W T K A	297.0	299.4	82	9.12	4.11	3.82
7 May	4 Y J	280.4	281.0	66	3.75	1.32	0.91
9 May	4 Y N	290.5	290.9	68	4.78	1.78	1.90
12 May	4 Y M	278.7	280.1	84	3.03	1.01	0.99
17 May	4 Y N	289.9	291.5	76	5.34	2.04	1.93
21 May	4 Y J	282.4	283.7	92	3.30	1.15	1.58
6 Jun	4 Y N	290.4	292.9	68	5.09	1.92	1.72
12 Jun	4 Y V	294.3	292.7	98	8.18	3.54	3.46
18 Jun	4 Y J	282.5	285.1	76	3.96	1.41	1.35
21 Jun	4 Y E	294.6	294.4	89	8.75	3.87	3.71
21 Jun	4 Y N	291.1	293.1	82	5.92	2.33	1.73
2 Jul	4 Y J	283.9	284.9	85	4.38	1.61	1.82
3 Jul	4 Y D	290.6	293.1	92	5.59	2.18	2.80
12 Jul	4 Y E	298.6	297.7	77	8.06	[illegible]8	3.50
17 Jul	4 Y V	295.7	297.9	88	6.44	2.60	2.64
31 Jul	4 Y D	293.5	293.3	95	7.12	2.96	2.58
2 Aug	4 Y N	294.8	296.1	69	5.83	2.28	1.83
6 Aug	4 Y B	280.7	282.7	82	3.17	1.08	1.18
7 Aug	4 Y C	283.6	283.7	91	4.87	1.83	1.78
9 Aug	4 Y N	295.0	295.9	76	4.96	1.83	1.74
11 Aug	4 Y D	294.1	293.5	80	5.96	2.36	2.07
23 Aug	4 Y I	282.4	285.1	87	4.67	1.74	1.72
23 Aug	4 Y N	294.7	294.7	75	4.98	1.88	1.88
31 Aug	4 Y K	290.8	293.3	91	6.50	2.63	3.25
4 Sep	4 Y P	286.0	283.7	98	5.64	2.20	2.65
4 Sep	4 Y V	295.4	298.7	80	8.96	4.02	4.14
16 Sep	4 Y C	281.8	281.9	81	4.38	1.61	1.37
27 Sep	4 Y N	295.1	294.1	73	7.31	3.06	2.27
4 Oct	4 Y N	290.5	294.9	79	5.42	2.08	2.46
8 Oct	4 Y V	294.4	295.3	65	5.58	2.16	1.50
13 Oct	4 Y J	286.4	285.9	87	5.26	2.00	1.87
18 Oct	4 Y N	294.1	294.9	73	4.90	1.84	2.08
23 Oct	4 Y P	279.3	282.1	71	3.18	1.08	0.83
25 Oct	4 Y N	290.2	294.3	73	5.73	2.24	2.12
27 Nov	4 Y P	282.0	280.9	80	3.67	1.27	0.91
3 Dec	4 Y V	291.8	291.3	66	5.00	1.87	1.58
5 Dec	4 Y E	292.5	292.2	72	6.43	2.59	2.09
26 Dec	4 Y E	291.2	292.7	85	6.30	2.53	2.01
26 Jan	4 Y J	280.6	281.5	75	3.07	1.05	1.12

30–50° on both the hemispheres for four seasons are derived by Nieman (1977).

Some of these mean profiles of the temperature and relative humidity are shown in Fig. 2. It is interesting to note that in all cases the mean relative humidity profiles are very similar and decrease from surface up to ~200 mb almost linearly as a function of height. These data are heavily weighted by

TABLE 1b. As in Table 1a but with radiosonde data from island stations.

Day	Station	T_{IRIS}	T_{ISLE}	RH(p_0)	l_{IRIS}	w_{IRIS}	w_{ISLE}
20 Apr	78118 Turks	298.3	296.9	72	5.56	2.16	2.23
20 Apr	78486 S. Domingo	298.4	296.8	75	6.03	2.41	2.53
23 Apr	78970 Trinidad	299.3	298.1	90	7.52	3.18	3.57
25 Apr	78526 San Juan	298.0	299.3	67	7.13	2.97	3.05
29 Apr	91245 Wake	298.7	301.8	71	7.38	3.10	3.05
3 May	78806 Howard	296.9	297.7	93	10.46	5.00	5.07
8 May	91165 Lihue	297.2	300.2	72	6.08	2.43	2.57
11 May	78118 Turks	297.4	297.1	66	6.14	2.45	2.57
13 May	91285 Hilo	299.7	298.8	69	6.16	2.46	2.61
18 May	78526 San Juan	298.9	298.8	71	7.30	3.06	2.71
18 May	78988 Plesman	299.5	300.9	81	10.30	4.90	5.15
21 May	78970 Trinidad	298.7	300.5	83	10.00	4.70	4.17

TABLE 2. Mean line strength $\bar{l}$ and total precipitable water $\bar{w}$ as a function of the surface temperature T_s for different mean atmospheres.

	T_s (K)	$\bar{l}$ (°C)	$\bar{w}$ (g cm^{-2})
Tropics 10°N–10°S	298.9	9.6	4.3
Subtropics summer 10–30°N	298.6	9.5	4.3
Subtropics winter 10–30°N	291.9	7.2	3.2
Temperate summer 30–50°N	292.8	7.5	3.4
Temperate winter 30–50°N	273.9	3.1	1.0

land stations. However, from a limited sample of ship and island radiosonde data taken over North Atlantic and North Pacific (see Tables 1a and 1b) during the year 1970, we find a mean value of 81% relative humidity near the surface, and a similar decrease with height. For this reason we have adopted, for the radiative transfer simulations, one relative humidity profile, as shown in Fig. 2, for all oceanic regions.

From the synthetic spectral data, calculated with the simulation program, using the above mean atmospheric conditions, we have obtained the brightness temperature at eight peaks (see Fig. 1) in the 9 μm region—1127.7, 1140.2, 1158.3, 1168.0, 1181.9, 1193.1, 1202.8, 1232.0 cm^{-1}—and averaged them. Similarly, the brightness temperature at eight adjacent valleys—1136.1, 1149.9, 1165.3, 1173.6, 1186.1, 1197.2, 1211.2, 1225.1 cm^{-1}—are averaged. The difference between these two averages gives an estimate of the line strength representative of the spectral region 1125–1235 cm^{-1}. The corresponding brightness temperature $\bar{l}$ is shown in Table 2 for the various mean atmospheres. In Fig. 3 we show graphically the relationship between $\bar{l}$ and $\bar{w}$ for the various mean atmospheres. Further in Fig. 4 the dependence of $\bar{l}$ and $\bar{w}$ on the surface temperature T_s is shown. Both $\bar{w}$ and $\bar{l}$, as seen from the Fig. 4, increase with T_s in an analogous fashion. The increase in $\bar{w}$ may be readily understood in terms of the moisture holding capacity of the warmer atmospheres. The growth of $\bar{l}$ needs some explanation.

