Exploiting diurnal variations to evaluate the ISCCP-FD flux calculations and radiative-flux-analysis-processed surface observations from BSRN, ARM, and SURFRAD

Yuanchong Zhang,[1] Charles N. Long,[2] William B. Rossow,[3] and Ellsworth G. Dutton[4]

Received 26 June 2009; revised 5 March 2010; accepted 29 March 2010; published 6 August 2010.

[1] Using a meteorological similarity comparison method (MSCM), we performed a mutual and simultaneous evaluation of the surface radiative flux datasets from the International Satellite Cloud Climatology Project-FD and the new radiative-flux-analysis-processed surface observations (RFA-PSO). For downward shortwave (SW), diffuse (Dif), and direct (Dir) fluxes, matching cloud fraction (CF) reduces the flux difference between FD and RFA-PSO by up to a factor of 2. Decreasing the aerosol optical depth values used in the FD calculations accounts for much of the remaining difference. For downward longwave (LW) flux, matching either surface air temperature or CF reduces the flux difference to nearly zero. For the total downward SW diurnal variations, there is excellent agreement for both clear and cloudy sky, but less good agreement for the Dif and Dir components. The latter agree much better for clear sky when the FD aerosol optical depth is reduced and for cloudy sky when matching CF and cloud optical depth jointly. For LW diurnal variations, the agreement is best for clear sky, but FD has a larger amplitude by 3–7 W/m^2 for cloudy sky because of differing sensitivities to cirrus and low clouds in the two datasets. These results confirm that the source of the FD surface flux uncertainty of ~10–15 W/m^2 is the input quantities, not the radiative transfer model. An important limitation of the RFA-PSO cloud parameters (not the fluxes) is the inhomogeneous diurnal sampling and the retrieval difficulties with broken clouds (SW) and cirrus clouds (LW).

Citation: Zhang, Y., C. N. Long, W. B. Rossow, and E. G. Dutton (2010), Exploiting diurnal variations to evaluate the ISCCP-FD flux calculations and radiative-flux-analysis-processed surface observations from BSRN, ARM, and SURFRAD, *J. Geophys. Res.*, *115*, D15105, doi:10.1029/2009JD012743.

1. Introduction

[2] Rotation of the Earth causes a diurnal cycle of solar heating that is manifested through various physical processes as variations of the state of the surface and atmosphere in all but the highest polar regions of the Earth (where solar variations are more strongly seasonal). Particularly important processes are those affecting evaporation of water from the surface and triggering the stronger convective motions that produce much of the precipitation. The diurnal cycle is a notable signal in Earth's weather and climate, so whether it is correctly represented in both magnitude and phase is an important test of how well a general circulation model (GCM) simulates the relevant processes that are involved in the exchanges of energy and water between the surface and atmosphere [*Bosilovich and Schubert*, 2002]. The test would begin with how well the diurnal variations of radiative fluxes can be determined to provide a more subtle evaluation of the radiative transfer model. In the case of shortwave fluxes (SW = 0.2–5.0 μm wavelength), deviations from the purely geometric variations (solar zenith angle) of the fluxes over the course of the day have a strong dependence on the diurnal variations of clouds. Moreover, the diurnal variations of the ratio of the direct and diffuse flux components for clear and cloudy sky may reveal more subtle dependencies on the proportions of absorption and scattering by clouds, aerosols, and gases. This ratio directly shows whether the amount of scattering by aerosols (primarily under clear skies) and clouds is correctly specified. In the case of longwave fluxes (LW = 5.0–200 μm wavelength), the diurnal variations of the downward and upward fluxes for both amplitude and phase depend mostly on the associated diurnal variations in atmospheric temperature (with weaker dependence on humidity and cloud variations) and surface skin temperature, respectively.

[1]Department of Applied Physics and Applied Mathematics, Columbia University, NASA Goddard Institute for Space Studies, New York, New York, USA.
[2]Pacific Northwest National Laboratory, Richland, Washington, USA.
[3]Cooperative Remote Sensing Science and Technology Center (CREST), City University of New York, NASA Goddard Institute for Space Studies, New York, New York, USA.
[4]National Oceanic and Atmospheric Administration-Global Monitoring Division, Earth System Research Laboratory, Boulder, Colorado, USA.

0148-0227/10/2009JD012743

[3] The International Satellite Cloud Climatology Project-FC (ISCCP-FC, called FC hereafter) [*Zhang et al.*, 1995] provided an early opportunity to study diurnal radiative flux variations. *Rossow and Zhang* [1995] compared FC's monthly mean 3-hourly (in local solar time) SW albedo and LW upward fluxes at the top of atmosphere (TOA) with the Earth Radiation Budget Experiment [*Ramanathan et al.*, 1989]. The results were generally acceptable, but they identified a lack of diurnal atmospheric temperature variations and a likely clear-sky bias in surface temperature diurnal variations as problems with surface LW fluxes. The current product, ISCCP-FD (FC's successor, called FD hereafter) [*Zhang et al.*, 2004], has incorporated a diurnal adjustment scheme for near-surface atmospheric temperatures and surface skin temperature for land areas. The scheme is generally better at representing LW diurnal variations on the basis of the comparisons with surface measurements [*Zhang et al.*, 2004].

[4] However, a systematic and more detailed evaluation for our surface radiative fluxes has not yet been performed, partly because of the lack of high-quality surface observations with enough information about the coincident meteorological conditions to allow explanation of the relationship between the surface and satellite-based flux determinations. In addition to nearly continuous radiative flux data records over long enough time periods to sample the full range of meteorological variability, we needed coincident aerosol, cloud, and atmospheric properties for a large enough number of the stations in different climate regimes. This situation has changed, however, with the emergence of a methodology for estimating the cloud effects on the surface radiative fluxes, the radiative flux analysis (RFA) method, as described in a series of papers [*Long and Ackerman*, 2000; *Long and Gaustad*, 2004; *Long*, 2004, 2005; *Long et al.*, 2006; *Barnard and Long*, 2004; *Barnard et al.*, 2008; *Long and Turner*, 2008]. The RFA method has now been applied to high time resolution surface radiative flux datasets from the Baseline Surface Radiation Network (BSRN) [*Ohmura et al.*, 1998], the Atmospheric Radiation Measurement (ARM) [*Ackerman and Stokes*, 2003], and the National Oceanic and Atmospheric Administration's (NOAA's) Surface Radiation budget network (SURFRAD) [*Augustine et al.*, 2000] sites. We call this product RFA-Processed Surface Observations (RFA-PSO), simplified to PSO here. PSO has been released as part of the Global Energy and Water Cycle Experiment (GEWEX) Radiative Flux Assessment (see http://gewex-rfa.larc.nasa.gov/GEWEX-RFA/). The importance of PSO lies in that, in addition to the directly observed flux components, it contains estimates of the clear-sky counterparts (beyond the limited direct, true clear-sky measurements) and their associated meteorological properties (see section 2). This combination of all-sky and clear-sky fluxes with the meteorological information provides the long-awaited opportunity to evaluate the FD surface radiative fluxes in greater detail, and conversely, it also provides a useful test of the PSO product, and therefore the RFA method, which is still being refined. Because the two data products are entirely independent, their agreement is actually a mutual validation while the investigation of physical causes for their differences can lead to further improvements for both the radiation model and RFA method.

[5] Although some more general evaluation is always necessary whenever datasets are compared for the first time, we focus more on the diurnal variation comparisons using a Meteorological Similarity Comparison Method (MSCM) that exploits the availability of the coincident meteorological information to refine the comparison. We first describe the two datasets in section 2 and introduce MSCM in section 3. We compare the diurnal variations of the SW fluxes in section 4 and the diurnal variations of the LW fluxes in section 5. Section 6 summarizes our results and conclusions.

2. Datasets

[6] The RFA-PSO product provides 15-minute averages of surface radiative flux measurements and, if allowed by the RFA retrieval, of the associated meteorological parameters. All quantities are arithmetic averages over all available times at the original 1- to 3-minute measurement intervals for all days in 2004 at 20 stations. Monthly averages of the15-minute mean values for the period from 1992 to 2006 are also available at 30 stations. The measured and derived parameters in the PSO product include all- and clear-sky, downwelling and upwelling, SW and LW fluxes, SW-derived cloud cover and effective visible cloud optical depth (only for original 1- to 3-min measurements where conditions are overcast, cloud cover $\geq$0.90), LW-derived cloud cover and effective cloud base height, surface air temperature, humidity, and wind speed. In addition, surface downwelling diffuse and direct SW fluxes for both all and clear sky are also provided (see ftp://typhoon.larc.nasa.gov/pub/gewex-rfa/data/flux_analysis/1_Flux_Analysis_notes_RFA.txt, referred to Chuck Long's Derived Parameters at http://gewex-rfa.larc.nasa.gov/GEWEX-RFA/; user registration is required for accessing this PSO product).

[7] Fifteen stations are selected for this study because they have the most complete information and most nearly complete time records. We use the 2004 dataset, which has the full 15-minute time record. The stations used are listed in Table 1. Also indicated in Table 1 are the 10 stations with aerosol optical depth (AOD) measurements available. Seven of them are from the SURFRAD and BSRN (SURFRAD AOD values are provided compliments of J. Augustine [*Augustine et al.*, 2008]), and the rest (Nauru, Manus, and Darwin) are from ARM/AERONET [*Holben et al.*, 1998] sites but newly processed with more stringent thresholds for cloud screening by C. Flynn (personal communication). For comparison with FD, we have converted the original spectral AOD values at 415, 500, and 615 nm wavelengths to AOD at 550 nm wavelength using the Angstrom power law formula, because 550 nm is the reference wavelength that FD uses to specify aerosol radiative properties. Table 2 gives the definitions of parameter acronyms used in this study.

[8] For ARM and SURFRAD, the 2σ uncertainty of PSO monthly mean fluxes is estimated to be well below 10 W/m^2 for all-sky total, diffuse, and direct downward SW and LW fluxes; the inferred clear-sky fluxes should also have uncertainties <10 W/m^2. The uncertainty is estimated to be 10% or better for SW-derived cloud fraction (CFsw) and optical thickness (CldTau), 0.01–0.02 for AOD by *Augustine et al.* [2008], and 0.6 K for surface air temperature and 2%–3% for humidity [*Ritsche*, 2006]. For BSRN the uncertainty

Table 1. The 15 Stations Selected from BSRN, ARM, and SURFRAD Networks for This Study[a]

Station Acronym	Station Name (Owner)[b]	Quality Rate Network[c]	Station Latitude/ Longitude	FD Cell[d] Latitude/Longitude	AOD[e]
NYA	Ny Alesund, Spitsbergen (GM/NY)	B-BSRN	78.9°N/11.9°E	78.8°N/6.4°E	N/A
FPE	Fort Peck, MT [USA]	A-SURFRAD	48.5°N/254.8°E	48.8°N/255.8°E	AV
PAY	Paverne [Switzerland]	A-BSRN	46.8°N/6.9°E	46.2°N/5.4°E	N/A
PSU	Rock Springs, PA [USA]	A-SURFRAD	40.7°N/282.1°E	41.2°N/281.7°E	AV
BOS	Boulder, CO [USA]	A-SURFRAD	40.2°N/254.6°E	41.2°N/255.0°E	AV
BON	Bondville, IL [USA]	A-SURFRAD	40.1°N/271.4°E	41.2°N/271.7°E	AV
DRA	Desert Rock, NV [USA]	A-SURFRAD	36.6°N/243.9°E	36.2°N/243.6°E	AV
BIL	Billings, OK [USA]	B-ARM	36.6°N/262.5°E	36.2°N/262.2°E	N/A
TAT	Tateno [Japan]	B-BSRN	36.0°N/140.1°E	36.2°N/141.2°E	N/A
GCR	Goodwin Creek, Mississippi [USA]	A-SURFRAD	34.2°N/270.1°E	33.8°N/271.5°E	AV
NAU	Nauru Island [USA]	B-ARM	0.5°S/166.9°E	1.2°S/166.2°E	AV
MAN	Momote, Manus Island, Papua New Guinea [USA]	B-ARM	2.1°S/147.7°E	1.2°S/148.8°E	AV
DAR	Darwin [Australia]	B-ARM	12.5°S/130.9°E	13.8°S/129.9°E	AV
GVN	George von Neumaver, Ant. [GM]	B-BSRN	70.7°S/351.8°E	71.2°S/348.3°E	N/A
SPO[f]	South Pole, Antarctica [USA]	B-BSRN	89.8°S/258.0°E	88.8°S/300.0°E	AV

[a]All the sites may have practical problems such as tracker fails or human errors that make observations unavailable. Except GCR, all have both downward and upward fluxes and some meteorological parameters, but upward SW and LW of NAU and MAN are not used because they are not reliable.
[b]Owners of different countries, where GM = Germany and NY = Norway.
[c]Quality rate (based on Long's personal judgment): A = very good; B = pretty good or good, and the network a station belongs to.
[d]FD's 280-km equal area grid cell's central position.
[e]AV means available for aerosol optical depth (AOD) data (at a total of 10 stations) and N/A means not available.
[f]The only reason to use longitude = 258.0°E is that this longitude is used by RFA-PSO, so the cell of FD is also picked using this longitude. Theoretically speaking, arbitrary values can be used at the South Pole, but practically we have to use the consistent values for comparison.

estimates are 5 Wm^2 for downwelling SW and 10 W/m^2 for downwelling LW, but these estimates are determined from standard deviations of the calibration coefficients [*Ohmura et al.*, 1998], so they are incomplete. A more complete evaluation is needed because the actual uncertainty for field measurements will be somewhat larger than for instrument calibration. Some of these sources of error are illustrated here. For detailed discussions on RFA-PSO data uncertainties, see the Appendix.

[9] The ISCCP-FD dataset consists of the upwelling, downwelling, shortwave, and longwave radiative flux profiles (TOA to surface) for all, clear, and overcast sky, along with all the input parameters used in the flux calculations (except for aerosols). The fluxes are calculated using a complete, detailed, and physically consistent radiative transfer model from the NASA Goddard Institute for Space Studies General Circulation Model (with some modifications; see *Zhang et al.* [2004]) with improved observations of the physical properties of the surface, atmosphere, and clouds based on the ISCCP D-series cloud climatological datasets [*Rossow and Schiffer*, 1999] and several other satellite data products. The FD flux product is global, 3-hourly on a 280 km equal area grid map covering the time period from July 1983 through December 2007 (to be extended as more ISCCP products become available). FD is provided in four subdatasets with the same spatial and temporal resolutions (see http://isccp.giss.nasa.gov/projects/flux.html), of which only the surface (SRF) dataset is needed here. However, to be able to compare to PSO's downwelling diffuse and direct SW fluxes, as well as the other meteorological parameters that the standard SRF FD does not report, we produced a special FD product with additional input and output parameters saved at 3-hr intervals for all of 2004. The code and input datasets are identical to those used for the standard FD.

