JOURNAL OF GEOPHYSICAL RESEARCH: ATMOSPHERES, VOL. 118, 9552–9564, doi:10.1002/jgrd.50697, 2013

Correcting surface solar radiation of two data assimilation systems against FLUXNET observations in North America

Lei Zhao,[1] Xuhui Lee,[1] and Shoudong Liu[2]

Received 15 November 2012; revised 19 June 2013; accepted 31 July 2013; published 5 September 2013.

[1] Solar radiation at the Earth's surface is an important driver of meteorological and ecological processes. The objective of this study is to evaluate the accuracy of the reanalysis solar radiation produced by NARR (North American Regional Reanalysis) and MERRA (Modern-Era Retrospective Analysis for Research and Applications) against the FLUXNET measurements in North America. We found that both assimilation systems systematically overestimated the surface solar radiation flux on the monthly and annual scale, with an average bias error of +37.2 Wm^{-2} for NARR and of +20.2 Wm^{-2} for MERRA. The bias errors were larger under cloudy skies than under clear skies. A postreanalysis algorithm consisting of empirical relationships between model bias, a clearness index, and site elevation was proposed to correct the model errors. Results show that the algorithm can remove the systematic bias errors for both FLUXNET calibration sites (sites used to establish the algorithm) and independent validation sites. After correction, the average annual mean bias errors were reduced to +1.3 Wm^{-2} for NARR and +2.7 Wm^{-2} for MERRA. Applying the correction algorithm to the global domain of MERRA brought the global mean surface incoming shortwave radiation down by 17.3 W m^{-2} to 175.5 W m^{-2}. Under the constraint of the energy balance, other radiation and energy balance terms at the Earth's surface, estimated from independent global data products, also support the need for a downward adjustment of the MERRA surface solar radiation.

Citation: Zhao, L., X. Lee, and S. Liu (2013), Correcting surface solar radiation of two data assimilation systems against FLUXNET observations in North America, *J. Geophys. Res. Atmos.*, *118*, 9552–9564, doi:10.1002/jgrd.50697.

1. Introduction

[2] Solar energy at the earth's surface (S) is an important driver of various interactions between the land and the atmosphere. To improve our understanding of how meteorological processes distribute energy in the climate system, an accurate assessment of this variable is required. Because this variable is not among the measurements made at standard surface weather stations, models of various complexities are used to estimate its spatiotemporal distributions. Atmospheric reanalysis is one such class of models. In reanalysis model systems, solar radiation at the surface is calculated with radiative transfer models (RTMs). The reanalysis modeled products are superior to discrete surface observations because of large and continuous spatial and temporal coverages. These features are especially attractive for people interested in climate and ecological patterns on the regional and global scales.

[3] Implementation of RTMs can be divided into two categories: stand-alone mode and application embedded in global climate models (GCMs) or reanalysis systems. The stand-alone models provide accurate single column calculations, while the applications in GCMs and reanalyses are suited for regional and global scales with minimal loss of accuracy. Previous efforts have taken advantages of both types of the RTMs to produce the radiation fluxes at the top of atmosphere (TOA), within the atmosphere and at the surface [*Rossow and Zhang*, 1995; *Zhang et al*., 1995; *Kiehl and Trenberth*, 1997; *Zhang et al*., 2004; *Hatzianastassiou et al*., 2005]. Of interest here is the discrepancy in the estimate of the global mean S. According to the ISCCP-FD (fluxes using the International Satellite Cloud Climatology Project-D input data), the global mean S value is 188.5 Wm^{-2} for the period of 2000 to 2004 [*Trenberth et al*., 2009]. Using the same cloud climatology in the ISCCP-D series but a different RTM and ancillary data sets [*Zhang et al*., 2004], *Hatzianastassiou et al*. [2005] estimated a mean value of 171.6 Wm^{-2}. These two estimates differ by 17 Wm^{-2} or 9.8%, indicating uncertainties in the various RTMs or input data sets used.

[4] Validation of reanalysis model products against surface observations is an active area of research. This is because reanalysis represents only the best "guess" of various atmospheric and hydrological variables through a combination of model predictions and a variety of observations, the latter of which serve to constrain the model calculations. One source of error is the spatial and temporal mismatch between

Additional supporting information may be found in the online version of this article.

[1]School of Forestry and Environmental Studies, Yale University, New Haven, Connecticut, USA.

[2]Yale-NUIST Center on Atmospheric Environment, Nanjing University of Information Science and Technology, Nanjing, China.

Corresponding author: L. Zhao, School of Forestry and Environmental Studies, Yale University, Rm. 300, 21 Sachem St., New Haven, CT 06511, USA. (lei.zhao@yale.edu)

2169-897X/13/10.1002/jgrd.50697

the model analytical framework and the observational data sets. In the case of radiative fluxes, neither cloud observations nor surface radiation measurements are directly assimilated in the model system. (MERRA assimilates clouds indirectly by adjusting moisture over the oceans.) Furthermore, as noted above, the RTMs imbedded in the modeling system have their own inherent uncertainties.

[5] Several investigators have compared the surface radiation variables produced by NARR and MERRA against field observations. Using the data obtained at a FLUXNET site in Oklahoma, USA, *Kennedy et al*. [2011] concluded that both MERRA and NARR have positive biases (NARR: 47 Wm^{-2}; MERRA: 19 Wm^{-2}) for *S* and negative biases for the surface downward longwave radiation. They attributed the significant positive bias in NARR to a combination of too low cloud amounts in the model domain and too weak light extinction by aerosols and water vapor. *Markovic et al*. [2009] found similar positive biases (40 Wm^{-2}) in the annual mean *S* for NARR at six sites in six states from west to east US and attributed them to a negative bias in cloud fraction. *Walsh et al*. [2009] evaluated surface radiative fluxes and cloud fraction from NARR and three other reanalyses against ground measurements at Barrow, Alaska in the Arctic, and found positive biases of +4 to +43 W m^{-2} in monthly *S* and the associated negative biases of cloud fraction. Over the same region, *Zib et al*. [2012] found an annual mean high bias of 3.9 Wm^{-2} for MERRA at two BSRN (Baseline Surface Radiation Network) sites. They also related the bias errors to biases in the modeled cloud fraction although the radiation biases still exist for some time periods despite that the observed cloud fraction is correctly reproduced by the model. In a recent evaluation study involving six reanalysis products at nine field sites on the Tibetan Plateau, *Wang and Zeng* [2012] reported an overestimation of up to 40 W m^{-2} for *S* in MERRA. According to their study, the bias errors appear to be related to latitude.

[6] The few investigations conducted over large spatial scales have confirmed the tendency for reanalysis models to overestimate the surface solar radiation. Utilizing hourly observations at 33 FLUXNET sites in the US and Canada, *Decker et al*. [2012] evaluated a number of surface micrometeorological and flux variables produced by the NCEP/NCAR, CFSR, ERA-40, ERA-Interim, GLDAS, and MERRA reanalysis systems, concluding that all of them overestimate *S* by varying amounts of 10 to 50 Wm^{-2} with ERA-Interim having the best accuracy. In another large-scale evaluation study, *Wild et al*. [1998] compared surface solar energy of the reanalysis product ECMWF (European Center for Medium Range Weather Forecasts) against the monthly mean radiation observations at 720 GEBA (Global Energy Balance Archive) sites distributed worldwide. In terms of annual means, an underestimation at low latitudes in the Northern Hemisphere and an overestimation in the rest of the world were found in the study. But overall, the reanalysis bias errors are substantially smaller than in GCM estimates [*Wild et al*., 1995].

[7] The published studies show that the annual mean bias errors of reanalyzed *S* are in the range of +5 to +60 W m^{-2}. These positive bias errors have undesirable consequences for atmospheric and ecological applications. First, these bias errors are 5–30% of the observed net all-wave surface radiation balance at midlatitude forests [*Rotenberg and Yakir*, 2010; *Lee et al*., 2011]. Use of the reanalyzed surface solar radiation to drive land surface model calculations will result in extra energy going to the atmosphere via sensible and latent heat fluxes assuming other radiation terms remain unchanged. An excess of net radiation at the surface, for example, could cause excessive surface evaporation especially under moist climate [*Betts et al*., 1996]. Under the assumption that surface evapotransporation is balanced by precipitation, this would lead to excessive precipitation in climate models. In reanalysis systems, excessive surface evapotransporation calculated by the model may not be balanced by precipitation assimilated from observations, thus forcing the model to drift to a dry soil state [*Viterbo and Courtier*, 1995]. Second, the reanalysis *S* is used to predict surface ozone formation [*Mickley et al*., 2004; *Leung and Gustafson*, 2005; *Nolte et al*., 2008; *Weaver et al*., 2009; *Hickman et al*., 2010]. The study by *Nolte et al*. [2008] suggests that a high bias error of 20 W m^{-2} in *S* can cause an increase of surface ozone concentration by about 3–5 ppb in eastern Texas and along much of the east coast of the US. Third, because solar radiation is the primary driver of plant photosynthesis, bias errors in *S* are problematic for ecological models. Finally, the reanalysis radiation data are used to assess the global energy budget of the climate system [*Wild*, 2009; *Bosilovich et al*., 2011; *Stephens et al*., 2012a; *Stephens et al*., 2012b]. The reported bias errors in *S* are comparable in magnitude to the global mean surface sensible heat flux [*Trenberth et al*., 2009], emphasizing the need for a quantitative correction algorithm.

[8] The first objective of this study is to evaluate the two reanalysis *S* products, NARR and MERRA, against the FLUXNET observations in North America. Our strategy is to perform the evaluation at a large number of FLUXNET sites across different climate regimes, over a wide range of elevations, and at multiple time scales of months to multiple years. Simultaneous evaluation of two reanalysis products may help us to uncover model errors related to the data sources and the approach taken by the models. FLUXNET is a global network of eddy flux towers which maintain continuous measurements of surface meteorological variables and land-atmosphere fluxes [*Baldocchi et al*., 2001]. Because our analysis is conducted with monthly averages, it is essential that the sites have long records to produce large enough numbers of gap-free months for rigorous comparison. Many FLUXNET sites satisfy this criterion (Table 1). In addition, long time series allow us to examine the capability of the reanalyses to capture the interannual variations in *S*.