The behavior of $\bar{l}$ can be explained with the help of the weighting functions, $d\tau/dz$ associated with the valleys and peaks in the spectra.

The weighting function associated with the valleys has a distinct maximum which moves down as the total water content decreases. The weighting function at the peaks, on the other hand, does not show any maximum. The strength of the line l is thus crucially dependent on the superincumbent water vapor in the troposphere.

Over the oceans, the relationships shown in Figs. 3 and 4 represent basic correlations among the surface temperature T_s, w and l. These correlations may be broken when atmospheric temperature and water vapor profiles are perturbed from mean conditions. Such perturbations are present over the global oceans.

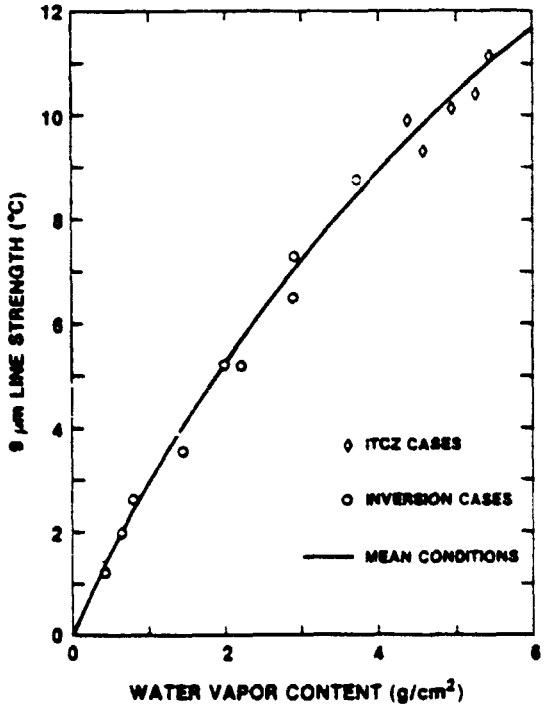


FIG. 3. The relationship between the 9 μm line strength and the total precipitable water for mean atmospheric conditions, for inversion conditions (from Tables 3a–3c), and ITCZ cases (see Table 4).

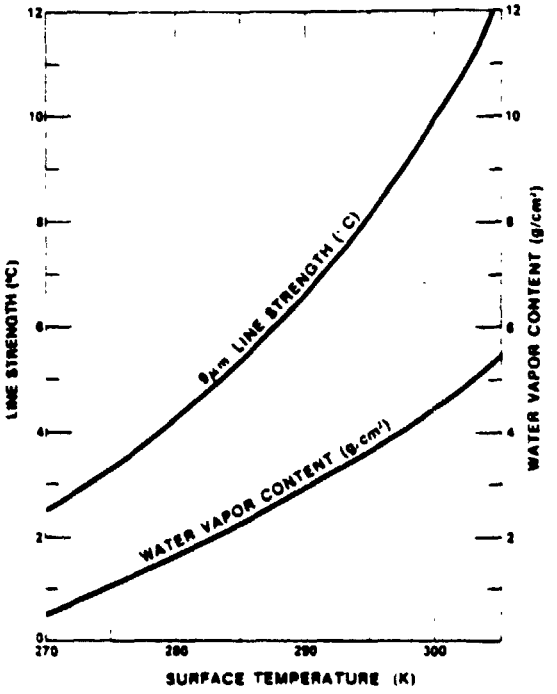


FIG. 4. Dependence of the 9 μm mean line strength $\bar{l}$ and the total water vapor content $\bar{w}$ on the surface temperature T_s over the oceans.

In the following discussion these climatologically significant perturbations are examined in some detail.

b. Trade wind inversion

Trade wind inversion is a climatological feature associated with oceanic subtropical anticyclones. In the Northern Hemisphere the axis of these anticyclones tilts from NE to SW (see, e.g., Haurwitz and Austin, 1944) and thereby the vertical motion associated with these anticyclones has an organized pattern. On the NE side there is a pronounced subsidence, while on the SW the vertical motion is weak. This organized motion of the subtropical anticyclones produces a strong inversion on the NE side. The observations of the Meteor Oceanographic Expedition (Ficker, 1936) and the Indian Ocean Expedition (Ramage, 1966, 1971) reveal the general nature of the trade wind inversion over the North and South Atlantic, and the Arabian Sea. Riehl (1954) shows the NE–SW cross sections of the trade inversion regime derived from radiosonde data gathered during July–October 1945 over the North Pacific Ocean. The more recent experiments related to GATE, in particular BOMEX and ATEX (Augstein *et al.*, 1973; Dunckel *et al.*, 1974), reinforce the observations of these earlier studies.

In the trade wind inversion the stronger the subsidence the lower the height h of the inversion and the larger the temperature increase ΔT from bottom to top of the inversion due to adiabatic warming (Riehl, 1954). The change in relative humidity ΔRH is primarily due to the capping of convection by the inversion, which limits the height of the lower humid layer, while above the inversion sinking of dry air and dynamic warming lower the relative humidity. In this manner a correlation between the temperature profile and the water vapor profile in the lowest layers of the troposphere is developed in oceanic regions where the inversion conditions prevail.

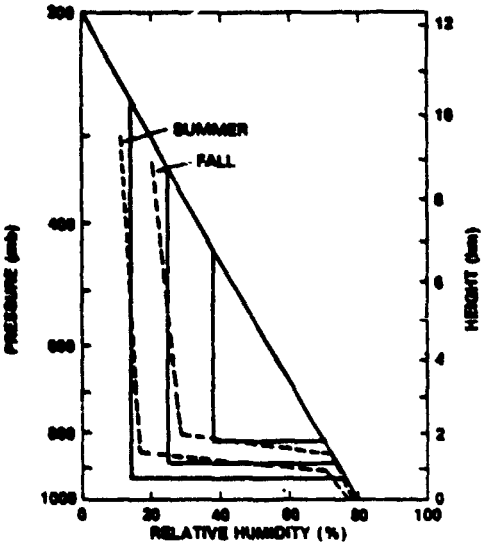


FIG. 5. Adopted mean relative humidity profile (shown also in Fig. 2) and the models of the relative humidity profile for three conditions (see Table 3) of the inversion. The average profiles for June to August and September to October from the Ship 4 Y N radiosonde data (Table 1a) are also shown (dashed lines).

Utilizing these climatological correlations, we have empirically modeled the characteristics of the trade wind inversion layer over the oceans, with one parameter (ΔRH) which is coupled to h and ΔT. In Table 3 three values of the inversion parameter ΔRH and the associated h and ΔT used in our calculations are listed. The adopted mean humidity profile shown in Fig. 2 is suitably modified to represent the three inversion conditions. Above the level of inversion, the relative humidity profile is assumed to have a constant value until it meets the undisturbed profile as shown in Fig. 5.

