A comparison of satellite based, modeled derived daily solar radiation data with observed data for the continental US

Jeffrey W. White[a,*], Gerrit Hoogenboom[b†], Paul W. Wilkens[c], Paul W. Stackhouse Jr.[d], James M. Hoell[e]

[a] U.S. Arid Land Agricultural Research Center, USDA-ARS,
21881 N. Cardon Lane, Maricopa, AZ 85138
[b] University of Georgia, Dep. Biol. & Ag. Eng, 1109 Experiment St., Griffin, GA 30223-1797
[c] International Fertilizer Development Center, PO Box 2040, Muscle Shoals, AL 35662
[d] NASA Langley Research Center, 21 Langley Blvd., Mail Stop 420, Hampton, VA 23681-2199
[e]SSAI/Langley Research Center, One Enterprise Pkwy., Hampton, VA 23666-5845

* *Corresponding author:* Tel. +1 520 316 6368; fax +1 520 316 6330.
E-mail: jeffrey.white@ars.usda.gov
† Current address: AgWeatherNet, Washington State University, 24106 North Bunn Road, Prosser, WA 99350-8694

Key words
Data quality, Decision Support Systems, remote sensing, weather data

Abstract

Many applications of simulation models and related decision support tools for agriculture and natural resource management require daily meteorological data as inputs. Availability and quality of such data, however, often constrain research and decision support activities that require use of these tools. Daily solar radiation (SRAD) data are especially problematic because the instruments require electronic integrators, accurate sensors are expensive, and calibration standards are seldom available. The Prediction Of Worldwide Energy Resources (NASA/POWER; power.larc.nasa.gov) project at the NASA Langley Research Center estimates daily solar radiation based on data that are derived from satellite observations of outgoing visible radiances and atmospheric parameters based upon satellite observations and assimilation models. The solar data are available for a global 1° x 1° coordinate grid. SRAD can also be estimated based on attenuation of extraterrestrial radiation (Q_0) using daily temperature and rainfall data to estimate the optical thickness of the atmosphere. This study compares daily solar radiation data from

NASA/POWER ($SRAD_{NP}$) with instrument readings from 295 stations ($SRAD_{OB}$), as well as with values that were estimated with the WGENR solar generator. WGENR was used both with daily temperature and precipitation records from the stations reporting solar data and records from the NOAA Cooperative Observer Program (COOP), thus providing two additional sources of solar data, $SRAD_{WG}$ and $SRAD_{CO}$. Values of $SRAD_{NP}$ for different grid cells consistently showed higher correlations (typically 0.85 to 0.95) with $SRAD_{OB}$ data than did $SRAD_{WG}$ or $SRAD_{CO}$ for sites within the corresponding cells. Mean values of $SRAD_{OB}$, $SRAD_{WG}$ and $SRAD_{NP}$ for sites within a grid cell usually were within 1 $MJm^{-2}d^{-1}$ of each other, but NASA/POWER values averaged 1.1 $MJm^{-2}d^{-1}$ lower than $SRAD_{OB}$. The magnitude of this bias was greater at lower latitudes and during summer months and may be at least partially explained by assumptions in ambient aerosol properties. Overall, the NASA/POWER solar radiation data are a promising resource for regional modeling studies where realistic accounting of historic variation is required.

Abbreviations

COOP, NOAA National Weather Service Cooperative Observer Program; NASA/POWER, NASA Prediction Of Worldwide Energy Resources; Q_0, daily integral of extraterrestrial insolation; RMSE, root mean squared error; SRAD, daily integral of solar radiation; T_{max}, daily maximum temperature; T_{min}, daily minimum temperature

1. Introduction

Many agricultural and natural resource management efforts involve spatial scales above the field and farm levels. Applications range from monitoring regional water use, to identifying promising zones for production of new crops, to targeting of specific cultivars or crop traits, to determining the potential impact of climate change and potential options for adaptation. Spatial assessments often consider climatic variation and increasingly, long-term records of daily weather data are required to examine climatic risks or trends related to climate change. Such analyses, however, are usually constrained by the availability and quality of the observed long-term meteorological data. Weather stations may not be available in the regions of interest, and individual stations may lack data for long time intervals. Weather data *per se* may show local variation due to positioning of the station and the instrument, instrument calibration drift, change in instrumentation, and other factors (Younes et al., 2005; Davey and Pielke, 2005). Solar radiation data have long been recognized as especially problematic (Durrenberger and Brazel, 1976; Stoffel et al., 2000). Radiation must be correctly integrated at low sun elevation angles and over all wavelengths. Radiometers using thermopiles are expensive, while lower-cost silicon pyranometers are less accurate. Both types of sensors require electronic circuitry to integrate readings over time and are sensitive to ambient temperatures. Sensor calibration is difficult because accurate reference values (besides 0) cannot be produced through simple techniques; thus sensors are usually cross-calibrated to radiometers whose calibrations are traceable to standards such as those maintained by the National Institute of Standards and Technology.

