AgRISTARS

YM-NI-04198
JSC-17784

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

Yield Model Development

November 1981

DEVELOPMENT OF A SURFACE ISOLATION ESTIMATION TECHNIQUE SUITABLE FOR APPLICATION OF POLAR ORBITING SATELLITE DATA

P. A. Davis and L. M. Penn

Research and Data Systems, Inc.
9420 Anapolis Road
Lanham, MD 20801

Lyndon B. Johnson Space Center

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
YM-N1-04198, JSC-17784		

4. Title and Subtitle	5. Report Date
Development of a Surface Insolation Estimation Technique Suitable forApplication of Polar Orbiting Satellite Data	November 1981
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
P. A. Davis, L. M. Penn	
	10. Work Unit No.

9. Performing Organization Name and Address	11. Contract or Grant No.
Research and Data Systems, Inc. 9420 Annapolis Road Lanham, MD 20706	NA-80-SA C-00741
	13. Type of Report and Period Covered

12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
NOAA/NESS World Weather Bldg. Washington, DC 20233	

15. Supplementary Notes

16. Abstract

A technique is developed for the estimation of total daily insolation on the basis of data derivable from operational polar-orbiting satellites. Although surface insolation and meteorological observations are used in the development, the algorithm is constrained in application by the infrequent daytime polar-orbiter coverage.

17. Key Words (Suggested by Author(s))	18. Distribution Statement

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Non		67	

*For sale by the National Technical Information Service, Springfield, Virginia 22161

RDS RESEARCH AND DATA SYSTEMS, INC.

9420 ANNAPOLIS ROAD, LANHAM, MD. 20801 (301) 459-0001

DEVELOPMENT OF A SURFACE INSOLATION ESTIMATION TECHNIQUE SUITABLE FOR APPLICATION OF POLAR-ORBITING SATELLITE DATA

DEVELOPMENT OF A SURFACE INSOLATION ESTIMATION TECHNIQUE SUITABLE FOR APPLICATION OF POLAR-ORBITING SATELLITE DATA

FINAL REPORT

Prepared For

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL EARTH SATELLITE SERVICE
WORLD WEATHER BUILDING
WASHINGTON, D.C. 20233

ATTENTION: DR. D. TARPLEY

Under

CONTRACT NO. NA-80-SAC-00741

SEPTEMBER 1981

Prepared By

P. A. DAVIS
L. M. PENN

RESEARCH AND DATA SYSTEMS, INC.
9420 ANNAPOLIS ROAD
LANHAM, MARYLAND 20706

ABSTRACT

A technique is developed for the estimation of total daily insolation on the basis of data derivable from operational polar-orbiting satellites. Although surface insolation and meteorological observations are used in the development, the algorithm is constrained in application by the infrequent daytime polar-orbiter coverage.

A linear regression approach is used to relate observed insolation to the optical thickness, expressed through water vapor slant path and effective cloud transmittance (requiring cloud amount and type). Regression analyses relating these parameters to the total daily insolation result in daily estimates for each hour. Estimates corresponding to the times of satellite passage are weighted, according to their proximity to local solar noon, and averaged to provide a single daily insolation estimate.

Prior to final tuning of the coefficients, an interactive test of the procedure was conducted on independent data, with GOES satellite imagery as the data source for polar-orbiter times. Results were successful in terms of the ratio of the standard deviation of the residuals to the observed mean (less than 0.2 threshold) and correlation coefficients exceeding 0.80. After tuning the cloud parameterizations and developing a final set of regression coefficients, checks on independent surface data revealed that additional improvements can be made by introducing a bias reduction technique.

An automated procedure for cloud type classification based on the multispectral infrared sounder data was stipulated from one data swath and tested on another. This type of cloud classification from operational data shows considerable promise for future application.

This document has been prepared as a part of the FY81 Yield Model Development Project.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ACKNOWLEDGMENT

The contributions of Gary C. Chatters, especially in setting up and conducting pertinent initial processing-analysis routines, are greatly appreciated. We are also grateful to Dr. Dan Tarpley, the Contract Technical Officer, for his consultation and inputs.

1.0 INTRODUCTION

1.1 Agricultural Requirements

The Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing (AgRISTARS) program plan calls for estimates of solar radiation incident at the surface (insolation) based on operational satellite data. Insolation and other quantities are needed both for early warning of changes affecting resource quality and for commodity production forecasts. To meet requirements, techniques for estimation of insolation must be developed and then put into operational application for user testing over significant agricultural areas of the world.

Insolation is an AgRISTARS quantity more closely associated with input to crop yield models (for commodity production forecasts) than for other applications, but the insolation also is a contributor to the estimation of other quantities (e.g., soil moisture) that relate to early warning of changes of significance to agriculture. Global coverage requirements call for at least one technique that will provide local estimates of daily total insolation for any region of the globe.

Most of the recent effort to extract information on insolation from satellite data has concentrated on the analysis of radiance measurements of reflected solar radiation obtained from geosynchronous satellites. These platforms offer frequent (half-hour) high-resolution measurements of solar radiation failing to reach the surface. Although it is possible to cope with the data handling problems for routine processing of high-resolution digital data, not all geographical areas of interest are presently covered. Data from the operational polar-orbiting satellites provide an alternative with regular global coverage. The key question is whether or not the infrequent observations from the polar-orbiter will suffice for a specification of the daily total insolation, which depends critically on the cloudiness through the central sunlit hours. Utilization of the multispectral capabilities of the polar orbiters might compensate for the infrequency of observations at a given location.

It is contemplated that the insolation algorithm will initially be implemented interactively, in conjunction with precipitation estimation techniques. Optimum data characteristics (such as high resolution, or noontime data) may not be available. Therefore, the estimation technique must be flexible enough to make the best use of whatever information is available operationally.

1.2 Objectives

The major objective of this study has been the development of a technique for estimation of the daily total insolation using data acquired from operational polar-orbiting satellites. As a first phase, an algorithm for the technique is designed on the basis of observations at the surface. Adapted parameterizations are to be made from operational satellite data -- possibly from multispectral atmospheric sounder data. Alternatively, the parameters of the algorithm are estimated by interactive analysis of routine satellite imagery of moderate resolution, supplemented by occasional surface information or products from the operational satellite sounder. It is intended that the derived algorithm be made available for routine testing.

1.3 Scope

This report includes a very limited discussion of background, data sources, and preliminary analyses in the establishment of the approach. A discussion of the approach reveals limitations that have a bearing on the results obtained and on the prospects for future improvements.

Results (after modifications) of the regression analyses are summarized for each season from 2½ years of data over the United States. Tabulations of the latest weighting coefficients enable application of the method. A brief account of independent data tests is given, as well as a discussion of the feasibility study of the direct estimation of cloud parameters from multispectral infrared sounder data.

2.0 BACKGROUND

Methods that have been invoked to provide estimates of the total insolation on the basis of satellite-derived information have emphasized either an empirical approach or a simplified physical model. Rigorous radiative transfer computations, over the entire solar spectrum and including all orders of scattering, are not practicable for routine application. Furthermore, all pertinent optical information for the existing atmosphere and clouds at arbitrary times and locations are not available in any case. Thus, both empirical and physical models that are applied will make use of parameters that are likely to be based on climatological information. Whenever possible, however, the significant variations of clouds and water vapor will be depicted on the basis of operational information. Cloud properties and total water vapor, objectively described, become the important components of any technique. Remote sensing data might be used to infer such properties of interest indirectly, or might be needed to obtain some related information directly (e.g., cloud reflectance).

Early efforts to make use of satellite data in estimation of surface insolation (cf. Fritz, et al.[1]) demonstrated the relationship between satellite measurements and atmospheric attenuation. The significance of cloud amount has been evident in all efforts (cf. Hanson et al.[2], Quinn[3], Ellis and Vonder Haar[4]) to establish parameterized expressions. All results are dependent on the space-time scales and solar angles considered and, on the variable atmospheric attenuation; they are linked indirectly to surface reflectance. Not all recent models use satellite data but depend rather on conventional surface observations and parameterizations (cf. Atwater and Brown[5], Atwater and Ball[6]). These studies reveal the importance of cloud type information. With adaptations, satellite data could be used for input information if cloud types could be satisfactorily characterized. Another numerical model that has been applied extensively is that of Lacis and Hansen[7]. Again, for the direct adaptation of satellite data, translations to optical parameters are necessary.

Current approaches most characteristic of the use of solar reflectance measurements from geosynchronous satellites are the empirical regression model (Tarpley[8]) and the simplified physical model (Gautier[9]). Both methods require brightness measurements for both cloudy and cloud-free backgrounds. The Gautier model requires specification of atmospheric attenuation coefficients and cloud absorption coefficients, whereas the Tarpley model coefficients are based on least-squares regression.

The basic purpose of monitoring cloud reflectance is to determine the reduction in insolation. The more frequently this is done, the better will be the estimate of the daily insolation especially when the observations are made close to the central part of the day.

An alternative to measuring reflectance is to infer (from measurements) the transmittance of available radiation to the surface. This approach is necessary when insufficient daytime reflectance measurements are available, as is the case for this study. Ultimately, it is necessary to account for transmittance in any case when working with one-sided satellite measurements of upwelling radiance. (Another alternative is to base the estimate of insolation on a measured response of the surface, through its temperature, but such a program of remote measurement requires information on cloudiness and surface characteristics as well. Furthermore, it is preferable instead to estimate the impact of insolation on the thermal response).

Previous techniques have revealed the difficulties both of the statistical approach to the data and of the digital handling of the satellite data. Some bias in results is anticipated for estimates of extremes. As for the development of the application of satellite data, it will require good quality ground truth data of the insolation at localities that are representative of areas covered by remote sensing measurements.

3.0 DATA

3.1 SOLMET Data

SOLMET, hourly Solar Radiation - Surface Meteorological Observations, is an archived dataset, available on magnetic tape produced by the Environmental Data and Information Service of the National Climate Center.[10] SOLMET combines hourly incoming solar radiation data with the corresponding meteorological observations over the 38-station NOAA Solar Radiation Network into one comprehensive dataset. Two and one-half years of data from twenty of the SOLMET stations were used in this analysis. Figure 1 illustrates these locations geographically, while Table 1 provides pertinent station information.

For the insolation, use was made of the measured and edited global (total direct and diffuse) solar radiation on a horizontal surface. For the extraterrestrial radiation, listings of the computed available solar energy received by a horizontal surface at the top of the atmosphere (based on a solar constant of 1377 $J/m^2 \cdot s$), were used. The data are flagged to indicate source, if other than observed irradiance, or to indicate corrections. The meteorological data used consisted of surface pressure and dew point as well as detailed cloud layer type and amount information. (Visibility observations were not used in this study). Table 2 summarizes most of the SOLMET data actually incorporated into the analyses.

3.2 Satellite Data

The SOLMET data with the supporting hourly (and 3-hourly) surface meteorological observations cover only very small portions of the area under routine satellite surveillance. Each polar-orbiter can view almost any area twice a day (four times for the system pair of operational polar orbiting satellites). Of particular interest are the TIROS-N, NOAA series of satellites[11]. The operational satellites contain two scanning systems with visible and infrared radiometric measurements: (1) the Advanced Very High Resolution Radiometer (AVHRR), and (2) the High-Resolution Infrared Radiometer Sounder (HIRS). The latter scanner is part of the TIROS Operational Vertical Sounder (TOVS) system and plays a major role in the generation of TOVS retrieved products.

The AVHRR data are available in five different spectral ranges, two of which are in the short wavelength region (0.58-0.68 um and 0.725-1.10 um) and three in the infrared region (3.7 um, 11 um, 12 um). Spatial resolution near nadir is about 1.1 km. It was intended that daytime short wavelength AVHRR data

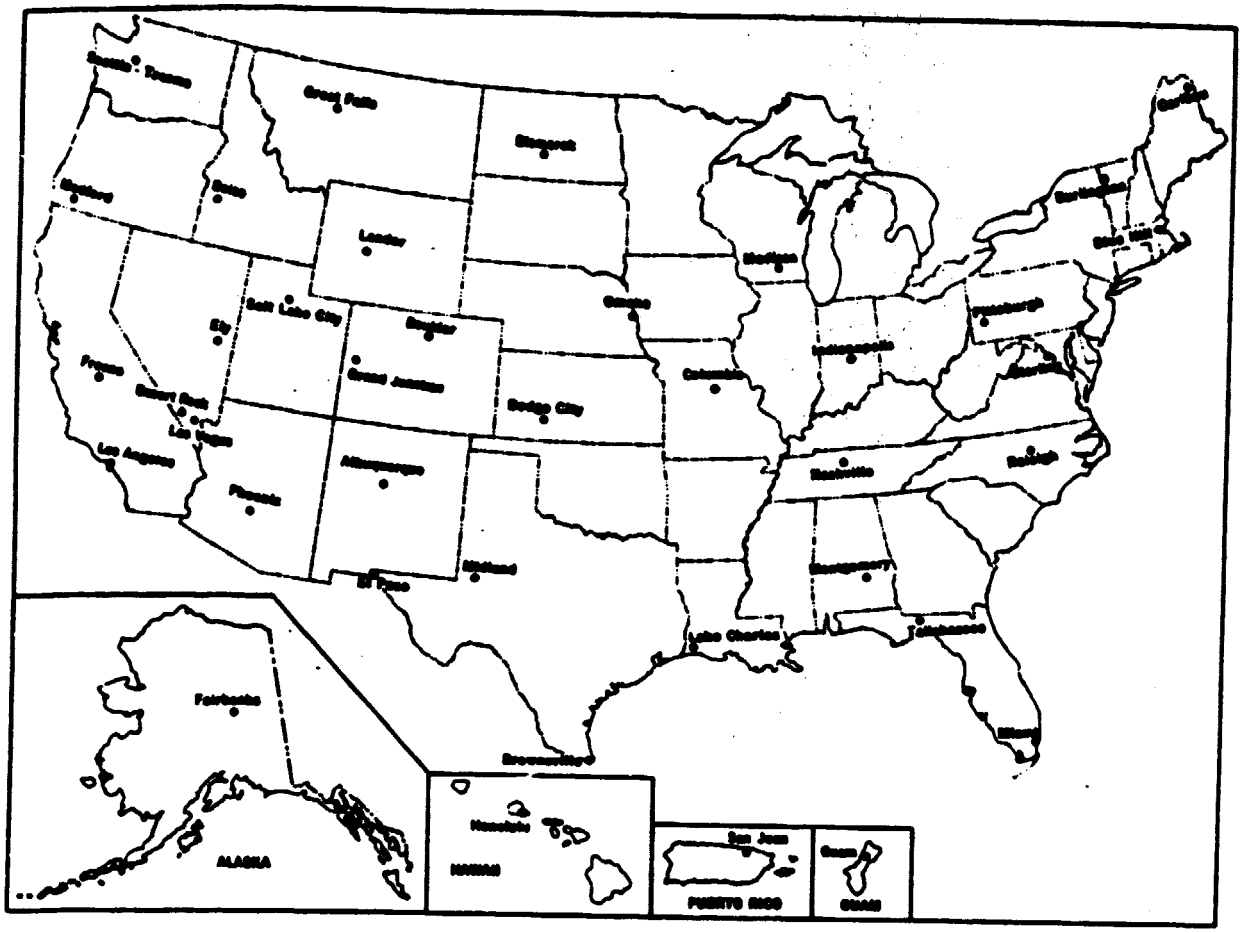


FIGURE 1 NOAA SOLAR RADIATION NETWORK

TABLE 1

SOLMET STATIONS USED IN STUDY

STATION #	NAME	LAT	LON	ELEV
03937	Lake Charles, LA	30.07N	93.13W	19
03945	Columbia, MO	38.49N	92.13W	277
12839	Miami, FLA	25.49N	80.17W	8
12919	Brownsville, TX	25.54N	97.26W	12
13722	Raleigh, NC	35.52N	78.47W	137
13895	Montgomery, ALA	32.18N	86.24W	68
13897	Nashville, TENN	36.07N	86.41W	186
13985	Dodge City, KAN	37.46N	99.58W	795
14607	Caribou, ME	46.52N	68.01W	195
14742	Burlington, VT	44.28N	73.09W	112
14837	Madison, WIS	43.08N	89.20W	271
23023	Midland, TX	31.57N	102.11W	872
24011	Bismarck, ND	46.46N	100.46W	511
24143	Great Falls, MON	47.29N	111.22W	1118
26411	Fairbanks, Alaska	64.49N	147.52W	143
93734	Sterling, VA	38.59N	77.28W	87
93805	Tallahassee, FL	30.23N	84.22W	18
93819	Indianapolis, IND	39.44N	86.16W	244
94823	Pittsburg, PA	40.30N	80.13W	371
94918	Omaha, NEB	41.22N	96.01W	404

TABLE 2

TYPES OF SOLMET DATA USED IN STUDY

WBAN STATION NUMBER
YEAR
MONTH
DAY
SOLAR HOUR
SOLAR MINUTE
ESTRATERRESTIAL RADIATION
DIRECT RADIATION & FLAG
INDIRECT RADIATION & FLAG
OBSERVED TOTAL INSOLATION AND FLAG
STATION PRESSURE
STATION DEW POINT
SKY CONDITION (clear, scattered, broken, overcast, or obscured)
TOTAL SKY COVER
LOWEST CLOUD LAYER AMOUNT AND TYPE*
SECOND CLOUD LAYER AMOUNT AND TYPE*
THIRD CLOUD LAYER AMOUNT AND TYPE*
FOURTH CLOUD LAYER AMOUNT AND TYPE*
TOTAL OPAQUE SKY COVER

---

*Given every three hours, based on 16 generic cloud types or obscuring phenomena.

at the high resolution would be used only as a means of designing and checking cloud specifications, and may be available only once per day. For regular analysis it is more likely that AVHRR data that are stored at reduced resolution (Global Area Coverage), with a subpoint resolution of 4 km for both the infrared and available visible data, would be used. The reduced resolution, with only a third of the original scanlines used, consists of 409 spots in each scanline. After identification of AVHRR pixels within some specified grid box, thresholding and a histogram analysis could be performed for cloud amount, brightness, and temperature. The data may be analyzed interactively from multichannel imagery.

HIRS data, with an optimum spatial resolution of about 17.5 km, represent a potential source of data for routine analysis of cloudiness within the general areas of interest. Each scan line contains 56 scanspots of data for each channel. Every 40th scanline there usually is a 3-line data gap for calibration. The HIRS instrument includes 19 infrared radiometric channels spaced from the near-infrared window region around 3.7 um through the central portion of the 15 um band of carbon dioxide. A single short wavelength channel is centered around 0.7 micrometers. This channel is available, under sufficient sunlight, as a general source of brightness data to supplement daytime inferences made from the infrared data. However, it is likely that AVHRR data would still be required for the initial development of cloud amount specifications. If HIRS and AVHRR data are applied, say, to a 2.5 x 2.5 degree grid box with a subgrid of 0.5 degrees (25 subdivisions), then several HIRS datapoints can be averaged for each subdivision. Within each subdivision, it would be possible to obtain statistics of AVHRR data (over 150 points archived at 4-km resolution).

AVHRR and HIRS data available on digital tape include location and calibration information. However, limb correction algorithms must be applied for standardization prior to analysis (spectral variations in infrared slant path attenuation with zenith angle are thereby accounted for). HIRS data swaths from successive orbits begin to overlap poleward of about 35 degrees latitude. Equatorward of that latitude it is possible to have data gaps to the extent that a particular spot may not be covered by the orbital swaths for that mode (ascending or descending) on that day. Also, on successive days, the satellite subtrack shifts about 8 degrees of longitude relative to a fixed equator crossing.

With an additional reduction in spatial coverage, standard TOVS products are routinely available for specified retrieval boxes. For each box these products include clear-column HIRS radiances, retrieved temperatures and retrieved mixing ratios appropriate to specified pressure levels. The total precipitable water may be inferred from the mixing ratios.

4.0 APPROACH

4.1 Planned Analyses and Tests

4.1.1 General Constraints

The approach adopted for the use of data obtainable operationally from the polar-orbiting NOAA satellites has evolved on the premise that the preferred repetitive high-resolution visible reflectance data will not be available at optimum times during the day. This limitation of the available polar-orbiter data does not alter the major objective of any successful approach, namely, to account for the impact of cloud attenuation on the daily insolation. Toward that end it is most desirable to deal with a physical model that accounts for the radiative processes. A rigorous radiative transfer treatment would not be acceptable on operational grounds, both from the point of view of available information and computational demands. Greatly simplified physical models can be introduced as long as adequate parameterizations are incorporated. Components of the model should be "observable" from the satellite platform. However, when reflectances themselves are not incorporated in the model, then the link between the remote-sensing data and the total transmittance of solar radiation to the lower boundary surface must be achieved indirectly through more parameterization.

4.1.2 Regression Analysis, Dependent Data

The difficulty of satisfactorily accounting for atmospheric transmittance from one-sided measurements dictated the initiation of a regression approach, using actual observations near the surface. Initially, the independent variables in the regression approach are selected from surface data, but are chosen to represent information that can be inferred from polar-orbiting satellite data. The regression expression, on the other hand, is selected so as to represent (in a very generalized way) the physical processes. Success or failure of the regression approach depends on the suitability of the available data and on the suitability of the physical representativeness of the regression expression.