This model is consistent with some recent radiosonde data, referred to in Table 1a, of the ship station 4 Y N (30°N, 140°W) in the North Pacific where the trade wind regime is strong during summer. The average relative humidity profile of the ship 4 Y N cases, listed in Table 1a, for the months of June–August and September–October are shown in Fig. 5 for comparison. These profiles support the inversion model.

Applying these three inversion characteristics to the tropical, midlatitude summer and winter atmospheres we have calculated I, the strength of the lines in the 9 μm region. In Table 3 the strength of I and the total precipitable water w for the three inversion conditions are listed. These tables show the dependence of I on the boundary-layer parameters ΔRH, h and ΔT in a joint fashion.

TABLE 3. Computed 9 μm line strength corresponding to three inversions for tropics (a), midlatitude summer (b) and midlatitude winter (c).

h (km)	ΔT (°C)	ΔRH (%)	I (°C)	w (g cm^{-2})	$(\bar{w} - w)/\bar{w}$
		(a) Tropical mean atmosphere (T_s = 298.9 K, $\bar{I}$ = 9.62°C, $\bar{w}$ = 4.28 g cm^{-2})			
1.70	2	32	8.75	3.71	0.13
1.15	4	48	7.30	2.91	0.32
0.75	6	62	5.22	1.98	0.54
		(b) Midlatitude summer (T_s = 292.6 K, $\bar{I}$ = 7.48°C, $\bar{w}$ = 3.35 g cm^{-2})			
1.70	2	32	6.51	2.89	0.14
1.15	4	48	5.18	2.22	0.34
0.75	6	62	3.54	1.46	0.56
		(c) Midlatitude winter (T_s = 273.8 K, $\bar{I}$ = 3.10°C, $\bar{w}$ = 0.96 g cm^{-2})			
1.70	2	32	2.62	0.81	0.16
1.15	4	48	1.97	0.65	0.35
0.75	6	62	1.22	0.43	0.57

TABLE 4. Calculated 9 μm line strength, the total water content w, and the index $(\bar{w} - w)/\bar{w}$ for the ITCZ cases.

Source	T_s (K)	l (°C)	w (g cm^{-2})	$\bar{w}$ (g cm^{-2})	$(\bar{w} - w)/\bar{w}$
Jordan (1958)	299.15	9.92	4.38	4.34	−0.01
Augstein *et al.* (1974)	299.15	10.13	4.94	4.33	−0.14
Estoque (1975)	297.65	10.39	5.26	4.09	−0.28
Estoque and Douglas (1978)	298.65	9.31	4.58	4.25	−0.08
Godbole and Ghosh (1975)	299.65	11.13	5.43	4.41	−0.23

For a given surface temperature the total water content in each one of these models varies as the boundary-layer parameters change. The water content is at maximum when the inversion is absent and least when the inversion is at the lowest levels as shown in the tables. As the total water is an integral property reflecting the inversion characteristics we may relate it to l.

In Fig. 3 the values of l and w taken from Table 3 are plotted. It can be seen from this figure that all the points are very close to the curve corresponding to mean atmospheric conditions. As a consequence of this result it is possible to infer w from l in the presence of the inversion condition.

We have examined the sensitivity of l separately to the temperature profile and the water vapor profile. Such an examination shows that ~90% of the change in l is due to the variation in the water vapor profile, while the temperature profile accounts for ~10%. This analysis shows the overwhelming importance of the water vapor profile. Any inference of the temperature profile from l thus has to be deduced by way of statistical correlation with respect to water vapor profile.

The surface temperature T_s on the ocean can be measured independently from the 11 μm IRIS window data. Given the surface temperature we can estimate the total water content $\bar{w}$ corresponding to mean atmospheric conditions from Fig. 4. Then we can formulate a parameter $(\bar{w} - w)/\bar{w}$ to indicate the strength of the inversion.

c. *Intertropical convergence zone*

Another well known phenomenon over the tropical oceans is the intertropical convergence zone (ITCZ). This is the region where the low-level winds converge to produce a rising motion which carries water vapor upward and produces above average humidity conditions aloft. The temperature inversion above the boundary layer in the ITCZ is generally absent or weak (Estoque, 1975). Thus the temperature profile in the ITCZ areas does not show a significant departure from mean atmospheric conditions. However, the water vapor profile reflects the effects of rising motion, resulting in a significant increase in total water content. For this reason we have not developed a simple model of the coupling of the temperature and water vapor profiles in the ITCZ.

Observations of the temperature and humidity corresponding to the ITCZ conditions are given in the studies of Augstein *et al.* (1974), Estoque (1975), Godbole and Ghosh (1975) and Estoque and Douglas (1978). The relative humidity profile in the ITCZ shows a significantly larger value than the climatological average between about 800 and 600 mb. In this region of the atmosphere a significant deficit in humidity is noticed when inversion conditions prevail. Utilizing the radiative transfer program we have calculated the 9 μm line strength for all these ITCZ cases which are shown in Table 4. In the same table we give also the results for an average profile for the West Indies region (Jordan, 1958).

The values of l and the corresponding w for the ITCZ cases are also plotted in Fig. 3. This figure clearly shows that the ITCZ l values conform reasonably well with the general l vs w relationship that was derived from mean atmospheric conditions and trade wind inversion models. In all the ITCZ cases the parameter $(\bar{w} - w)/\bar{w}$ is negative (see Table 4), suggesting that the water vapor content and the depth of the humid layer exceed average conditions.

We can conclude from the preceding discussion of spectral simulations that the relationship between l and w, shown in Fig. 3, is applicable for the various atmospheric conditions we have considered. The parameter $(\bar{w} - w)/\bar{w}$ in a general fashion reflects excess or deficit of water vapor content in the atmosphere with respect to some mean conditions. For positive values this parameter, together with some statistical information, can be related to the strength of the trade wind inversion. Negative values of this parameter indicate the presence of deep moisture convection such as that produced by ITCZ.

d. *Comparison with radiosonde data*

The strength of the water vapor lines l as shown in the previous section has a quasi-linear dependence on the total water. This is not surprising as l is a water vapor spectral feature in the window region.

To further validate this result we have used ground truth data obtained from ship and island radiosonde stations.

The Nimbus 4 IRIS obtained spectral measure-

ments for a period of about 9 months from April–December 1970. The satellite 9 μm line strength measurements, within $\pm 1°$ latitude and longitude from the stations, are used to estimate the total water content according to the relationship shown by solid line in Fig. 3. The cloud contamination was avoided by accepting only the data in which the surface temperature reported by the ship (or island) was within ~1.5°C with respect to satellite derived SST. In this fashion we were able to gather 52 cases for comparison, shown in Tables 1a and b. The radiosonde data that went into this sample contained atmospheric conditions ranging from inversion to convectively active situations. This comparison, presented in Fig. 6, reveals that the total water content estimated from the 9 μm line strength agrees with the radiosonde measurements within ~±15%.