The Prediction Of Worldwide Energy Resources (NASA/POWER) project at the NASA Langley Research Center provides daily data for surface solar radiation and other weather variables on a 1° x 1° geographic coordinate grid for the entire globe (Stackhouse, 2010a; See Table 1 for an overview of the data available from the POWER archive.). The solar data are inferred from satellite observations of the outgoing top of atmosphere (TOA) radiances via an updated version of the Pinker and Laszlo (1992) radiative transfer based algorithm that was used to produce the fluxes for the Global Energy and Water Cycle Experiment (GEWEX) Surface Radiation Budget (SRB) project solar algorithm v2.81 (Gupta et al., 2006). Within this algorithm, a calculated TOA albedo is matched to an inferred TOA albedo from measured dark clear-sky background and instantaneous (every 3 hours) clear-sky and cloudy-sky satellite visible radiances using a radiative transfer model (through the use of lookup tables) on a 1°x1° degree grid. Using the background clear-sky radiance and information about the atmosphere (e.g., water vapor and ozone) the radiative transfer model infers an absolute surface albedo for a particular time and location. This step assumes a background aerosol and an assumed spectral albedo shape based upon the most prevalent surface type of the area. Then the surface albedo and the other input information are used

to infer the cloud and aerosol optical depths needed to match the observed clear and cloudy sky TOA albedos within a certain tolerance. The algorithm computes solar irradiance for clear and cloudy conditions using the inferred clear and cloud optical depths respectively and the other atmospheric information. The total solar irradiance is computed as the weighted sum of the clear and cloud fluxes and is integrated over the day to produce the daily averaged fluxes. The NASA/GEWEX SRB solar data currently in the POWER archive span the time period from July 1, 1983 through December 2007. Appended to this time series are irradiances estimate from the NASA CERES (Clouds and Earth Radiant Energy System) FLASHFlux (Fast Longwave and SHortwave radiative Fluxes from CERES and MODIS) data sets (Stackhouse, 2010b) spanning from Jan 2008 through within one week of the current date (for algorithm description see, Kratz et al., 2010; the latter data are not considered in this paper). These data may be downloaded with other variables in a format compliant with the standards of the International Consortium for Agricultural Systems Applications (ICASA; www.ICASA.net, Hunt et al., 200, 2006), which facilitates use in decision support tools such as the Decision Support System for Agrotechnology Transfer (DSSAT; Hoogenboom et al., 2010).

Although initially developed for applications related to solar energy, energy consumption, and energy conservation, the NASA/POWER data appear suitable for agriculture and natural resource management (White et al., 2009; Bai et al., 2010). Their coarse geographic scale, however, may limit their usefulness: one degree of longitude is approximately 110 km at the equator and 80 km at 45° latitude. In assessing the effect of spatial resolution of precipitation and radiation data on regional yield forecasts, de Wit et al. (2005) concluded that a 50 × 50 km grid provided an adequate resolution. Similarly, for climate change research in the contiguous US, Janis et al. (2004) concluded that a network of 327 stations was adequate to monitor a 0.10°C decade^{-1} temperature trend. Besides spatial resolution, there remains the question of whether the data assimilation process introduced important bias or other error in the data.

Faced with a lack of reliable solar radiation data, numerous researchers have opted for generating values using data on latitude, temperature and precipitation as inputs. The procedures first estimate the daily extraterrestrial insolation (Q_0) based on latitude, date and the solar constant. This value is then reduced based on atmospheric transmittance or similar considerations (e.g., Richardson, 1981; Bristow and Campbell, 1984; Richardson and Wright, 1984; Hodges et al., 1985; Cooter and Dhakhwa, 1995; Liu and Scott, 2001). Transmittance is typically estimated from region-specific relations that consider the diurnal range of air temperature, which may further be varied depending on whether precipitation occurred that day. This paper compares the NASA/POWER solar radiation data with data from weather stations reporting instrument-based observations and from an implementation of the WGENR solar radiation generator (Garcia y Garcia and Hoogenboom, 2005; Garcia y Garcia et al., 2008).

