**SANDIA REPORT**
SAND2015-8023
Unlimited Release
Printed September 2015

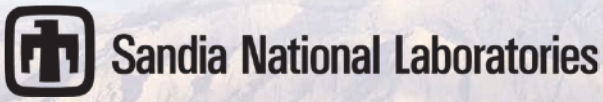
# Analysis of Global Horizontal Irradiance in Version 3 of the National Solar Radiation Database

Clifford W. Hansen, Curtis E. Martin, Nathan Guay

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.





Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.



Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2015-8023
Unlimited Release
Printed September 2015

# Analysis of Global Horizontal Irradiance in Version 3 of the National Solar Radiation Database

Clifford W. Hansen, Curtis E. Martin, Nathan Guay
Photovoltaic and Distributed Systems Department, Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-1033

**Abstract**

We report an analysis that compares global horizontal irradiance (GHI) estimates from version 3 of the National Solar Radiation Database (NSRDB v3) with surface measurements of GHI at a wide variety of locations over the period spanning from 2005 to 2012. The NSRDB v3 estimate of GHI are derived from the Physical Solar Model (PSM) which employs physics-based models to estimate GHI from measurements of reflected visible and infrared irradiance collected by Geostationary Operational Environment Satellites (GOES) and several other data sources. Because the ground measurements themselves are uncertain our analysis does not establish the absolute accuracy for PSM GHI. However by examining the comparison for trends and for consistency across a large number of sites, we may establish a level of confidence in PSM GHI and identify conditions which indicate opportunities to improve PSM.

We focus our evaluation on annual and monthly insolation because these quantities directly relate to prediction of energy production from solar power systems. We find that generally, PSM GHI exhibits a bias towards overestimating insolation, on the order of 5% when all sky conditions are considered, and somewhat less (~3%) when only clear sky conditions are considered. The biases persist across multiple years and are evident at many locations. In our opinion the bias originates with PSM and we view as less credible that the bias stems from calibration drift or soiling of ground instruments.

We observe that PSM GHI may significantly underestimate monthly insolation in locations subject to broad snow cover. We found examples of days where PSM GHI apparently misidentified snow cover as clouds, resulting in significant underestimates of GHI during these days and hence leading to substantial understatement of monthly insolation. Analysis of PSM GHI in adjacent pixels shows that the level of agreement between PSM GHI and ground data can vary substantially over distances on the order of 2 km. We conclude that the variance most likely originates from dramatic contrasts in the ground's appearance over these distances.

## ACKNOWLEDGMENTS

We obtained measured irradiance and other meteorological data from the following sources:

- The National Renewable Energy Laboratory's Solar Resource Radiation Laboratory.
- First Solar, Inc.
- The University of Arizona's Arizona Meteorological Network (AZMET).
- The University of Georgia's Georgia Automated Environmental Monitoring Network.
- The University of Oregon's Solar Radiation Monitoring Laboratory.
- The National Oceanographic and Atmospheric Agency's Earth System Research Laboratory's Surface Radiation (SURFRAD) Network.

# CONTENTS

# FIGURES



# TABLES

# NOMENCLATURE

| | |
|---|---|
| DHI | diffuse horizontal irradiance |
| DOE | Department of Energy |
| DNI | direct normal irradiance |
| GHI | global horizontal irradiance |
| GOES | Geostationary Operational Environmental Satellite |
| GSIP | Global Solar Insolation Project |
| NREL | National Renewable Energy Laboratory |
| NSRDB | National Solar Radiation Database |
| POA | plane-of-array |
| PSM | Physical Solar Model |
| PV | photovoltaic |
| SNL | Sandia National Laboratories |

# 1. INTRODUCTION

The National Solar Radiation Database (NSRDB) comprises several iterations each of which contains serially complete, gridded data for GHI and other quantities over the continental United States. The NSRDB v1 [1] contains hourly values of GHI from 1961 to 1990 at 239 locations; roughly 7% of these data are actual GHI measurements, with the remainder being estimated using the METSTAT model. The NSRDB v2 [2], [3] contains a 10km gridded, hourly data set of GHI estimated from GOES imagery from 1991 to 2010 using algorithms developed by R. Perez and others [4].

The most recent NSRDB v3 [5], [6] contains a gridded 0.038 degree latitude by 0.038 degree longitude (roughly 4km square, or 16 $km^2$), 30-minute data set of GHI. The GHI in the NSRDB v3 are estimated by the Physical Solar Model (PSM) [5] from GOES imagery and other data sources (e.g., aerosol optical depth) The PSM was developed at NREL by combining the Global Solar Insolation Project (GSIP) model [7] with NREL-developed algorithms [8], [9].

We report an analysis that compares NSRDB GHI, computed in 2014 using data spanning from 2005 to 2012, with surface measurements of GHI at a wide variety of locations over the same period. We focus our comparison primarily on quantities relevant to predicting energy production from solar power systems, i.e., annual and monthly insolation.

Strictly speaking, we cannot determine the absolute accuracy of NSRDB GHI by this comparison because the ground measurements themselves are uncertain. However, by comparing NSRDB GHI with ground measurements at a large number of sites, and by examining the results for consistency, we may establish a level of confidence in NSRDB GHI data and may identify opportunities for PSM improvement.

Our report is organized as follows:

- Section 2 describes the data used in our analysis;
- Section 3 describes our methodology for comparing NSRDB GHI to ground measurements;
- Section 4 presents the results of the analysis;
- Section 5 summarizes our conclusions and recommendations.

# 2. DATA AND DATA QUALITY

This section describes the two main categories of data used in this study. First, we briefly describe the satellite-based PSM data product, which is the focus of the analysis. We then summarize the various sources of ground measurement data that we compared to the NSRDB GHI estimates.

## 2.1. PSM GHI Data

The PSM data product is provided by NREL. This data product is derived by combining visible and infrared imagery from the Geostationary Operational Environmental Satellite (GOES) system with data from the National Snow and Ice Data Center (NSIDC) and results from weather and clear-sky models[1]. In routine operations, two GOES satellites provide imagery spanning the continental United States every 15 minutes. The native spatial resolution of GOES imagery is approximately 1 km for visible imagery and 4 km for infrared imagery. PSM uses half-hourly visible and infrared GOES images to produce its 30-minute data set. The imaging instruments on the GOES satellites are scanning instruments which require several minutes to collect a complete image. Depending on the size of the area to be imaged, these raster scan times are typically 7 or 15 minutes. As a result, there is an unknown offset between the timestamp applied to the image (typically at the beginning of the raster scan) and the actual time at which the instrument sampled any given location (pixel) in the image.

Table 1 summarizes the data sets included in the NSRDB v3. The data sets are provided as co-registered three-dimensional arrays with a spatial resolution of approximately 4 $km^2$ (0.38 degree latitude by 0.38 degree longitude) and a temporal resolution of 30 minutes. Time and location (latitude, longitude, elevation and time zone) are provided for each sample. Although there is a delay between the time stamp applied to each satellite image and the actual sample time for each pixel, we treat the NSRDB data as "snapshots," all taken at the instant of the time stamp, regardless of the pixel location.

For each ground measurement site identified below, we obtained NSRDB data from a 5x5 pixel array (an area approximately 20 km x 20 km) surrounding the site for the period 2005-2012. We focused this study exclusively on the NSRDB GHI data set because this quantity is of primary interest for solar energy modeling, and due to the lack of sufficient ground measurements to evaluate the other data sets.

## 2.2. Ground Data

We compare NSRDB GHI data to GHI measurements collected by seven different agencies at a total of 78 different ground stations (see Table 2). These data are described in greater detail in this section. At every site we separated daytime from nighttime values by computing the solar elevation and considering only data with positive solar elevation angles (i.e., with the sun at or above the geometric horizon).

[1] https://nsrdb.nrel.gov/current-version

In almost all cases, each data set included some time periods during which the irradiance data were missing, or during which data quality flags indicated that the data were unreliable. We omitted these time periods. In addition, we omitted data where irradiance values were unreasonably high or low, where data were inconsistent, and by visual inspection (superimposing a plot of the ground time series data on a plot of the corresponding NSRDB GHI data) the ground data appeared clearly anomalous. Where there were clear and systematic differences between the NSRDB GHI and ground data (on the order of 10% or more), we compared ground data between earlier or later time periods, with clear sky irradiance models, and, where possible, with other nearby sites, to judge whether the ground data were reliable.

**Table 1. NSRDB Data Elements**

| **Data Set** | **Units** |
|---|---|
| Global Horizontal Irradiance (GHI) | $W/m^2$ |
| Diffuse Horizontal Irradiance (DHI) | $W/m^2$ |
| Direct Normal Irradiance (DNI) | $W/m^2$ |
| Clear Sky GHI | $W/m^2$ |
| Clear Sky DHI | $W/m^2$ |
| Clear Sky DNI | $W/m^2$ |
| Solar Zenith Angle | degrees |
| Surface Air Temperature | Kelvin |
| Surface Pressure | mbar |
| Surface Relative Humidity | percent |
| Total Precipitable Water | cm |
| Dew Point | Kelvin |
| Aerosol Optical Depth | unitless |
| Cloud Type | integer code (14 values) |
| Wind Direction | degrees |
| Wind Speed | m/s |

**Table 2. Ground data sources**

| Data Source | Number of sites | Time Span | Frequency |
|---|---|---|---|
| AZMET | 23 | 2005-2012 | Hourly |
| First Solar | 4 | 2010-2012 | 1 minute |
| NREL Hawaii | 17 | 2010-2011 | 1 second |
| NREL SRRL | 1 | 2005-2012 | 1 minute |
| SNL PSEL | 1 | 2005-2012 | 1 minute |
| SURFRAD | 6 | 2005-2012 | 1 or 3 minutes |
| Univ. of Oregon | 8 | 2005-2012 | 5 or 15 minutes |
| Univ. of Georgia | 35 | 2005-2012 | 15 minutes |

We selected ground data to represent a wide range of climate zones as distinguished by the Köppen-Geiger climate classification [10]. Figure 1 displays the ground stations superimposed on a map of climate classes. Table 3 summarizes the number of site-years of ground data within each climate class.