During the design stage of the regression analysis (with dependent data) observations of total daily insolation, along with computed insolation incident on a horizontal surface at the top of the atmosphere, are used to describe the dependent variable. Cloud observations (amount and type) and surface dew point temperatures are used to describe independent variables indicative of

transmittance through cloudy and cloud-free atmospheres (with variable water vapor). Experience with these data types also provides the information needed for improving the model or for altering the dataset when possible. The program starts with examination of hourly data, since both the dependent and independent data sources are available at that frequency. Analysis of hourly data also makes it easier to assess the role of changes in the solar zenith angle. Next, the total daily insolation is examined in relation to specific hourly sources of independent data, since the data will be acquired only at several specific hours (from polar-orbiting satellites).

4.1.3 Independent Data Tests

Once the regression analyses with the dependent data are established, expressions must be tested. One type of test is to make use of the SOLMET data themselves for periods not previously included in the design analysis with dependent data. For that program, two months of data during the agriculturally significant spring period were reserved for testing.

After progressing to the stage of incorporation of operational satellite data as the source of independent data, the first tests are in terms of an interactive analysis of simple satellite imagery, primarily from single channel infrared radiometric measurements, as supplemented during the day by visible reflectance measurements. These measurements are typical of operational scanners aboard the polar-orbiting satellites, although the high resolution scanners on the most recent polar-orbiters are not limited to a single visible and a single infrared channel. Nevertheless, to simulate polar-orbiter coverage, satellite images were selected from full-disc infrared GOES data, along with occasional visible imagery, to represent imagery that might be available to the operational analyst for interactive scrutiny. Generally, only three images were selected to represent the early morning-early evening passes and the mid-afternoon pass of the NOAA series polar orbiters. For this test, cloud parameters had to be estimated for local areas immediately around specific pyranometer stations, while the moisture parameter was estimated from NMC hemispheric (6-hour) reports on surface charts.

In addition to the interactive test of the method by using satellite imagery, the operational sounder data are examined as the source for estimating cloud parameters as input to the insolation model. In particular, seven of the

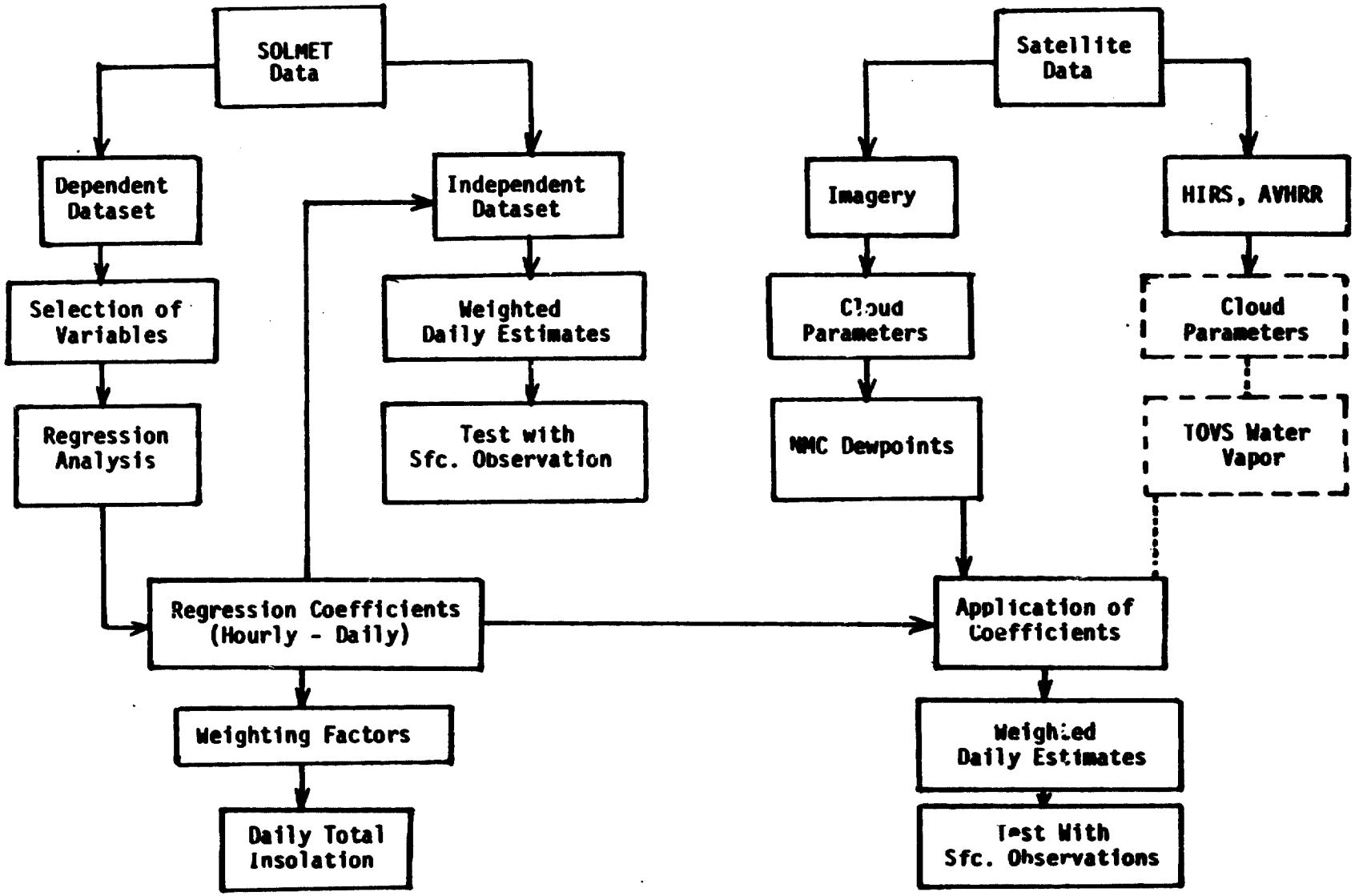


FIGURE 2 BLOCK DIAGRAM OF STUDY PROCEDURE

infrared HIRS channels were introduced into a multiple discriminant analysis program designed to estimate cloud type. The cloud type is used to specify the cloud transmittance. (When ready for operational application, the cloud specification and the precipitable water derived from the sounder data may be tested in the regression expression). The organization of the analysis procedures is outlined in Figure 2.

4.2 Model

4.2.1 Basic Expression

On physical grounds it was desired to model a linear expression since multiple linear regression was to be used to define the algorithm. At any given time the total direct-plus-diffuse solar radiation F_s incident at the surface can be expressed in terms of solar flux F_o incident on a horizontal surface at the top of the atmosphere and an "equivalent" total flux transmittance through the cloud-free or cloudy atmosphere. Independent variables are introduced through the transmittance, but precaution is taken to avoid an excess of variables that are not independent of each other. If τ_a represents cloud-free atmospheric transmittance and τ_c represents the cloudy atmospheric transmittance, then

$$F_s = F_o \ \tau_a \ \tau_c = I_o (\bar{R}/R)^2 \cos Z \ \tau_a \ \tau_c \qquad (1)$$

where I_o is the solar constant (1377 joules m^{-2} s^{-1}), $\bar{R}$ is the mean earth-sun distance, and Z is the solar zenith angle. The indicated cos Z accounts only for the horizontal surface at the top of the atmosphere; slant paths involving the solar zenith angle are implicit in the transmittance expressions. Each of the transmittance terms could be factored further to separate, say, molecular scattering or absorption from aerosol scattering or absorption.

One method for retaining an expression for linear regression with terms that can be associated with physical processes is to consider the negative of the natural logarithm of Eq. (1)

$$-\ln (F_s/F_o) = -\ln(\tau_a) - \ln(\tau_c) \qquad (2)$$

Thus, by dealing with the log of the total transmittance (F_s/F_o), a linear expression results in terms of the logarithms of individual atmospheric transmittance terms. To the extent that the transmittance terms can be represented by exponentials, the right-hand side of Eq. (2) represents the optical thickness.

4.2.2 Variables and Parameterizations

(a) Molecular Atmosphere: To associate the attenuation with the regression coefficients, the cloud-free transmittance τ_a may be approximated by the product of two exponentials e^{-a_0} and $e^{-a_1 U}$. The path length U is postulated as a pressure-scaled slant path through the total atmospheric precipitable water, as estimated from the surface vapor pressure e_s:

$$U = (b_0 + b_1 e_s)(p_s/p_r) M \tag{3}$$

where the constants b_0 and b_1 are assigned for an average representation of the total precipitable water (cf. Smith[12]) but could be expanded to allow for seasonal latitudinal variations in the relationship to the surface dew point or, in this case, the vapor pressure. The ratio of surface pressure to a reference pressure, as a scaling factor, effectively accounts as well for large changes due to station elevation. The slant path magnification factor M replaces the sec Z so as to avoid problems with the plane-parallel assumption at very low scan angles. Following Rogers, the relationship is expressed as

$$M = k \left[(k^2 - 1) \cos^2 Z + 1\right]^{-\frac{1}{2}} \tag{4}$$

with k = 35.

(b) Clouds: In order to maintain a simple initial cloud classification, each of the surface observations of cloudiness is placed into one of seven total categories: CLEAR, HIGH, MID, FOG, LOW, MIXED and THICK. The categorical ranking is in terms of an assumed increasing attenuation (decreasing cloud transmittance). The MID category actually includes high cloud also, if they are present with the middle clouds. The initial classification of FOG or OBSCURATION, LOW and MIXED were subsequently altered on the basis of transmittances inferred from the SOLMET data and the initial regression model. The MIXED category initially included FOG or LOW CLOUDS with MID and/or

HIGH clouds. Subsequently, most of the MIXED clouds were reclassified according to the dominant layer determining the cloud ceiling (from below), with a subsequent increase in the transmittance for the remaining MIXED category. Also, in the THICK category (including nimbostratus and cumulonimbus), the cumulonimbus category was reclassified as low for scattered amounts, since their effects may have been minimal over the observing station. Obscuration conditions were added to the THICK category, whereas LOW subsequently was separated into cumulus (CU) and stratiform (ST).

To consider the cloudy transmittance in more detail, the insolation may be described by

$$F_s = F_o \left[(1-C)\, \tau_a + C\, \tau_a^{*} \psi\right] \tag{5}$$

where C is the cloud amount and ψ is the cloud transmittance. For use with SOLMET data the cloud amount is described as the observed opaque cloud amount plus an additional 35 percent of the difference between the total cloud amount and the opaque amount. The transmittance τ_a^{*} outside the cloud differs somewhat from the cloud-free transmittance τ_a, largely as a result of scattering. However, the difference can be absorbed as a factor of ψ, perhaps by letting ψ become ψ^{*}. Then Eq. (5) becomes

$$F_s/F_o = \tau_a \left[1 - C(1 - \psi^{*})\right] \tag{6}$$

The bracket term, raised to some power (close to unity) to accomodate the approximation and as a convenience for regression, defines the cloudy transmittance τ_c and

$$F_s/F_o = \tau_a\, \tau_c = a \left[1 - C(1 - \psi^{*})\right]^{a2} \tag{7}$$

The assumed form for ψ is a linear function of the cosine of the solar zenith angle with a transmittance factor, Y, to represent the transmittance for a particular cloud type at a particular solar zenith angle:

$$\psi = Y(d_o + d_1 \cos Z) \tag{8}$$

Actually, the overall representation of the total transmittance over a finite time period need not necessarily be restricted to a simple cloud "type" with a given amount, as implied by Eq. (7). In fact, the cloudy transmittance term could be a product term of more than one type. Nevertheless, for this study it is assumed that the cloud classification and the transmittance relationship can be applied to a time period of one hour. Therefore, the cos Z that appears in the modeled expressions is considered to be represented by the average cos Z for the given hour.

(c) Daily Averages: When considering the total daily insolation, it becomes necessary to consider the average daily cos Z, which can be expressed in terms of the half-day hour angle H as

$$\overline{\cos Z} = \sin \phi \sin \delta + \cos \phi \cos \delta \sin H/H \qquad (9)$$

where ϕ is the latitude and δ the declination. (The angular velocity of the earth $\omega = dh/dt = 2$ rad day^{-1}). The maximum solar zenith angle is given by the difference ($\phi - \delta$), whereas the hour angle to either sunrise or sunset is given by cos H = -tan ϕ tan δ. If, during the evaluation of Eq. (7) over a period of a day, the numerator and denominator of the left-hand side are summed or integrated separately, then a restriction is implicitly imposed on the average daily transmittance corresponding to the observed-computed left-hand side. In particular

$$\int F_s \, dt / \int F_o \, dt = \hat{\tau}_a \, \hat{\tau}_c$$

where

$$\hat{\tau}_a \, \hat{\tau}_c = \int \tau_a \, \tau_c \, F_o \, dt / \int F_o \, dt$$

$$= \frac{1}{\Delta t} \int_{t}^{t+\Delta t} \tau_a \, \tau_c \, \frac{\cos Z}{\overline{\cos Z}} \, dt \qquad (10)$$

Instead of introducing the weighting factor indicated in Eq. (10) each of the averaged transmittances was approximated by assignment of a daily average cos Z in place of the cos Z indicated in their definitions (Eqs. 4 and 8).

4.2.3 Regression Coefficients and Weighting Factors

With the parameterizations described, the regression expression for any hour becomes

$$-\ln (F_s/F_o) = a_o + a_1 U - a_2 \ln \tau_c \qquad (11)$$

whereas for the daily total

$$-\ln (\bar{F}_s/\bar{F}_o) = a_o + a_1 \hat{U} - a_2 \ln \tau_c \qquad (12)$$

where the caret indicates that the daily average $\overline{\cos Z}$ has been used instead of the instantaneous cos Z. In terms of the desired insolation itself, the hourly or daily insolation

$$F_s = F_o \exp \left\{ -(a_o + a_1 U - a_2 \ln \tau_c) \right\} \qquad (13)$$

where the coefficients have been determined by linear regression in terms of $-\ln (F_s/F_o)$.

When observations from several different hours (k_i) are to be combined in a single estimate of the daily total insolation, then additional regression coefficients are introduced in the analysis for each direct transmittance

$$-\ln (\bar{F}_s/\bar{F}_o) = a_o + a_1 \bar{U} + a_2 \ln \hat{\tau}_c (k_1) + a_3 \ln \hat{\tau}_c (k_2) + a_4 \ln \hat{\tau}_c (k_3) \qquad (14)$$

where only a single average U term is used for the day, but separate terms are maintained for the cloud transmittances because it is not likely that they should all be weighted equally. Proper relative weighting is handled by the regression coefficients.

An alternative procedure to that represented by Eq. (14) for prespecified hours is to form a weighted average of separate daily insolation estimates (one for each observation time) based on information obtained at different hours, characterized by polar-orbiter overpasses. If $F_s(k_1)$ represents the estimated daily insolation based on data for the hour k_1,

$$\bar{F}_s(k_1) = \bar{F}_o \exp - \left\{ a_o + a_1 \hat{U}(k_1) - a_2 \ln \hat{\tau}(k_1) \right\},$$

Then the final daily estimate might be based on three separate initial estimates

$$\bar{F}_s = \frac{W_1 \bar{F}_s(k_1) + W_2 \bar{F}_s(k_2) + W_3 \bar{F}_s(k_3)}{W_1 + W_2 + W_3} \tag{15}$$

where the W's represent weighting factors. More generally, if E is used to define the estimated daily insolation and the subscript k identifies predictor hour,

$$E = \sum_{k=1}^{N} (W_k E_k) / \Sigma W_k \tag{16}$$

where the summation is over the number N of daily estimates based on individual hourly data. The weighting factor W_k will increase to a maximum for the hour with the daily maximum cosine of the solar zenith angle, cos Z_{MX}. This also represents the time with the maximum correlation between the estimated E_k and the observed F_s, at least for the dependent data used in the regression analysis. Therefore, as a first approximation it is possible to use the correlation coefficients derived from the regression analysis as the weighting factor. However, the general form that was adopted is expressed in terms of the solar zenith angle for that hour, cos Z_k, as well as the maximum solar zenith angle

$$W_k = A \cos Z_{MX} + \cos Z_k \tag{17}$$

where the constant A generally exceeds unity.

5.0 RESULTS

5.1 Observed Hourly-Daily Insolation Correlations

The SOLMET data files, with hourly insolation observations, are ideal for the determination of the correlation between an hourly total and the total daily insolation, the variable of most interest. Such correlations, apart from the obvious contribution arising from the fact that the hourly total is also part of the daily total, are especially informative about the relative significance of observation times to be used in the daily estimation. Only fully sunlit hours are included in the correlations, which have been determined from data for 20 different stations during each seasonal period. Biased cloud distributions are about the only factor that could result in an asymmetry of the insolation distribution relative to solar noon. All of the results are summarized in Figure 3. Some slight skewness toward the afternoon hours is apparent. Otherwise, the correlation curves are quite predictable, with the most peaked distribution for the shorter winter days.

It is apparent that an observation near solar noon carries the most information about the daily total. Obviously, a satellite observation near noon or early afternoon is most useful. A number of hours close to noon, say within three hours, also are highly correlated with the daily total. The correlation drops off near sunrise and sunset, even if only full sunlit hours are considered.

Any individual station skewness in distribution is masked in Figure 3 because each curve is based on data from 20 scattered stations. Seasonal variability in distribution appears largest in the afternoon. The correlations in Figure 3 may be interpreted as limiting conditions for estimation since they represent what would result if perfect hourly estimates were available (for any hour). To a first approximation, the square of the correlations near sunrise and sunset also suggest the magnitude of the explained variance that might be associated with some skill. Thus, as will be seen later, regression results near midnight are associated with reduced explained variance that show little value relative to the estimation of total daily insolation.

5.2 Preliminary Computations

5.2.1 Hourly-Hourly Regressions

Regression coefficients can be established between hourly insolation and hourly estimates based on cloud and water vapor slant path

