NASA CONTRACTOR
REPORT

NASA CR-61345

DEVELOPMENT OF A GLOBAL CLOUD MODEL FOR SIMULATING EARTH-VIEWING SPACE MISSIONS

By James R. Greaves, David B. Spiegler, and
James H. Willand

Allied Research Associates, Inc.
Virginia Road
Concord, Massachusetts

April 7, 1971

Prepared for

NASA-GEORGE C. MARSHALL SPACE FLIGHT CENTER
Marshall Space Flight Center, Alabama 35812

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-61345		

4. Title and Subtitle	5. Report Date
DEVELOPMENT OF A GLOBAL CLOUD MODEL FOR SIMULATING EARTH-VIEWING SPACE MISSIONS	April 7, 1971
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
James R. Greaves, David B. Spiegler, and James H. Willand	9G78-F
	10. Work Unit No.
9. Performing Organization Name and Address	11. Contract or Grant No.
Allied Research Associates, Inc. Virginia Road Concord, Massachusetts	NAS8-25812
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	CONTRACTOR REPORT
NASA Washington, DC 20546	14. Sponsoring Agency Code

15. Supplementary Notes Technical Coordinator: S. C. Brown, Aerospace Environment Division, Aero-Astrodynamics Laboratory, Marshall Space Flight Center, Alabama.

16. Abstract: This study was made to improve and expand the existing cloud model described in NASA CR-61226, "World-Wide Cloud Cover Distributions for Use in Computer Simulations," and to add to the model other statistical data of importance to electromagnetic energy propagation through the atmosphere.

An intensive data extraction procedure was carried out for six of the original 29 regions. Using simultaneous cloud-amount observations, a relationship was developed between ground and satellite-derived cloud frequency distributions. This relationship was used to introduce internal consistency between the unconditional and conditional statistics and between the temporal and spatial conditionals. As a byproduct of this procedure, the conditional data as listed in the Revised Data Bank are now stratified by month rather than by six-month seasons. A separate listing of the 1300 local-time unconditionals as modified to simulate a satellite-observed distribution was added to the data bank.

The intensive data extraction procedure made possible the tabulation of temporal conditionals in 24-hour intervals from 24 hours to 240 hours, and the tabulation of spatial conditionals in 60-mile increments from 60 miles to 600 miles. Using these data, a Markov scaling technique was developed to scale the 24-hour and 200-mile conditional arrays to any other time or distance scale. The original linear scaling technique with its inherent "stuffing" procedure was thereby eliminated. The new data extraction procedure also made possible an evaluation of the cloud-amount homogeneity of the six selected regions. A new area enlarge procedure was developed which carries out a repetitive doubling of the representative area size for both the conditional and unconditional arrays. A second technique which can be used to increase or decrease the area size for unconditional distributions is also provided.

Some preliminary indications were made of how varying spatial resolutions would affect the simulation results. Recommendations are made as to the types of cloud data which are now required to extend the simulation results to other types of sensors such as infrared and microwave radiometers and to other spatial resolutions.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Cloud Model, Temporal and Spatial Conditionals, Markov Scaling Technique, Area Enlarge Procedure	Unclassified - Unlimited E. D. Geissler Director, Aero-Astrodynamics Laboratory

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
UNCLASSIFIED	UNCLASSIFIED	141	$3.00

*For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151

FOREWORD

The research described in this report was performed by the Geophysics and Aerospace Division of Allied Research Associates, Inc., under sponsorship of the George C. Marshall Space Flight Center, Aerospace Environmental Division of the National Aeronautics and Space Administration, Contract No. NAS 8-25812.

Any study such as this requires the dedicated efforts of a number of scientists. Mr. Robert Smiley prepared the mathematical derivations found in Sections 4.1.1 and 4.2.2. Mr. James Barnes is responsible for Sections 3.2.1 and 3.3.3 and for selecting the six new data extraction areas and coordinating the data reading activities. The actual data extraction was carried out by Mr. Hans Ackerman through the cooperation of Mr. Harold Broderick of the National Environmental Satellite Service. Miss Mary Grace Fowler provided programming support for Sections 6 and 7. Mr. Frank Lewis, Chief of the Computer Systems Branch of the Techniques Development Laboratory of NOAA graciously made available the precipitable water tape for developing the water vapor frequency distributions in Section 7.

TABLE OF CONTENTS

		Page
SECTION 1	INTRODUCTION	1
SECTION 2	THE PREVIOUS STUDY	3
	2.1 The Data Base	4
	2.1.1 Cloud-Climatic Regions	4
	2.1.2 Unconditional Distributions	4
	2.1.3 Conditional Distributions	6
	2.1.4 Limitations	7
	2.2 Statistical Adjustment Techniques	9
	2.2.1 Time and Distance Scaling	9
	2.2.2 Area Adjustment	9
	2.2.3 Limitations	10
	2.3 Mission Simulations	11
	2.3.1 Monte Carlo Procedure	11
	2.3.2 Limitations	14
	2.4 Summary	14
SECTION 3	DEVELOPMENT OF REVISED DATA BANK	17
	3.1 New Ground-Observed Data	17
	3.2 New Satellite-Observed Data	17
	3.2.1 Description of Selected Cloud Regions	19
	3.2.2 Data Extraction Procedures	20
	3.2.3 Two-Reader Comparison	22
	3.3 Interval Data Consistency	24
	3.3.1 Correction for Different Data Sources	24
	3.3.2 Regional Homogeneity	29
	3.3.3 Seasonal Reversals	29
	3.4 Ground/Satellite Cloud Observation	34
	3.5 Description of Revised Data Bank	37

TABLE OF CONTENTS, contd.

		Page
SECTION 4	REFINING THE STATISTICAL ADJUSTMENT TECHNIQUES	41
	4.1 Time and Distance Scaling	41
	4.1.1 Markov Scaling	41
	4.1.2 A Test of Markov Scaling	42
	4.1.3 Diurnal Variations	54
	4.2 Area Adjustment	61
	4.2.1 Enlarging the Representative Area Size	61
	4.2.2 Area Adjustment of Unconditional Distributions	69
SECTION 5	SIMULATION PROCEDURES	77
SECTION 6	SEVERE WEATHER STATISTICS	83
	6.1 Thunderstorm Statistics	83
	6.2 Tornado and Severe Thunderstorm Statistics for the United States	91
	6.3 Tropical Cyclone Statistics	101
SECTION 7	DISTRIBUTION OF WATER VAPOR CONTENT	107
	7.1 Global Distributions - Annual	107
	7.2 North American Distribution - Monthly	111
	7.3 Frequency Distributions of Moisture for the United States	111
SECTION 8	CURRENT STATUS OF GLOBAL CLOUD MODEL	129
REFERENCES		131

LIST OF ILLUSTRATIONS

Figure		Page
2-1	Cloud Cover Regions	5
2-2	Monte Carlo Simulation Routine	12
2-3	Monte Carlo Simulation Results	13
3-1	New Data Extraction Areas	18
3-2	Data Extraction Templates	21
3-3	Two-Reader Cloud Amount Agreement	23
3-4	Regional Homogeneity; Regions 4 and 5	30
3-5	Regional Homogeneity; Regions 11 and 13	31
3-6	Regional Homogeneity; Regions 16 and 22	32
3-7	Computer Produced Unconditionals for Regions 13 and 22	33
3-8	Applications of Transfer Matrix	38
4-1	Decay of Spatial Conditionals; Region 4, Annual Mean	44
4-2	Decay of Temporal Conditionals; Region 4, Annual Mean	45
4-3	Decay of Spatial Conditionals; Region 11, Season 1	46
4-4	Decay of Temporal Conditionals; Region 11, Season 1	47
4-5	Decay of Spatial Conditionals; Region 11, Season 2	48
4-6	Decay of Temporal Conditionals; Region 11, Season 2	49
4-7	Decay of Spatial Conditionals; Region 11, Season 3	50
4-8	Decay of Temporal Conditionals; Region 11, Season 3	51
4-9	Decay of Spatial Conditionals; Region 11, Season 4	52
4-10	Decay of Temporal Conditionals; Region 11, Season 4	53
4-11	Diurnal Correspondence Between Times A and B	59
4-12	Area Enlarge Configuration	63
4-13	Analytical Representation of the Probability of Fractional Cloud Cover	70
4-14	Comparison of Observed and Normal Probability Curve Distributions	72
4-15	Comparison of Unconditional Probabilities for a 25:1 Area Increase	74
4-16	Comparison of Unconditional Probabilities for a 25:1 Area Decrease	75
5-1	Estimated and True Cloud Cover	78
5-2	Changes in Simulation Results as a Function of R	80
6-1	Thunderstorm Frequency Regions	85
6-2	Seasonal Analysis of Thunderstorm Frequency; December, January and February	87

LIST OF ILLUSTRATIONS, contd.

Figure		Page
6-3	Seasonal Analysis of Thunderstorm Frequency; March, April and May	88
6-4	Seasonal Analysis of Thunderstorm Frequency; June, July and August	89
6-5	Seasonal Analysis of Thunderstorm Frequency; September, October and November	90
6-6	Tornado Distributions; 1880 to 1942	92
6-7	Tornado Distributions; 1955 to 1967	93
6-8	Severe Local Storm Regions; Annual	95
6-9	Tornado Frequencies; Winter	96
6-10	Tornado Frequencies; Spring	97
6-11	Tornado Frequencies; Summer	98
6-12	Tornado Frequencies; Fall	99
6-13	Maximum Monthly Average of Tornado Days	100
6-14	Mean Tropical Cyclone Tracks	102
6-15	Climatological Probability of Tropical Cyclones in the North Atlantic	105
7-1	Mean Annual Precipitable Water; Northern Hemisphere	108
7-2	Vertical Cross Section of Mean Specific Humidity; Northern Hemisphere	109
7-3	Time Averaged Global Distribution of Precipitable Water; 1958	110
7-4	Mean Precipitable Water Distribution for North America; January	112
7-5	Mean Precipitable Water Distribution for North America; April	113
7-6	Mean Precipitable Water Distribution for North America; July	114
7-7	Mean Precipitable Water Distribution for North America; October	115
7-8	Precipitable Water Frequency Distributions; Mid-Season Months for Key West, Florida	117
7-9	Precipitable Water Frequency Distributions; Mid-Season Months for Athens, Georgia	118
7-10	Precipitable Water Frequency Distributions; Mid-Season Months for Caribou, Maine	119
7-11	Precipitable Water Frequency Distributions; Mid-Season Months for Topeka, Kansas	120

LIST OF ILLUSTRATIONS, contd.

Figure		Page
7-12	Precipitable Water Frequency Distributions; Mid-Season Months for Tucson, Arizona	121
7-13	Precipitable Water Frequency Distributions; Mid-Season Months for Winnemucca, Nevada	122
7-14	Precipitable Water Frequency Distributions; Mid-Season Months for San Diego, California	123
7-15	Precipitable Water Frequency Distributions; Mid-Season Months for Tatoosh, Washington	124
7-16	Analysis of Precipitable Water Frequency Distributions; Lake Charles, Louisiana	125
7-17	Analysis of Precipitable Water Frequency Distributions; Yucca Flats, Nevada	126

LIST OF TABLES

Table		Page
2-1	Cloud Category Designation	6
2-2	Example of Tabulated Cloud-Cover Distributions	8
3-1	Two-Reader Comparison	22
3-2	Consistency Check (Region 11, Month 8)	26
3-3	Discrepancies in Unconditional Distributions	26
3-4	Ground Station Data Sources	34
3-5	Transfer Matrix	35
4-1	Spatial Conditional Pairs per Day	43
4-2	Original Conditional Arrays; Region 11, Season 1	54
4-3	Detailed Spatial Scaling Results; Region 11, Season 1	55
4-4	Detailed Temporal Scaling Results; Region 11, Season 1	56
4-5	Error Comparison of Linear and Markov Scaling	57
4-6	Calculation of a Pseudo-Conditional Distribution for Diurnal Variation	60
4-7	Cloud Group Location Matrix	64
4-8	Area Enlarging of Unconditionals	67
4-9	Area Enlarging of Temporal Conditional; Region 11, Season 1	68
4-10	Error Comparison of Old and New Area Enlarge Procedures	69
5-1	One-Pass Simulation	79
6-1	Average Frequency of Tropical Cyclones (By Months)	103

1. INTRODUCTION

The great success of this country's meteorological satellite program has led to the concept of using spacecraft for observations of the earth itself. With an eye toward providing more effective monitoring and control of the world's diminishing natural resources, NASA is currently developing a series of experimental Earth Resource Technology Satellites (ERTS). The ERTS program is intended to be the forerunner of an operational satellite system which will monitor the earth's resources and provide much needed data to scientists in such diverse fields as geography, agriculture, hydrology, geology and oceanography.

To date, however, relatively little attention has been paid to the earth's atmosphere, through which all observations must be made. The effect of the intervening atmosphere is to degrade, or in many instances preclude, nearly all observations which might be made from space, whether they be in the visible, infrared, or microwave portions of the electromagnetic spectrum. A statistical model of atmospheric effects on a global scale is a necessary first step in the development of realistic procedures for planning and evaluating potential earth-oriented space missions. The feasibility of, and support requirements for proposed earth resources experiments can then be determined through computer mission simulations.

The Aero-Astrodynamics Laboratory of NASA's George C. Marshall Space Flight Center has initiated a series of studies intended to provide a useful model of atmospheric effects. The study to be reported on in this document is concerned primarily with the development of an effective global cloud model. This work is a follow-on to an earlier study which established a preliminary worldwide cloud data bank and developed techniques for manipulating and applying these data (Sherr et al., 1968). The accomplishments and limitations of the previous study will be discussed in detail in Section 2.

From actual application of the first-effort cloud model to specific mission analysis problems, it was recognized that improvements, both in the statistics and in the flexibility of the cloud model, were required. The relatively small size of the original data base led to weakness in the conditional statistics, even though distributions were developed for only winter and summer seasons. The use of ground-observed cloud amounts in the development of the unconditional statistics and satellite-observed cloud amounts in the development of most of the conditional statistics led to inconsistencies between the unconditional and conditional probability distributions and occasionally between the spatial and temporal conditionals.

The preliminary techniques to scale the statistics for any distance and time separation (including diurnal changes), and to change the representative area size required refinement and verification. Finally, the initial simulation results indicated only what coverage a satellite with the spatial resolution of a Nimbus or ESSA satellite (the data sources) would achieve under varying viewing conditions. No correction was made for different sensor types or different resolution capabilities. All of these problem areas have been approached in the current study. In addition, preliminary severe weather and water vapor content statistics have been added to the overall data base.

It is cautioned that the results presented here should not be taken as the final global cloud model. The data base has been significantly improved and made internally consistent, statistical manipulation techniques have been updated and refined, and the problems of modifying the simulation results for different sensor resolutions have been reviewed, but deficiencies in the overall cloud model still exist. Some problem areas, such as the development of a truly adequate correlation between ground and satellite-observed cloud amounts, or the correction of simulation results for different sensor resolutions cannot be resolved without the introduction of cloud type and structure statistics. The cloud-amount data base for many remote world areas still remains only marginally adequate. (Fortunately, most applications of the existing cloud model in the immediate future will be limited to areas such as the mainland United States where the data base is adequate.) The inherent complexities of real-world cloudiness patterns may preclude the development of any universally applicable closed-form statistical technique for adjusting area size. The techniques for compiling pass-to-pass coverage in mission simulations require still further modification and refinement. Throughout this study it has been our philosophy to make the best possible use of the funds, time and data available, recognizing that the correction of certain deficiencies would have to await further study and increased data availability.

2. THE PREVIOUS STUDY

The task of developing a global cloud model may be considered as consisting of three major phases:

1) The Development of a Data Base

In this phase decisions are made regarding data sources, data extraction techniques, data formats and how the data are to be stratified. Internal consistency checks as well as known "boundary conditions" should be applied during this phase to insure that the data base is as realistic as possible.

2) The Development of Statistical Adjustment Techniques

Clearly, it would be impossible to develop a data bank which incorporated all possible combinations and permutations of time and distance scales as well as representative area sizes. It is necessary, therefore, to first select a particular time and distance scale and a representative area size and then develop techniques for modifying the data base to accommodate changes in these parameters. Techniques for handling special situations such as the imposition of diurnal variations on the temporal conditional statistics or the passage of a satellite from one homogeneous cloud region to another must also be developed.

3) The Application of the Model to Specific Mission Problems

Procedures must be developed so that given a particular set of orbital characteristics, the cloud cover conditions which would be encountered on such a mission can be simulated. If the cloud-free portion of each data frame is to be composited from pass to pass, some information regarding cloud structure (layer or cellular) is desirable. To be truly useful, some procedures are needed to modify the simulation results for varying sensor types and particularly for sensor resolution.

Many, but certainly not all of these individual tasks were accomplished in the original study. Some of the statistical adjustment techniques developed at that time were necessarily of a preliminary nature. The purpose of this section is to outline the earlier work and to point out some of the limitations still remaining at the close of that study. Subsequent sections will then review the progress which has been made during the current study in removing these limitations.

2.1 The Data Base

2.1.1 Cloud-Climatic Regions

The practical use of cloud statistics in computer simulation routines dictates the subdivision of the earth into nominally homogeneous cloud-climatic regions. The number of such regions was arbitrarily set by consideration of the data volume that must be handled by the computer and by the amount of suitable data available. Since tabulations of the diurnal variation of cloud cover, of spatial conditional cloud distributions, and of temporal conditional distributions were required, the number of regions was kept relatively small.

Beginning with basic climatological classification systems, boundaries determined by temperature criteria alone were deleted, and greater emphasis was placed on those boundaries determined by precipitation differences. Further adjustments were then made on the basis of mean monthly cloud maps, cloud summaries, and satellite data. The locations of the resulting cloud regimes are shown in Figure 2-1. The 29 cloud climatologies are distributed over some 80 separate geographic world areas covering both hemispheres.

2.1.2 Unconditional Distributions

After completion of the initial climatological region selection, data were obtained from approximately 100 observing stations distributed throughout the world. In most cases the recorded cloud observations extended over a period of 10 to 15 years or more. For as many regions as possible, single representative stations were selected, and unconditional cloud cover distributions were derived from the data summaries for this station. Occasionally, because of lack of adequate data, the cloud climatologies for certain southern hemisphere regions were taken as being "seasonal reversals" of similar northern hemisphere regions. (i.e., The northern hemisphere data were shifted by six months and applied to the corresponding southern hemisphere region.) For a very few regions, where representative data could not be obtained, the statistics were modified from those of other regions on the basis of climatological considerations.

Five cloud cover categories were designated as displayed in Table 2-1. As can be seen, there are separate categories for clear sky and overcast conditions. Although a linear classification scheme might have been preferable, the format of the available cloud climatologies, particularly those for foreign countries, precluded any further breakdown of the cloud categories.

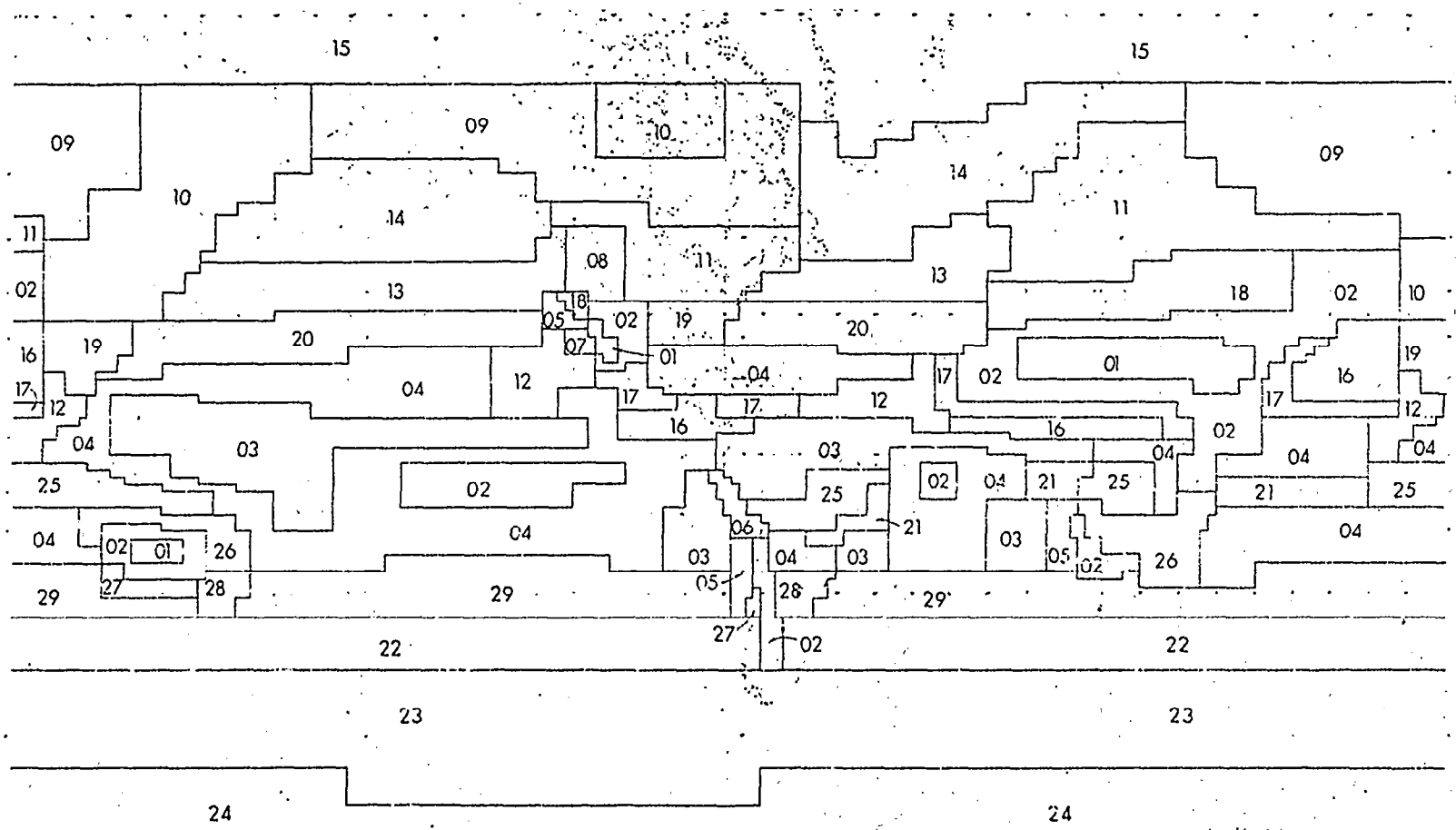


Figure 2-1 Cloud Cover Regions

TABLE 2-1

CLOUD CATEGORY DESIGNATION

Category	Tenths	Eighths (Octas)
1	0	0
2	1, 2, 3	1, 2
3	4, 5	3, 4
4	6, 7, 8, 9	5, 6, 7
5	10	8

Other than cloud climatological region, the nonrandom factors having the greatest effect on cloud cover are season and time of day. Accordingly, unconditional cloud distributions were provided for each month of the year and at three-hour intervals in the solar day. In cases in which an equivalent cloud climatological region occurs in both hemispheres, the seasons were inverted by shifting six months, and a new region designation is provided. Validation tests have shown this to be a reasonable procedure and better than accepting data from a location known to be unrepresentative.

2.1.3 Conditional Distributions

Cloud statistics, conditional with regard to time and space, were compiled for each climatological region. From the temporal conditional distributions the cloud amount probability distribution for "tomorrow" can be determined, given a cloud amount "today." Similarly, from the spatial conditional distributions the cloud amount probability distribution for a location at a specified distance from a base location can be determined, given a cloud amount at the base location.

Satellite data were heavily used to derive the conditional statistics, because the effort involved in summarizing raw conventional cloud data from various parts of the world would have been prohibitive. Satellite observations were obtained for most of the climatological regions; distributions for the remaining regions were adapted from the statistics available for apparently comparable regions.

Because sun-synchronous satellites (Nimbus) were used as the primary data source, the temporal conditional distributions were compiled for a time

interval of 24 hours. The spatial conditional data were tabulated for a 200 n.mi. distance separation. In both cases it was necessary to compute the probabilities on a six-monthly basis in order to increase the sample size.

Table 2-2 presents an example of the unconditional and conditional cloud-amount data as tabulated in the original study. In this case, the Month 1 (January) statistics for Cloud-Climatic Regions 1, 2 and 3 are shown.

2.1.4 Limitations

The number of cloud climatic regions was limited by consideration of the data volume that had to be handled by the computer and by the amount of suitable data available. The number of regions was kept relatively small. The entire United States, for example, is effectively covered in only four or five regions, although sufficient data exist for a much finer regional breakdown.

The satellite-derived data base for the compilation of conditional statistics is generally weak. The cloud climatologies for nine of the southern hemisphere regions were taken as being seasonal reversals of similar northern hemisphere regions. For certain oceanic areas where representative data could not be obtained, statistics were modified from those of other regions based on climatological considerations. It was necessary to compute the conditional probabilities on a six-monthly rather than seasonal or a monthly basis in order to produce an adequate sample size for statistical manipulations. Since the original study, nearly three years of additional satellite data have become available, and it is now possible to compile conditional statistics on at least a seasonal basis. By using the unconditional statistics as a "weighting factor" it is even possible to introduce monthly variations in the conditional arrays.

The use of ground-observed cloud amounts in compiling the unconditional statistics, and satellite-observed cloud amounts in compiling the conditional arrays has led to inconsistencies between these two data types. Subsequent statistical manipulations involving both conditional and unconditional statistics will not be valid unless these data arrays are first made compatible.

TABLE 2-2

EXAMPLE OF TABULATED CLOUD COVER DISTRIBUTIONS

CLIMATOLOGICAL REGION NUMBER 1 STATISTICS FOR MONTH 1

	UNCONDITIONAL PROBABILITIES TIME (LST)									CONDITIONAL PROBABILITIES 24 HOUR TEMPORAL						CONDITIONAL PROBABILITIES 200 NM SPATIAL				
	01	04	07	10	13	16	19	22		1	2	3	4	5		1	2	3	4	5
1	.62	.63	.49	.44	.42	.39	.46	.59	1	.85	.08	.05	.02	.0	1	.81	.07	.03	.09	.0
2	.14	.12	.17	.16	.18	.19	.18	.15	2	.78	.12	.05	.05	.0	2	.70	.10	.0	.20	.0
3	.06	.07	.10	.09	.07	.10	.10	.08	3	.75	.10	.10	.05	.0	3	.57	.29	.0	.14	.0
4	.11	.10	.14	.20	.22	.20	.17	.10	4	.80	.05	.05	.05	.05	4	.50	.0	.07	.29	.14
5	.07	.08	.10	.11	.11	.12	.09	.08	5	.80	.08	.05	.05	.02	5	.0	.0	.0	.01	.99

GIVEN

CLIMATOLOGICAL REGION NUMBER 2 STATISTICS FOR MONTH 1

	UNCONDITIONAL PROBABILITIES TIME (LST)									CONDITIONAL PROBABILITIES 24 HOUR TEMPORAL						CONDITIONAL PROBABILITIES 200 NM SPATIAL				
	01	04	07	10	13	16	19	22		1	2	3	4	5		1	2	3	4	5
1	.37	.38	.22	.18	.17	.17	.23	.31	1	.73	.06	.04	.11	.06	1	.63	.11	.02	.21	.03
2	.19	.16	.20	.21	.21	.19	.22	.20	2	.70	.10	.10	.05	.05	2	.50	.0	.0	.0	.50
3	.10	.08	.11	.10	.12	.13	.12	.12	3	.40	.13	.0	.14	.13	3	.20	.0	.20	.20	.40
4	.72	.23	.33	.35	.32	.33	.26	.24	4	.43	.04	.06	.25	.22	4	.22	.0	.22	.22	.34
5	.12	.15	.14	.16	.18	.18	.17	.13	5	.29	.07	.03	.35	.26	5	.25	.0	.12	.13	.50

GIVEN

CLIMATOLOGICAL REGION NUMBER 3 STATISTICS FOR MONTH 1

	UNCONDITIONAL PROBABILITIES TIME (LST)									CONDITIONAL PROBABILITIES 24 HOUR TEMPORAL						CONDITIONAL PROBABILITIES 200 NM SPATIAL				
	01	04	07	10	13	16	19	22		1	2	3	4	5		1	2	3	4	5
1	.17	.18	.15	.06	.01	.01	.04	.09	1	.17	.16	.17	.42	.08	1	.30	.20	.20	.20	.10
2	.24	.25	.21	.19	.11	.13	.18	.22	2	.05	.18	.24	.42	.11	2	.0	.25	.25	.50	.0
3	.12	.13	.13	.16	.16	.17	.13	.13	3	.02	.15	.27	.39	.17	3	.0	.0	.46	.46	.08
4	.27	.23	.33	.32	.41	.39	.34	.34	4	.03	.09	.21	.50	.17	4	.0	.0	.05	.73	.22
5	.20	.21	.29	.29	.31	.30	.31	.22	5	.05	.04	.18	.52	.21	5	.0	.0	.0	.41	.59

GIVEN

2.2 Statistical Adjustment Techniques

2.2.1 Time and Distance Scaling

The conditional distributions as tabulated in the data bank are given for a 200 n.mi. spatial distance (assumed to be directionally independent) and a 24-hour time separation. For simplicity in computer simulations, the assumption was made that the conditional probabilities decay linearly with distance or time. Thus, modifications for distances less than 200 miles or times less than 24 hours involve only a straightforward linear interpolation. Beyond 200 miles or 24 hours, two additional conditions are imposed:

1) For values on the diagonal, the interpolated value must be greater than the unconditional probability of the diagonal value; i.e., $P(x|x)$ must be greater than $P(x)$. If this test fails, the horizontal line of the 5 x 5 conditional matrix is replaced with the vertical column of unconditional statistics.

2) For off-diagonal values, the interpolated value must remain below the unconditional probability of the given cloud group. i.e., $P(x|y)$ must be less than $P(x)$. If this test fails, the entire horizontal line is again replaced by the unconditional statistics.

These restrictions insure that the data will return to the unconditional level for large distance and time separations (where little conditionality remains). They also preclude such undesirable products of manipulation as having the probability of clear skies at point A greater, when it is known to be not clear at a nearby point B, than when the situation at B is unknown.

An additional complication occurs when dealing with the temporal conditional statistics. In this case, it may be necessary to correct the interpolated matrices for diurnal variations in cloudiness. The technique which was derived to handle this situation effectively "weights" the regularly derived conditional array by the change in the unconditional distribution between the original and the new time. This approach satisfies the intuitive notion that diurnal change is superimposed on the more gross synoptic scale variations.

2.2.2 Area Adjustment

The general features of the change of cloud cover distribution with the size of the sample area can be readily visualized. The cloud cover over a true point

can have but one of two values, clear or overcast. The cloud cover over the entire earth seems to stay reasonably constant at about 40%. Areas of intermediate size have cloud distributions which pass from the U-shape characteristic of small areas to more bell-shaped distributions at rates which depend on the prevalence of large-scale cloud systems. The temperature zones, in which large cloud systems are the rule, show characteristically U- or J-shaped distributions at the 30-mile scale size of the ground observer. Tropical regions may already exhibit bell-shaped distributions at this scale. Similar transitions in distribution shape occur with the conditional arrays, although in this case the change is from valley to ridge configurations.

The procedure developed for increasing the representative area size utilized a Markov chain of individual data blocks across the diameter of the new area size. The joint distribution of the combined area in the diametric strip was thus determined and summed internally over like cloud amounts. On the assumption that cloud cover, rather than being randomly distributed, simply appears as a gradient across the area, the distribution of cloud cover in the diametric strip was taken as the distribution for the entire area.

2.2.3 Limitations

The procedure of replacing individual rows of the conditional arrays by the unconditional distributions (when scaling for time or distance) has come to be known as "stuffing." Insofar as the stuffing procedure represents a return to the unconditional statistics (i.e., neglecting whatever we may know about the conditions at some other point), its use for short intervals in space and time should be avoided. Of course, beyond certain limits the conditionality expires and the unconditionals are required. The premature return to the unconditional assumption is a drawback of the present system and is due in part, at least, to the use of a relatively unrealistic linear decay rate.

Although the existing techniques for the enlargement of the representative area size seem to work reasonably well for the unconditional statistics, experiments in adjusting the conditional statistics have revealed some discrepancies. The most vulnerable portions of the earlier enlargement scheme are the reliance on linear decay rates in estimating cloud distributions across a diametric strip and the extension of that distribution to the entire area. For future applications, where a wide variety of fields-of-view may be anticipated, a new technique should be sought for the area adjustment of the conditional statistics.

2.3 Mission Simulations

2.3.1 Monte Carlo Procedure

Given the basic data bank and a set of procedures to convert these data to the format required for specific earth observations, it is possible to simulate the performance of a sensor system, or of an entire earth observation program. The overall logic employed in a Monte Carlo routine to simulate repetitive looks at a particular area is presented in Figure 2-2. For computational convenience, all tables are first organized as cumulative probabilities in ascending order of cloud cover. Mission iteration number, Q, and pass number, n, are initialized. A random number, RAN, uniformly distributed in the interval 0 to 1 is generated and the first draw is made from the unconditional table by finding which cloud group probability interval contains RAN. (Recall that the probabilities are now cumulative in ascending order of cloud cover.) If the cloud group selected, G(n), is number 1 (clear), 100% coverage has been achieved and a suitable tabulation is made and another mission is initiated. If the cloud group is other than clear, a coverage percentage B is assigned and the pass number is incremented by 1. A new cloud cover is drawn from a temporal conditional table, using the row designated by the cloud cover drawn on the previous pass, and the column designated by a newly selected RAN. Again, if cloud group 1 occurs, (G(n) = 1), 100% cumulative coverage is tabulated for the pass, and a new mission is initiated. If G(n) is not 1, an incremental coverage, ΔB, is added to the existing coverage, B(n). The incremental coverage may add no new information at all, or it may result in total coverage. If it does, the 100% coverage branch is followed. If not, the coverage achieved is recorded as a function of the number of passes. This process is repeated for the N passes allowed in the total mission. The entire simulation is iterated NOQ times.

An example of the sort of results possible through the application of these techniques is presented in Figure 2-3. The graph shown in this figure results from a simulation of daily noontime observations from a sun-synchronous spacecraft with a 100 n.mi. field of view over the Gulf Stream region for the month of June. Percent coverage is plotted vs. the probability of obtaining that coverage as a function of the number of consecutive passes. Clearly, a number of other presentation formats are possible. Plotting the number of satellite passes required to achieve some particular coverage vs. the probability of achieving that coverage has also been found to be very useful. (Brown, 1969.)

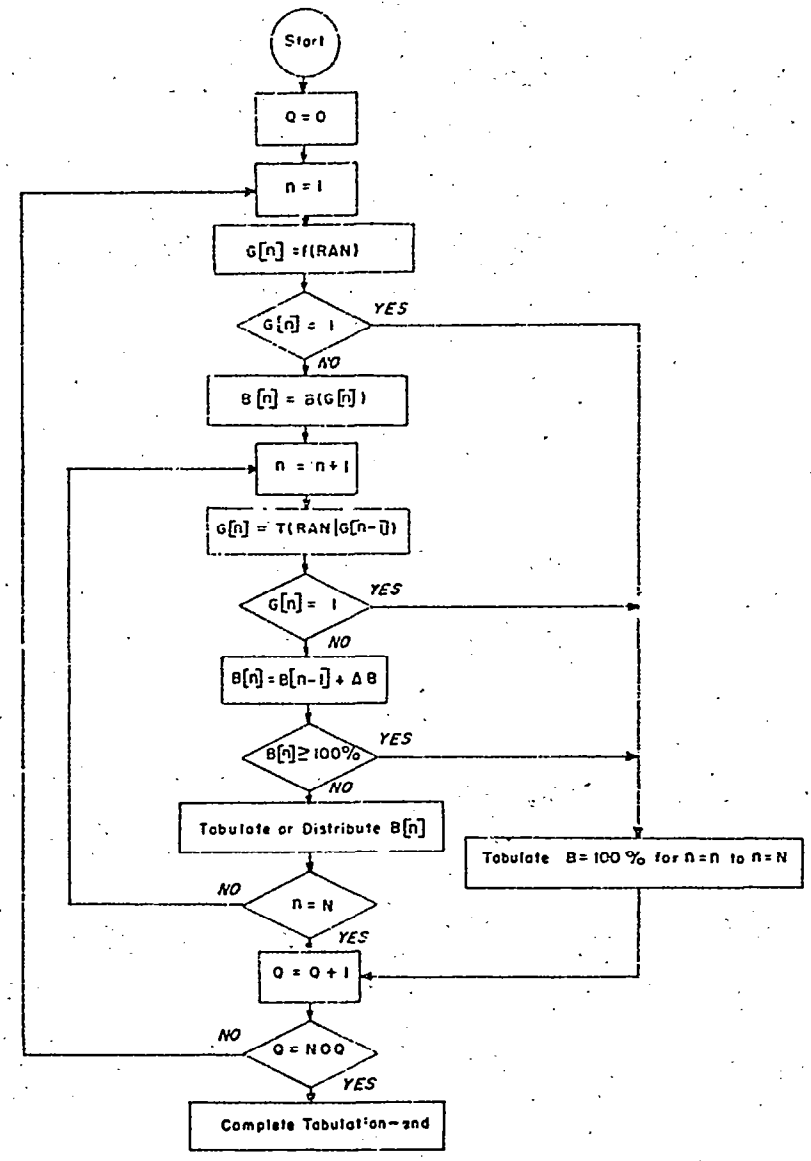


Figure 2-2 Monte Carlo Simulation Routine

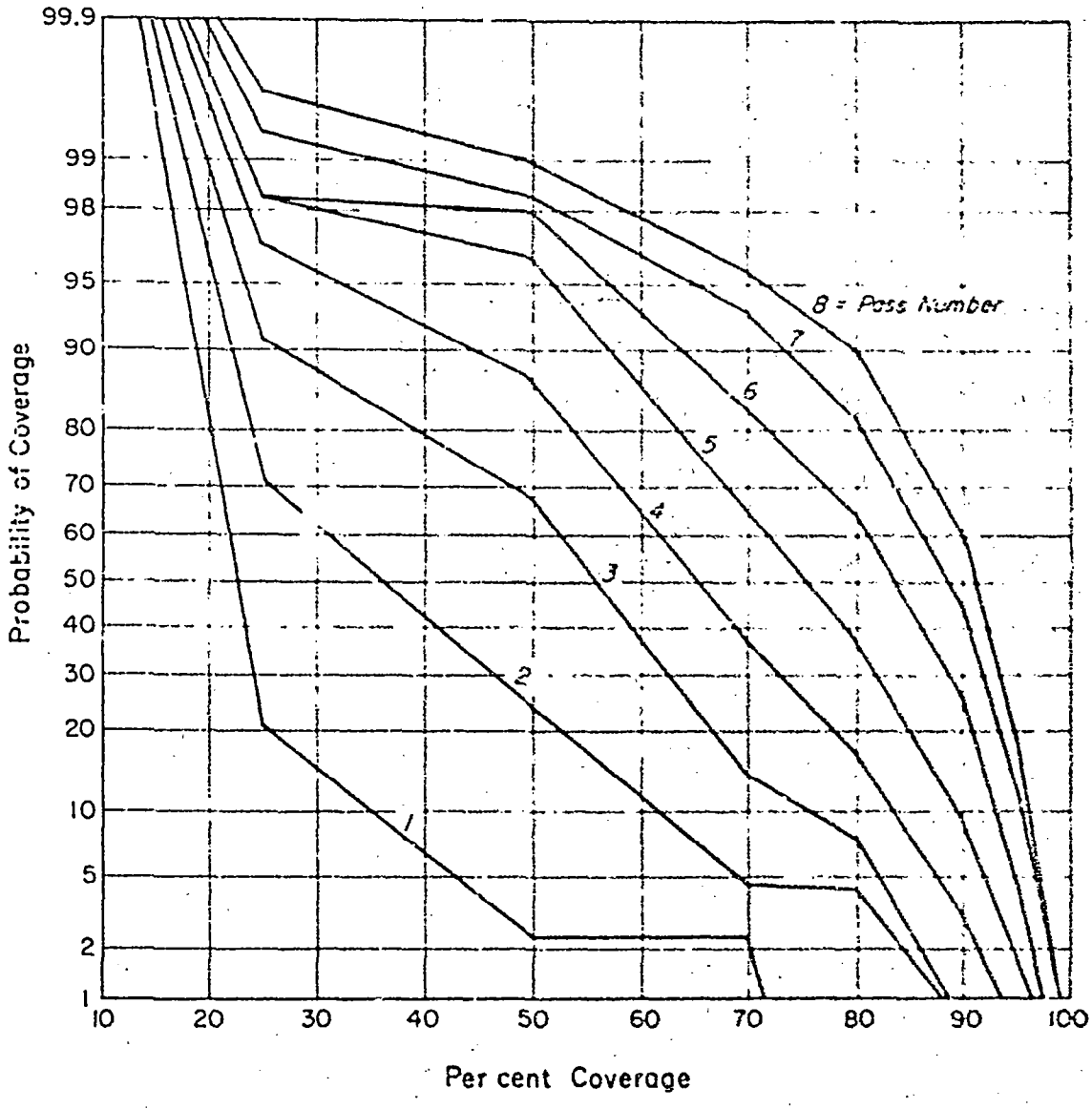


Figure 2-3 Monte Carlo Simulation Results

The resultant graphs of the simulation procedure show only the results of compiling clear sky information from pass to pass. For many applications (e.g., sea-surface temperature determination) the time interval between successive passes may be too great to permit the useful compiling of surface data. For these applications, it would be more important to determine the probability of achieving a certain percent of coverage on a single pass after a prescribed number of attempts. In this case, a Monte Carlo procedure is not even necessary since the desired results may be obtained through a simple binomial expansion of the original unconditionals in the cumulative format (see Section 5).