[9] Our second objective is to develop an algorithm for reducing the systematic errors. A computationally efficient postreanalysis correction algorithm is desired by the end users engaged in the applications discussed above. The routine was applied to the MERRA data to obtain a sense of its bias in the global mean *S* estimate. Although the algorithm is restricted to MERRA and NARR, the insights gained may be useful for developing correction algorithms for other reanalysis products.

[10] As in *Decker et al*. [2012] and *Wild et al*. [1998], we validated the reanalyzed *S* products against a large number of observational sites. We expanded the work of *Wild et al*. [1998] by focusing not only on annual mean *S* but also on its seasonal and interannual variations. Instead of using hourly data as in *Decker et al*. [2012], we conducted the

Table 1. Annual Mean Surface Incoming Solar Radiation Fluxes and Ancillary Information on the Calibration and Validation Sites

Site Code	Site Name	State/ Prov	Lat	Long	Elev (m)	Years Included	Annual Means Observed (Wm^{-2})	Annual Mean Bias NARR (Wm^{-2})	Annual Mean Bias MERRA (Wm^{-2})	Reference
					Calibration					
CA-Obs	Old Black Spruce	SK	53.99	−105.12	629	10	134.6	29.2	18.6	*Jarvis et al.* [1997]
CA-Ojp	Old Jack Pine	SK	53.92	−104.69	579	10	133.3	30.8	19.9	*Baldocchi et al.* [1997]
CA-Oas	Old Aspen	SK	53.63	−106.20	601	10	136.4	29.6	19.5	*Blanken et al.* [1997]
CA-Ca1	BC Douglas-fir 1949	BC	49.87	−125.33	320	12	127.2	34.6	21.0	*Humphreys et al.* [2003]
US-UMB	UMBS	MI	45.56	−84.71	234	8	150.8	39.3	22.5	*Schmid et al.* [2003]
CA-Cbo	Borden	ON	44.20	−79.93	217	14	149.2	39.8	26.0	*Lee et al.* [1999]
US-NR1	Niwot Ridge	CO	40.03	−105.55	3050	5	188.0	59.8	37.0	*Monson et al.* [2002]
US-MMS	Morgan Monroe State Forest	IN	39.32	−86.41	275	9	167.6	41.0	22.7	*Schmid et al.* [2000]
US-Ton	Tonzi Ranch	CA	38.43	−120.97	177	7	218.8	23.6	6.0	*Ma et al.* [2007]
US-Var	Vaira Ranch	CA	38.41	−120.95	129	8	213.1	26.9	9.0	*Ma et al.* [2007]
US-WBW	Walker Branch	TN	35.96	−84.29	343	10	174.3	35.6	22.8	*Wilson and Meyers* [2001]
US-Aud	Audubon Research Ranch	AZ	31.59	−110.51	1469	6	237.2	35.0	11.1	*Krishnan et al.* [2012]
US-SP2	Mize	FL	29.76	−82.24	43	7	184.2	41.3	18.9	*Gholz and Clark* [2002]
US-SP3	Donaldson	FL	29.75	−82.16	36	7	182.1	43.1	21.1	*Gholz and Clark* [2002]
					Validation					
CA-Qfo	Quebec Mature Boreal Forest	QC	49.69	−74.34	390	1	127.9	30.9	24.7	*Bergeron et al.* [2007]
CA-Ca3	BC Douglas-fir 1988	BC	49.53	−124.90	120	1	133.5	35.5	17.5	*Humphreys et al.* [2006]
US-Ho1	Howland Forest Main	ME	45.20	−68.74	60	2	150.7	37.7	28.5	*Hollinger et al.* [1999]
US-Bkg	Brookings	SD	44.35	−96.84	510	4	175.3	35.4	12.4	*Gilmanov et al.* [2005]
US-Bo1	Bondville	IL	40.01	−88.29	219	7	168.7	39.6	23.7	*Hollinger et al.* [2005]
US-Slt	Silas Little Experimental Forest	NJ	39.91	−74.60	30	4	159.5	48.5	27.6	*Clark et al.* [2010]
US-MOz	Missouri Ozark	MO	38.74	−92.20	219	4	180.0	39.4	16.4	*Gu et al.* [2006]
US-Dk2	Duke Forest Hardwoods	NC	35.97	−79.10	168	5	183.4	35.7	15.3	*Oren et al.* [2006]
US-NC2	North Carolina Loblolly Pine	NC	35.80	−76.67	12	4	170.3	40.3	27.2	*Noormets et al.* [2010]
US-Fmf	Flagstaff Managed Forest	AZ	35.14	−111.73	2160	3	230.2	36.1	13.4	*Dore et al.* [2008]

comparison using monthly means which have smaller random errors than the hourly data, allowing better isolation of systematic errors.

2. Materials and Methods

2.1. Sites and Data

2.1.1. Surface Observations

[11] The Ameriflux (level 2) and FLUXNET-CANADA (http://public.ornl.gov/ameriflux/index.html; http://fluxnet.ccrp.ec.gc.ca/e_about.htm) are two regional networks of FLUXNET, consisting of eddy-covariance sites in North America, Central America, and South America. In this study, we selected 24 sites, spanning a large geographic range of the US and Canada (Figure 1). Of these, 14 sites have long measurement records and were used as calibration sites (sites used to develop the correction algorithm). The other 10 sites were used as validation sites (sites used to independently validate the corrected *S*). These sites all measured the four components of the surface radiation balance and had high (>90%) data coverage. All the sites in this study used a pyranometer to measure *S*, but the sensor type varies. The

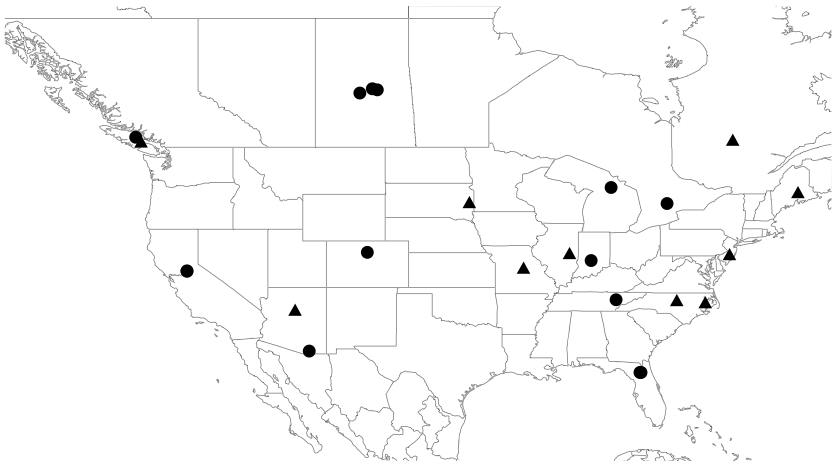

Figure 1. Map of the selected Ameriflux and FLUXNET-CANADA sites. Circles denote the calibration sites, and triangles denote the validation sites.

uncertainty of measured S in Ameriflux is ±3.55 W m^{-2} according to a cross-site comparison with a roving Ameriflux standard pyranometer [*Schmidt et al*., 2012].

[12] BSRN [*Ohmura et al*., 1998], established in 1992, is a global network of continuous measurements of radiative fluxes at the Earth's surface. It has 56 stations, covering a latitudinal range from 80°N to 90°S. BSRN monitors the global solar irradiance primarily by a combination of diffuse sky irradiance measured by a shaded pyranometer and direct solar irradiance measured by an absolute cavity radiometer, with a target accuracy of 5 W m^{-2} [*Ohmura et al*., 1998]. In this study, we selected 42 BSRN sites for further validation of the correction algorithm outside of North America (supporting information Figure S1 and Table S1); at these sites, measurements are complete for at least one year. In addition to the BSRN sites, we also include on the validation list the data from the published literature in which the observed annual mean S is reported.

2.1.2. NARR

[13] The North American Regional Reanalysis (NARR), carried out by the National Center for Environmental Prediction (NCEP), is a long-term, near real-time, high-resolution, high-frequency, atmospheric and land surface reanalysis product [*Mesinger et al*., 2006]. This regional reanalysis provides a much improved data set of land hydrology and land-atmosphere interactions compared to the earlier global reanalysis data set NCEP-NCAR. NARR covers the period from 1979 up to the present, with the data archived at 3 hourly, daily, and monthly time scales. The grid resolution is approximately 32 km. Documented in the NARR outputs are meteorological, hydrological, and ecological variables. Here we used the surface incoming shortwave radiation flux data.

[14] The radiative fluxes at the surface are computed through the radiation scheme embedded in the NCEP Eta model, which has a shortwave [*Lacis and Hansen*, 1974] and a longwave package [*Fels and Schwarzkopf*, 1975]. The shortwave absorption scheme considers the amount and type of cloud, the humidity, the solar elevation angle, and the vertical distribution of ozone within the stratosphere. The cloud information is obtained from the cloud microphysics in the Eta model [*Ferrier et al*., 2002].

2.1.3. MERRA

[15] MERRA [*Rienecker et al*., 2011], maintained by NASA Global Modeling and Assimilation Office, is the second generation reanalysis data set, which uses the Goddard Earth Observing System Data Assimilation System-Version 5 (GEOS-5). GEOS-5 includes GEOS-5 atmospheric circulation model and the grid point statistical interpolation. MERRA implements a procedure called incremental analysis updates [*Bloom et al*., 1996] to slowly converge modeled calculations toward the observations. A key feature of this global reanalysis is that it takes advantage of a variety of recent satellite observations to improve the estimates of earth's energy and water cycles. Same with NARR, MERRA spans the satellite era, from 1979 to the present. Most of the MERRA outputs are archived hourly at its native spatial grid resolution of 2/3° × 1/2°.

[16] The radiative transfer model developed at the Goddard Climate and Radiation Branch at NASA is utilized in MERRA to generate the radiative fluxes at the top of the atmosphere (TOA) and at the surface. The shortwave radiation scheme, documented in *Chou and Suarez* [1999], resolves the absorption by water vapor, ozone, oxygen, carbon dioxide, and aerosols. The longwave radiation scheme, documented in *Chou et al*. [2001], calculates the absorption by water vapor, trace gases, clouds, and aerosols. A prognostic cloud scheme embedded in MERRA assumes that clouds are maximum-randomly overlapped.

