Assessment of the Global Monthly Mean Surface Insolation Estimated from Satellite Measurements Using Global Energy Balance Archive Data

ZHANQING LI
Canada Centre for Remote Sensing, Ottawa, Ontario, Canada

CHARLES H. WHITLOCK AND THOMAS P. CHARLOCK
NASA Langley Research Center, Hampton, Virginia

(Manuscript received 26 October 1993, in final form 12 May 1994)

ABSTRACT

Global datasets of surface radiation budget (SRB) have been obtained from satellite programs. These satellite-based estimates need validation with ground-truth observations. This study validates the estimates of monthly mean surface insolation contained in two satellite-based SRB datasets with the surface measurements made at worldwide radiation stations from the Global Energy Balance Archive (GEBA). One dataset was developed from the Earth Radiation Budget Experiment (ERBE) using the algorithm of Li et al. (ERBE/SRB), and the other from the International Satellite Cloud Climatology Project (ISCCP) using the algorithm of Pinker and Laszlo and that of Staylor (GEWEX/SRB). Since the ERBE/SRB data contain the surface net solar radiation only, the values of surface insolation were derived by making use of the surface albedo data contained in the GEWEX/SRB product. The resulting surface insolation has a bias error near zero and a root-mean-square error (RMSE) between 8 and 28 W m^{-2}. The RMSE is mainly associated with poor representation of surface observations within a grid cell. When the number of surface observations are sufficient, the random error is estimated to be about 5 W m^{-2} with present satellite-based estimates. In addition to demonstrating the strength of the retrieving method, the small random error demonstrates how well the ERBE derives the monthly mean fluxes at the top of the atmosphere (TOA). A larger scatter is found for the comparison of transmissivity than for that of insolation. Month to month comparison of insolation reveals a weak seasonal trend in bias error with an amplitude of about 3 W m^{-2}. As for the insolation data from the GEWEX/SRB, larger bias errors of 5–10 W m^{-2} are evident with stronger seasonal trends and almost identical RMSEs.

1. Introduction

As the major component of surface heat balance, solar radiative flux has a large impact on the energy exchange between the surface and the atmosphere. Although surface solar radiation has been measured on the ground for a long time, the measurements are mostly over inhabited areas. The fact that there are few observations over ocean and remote regions and that the observations are nonuniform over land has posed some problems for climate studies, especially for climate modeling. Since the 1980s, substantial progress has been made in the retrieval of the surface radiation budget (SRB) from satellite measurements, especially for the shortwave (SW) components, as summarized in Schmetz (1989) and Pinker et al. (1995). In order to take advantage of the global coverage offered by satellite observations, the World Climate Research Program (WCRP) recommended the development of a satellite-based global climatology of the SRB with an accuracy of 10 W m^{-2} for monthly mean fluxes in 250 $\times$ 250 km^2 grids (Suttles and Ohring 1986).

To meet this goal, the Global Energy and Water Cycle Experiment (GEWEX) SRB Project was established to retrieve fluxes of the SRB over the globe from operational satellite measurements in the late 1980s. A number of satellite-based SW algorithms were tested using the data collected from field experiments and global surface observations (Whitlock et al. 1990). Two algorithms, namely, Pinker and Laszlo (1992; hereafter referred to as the Pinker algorithm) and Staylor (Darnell et al. 1992), were initially selected for deriving the global shortwave SRB from the International Satellite Cloud Climatology Project (ISCCP) C1 data (Whitlock et al. 1993). ISCCP provides more than 8 yr of cloud retrieval and radiance observation at high spatial and temporal resolution (Schiffer and Rossow 1983; Rossow et al. 1991). While ISCCP satellite radiance measurements are the principal input data for the GEWEX SRB Project, other datasets were also employed. For example, the albedo data used in the Staylor algorithm were estimated from the Earth Radiation Budget Ex-

Corresponding author address: Dr. Zhanqing Li, Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Ontario K1A 0Y7, Canada.

periment (ERBE). The first official release of the GEWEX/SRB (Version 1.1) was made recently by the NASA Langley Research Center for WCRP (Charlock et al. 1993). The data include surface downward irradiance (insolation), albedo, and net solar radiation for the period March 1985 through December 1988.

In addition to the GEWEX/SRB dataset, global climatology of monthly mean surface net solar radiation was also developed from monthly mean ERBE TOA fluxes (Li and Leighton 1993) using the algorithm of Li et al. (1993a; hereafter referred to as the Li algorithm). The ERBE measurements were made from late 1984 to early 1990. Although the ERBE data cover a shorter period than ISCCP data, the former are well-calibrated broadband measurements. The narrowband measurements from the ISCCP lack onboard calibration. A potential source of error, namely, narrowband to broadband conversion (Li and Leighton 1992), is therefore eliminated when ERBE data are used. Moreover, the ERBE flux retrievals were corrected for bidirectional effects. Note that the ERBE/SRB of Li and Leighton (1993) contains surface net solar radiation only.

By the nature of retrieval, these satellite-based estimates need rigorous quality evaluation. The most straightforward evaluation approach is to compare the estimates with in situ measurements. So far, among the components of solar radiation, only insolation observations may be qualified for such a purpose. The present ground observations of surface solar net flux and albedo are very limited, and their spatial representation is poor. This paper presents results of the comparison of satellite-based estimates of monthly mean surface insolation with ground-truth observations from the Global Energy Balance Archive (GEBA). Since the ERBE/SRB does not include insolation, surface albedo from the GEWEX/SRB is combined with net solar radiation from ERBE/SRB to produce the ERBE/GEWEX insolation. While insolation data from both datasets are validated, the emphasis is placed on the ERBE/GEWEX data, as preliminary results of validation for the GEWEX/SRB have been reported (DiPasquale and Whitlock 1993). Note that the available global satellite-based SRB datasets at present are not limited to the above mentioned sets. For example, Bishop and Rossow (1991) and Breon et al. (1994) also derived global surface solar irradiance by applying their algorithms to the ISCCP data and the ERBE data, respectively.

2. Data and algorithm

b. GEWEX/SRB (Version 1.1)

The main dataset employed in the generation of the GEWEX/SRB is the ISCCP C1, a subset of ISCCP data at a reduced spatial resolution. The ISCCP C1 gridded system has 6596 equal-area grids with a cell size of approximately 280 × 280 km^2. The ISCCP data were measured by radiometers with nominal resolutions of 4 to 8 km. Within each 280 × 280 km^2 equivalent area of an ISCCP C1 grid cell, the pixels were sampled every 25 km. A GEWEX/SRB estimate was therefore based on satellite data that cover a fraction of the area of each grid cell. ISCCP C1 data contain 3-hourly and daily values of 132 parameters. While about 20 of these parameters were used in the estimation of the SRB, the major input is narrowband shortwave radiance taken from many operational weather satellites including polar-orbiting satellites (NOAA series) and geostationary satellites (GOES, GMS, and Meteosat). Therefore, the data have good spatial and temporal coverage around the earth throughout the day. Due to the lack of onboard absolute calibration, offline relative calibration was implemented with reference to the AVHRR measurements of *NOAA-7*.

The GEWEX/SRB of version 1.1 includes two sets of data produced by the algorithms of Pinker and Staylor. Three basic steps are followed for both algorithms. First, relationships are established between the broadband (0.2–4.0 μm) transmissivity and the reflectivity at the TOA under various conditions pertaining to the surface, atmosphere, and cloud. Second, parameters characterizing these conditions are obtained, such as surface albedo and optical properties of the atmosphere and clouds. Third, transmissivity is determined from these parameters and relationships. The major difference between the Pinker and Staylor algorithms lies in the way the relationships are established. Pinker's algorithm makes use of detailed radiative transfer calculations, generates a surrogate TOA broadband flux from narrowband radiance, and relies heavily on the ISCCP calibration. Staylor's algorithm employs simplified parameterizations and is partly self-calibrating. Both use H_2O and O_3 data from the ISCCP TOVS record. Since both algorithms are quite nonlinear with respect to input parameters, they are not applicable to temporally (e.g., monthly) averaged input data. The monthly mean insolation data of the GEWEX/SRB were obtained from the daily mean values that were in turn based on retrievals that used 3-hourly satellite radiance data.