2.3.2 Limitations

In assigning a value to the incremental coverage, ΔP_i, in the Monte Carlo simulation procedure, an assumption of randomly distributed cloudiness is made. In an area of major cloud systems, the assumption of randomly distributed cloudiness becomes highly questionable. Some attempt has been made to modify the procedure as a function of whether the area in question was predominantly cellular or stratiform cloudiness (Brown, 1970). Further effort to incorporate cloud structure into the simulation procedure is required.

The resultant curves show only the net amount of cloudiness present as observed by an orbiting camera system with a 1.5 to 2 mile spatial resolution (Nimbus or ESSA). No attempt is made to infer the amount of "useful" cloud-free area as a function of different sensor types (microwave, infrared) or different spatial resolutions. To do so would require the incorporation of information as to the frequency distributions of cloud type. From the cloud type, in turn, we would have to infer such parameters as the horizontal and vertical cloud structure, the water drop size and the water content. In time, all of this information must be incorporated into the overall cloud model.

2.4 Summary

To review, the status of global cloud model at the close of the previous study was as follows:

(1) A data bank had been established which provided both unconditional and conditional cloud-amount tabulations. The data were particularly weak in the conditional arrays (necessitating a six-monthly stratification) and in certain southern hemisphere regions. The use of ground and satellite observed cloud-

amounts introduced inconsistencies into the data base. Only 29 cloud climatic regions were defined, although in certain areas, such as the Continental United States, a much finer breakdown is possible.

2) Preliminary techniques had been developed to adjust the time, distance and area scales of the existing data bank. The linear decay assumption in the time and distance scaling was not sufficiently realistic and lead to premature "stuffing." The area adjustment technique was also based on the assumption of a linear decay rate and, in practice, yielded questionable results.

3) A Monte Carlo program had been developed to simulate anticipated cloud amounts. The validity of the results were dependent upon the accuracy of both the data base and the statistical manipulation techniques. Moreover, no correction was made for varying sensors or sensor resolutions.

The following sections present a detailed summary of the work which has been accomplished during the current study to update and improve the global cloud and atmospheric model. In presenting our results, we shall follow the same general organizational structure as that established above, with separate sections on the data base, statistical techniques, and simulation procedures. Subsequent sections shall then discuss the introduction of severe weather and water vapor statistics.

3. DEVELOPMENT OF A REVISED DATA BANK

3.1 New Ground-Observed Data

A complete new set of ground-observed unconditional data has been included for Region 5, a Desert Marine climatic zone occurring over oceans off the west coasts of continents. Ten years of record from the San Clemente Island station off the California coast were used. The previous data set had been extracted from the Los Angeles Weather Bureau records and had required some modification by time of day and time of year. The actual processing of the San Clemente data was carried out in a separate study performed by ARA for the Scripps Institute of Oceanography (Duntley et al., 1970). These new data are included in the Revised Data Bank presented in Appendix C.

It had been hoped at the start of this study that some of the cloud cover work being performed by ARA for the Woods Hole Oceanographic Institution could be used to expand the data base for Region 6, an area occurring only off the coast of Peru. Unfortunately, only the lower resolution ATS data were used and then only for a one or, at most, a two year period of record. Given these limitations, it was decided to let the Region 6 data as recorded at Talara, Peru, remain unchanged.

The possibility of using the U. S. Navy Marine Climatic Atlas of the World as an additional unconditional data source for selected ocean regions was investigated. These data are presented by month in cumulative frequency graphs of cloud amount for individual data stations. (U. S. Navy, 1955.) Because they are not stratified by time of day, however, relatively little use can be made of these data other than as general guidelines to mean cloud amounts. As such, they may be useful in future studies to refine the boundaries of those cloud climatic regions occurring over remote ocean areas.

3.2 New Satellite-Observed Data

Since the completion of the previous study, nearly three years of additional satellite data have become available. Rather than provide only a cursory examination of all 29 regions it was decided to conduct a more intensive data extraction program for a limited number of areas. These areas include Regions 4, 5, 11, 13, 16 and 22, and are depicted as shaded areas in Figure 3-1. As will be seen in later sections, besides furnishing a reliable data base, the more intensive data extraction techniques provided a sound basis for the development of realistic time, distance and area scaling techniques.

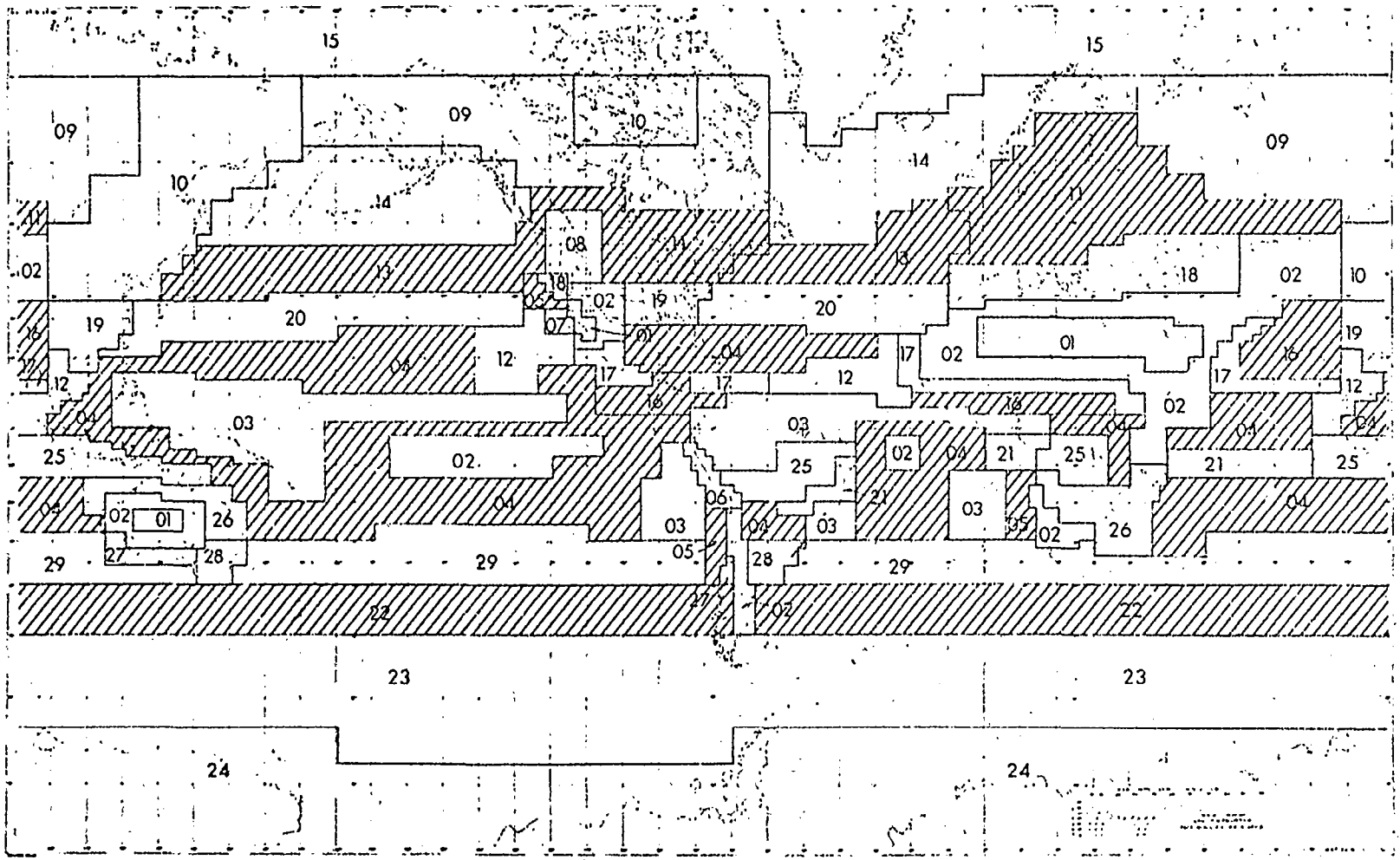


Figure 3-1 New Data Extraction Areas

3.2.1 Description of Selected Cloud Regions

The two primary cloud regions selected for an intensive data extraction were Regions 11 and 4, representative of mid-latitude and tropical cloudiness, respectively. For each region, three years of data (1967 through 1969) were extracted. The additional regions, for which two years of data (1968 and 1969) were extracted, are Regions 5, 13, 16 and 22.

Region 11 (Base Station; 40°N, 90°W)

Region 11 is of particular interest because a large portion of the United States falls within this area. The base station was selected so that the data extraction template (described in next section) would encompass the ground stations at Peoria, Illinois, and Pittsburgh, Pennsylvania. (These data were then used in developing ground/satellite cloud-amount comparisons - see Section 3.4.)

Region 4 (Base Station; 20°N, 65°W)

Region 4 is typical of a tropical ocean area with a moderate amount of cloudiness throughout the year.

Region 5 (Base Station; 32°N, 130°W)

Region 5, an oceanic area with predominately stratiform cloudiness, is a valuable addition to the original Data Bank, since the previously derived statistics for this region had been assigned a low-confidence factor. The base point for Region 5 was placed such that the template would encompass the San Diego ground station.

Region 16 (Base Station; 8°N, 0°W)

Cloud Region 16 is a tropical region with an extreme seasonal variation in cloud amount. It may be possible, therefore, to apply the Region 16 conditional statistics for various seasons to other tropical regions.

Region 13 (Base Station; 44°N, 42°W)

Cloud Region 13 is representative of mid-latitude, oceanic cloudiness. The satellite-derived statistics can be compared with those derived from Ship D, at 44°N, 41°W.

Region 22 (Base Station; 44°S, 30°E)

The statistics from the southern hemisphere Region 22 can be compared with those derived for Region 13, the northern hemisphere counterpart. Since data for these regions now cover a full two-year period, it is possible to examine more carefully the monthly variations in cloudiness and to determine whether Region 22 is actually a "seasonal reversal" of Region 13.

3.2.2 Data Extraction Procedures

The digital brightness mosaics from the ESSA satellite pictures were selected as the primary data source. The data in this particular configuration have the considerable advantage of a uniform projection geometry on a global scale without the shading effects present in individual picture frames. The actual data extraction was carried out at the National Environmental Satellite Service (NESS) in Suitland, Maryland.

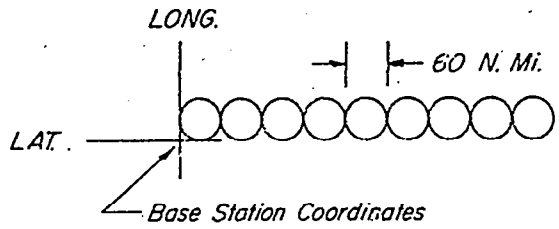
A template was prepared consisting of nine 60 n.mi. circles aligned in an east-west direction (see Figure 3-2). The first (left-most) circle was placed so that its left and lower edges were tangent to the selected longitude and latitude (respectively) of the "base station." Cloud amounts were then recorded for each of the individual circles or stations beginning at the base station and continuing toward the east. In order to permit the calculation of 540 and 600 n.mi. spatial conditionals, the original configuration of nine consecutive circles used for Regions 4 and 11 was altered somewhat for the remaining four regions. These two data extraction formats provided a number of significant advantages:

1) They provided an internal check as to the homogeneity of the cloud region (at least across the east-west extent of the nine circles).

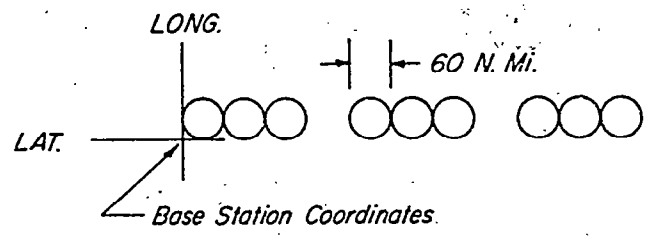
2) They significantly increased the data base for the calculation of both the unconditionals and the temporal conditionals.

3) They made possible the calculation of spatial conditional statistics in 60 n.mi. increments from 60 n.mi. up to 480 n.mi. using the initial configuration, and up to 600 n.mi. using the final configuration.

The use of the nine-circle templates sufficiently increased the data base so that conditional statistics could be directly tabulated by month for the six selected regions. It will be seen in Section 3.3.1 that it was also possible to tabulate the conditional distributions of all other regions by month, using the unconditional arrays to weight the existing conditionals.



Initial Configuration (Used for Regions 4 & 11)



Final Configuration (Used for Regions 5, 13, 16, & 22)

Figure 3-2 Data Extraction Templates

3.2.3 Two-Reader Comparison

As an internal consistency check, a second reader was employed to duplicate four months of cloud amount data from the ESSA mosaics. The specific months selected were: Region 4, August and September 1969; Region 11, July 1969; Region 13, April 1968. Unconditional frequency distributions were tabulated from each reader's data sheets for the combined four-month period. The results are shown in Table 3-1. The middle cloud amounts seemed to present the most difficulty, but the overall agreement is quite acceptable.

TABLE 3-1

TWO-READER COMPARISON

Cloud Category	Reader "A" (%)	Reader "B" (%)	Difference (%)
1	22	20	2
2	40	44	4
3	16	13	3
4	17	17	0
5	5	6	1
		Mean per-category difference:	2%

It is of at least academic interest to examine the two-reader agreement as a function of cloud amount. Figure 3-3 presents a graph of percent agreement vs. cloud amount (in tenths) for varying error margins. Although it was anticipated that the agreement would be the poorest for middle cloud amounts, the less than 20% exact agreement for 3, 4 and 5 tenths cloud cover is surprising. It can be seen, however, that even for these cloud amounts, nearly 60% of the readings were within one-tenth of each other. Again it can be seen that overcast and clear conditions produce the most favorable comparisons. This same sort of bias has frequently been noted in ground-observed cloud cover data.

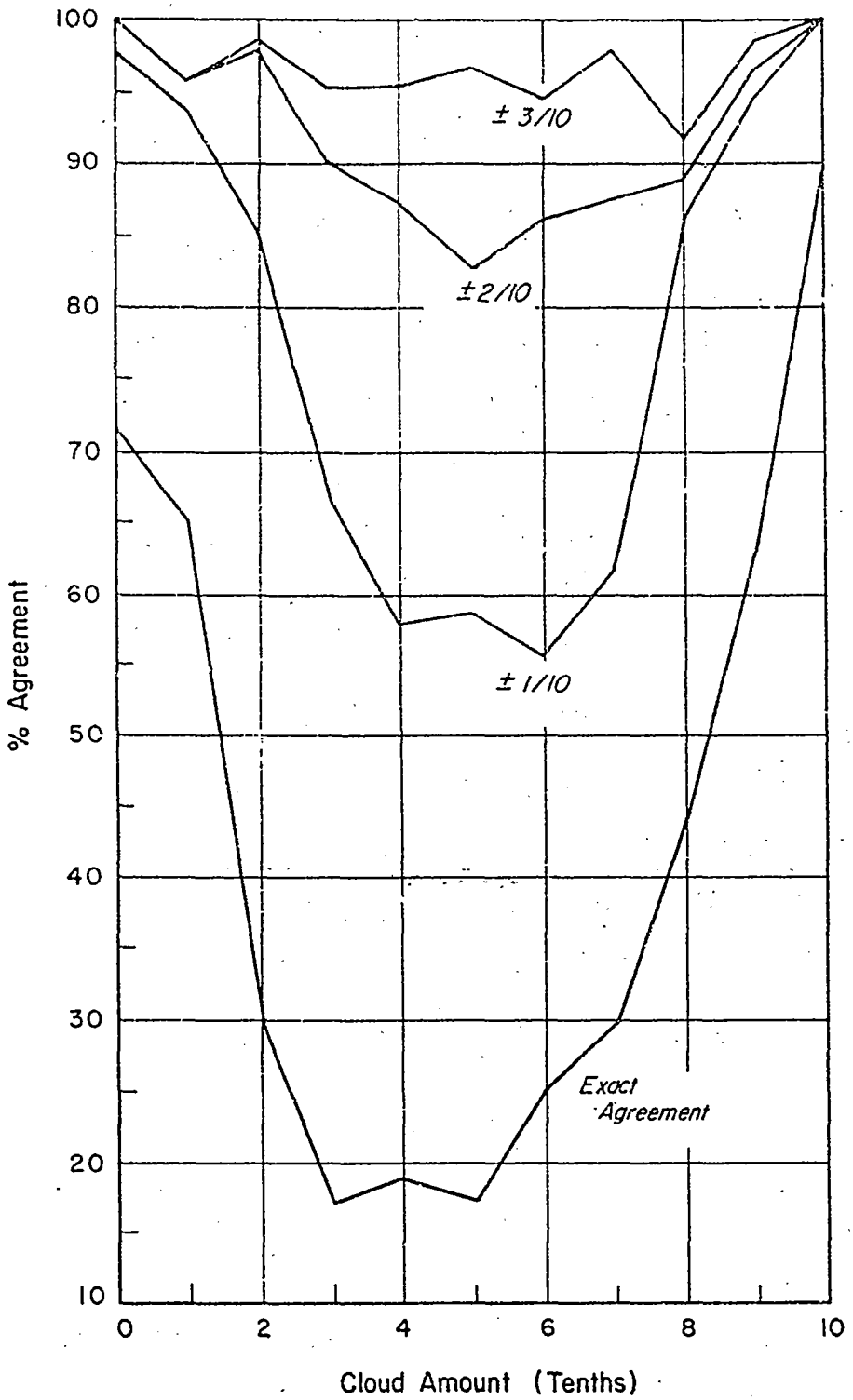


Figure 3-3 Two-Reader Cloud Amount Agreement

3.3 Internal Data Consistency

It has been noted that the cloud data bank at the close of the previous study had a number of internal inconsistencies. These seemed to stem largely from the use of different data sources in the compilation of the unconditional and conditional statistics. Fortunately, a number of well-established statistical techniques exist to remove or at least minimize the discrepancies.

3.3.1 Correction for Different Data Sources

Theoretical Consideration

In stratifying our statistical base by cloud climatic region and by month, we are making the following implicit assumption:

> Any two stations within the same cloud region and the same month will have identical unconditional cloud-amount frequency distributions.

A "station" in our case, consists of any arbitrarily sized circular area. The two stations may be separated by space or time or both. We can calculate the joint distribution of cloud amounts, P(a, b), between any two stations, a and b, as follows:

$$P(a,b) = P(a)\cdot P(b|a) \qquad (3.1)$$

where

$P(a)$ is the unconditional cloud frequency distribution at a

and

$P(b|a)$ is the conditional dependence of the cloud amount at b given the cloud amount at a.

It can be shown that

$$\left.\begin{aligned} \sum_{R} P(a,b) &= P(a) \\ \sum_{C} P(a,b) &= P(b) \end{aligned}\right\} \qquad (3.2)$$

where the R and C represent row and column sums, respectively. By our basic assumption, however, P(a) = P(b). Thus, when the tabulated 1300 local-time unconditionals (closest to the derivation time of the conditional arrays) are multiplied through the spatial and temporal conditional arrays, the resultant joint

distributions should have nearly identical row and column sums and they should equal the 1300 local-time unconditionals. The extent to which this is not the case (and it rarely is) is an indication of some discrepancy in the original data base. The discrepancy can arise from a number of sources:

1) The conditional and unconditional statistics are drawn from different populations.

2) The conditional and unconditional data are representative of different area sizes.

3) There are errors made in extracting cloud amount.

4) The cloud "regions" are not truly homogeneous.

Because the newly acquired data are derived solely from satellite observations, the row and column sums of their corresponding joint distributions are nearly identical. A representative example is shown in Table 3-2 for the Region 11, August data where the $P(b|a)$ is the raw 24-hour temporal conditional data. The row and column sums in this case differ by, at most, 2%. Since only the latter two error sources listed above are present in the newly acquired data, these small differences imply the relative unimportance of those error sources.

A Statistical Adjustment Technique

The question remains, of course, as to what can be done for the remaining 15 regions (not counting the eight seasonal reversals) where new data are not available. In order to get some feel for the extent of the problem, a computer program was written to extract from the existing conditional tables, those unconditionals which when multiplied through the conditionals would yield the proper row and column sums. It was learned that there were inconsistencies not only between unconditional and conditional arrays, but also between temporal and spatial conditionals. Table 3-3 presents a reasonably typical case, again for the Region 11 August data. The same sort of discrepancies were found for all regions and all months. (In certain specialized cases such as Region 13 where the unconditionals and temporal conditionals both came from the same data base, the agreement was nearly perfect.) In any event, it became evident that some correction procedure was required for the remaining 15 regions.

Several statistical techniques exist for adjusting joint frequency distributions to yield specific marginal totals, assumed to be known from other sources. (See in particular Chapter 8, "Adjusting Sample Frequencies to Expected Marginal Totals," in Deming, 1964.) In general terms, we are seeking a procedure to

TABLE 3-2

CONSISTENCY CHECK (REGION 11, MONTH 8)

P(a)		P(b|a)						P(a, b)					Row Sum
.26	X	.50	.25	.04	.18	.03	=	.1300	.0650	.0104	.0468	.0078	.26
.30		.22	.30	.13	.28	.07		.0660	.0900	.0390	.0840	.0210	.30
.10		.12	.31	.21	.31	.05		.0120	.0310	.0210	.0310	.0050	.10
.28		.18	.29	.09	.35	.09		.0504	.0812	.0252	.0980	.0252	.28
.06		.24	.28	.08	.34	.06		.0144	.0168	.0048	.0204	.0036	.06
							Column Sum:	.27	.28	.10	.28	.06	

TABLE 3-3

DISCREPANCIES IN UNCONDITIONAL DISTRIBUTIONS

(REGION 11, MONTH 8)

Ground-Observed Unconditionals (1300)	From Temporal Conditionals	From Spatial Conditionals
.07	.43	.26
.20	.13	.14
.18	.10	.13
.37	.19	.28
.18	.15	.19

operate on the existing joint distributions (as derived from the conditional arrays) to produce the desired marginal sums or "boundary conditions," while at the same time preserving as much as possible of the original cell distribution. In a least squares sense, we wish to enforce the boundary conditions and simultaneously minimize the sum

$$S = \sum_i \frac{(\Delta m_i)^2}{m_i} \qquad (4.11)$$

where

m_i is the original entry in the ith cell,

Δm_i is the adjustment required in the ith cell, and the sum is taken over all cells.

A rigorous solution to the least squares approach for a two-dimensional array does exist, but is extremely complex. Fortunately, much simpler iterative procedures exist which yield nearly identical results (Deming, 1943).

A brief outline of the iterative procedure which we have used to "normalize" the existing joint distributions is as follows:

1) Multiply each member of the ith row by the ratio $R_i/RSUM_i$ where R_i is the desired marginal total and $RSUM_i$ is the original row sum. This step simply normalizes each row so that the row sums equal R_i.

2) Multiply each member of the jth column by the ratio $R_j/CSUM_j$ where $CSUM_j$ is the current column sum. At this point the column sums are correct.

3) Return to Step 1 and continue until each of the row and column sums are simultaneously within some prescribed range (0.005 in our program) of the desired marginal totals. Convergence usually occurs within a very few iterations.

After convergence of the iterative procedure, the new conditional tables may be derived by dividing through each row by the final marginal sum.

The Desired Marginal Sums

Given that statistical techniques exist which may be used to normalize our joint distribution tables, one important question remains: To what set of unconditionals do we normalize the joint distribution? There are three possibilities:

1) Unconditionals as extracted from the conditionals - There are many problems here. The unconditionals as extracted from the spatial and temporal conditionals do not agree with each other, so that some sort of mean would have to be established. The resultant conditionals would still be stratified by six-month intervals. The data base for many of the conditional statistics is quite poor so that the extracted unconditionals would be unreliable.

2) The existing 1300 local-time ground-observed unconditionals - Here, at least, the data base is larger. It is difficult to say, however, just what the new conditional tables would represent. They would be an unusual combination of ground and satellite observed statistics.

3) Simulated satellite unconditionals - If a way can be found to accurately modify the existing 1300 local-time ground-observed unconditionals, valid for an area approximately 30 miles across, to the equivalent of satellite-observed unconditionals, valid for a 60-mile area, we would have the ideal unconditional set. They would be derived from a firm data base and would be completely compatible with the existing satellite-derived conditional distributions.

One of the tasks of this study was to develop a relationship between cloud frequency distributions obtained from ground observations and from satellite observations. In Section 3.4, a "transfer matrix" is derived which may be used to convert from ground-observed to satellite-observed frequency distributions. After considering the pros and cons of all three marginal sum possibilities, it was decided to use the "transfer matrix" to modify the existing 1300 local-time ground-observed unconditionals. The resultant simulated satellite unconditionals were then used to normalize the conditional arrays to produce a consistent data set. With the exception of those regions for which new data were available (Regions 4, 5, 11, 13, 16 and 22), all of the spatial and temporal conditional tables which appear in the Revised Data Bank discussed in Appendix C were generated in this manner.

Conditional Directionality

Before closing this section it is worthwhile to note that if the spatial conditionals were truly independent of direction we would have yet another boundary condition. It can be shown under this assumption (or the assumption that time runs equally well backwards or forwards) that the corresponding joint distribution matrix should be symmetric about the diagonal. The joint distribution matrix in Table 3-2 indicates that in at least some cases the time reversal assumption

may be valid. Lacking any firm evidence that cloud-amount conditionality is, in general, independent of direction in space or time, we decided not to apply this additional boundary condition. Strictly speaking then, the spatial conditionals as tabulated in the Revised Data Bank are west-to-east conditionals. The effects of the lack of north-south conditionals are considerably lessened by the east-west alignment of most of the cloud climatic regions.

3.3.2 Regional Homogeneity

One of the advantages of the data extraction procedure devised for the current study was that it afforded a means of checking the homogeneity of the six selected regions. In Figures 3-4, 3-5 and 3-6 we have arbitrarily chosen the observational frequency of clear skies as an indicator of regional homogeneity. (Season 1 represents Months 12, 1 and 2; Season 2, Months 3, 4 and 5; Season 3, Months 6, 7 and 8; Season 4, Months 9, 10 and 11.) It may be seen that Regions 13, 22 and, to a great extent, 4 show little variation in clear sky frequency (at least across the nine data extraction circles) throughout the year. Regions 11 and 16 are relatively homogeneous during some seasons, but are more variable during others. Region 5, which encompasses ocean and land conditions shows the greatest variability. (As will be seen in Section 3.4, this variability necessitated the dropping of the San Diego data from the ground/satellite cloud-amount calculations.) These results suggest that it would be worthwhile at some future time to refine the cloud climatic regions where the data are available to do so.

3.3.3 Seasonal Reversals

The unconditional statistics tabulated for cloud Regions 13 and 22 provided an opportunity to reexamine the "seasonal reversal" hypothesis between these two regions. The cloud frequency distributions by month for the two regions are presented in Figure 3-7. These computer-produced graphs show that the overall distributions are rather similar, with some months, such as April, May and November, being nearly identical for the two regions. For each region, the frequency of cloud-free conditions is 10% or greater in only three months of the year.

The frequency distributions for each region vary somewhat from month to month, with the mean cloud amounts ranging from about 65 to 45%. The variation is not exactly seasonal, however; for Region 22, April, May and June have the lowest mean cloud amounts whereas January and August have approximately the

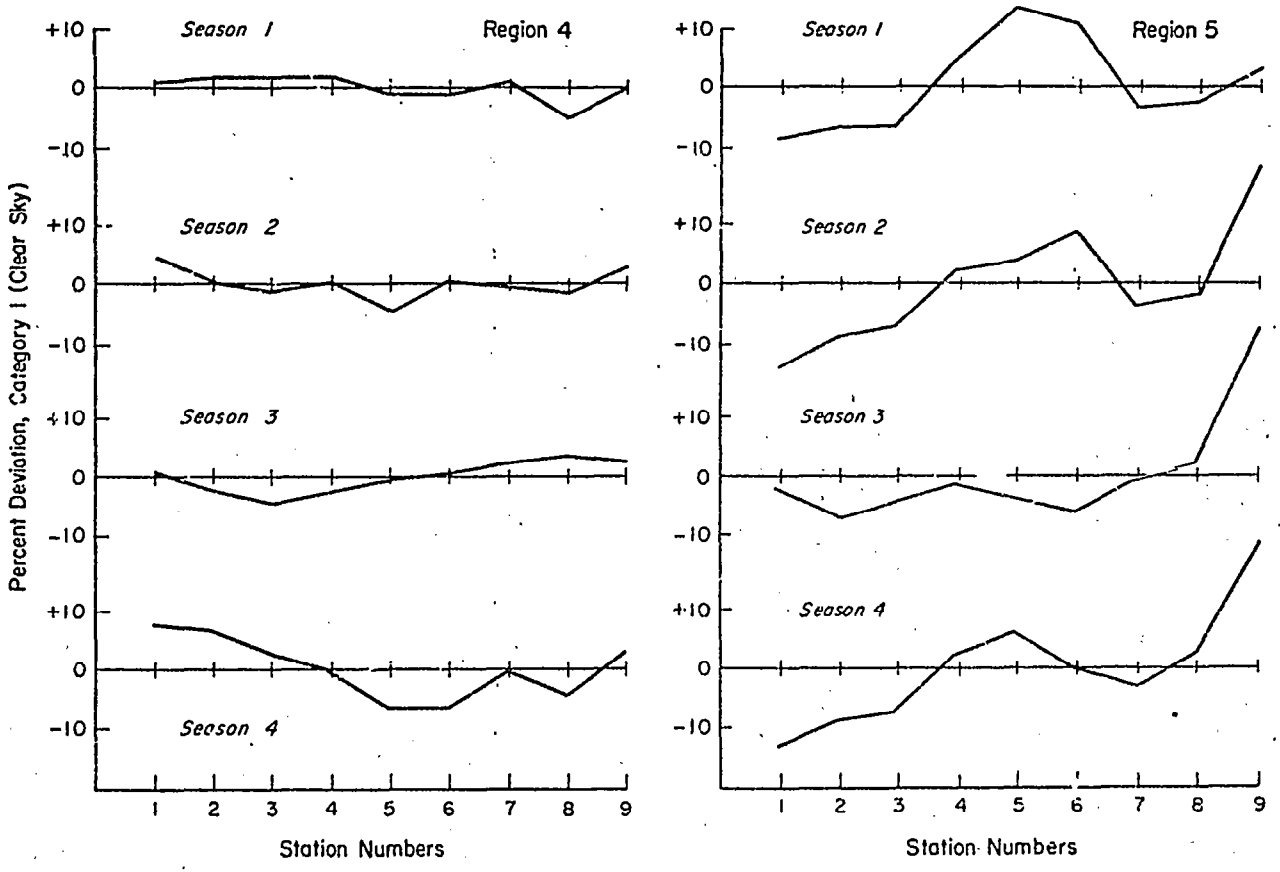


Figure 3-4 Regional Homogeneity; Regions 4 and 5

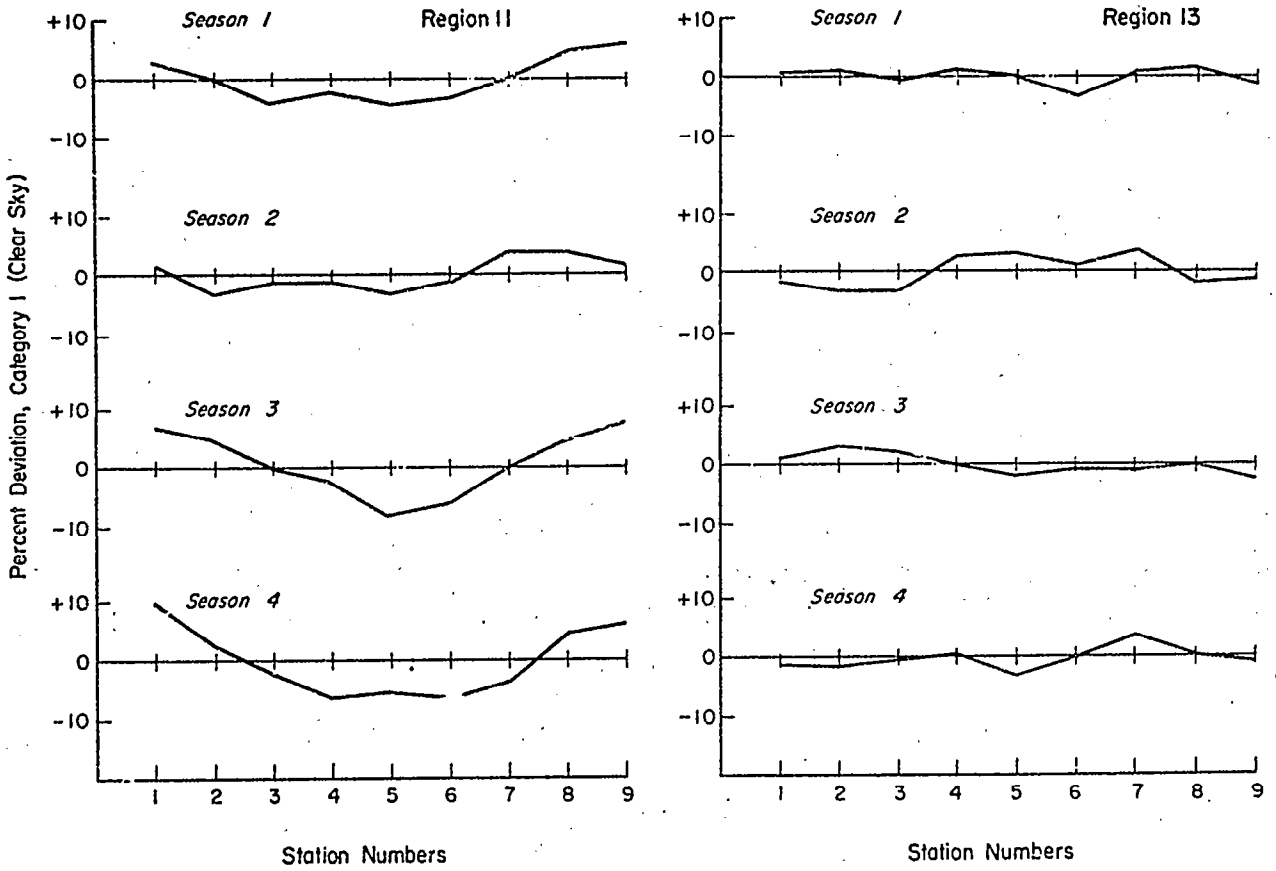


Figure 3-5 Regional Homogeneity; Regions 11 and 13

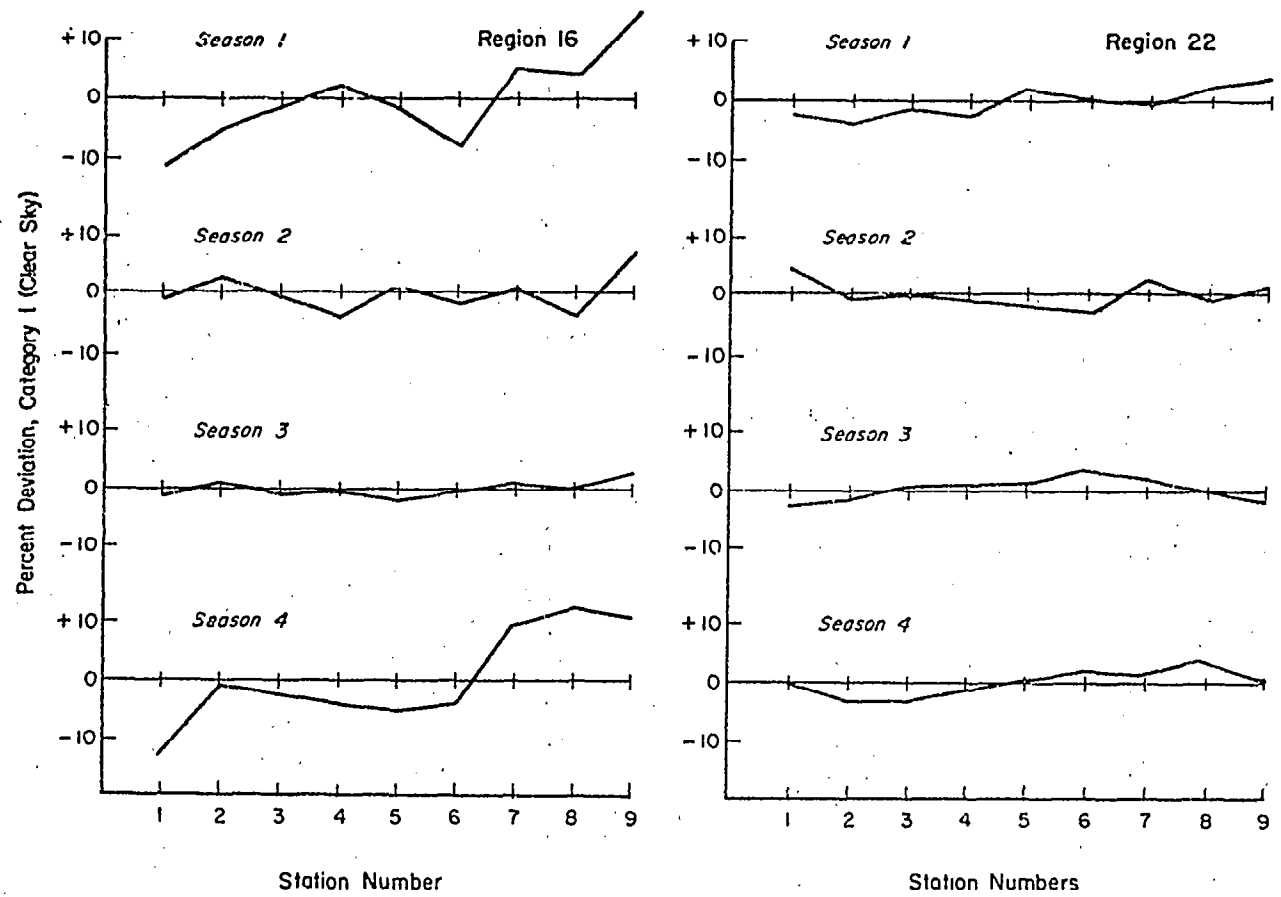


Figure 3-6 Regional Homogeneity; Regions 16 and 22

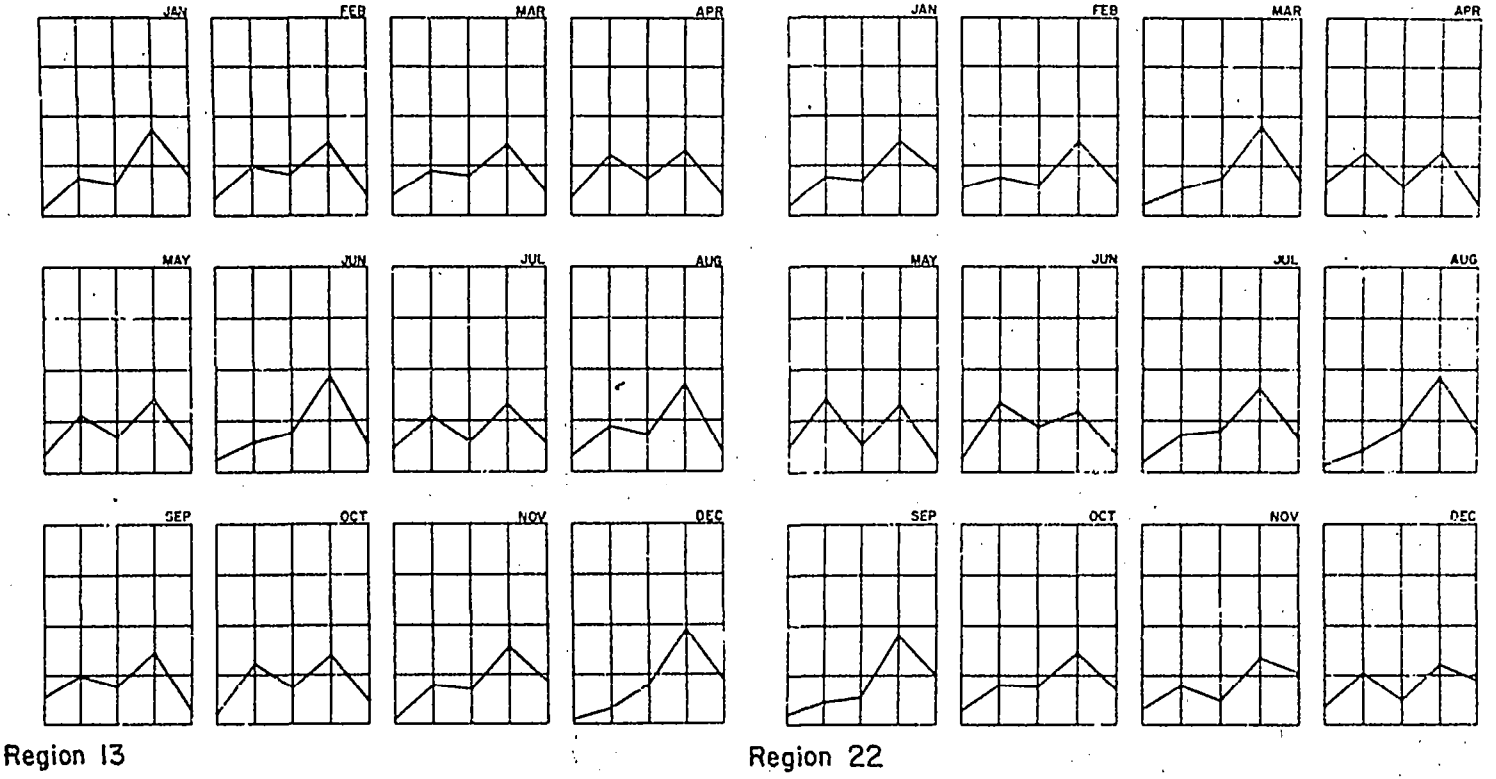


Figure 3-7 Computer Produced Unconditionals for Regions 13 and 22

same values. Other investigators have also reported that the cloudiness at high southern latitudes does not vary significantly throughout the year. For example, Van Loon (1966) gives the cloud cover at 50°S as about 80% throughout the year and at 35°S about 65%. These values agree fairly well with those given in the original data bank and in the satellite-derived unconditional statistics.