[10] The uncertainty for FD surface flux components is estimated from previous studies to be 10–15 W/m^2 for regional monthly means [*Zhang et al.*, 2004]. For the other input parameters (used in the flux calculation), uncertainties are estimated to be less than or equal to ~10% for cloud fraction, less than or equal to ~10% for cloud optical thickness, 2–3 K for surface air temperature, and 20–25% for precipitable water [*Rossow and Schiffer*, 1999]. The latter two parameters are from the operational Television

Table 2. Definition of Acronyms for Parameters Used ion This Study

Parameter	Definition
SWdn	All-sky surface downward broadband SW flux
CSWdn	Clear-sky surface downward broadband SW flux
SWup	All-sky surface upward broadband SW flux
CSWup	Clear-sky surface upward broadband SW flux
Dir	All-sky surface downward broadband direct flux
CDir	Clear-sky surface downward broadband direct flux
Dif	All-sky surface downward broadband diffuse flux
CDif	Clear-sky surface downward broadband diffuse flux
AOD	Column aerosol optical depth, but nominal 550-nm value is actually used here
CFsw	SW-based cloud fraction (CF) for PSO; for FD, it is just CF matched to PSO CFsw
CldTau	Cloud optical depth
SWnet	All-sky net surface SW flux
CSWnet	Clear-sky net surface SW flux
LWdn	All-sky surface downward broadband LW flux
CLWdn	Clear-sky surface downward broadband LW flux
LWup	All-sky surface upward broadband LW flux
CLWup	Clear-sky surface upward broadband LW flux
CFlw	LW-based cloud fraction (CF) for PSO; for FD, it is just CF matched to PSO CFlw
CldHgt	Cloud base height
Ta	Surface air temperature
SLSH, SASH	Surface layer specific humidity (for FD) and Surface air specific humidity (for PSO)
LWnet	All-sky net surface LW flux
CLWnet	Clear-sky net surface LW flux
FDrv	FD-like but calculated with the only change of input aerosols with total column AOD ≈ 50% of the original in the standard FD input aerosols.

Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) product [*Zhang et al.*, 2006].

[11] There are fundamental differences between FD and PSO. The former uniquely determines flux values using forward modeling from all the input atmospheric, cloud, and surface parameters, whereas the latter retrieves some of the cloud parameters (and clear-sky fluxes at all times) from the measured flux values using the RFA retrieval method. The FD results, being "forward" calculations, produce unique relationships, whereas the RFA retrieval may not be unique in all situations. In other words, the functional relationships between the fluxes and the meteorological parameters in the two are essentially determined in "opposite" directions. As a result, FD flux errors are associated with input parameter and model errors, whereas PSO retrieved parameter errors are associated with errors in the flux measurements and the retrieval method. Moreover, the spatiotemporal sampling of the two products is very different. FD fluxes are assumed to represent 3-hr averages (intervals centered on UTC = 0, 3, … 21 hr), but they are based on single time samples within the 3-hr interval. SW fluxes are calculated using 3-hr-averaged cosine solar zenith angle while holding all other quantities constant, and LW fluxes are calculated with all quantities constant. The FD 280-km grid cell mean is based on a sampling of the original satellite pixels at 30 km intervals, so it represents only a few percent of the area at each time. However, the FD fluxes are physically consistent with the cloud and other meteorological information through the radiation model. In contrast, the PSO dataset used in this work is a 3-hr average produced from the 15-minute averages at a single site ("point" but the hemispherical view of the surface instruments combined with the temporal averaging represents scenes on average ~30–50 km across). All the 15-min samples are not always available for a given 3-hr period because of a number of practical problems, such as instrumental failures, human errors, and retrieval limitations. Consequently, the 3-hr mean values may not be statistically homogeneous over the time interval and may not provide consistent time sampling for all parameters. For instance, ~5% of the 3-hourly mean SW downwelling (SWdn) fluxes and 10% of LW downwelling fluxes (LWdn) do not have complete 15-min sampling. Likewise, 9% of diffuse fluxes (Dif) and 44% of SW-derived cloud fractions (CFsw) are incompletely sampled. The same kind of inhomogeneity also affects the 15-min mean parameters when they are averaged from the original 1- to 3-min data. These fundamental differences between the two datasets must be carefully taken into account in the mutual evaluation (see section 3).

[12] To reduce the scatter in the comparisons caused by the sampling differences, we focus on the comparison of monthly means [*Rossow and Zhang*, 1995], and therefore we compute monthly averaged 3-hourly means, shortened to monthly 3-hourly, at Local Solar Time (LST) to describe the diurnal cycle. We have also examined finer time variations, some of which are mentioned, to confirm any conclusions reached. This is particularly important when the effects of spatiotemporal mismatches are more severe, as we show here. The central latitude and longitude of an FD cell are used to convert UTC to LST. All monthly mean values are based on matched samples of the 3-hourly means between FD and PSO; in other words, monthly averages are based only on the 3-hourly means when both FD and PSO values are available.

[13] The effects of the PSO sampling are best illustrated by the fact that the monthly mean values based on the monthly averaged 15-min values, monthly averaged 3-hourly means and daily means are sometimes different. On the other hand, since the FD sampling is complete (by design), all three monthly means are identical. However, when FD results are matched to PSO, as in this study, the different monthly means exhibit differences, depending on how complete the PSO time sampling is for a specified parameter. From our review of all the results, this is not a major issue, so we usually do not distinguish the different monthly averages unless necessary for clarity.

3. Methodology

[14] Usually satellite-derived surface radiative fluxes are evaluated by comparison with surface-measured fluxes [e.g., *Rossow and Zhang*, 1995; *Zhang et al.*, 2004], assuming the latter to be the "truth"; but the surface-based dataset is just another dataset with its own sources of uncertainty. Under the assumption that the uncertainty of the surface-based measurements is smaller than that of the satellite-based fluxes, such a comparison does provide some kind of uncertainty estimate for the satellite-based fluxes, even though it is probably an overestimate, but it does not provide any insight into the causes of errors. In particular, we cannot determine how much of the uncertainty comes from errors of the radiative transfer model, the input datasets, or the surface observations themselves. Although we use the traditional method here as a first evaluation, a more meaningful and accurate validation has to involve simultaneous and mutual evaluations of both datasets to obtain a more accurate estimate of uncertainties.

[15] Assuming a particular flux component, f (e.g., SWdn), is a function of the set of parameters specifying the radiative properties of the environment, **ev**, then

$$f = F(\mathbf{ev}), \tag{1}$$

where F is the radiative transfer function represented by a radiative transfer code and **ev** is shown as a vector to indicate that it represents many meteorological parameters, such as temperature, humidity, aerosols, and clouds. If we know the true environmental properties, $\mathbf{ev_o}$, then the true flux is $f_o = F(\mathbf{ev_o})$ because F is the assumed true radiative transfer function in equation (1), whereby using "true," we mean $\mathbf{ev_o}$, f_o, and F represent unique values without any errors. FD uniquely determines each flux, $f_1 = F_1(\mathbf{ev_1})$, where F_1 represents the radiative transfer model actually used, which is assumed to depart from F by an amount depending on how good the model is and $\mathbf{ev_1}$ is obtained from various other datasets, each with their own errors. PSO retrieves some of the environmental parameters from the measured fluxes, whereas some others are measured by other systems or assumed; together, we can represent these data as $\mathbf{ev_2} = F_2^{-1}(f_2) = G_2(f_2)$, where the inverse radiative transfer function, G_2, may not be unique in all situations, especially since only a few parameters of the environment are retrieved. Generally $F \neq F_1 \neq F_2$ and $\mathbf{ev_0} \neq \mathbf{ev_1} \neq \mathbf{ev_2}$, so given that $f_0 \neq f_1 \neq f_2$, we are not able to separate the

different reasons for the flux differences. The usual evaluation method only compares f_1 with f_2, which is not very informative, and assigns all of the discrepancies to errors in f_1, which is an overestimate. In fact, f_o, $\mathbf{ev_o}$ and their relationship $f_o = F(\mathbf{ev_o})$ is never exactly known because of the uncertainties in the several measurement systems involved, but we can try to reduce overall uncertainties, especially for the model results, by a mutual evaluation approach that uses more information.

[16] If we write

$$\Delta f(\mathbf{ev_1},\ \mathbf{ev_2}) = f_1 - f_2 = F_1(\mathbf{ev_1}) - F_2(\mathbf{ev_2}), \qquad (2)$$

we can propose an alternate evaluation approach (the Meteorological Similarity Comparison Method, MSCM) that exploits equation (2) to go beyond the usual approach, provided that the surface flux measurements are accompanied by coincident meteorological parameters for $\mathbf{ev_2}$. This is, in fact, the design concept for the Baseline Surface Radiation Network. We can then search for subsets of the matched satellite and surface flux datasets where the differences between $\mathbf{ev_1}$ and $\mathbf{ev_2}$ are minimized to give us more meaningful uncertainty estimates for the FD and PSO fluxes. This matching process filters out bad mismatches between the two datasets that can arise because of the differing spatiotemporal sampling.

[17] In an ideal case, if $\Delta f \rightarrow 0$ when $\mathbf{ev_1} \rightarrow \mathbf{ev_2}$ and $F_1 \rightarrow F$ and $F_2 \rightarrow F$, then we could claim that the only errors in f from the satellite are those coming from the errors in $\mathbf{ev_1}$ and provide a direct uncertainty estimate. We have already evaluated the possible errors in FD from the input parameters [*Zhang et al.*, 2004, 2006, 2007]. However, another possibility is that $\Delta f \rightarrow 0$ when $\mathbf{ev_1} \rightarrow \mathbf{ev_2}$ but only that $F_1 \rightarrow F_2$. Thus, if the two sets of measurements were based on the same radiative transfer relationships (model), we would not know about the errors in f coming from the model itself. In the case of the comparison of FD and PSO, we know that $F_1 \neq F_2$ (as some of our results show), so it might be the case that $f_1 \neq f_2$, even if $\mathbf{ev_1} = \mathbf{ev_2}$. However, if Δf ($\mathbf{ev_1}$,$\mathbf{ev_2}$) is small when $\mathbf{ev_1} = \mathbf{ev_2}$, as is the case, and we can identify and separate out the cases where F_1 and F_2 differ most, then we can isolate the errors in f_1 that come from F_1 separately from the errors that come from $\mathbf{ev_1}$. Likewise, we can do the same for the errors in f_2 that come from F_2 and the parameters in $\mathbf{ev_2}$, most of which are not directly measured but retrieved from f_2.

[18] Thus, the MSCM is a mutual validation method that essentially matches the FD and PSO fluxes and the atmospheric and surface physical properties, which most affect these fluxes, as closely as possible. In practice $\Delta f(\mathbf{ev_1},\mathbf{ev_2})$ is a minimum for an optimum but nonzero value of $\Delta\mathbf{ev} = \mathbf{ev_1} - \mathbf{ev_2}$ because we do not have enough information to consider all of the parameters that affect the fluxes. We evaluate $\Delta f(\mathbf{ev_1},\mathbf{ev_2})$ either by sorting until $\Delta\mathbf{ev}$ produces a minimum or alternatively by directly changing $\mathbf{ev_1}$ to $\mathbf{ev_2}$ and repeating the FD calculations. We approximate $\mathbf{ev}$ as a function of only one or two meteorological variables and use a trial-and-error approach to find the optimum values that minimize Δf. In this way, we can eliminate those flux value matchups that cause an overestimate of the uncertainties in the usual comparison approach but are not actually associated with true errors of the radiation model or the surface flux measurements. Rather, these cases are associated with the differences in spatiotemporal matching that cause differences in environmental conditions. We can then evaluate the uncertainties in FD associated with the radiative transfer model and in PSO associated with the retrieval method (RFA).

4. Shortwave Flux Evaluation

[19] Our previous comparisons of SW fluxes from satellite-based calculations (FD) and direct surface measurements (BSRN) have shown generally good agreement (*Rossow and Zhang* [1995], *Zhang et al.* [2004]) with biases generally <10 W/m^2 and RMS differences >50 W/m^2 for 3-hourly and 10-15 W/m^2 for monthly means for regions ~280 km in size. Much of the scatter in the comparisons appeared to be associated with the effects of the sparse space-time sampling of the clouds in the ISCCP cloud products, together with the effects of comparing area means and point measurements [*Rossow and Zhang*, 1995]. A large part of the remaining bias was thought to be associated with biases of the AOD from the aerosol climatology used in the calculations. However, limitations in the coincident and collocated meteorological information available with the surface flux measurements at that time, especially concerning the properties of clouds and aerosols, precluded further investigation of the flux differences and confirmation of these conclusions. Using the RFA-PSO product, we can now investigate and identify the causes of the flux differences between FD and PSO.

4.1. General Comparison and Sampling Effects

[20] Comparisons of various monthly mean values aggregated over the 15 stations for 2004 are summarized in Table 3. All differences reported here and throughout this work are taken as FD minus PSO. The mean difference (standard deviation of the differences, SD) for downwelling all-sky surface shortwave flux (SWdn) is −1.1 (20.9) W/m^2. This is similar to the result of +2.0 (18.5) W/m^2 reported in *Zhang et al.* [2004] for comparisons to 1,970 monthly mean samples at 35 BSRN stations over 1992–2001. For the diffuse (Dif) and direct (Dir) fluxes, the mean (SD) differences are 29.8 (24.2) and −30.4 (24.8) W/m^2, respectively. The mean Dir/Dif ratios for the FD and PSO data are 0.53 and 1.12, respectively, yielding a mean (SD) difference of −0.58 (0.57). The mean difference for SWdn for all available matches of all 3-hourly means is about the same as the monthly bias, but the SD of the differences (85.2 W/m^2) is much larger as expected because of poorly sampled cloud variability. Dif and Dir show similar results for 3-hourly values. The fact that the bias in SWdn is very small and the biases in Dir and Dif are larger but of opposite signs indicates that the differences between FD and PSO are in the partitioning of the insolation into direct and scattered fluxes.