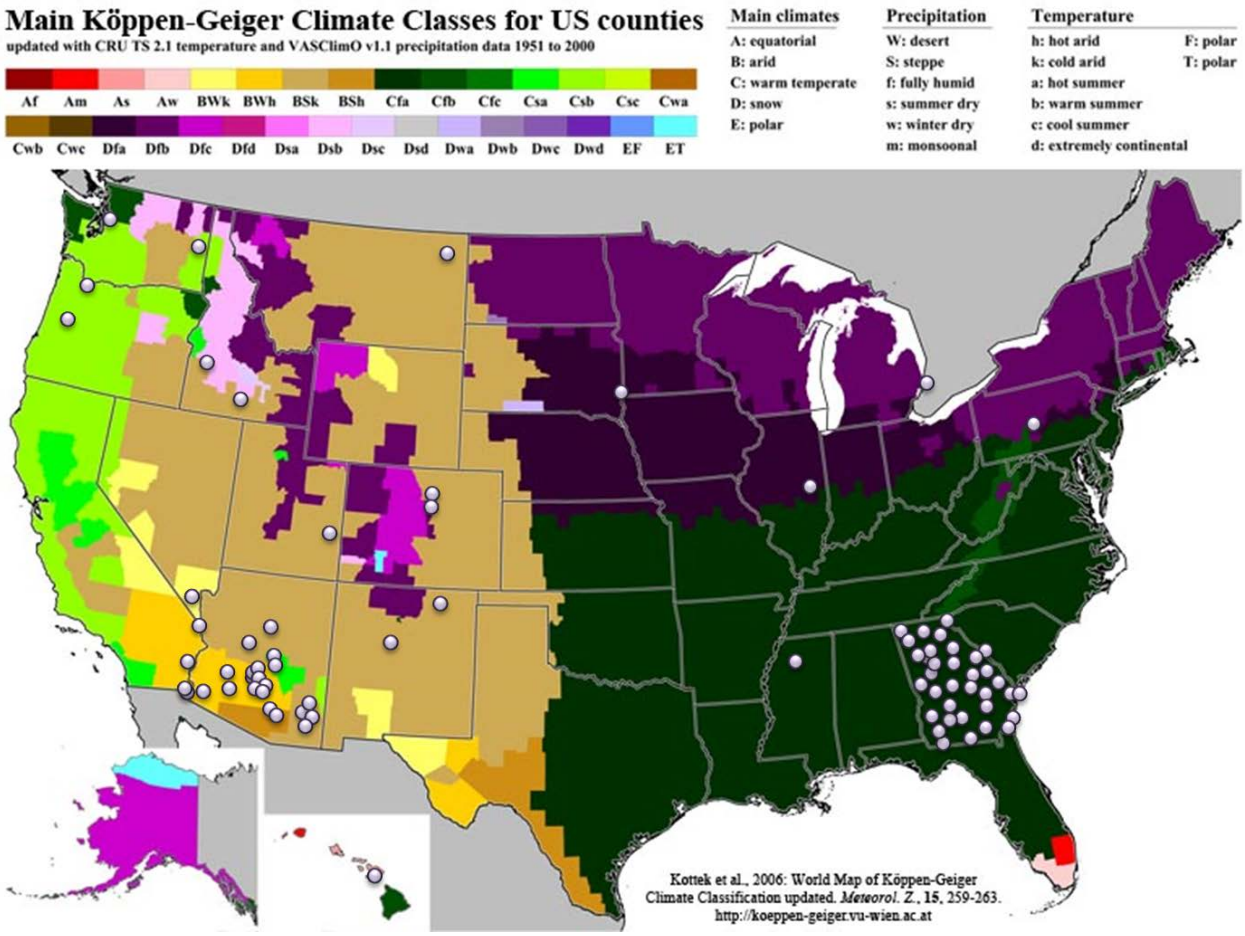



**Figure 1. Ground data sites with map of climate zones.**

**Table 3. Climate zones represented in ground data**

| Climate | Number of site-years | Data source |
|---|---|---|
| Equatorial/dry summer (As) | 17 [1] | NREL Hawaii |
| Arid/steppe/hot arid (BSh) | 28 | AZMET |
| Arid/steppe/cold arid (BSk) | 86 | AZMET, First Solar, PSEL, Univ. of Oregon, SURFRAD |
| Arid/desert/hot arid (BWh) | 83 | AZMET, First Solar |
| Arid/desert/cold arid (BWk) | 2 | First Solar |
| Temperate/humid/hot (Cfa) | 242 | Univ. of Georgia, SURFRAD |
| Temperate/dry hot summer (Csa) | 8 | AZMET |
| Temperate/dry warm summer (Csb) | 39 | AZMET (1 site), Univ. Of Oregon |
| Snow/humid/hot (Dfa) | 16 | SURFRAD |
| Snow/humid/warm (Dfb) | 19 | First Solar, NREL SRRC, SURFRAD |

[1] 17 sites with data from the same year.

### 2.2.1. AZMET

We used hourly global horizontal irradiance (GHI) measurements, reported in watts per square meter, collected at 23 sites of the Arizona Meteorological Network (AZMET) [11] using LI-COR pyranometers with annual calibration of the instruments but without intra-year frequent cleaning.. AZMET data are reported as end-of-hour averages; we used data collected from 2005 through 2012 for this analysis, as summarized in Table 4.

We examined a larger number of sites and years than are indicated in Table 4 to select the data we used. We first discarded sites with data that had missing values or unreasonably large or small values at any time during the period of interest. AZMET also reports daily insolation totals. We also discarded sites with inconsistencies between annual insolation totaled from the hourly and the daily data.

**Table 4. AZMET data sites**

| Site | Climate | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---|---|---|---|---|---|---|---|---|---|
| Aguila | BSk | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Bonita | BSk | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Bowie | BSk | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Coolidge | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Desert Ridge | BSh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Flagstaff | Csb | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| Harquahala | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Kansas Settlement | BSk | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Marana | BSh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| Maricopa | BWh | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| Mesa | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mohave2 | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Paloma | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Payson | Csa | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Phoenix Encanto | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Phoenix Greenway | BSh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Prescott | BSk | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Queen Creek | BWh | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Roll | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Safford | BSk | ✓ | ✓ | | ✓ | | ✓ | ✓ | ✓ |
| Tucson | BSh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Yuma Mesa | BWh | ✓ | ✓ | ✓ | ✓ | | | | |
| Yuma Valley | BWh | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

### 2.2.2. First Solar Irradiance Data

We used one minute GHI measurements collected by First Solar, Inc. using CMP-11 secondary standard pyranometers at four locations in North America [12]; we used data as summarized in Table 5.

**Table 5. First Solar data sites**

| Site Name | Location | Climate | Time Span |
|---|---|---|---|
| Plant 1 | California | BWh | 2009-2013 |
| Plant 2 | New Mexico | BSk | 2012 |
| Plant 5 | Ontario, Canada | Dfb | 2011-2012 |
| Plant 6 | Nevada | BWk | 2011-2012 |

### *2.2.3. NREL Hawaii Irradiance Data*

As part of the Solar Resource & Meteorological Assessment Project, the National Renewable Energy Laboratory (NREL) has recorded GHI measurements using 17 LI-COR LI-200 pyranometers, all of which are installed within an area of one square kilometer at the southwest corner on the island of Oahu (near Kalaeloa Airport) [13]. GHI measurements were collected every second from 5:00 to 20:00 HST from March 2010 to October 2011. Samples marked by NREL as unreliable were omitted from the analysis. Otherwise, no anomalous samples were identified. These data are used only for assessing the variability of NSRDB GHI within a single pixel and between neighboring pixels.

### *2.2.4. NREL SRRL Irradiance Data*

NREL has recorded solar irradiance measurements using multiple instruments at the Solar Radiation Research Laboratory (SRRL) near Denver, Colorado for many years. The SRRL site is situated in a Dfb climate. For this study, we used GHI data collected every minute of 2005 through 2012 using a Kipp & Zonen CM22 precision pyranometer. Samples marked by NREL as unreliable were omitted from the analysis. Otherwise, no anomalous samples were identified.

### *2.2.5. SNL PSEL Irradiance Data*

Sandia National Laboratories (SNL) maintains a solar irradiance measurement station at its Photovoltaic Systems Evaluation Laboratory (PSEL) in Albuquerque, NM. This site is located in a BSk climate. For the overall time period of this study (2005-2012) the finest time resolution available for data from this site is nominally one minute. However, the temporal sampling is somewhat irregular, so every fifth or sixth sample is missing. In order to facilitate detection of clear sky conditions, we used linear interpolation to fill in these missing samples. This interpolation process does not significantly alter the comparisons between ground and satellite measurements at the 30-minute time scale.

### *2.2.6. SURFRAD Irradiance Data*

SURFRAD [14] is a network of ground-based solar measurement stations operated by the National Oceanic and Atmospheric Administration (NOAA) which has been used for multiple

studies of solar resources (see, for example, [8]). The stations are distributed across North America in diverse climatic regions. SURFRAD GHI measurements are reported every three minutes. They are taken with Spectrolab SR-75 pyranometers [15], which are calibrated on a yearly basis. We used data spanning the time period of 2009-2012 for the purposes of this study. Samples marked by NOAA as unreliable were omitted from the analysis. Otherwise, no anomalous samples were identified.

**Table 6. SURFRAD data sites**

| Site Name | Climate | Time Span |
|---|---|---|
| Bondville, IL | Dfa | 2009-2012 |
| Sioux Falls, SD | Dfa | 2009-2012 |
| Boulder, CO | BSk | 2009-2012 |
| Fort Peck, NV | BSk | 2009-2012 |
| Goodwin Creek, MS | Cfa | 2009-2012 |
| Penn St. Univ., PA | Dfb | 2009-2012 |

### *2.2.7. University of Georgia Irradiance Data*

The Georgia Automated Environmental Monitoring Network is a collection of 81 ground-based meteorological stations distributed across Georgia, operated by the University of Georgia for agricultural and environmental applications. We obtained data from 33 of these sites for this study. GHI data is provided as the average of one second measurements over the previous fifteen minutes from LI-200X pyranometers [16]. Calibration is performed by comparison with similar, tertiary standard instruments. Maintenance is reported as occurring at about eight week intervals [16]. Table 7 summarizes the years of data at each site used in the study. Years missing a significant number of sample times were removed as were years during which annual insolation disagreed when determined from separately-reported 15-minute and daily insolation values. All University of Georgia sites are located in a Dfa climate.

**Table 7. University of Georgia data sites**

| Site | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---|---|---|---|---|---|---|---|---|
| Albany | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Atlanta | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| Attapulgus | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Baxley | | | | | ✓ | ✓ | ✓ | ✓ |
| Williamson | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Brunswick | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Byron | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Clarks Hill SC | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Cordele | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Dallas | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Dearing | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Dublin | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Dunwoody | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Eatonton | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Ellijay | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ |
| Rome | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ |
| Gainesville | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Homerville | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Howard | | | | | | ✓ | ✓ | ✓ |
| Jonesboro | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| LaFayette | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Midville | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Newton | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Savannah | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Shellman | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Statesboro | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Tennille | | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Tifton | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Tiger | | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| Valdosta | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Vidalia | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Watkinsville-USDA | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| Woodbine | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

Much of the ground data in the University of Georgia appears inconsistent with the NSRDB GHI as can be seen by examining time series plots. Figure 2 compares ground data and NSRDB GHI at the Attapulgus site. The comparison is typical of many sites in the Georgia data, where the NSRDB GHI remains fairly consistent from year to year but ground GHI values can vary significantly between years. As a result of these common inconsistencies, we consider the Georgia data to be a less reliable reference and complete the bulk of our analysis without it.