These results demonstrate that the three principal southern hemisphere oceanic cloud regions (Regions 22, 23 and 29) are most likely not exact "seasonal reversals" of the corresponding northern hemisphere regions. However, given the general lack of any reliable data for the southern hemisphere oceanic regions, we believe that no advantage would result from any revision of the current seasonal reversal hypothesis. For the southern hemisphere land areas, we can expect greater annual variations in mean cloud amount, and therefore a greater applicability of the seasonal reversal technique.

3.4 Ground/Satellite Cloud Observations

As has been noted, one of the tasks of this study was to develop a relationship between ground and satellite cloud-amount observations. This was achieved by selecting the base stations for Regions 5, 11 and 13 so that one or more of the data extraction circles would fall directly over an existing ground-observing station. Simultaneous ground and satellite cloud-amount observations were then recorded for each station on a day by day basis. The Local Climatological Data sheets issued by the Environmental Data Service of ESSA were used as the ground-observed data source for Regions 5 and 11. For Region 13, a punched-card record of daily observations was used. Table 3-4 summarizes the ground stations which were used, the corresponding satellite station, and the period of record.

TABLE 3-4

GROUND STATION DATA SOURCES

Region	Station	Satellite Station No. (Extraction Circle)	Period of Record
5	San Diego, California	9	1968-1969
11	Peoria, Illinois	1	1967-1969
11	Pittsburgh, Pennsylvania	8	1967-1969
13	Weather Ship "D"	1	1968-1969

In analyzing the simultaneous ground and satellite cloud-amount data, we attempted to set up a "transfer matrix" which when used to operate on a ground-observed unconditional data set would yield the equivalent of a satellite-observed distribution. To accomplish this, a joint distribution table of ground and satellite cloud amounts was generated using the individual day by day observations. The correlations between ground and satellite observations were found to be extremely poor for the San Diego station. Apparently the proximity of ocean, coastal and mountainous conditions in this region are such that any slight misalignment of the data extraction circle would yield highly erroneous results. For this reason the San Diego data were omitted from any further analysis. The final transfer matrix was derived by dividing through the joint distribution table (composed of Pittsburgh, Peoria and Ship "D" data) by the marginal totals. The resulting matrix is shown in Table 3-5. This matrix may be used to calculate a simulated satellite unconditional distribution, $P_s(j)$, from the original ground-observed cloud amounts, $P_g(i)$, as follows:

$$P_s(j) = \sum_{i=1}^{5} P_g(i) \cdot T(i,j) \qquad (3.4)$$

where

T(i, j) is the transfer matrix.

TABLE 3-5

TRANSFER MATRIX

		Satellite (j)				
		1	2	3	4	5
Ground (i)	1	.68	.17	.07	.08	.0
	2	.46	.28	.09	.15	.02
	3	.30	.27	.13	.24	.06
	4	.17	.24	.15	.35	.09
	5	.06	.10	.07	.43	.34

There are a number of observations which might be made regarding the T(i, j) transfer matrix derived above:

- It represents a tabulation of daily observations for a combined eight-year period, so that the joint distribution matrix has a population of nearly 3000 separate entries.
- Because the observations were compiled on a day by day basis, the final matrix is relatively independent of region, season or distribution shape. (e.g., The transformation from a 30% ground-observed cloud cover to some equivalent satellite-observed cloud amount should be independent of where or when the 30% coverage occurs or how often it occurs.) The only exceptions to this general rule would occur as a result of varying cloud types. (e.g., The overcast-to-clear transformation would occur more frequently in an area of predominant thin cirrus overcast than in an area with predominantly lower and thicker overcasts.) Lacking the cloud-type data necessary for any refinements in the transfer matrix, we decided to simply note that some problems may exist and to use the matrix as originally derived.
- The individual members of the T(i, j) matrix seem to be inuitively reasonable. The fact that a satellite with a 60-mile scan spot would record clear skies on only 68% of the occasions that a ground observer with a 30-mile observation circle records clear skies is in keeping with the decrease in clear sky observations when quadrupling the area size. The overcast-to-overcast reduction factor is even greater (34%) due to the transparency (to satellites) of certain types of overcast. The fact that a satellite will record clear sky conditions on nearly half (46%) of the occasions that a ground observer records Category 2 may be attributed to the 2 n.mi. resolution of the ESSA data source.
- One of the potential error sources in tabulating the T(i, j) stems from misalignments of the data extraction circle. We have assumed that, with nearly 3000 observations, these errors tended to cancel out.

The transfer matrix, T(i, j), derived above was used to operate (via Eq. (3.4)) on all of the original ground-observed 1300 local-time unconditionals, $P_g(i)$, to yield the equivalent satellite cloud amounts, $P_s(j)$. These simulated satellite unconditionals are tabulated by month and by region in the Revised Data Bank in Appendix C. Figure 3-8 shows examples selected from Regions 4, 16 and 22 (not used in the deviation of the transfer matrix) of how the "transformed" 1300 local-time unconditionals match the recently recorded satellite unconditionals. (Recall that we are going from a ten-year data set to a different two-year data set.)

3.5 Description of Revised Data Bank

The Revised Data Bank is listed in Appendix C. In Section C.1, the unconditional and conditional distributions are listed in exactly the same format as that used in the original report (Sherr, 1968), and as shown in Table 2-2. In Section C.2, the 1300 local-time ground-observed unconditional distributions as modified to simulate satellite-observed unconditionals are listed by month and by region.

Unconditional Distributions

For all regions except Region 5, the unconditional probabilities as listed in Section C.1 of Appendix C are unchanged from those listed in the original data bank. The original Region 5 data were replaced by an unconditional data set newly extracted from ten years of San Clemente records. As before, these data are valid for a representative area size with a 30 n.mi. diameter.

Conditional Distributions

The conditional distributions for all regions as listed in Section C.1 of Appendix C were derived as follows:

1) Using Eq. (3.4) and the T(i, j) transfer matrix derived above, the original 1300 local-time ground-observed unconditionals, $P_g(i)$, were modified to simulate satellite-observed unconditionals, $P_s(j)$, valid for a 60 n.mi. area.

2) For Regions 4, 5, 11, 13, 16 and 22 the original temporal and spatial conditional arrays were replaced by the newly extracted conditional data.

3) The simulated satellite unconditionals, $P_s(j)$, were then multiplied through the conditional arrays, $P(i \mid j)$, for all regions to produce joint distribution matrices, P(i, j).

$$\text{i.e.,}\quad P(i,j) = P_s(j) \cdot P(i \mid j). \tag{3.5}$$

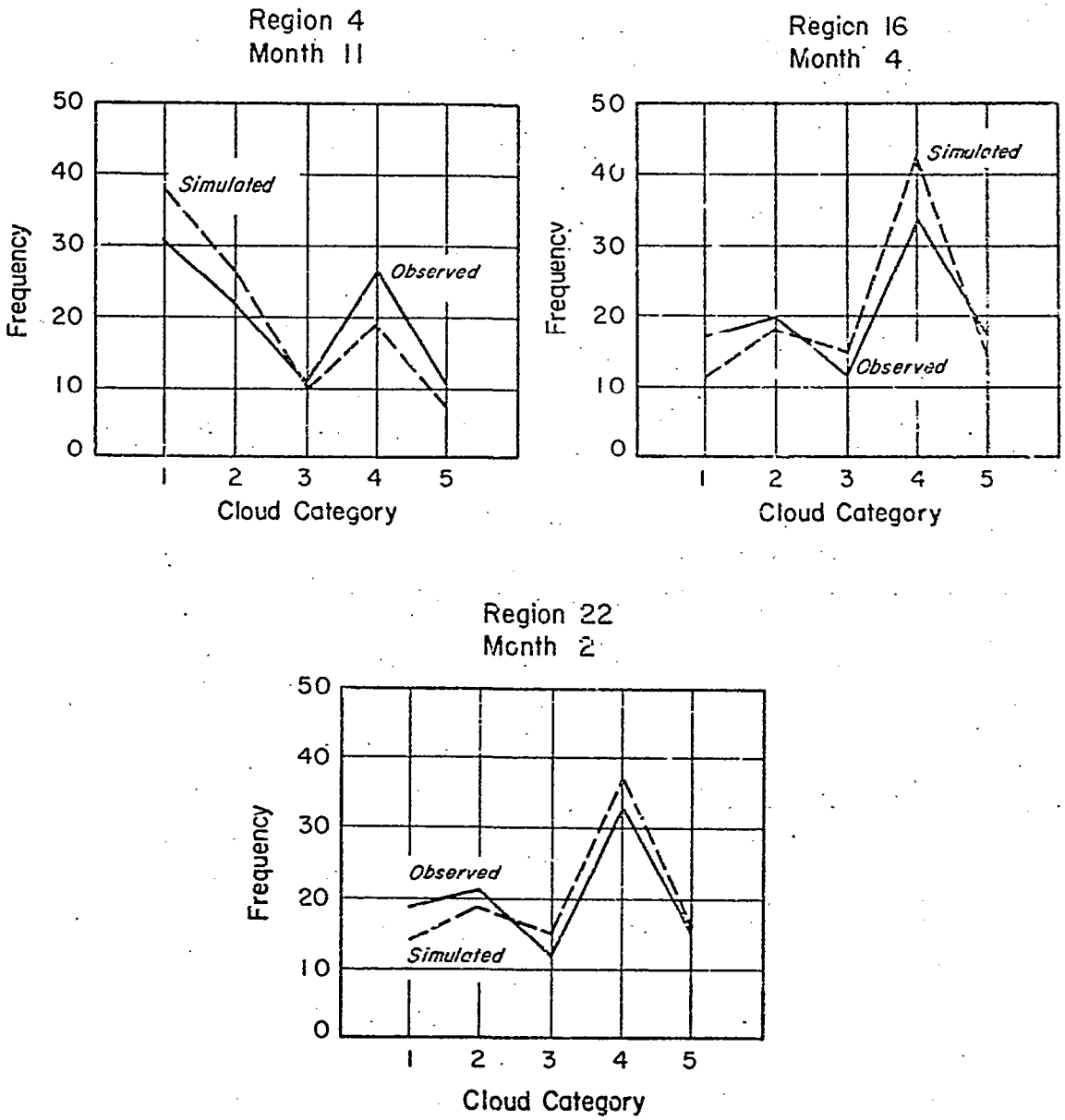


Figure 3-8 Applications of Transfer Matrix

(The newly extracted satellite unconditional data for the six new regions were quite similar to the simulated data, but exhibited greater fluctuations from month to month, probably as a result of the smaller data base. For this reason we felt it preferable to use the simulated unconditionals in the step above.)

4) Using the iterative procedure described in Section 3.3.1, the joint distribution matrices were normalized to yield the simulated-satellite unconditionals as the row and column sums.

5) The normalized joint distribution matrix was then divided through by the final row sums to yield the conditional distributions listed in Appendix C.

Simulated Satellite Unconditionals

The simulated satellite unconditionals as derived from the 1300 local-time ground-observed unconditionals are listed in Section C.2 of Appendix C by month and region.

4. REFINING THE STATISTICAL ADJUSTMENT TECHNIQUES

4.1 Time and Distance Scaling

4.1.1 Markov Scaling

As was noted in Section 2.2.3, the use of a linear decay rate for temporal or spatial conditionality often led to a premature return to unconditional levels through the procedure known as "stuffing." This section considers the application of the assumption of a simple Markov chain process to the problem of scaling the conditional probability of the cloud cover for a specific area size and a reference distance (or time) to any other distance (or time). The following discussion refers specifically to the spatial scaling problem, but is also directly applicable to temporal scaling for integer multiples of a day. It can be extended to apply to other times by making use of the diurnal variation of the corresponding unconditional probabilities in the technique described in Section 4.1.3.

Consider first the cloud cover condition at some observed location b located a distance ℓ away from a reference location a. In general, the cloud cover condition at b is related to the conditions at a by a conditional probability matrix of the form $P(b|a)$. For distances from a which are N times ℓ away from this reference location, the corresponding general expression for the conditional probability $P_N(b|a)$ is of the form (using N = 4 as an example):

$$P_4(e|a) = \sum_d \sum_c \sum_b P(e|d, c, b, a)\, P(d|c, b, a)\, P(c|b, a)\, P(b|a). \qquad (4.1)$$

If it is now assumed that cloud cover conditional probabilities follow a simple or first order Markov chain, the various terms in Eq. (4.1) become simplified as follows:

$$P_4(e|a) = \sum_d \sum_c \sum_b P(e|d)\, P(d|c)\, P(c|b)\, P(b|a). \qquad (4.2)$$

Finally, assuming that the first order conditional probability matrix (for the distance ℓ) does not vary along the N elements of the chain, all of the P's on the right-hand side of Eq. (4.2) become identical and the entire right-hand side of the equation is simply the N-th power of the basic $P(b|a)$ matrix; or in general

$$P_N(b|a) = \left[P(b|a) \right]^N. \qquad (4.3)$$

The above equation applies for any value of N, whether integer or not, as can be visualized, for example, by considering N to be composed of the quotient of two integers $N = N_1/N_2$, and referring all quantities to a smaller reference distance t/N_2. However, it might be noted that if noninteger powers are used with Eq. (4.3), the resulting calculated conditional probabilities for the cloud cover statistics considered in the present program did not always turn out to be positive real numbers. In some cases, particularly for small fractional powers, calculated probabilities sometimes turned out to have small imaginary parts or small real negative values. These small discrepancies appeared to be largely associated with round-off errors in the input cloud statistics data and have been arbitrarily smoothed out in a computer program which has been developed for raising a conditional probability matrix to a power. The input to this program includes the conditional array to be scaled, the power to which the matrix is to be raised and the unconditional probability levels to which the matrix is to converge. A FORTRAN listing of this subroutine may be found in Section B.1 of Appendix B.

4.1.2 A Test of Markov Scaling

One of the primary reasons for developing the particular data extraction template described in Section 3.2.2, was to provide a data base which would be sufficiently large to permit a careful analysis of the decay of conditionality with time or space. Temporal conditional arrays were compiled by noting the day to day change in cloud cover for each of the data extraction circles or stations within a region, and then combining the individual station data to produce final conditional arrays stratified by month or by season. Temporal conditional distributions were calculated in 24-hour increments from 24 hours to 240 hours. The spatial conditional statistics were compiled by considering each circle in turn as the base station and then tabulating the change in cloud cover from that location to every other circle to the right of that location. Table 4-1 shows, for both template formats, the number of spatial conditional pairs which could be extracted from each day's data. The spatial conditionals were tabulated in 60 n.mi. increments from 60 n.mi. to either 480 n.mi. (initial template) or 600 n.mi. (final template). In both cases, the reliability of the data tends to diminish (fewer observations) with increasing distance.

In order to make the analysis of the decay of conditionality a more manageable task, a new cloud-amount categorization was defined. The new system included only three categories, 1, 3′ and 5, where 1 and 5 were clear and overcast as before, and 3′ included Categories 2, 3 and 4 of the original classification.

TABLE 4-1

SPATIAL CONDITIONAL PAIRS PER DAY

Distance (n.mi.)	Number of Pairs (Initial Template)	Number of Pairs (Final Template)
60	8	6
120	7	5
180	6	4
240	5	6
300	4	4
360	3	3
420	2	2
480	1	3
540	0	2
600	0	1

Thus a conditional array consisted of only nine numbers in a three-by-three format, rather than 25 numbers. Figures 4-1 through 4-10 show the observed decay of conditionality with distance and time for Regions 4 and 11. Graphs of certain cloud category combinations were omitted when the populations were so small as to render the statistics meaningless. The data are stratified by season (Season 1 for Months 1, 2 and 12; Season 2 for Months 3, 4 and 5; Season 3 for Months 6, 7, 8 and 9; Season 4 for Months 10 and 11) for Region 11 and are presented as an annual summation for Region 4. (The particular seasonal breakdown for Region 11 and the decision to use an annual summary for Region 4 resulted from a review of computer-produced graphs of monthly unconditional distributions such as those shown in Figure 3-7 in Section 3.3.3.)

Of particular note in these figures is the lack of oscillation or "antipersistence" about the unconditional levels such as was found in the previous study. All of the decay lines begin at one or zero and seem to approach the unconditional levels asymtotically. Perturbations at large distances or times are believed to be due to a diminished data base (recall Table 4-2).

In order to evaluate the Markov scaling routine, the observed 180 n.mi. spatial conditional was taken as a "given," and then scaled backward to 60 and 120 n.mi. and forward to 240, 300, 360 and 420 n.mi. The observed 48-hour temporal conditional was scaled to 24 hours and the observed 24-hour conditional

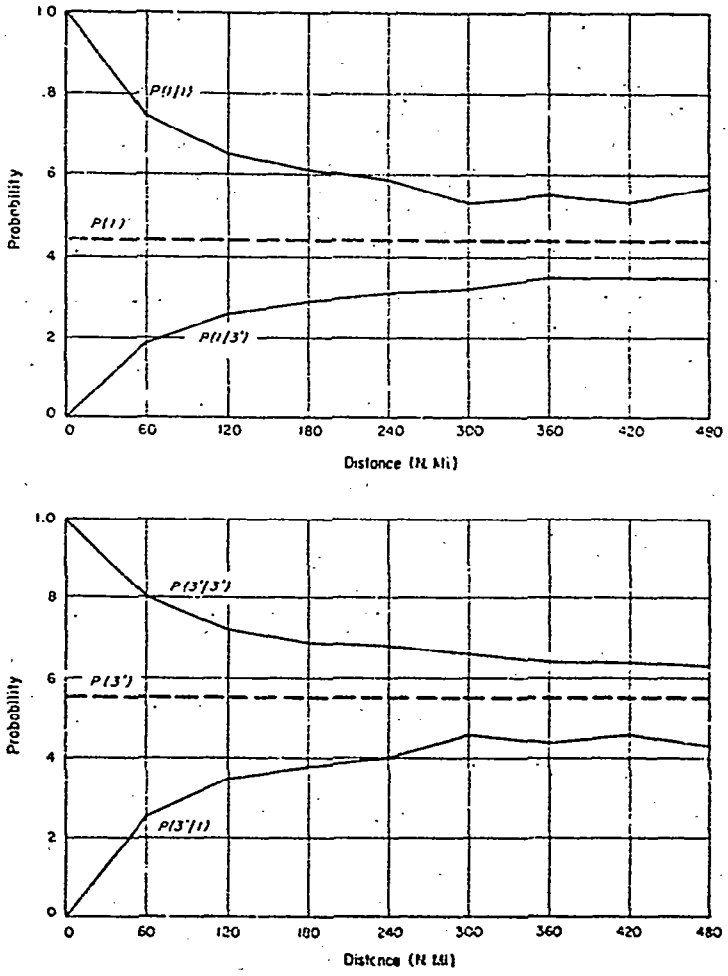


Figure 4-1 Decay of Spatial Conditionals; Region 4, Annual Mean

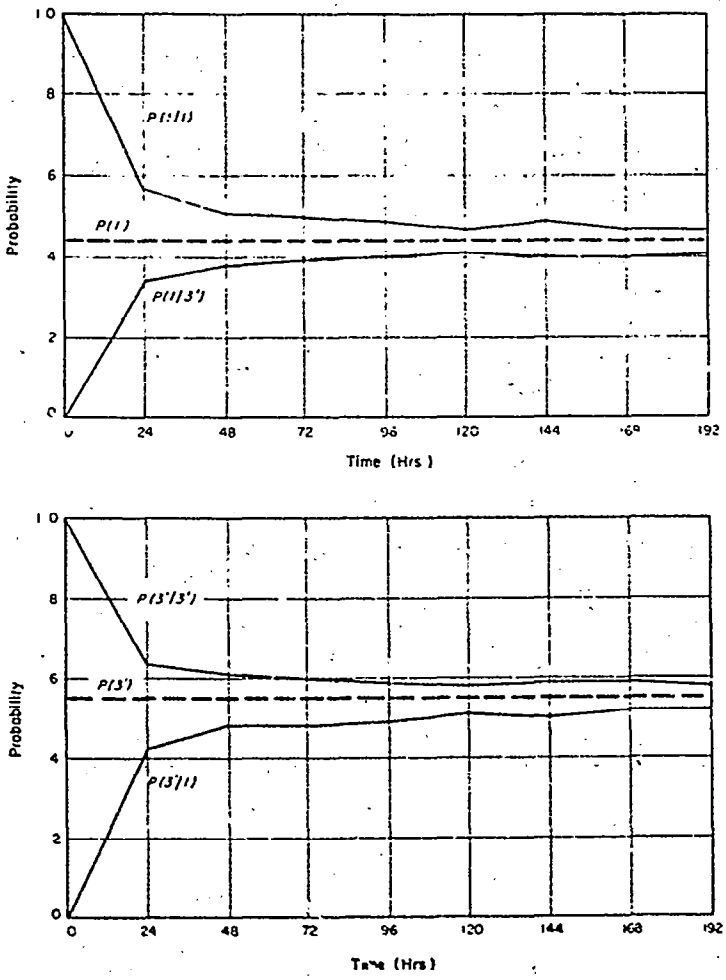
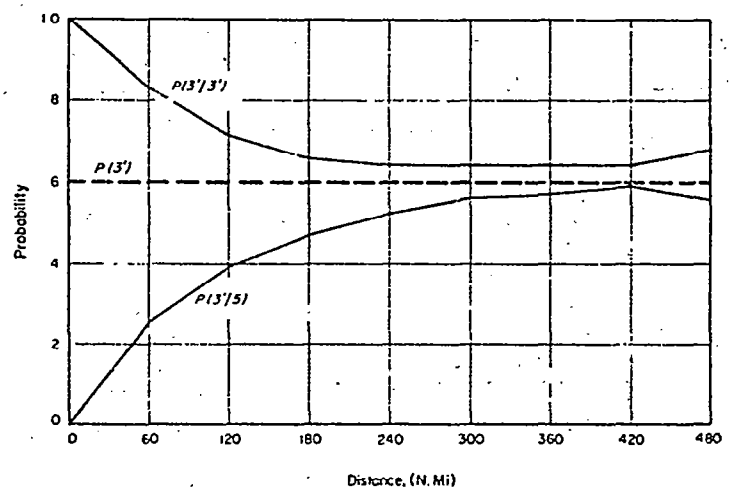


Figure 4-2 Decay of Temporal Conditionals; Region 4, Annual Mean



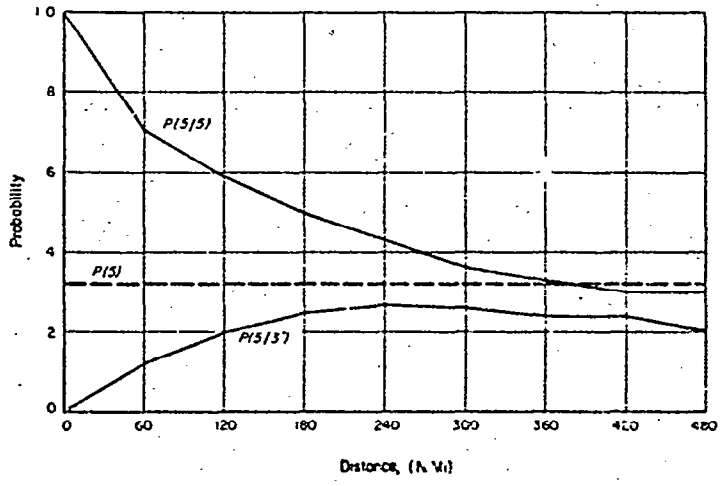


Figure 4-3 Decay of Spatial Conditionals; Region 11, Season 1

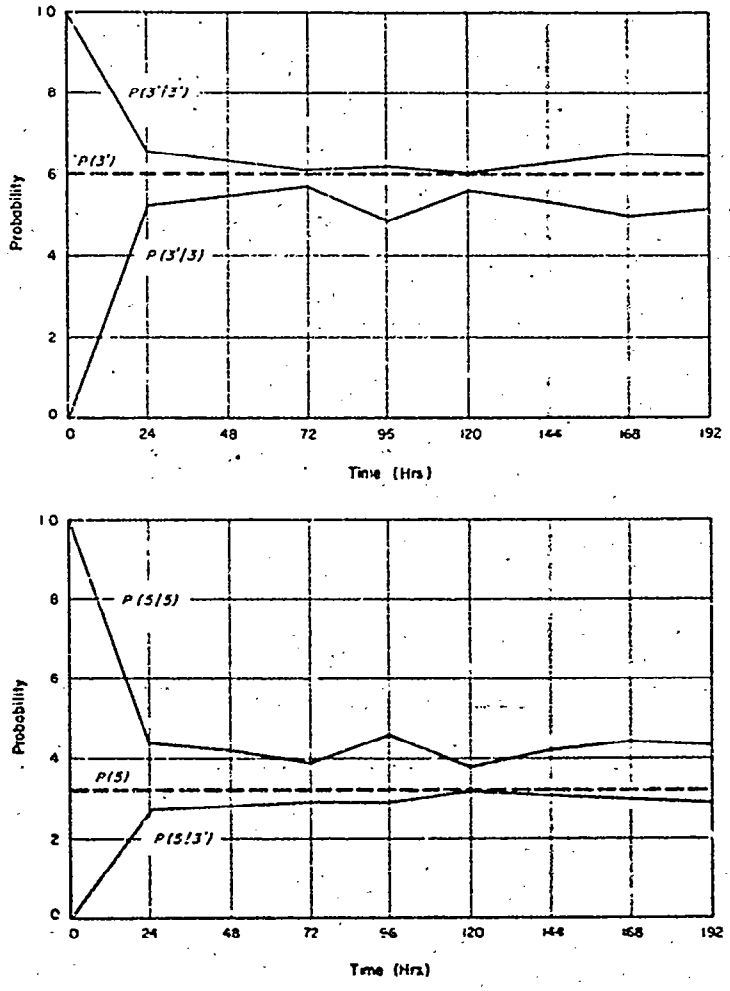


Figure 4-4 Decay of Temporal Conditionals; Region 11, Season 1

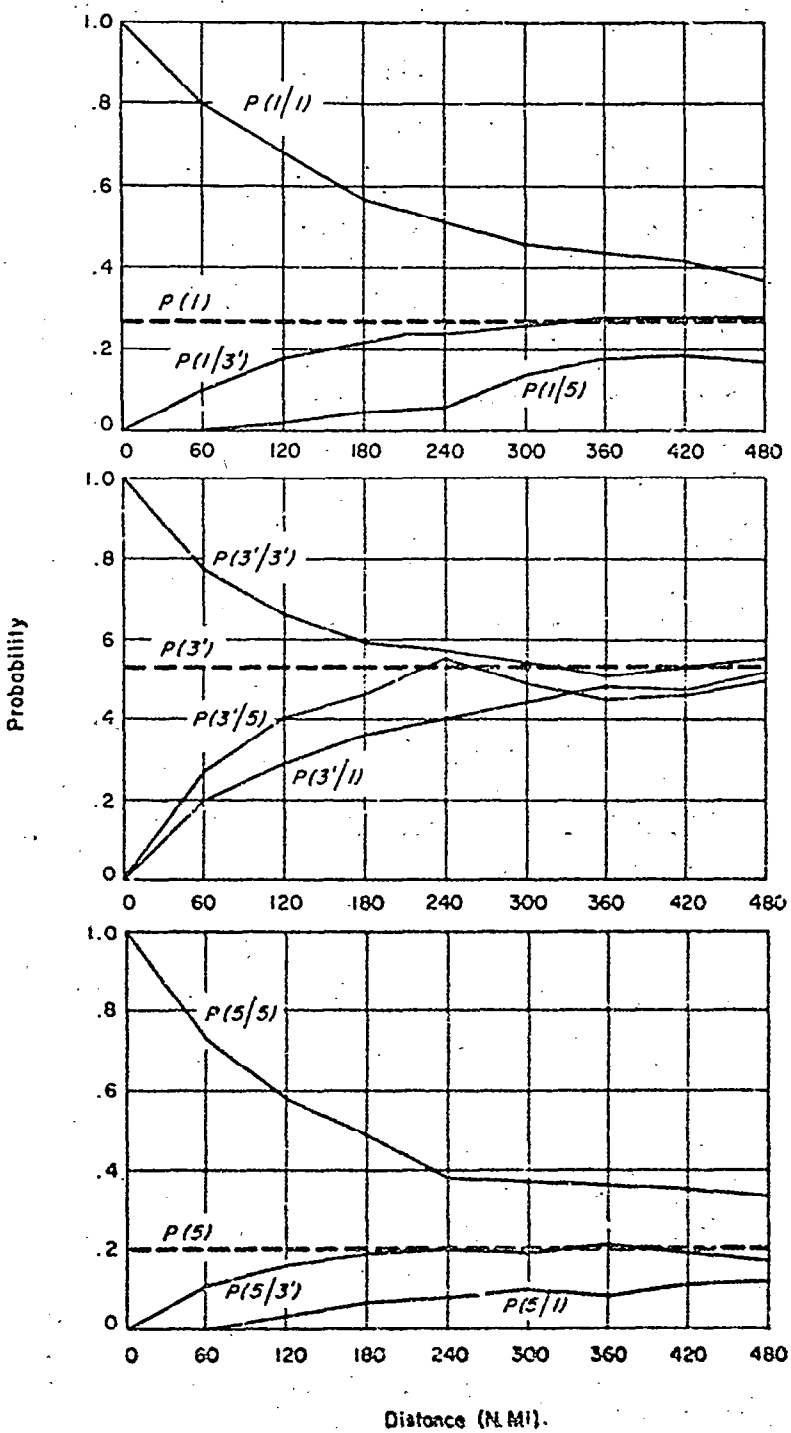


Figure 4-5 Decay of Spatial Conditionals; Region 11, Season 2

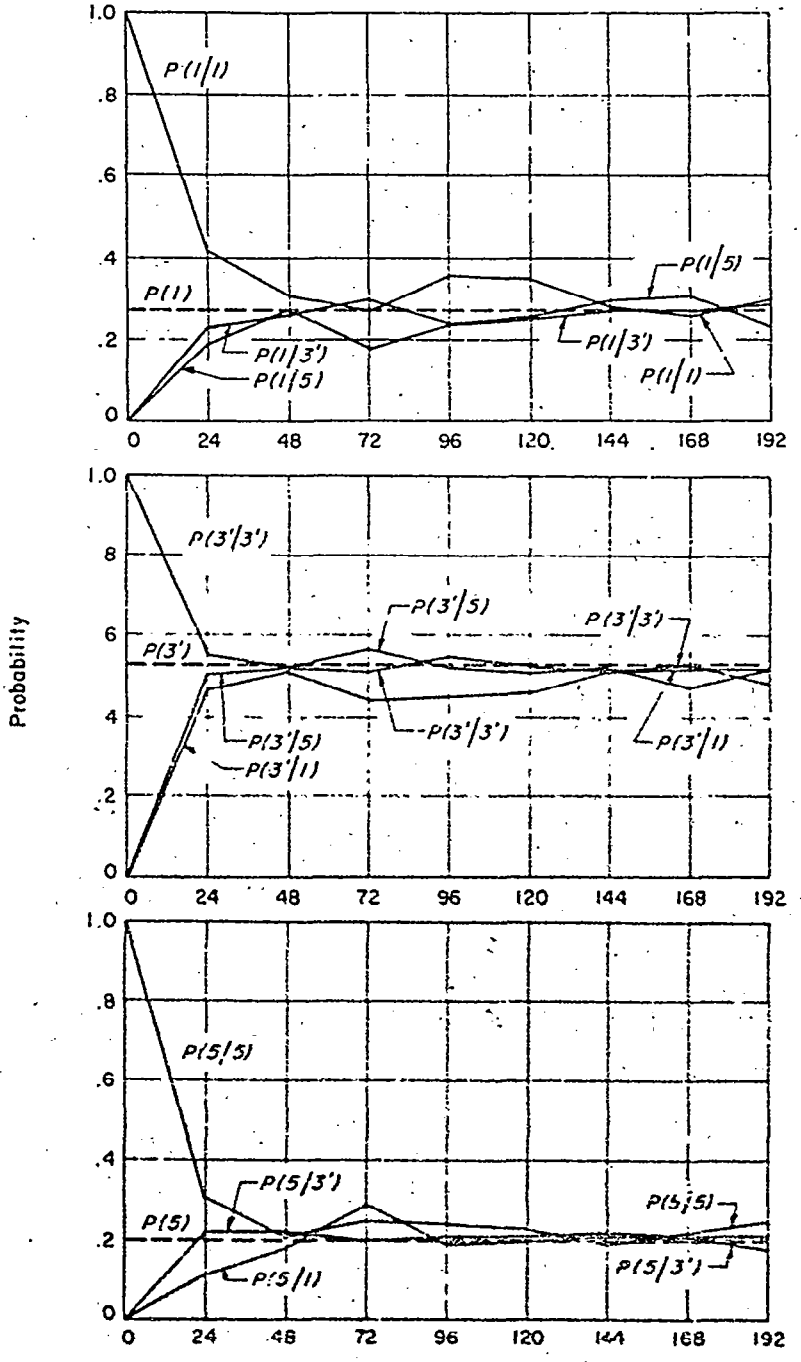


Figure 4-6 Decay of Temporal Conditionals; Region 11, Season 2

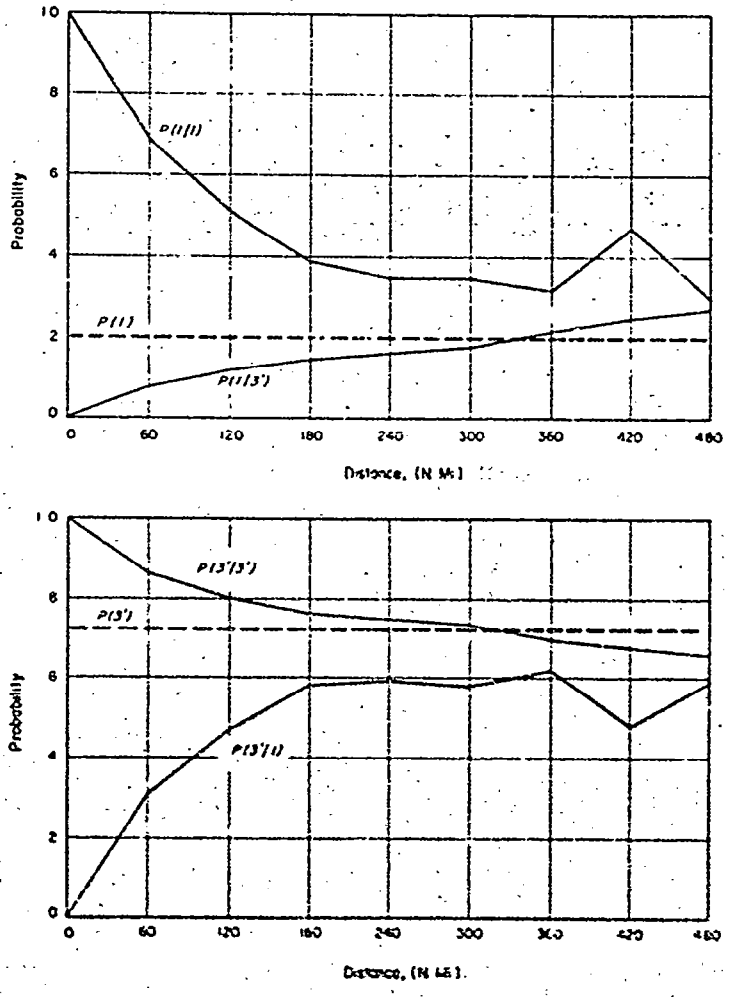


Figure 4-7 Decay of Spatial Conditionals; Region 11, Season 3

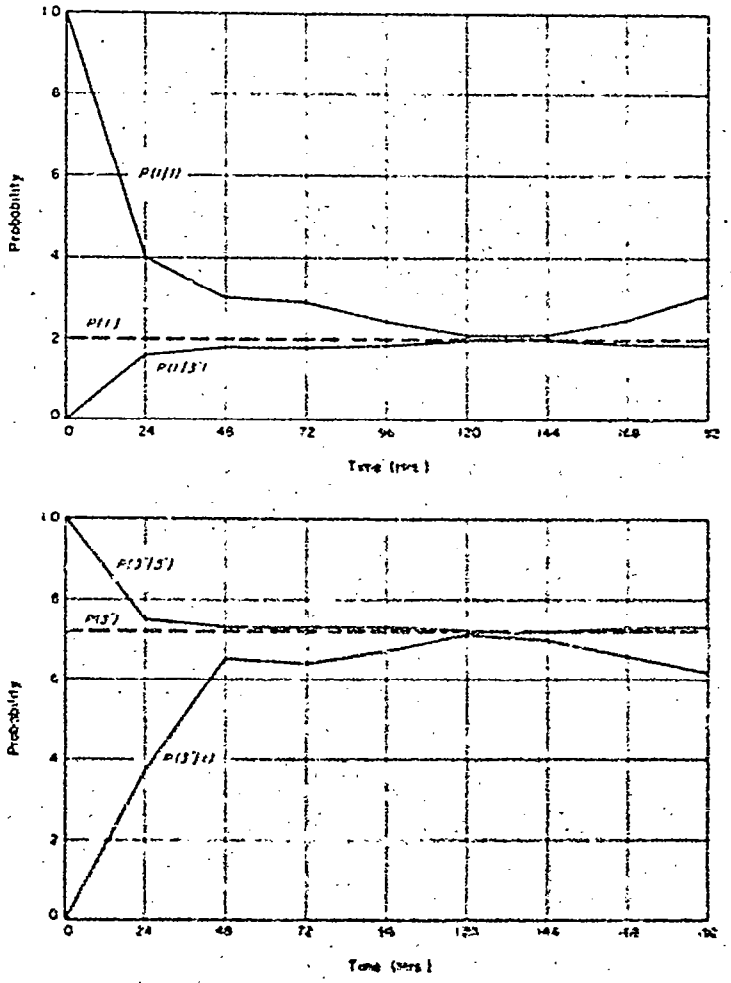


Figure 4-3 Decay of Temporal Conditionals; Region 11, Season 3

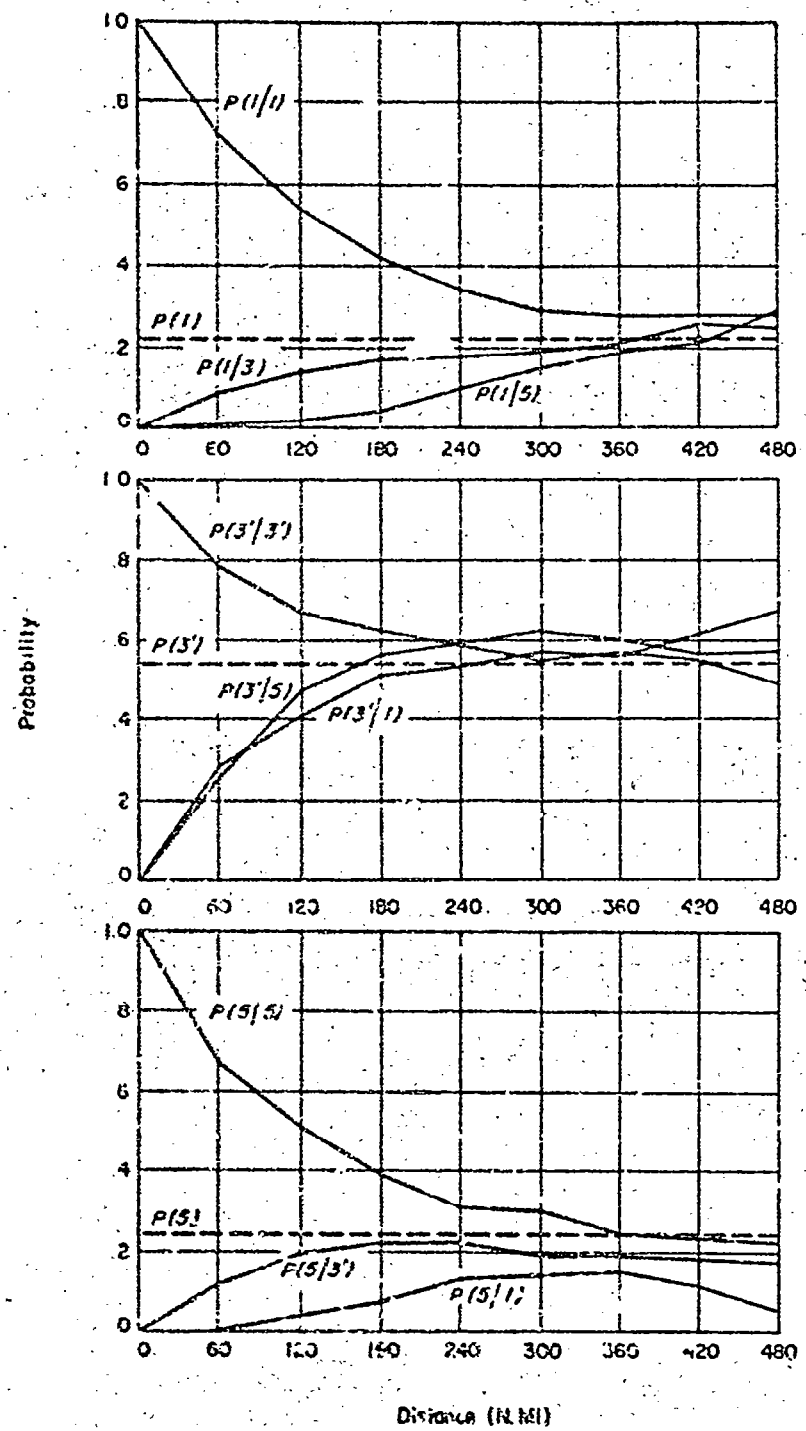
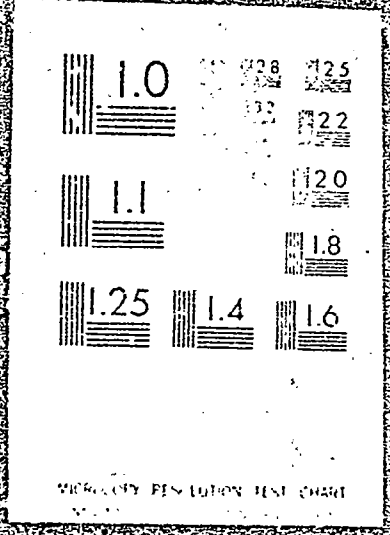


Figure 4-9 Decay of Spatial Conditionals; Region 11, Season 4



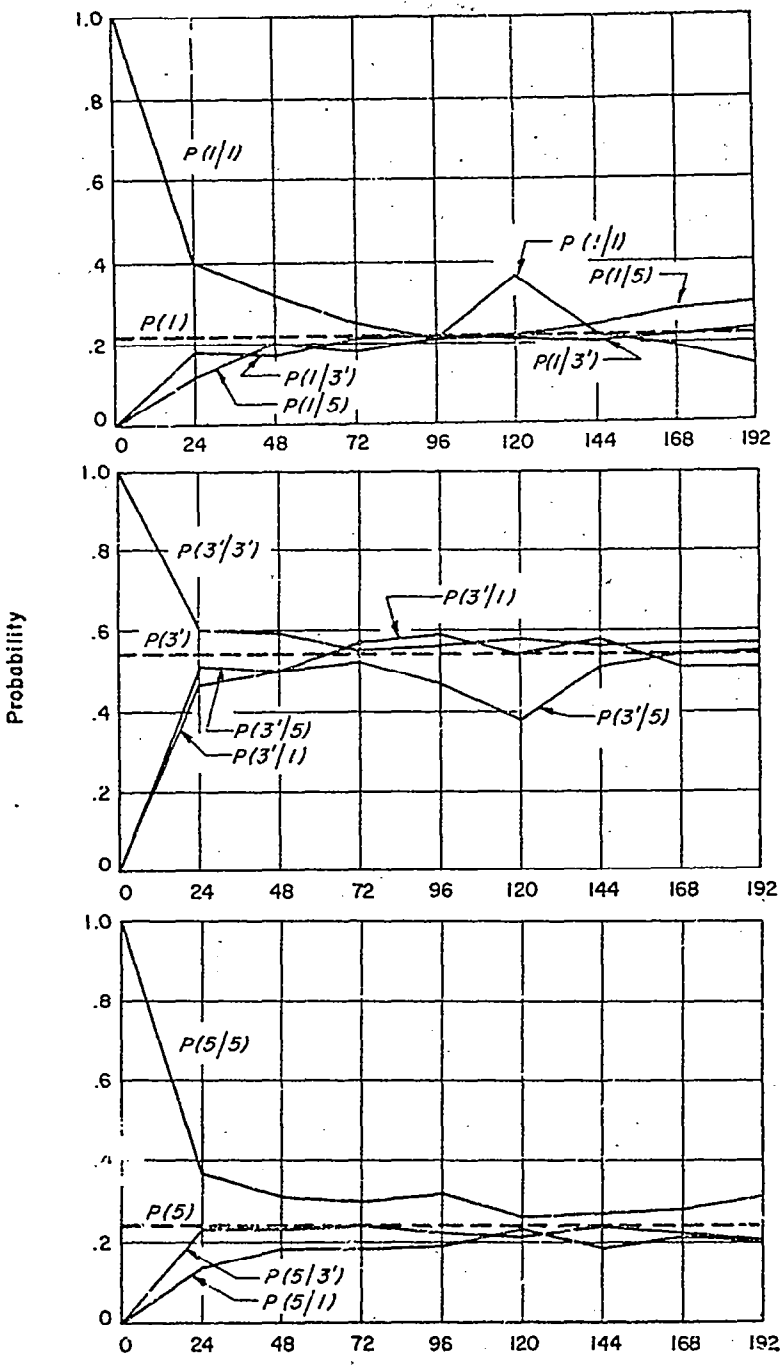


Figure 4-10 Decay of Temporal Conditionals; Region 11, Season 4

was scaled to 48 and 72 hours. These scaled values were then compared with the observed arrays and with values scaled in a linear fashion (including stuffing). Table 4-2 shows the original spatial and temporal conditional arrays for Region 11, Season 1, before scaling.