2. Materials and Methods

Daily data for solar radiation from NASA/POWER ($SRAD_{NP}$) were obtained from the web site (power.larc.nasa.gov; Stackhouse, 2010a), which allows for downloading of the data in several ASCII based formats. The dataset covered the continental US on a 1° x 1° latitude and longitude grid, representing 867 grid cells. The time interval considered was from 1 Jul. 1983 to 31 Dec. 2004. The NASA/GEWEX SRB solar version v2.81, as identified above, corresponds to the dataset used in our previous comparison of temperature data (White et al., 2008).

Observed solar radiation data ($SRAD_{OB}$) were obtained mainly from Internet sources such as state or regional climate networks (Supplement Table 1), which typically report data from automated weather stations using silicon pyranometers that output a current signal. Instantaneous values are registered and integrated digitally. For the widely used LI-200 Pyranometer[1], LI-COR (2005) states that these sensors are calibrated against an Eppley Precision Spectral Pyranometer (PSP) using natural daylight, and the maximum absolute error is typically ± 3%, with a maximum of ± 5%. Datasets from stations were rejected if they provided less than two years of data between 1983 and 2004, in order to match the period represented in our NASA/POWER dataset. Data reported for several stations were clearly incorrect, including values much larger than Q_0, negative values and values with a large systematic bias. Where detected, problem values were excluded based on the following criteria: $SRAD_{OB}$ greater than Q_0, $SRAD_{OB}$ less than 0.2 $MJm^{-2}d^{-1}$, or time series of $SRAD_{OB}$ showing large, systematic deviations from patterns observed in other years, values of Q_0 or nearby locations. After the quality control process, a total of 295 stations were available, which were located in 181 grid cells of the NASA/POWER dataset (Fig. 1).

The WGENR program (Hodges et al., 1985; Garcia y Garcia and Hoogenboom, 2005), which uses the Richardson approach (Richardson and Wright, 1984) to estimate daily values of SRAD, was used to obtain two additional estimates of solar radiation from observed values for daily maximum and minimum temperatures and daily precipitation. The first estimate ($SRAD_{WG}$) was obtained using daily temperature and precipitation records from the datasets of the observed values of solar radiation and thus coincided with the source locations of the $SRAD_{OB}$. A parallel set of daily data ($SRAD_{CO}$) were estimated from daily temperature and precipitation records from 855 individual ground stations from the National Weather Service Cooperative Observer Program (COOP). The COOP stations were selected based on

[1] Mention of a trademark, proprietary product, or vendor is for information only and does not constitute an endorsement by the USDA, NASA or the University of Georgia.

their being nearest to the centroid of a given grid cell of the NASA/POWER data set (White et al., 2008). Use of the COOP observations provided $SRAD_{CO}$ for 855 grid cells of the NASA/POWER dataset.

Comparisons of the solar radiation were based primarily on Pearson product-moment correlations calculated for daily data from individual stations, which were calculated using the Correlation procedure (PROC CORR) of the SAS 9.2 TS (SAS Institute Inc., Cary, NC, USA). Since annual variation in SRAD is dominated by readily calculated variation in Q_0 (e.g., Bristow and Campbell, 1984), the analyses also considered relations with Q_0.

3. Results

3.1. Comparisons at a single location

A single, arbitrarily selected location, Immokalee, Florida (latitude 26.46°, elev. 11 m) was used to illustrate comparisons for one location. Based on correlations (Table 3), the $SRAD_{NP}$ showed the best agreement with $SRAD_{OB}$, with a correlation of 0.86 ($P < 0.001$). While the overall good agreement is borne out by Fig. 2, values of $SRAD_{NP}$ were consistently lower than for $SRAD_{OB}$, and the means of all daily values were 19.5 $MJd^{-1}m^{-2}$ for $SRAD_{OB}$ and 18.0 $MJd^{-1}m^{-2}$ for $SRAD_{NP}$. Mean values of $SRAD_{WG}$ and $SRAD_{CO}$ were 18.0 and 18.1 $MJd^{-1}m^{-2}$, respectively. Values of $SRAD_{OB}$ included 7 daily values out of 3151 that exceeded 95% of Q_0, but excluding these values had minimal effect on the correlations and means.