[21] The left side of Figure 1 (1L thereafter and similarly for all figures, R for the right side) provides the framework for investigating the partitioning of the SW flux into Dir and Dif fluxes by showing the ratio as a function of SW-derived cloud cover fraction (CFsw) for FD and PSO, respectively. Note that because CFsw can be derived only from SWdn when the sun is 10° or more above the horizon [*Long et al.*, 2006], only ~80% of the nonzero SWdn values (50% of all the SWdn values or 20,280 of all 41,005 samples) in the

Table 3. Comparison of Monthly Mean Surface SW Flux Components and Their Associated Meteorological Parameters[a] Between FD (X) and PSO (Y)[b]

Variable Name	FD (X)	PSO (Y)	Mean Difference	SD	Correlation Coefficient	Slope	Intercept	Nmdv
SWdn	173.78 (178.23)[c]	174.92 (174.92)	−1.135 (3.310)	20.865 (20.970)	0.9733 (0.9741)	0.94 (0.93)	10.93 (9.38)	14.71 (14.59)
CSWdn	231.45 (237.89)	238.84 (238.84)	−7.394 (−0.956)	11.442 (9.031)	0.9945 (0.9969)	0.99 (0.97)	10.45 (7.71)	8.08 (6.07)
SWup	31.64	50.14	−18.497	25.895	0.8908	0.99	18.74	18.38
CSWup	40.34	67.72	−27.381	28.555	0.9082	1.05	25.24	19.54
Dif	101.55 (96.74)	71.72 (71.72)	29.832 (25.014)	24.152 (21.843)	0.8769 (0.8906)	0.67 (0.71)	3.19 (2.72)	15.04 (13.97)
CDif	58.98 (42.53)	35.73 (35.73)	23.251 (6.794)	21.014 (11.461)	0.7910 (0.8357)	0.40 (0.64)	12.16 (8.71)	8.93 (7.29)
Dir	71.90 (81.15)	102.32 (102.32)	−30.423 (−21.173)	24.828 (21.388)	0.9234 (0.9424)	1.08 (1.03)	24.87 (18.89)	16.65 (14.87)
CDir	172.47 (195.36)	203.10 (203.10)	−30.634 (−7.739)	30.175 (17.971)	0.9473 (0.9827)	0.98 (0.95)	34.26 (16.56)	21.52 (12.61)
AOD	0.145 (0.071)	0.103 (0.103)	0.042 (−0.032)	0.050 (0.053)	0.8181 (0.8181)	0.78 (1.58)	−0.01 (−0.01)	0.04 (0.02)
CFsw	0.65	0.56	0.094	0.108	0.8016	0.95	−0.06	0.08
CldTau	7.21	18.59	−11.372	4.790	0.4498	0.61	14.22	3.89
Albedo	21	32	−10.4	14.1	0.8108	0.81	14	10
Clr-sky Albedo	21	33	−11.9	12.0	0.8679	0.93	13	9
Dir/Dif	0.53 (0.67)	1.12 (1.12)	−0.582 (−0.446)	0.573 (0.522)	0.6119 (0.6926)	1.01 (1.03)	0.58 (0.42)	0.40 (0.36)
Cdir/Cdif	2.38 (3.64)	4.16 (4.12)	−1.779 (−0.520)	1.286 (1.448)	0.5453 (0.5870)	0.43 (0.38)	3.13 (2.76)	0.90 (0.88)
SWnet	133.07	118.99	14.082	21.672	0.9698	0.86	4.50	13.80
CSWnet	177.81	161.01	16.793	26.411	0.9675	0.92	−1.79	18.37

[a]See Table 2 for acronym meanings. All SW components are in W/m^2, except AOD, CFsw, CldTau, and the ratio for Direct to Diffuse (Dir/Dif and CDir/CDif), which have no unit, and Albedo is in %.

[b]Mean difference is for FD minus PSO throughout this work. Nmdv is the RMS distance of all points from the regression line, and SD, standard deviation. All the values are based on arithmetic average except CldTau, which is radiative average value (essentially logarithmic, but we use ISCCP count values [*Rossow et al.*, 1996], which are based on radiation model results) from 3-hourly (FD) or 15-minute (PSO) mean. The latter is based on arithmetic average from the original 1- or 3-minute mean, however.

[c]All the values in parentheses are from FDrv, the FD-like but calculated with the only change of input aerosol with AOD 50% of the original in the standard FD.

PSO product have associated CFsw values, of which a majority (~70%) are for LST = 9, 12, and 15 hr. The results are shown for local solar noon conditions but are generally similar for all the 3-hourly mean values. For clarity, the ratio is aggregated over each CFsw bin range (= 0.05) for a total of 20 bins. The similarity of the dependence of Dir/Dif on CFsw in both datasets indicates a similar self-consistency between the partitioning of the SWdn flux and the primary

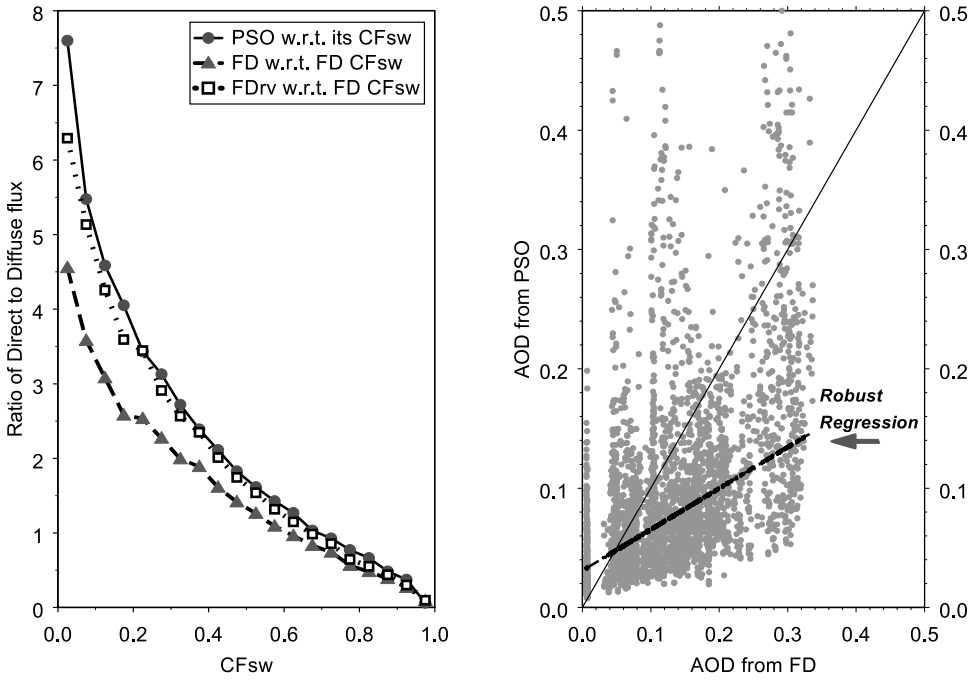

Figure 1. Left: Direct-to-diffuse flux ratios (Dir/Dif, Y) versus the associated cloud fraction (CFsw, X) at local solar noon from all 15 stations for 2004 for RFA-PSO, ISCCP-FD, and FDrv, respectively. The ratio values are obtained by averaging over each CFsw bin (= 0.05) for a total of 20 bins. Right: Scatterplot for the column aerosol optical depth (AOD) at 550 nm, Y for PSO, and X for FD, based on 3-hourly values at 10 stations (Table 1). The robust linear regression line is also shown.

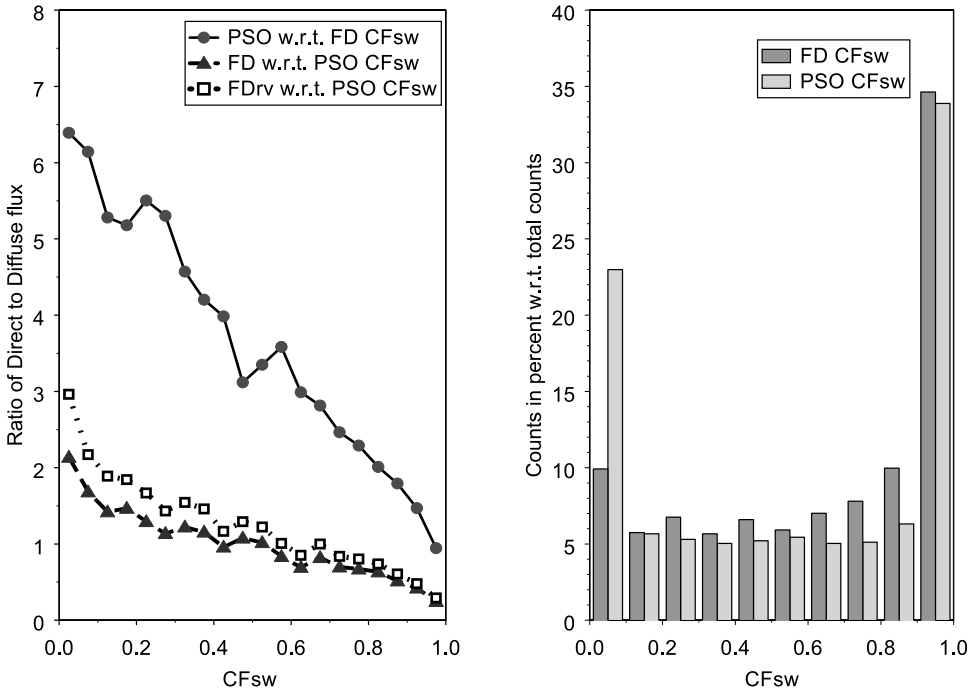

Figure 2. Left: same as Figure 1 (left), but all the ratios are versus CFsw of other datasets (i.e., PSO's ratio versus FD's CFsw and vice versa). Right: histogram of the corresponding 3-hourly (at local noon) CFsw in percentage from FD and PSO.

physical parameter affecting scattering. The quantitative agreement for the value of Dir/Dif under overcast conditions (CFsw near unity) is very good, but the FD values are systematically lower than PSO values as clear-sky conditions are approached (decreasing CFsw), suggesting the possibility of too much aerosol scattering in the FD calculations. This possibility is supported by the AOD mean bias = 0.042 in Table 3 based on 10 stations with AOD available. Figure 1R compares 3-hourly AOD values from FD and PSO with a robust fit line, which also indicates an overestimate of AOD by FD. The reason we use robust fit [e.g., *Venables and Ripley*, 1997] is to reduce the effect of the few extreme outliers in the PSO dataset; a small number of AOD values go up to 1.7 (not shown in Figure 1R), compared to a maximum of 0.3 for the FD AOD values. To determine the effect of a high bias in AOD, we use the fit results in Figure 1R to revise the FD dataset, called FDrv, by reducing all the AOD values to 50% of their original values and repeating the flux calculations for all 15 stations with all other inputs unchanged. The mean AOD difference now becomes −0.032, because the outliers are included in the PSO average. The resultant Dir/Dif values for FDrv are also shown in Figure 1L, illustrating substantially improved agreement between FDrv and PSO.

[22] A more precise mutual evaluation must account for other meteorological information (section 3). To illustrate how different the two products are, Figure 2L shows the Dir/Dif ratio from each dataset against the matched values of CFsw from the other product. The PSO ratios are now much larger even for overcast conditions, and the change of AOD does not improve the comparison much. To explain this, we compare the frequency distribution of CFsw for the two products in Figure 2R, which shows that the pointlike dataset (PSO) has more completely clear cases than the arealike dataset (FD). In other words, many of the cloudy scenes according to FD are clear over the PSO surface site [cf. *Rossow et al.*, 1993]. Thus, the CFsw distribution of FD moves the PSO Dir/Dif values from the clear sky to the cloudy category, causing a general increase of the ratio in Figure 2L.

[23] Figure 2 also implies that the results in Figure 1L are actually mixtures of different scenes. Such mixing exaggerates the discrepancies between the two products. Figure 3 shows the differences of SWdn and Dir/Dif for matched 3-hourly FD and PSO values as functions of their respective CFsw values. As expected, the smallest flux differences occur when the two datasets have the same CF_{SW} (zero difference contour is near the one-to-one line) and the largest flux differences occur when the two datasets disagree most about CFsw, as can happen for point-area comparisons like this [cf. *Rossow et al.*, 1993]. That is, some of the scatter in the comparison results comes from mismatched cloud cover conditions that are inherent in this type of comparison [section 2, see also *Rossow and Zhang*, 1995]. With MSCM, a more precise comparison is possible if we sort the flux values on the basis of the degree of agreement between the coincident meteorological information.

[24] We now approximate $\mathbf{ev_1}$ and $\mathbf{ev_2}$ as functions of a single variable, Cfsw, apply gradually stricter constraints on the allowed (absolute) differences in 3-hourly CFsw (ΔCFsw varies from 1.0, no restriction, to 0.01), and evaluate the (daily mean based) monthly means on the basis of this reduced sample. Note that with the restriction of ΔCFsw to 0.01, the sample is reduced to ~7% of the original and the remaining flux values are concentrated at LST = 9, 12, and 15 hr. Also note that if all cases where ΔCFsw >0.10 are

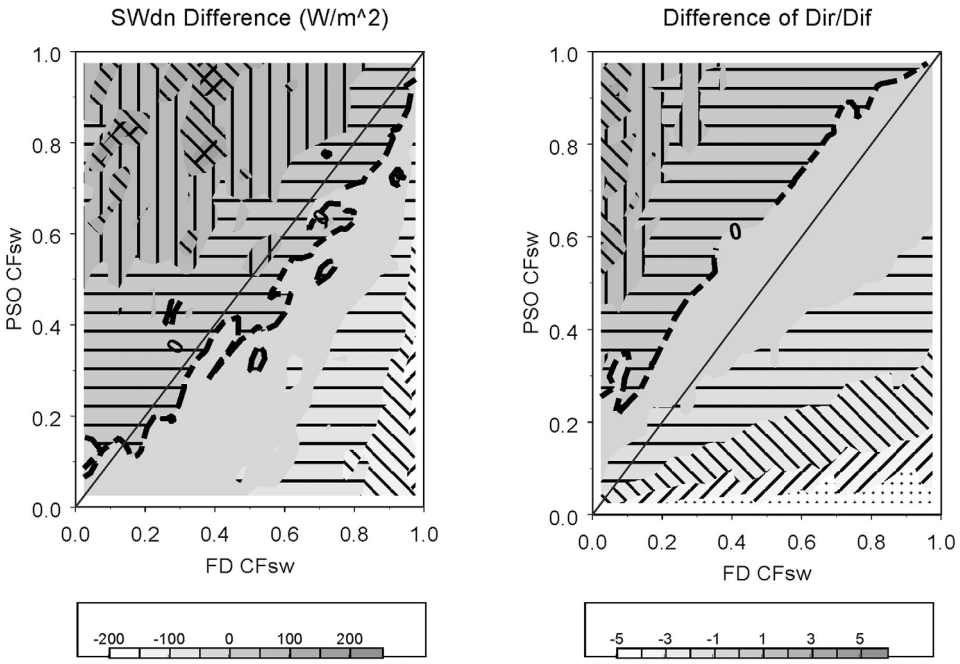

Figure 3. Contours of the differences (FD minus PSO throughout all figures) of downwelling SW flux (SWdn in W/m^2, left) and Dir/Dif ratio (right) versus CFsw of FD (X) and PSO (Y). The zero contour line (thick dashed) is displayed.

removed, almost all of the partially cloudy cases are removed and (almost all of) the remaining cases are either totally clear or overcast. The latter are five times more frequent than the former. Figure 4 shows the evolution of the mean and SD of the flux differences (in relative percentage) for Dir, Dif, and SWdn with respect to $|\Delta CFsw|$. The mean differences for Dir and Dif are reduced to −17.9 and 24.7% at $\Delta CFsw$ = 0.05 from −36.8 and 35.6% (at $\Delta CFsw$ = 1.0), respectively, a 2.33-fold reduction; but for SWdn, the difference actually increased by 8% (from −0.4 to 7.6%), still within the

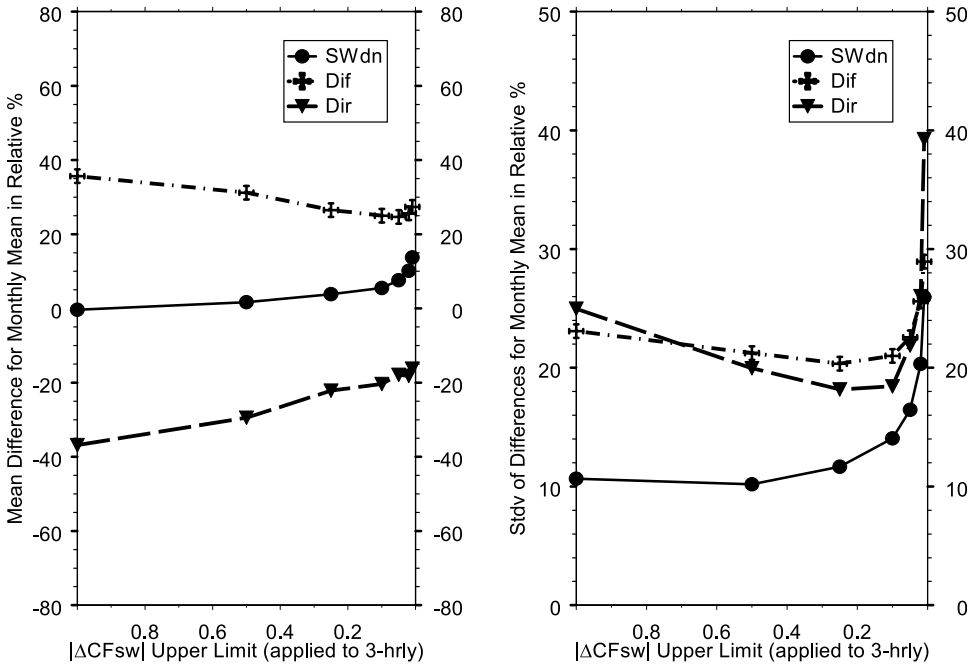

Figure 4. The mean (left) and standard deviation (right) of the differences for monthly mean SWdn (solid circle with solid line), Dif (cross with dotted-dashed line), and Dir (solid triangle with dashed line) in relative percentage versus the upper limit of absolute CFsw difference (X) between FD and PSO. The upper limit is applied to 3-hourly fluxes, which is used to produce the monthly average.