In the development of the GEWEX/SRB, surface albedo serves as both input data for retrieving surface insolation and output data for computing surface net solar radiation. Global datasets of surface albedo used in these algorithms were first derived from clear-sky planetary albedo and then corrected for the effect of clouds. Pinker's albedos are based on narrowband ISCCP data (Pinker and Laszlo 1992), while Staylor's are from broadband measurements of the ERBE (Staylor and Wilber 1990). To derive broadband albedos from narrowband albedos, the Pinker algorithm incorporates the spectral dependence model of Briegleb et al. (1986). The effect of clouds on the Pinker surface albedos is accounted for through radiative transfer calculations. All-sky Staylor surface albedos are computed from the albedos for clear sky and for overcast sky by weighting them according to daily mean transmissivity as a measure of cloudiness. The daily overcast albedos for ocean are set to be 6.5%, and those for land are calculated from clear-sky albedos and the daily mean cosine of solar zenith angle (SZA).

b. ERBE/SRB and ERBE/GEWEX

As mentioned in the introduction, the ERBE/SRB data were derived from the monthly mean ERBE S-4 product. The ERBE broadband radiometers were aboard three spacecrafts, including two NOAA satellites and one Earth Radiation Budget Satellite (ERBS). At times, however, measurements were made from only two spacecrafts, as the two NOAA satellites had a short overlapping observation from October 1986 to January 1987. The polar-orbiting NOAA satellites provide measurements at about the same local solar times. The inclined orbit of the ERBS, however, allows observations at different local times over a period of 36 days. The number of the ERBE shortwave measurements per day made over the same target varies from 1 to 5. As the sampling is limited in time, the diurnal variability of the earth's radiation budget was modeled to derive daily means from the instantaneous values using a temporal-averaging scheme (Brooks et al. 1986). A set of angular dependence models (ADMs) were employed to account for the variation of planetary albedo with viewing angles and SZA for various scene types (Suttles et al. 1989). The uncertainties in the ERBE S-4 data were estimated to be on the order of 5 W m^{-2} (Barkstrom et al. 1989). The number is not a definitive estimate of the uncertainty, however, but simply the standard deviation within which the true value of a measurement might occur.

The Li algorithm directly estimates surface net solar flux from a TOA-reflected flux. The algorithm comprises linear relationships between the two quantities, with its slope and offset being functions of precipitable water and SZA. The algorithm is based on the results of comprehensive radiative transfer simulations for a variety of surface, cloud, and atmospheric conditions (Li et al. 1993a). Although different sets of coefficients were obtained for clear skies and various cloud types, sensitivity tests show weak dependency on sky condition. Using both the matched ERBE pixel data (S-8) and tower measurements, it was demonstrated that the coefficients for clear skies can be equally applicable to cloudy conditions, regardless of surface type (Li et al. 1993b). Therefore, the clear-sky coefficients were employed to generate the ERBE/SRB data (Li and Leighton 1993). It is partly the objective of the this study to further validate this by comparing the resulting SRB estimates with ground-truth observations. Since the algorithm is linear with respect to the TOA-reflected flux, it is applicable to the monthly mean TOA data. It was proven analytically (Li and Leighton 1993) and experimentally (Li et al. 1993b) that the algorithm is applicable to the temporally means of the cosine of SZA and precipitable water. Thus, monthly mean surface net solar radiation data were derived from monthly mean values of the TOA-reflected flux, precipitable water, and cosine of SZA (Li and Leighton 1993). The TOA data were from ERBE S-4; precipitable water data were from ECMWF humidity analyses; and the cosines of SZA were calculated according to date and latitude. Note that both the ECMWF and TOVS humidity data may contain systematic errors in some regions (Liu et al. 1992).

Therefore, assessment of monthly mean estimates has two aspects of significance. One is to further test the validity of the algorithm to estimate monthly mean surface net solar radiation from monthly mean input data using a set of coefficients derived for clear-sky conditions. The other is to test how well the ERBE derives monthly mean TOA fluxes from limited radiance measurements by applying an averaging scheme to account for diurnal variations and a set of ADMs to correct for angular effects. The tests are especially important, considering that the Li algorithm, the ERBE-averaging scheme, and the ERBE ADMs will be used to derive monthly mean surface solar fluxes in the Clouds and Earth's Radiant Energy System (CERES), an important subsystem of the Earth Observing System (EOS) (Wielicki et al. 1994).

To derive surface insolation from the estimated net solar radiation, a global dataset of surface albedo is required. In addition to the two surface albedo datasets contained in the GEWEX/SRB, such a dataset was also developed by Li and Garand (1994) from the clear-sky ERBE TOA measurements. The data are not corrected to account for the effect of clouds, however. Good comparison is found between the clear-sky surface albedos obtained by Staylor and Wilber (1990) and those by Li and Garand (1994), despite the different retrieving methods used. In contrast to Pinker's surface albedo, the determination of Staylor's does not require spectral correction, as they were based on the broadband ERBE data. Therefore, the Staylor's surface albedo was used to convert the ERBE net solar radiation into insolation. The resulting ERBE/GEWEX insolation will be compared with surface insolation measurements. It should be noted that surface albedo has a direct impact on the result of validation for the ERBE/GEWEX insolation data. The effect of surface albedo on validation of the GEWEX/SRB insolation is secondary, since surface albedo is used as a boundary condition to derive insolation. Unfortunately, no validation is done to evaluate the quality of the satellite-based surface albedos due to the lack of ground-truth measurements for areal mean surface albedo. Selection of the Staylor surface albedo data stems mainly from the considerations discussed above.

The principal satellite datasets and algorithms used here for retrieving the shortwave components of SRB

TABLE 1. Satellite products, input datasets, and algorithms (in parentheses) related to this study.

Products	Albedo	Insolation	Net
GEWEX/SRB (Pinker)	ISCCP (Pinker)	ISCCP (Pinker)	ISCCP (Pinker)
GEWEX/SRB (Staylor)	ERBE (Staylor)	ISCCP (Staylor)	ISCCP + ERBE (Staylor)
ERBE/GEWEX (Li/Staylor)	ERBE (Staylor)	ERBE (Li + Staylor)	ERBE (Li)

are summarized in Table 1. For most grid cells, the ISCCP sampling is more dense temporally, while the ERBE sampling is more dense spatially.

c. GEBA

The GEBA is a worldwide database containing surface-based measurements of energy fluxes. It was prepared by Ohmura and Gilgen (1991) at the Swiss Federal Institute of Technology under a project of the World Climate Program-Water. The GEBA contains 150 000 station months worth of data for up to 1600 sites. The main sources of data are periodicals, monographs, data reports, unpublished data, and the World Radiation Data Center (WRDC) at St. Petersburg, Russia, where worldwide surface radiation observations have been archived. After data are extracted, they are digitized (if necessary), standardized, and formalized. Quality control is then applied to the data in order to ascertain its integrity and consistency. Monthly mean fluxes for each year are archived together with information pertaining to location, altitude, geographical region and country, surface characteristics, observation period and method, instrument, and data source.

The solar radiation data of GEBA include direct and diffuse downward irradiance (insolation) and reflected irradiance (albedo). Insolation has been measured most frequently and widely (Hunt et al. 1986). Currently, there exists more than 1500 stations observing insolation worldwide, 600 of which have continuous measurements for more than 10 years. The accuracy of the operational observation of instantaneous value of surface insolation is generally not better than 3% (Schmetz 1989). Surface albedo is usually the most poorly observed component. Although surface albedo observations are made at many meteorological stations, their areal representation is limited, as the stations are often covered with short grass (Ohmura and Gilgen 1993). This makes surface-based albedo measurements of limited utility for this study. Validation is, therefore, restricted to surface insolation.

The insolation data used in the study are those considered as being accurate from the GEBA (GEBA category 88) during the period for which satellite SRB data exist. The data in category 88 were subject to quality control to ascertain that they are 1) "physically possible," falling within certain upper and lower bounds, and 2) "physical probable," remaining consistent with historical data from the same region for the same month. The GEBA data involved in this comparison are not in the original format as described in Ohmura and Gilgen (1991) but are given in ISCCP C1 grids. Based on site coordinates, the GEBA data are assigned into ISCCP cells. All the GEBA measurements falling within an ISCCP cell are averaged and the mean value is assumed to represent the grid-mean flux. The GEBA data are regridded at the NASA Langley Research Center (Whitlock et al. 1993) and archived in the GEWEX/SRB product for comparison with the satellite-based estimate of the SRB. To distinguish the gridded GEBA SRB data from the original site-based GEBA data, the former is referred to as the GEBA-ISCCP data, and the ISCCP cells that contain at least one qualified GEBA site are referred to as the GEBA-ISCCP cells.

Figure 1 shows the geographical distribution of the GEBA-ISCCP cells and the numbers of surface sites per ISCCP cell for June 1986. Note that the distribution is not uniform; the majority of the stations are on land. The most dense observation network is found in western Europe. Serious data shortages exist in parts of Latin America, the United States, Arabian countries, and East Asia. Of the GEBA-ISCCP cells, approximately 75% contain only one GEBA site. There are three cells in Germany that contain seven or more sites. It is noteworthy that the number of GEBA-ISCCP cells and the number of surface sites located in a cell do not change very much throughout the period under study. The surface stations that operate during this period, however, may differ considerably from those that were in operation, as shown in Fig. 1 of Ohmura and Gilgen (1993).