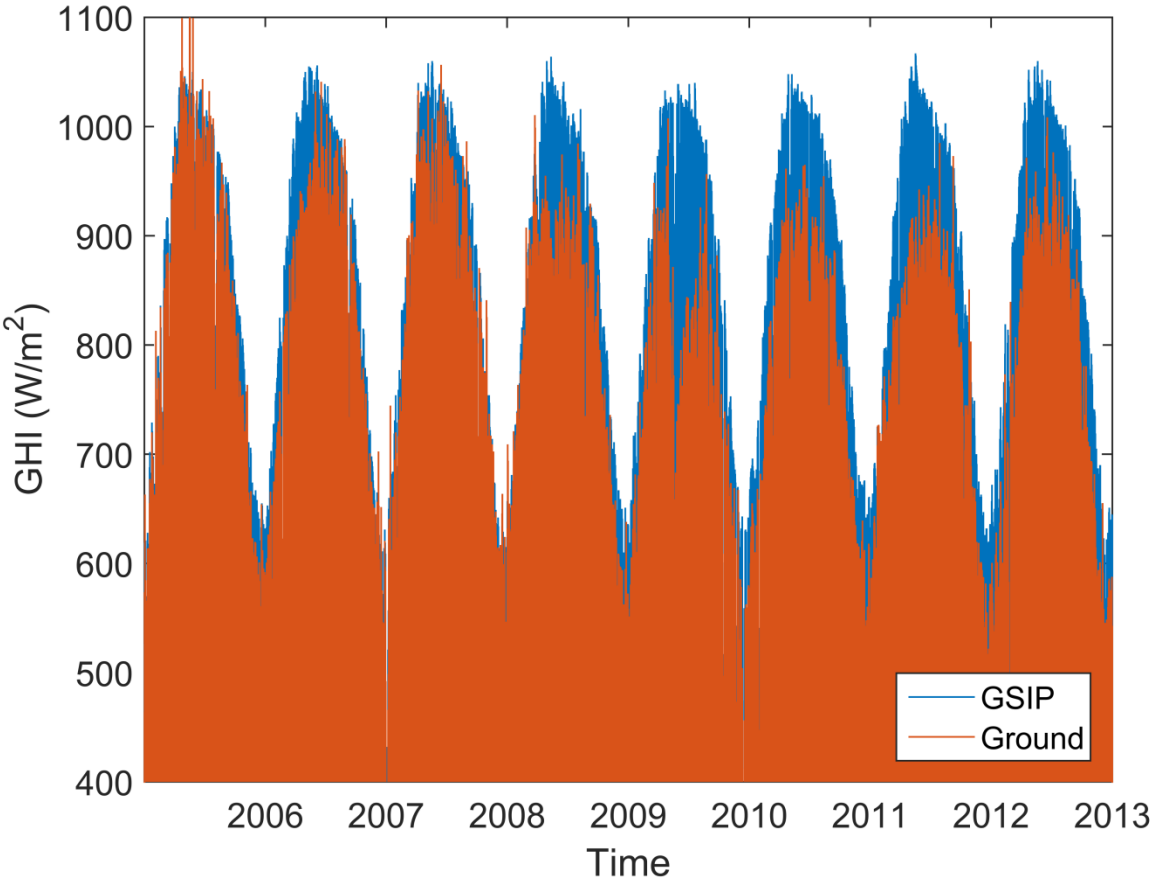



**Figure 2. Comparison of NSRDB and University of Georgia ground data at Attapulgus.**

### *2.2.8. University of Oregon Irradiance Data*

The University of Oregon Solar Radiation Monitoring Laboratory operates a network of 40 solar measurement stations in the Northwest United States. We selected seven sites covering a range of geographic and climate conditions and which were generally free of near-field shading (as determined by visual inspection). Table 8 summarizes the years of data used for the eight sites chosen.

**Table 8. University of Oregon data sites**

| Site | Climate | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---|---|---|---|---|---|---|---|---|---|
| Boise, ID | Csb | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Twin Falls, ID | BSk | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Eugene, OR | Csb | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Portland, OR | Csb | ✓ | ✓ | | | | | | |
| Moab, UT | Csb | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Cheney, WA | Csb | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Seattle, WA | BSk | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

# 3. METHODOLOGY

Here we describe our approach for comparing the ground measured GHI with the NSRDB GHI. Our analysis approach consists of registering the satellite and ground data sets in both space and time, identifying periods of time with clear sky conditions, and computing various comparative statistics. Each step is detailed below.

## 3.1. Registration

While NSRDB GHI data is typically interpreted as the irradiance received everywhere within the 4 x 4 km pixel at the instant of the time stamp, ground data is most often a time-averaged value of GHI measured at some regular interval (e.g., one minute, one hour) using a sensor (in our analysis, a pyranometer) whose aperture is typically about 1 $cm^2$. Thus we view a value of GHI from a ground instrument as a time-average of GHI at a point. In order to compare the data between the satellite and the ground measurement stations, the two data sets must be registered in both time and space. Spatial registration is accomplished by selecting the NSRDB pixel that "contains" the ground measurement site, i.e., the satellite pixel whose center is closest to each ground measurement location. Effects of spatial mis-registration are briefly touched on in Section 4.5.2.

Spatial registration yields two time series of GHI (NSRDB GHI and ground measurement) to compare at each ground measurement location. We synchronized these two time series by averaging the time series with the higher sampling rate to match the lower sampling rate. Our ground data (described in Section 2) includes values averaged over 1-, 5-, 15- and 60-minute intervals. For comparison of NSRDB GHI to 1-, 5-, and 15-minute ground-measured GHI data, we averaged the ground data using blocks centered on the time stamp for each NSRDB GHI value. For comparison of NSRDB GHI to 60-minute ground data (AZMET) we computed a weighted average of the NSRDB GHI values, giving the NSRDB GHI value in the center of the 60-minute period twice the weight of the values at each end. This approach ensures that the resulting NSRDB GHI time averages are taken over the same windows as the AZMET ground data, and that each NSRDB GHI value is used equally in creating the average time series.

## 3.2. Clear Sky Determination

To assess whether agreement between NSRDB GHI and ground measured GHI depends on sky condition, we identify days representing mostly clear sky conditions from the ground data. Because sky condition reports for each of the ground measurement sites were unavailable, we adapted a method for identifying clear sky periods developed by Reno et al. [17] [18]. We did not attempt to identify clear sky periods for either the hourly AZMET data or the NREL Hawaii data.

We first classified each data point in the ground data as clear or not clear. Using a window to characterize the shape of the GHI curve in the neighborhood of each data point, the Reno method uses five criteria to determine if clear sky conditions are present. Four of these criteria are based

on comparisons of the GHI measurements in the window with model-based estimates of the clear sky GHI for the same location and time period: 1) the difference between the average measured and clear sky GHI, 2) the difference between the maximum measured and clear sky GHI, 3) the difference between the path length of the measured and clear sky GHI curves, and 4) the maximum of the absolute value of the deviation of the slope of consecutive measured GHI values from that of the clear sky GHI. The fifth criterion does not depend on the clear sky model and is the standard deviation of the slopes between measured GHI values normalized by the average measured GHI. To determine if a sample represents clear skies, a threshold is associated with each criterion. For one minute samples and ten minute windows, Reno et al. suggest the following threshold values: 1) mean GHI and 2) maximum GHI within ±75 W/m$^2$, 3) measured GHI path length within +10 and -5 of the clear sky line length, 4) maximum deviation of measured GHI slope from clear-sky slope not exceeding 8 W/m$^2$ per minute, and 5) normalized variance of measured GHI slope less than 0.005. A point in time is identified as having clear sky conditions if the time point is contained within a window for which all five criteria meet their thresholds [17]. A day is classified as having mostly clear conditions if most of the time points are identified as clear.

We performed our clear sky determination using the finest time resolution available at each ground site. Table 9 summarizes the thresholds for the algorithm in [17] that are used for the various sample rates used in this study. For those sites with a 5- or 15-minute sampling interval, we adjusted the time window and thresholds for the last three criteria described above because they depend directly on the rate of change in GHI which depends on the data interval and the number of samples in the window. To compensate for the longer data intervals, we shortened the window size from ten to five samples for these data sources. We chose to use five samples in order to keep the window as short in time as possible while keeping enough samples to get a good sense of the local shape of the GHI curve.

Next we classified days as either clear or not clear. For one minute data, we identify a day as clear if 80% of the daylight ground measurements were classified as clear. For data with less frequent sampling rates, we used the count of not clear daylight periods in a day to determine if the day was clear. For five-minute data, we declared a day as clear if less than 24 of its daylight periods were not clear; for 15-minute data, we declared a day as clear if less than 7 of its daylight periods were not clear. The day shown in Figure 3 is classified as not clear.

**Table 9. Thresholds for determination of clear sky samples and clear days**

| | Data Interval (minutes) | | |
|---|---|---|---|
| | 1 | 5 | 15 |
| Criterion for clear-sky determination | | | |
| Mean GHI (W/m$^2$) | 75 | 75 | 75 |
| Maximum GHI (W/m$^2$) | 75 | 75 | 75 |
| Path Length (units?) | $-5 < x < 10$ | $-20 < x < 10$ | $-20 < x < 10$ |
| Deviation of slope (W/m$^2$ per sample) | 8 | 2.0 | 1.0 |
| Variance of slope (unitless) | 0.005 | 0.002 | 0.002 |
| Neighboring clear samples | 9 | 4 | 4 |
| Limit on not-clear samples for clear days | 20% | 23 | 6 |

We verified our approach for data with the five- and 15-minute intervals by averaging one minute data at the PSEL to produce equivalent five- and 15-minute data sets. We then applied our modified approach to these resampled data sets. By visually comparing the classification results obtained using the five- or 15-minute resampled data to those obtained from the one-minute data, we determined that the modified thresholds and classification rules produced equivalent results. Figure 3 shows an example of a partly cloudy day from the New Mexico PSEL location. Note that the same time periods are classified as cloudy in both cases.