3.2. Comparisons over all sets of observed solar radiation data

Across the 295 locations considered, $SRAD_{NP}$ exhibited higher correlation with daily variation in $SRAD_{OB}$, with many correlations of 0.9 or higher (Fig. 1A and Table 3), than $SRAD_{WG}$ or $SRAD_{CO}$. Correlations between $SRAD_{OB}$ and $SRAD_{WG}$ were typically between 0.8 and 0.9 (Fig. 1B and Table 3), while correlations between $SRAD_{OB}$ and $SRAD_{CO}$ were slightly lower (Table 3). Interestingly, correlations of Q_0 with $SRAD_{OB}$ were similar in value to those from the two weather generators (Table 3). Values of root mean square error (RMSE) for prediction of $SRAD_{OB}$ by $SRAD_{NP}$ were generally 2 to 3 $MJm^{-2}d^{-1}$ (Fig. 3a) while RMSE values for $SRAD_{OB}$ and $SRAD_{CO}$ were 4 to 5 $MJm^{-2}d^{-1}$ (Fig. 3b and 3c). These results suggested that $SRAD_{NP}$ data represented day-to-day variation in $SRAD_{OB}$ better than values from WGENR, a conclusion also supported by density plots comparing $SRAD_{OB}$ with $SRAD_{NP}$ and $SRAD_{CO}$ (Fig. 4). Thus, the data assimilation process used with the NASA/POWER data appeared superior to approaches that try to recreate variability in SRAD by considering daily maximum and

minimum temperatures and precipitation patterns (e.g., wet or dry days) as done in weather generators. As an aside, we note that Figure 4B also evidenced a problem with WGENR in that it appeared to produce an excess of values around 6 $MJm^{-2}d^{-1}$ and which coincided with days with rainfall (data not shown).

Comparisons of means indicated that the NASA/POWER values were slightly lower than observed values, with a mean across all stations of 16.2 for $SRAD_{NP}$ vs. 17.4 for $SRAD_{OB}$ (Table 2 and Fig. 1C). This difference was greatest in the summer months and was more pronounced at lower latitudes (Fig. 5A and 5B). However, when expressed on a relative basis (Fig. 5C), the differences were more pronounced in winter months. We note that $SRAD_{WG}$ and $SRAD_{CO}$ also tended to have mean values less that $SRAD_{OB}$ (Table 3 and Fig 1D).

A possible source of discrepancies in $SRAD_{NP}$ values relative to $SRAD_{OB}$ might relate to elevations of individual weather stations as compared to the elevation for NASA/POWER grid cell, which is the average of the topography associated with the 1-degree cell. Since the thickness of the atmosphere decreases with elevation, clear-sky transmittance increases with elevation. The elevation of individual weather stations differed from the average elevation of the associated grid cell from the NASA/POWER dataset by as much as 800 m. Comparisons of correlations of $SRAD_{OB}$ with $SRAD_{NP}$ showed a weak trend related to differences in elevation (Fig. 6A). There was a slight relation between elevation difference and difference in mean values of $SRAD_{OB}$ and $SRAD_{NP}$ (Fig. 6B), but given that the mean elevation difference was only 219 m (Table 3), the net bias due to elevation differences would be less 0.4 $MJm^{-2}d^{-1}$.

3.3 Comparisons of NASA/POWER data with data generated using NOAA COOP data

The availability of the large set of data for paired NASA/POWER grid cells and NOAA COOP locations allowed a more detailed comparison for the continental US. The correlations between $SRAD_{NP}$ and $SRAD_{CO}$ were largest in the western USA and lowest in southeastern regions (Fig. 7A). The overall mean value of $SRAD_{NP}$ was 15.0 $MJd^{-1}m^{-2}$ vs. 15.9 $MJd^{-1}m^{-2}$ for $SRAD_{CO}$ (Table 2), again suggesting that values of $SRAD_{NP}$ are lower than other estimates. Mean values of individual cells diverged by as much as 3 MJ $d^{-1}m^{-2}$ (Fig. 7B). The best agreement for means occurred in California, Oregon and southern to eastern states. In the Rocky Mountain region, mean values of $SRAD_{CO}$ were large relative to $SRAD_{NP}$, while $SRAD_{CO}$ values were low for certain coastal regions.

3.4 Data quality issues

In the initial analyses, several locations showed correlations between $SRAD_{OB}$ and $SRAD_{NP}$ that were less than 0.8, suggesting possible problems with the observed values. Data from Semmes, AL (Supplement Figure 1A) are indicative. For this location, $SRAD_{OB}$ showed a steady decline from 1995 through 2004, so the location was excluded from the analyses. Plots comparing $SRAD_{OB}$ to Q_0 proved difficult to evaluate visually, so data were re-plotted as the ratio of $SRAD_{OB}$ to Q_0 (Supplement Figure 2). Although not analyzed quantitatively in this study, it appeared that the maximum value of this ratio, which would correspond to very dry, clear-sky conditions, is approximately 0.8 (as assumed in WGENR), and this value is marked by a reference line on the graphs.