Tables 4-3 and 4-4 show the detailed scaling results for Region 11, Season 1. It can be seen that in every instance the Markov scaled results represent a measureable improvement of the linear scaling. Table 4-5 lists the per-element (of the three by three array) error made in scaling the conditional arrays by both the Markov and linear techniques. Note that the poorest results occur when scaling from the observed 48-hour conditional array to 24 hours. This is due to the very rapid decay of conditionality to almost the unconditional level before 48 hours. The extrapolation results from 24 hours to 48 hours and 72 hours are much improved.

TABLE 4-2

ORIGINAL CONDITIONAL ARRAYS; REGION 11, SEASON 1

.33	.59	.08
.09	.66	.25
.03	.47	.50

180 n.mi. Spatial Conditional

.34	.56	.10
.07	.66	.27
.03	.52	.45

24 hour Temporal Conditional

.27	.54	.19
.10	.62	.28
.04	.54	.42

48 hour Temporal Conditional

4.1.3 Diurnal Variations

As a part of the previous study, a technique was developed for adjusting the linearly scaled temporal conditional distributions for diurnal variations (Section 6.6 of Sherr et al, 1968). In effect the linearly scaled conditionals were weighted by the change in the unconditional distribution between the two reference times. This satisfies the intuitive notion that diurnal change is superimposed on more gross synoptic scale variability. The same technique can be applied to the new data base, except that now the Markov scaled temporals are adjusted. The following derivation of the diurnal adjustment procedure is essentially a revised version of the original derivation.

TABLE 4-3

DETAILED SPATIAL SCALING RESULTS; REGION 11, SEASON 1

	60 n.mi.	120 n.mi.	240 n.mi.	300 n.mi.
Observed:	.64 .36 0 .05 .83 .12 0 .25 .75	.43 .54 .03 .08 .72 .20 .01 .39 .60	.27 .60 .13 .10 .63 .27 .05 .52 .43	.23 .61 .16 .10 .64 .26 .08 .56 .36
Markov Scaled:	.65 .35 0 .05 .82 .13 0 .25 .75	.46 .52 .02 .08 .71 .21 .01 .39 .60	.22 .64 .14 .08 .64 .28 .04 .52 .44	.17 .64 .19 .08 .63 .29 .05 .55 .40
Linear Scaled:	.77 .20 .03 .03 .89 .08 .01 .16 .83	.56 .39 .05 .06 .77 .17 .02 .31 .67	.08 .60 .32* .08 .60 .32 .08 .60 .32	.08 .60 .32* .08 .60 .32 .08 .60 .32

*Stuffed.

TABLE 4-4

DETAILED TEMPORAL SCALING RESULTS; REGION 11, SEASON I

	24 Hours			48 Hours			72 Hours		
Observed:	.34	.56	.10	.27	.53	.19	.18	.60	.21
	.07	.65	.27	.10	.63	.28	.10	.61	.29
	.03	.52	.44	.04	.55	.42	.04	.57	.39
Markov* Scaled:	.48	.42	.10	.16	.60	.23	.11	.60	.29
	.08	.70	.21	.08	.61	.31	.08	.60	.32
	.01	.42	.57	.06	.59	.35	.07	.60	.33
Linear* Scaled:	.63	.27	.10	.08	.60	.32 **	.08	.60	.32 **
	.05	.81	.14	.08	.60	.32	.08	.60	.32
	.02	.28	.70	.08	.60	.32	.08	.60	.32

*24 hour results scaled from 48 hours; 48 and 72 hour results scaled from 24 hours.

**Stuffed.

TABLE 4-5

ERROR COMPARISON OF LINEAR AND MARKOV SCALING

	Region 4		Region 11									
	Annual		Season 1		Season 2		Season 3		Season 4		Average	
Distance (n. mi.)	Markov	Linear	Markov	Linear	Markov	Linear	Markov	Linear	Markov	Linear	Markov	Linear
60	.038	.107	.004	.069	.012	.062	.017	.089	.025	.069	.018	.072
120	.030	.093	.009	.062	.007	.044	.020	.073	.023	.053		
240	.023	.098	.018	.078	.019	.072	.016	.063	.030	.066	.035	.069
300	.026	.099	.044	.067	.023	.083	.034	.057	.033	.056		
360	.026	.072	.043	.071	.042	.079	.053	.049	.032	.039		
420	.039	.051	.062	.073	.038	.069	.079	.084	.043	.046		
Time (Hrs.)												
24*	.059	.169	.072	.156	.063	.162	.053	.160	.078	.163	.065	.162
48**	.026	.049	.047	.074	.009	.016	.024	.044	.033	.049	.032	.045
72**	.031	.037	.037	.046	.048	.046	.040	.056	.023	.029		

*Scaled from 48 hours.

**Scaled from 24 hours.

A joint probability distribution is first formed between $P_A(i)$, the unconditional cloud-amount distribution at time A, and $P_B(i)$, the unconditional distribution at time B. The assumption is made that an event at time B corresponding to a specific event at time A is the one occurring at the same cumulative probability level. The situation may be clarified by referring to Figure 4-11. Here the unconditional distributions for times A and B are represented graphically. Category 1 (clear skies) occurs with a relative frequency of 20% at time A and 30% at time B. Our assumption implies that an event of clear skies at time B occurs on every occasion that Category 1 is observed at time A, and approximately 20% of the time that Category 2 is observed at time A. If either of Categories 3, 4 or 5 are recorded at time A, Category 1 cannot be recorded at time B. Hence the equation:

$$P_B(1) = P_A(1) + .2\, P_A(2). \tag{4.4}$$

The remaining four equations in Figure 4-11 are determined in a similar manner. It is the coefficients of these five equations that are used to "weight" the scaled temporals.

These coefficients may also be determined in a nongraphical manner. The cloud categorization intervals fall at different cumulative probabilities in the distributions of events at times A and B. Thus it is necessary to divide up the intervals of the distribution at time A and assign them to intervals of the distribution at time B, assuming uniform distribution within an interval. To form the joint probability matrix shown in Table 4-6b we find the fractional part of $P_A(1)$ that is contained in (jointly distributed with) $P_B(1)$. In the example in Table 4-6a, all of $P_A(1)$, 0.2, is contained in $P_B(1)$. Thus, 0.2 is entered in the joint probability matrix at position A = 1, B = 1 (cell number of joint table). Since $P_B(1)$ is greater than $P_A(1)$, the additional 0.1 in $P_B(1)$ could not have occurred jointly with $P_A(1)$. Therefore, it is placed in the joint probability matrix at position A = 2, B = 1.

In a similar way, we rate (jointly distribute) $P_A(2)$ with $P_B(2)$ and find that only 0.3 are contained in both. Therefore, 0.3 is located in the joint matrix at A = 2, B = 2. Again there is an additional part to be allocated; this time 0.1 of $P_A(2)$ must have occurred with $P_B(3)$; it is thus entered in the matrix at A = 2, B = 3.

This process is continued for all categories as shown. These individual entries, divided by the marginal total become the entries in a pseudo conditional, PSCON (B|A). Note that the PSCON columns are the coefficients of the equations

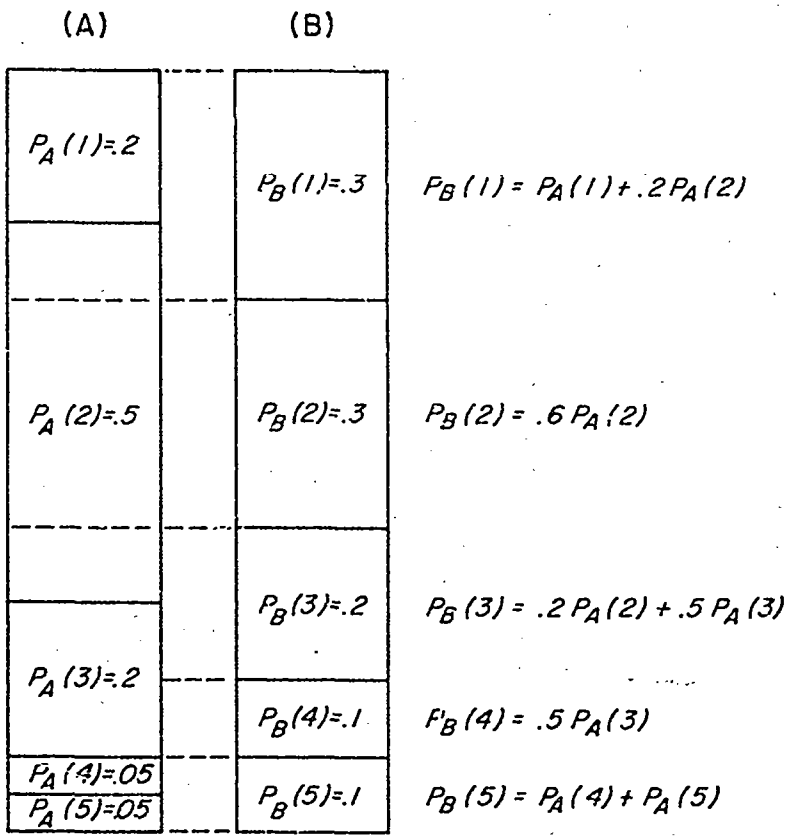


Figure 4-11 Diurnal Correspondence Between Times A and B

TABLE 4-6

COMPUTATION OF A PSUEDO-CONDITIONAL DISTRIBUTION FOR DIURNAL VARIATION

(a)

Cloud Category	Time (A) Probability	Time (A) Rated Probability	UNCON Joint Cell Number	Time (B) Rated Probability	Time (B) Probability
1	.2	.2	1-1	.2	.3
			2-1	.1	
2	.5	.1			.3
		.3	2-2	.3	
		.1	2-3	.1	
3	.2	.1	3-3	.1	.2
		.1	3-4		
4	.05	.05	4-5	.1	.1
				.05	
5	.05	.05	5-5	.05	.1

(b)

JOINT PROBABILITY

(A) \ (B)	1	2	3	4	5	Total
1	.2	0	0	0	0	.2
2	.1	.3	.1	0	0	.5
3	0	0	.1	.1	0	.2
4	0	0	0	0	.05	.05
5	0	0	0	0	.05	.05

(c)

PSCON (B/A)

(A) \ (B)	1	2	3	4	5
1	1.0	0	0	0	0
2	.2	.6	.2	0	0
3	0	0	.5	.5	0
4	0	0	0	0	1.0
5	0	0	0	0	1.0

derived in Figure 4-1. A subroutine to calculate the PSCON matrix from the original unconditional distributions is listed in Section B.2 of Appendix B.

Finally, we form the diurnally adjusted conditional distribution, DITCON (i|j) by:

$$\text{DITCON}(i|j) = \sum_k \text{PSCON}(i|k) \cdot \text{TCON}(k|j) \qquad (4.5)$$

where

TCON (k|j) is the Markov scaled temporal conditional.

Boundary Crossing

It is worth noting that exactly the same procedure may be used to simulate a satellite crossing from one homogeneous cloud region to another. In this case, the PSCON is derived from the unconditionals (at the appropriate time) in Regions A and B. The TCON is replaced by SCON, the spatial conditional array (possibly Markov scaled) defined in Region A. The final DITCON is then the spatial conditional which links Regions A and B.

4.2 Area Adjustment

4.2.1 Enlarging the Representative Area Size

As was noted in Section 2.2.3, the original area enlargement procedure relied upon a number of assumptions, including a linear decay of conditionality and a one-directional variability of cloud amounts. In this section we derive a new area enlargement procedure which avoids the necessity of these assumptions. From at least a mathematical point of view, the new procedure is more easily defended than its predecessor. It yields results which are consistent with our intuitive notions of how statistical distributions should change with increasing area size, and when applied to particular cases, its performance represents a measurable improvement over the original technique. On the minus side, the newly developed technique essentially carries out a repetitive doubling of the original (60 n.mi. diameter) area size so that frequency distributions of cloud amount for intermediate sized areas must be found through element by element interpolation (linear with area). In the next section we shall discuss a somewhat more sophisticated technique which can be used to enlarge or diminish the representative area size of the unconditional distributions. We hope at some future point to apply the techniques of the second method to area adjustment of conditional distributions as well, and thereby produce a more flexible "third generation" area adjustment procedure.

Derivation of Enlargement Procedure

As an example of the new enlargement procedure, we shall outline the steps which must be taken to increase the representative area size of any set of conditional and unconditional data. Consider the situation depicted in Part 1 of Figure 4-12. Two areas, a and b, are separated by an amount X in space or time. Because the conditional data are tabulated for a 60 n.mi. diameter area, we shall begin with a 60-mile area in our example. The tabulated 30-mile ground-observed unconditionals can be converted to 60 mile satellite unconditionals using the Transfer Matrix derived in Section 3.4). We wish to derive the unconditional cloud-amount distribution for new enlarged (doubled) area A, and the conditional relationship (temporal and/or spatial) between B and A.

We know, or can readily calculate the following:

$P(a)$	The unconditional cloud-amount frequency distribution for area a. Assuming that a and b are within the same region and/or month, $P(a) = P(b)$.
$P_X(b\|a)$	This may either be taken directly from the Revised Data Bank (if X = 24 hours or 200 miles) or scaled to the proper time or distance, X.
$P_{60}(c\|a)$ $P_{60}(d\|b)$	Obtained by scaling the 200-mile spatial conditional array in the Revised Data Bank to 60 n.mi.

We wish to find:

$P(A)$	The unconditional distribution for enlarged area A, and
$P_X(B\|A)$	the conditional relationship between B and A. Again we assume $P(A) = P(B)$.

Using the Markov dependency assumption, we can write:

$$P(a, b, c, d) = P(a)\, P_{60}(c|a)\, P_X(b|a)\, P_{60}(d|b) \qquad (4.6)$$

where

$P(a, b, c, d)$ is the joint probability of events in all four areas.

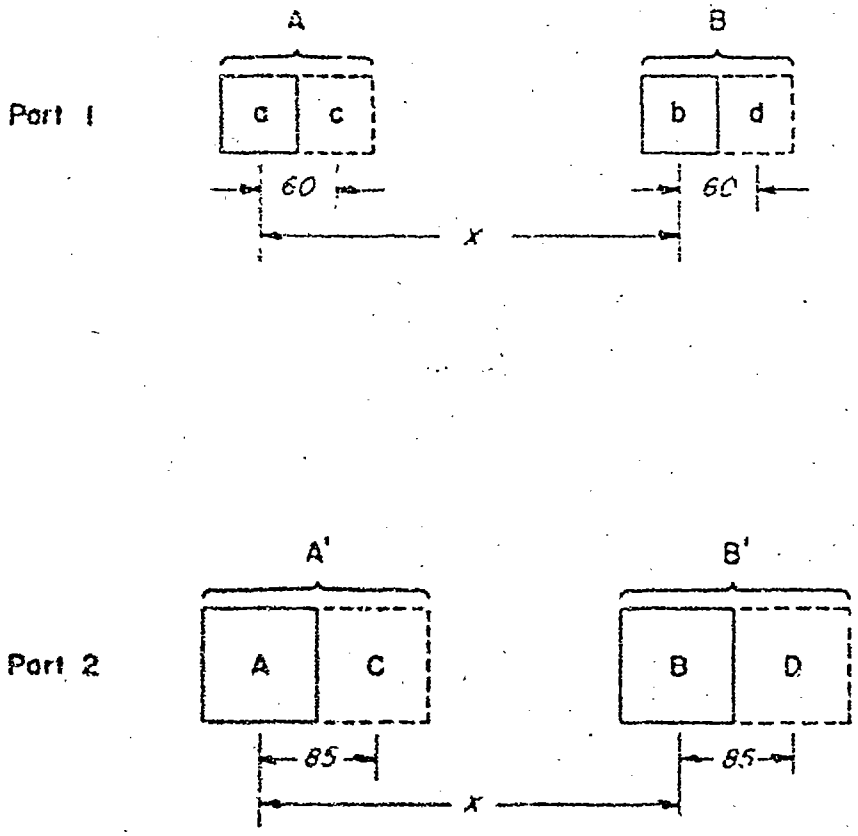


Figure 4-12 Area Enlarge Configuration

Defining the cloud cover in A to be the average of the cloud cover a and c while the cloud cover in B is the average of b and d, we may write:

$$P(A, B) = P(\overline{ac}, \overline{bd}) \tag{4.7}$$

To find $P(\overline{ac}, \overline{bd})$, the KWHERE location matrix shown in Table 4-7 and defined in the previous study is used four-dimensionally.

TABLE 4-7

CLOUD GROUP LOCATION MATRIX

(KWHERE)

Cloud Group	1	2	3	4	5
1	1	2	2	3	3
2	2	2	2	3	4
3	2	2	3	4	4
4	3	3	4	4	4
5	3	4	4	4	5

(The KWHERE matrix simply indicates the appropriate cloud category for an area composed of two equal areas whose cloud categories are the row and column margins of the KWHERE matrix.) The desired unconditional distribution may be found from a row sum of the resultant P(A, B).

$$\text{i.e.} \quad P(A) = \sum_{R} P(A, B). \tag{4.8}$$

Note that with internally consistent data we would have:

$$P(A) = P(B) = \sum_{C} P(A, B). \tag{4.9}$$

Finally, we can determine the desired $P_X(B|A)$ by dividing through the joint distribution by the row sum.

$$P_X(B|A) = \frac{P(A, B)}{P(A)}. \tag{4.10}$$

A subroutine which takes the input parameters on the right side of Eq. (4.6) and returns with the desired P(A) and $P_X(B|A)$ is listed in Section B.3 of Appendix B.

As an aside, we shall simply list here a Monte Carlo procedure which, if repeated often enough, will duplicate the four-dimensional application of the KWHERE matrix.

1) Average the P(a) distribution in a cumulative format and set P(A, B) = 0.

2) Select a cloud cover for $\underline{a}$ by choosing a random number in the interval 0 to 1 and entering P(a).

3) Determine cloud covers for $\underline{b}$ and $\underline{c}$ from $P_X(b|a)$ and $P_{60}(c|a)$.

4) Determine cloud cover for $\underline{d}$ from $P_{60}(d|b)$.

5) Determine $A = \overline{ac}$ and $B = \overline{bd}$ from KWHERE matrix.

6) Add <u>one</u> to the appropriate A, B location in P(A, B).

7) Repeat Steps 1 through 6 many times.

8) Divide the final P(A, B) by the row sum P(A) to get $P_X(B|A)$.

We have now doubled the representative area size of $P_X(b|a)$ to yield $P_X(B|A)$. If we now wish to redouble the area size, it makes a difference whether "X" in Figure 4-12 represents a distance or a time separation. If a distance, the procedure is straightforward. Scale the new $P_X(B|A)$ to a distance $\sqrt{2} \times 60 \cong 85$. (The radius or diameter increases as the square root of the area.) Again using the Markov dependency assumption, we can write (see Part 2 of Figure 4-12):

$$P(A, B, C, D) = P(A)\, P_{85}(C|A)\, P_X(B|A)\, P_{85}(D|B) \qquad (4.11)$$

where now:

$$P(A', B') = P(\overline{AC}, \overline{BD}). \qquad (4.12)$$

Eq. (4.11) and (4.12) are identical in form to Eq. (4.6) and (4.7) so that we can now continue as before.

If "X" represents a time separation, an intermediate step is required. Before we can calculate either $P_{85}(C|A)$ or $P_{85}(D|B)$ (numerically identical) in Eq. (4.11) above, we must first enlarge $P_{200}(b|a)$ (from the Revised Data Bank) to $P_{200}(B|A)$ and then scale down to 85 n.mi. Thus, enlarging temporal conditionals beyond the first doubling requires the simultaneous enlargement of spatial conditionals.

Analysis of Results

The newly derived area enlarging procedure was used to triple the representative area size of the unconditional and temporal conditional distributions for Regions 4 and 11. These data were stratified by year and season, respectively. (The same seasonal definition as used in Section 4.1.2 above.) The results were then compared with observed values and with similar results derived from the previous area enlargement technique. To simulate a data base, the nine data extraction circles were grouped into three sets of three contiguous circles, and new unconditional and temporal conditional data were tabulated. (Due to the unusual shapes of the new areas, calculations of spatial conditionals were not attempted.)

Table 4-8 compares the performance of the two area enlargement techniques in simulating unconditional distributions for an enlarged area for both regions and all seasons. On the average, the per-element error using the new procedure is only one-third as large as that using the old method. In Table 4-9, the detailed results of area enlarging the 24-hour temporal conditional for Region 11, Season 1, are presented. The mean per-element errors for both methods are listed in Table 4-10 by region and season. The new results represent a modest (25%) improvement over the earlier method. Some extenuating circumstances are worth mentioning:

1) The oblong shape of the three contiguous data circles may not be entirely representative of a more uniform tripling of area size.

2) Due to assumptions inherent in the method, the original area-enlarge technique should perform the best for small area increments and become worse as the representative area increases. The new method is not limited by the same assumptions.

3) An exact solution is probably not possible. In any event, a more accurate approach would have to take into account the spatial distributions of the predominant cloudiness.

TABLE 4-8

AREA ENLARGING OF UNCONDITIONALS

	1	2	3	4	5	E
A	.45	.37	.08	.09	.01	-
B	.26	.56	.10	.08	0	-
C	.34	.48	.10	.07	.01	.036
D	.32	.52	.08	.08	0	.024

Region 4, Annual

	1	2	3	4	5	E
A	.08	.13	.08	.39	.32	-
B	.03	.14	.12	.52	.19	-
C	.05	.12	.16	.44	.23	.036
D	.04	.13	.13	.50	.20	.012

Region 11, Season 1

	1	2	3	4	5	E
A	.28	.15	.07	.30	.20	-
B	.19	.22	.11	.37	.11	-
C	.20	.17	.17	.31	.15	.044
D	.18	.21	.13	.36	.12	.012

Region 11, Season 2

	1	2	3	4	5	E
A	.20	.27	.13	.32	.08	-
B	.10	.34	.18	.36	.02	-
C	.13	.27	.22	.33	.05	.040
D	.10	.34	.17	.36	.03	.004

Region 11, Season 3

	1	2	3	4	5	E
A	.22	.14	.11	.29	.24	-
B	.12	.21	.12	.44	.11	-
C	.15	.16	.19	.34	.16	.060
D	.12	.22	.16	.39	.11	.010

Region 11, Season 4

Legend

A Original unconditional distribution

B Unconditionals for enlarged area

C Results of old technique

D Results of new technique

E Mean per-element error

TABLE 4-9

AREA ENLARGING OF TEMPORAL CONDITIONALS; REGION 11, SEASON I

Original 24-Hour TCON :

.34	.20	.08	.28	.10
.14	.21	.09	.33	.23
.09	.18	.16	.33	.24
.05	.14	.09	.42	.30
.03	.07	.05	.40	.45

Enlarged 24-Hour TCON :

.16	.40	.16	.28	0
.08	.29	.13	.38	.12
.03	.22	.16	.51	.08
.02	.10	.13	.58	.17
.01	.04	.06	.53	.36

Result, Old Technique :

.21	.23	.21	.27	.08
.08	.18	.19	.43	.12
.07	.14	.18	.44	.17
.04	.10	.14	.47	.25
.02	.08	.13	.44	.33

Mean Per-Element Error = .0536

Result, New Technique :

.16	.24	.17	.36	.07
.07	.19	.16	.48	.10
.06	.16	.15	.49	.14
.03	.11	.12	.54	.20
.01	.07	.10	.56	.26

Mean Per-Element Error = .0424

TABLE 4-10

ERROR COMPARISON OF OLD AND NEW AREA ENLARGE PROCEDURES

Regions	Old Technique	New Technique
Region 4, Annual	.0410	.0340
Region 11, Season 1	.0536	.0424
Region 11, Season 2	.0552	.0280
Region 11, Season 3	.0420	.0372
Region 11, Season 4	.0708	.0516

4.2.2 Area Adjustment of Unconditional Distributions

The procedure described in the previous section provides for the calculation of probabilities for areas larger than the observed area, but does not permit calculations for areas smaller than an observed area. The following more approximate procedure was therefore developed which provides estimates of unconditional probabilities for areas either larger or smaller than the observed area.

The procedure is based upon a replacement of the original discrete cloud-amount probability distribution function (as represented by the five cloud categories) by a continuous distribution function. Some assumption is then made as to the variation of this function with changing area size. We have selected a simple normal probability curve, with mean μ and standard deviation σ, as the continuous distribution function. There are a number of reasons for this choice.

1) It can be shown that the assumption of a normal distribution is strictly true for cloud sizes smaller than the field of view. (It will be seen that the adopted technique produces acceptable results even for U or J shaped initial conditions.)

2) The normal probability curve is well known, widely tabulated and requires only two parameters to be completely specified.

Figure 4-13 illustrates the application of the normal approximation to cloud-amount distributions. In this representation, the area I to the left of the $x = 0$ boundary is interpreted as the probability of zero cloud cover P_1 and the area II to the right of the $x = 1$ boundary is interpreted as the probability of 100% cloud cover P_5: the probabilities of intermediate degrees of cloud cover P_2, P_3 and P_4 are obtained

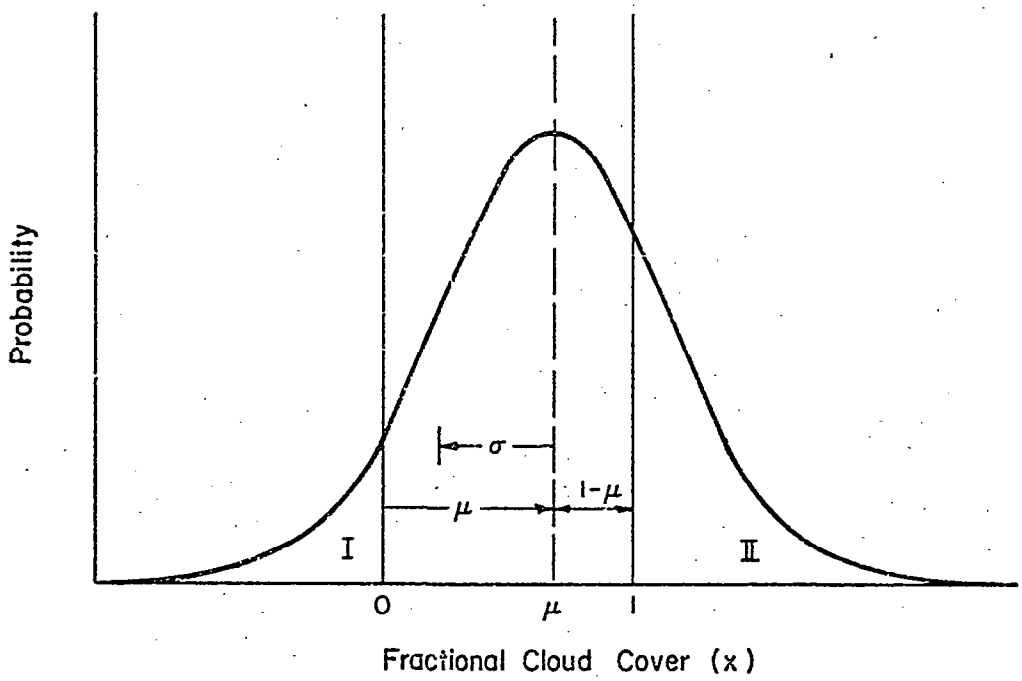


Figure 4-13 Analytical Representation of the Probability of Fractional Cloud Cover

by integration of the normal probability curve over the appropriate ranges of x. The degree of realism of this representation is indicated in Figure 4-14, which presents a typical comparison of actual probabilities for the various cloud cover groups with calculated probabilities, based on representation of these observed values by a normal probability curve.

As the second point of this area increase/decrease procedure, consider the variation with area change of the probability P_α of the cloud group composed of all cloud groups neither completely clear nor completely overcast. (i.e., $P_\alpha = P_2 + P_3 + P_4$.) The probability of this cloud group will obviously increase with increasing area size. This increase will be linear with increasing diameter of area so long as the observation area is not too large compared with individual cloud areas. These considerations suggest that the variation of P_α with the diameter D of an observed area could be represented as

$$dP_\alpha/dD = P_\alpha/D. \tag{4.13}$$

Eq. (4.13) must be modified, however, for large observation areas (where P_α becomes large) to conform to the limiting condition that P_α must approach but not exceed unity. This can be reasonably well accomplished by the following simple modification of Eq. (4.13):

$$dP_\alpha/dD = (P_\alpha/D)(1 - P_\alpha) \tag{4.14}$$

which has the solution

$$\left.\begin{aligned} P_\alpha &= R/\left[R + (1 - P_{\alpha o})/P_{\alpha o}\right] \\ \text{where} \qquad R &= D/D_o \end{aligned}\right\} \tag{4.15}$$

and the subscript o refers to starting or reference conditions.

The above results are utilized as follows. From a given set of unconditional probabilities for an initial diameter D_o, parameters of the corresponding normal distribution (μ and σ) and $P_{\alpha o}$ are calculated. Then, from Eq. (4.15), for any other larger or smaller diameter D, the parameter P_α is calculated. The corresponding standard deviation of the appropriate normal probability curve is then related to P_α by the equation

$$2P_\alpha = \operatorname{erf}\left[\mu/(\sqrt{2}\sigma)\right] + \operatorname{erf}\left[(1 - \mu)/(\sqrt{2}\sigma)\right] \tag{4.16}$$

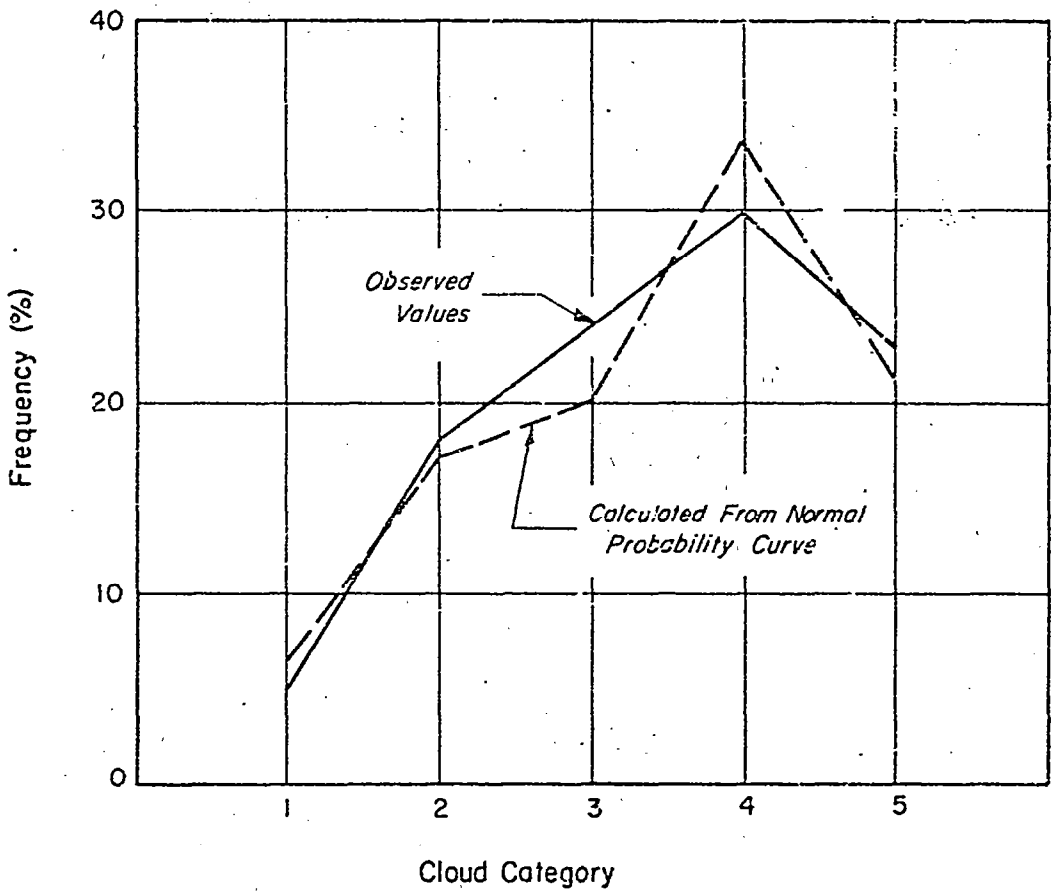


Figure 4-14 Comparison of Observed and Normal Probability Curve Distributions

where erf represents the error function, defined as the integral over a standardized normal distribution:

$$\text{erf}(x) = \left(2/\sqrt{\pi}\right)\int_0^x \exp(-t^2)\, dt$$

This equation can then be easily solved numerically for σ. Finally, values of P_1, ... , P_5 for the new area size are obtained by using this value of σ together with tables of the normal probability function or error function. (We have programmed and used this procedure in the Telcomp computer language.)

Some indication of the accuracy of the above procedure is given in Figure 4-15 for a 25:1 area increase ratio (corresponding to a diameter ratio of $5^o/1^o$) and in Figure 4-16 for a 25:1 area decrease ratio, based on data compiled during the previous study. Also shown in Figure 4-15 are the original unconditional distributions for a 1^o area. (It is interesting to note the transition in the original distribution shape from Region 9 (high latitude) to 11 (mid-latitude) to 19 (sub-tropical) to 4 (tropical) as the curve passes from U to bell shape. This effect was noted previously in Section 2.2.2.) It may be noted that for all cases the area scaled data are in fair to good agreement with the observed data, both for area increase and area decrease ratios of 25:1.