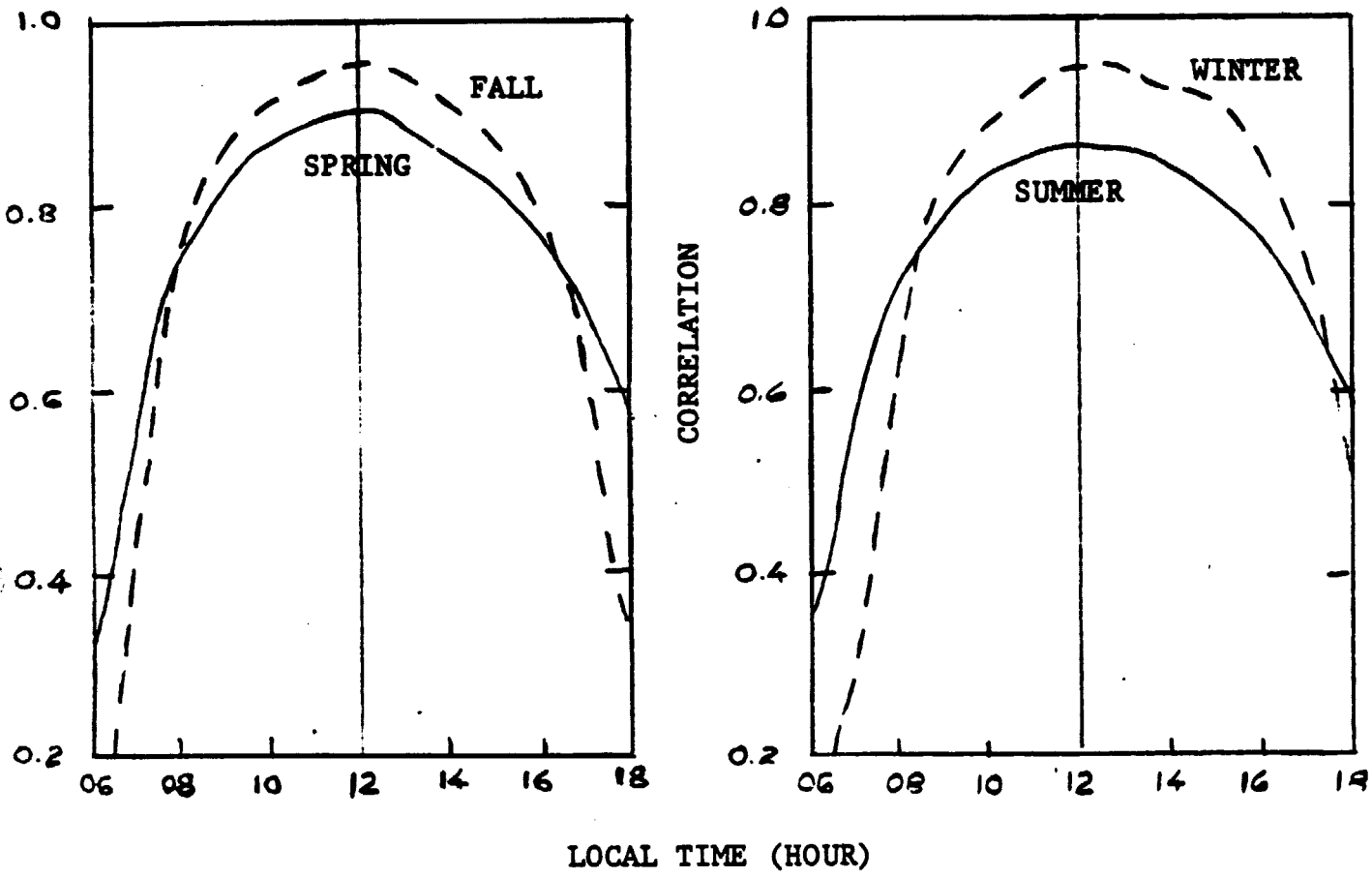


FIGURE 3 DIURNAL VARIATION OF CORRELATION BETWEEN HOURLY AND DAILY TOTAL INSOLATION FOR TWENTY STATIONS

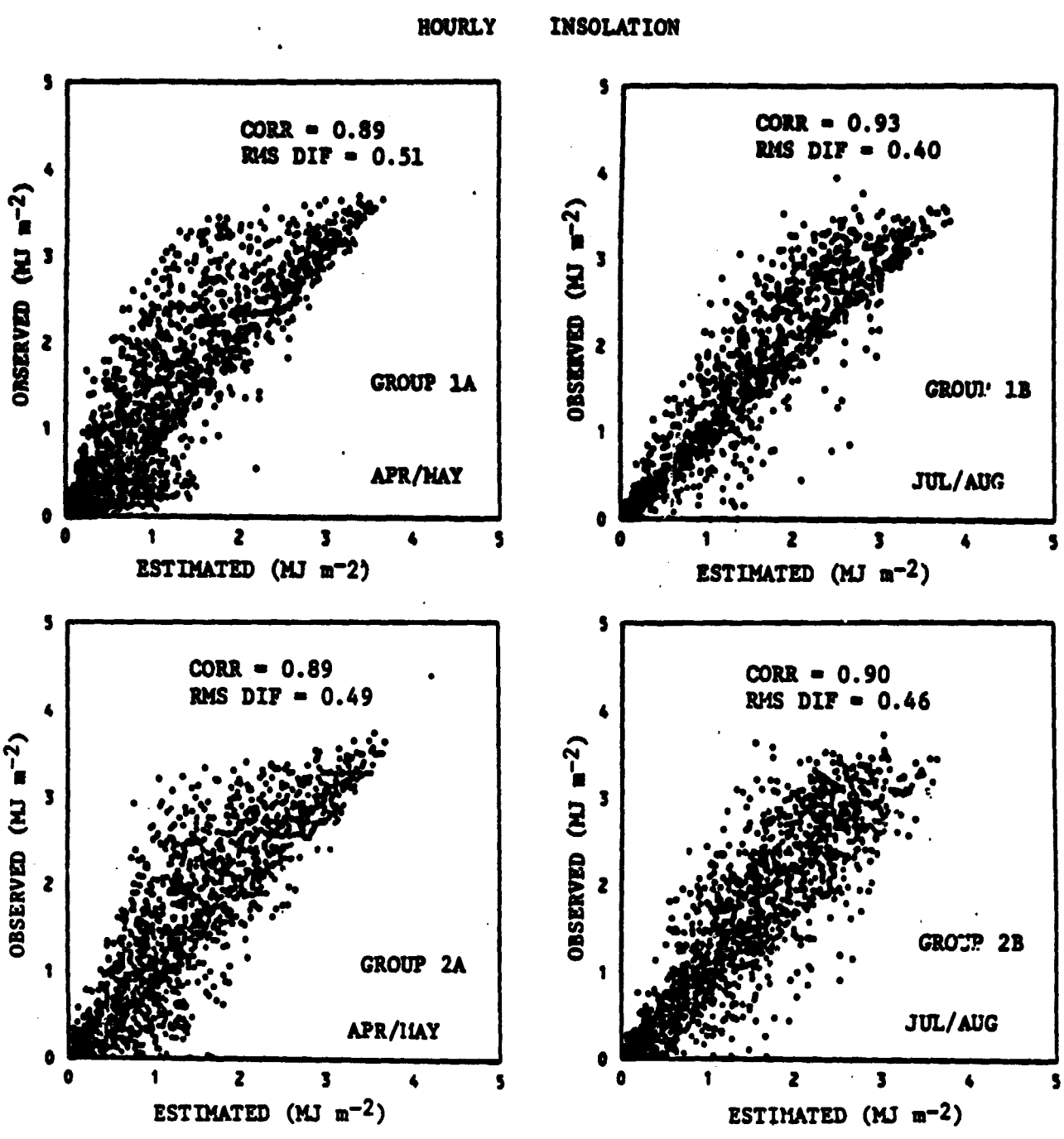


FIGURE 4 OBSERVED AND ESTIMATED HOURLY INSOLATION
Group 1 = Bismarck, Dodge City, Madison
Group 2 = Montgomery, Lake Charles, Nashville

TABLE 3

REGRESSION COEFFICIENTS FOR PRELIMINARY HOURLY INSOLATION ESTIMATES

(Standard errors in parentheses)

GROUP	INTERCEPT (a_0)	WATER VAPOR COEFFICIENTS (a_1)	CLOUD TRANSFER COEFFICIENTS (a_2)
1A	.1765 (.0321)	.0311 (.0051)	1.2774 (.0338)
1B	.1059 (.0256)	.0365 (.0032)	.9771 (.0319)
2A	.1856 (.0283)	.0272 (.0030)	1.1240 (.0317)
2B	.1278 (.0242)	.0278 (.0016)	.8866 (.0251)

1: Bismarck, Dodge City, Madison
A = APR, MAY B = JUL, AUG

2: Montgomery, Lake Charles, Nashville
A = APR, MAY B = JUL, AUG

parameters. Figure 4 illustrates the initial sets of results[14] derived from two separate groups of 3 stations each during spring and summer months. (Group 1: Bismarck, Dodge City, Madison; Group 2: Montgomery, Lake Charles, Nashville). It is apparent from these results (obtained without any fine tuning of initial parameters) that there is little geographical dependence, but seasonal differences do appear. Table 3 shows the corresponding regression coefficients and their standard errors. All of the coefficients are significant; magnitudes are representative of those for other tests. Coefficients for the cloudy transmittances, as well as intercepts, are larger in the spring than the summer. It was anticipated that the cloudy atmosphere transmittance coefficients would be close to unity.

The results in Figure 4 show that the summer estimates and observations are somewhat better correlated, with less bias, than in spring. The bias in the spring data shows overestimates for lowest magnitudes and underestimates for higher magnitudes. Part of the failure to successfully estimate the extremes results merely from the statistical nature of the coefficients. Some of the low-magnitude overestimation also could result from underestimates of cloud absorption in certain clouds, whereas in other circumstances the cloud transmittance is underestimated. Improper specifications of the physical processes affecting the linear regression model are likely sources of nonlinearity. On the other hand, even with the proper regression expression in handling the logarithm of the insolation, some distortion is introduced in the average of the insolation estimates that are based on the least-squares coefficients for the logarithm of insolation.

5.2.2 Hourly-Daily Estimates

If hourly estimates of insolation are obtained throughout the day, they may be summed to provide estimates of the daily total insolation. Figure 5 presents such sample results for one group of stations shown in Figure 4, for both spring and summer. Presumably, such estimates are superior to those daily estimates that do not include estimates for every hour. As can be seen in Figure 5, there are far fewer points (only one per day) than in hourly-hourly comparisons. Furthermore, the averaging over the day results in a degradation of significance of the U coefficient, which is dependent on the diurnal variations in the slant path (the average daily zenith angle does not vary significantly within the data sample). As before, the spring estimates are somewhat more biased than the summer estimates.

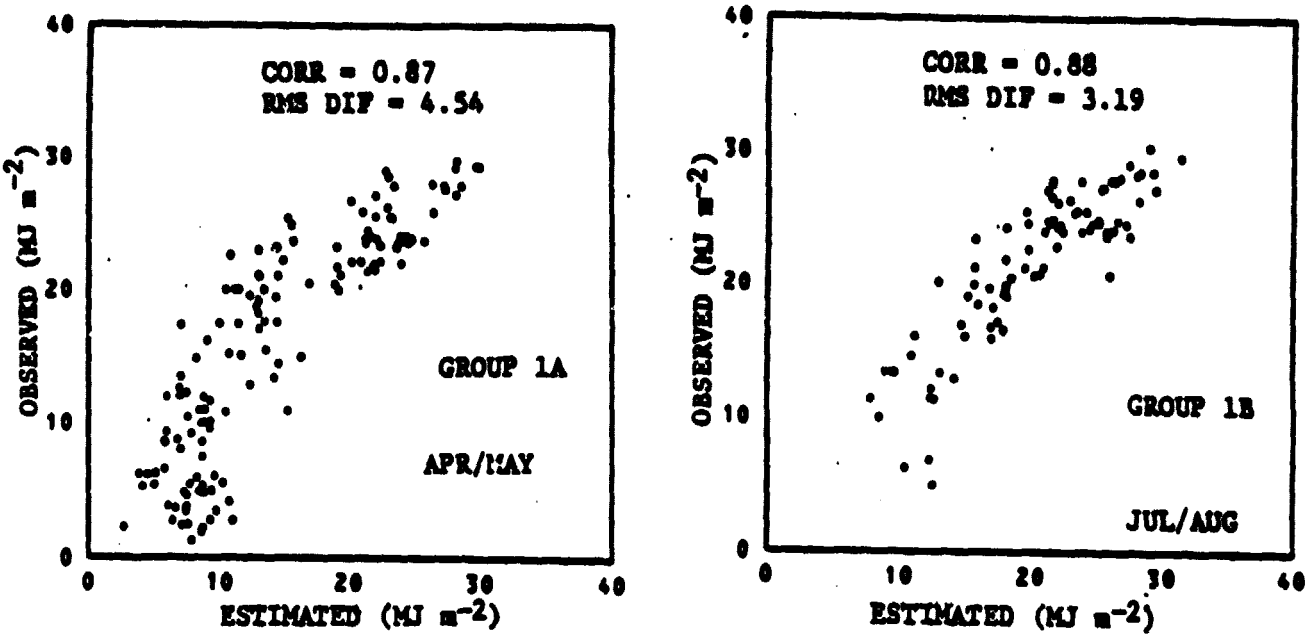


FIGURE 5 DAILY INSOLATION, OBSERVED AND ESTIMATED FROM HOURLY SUMMATIONS

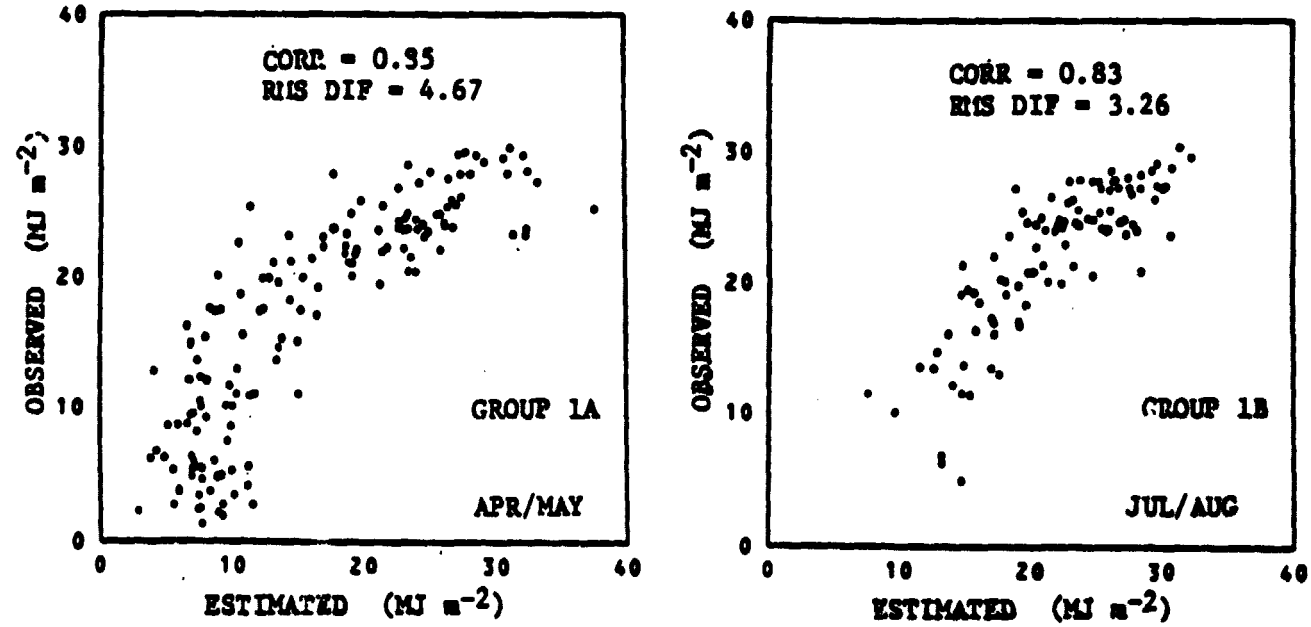


FIGURE 6 DAILY INSOLATION, OBSERVED AND ESTIMATED FROM REGRESSION FOR COMBINATION OF PARAMETERS AT 0730 AND 1500 LOCAL TIME

If estimates are not available for each hour, so that the hourly estimates can be summed to obtain the daily total, the cloud and moisture observations at any hour can be used to estimate the total daily insolation on the basis of hourly-daily regression (see Eq. 12). Or, if suitable weighting coefficients were available, individual hourly estimates could be weighted to provide a daily estimate (see Eq. 16). Another means for obtaining daily estimates from hourly data is to use regression coefficients that have been designed to link parameters from two or more specific hours directly to the daily insolation (see Eq. 14). For the first test of this latter procedure, water vapor and cloud parameters for 0730 and 1500, approximating NOAA-6 and TIROS-N daytime passes, were used to specify the daily insolation. Regression coefficients for the cloud transmittance terms are dominant but smaller than those associated with a single hour. Results of the preliminary computations corresponding to the data in Figure 5 are illustrated in Figure 6. Despite a slight reduction in the correlation and the slight increase in residual differences the overall distributions of estimates and observations are similar to those for the hourly summations. This result was encouraging, since the polar orbiters do not provide the frequency of observation or the optimum times of observation that are included in hourly summations. However, it does appear that the bias in spring is more pronounced than for the hourly summations, although some overestimates of the daily insolation are now made at the high-magnitude end of the range.

5.2.3 Apparent Sensitivities

The preliminary computations indicated that the regression analysis approach was feasible, but that improvements in parameterization would be worthwhile. Regression coefficients for the precipitable water term appear to be significant only when significant variations in the average cosine of the solar zenith angle occur. Surface dewpoint probably is an adequate estimator of effective water vapor path. Clouds dominate the results (also related indirectly to surface dew point) and their parameterization is the most logical area for improvement.

Initial computations for selected U.S. stations east of 105W longitude and between 30N and 47N latitude failed to show any pronounced latitudinal or regional influence on the regression coefficients. On the otherhand, seasonal variations appear to be significant, both in terms of the magnitudes of the

regression coefficients and the joint distribution of estimations and observations. The distributions show some bias in the estimates, especially with departure from local noon.

The question of the dependency of results on individual stations was checked further by computing regression coefficients separately for each of 20 stations, and coefficients also for all 20 stations combined. Table 4 summarizes the most important regression coefficient (for the cloud transmittance term) for the most significant hour, 1200-1300 local time, appropriate to two 4-month periods, April-May-October-November and June-July-August-September. The average coefficient is largest in the spring/fall period. Only two of the stations show anomalies of twenty percent or more in the same sense for both seasons. Fairbanks has low coefficients and Sterling shows anomalously high coefficients.

5.3 Modifications

5.3.1 Regression Expression

With the dataset expanded to 20 stations and SOLMET data over a 2½ year period, checks were made on any improved correlation between estimation and observation as a result of alterations in the regression expression. Some improvement was found by the simple introduction of the cloud amount C as a factor in the logarithmic expression (Eq. 12):

$$-\ln (F_s/F_o) = a_0 + a_1 \hat{U} + a_2 C(-\ln \hat{\tau}_o) \qquad (18)$$

Eq. (18) was used in all subsequent regression analyses. The improvement, while not explained on physical grounds, may have been related to the nonhomogeneity of cloud conditions (i.e., departure from an idealized single cloud cover).

The terms with carets in Eq. (18) indicate that a single average cosine of the solar zenith angle is used to describe the path for the daily estimate. A comparison of results using $\overline{\cos Z}$, the daily average cosine, or $\cos Z_{MX}$, the maximum cosine, showed very little difference in results. Consequently, most of the final regressions that were performed used $\cos Z_{MX}$ instead of $\cos Z$. This change is justified in part by Eq. (10).

The impact of introducing cloud amount C in the last term of Eq. (18) is illustrated in Figure 7, which shows the distributions for estimates and observations both without (7(a)) and with (7(b)) the factor C. Only the results from

TABLE 4

SUMMARY OF DOMINANT INSOLATION-ESTIMATION REGRESSION COEFFICIENT COMPUTED FOR INDIVIDUAL STATIONS AT 1200-1300 LOCAL TIME

	CLOUD TRANSMITTANCE REGRESSION COEFFICIENT	
STATION	SPRING/FALL	SUMMER
Lake Charles	1.020	0.746
Columbia	1.056	0.804
Miami	0.992	0.626
Brownsville	0.800	0.694
Raleigh	1.174	0.680
Montgomery	1.080	0.586
Nashville	1.119	0.812
Dodge City	1.013	0.802
Caribou	0.873	1.074
Burlington	1.064	1.017
Madison	1.035	0.949
Midland	1.067	0.963
Bismarck	0.804	0.888
Great Falls	0.865	0.960
Fairbanks	0.499	0.629
Sterling	1.210	1.088
Tallahassee	0.806	0.598
Indianapolis	1.154	0.809
Pittsburg	1.057	0.818
JOINT COEFFICIENT	1.000	0.836

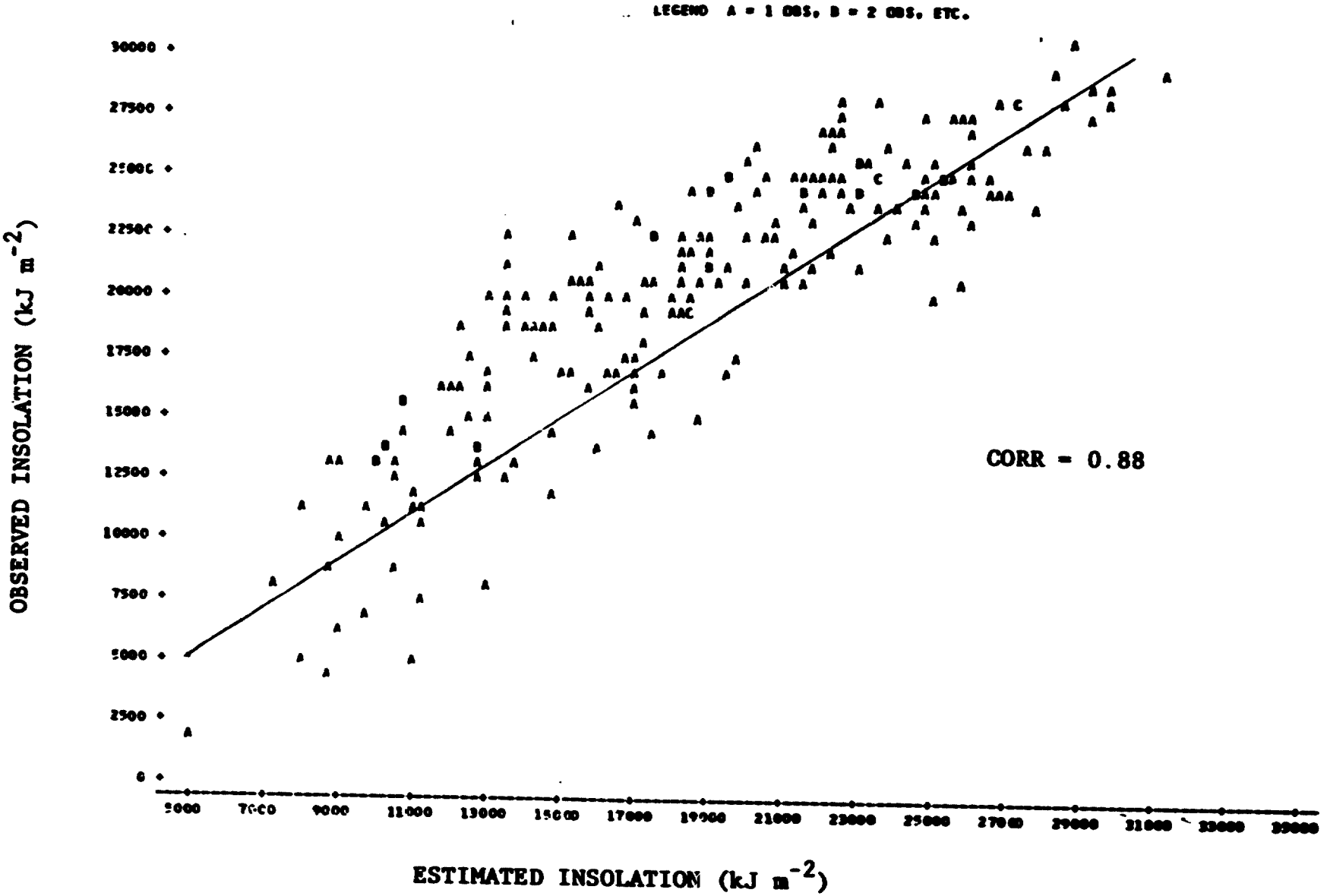


FIGURE 7 OBSERVED AND ESTIMATED DAILY INSOLATION WITH ESTIMATES BASED ON TWO TYPES OF CLOUD ATTENUATION REPRESENTATION
(a) ln τ_c