3. Validation

The validation employs nine months of data for which all three datasets were available: the GEBA-ISCCP, the ERBE/GEWEX, and the GEWEX/SRB. They include April, July, and October of 1985; January, April, July, and October of 1986; and January and April of 1987. The comparison is made for each of five combinations of the GEBA data that were classified based on the number of surface sites per cell and data quality (Whitlock et al. 1993). They are 1) the total global GEBA set that includes the data from all the GEBA-ISCCP cells; 2) the GEBA subset that contains the data from all the GEBA-ISCCP cells except those situated in mountains, at high latitudes, and at potentially foggy coastal sites; 3) the multisite GEBA subset composed of the entire multisite GEBA subset; 4) the

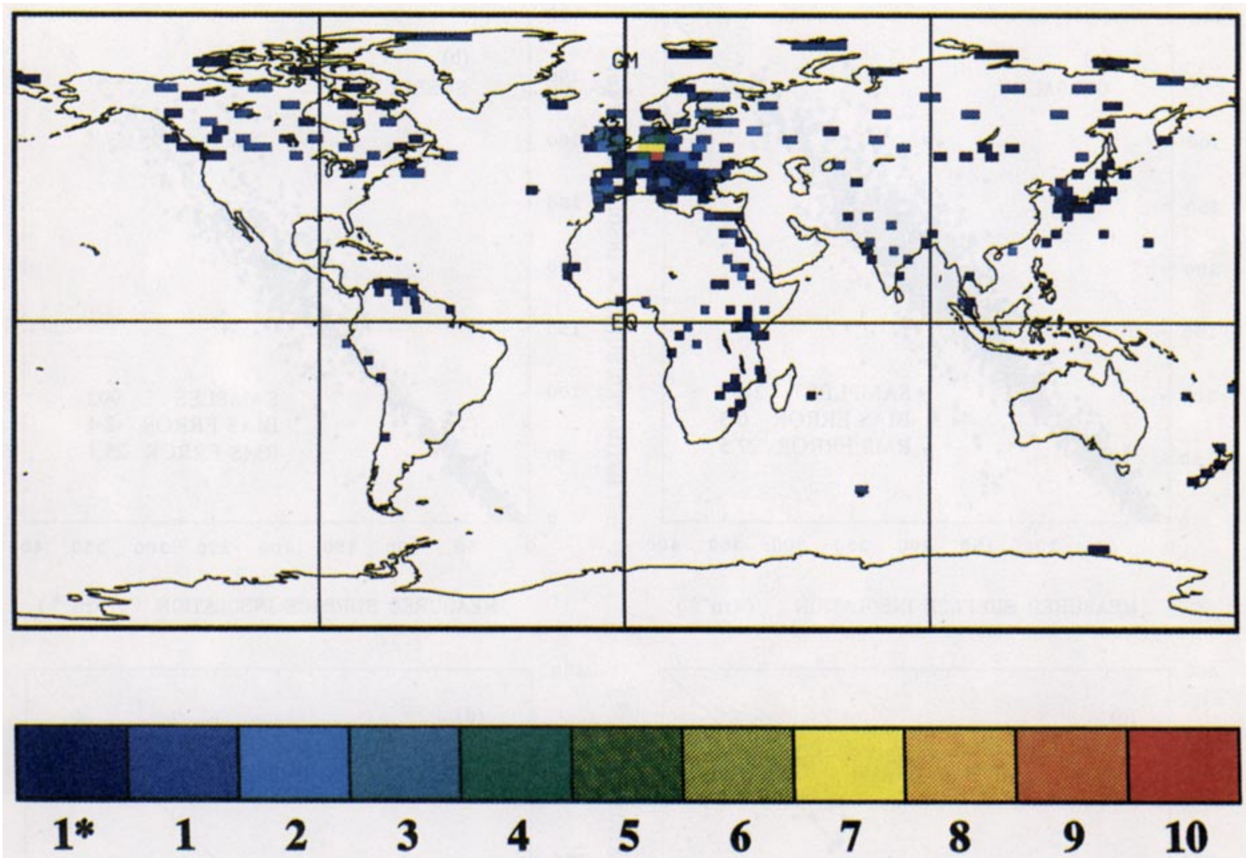

FIG. 1. Geographical distribution of the ISCCP C1 cells in which there is at least one surface site measuring surface insolation in the month of June 1986. Different colors represent various numbers of surface sites situated in an ISCCP cell. The cells denoted by 1* are located either in mountains, at high latitudes, or at potentially foggy coastal sites.

European multisite GEBA subset that contains multisite European data only; and 5) the German subset that consists of the data from the ISCCP cells with at least seven GEBA sites located in the former West Germany. The typical numbers of GEBA-ISCCP cells for one month in the categories 1 through 5 are 250, 100, 25, 20, and 3, respectively.

a. Overall comparison

Figure 2 shows the ERBE/GEWEX insolation versus the GEBA insolation intercomparisons for four of the five combinations. The comparison for the European multisite subset is not included because it is very similar the multisite subset. Note that the monthly mean values of the ERBE/GEWEX surface insolation agree well with the surface observations from the GEBA-ISCCP data. The four comparisons show the bias errors of 0.3, −2.4, −2.4, and 0.8 W m^{-2}, and the root-mean-square errors (RMSEs) of 27.5, 25.1, 14.9, and 8.1 W m^{-2} for the four datasets, respectively. Dashed lines in Fig. 2 represent least-square linear regression relationships between observed and estimated insolation with intercepts, slopes, and correlation coefficients presented in Table 2. Note that the perfect agreement denoted by 1:1 solid line would have an intercept of 0, slope of 1.0, and correlation coefficient of 1.0. It appears that the regression lines deviate slightly from 1:1 line and the bias error depends slightly on surface insolation. It tends to be negative when insolation is small and positive when insolation is high.

Because of the assumption that cell-averaged flux is equal to the mean of the fluxes measured at all the sites within a cell, the highest confidence should be given to the comparison with the German data, which exhibits extraordinarily good agreement. It should be stated that for this comparison, an apparent incorrect number for the surface albedo was identified and corrected in the month of April 1986 at cell 5733, which corresponds to a latitude of 48.75°N and longitude of 9.47°E. The original albedo is 0.5508, which is inconsistent with results for other years. Values for April 1985 and April 1987 are 0.1432 and 0.1444, respectively. The 1986 value was based on an automated gap-filling interpolation procedure, because clear-sky ERBE data were missing for April 1986 in this region. The interpolation apparently produced an excessively high value. The albedo value is therefore corrected by re-