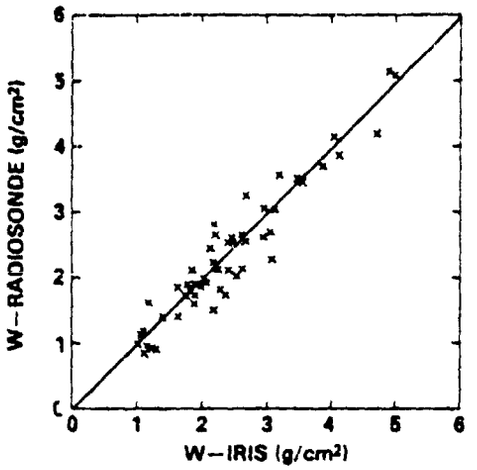


FIG. 6. Comparison between the water vapor content derived from the 9 μm line strength and the water vapor content obtained from radiosonde data.

4. Information on global oceans derived from Nimbus 4 IRIS

The field of view of IRIS was about 95 km in diameter and this instrument had only subsatellite viewing geometry. As the successive orbits of the satellite are ~26° longitude apart and as some data are cloud contaminated, it was not possible to construct a global map of the line strength I for each day. It was found necessary to combine ~90 days of IRIS data to get a satisfactory global map. For this reason we have divided the IRIS data into three periods: April–June, July–September and October–December. The cloud contamination is eliminated with the help of the brightness temperature in the 11 μm window region. It is assumed that when the window temperature exceeds 290 K in the 20°N–20°S latitude belt over the oceans clouds are absent. Similarly, thresholds of 285 and 280 K are applied to the belts 20–30° and 30–50°, on either side of the equator, respectively. Thus the data used in this study emphasize the clear-sky conditions and do not represent cloudy areas. In addition, the composite of data may not lead to a fair spatial and seasonal average.

From the relationship shown in Fig. 3 between I and w we have constructed three maps of total water vapor distribution over the global oceans from about 50°N to 40°S. As an example, only the map for the April–June period is presented in Fig. 7. The full sets of maps derived from the IRIS data can be found in Prabhakara *et al.* (1978).

We are not able to compare this information with a global map of total water derived from conventional data at this time. However, Grody (1978) has derived such global maps of water vapor using the data from a scanning microwave spectrometer (SCAMS) that was flown on Nimbus 6. We have compared one of these maps of total water content for the period 18 August–4 September 1975 with the map constructed from IRIS data for the three-month period July–September 1970. The distribution of the total water vapor in both these maps is quite similar. This result supports our method for remote sensing of total water vapor over the global oceans.

Using the 11 μm IRIS window data and utilizing the split window technique described by Prabhakara *et al.* (1974) for each one of these periods, a sea surface temperature (SST) map is derived. As an example the SST map for April–June is shown in Fig. 8.

These maps of w and SST essentially constitute the basic observational information for the subsequent discussion. The data between the longitude region of ~100–120°E are missing due to some telemetry limitation. As a consequence information over this part of the Indian Ocean is not contained in our analysis. With the help of the surface temperature maps and the corresponding maps of the total water content w, we can construct maps of $(\tilde{w} - w)/\tilde{w}$ for the three periods covered by the IRIS observation. In Figs. 9, 10 and 11 three maps of this index for the periods April–June, July–September and October–December are shown.

In a general fashion these maps reflect the gross structure of the boundary layer over the oceans for different seasons. The trade wind inversion associated with the subtropical anticyclone over the North and South Atlantic and Pacific oceans are clearly delineated by the positive values of $(\tilde{w} - w)/\tilde{w}$. Over the Arabian sea strong inversion conditions are revealed during the period April–June as well as during October–December. Apparently, the inversion conditions are absent in the monsoon period, July–September, giving way to convective activity. A vast region of subsidence in the equatorial Pacific is revealed with some seasonal changes in strength and position. This interesting phenomenon in the equatorial Pacific is known from ship observations (see Riehl, 1954).