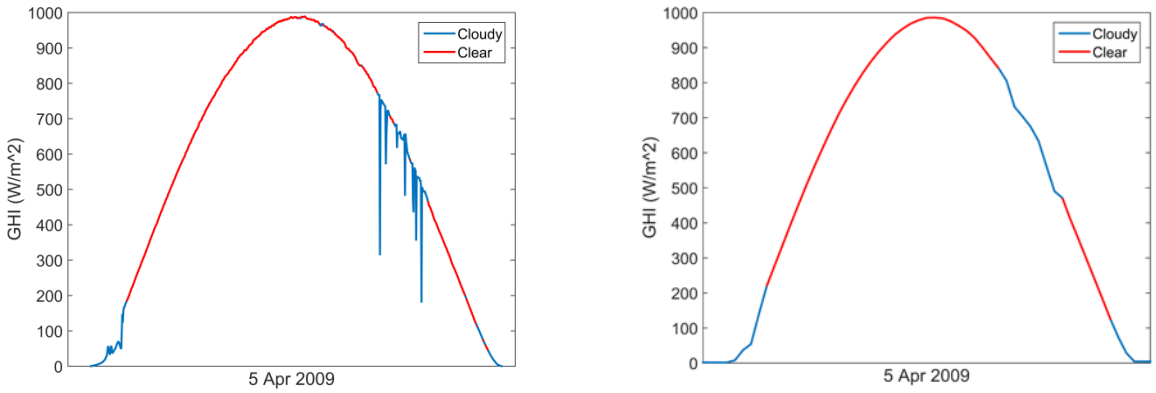



**Figure 3. Clear sample identification using 1-minute (left) and 15-minute (right) data**

## 3.3. Statistics

With the satellite and ground data registered in time and space, we computed statistics to compare the two data sets both for irradiance as well as for insolation (i.e. the time integral of irradiance).

### *3.3.1. Difference statistics*

We computed three statistics based on point-wise comparisons of the satellite and ground data sets at each site. Given $N$ GHI values $G_{satellite}(t_i)$ and $G_{ground}(t_i)$ at a given site, we compute the following statistics:

$$\text{mean bias difference: } MBD = \frac{1}{N}\sum_{i=1}^{N} G_{satellite}(t_i) - G_{ground}(t_i) \tag{1}$$

$$\text{mean absolute difference: } MAD = \frac{1}{N}\sum_{i=1}^{N} \left| G_{satellite}(t_i) - G_{ground}(t_i) \right| \tag{2}$$

$$\text{root mean square difference: } RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left( G_{satellite}(t_i) - G_{ground}(t_i) \right)^2} \tag{3}$$

In all three cases, we compute the average over all daylight periods for which we have a valid data sample from both the satellite and ground data sets.

### *3.3.2. Insolation ratios*

To identify climate or seasonal conditions that may affect the performance of the NSRDB algorithm, we compute an insolation ratio as follows:

$$p(site) = 100 \cdot \frac{\sum_{t_i \in \omega} G_{satellite}(t_i, site)}{\sum_{t_i \in \omega} G_{ground}(t_i, site)}, \quad (4)$$

where $\omega$ represents the subset of the time points that correspond to the condition of interest (e.g., time of day, climate, and/or sky condition). The insolation ratio represents the accumulated irradiance estimated by NSRDB as a percentage of the accumulated irradiance as measured by the ground instrument. Relative to the ground measurement, percent error in the NSRDB GHI value is easily obtained by subtracting 100 from $p$. As with the difference statistics, the insolation ratio includes only daylight samples for which we have a valid data point in both data sets.

Creating box plots of $p$ using different subsets $\omega$ allows assessment of performance of the PSM algorithm in various conditions. A box plot of values of $p$ computed for each year, for example, may reveal variations either in the satellite or in the data sources. By using each month as a different subset, we can look for seasonal trends. By computing statistics using only samples from clear conditions, we can determine if the PSM algorithm performs differently depending on the sky condition.

# 4. RESULTS

We present our comparison of with ground data in three parts:

- We first consider all locations and all times and examine the comparison across years, by month and by climate zone.
- We next consider a subset of the data representing only days with clear skies.
- We finally investigate the comparison over short spatial scales.

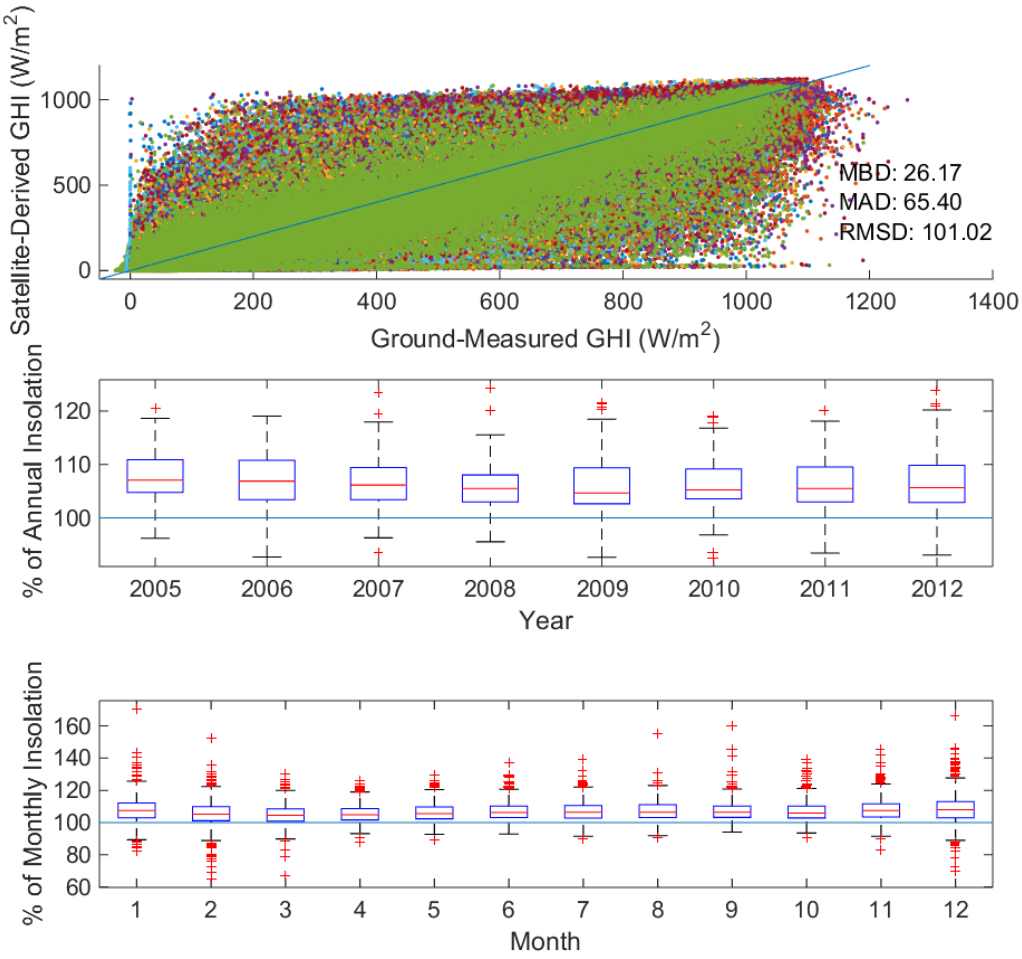
## 4.1. Analysis over all locations and conditions

Figure 4 shows three summary comparisons between the NSRDB and ground data, in which all data from all ground sites is included. The top panel shows a scatter diagram of the satellite-derived GHI plotted against the ground-measured GHI, the middle panel shows a boxplot representing the annual insolation ratios for each year, and the bottom panel shows a boxplot depicting the insolation ratios by month of the year. In the scatter plot, it is evident that the data generally follow a 1:1 relationship, although there is significant spread. Much of the scatter is due to comparing time-averaged point data (ground data) to spatially-averaged moment-in-time data (satellite values).

Three features of the scatter plot in Figure 4 merit special attention: the set of points in the lower left corner for which the ground GHI is negative, and the points lying on or along each of the two axes, where one source reports significant GHI while the other indicates near-zero GHI. The set of points with negative ground GHI is due to the well-known night zero depression phenomenon common to thermopile pyranometers. Although we exclude most night-time samples from our analysis, a small number of early morning and/or late evening samples in which this phenomenon is evident remain in the data for a several sites with thermopile instruments (i.e., SRRL and SurfRad sites). These samples are less than 1% of the analyzed data for any site; we do not believe they materially affect the results of our analysis.

The vertical line of points indicates samples in which the ground instrument reports a value of zero instead of a “not-a-number” flag. Zero values likely result from non-functioning or non-reporting instruments rather than from the absence of GHI. These samples also comprise less than 1% of samples for each site and we did not exclude them from our analysis. We do not believe their inclusion affects our conclusions.

The horizontal line or points represents sample times for which the NSRDB GHI is nearly zero while the ground measurement exceeds 200 W/m$^2$. These samples occur most frequently during the summer months at sites north of 37° N and east of 117° W. They appear to be distributed uniformly across time of day. These data are relatively infrequent (at greatest less than 2%, which occurs at the SRRL site). We included these data in our analysis and investigated possible reasons for the large discrepancy between the NSRDB GHI and ground data. For one site we found a pattern which may indicate the underlying cause (see Sect. 4.5.2), but at other locations we found no systematic patterns. In our opinion these data indicate times when the PSM algorithm is not performing and which merit further, detailed investigation.




Note: maximum whisker length is 1.5 × interquartile range.

**Figure 4. Ratios of NSRDB GHI to ground GHI, all locations and years.**

The mean bias difference of 26.71 between ground and NSRDB GHI indicates that the NSRDB values tend to be higher than the values from ground measurements. This tendency is evident in the middle panel of Figure 4. For each year, the median annual insolation as measured by NSRDB is on the order of 5-10% above the median annual insolation as measured at the ground sites. The boxplot in the bottom panel appears to show a slight seasonal pattern in the amount of bias between the satellite and ground, as well as the amount of variation in the insolation ratios (the whiskers in winter and summer months appear to be longer than in spring and fall).

We examined the comparison between NSRDB and ground data at each site to judge the extent to which the differences summarized in Figure 4 may result from bias in the ground data rather than in the NSRDB GHI values. Figure 5 shows a box plot representing the distribution of the annual insolation ratios at each site. Figure 6 shows a similar plot where the annual insolation ratios are aggregated over all sites in the same data source. These two figures show that the difference between the satellite estimates and ground measurements depends systematically on the source of the ground measurements.

In particular, at the AZMET sites we note a consistent bias but generally only minor variation among the individual AZMET sites. The AZMET data are measured using instruments of lower quality with annual calibration of the instruments but without frequent cleaning. Consequently we may anticipate that the AZMET values may exhibit a small negative bias, due to sensor soiling, which we judge would have a minor effect based on operational experience at Sandia's PSEL. Thus, the small, consistent bias at AZMET sites evident in Figure 4 could result from soiled ground instruments rather than a deficiency in PSM's algorithm.

For Georgia data, however, we observe considerable bias, significant variation among the sites, as well as substantial year-to-year variation at many individual sites. The Georgia data are also measured with lower quality instruments and their maintenance is unknown. Because climate conditions are consistent across the Georgia sites, it seems unlikely that NSRDB GHI would exhibit wide variation in accuracy among these locations. We view the most likely explanation for the results in Figure 5 to be a lack of quality in the ground data. Unmaintained sensors can be anticipated to show a negative bias in their measurements; instruments without periodic calibration are not expected to maintain consistent accuracy year-to-year. Moreover, site-specific investigations have shown that instrument siting may not have always received careful attention. Consequently, we view the comparison between NSRDB GHI and the Georgia data as more revealing of problems with the Georgia data rather than providing insight into the performance of PSM.