For seven weather stations, low correlations of $SRAD_{OB}$ with $SRAD_{NP}$ were not associated with errors visible in time series plots of $SRAD_{OB}$, and the data showed mean values and patterns of variation similar to neighboring sites. Discrepancies in observation time are problematic in reporting of daily temperature data, so we tested whether correlations of $SRAD_{OB}$ with $SRAD_{NP}$ improved if it was assumed that the reported date of observation was a day later than the actual date. This assumption improved the correlations for these stations from a mean value 0.62 to a mean of 0.92 (Table 4), implying that the weather stations were reporting solar data with a one day offset (delay).

4. Discussion

The comparisons of the different sources of solar radiation data suggested that data from NASA/POWER reproduced variability at a daily time scale better than either set of generated values (Table 3 and Fig. 1). Comparisons of mean values (Tables 2 and 3 and Fig. 1), however, indicated that $SRAD_{NP}$ values were often lower than observed values or values derived from WGENR. Figures 1, 5B, 5C and 7 suggest that there is a regional component to the bias, although the variation in Fig. 7 also may result from bias in $SRAD_{CO}$ values. The subsequent version of the NASA/POWER solar irradiance data set showed reductions in this bias that are directly attributable to a reduction in the background aerosol specification. The new aerosol climatology based upon an upgraded NCAR Model of Atmospheric Transport and Chemistry (MATCH, Rasch *et al*., 1997) that resulted in the initial background aerosol optical depths being reduced from about 10% in the southeastern US to over 50% in the northwestern US. Although the background aerosol does not determine the final aerosol optical depth for a given space and time (as noted above), it governs the optimization of the match between the inferred and computed TOA albedos by constraining the surface albedo and in this case, it did lead systematically to an increase of solar irradiance values in the continental US. Other potential sources of the bias were examined. Differences in elevations of station locations and of mean elevations of NASA/POWER grid cells appeared at best to explain only a small portion of the bias (Fig. 5B), and elevation differences also

appeared to have little influence on correlations between $SRAD_{OB}$ and $SRAD_{NP}$ (Fig. 5A). Another possible explanation for the bias concerns locations of the weather stations. Stations associated with airports or agricultural research centers may have been located in open areas with a clear field of view and low probability of cloud cover, while NASA/POWER grid cells may have included mountainous areas that experienced lower daily radiation due to greater cloud cover. Reflection from clouds can increase irradiance on a scale of minutes, but such effects tend to be cancelled out when the sun is obscured by clouds (Pfister et al., 2003) and thus seem unlikely to contribute to the bias. The occurrence of sporadic excessively high values of $SRAD_{OB}$ also might have biased the mean value of $SRAD_{OB}$, or it may evidence a tendency of some automated instruments to overestimate SRAD. As a partial test for this problem, the mean of $SRAD_{OB}$ was re-calculated after limiting all values of SRAD to a maximum of 0.8 of Q_0. The resulting mean was 17.25 $MJm^{-2}d^{-1}$ as compared to 17.32 $MJm^{-2}d^{-1}$ for the dataset as used in the rest of the paper, where values greater than Q_0 were simply excluded. Thus, we suggest that errors in the prescription of the background aerosol for the GEWEX SRB solar irradiance v2.81 are the largest contributor to the noted systematic bias.

The NASA/POWER data accurately reproduced the variability in solar radiation data, did not show major variation related to effects of elevation, and are readily available via the Internet for a time span from 1984 onward. They thus show excellent potential as a source of solar radiation data for diverse applications. However, the differences in mean values of $SRAD_{OB}$ and $SRAD_{NP}$ were a concern, but appear to be at least partially addressed in the GEWEX SRB solar v3.0 that is now available. If analysis of the new version largely explains the differences noted in this paper, the NASA/POWER data would appear to be superior to data from weather generators.

A previous paper comparing NASA/POWER daily temperature data to COOP data found larger differences between these two sources (White et al., 2008). Several hypotheses can be forwarded to explain why the NASA/POWER solar data may show greater consistency than the temperature data. The first and foremost is that the estimate of solar radiation is far more dependent upon the structure and variability of the actual cloud fields than on accuracy in the underlying meteorological fields from the atmospheric assimilation data sets. Those cloud fields are directly observed by the satellite measurements. In fact, only the total water vapor profile is required for the solar radiation calculation and even 10-30% errors in the water vapor profile amount (which are probably correlated with surface temperature errors in the assimilation) correspond to solar radiation errors $< 1\%$. Time of observation bias should have been less of a problem since observed solar radiation data presumably were based on a midnight to midnight integration period. Instrument siting errors may have less effect on solar radiation data than on temperature data. We emphasize that the algorithms used to create the datasets are subject to periodic review and improvement as noted by the effect of the improved aerosol inputs described above.