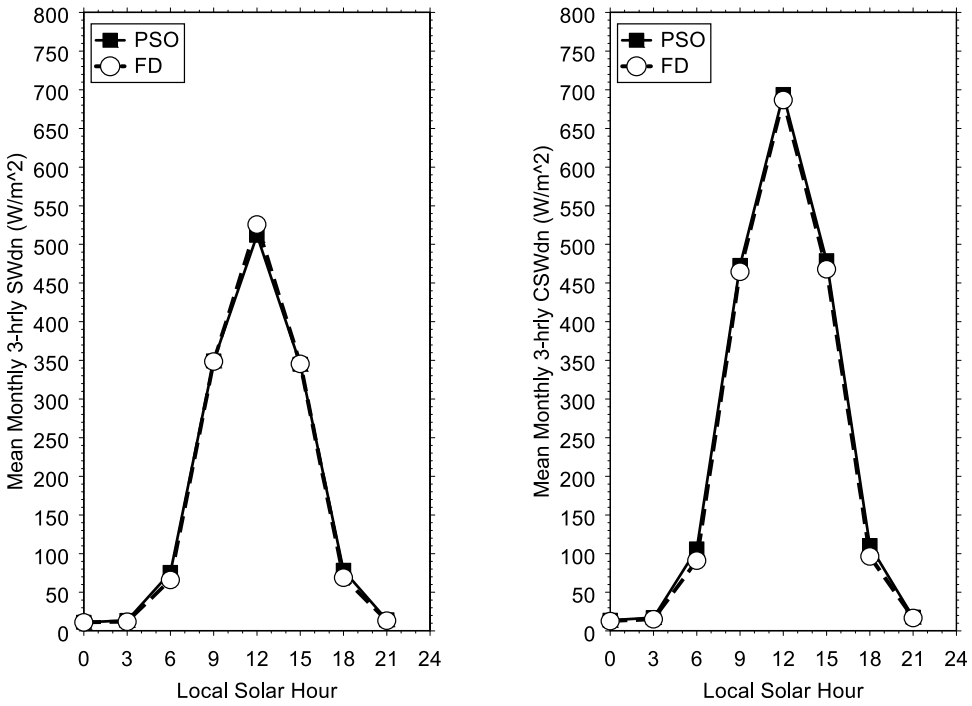

Figure 5. Diurnal variations of surface downwelling SW flux under all sky (SWdn, left, in W/m^2) and clear sky (CSWdn, right, in W/m^2) in local solar time (LST) for mean monthly–mean 3-hourly, averaged over all 2004 months and 15 stations. Solid square (with solid line) is for PSO and open circle (with dashed line) for FD.

uncertainty estimates. The SD values are reduced to 18.4 and 21.0% at ΔCFsw = 0.1 (~7 and ~2% reduction) for Dir and Dif, respectively; but for SWdn, after slight improvement, SD increases to 14.1% (at ΔCFsw = 0.1, 4% increase) from the original 10.7%. If FDrv is used, the corresponding mean (SD) of the differences decrease (except SWdn) to −4.8 (13.8), 21.2 (20.1), and 10.3 (14.0)% from −24.7 (21.4), 30.9 (21.9), and 2.2 (10.4)% for Dir, Dif, and SWdn, respectively (not shown). There are no substantial changes from daily-based to LST-based monthly means. Similar changes appear for the 3-hourly fluxes, but SDs are 2–3 times larger than for monthly. The mean (SD) of the differences are −18.8 (47.8), 27.5 (49.7), and 9.6 (35.4)% from −35.1 (82.6), 34.6 (58.4), and −0.4 (34.0)% for the three flux components, respectively, at ΔCFsw = 0.05 (0.1) as above (not shown). Because the PSO sampling of CFsw is concentrated near-noon LST, the monthly mean fluxes are based on an incomplete diurnal sampling.

[25] The optimum value of ΔCF_{SW}, which produces the minimum flux difference, is ~0.05, not 0.01 (Figure 4). This occurs because there are other quantities that also affect the flux comparison. In particular, CldTau, cloud types and smaller sampling populations begin to play more important roles in the flux comparisons between FD and PSO when ΔCF_{SW} is restricted to a smaller effect. In fact, we show in the next section that the magnitude of the flux differences can be reduced even further by applying additional matching constraints on CldTau and that the apparent increase in the SWdn bias of FD relative to PSO for overcast conditions (when ΔCFsw is small, most of the samples are overcast) arises because the average CldTau for PSO is biased high relative to FD.

[26] All the PSO clear-sky SW flux components (CSWdn, CDif and CDir) are estimated from the clear-sky identification and fitting algorithm described by *Long and Ackerman* [2000]. We can also extract the PSO flux values for truly clear scenes, when the scene is flagged as clear at the station, but have found no drastic differences with the derived PSO clear-sky values, so we do not report them. The general comparisons for the clear-sky fluxes for both FD and FDrv with PSO are also summarized in Table 3. FDrv exhibits substantially improved clear-sky downwelling fluxes, especially for the mean differences: from 23.3, −30.6 and −7.4 to 6.8, −7.7 and −1.0 W/m^2 for CDif, Cdir, and CSWdn, respectively.

4.2. Diurnal Comparison

[27] Figure 5 shows the average diurnal variation of the downwelling SW fluxes for all and clear sky over eight 3-hourly local solar times from FD and PSO. We have also looked at the 1-hourly diurnal cycle from PSO, but there is no particular difference between it and the corresponding 3-hourly results, except that the variation is smoother. Although the average fluxes agree quite well, this figure shows a systematic tendency for FD to slightly underestimate the clear-sky fluxes at all times of day relative to PSO, consistent with the clear-sky bias shown in the previous section.

[28] To investigate the effects on the diurnal cycle of SW fluxes by CFsw and CldTau, the 3-hourly ("normal" hereafter) datasets we have used thus far cannot be used, because the fluxes and other parameters have different time sampling. PSO determines optical depth only for overcast skies (sky cover ≥0.90) at the original 1- to 3-minute level, which

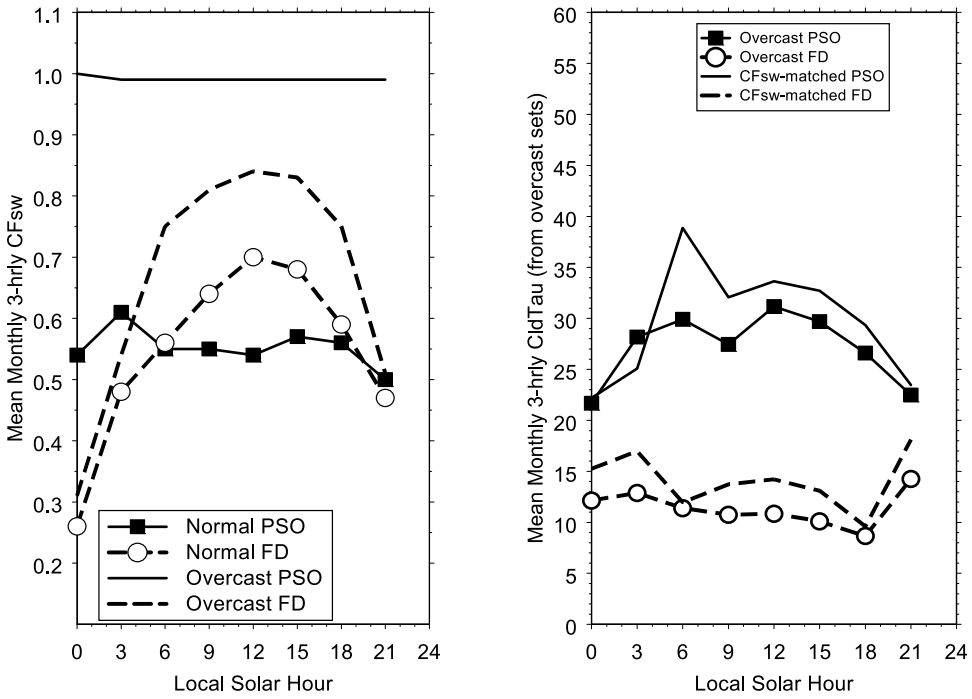

Figure 6. Left: Diurnal variations of CFsw from the normal and overcast datasets; solid square (with solid line) is for normal PSO, open circle (with dashed line) for normal FD, and sole solid and dashed lines are for overcast PSO and FD, respectively. Right: cloud optical thickness (CldTau) from the overcast datasets (see section 4.2); solid square (with solid line) is for PSO, and open circle (with dashed line) for FD Also shown are CFsw matched CldTau for ΔCFsw $\leq$0.10 from the two overcast datasets in sole solid and dashed lines for PSO and FD, respectively.

leads to inhomogeneous sampling for 15-minute and longer time averages. We produced an alternative 3-hourly dataset for PSO in which all parameters (including fluxes) have the same sampling as CldTau at the 15-min level; we also redid the FD set to match. We call these new datasets the overcast (or CldTau-matched) datasets. The overcast dataset is 44% of the normal datasets. Figure 6L compares the diurnal variations of CFsw between FD and PSO for both the normal and overcast sets. FD shows a notable diurnal cycle with increased magnitude for the overcast CFsw, but PSO has virtually no diurnal cycle for either subset. Remember that the PSO retrieval scheme greatly reduces the number of samples away from midday, which may explain a lack of diurnal variations near dawn or dusk. The overcast CFsw is ~1.0 because of the restricted conditions on CldTau in the PSO product, but Figure 6L also shows that the FD CFsw is different from the PSO for overcast scenes, which is not obvious in Figure 2R, where the overcast populations are not so different. In contrast, Figure 6R shows the reverse situation for CldTau: PSO has obvious diurnal variations but FD does not for the overcast set. Again, the strong reduction of samples away from midday in the PSO product may explain this behavior. In any case, Figure 6 indicates that the two analyses (ISCCP FD and RFA-PSO) are partitioning the solar radiation information into these two cloud parameters in different ways.

[29] The overcast sets give the mean diurnal variations for SWdn and Dif shown in Figure 7. The agreement in the variation with time of day is very good on average, despite the poor agreement of the mean CFsw and CldTau values (Figure 6). FD has larger SWdn and Dif by up to >80 W/m^2 at LST = 12 hr. The Dir difference is generally $\leq$10 W/m^2 except for a difference of ~40 W/m^2 at LST = 0 hr at a very high-latitude location with extreme solar zenith angles (not shown). These differences can be reduced if we apply matching restrictions on both CFsw and CldTau. Figure 8 shows the differences for SWdn (left) and Dif (right) as functions of the differences of CFsw (X) and CldTau (Y) from the overcast 3-hourly sets. Figure 8R suggests that when both the CFsw and CldTau differences are nearly zero (difference for CFsw $\approx$ 0 and CldTau ≈ -5), the Dif difference is also nearly zero, and when both the CFsw and CldTau differ, the Dif difference values diverge, increasing toward the right bottom corner where the CFsw (CldTau) difference becomes positive (negative) maximum. If the CldTau difference is held constant, increasing the CFsw difference (most changes are from FD because PSO CFsw $\geq$0.90) increases the Dif difference, which is physically plausible. On the other hand, if the CFsw difference is held constant, increasing the CldTau difference decreases the Dif difference because higher (lower) CldTau for FD (PSO) means less (more) Dif for FD (PSO). For Figure 8L, the SWdn difference for the overcast cases depends mostly on the CldTau difference (as expected) with the zero-difference contour at around -5 for the CldTau difference (including ΔCFsw $\approx$ 0). The CldTau influence is dominant for this subset because the strong monotonic decrease of both Dif and Dir when CldTau increases (not shown) while CFsw is constrained to $\geq$0.90, the overcast scenes for PSO, limiting the variations of CFsw to only a small effect. Again, we

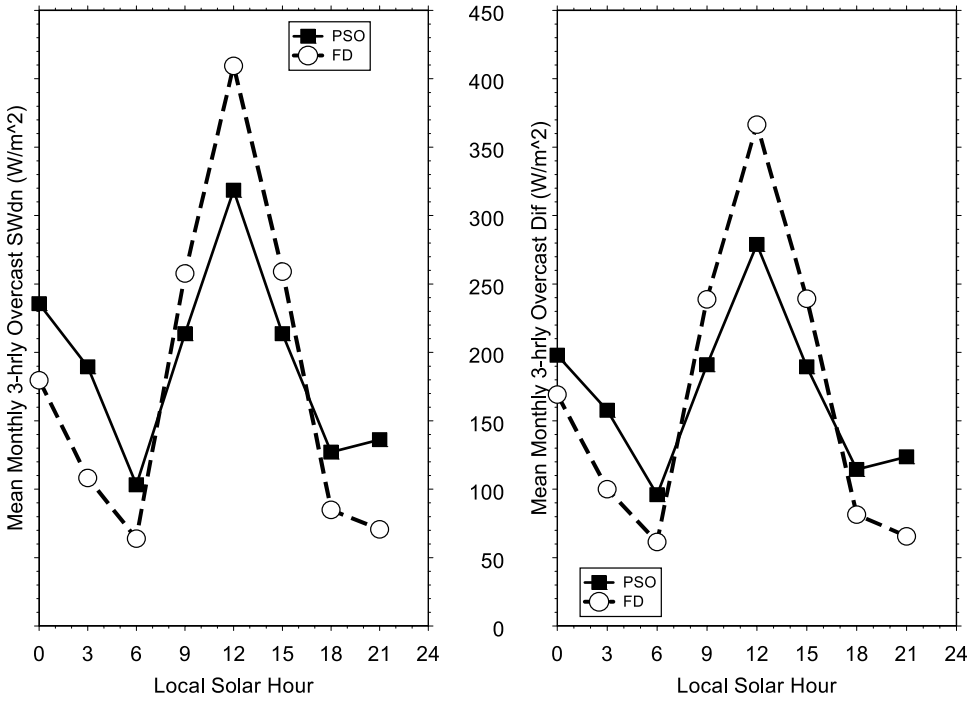

Figure 7. Diurnal variations of SWdn (left, in W/m^2) and Dif (right, in W/m^2) for the overcast datasets. Solid square (with solid line) is for PSO and open circle (with dashed line) for FD.

see the effect of an area-point comparison where, even though the surface site indicates overcast conditions, the larger area is not always so cloudy.