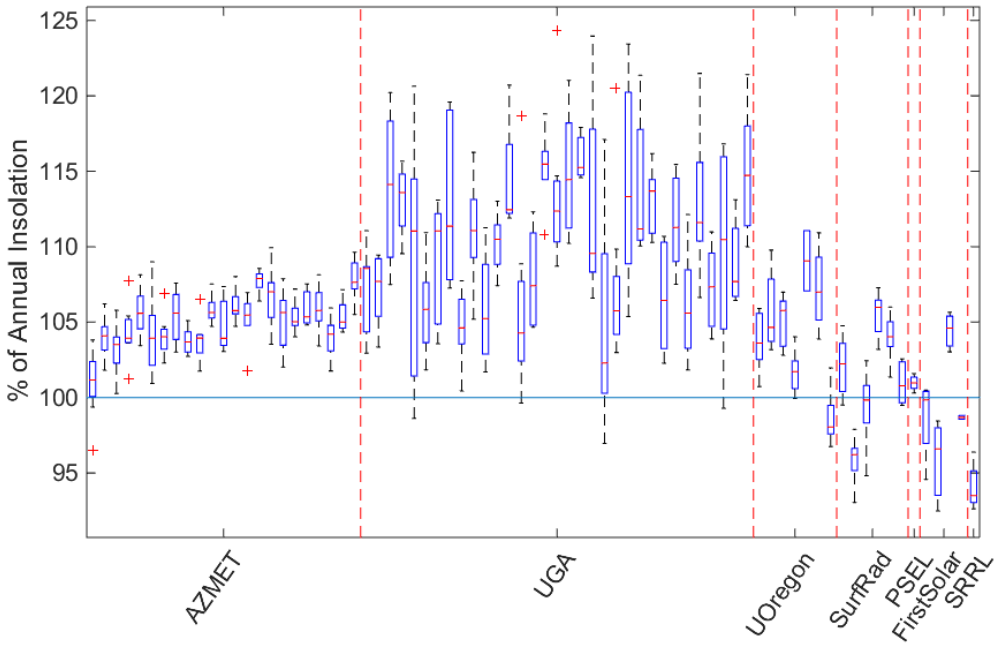



**Figure 5. Annual insolation ratios (NSRDB GHI to ground GHI) by ground site**

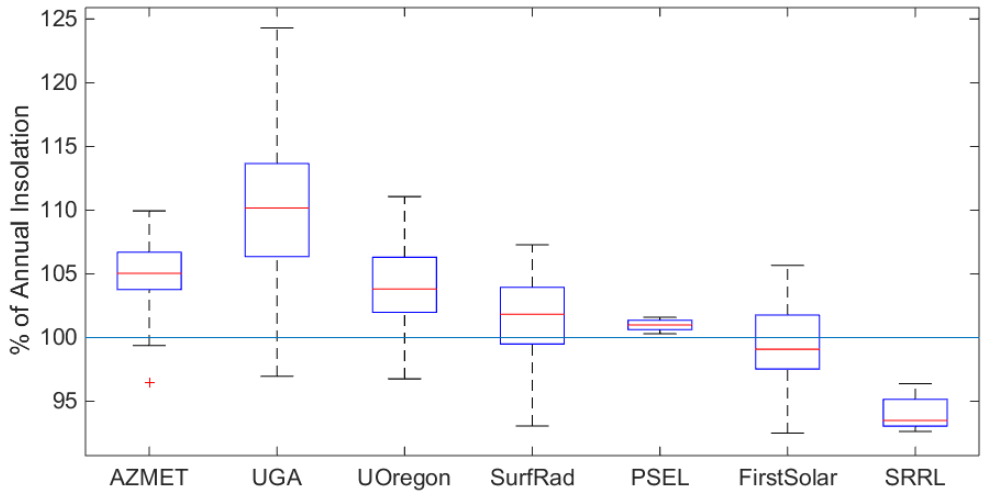



**Figure 6. Annual insolation ratios by ground data source**

Removing the University of Georgia data from the analysis yields significantly improves the comparison between NSRDB GHI and ground data, as shown in Figure 7. The data appear to cluster slightly more tightly about the 1:1 line in the scatter diagram. The statistics and the box plots indicate that the satellite measurements remain generally high relative to the remaining ground sites, though substantially less so (MBD drops from 26.71 to 12.84, RMSD from 102.22 to 85.11).

Figure 8 shows the distributions of monthly insolation ratios for sites in six of the nine climates represented in the set of ground sites. Omitting the University of Georgia sites, at three of the nine climates, we have only a single site-year of data, so a distribution of monthly insolation ratio cannot be computed. Arid climates (B) are shown on the left-hand side of Figure 8, while temperate (C) and continental (D) are on the right. As a depiction of the interquartile range, the height of each box in Figure 8 provides a sense of the amount of variation (among sites and across years) in the insolation ratios for each month.

A clear seasonal pattern is evident in the variation of monthly insolation ratios for temperate and continental climates on the right side of Figure 8. Specifically, much greater variation is noted in the winter months than that observed in the summer months for climate types Csb (coastal Northwest), Dfa and Dfb (Midwestern states). In contrast, the variation in monthly insolation ratios does not appear to follow such a distinct pattern for the arid climates. These results suggest that PSM may not estimate GHI as accurately when snow or ice cover is present. Other satellite-to-GHI algorithms have adapted to use both infrared and visible channels to distinguish between snow cover and clouds [19], a feature which may not be fully mature in PSM.