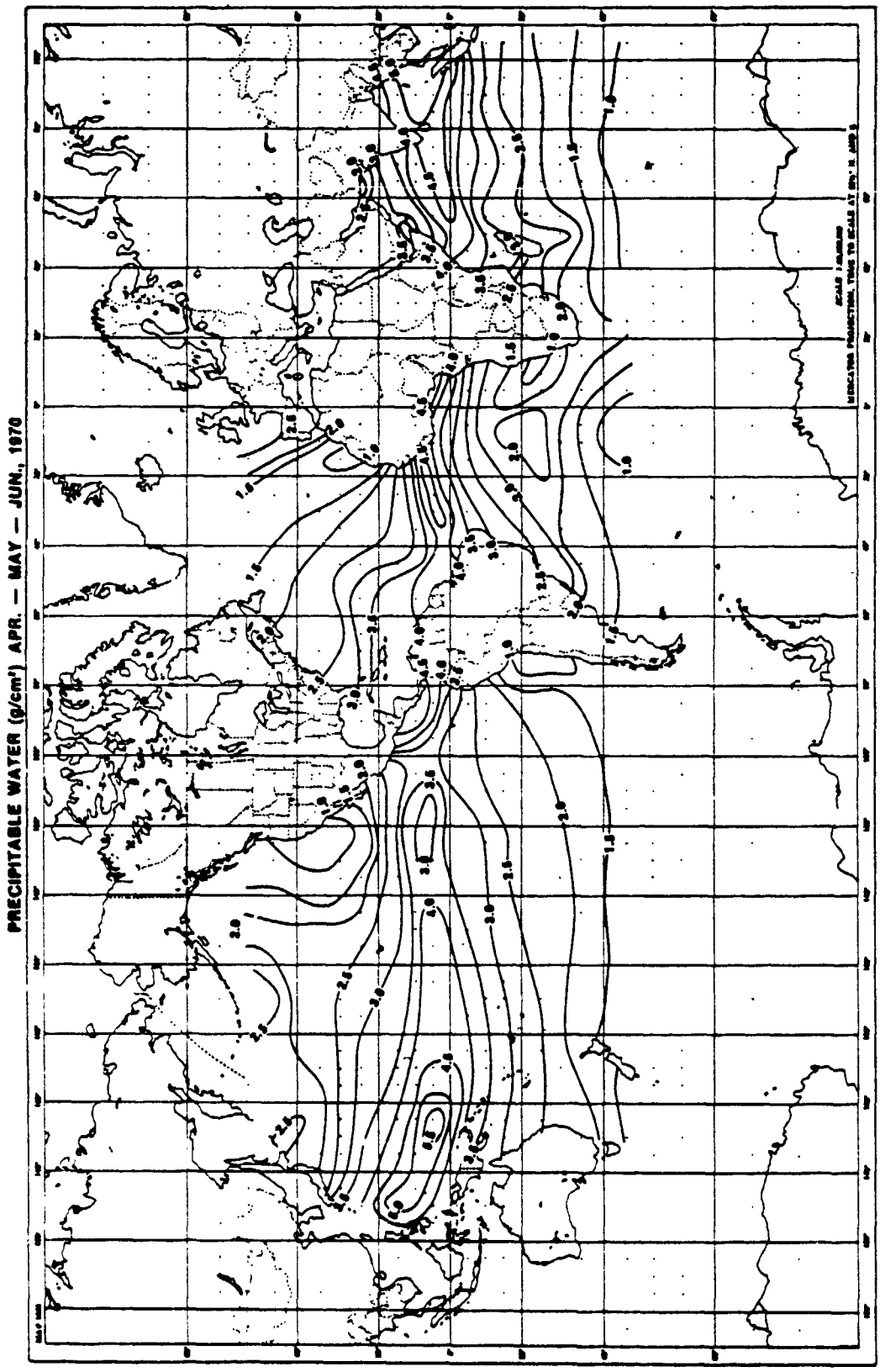


FIG. 7. Distribution of the total water vapor content (g cm^{-2}) over the global oceans (50°N–40°S) derived from the Nimbus 4 IRIS data over the period April, May and June, 1970.

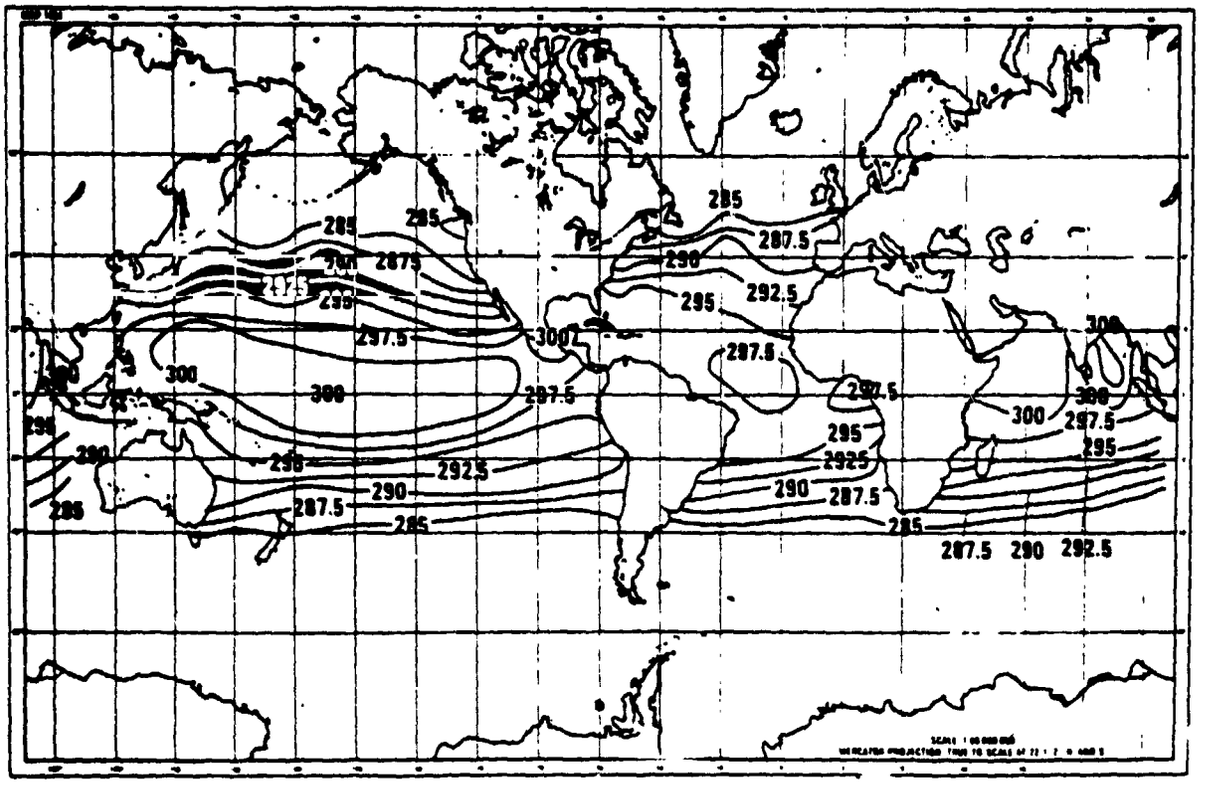


FIG. 8. Distribution of the surface temperature over the global oceans (50°N–40°S) derived from the 11 μm window region observations of the Nimbus 4 IRIS for the period April, May and June, 1970.

The regions with negative values of $(\tilde{w} - w)/\tilde{w}$, indicated by stippling, show the presence of convectively active areas as explained earlier. Such regions are present where the ITCZ generally prevails. Convectively active areas are also revealed along the course of the Gulf Stream and Kuroshio currents. In the Southern Hemisphere a significantly large convectively active zone appears in the October–December period. This zone, which is about 10° latitude in width, spans from ~20°S, near the east coast of Australia, to ~35°S near Chile in South America.

The significant features of the boundary layer characteristics over the oceans revealed by the remote sensing technique, are in general agreement with the known climatology.

The inferred information on the boundary-layer structure depends on one important assumption, i.e., the mean water vapor profile over the global oceans can be modeled in a simple fashion. Essentially it is assumed that one mean relative humidity profile showing a monotonic decrease from the surface to 200 mb, represents the average condition applicable to all latitudes for different seasons. Departures from such mean conditions are then associated with the stable condition of inversion or the convectively active state. For this reason the results obtained in this study are somewhat model-dependent. Particularly if the mean relative humidity profile changes significantly as a function of latitude and season, the patterns shown in Figs. 9, 10 and 11 will be distorted. From the climatological data of Nieman it appears that the mean relative humidity profile changes within ~10%, reflecting a 10% error in $\tilde{w}$. The error introduced in $(\tilde{w} - w)/\tilde{w}$ from such a source has the same magnitude, i.e., 0.1. Since the range of $(\tilde{w} - w)/\tilde{w}$ shown in the maps is much larger than 0.1 we are able to resolve the patterns in a meaningful way.

The random errors in the remotely sensed SST and w are considerably smoothed in the three-month averaging process. As a consequence the error introduced in $(\tilde{w} - w)/\tilde{w}$ from this source is significantly less than 0.1.