[30] Thus, reducing the differences for both CFsw and CldTau to optimum values (not zero because of other unaccounted factors) decreases the SW flux differences to a minimum. This is the two-variable approximation for $\mathbf{ev_1}$ and $\mathbf{ev_2}$ as discussed in section 3. When applying the joint CFsw and CldTau constraints to the overcast 3-hourly fluxes for $ev_1 \rightarrow ev_2$ as MSCM is applied, the best (approximate minimum value) mean (SD) of the differences for Dir, Dif, and SWdn appear at ΔCFsw $\leq$0.20 and ΔCldTau $\leq$10 (with

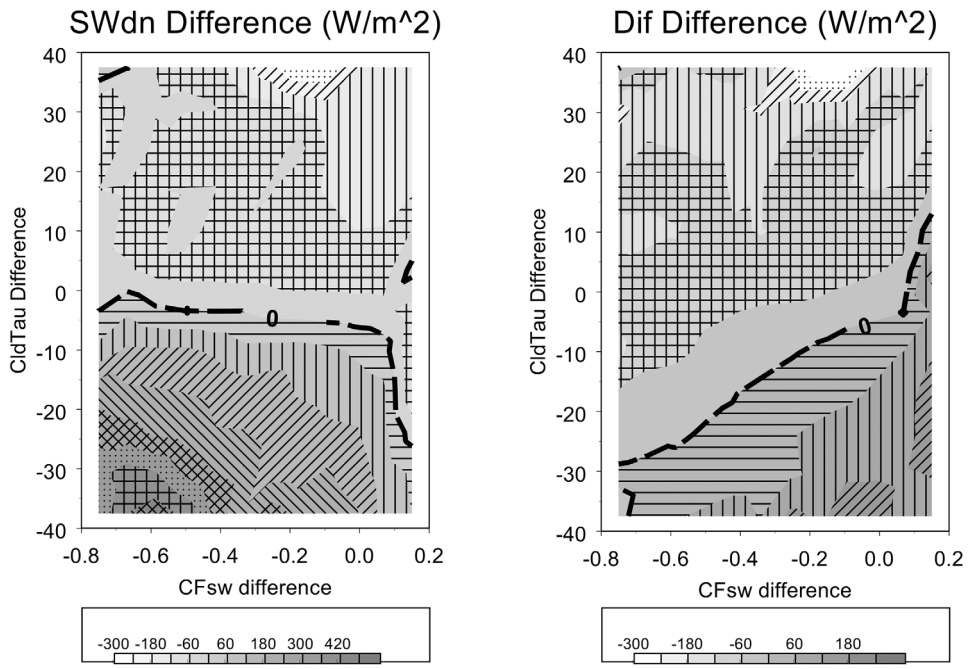

Figure 8. Contours of the differences of SWdn (left, in W/m^2) and Dif (right, in W/m^2) versus the difference of cloud fraction (X) and difference of cloud optical depth (CldTau, Y) based on the overcast datasets.

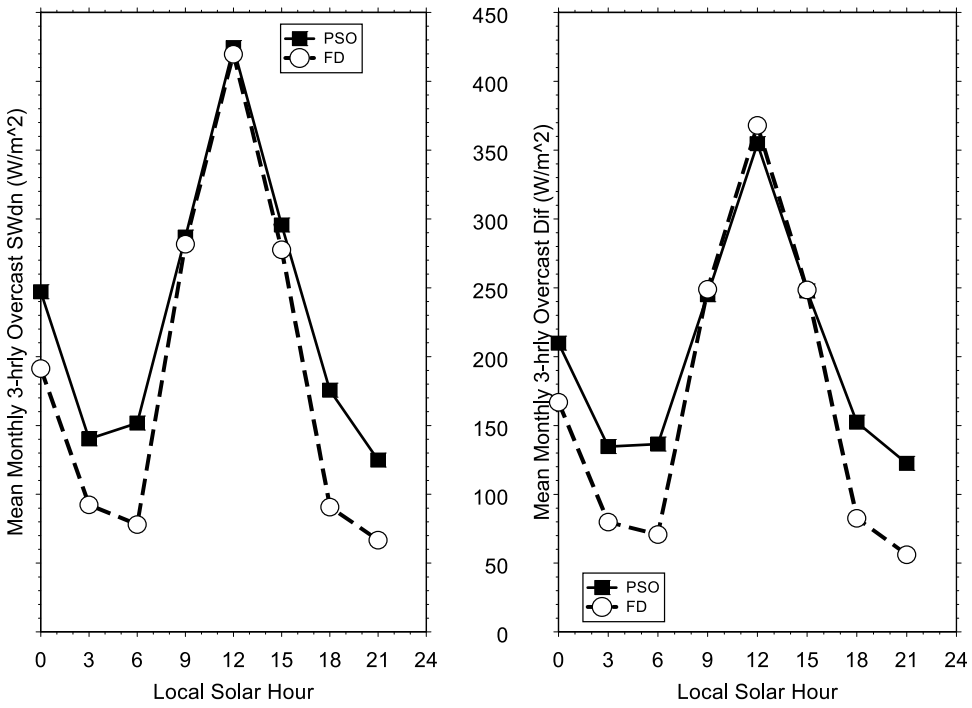

Figure 9. Diurnal variations of SWdn (left, in W/m^2) and Dif (right, in W/m^2) under the double constraints for ΔCFsw $\leq$0.20 and ΔCldTau $\leq$10, based on the overcast datasets (see section 4.2). Solid square (with solid line) is for PSO and open circle (with dashed line) for FD.

30% of overcast set samples): −14 (58), −14 (82), and −28 (104) W/m^2, respectively, compared with 53 (127), 12 (118), and 64 (173) W/m^2 before applying the constraints. Figure 9 shows the diurnal cycles for SWdn and Dif with the best joint constraints. FD and PSO have near-perfect agreement for LST = 9 − 15 hr, which is the majority (~70%) of the data. The Dir/Dif ratio difference also comes within ~0.05 for all local hours except LST = 3 and 21 hr (only 1% of all values for the polar regions) and the Dir differences $\leq$15 W/m^2 (not shown).

[31] In Figure 6R, we also show an additional pair of CldTau values obtained when ΔCFsw $\leq$0.1 ("CFsw-matched CldTau" in the figure). Both pairs show that the PSO's CldTau is systematically higher than FD by 10–20, which helps explain the increased bias of SWdn in Figure 4 as ΔCFsw is restricted and concentrated to overcast scenes. To estimate CldTau, the RFA method [*Long et al.*, 2006] must first estimate sky cover. To do so, it first eliminates the overcast cases when 10% or less of the total SW comes from the direct component. This procedure requires CldTau to be roughly 5–6 or more. For cases with lower optical depth, the enhancement of the diffuse over clear-sky equivalent indicates less than complete sky cover. Thus there is a tendency for the sky cover of optically thin (CldTau <3) overcast to be underestimated. Given that the PSO CldTau is only retrieved for (detectable) local overcast scenes with CFsw $\geq$0.9, assuming a single uniform cloud layer, cases of optically thin overcast of roughly 2 or less are effectively eliminated during the optical depth retrievals. This results in a high-biased aggregate PSO CldTau. Second, the linear averaging of CldTau from 1 or 3 minutes to 15 minutes means it also exaggerates CldTau values in terms of its radiative effects. This effect can be as much as a factor of 2–3, depending on the space-time scales involved. In contrast, CldTau values in FD have been averaged giving equal weight radiatively. The larger differences in CldTau indicate that improvement in the comparison by matching CFsw alone is limited; stricter matching can produce less agreement as shown in Figures 4 and 9. As discussed in section 3, the trial-and-error application of MSCM only gives optimum (but nonzero in difference) values of $\mathbf{ev_1}$ and $\mathbf{ev_2}$ that minimize the flux differences.

[32] Figure 10 shows the diurnal variations of CDir and CDif comparison with PSO for FD and FDrv, respectively. The FDrv values agree much better with the PSO values for both the clear-sky direct and diffuse components and their diurnal variations. Thus, the discrepancies in clear-sky SW fluxes (including CDir and CDif) and their diurnal variations can be mostly explained by adjusting the AOD used in the FD calculations.

[33] Many previous studies have found model calculations of clear-sky diffuse fluxes to overestimate observed values [e.g., *Halthore and Schwartz*, 2000]. *Nowak et al.* [2008] also reviewed this model-observation discrepancy, saying that modeled clear-sky downward SW broadband fluxes at the surface during the 1991 FIRE experiment were up to 10% larger than corresponding measurements, of which 4% could be attributed to disagreements in diffuse SW flux among the surface instruments. All the stations for Dif and Dir measurement are equipped with a ~5.5° shading disk, with a field of view (FOV) matched to the FOV of the pryheliometer such that when the Dir and Dif are summed, they give the SWdn. Because the 5.5° FOV is larger than the extent of the solar disk, a significant share of the diffuse flux

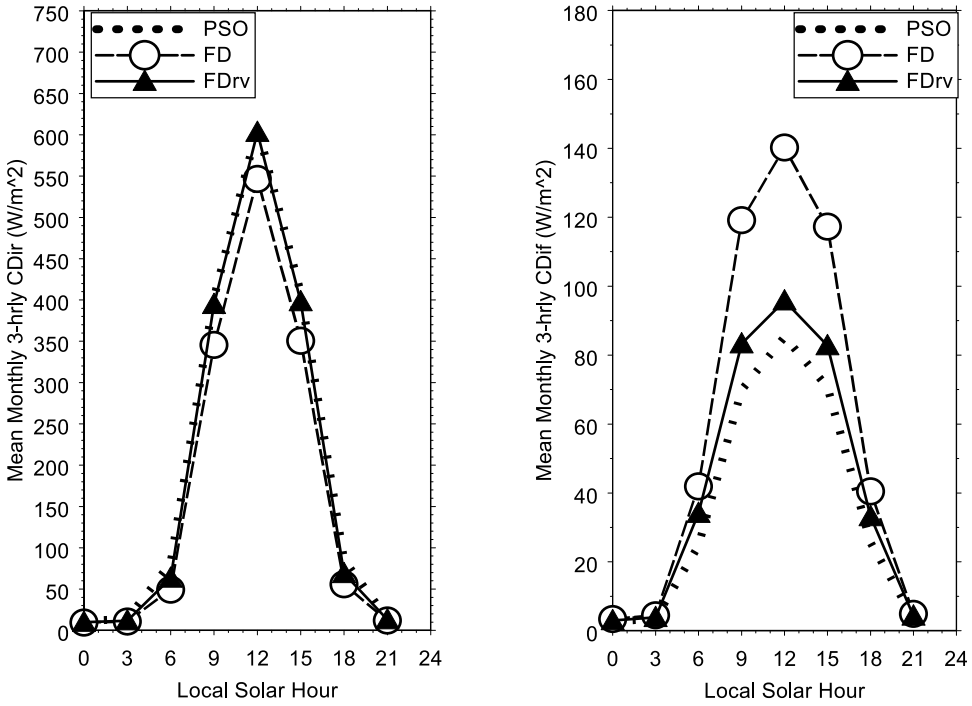

Figure 10. Diurnal variations of clear-sky direct (CDir, left, in W/m^2) and clear-sky diffuse (CDif, right, in W/m^2). Thick dotted line is for PSO, open circle (with dashed line) for FD, and solid triangle (with solid line) for FDrv.

is counted as direct by this large aperture and thus is missed by the shaded diffuse instruments. The underestimate (overestimate) of diffuse (direct) is <10 W/m^2 for low aerosol optical depths and in the extreme cases (AOD ≥1), it may reach >100 W/m^2 [*Grassl*, 1971]. However, for undetected cirrus clouds, PSO mislabels these scenes as clear sky. Therefore, the reduction of AOD values actually needed to match to PSO is likely to be <50% because there are a few other causes of the differences with FD, including PSO's slight underestimate of diffuse irradiance. More detailed study of this issue is needed because the MSCM-based corrections of AOD are primitive.

5. Longwave Flux Evaluation

[34] Our previous comparisons of FC (FD's precursor) and FD surface LW fluxes with surface measurements showed larger biases and RMS differences than for SW fluxes that appeared to be associated mostly with the poor accuracy of the atmospheric temperature (and humidity) dataset used in the calculations, rather than with errors associated with cloud sampling or cloud base heights [*Rossow and Zhang*, 1995; *Zhang et al.*, 2004]. In particular, we found that there was too much variability of the FC downwelling LW fluxes in the tropics, associated with too much atmospheric temperature variability attributed to measurement errors, but too little variability in FC downwelling fluxes at higher latitudes, caused in part by a lack of diurnal variation of the atmospheric temperatures used in the calculations [*Rossow and Zhang*, 1995]. To reduce the latter problem in FD, we introduced a climatology-based diurnal model for atmospheric temperatures and also estimated a reduction in the diurnal surface air and skin temperature diurnal amplitudes for cloudy situations. Note that errors in the atmospheric properties also propagate into errors in the surface skin temperature retrieved in the ISCCP analysis [cf. *Zhang et al.*, 2007]. Although we compared the resulting diurnal cycles in surface air and skin temperatures to other datasets in *Zhang et al.* [2007], we can make a more precise evaluation of both FD and PSO using the RFA results with MSCM.