2.2. Method

2.2.1. Spatial Interpolation

[17] In order to account for the spatial mismatch between modeled grids and surface observations, the reanalysis data were horizontally interpolated, using a bilinear interpolation technique, to the measurement site from the center of the four surrounding grid cells with a weighting factor that is inversely proportional to the distance. The same technique was used by *Wild et al*. [1998] in their validation of the ECMWF reanalysis. The interpolation was done at the 3 hourly intervals for NARR and at hourly intervals for MERRA, and daily mean values were computed with the interpolated data. The interpolation should eliminate errors arising from latitudinal mismatch especially for measurement sites located near the edge of a grid cell. The slight mismatch between the interpolated model grid elevation and that of the measurement site (500 m at most) has negligible consequences and is ignored.

2.2.2. Monthly Average

[18] The space-time sampling mismatch between model products and surface measurements can cause large uncertainties in validation studies [*Rossow and Zhang*, 1995; *Zhang et al*., 2004]. This is because surface measurements are for single points in space, whereas modeled fluxes are for the area of a grid cell. At hourly time steps, large random errors are unavoidable due to variations of cloud within the grid cell. Using longtime temporal averages to compare the cell mean and the point measurement should reduce the errors caused by the problem [*Zhang et al*., 2004]. For this reason, we conducted the comparison using monthly averages.

[19] The original observational data are half-hourly. We excluded outliers (daytime negative values) in the original data. In order to avoid introducing new uncertainties, we did not perform any gap filling for missing data or daytime outliers. Instead, we excluded the day if one or more half-hourly observation was missing during the daytime and excluded the whole month if five or more days were missing in that month. (Missing nighttime values were substituted by zero.) To ensure proper comparison, we also excluded the same days to calculate the reanalyzed monthly means.

2.2.3. Dependence of Bias Error on Clearness Index

[20] We utilized the concept of clearness index (k_t) to develop an algorithm for correcting the reanalysis bias errors. This index is defined as the ratio of global solar radiation received at the surface to the extraterrestrial radiation at the TOA (S_e) [*Gu et al*., 1999],

$$k_t = \frac{S}{S_e}, \tag{1}$$

where S_e is given by

$$S_e = S_{sc}[1 + 0.033\cos(360t_d/365)]\sin\beta, \tag{2}$$

[21] In equation (2), S_{SC} is the solar constant, t_d is the day of the year, and β denotes the solar elevation angle. The

hourly values of S_e given by equation (2) were converted into monthly means using the same method as for the reanalyzed S. The clearness index was used as an independent variable to relate with the model errors in this study.

[22] A bias ratio (b_m) was used to measure the relative bias error, as:

$$b_m = \frac{S_m - S}{S_m}, \tag{3}$$

where S_m denotes the modeled solar radiation at the surface, and S denotes the observed solar radiation. Using the bias ratio rather than the actual bias nondimensionizes the error and helps to eliminate the possible latitudinal dependence of the bias [*Wild et al*., 1998].

[23] The reanalysis systems have much better performance under clear-sky conditions than under cloudy sky conditions. *Wild et al*. [1998] have demonstrated the accurate performance of the radiation scheme in ECMWF under clear-sky conditions. Similarly, MERRA is also able to calculate the clear-sky surface global radiation flux reasonably well [*Kennedy et al*., 2011]. Under cloudy skies, both NARR and MERRA underestimate the cloud fraction [*Kennedy et al*., 2011; *Zib et al*., 2012], contributing to overestimation of S. Here we used the following simple linear equation to include the dependence of bias on sky conditions

$$b_m = a \cdot k_t + b \tag{4}$$

where a and b are empirical coefficients.

2.2.4. Dependence of Bias Error on Elevation

[24] Site elevation can also introduce model bias errors [*Frauenfeld et al*., 2005; *Zhao et al*., 2007; *Wang and Zeng*, 2012], for two reasons. First, if elevation at the model grid does not match that at the observational point, the sunlight optical path is not correctly resolved by the model, leading to underestimation or overestimation of the sunlight extinction. However, among all the observation sites chosen for this study, the elevation mismatch is at most 0.5 km, and according to the observed elevation gradient of S in the eastern US [*Richardson et al*., 2004], the resulting bias in S should be no greater than 2%. Second, our analysis suggests that either the lack of fully resolved orographic clouds or the bias in higher cloud liquid water path (LWP) results in increased bias errors at higher elevation sites. Empirically, this second elevation effect is captured here by modifying the coefficient b in equation (4) to be a function of the site elevation z_e (< 5 km)

$$b = c_0 {z_e}^2 + c_1 z_e + c_2, \tag{5}$$

where c_0, c_1, and c_2 are empirical coefficients.

[25] To determine the regression coefficients in equations (4) and (5), we first applied the geometric mean regression to equation (4), using monthly data from all the calibration sites, to obtain the slope coefficient a separately for NARR and MERRA. Next we applied equation (4) to the monthly data at individual sites with the fixed slope coefficient to determine the intercept coefficient b. Finally, the coefficients c_0, c_1, and c_2 in equation (5) were found by regression of the site b value against its elevation. To avoid abnormal behavior of the quadratic curve beyond the valid data range, we set a threshold of 5 km in site elevation, beyond which b remains constant.

2.2.5. Postreanalysis Correction

[26] The above set of equations cannot be used directly to correct the model bias errors because the clearness index is an unknown variable without actual measurement of the surface solar radiation. However, a solution for the corrected monthly mean S can be derived from equations (3) and (4), as

$$S = \frac{(1-b)S_m S_e}{aS_m + S_e}, \tag{6}$$

where a is fixed for each of the two reanalysis products and b depends on the surface elevation of the model grid according to equation (5). To avoid negative bias ratio, we set the thresholds for k_t as $k_t > 0.75$ for NARR and $k_t > 0.70$ for MERRA, or,

$$\frac{(1-b)S_m}{aS_m + S_e} > 0.75 \quad \text{for NARR}, \tag{7}$$

$$\frac{(1-b)S_m}{aS_m + S_e} > 0.70 \quad \text{for MERRA}. \tag{8}$$

No correction was conducted above these thresholds.

3. Results and Discussion

3.1. Relations of Bias Errors to Clearness Index and Elevation

[27] Sky condition is the primary factor that drives the variations of model bias error for both NARR and MERRA. Applying the regression analysis to 1253 monthly observations pooled together from the 14 calibration sites, we obtained the following empirical relationships:

$$b_m = -0.89 \cdot k_t + 0.65 \text{ for NARR}, \tag{9}$$

$$b_m = -0.82 \cdot k_t + 0.55 \text{for MERRA}. \tag{10}$$

[28] The slopes of these regression equations are large, reflecting large sensitivity to the clearness index. The sensitivity to the clearness index is slightly larger for NARR than for MERRA. The intercepts of the regression equations represent the model bias ratios under the theoretical limit of $k_t = 0$. In this limit, NARR has larger bias errors than MERRA. These intercepts also give the upper limits of the model bias ratios. According to the regression R^2 values, the clearness index explains 62% of the variations of the bias errors in NARR and 36% in MERRA. MERRA has a smaller R^2 value than NARR, in part due to the few outliers (Figure 2b) which will be discussed in section 3.3; excluding these outliers, the R^2 would improve to 0.52.

[29] The negative relationships shown in Figure 2 confirm that the model bias errors are larger under cloudy skies than under clear skies. This indicates that the NARR and MERRA model systems have poor capability of describing cloudiness. When $k_t \geq 0.7$, both reanalysis systems have bias ratios near zero. When $k_t \leq 0.2$, NARR and MERRA have bias ratios of around 0.5 and 0.4, respectively. Like other reanalysis systems, NARR and MERRA parameterize cloud with cloud microphysics packages instead of directly assimilating cloud observations. Our results suggest that these parameterizations have a tendency to underestimate cloud amount. Similarly, *Walsh et al*. [2009], *Kennedy et al*.

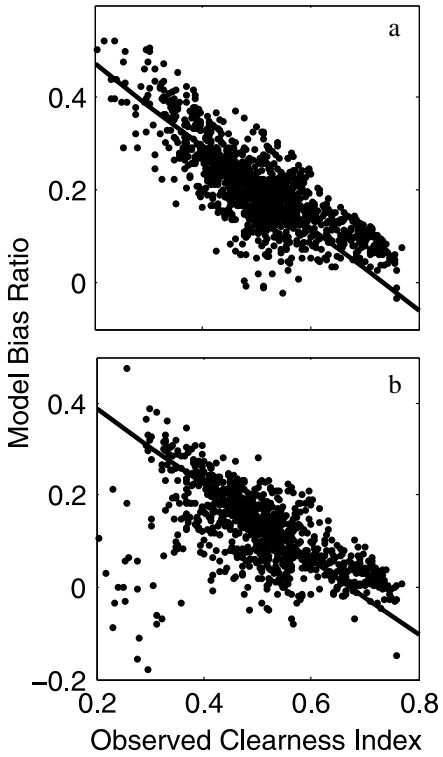

Figure 2. Relationship between monthly mean clearness index, S / S_e, and monthly mean model bias ratio, ($S_m - S$) / S_m, for the calibration sites. (a) NARR, regression equation $y = -0.89 \times x + 0.65$ ($R^2 = 0.62$, $n = 1253$); (b) MERRA, regression equation $y = -0.82 \times x + 0.55$ ($R^2 = 0.36$, $n = 1253$).

[2011], and *Zib et al*. [2012] found that a negative bias in the modeled cloud fraction is associated with a positive bias in the reanalyzed surface solar radiation. *Wu et al*. [2012] found significant underestimation of cloud fraction (20–40%) in three reanalyses over the Southern Great Plains, USA and this underestimation is related to underestimation of other cloud properties such as cloud albedo and surface relative shortwave cloud forcing.