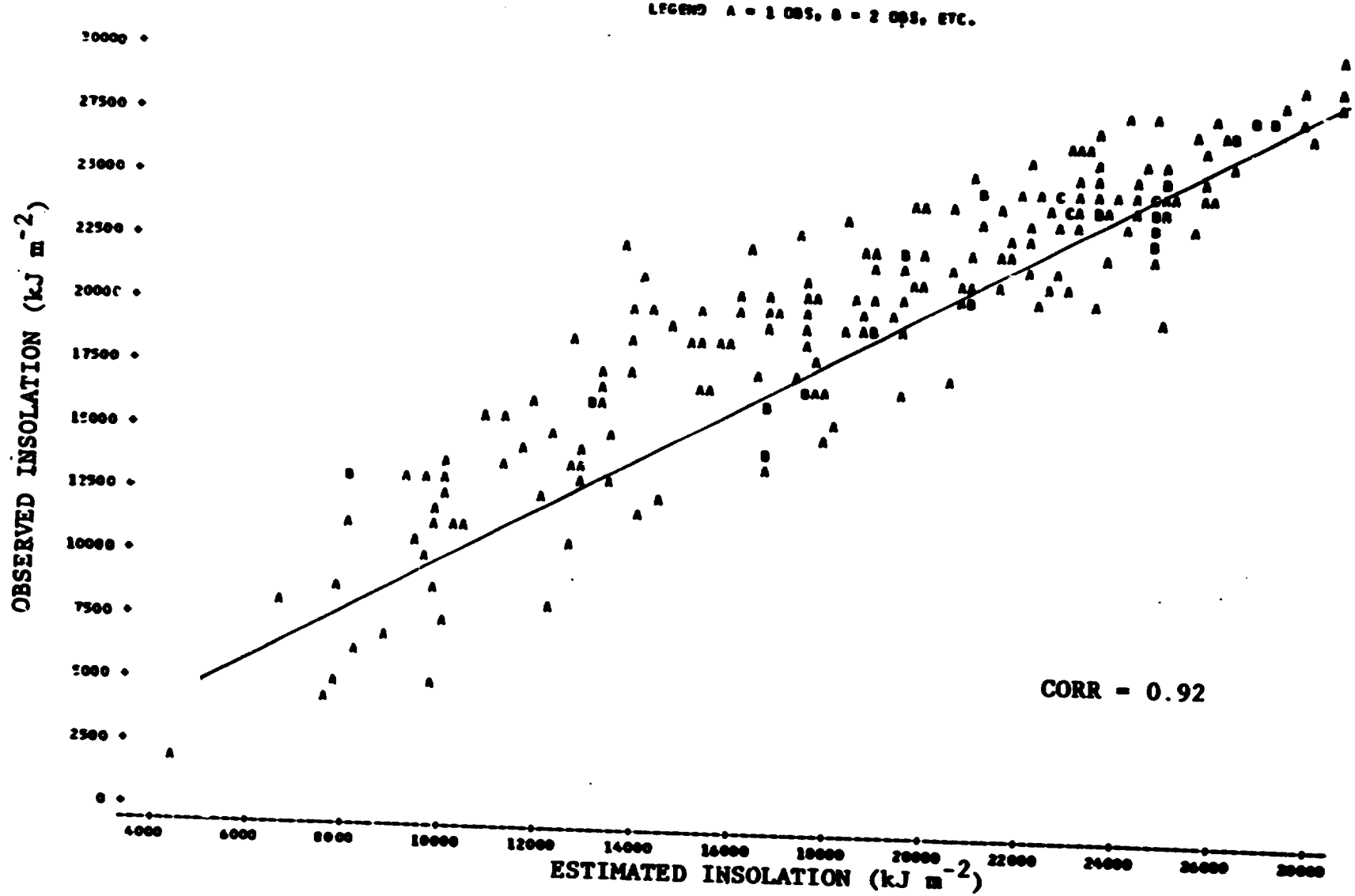


FIGURE 7 OBSERVED AND ESTIMATED DAILY INSOLATION WITH ESTIMATES BASED ON TWO TYPES OF CLOUD ATTENUATION REPRESENTATION
(b) c ln τ_c

the 6-station dataset for summer are shown. The correlation coefficient for this case is increased from 88 to 92 percent.

5.3.2 Cloud Transmittance Parameterization

For the earliest (preliminary) computations with a small dataset, regression computations in accordance with Eq. (12) used a given set of cloud categories, each with a given γ for computing the transmittance from Eq. (8) with d_0 and d_1 set at 0.6 and 0.7, respectively:

$$\psi = \gamma(0.6 \cos Z + 0.7) \quad .$$

The choice of d_0 and d_1 requires that γ is the transmittance for $\cos Z = 0.5$. Initially adopted values of γ are indicated in the first list of γ's in Table 5.

After the preliminary computations with the initial cloud categories and γ's, the estimates of insolation were examined by cloud category. On the basis of the bias, if any, observed in each category the γ's were altered accordingly. However the constants d_0 and d_1 (Eq. 8) also were altered so as to define γ as the cloud transmittance for an overhead sun ($\cos Z = 1.0$). Thus, the second set of transmittances that appear in Table 5 were altered both to compensate for observed biases and to normalize to an overhead sun. This second set of transmittances (along with $\psi = \gamma\ (0.7 \cos Z + 0.3)$) was used for the interactive data tests, using satellite imagery for cloud parameter specification.

Computations with the second set of cloud transmittances were separated into three categories: CLEAR ($\leq$0.1 cloud amount), CLOUDY ($\geq$0.9 cloud amount), and the REMAINDER. Most of the data fell in the latter category. Regression results for the latter category were similar to those obtained previously. Regression analyses for the CLOUDY category were surprising in that the explained variance was very low. As a result of the relatively poor job of fitting the CLOUDY cases, alterations were sought for improvement of the parameterization. To obtain a basis for a modified parameterization, the regression coefficients and the observations were introduced into the regression expression. Solution for the γ's was achieved by inverting the regression expression. At the same time, surface reports were used to divide the LOW cloud category into stratiform and cumuliform. Much of the MIXED cloud category was

placed either into one of the LOW categories or the MID category, according to the observed ceiling. The redefined γ's that appear as the final listing in Table 5 were based only on a single computational step rather than an interactive procedure. They represent effective empirical transmittances, dependent on the surface observations, that were designed to yield improved estimates of daily insolation from the model adopted.

Figure 8 illustrates several of the effective cloud transmittances (total direct and diffuse) as a function of the cosine of the solar zenith angle. For comparison, some of the theoretical computations by Liou[15] are included in Figure 8. The large discrepancy between transmittances for the cumulus category results to some extent from the fact that the observations are available only for partial cloud covers and include enhanced scattering from cloud walls.

In addition to the changes in cloud categories, some of the transmittances were increased by a significant amount. An objective analysis of the manner in which all of the cloud categories would be identified in the satellite data was not completed. Nevertheless, the impact of revised cloud categories of the SOLMET data was to improve the insolation estimates.

5.3.3 Final Adjustment

Inasmuch as the regression coefficients were determined in a least squares procedure applied to the logarithm of the insolation, it was considered likely that a statistical bias could result in average underestimates of the insolation itself (when applying the coefficients). A possible method to avoid such bias is to apply a nonlinear least squares regression analysis directly to the insolation form of the equation (see Eq. 13), in which the coefficients all appear within the argument of an exponential term. In this approach the attempt is to obtain coefficients that minimize the residual insolation standard deviation. Figure 9A shows the resulting distribution of estimated and observed daily insolation based on nonlinear regression with data for the best local hour, 1200-1300. For comparison, Figure 9B shows the distribution for estimates derived after the usual linear regression on the logarithm of the insolation. It appears that the nonlinear technique did not result in a superior estimate (slightly worse in this case). Thus it is likely that any residual estimation bias, if not from physically inadequate parameterization, might arise from the statistical association of data for any one hour with that for the entire day. For example, if three separate hours are involved in separate estimations of the daily insolation and all show cloudy

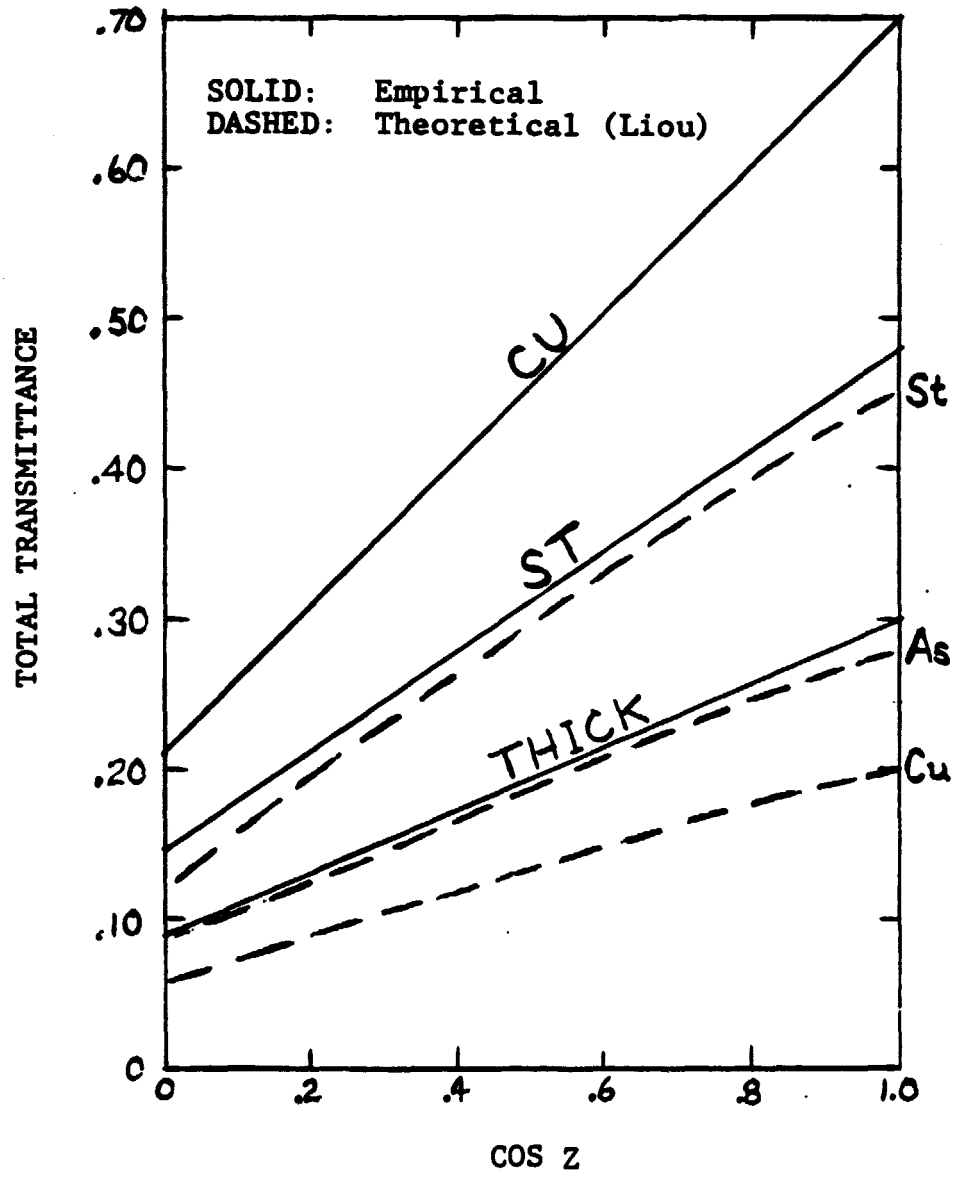


FIGURE 8 VARIATION OF EFFECTIVE CLOUDY-SKY TRANSMITTANCE WITH COSINE OF SOLAR ZENITH ANGLE

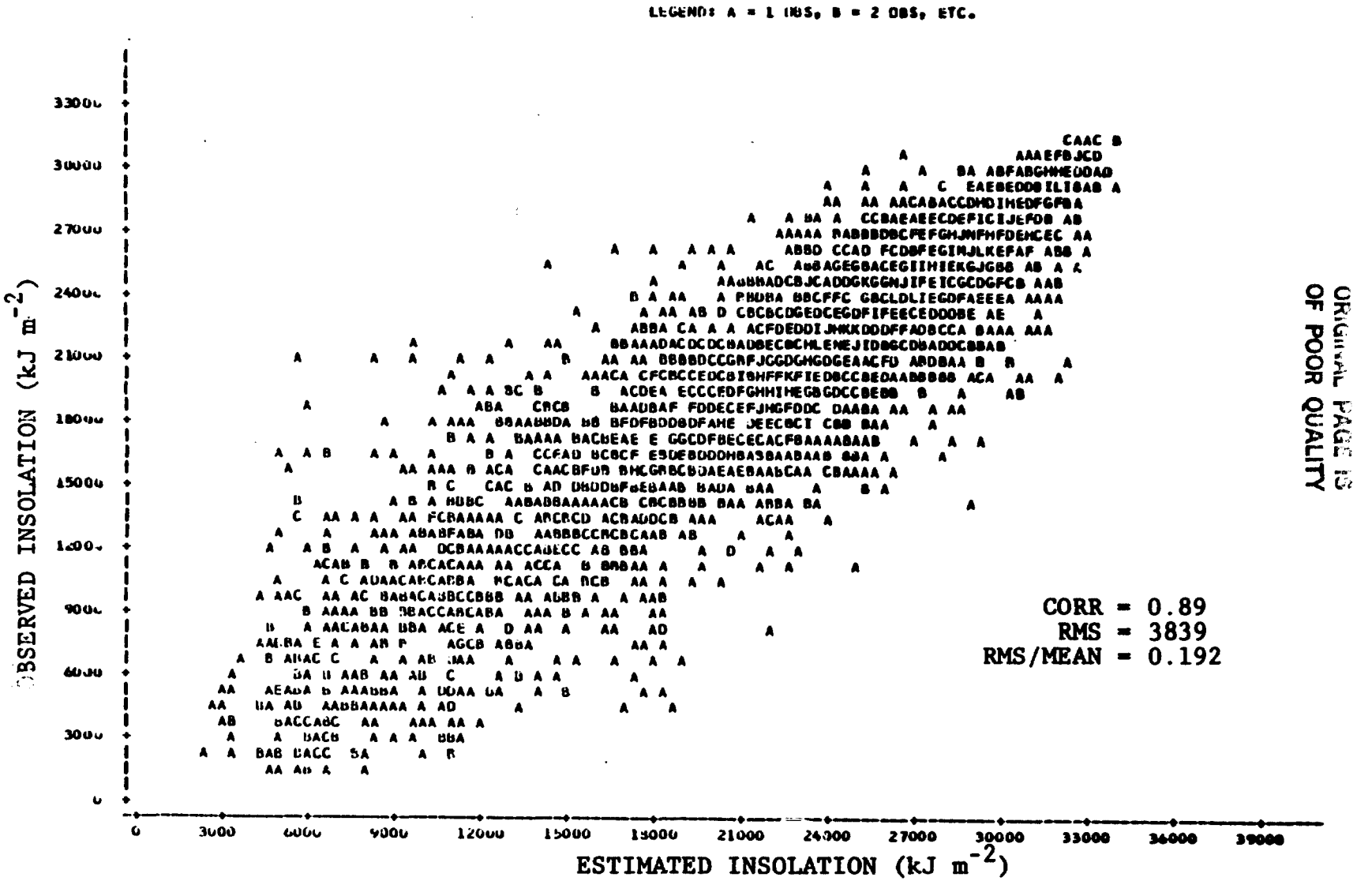


FIGURE 9 OBSERVED AND ESTIMATED DAILY INSOLATION, SUMMER 1200-1300 LST
(a) Nonlinear Regression of Transmittances

ORIGINAL PAGE IS
OF POOR QUALITY

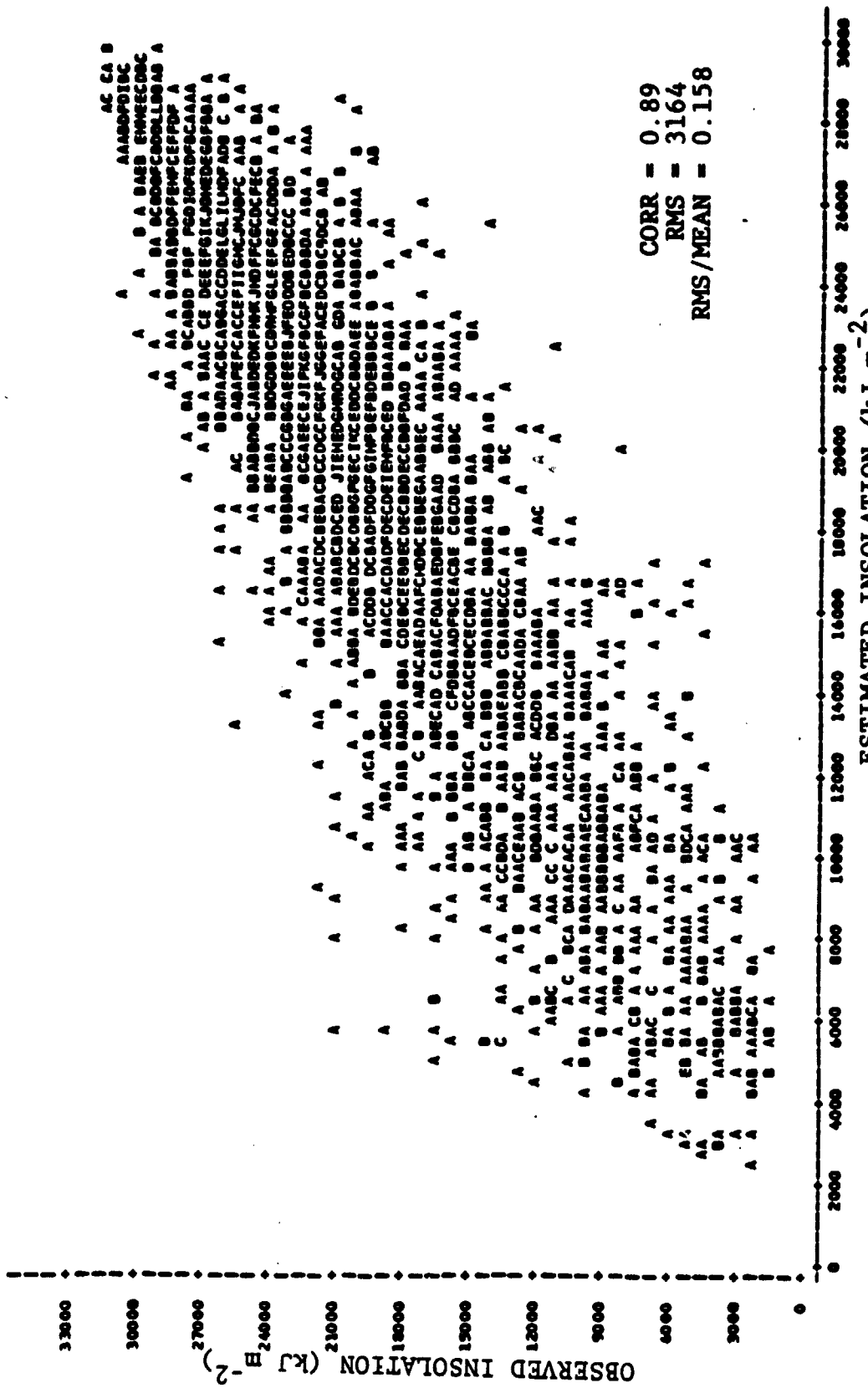


FIGURE 9 OBSERVED AND ESTIMATED DAILY INSOLATION, SUMMER 1200-1300 LST
(b) Linear Regression of Optical Thickness

conditions, then each estimate will be slightly biased because each regression will include clear data for other hours rather than considering the specific joint observation at the other two estimation hours. If treated simultaneously in a joint regression, the fact that all three hours were cloudy would indicate a much better chance for the entire day being cloudy and would result in lower magnitude than for the individual hourly treatments. Similarly, the single equation might give higher magnitudes for clear sky conditions than does the weighted sum of individual estimations. These possibilities suggest that the weighting factors to be applied to individual hourly estimates of daily insolation should take into account the known prevailing sky conditions at all (three) times of daily estimation.

If residual estimation bias is sufficiently large, it can be removed by an approximate correction applied directly to the estimate for a given hour. This can be accomplished in the following manner where E_k is the original estimate and E_k' is the revised estimate:

$$E_k' = c_0 + c_1 E_k + c_2 E_k^2 + \cdots \qquad (19)$$

In effect, only a slight linear rotation may be required. The coefficients in (19) are determined by specifying values of E_k' from the observed F_s in the dependent datasets. Subsequently, the revised estimates of F_s, through E_k', can be applied to the estimates that are obtained for the weighted combination of data from three different hours. Alternatively, Eq. (19) could be applied directly to a weighted estimate of the daily insolation.

TABLE 5

INITIAL AND REVISED CLOUD CATEGORIES AND TRANSMITTANCE FACTORS

CATEGORY	INITIAL* DESCRIPTOR		SECOND**	FINAL** DESCRIPTOR	
1	CLEAR	1.00	.98	CLEAR	.99
2	HIGH	.66	.80	HIGH	.75
3	MID w/wo HIGH	.50	.63	CU	.70
4	FOG; OBSCURATION	.40	.50	MIXED	.65
5	LOW	.33	.40	MID w/wo HIGH	.60
6	MIXED	.22	.35	ST	.48
7	THICK	.11	.22	THICK	.30

* $\psi = 0.6 + 0.7$
** $\psi = 0.7 + 0.3$

5.4 Interactive Tests From Satellite Photographs

To conduct a limited "operational" test of the technique for estimating total daily insolation, hard copy prints of restricted satellite imagery were supplied to the interactive analyst. Prints were generally available at resolutions of about 8km or lower, with three to four infrared images considered acceptable and with only a single visible print allowed in the test. Use of the low resolution single-channel GOES infrared data was expected to yield estimates inferior to those determined from SOLMET data. Nevertheless, the tests indicate the overall adequacy of the approach. Of course, when the model with the SOLMET data was formulated, it was not geared to the interactive analyst working with data at relatively low resolution. Independent data for testing the daily insolation estimation technique were acquired for a spring period (15 April to 15 May 1979) and a summer period (August,1979). Ten SOLMET stations were used, distributed in high plains (Midland, Dodge City), the prairie (Bismark,Omaha, Columbia), the southeast (Nashville, Montgomery, Raleigh) and the Ohio Valley (Indianapolis, Pittsburg). For each station on each day, observations of cloud amount and type were made from GOES full-disk IR imagery at times approximating those of the two polar orbiting satellites (3 a.m., 7 a.m., 3 p.m., 7 p.m.). One early afternoon visible image was used as well. In addition, dewpoints were recorded from hemisphere surface maps for each observation time, for use in the precipitable water term.