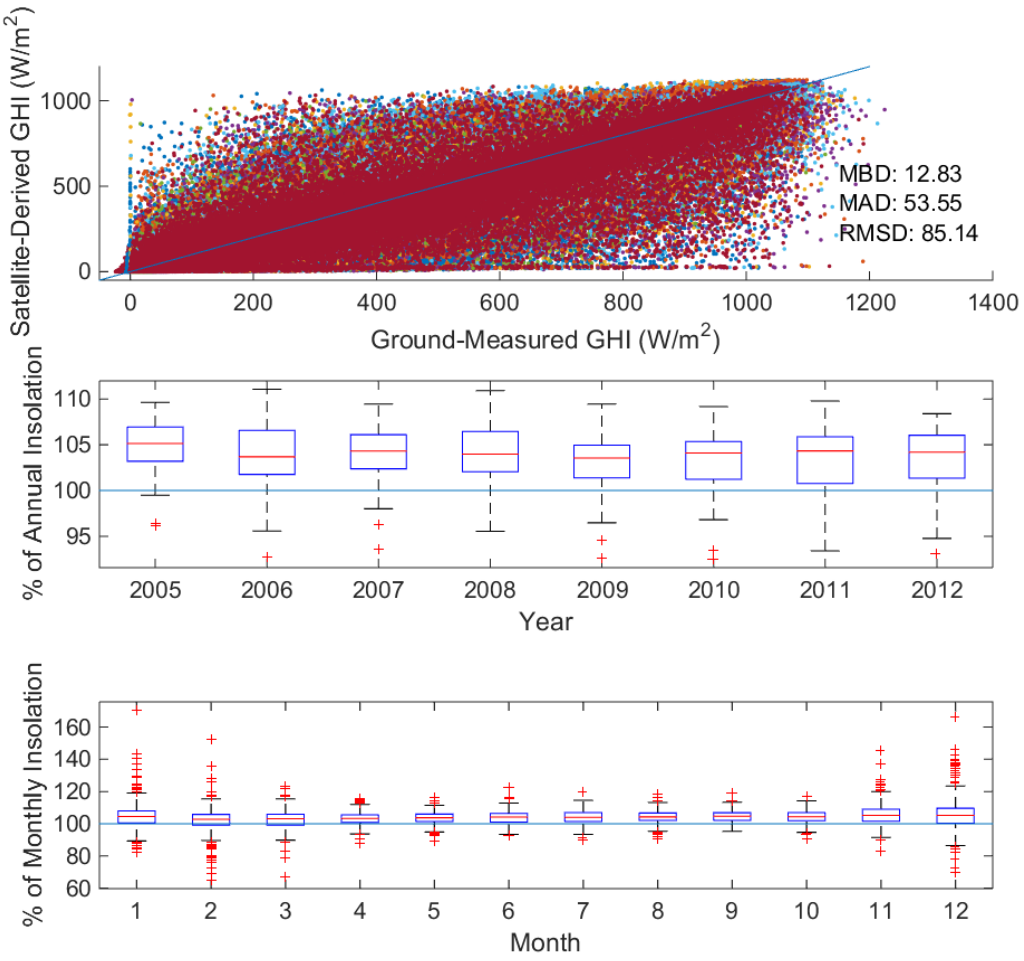



**Figure 7. Ratios of NSRDB to ground GHI omitting the University of Georgia sites.**

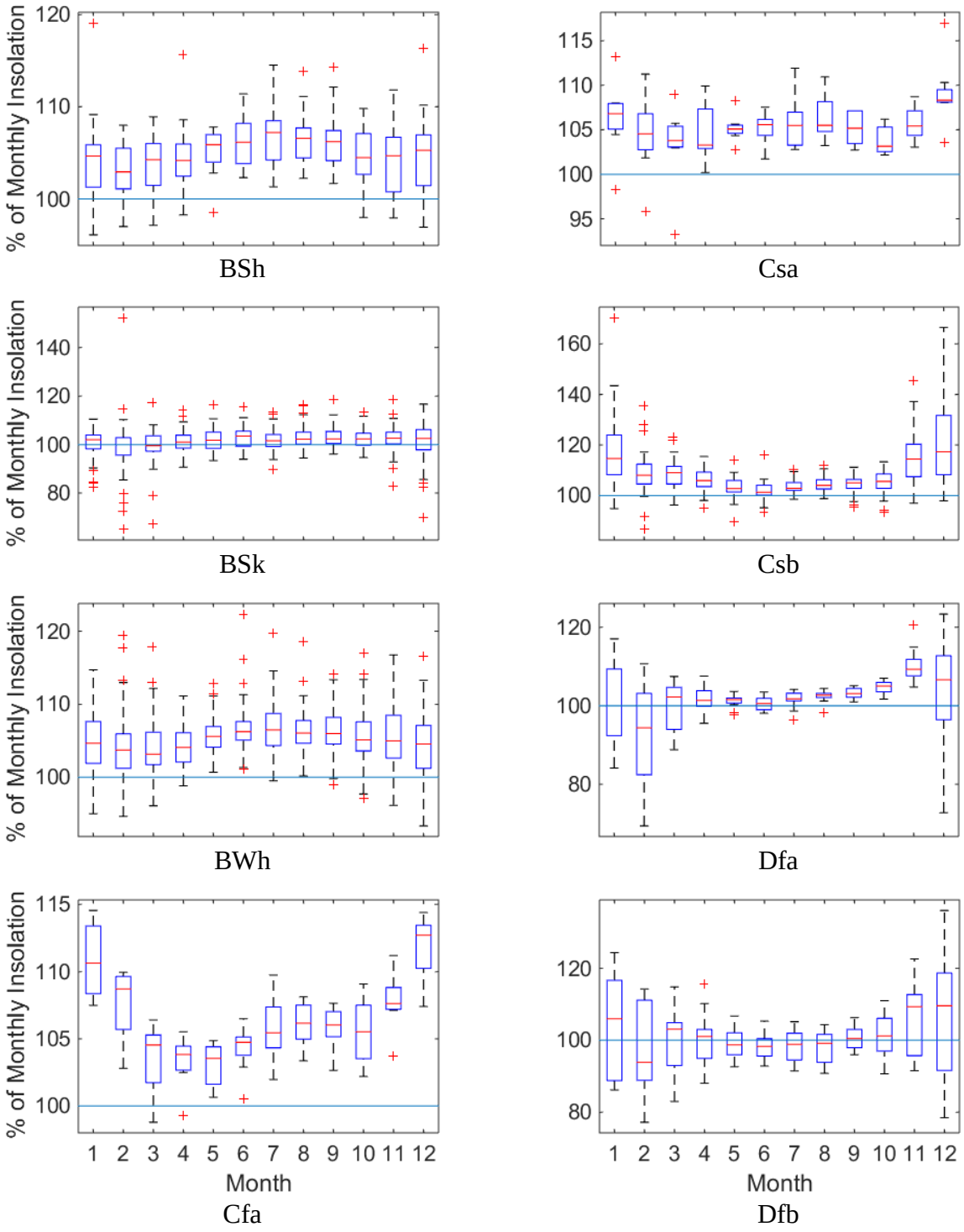



**Figure 8. Distributions of monthly insolation ratios (NSRDB GHI to ground GHI) for eight climate types**

## 4.2. Comparison During Clear Sky Conditions

We analyzed PSM performance during clear sky periods by comparing NSRDB GHI with ground data measured when we determined clear sky conditions were persistent. We identified each day as having either clear or not clear conditions using the algorithm described in Sect. 3.2. For clear days we compared NSRDB GHI with 30-minute averaged ground data for time samples when the ground data indicated clear conditions (a day identified as clear could have a minority of time samples with not clear conditions — see Table 9.)

Clear conditions for ground data mean only that the data indicate that no clouds were interposed between the sun and the sensor; clear conditions do not imply the absence of cloud shadows in the area near the sensor. Because the NSRDB GHI data comprise spatial averages at a moment in time, it is possible that the ground data could show clear conditions while clouds are present within the NSRDB GHI pixel and thus introduce a bias in the comparison (where NSRDB GHI will appear less than the ground values). We reduced the likelihood of including data samples with this discrepancy in two ways:

1. The algorithm which identifies clear times requires that clear sky criteria are also met for some period of time before and after the time sample of interest. Accordingly a time stamp identified as clear corresponds to a period of time where the data indicate the absence of cloud shadows.
2. We use only ground data on clear days when the great majority of time samples are identified as clear.

Figure 9 compares NSRDB GHI to ground measurements during clear sky conditions on clear days. As discussed earlier, because the AZMET ground data were sampled hourly, we did not identify clear sky times or days for AZMET sites. As a result, the AZMET data are omitted from this comparison for clear sky conditions.

As one might expect, the data depicted in Figure 9 shows considerably less spread than was observed when all sky conditions were included. The mean bias difference for clear conditions is not much different from all sky conditions, but the mean absolute and root mean square differences are significantly reduced. Though reduced in magnitude from the all-conditions case, the annual insolation ratios continue to show a positive bias for NSRDB GHI. We accept as the more likely explanation that the positive bias originates with the NSRDB GHI clear sky algorithm because we view as less credible the explanation that monthly insolation is systematically underestimated at all of the ground locations. Monthly insolation ratios for clear-sky conditions continue to show higher variation in the winter than in the summer.

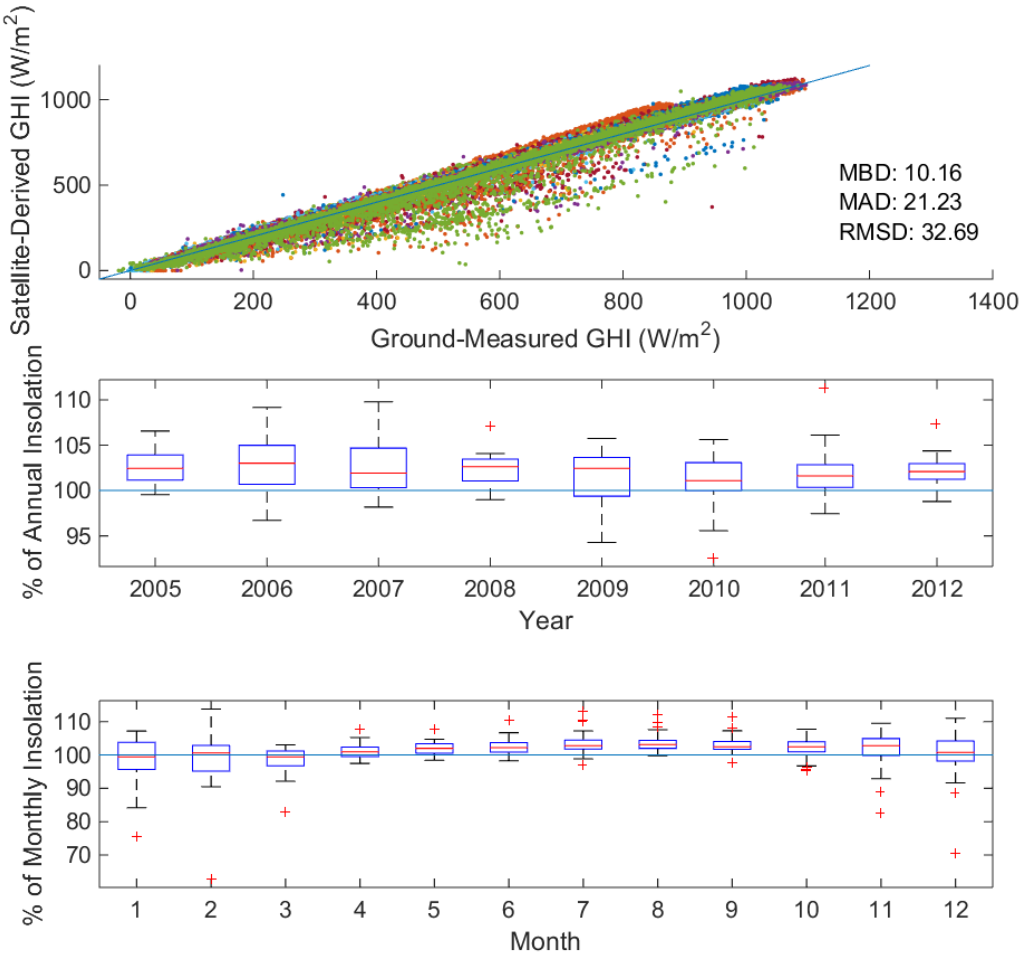



**Figure 9. Annual insolation ratios (NSRDB GHI to ground GHI) for clear sky periods**

Figure 9 shows significant outliers for winter months, which we examine in some detail by separating the data by climate class. Having omitted both the AZMET and University of Georgia data, only four climates are represented with sufficient data (three or more clear days within a month) to examine the variation in monthly insolation ratios by climate. The distributions of monthly insolation ratios for these remaining climates are shown in
Figure 10. Several observations may be made regarding the plots in
Figure 10 and comparison with the corresponding plots in Figure 8:

- NSRDB GHI tends to exceed the ground measurements for all climate conditions during most months. Greater differences in monthly insolation are generally observed during summer months, whereas greater variation in the difference is generally observed during winter months. Without data to corroborate the ground measurements we are uncertain if the bias results from PSM or from the ground instrumentation.
- At climates subject to winter snow or ice conditions (primarily Dfa and Dfb but also BSk) PSM predicts lower monthly insolation than is measured at the ground sites for winter months but greater monthly insolation for summer months. Moreover, outliers with NSRDB GHI significantly lower than ground measurements are observed for winter months. These patterns are not evident for the sites in the Cfa and Csb climate classes, all of which are in the northwest United States where snow or ice cover is unlikely to persist

and for which ground data are sourced from the University of Oregon. For these sites, the NSRDB GHI tend to be about 5% higher than the corresponding ground measurements.