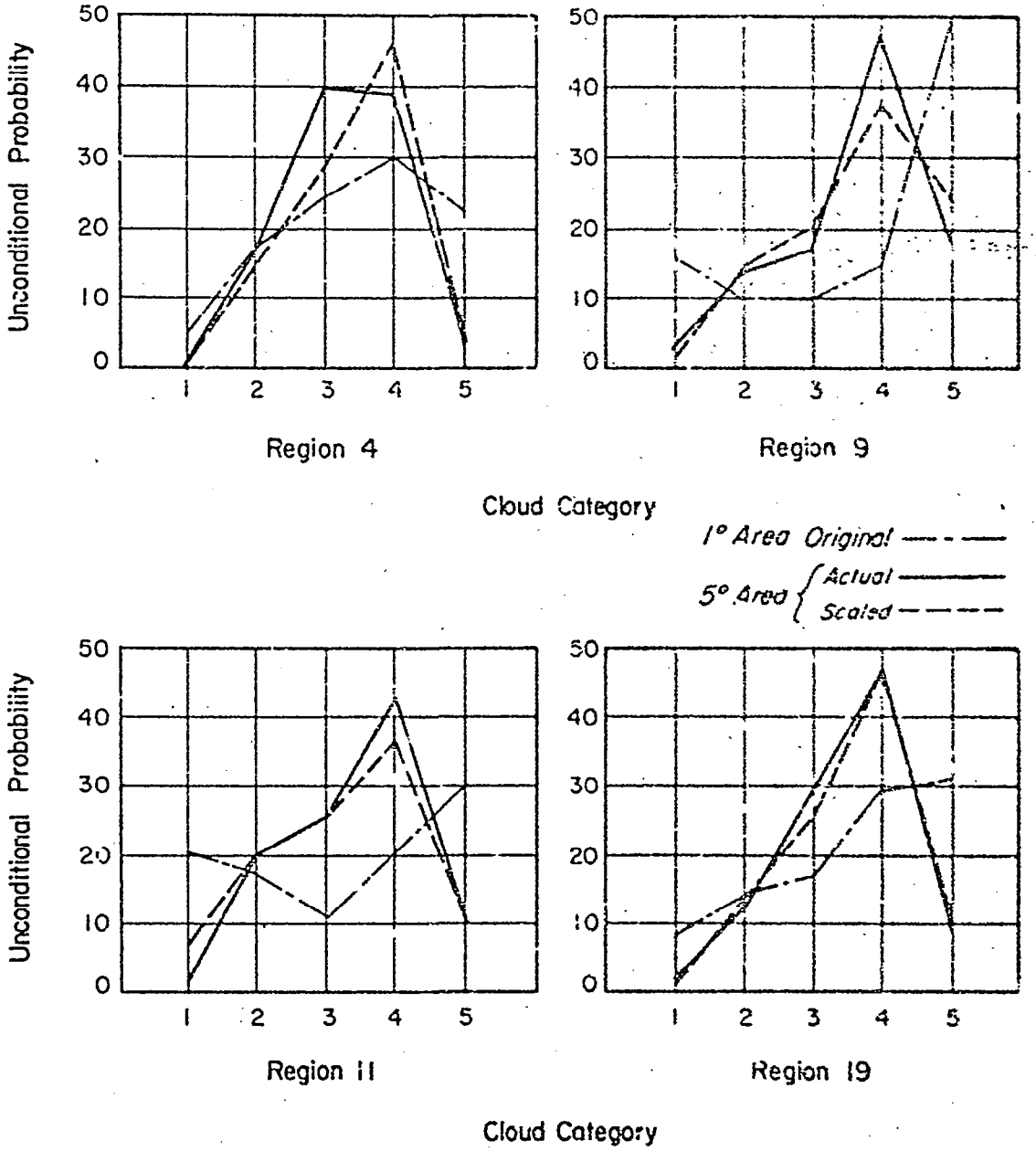


Figure 4-15 Comparison of Unconditional Probabilities for a 25:1 Area Increase

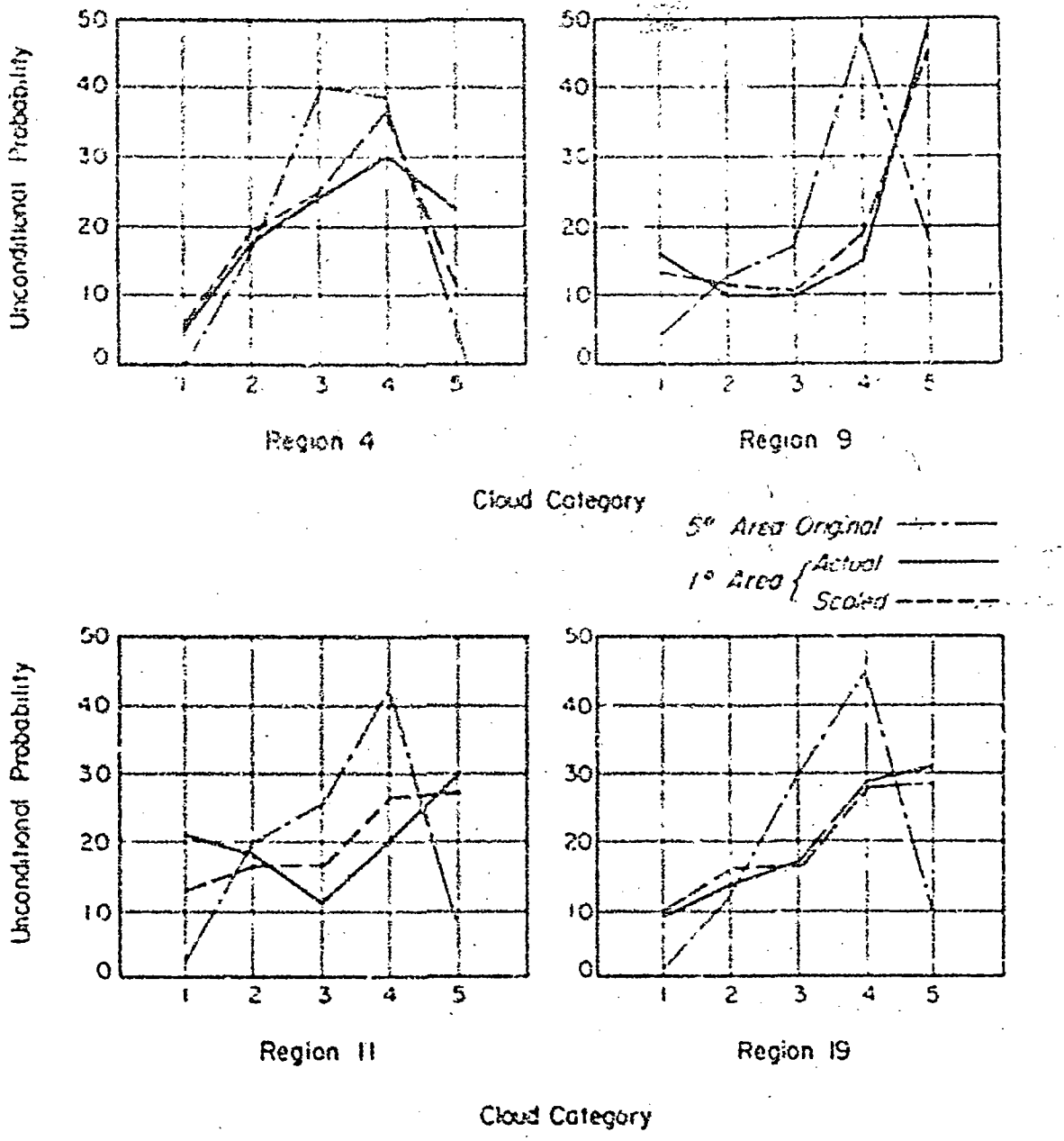


Figure 4-16 Comparison of Unconditional Probabilities for a 25:1 Area Decrease

PRECEDING PAGE BLANK NOT FILMED

5. SIMULATION PROCEDURES

The basic structure of a Monte Carlo simulation procedure was outlined in Section 2.3.1. Although designed specifically to simulate repeated looks at a given area, the same general procedure is readily adaptable to more complex situations such as simulations of cloud cover in orbital swaths, or varying time intervals between observations. An example of one of the many possible presentation formats was presented in Figure 2-3. In that figure, the percent of cloud-free coverage is plotted against the cumulative probability of achieving that coverage as a function of the number of satellite passes or observations. In reviewing results such as these, one important fact should be kept in mind: What we are really simulating is the potential performance of a vidicon camera system with the spectral and spatial resolution characteristics of the data source - i.e. ESSA or Nimbus. Ways must be found to adjust these results for varying sensor types (visible, infrared, microwave) and for varying spatial resolutions.

By and large, the adjustments for different sensor types must await the inclusion in the cloud model of other data types such as cloud height and thickness for infrared measurements and cloud water content and drop size for microwave measurements. The adjustments for varying spatial resolutions also require information as to the horizontal distribution (spatial frequencies) of cloud cover and therefore of cloud type, but here at least, some preliminary ground work has already been done.

Shenk and Salomonson (1971) have studied the effects of sensor spatial resolution on satellite estimates of cloud cover using simulated cloud data. Their results are presented in terms of the ratio, R, of areal cloud size to areal resolution element size. Having some estimate of the value of this ratio is important in simulation work, since as the resolution element of a sensor becomes smaller, more and more of the available clear areas become usable. (For the purposes of this discussion, we define a "usable" data element as being totally cloud-free. Obviously for certain applications this is unrealistic, but it will serve to clarify the nature of the relationship between the sensor resolution element and mean cloud size or its inverse, the mean "hole" size.) Beyond a certain point - i.e., when the resolution element reaches the size of the smallest clear area - nothing will be gained from a further improvement in sensor resolution.

In order to demonstrate how the ratio R can affect the Monte Carlo simulation results, we used the curves presented in Figure 5-1 which show the rela-

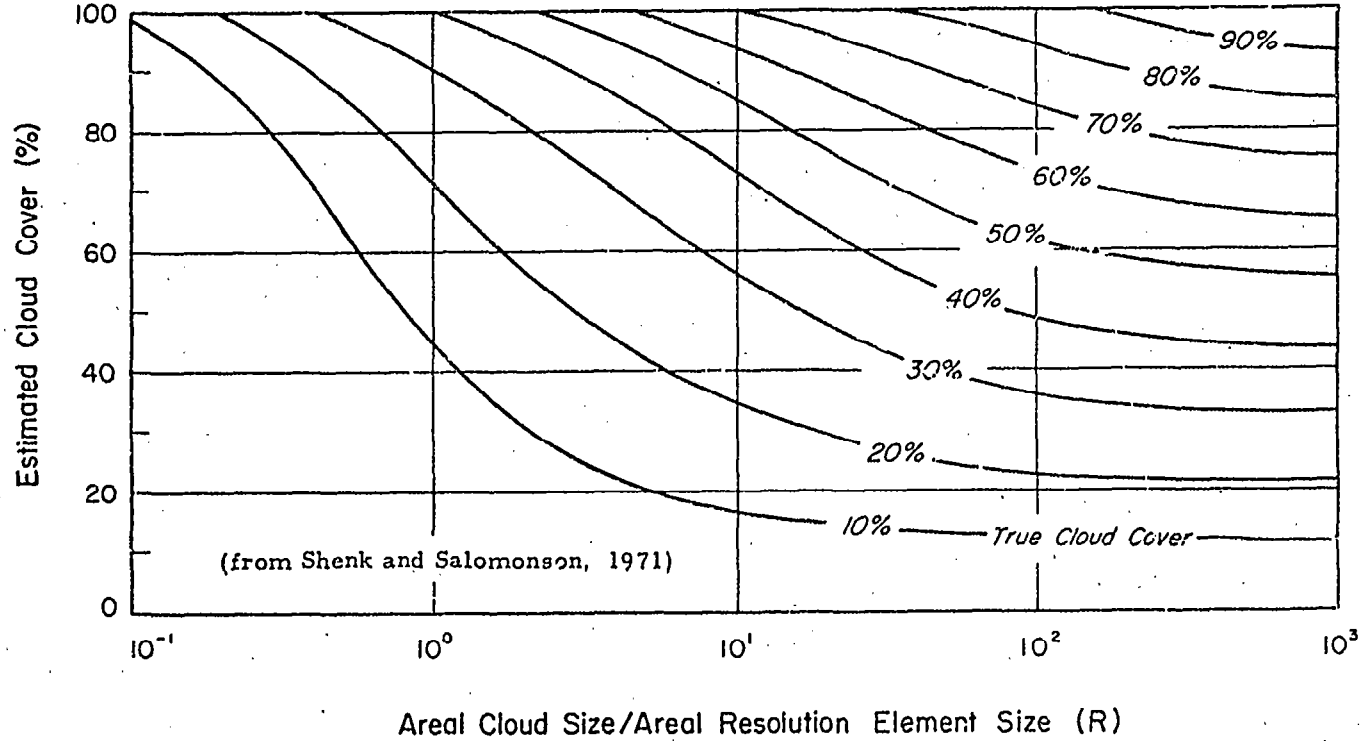


Figure 5-1 Estimated and True Cloud Cover

tionship between true and estimated cloud cover as a function of R. (Figure 5-1 is taken from a preliminary version of the Shenk and Salomonson report.) This nomogram was generated using three simulated cloud patterns: (1) a regularly spaced pattern of homogeneous dots arranged in rows and columns, (2) a randomly arranged, irregularly spaced pattern of the same dots, and (3) a heterogeneous cloud size distribution irregularly spaced. From this figure, it can be seen that with a resolution element one-tenth the area of the mean cloud size (R = 10), a true 30% cloud cover has the same effect as a 58% cloud cover. Thus, with R = 10 only 42% of the resolution elements will be usable (cloud-free), even though 70% of the total area is cloud-free.

In Figure 5-2 we have plotted (R = 10 curve) the anticipated coverage of a one-pass mission over an area with an unconditional cloud-amount frequency distribution as indicated in Table 5-1. The unconditional distribution P(i) is first arranged in a cumulative format and then plotted against the corresponding (cumulative) coverage.

TABLE 5-1

ONE-PASS SIMULATION

Cloud Group	Cloud Amounts (1/10ths)	Corresponding Coverage (%)	P(i)	Cumulative Probability	Cumulative Coverage (%)
1	0	100	.20	.20	100
2	1, 2, 3	70 - 90	.13	.33	≥ 70
3	4, 5	50 - 60	.07	.40	≥ 50
4	6, 7, 8, 9	10 - 40	.24	.64	≥ 10
5	10	0	.36	1.00	≥ 0

Note that by a simple binomial expansion of the points in this curve we can simulate the probabilities of "single looks" after any number of passes. For example, if we have a 33% chance of obtaining 70% or more coverage for one pass, we can calculate the probability of obtaining that same coverage on at least one pass after N passes as follows:

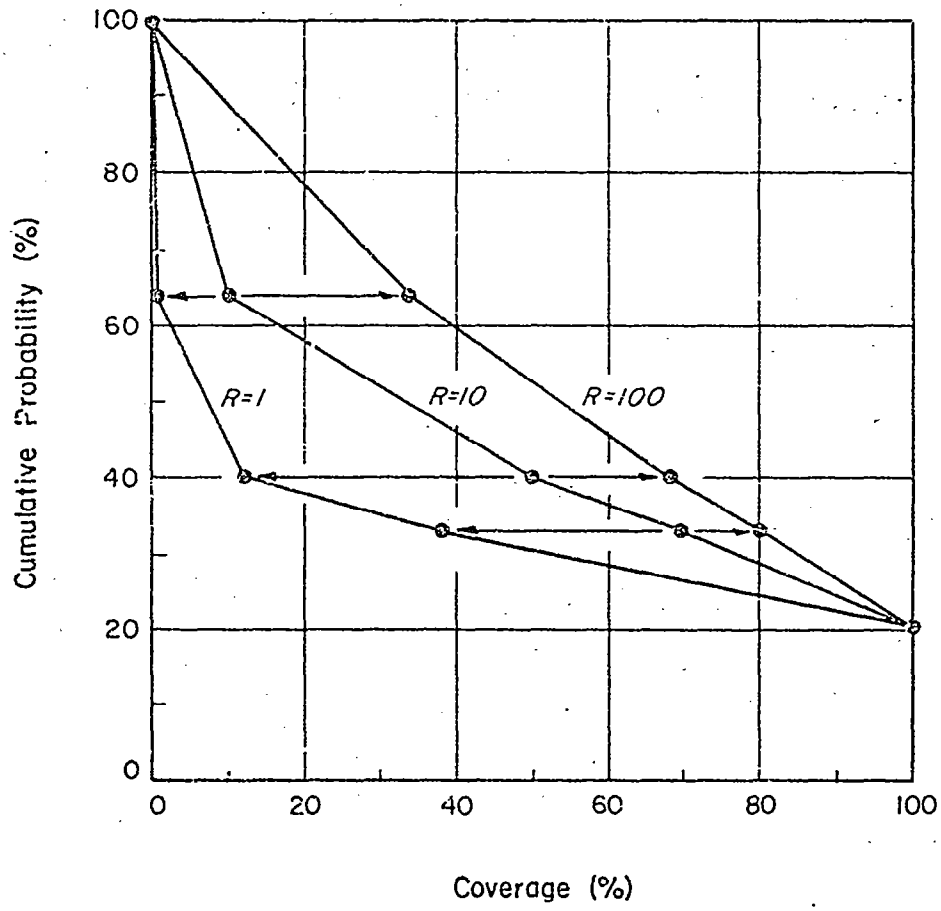


Figure 5-2 Changes in Simulation Results as a Function of R

$$P = \sum_{k=1}^{N} \frac{N!}{k!\,(N-k)!} (.33)^k (1 - .33)^{n-k} \qquad (5.1)$$

where

P is the desired probability.)

We have arbitrarily assumed that R = 10 for the mission being simulated. We now wish to examine the anticipated change in simulation results for an order of magnitude increase of decrease in the R ratio. (A $\sqrt{10}$ change in sensor resolution.) These changes are shown in Figure 5-2. Consider the coverage = 50% point in Figure 5-1 as an example. The R = 10 and Estimated Cloud Cover = 50% intersection point occurs at a True Cloud cover value of approximately 27%. Following the 27% line back to R = 1 or forward to R = 100, one can read the new estimated cloud cover. The corresponding coverage (one minus the cloud cover) is then plotted in Figure 5-2. As can be seen, the spatial resolution can significantly affect the simulation results.

In the discussion above we have made some fairly unrealistic assumptions and have totally neglected other considerations such as the transparency of certain cloud types to some satellite sensors. Nevertheless, it is clear that any simulation procedure must take into account the spatial and spectral characteristics of the sensor being simulated. The denominator of the R value (the sensor resolution) is known. What is unknown is the numerator - the mean cloud size. This will vary from region to region and from season to season. Before our simulation results can safely be extended to other satellite systems, new cloud-type and cloud spacing data will be required. The effects of a given cloud cover on various sensors can then be reduced to a common denominator.

6. SEVERE WEATHER STATISTICS

Task 5 of the contract called for the development of severe weather statistics for as many cloud regions as possible. Severe weather is defined here as the occurrence of any of the following:

- Thunderstorms
- Winds ≥ 50 knots associated with thunderstorms
- Tornadoes
- Tropical cyclones

We recognize that other weather phenomena may be termed "severe;" e.g., blizzards, extreme heat or cold. It is those severe weather phenomena listed above, however, that impact most seriously on electromagnetic energy propagation through the atmosphere. This working definition was coordinated with NASA/MSFC.

To compute climatological severe weather statistics for the globe (or even for North America) from raw observational data would require the acquisition, processing, and analysis of an enormous amount of data. Such a task is well beyond the scope of this study. Thus, our approach was to acquire available climatological studies and analyze the results with the objective of eliminating any discrepancies between studies.

6.1 Thunderstorm Statistics

The most recent comprehensive global collection of thunderstorm data was that compiled by the World Meteorological Organization (1953 and 1956). The WMO requested such data from the various meteorological services over the world. Before discussing the thunderstorm statistics, it is necessary to recognize that there are some serious limitations to the data.

In the first place, a thunderstorm day is defined (by international agreement) as a local calendar day on which thunder is heard. This does not take into account the number of thunderstorms occurring on that day, or the intensity or duration of the storms. The necessity for thunder to be heard to count as a thunderstorm day limits the area covered by each observation station to a circular area about 20 km radius. It is possible that distant thunder may be ignored by an observer, particularly in the tropics during the rainy season when thunderstorms are common.

Other factors affecting the accuracy of the data are:

- There is wide variation in the number of years of record for both land and ocean regions.
- Many ocean regions have very sparse data coverage and data distribution is erratic in space and time.
- The thunderstorm day figures for ocean areas are based on ships' observations of thunder heard or lightning seen and are converted from station to area (5° latitude by 5° longitude) based on the period of record, the number of observations and theoretical probability formulas. (The reader is referred to the original WMO publication for a detailed discussion of the method for computing frequency over the oceans (WMO, 1956).
- Charts of the global thunderstorm frequency were necessarily smoothed because isolines did not always match in regions along the boundaries of neighboring countries.

Despite the formidable data limitations, the WMO analyses provide a gross picture of monthly, seasonal, and yearly thunderstorm frequency over the globe. It should be recognized that the nature of the phenomena does not lend itself to the relatively simple homogeneous region concept used for the cloud statistics. There are very sharp gradients of thunderstorm frequency over many tropical and subtropical regions. Nevertheless, Figure 6-1 shows a stepwise latitude-longitude depiction of various annual percent frequencies of thunderstorm occurrence to allow for the possibility of computer application. Figure 6-1 is based primarily upon the WMO data, although adjustments were made in the tropical South Atlantic, tropical Southeast Pacific and tropical South Indian Oceans where satellite data from our cloud statistics study indicate that these are moderate to strong convective regions. Our adjustment reflects a modest increase in frequency for these areas based upon the information implicit in describing the predominant cloud types in these regions as convective. As will be seen, the actual frequency may still be considerably higher than shown.

Superimposition of the cloud homogeneous regions upon regions of similar thunderstorm frequencies show that the regions are compatible in at least some areas of the globe; e.g., northern South America where high thunderstorm frequency agrees with the descriptions of cloud Regions 2 and 25 as regions with predominantly convective-type cloudiness. In other areas it is not compatible, as different percent

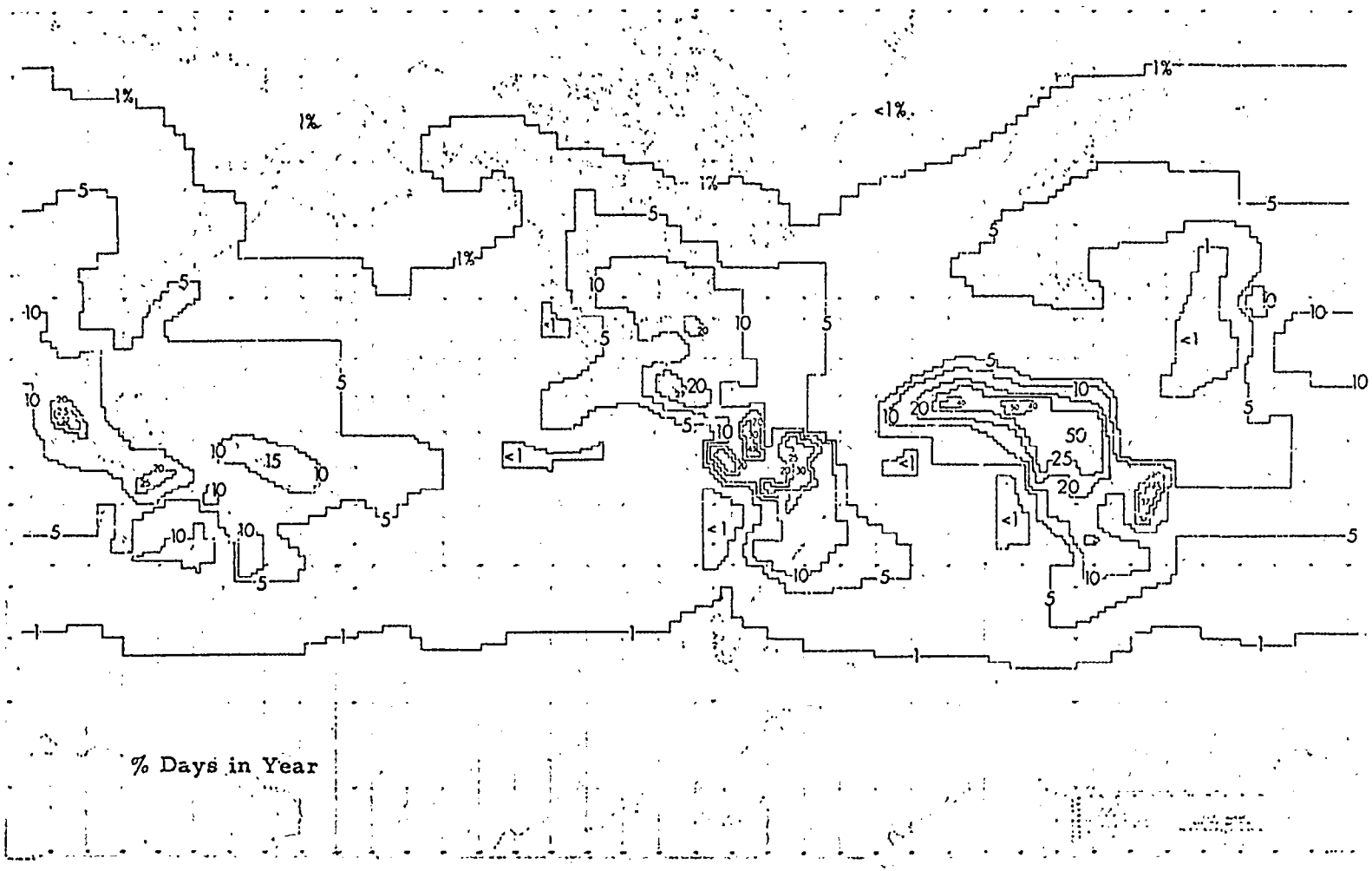


Figure 6-1 Thunderstorm Frequency Regions

frequency boundaries are perpendicular to and cross various homogeneous cloud regions; e.g., as in the western half of the North Atlantic Ocean.

Refinements were made in the frequencies over the Australian continent based upon data received from the Australian Bureau of Meteorology (1964). Refinements were considered unnecessary for Canada after examining a Canadian paper on the subject (Kendall and Petrie, 1962).

Figures 6-2 through 6-5 show the seasonal analyses of thunderstorm frequency (unadjusted). As expected, over most regions, the frequency is highest during the warmer months of the year; e.g., in the United States and Japan about one-half the annual number of thunderstorms occur in June, July and August, while in Australia 50% of the annual number are observed in December, January and February. In tropical latitudes, the thunderstorm frequency generally follows the movement of the intertropical convergence zone (ITCZ) with maximum occurrence in Summer and Fall (northern hemisphere) between 5° and 15°N latitude. This is particularly true over Africa, the Western Caribbean and Malaysia.

Beginning in 1967, the National Hurricane Center (NHC) has used satellite information in combination with conventional data to identify and track tropical disturbances. At NHC, a disturbance has been defined as as a migratory system that persists for at least 24 hours and covers an area at least 100 to 300 miles in diameter, with apparently intense convection (Simpson et al, 1969). Their analyses identified 61, 110 and 111 disturbances in 1967, 1968 and 1969, respectively (Simpson et al, 1968, 1969 and Frank, 1970). Results indicate that the majority of disturbances that enter the Caribbean Sea from the east during Summer and Fall originate over Africa. Carlson (1969) in a further analysis of 1968 African disturbances indicates the frequency to be one per 3.2 days. Because these are convective disturbances that form or pass through the regions of very high thunderstorm frequency over Africa (likely contributing to the observed high frequency) and because evidence from satellite data indicates these disturbances can frequently be tracked across the Atlantic, it suggests that the thunderstorm frequencies shown for Summer, Fall and annually in Figures 6-4, 6-5 and 6-1 are probably much too low for the tropical Atlantic between the African Coast and the Caribbean Sea.

To further complicate the problem of accurately determining thunderstorm frequency, "cloud clusters" that have been observed by satellite over tropical regions have been shown to be largely nonconservative; i.e., two-thirds of those observed in 1969 persisted for two days or <u>less</u>. Cloud clusters are defined as a cloud mass of at least 3° latitude diameter and a cloud cover greater than 50%. Undoubtedly, many cloud clusters contain thunderstorm cells and the observation

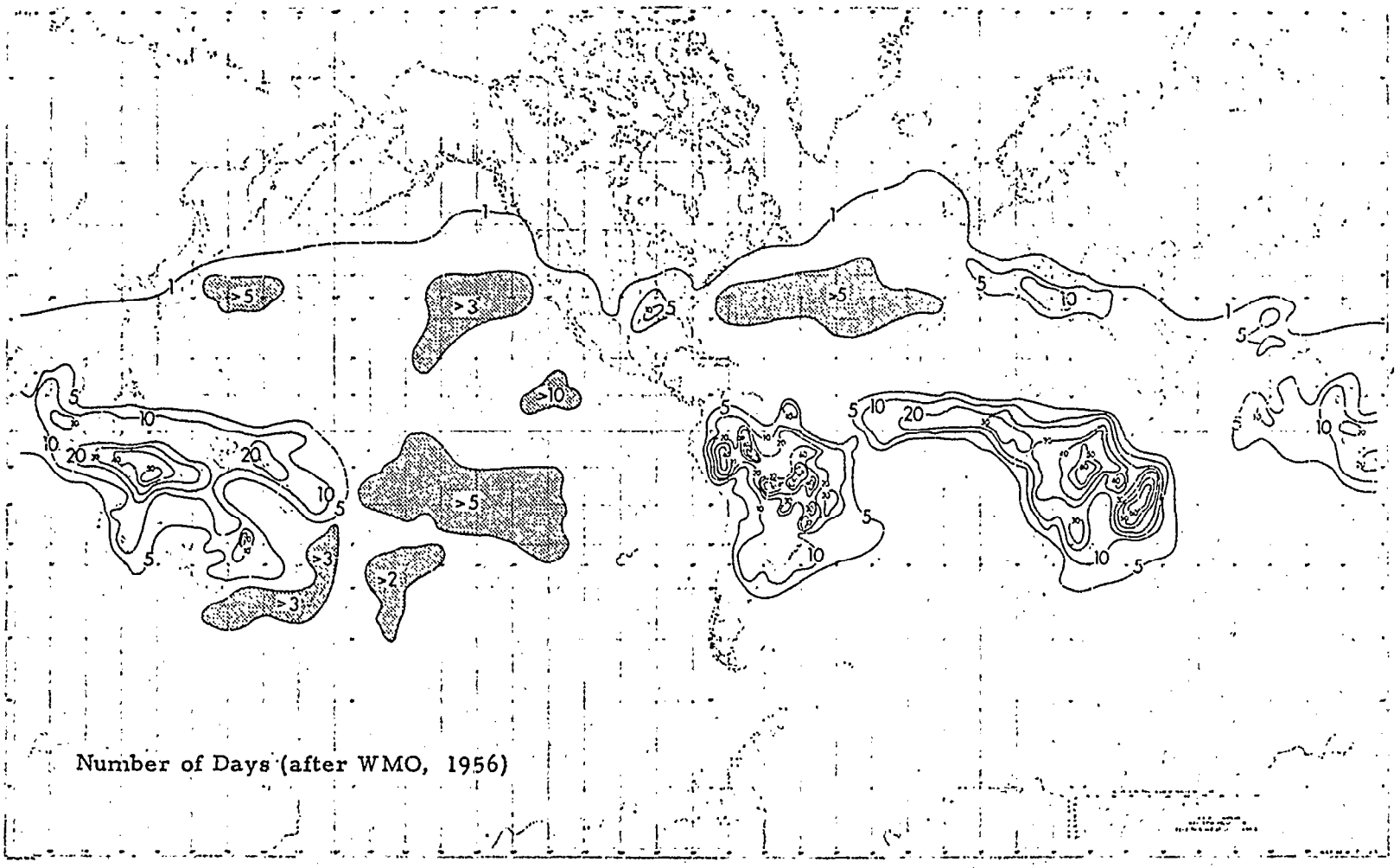
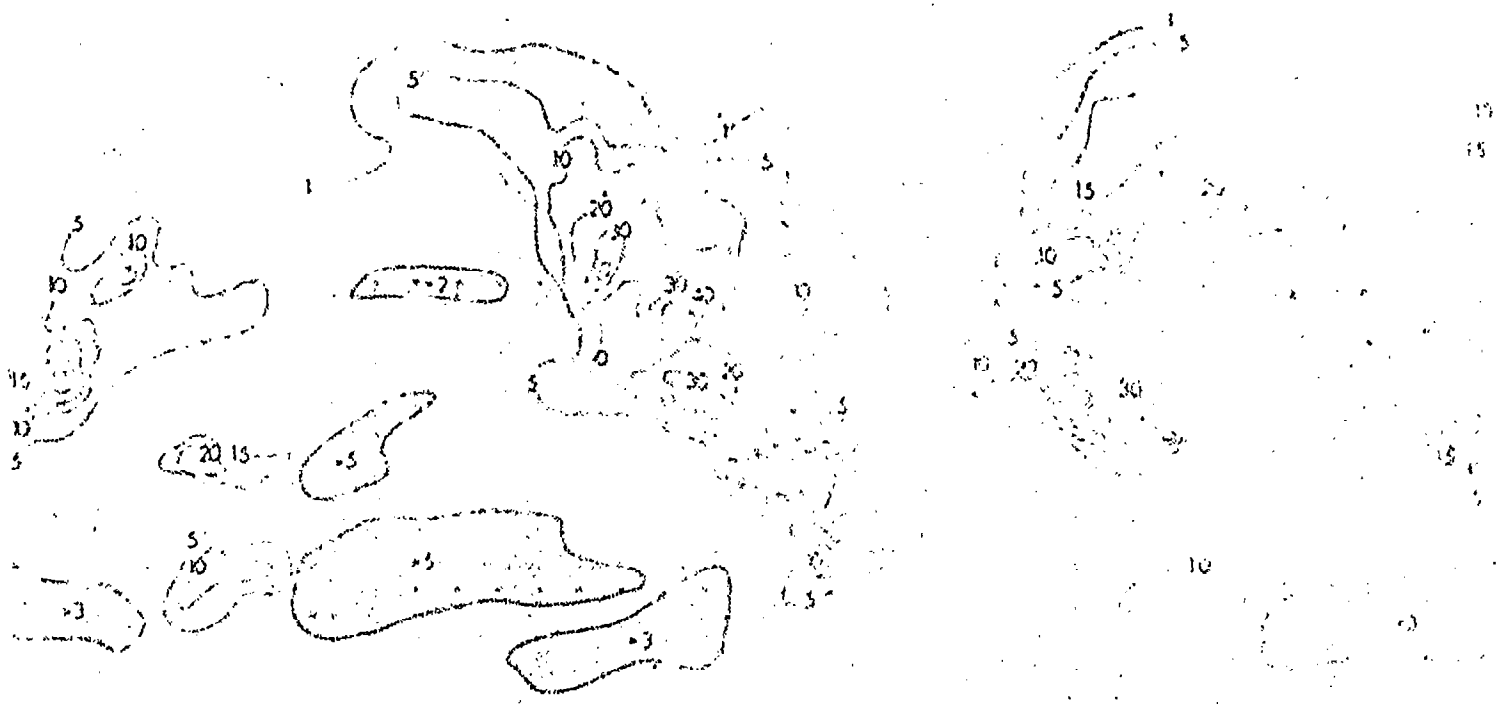


Figure 6-2 Seasonal Analysis of Thunderstorm Frquency; December, January and February

Number of Days (after WMO, 1956)

Figure 6-3 Seasonal Analysis of Thunderstorm Frequency; March, April and May



Number of Days (after WMO, 1956)

Figure 6-4 Seasonal Analysis of Thunderstorm Frequency; June, July and August

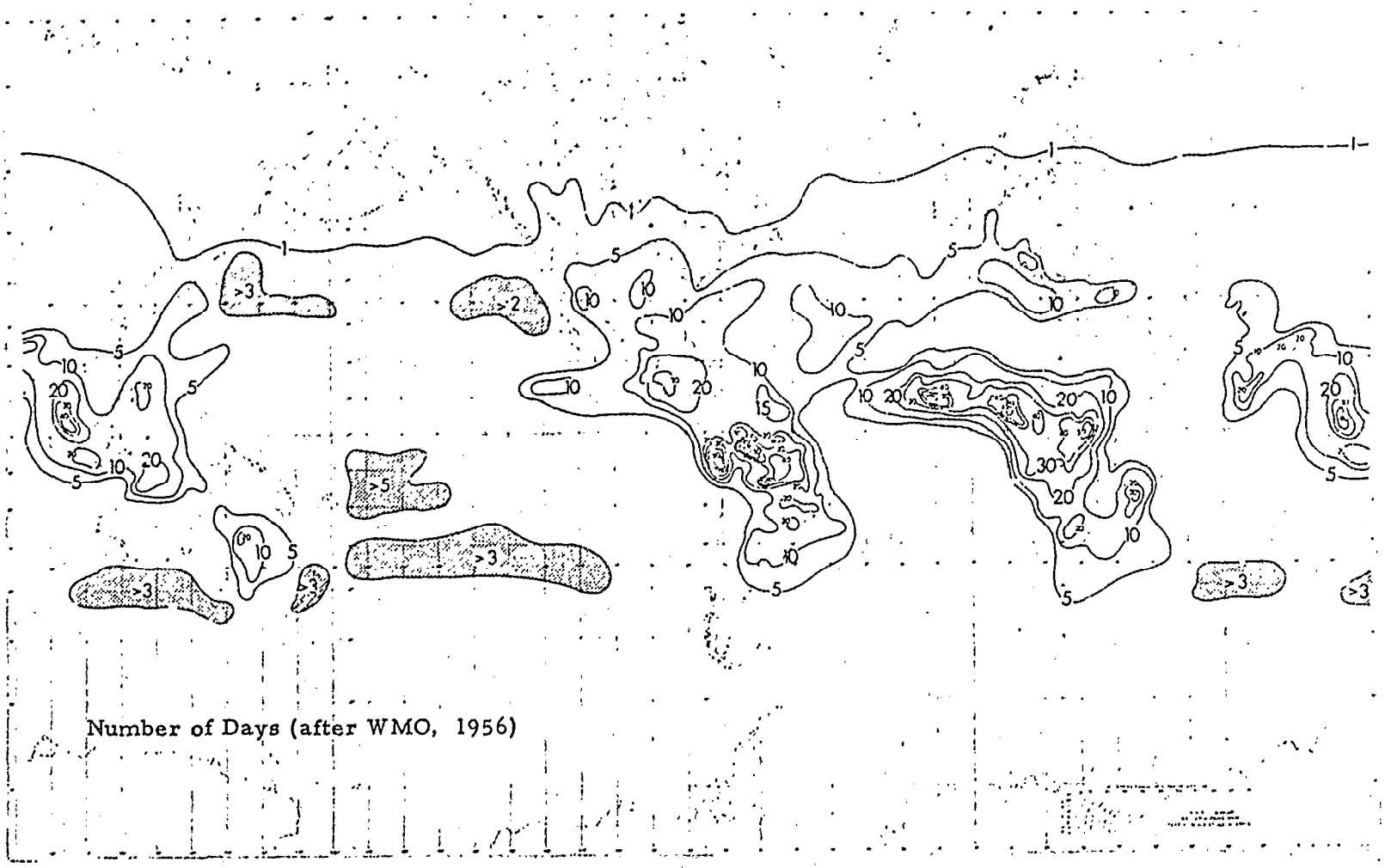


Figure 6-5 Seasonal Analysis of Thunderstorm Frequency; September, October and November

that 175 different cloud cluster systems occurred south of 30°N in the Atlantic during the July, August and September 1969 period enforces the notion that thunderstorm frequency is much higher than depicted in the WMO analyses in these regions.

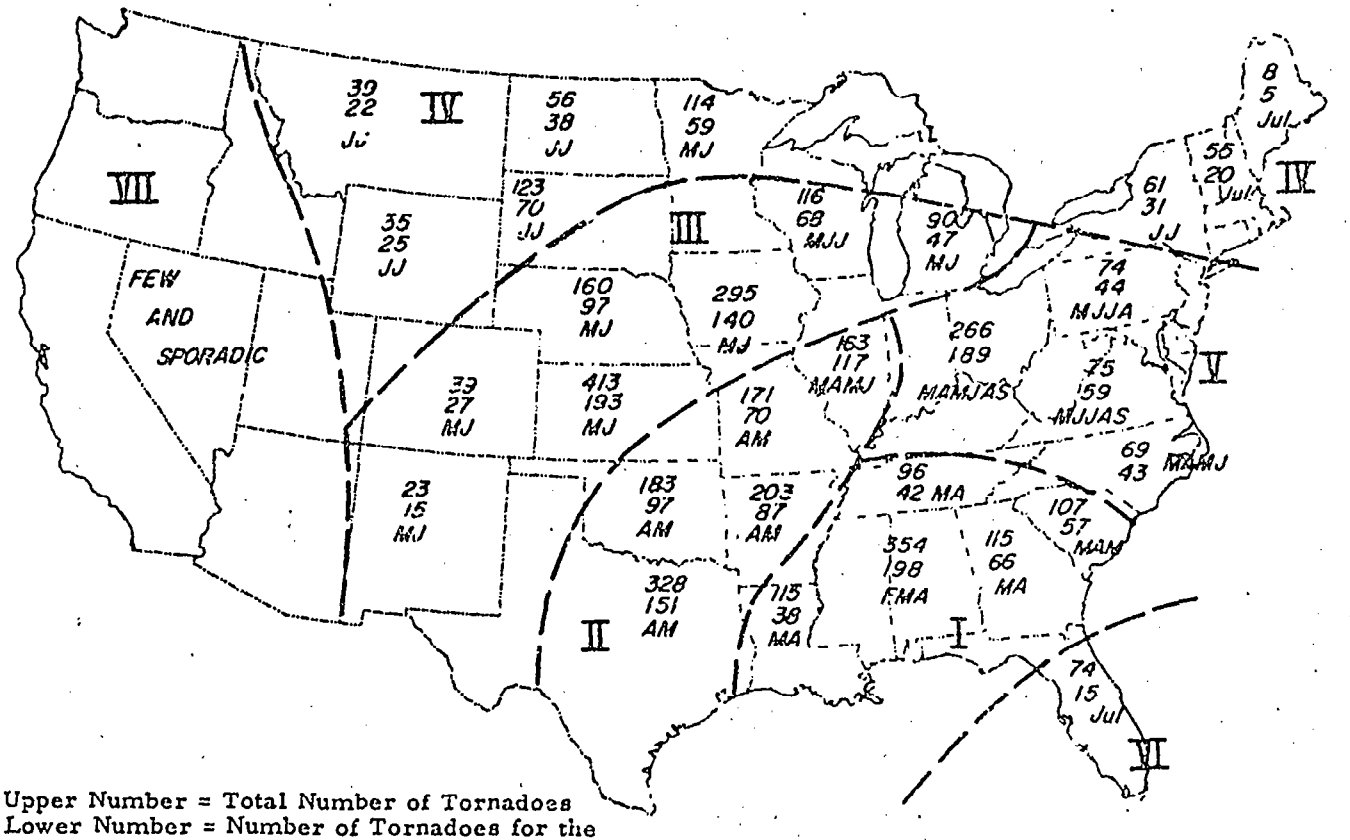
It is recommended that a separate study be performed to more accurately determine frequency of thunderstorms in the tropical ocean areas - particularly in view of the impact of thunderstorm occurrence on the electromagnetic sensor data.

6.2 Tornado and Severe Thunderstorm Statistics for the United States

Thunderstorm frequency data for the United States was supplied to the WMO by our National Weather Service (formerly Weather Bureau) but no update of this statistical information has been compiled in the past 15 years.

There has been, however, a comprehensive collection and analysis of severe local storm occurrences in the United States for the period 1955-67 by the National Severe Storms Forecast Center (Staff, SLS Unit, 1969). Their definition of severe storms is the occurrence of wind gusts in excess of 50 knots (related to convective phenomena) and/or the occurrence of hail ≥ 3/4" in diameter. This definition, of course, includes tornadoes but does not include ordinary thunderstorms that produce less severe effects. Nevertheless, the summary does provide useful information regarding the frequency of severe thunderstorms and tornadoes in the United States. It is not the intention here to reproduce their results in detail, but rather to discuss the frequency of severe local storms, in terms of the previously defined homogeneous cloud regions.