5.4.1 Spring Tests

SOLMET insolation data were only available for 196 of the 300 "station-days" for which data were collected. Three of the four observation times were used (early a.m., late p.m. and evening). Each of the three observations was applied to the particular predictive equation corresponding to the time of the observation. This yielded three differing predictions* of the daily total insolation. The three values were combined into a single estimate through weighting

---

*Implicit in the development of each estimation equation is the tacit assumption that conditions observed at that hour will prevail throughout the day. A daily average of the cosine of the solar zenith angle is included in the parameterization. Regression results appear to be relatively insensitive to changes in the effective solar zenith angle included.

techniques (the sum of the products of estimates and weighting factors divided by the sum of the weighting factors). The weighting procedure emphasizes the values closest to noon, the best time for estimates of daily totals. One technique used the correlation coefficients, associated with the appropriate hourly equations produced in the developmental stage, as the weights. A second technique modified each of these weights by multiplication with the additional factor $[1 + (\cos Z/\cos Z_{MX})]$ where the cosine of the solar zenith angle, cos Z, refers to the average for the particular hour whereas cos Z_{MX} refers to the solar noon value.

As an additional test, use was made of a single predictive equation with regression coefficients based on the incorporation of all three observations at the three different times in one regression expression. Only one estimate of the daily total insolation is obtained, and no weighting is involved in this technique.

5.4.2 Summer Tests

The summer tests were conducted in the same manner as the spring tests, and with about the same ratio of actual data matches to the total numbers of estimates (198 to 310 station days for summer). In addition to the three weighting methods (combining daily estimates from three separate observation times, and the single predictive equation), a third weighting method was introduced for the summer tests. This last weighting method was intended to be free of the specific regression results and thus did not involve the correlation coefficients. It did include the cosine of the maximum solar zenith angle (a function of latitude and declination) and the cosine of the solar zenith angle for the hour of estimation. Therefore, the non-correlative weighting method employed the weighting factor

$$[1.1 \cos Z_{MX} + \cos Z]$$

In this expression the cosine of the solar zenith angle could be allowed to become negative, since the coefficient of cos Z_{MX} would prevent the total weighting factor from becoming negative.

During the interactive analysis for the summer season it was possible to introduce significant errors in transition from a satellite image to a representative parameterization of cloudiness at a local point. Most of the thick

cloudiness was of the convective type with considerable variation over short distances. This may have contributed to some degradation in quality of the summer results as compared to those of spring.

5.4.3 Results From Interactive Tests

Table 6 summarizes the spring and summer test results obtained with the different weighting methods. In both tests the three-equation (individual hourly regressions) models with weighting of each daily estimate performed better than the single equation model (single regression for combination of hours), at least in terms of correlation and the standard deviation of the residual. The best result in both tests stems from the modified correlation weighting procedure, but a test with the non-correlative weighting showed about the same results

Table 6

RESULTS FROM INTERACTIVE TEST WITH SATELLITE IMAGERY

	CORRELATION Estimated-Observed		RATIO: Residual RMS To Observed Mean	
WEIGHTING METHOD	Spring	Summer	Spring	Summer
TEST 1 (Correlation weighting)	0.903	0.819	0.191	0.184
TEST 2 (Modified correlation weighting)	0.904	0.836	0.185	0.178
TEST 3 (Single pridictive equation)	0.888	0.794	0.239	0.184
TEST 4 (Non-correlative weighting)	——	0.831	——	0.179

In general, it was anticipated that clear-day insolation would be underestimated somewhat, while cloudy-day insolation would be overestimated. For the interactive tests, the second or intermediate set of cloud categories and transmittance factors was used (see Table 5). Some difficulty was expected with an interactive "estimate" of cloud amount; such estimates usually involve a regional averaging because of the impracticality of pinpointing a useful local cloud amount from a full disk-image. With the imagery available, cloud amount was specified only to the nearest quarter of the sky. Nevertheless, even though optimum data were not used and not all of the recent model improvements were included, the summary in Table 6 indicates that the interactive tests were quite successful.

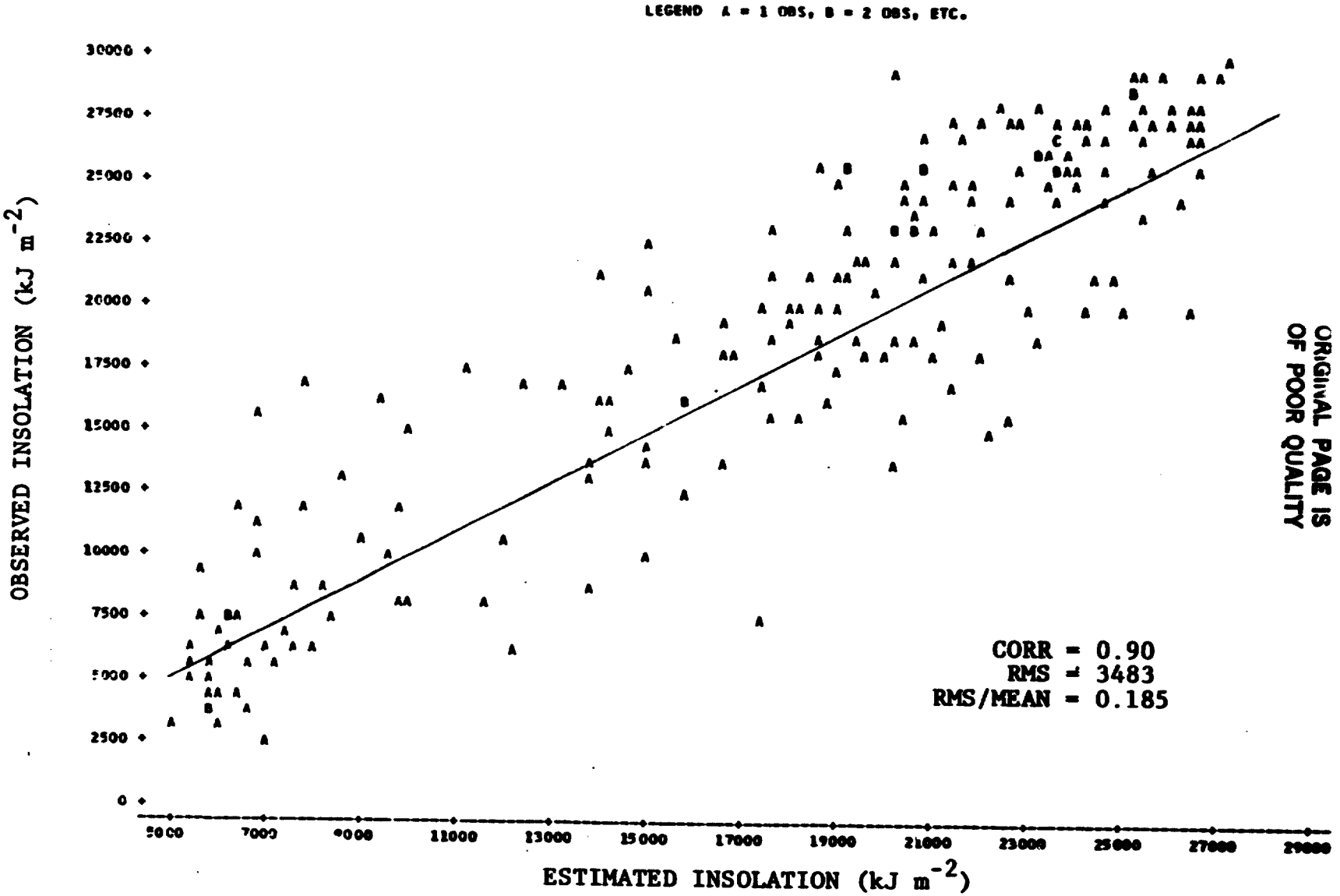


FIGURE 10 OBSERVED DAILY INSOLATION AND WEIGHTED ESTIMATES BASED ON THREE DAILY SATELLITE IMAGES, APRIL/MAY 1980

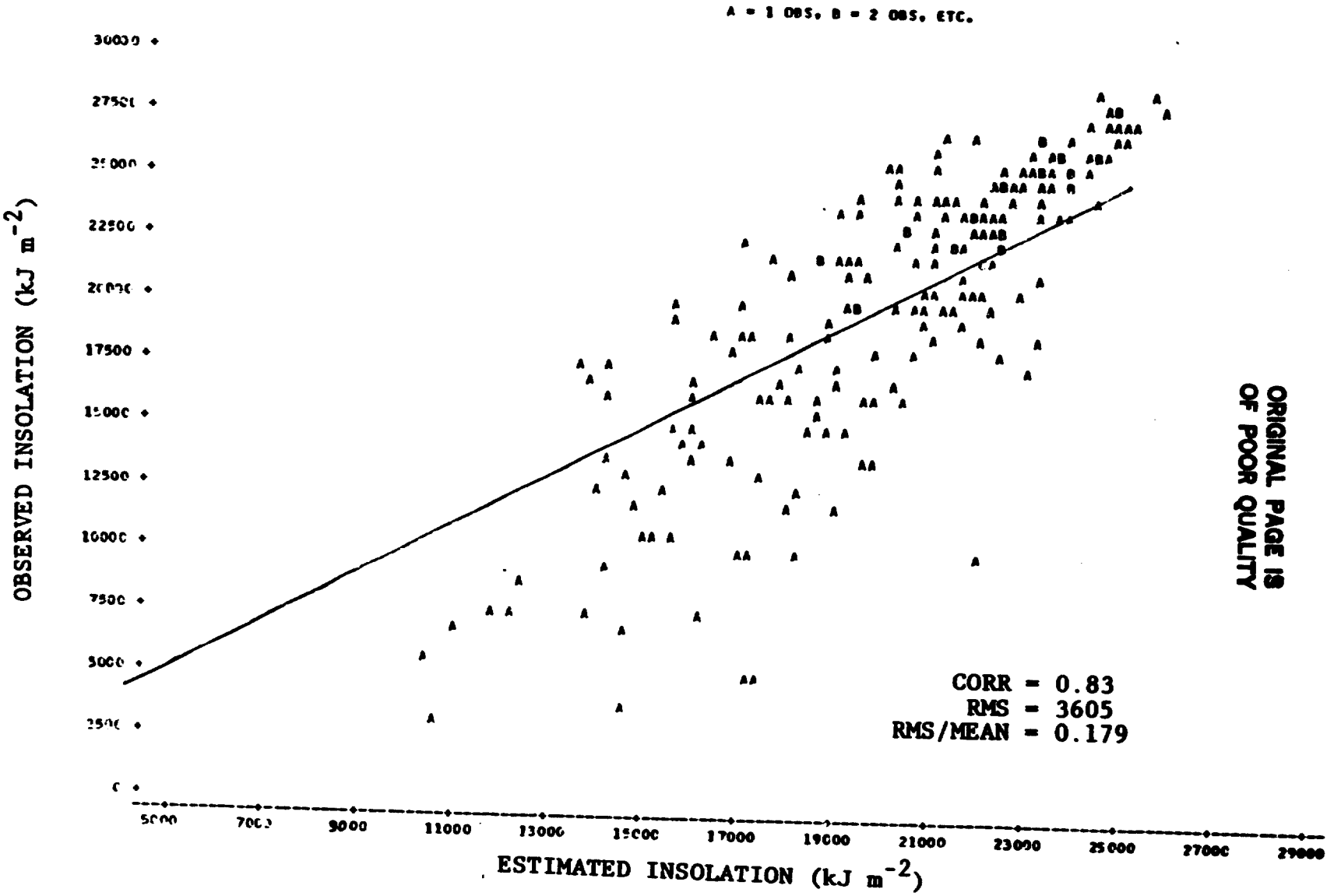


FIGURE 11 OBSERVED DAILY INSOLATION AND WEIGHTED ESTIMATED BASED ON THREE DAILY SATELLITE IMAGES, AUGUST 1980

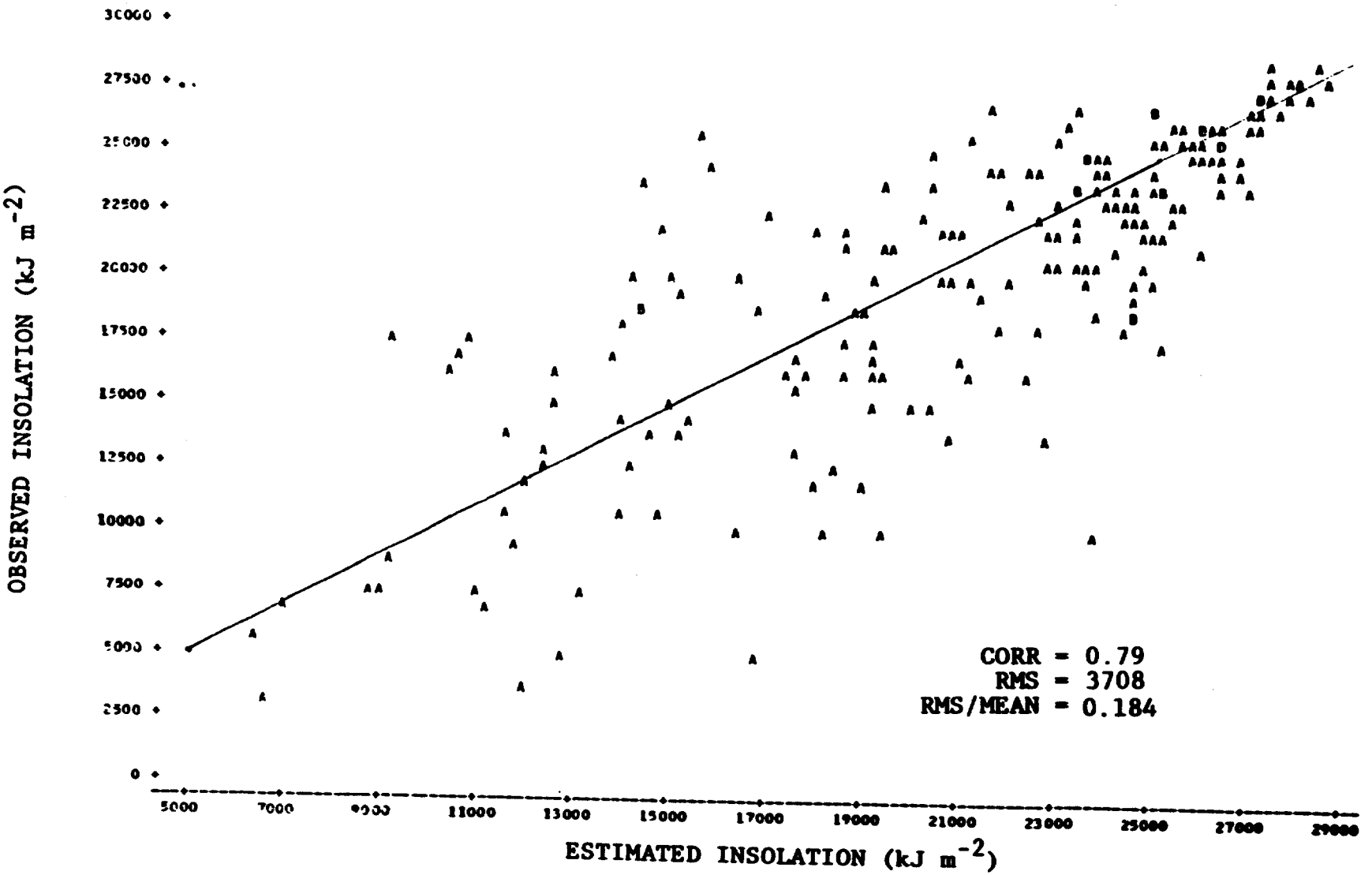


FIGURE 12 OBSERVED DAILY INSOLATION AND ESTIMATES BASED ON SINGLE PREDICTIVE EXPRESSION INCORPORATING DATA FROM THREE DAILY SATELLITE IMAGES, AUGUST 1980

Figure 10 illustrates results for the spring test and Figure 11 shows the summer results. Although the spring results show much less bias than those of the summer test, and show higher correlation, the ratio of the error to the observed mean is actually a bit smaller for the summer test. With the higher summertime mean this is so despite the more obvious bias in the summer results (somewhat exaggerated by the scales). Figure 12 shows another estimate for the summer test data, this time using the three hour combined regression equation (single predictive equation). This result shows less-biased estimates (in the extremes) than those in Figure 11, although the correlation is lower and the RMS a little larger (see Table 6). Not shown is the fact that the average deviation (absolute) of estimate from observation was actually smaller for the single predictive equation. Future development should reexamine the possibility of adjusting the single equation approach, with modification, as an alternative to improving the weighting factors for individual hours.

5.5 Final Regression Model Results

SOLMET data were split into three groups for final analysis: December-January-February-March(winter), April-May-October-November(spring-fall), and June-July-August-September(summer). The spring-fall combination subsequently was separated into spring and fall subgroups, each about half as large as the winter and summer groups. Attenuation regression coefficients, along with their standard errors, are listed for each hour of the day in Tables 7a-d. The number of observations involved in each hourly-daily analysis showed a minimum after midnight and a maximum in mid-afternoon for all seasons. In terms of the estimated insolation itself, after exponentiation with the derived coefficients, the correlation coefficients between estimated and observed insolation are determined, as are the standard deviation of the residual differences. These terms are included in the tables.

Seasonal changes in the estimation observation statistics are evident. Correlation coefficients, disregarding the nighttime values, are much the same in spring and summer, highest in fall, and almost as high in winter. The standard deviation of residual differences, or standard residual, shows a general inverse relationship to the correlation. It was smallest during the fall and winter. During the important spring and summer seasons, the ratio of the standard residual to the mean observed insolation becomes smaller in summer with the higher observed means.