5.1. General Comparison, Sampling and Cloud-Aerosol Effects

[35] Table 4 shows comparisons of FD and PSO monthly mean LW-related quantities. The mean (SD) differences for downwelling all-sky (LWdn) and clear-sky (CLWdn) and upwelling all-sky (LWup) and clear-sky (CLWup) LW fluxes, respectively, are 10.0 (15.6), 5.5 (13.3), 5.9 (17.8) and −1.1 (21.4) W/m^2, or, in relative terms, 3.2 (4.9), 1.9 (4.6), 1.7 (5.0), and −0.3 (6.0)%. Such differences are consistent with our previous evaluations with respect to BSRN data, which yielded mean (SD) LWdn differences of 2.2 (19.0) W/m^2 or 0.7 (6.3)%, based on 1,831 monthly mean samples over 35 BSRN stations for 1992–2001 [*Zhang et al.*, 2004]. However, the current LWdn mean difference is somewhat larger than previously owing to a spurious increase in the TOVS near-surface air temperatures by about 2–3 K in late 2001 that affected all subsequent years, including the particular one (2004) used here (see next paragraph).

[36] We used the trial-and-error MSCM to sort the spatiotemporally matched LWdn to see whether the flux differences can be explained. The two variables that seem to show clear relationships with LWdn differences are the

Table 4. Comparison of Monthly Mean Surface LW Flux Components and Associated Meteorological Parameters[a] Between FD (X) and PSO (Y)

Variable Name	FD (X)	PSO (Y)	Mean Difference	SD	Correlation Coefficient	Slope	Intercept	Normal Deviation
LWdn	321.02	311.06	9.957	15.551	0.9870	1.09	−38.89	9.37
CLWdn	293.63 (293.05)[b]	288.12 (288.12)	5.506 (4.926)	13.266 (13.372)	0.9883 (0.9881)	1.00 (1.00)	−6.67 (−6.28)	9.36 (9.43)
LWup	360.68	354.74	5.944	17.858	0.9816	1.02	−12.63	12.46
CLWup	358.08	359.14	−1.062	21.374	0.9741	1.00	0.28	15.10
CFlw	0.63	0.45	0.177	0.109	0.7633	0.81	−0.06	0.08
CldHgt	4.74	2.76	1.982	1.688	0.5290	0.26	1.52	0.81
Ta	284.68	281.41	3.264	7.398	0.9715	1.35	−103.75	3.00
SLSH or SASH	7.98	8.01	−0.037	4.165	0.7610	0.94	0.51	3.03
LWnet	−54.32	−59.13	4.816	14.610	0.8874	0.80	−15.82	10.26
CLWnet	−81.64	−88.22	6.582	19.732	0.6891	0.64	−35.74	14.68

[a]All the listed parameters are for monthly mean values for 15 stations for 2004, similar to SW as explained in Table 3. All LW components are daily mean removed (so their mean is zero) in W/m^2. CFlw is unitless. CldHgt is in km, Ta in Kelvin, SLSH (FD's surface layer humidity) and SASH (PSO's surface air humidity) in g/kg.

[b]In the parentheses are for FDrv versus PSO; FDrv is defined in Table 3.

surface air temperature (Ta) and cloud fraction (CFlw). For PSO, CFlw is "LW effective sky cover," which primarily represents low and middle cloudiness [see *Dürr and Philipona*, 2004; *Long and Turner*, 2008], whereas for FD CFlw is the same as CFsw but now matched in time to PSO at the 3-hourly level. Note that in contrast to CFsw, PSO has CFlw values for virtually all LWdn values, so sampling of the LW diurnal cycle is homogeneous. Figure 11 shows LWdn difference as functions of Ta (left) and CFlw (right) for FD (X) and PSO (Y) as before. For these 3-hourly averaged flux values, the Ta values from FD are systematically 2–3 K larger than PSO, which produces up to 25 W/m^2 difference. However, when the values of CFlw and Ta agree, the values of LWdn agree as well (zero contour located along the 1-to-1 line). When we apply MSCM to match CFlw, LWdn agreement improves from ≤10 W/m^2 to near zero as ΔCFlw decreases from 1.0 to 0.01, but there is virtually no change of SD of the differences (not shown), indicating that the time variations in the flux differences are controlled more strongly by another quantity. The time variation of LWdn differences are affected more by Ta disagreements as shown in Figure 12. As the Ta difference is more strongly constrained from 50 down to 0.2 K, the mean (SD) flux difference decreases from 10 (35) to 1 (~25) W/m^2 for LWdn, and similarly for CLWdn, going from 5.5 (26) to < 1 (< 19) W/m^2. This shows that Ta is a more important parameter in $\mathbf{ev_1}$ and $\mathbf{ev_2}$ than CFlw for the LW fluxes. Because Ta is the leading parameter affecting the LW fluxes, the variations in the LW agreement have a stronger geographic and seasonal dependence on Ta differences than CFlw. In particular, we have investigated the cases where FD's Ta values are very highly biased (up to 50 K). These

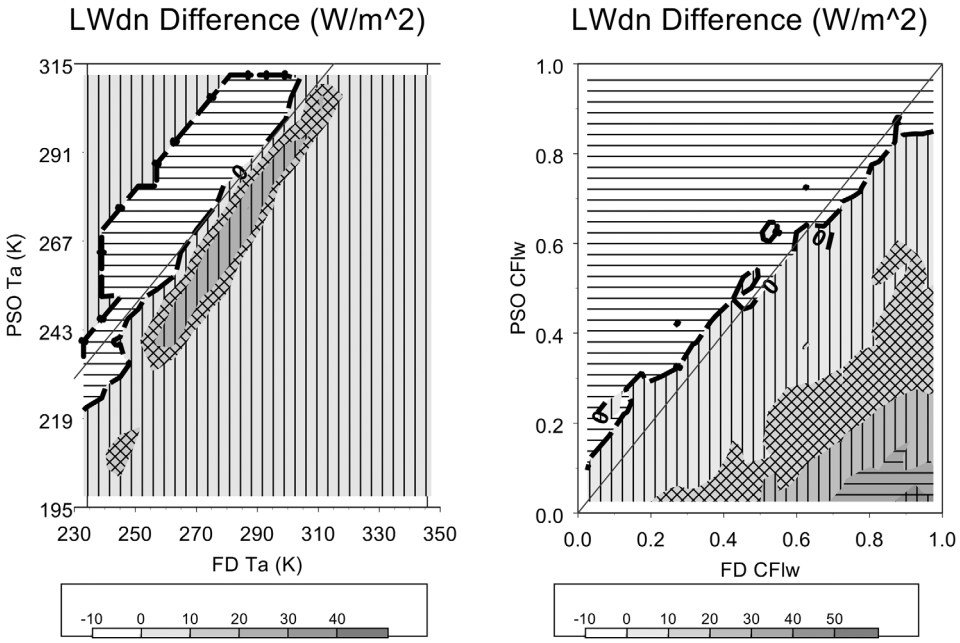

Figure 11. Contours of the differences of downwelling LW flux (LWdn, in W/m^2) versus surface air temperature (Ta, left) and versus cloud fraction (CFlw, right) for FD (X) and PSO (Y).

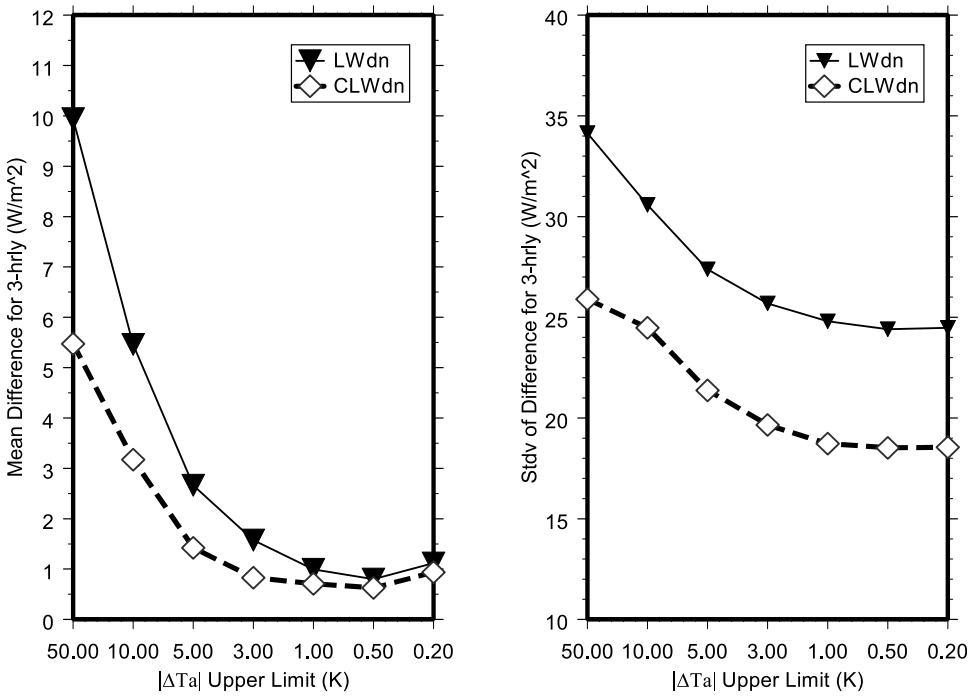

Figure 12. Mean (Y, left) and standard deviation (Y, right) of the difference for 3-hourly, all-sky (triangle with solid line) and clear-sky (open diamond with dash line) downweling LW (in W/m^2) versus upper limit of (absolute) surface air temperature allowed (Ta, X).

cases are all from the South Pole (SPO) station comparisons and are caused by the lack of temperature inversions in the TOVS temperature profiles in winter. The high bias of the Ta values used in FD, 2–3 K is larger in winter overall (up to 12 K at LST = 15 for Bondville (BON) station in January and at SPO, 27 K, not shown) than in summer (usually <6 K, not shown); the agreement at most other locations is better than ±3 K but slightly larger at higher latitudes than at lower latitudes.

[37] Aerosols are usually considered important only for solar radiation, and there are very few publications concerning their effects on LW fluxes, but they do modify the terrestrial radiation, particularly close to the ground, because they absorb and reemit the earth's infrared radiation [*Goody and Yung*, 1989] and, in general, increase surface downwelling LW [e.g., *Miller and Tegen*, 1999]. The comparison of clear-sky downwelling LW (CLWdn) from FD and PSO reveals a small but systematic effect of the aerosols. FDrv does improve the CLWdn agreement by ~10% (from 5.5 to 4.9 W/m^2) in mean difference as listed in Table 4, but has little effect on cloudy-sky LW fluxes.

5.2. Diurnal Flux Comparison

[38] To study the diurnal variations of LW fluxes and associated physical variables, we use their monthly mean 3-hourly LST averaged over all the stations and months in 2004 with the "daily mean" removed, i.e., the diurnal anomaly, shortened to "mean monthly diurnal." Because we do not find substantial disagreements between FD and PSO here and we have already shown that the uncertainties come from FD's inputs [*Zhang et al.*, 2006, 2007], we did not apply the MSCM here.

[39] We first compare the mean monthly diurnal variations of the parameters that affect the FD and PSO LW fluxes. Figure 13 compares the mean diurnal anomalies of Ta (and also the surface skin temperature, Ts, from FD, which PSO does not supply), CFlw, the surface layer or surface air specific humidity (SLSH or SASH for FD and PSO, respectively) and the cloud base heights (CldHgt). Figure 13a shows that the phase and amplitude of the Ta diurnal cycles for FD and PSO are in excellent quantitative agreement. The diurnal cycle of Ts from FD exhibits a peak at LST = 12 hr, larger than and 3 hr ahead of the peak value of Ta at LST ≈15 hr, which is physically plausible because Ta is increased by exchanges of energy from the solar-warmed surface (the precise lag time can be <3 hr in the averages used). On the other hand, as with CFsw (Figure 6L), CFlw from FD shows a diurnal variation (with peak cloudiness at LST = 12–15 hr), whereas CFlw from PSO shows no apparent diurnal variation (Figure 13b). Note that although CFlw from FD is the same as CFsw, the diurnal variations shown in Figures 6 and 13 are different due to the matching of FD to the inhomogeneous diurnal sampling in PSO for SW parameters. The FD diurnal cycle of cloudiness differs from PSO because of differing cirrus detection sensitivities [cf. *Jin et al.*, 1996 for ISCCP and *Dupont et al.*, 2008 for PSO] and because high-level clouds have a different diurnal phase than low-level clouds [cf. *Cairns*, 1995].

[40] Figure 13c shows fair agreement in the small diurnal variations of near-surface water vapor. Figure 13d shows nearly opposite phases of the diurnal variations of CldHgt. There is also an overall 2 km high bias in CldHgt for FD (Table 4). Both of the CldHgt differences can be explained by the different sensitivities of the two analyses to cirrus clouds: in particular, high clouds are not included in the

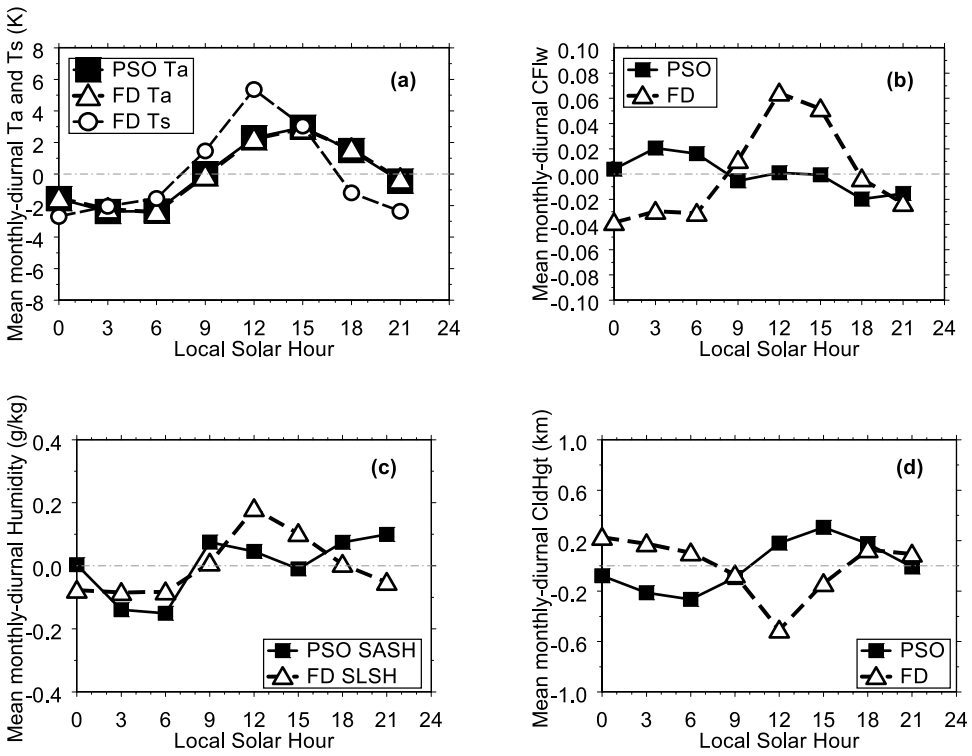

Figure 13. Diurnal variation for (a) surface air temperature (Ta) and surface skin temperature (Ts, for FD only): solid square (with solid line) is for PSO, and open triangle (with dashed line) for FD, with additional Ts from FD (open circle with dashed line); (b) cloud fraction (CFlw), (c) surface air (PSO) and layer-mean (FD) specific humidity (SASH/SLSH), and (d) cloud base height (CldHgt). Note that "daily mean" is removed for all mean monthly-diurnal LW fluxes and their associated parameters (see section 5.2).