[30] Underestimation of the cloud liquid water path (LWP) is another possible source of error. The amount of cloud liquid water influences the reflection and absorption of shortwave radiation. *Zib et al*. [2012] found that the radiation biases still exist for some time periods despite that the observed cloud fraction is correctly reproduced by the model. *Cullather and Bosilovich* [2012] reported that MERRA cloud LWP is about 45% of that of microwave retrievals in the Arctic and this bias is consistent with biases in the surface net radiation. *Zhao and Wang* [2010] evaluated ECMWF cloud LWP against long-term observations at Barrow, Alaska in the Arctic during 1999–2007 and found that the model on average underestimates LWP by 30 g m^{-2}. Outside of the Arctic, *Duynkerke and Teixeira* [2001] found that the ECMWF reanalysis ERA-15 strongly underestimates marine stratocumulus cloud cover and LWP off the coast of California. Similar results were found in ERA-40 by *Stevens et al*. [2007].

[31] Figure 3 illustrates that the model bias errors tend to be larger at higher elevation sites. The regression equations are

$$b = -0.010{z_e}^2 + 0.070z_e + 0.63 \text{ for NARR}, \quad (11)$$

$$b = -0.0087{z_e}^2 + 0.065z_e + 0.51 \text{ for MERRA}, \quad (12)$$

In this figure, we have added the available high elevation FLUXNET validation and BSRN sites (elevation greater than 1 km). Even though the regressions were established with only two high elevation FLUXNET calibration sites, they captured the overall elevation dependence reasonably well. The choice of the quadratic fitting function is somewhat arbitrary. Use of a linear fit function would yield similar results (supporting information Table S2).

[32] The elevation-dependent intercept brings much improvement to the algorithm for high elevation sites, regardless of the form of the fitting function for the intercept parameter. For example, at the Niwot Ridge site in Colorado (site ID, US-NR1; elevation, 3050 m), the monthly mean NARR and MERRA were biased high by 62.6 and 39.4 Wm^{-2}, respectively (Table 2). Using the intercepts in equations (9) and (10) (without accounting for site-specific elevation) reduced the bias errors slightly to 61.5 and 35.7 Wm^{-2}, respectively. Using the elevation-dependent intercept parameters (equations (11) and (12)), the bias errors were reduced to 6.4 and 6.6 Wm^{-2} (Table 2). At a site in Tibet (elevation 4620 m, supporting information Table S1 [*Wang and Zeng*, 2012]), the monthly MERRA was biased high by 19.8 Wm^{-2}. Using the correction with the intercept given in equation (10) and with the elevation-dependent intercept (equation (12)) reduced the bias error to 18.5 and 2.0 Wm^{-2}, respectively. With the elevation-dependent intercept, significant improvements were also seen at Audubon Research Ranch, Arizona (elevation 1469 m) and Flagstaff Managed Forest, Arizona (elevation 2160 m; Table 1) and at five BRNS and the literature sites with elevation >1000 m (supporting information Table S1).

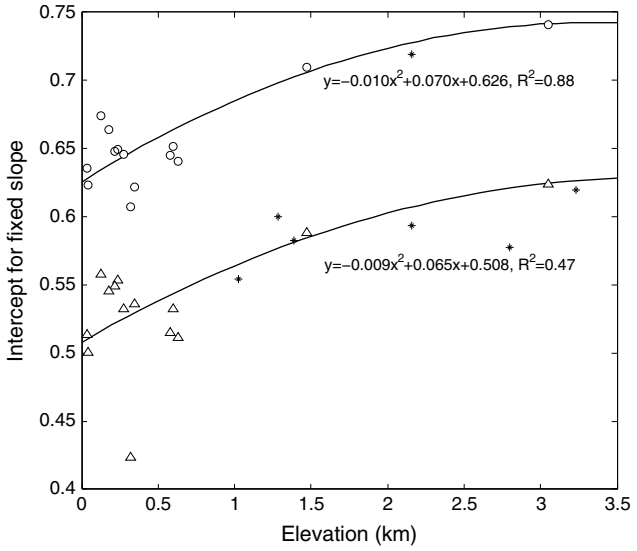

Figure 3. Relationship between the intercept coefficient of equation (4) and the site elevation for the calibration sites. The slope coefficient of equation (4) is fixed at −0.89 for NARR and −0.82 for MERRA. Circles: NARR; Triangles: MERRA; Star: available high elevation sites from BSRN and validation sites.

Table 2. Statistics of Monthly Mean Surface Incoming Solar Radiation Fluxes in NARR and MERRA Before and After Correction

	NARR			Corrected NARR			MERRA			Corrected MERRA		
Site Code	ME (Wm^{-2})	RMSE (Wm^{-2})	R^2	ME (Wm^{-2})	RMSE (Wm^{-2})	R^2	ME (Wm^{-2})	RMSE (Wm^{-2})	R^2	ME (Wm^{-2})	RMSE (Wm^{-2})	R^2
Calibration												
CA-Obs	31.3	35.1	0.98	−11.3	21.6	0.95	20.9	28.3	0.98	1.1	16.4	0.97
CA-Ojp	32.7	37.4	0.98	−7.8	20.5	0.94	22.0	29.1	0.98	3.1	16.8	0.98
CA-Oas	32.9	37.1	0.98	−7.7	19.5	0.95	23.8	30.6	0.98	5.6	17.3	0.98
CA-Ca1	33.6	37.5	0.98	−9.7	19.2	0.95	20.5	29.9	0.98	−0.3	22.2	0.97
US-UMB	35.6	37.6	0.99	6.4	17.8	0.98	20.9	22.2	0.99	7.2	13.5	0.98
CA-Cbo	43.4	46.1	0.98	4.9	19.3	0.95	28.5	31.3	0.98	10.4	17.1	0.97
US-NR1	62.6	70.8	0.96	6.4	22.0	0.96	39.4	46.5	0.97	6.6	23.0	0.96
US-MMS	40.9	42.8	0.98	1.5	15.6	0.95	22.4	25.7	0.97	2.5	16.3	0.95
US-Ton	22.1	23.9	0.99	9.0	22.2	0.97	4.7	10.8	0.99	−1.1	9.9	0.99
US-Var	26.5	28.2	0.99	14.9	24.9	0.98	8.6	12.7	0.99	3.4	10.4	0.99
US-WBW	36.3	38.1	0.97	−11.9	22.9	0.90	23.5	26.9	0.96	1.4	20.3	0.90
US-Aud	34.9	39.8	0.94	7.2	24.8	0.94	11.1	15.2	0.98	−3.2	24.4	0.89
US-SP2	37.8	40.9	0.90	−0.6	32.4	0.63	16.7	25.8	0.84	−2.7	34.5	0.58
US-SP3	43.4	47.3	0.86	6.6	35.5	0.59	22.0	31.5	0.78	3.3	36.6	0.53
Average	36.7	40.2	0.96	0.6	22.7	0.90	20.4	26.2	0.96	2.7	19.9	0.90
Validation												
CA-Qfo	30.9	35.9	0.99	−17.9	21.3	0.98	24.7	34.4	0.98	1.5	18.0	0.98
CA-Ca3	35.5	35.4	0.98	1.6	25.6	0.94	17.5	24.5	0.96	2.0	31.5	0.93
US-Ho1	37.7	43.8	0.97	−0.6	21.2	0.94	28.5	32.5	0.98	12.9	19.3	0.99
US-Bkg	35.4	40.7	0.97	8.0	24.2	0.94	12.4	20.1	0.97	−3.0	16.5	0.96
US-Bo1	39.6	43.2	0.95	4.5	19.3	0.93	23.7	30.8	0.93	8.3	23.5	0.91
US-Slt	48.5	53.0	0.98	19.6	31.4	0.96	27.6	30.5	0.98	12.7	18.8	0.97
US-MOz	39.4	41.8	0.98	12.0	23.5	0.96	16.4	19.9	0.98	1.5	14.0	0.96
US-Dk2	35.6	38.0	0.97	−0.6	20.5	0.90	15.4	19.8	0.95	−2.4	21.5	0.88
US-NC2	40.0	43.4	0.98	2.6	19.8	0.91	27.0	31.0	0.96	12.6	24.8	0.89
US-Fmf	38.5	48.9	0.87	0.8	23.0	0.96	14.7	24.0	0.94	−4.2	16.6	0.98
Average	38.1	42.4	0.96	3.0	23.0	0.94	20.8	26.7	0.96	4.2	20.5	0.95

3.2. Annual and Monthly Bias Errors Before and After Correction

[33] Equation (6) was used to obtain the corrected monthly mean S at both the calibration and validation sites. In this postreanalysis correction, the coefficient a was −0.89 for NARR and −0.82 for MERRA, and the coefficient b was given by equations (11) and (12) as a function of site elevation. The algorithm reduced the annual mean bias errors from +37.2 Wm^{-2} (range 23.6 to 59.8 W m^{-2}) and +20.2 Wm^{-2} (range 6.0 to 37 W m^{-2}), which are much larger than the measurement uncertainties of 3–5 Wm^{-2}, to +1.3 Wm^{-2} (range −17.9 to 19.6 W m^{-2}) and +2.7 Wm^{-2} (range −3.9 to 16.6 W m^{-2}) for NARR and MERRA, respectively (Table 1). Figure 4a shows the overestimation of surface solar radiation on the annual scale in NARR and MERRA. After correction, the annual mean values are evenly distributed along the 1:1 line as illustrated in Figure 4b.

[34] The monthly statistics for the individual sites are listed in Table 2. Because the data samples are slightly different, the monthly mean bias errors are not identical to the annual bias errors shown in Table 1. Averaging among all the sites, the correction algorithm reduced the magnitudes of both the monthly mean bias errors and RMSEs (root mean-square errors) for the two reanalyses. The values of R^2 were slightly reduced after the correction. For the calibration and the validation sites, the average R^2 were both 0.96 before correction for NARR and MERRA; after the correction, the average R^2 for the calibration sites were reduced to 0.90 for both NARR and MERRA, and for the validation sites reduced to 0.94 and 0.95, respectively. The larger reduction of the average R^2 for the calibration sites was caused by overcorrection at the two Florida sites, Mize (US-SP3) and Donaldson (US-SP2) near Gainesville. The calibration and validation sites behave very similarly in terms of the magnitudes of monthly mean bias errors, RMSE and R^2 before and after the correction, demonstrating robustness of the correction coefficients in equations (6), (11), and (12).