TABLE 7a

SPRING

SUMMARY OF DAILY INSOLATION COEFFICIENTS DETERMINED FROM REGRESSION ANALYSIS FOR AVERAGE BACKGROUND (INTERCEPT), WATER VAPOR SLANT PATH, AND CLOUD ATTENUATION PARAMETERS

Standard errors are in parentheses

HOUR	NUMBER OF OBS	(a_0) INTERCEPT	(a_1) PATH	(a_2) CLOUD	CORR	STD RES
01	641	0.7214 (0.057)	-0.0205 (0.019)	0.5505 (0.054)	0.46	6620
02	641	0.6706 (0.056)	-0.0185 (0.019)	0.6116 (0.055)	0.49	6484
03	1141	0.6064 (0.041)	0.0089 (0.014)	0.6991 (0.042)	0.53	6354
04	1187	0.5872 (0.039)	0.0128 (0.014)	0.6967 (0.040)	0.55	6288
05	1187	0.5409 (0.039)	0.0247 (0.014)	0.7417 (0.041)	0.56	6182
06	1356	0.4771 (0.034)	0.0335 (0.012)	0.8127 (0.037)	0.62	5912
07	1502	0.4508 (0.031)	0.0278 (0.011)	0.9090 (0.034)	0.68	5537
08	1464	0.4118 (0.030)	0.0249 (0.010)	1.0642 (0.034)	0.74	5074
09	1487	0.3654 (0.027)	0.0323 (0.009)	1.1678 (0.031)	0.79	4583
10	1521	0.3390 (0.025)	0.0345 (0.009)	1.2456 (0.029)	0.83	4170
11	1504	0.2965 (0.023)	0.0377 (0.008)	1.3682 (0.028)	0.86	3784
12	1549	0.2666 (0.022)	0.0436 (0.007)	1.4058 (0.027)	0.88	3544
13	1585	0.2427 (0.021)	0.0497 (0.007)	1.4344 (0.026)	0.89	3341
14	1575	0.2047 (0.020)	0.0603 (0.007)	1.5110 (0.026)	0.89	3356
15	1591	0.2176 (0.021)	0.0671 (0.008)	1.4454 (0.027)	0.88	3595
16	1604	0.2442 (0.022)	0.0667 (0.008)	1.3848 (0.027)	0.86	3871
17	1556	0.2584 (0.024)	0.0660 (0.008)	1.3701 (0.030)	0.83	4239
18	1541	0.3205 (0.025)	0.0525 (0.009)	1.2659 (0.032)	0.78	4695
19	1546	0.3853 (0.027)	0.0389 (0.010)	1.1641 (0.033)	0.74	5048
20	1417	0.4269 (0.030)	0.0298 (0.011)	1.1182 (0.036)	0.70	5414
21	1332	0.4744 (0.033)	0.0230 (0.012)	1.0388 (0.038)	0.67	5691
22	1318	0.5208 (0.034)	0.0167 (0.012)	0.9693 (0.038)	0.64	5891
23	1179	0.4394 (0.037)	0.0670 (0.010)	0.9018 (0.045)	0.59	6233
24	746	0.5956 (0.044)	-0.0037 (0.016)	0.7598 (0.049)	0.58	5998

HOUR = Local time at end of hour
CORR = Correlation between estimated and observed daily insolation
STD RES = Standard deviation of residuals (kj m^{-2})

TABLE 7b

SUMMER

SUMMARY OF DAILY INSOLATION COEFFICIENTS DETERMINED FROM REGRESSION ANALYSIS FOR AVERAGE BACKGROUND (INTERCEPT), WATER VAPOR SLANT PATH, AND CLOUD ATTENUATION PARAMETERS

Standard errors are in parentheses

HOUR	NUMBER OF OBS	(a_0) INTERCEPT	(a_1) PATH	(a_2) CLOUD	CORR	STD RES
01	1325	0.4217 (0.043)	0.0556 (0.011)	0.4996 (0.032)	0.59	5610
02	1325	0.4136 (0.043)	0.0557 (0.011)	0.5279 (0.032)	0.60	5563
03	2197	0.5189 (0.031)	0.0329 (0.008)	0.5885 (0.026)	0.60	5664
04	2283	0.5035 (0.030)	0.0345 (0.008)	0.6019 (0.025)	0.62	5535
05	2289	0.4669 (0.029)	0.0403 (0.008)	0.6617 (0.025)	0.64	5372
06	2752	0.4114 (0.024)	0.0421 (0.006)	0.7400 (0.022)	0.70	4970
07	3045	0.4270 (0.023)	0.0308 (0.006)	0.7751 (0.020)	0.73	4796
08	2961	0.3918 (0.022)	0.0317 (0.006)	0.9174 (0.020)	0.76	4509
09	3028	0.3676 (0.020)	0.0311 (0.005)	1.0081 (0.019)	0.80	4156
10	3077	0.3533 (0.018)	0.0285 (0.005)	1.1042 (0.018)	0.83	3803
11	3049	0.3299 (0.017)	0.0284 (0.004)	1.2350 (0.018)	0.87	3436
12	3203	0.3074 (0.015)	0.0311 (0.004)	1.3056 (0.016)	0.89	3160
13	3254	0.3159 (0.015)	0.0312 (0.004)	1.2673 (0.016)	0.89	3164
14	3225	0.2856 (0.015)	0.0376 (0.004)	1.3304 (0.017)	0.89	3094
15	3274	0.2744 (0.015)	0.0436 (0.004)	1.3039 (0.017)	0.88	3247
16	3283	0.2975 (0.016)	0.0422 (0.004)	1.2235 (0.018)	0.86	3509
17	3173	0.2994 (0.017)	0.0454 (0.005)	1.1911 (0.019)	0.83	3802
18	3156	0.3318 (0.019)	0.0457 (0.005)	1.0578 (0.021)	0.78	4296
19	3152	0.3832 (0.021)	0.0395 (0.006)	0.9447 (0.022)	0.74	4704
20	2823	0.4103 (0.024)	0.0388 (0.006)	0.8952 (0.025)	0.69	5072
21	2706	0.4239 (0.026)	0.0426 (0.007)	0.7785 (0.025)	0.66	5297
22	2692	0.4661 (0.027)	0.0372 (0.007)	0.7027 (0.025)	0.64	5483
23	2331	0.3918 (0.029)	0.0665 (0.007)	0.6531 (0.029)	0.60	5718
24	1589	0.4637 (0.033)	0.0432 (0.009)	0.5447 (0.030)	0.61	5648

HOUR = Local time at end of hour
CORR = Correlation between estimated and observed daily insolation
STD RES = Standard deviation of residuals (kj m^{-2})

TABLE 7c

FALL

SUMMARY OF DAILY INSOLATION COEFFICIENTS DETERMINED FROM REGRESSION ANALYSIS FOR AVERAGE BACKGROUND (INTERCEPT), WATER VAPOR SLANT PATH, AND CLOUD ATTENUATION PARAMETERS

Standard errors are in parentheses

HOUR	NUMBER OF OBS	(a_0) INTERCEPT	(a_1) PATH	(a_2) CLOUD	CORR	STD RES
01	603	0.6111 (0.071)	0.0583 (0.019)	0.4744 (0.050)	0.62	4310
02	603	0.5886 (0.070)	0.0602 (0.019)	0.5154 (0.051)	0.63	4271
03	999	0.6279 (0.055)	0.0408 (0.014)	0.5958 (0.041)	0.65	4082
04	1034	0.5751 (0.052)	0.0477 (0.013)	0.6383 (0.038)	0.69	3913
05	1003	0.5550 (0.052)	0.0499 (0.013)	0.6697 (0.038)	0.70	3786
06	1111	0.5355 (0.045)	0.0462 (0.012)	0.6952 (0.034)	0.75	3713
07	1311	0.4960 (0.039)	0.0435 (0.011)	0.7789 (0.030)	0.79	3425
08	1292	0.4463 (0.039)	0.0423 (0.010)	0.8522 (0.030)	0.81	3177
09	1395	0.3868 (0.032)	0.0394 (0.009)	0.9590 (0.027)	0.85	2821
10	1417	0.3586 (0.031)	0.0367 (0.008)	1.0185 (0.025)	0.89	2461
11	1385	0.3210 (0.029)	0.0394 (0.008)	1.1035 (0.024)	0.91	2155
12	1447	0.2882 (0.026)	0.0442 (0.007)	1.1678 (0.022)	0.93	1927
13	1471	0.2858 (0.025)	0.0449 (0.007)	1.1818 (0.021)	0.94	1869
14	1448	0.2787 (0.025)	0.0473 (0.007)	1.1961 (0.022)	0.93	1941
15	1474	0.3010 (0.026)	0.0476 (0.007)	1.1537 (0.023)	0.92	2133
16	1493	0.3395 (0.027)	0.0458 (0.008)	1.0989 (0.023)	0.90	2374
17	1420	0.3505 (0.030)	0.0500 (0.008)	1.0811 (0.026)	0.87	2635
18	1378	0.3624 (0.034)	0.0562 (0.009)	1.0184 (0.028)	0.85	2917
19	1372	0.3989 (0.036)	0.0540 (0.007)	0.9543 (0.028)	0.83	3105
20	1221	0.4403 (0.042)	0.0494 (0.011)	0.9042 (0.032)	0.79	3293
21	1165	0.4829 (0.046)	0.0438 (0.011)	0.8298 (0.034)	0.77	3421
22	1169	0.5656 (0.047)	0.0303 (0.012)	0.7652 (0.035)	0.75	3573
23	1078	0.5366 (0.055)	0.0721 (0.011)	0.6410 (0.041)	0.69	3913
24	651	0.5807 (0.063)	0.0482 (0.017)	0.6052 (0.048)	0.68	3926

HOUR = Local time at end of hour
CORR = Correlation between estimated and observed daily insolation
STD RES = Standard deviation of residuals (kj m^{-2})

TABLE 7d

WINTER

SUMMARY OF DAILY INSOLATION COEFFICIENTS DETERMINED FROM REGRESSION ANALYSIS FOR AVERAGE BACKGROUND (INTERCEPT), WATER VAPOR SLANT PATH, AND CLOUD ATTENUATION PARAMETERS

Standard errors are in parentheses

OUR	NUMBER OF OBS	(a_0) INTERCEPT	(a_1) PATH	(a_2) CLOUD	CORR	STD RES
1	1658	0.5878 (0.035)	0.1362 (0.013)	0.3258 (0.032)	0.56	4673
2	1658	0.5705 (0.035)	0.1349 (0.013)	0.3546 (0.033)	0.57	4645
3	2800	0.5968 (0.027)	0.0906 (0.009)	0.4197 (0.025)	0.62	4460
4	2880	0.5688 (0.026)	0.0887 (0.009)	0.4592 (0.024)	0.65	4398
5	2812	0.5259 (0.027)	0.0929 (0.009)	0.5118 (0.024)	0.67	4261
6	2993	0.4865 (0.024)	0.0988 (0.009)	0.5403 (0.022)	0.70	4123
7	3251	0.4407 (0.022)	0.0984 (0.008)	0.5903 (0.020)	0.74	3913
8	3227	0.3873 (0.021)	0.0938 (0.008)	0.6832 (0.020)	0.77	3662
9	3490	0.3315 (0.018)	0.0894 (0.007)	0.7745 (0.018)	0.82	3290
0	3552	0.3170 (0.017)	0.0851 (0.006)	0.8169 (0.017)	0.85	3035
1	3461	0.3017 (0.016)	0.0781 (0.006)	0.8978 (0.017)	0.88	2729
2	3520	0.2748 (0.015)	0.0815 (0.006)	0.9527 (0.016)	0.90	2470
3	3571	0.2586 (0.014)	0.0814 (0.005)	0.9963 (0.015)	0.91	2314
4	3512	0.2513 (0.015)	0.0850 (0.006)	1.0184 (0.015)	0.92	2280
5	3582	0.2563 (0.015)	0.0923 (0.006)	0.9907 (0.015)	0.91	2410
6	3623	0.2795 (0.015)	0.0959 (0.006)	0.9340 (0.016)	0.89	2646
7	3492	0.2901 (0.017)	0.0970 (0.006)	0.9078 (0.017)	0.87	2902
8	3470	0.3180 (0.018)	0.1007 (0.007)	0.8362 (0.018)	0.84	3221
9	3476	0.3636 (0.018)	0.0946 (0.007)	0.7803 (0.018)	0.82	3435
0	3124	0.3928 (0.021)	0.0906 (0.008)	0.7509 (0.020)	0.78	3619
1	2982	0.4366 (0.023)	0.0861 (0.008)	0.6923 (0.021)	0.75	3763
2	2996	0.4847 (0.023)	0.0752 (0.008)	0.6492 (0.021)	0.73	3893
3	2825	0.4983 (0.025)	0.0688 (0.007)	0.6368 (0.023)	0.71	3985
4	1882	0.5209 (0.028)	0.0688 (0.011)	0.5385 (0.028)	0.69	4078

HOUR = Local time at end of hour
CORR = Correlation between estimated and observed daily insolation
STD RES = Standard deviation of residuals (kj m^{-2})

The standard errors of the regression coefficients are included parenthetically after each coefficient in the tables. Only the few of the nighttime water vapor path coefficients during spring (especially the negative values) are insignificant. Otherwise, the daytime water vapor path coefficients are smallest in summer and largest in winter. The most important coefficient, for the cloud transmittance term, is largest in spring and declines with season to smallest values in winter. Intercepts tend to be smallest in spring and summer and larger in fall and winter.

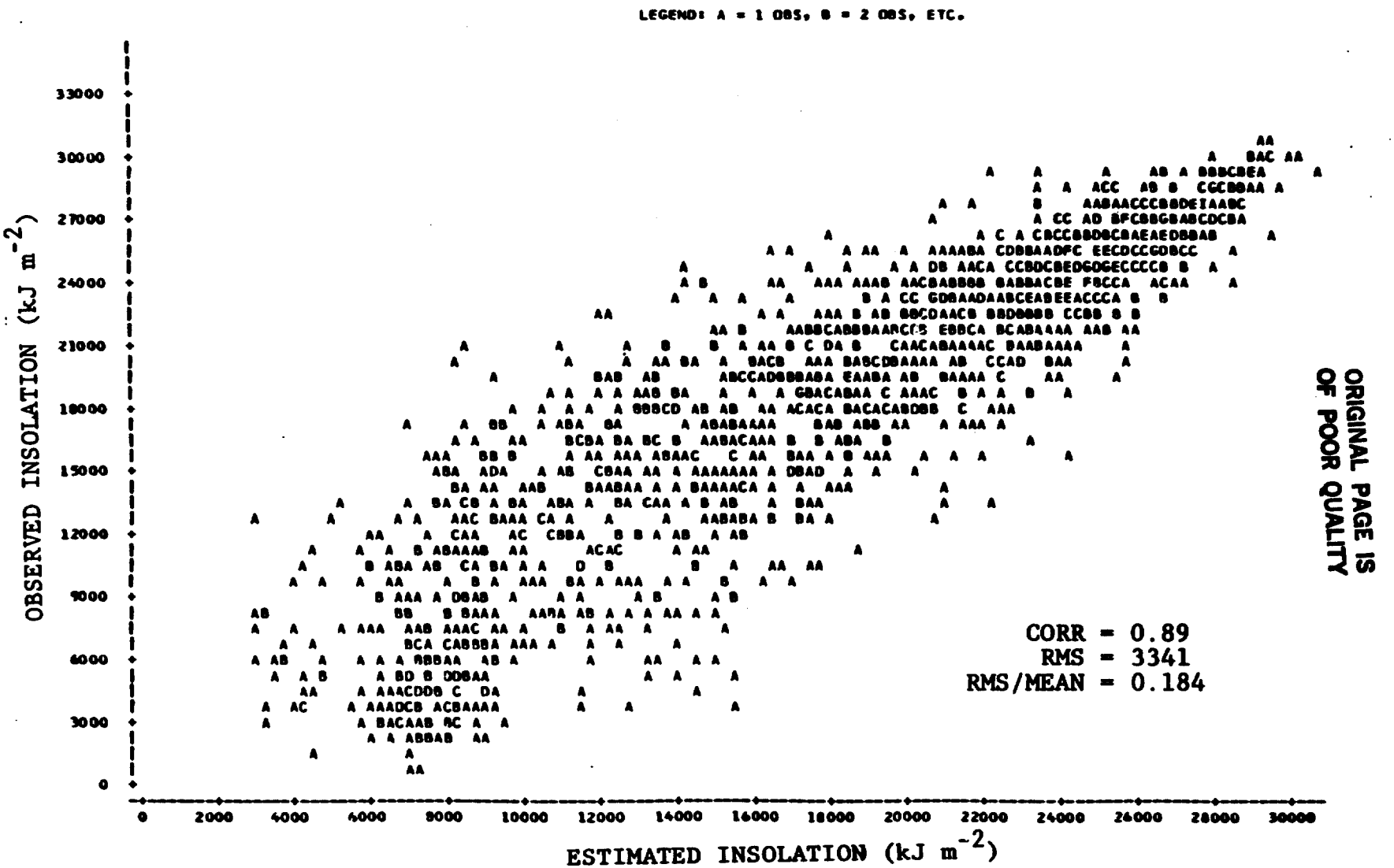
All coefficients vary diurnally. Near midnight, the explained variances and correlation cofficients reach low levels, in many instances dropping below the variance associated simply with the changing solar zenith angle. Below such a threshold it is considered that the regression relationships for that hour contain little useful information for estimation of the total daily insolation.

Figures 13-16 illustrate distributions of estimated and observed daily insolation, based on the regression results that are available for the hour 1200-1300 during each seasonal period. These illustrations give a good view of the scatter and residual bias in the data distributions for this optimal hour. In application, the daily estimate would be based on a combination of estimates from three different times. A distribution of such results would resemble the distributions illustrated.

5.6 Independent Data Tests

5.6.1 Independent SOLMET Data

Two months of SOLMET data, for an April and a May, were withheld from the final regression analysis (dependent data) so that they could be used as a source of independent data for checking results. Two tests were made with these data, by using information taken from three different hours (simulating data coverage from two polar-orbiting satellites). The only difference between tests was a one-hour shift in the time of the central data input, from the 1300-1400 local hour to 1400-1500. In each case the results were weighted in accordance with Eqs.(16) and (17). For these tests the latest or "final" version of the cloud categories and transmittance factors (see Table 5) was used. For one of the tests, an adjustment to remove bias, in accordance with two terms of Eq. (19), was attempted with factors based on the empirical dependent data for spring.