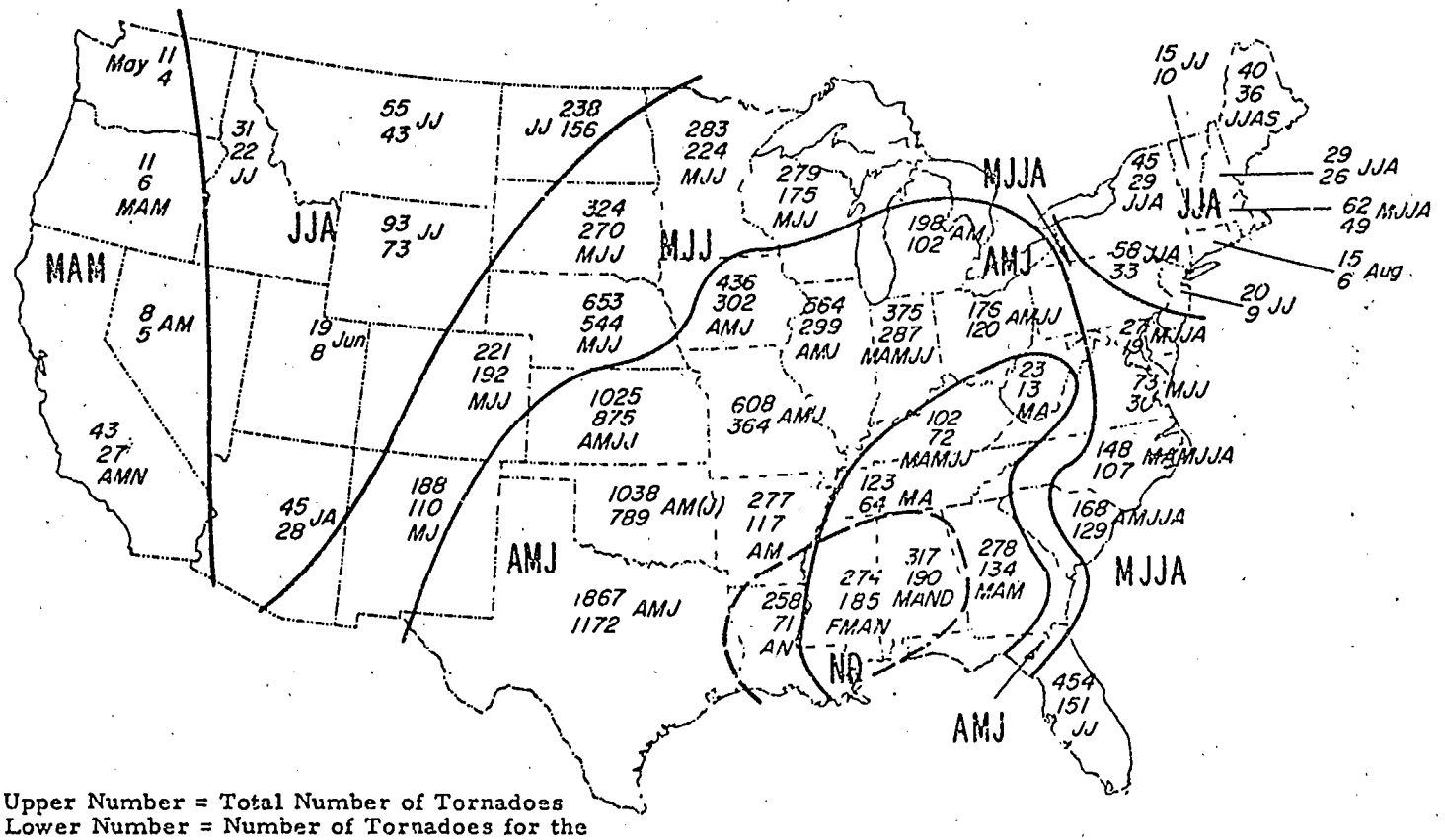
Because severe local storms are micro or meso scale phenomena and because their duration is generally of the order of minutes to an hour or so at most, many occurrences may go unobserved or, if observed, sometimes go unreported. Thus, there is a tendency for higher frequencies of occurence near populated regions, although this tendency has diminished with time. (Suburban and rural areas are now more populated and there is increasing public interest in reporting severe local storm occurrences.) Comparison of Figures 6-6 and 6-7 which show tornado distributions (as reported) for the period 1880 to 1942 (Showalter and Fulks, 1943) and 1955 to 1967 is a striking illustration - not of tremendously increased tornado activity - but of increased awareness and responsibility in reporting tornadoes. The reliability of the statistics has increased in the past two decades.



Upper Number = Total Number of Tornadoes
Lower Number = Number of Tornadoes for the Months Indicated
Letters = Consecutive Months of Maximum Activity (e. g. FMA is February, March and April)

(after Showalter and Fulks, 1943)

Figure 6-6 Tornado Distributions; 1880 to 1942



Upper Number = Total Number of Tornadoes
Lower Number = Number of Tornadoes for the Months Indicated
Letters = Consecutive Months of Maximum Activity
(e. g. FMA is February, March and April)

Figure 6-7 . Tornado Distributions; 1955 to 1967

It is readily apparent that the highest frequency of tornadoes occurs in the area from Texas northward through the Central Plains and eastward through the southern Midwest. The Central Gulf States also exhibit relatively high frequencies with fewest occurrences over the West Coast states, Rockies, and the extreme Northeast.

On the basis of the number of tornadoes and severe thunderstorms observed and the time of the year with maximum frequency, homogeneous severe local storm regions were defined for the United States. These regions are shown in Figure 6-8 upon which is superimposed the homogeneous cloud regions. The average number of tornadoes for each of the six severe storm regions was computed from data for individual states for four seasons and is shown in Figures 6-9 through 6-12. Tornado Region 3 lies almost entirely within cloud Region 19. In Region 3 maximum tornado frequency is in the Spring with almost uniform distribution through the remainder of the year. The western quarter of cloud Region 19, however, is in Region 1, an area of very high tornado frequency. Within cloud Region 11 are portions (or all) of four different tornado regions. Thus, the large variability of tornado frequency within some cloud regions leads to the recommendation that the cloud regions as they are presently defined not be used to identify homogeneous regions of severe weather.

Another tornado statistic of interest is shown in Figure 6-13. It indicates the average number of tornado days during the month of the year that has the maximum number of tornado days. Texas and Kansas exhibit the highest average number of tornado days.

Severe thunderstorm patterns show similar (to tornado patterns) relative frequencies between regions with the maximum number of occurrences in severe storms Region 1. Percentage of days with thunderstorms with wind gusts over 60 knots range from a high of 9 to 12% in Kansas, Oklahoma and Texas to 4 to 8% in the Midwest to less than 1% in part of the Northeast and Far West. The reader is referred to the original report for complete details (Staff, SLS Unit, 1969).

We would like to reemphasize at this point, that the confidence factor regarding the absolute values for frequency of severe weather phenomena as defined in this report is not high. The more useful part of the information concerns the relative frequency of occurrences between different regions of the globe. Here one can be reasonably confident that the relative differences, in most cases, will probably be maintained with a larger data sample. While this

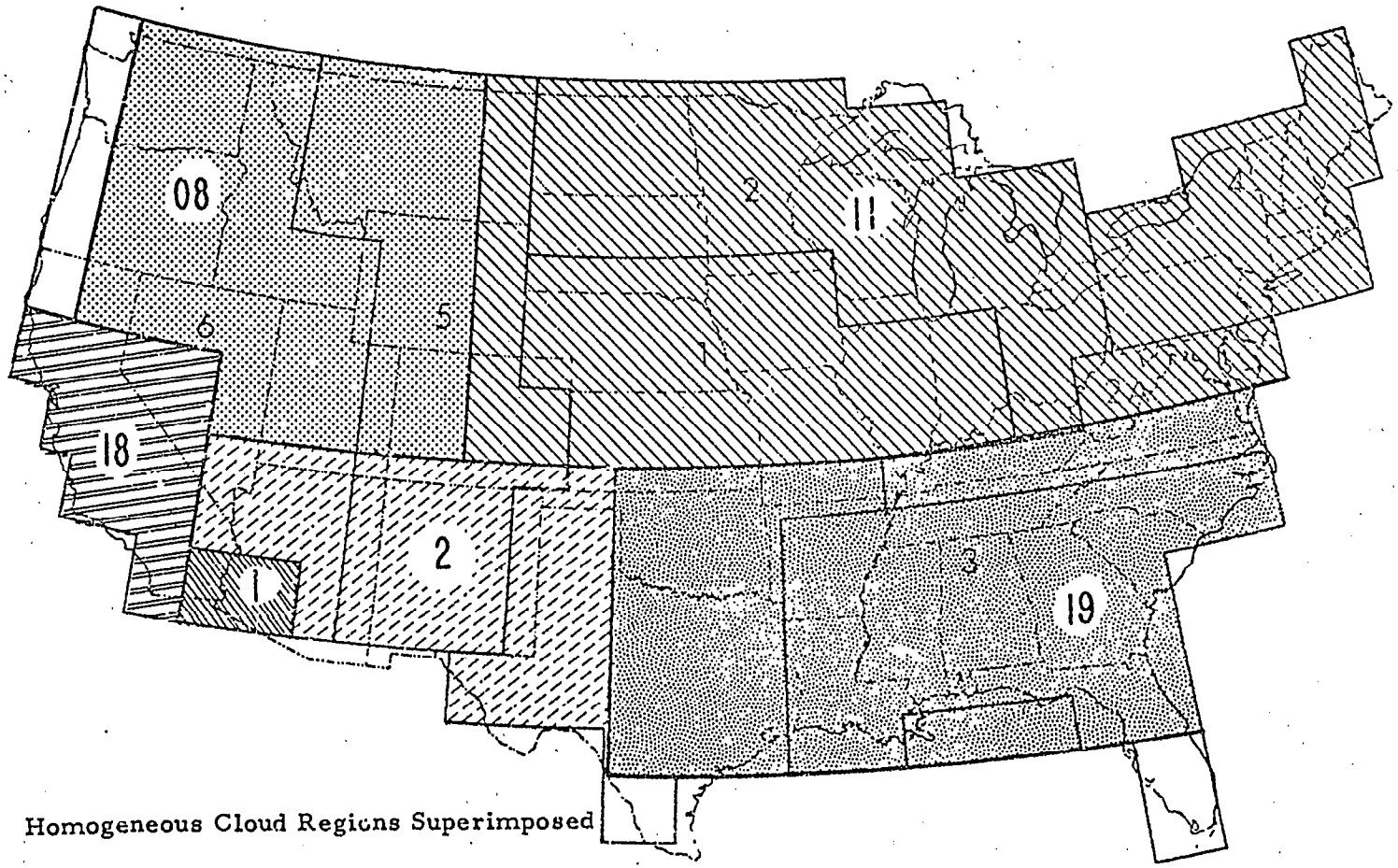


Figure 6-8 Severe Local Storm Regions; Annual

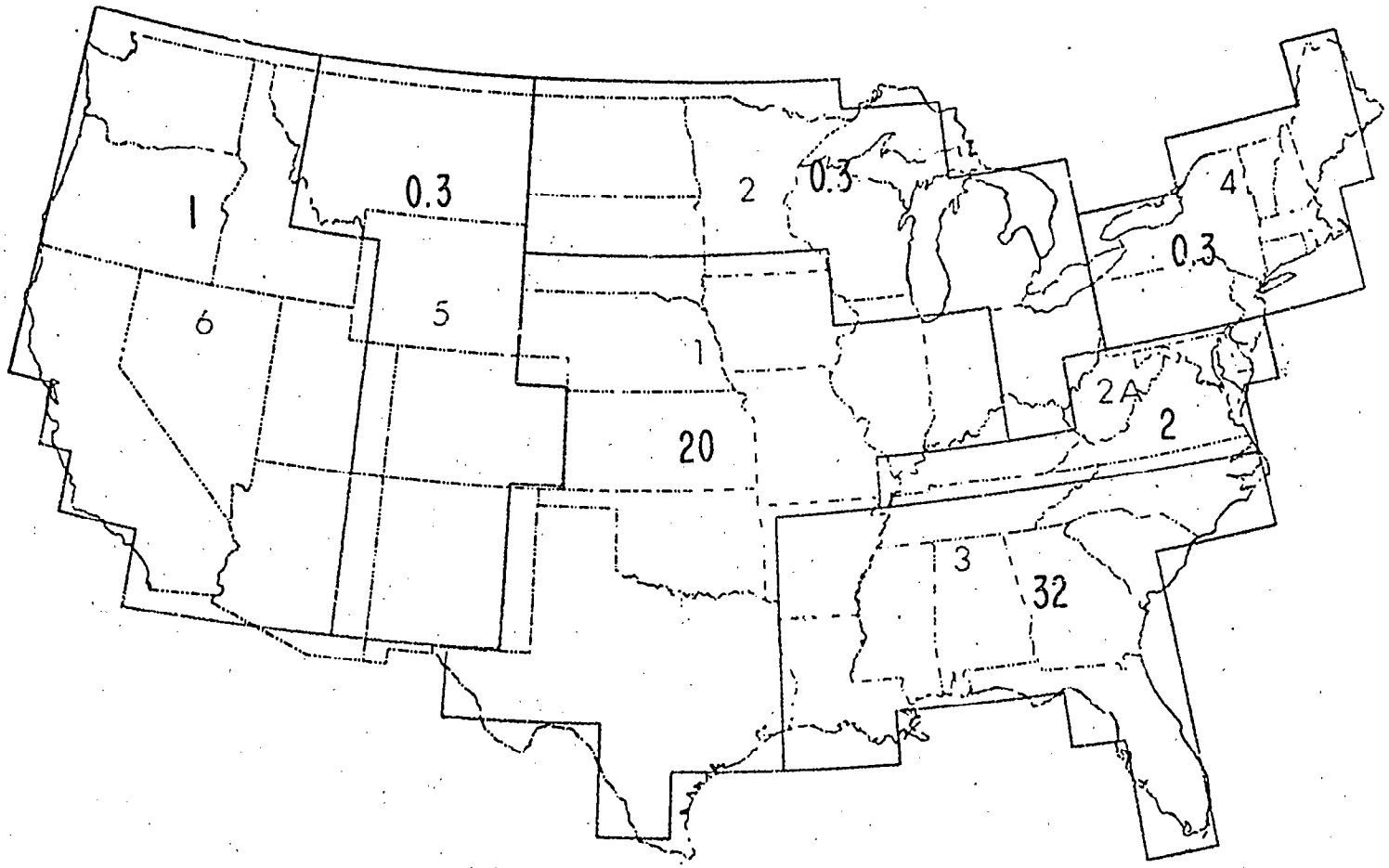


Figure 6-9 Tornado Frequencies; Winter

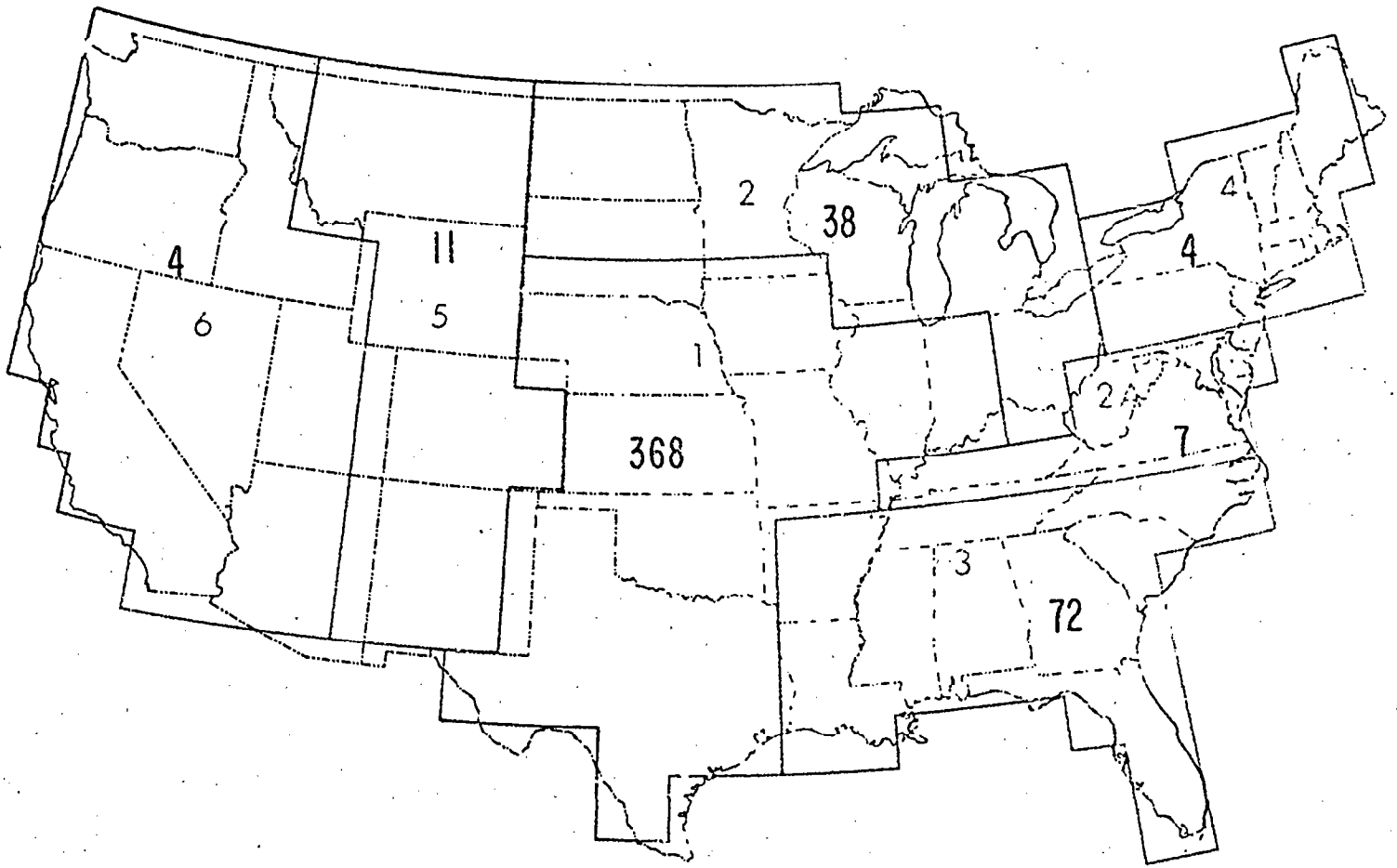


Figure 6-10 Tornado Frequencies; Spring

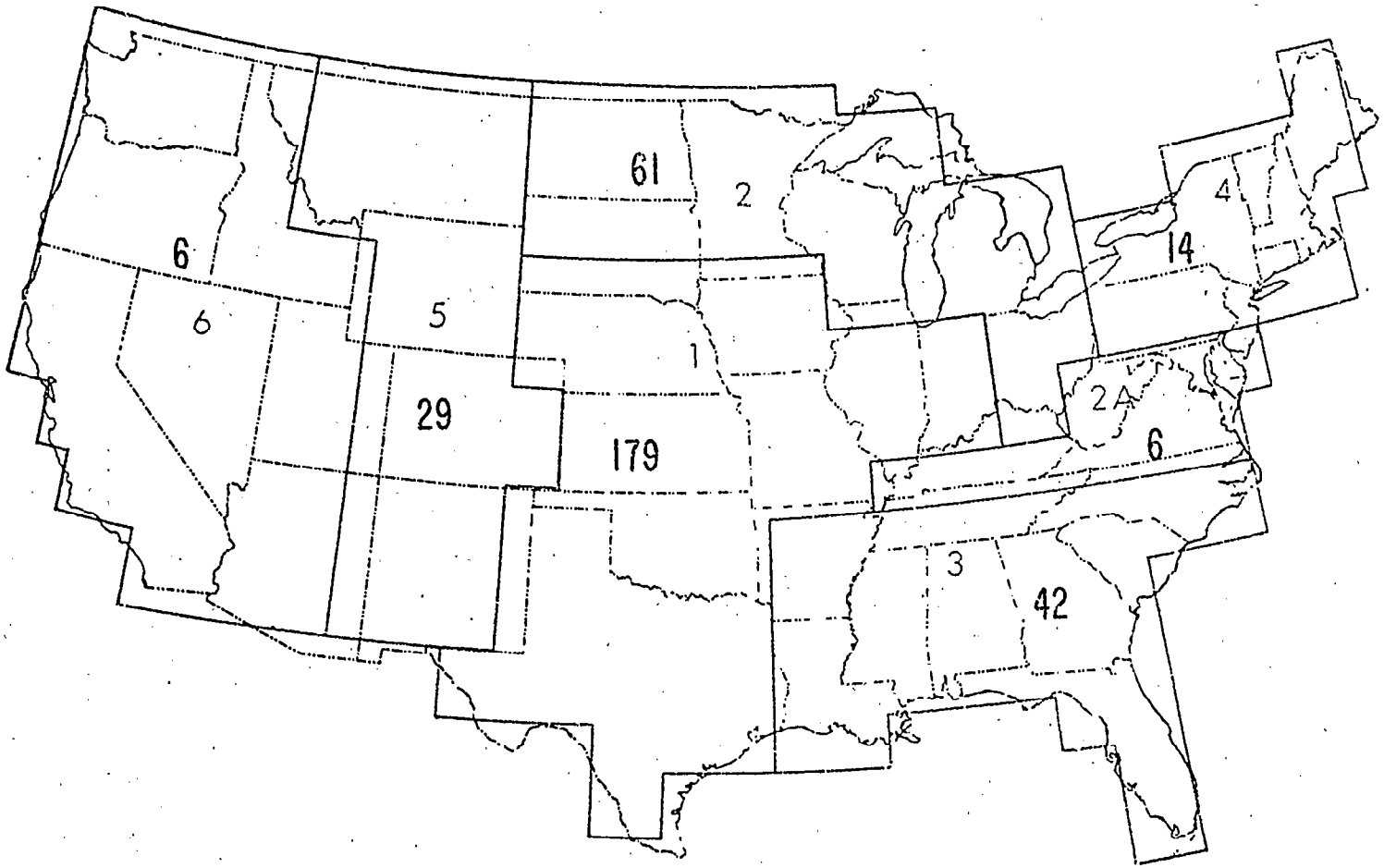


Figure 6-11 Tornado Frequencies; Summer

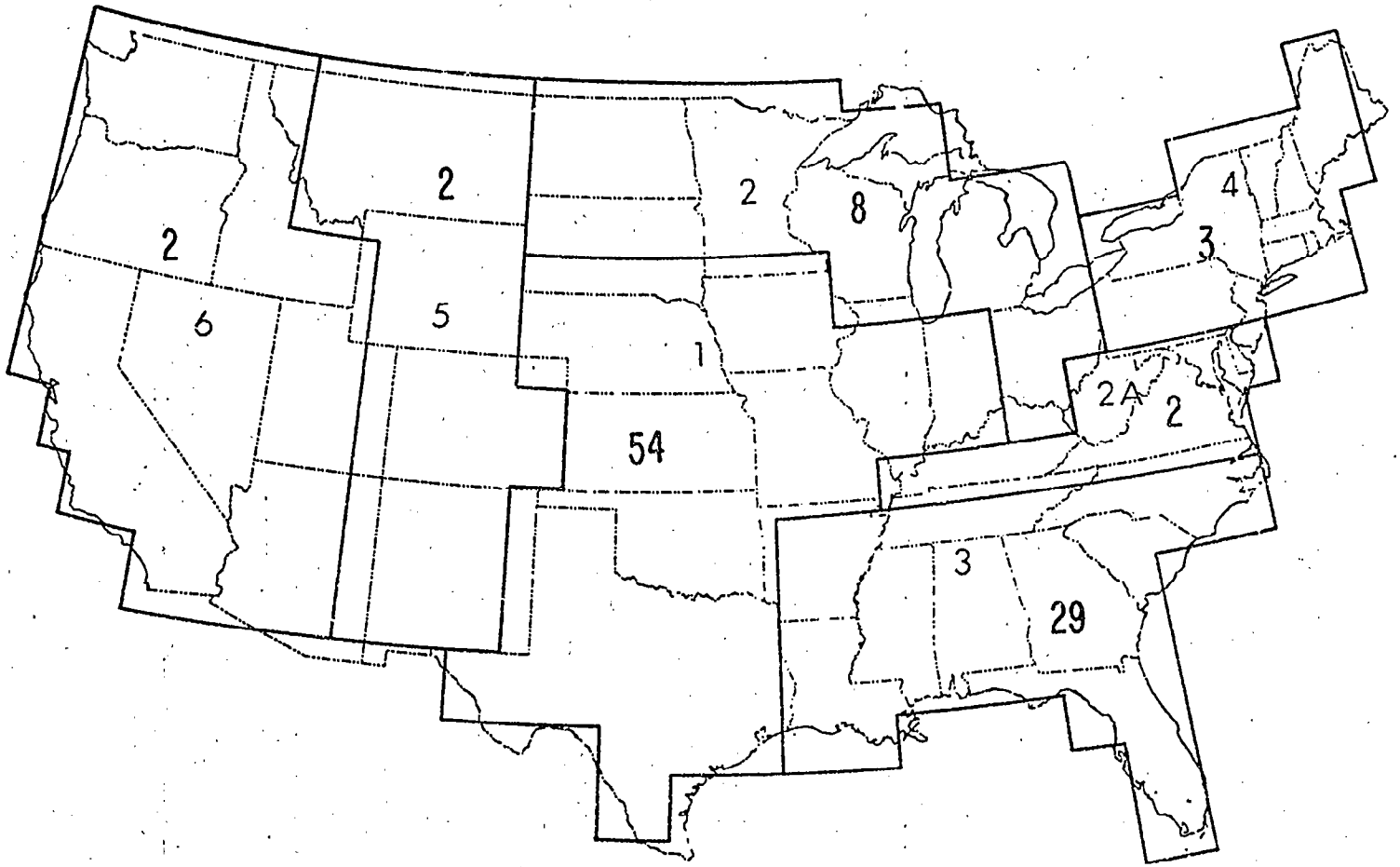


Figure 6-12 Tornado Frequencies; Fall

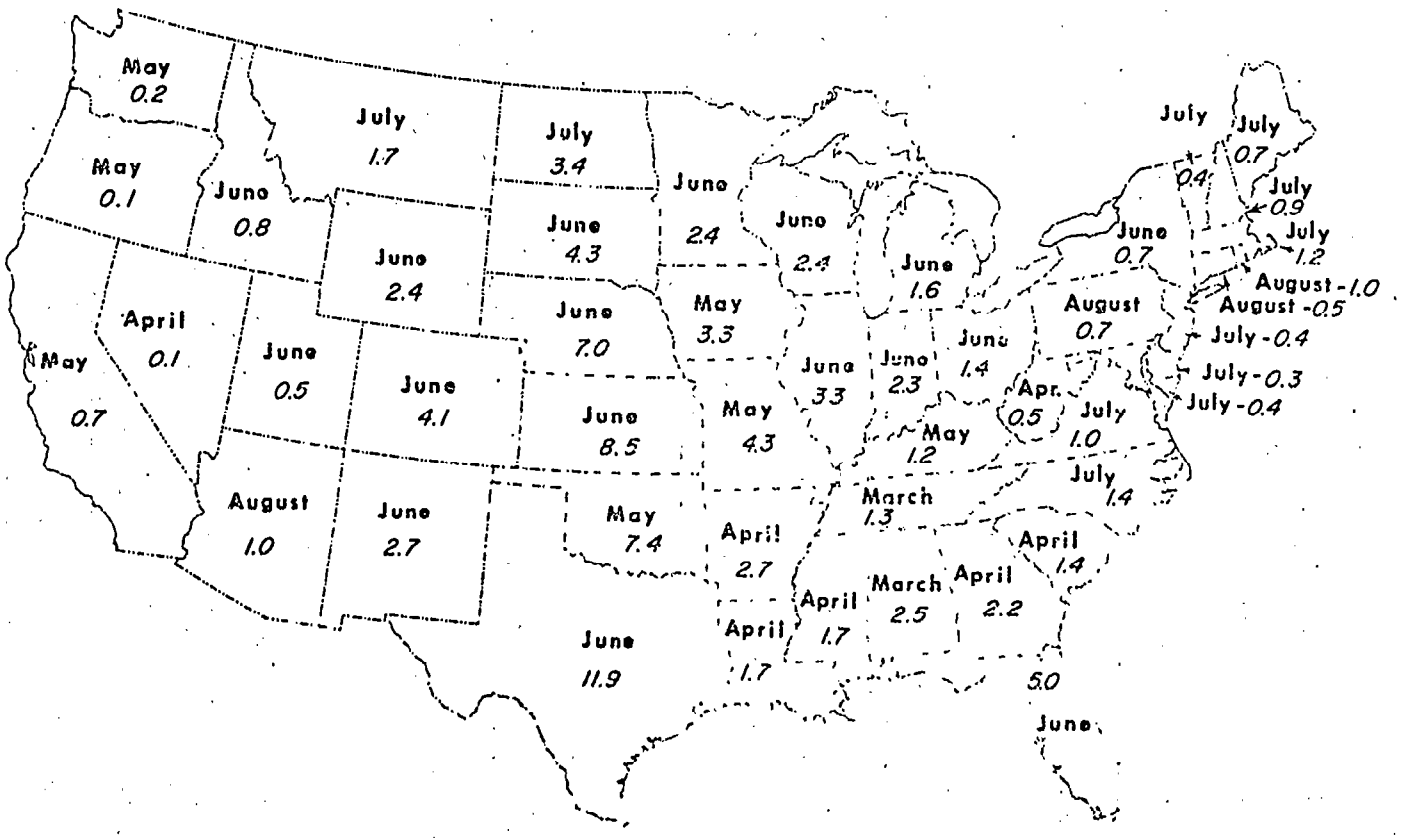


Figure 6-13 Maximum Monthly Average of Tornado Days

...l information, we believe that new studies based on a combination of satellite data and conventional data are needed to accurately determine the probability of thunderstorm occurrence over any area of the globe for any time of day for any month of the year. Such studies together with the global cloud statistics and four-dimensional atmospheric model statistics will provide almost all the required information on expected atmospheric effects on electromagnetic sensors.

6.3 Tropical Cyclone Statistics

Since the beginning of the twentieth century, the reliability of tropical cyclone records has improved. Records of tropical cyclone occurrences have become more reliable to the point where in the past few years chances are small that a tropical cyclone went undetected.

The improved detection has been due to:

1) Increase of routine ship traffic in tropical North Atlantic and North Pacific waters.

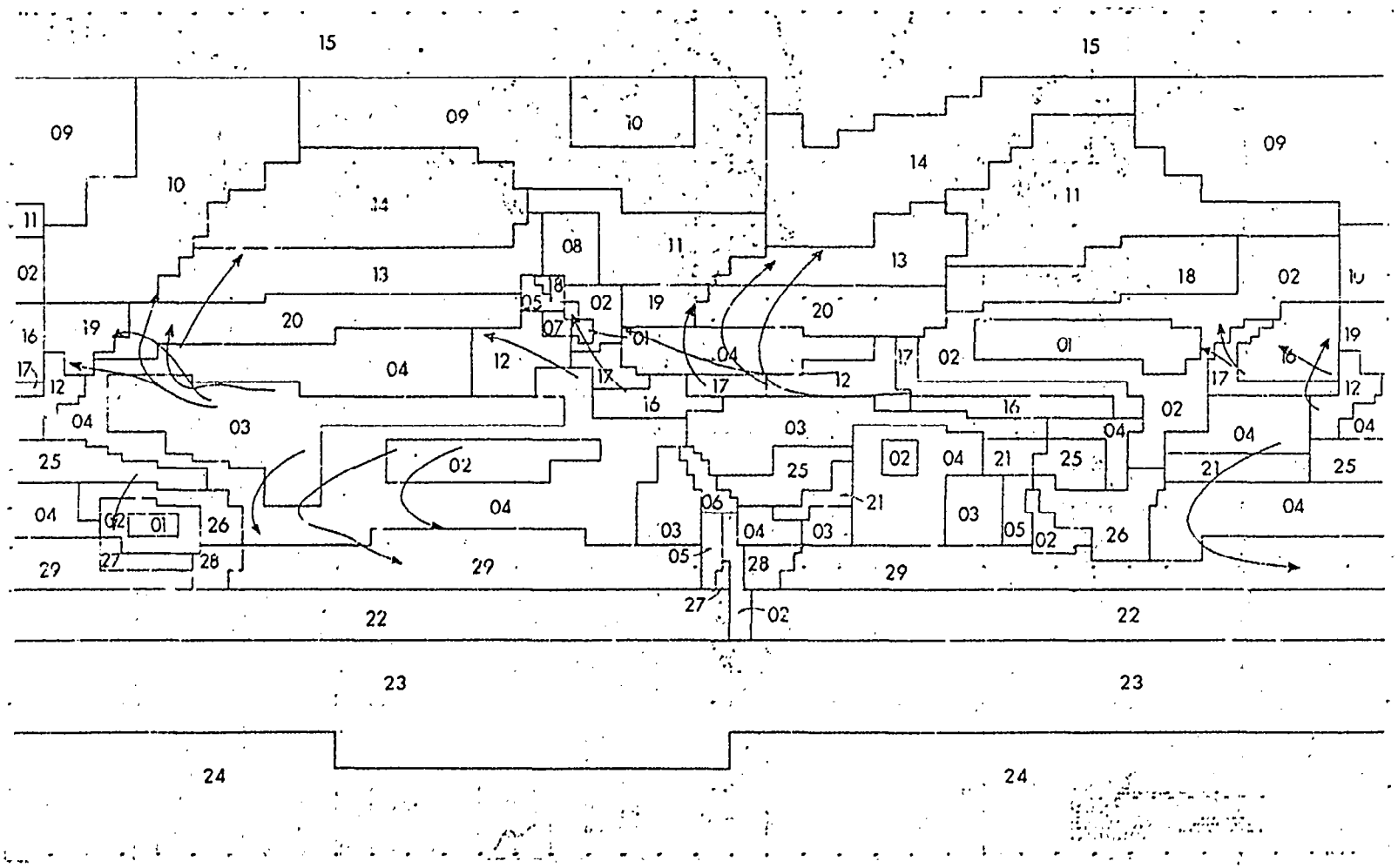
2) Systematic aircraft reconnaissance of tropical storms after World War II in the Atlantic and North Pacific.

3) The advent of the meteorological satellite in 1960 with daily global coverage for the past four years.

It is likely that climatological statistics of tropical cyclone frequency will be revised in the light of the better routine global observational capability provided by satellites. An example of an area of the world where previous estimates of the annual average number of tropical storms is in error by a large amount is the eastern North Pacific Ocean where shipping, aircraft reconnaissance and satellite coverage were quite limited until recently. More specific details on the previous error in annual tropical cyclone frequency in this region will be given later in this section.

Tropical cyclones occur in the regions of the world shown by the mean tracks for these storms in Figure 6-14. These mean tracks are superimposed upon the homogeneous cloud climatological regions. Tracks of individual storms vary considerably from the mean.

The average frequencies of tropical cyclones by months for various regions are listed in Table 6-1. The figures in Table 6-1 were compiled from various sources and these are indicated for each region. The Northwest Pacific Ocean has the largest average annual frequency (21.1) by far. They have been known to



(after Dunn and Miller, 1960)

Figure 6-14 Mean Tropical Cyclone Tracks

TABLE 6-1

AVERAGE FREQUENCY OF TROPICAL CYCLONES (BY MONTHS)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Total
North Atlantic Ocean 1901 - 1963 Computed from Cry (1965)	*	*	*	*	*	0.5	0.6	1.8	2.7	1.7	0.3	0.2(A)	7.8
Eastern North Pacific Ocean East of 140°E	1910 - 1940				0.1	0.8	0.7	1.0	1.9	1.0	0.1		5.7
	1960 - 1969(B)				0.2	1.6	1.6	2.1	3.9	2.0	0.1		10.8
North Pacific Ocean Long. 170° Westward 1901 - 1940 Dunn and Miller (1960)	0.4	0.2	0.3	0.4	0.7	1.0	3.2	4.2	4.6	3.?	1.7	1.2	21.1
North Indian Ocean - Bay of Bengal Dunn and Miller (1960)	0.1	0.0	0.2	0.2	0.5	0.6	0.8	0.6	0.7	0.9	1.0	0.4	6.0
North Indian Ocean - Arabian Sea Dunn and Miller (1960)	0.1	0.0	0.0	0.1	0.2	0.3	0.1	0.0	0.1	0.2	0.3	0.1	1.5
South Indian Ocean - Madagascar Eastward to 90°E Dunn and Miller (1960)	1.3	1.7	1.2	0.6	0.2							0.1	5.1
South Indian Ocean - Northwest Australia (38 years of data) Brunt and Hogan (1956)	0.5	0.6	0.6	0.1								0.3	2.1(C)
South Pacific Ocean - Northeast Australia (50 years of data) Brunt and Hogan (1956)	0.9	0.8	0.9	0.3								0.2	3.1(D)
South Pacific Ocean - East of Fiji Islands 1940-1956 Gabites (1956)	1.1	1.2	0.1	0.3	0.2						0.2	0.5	4.6

(A) Total December-May inclusive.
(B) Estimated for months based on ratio from total number.
(C) Annual summaries for period 1962-1963 through 1966-1967 indicate average of 9 per year.
(D) Annual summaries for period 1962-1963 through 1966-1967 indicate average of 6 per year.

Director of Meteorology 1965, 1966, 1968a, 1968b and 1969.

occur in all months of the year, but maximum frequency is in late Summer and Fall - similar to other northern hemisphere regions where tropical cyclones occur.

In the North Atlantic area, which includes the Caribbean Sea and the Gulf of Mexico, the long-term average is near eight per year, but the annual variation is large, ranging from one in some years to a maximum of 21 in 1933. Five-year running averages computed by Dunn and Miller (1960) show that the averages were near 6 per year from 1901 to 1930 and around 10 per year from 1931 to 1960. For the past decade, the average was 9 per year. The climatological probability of tropical cyclones in the North Atlantic by month is shown in Figure 6-15. This, of course, is the probability of occurrence within any part of this large region. The probability for a tropical cyclone at any particular location on any particular day of the month is considerably lower than the probabilities shown in Figure 6-15. Thus, it is easy to see that tropical cyclones are a relatively rare event for a particular area; e.g., relative to the number of extratropical cyclones. However, as was mentioned in Section 6.1, tropical disturbances (from which hurricanes form) are common in ocean areas during the tropical cyclone season (Simpson et al., 1969 and Frank, 1970). These relatively weak hurricane seedlings, while not possessing the required surface wind speeds and circulation characteristics of tropical storms or hurricanes, often cover about the same area (as given by the cloud mass) as does the stronger disturbance. From the point of view of electromagnetic degradation, the cloud pattern and type (mostly convective with cirrus blowoffs) may have nearly the same effect for both the tropical disturbance and the tropical storm or hurricane. The major difference is that in the hurricane there are spiral cloud bands of very intense convection that feed into the hurricane circulation. (Within these "feeder bands," attenuation of sensor data will be most pronounced.) Convective bands, somewhat less intense, are often noted within the ordinary tropical disturbance. It is apparent to us that new studies based upon the last 5 years of daily global satellite data are required to assemble statistics on tropical disturbances and cloud clusters in tropical and equatorial regions.

Returning to the statistics shown in Table 6-1 for other regions of the globe, it has been shown by Sadler (1964), Denney (1969) and Baum (1970) that tropical cyclone frequency in the the eastern North Pacific is considerably higher than originally reported by Dunn and Miller (1960). Sadler believes that the number of tropical storms per year may be near 30. In 1970, 20 tropical storms were observed in this region. A much longer period of record with the daily

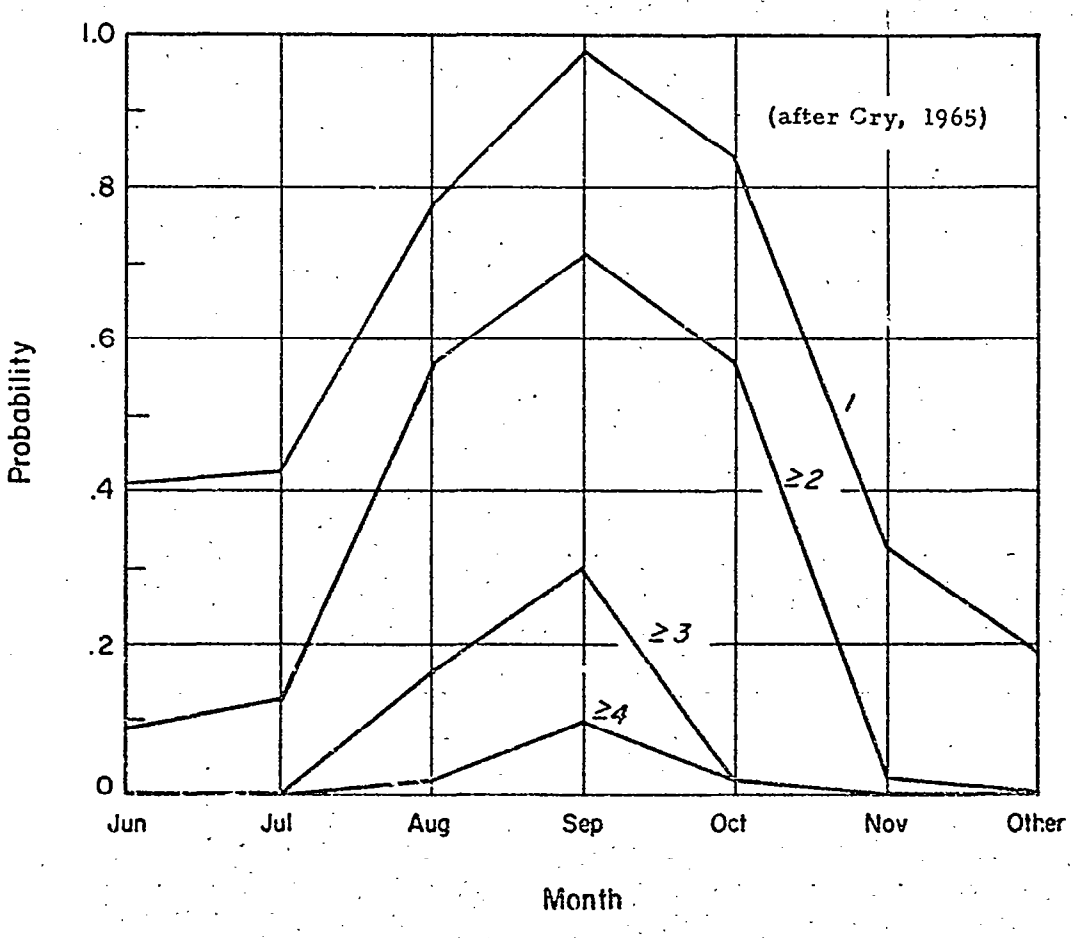


Figure 6-15 Climatological Probability of Tropical Cyclones in the North Atlantic

satellite coverage is required before a reliable climatology is determined not only for the eastern North Pacific, but for the South Pacific and South Indian Oceans as well. We believe that in these regions, where shipping is less than in the North Atlantic and North Pacific, numerous storms may not have been detected in the past. For example, in the Northwest Australian region where the annual average frequency of tropical cyclones was listed as 2.1 based on 38 years of record (Brunt and Hogan, 1956) the summaries for the 5 recent seasons (Director of Meteorology, 1965, 1966, 1968a, 1968b and 1969) show an average of 9.0.