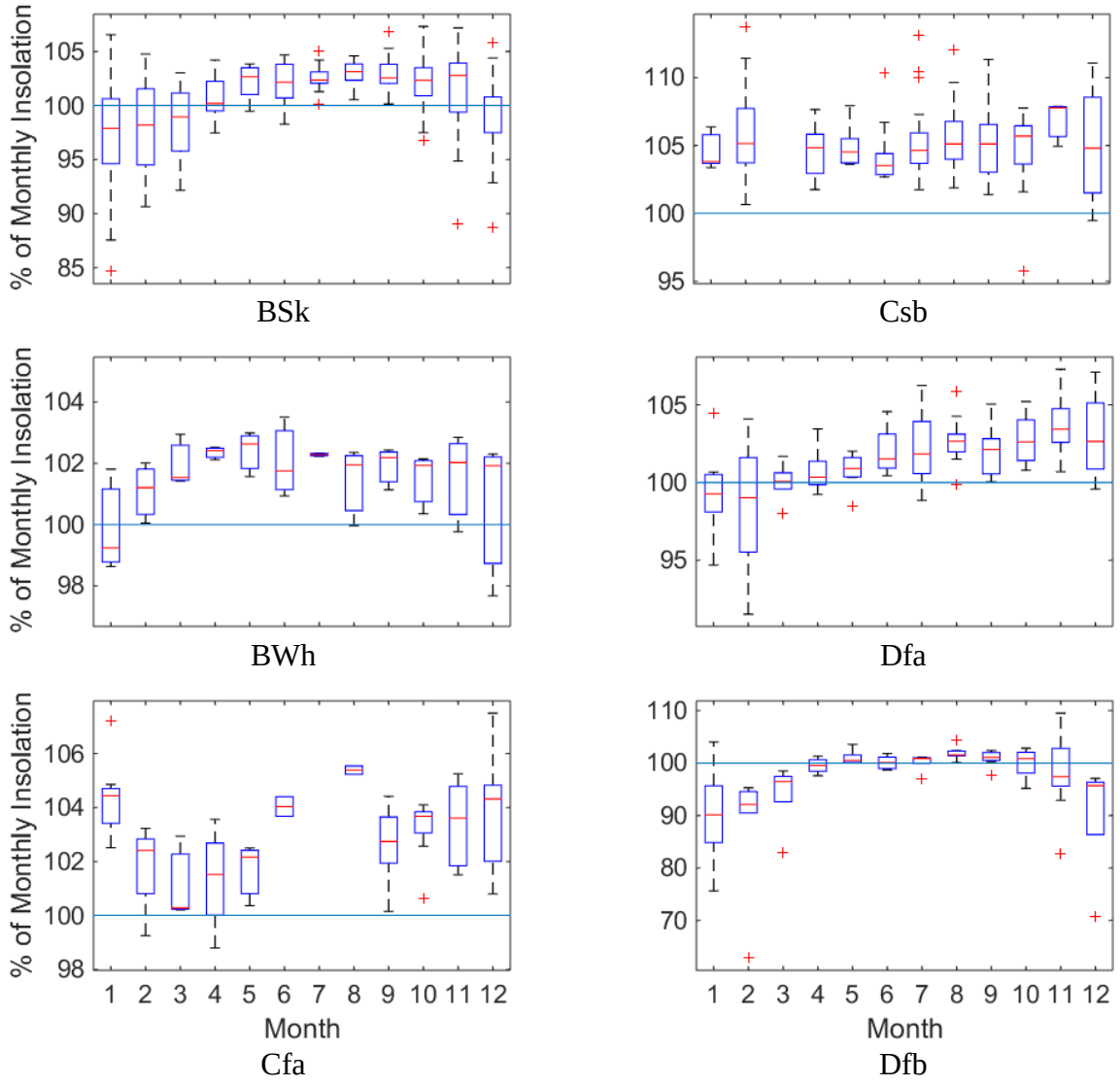



**Figure 10. Distributions of monthly insolation ratios during clear-sky conditions**

To investigate the wintertime outliers, we examined the time series from NREL's SRRL (Dfb climate class) which we view as a reliable source of ground data for GHI. Figure 11 displays the time series of ground measurements of GHI for clear days and the corresponding time series for NSRDB GHI for months when the NSRDB insolation and insolation from ground data are substantially different (and present low outliers in Figure 10). Excellent agreement is seen for some, but not all days. In particular, we note a sequence of days from 16 to 20 November 2009 during which the ground data indicate clear weather conditions but NSRDB data indicate otherwise. Analysis presented next indicates that the difference originates with PSM and is manifest when snow is on the ground.

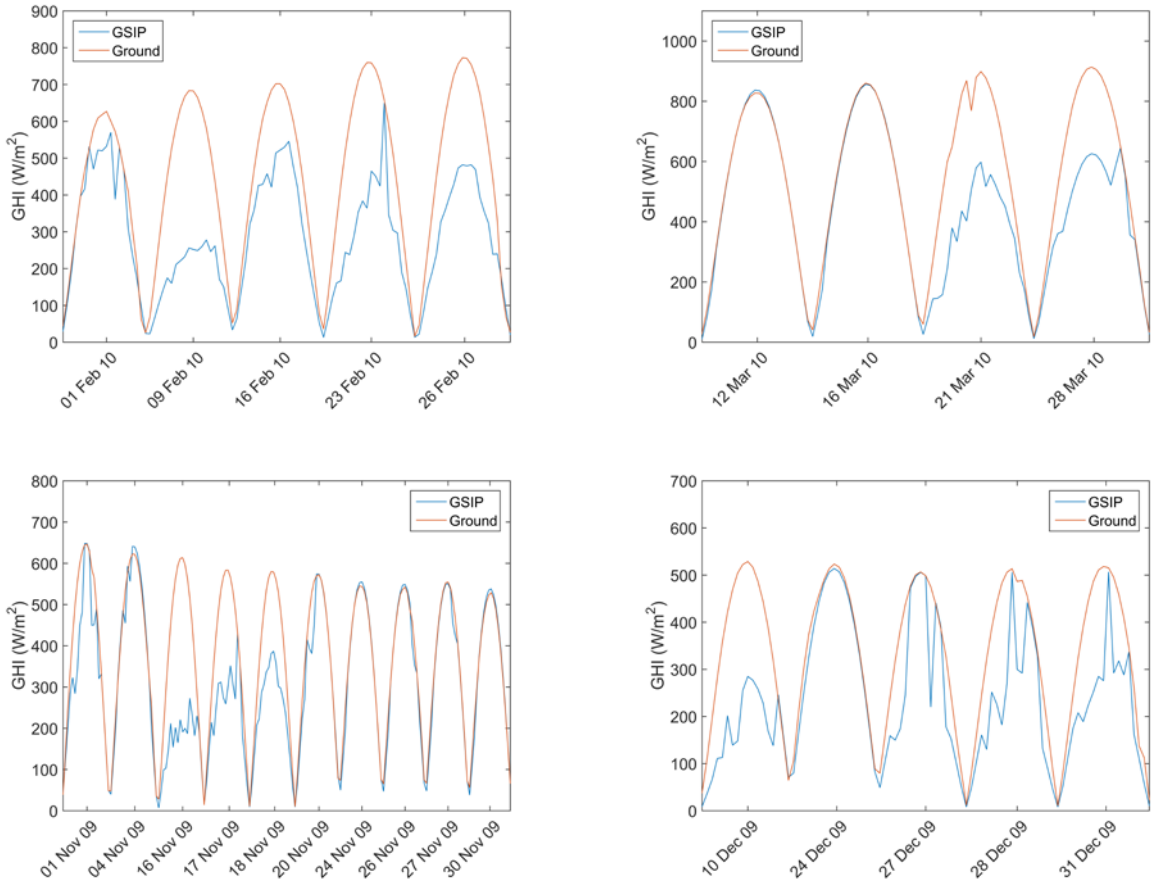



**Figure 11. SRRL (climate class Dfb) GHI profiles for outliers during clear sky conditions**

Figure 12 shows the ground GHI time series for the nine day period 14–22 November 2009; days identified as clear and not clear are indicated by green and red lines, respectively. To corroborate this identification we obtained cloud ceiling measurements from the NOAA weather station nearest to the SRRL (at Rocky Mountain Metropolitan Airport in Broomfield, CO, approximately 15 miles north). The cloud ceiling values of 22000ft during 16–20 November 2009 correspond to unlimited ceilings which corroborates our identification of these days as having clear skies. Weather records[2] indicate that approximately 10 inches of snow fell on the area during 14–15 November 2009. The ground and ceiling data agree that the storm was followed by five days with clear skies. In contrast, the NSRDB GHI values increase slowly each day, and do not agree with the ground measurements until the fourth or fifth day after the storm.

This delayed recovery suggests that the PSM algorithm does not distinguish well between clouds and snow on the ground. Figure 13 shows the GOES visible band images taken at approximately noon local time on each day shown in Figure 12. Green lines in Figure 13 represent state borders, red lines denote 40° N latitude and 105° W longitude, and the blue dot indicates the approximate SRRL location. Widespread clouds are evident on 14 and 15 November. A similarly broad region of high reflectivity is apparent on the 16th. However, comparison between subsequent

[2] http://www.goldencoweather.com/wxsnowdetail.php?year=2010

frames suggests that this region is snow cover which diminishes over time, rather than clouds. The steady decrease in the extent and brightness of the snow cover appears to correlate with the increase in NSRDB GHI noted in Figure 12.