3.3. Temporal and Spatial Variations in Bias Errors

[35] The time series plots in Figure 5 illustrate that the corrected S tracked the observed interannual variabilities reasonably well at the six selected calibration sites (Old Aspen, Saskatchewan; Borden, Ontario; Morgan-Monroe State Forest, Indiana; Walker Branch, Tennessee; Donaldson, Florida; Vaira Ranch, California). These sites span a large latitudinal/longitudinal band, including both good (Figures 5a, 5b, and 5c) and one of the two worst sites (Donaldson, FL, Figure 5e) in terms of the algorithm performance. The monthly mean errors of the good sites were significantly reduced, with the correlation coefficients of monthly variability retained (Table 2), and the correlation coefficients of interannual variability were improved (Table 3). The two worst sites are Mize and Donaldson, both in Florida in the subtropical Mediterranean climate regime. Although the monthly mean error of Donaldson site was reduced, the correlation coefficients of monthly variability significantly decreased from 0.86 and 0.78 to 0.59 and 0.53 for NARR and MERRA, respectively.

[36] Table 3 lists the correlation coefficient of the observations with the modeled annual mean S before and after the

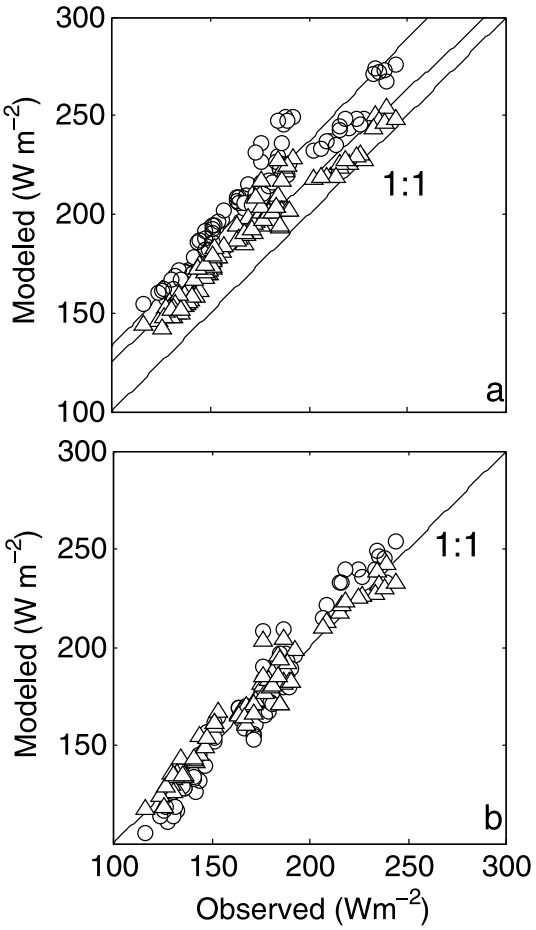

Figure 4. Comparison of the annual mean surface incoming shortwave radiation flux at the calibration sites. (a) before correction; (b) after correction. Circles: NARR; Triangles: MERRA. Each data point represents an annual mean value for a site year.

correction for the 14 calibration sites. (The validation sites are not listed as they do not have sufficient records for the computation of the correlation coefficient.) The two reanalysis systems generally capture the interannual variations (linear correlation $R > 0.6$), again with the two Florida sites being notable exceptions where both systems have negative correlation coefficients with the observations and the correction algorithm was unable to rectify this problem. Averaged across all the calibration sites, the correction algorithm improved the correlation coefficient slightly by 0.06 for NARR, and 0.08 for MERRA.

[37] Figure 6 illustrates the seasonal pattern of the bias errors before and after correction for the six selected sites as in Figure 5. Before correction, the seasonal cycles of bias were site dependent. Old Aspen, Saskatchewan and Borden, Ontario had maximum bias errors in June (Figures 6a and 6b). Morgan Monroe State Forest, Indiana, Walker Branch, Tennessee, and Donaldson, Florida had peak bias errors in April (Figures 6c–e). Vaira Ranch, California did not have much seasonal variation (Figure 6f). Generally the cold season bias was smaller than the warm season errors.

[38] After correction, the bias at the two northern sites (Old Aspen and Borden) had no obvious seasonal cycle. The algorithm overcorrected the modeled surface solar radiation in the summer at Donaldson (Figure 6e), and to a lesser degree at Morgan Monroe State Forest and Walker Branch (Figures 6c and 6d). At Donaldson, overcorrection caused negative monthly mean biases of up to 40 Wm^{-2} to both reanalyses. Similar magnitude of overcorrection was also found for Mize (site ID US-SP2) which is 8 km away from Donaldson. On the annual time scale, the warm season negative bias was compensated by the cold season positive bias, resulting in, fortuitously, much reduced mean bias (Table 2). At Vaira Ranch, our correction procedure did not bring improvements to S from May to October (Figure 6f). During these months, the clearness index was 0.70–0.75, which are beyond the k_t threshold of correction.

[39] The difficulty encountered at Varia Ranch and Donaldson indicates that factors other than cloudiness and elevation also contribute to the model bias errors. Measurement errors, according to the cross-site comparison of the radiometers at these sites against an Ameriflux roving standard, are too small to explain the anomalous results [*Schmidt et al*., 2012]. A possible explanation is incorrect description of aerosols at these locations in the models, noting that even the original NARR and MERRA data fail to capture the interannual variations of S. Despite this limitation, our simple correction algorithm has resulted overall reduction of the RMSEs for the sites we examined (from 41.1 to 22.8 W m^{-2} for NARR and from 26.4 to 20.1 W m^{-2} for MERRA; Table 2).

[40] The outliers in the left corner of Figure 2b came from the Douglas Fir 1949 site on Vancouver Island, British Columbia. The original MERRA data had very low and even slightly negative bias ratios. These points all came from the winter months with low monthly mean S (< 30 W m^{-2}; lowest among the sites we examined). Our algorithm caused slight overcorrection during these months. But the performance on the annual scale was not adversely affected, as evidenced by the large reduction of the mean bias error from 21.0 to -1.8 W m^{-2} (Table 1). This wintertime overcorrection was also found in Douglas Fir 1988 site, which is about 50 km away, and for the same reason, the algorithm was able to bring improvement, reducing the annual mean bias error from 17.5 to 2.0 W m^{-2} (Table 1).

3.4. Comparison Between the Two Reanalysis Products

[41] The error structures of the two products share several similar features. The bias ratios of both reanalyses show dependence on clearness index and surface elevation (Figures 2 and 3). In addition, the two products show similar seasonal variations in the bias errors (Figure 6). The reasonable correlations with the observed annual mean S suggest the similar interannual variations in the bias errors for the two reanalyses (Figure 5 and Table 3).

[42] There are also a number of differences. In general, MERRA shows better agreement with observations than NARR (Table 1 and Figure 4a). The mean errors of NARR and MERRA for all the sites were +37.3 and +20.5 Wm^{-2}, respectively, and their RMSEs were 41.1 and 26.4 Wm^{-2}, respectively (Table 2). The significantly positive bias in NARR likely resulted from a combination of underestimation of cloud and a lack of aerosols and water vapor in the atmospheric column [*Kennedy et al*., 2011]. After correction, the average bias errors of NARR and MERRA for all the sites

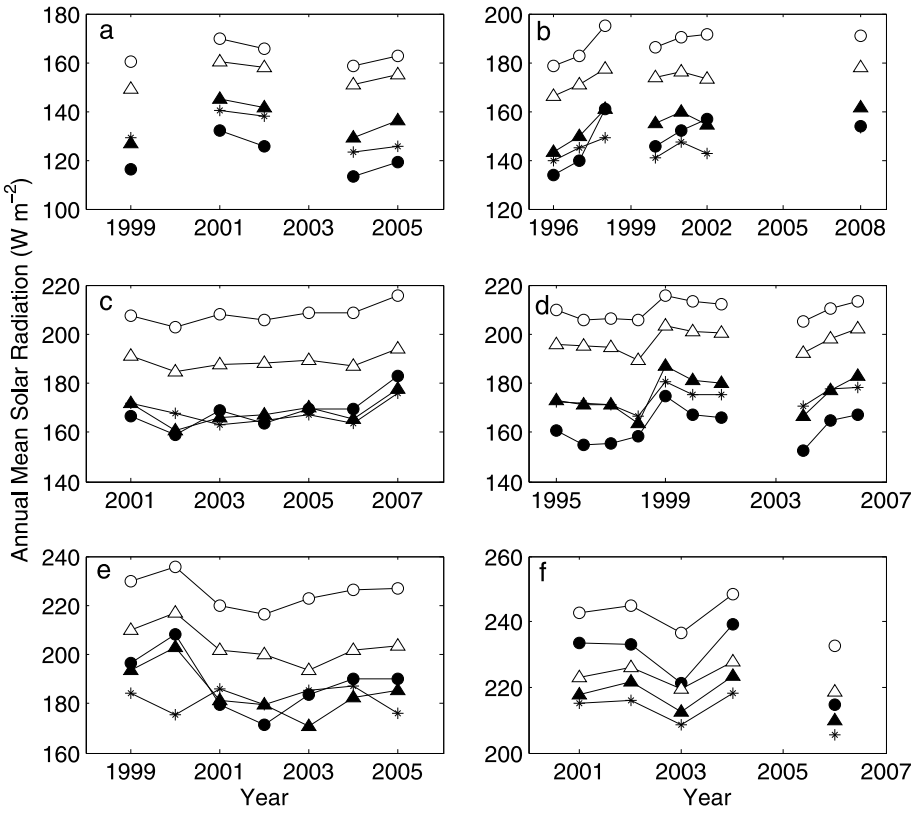

Figure 5. Annual mean surface incoming shortwave radiation flux at six selected sites. Open circles: NARR before correction; Black circles: NARR after correction; Open triangles: MERRA before correction; Black triangles: MERRA after correction; Stars: observations. (a) Old Aspen, Saskatchewan (site ID CA-Oas); (b) Borden, Ontario (site ID CA-Cbo); (c) Morgan-Monroe State Forest, Indiana (site ID US-MMS); (d) Walker Branch, Tennessee (site ID US-WBW); (e) Donaldson, Florida (site ID US-SP3); (f) Vaira Ranch, California (site ID US-Var).

were brought down to 1.6 and 3.3 Wm^{-2}, respectively and the RMSEs were 22.8 and 20.1 Wm^{-2}, respectively.