Aside from the number of storms, the life of the storm and its areal coverage are of interest. Both the life and the size vary widely between storms and are a function of the area of formation and general circulation features influencing its history. Average life for the Atlantic storms is 9 days, but may be as little as one or two days and some have been tracked for three to four weeks. August storms have an average 12-day life span; July and November storms about 8 days.

The size of the storm varies from about 50 to 100 miles diameter to as much as 1000 miles in diameter for the large Pacific typhoons. Average size is about 300 to 500 miles diameter for mature storms.

Intensity of the wind circulation is also subject to wide individual variations. Tropical storms must have maximum winds $\geq$ 35 knots, but the more intense typhoons and hurricanes have wind speeds up to about 175 knots.

To summarize briefly, the occurrence of tropical cyclones is a relatively rare event, although not as rare as believed before the satellite provided daily global coverage leaving little chance for an undetected storm. Tropical disturbances from which tropical cyclones form have been shown to be a frequent occurrence. These disturbances are characterized by intense convection and have dimensions similar to tropical storms. Statistics for these tropical disturbances over all tropical regions of the globe are lacking. At the present time, enough daily satellite data exist to perform a statistical analysis based on 5 years of record. It is recommended that such a study be performed.

7. DISTRIBUTIONS OF WATER VAPOR CONTENT

Knowledge of the amount of water vapor present in the atmosphere is of major importance in estimating the attenuation of electromagnetic radiation between the earth's surface and a satellite or airborne sensor. In order to provide at least a preliminary basis for estimating the expected attenuation, we have acquired data from various sources to determine global distributions of water vapor content (as given by vertically integrated precipitable water).

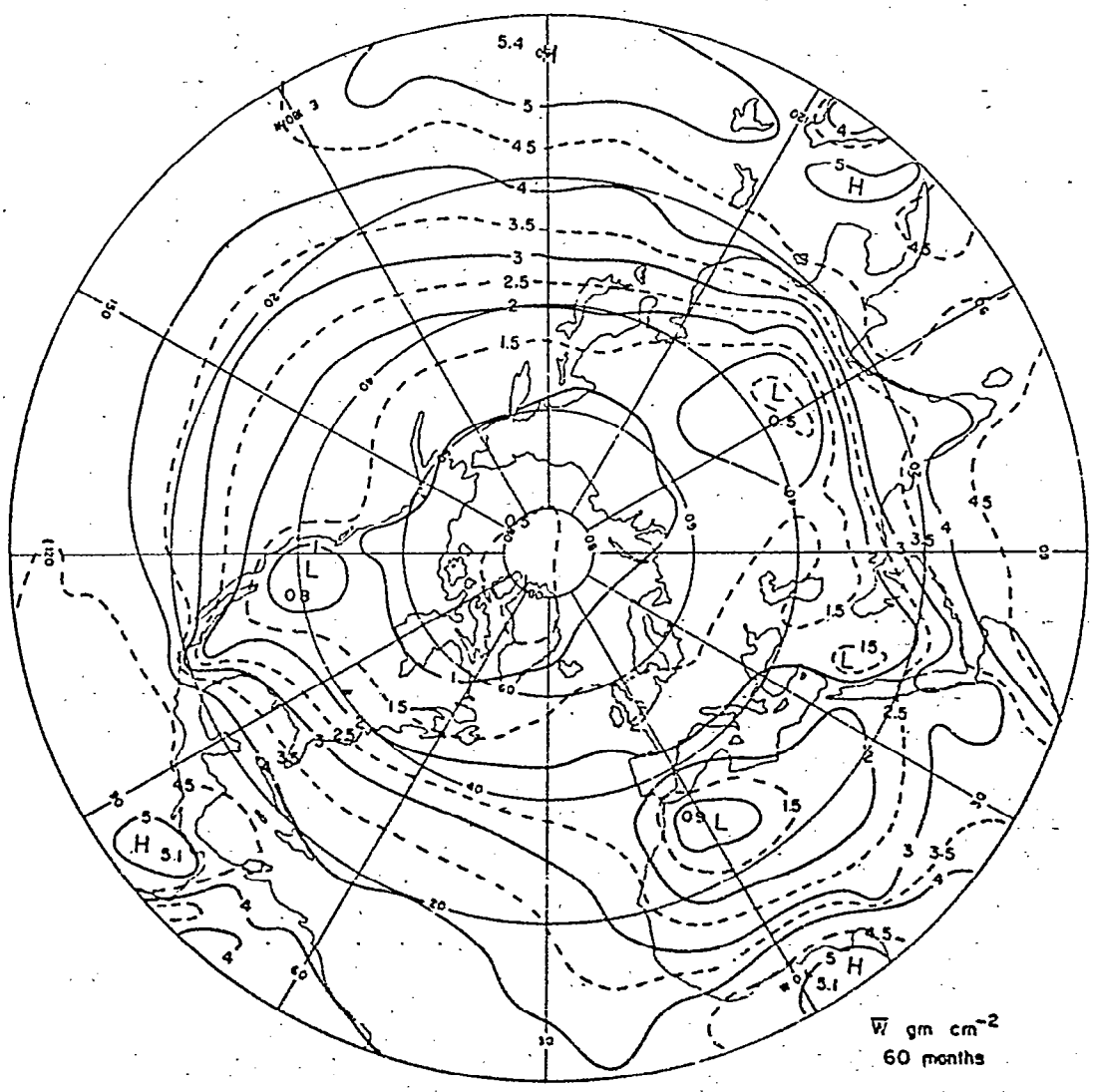
7.1 Global Distributions - Annual

In the literature, a recent study reported by Piexoto (1970) discusses water vapor distributions for the northern hemisphere based on a five-year data sample. The mean annual precipitable water (for this sample) shown in Figure 7-1 is from his paper. It shows, in general, that the water vapor gradually increases from pole to equator as one might expect from the dependency of water vapor content on temperature. Exceptions to the general zonal characteristic of the mean water vapor distributions are dry conditions that are found over high terrain (e.g., Rocky Mountains, Himalayan Mountains, etc.) and over desert regions as in the southwest United States, the Sahara, Arabia, Iran, etc. Regions of highest precipitable water are over tropical and equatorial ocean areas and over equatorial Africa.

Figure 7-2, also from Piexoto's paper, is a vertical cross section of the mean specific humidity across latitudes from south to north. It shows highest values within 10° of the equator near the surface and decreases sharply with increasing latitude and almost exponentially with elevation. Thus, approximately 90% of the total water vapor in the atmosphere is below the 500-mb level.

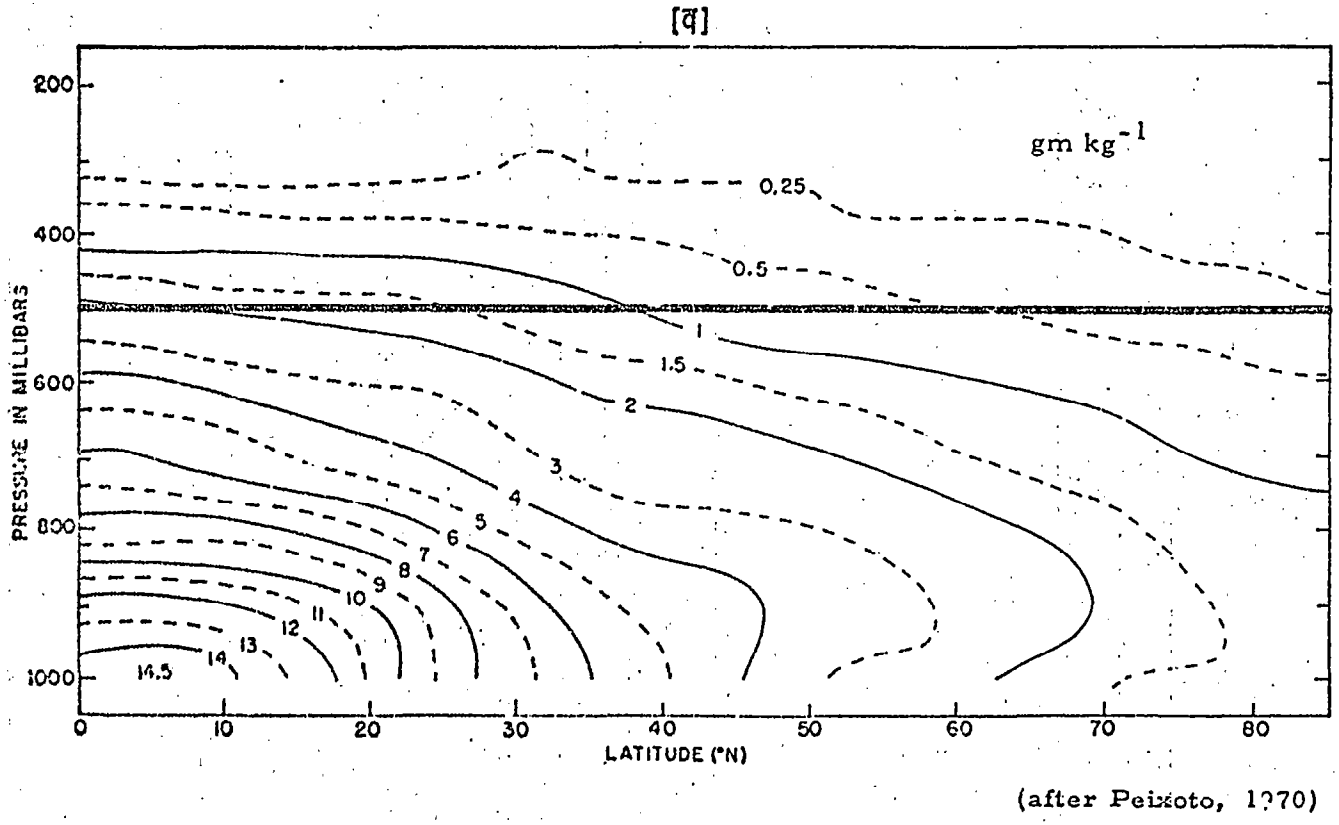
Detailed moisture data for the southern hemisphere was studied only for the IGY period (Calendar Year 1958). Figure 7-3 (after Starr et al, 1969) shows the global time-averaged distribution of precipitable water for 1958. Comparison of the northern hemisphere portion of Figure 7-3 with Figure 7-1 shows marked similarities. This suggests that the mean annual precipitable water distributions are relatively conservative. Thus, we believe the distribution for the southern hemisphere shown in Figure 7-3 for 1958 may be a close approximation to the mean annual precipitable water for a longer period of record.

Highest values in the southern hemisphere appear between 5° and 25° latitude over South America and within 10° of the equator over East Africa and near New Guinea (also regions of maximum thunderstorm frequency as discussed



(after Peixoto, 1970)

Figure 7-1 Mean Annual Precipitable Water; Northern Hemisphere



(after Peixoto, 1970)

Figure 7-2 Vertical Cross Section of Mean Specific Humidity; Northern Hemisphere

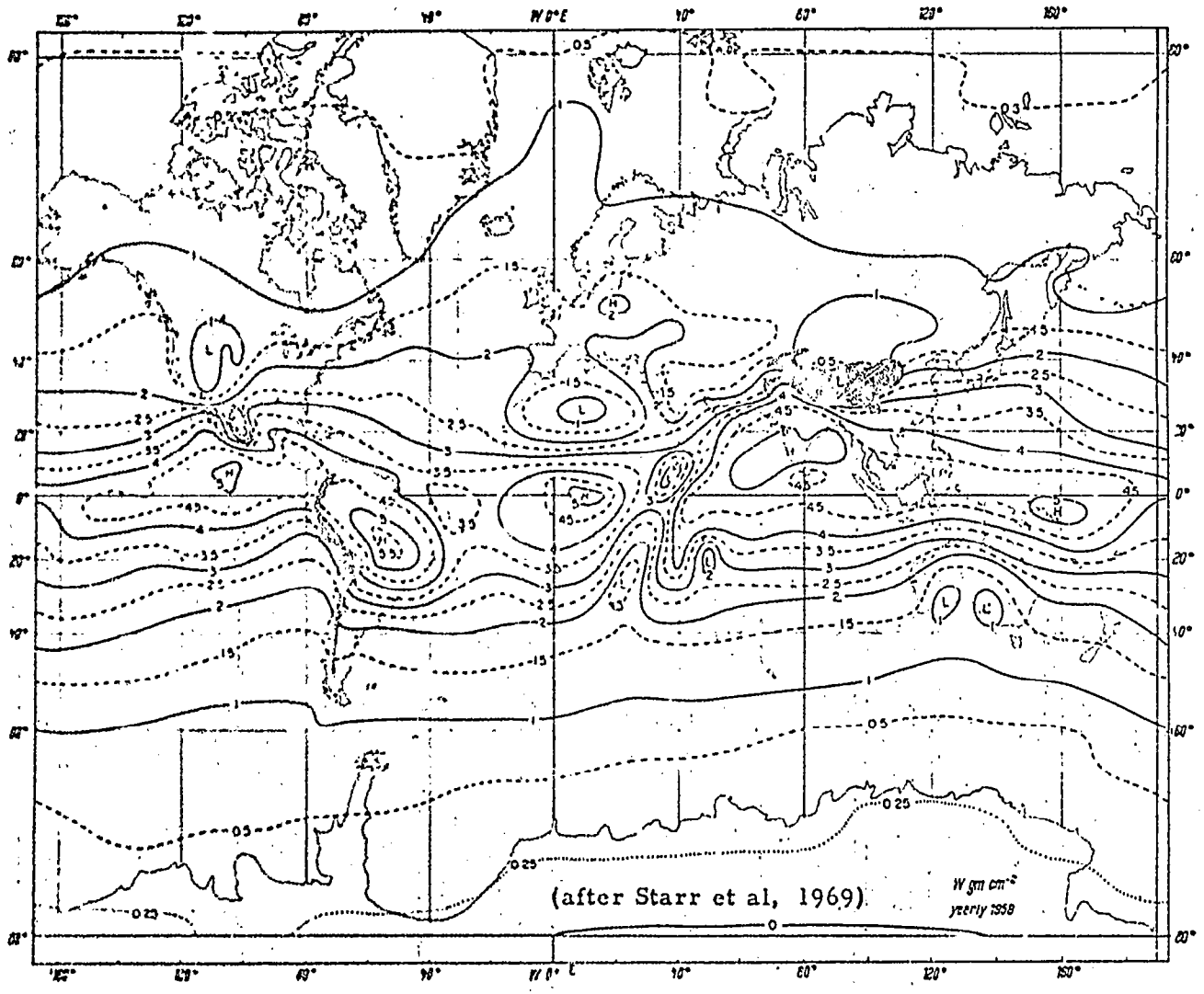


Figure 7-3 Time Averaged Global Distribution of Precipitable Water; 1958

in Section 6.1). Minimum values occur over the desert regions of Australia and there is a gradual north to south decrease to a worldwide low value of less than .25 gm cm^{-2} over Antarctica.

7.2 North American Distributions - Monthly

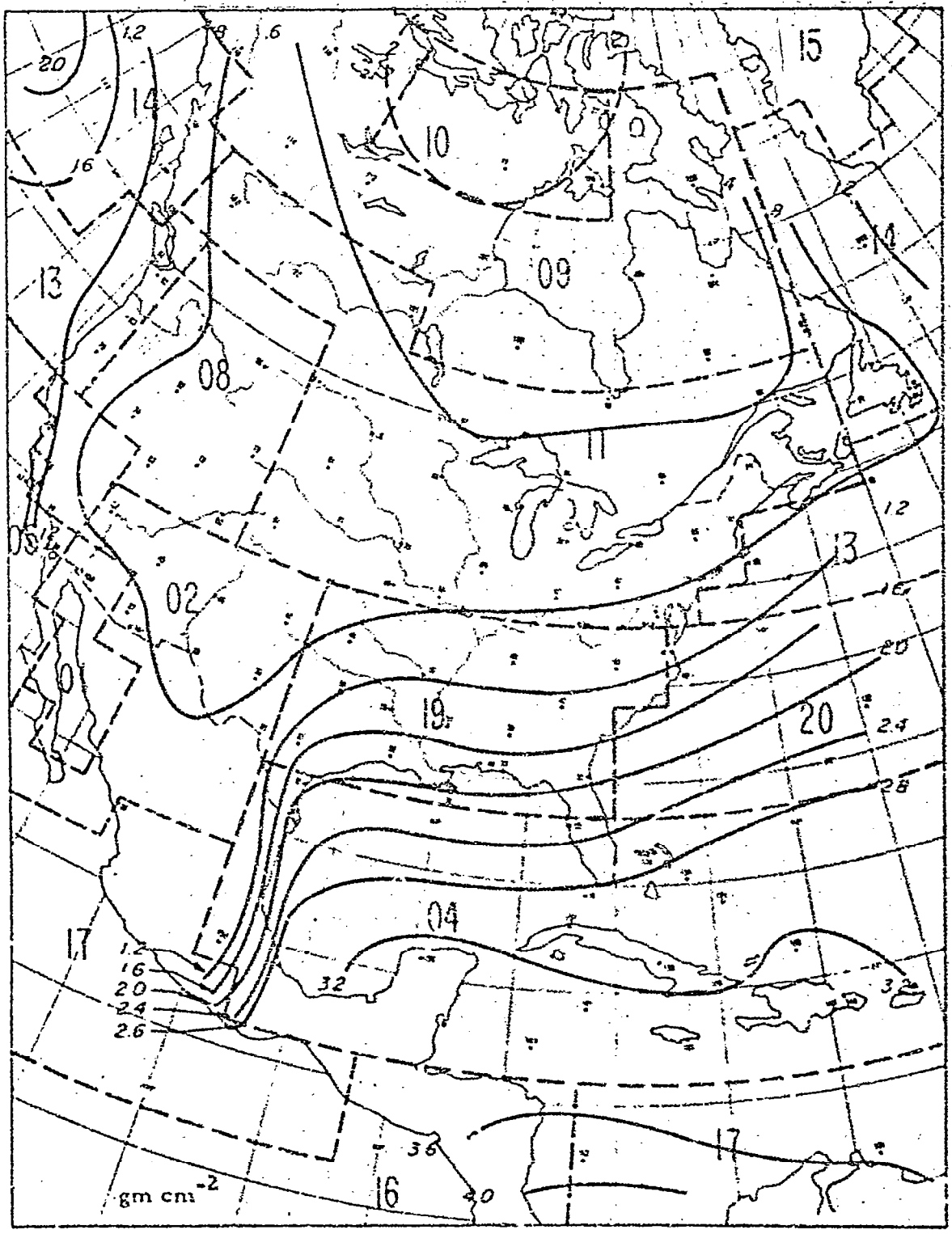
Monthly distributions were available only for North America. Mean monthly precipitable water charts based on five years of data were plotted from a tabulation of moisture data over North America (ESSA, 1966). The charts were analyzed and the distribution patterns for midseason months are shown in Figures 7-4 through 7-7. Homogeneous cloud regions are superimposed on the figures. The mean patterns behave in a regular fashion. Highest moisture values are present in Summer, gradually lowering through Fall in the mid and high latitudes and lowest in Winter, increasing again especially in the mid and high latitudes in Spring. In Winter (January) strong moisture gradients are evident across cloud Regions 19 and 20 and the northwest portion of Region 4. In the Summer (July) Regions 4, 19 and 20 are more homogeneous and the gradients exist further north in the southern portions of Regions 11 and 13.

The relatively dry conditions across Central Mexico (due to the high terrain) cause large gradients between Regions 17 and 14 during most of the year. Other exceptions to the normal south to north decrease in moisture are dry conditions in the southwest desert and the mountain regions of western North America.

In general, homogeneous cloud regions do not necessarily reflect homogeneous moisture regions. (Although at some times of the year, Winter in the high latitudes and Summer in the tropical and subtropical latitudes, absolute moisture is relatively homogeneous within the cloud regions.) The primary reason is that absolute moisture is a function of the temperature and the temperature pattern displays the normal south to north decrease. Clouds, on the other hand, are dependent on the relative humidity which is a function of the difference between the temperature and dewpoint.

7.3 Frequency Distributions of Moisture for the United States

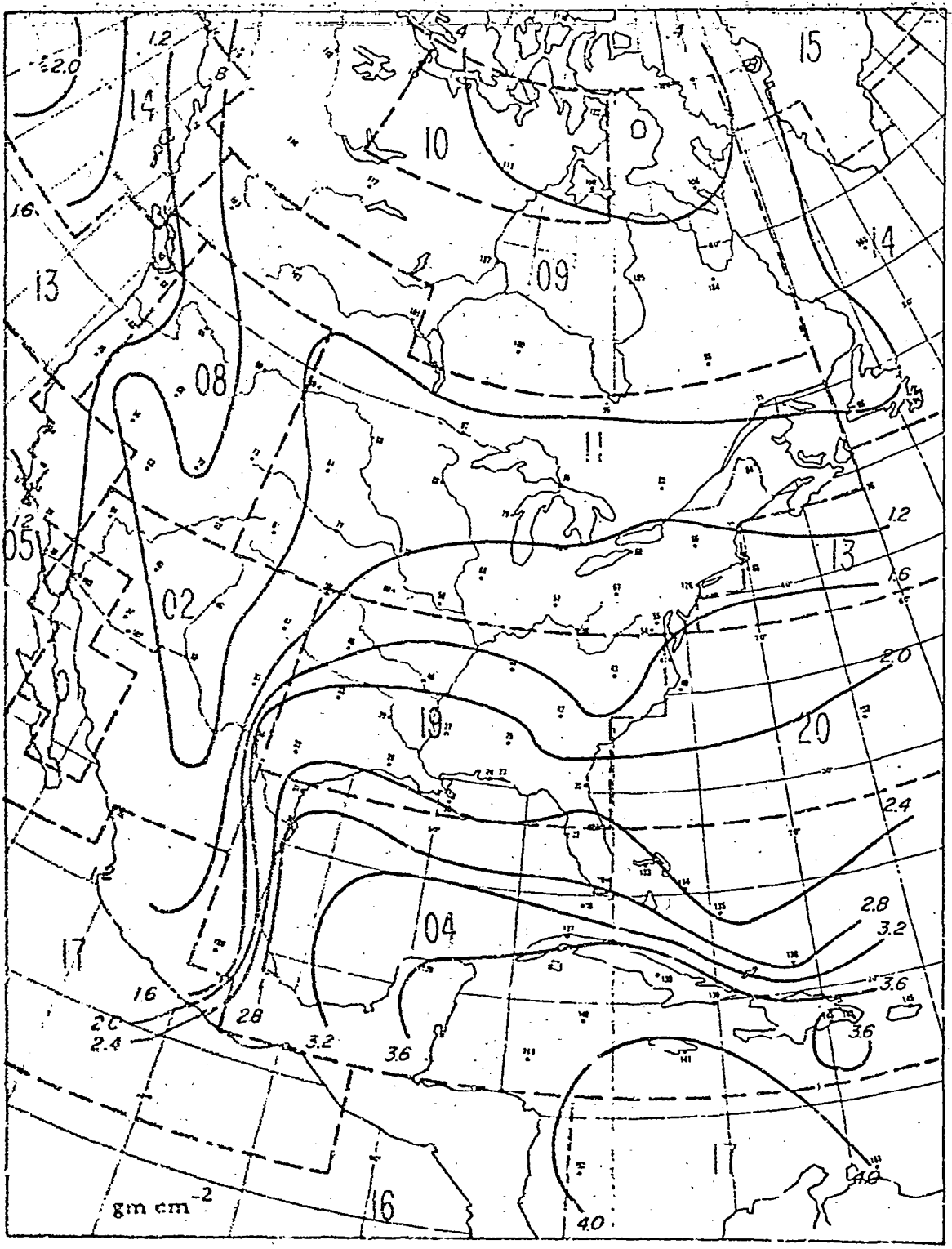
Inquiries revealed that a source of daily precipitable water values for 69 radiosonde stations was available on magnetic tape from the National Weather



Homogeneous Cloud Regions Superimposed

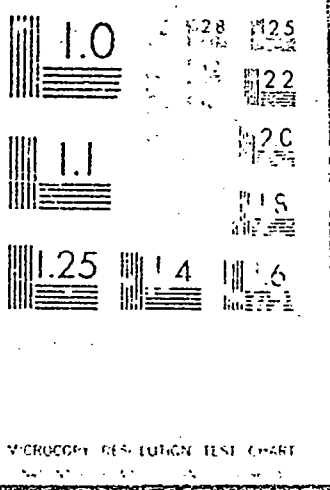
Figure 7-4 Mean Precipitable Water Distribution for North America; January

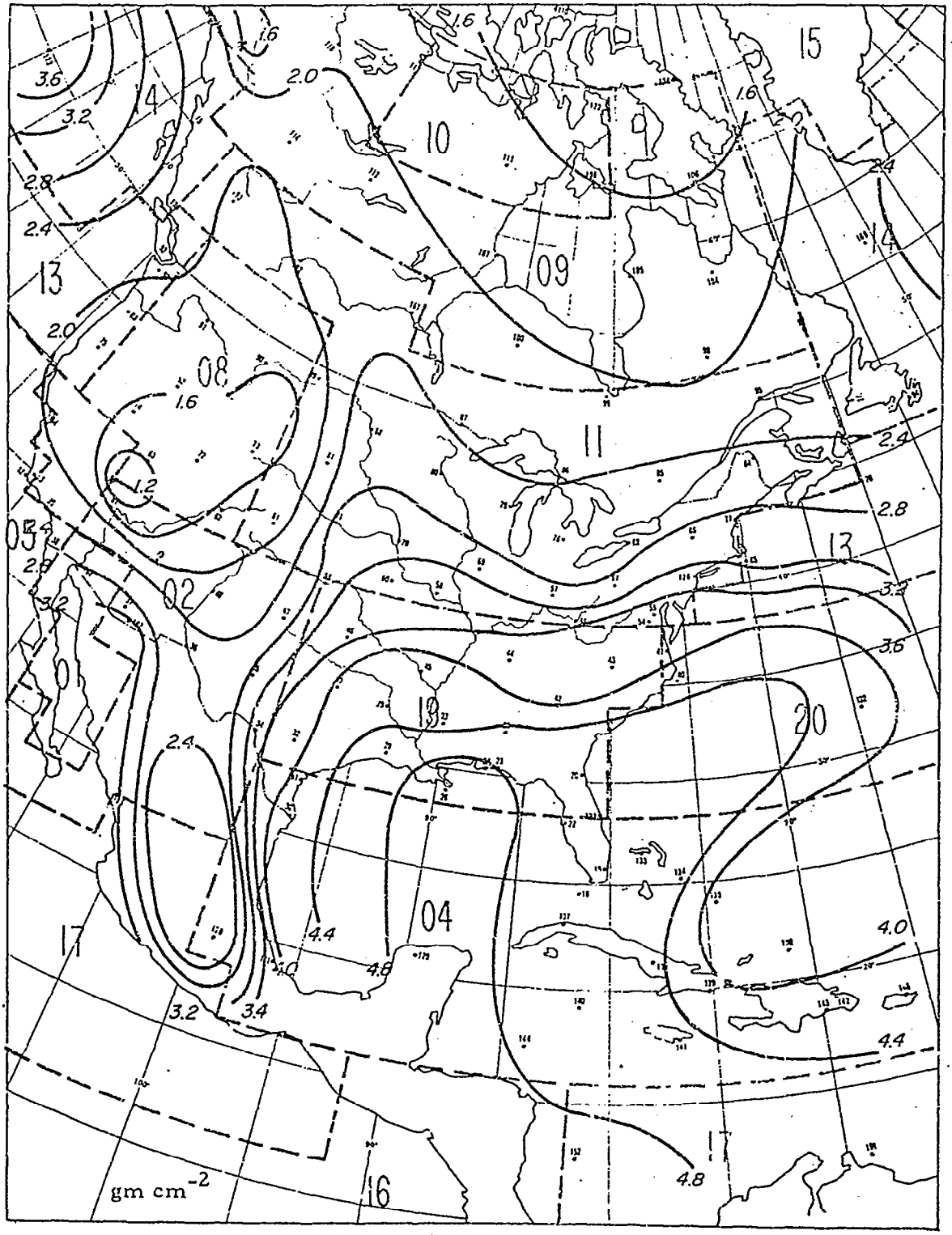
7617



Homogeneous Cloud Regions Superimposed

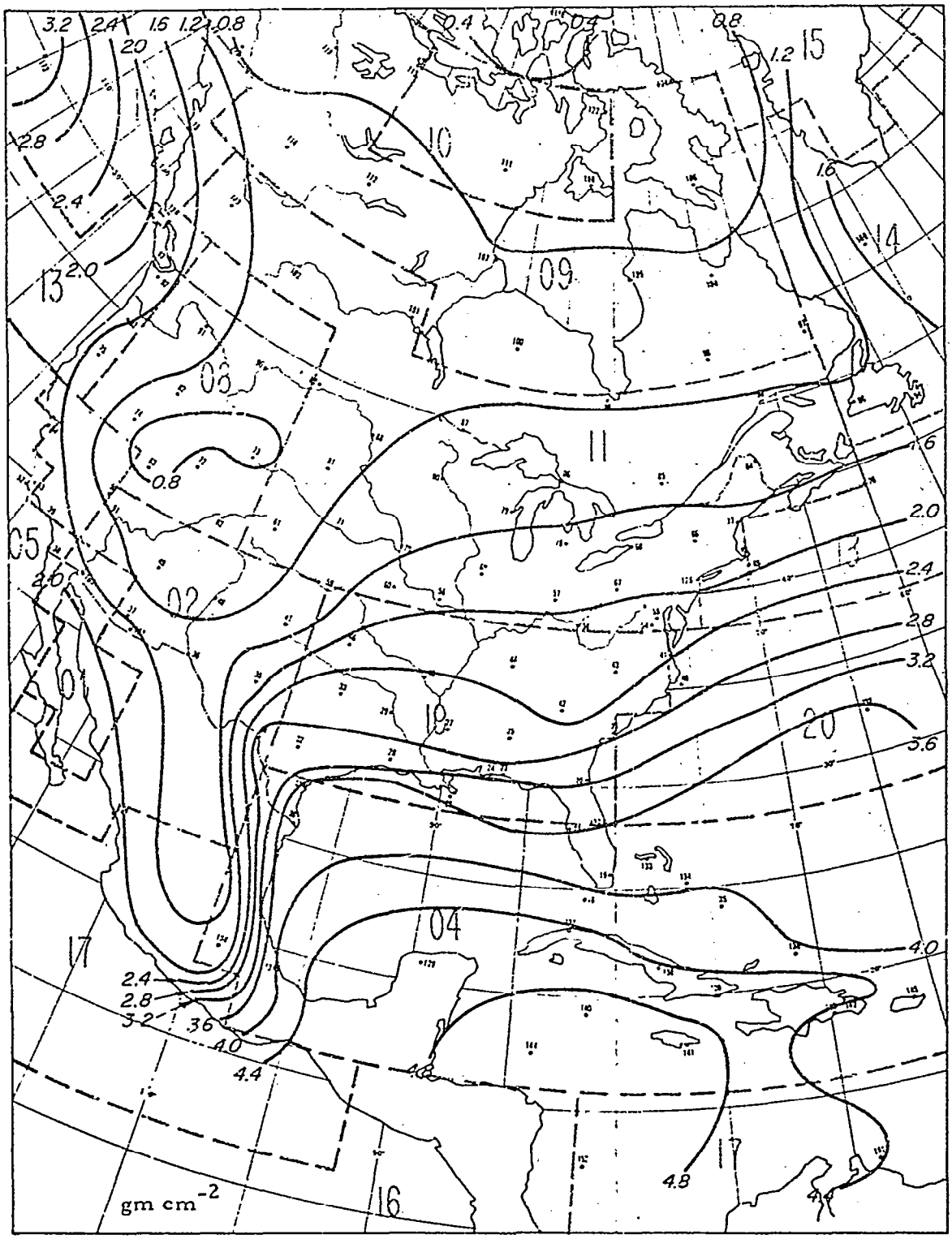
Figure 7-5. Mean Precipitable Water Distribution for North America; April





Homogeneous Cloud Regions Superimposed

Figure 7-6 Mean Precipitable Water Distribution for North America; July



Homogeneous Cloud Regions Superimposed

Figure 7-7 Mean Precipitable Water Distribution for North America; October

7615

Service. In general, about eight years of twice-daily data are available for most stations.

We defined 10 categories of precipitable water amount (.5 gm cm^{-2} intervals from 0 to 4 gm cm^{-2}, 4 to 5 gm cm^{-2} and greater than 5 gm cm^{-2}) and wrote a computer subroutine to generate frequency distributions of precipitable water (averaged for each day) within each category for all stations for each month of the year. The detailed computer output from this program is given in Appendix D, but some representative distributions were plotted for eight stations for the midseason months. These are shown in Figures 7-8 through 7-15. Tropical, subtropical, mid-latitude, mountain, desert, and coastal regions are represented by the figures. Results clearly show an annual cycle for all stations. The frequencies are somewhat more peaked than might be expected during transition months for some of the stations; e.g., Tucson, Caribou, Winnemucca, and San Diego, in April. Not surprisingly, winter and summer distributions are peaked at most stations in the low and high moisture categories, respectively. Interestingly, Tucson, a desert station, exhibits rather high frequency of high precipitable water amounts for the Summer, while Tatoosh, a location with heavy annual precipitation shows the greatest frequencies in the low to moderate precipitable water categories. Winnemucca at an elevation near 3 km has very dry conditions most of the year. Key West, having a tropical maritime-type climate much of the time during the year has the highest frequencies in the high moisture categories. Mid-latitude stations have generally flatter distributions.

An analysis of the frequency (probability) distributions in Appendix D is a simple method for determining maximum and minimum as well as showing the rate of change over the year. Examples of such an analysis are given in Figures 7-16 and 7-17 for Lake Charles, Louisiana, and Yucca Flats, Nevada, respectively. Little change in the distributions is seen in the relatively dry condition over the year for Yucca Flats, at an elevation above 2 km, while major changes take place through the year at Lake Charles.

Although daily precipitable water data for the remainder of North America were unavailable, one can infer the distributions for both Canada and the tropical regions over the Caribbean and in Central Mexico. This can be accomplished by studying both the mean monthly values for North America and the frequency distributions at appropriate locations near the region and applying adjustments based on meteorological experience. For example, Key West, Florida, has relatively high mean values for most of the year and is in pure maritime tropical air most of the Summer. Thus, the frequency distribution for Kingston, Jamaica, can be

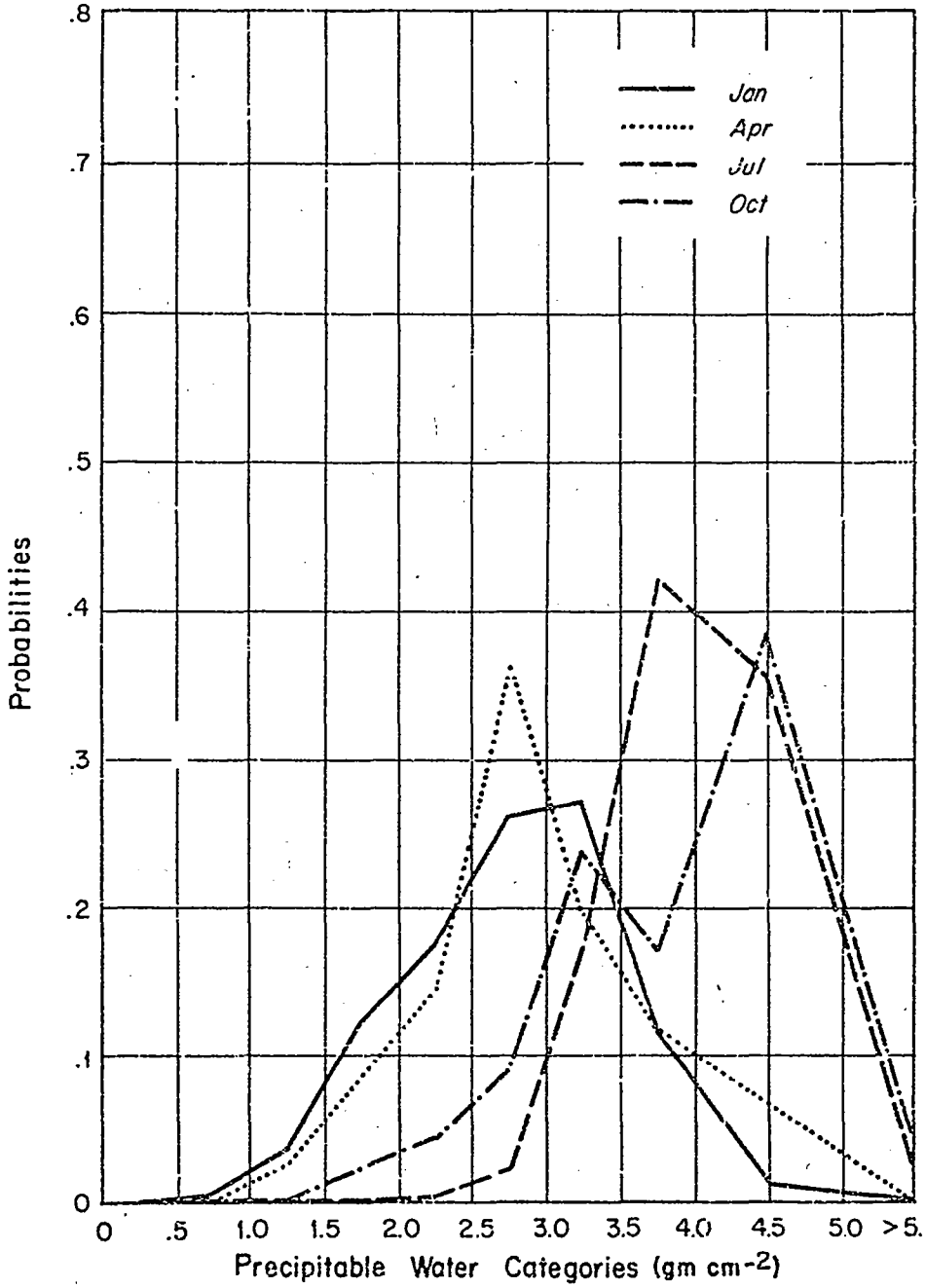


Figure 7-8 Precipitable Water Frequency Distributions; Mid-Season Months for Key West, Florida

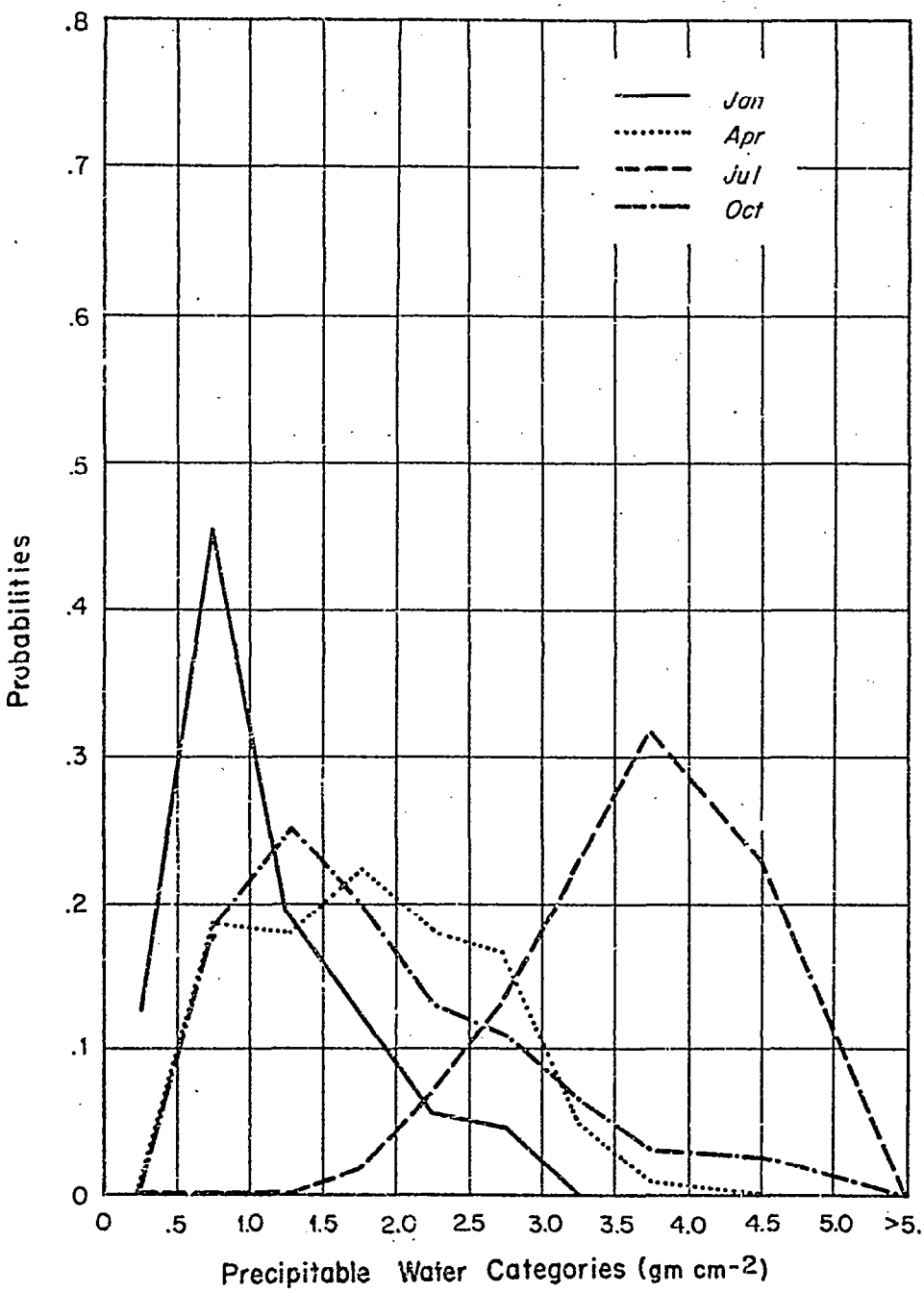


Figure 7-9 Precipitable Water Frequency Distributions; Mid-Season Months for Athens, Georgia