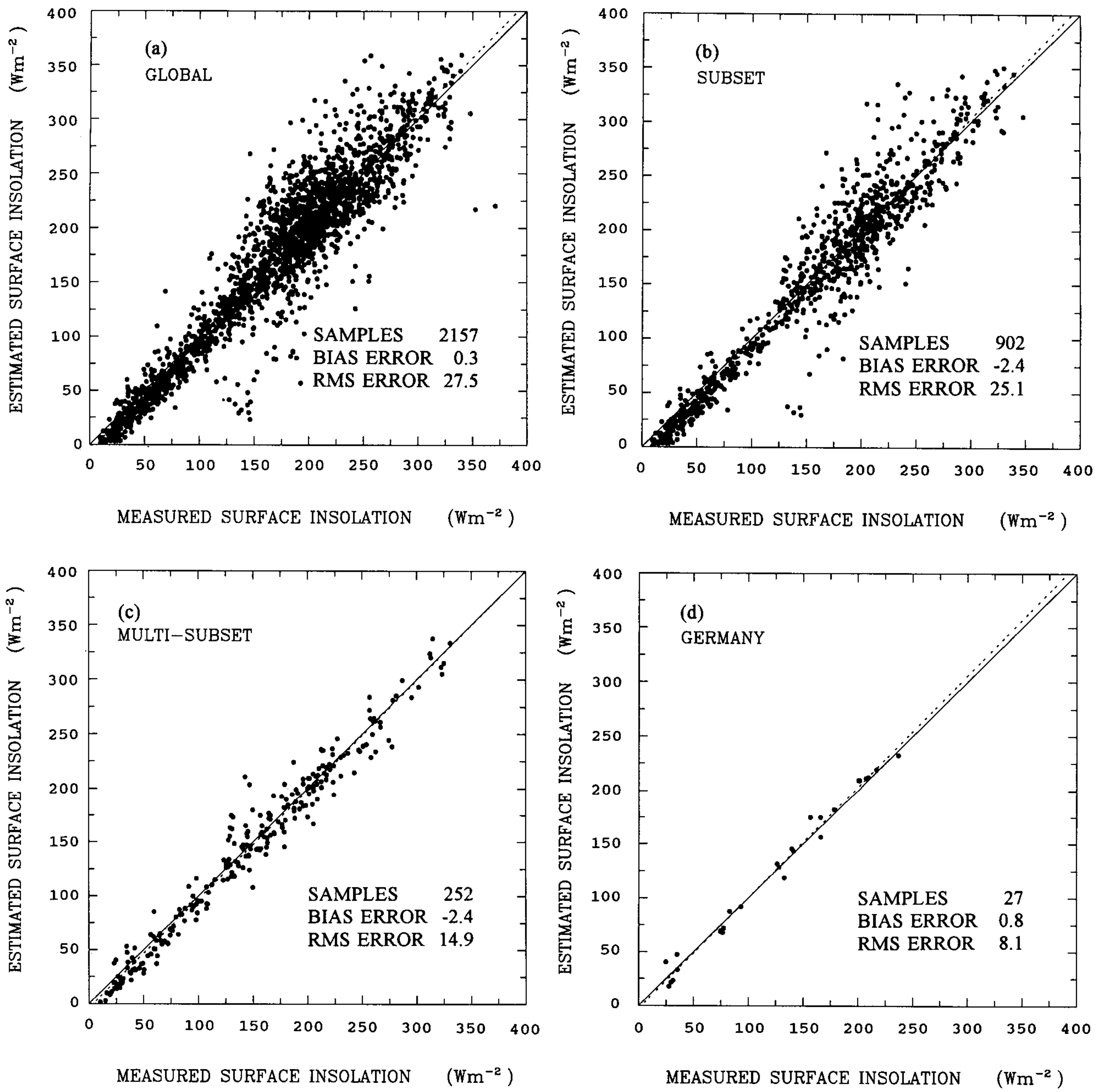

FIG. 2. Comparison of the ERBE/GEWEX surface insolation derived from ERBE satellite data with the insolation measured at the surface radiation network from the Global Energy Balance Archive (GEBA). ERBE/GEWEX insolation is obtained by combining the surface net solar flux and albedo estimated using the algorithms of Li et al. and Staylor, respectively. Comparisons are made for (a) the total global GEBA dataset; (b) GEBA subset; (c) multisite GEBA subset; and (d) German subset. Solid and dashed lines correspond to 1:1 line and regression relationship between estimated and observed insolation, respectively.

placing it with the mean of 0.1432 and 0.1444. Likewise, it is possible that some of the points appearing in Fig. 2a that deviate considerably from 1:1 line may be due to the use of incorrect surface albedos. In addition to errors in interpolation for missing clear-sky measurements, incorrect surface albedo may also result from misidentification of clear pixels, inability to account for spatial and temporal variations of surface albedos, and errors in the inversion algorithm. In fact, the surface albedos derived from the clear-sky ERBE measurements over snow–ice surfaces at high latitudes are questionable due to the problem of the ERBE scene identification (Li and Leighton 1991).

If the comparison for the German dataset is typical, the uncertainty of the satellite-based estimates of monthly mean shortwave radiative flux over the grid

TABLE 2. Values of intercept, slope, and correlation coefficient for the linear regression between observed insolation (response) and estimated ERBE/GEWEX insolation (variable) for four GEBA datasets.

Data categories	Intercept	Slope	Correlation coefficient
Global	−5.74	1.04	0.95
SUBSET	−9.44	1.04	0.96
MULTI-SITE	−3.89	1.01	0.98
GERMAN	−2.65	1.03	0.99

cells may be comparable to that of surface-based observations. Despite relatively poor representation of surface measurements in all datasets except Germany, bias errors for these comparisons are of a certain significance, as they are derived from a large number of samples. The uncertainty of the bias error decreases with the number of samples. As will be demonstrated later, the RMSE is very sensitive to the problem of geographical representation of a surface observation. To the extent that the RMSE is due to the inadequate areal representation, the value of the RMSE does not indicate the random error of a satellite-based estimate. Under this circumstance, the bias error of the comparison is more meaningful than the RMSE, unless the surface insolation data or the surface albedo data are biased. There seems to be no reason to believe that the global surface observations of insolation from the GEBA are biased, except that the spectral interval of the surface pyranometer differ slightly from that of the shortwave radiometer of the ERBE. The nominal interval of the ERBE radiometer lies between 0.2 and 5.0 μm, while ground-truth measurements are for the wavelength ranging from 0.3 to 3.0 μm. Depending on cloud thickness, a positive bias between 0% and 1% may be expected (Whitlock et al. 1993), which is approximately equivalent to 0 to 2 W m^{-2} for the present comparison. Due to the severe shortage for ground-truth data of surface albedo, it is not known whether the surface albedo data used are biased.

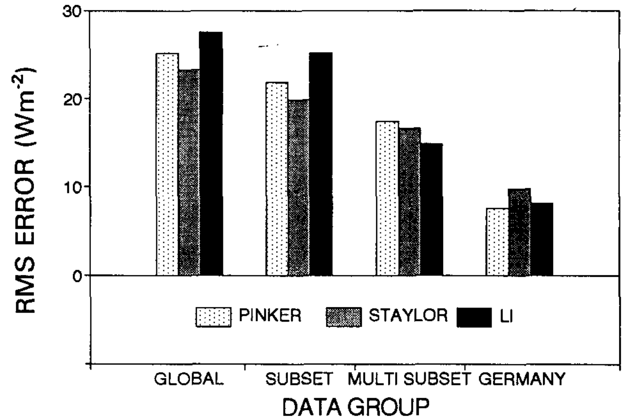

FIG. 4. Same as Fig. 3 but for rms error.

Likewise, the GEWEX/SRB insolation products were also validated against the same ground-truth data. Figures 3 and 4 present the bias and RMSEs for four subsets of insolation from each of the two GEWEX/SRB products, compared to those of the ERBE/GEWEX insolation data. Estimates of surface insolation from the GEWEX/SRB products are systematically higher than those observed at the surface from the GEBA data by 5.1–10.1 W m^{-2}, as compared to −2.4–0.8 W m^{-2} for estimates from the ERBE/GEWEX insolation data. On the other hand, the RMSEs for insolation estimates from the four datasets are of the same order of magnitude. Linear regression analyses were also conducted between the estimated insolation from the GEWEX/SRB and the observed insolation from the GEBA. The values of intercept and slope of regression equations and the correlation coefficients are presented in Table 3 for the estimates obtained using Pinker algorithm and in Table 4 for those using Staylor algorithm. Comparing these tables to Table 2 shows that the values of slope are closer to unity for the GEWEX/SRB insolation than for the ERBE/GEWEX insolation. Overall, the magnitudes of the intercept given in Tables 3 and 4 are larger than those presented in Table 2. These findings further demon-

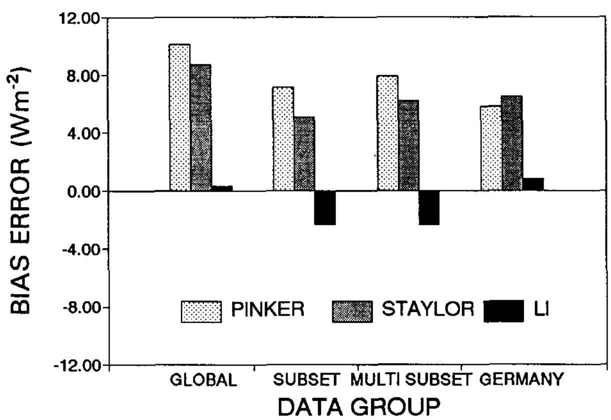

FIG. 3. Bias errors in the satellite estimates of surface insolation that were obtained using the retrieving algorithms of Li, Pinker, and Staylor. Note that the surface albedos of Staylor are employed in the estimation of surface insolation by Li.

TABLE 3. Same as Table 2, but for the estimates of GEWEX/SRB surface insolation obtained using the Pinker and Laszlo algorithm.

Data	Intercept	Slope	Correlation coefficient
Global	10.80	1.00	0.95
Subset	6.14	1.01	0.96
Multisite	9.75	0.99	0.98
Germany	14.45	0.93	0.99

TABLE 4. Same as Table 3, except using the Staylor algorithm.

Data	Intercept	Slope	Correlation coefficient
Global	11.02	0.99	0.96
Subset	6.34	0.99	0.97
Multisite	9.81	0.98	0.98
Germany	8.42	0.98	0.99

strate that the GEWEX/SRB insolation products contain systematic error; that is, the bias errors have little to do with the magnitude of surface insolation. Recently, Li (1995) found that the systematic errors in the GEWEX/SRB data are due to the use of the Lacis and Hansen (1974; hereafter LH) parameterization scheme to compute water vapor absorption by Pinker and Staylor. According to the studies of Kratz and Cess (1985) and Ramaswamy and Freidenrecich (1992), the LH scheme produces solar water vapor absorptions that are too low. For example, the difference between LH and line by line results amounts to 15.8 W m^{-2} for the midlatitude summer atmosphere at solar zenith angle of 30° with a surface albedo of 0.2 (Ramaswamy and Freidenrecich 1992). It is interesting to note that the correlation coefficients given in Tables 2, 3, and 4 for the same data category are almost identical. This finding, together with the finding that the RMSEs are similar for the ERBE/GEWEX insolation and the GEWEX/SRB insolation data, may indicate that the cause of random errors of the three satellite-based products is similar. It has little to do with the physical aspects of the three algorithms that bear little resemblance to each other.