[43] Both uncorrected NARR and MERRA capture well the observed seasonal ($R^2 > 0.8$ for most sites) and interannual variations ($R > 0.6$ for most sites). MERRA showed no superiority compared to NARR in terms of capturing the observed seasonal and interannual variability. The correction procedure did not improve the capability of capturing seasonal variability and improved slightly interannual variability for the two reanalysis products (Table 3).

3.5. Implication for the Global Radiation and Energy Balances

[44] Our results indicate that MERRA overestimates the global mean *S*. In order to apply our algorithm on the global domain, examining the algorithm's performance outside of North America is required in addition to the performance evaluation done for the North American sites discussed above. Towards this goal, we selected 42 BSRN sites which have at least one year continuous *S* measurement and 8 sites from the literature (supporting information Figure S1 and Table S1). The algorithm was able to adjust the modeled annual mean *S* towards the 1:1 line for the sites both in and outside North America (Figure 7). Even though the algorithm was developed from the calibration sites on land, it improved the results at the BSRN ocean sites: the average bias error was 13.9 W m^{-2} before correction and −3.0 W m^{-2} after correction for these sites (supporting information Figure S2). Aerosol loading varies between land and ocean sites. The improvement at the ocean sites suggests that the clearness index as an independent variable has some capacity to implicitly account for the aerosol impact. For the sites outside North

Table 3. Correlation Coefficients Between the Modeled and the Observed Annual Mean Radiation at the Calibration Sites

Site Code	NARR	Corrected NARR	MERRA	Corrected MERRA
CA-Obs	0.85	0.98	0.74	0.82
CA-Ojp	0.65	0.52	0.61	0.67
CA-Oas	0.87	0.95	0.79	0.85
CA-Ca1	0.94	0.94	0.80	0.88
US-UMB	0.90	0.91	0.95	0.96
CA-Cbo	0.83	0.79	0.85	0.90
US-NR1	0.49	0.57	0.17	0.22
US-MMS	0.59	0.61	0.41	0.79
US-Ton	0.98	0.98	0.93	0.95
US-Var	0.86	1.00	0.83	0.97
US-WBW	0.90	0.84	0.94	0.96
US-Aud	0.24	0.25	0.31	0.22
US-SP2	−0.20	0.32	−0.13	0.12
US-SP3	−0.41	−0.35	−0.51	−0.50
Average	0.61	0.67	0.55	0.63

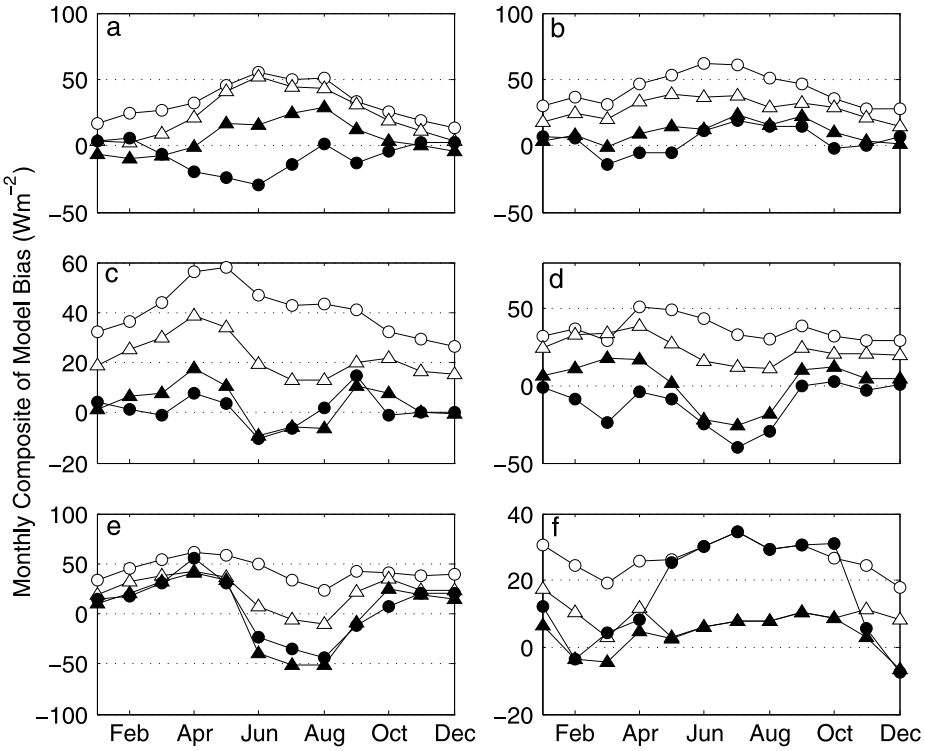

Figure 6. Same as Figure 5 except for monthly composite bias errors.

America, the mean bias error was 14.5 W m^{-2} before correction and −2.1 W m^{-2} after correction. Excluding the four obvious outliers, the mean bias error was 16.7 W m^{-2} and 1.3 W m^{-2} before and after correction, respectively. The four outlier sites are in the tropics (site IDs Llorin, Kwajalein, Momote and Tapajos). At these sites the original MERRA *S* matched reasonable well with the observations, and the correction algorithm caused a low bias. If all the 12 tropical sites are considered (Table S1), the MERRA bias is 15.4 W m^{-2} before correction and 0.7 W m^{-2} after correction.

[45] The reasonably good performance of the algorithm both in North America and elsewhere provides the basis of using it on the global domain. According to the default MERRA data, the global annual mean *S* is 192.8 Wm^{-2} for the year 2000–2004, the same years as used by *Trenberth et al*. [2009]. After correction, it was reduced to 175.5 Wm^{-2}. The default MERRA overestimates the global *S* by 17.3 Wm^{-2}.

[46] Obviously, if *S* is adjusted in the light of this study, equivalent adjustments on other terms of the radiation and energy balances are required for energy closure. In Table 4, we summarize the revised global surface radiation and energy balance and compare it with the estimates of *Trenberth et al*. [2009]. Briefly, in *Trenberth et al*. [2009], the global mean *S* is an improved estimate from that of *Kiehl and Trenberth* [1997] using ISCCP-FD and CERES and with an improved calculation of the absorption by atmospheric aerosol and water vapor. Their *S* value is very close to the CERES satellite estimate (186.7 Wm^{-2}) for the period 2000–2010 [*Kato et al*., 2013]. The global mean albedo is derived from field and satellite observations. The outgoing longwave radiation is derived from satellite observations of emissivity and the surface temperature. The sensible heat flux is the average of three reanalysis products spanning the range of 15.7 and 18.9 W m^{-2}. The latent heat flux is estimated under the assumption that precipitation is equal to global evaporation; this number is uncertain because considerable uncertainty exists in precipitation measurements, especially over the oceans. The incoming longwave radiation is computed as a residue of the surface energy balance.

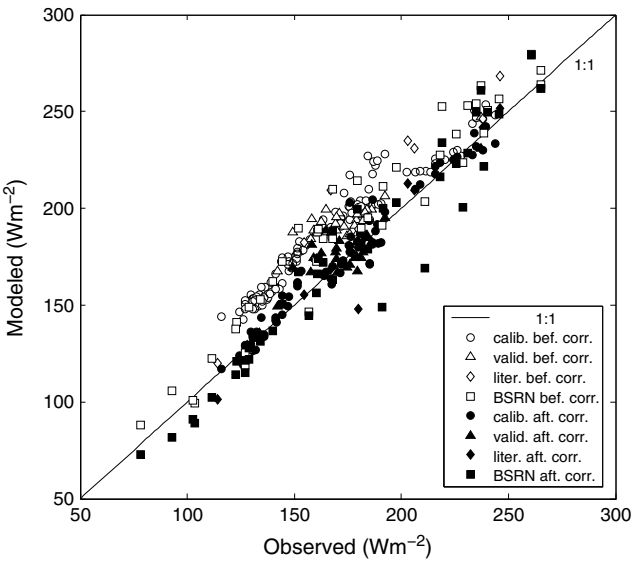

Figure 7. Comparison between observed and modeled annual mean surface incoming shortwave radiation flux before and after correction at the calibration, the validation, the BSRN, and the sites from the literature. Each data point represents measurement at one site.

Table 4. Comparison of Two Estimates of the Global Radiation and Energy Balances[a]

	S	S_u	$L\downarrow$	$L\uparrow$	H	LE
Trenberth et al. [2009]	184	23	333	396	17	80
This study	175	22	344	396	20	76

[a]Units are W m^{-2}. S: incoming solar radiation; S_u: reflected solar radiation; $L\downarrow$: incoming longwave radiation; $L\uparrow$: outgoing longwave radiation; H: surface sensible heat flux; LE: surface latent heat flux.

[47] In our revised depiction, the energy balance terms were estimated from independent sources and were independent of the energy balance constraint. The good energy balance closure (within 5 Wm^{-2}) serves as confirmation that the MERRA global S value was indeed biased high. Our assessment was based on the following considerations:

[48] 1. We estimated the reflected shortwave radiation by adopting the same albedo of *Trenberth et al*. [2009].

[49] 2. The incoming longwave radiation $L\downarrow$ was provided by *Stephens et al*. [2012b] according to the synthesis products. These authors found a systematic underestimation of reanalyzed $L\downarrow$ and attributed the bias also to the underestimation of modeled cloudiness. Globally this underestimation is on the order of 10 Wm^{-2}.

[50] 3. As for the outgoing longwave radiation, we adopted *Trenberth et al*.'s [2009] number because it is derived from observations, not from reanalysis products.