Additional improvements in the satellite calibration, sampling size and solar algorithm is anticipated and in 2012, it is planned to reduce the grid cell size to 0.5°.

In working with datasets from automatic weather stations, accessed via the Internet, numerous problems with data quality were encountered. Together, these reinforce concerns over the management of daily weather data (e.g., Davy and Pielke, 2005; Holder et al., 2006; Pielke et al., 2007). Various data checking procedures exist, but they do not appear to be used by all data providers. The variable quality control provides another argument in favor of using a single, well-documented data source such as offered by NASA/POWER.

5. Conclusions

Considering the constraints inherent with its coarse grid size of 1° x 1° of latitude and longitude, the NASA/POWER solar radiation data compare favorably with data reported from automatic weather stations. In terms of representing historic variability on time scales of a few days, they appeared superior to values estimated using the WGENR weather generator. However, the means of the $SRAD_{NP}$ data were often 1 to 2 $MJm^{-2}d^{-1}$ lower than $SRAD_{OB}$, and this discrepancy merits further investigation.

NASA/POWER data are available for over 25 years with global coverage and are continuously being updated and improved. They represent a valuable source of solar radiation data for research concerned with regional to global geographic scales.

Acknowledgements

We acknowledge the valuable assistance of Belinda Wernau and Samuel Wright in processing of the station data. Portions of this work were funded by a grant from the NASA Langley Research Center under Contract NN07AA00C with Science Systems and Applications, Inc. and a Special Research Grant from the US Department of Agriculture –National Institute for Food and Agriculture (NIFA).

References

Bai, J., Chen, X., Dobermann, A., Yang, H., Cassman, K. G., Zhang, F., 2010. Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. Agron. J. 102, 9 - 16.

Bristow, K.L., Campbell, G.S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agric. For. Meteorol. 31, 159-166.

Cooter, E.J., Dhakwa, G.B., 1995. A solar radiation model for use in biological applications in the South and Southeastern USA. Agric. For. Meteorol. 78, 31-51.

Davey, C.A., Pielke, Sr. R.A., 2005. Microclimate exposures of surface-based weather stations. Bull. Amer. Meteorol. Soc. 86, 497-504.

de Wit, A.M.R., Boogaard, H.L., van Diepen.C.A., 2005. Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. Agric. For. Meteorol. 135, 156-168.

Durrenberger, R.W., Brazel, A.J., 1976. Need for a better solar radiation data base. Science 193, 1154-1155.

Garcia y Garcia, A., Hoogenboom, G., 2005. Evaluation of an improved daily solar radiation generator for the southeastern USA. Climate Res. 29, 91-102.

Garcia y Garcia. A., Guerra, L.C., Hoogenboom, G., 2008. Impact of generated solar radiation on simulated crop growth and yield. Ecol. Model. 210, 312-326.

Gupta, S. K., Stackhouse, P.W., Jr., Cox, S.J., Mikovitz, J. C., Zhang, T., 2006: 22-Year Surface Radiation Budget Data Set. *GEWEX News*, **V 16**, No. 4, November, p. 12-13.

Kratz, D. P., Gupta, S. K., Wilber, A. C., Sothcott, V. E., 2010. Validation of CERES Edition 2B Surface-Only Flux Algorithms. J. Appl. Meteorol. Clim., 49, 164-180.

Hodges T, French, V., LeDuc, S.K., 1985. Estimating solar radiation for plant simulation models. AgRISTARS Tech. Rep. JSC-20239; YM-15-00403, 21 pp.

Holder, C., Boyles, R., Syed, A., Niyogi, D., Raman, S., 2006. Comparison of collocated automated (NCECONet) and manual (COOP) climate observations in North Carolina. J. Atmos. Ocean. Tech. 23, 671-682.

Hoogenboom, G., Jones, J. W., Wilkens, P. W., Porter, C. H., Batchelor, W. D., Hunt, L. A., Boote, K. J., Singh, U., Uryasev, O., Bowen, W. T., Gijsman, A., Toit, A. du, White, J. W., and Tsuji, G. Y. 2010. Decision Support System for Agrotechnology Transfer Version 4.5. Honolulu, HI: University of Hawaii. CD-ROM software.