b. *Inference of real random error*

In the comparison of the instantaneous estimates of the SRB, Li et al. (1993) argued the RMSE is due mainly to the discrepancy of satellite and surface observations. The argument can be well demonstrated if there are various numbers of surface observations corresponding to a satellite estimate. If the RMSE is, to a large extent, caused by inadequate surface sampling, it should decrease with the increasing number of surface sites. To illustrate this, the GEBA subset data are further classified into four categories based solely on the number of surface sites that are equal to 1, 2, 3, and 4 or more, respectively. It follows from the comparisons shown in Fig. 5 that the scatter of the points diminishes monotonously as the number of surface sites increases. The bias errors remain small but fluctuate more, owing to fewer samples and thus larger statistical uncertainties. To infer the "real" random error that would result from comparison against ground truth with an unlimited number of surface observations, the RMSE is fitted as a function of the number of surface sites (Fig. 6). Note that the noninteger number of 6.3 is regarded as the effective number of surface sites for the data shown in Fig. 5d. It is obtained from the number of surface sites weighted by the numbers of ISCCP cells that contain the corresponding number of surface sites. Figure 6 shows that the RMSE decreases quickly with the number of surface sites (N), and the variation of RMSE can be well fitted by

$$\mathrm{RMSE} = 4.1 + \frac{24.2}{N}. \quad (1)$$

The equation suggests that the RMSE would be equal to 4.1 W m^{-2} if the number of surface sites in an ISCCP cell were infinite. The estimate of the random error is, of course, subject to statistical uncertainty. For a significance level of 95%, the real random error is estimated to fall within an confidence interval of S_1 and S_2 (Freund 1973),

$$S_1 = \mathrm{RMSE}\bigg/\left(1 + \frac{1.96}{\sqrt{2N}}\right) \quad (2)$$

$$S_2 = \mathrm{RMSE}\bigg/\left(1 - \frac{1.96}{\sqrt{2N}}\right), \quad (3)$$

where N denotes the number of measurements. For each data subset presented in Fig. 5, S_1 and S_2 can be computed according to its RMSE and N. Terms S_1 and S_2 are then fitted against the number of surface sites:

$$S_1 = 2.4 + \frac{24.3}{N} \quad (4)$$

$$S_2 = 6.6 + \frac{23.4}{N}. \quad (5)$$

The multiple correlation coefficients are larger than 0.99 for both fittings. Equations (4) and (5) reveal that the real random error ranges from 2.4 W m^{-2} and 6.6 W m^{-2} for a probability of 95%.

It is therefore concluded that the imperfect spatial and temporal matching accounts for a larger part of the difference between satellite-estimated and surface-observed insolation. This was also demonstrated by Moser and Raschke (1983). The real random error may arise from both algorithm and input data. For the algorithm, the major sources of random error include perturbation due to clouds and aerosol. For the satellite input data, random error may occur in temporal sampling (and spatial sampling for ISCCP C1), angular corrections, etc. The coincidence that the random error obtained here is very close to the uncertainty in the ERBE monthly mean TOA fluxes (Barkstrom et al. 1989) may indicate that the error in input data is the major cause of the random error for the estimation of monthly mean surface solar flux.

Since the RMSE is due both to the real random error of the estimates and to the representative error of surface observations, it would be more appropriate to call

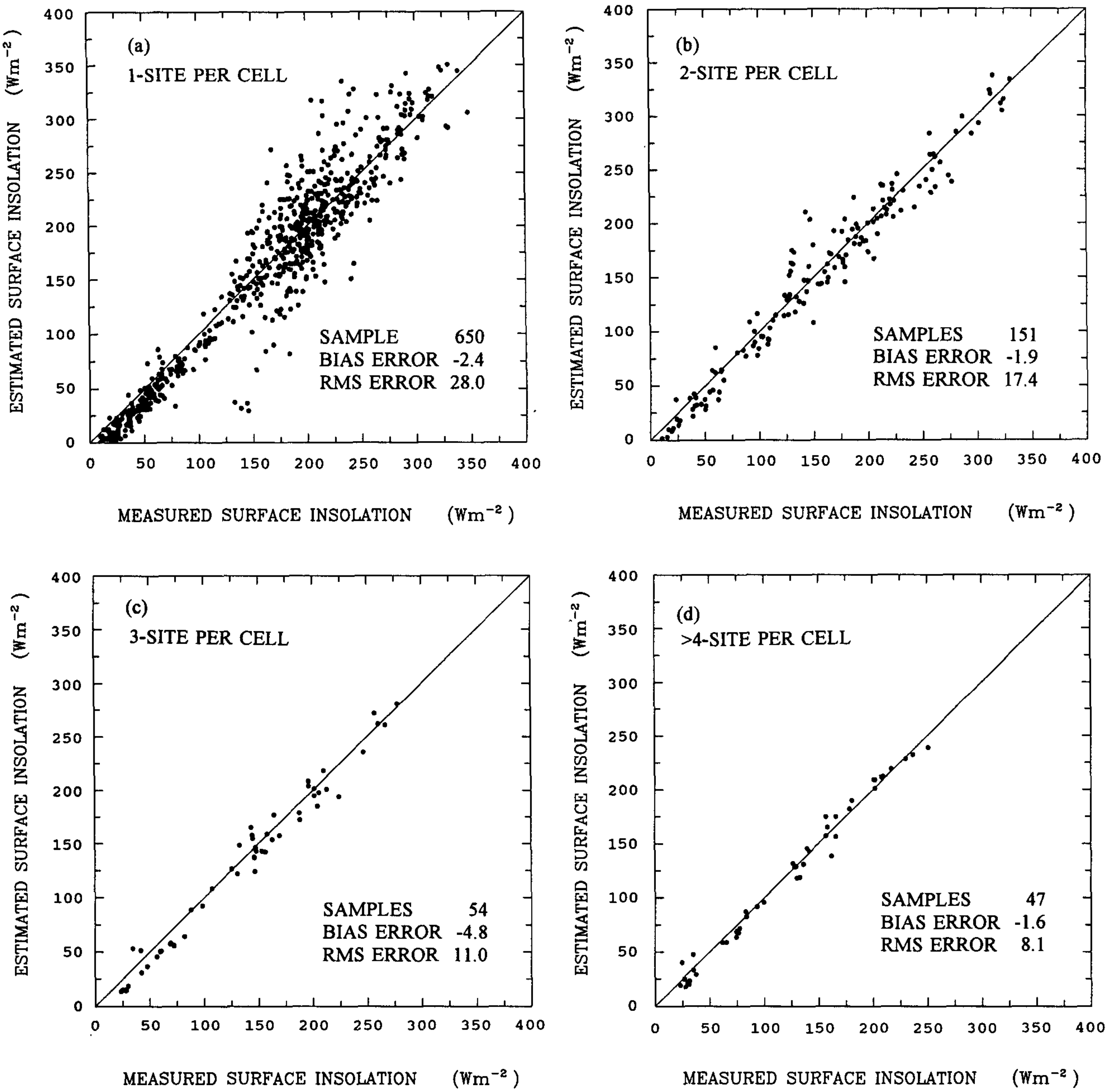

FIG. 5. Same as Fig. 2 but for four categories of the GEBA subset data classified solely on the number of surface sites per cell as indicated on the plots.

the RMSE as the rms difference of the two datasets. The rms difference can, therefore, be expressed as

$$(\text{rms difference})^2 = (\text{random error})^2 + (\text{representative error})^2,$$

where the random error of the satellite-estimated insolation is on the order of 5 W m^{-2}, and the representative error would vary with $N^{-1/2}$ if the N observations were independent. It is, however, likely that the N observations obtained over the same ISCCP C1 cell are correlated. Moreover, the degree of dependency varies with the number of surface radiation stations. As the density of surface stations decreases, the N observations tend to be more independent. Due to such a complexity, it is difficult to determine statistically the power of N. The power is thus computed from optimization computation. Since the optimization computation leads to a power close to -1 (-0.91), it is rounded off to -1, from which the two coefficients in (1), (4), (5) were derived.