5. Summary and conclusions

Utilizing the Nimbus 4 IRIS spectral measurements we have remotely sensed two pieces of information over the global ocean for three different seasons. The two pieces of information are 1) the corrected sea surface temperature (SST) obtained from the differential absorption in the 11 and 13 μm region, and 2) the precipitable water vapor w ob-

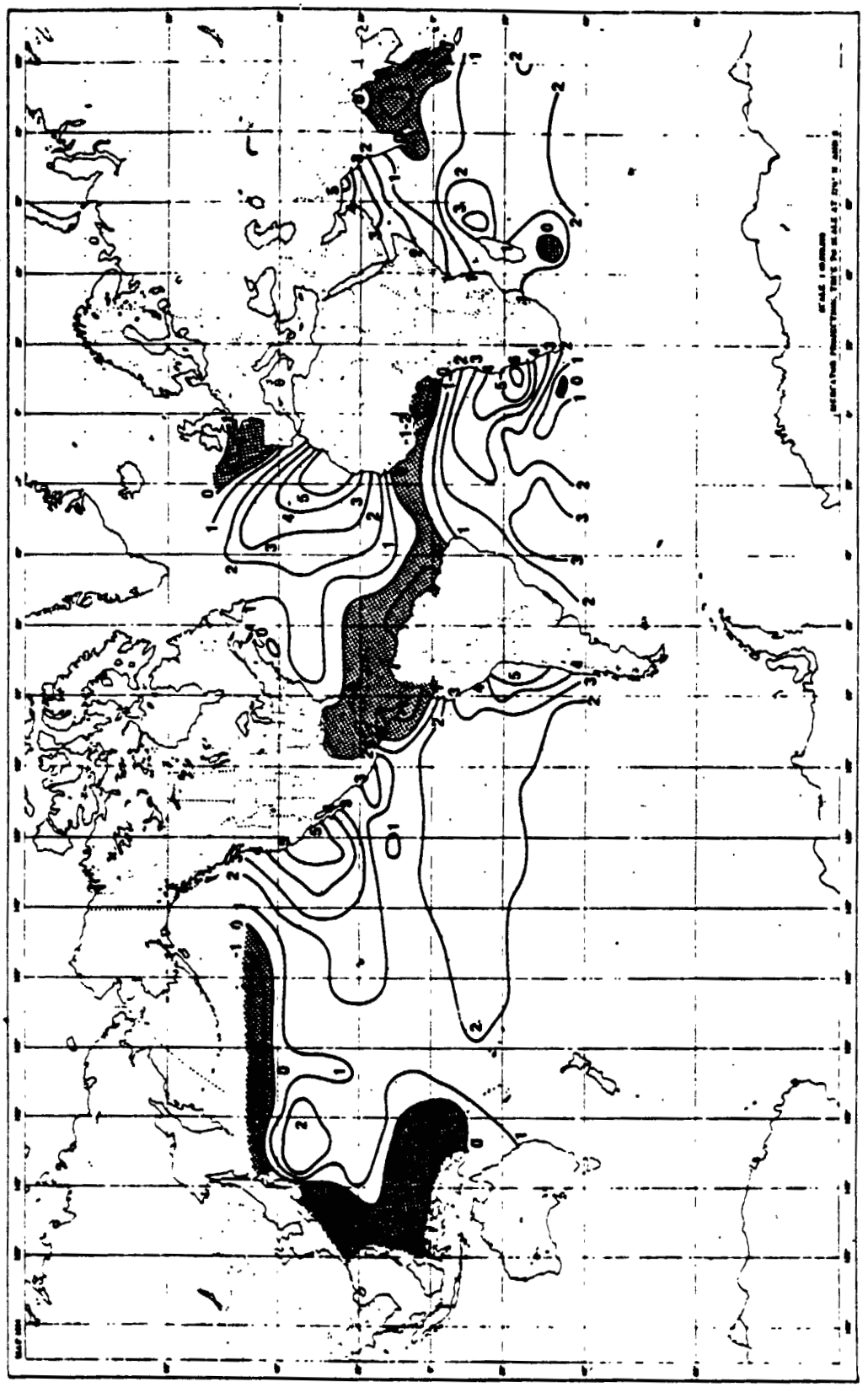


FIG. 9. Distribution of the index, $(\bar{w} - w\bar{w} \times 10$, over the global oceans derived from the Nimbus 4 IRIS data for the period April, May and June, 1970. Positive values of this index indicate inversion and negative values convection (shown by stippling).

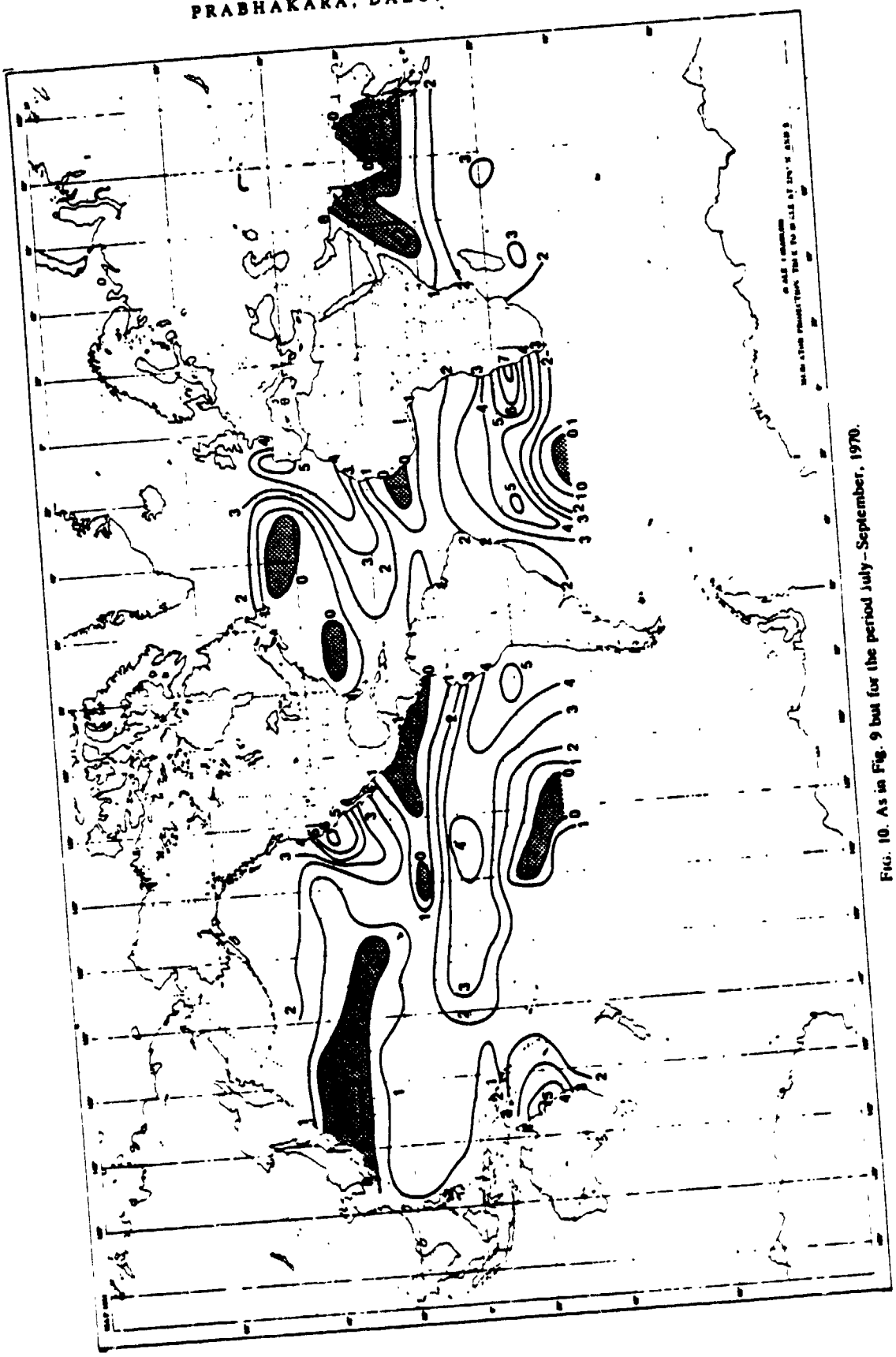


FIG. 10. As in Fig. 9 but for the period July–September, 1970.