Hunt, L.A., White, J.W., Hoogenboom, G., 2001. Agronomic data: advances in documentation and protocols for exchange and use. Agric. Syst. 70, 477-492.

Hunt, L. A., Hoogenboom, G., Jones, J. W., and White, J. W., 2006. ICASA Version 1.0 Data Standards for Agricultural Research and Decision Support. http://www.icasa.net/standards/index.html. (Verified 12 Sept., 2010)

Janis, M.J., Hubbard, K.G., Redmond, K.T., 2004. Station density strategy for monitoring long-term climatic change in the contiguous United States. J. Climate 17, 151-162 .

LI-COR. 2005. LI-COR terrestrial radiation sensors: Instruction manual. LI-COR, Linclon, NE.

Liu, D.L., Scott, B.J., 2001. Estimation of solar radiation in Australia from rainfall and temperature observations. Agric. For. Meteorol. 106, 41-59.

NOAA. 2006. National Weather Service's Cooperative Observer Program. www.nws.noaa.gov/om/coop/ (Verified 12 Sept., 2010).

Pfister, G., McKenzie, R.L., Liley, J.B., Thomas, A., Forgan, B.W., Long, C.N. 2003. Cloud coverage based on all-sky imaging and its impact on surface solar irradiance. J. Appl. Meteorol. 42, 1421-1434.

Pielke Sr., R.A., Davey, C., Niyogi, D., Fall, S., Steinweg-Woods, J., Hubbard, K., Lin, X., Cai, M., Lim, Y.-K., Li, H., Nielsen-Gammon, J., Gallo, K., Hale, R., Mahmood, R., Foster, S., McNider, R.T., Blanken, P., 2007. Unresolved issues with the assessment of multi-decadal global land surface temperature trends. J. Geophys. Res. 112, D24S08.

Pinker, R., Laszlo, T. I., 1992. Modeling surface solar irradiance for satellite applications on a global ccale, J. Appl. Met., 31, 194-211.

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resources Res. 17, 182-190.

Richardson, C.W., Wright. D. A., 1984. WGEN: A model for generating daily weather variables. USDA, Agricultural Research Service. Pub. No. ARS-8. 83 pp.

Stackhouse, P.W., Jr., 2010a. Prediction Of Worldwide Energy Resources. http://power.larc.nasa.gov/ (Verified 12 September 2010).

Stackhouse, P.W., Jr., 2010b. CERES FLASHFlux. http://flashflux.larc.nasa.gov/ (Verified 12 September 2010).

Stoffel, T., Redo, I., Myers, D., Renne, D., Wilcox, S., Treadwell, J., 2000. Current issues in terrestrial solar radiation instrumentation for energy, climate and space applications. Metrologia 37, 399–402.

White, J.W., Hoogenboom, G., Stackhouse, P.W., Jr., Hoell, J.M., 2008. Evaluation of daily temperature data for the continental US modeled from satellite data. Agric. For. Met. 148, 1574-1584.

Younes, S., Claywell, R., Muneer, T., 2005. Quality control of solar radiation data: present status and proposed new approaches: measurement and modelling of solar radiation and daylight-challenges for the 21st century. Energy 30, 1533-1549.

Captions for Figures

Figure 1. Comparisons of correlations and differences between mean values of daily solar radiation data from different sources. A. Correlation between $SRAD_{OB}$ and $SRAD_{NP}$. B. Correlation between $SRAD_{OB}$ and $SRAD_{WG}$. C. Difference between $SRAD_{OB}$ and $SRAD_{NP}$. D. Difference between $SRAD_{OB}$ and $SRAD_{WG}$.

Figure 2. Comparisons of solar radiation values (in units of $MJm^{-2}d^{-1}$) from the various sources of SRAD data for Immokalee, Florida from 1 Jan. 1998 to 31 Dec. 2004. The diagonal lines represent a 1:1 relation. A. $SRAD_{OB}$ vs $SRAD_{NP}$. B. $SRAD_{OB}$ vs $SRAD_{WG}$. C. $SRAD_{OB}$ vs $SRAD_{CO}$. D. $SRAD_{CO}$ vs $SRAD_{NP}$. All values of r^2 are significant at the $P < 0.001$ level.

Figure 3. Root mean square errors (RMSE) for prediction of observed solar radiation from the other sources. A. $SRAD_{OB}$ vs $SRAD_{NP}$. B. $SRAD_{OB}$ vs $SRAD_{WG}$. C. $SRAD_{OB}$ vs $SRAD_{CO}$. D. $SRAD_{OB}$ vs Q_0.