LW-based PSO retrievals, which shifts the average base height downward and changes the average diurnal variation [*cf*. *Cairns*, 1995]. The CldHgt determinations in both datasets are very different: PSO uses LW fluxes over the day that are most sensitive to lower-level cloud variations with an assumption of single-layer liquid clouds and a fixed temperature lapse rate (see Appendix), whereas FD uses a climatological set of vertical cloud structures that vary with the cloud types present in the scene [*Rossow et al.*, 2005]. *Zhang et al.* [2004] also showed that comparisons of average cloud base heights tend to exaggerate the differences between FD and PSO because of the inclusion of more (thin) upper-level cloudiness in the former. The small effect that these differences, except for Ta, have on the LWdn comparison supports this interpretation (see next paragraph).

[41] Figures 14a and 14b show the average diurnal cycle of LWdn and CLWdn, respectively. FD has a larger amplitude for LWdn than PSO by ~3 W/m^2 with a slight phase difference, shifted more towards LST = 12 than 15 hr (given the 3-hourly window the difference in phase is actually <3 hr). Since the CLWdn phase and amplitude agree near-perfectly (Figure 14b), the differences in LWdn are likely explained by the disagreements in the diurnal variations of CFlw and CldHgt: The slight amplification in LWdn is consistent with increased CFlw and decreased CldHgt in FD (Figures 13b and 13d). The differences in amplitude and phase for LWdn, averaged over all stations, are nearly independent of season: The amplitude difference is slightly smaller in wintertime (not shown). The geographic dependence is more complex: A similar amplitude-phase difference is exhibited at the mid- to high-latitude stations, but in the tropics, the diurnal amplitude difference is smaller and the phase of FD is shifted to slightly earlier in the day (not shown). Figures 14c and 14d show the same comparison for upwelling LW (LWup) and the net LW loss (opposite sign to LW net flux), respectively. For all sky, the latter quantity is sensitive to the phase differences of LWdn and LWup. LWup is primarily determined by Ts: The FD and PSO values are in phase with near-coincident peaks at LST = 12 hr but there are various differences at other LST, up to >10 W/m^2 (LST = 3). Although PSO did not report values of Ts, such flux differences suggest a slightly low bias of the FD Ts values by <1 K during daytime and a high bias of about the same amount just before dawn. This difference is nearly uniform over seasons but varies somewhat geographically: The midlatitude stations exhibit similar difference patterns but with much smaller flux differences (~5–10 W/m^2), whereas the tropical stations show a pattern with larger differences (~10–20 W/m^2) (recall that we are comparing a point to an area and the tropical sites are islands). The resulting diurnal variation of the net loss by LW is smaller in FD than that from PSO.

[42] To evaluate the diurnal adjustment scheme applied to temperatures in FD, Figures 15a and 15b compare PSO-defined clear sky, referred to as "true clear sky," and cloudy sky (when PSO CFlw >0.7) for LWdn and Ta, respectively.

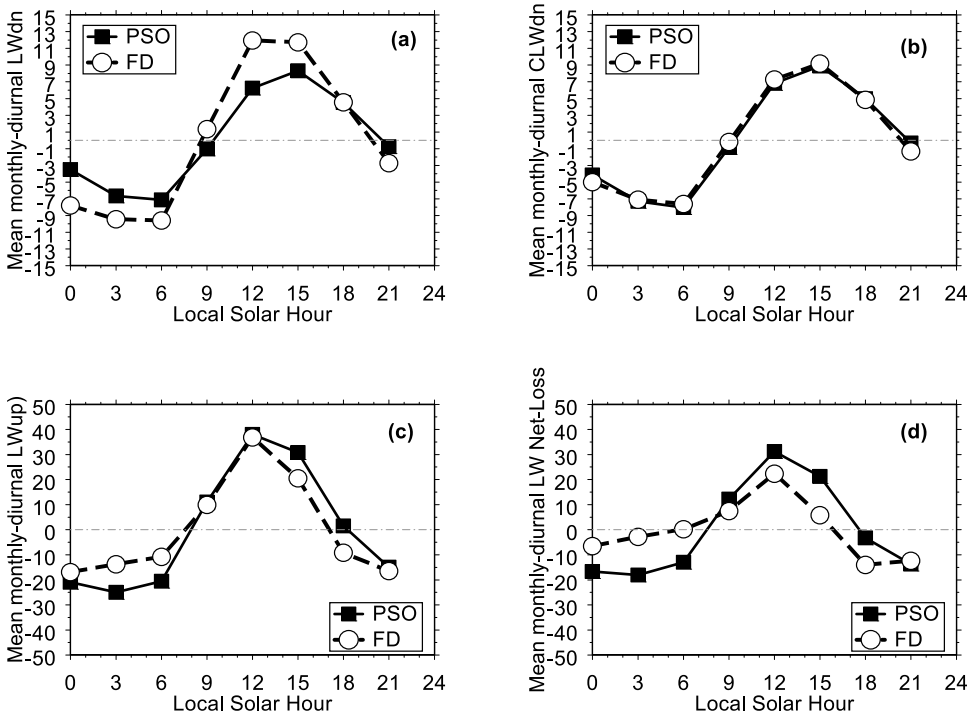

Figure 14. Diurnal variations of (a) downward LW (LWdn), (b) clear-sky downward LW (CLWdn), (c) upwelling LW (LWup), and (d) net-loss LW. Solid square (with solid line) is for PSO, open circle (with dashed line) for FD. All are in W/m^2.

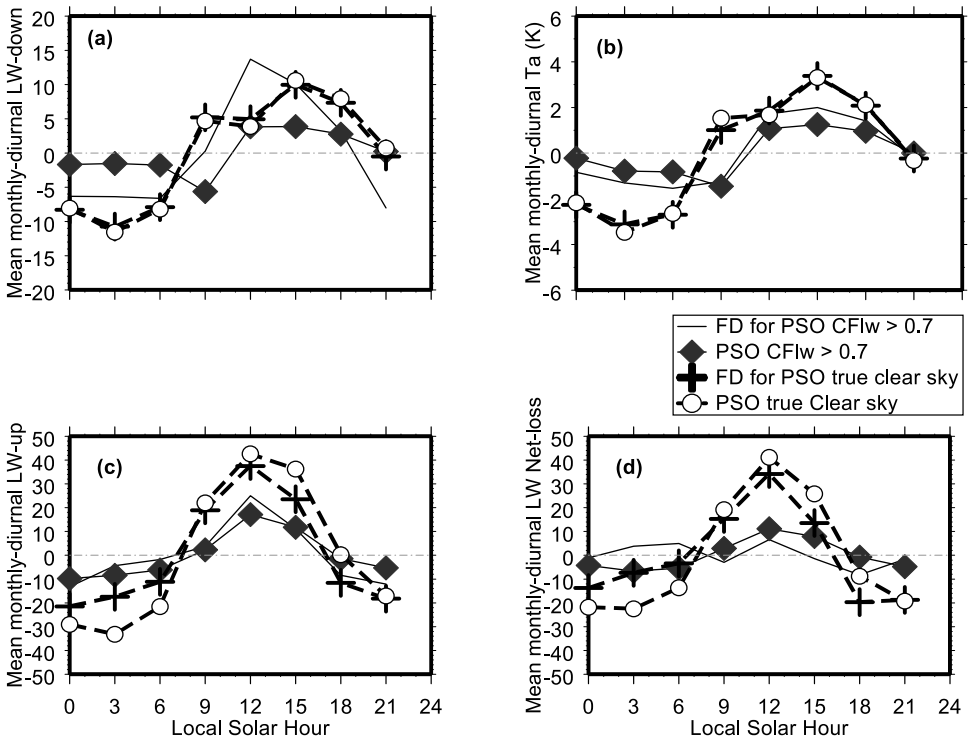

Figure 15. Diurnal variations of (a) downward LW, (b) Ta, (c) upward LW, and (d) net-loss LW under cloudy sky (defined as PSO CFlw > 0.70) and true-clear sky as observed by PSO. Solid diamond (with solid line) is for PSO cloudy sky, open circle (with dashed line) for PSO true-clear sky, solid line is for FD cloudy sky, and cross (with dashed line) for FD true-clear sky. All are in W/m^2 except Ta, which is in K.

FD and PSO are in excellent agreement for both LWdn and Ta for true clear sky. The agreement is not as good for cloudy sky: FD has larger amplitude of >5 W/m^2 and <1 K than PSO, for LWdn and Ta, respectively, which may be caused by FD's use of a climatology of diurnal amplitude and a crude representation of the cloud effects [*Rossow et al.*, 2005] as well as the differing sensitivity to cirrus clouds for both. Figure 15c shows the reverse situation for LWup and net LW loss: FD and PSO are in good agreement for cloudy sky but not for true clear sky, for which FD has slightly smaller amplitude for upward LW than PSO by up to >10 W/m^2 (LST = 3 hr) but generally by ~5 W/m^2. As discussed before, FD is more likely to have cloudy scenes than PSO (Figure 2R). The combined effects appear in Figure 15d: FD has smaller amplitude than PSO for both cloudy and true clear-sky scenes with larger differences in near dawn for net LW loss. Nevertheless, the agreement shows that the FD diurnal adjustment scheme for land Ta and Ts was an improvement over the previous FC results: The larger amplitude for all- and cloudy-sky LWdn and Ta and smaller amplitude for LWup for all the scenes (all, cloudy, and clear sky) suggests that further refinement may be possible.

[43] Putting all the factors together, net LW loss from FD is about ≤10 W/m^2 smaller in diurnal amplitude than PSO for all, cloudy, and true clear sky, caused by both clouds and surface skin temperature differences. Table 4 shows that the average loss over the diurnal cycle for FD is nearly 5 W/m^2 less than PSO averaged over all stations and seasons, showing less heat loss during the peak temperature period near local solar noon and more heat loss in the predawn hours, consistent with high-biased Ta and low-biased Ts for FD.

6. Summary

[44] The usual flux product evaluation does not have available or use much of the associated meteorological information and, even if it does, the information is usually not as systematic or complete as in the RFA-PSO product. Our study illustrates a way to evaluate a satellite flux product, namely, MSCM that is more meaningful and provides simultaneous evaluations of both the ISCCP-FD and RFA-PSO products. The optimum matching criteria vary for different flux components, reflecting the multiplicity of factors that produce apparent flux discrepancies between the two products.

[45] In general, we confirmed our previous results for the monthly averaged, all sky SW and LW fluxes: The uncertainty remains 10–15 W/m^2 [*Zhang et al.*, 2004]; if we attribute some of the error to the PSO values, the FD uncertainty is smaller. However, we have shown that a substantial portion of that uncertainty comes from errors in the input quantities and not from the radiative transfer model used in the FD calculations. By reducing the CFsw differences to an optimum (small) constraint using MSCM, the mean differences between FD and PSO are improved, by up to a factor of 2 for Dir in the most dramatic case. But we cannot draw more concrete conclusions because of inhomogeneities in the diurnal sampling of PSO. When we apply MSCM to both CFsw and CldTau, more improvement is obtained though this is limited to the overcast scenes by the PSO retrieval. In the clear-sky case, FDrv substantially improved the overall flux comparisons, particularly for CDif and CDir. Nevertheless, it is clear that the actual uncertainty in the FD values of SWdn can be reduced with better aerosol information and a denser sampling of cloud variations. The PSO values represent accurate measurements of the flux, but their statistics, even monthly averages, have to be interpreted carefully because of the inhomogeneous diurnal sampling, as well as the usual problem of representing an area statistic with a point measurement.

[46] Although both SWdn and CSWdn show excellent agreement of the diurnal variations between FD and PSO, the Dif and Dir components disagree somewhat. Here there are two different stories. For clear sky, the dominant influence is the AOD difference; when we reduce the AOD values by half in FDrv, the total flux agreement becomes even better and the ratio Dir/Dif is improved substantially. However, as we showed, the area-point difference causes some mixing of clear and partly cloudy scenes, so the needed aerosol correction is not really this large. For cloudy sky, applying MSCM with a joint constraint on both CFsw and CldTau differences leads to a nearly perfect agreement for Dif and Dir (and therefore their sum SWdn), but this result is confined to LST = 9–15 hr because of the limitations of the PSO retrieval. So we conclude that the diurnal sampling differences also help explain part of the differences in the total fluxes, but we need further tests of the FD product at more extreme sun angles.

[47] For LW downwelling fluxes, when MSCM is applied to 3-hourly Ta, the mean difference is reduced to nearly zero from ~10 W/m^2 and the SD of differences is also improved to ≤25 W/m^2 from ~35 W/m^2 for LWdn and CLWdn. These values are about or slightly better than the combined uncertainties from FD and PSO (<20–25 W/m^2, based on uncertainty values from both for all ARM, SURFRAD, and BSRN sites). When CFlw is matched, the mean difference of LWdn is also reduced to nearly zero from ~10 W/m^2. The CLWdn is also slightly improved by FDrv. Thus, eliminating disagreements between temperatures and cloud cover nearly eliminates the disagreement of the FD and PSO LW fluxes: the former can therefore be improved with improved temperature data as input and denser cloud sampling and the latter is subject to the same limitation in representing area fluxes.

[48] FD and PSO generally exhibit good agreement for both the phase and amplitude of the LW diurnal variations. For downwelling LW diurnal variation comparison, FD shows good agreement with PSO for RFA-defined true clear sky but overestimates the amplitude for cloudy sky by 3–7 W/m^2, which is probably caused by different sensitivities to cirrus clouds. FD underestimates the diurnal amplitude of the LW upwelling flux for all and clear sky but generally agrees for overcast (CFlw >0.7). With the combined effect of downwelling and upwelling LW fluxes, FD underestimates the diurnal variation of the net LW loss for all the scenes by up to 10 W/m^2, although the daily mean net loss is more accurate. Such differences are smaller than the combined uncertainties from both of FD and PSO. The main conclusion here is that the PSO product's diurnal cycle accuracy is limited by its insensitivity to cirrus clouds, whereas the FD accuracy limit is associated with the diurnal variations of low clouds that are not completely sampled from the satellite perspective.

[49] We have also compared the diurnal variations between FD and PSO for individual stations and months, but we only summarize the main results here by showing the aggregate of all the statistics for all 15 stations and all 12 months because these summaries are consistent with the more detailed comparisons. Therefore, both FD and PSO accurately portray the diurnal flux variations.