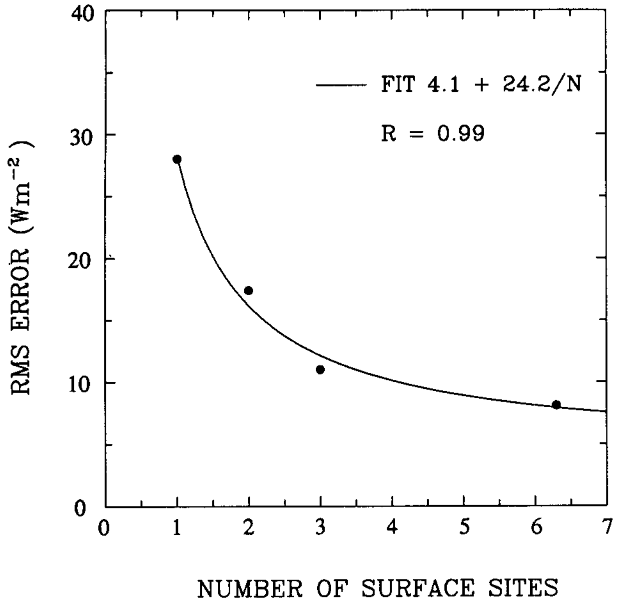

FIG. 6. Variation of the root-mean-square error with the number of surface sites per cell. The solid line is a fit by the function given in the plot with a multiple correlation coefficient of 0.99.

c. Month to month comparison

Month to month comparison of the monthly mean fluxes estimated from satellite data and surface observations is useful to evaluate the ability of satellite observations to monitor intraannual and interannual variations of the surface insolation and to reveal any seasonal trend in the accuracy of the estimates. Figure 7 shows comparisons of surface insolation corresponding to the three cells in the German subset using the GEBA insolation and the ERBE/GEWEX insolation data. For these comparisons, the signal to noise ratios may be marginally large enough as a result of the representation of the surface grid-mean values. It seems that the satellite estimates follow moderately well the course of the interannual and intraannual variations of the surface observations, though a weak trend of estimation error is discernible. Of the six months of July for the three cells, for example, there are five months in which the estimated values of insolation are slightly higher than the observed ones. On the other hand, estimated values are somewhat smaller for four of the six months of January that are presented here.

Figure 8 shows the mean monthly bias errors averaged over the global GEBA-ISCCP cells for all the three insolation products. Although the representative error is larger for the total dataset than for the German subset, averaging over a large number of cells can effectively eliminate the representative error. It appears that the bias errors in all the three insolation products have seasonal trends with sinusoidal shapes. The amplitude of the insolation estimate for the GEWEX/SRB is about 5 W m^{-2}, and for the ERBE/GEWEX it is around 3 W m^{-2}. For the GEWEX/SRB insolation estimate, maximum and minimum bias errors occur in July and January, respectively. This may suggest that the bias error trend in the GEWEX/SRB insolation estimate is associated with the seasonal variation of the TOA insolation. For the ERBE/GEWEX insolation, the seasonal trend is not as evident as that for the GEWEX/SRB. In any case, the seasonal trend explains only part of the bias errors in the estimates of surface insolation from the two GEWEX/SRB products. The bias errors in the GEWEX/SRB insolation estimates are positive for all months, whereas those in the ERBE/GEWEX include both positive and negative

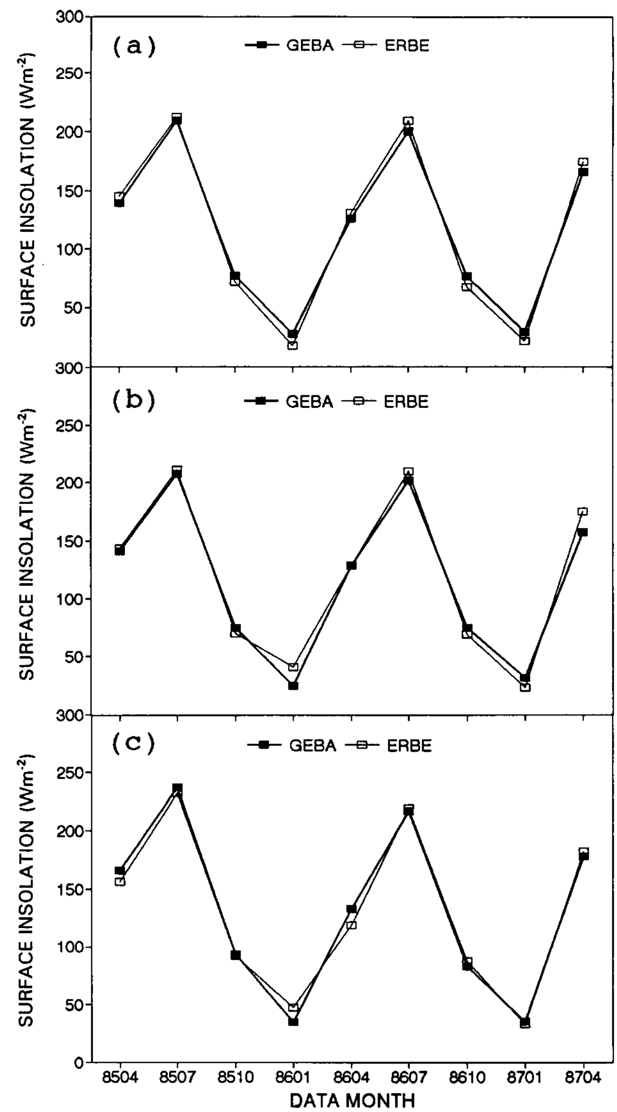

FIG. 7. Month to month comparison of the ERBE/GEWEX surface insolation with the measured insolation from GEBA for three ISCCP cells located in Germany denoted by the ISCCP numbers of (a) 5733, (b) 5827, and (c) 5828.

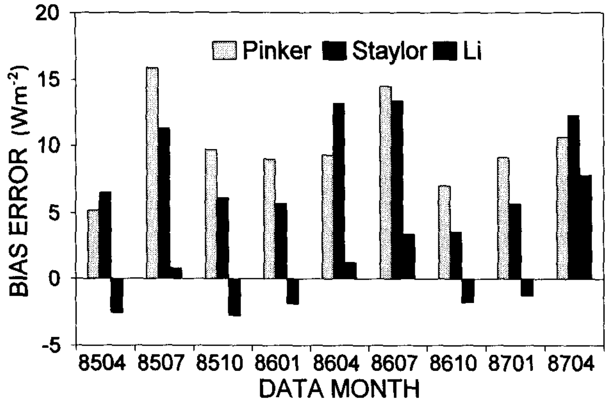

FIG. 8. Bias errors in the monthly mean estimates of surface insolation averaged over the global GEBA-ISCCP cells.

values of much smaller magnitudes. This reinforces the finding that the GEWEX/SRB insolation data are systematically overestimated.

d. Decoupling the agreement

Surface insolation depends on both TOA insolation and absorbing–scattering by the column of the atmosphere. The daily mean TOA insolation is altered by the daily mean cosine of SZA and daytime length, both of which can be computed according to latitude and date. Clouds play a dynamic role in absorbing and scattering the solar radiation reaching the atmosphere. Good correlation between the observed and estimated surface insolation is partially due to the dependency of surface insolation on the TOA insolation. When the TOA insolation is high, surface insolation tends to be high also. This can be clearly seen in Fig. 9, which shows the relationship between TOA insolation and surface insolation (a) measured on the ground and (b) estimated from space. It follows that the relationship between the two types of insolation is approximately linear. The scatter around the regression line reflects the effect of clouds, which strengthens with increasing TOA insolation. Also noteworthy is the similar appearance of the two scatterplots. To remove the dependency on TOA insolation, surface insolation is divided by TOA insolation. The resulting ratio denotes the monthly mean atmospheric transmissivity, which is compared in Fig. 10. In comparison with Fig. 2, the agreement for transmissivity is worse than for insolation. Apart from the effect of removing the dependency on TOA insolation, there are three plausible reasons for the degradation. First, for low insolation, the disagreement is seemingly enlarged. For example, the moderate agreement found in Fig. 2d for low values of surface insolation is transformed into very poor agreement for transmissivity comparison shown in Fig. 10d. Second, the comparison for transmissivity is more sensitive to the representative error than insolation. To a large extent, the representative error is associated with the presence of clouds. The apparent effect of clouds on transmissivity is larger than that on insolation, because the effect of SZA on transmissivity is alleviated. Third, the scatter around the regression line is also due to aerosols, H_2O, surface reflection, or anything that absorbs or scatters but is not parameterized or observed perfectly. Any of these effects are amplified for transmissivity comparison. Nevertheless, it still appears true that the