ORIGINAL PAGE IS
OF POOR QUALITY

FIGURE 13 OBSERVED DAILY INSOLATION AND ESTIMATES BASED ON REGRESSION FOR 1200-1300 LST, SPRING

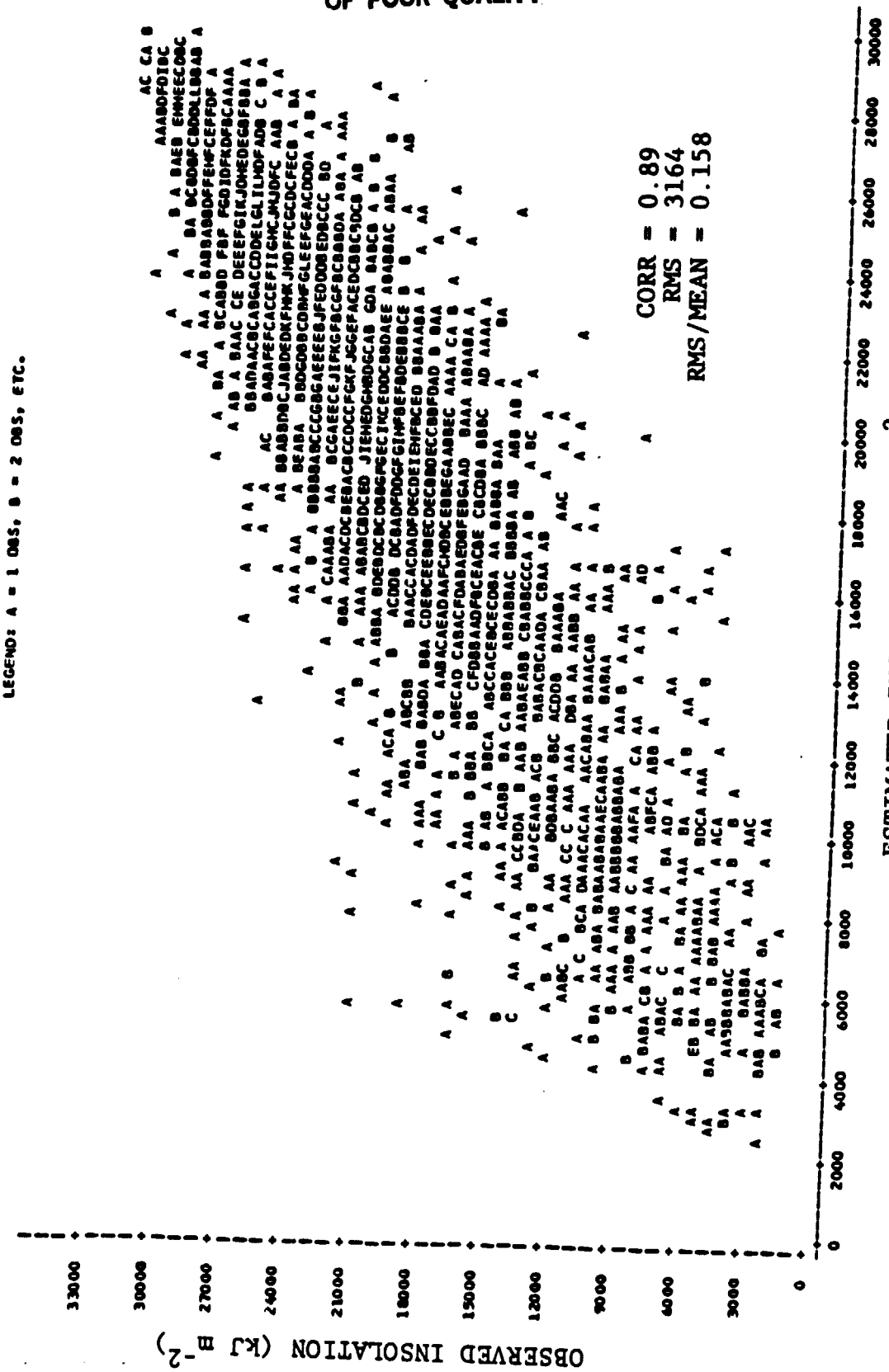


FIGURE 14 OBSERVED DAILY INSOLATION AND ESTIMATES BASED ON REGRESS. ON FOR 1200-1300 LST, SUMMER

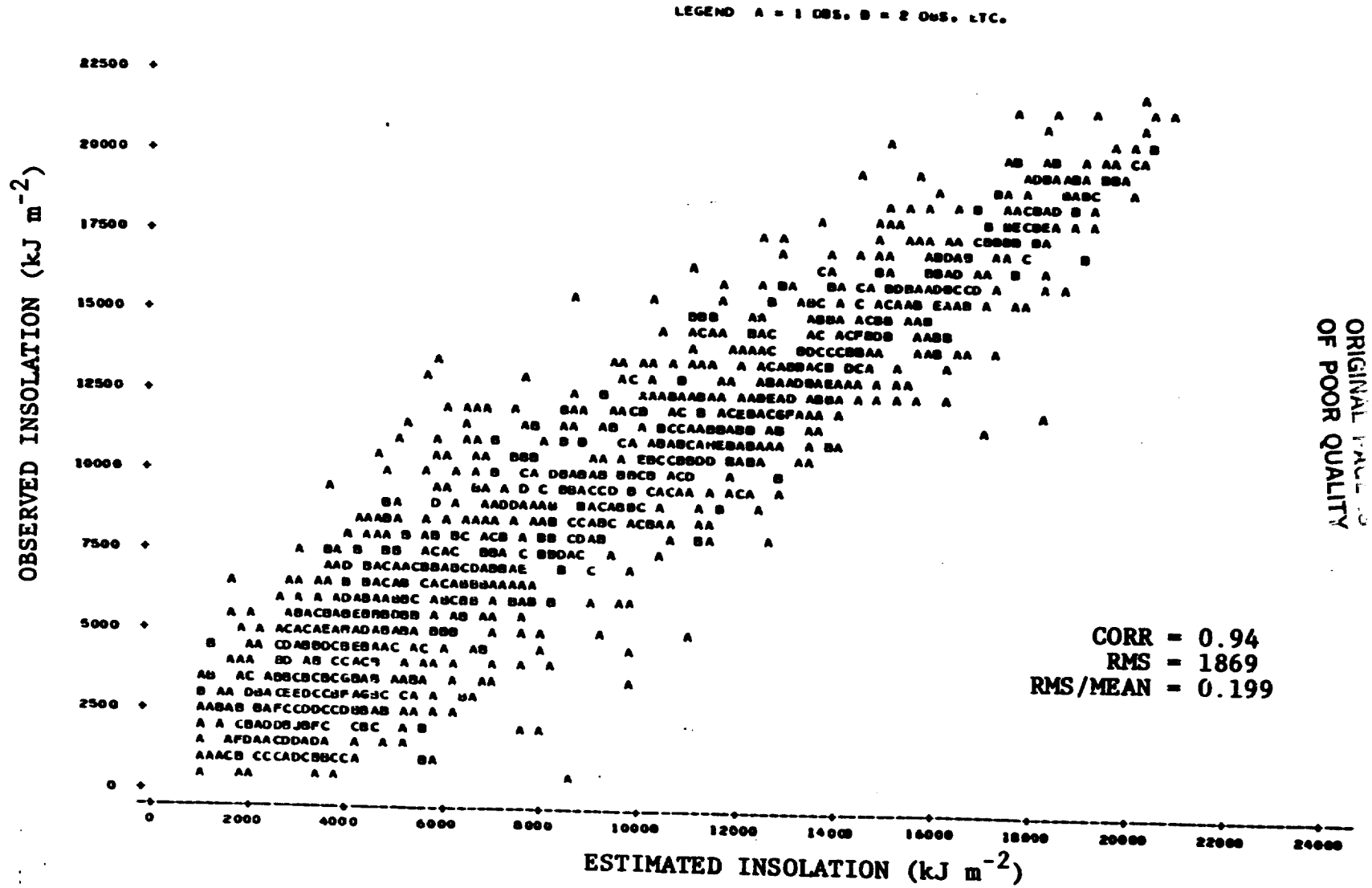


FIGURE 15 OBSERVED DAILY INSOLATION AND ESTIMATES BASED ON REGRESSION FOR 1200-1300 LST, FALL

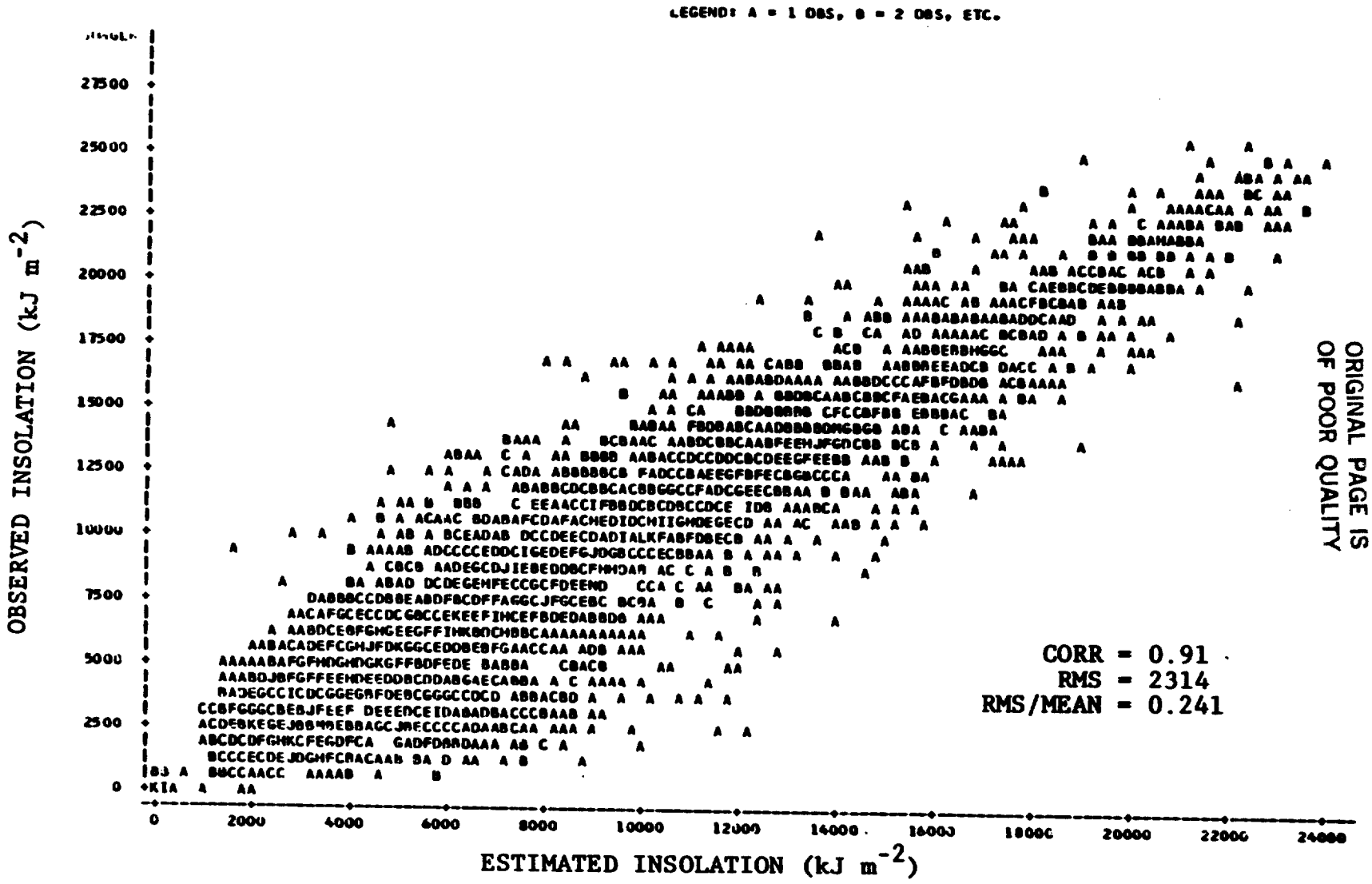


FIGURE 16 OBSERVED DAILY INSOLATION AND ESTIMATES BASED ON REGRESSION FOR 1200-1300 LST, WINTER

Results for the tests with independent SOLMET data are shown in Figure 17a, b, and c. In the statistical sense it must be concluded that the tests were certainly satisfactory, even though satellite inputs were not actually used here. The shift in the midafternoon data source by one hour had a slight but not especially significant impact on the results. The adjustment to remove the regression bias improved the RMS by 8 percent and visibly improved the symmetry of the distribution. The bias correction here was linear, for the three constituent hours. A non-linear or step-function correction might improve the results further.

5.6.2 Independent Digital Infrared Satellite Data

The aim of using digital multispectral data from an operational polar-orbiting satellite is an automated technique for insolation estimation involving cloud attenuation information. The initial step of specifying cloud types, for the parameterization of cloud transmittances was designed and tested as discussed in the Appendix.

Of most importance for an automated operational procedure is the availability of an objective cloud type specification. Data from seven channels of the operational HIRS sounder were selected for analysis toward this development. It was assumed that high resolution AVHRR data might be used with a thresholding method to define a corresponding cloud amount. Ultimately, both cloud type and amount for a given resolution (say 40 km) could be passed along to the interactive analyst, who could assimilate this information with whatever imagery was available for making final estimates.

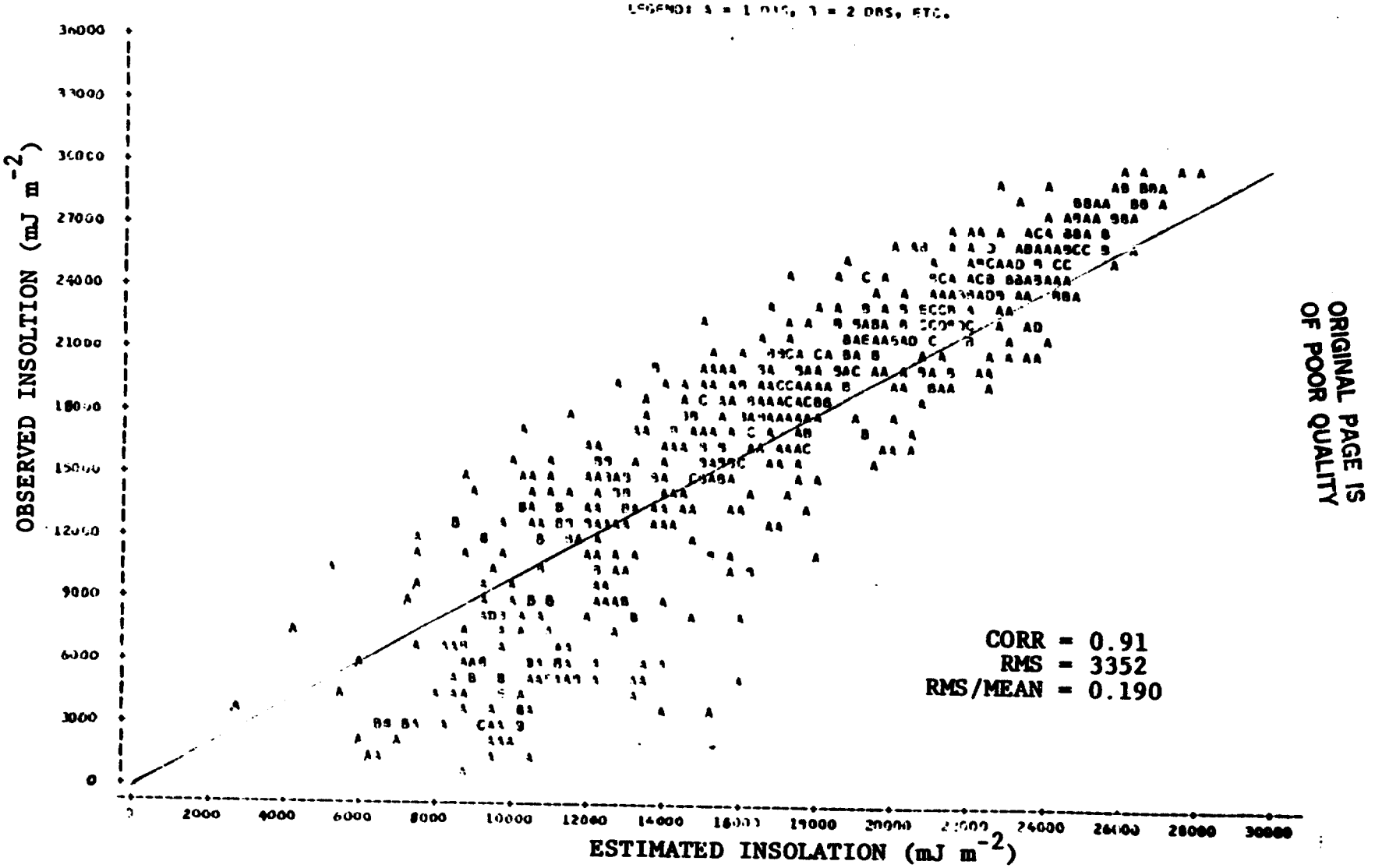


FIGURE 17 OBSERVED DAILY INSOLATION AND WEIGHTED ESTIMATES BASED ON THREE INDEPENDENT DAILY OBSERVATIONS, SPRING
(a) 0800, 1400, 1900

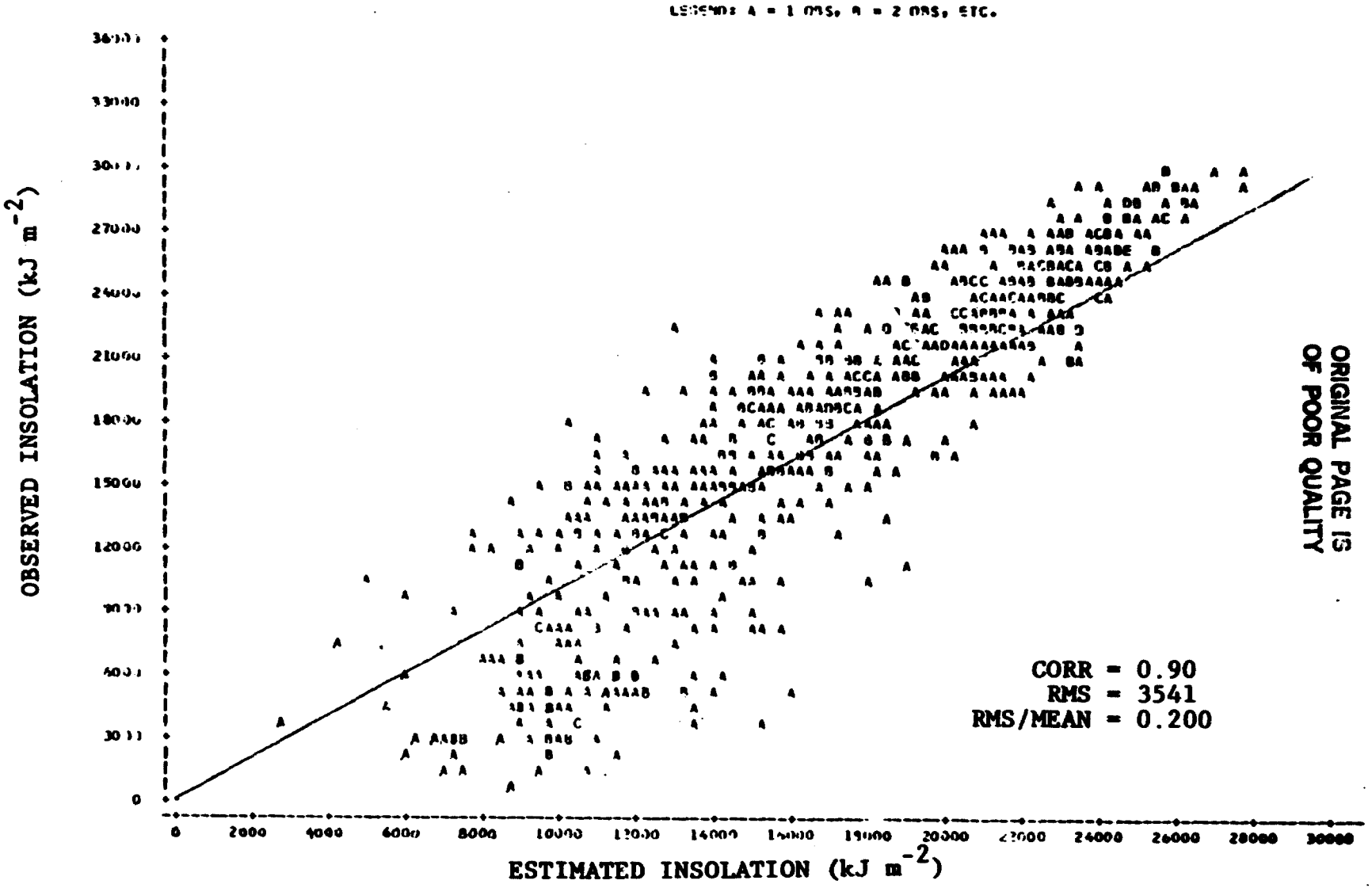


FIGURE 17 OBSERVED DAILY INSOLATION AND WEIGHTED ESTIMATES BASED ON THREE INDEPENDENT DAILY OBSERVATIONS, SPRING 1979
(b) 0800,1500,1900

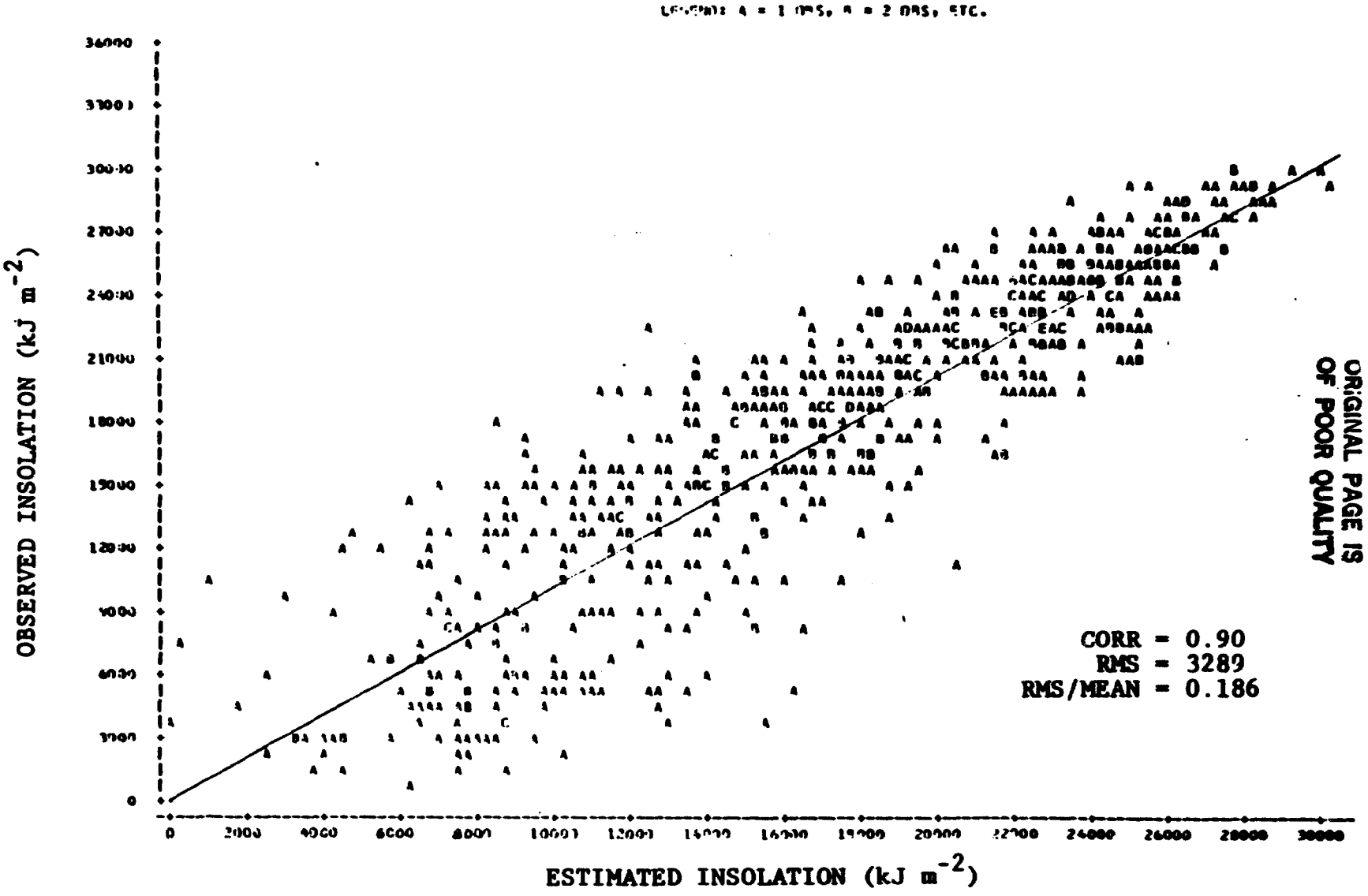


FIGURE 17 OBSERVED DAILY INSOLATION AND WEIGHTED ESTIMATES BASED ON THREE INDEPENDENT DAILY OBSERVATIONS, SPRING 1979
(c) 0800,1500, 1900---bias adjusted

6.0 CONCLUSIONS AND RECOMMENDATIONS

Results obtained from analyses of SOLMET data as well as tests have demonstrated the feasibility of implementing an operational algorithm for estimation of insolation on the basis of polar-orbiting satellite data. Surface data or other satellite data could be used as well. Cloud amount and cloud type are the primary predictors. Total water vapor is much less significant statistically, but the "water vapor" term is scaled to include surface pressure (or alternatively, height above sea level) as well as slant path (through sec Z). When diurnal variations are excluded in analysis, there is much less variance in the water term. To some extent there is also a built-in compensation on the optical thickness: the larger the average slant path due to latitude, the smaller is the optical thickness due to water vapor.

Generally, the selection of parameters has not led to any obvious conclusion on the spatial bias, if any. Some examination of regression coefficients has shown no coherent patterm of a geographical or site-specific dependency. On the other hand, any reconsideration of spatial bias will require a closer look at the geographical variation of diurnal and seasonal insolation distributions, which can depart from normal distributions. A seasonal separation of regression coefficients has already been adopted.

The present technique makes use of a set of regression coefficients for each hour (during any given season) to provide an estimate of the total daily insolation for the closest daytime period. For application to any location covered by a polar-orbiting satellite, all that is required, in addition to the pertinent hourly regression coefficients, is the time, latitude, and solar declination (assuming that the incident solar flux across a horizontal surface at the top of the atmosphere is known). Individual hourly estimates are combined through weighting factors that depend on the existing and maximum solar zenith angles.

The present method requires that the analyst supply three pieces of information: the gross cloud type, the cloud amount, and the total precipitable water or surface dewpoint. Initially the cloud type and amount will be determined by an analyst using an interactive system displaying satellite imagery. The feasibility of this approach was demonstrated in the interactive tests of Section 5.4. If not available from satellite sounder profiles of mixing ratio, the precipitable water quantity can be acquired from NMC analyses or from climatological values if no other source is available. For each time of satellite

coverage, tabulated regression coefficients are applied to the three parameters to obtain estimates of daily insolation. Results for the several hours are weighted to arrive at a single daily estimate. The interactive analyst can use his knowledge of atmospheric disturbances and their characteristic cloud signatures to alter weighting fctors if he believes the available cloud observations do not accurately represent the day as a whole for any particular location. Similarly, the bias introduced by the regression process which results in the overestimation of insolation on mostly cloudy days and the underestimation of insolation on mostly clear days can be removed if the analyst detects either of these two situations and compensates accordingly.