Figure 4. Density plots comparing values of solar radiation ($MJ\ m^{-2}d^{-1}$) for 295 stations in the continental US. A. Weather station vs. NASA/POWER. B. Weather station vs. WGENR-generated. Count ranges are 1 = 1 to 10 paired values; 2 = 11 to 100; 3 = 101 to 1000; 4 = 1001 to 10,000.

Figure 5. Annual variation in solar radiation based on seven-day averages. A. Daily means across all locations for $SRAD_{OB}$, $SRAD_{NP}$ and $SRAD_{WG}$. B. Mean difference between $SRAD_{OB}$ and $SRAD_{NP}$ for five latitude bands from less than 30°N to greater than 45°N. C. Relative error $[(SRAD_{NP} - SRAD_{OB})/SRAD_{OB}]$ for the five latitude bands.

Figure 6. Relation between difference in elevation of weather station and of the corresponding grid cell for the NASA/POWER dataset and two indicators of reliability for all stations. A. Correlation between $SRAD_{OB}$ and $SRAD_{NP}$. B. Difference between mean of $SRAD_{OB}$ and mean of $SRAD_{NP}$.

Figure 7. Comparison of solar radiation data from NASA/POWER and estimated with WGEN using daily weather data from NOAA COOP stations. A. Correlation between the two sources for each grid cell. All correlations are significant at the $P < .001$ level. B. Mean difference between the NOAA COOP and NASA/POWER sources.

Supplement Fig. 1. Variation in solar radiation for five stations where the data showed patterns suggesting problems in instrumentation. A. Semmes, AL. B. Avondale, CO. C. Novelty, MO. D. Calipatria/Mulberry, CA. E. Manteca, CA.

Supplement Fig. 2. Variation in the ratio of observed solar radiation to extraterrestrial (Q_0) for five stations where the data showed patterns suggesting problems in instrumentation. The horizontal reference line at 0.8 indicates an approximate upper limit for any location. A. Semmes, AL. B. Avondale, CO. C. Novelty, MO. D. Calipatria/Mulberry, CA. E. Manteca, CA.

Table 1.

Daily meteorological variables available on a global 1° grid through the NASA/POWER project.

Variable	Source	Time span	Availability from present date
Daily Maximum and minimum temperatures, Daily averaged temperature	Goddard Earth Observing System (GEOS) assimilation model version 4 Goddard Earth Observing System (GEOS) assimilation model version 5	January, 1983 to December 2007 January 2008 to present	online ≤ 1 week[b]
Precipitation	Satellite & ground observations from the Global Precipitation Climatology P project (GPCP)	January, 1997 to present[a]	≤ 2 months
Solar radiation	Satellite observations GEWEX SRB v3.0* FLASHFlux	July 1983 to December 2007 January 2008 to Present	online ≤ 1 week [b]
Dewpoint temperature	Goddard Earth Observing System (GEOS) assimilation model version 4 Goddard Earth Observing System (GEOS) assimilation model version 5	January, 1983 to December 2007 January 2008 to present	online ≤ 1 week[b]

[a] The most current GPCP file is August 2009.

[b] Data files are updated daily.

* v3.0 came available after the analysis presented here was completed; v2.81 corresponding to White et al., (2008) was used for this study.

Table 2.

Mean, minimum and maximum values of solar radiation for the four solar data sources. The NASA/POWER and COOP data are for 855 locations on a 1° latitude and longitude grid covering the continental US and representing a time series from 1983 through 2005. Elevations correspond to mean values of grid cells for the NASA/POWER dataset and to reported values for COOP stations.

Data source	Mean	Minimum	Maximum
Sites with automated stations (N = 295)			
Automated stations	17.4[a]	0.2[a]	43.0[a]
NASA/POWER	16.2	0.3	34.1
Generated based on station data	16.5	1.3	33.2
Generated based on COOP data	16.8	1.3	33.2
Difference between NASA/POWER and Automated Stations	-1.2	-30.4	29.6
Elevation difference (m): NASA/POWER –Automated Stations	+219	-847	368
Entire US based on COOP stations (N = 855)			
NASA/POWER	15.0	0.1	34.2
Generated based on COOP data	15.9	1.3	33.2
Difference between NASA/POWER and COOP-based data	-0.9	-31.1	26.1
Elevation difference (m): NASA/POWER - COOP	85	-1580	1270

[a] Minimum and maximum values of SRAD from automated stations reflect minimum and maximum values imposed in processing data.