[50] By using MSCM to exploit the functional relationship between radiative fluxes and their associated meteorological parameters, we have extended our previous evaluations for FD and obtained more accurate and meaningful estimates of the uncertainty of both data products. The real uncertainty values are likely smaller than separately estimated for both the general and diurnal comparisons, and the remaining discrepancies can be ascribed to the different natures of the two products (area versus point measurement, direct calculation versus retrieval) and the errors in specifying the relevant environmental parameters, $\mathbf{ev_1}$ for FD and $\mathbf{ev_2}$ for PSO. Therefore, the newly emerging RFA-processed product is more useful for a more accurate and insightful evaluation of radiative transfer models and satellite-based radiation products. As noted in the August 2006 issue of *GWEX News* [*Dutton*, 2006], the development of the RFA products significantly increases the value of surface radiation and meteorological measurements beyond that of the individual measurements themselves and makes possible far more detailed satellite and model comparison studies as demonstrated in this work. However, the RFA method still needs to be improved to produce more accurate meteorological parameters over a larger range of conditions and with more uniform coverage of the diurnal cycle (especially in SW). The method also needs to unify its cloud and aerosol parameters between SW and LW. An improved RFA product would make possible better evaluations between the satellite-derived and surface-observed fluxes as well as advances in understanding of the interactions between radiative fluxes and the associated meteorological parameters.

[51] The preparation for a next-generation, higher-resolution ISCCP cloud product with improved temperature/humidity profile, a refined cloud vertical structure model, a better diurnal adjustment scheme, and possibly a better aerosol climatology are under way. A set of more detailed and precise meteorological parameters, coincident and collocated with observed surface fluxes, such as RFA products, is needed for evaluation.

Appendix A: Uncertainties in RFA-PSO Dataset

[52] For the ARM measured data used in this study, estimates of the 2σ uncertainties of the measurements are 3% or 4 Wm^{-2}, 6% or 20 Wm^{-2}, 6% or 10 Wm^{-2}, and 2.5% or 4 Wm^{-2} (whichever is larger) for the downwelling direct normal SW, diffuse SW, total SW, and LW, respectively [*Stoffel*, 2005] for the 1-min data deduced from metrological considerations. However, metrological considerations tend to be pessimistic, given they are based on all things that could possibly go wrong with the measurement system. *Shi and Long* [2002] show typical 1-min long-term agreement between collocated instruments to be 9 Wm^{-2}, 14 Wm^{-2}, and 5 Wm^{-2} for the diffuse, direct normal, and LW, respectively. The 2σ uncertainties for the 3-min SURFRAD radiation budget measurements are given by *Augustine et al.* [2000] for the total and diffuse SW as 2%–5%, for the direct SW as 2%–3%, and for the LW as 9 Wm^{-2}. However, it is well known that longer-term averaging, such as monthly means, substantially reduces the random error contribution to these uncertainties and exhibits little bias, such that the monthly average uncertainties can be considered to be well below the 10 Wm^{-2} level. For the RFA clear-sky total, diffuse, and direct SW estimates, *Long and Ackerman* [2000] estimated the additional all-sky uncertainty due to interpolation for cloudy periods to be about equivalent to that of the measurements themselves, giving a total standard error only ~40% greater than the measurement uncertainty itself, and monthly mean uncertainties again below the 10 Wm^{-2} level. As Long and Ackerman also showed, the interpolation uncertainty is largely due to random errors, with little if any bias compared to the measurements themselves. *Long and Turner* [2008] estimated the clear-sky LW uncertainty to be about 4–5 Wm^{-2} for the 1-min data, again with little if any bias from the measurements themselves. Thus, for both the SW and LW, virtually any bias in the clear-sky estimates comes from the measurements themselves rather than from the fitting and interpolation process that produces the clear-sky estimations.

[53] For the SW-derived sky cover, *Long et al.* [2006] gave 1-min uncertainty as better than 10% sky cover, with precision listed as better than 3% sky cover. *Dürr and Philipona* [2004] listed the RMS 1-min uncertainty for their LW-derived effective sky cover as 1 okta, or 12.5% sky cover. *Barnard and Long* [2004] estimated the 1-min cloud visible optical depth uncertainty as 10% or better. These estimates apply to values reported, but as we discuss in this paper, there are important limitations on these quantities coming from both time sampling and methodology issues. The RFA estimations of clear-sky upwelling SW and LW are still under development and have not yet been vetted through the peer review process. While the uncertainties due to interpolation for cloudy periods in these estimates again appear to be on the order of the measurement uncertainties, to date the actual uncertainties have not been rigorously determined. The least certain of the RFA quantities used in this study is the cloud radiating surface height, in this study termed "CldHgt," which is inferred using the difference between the surface air temperature and the RFA-determined cloud radiating temperature (assuming a single-layer, plane parallel cloud and independent pixel approximation arguments) and an assumed atmospheric temperature lapse rate of 10 °C per km, generally used for all-around atmosphere by surface observers. Thus the RFA "CldHgt" is very crude, though again in the aggregate for monthly averages is likely to be reasonable in the mean, but the 1-min uncertainty is likely large. Again, there are possible larger biases in the relative in sensitivity of RFA to upper clouds.

[54] The air temperature and humidity measurements (used in the calculation of clear-sky downwelling LW) are from the ARM Surface Meteorological Observation Systems, which use a Campbell Scientific Model HMP35C temperature and humidity probe (manufactured by Vaisala) in an aspirated enclosure. The estimated uncertainty is 0.6 °C for air temperature, and 2%–3% for relative humidity (RH) [*Ritsche*, 2006].

[55] BSRN estimates their accuracies at 5 Wm^2 for downwelling SW and 10 Wm^2 for downwelling LW, but

these error estimates are not complete; they represent the standard deviations of the calibration coefficients [*Ohmura et al.*, 1998]. In other words, these values really represent the precision with which these instruments can be calibrated, but not the accuracy of the field measurements under all-sky conditions. Additionally, the value for direct normal irradiance in the table is from cavity radiometers. The "BSRN method" for obtaining the direct beam irradiance is a combination of cavity radiometer and pyrheliometer. *Ohmura et al.* [1998] stated, "In effect, the thermopile pyrheliometer is being calibrated during every successful measurement from the cavity radiometer" (p. 2120). But very few BSRN stations use this methodology.

[56] **Acknowledgments.** We thank the principal investigators, Rick Wagener, Brent N. Holben, and Ross Mitchell for their efforts in establishing and maintaining Nauru, Manus, and Darwin AERONET sites, respectively, and Connor Flynn for his data processing for these sites. The work by two authors (Y. Zhang and W. B. Rossow) is supported by NASA grant NNXD7AN04G, the MAP program directed by Dr. Donald Anderson, and grant NNXD7AO90G, the NEWS project directed by Dr. Jared Entin. C. N. Long acknowledges the support of the Climate Change Research Division of the U.S. Department of Energy as part of the Atmospheric Radiation Measurement (ARM) Program. Recognition is also extended to those responsible for the operation and maintenance of the instruments that produced the measurements used in this study; their diligent and dedicated efforts are often underappreciated.

[57] Except AOD data, the RFA-PSO datasets were obtained from the NASA Langley Research Center Atmospheric Science Data Center. We also express our appreciation for the comments and suggestions by several anonymous reviewers.

References

Ackerman, T., and G. Stokes (2003), The Atmospheric Radiation Measurement Program, *Phys. Today*, *56*, 38–45.

Augustine, J. A., G. B. Hodges, E. G. Dutton, J. J. Michalsky, and C. R. Cornwall (2008), An aerosol optical depth climatology for NOAA's national surface radiation budget network (SURFRAD), *J. Geophys. Res.*, *113*, D11204, doi:10.1029/2007JD009504.

Augustine, J. A., J. J. Deluisi, and C. N. Long (2000), SURFRAD: A national surface radiation budget network for atmospheric research, *Bull. Am. Meteorol. Soc.*, *81*(10), 2341–2357.

Barnard, J. C., and C. N. Long (2004), A simple empirical equation to calculate cloud optical thickness using shortwave broadband measurements, *J. Applied Meteor.*, *43*, 1057–1066.

Barnard, J. C., C. N. Long, E. I. Kassianov, S. A. McFarlane, J. M. Comstock, M. Freer, and G. M. McFarquhar (2008), Development and evaluation of a simple algorithm to find cloud optical depth with emphasis on thin ice clouds, *Open Atmos. Sci. J.*, *2*, 46–55, doi:10.2174/1874282300802010046.

Bosilovich, M. G., and S. D. Schubert (2002), Water vapor tracers as diagnostics of the regional hydrologic cycle, *J. Hydromet.*, *3*, 149–165.

Cairns, B. (1995), Diurnal variations of cloud from ISCCP data, *Atmos. Res.*, *37*, 133–246, doi:10.1016/0169-8095(94)00074-N.

Dupont, J.-C., M. Haeffelin, and C. N. Long (2008), Evaluation of cloudless-sky periods detected by shortwave and longwave algorithms using lidar measurements, *Geophys. Res. Lett.*, *35*, L10815, doi:10.1029/2008GL033658.

Dürr, B., and R. Philipona (2004), Automatic cloud amount detection by surface longwave downward radiation measurements, *J. Geophys. Res.*, *109*, D05201, doi:10.1029/2003JD004182.

Dutton, E. (2006), 9th BSRN scientific review and workshop, 29 May–2 June 2006, Lindenberg, Germany, *GEWEX News*, *16*(3), 18.

Goody, R. M., and Y. L. Yung (1989), *Atmospheric Radiation: Theoretical Basis*, 2nd ed., 519 pp., Oxford University Press, New York, 1989.

Grassl, H. (1971), Calculated circumsolar radiation as a function of aerosol type, field of view, wavelength, and optical depth. *Appl. Opt.*, *10*(11), 2542–2543.

Halthore, R. N., and S. E. Schwartz (2000), Comparison of model-estimated and measured diffuse downward irradiance at surface in cloud-free skies, *J. Geophys. Res.*, *105*(D15), 20,165–21,077, doi:10.1029/2000JD900224.

Holben, B. N., et al. (1998), AERONET: A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, *66*, 1–16.

Long, C. N. (2004), The next generation flux analysis: Adding clear-sky LW and LW cloud effects, cloud optical depths, and improved sky cover estimates, in *Proceedings of the Fourteenth ARM Science Team Meeting*, Albuquerque, NM, March 22–26, 2004.

Long, C. N. (2005), On the estimation of clear-sky upwelling shortwave and longwave, in *Proceedings of the Fifteenth ARM Science Team Meeting*, Daytona, FL, March 14–18, 2005.

Long, C. N., and T. P. Ackerman (2000), Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects, *J. Geophys. Res.*, *105*(D12), 15,609–15,626, doi:10.1029/2000JD900077.

Long, C. N., T. P. Ackerman, K. L. Gaustad, and J. N. S. Cole (2006), Estimation of fractional sky cover from broadband shortwave radiometer measurements, *J. Geophys. Res.*, *111*, D11204, doi:10.1029/2005JD006475.

Long, C. N., and K. L. Gaustad (2004), The shortwave (SW) clear-sky detection and fitting algorithm: algorithm operational details and explanations, *Tech. Rep. ARM TR-004*, Atmos. Radiat. Measure. Program, U.S. Dept. of Energy, Washington, DC.

Long, C. N., and D. D. Turner (2008): A method for continuous estimation of clear-sky downwelling longwave radiative flux developed using ARM surface measurements, *J. Geophys. Res.*, *113*, D18206, doi:10.1029/2008JD009936.

Miller, R. L., and I. Tegen (1999), Radiative forcing of a tropical direct circulation by soil dust aerosols, *J. Atmos. Sci.*, *56*, 2403–2433, doi:10.1175/1520-0469.

Nowak, D., L. Vuilleumier, C. N. Long, and A. Ohmura (2008), Solar irradiance computations compared with observations at the Baseline Surface Radiation Network Payerne site, *J. Geophys. Res.*, *113*, D14206, doi:10.1029/2007JD009441.

Ohmura, A., et al. (1998), Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research, *Bull. Am. Meteorol. Soc.*, *79*, 2115–2136.

Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad, and D. Hartmann (1989), Cloud-radiative forcing and climate: Results from the Earth Radiation Budget Experiment, *Science*, *243*, 57–63.

Ritsche, M. T. (2006), Surface Meteorological Observation System (SMOS) Handbook, *ARM Tech. Rep. ARM TR-031*, 28 pp.

Rossow, W. B., Y.-C. Zhang, and J. Wang (2005), A statistical model of cloud vertical structure based on reconciling cloud layer amounts inferred from satellites and radiosonde humidity profiles. *J. Climate*, *18*, 3587–3605, doi:10.1175/JCLI3479.1.

Rossow, W. B., and R. A. Schiffer (1999), Advances in understanding clouds from ISCCP, *Bull. Am. Meteorol. Soc.*, *80*, 2261–2287.

Rossow, W. B., A. W. Walker, D. E. Beuschel, and M. D. Roiter (1996), International Satellite Cloud Climatology Project (ISCCP) Documentation of New Cloud Datasets. *WMO/TD 737, World Climate Research Programme*, 115 pp.

Rossow, W. B., and Y.-C. Zhang (1995), Calculation of surface and top of atmosphere radiative fluxes from physical quantities based on ISCCP data sets 2. Validation and first results, *J. Geophys. Res.*, *100*(D1), 1167–1197, doi:10.1029/94JD02746.

Rossow, W. B., A. W. Walker, and L. C. Garder (1993), Comparison of ISCCP and other cloud amounts, *J. Climate*, *6*, 2394–2418, doi:10.1175/1520-0442(1993).

Shi, Y., and C. N. Long (2002), Best Estimate Radiation Flux Value Added Product: Algorithm Operational Details and Explanations, *Atmospheric Radiation Measurement Program Technical Report, ARM TR-008*, 58 pp.

Stoffel, T. (2005), Solar Infrared Radiation Station (SIRS) Handbook, *Atmospheric Radiation Measurement Program Technical Report, ARM TR-025*, 29 pp.

Venables, W. N., and B. D. Ripley (1997), *Modern Applied Statistics with S-Plus*, 548 pp., Springer, New York, 1997.

Zhang, Y.-C., W. B. Rossow, and A. A. Lacis (1995), Calculation of surface and top of atmosphere radiative fluxes from physical quantities based on ISCCP data sets, 1. Method and sensitivity to input data uncertainties, *J. Geophys. Res.*, *100*(1), 1149–1165, doi:10.1029/94JD02747.

Zhang, Y., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko (2004), Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data, *J. Geophys. Res.*, *109*, D19105, doi:10.1029/2003JD004457.

Zhang, Y., W. B. Rossow, and P. W. Stackhouse Jr. (2006), Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the near-surface atmosphere, *J. Geophys. Res.*, *111*, D13106, doi:10.1029/2005JD006873.

Zhang, Y., W. B. Rossow, and P. W. Stackhouse Jr. (2007), Comparison of different global information sources used in surface radiative flux cal-

culation: Radiative properties of the surface, *J. Geophys. Res.*, *112*, D01102, doi:10.1029/2005JD007008.

E. G. Dutton, National Oceanic and Atmospheric Administration ESRL R/GMD, 325 Broadway, Boulder, CO 80305, USA.

C. N. Long, Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352, USA.

W. B. Rossow, Remote Sensing (CREST), The City College of New York, Steinman Hall (T-107), 140th St. and Convent Ave., New York, NY 10031, USA.

Y. Zhang, Department of Applied Physics and Applied Mathematics, Columbia University, NASA Goddard Institute for Space Studies, 2880 Broadway, Rm. 320-B, New York, NY 10025, USA. (yzhang@giss.nasa.gov)