It is likely that algorithm improvements can be accomplished with improved parameterization and weighting. As long as surface data are used in the estimation process, surface dewpoints can be used (as estimators of total precipitable water). However, an alternative is required from satellite data. At the present time it appears that the retrieved mixing ratios that are included in the TOVS product should be applied, but other measures might be found also. More work is required on the cloud parameters, with possible consideration of an automated treatment of the larger scale cloud organization (helpful for filling in gaps between polar-orbiter data coverage). Reflectance data, both of cloud and of the surface, should be introduced even if limited to a single time each day. When available, the diurnal range of temperature can be used as an additional estimator of daily insolation or at least as a check on the estimated insolation.

One step in the improvement process is to engage in iterative regression to a greater extent that was done here. In other words, first solutions of coefficients are used to infer cloud parameters, and the inferred cloud parameters are then used to obtain new coefficients, with the process continuing until a convergence is obtained. Furthermore, if the final results tend to be nonlinear, departures of the estimates from the original observations can be used in another regression to eliminate the bias.

As the polar-orbiting satellite data are used in the algorithm, it will be beneficial in the learning stage to match the satellite data with the SOLMET data. This will provide the most direct link to the desired parameters. When interactive procedures are used, it may be most profitable for the analyst to estimate parameters over the sites of interest at other times than those of satellite passage. In fact, the times can be fixed near noon, or at critical departures from noon so that fixed weights can be applied.

Other satellite data, at high resolution, could be used to estimate a specific cloud amount and temperature. Results for each grid element could be stored with the TOVS product summary. Surface reflectance data should be added to permit an additional correction to the insolation estimation when necessary (e.g., with a fresh snow cover).

Finally, the SOLMET stations with both direct and diffuse solar radiation measurements should be examined separately to enable a direct-diffuse irradiance separation. This separation should lead ultimately to a better insolation estimation algorithm, especially for cloudy and partly cloudy conditions.

It is recommended that action be taken to complete the development of the use of the HIRS data in the specification of cloud parameters. Results or summaries derived therefrom, even if not used in an automated analysis, would aid in the decision making process of the interactive analyst.

REFERENCES

1. S. Fritz, P.K. Rao, and M. Weinstein, 1964: Satellite Measurements of Reflected Solar Energy and Energy Received at the Ground, J. Atmos. Sci., 21, 141-151.

2. K.J. Hanson, T.H. Vonder Haar, and V.E. Suomi, 1967: Reflection of Sunlight to Space and Absorption by the Earth and Atmosphere Over the United States During Spring, 1962, Mon. Wea. Rev., 95, 354-360.

3. W.H. Quinn, 1971: Studies of Parameterization of Solar Irradiance at the Earth's Surface, Proc. of the Miami Workshop on Remote Sensing; March 29-31, 1971, Miami, Florida (pp. 21-38).

4. J.S. Ellis and T.H. Vonder Haar, 1976: Application of Meteorological Satellite Visible Channel Radiances for Determining Solar Radiation Reaching the Ground, Proc. Seventh Conf. on Aerospace and Aeronautical Meteor. and Symp. on Remote Sensing from Satellites, Melbourne, Florida.

5. M.A. Atwater and P.S. Brown, Jr., 1974: Numerical Computations of the Latitudinal Variation of Solar Radiation for an Atmosphere of Varying Opacity, J. Appl. Meteor., 13, 289-297.

6. M.A. Atwater and J.T. Ball, 1978: Intraregional Variations of Solar Radiation in the Eastern United States, J. Appl. Meteor., 17, 1116-1125.

7. A.A. Lacis and J.E. Hansen, 1974: A Parameterization for the Absorption of Solar Radiation in the Earth's Atmosphere, J. Atmos. Sci., 31, 118-133.

8. J.D. Tarpley, 1979: Estimating Incident Solar Radiation at the Surface from Geostationary Satellite Data, J. Appl. Meteor., 18, 1172-1181.

9. C. Gautier, G. Diak and S. Masse, 1980: A Simple Physical Model to Estimate Incident Solar Radiation at the Surface from GOES Satellite Data, J. Appl. Meteor., 19, 1005-1012.

10. SOLMET, VOLUME I - USER'S MANUAL, Hourly Solar Radiation Surface Meteorological Observations, TD-9724, August 1978, U.S. Department of Commerce, National Oceanic and Atmsopheric Administration, Environmental Data and Information Service, National Climatic Center, Asheville, North Carolina (1980).

11. L. Lauritson, G.N. Nelson and F.W. Porto, 1979: Data Extraction and Calibration of TIROS-N/NOAA Radiometers, NOAA Technical Memorandum NESS107, National Oceanic and Atmospheric Administration, Washington, D.C.

12. W.L. Smith, 1966: Note on the Relationship Between Total Precipitable Water and Surface Dew Point, J. Appl. Meteor., 5, 726-727.

13. C.D. Rodgers, 1967: The Radiative Heat Budget of the Troposphere and Lower Stratosphere. Rept. No. A2, Dept. of Meteorology, M.I.T., Cambridge, MA.

14. P.A. Davis and G.C. Chatters, 1981: Insolation Estimation with Data Derivable from Polar-Orbiting Satellites Workshop Proceedings, Satellites and Forecasting of Solar Radiation, Am. Section of the International Solar Energy Society, February 2-5, 1981, Washington, D.C.

15. K-N Liou, 1976: On the Absorption, Reflection and Transmission of Solar Radiation in Cloudy Atmospheres, J. Atmos. Sci., 33, 798-805.

APPENDIX: HIRS DATA ANALYSIS FOR CLOUD TYPE

As a first step to the analysis of HIRS data into classes related to cloud parameters of interest for the estimation of insolation, cloud targets had to be selected on the basis of known cloud types. Prints of GOES imagery were acquired to assist is the selection of the cloud targets, with initial emphasis on cloud type rather than specific amount. These selections represent the first phase of a supervised cluster analysis. In effect, the cloud specifications, in terms of previously established cloud type categories, constituted the cloud truth for the study.

From all available HIRS infrared data, seven channels, centered at wave numbers of 705, 732, 748, 901, 1218, 2191 and 2210 cm^{-1}, were selected initially for further analysis. An effort was made to avoid excessive duplication in channel response, but to maintain a minimum number of channels consistent with the described number of cloud types. Data from each selected channel were limb corrected; selected printouts were made in terms of scanlines and scanspots. The application of latitude and longitude grid lines to the printouts makes it possible to check the accuracy of data locations and matchups. A gray-scaled printout of infrared window data from HIRS can be checked readily against the gridded image of the GOES infrared window.

Initial cloud types were defined so as to avoid ambiguities as much as possible. Six adopted types were CLEAR, HIGH, MID, LOW, MIXED and THICK. A stepwise multiple discriminant analysis was applied to the HIRS data (using the seven channels as predictors to specify six cloud types). Table A-1 summarizes the results obtained from the SPSS discriminant analysis package as applied to a June 26,1980 data swath(NOAA-6 satellite). It is apparent that clear skies over cool water were interpreted as low clouds in a number of cases. The HIGH category shows same spread associated with variations in opacity. As might be expected, the MIXED category sometimes is assigned to a particular cloud type.

Figure A-1 is a plot in the space of the first two discriminant functions. Sets of coefficients derived for each function are applied to each observation in determination of the category. Separation of data points, each assigned a cloud category, is apparent in Figure A-1, but distances between groups are not large. The same discriminant functions were applied in an independent test to HIRS data

TABLE A-1

SUMMARY OF INITIAL CLOUD CLASSIFICATIONS FROM HIRS/2 MULTISPECTRAL SCANSPOT DATA

CLASSIFICATION RESULTS -

ACTUAL GROUP		NO. OF CASES	PREDICTED GROUP MEMBERSHIP 1	2	3	4	5	6
CLEAR	1	62	51 82.3	0 0.0	0 0.0	0 17.7	0 0.0	0 0.0
HIGH	2	24	1 4.2	19 79.2	2 8.3	2 8.3	0 0.0	0 0.0
MID	3	17	0 0.0	0 0.0	16 94.1	0 0.0	1 5.9	0 0.0
LOW	4	22	0 0.0	0 0.0	0 0.0	22 100.0	0 0.0	0 0.0
MIXED	5	43	0 0.0	1 2.3	2 4.7	1 2.3	39 90.7	0 0.0
THICK	6	32	0 0.0	0 0.0	0 0.0	0 0.0	1 3.1	31 96.9

PERCENT OF GROUPED CASES CORRECTLY CLASSIFIED 89.00

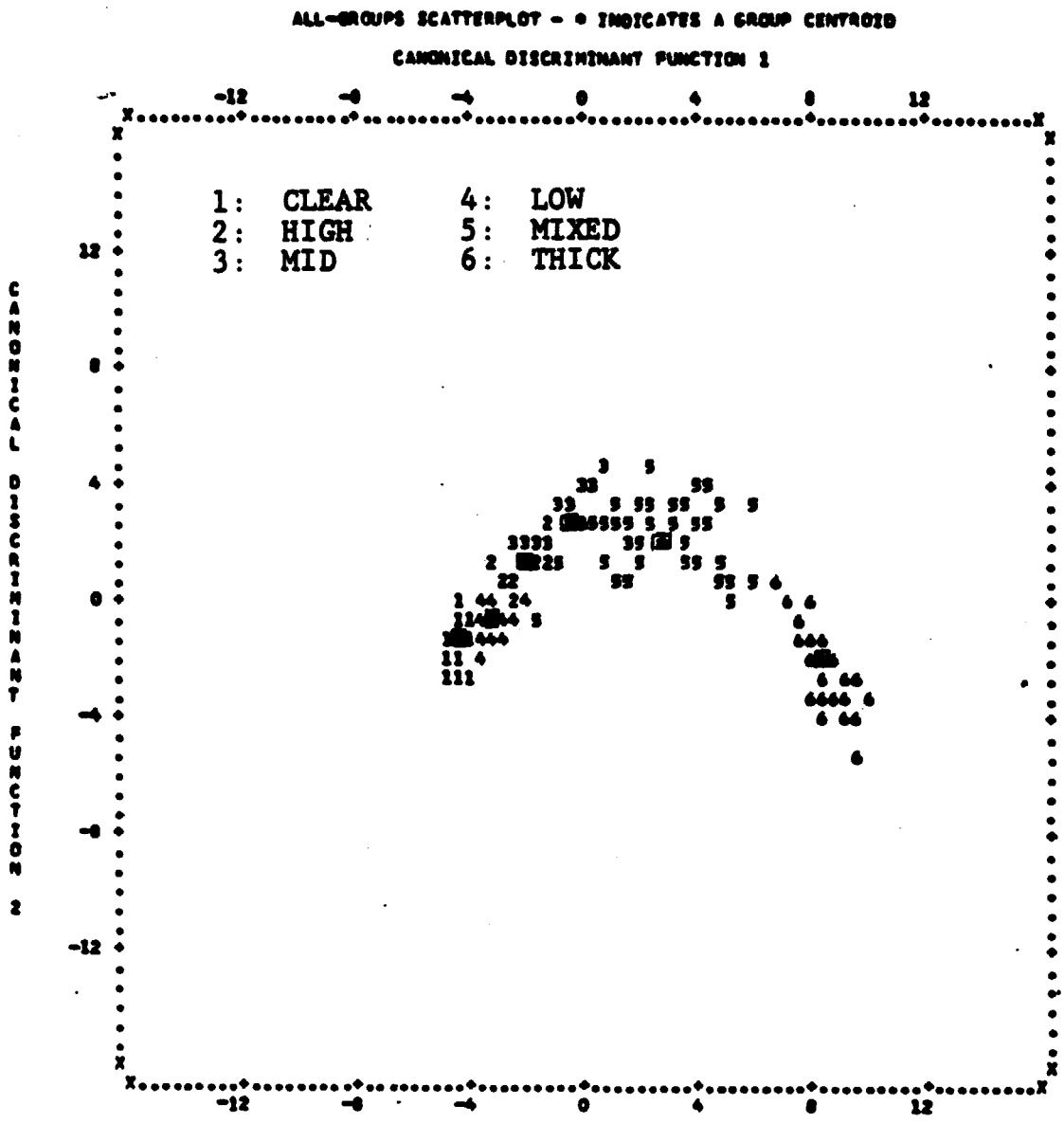


FIGURE A-1 CLOUD CLASSIFICATION DISTRIBUTION BASED ON HIRS DATA FOR FIRST TWO DISCRIMINANT FUNCTIONS .

from 0030-0215 GMT, 30 June 1980, over eastern North America and the coastal Atlantic. A complete distribution of predicted cloud types is shown in the printout displayed in Fig. A-2. There is little high-quality ground truth available for checking the predicted spatial distribution of types. Some spot checks are possible from Northern Hemispheric surface charts (not reproduced here). Several such reports, although not strictly coincident with HIRS data, substantiated significant classifications. Figure A-3 is a copy of a visible print from GOES imagery from 2315Z, 29 June 1980, approximately one and one-half hours prior to the satellite swath over the Eastern United States (0030 to 0045Z, 30 June 80). Later satellite imagery applies after sunset over the Eastern portion of the area. Several features apparent in the satellite imagery are well represented in Figure A-2. The large area of thunderstorms over northern Mississippi and Alabama is evident as a cluster of thick clouds (category 6) in Figure A-2. The precipitation producing clouds over New York extending into New England are clearly described. The large swirl of mostly low cloudiness over the Great Lakes is well represented as is the large clear area in the Ohio Valley. The clear area or "dry tongue" extending from Western New York southwestward into Kentucky can be seen readily. Although other features are difficult to discern in the illustrated copy of the GOES image, a qualitative evaluation of the cloud type estimation model using the original print indicates considerable success. No doubt improvement would result from the use of additional multispectral parameters. Therefore, this method should be fully developed for application.

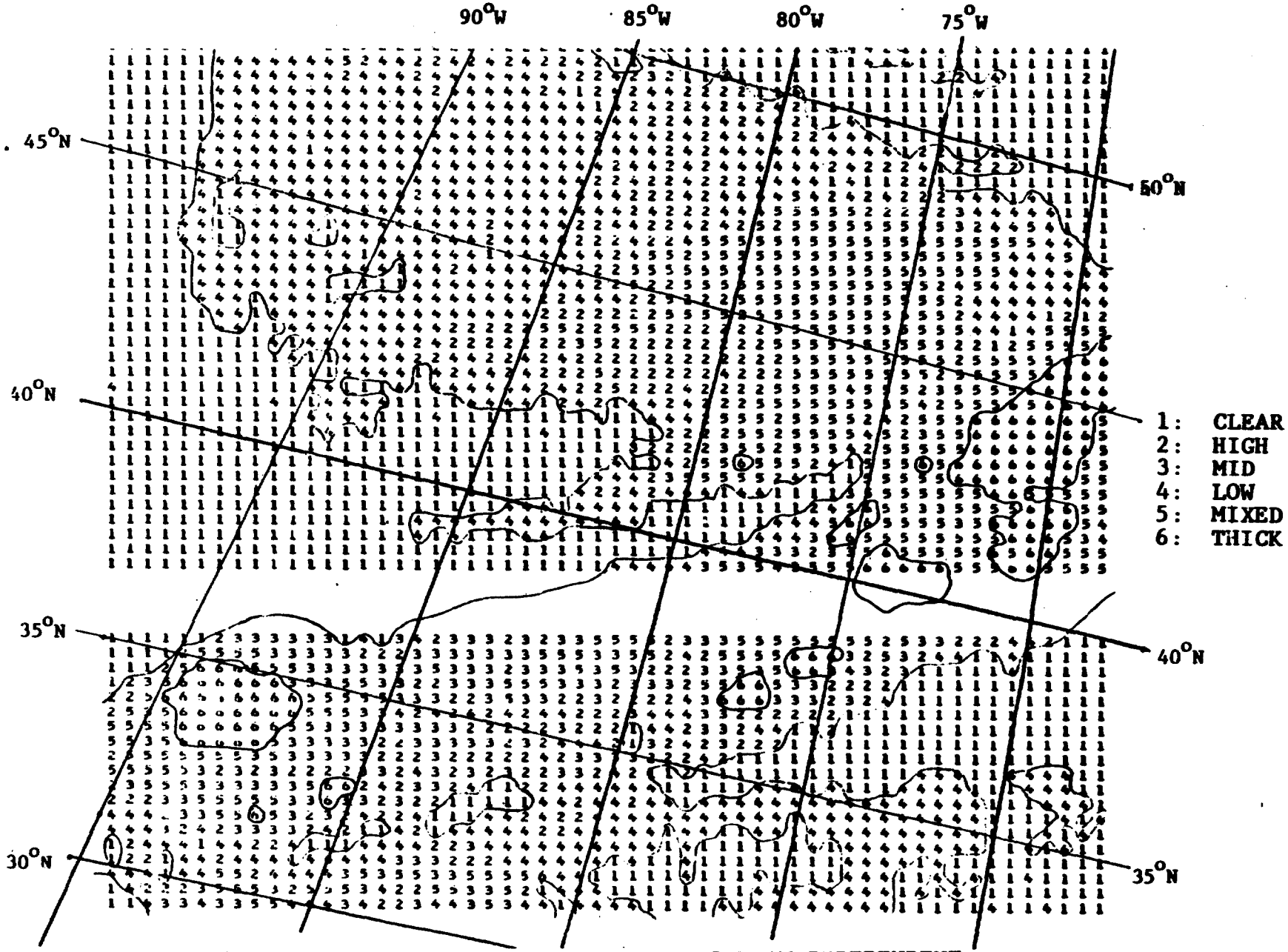
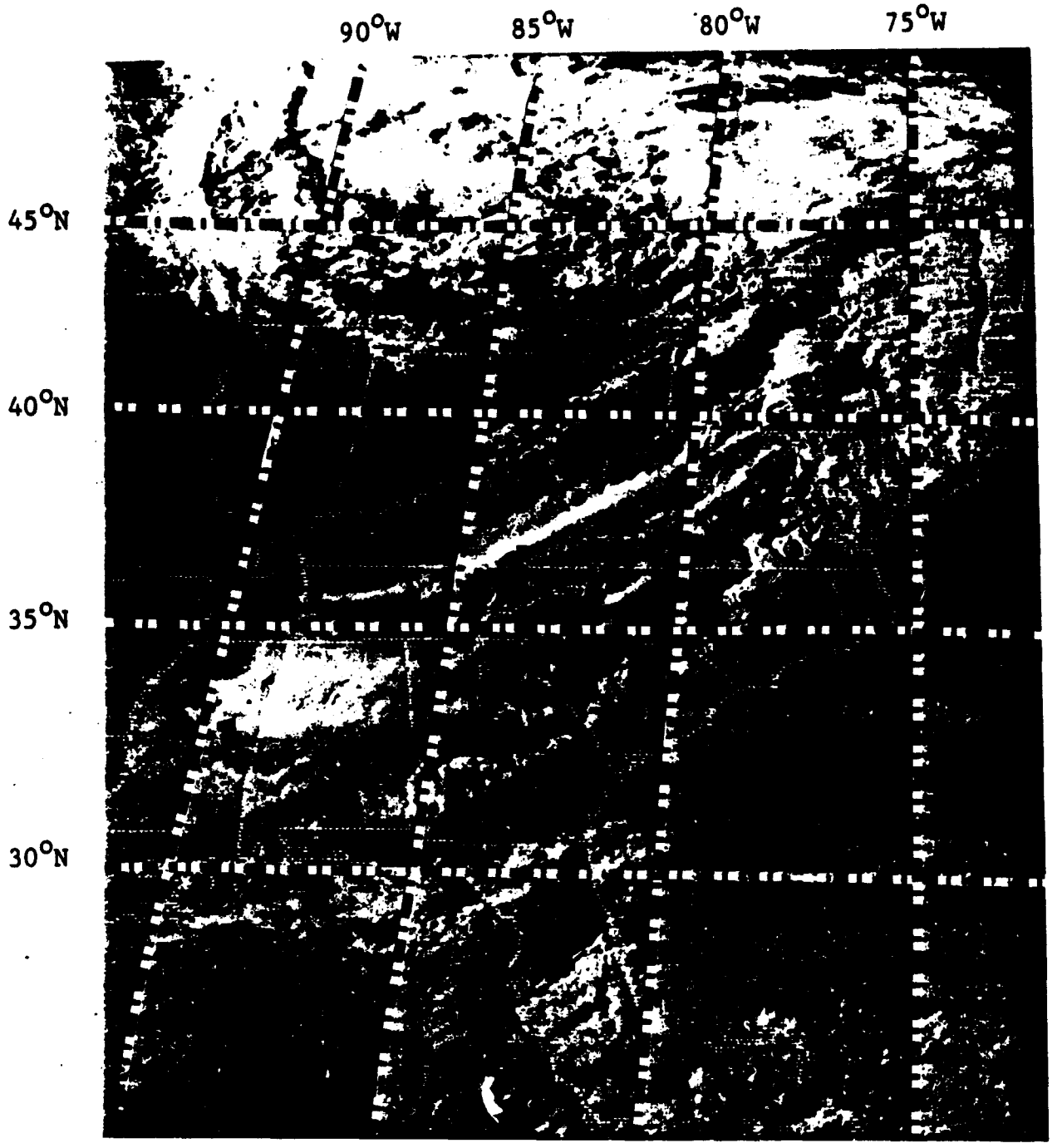


FIGURE A-2 ESTIMATED CLOUD CATEGORIES FROM INDEPENDENT HIRS SCANSPOTS, 29 JUNE 1980



ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

FIGURE A-3 COPY OF GOES VISIBLE IMAGE NEAR SUNSET, 29 JUNE 1980

NASA-JSC