[51] 4. The revised latent heat flux was an area-weighted average of a recent estimate of the terrestrial [*Mu et al*., 2011] and the ocean latent heat flux [*Yu et al*., 2008]. *Mu et al*. [2011] provides a global land surface evapotransporation data set over 2000–2006 based on MODIS and a global meteorological reanalysis, which has been validated at 46 Ameriflux tower sites. The ocean data set provides multidecadal estimates of air-sea fluxes over global oceans using bulk transfer formulation. (Here we used the ocean flux data from 2000 to 2004, the same period as our S estimate.) The global mean latent heat flux is insensitive to uncertainties in the land evaporation estimates and is largely determined by the ocean value. Replacing *Mu et al*.'s [2011] land value by *Jung et al*.'s [2011] value (39 W m^{-2}) changes the global latent heat flux slightly, to 76.5 W m^{-2}. Similarly, there is a relatively large spread among a variety of the land latent heat flux products examined by *Jimenez et al*. [2011]; when combined with the ocean flux of *Yu et al*. [2008], the resulting global mean value lies in a very narrow range of 75.2–77.0 W m^{-2}.

[52] 5. Similarly, the revised estimate of sensible heat flux was a combination of the land [*Jung et al*., 2011] and the ocean flux [*Yu et al*., 2008]. *Jung et al*. [2011] applied a machine learning technique-model tree ensembles to upscale FLUXNET observations to the global scale from 1982 to 2008. We adopted their long-term mean value of the terrestrial sensible heat flux.

[53] The downward adjustment in S is mostly compensated by an upward adjustment of the incoming longwave radiation by a similar amount. A recent study by *Stephens et al*. [2012a] suggests that much of the extra incoming longwave radiation (as compared to Trenberth's assessment) to the surface is offset by more latent heat flux from the surface. Our study suggests an alternative hypothesis that the compensation exists between S and $L\downarrow$without the need to adjust the other energy balance terms significantly.

4. Conclusions

[54] In this study, the surface incoming shortwave radiation S modeled by two data assimilation systems, NARR and MERRA, was evaluated against observations from 24 FLUXNET sites in the US and Canada at multiple time scales. NARR and MERRA systematically overestimated the surface solar radiation flux on both monthly and annual scales. Their bias errors were larger under cloudy skies than under clear skies and increased with increasing elevation. The two products show similar capability to reproduce the seasonal and interannual variations of S, and similar seasonal variations in the bias errors. MERRA generally shows better agreement than NARR with the flux tower measurements.

[55] A simple postreanalysis correction algorithm was proposed on the basis of the dependence of the bias on sky clearness and surface elevation. Results show that the correction algorithm worked well on the annual scale for the FLUXNET sites in North America; it reduced the annual mean bias errors from +37.2 Wm^{-2} and +20.2 Wm^{-2} to +1.3 Wm^{-2} and +2.7 Wm^{-2} for NARR and MERRA, respectively. The algorithm slightly improved the modeled interannual variability for the two products. The algorithm showed good performance as well for sites outside North America except for four tropical sites.

[56] There are a few limitations to this algorithm. The algorithm overcorrected S in Florida in the summer and the annual mean S at four sites near the equator. But this simple algorithm was able to reduce the overall mean bias errors and the RMSEs of the sites considered.

[57] The global mean S was 192.8 W m^{-2} for 2000 to 2004 according to MERRA. The correction algorithm reduced it by 9.0% to 175.5 W m^{-2}. This corrected S is 12.5 and 8.8 W m^{-2} lower than that given by *Stephens et al*. [2012a] and *Trenberth et al*. [2009]. It appears that various modeled products likely have similar problems in underestimating the atmospheric absorption of shortwave radiation.

[58] **Acknowledgments.** This research was supported in part by the Ministry of Education of China (grant PCSIRT), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PARD), the Yale Climate and Energy Institute, and a Yale University Graduate Fellowship. We thank three journal reviewers for their helpful comments.

References

Baldocchi, D., C. A. Vogel, and B. Hall (1997), Seasonal variation of energy and water vapor exchange rates above and below a boreal jack pine forest canopy, *J. Geophys. Res.*, *102*(D24), 28,939–28,951.

Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteorol. Soc.*, *82*(11), 2415–2434.

Bergeron, O., H. A. Margolis, T. A. Black, C. Coursolle, A. L. Dunn, A. G. Barr, and S. C. Wofsy (2007), Comparison of carbon dioxide fluxes over three boreal black spruce forests in Canada, *Global Change Biol.*, *13*(1), 89–107.

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo (1996), The land surface-atmosphere interaction: A review based on observational and global modeling perspectives, *J. Geophys. Res.*, *101*(D3), 7209–7225.

Blanken, P. D., T. A. Black, P. C. Yang, H. H. Neumann, Z. Nesic, R. Staebler, G. den Hartog, M. D. Novak, and X. Lee (1997), Energy balance and canopy conductance of a boreal aspen forest: Partitioning overstory and understory components, *J. Geophys. Res.*, *102*(D24), 28,915–28,927.

Bloom, S. C., L. L. Takacs, A. M. DaSilva, and D. Ledvina (1996), Data assimilation using incremental analysis updates, *Mon. Weather Rev.*, *124*(6), 1256–1271.

Bosilovich, M. G., F. R. Robertson, and J. Y. Chen (2011), Global energy and water budgets in MERRA, *J. Clim.*, *24*(22), 5721–5739.

Chou, M.-D., and M. J. Suarez (1999), A solar radiation parameterization for atmospheric studies, *NASA Tech. Rep. NASA/TM-1999-104606*, 40 pp, Series on Global Modeling and Data Assimilation.

Chou, M.-D., M. J. Suarez, X. Z. Liang, and M. M.-H. Yan (2001), A thermal infrared radiation parameterization for atmospheric studies, *NASA Tech. Rep. NASA/TM-2001-104606*, 56 pp, Series on Global Modeling and Data Assimilation.

Clark, K. L., N. Skowronski, and J. Hom (2010), Invasive insects impact forest carbon dynamics, *Global Change Biol.*, *16*(1), 88–101.

Cullather, R. I., and M. G. Bosilovich (2012), The energy budget of the polar atmosphere in MERRA, *J. Clim.*, *25*(1), 5–24.

Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. B. Zeng, and M. G. Bosilovich (2012), Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations, *J. Clim.*, *25*(6), 1916–1944.

Dore, S., T. E. Kolb, M. Montes-Helu, B. W. Sullivan, W. D. Winslow, S. C. Hart, J. P. Kaye, G. W. Koch, and B. A. Hungate (2008), Long-term impact of a stand-replacing fire on ecosystem CO(2) exchange of a ponderosa pine forest, *Global Change Biol.*, *14*(8), 1801–1820.

Duynkerke, P. G., and J. Teixeira (2001), Comparison of the ECMWF reanalysis with FIRE I observations: Diurnal variation of marine stratocumulus, *J. Clim.*, *14*(7), 1466–1478.

Fels, S. B., and M. D. Schwarzkopf (1975), Simplified exchange approximation: New method for radiative-transfer calculations, *J. Atmos. Sci.*, *32*(7), 1475–1488.

Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego (2002), Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta Model, paper presented at 19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction, Amer. Meteor. Soc., San Antonio, TX.

Frauenfeld, O. W., T. J. Zhang, and M. C. Serreze (2005), Climate change and variability using European Centre for Medium-Range Weather Forecasts reanalysis (ERA-40) temperatures on the Tibetan Plateau, *J. Geophys. Res.*, *110*, D02101, doi:10.1029/2004JD005230.

Gholz, H. L., and K. L. Clark (2002), Energy exchange across a chronosequence of slash pine forests in Florida, *Agric. For. Meteorol.*, *112*(2), 87–102.

Gilmanov, T. G., L. L. Tieszen, B. K. Wylie, L. B. Flanagan, A. B. Frank, M. R. Haferkamp, T. P. Meyers, and J. A. Morgan (2005), Integration of CO2 flux and remotely-sensed data for primary production and ecosystem respiration analyses in the Northern Great Plains: Potential for quantitative spatial extrapolation, *Global Ecol. Biogeogr.*, *14*(3), 271–292.

Gu, L. H., J. D. Fuentes, H. H. Shugart, R. M. Staebler, and T. A. Black (1999), Responses of net ecosystem exchanges of carbon dioxide to changes in cloudiness: Results from two North American deciduous forests, *J. Geophys. Res.*, *104*(D24), 31,421–31,434.

Gu, L. H., T. Meyers, S. G. Pallardy, P. J. Hanson, B. Yang, M. Heuer, K. P. Hosman, J. S. Riggs, D. Sluss, and S. D. Wullschleger (2006), Direct and indirect effects of atmospheric conditions and soil moisture on surface energy partitioning revealed by a prolonged drought at a temperate forest site, *J. Geophys. Res.*, *111*, D16102, doi:10.1029/2006JD007161.

Hatzianastassiou, N., C. Matsoukas, A. Fotiadi, K. G. Pavlakis, E. Drakakis, D. Hatzidimitriou, and I. Vardavas (2005), Global distribution of Earth's surface shortwave radiation budget, *Atmos. Chem. Phys.*, *5*, 2847–2867.

Hickman, J. E., S. L. Wu, L. J. Mickley, and M. T. Lerdau (2010), Kudzu (Pueraria montana) invasion doubles emissions of nitric oxide and increases ozone pollution, *Proc. Natl. Acad. Sci. U. S. A.*, *107*(22), 10,115–10,119.

Hollinger, D. Y., S. M. Goltz, E. A. Davidson, J. T. Lee, K. Tu, and H. T. Valentine (1999), Seasonal patterns and environmental control of carbon dioxide and water vapour exchange in an ecotonal boreal forest, *Global Change Biol.*, *5*(8), 891–902.

Hollinger, S. E., C. J. Bernacchi, and T. P. Meyers (2005), Carbon budget of mature no-till ecosystem in North Central Region of the United States, *Agric. For. Meteorol.*, *130*(1–2), 59–69.

Humphreys, E. R., T. A. Black, G. J. Ethier, G. B. Drewitt, D. L. Spittlehouse, E. M. Jork, Z. Nesic, and N. J. Livingston (2003), Annual and seasonal variability of sensible and latent heat fluxes above a coastal Douglas-fir forest, British Columbia, Canada, *Agric. For. Meteorol.*, *115*(1–2), 109–125.

Humphreys, E. R., T. A. Black, K. Morgenstern, T. B. Cai, G. B. Drewitt, Z. Nesic, and J. A. Trofymow (2006), Carbon dioxide fluxes in coastal Douglas-fir stands at different stages of development after clearcut harvesting, *Agric. For. Meteorol.*, *140*(1–4), 6–22.