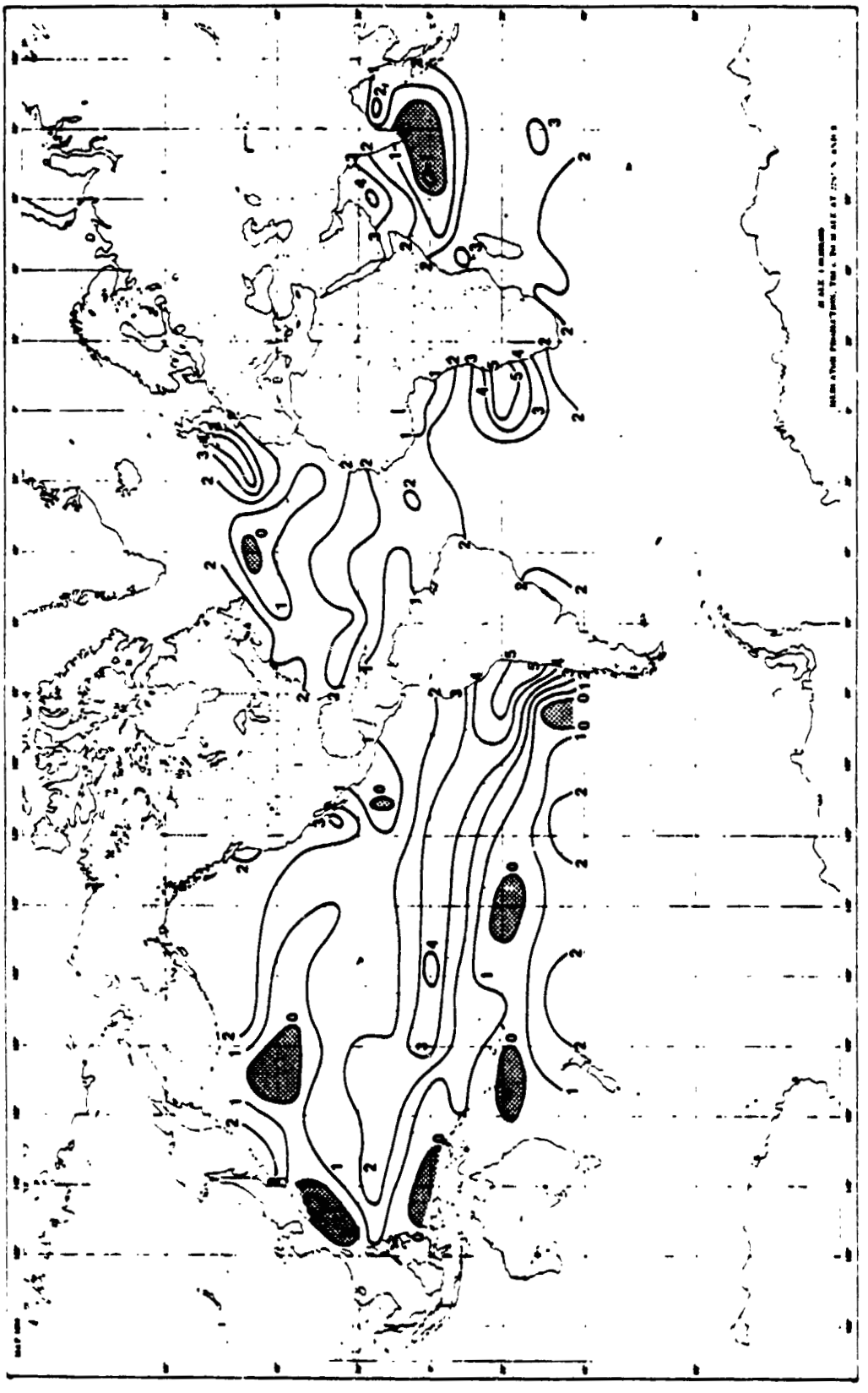


FIG. 11. As in Fig. 9 but for the period October–December, 1970.

tained from the depth of the water vapor lines in the 8–9 μm region.

The remotely sensed SST and w over the oceans can be used to infer the stratification of the atmosphere in the first few kilometers above the surface. This is possible because on the average the vertical distribution of water vapor in the atmosphere above water bodies can be represented with a simple model. Then some average value of precipitable water $\tilde{w}$ can be associated with a given surface temperature. The departure of the measured w from the corresponding $\tilde{w}$ relates to the atmospheric stratification in the boundary layer. For instance, w exceeds $\tilde{w}$ when convective conditions, such as the ITCZ, are present. When w is less than $\tilde{w}$ stable conditions of inversions prevail. Such an excess or deficit of w has been derived from IRIS data for three seasons of 1970. The gross characteristics of the boundary layer over the oceans are seen in these maps.

The trade wind inversion over the oceans is an integral part of the tropical circulation. The studies of Malkus (1956), Mak (1976) and Ogura *et al.* (1977) bring out the importance of this large-scale atmospheric phenomenon. The investigation of Reiter (1978) reveals an intimate connection between the sea surface temperature anomalies and the strength of the trade winds. Several studies (see, e.g., Namias, 1978) have revealed the significance of SST anomalies in producing seasonal climate change. From these studies it is clear that the information we are deriving from remote measurements over the global oceans on a seasonal basis, although crude, can be valuable in understanding the ocean atmosphere interaction. This knowledge is indispensable in developing seasonal climate prediction models.

The information, obtained with the help of satellite measurements, which we have presented here, is biased toward clear-sky conditions. Microwave remote sensing is capable of probing through clouds when they are not precipitating. Thus when the microwave technique to sense the sea surface temperature is developed, it appears that the bias introduced by clouds can be reduced.

REFERENCES

Augstein, E., H. Schmidt and F. Ostapoff, 1974: The vertical structure of the atmospheric boundary layer in undisturbed trade wind over the Atlantic Ocean. *Bound.-Layer Meteor.*, 6, 129–150.