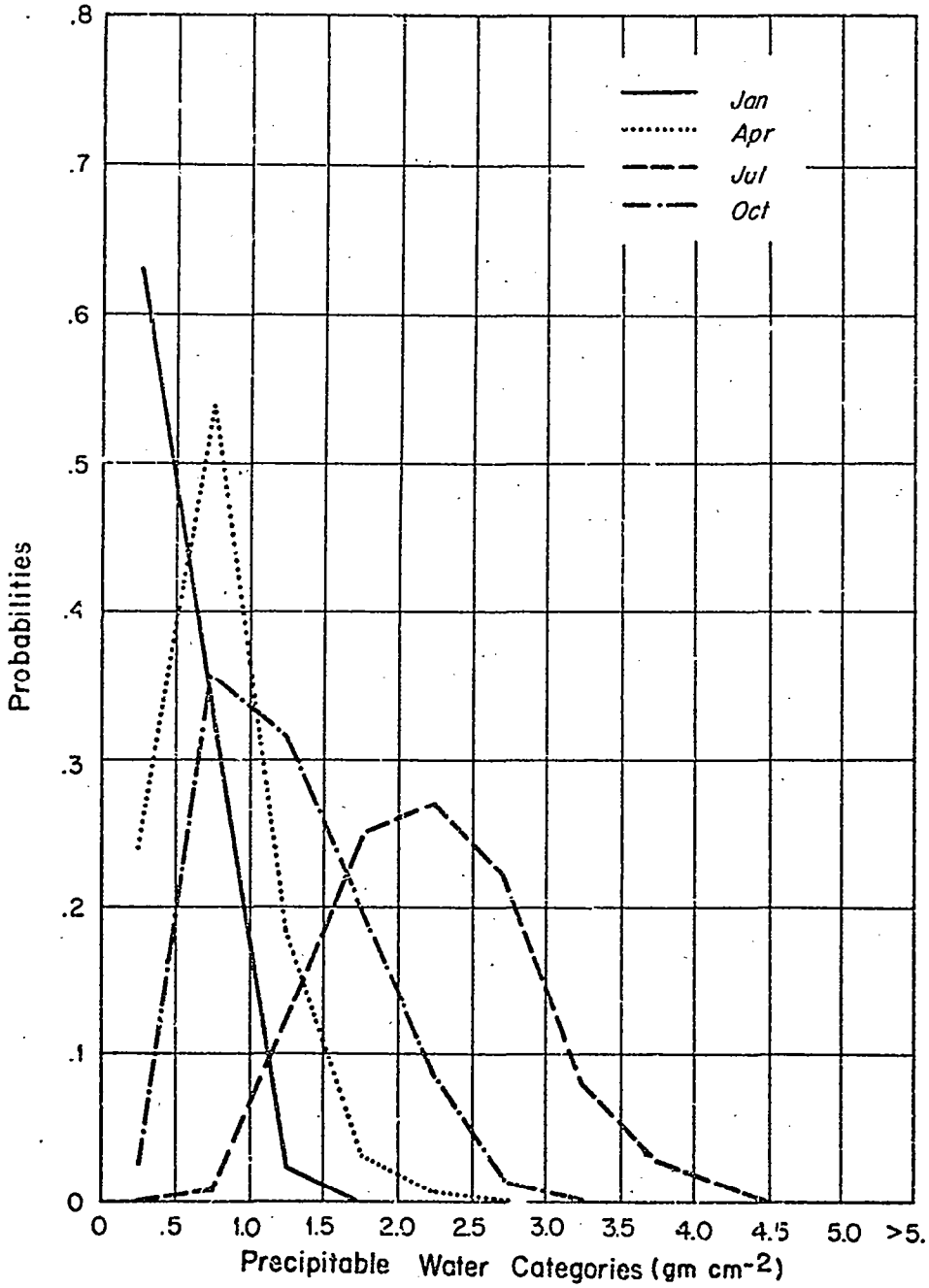


Figure 7-10 Precipitable Water Frequency Distributions; Mid-Season Months for Caribou, Maine

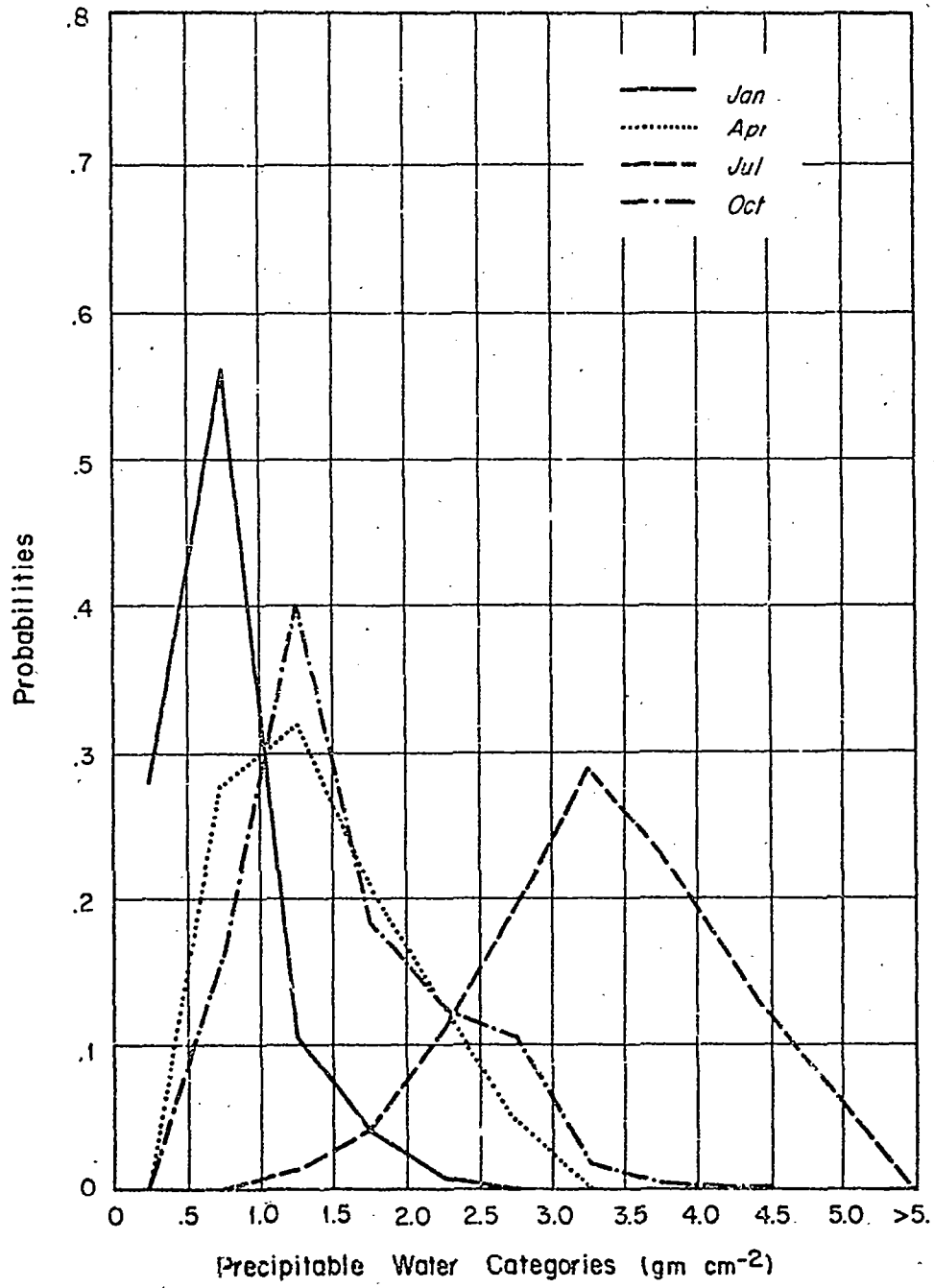


Figure 7-11 Precipitable Water Frequency Distributions; Mid-Season Months for Topeka, Kansas

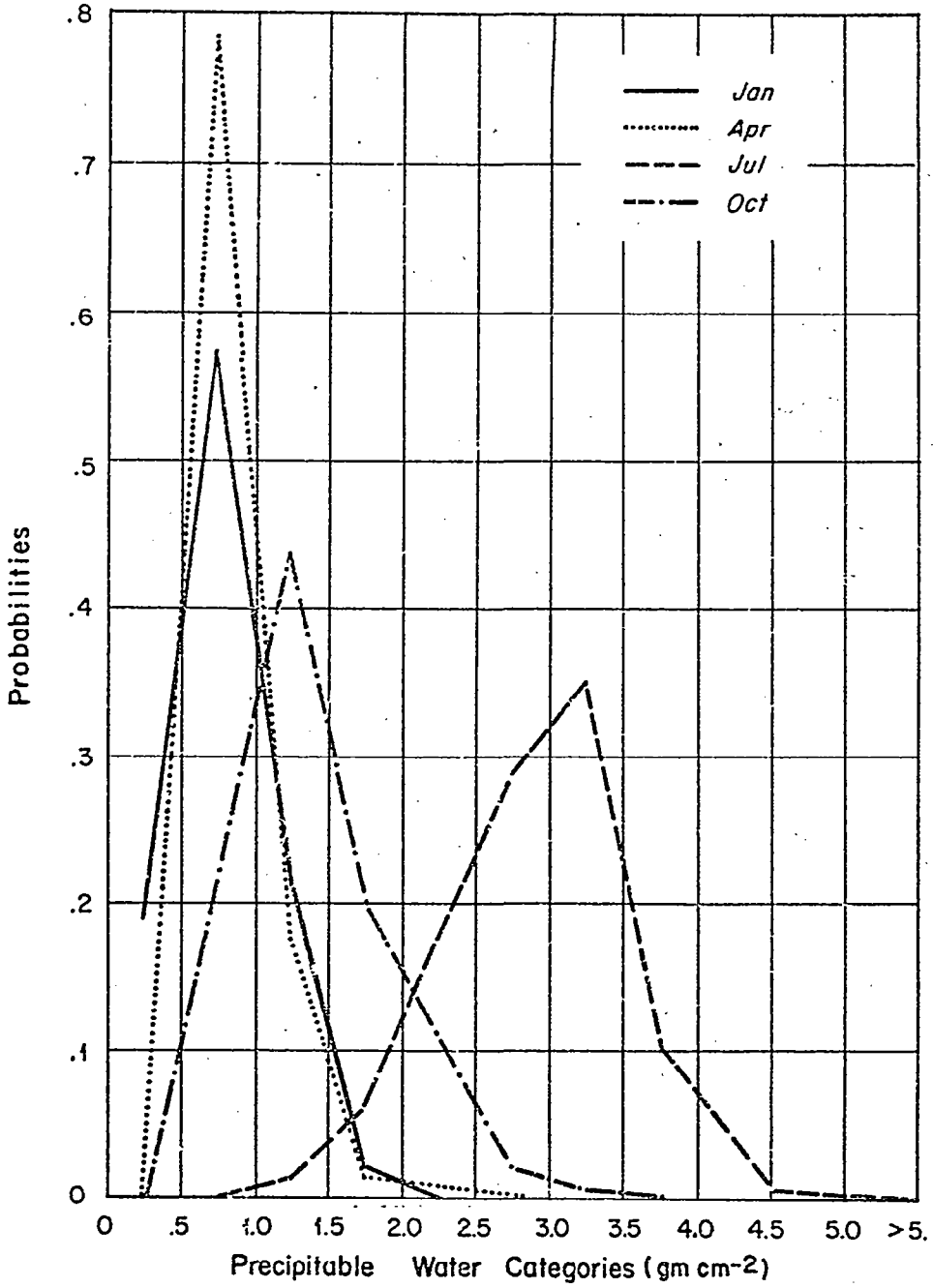


Figure 7-12 Precipitable Water Frequency Distributions; Mid-Season Months for Tucson, Arizona

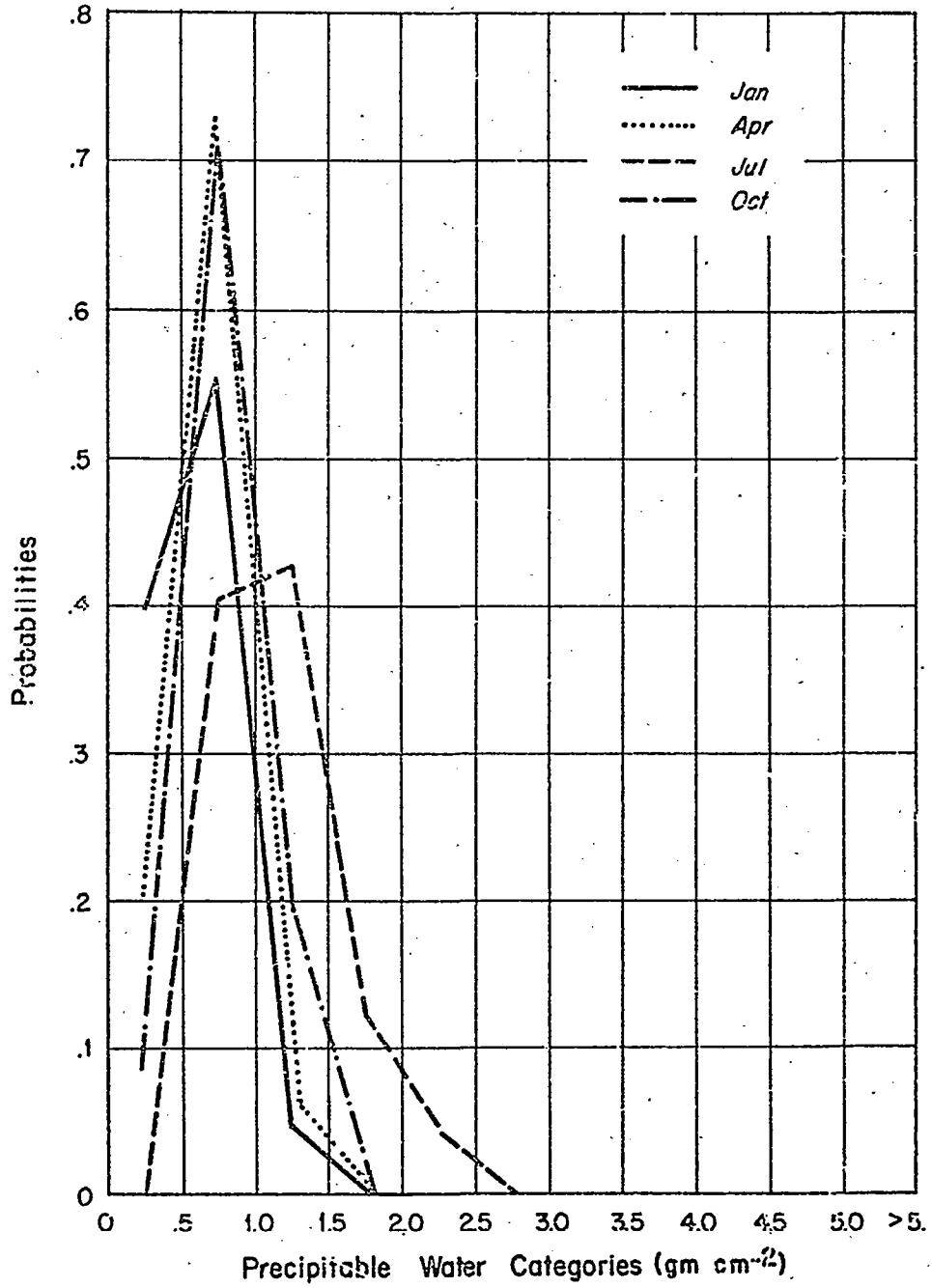


Figure 7-13 Precipitable Water Frequency Distributions; Mid-Season Months for Winnemucca, Nevada

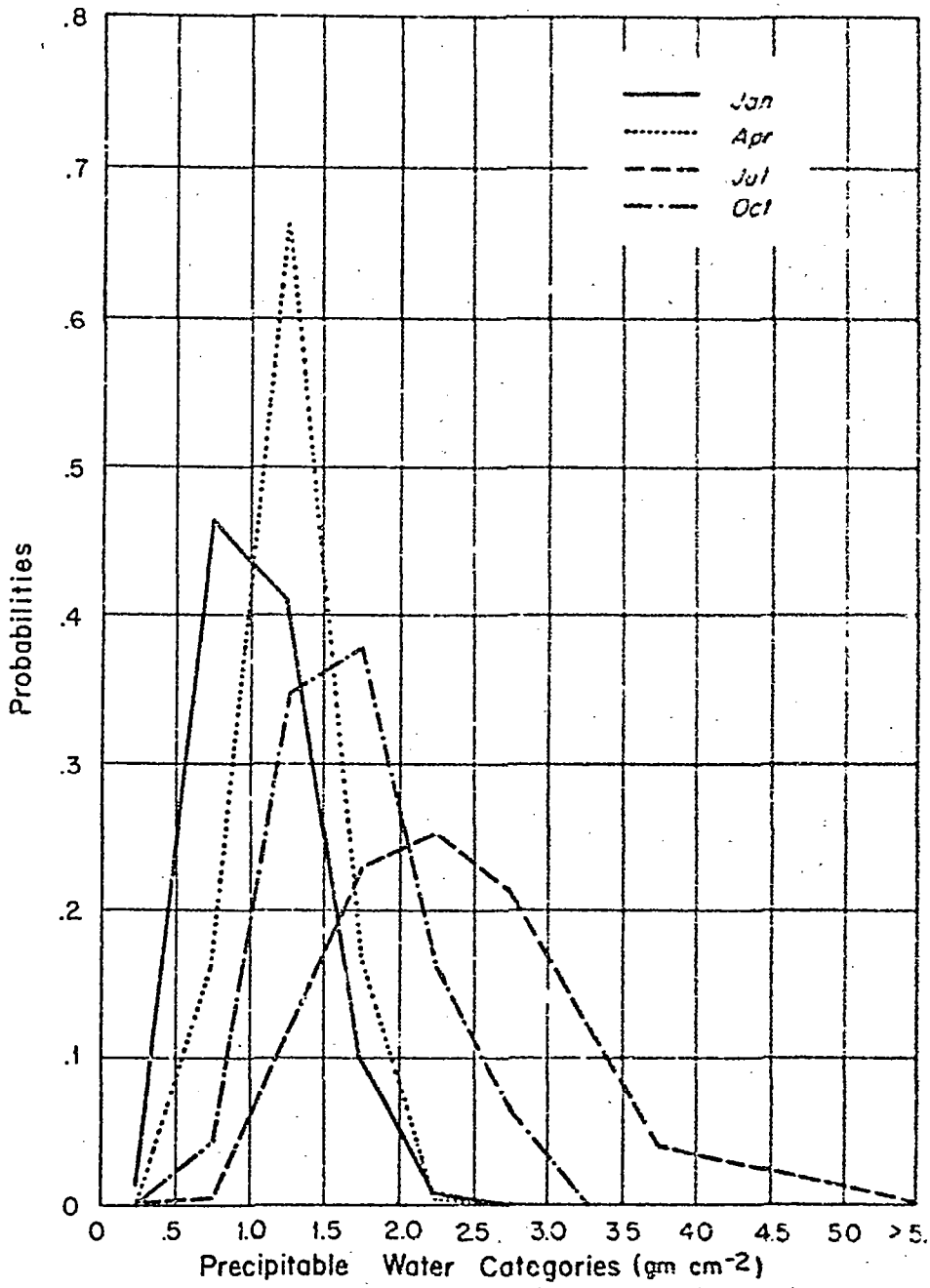


Figure 7-14 Precipitable Water Frequency Distributions; Mid-Season Months for San Diego, California

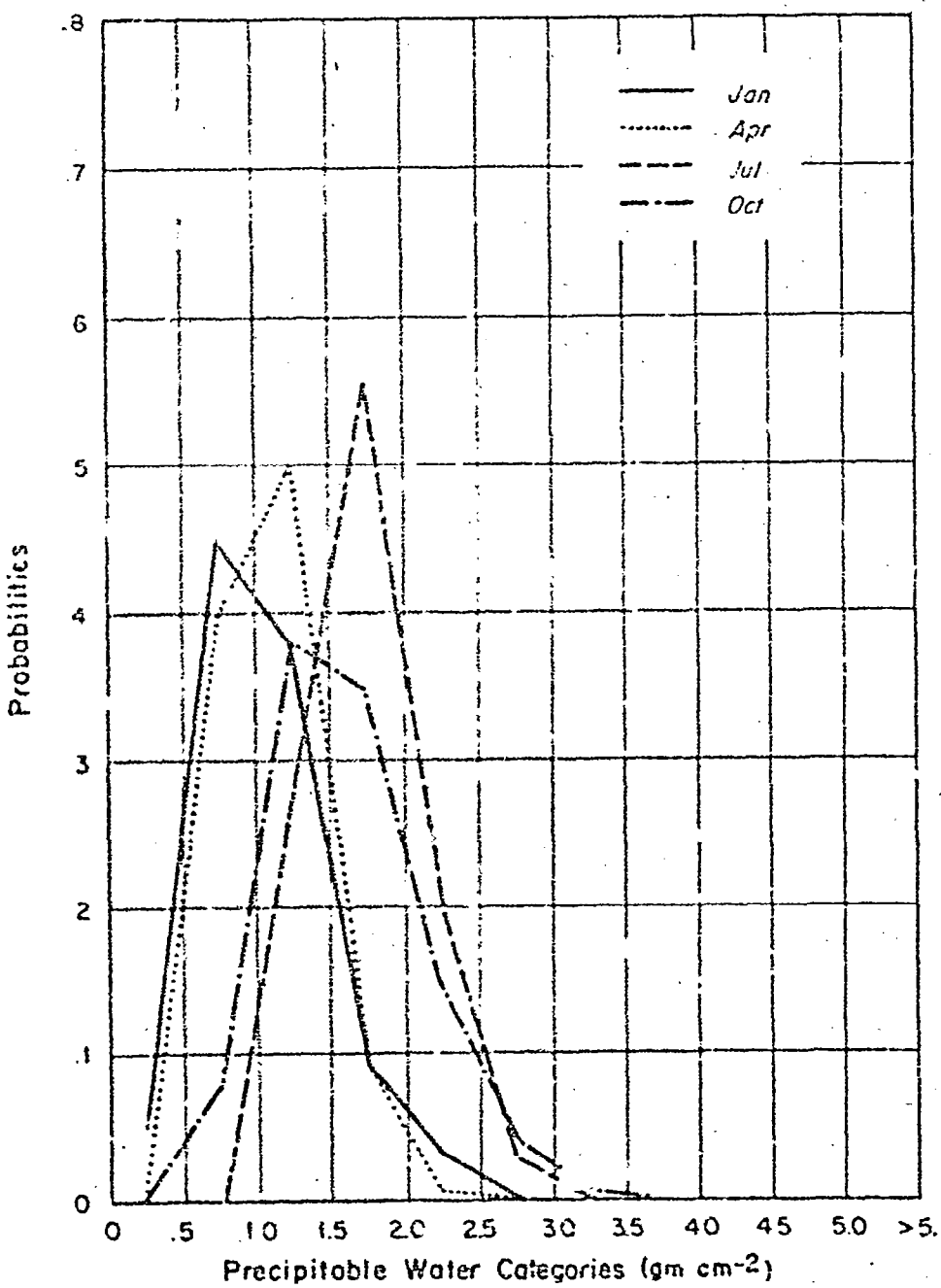


Figure 7-15 Precipitable Water Frequency Distributions; Mid-Season Months for Tatoosh, Washington

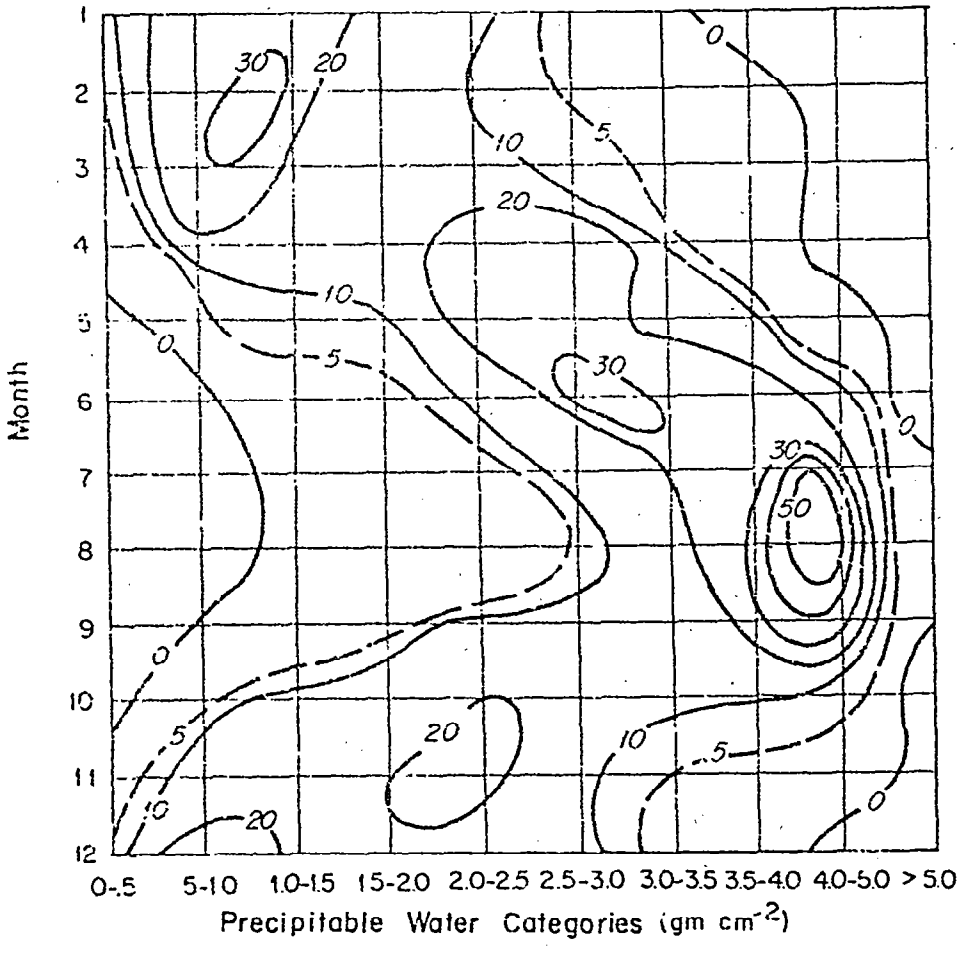


Figure 7-16 Analysis of Precipitable Water Frequency Distributions; Lake Charles, Louisiana

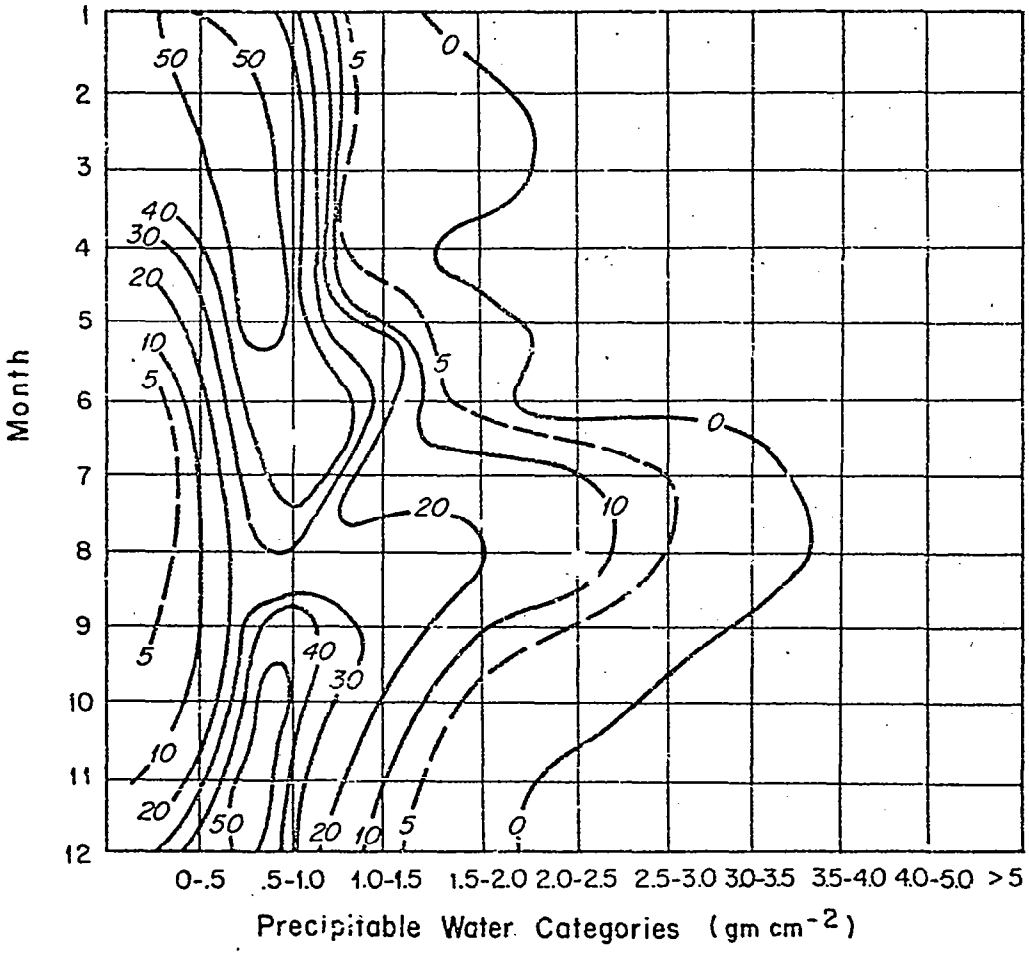


Figure 7-17 Analysis of Precipitable Water Frequency Distributions; Yucca Flats, Nevada

inferred by making the Spring and Fall distributions similar to Key West except adjusting it to have it somewhat higher, and more peaked. The Winter distribution at Kingston would probably be of the same shape as Key West except shifted slightly to the higher categories. Similarly, one could infer distributions for Canadian stations from northern border stations in the United States and could infer Central Mexico distributions from stations at high elevations in the southwest United States.

Careful examination of the data in Appendix D indicates widely differing frequency distributions between some stations with the same homogeneous cloud regions (e.g., International Falls, Minnesota, and Sterling, Virginia, in Region 11; and Valparaiso, Florida, and Oklahoma City, Oklahoma, in Region 19). Thus, the frequency distributions confirm our earlier contention that it is not advisable to consider the homogeneous cloud regions as homogeneous moisture regions.

8. CURRENT STATUS OF GLOBAL CLOUD MODEL

The ultimate goal of the research described in this report is to develop a global model of the effects of cloudiness on potential remote sensing missions. In time, the results will be combined with those of other studies concerned with developing atmospheric attenuation models to yield an overall model of the effects of the intervening atmosphere on earth observations from space. In that light, it is the purpose of this section to review what has been accomplished thus far and to suggest a future course of action.

The weak link of any global cloud model is lack of adequate data. For many parts of the world, these data simply do not exist. Now that global satellite coverage is being obtained on a more or less regular basis, however, we will be able to slowly accrue the necessary information for data-sparse regions. For other areas, such as the Continental United States, sufficient data already exist for refinement of the regional cloud boundaries. The assumption of homogeneous cloud climatologies has been found to be valid for some regions and seasons and less valid for others. Where possible the boundaries should be redrawn and new regions added. The "seasonal reversal" hypothesis has been found to be only marginally successful (at least for oceanic areas) and in time can be discarded as more data become available. Given the success of the Markov scaling routine, future satellite data extraction procedures can be carried out more efficiently with only two data extraction circles spaced 200 miles apart. Some thought should be given to the application of automated techniques to the currently time-consuming task of manual cloud-amount data extraction. The data bank as it now stands is internally consistent and is the most comprehensive tabulation of global cloud data available.

The impossibility of tabulating conditional cloud-amount statistics for all possible time and distance separations necessitates the development of techniques for adjusting the conditionals as tabulated to other time and distance separations. The Markov scaling technique developed and tested during the current study represents a significant improvement over the former linear scaling method, and within the limits of the data available, is probably the best possible approach. Techniques are also required to modify the existing data base so that it may be applied to other representative area sizes. During the current study a new area enlargement procedure has been developed which is more mathematically sound than its predecessor and is better able to simulate the effects of area enlargements. It is limited, however, to a repetitive doubling of area size (with possible interpolations)

and is unable to simulate area reductions. We are encouraged by the success of an area adjustment technique for unconditional distributions which utilizes a normal distribution assumption for middle cloud amounts. It should be possible to extend this technique to conditional arrays.

We are currently able to simulate the effective cloud cover for a sensor operating in the visible range with an ESSA or Nimbus resolution (1.5 to 2 n.mi.). Extension of the simulation results to other sensor resolutions requires some knowledge of the horizontal distribution or spatial frequencies of the cloudiness. Some preliminary work relating effective cloud cover to the cloud-size/resolution-size ratio has already been done. Extension of the simulation results to other sensor types such as infrared or microwave sensors requires in addition some knowledge of cloud thickness, height, water drop size and water content. One possible approach here would be to supplement the data bank with cloud-type information (possibly just layer vs. cellular for data-sparse regions) and to then relate the cloud structure and composition parameters to particular cloud types. The combination of these two classes of information (type frequencies and type characteristics) would make possible the introduction of a spectral dependency to our simulation results. Even with these limitations, the simulation procedures as they now stand provide a better estimate of earth viewing probabilities than any other techniques currently available.

In summary, the current status of the global cloud model is as follows:

- In certain areas the cloud-amount data base is still relatively weak. As time and funds permit, more satellite data should be extracted, particularly for southern hemisphere regions.
- Where the data are available, regional boundaries should be refined using ground-observed cloud amounts. The Continental United States is a prime area for future realignments.
- With the possible exception of further refinements to the area adjustment procedures, the statistical techniques for time, distance and area scaling of the tabulated data base appear adequate.
- The basic Monte Carlo simulation procedure is effective in predicting potential cloud-amount situations.
- The introduction of cloud structure and composition data is necessary to extend the current simulation results to other spatial and spectral scales. This area should have the highest priority for future efforts.

REFERENCES

Australian Bureau of Meteorology, 1964: Average Annual Thunderday Map of Australia (1954-1963), Dept. of the Interior, Melbourne, Australia.

Baum, R. A., 1970: "The Eastern Pacific Hurricane Season of 1969," Mo. Wea. Rev., 98(4), pp. 280-292.

Brown, S. C., 1969: "A Cloud-Cover Simulation Procedure," Astronautics and Aeronautics, 7(8), pp. 86-88.

Brown, S. C., 1970: "Simulating the Consequence of Cloud Cover on Earth-Viewing Space Missions," Bul. Am. Met. Soc., 51(2), pp. 126-131.

Brunt, A. T. and J. Hogan, 1956: "The Occurrence of Tropical Cyclones in the Australian Region," Proc. of the Tropical Cyclone Symposium, Brisbane, December 1956, Director of Meteorology, Melbourne, Australia, pp. 5-18.

Carlson, T. N., 1969: "Some Remarks on African Disturbances and their Progress Over the Tropical Atlantic," Mo. Wea. Rev., 97(10), pp. 716-726.

Cry, G. W., 1965: "Tropical Cyclones of the North Atlantic Ocean," U. S. Weather Bureau, Tech. Paper No. 55, Washington, D. C. 148 p.

Deming, W. E., 1964: Statistical Adjustment of Data, Dover Publications, New York, N. Y.

Denney, W. J., 1969: "The Eastern Pacific Hurricane Season of 1968," Mo. Wea. Rev., 97(3), pp. 207-224.

Director of Meteorology, 1965: Tropical Cyclones in the Northeastern and Northwestern Australian Regions for 1962-1963 Season, Melbourne, Australia.

Director of Meteorology, 1966: Tropical Cyclones in the Northeastern and Northwestern Australian Regions for 1963-1964 Season, Melbourne, Australia.

Director of Meteorology, 1968a: Tropical Cyclones in the Northeastern and Northwestern Australian Regions for 1964-1965 Season, Melbourne, Australia.

Director of Meteorology, 1968b: Tropical Cyclones in the Northeastern and Northwestern Australian Regions for 1965-1966 Season, Melbourne, Australia.

Director of Meteorology, 1969: Tropical Cyclones in the Northeastern and Northwestern Australian Regions for 1966-1967 Season, Melbourne, Australia.

Dunn, G. E. and B. I. Miller, 1960: Atlantic Hurricanes, Louisiana State University Press, 326 p.

REFERENCES, contd.

Duntley, S. Q., C. F. Edgerton and T. J. Petzold, 1970: Atmospheric Limitations on Remote Sensing of Sea Surface Roughness by Means of Reflected Daylight, Final Report Contract NAS 12-2126, Scripps Institute of Oceanography.

Environmental Science Services Administration, 1966: Tabulation of Monthly Values of Vertically Integrated Moisture and Moisture Transport for North America, May 1958-April 1963, Final Report 7477-244, Contract Cwb-11313, The Travelers Research Center, Inc., Hartford, Connecticut.

Frank, N. L., 1970: "Atlantic Tropical Systems of 1969," Mon. Wea. Rev., 98(4), pp. 307-314.

Gabites, J. F., 1956: "A Survey of Tropical Cyclones in the South Pacific," Proc. of the Tropical Cyclone Symposium, Brisbane, December 1956, Director of Meteorology, Melbourne, Australia, pp. 19-24.

Kendall, G. R. and A. G. Petrie, 1962: The Frequency of Thunderstorm Days in Canada, Meteorological Branch, Dept. of Transport, Toronto, Ontario.

Peixoto, Jose P., 1970: "Water Vapor Balance of the Atmosphere from Five Years of Hemispheric Data," Nordic Hydrology, 2, pp. 120-138.

Sadler, J. C., 1964: "Tropical Cyclones of the Eastern North Pacific as Revealed by TIROS Observations," J. of Appl. Met., 3(4), pp. 347-366.

Shenk, W. E. and V. V. Salomonson, 1971: A Simulation Study Exploring the Effects of Sensor Spatial Resolution on Estimates of Cloud Cover from Satellites, NASA Tech. Note (to be issued), Goddard Space Flight Center, Greenbelt, Maryland.

Sherr, P. E., A. H. Glaser, J. C. Barnes and J. H. Willand, 1968: World-Wide Cloud Cover Distributions for Use in Computer Simulations, Final Report Contract NAS 8-21040, Allied Research Associates, Inc.

Showalter, A. K. and J. R. Fulks, 1943: Preliminary Report on Tornadoes, U. S. Weather Bureau.

Simpson, R. H., N. Frank, D. Shideler and H. M. Johnson, 1968: "Atlantic Tropical Disturbances, 1967," Mon. Wea. Rev., 96(4), pp. 251-259.

Simpson, R. H., N. Frank, D. Shideler and H. M. Johnson, 1969: "Atlantic Tropical Disturbances, 1968," Mon. Wea. Rev., 97(3), pp. 240-255.

Staff, Severe Local Storms Unit, 1969: Severe Local Storm Occurrences 1955-1967, ESSA Tech. Memo WBTM-FCST12, (M. Pautz, ed) Tech. Proc. Br., Silver Spring, Maryland.

Starr, V. P., J. P. Peixoto and R. G. McKean, 1969: "Pole to Pole Moisture Conditions for the IGY," Pure and Applied Geophysics, 75, pp. 300-331 (formerly Geofisica pura e applicata).

Stephan, F. F., 1942: "An Iterative Method of Adjusting Sample Frequency Tables When Expected Marginal Totals are Known," Annals of Mathematical Statistics, 13(2), pp. 166-178.

REFERENCES, contd.

U. S. Navy, 1955: Marine Climatic Atlas of the World, NAVAER 50-1C-528, U. S. Govt. Printing Office, Washington, D. C.

Van Loon, H., 1966: "On the Annual Temperature Range over the Southern Oceans," The Geographical Review LVI, No. 4, pp. 497-515.

World Meteorological Organization, 1953: World Distribution of Thunderstorm Days, Part 1: Tables, WMO/OMM, No. 21. TP. 6, Geneva, Switzerland.

World Meteorological Organization, 1956: World Distribution of Thunderstorm Days, Part 2: Tables of Marine Data and World Maps, WMO/OMM, No. 21. TP. 21, Geneva, Switzerland.

MSFC—RSA, Ala

END

DATE

FILMED

JUL 23 1971