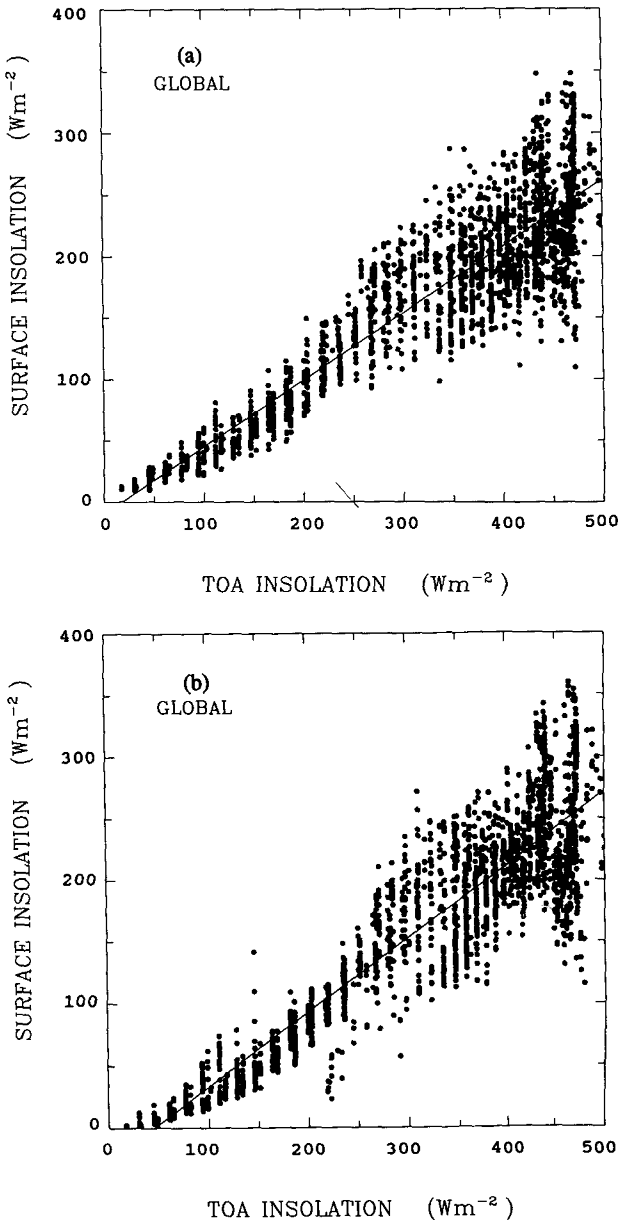

FIG. 9. Relationship between the monthly means of TOA insolation and surface insolation. Surface insolation data are obtained from both (a) ground-truth observation and (b) satellite-based estimation.

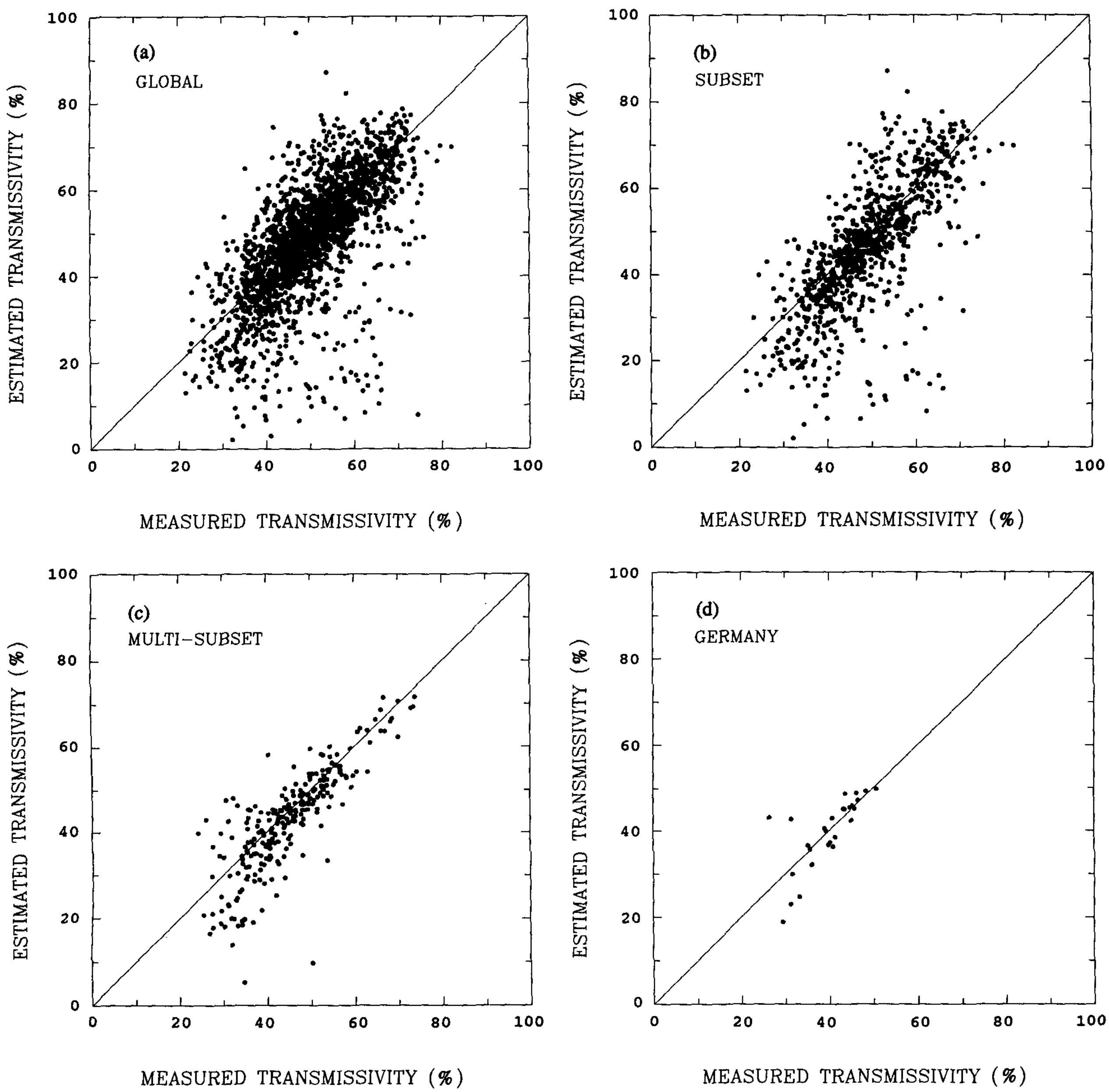

FIG. 10. Same as Fig. 2 but for monthly mean atmospheric transmissivity.

agreement improves as the number of surface stations increases. The correlation coefficients are equal to 0.67, 0.71, 0.82, and 0.74 for the comparisons with global dataset, subset, multisubset, and Germany subset, respectively. The drop of the correlation coefficient for the German comparison arises from the poor agreement for the low TOA insolation cases, which influence the result significantly owing to the small sample size. Bias errors in the transmissivity estimates are equal to −1.7%, −2.9%, −2.1%, and −0.1% for the global dataset, subset, multisite subset, and German subset, respectively.

4. Concluding remarks

The WCRP has established a SRB climatology project aimed at determining global SRB from satellite measurements because of the limitations in surface observations. The study validates monthly mean surface insolation data from two global SRB datasets against surface measurements from the Global Energy Balance Archive (GEBA). One SRB dataset was developed from the ERBE data (ERBE/SRB) using the algorithm of Li et al. (1993a), and the other from the ISCCP data (GEWEX/SRB) using the algorithm of Pinker and Laszlo and that of Staylor. The former contains net solar radiation data, while the latter includes both surface insolation and albedo data. The net solar radiation data from ERBE/SRB are combined with the Staylor surface albedo data from the GEWEX/SRB to derive surface insolation (ERBE/GEWEX insolation) data. Both the ERBE/GEWEX insolation data and the GEWEX/SRB insolation data are validated with emphasis on the former.

The validations were conducted for four combinations of nine months of the GEBA data, namely, global dataset, subset, multisite subset, and German subset. The number of surface sites per cell for these datasets range from 1 to 10. For the four combinations, the bias errors of the ERBE/GEWEX insolation are 0.3, −2.4, −2.4, and 0.8 W m^{-2}, and the rms errors are 27.5, 25.1, 14.9, and 8.1 W m^{-2}, respectively. For the same comparison with the GEWEX/SRB insolation data, the bias errors of the estimates obtained using the Staylor algorithm are 8.7, 5.1, 6.2, and 6.5 W m^{-2}, and those using the Pinker algorithm are 10.1, 7.1, 7.9, and 5.8 W m^{-2}. The rms errors for the GEWEX/SRB insolation estimates are about the same as those of the ERBE/GEWEX insolation estimates. The fact that rms error decreases drastically with the increasing number of surface sites per cell suggests that the rms error is caused mainly by inadequate spatial representation of point-surface measurements within a large region of a cell. Analysis of the variation of rms error with the number of surface sites showed that the rms error would be about 5 W m^{-2} (same as the uncertainty of the ERBE monthly mean TOA data) if the number of surface observations were very large. In addition, month to month comparisons suggested that satellite observation is capable of monitoring intra- and interannual variations of surface insolation. A weak seasonal trend exists, however, in the difference of surface insolation between ground-truth observations from the GEBA and satellite-based estimates from both the ERBE and the ISCCP data. The good agreements found in this study are partially attributed to the dependency of surface insolation on TOA insolation. After such a dependency is removed, the comparison for transmissivity deteriorates somewhat in terms of RMSE. Bias errors for transmissivity comparisons are still close to zero. Since the ERBE/SRB data were based on the Li et al. (1993a) algorithm using the coefficients derived for clear skies, the near-zero–bias error may imply that clouds do not significantly modify SW absorption in the atmosphere on a monthly basis. This is consistent with the findings that SW cloud forcing within the atmosphere layer (i.e., the difference in solar atmospheric radiation budget between clear and cloudy atmospheres) is very small ($\sim$5 W m^{-2}) (Li and Leighton 1993).