——, H. Riehl, F. Ostapoff and V. Wagner, 1973: Mass and energy transports in an undisturbed Atlantic trade-wind flow. *Mon. Wea. Rev.* 101, 101–111.

Bignell, K. J., 1970: The water vapor infrared continuum. *Quart. J. Roy. Meteor. Soc.*, 96, 390–403.

Burch, D. E., 1970: Investigation of the absorption of infrared radiation by atmospheric gases. Publ. 4-4784, Philco-Ford Corp., Aeronutronic Div., Newport Beach, CA, 27 pp.

Dunckel, M., L. Hasse, L. Krügermeyer and D. Schriever, 1974: Turbulent fluxes of momentum, heat and water vapor in the atmospheric surface layer at sea during ATEX. *Bound.-Layer Meteor.*, 6, 81–106.

Estoque, M. A., 1975: Structure of the mid-oceanic intertropical convergence zone. *J. Meteor. Soc. Japan*, 53, 317–321.

——, and M. Douglas, 1978: Structure of the intertropical convergence zone over the GATE Area. *Tellus*, 30, 55–61.

Ficker, H. von, 1936: Die Passatinversion. *Veröff. Meteor. Inst.* Universität Berlin, 1, No. 4, 33 pp.

Godbole, R. V., and S. K. Ghosh, 1975: The structure of the intertropical convergence zone and equatorial westerlies during MONEX-1973. *Tellus*, 27, 123–131.

Grody, N. C., 1978: High resolution passive microwave satellites, applications to synoptic meteorology and climatology. *Application Review Panel Report*, D. H. [illegible] and P. W. Rosenkranz, Eds., MIT Research Lab. of Electronics, 6.1–6.40.

Hanel, R. A., B. J. Conrath, V. G. Kunde, C. Prabhakara, I. Revah, V. V. Salomonson and G. Wolford, 1972: The Nimbus 4 infrared spectroscopy experiment, 1. Calibrated thermal emission spectra. *J. Geophys. Res.*, 77, 2629–2641.

Haurwitz, B., and J. M. Austin, 1974: *Climatology*. McGraw-Hill, 410 pp.

Jordan, C. L., 1958: Mean soundings for the West Indies area. *J. Meteor.*, 15, 91–97.

Kunde, V. G., and W. C. Maguire, 1974: Direct integration transmittance model. *J. Quant. Spectrosc. Radiat. Transfer*, 14, 803–817.

Mak, M. K., 1976: A model study of the downstream variation of the lower trade wind circulation. *Tellus*, 28, 97–107.

Malkus, J. S., 1956: On the maintenance of the trade winds. *Tellus*, 8, 335–350.

Manabe, S., and R. T. Wetherald, 1967: Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.*, 24, 241–259.

Namias, J., 1978: Multiple causes of the North American abnormal winter 1976–77. *Mon. Wea. Rev.*, 106, 279–295.

Nieman, R. A., 1977: A comparison of radiosonde temperature and humidity profile data bases. CSC/TM-77/6133, Contract NAS 5-11999, Computer Sciences Corporation, Silver Spring, MD, 48 pp.

Ogura, Y., J. Russel and H. R. Cho, 1977: A semi-empirical model of the trade-wind inversion. *J. Meteor. Soc. Japan*, 55, 209–221.

Prabhakara, C., G. Dalu and V. G. Kunde, 1974: Estimation of sea surface temperature from remote sensing in the 11 to 13 μm window region. *J. Geophys. Res.*, 79, 5039–5044.

——, ——, R. C. Lo and N. R. Nath, 1978: Inference of the boundary layer structure over the oceans from satellite infrared measurements. NASA Tech. Memo. 79653, NASA-Goddard Space Flight Center, Greenbelt, MD, 55 pp.

Ramage, C. S., 1966: The summer atmospheric circulation over the Arabian Sea. *J. Atmos. Sci.*, 23, 144–150.

——, 1971. *Monsoon Meteorology*. Academic Press, 296 pp.

Reiter, E. R., 1978: The interannual variability of the ocean-atmosphere system. *J. Atmos. Sci.*, 35, 349–370.

Riehl, H., 1954: *Tropical Meteorology*. McGraw-Hill, 392 pp.

Smith, W. L., 1969: A polynomial representation of carbon dioxide and water vapor transmission. ESSA Tech. Rep. NESC-47, National Environmental Satellite Center, Washington DC, 20 pp.

Telegadas, K., and J. London, 1954: A physical model of Northern Hemisphere troposphere for winter and summer. Sci. Rep. No. 1, Contract AF19(122)-165, Research Division, College of Engineering, New York University, 55 pp.

N81-21701

Paper 126

NASA Technical Memorandum 79653 September 1978

INFERENCE OF THE BOUNDARY LAYER STRUCTURE OVER THE OCEANS FROM SATELLITE INFRARED MEASUREMENTS

C. Prabhakara, G. Dalu*

Laboratory for Atmospheric Sciences (GLAS)

NASA/Goddard Space Flight Center

Greenbelt, Md. 20771

R. Lo, N. R. Nath

Computer Sciences Corporation

Silver Spring, Md. 20910

ABSTRACT

Remote infrared spectral measurements in the 8–13 μm window region, at a resolution of about 3 cm^{-1}, contain useful information about the water vapor and temperature stratification of the atmosphere within the first few kilometers above the water surface. Two pieces of information are retrieved from the spectral measurements: (1) precipitable water vapor in the atmosphere, from the depth of the line structure between 8 and 9 μm due to water vapor lines; and (2) sea surface temperature, from the variation of brightness temperature between 11 and 13 μm. Together, these two pieces of information can signify either the presence of a deep moist convective layer or the prevalence of stable conditions, such as caused by temperature inversions, which inhibit moist convection.

A simple infrared radiative transfer model of the 9 μm water vapor lines is developed to validate the method. With the help of this model and the Nimbus 4 Infrared Interferometer Spectrometer (IRIS) data, a gross picture of the planetary boundary layer for different seasons over the global oceans is deduced. The important regions of the trade wind inversion and the intertropical convergence zones over all the oceans are clearly identified with this method. The derived information is in reasonable agreement with some observed climatological patterns over the oceans.

*NAS/NRC Research Associate on leave from C.N.R., Istituto di Fisica della Atmosfera, Rome, Italy.

APPENDIX A

AUTHOR INDEX

AUTHOR INDEX (continued)

AUTHOR INDEX (continued)

AUTHOR INDEX (continued)

AUTHOR INDEX (continued)

AUTHOR INDEX (continued)

AUTHOR INDEX (continued)

AUTHOR INDEX (continued)

APPENDIX B

SUBJECT INDEX

PRECEDING PAGE BLANK NOT FILMED

SUBJECT INDEX (continued)

SUBJECT INDEX (continued)

SUBJECT INDEX (continued)

SUBJECT INDEX (continued)