Jarvis, P. G., J. M. Massheder, S. E. Hale, J. B. Moncrieff, M. Rayment, and S. L. Scott (1997), Seasonal variation of carbon dioxide, water vapor, and energy exchanges of a boreal black spruce forest, *J. Geophys. Res.*, *102*(D24), 28,953–28,966.

Jimenez, C., et al. (2011), Global intercomparison of 12 land surface heat flux estimates, *J. Geophys. Res.*, *116*, D02102, doi:10.1029/2010JD014545.

Jung, M., et al. (2011), Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *J. Geophys. Res.*, *116*, D02102, doi:10.1029/2010JD014545.

Kato, S., N. G. Loeb, F. G. Rose, D. R. Doelling, D. A. Rutan, T. E. Caldwell, L. S. Yu, and R. A. Weller (2013), Surface irradiances consistent with CERES-derived top-of-atmosphere shortwave and longwave irradiances, *J. Clim.*, *26*(9), 2719–2740.

Kennedy, A. D., X. Q. Dong, B. K. Xi, S. C. Xie, Y. Y. Zhang, and J. Y. Chen (2011), A comparison of MERRA and NARR reanalyses with the DOE ARM SGP data, *J. Clim.*, *24*(17), 4541–4557.

Kiehl, J. T., and K. E. Trenberth (1997), Earth's annual global mean energy budget, *Bull. Am. Meteorol. Soc.*, *78*(2), 197–208.

Krishnan, P., T. P. Meyers, R. L. Scott, L. Kennedy, and M. Heuer (2012), Energy exchange and evapotranspiration over two temperate semi-arid grasslands in North America, *Agric. For. Meteorol.*, *153*, 31–44.

Lacis, A. A., and J. E. Hansen (1974), Parameterization for absorption of solar-radiation in Earth's atmosphere, *J. Atmos. Sci.*, *31*(1), 118–133.

Lee, X., J. D. Fuentes, R. M. Staebler, and H. H. Neumann (1999), Long-term observation of the atmospheric exchange of CO2 with a temperate deciduous forest in southern Ontario, Canada, *J. Geophys. Res.*, *104*(D13), 15,975–15,984.

Lee, X., et al. (2011), Observed increase in local cooling effect of deforestation at higher latitudes, *Nature*, *479*(7373), 384–387.

Leung, L. R., and W. I. Gustafson (2005), Potential regional climate change and implications to US air quality, *Geophys. Res. Lett.*, *32*, L16711, doi:10.1029/2005GL022911.

Ma, S. Y., D. D. Baldocchi, L. K. Xu, and T. Hehn (2007), Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in California, *Agric. For. Meteorol.*, *147*(3–4), 157–171.

Markovic, M., C. G. Jones, K. Winger, and D. Paquin (2009), The surface radiation budget over North America: Gridded data assessment and evaluation of regional climate models, *Int. J. Climatol.*, *29*(15), 2226–2240.

Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*(3), 343–360.

Mickley, L. J., D. J. Jacob, B. D. Field, and D. Rind (2004), Climate response to the increase in tropospheric ozone since preindustrial times: A comparison between ozone and equivalent CO2 forcings, *J. Geophys. Res.*, *109*, D05106, doi:10.1029/2003JD003653.

Monson, R. K., A. A. Turnipseed, J. P. Sparks, P. C. Harley, L. E. Scott-Denton, K. Sparks, and T. E. Huxman (2002), Carbon sequestration in a high-elevation, subalpine forest, *Global Change Biol.*, *8*(5), 459–478.

Mu, Q. Z., M. S. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sens. Environ.*, *115*(8), 1781–1800.

Nolte, C. G., A. B. Gilliland, C. Hogrefe, and L. J. Mickley (2008), Linking global to regional models to assess future climate impacts on surface ozone levels in the United States, *J. Geophys. Res.*, *113*, D14307, doi:10.1029/2007JD008497.

Noormets, A., M. J. Gavazzi, S. G. McNulty, J. C. Domec, G. Sun, J. S. King, and J. Q. Chen (2010), Response of carbon fluxes to drought in a coastal plain loblolly pine forest, *Global Change Biol.*, *16*(1), 272–287.

Ohmura, A., et al. (1998), Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research, *Bull. Am. Meteorol. Soc.*, *79*(10), 2115–2136.

Oren, R., C. I. Hseih, P. Stoy, J. Albertson, H. R. McCarthy, P. Harrell, and G. G. Katul (2006), Estimating the uncertainty in annual net ecosystem carbon exchange: Spatial variation in turbulent fluxes and sampling errors in eddy-covariance measurements, *Global Change Biol.*, *12*(5), 883–896.

Richardson, A. D., X. Lee, and A. J. Friedland (2004), Microclimatology of treeline spruce-fir forests in mountains of the northeastern United States, *Agric. For. Meteorol.*, *125*(1–2), 53–66.

Rienecker, M. M., et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, *24*(14), 3624–3648.

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.

Rotenberg, E., and D. Yakir (2010), Contribution of semi-arid forests to the climate system, *Science*, *327*(5964), 451–454.

Schmid, H. P., C. S. B. Grimmond, F. Cropley, B. Offerle, and H. B. Su (2000), Measurements of CO2 and energy fluxes over a mixed hardwood

forest in the mid-western United States, *Agric. For. Meteorol.*, *103*(4), 357–374.

Schmid, H. P., H. B. Su, C. S. Vogel, and P. S. Curtis (2003), Ecosystem-atmosphere exchange of carbon dioxide over a mixed hardwood forest in northern lower Michigan, *J. Geophys. Res.*, *108*(D14), 4417, doi:10.1029/2002JD003011.

Schmidt, A., C. Hanson, W. S. Chan, and B. E. Law (2012), Empirical assessment of uncertainties of meteorological parameters and turbulent fluxes in the AmeriFlux network, *J. Geophys. Res.*, *117*, G04014, doi:10.1029/2012JG002100.

Stephens, G. L., J. Li, M. Wild, C. A. Clayson, N. Loeb, S. Kato, T. L'Ecuyer, P. W. Stackhouse, M. Lebsock, and T. Andrews (2012a), An update on Earth's energy balance in light of the latest global observations, *Nat. Geosci.*, *5*(10), 691–696.

Stephens, G. L., M. Wild, P. W. Stackhouse, T. L'Ecuyer, S. Kato, and D. S. Henderson (2012b), The global character of the flux of downward longwave radiation, *J. Clim.*, *25*(7), 2329–2340.

Stevens, B., A. Beljaars, S. Bordoni, C. Holloway, M. Kohler, S. Krueger, V. Savic-Jovcic, and Y. Y. Zhang (2007), On the structure of the lower troposphere in the summertime stratocumulus regime of the northeast Pacific, *Mon. Weather Rev.*, *135*(3), 985–1005.

Trenberth, K. E., J. T. Fasullo, and J. Kiehl (2009), Earth's global energy budget, *Bull. Am. Meteorol. Soc.*, *90*(3), 311–323.

Viterbo, P. A., and P. Courtier (1995), The importance of soil water for medium-range weather forecasting, paper presented at Implications for Data Assimilation, Workshop on Imbalance of Slowly Varying Components of Predictable Atmospheric Motions, World Meteorol. Org., Beijing, China.

Walsh, J. E., W. L. Chapman, and D. H. Portis (2009), Arctic cloud fraction and radiative fluxes in atmospheric reanalyses, *J. Clim.*, *22*(9), 2316–2334.

Wang, A., and X. Zeng (2012), Evaluation of multireanalysis products with in situ observations over the Tibetan Plateau, *J. Geophys. Res.*, *117*, D05102, doi:10.1029/2011JD016553.

Weaver, C. P., et al. (2009), A preliminary synthesis of modeled climate change impacts on U.S. regional ozone concentrations, *Bull. Am. Meteorol. Soc.*, *90*(12), 1843–1863.

Wild, M. (2009), Global dimming and brightening: A review, *J. Geophys. Res.*, *114*, D00D16, doi:10.1029/2008JD011470.

Wild, M., A. Ohmura, H. Gilgen, and E. Roeckner (1995), Validation of general-circulation model radiative fluxes using surface observations, *J. Clim.*, *8*(5), 1309–1324.

Wild, M., A. Ohmura, H. Gilgen, and J. J. Morcrette (1998), The distribution of solar energy at the Earth's surface as calculated in the ECMWF Re-Analysis, *Geophys. Res. Lett.*, *25*(23), 4373–4376.

Wilson, K. B., and T. P. Meyers (2001), The spatial variability of energy and carbon dioxide fluxes at the floor of a deciduous forest, *Bound.-Lay. Meteorol.*, *98*(3), 443–473.

Wu, W., Y. G. Liu, and A. K. Betts (2012), Observationally based evaluation of NWP reanalyses in modeling cloud properties over the Southern Great Plains, *J. Geophys. Res.*, *117*, D12202, doi:10.1029/2011JD016971.

Yu, L., X. Jin, and R. A. Weller (2008), Multidecade global flux datasets from the Objectively Analyzed Air-sea Fluxes (OAFlux) Project: Latent and sensible heat fluxes, ocean evaporation, and related surface meteorological variables, *OAFlux Project Technical Rep. OA-2008-01*, 64 pp, Woods Hole Oceanographic Institution, Woods Hole. Massachusetts.

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*(D1), 1149–1165.

Zhang, Y. C., 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.

Zhao, M., and Z. E. Wang (2010), Comparison of Arctic clouds between European Center for Medium-Range Weather Forecasts simulations and Atmospheric Radiation Measurement Climate Research Facility long-term observations at the North Slope of Alaska Barrow site, *J. Geophys. Res.*, *115*, D23202, doi:10.1029/2010JD014285.

Zhao, P., Z. J. Zhou, and J. P. Liu (2007), Variability of Tibetan spring snow and its associations with the hemispheric extratropical circulation and East Asian summer monsoon rainfall: An observational investigation, *J. Clim.*, *20*(15), 3942–3955.

Zib, B. J., X. Q. Dong, B. K. Xi, and A. Kennedy (2012), Evaluation and intercomparison of cloud fraction and radiative fluxes in recent reanalyses over the Arctic using BSRN surface observations, *J. Clim.*, *25*(7), 2291–2305.