The results of this study suggest that the algorithm of Li et al. (1993a) is able to derive monthly mean surface solar radiation budget estimates with acceptable accuracy and that the ERBE schemes that were used to obtain monthly mean TOA fluxes from individual radiance measurements work well. The conclusion, however, is based on two assumptions. One is that there is no trade-off between the errors in the estimation of the SRB and the errors in the application of diurnal variation models and the ADMs. The other is that there is no offset between the errors in the estimates of the net solar radiation and the errors in the values of surface albedo. Unfortunately, these assumptions cannot be justified with the current datasets. Moreover, the surface observations employed in the comparison were made almost exclusively over land. There is a severe shortage of ground-truth data over ocean and polar regions. It is not clear from this study why the two datasets compare differently with surface observations. In view of these limitations, quality analyses were also conducted by intercomparing the two SRB datasets in terms of the discrepancies for both input data and algorithm that were used to generate the SRB datasets (Li 1995). It was found that the systematic errors in the GEWEX/SRB lie in the computation of water vapor absorption. If both the algorithm of Li et al. (1993a) and that of Pinker and Laszlo (1992) used the same method for computing water vapor absorption and were applied to the same input data, the two algorithms would lead to very similar estimates of surface net solar radiation, the global means agreeing to within 1 W m^{-2}. In addition, the intercomparison study identifies some regional deficiencies for both datasets, particularly over the major desert and polar regions (Li 1995).

Acknowledgments. Comments from two anonymous reviewers and the editor, Prof. J. Coakley, Jr., helped to improve the clarity and quality of the paper. Constructive suggestions made by Prof. Coakley inspire the discussion in section 3d. We are grateful to T. Alberta for extracting some of the GEWEX/SRB parameters and to C. Moats for editorial assistance. Useful comments were also made by Drs. H. Barker and J. Cihlar. This work is supported by an operating grant from the Canadian Green Plan to the Canada Centre for Remote Sensing.

REFERENCES

Barkstrom, B., E. Harrison, G. Smith, R. Green, J. Kibler, R. Cess, and the ERBE Science Team, 1989: Earth Radiation Budget Experiment (ERBE) archival and April 1985 results. *Bull. Amer. Meteor. Soc.*, **70**, 1254–1262.

Bishop, J. K. B., and W. B. Rossow, 1991: Spatial and temporal variability of global surface solar irradiance. *J. Geophys. Res.*, **96**, 16 839–16 858.

Breon, F.-M., R. Frouin, and C. Gautier, 1994: Global shortwave energy budget at the earth's surface from ERBE observations. *J. Climate,* **7**, 309–324.

Briegleb, B., P. Minnis, V. Ramanathan, and E. Harrison, 1986: Comparison of regional clear-sky albedo inferred from satellite observations and model computations. *J. Climate Appl. Meteor.*, **25**, 214–226.

Brooks, D. R., E. F. Harrison, P. Minnis, and J. T. Suttles, 1986: Development of algorithms for understanding the temporal and spatial variability of the earth's radiation balance. *Rev. Geophys.*, **24,** 422–438.

Charlock, T. P., C. H. Whitlock, and T. Alberta, 1993: The GEWEX Surface Radiation Budget Project: Results available on line. *GEWEX News,* **3,** No. 1, 7–9 and No. 2, p. 11. [Available from International GEWEX Project Office, Washington DC]

Darnell, W. L., W. F. Staylor, S. K. Gupta, N. A. Ritchey, and A. C. Wilber, 1992: Seasonal variation of surface radiation budget derived from International Satellite Cloud Climatology Project C1 data. *J. Geophys. Res.,* **97,** 15 741–15 760.

DiPasquale, R. C., and C. H. Whitlock, 1993: First WCRP long-term satellite estimates of surface solar flux for the global and selected regions. *Proc. 25th Int. Symp. on Remote Sensing and Global Environmental Change,* Graz, Austria, ERIM/JOANNEAM Research/CIESIN,

Freund, J. E., 1973: *Modern Elementary Statistics.* Prentice-Hall, 532 pp.

Hunt, G., R. Kandel, and A. Mecherikunnel, 1986: A history of presatellite investigations of the Earth's radiation budget. *Rev. Geophys.,* **24,** 351–356.

Kratz, D. P., and R. D. Cess, 1985: Solar absorption by atmospheric water vapor: A comparison of radiation models, *Tellus, Ser. B,* **37,** 53–63.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.,* **31,** 118–132.

Li, Z., 1995: Quality analysis of the global solar SRB data: Intercomparison between two satellite-based products. *J. Geophys. Res.,* submitted.

——, and H. G. Leighton, 1991: Scene identification and its effect on cloud radiative forcing in the Arctic. *J. Geophys. Res.,* **96,** 9175–9188.

——, and ——, 1992: Narrowband to broadband conversion with spatially autocorrelated reflectance measurements. *J. Appl. Meteor.,* **31,** 421–432.

——, and ——, 1993: Global climatologies of solar radiation budgets at the surface and in the atmosphere from 5 years of ERBE data. *J. Geophy. Res.,* **98,** 4919–4930.

——, and L. Garand, 1994: Estimation of surface albedo from space: A parameterization for global application. *J. Geophys. Res.,* **99,** 8335–8351.

——, H. G. Leighton, K. Masuda, and T. Takashima, 1993a: Estimation of SW flux absorbed at the surface from TOA reflected flux. *J. Climate,* **6,** 317–330.

——, ——, and R. D. Cess, 1993b: Surface net solar radiation estimated from satellite measurements: Comparison with tower observations. *J. Climate,* **6,** 1764–1772.

Liu, W. T., W. Tang, and F. J. Wentz, 1992: Precipitable water and surface humidity over global oceans from special sensor microwave image and European Centre for Medium-Range Weather Forecasts. *J. Geophys. Res.,* **97,** 2251–2264.

Moser, W., and E. Raschke, 1983: Mapping of global radiation and of cloudiness from METEOSAT image data. *Meteor. Resch.,* **36,** 33–41.

Ohmura, A., and H. Gilgen, 1991: *Global Energy Balance Archive (GEBA).* World Climate Program—Water Project A7, Report 2: The GEBA Database, Interactive Application, Retrieving Data. Verlag der Fachvereine, 60 pp.

——, and ——, 1993: Re-evaluation of the global energy balance. *Geophys. Monogr.,* No. 75, IUGG, 93–110.

Pinker, R., and I. Laszlo, 1992: Modeling surface solar irradiance for satellite applications on a global scale. *J. Appl. Meteor.,* **31,** 194–211.

——, R. Frouin, and Z. Li, 1995: A review of satellite methods to derive surface shortwave irradiance. *Rem. Sens. Environ.,* in press.

Ramaswamy, V., and S. M. Freidenrecich, 1992: A study of broadband parameterizations of the solar radiative interactions with water vapor and water drops. *J. Geophys. Res.,* **97,** 11 487–11 512.

Rossow, W., L. Garder, P. Lu, and A. Walker, 1991: ISCCP documentation of cloud data. *WMO/TD,* **266,** WMO, 76 pp.

Schiffer, R. A., and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Programme. *Bull. Amer. Meteor. Soc.,* **64,** 779–984.

Schmetz, J., 1989: Towards a surface radiation climatology: Retrieval of downward irradiance from satellites. *Atmos. Res.,* **23,** 287–321.

Staylor, W. F., and A. C. Wilber, 1990: Global surface albedos estimated from ERBE data. Preprints, *7th AMS Conf. Atmosphere Radiation,* San Francisco, CA, Amer. Meteor. Soc., 231–236.

Suttles, J. T., and G. Ohring, 1986: Surface radiation budget for climate applications. *NASA Ref. Publ.,* 1169, 132 pp.

——, R. N. Green, P. Minnis, G. L. Smith, W. F. Staylor, B. A. Wielicki, I. J. Walker, D. F. Young, V. R. Taylor, and L. L. Stowe, 1989: Angular radiation models for earth-atmosphere system, Vol. 1, Shortwave radiation. NASA Ref. Publ., 1184, 84 pp.

Whitlock, C. H., and Coauthors, 1990: Comparison of surface radiation budget satellite algorithms for downwelled shortwave irradiance with Wisconsin FIRE/SRB surface truth data. Preprints, *Proc. 7th AMS Conf. Atmosphere Radiation,* San Francisco, CA, Amer. Meteor. Soc., 237–242.

——, T. P. Charlock, W. F. Staylor, R. T. Pinker, I. Laszlo, R. C. DiPasquale, and N. A. Ritchey, 1993: WCRP surface radiation budget shortwave data product description—Version 1.1. NASA Tech. Memo., **10 7747** 28 pp.

Wielicki, B. A., B. R. Barkstrom, and 18 CERES Team Members, 1994: Clouds and the Earth's Radiant Energy System (CERES) algorithm theoretical basis document: Overview of cloud retrieval and radiative flux inversion, Version 1.1. NASA Langley Research Center, Hampton, VA, 19 pp.