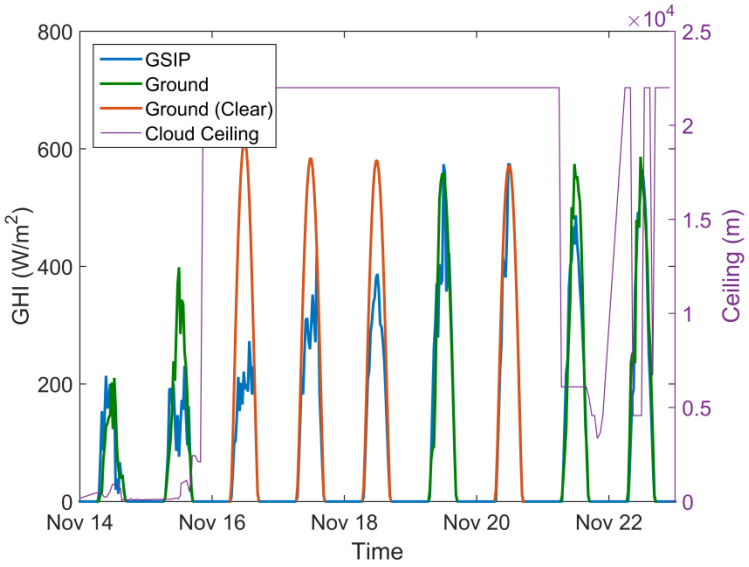



**Figure 12. SRRL GHI and cloud ceiling in November 2009**

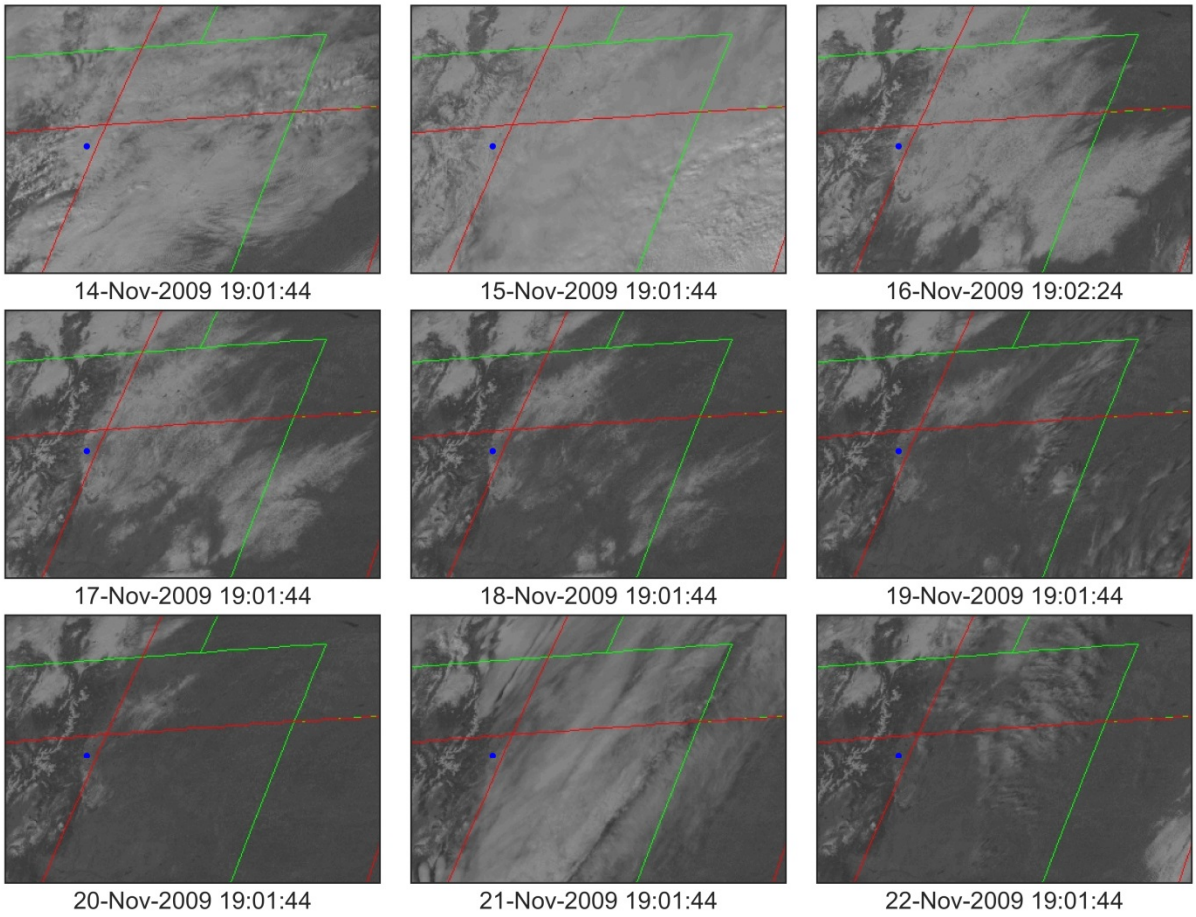


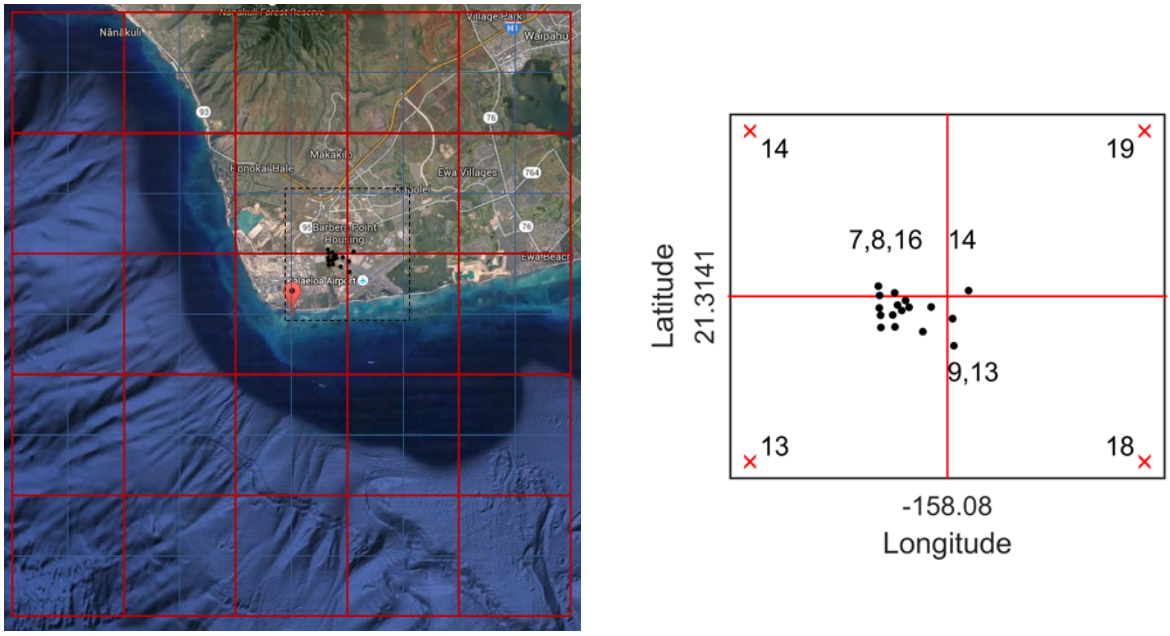

**Figure 13. Mid-day GOES images of northeast Colorado (times are UTC)**

## 4.3. Small-Scale Variability

The irradiance data produced by PSM yields GHI values at the resolution of a single pixel which represents a spatial average over an area that is roughly 4 km x 4 km. Here, we compare NSRDB GHI with ground measured GHI at a number of locations within a single pixel or across a few adjoining pixels. We use the NREL Hawaii data and one of the First Solar sites for this comparison.

### *4.3.1. Comparison with NREL Hawaii data*

The NREL Hawaii data comprise 17 sensors within one square kilometer which spans several GOES image pixels. These data allow evaluation of both intra-pixel and between-pixel variability. Data from each sensor is compared against the corresponding NSRDB GHI value determined by latitude and longitude. The sensors are clustered rather tightly near the intersection of 4 pixels with 11 of the 17 sensors falling within a single pixel. The sensor locations are shown in Figure 14.




Note: Left: red lines are nominal pixel boundaries. Right: × represents a pixel center, number near a red × is a pixel reference; other numbers refer to ground sensors within each pixel. Detailed instrument locations at http://www.nrel.gov/midc/oahu_archive/map.jpg.

**Figure 14. NREL Hawaii Sensor Locations.**

Figure 15 shows the monthly insolation for NSRDB GHI for a single satellite pixel (13) as a percentage of the monthly insolation measured by each of the 17 ground sensors (including sensors outside of pixel 13). Roughly 3% (of the ground measured insolation) variation is apparent among the ground sensors for any given month; this variation results from variation in sensor calibration rather than PSM, because in these results a fixed value of NSRDB insolation is compared with each sensor. However, this variation is minor compared to a bias in the satellite estimate which varies from month to month, suggesting a systematic source of bias in the

NSRDB GHI results. It is unclear why August departs from the general trend of greater bias in the summer as compared to the winter. It may be that weather conditions (i.e., occurrence of clear vs. non-clear conditions) during August were more similar to weather conditions during winter months than during summer months.

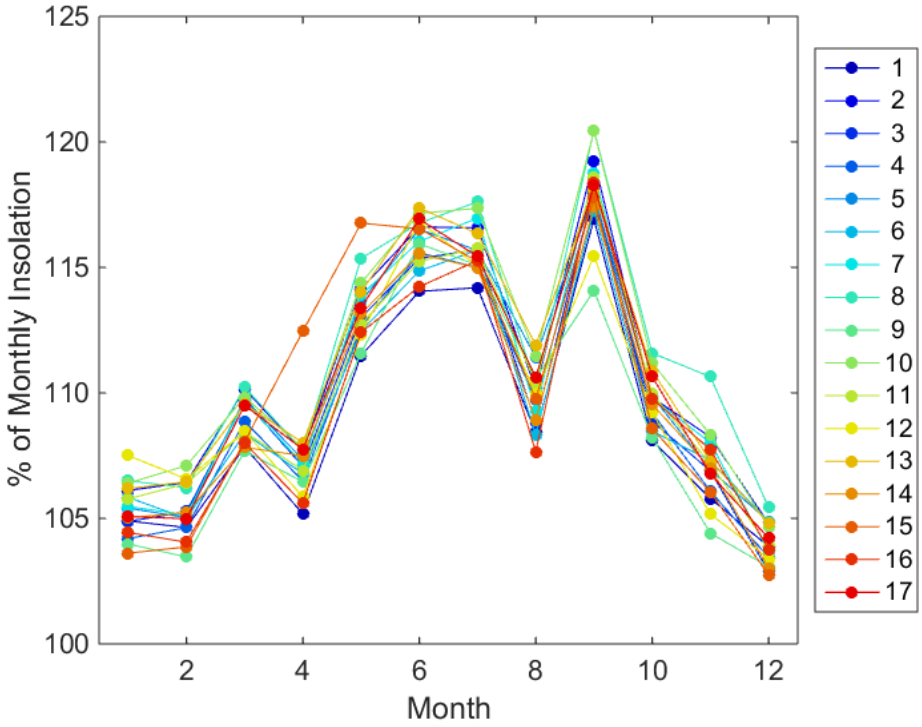



**Figure 15. Monthly insolation ratios for NREL Hawaii data**

Figure 16 shows a box plot of monthly insolation ratios obtained by comparing the ground GHI values with NSRDB GHI values for the pixel containing the ground sensor's location (i.e., pixel 14 for sensors 7, 8, and 16; pixel 18 for sensors 9 and 13; etc.). We note in Figure 16 that significantly less bias is evident for sensors 7, 8, 9, 14, and 16—five of the six sensor locations that fall outside of pixel 13 (the other location is sensor 13 which shows less bias but not significantly less than locations within pixel 13). Figure 13 indicates that the agreement between NSRDB and ground GHI is better outside of pixel 13. For each sensor, the variation indicated by the box plot arises from seasonal changes in the level of agreement between NSRDB and ground measurements, and show that the level of agreement for pixel 13 vs. the other pixels persists throughout the year.

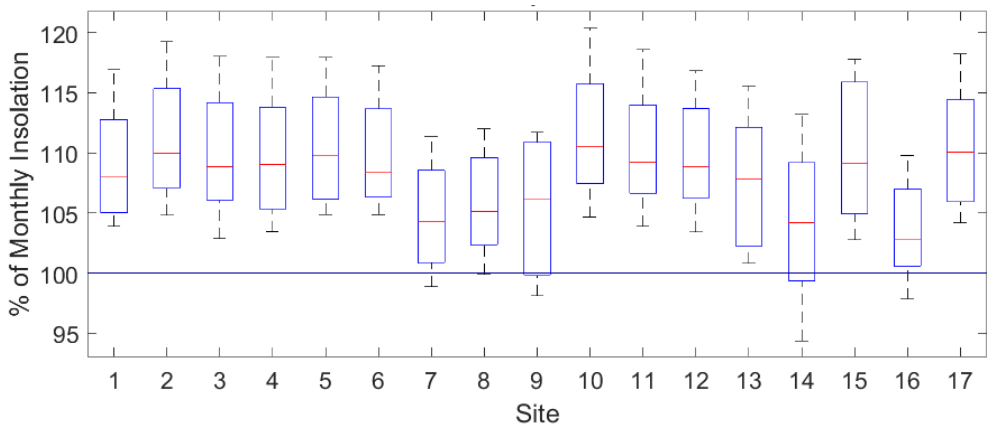



**Figure 16. "Within-pixel" comparison of monthly insolation ratios**

We also compared the ground measurements with each of the four NSRDB pixels indicated in Figure 11. Figure 17 through Figure 20 show box plots of monthly insolation ratios comparing all ground sensors with NSRDB GHI in each pixel. In all figures, the NSRDB insolation is consistently higher than the different ground values indicating a bias in NSRDB's values. However, the bias is significantly higher for pixels 13 and 18 (the southern pair) as compared to pixels 14 and 19 (the northern pair). The most credible explanation appears to be that NSRDB GHI is less consistent with the ground data in the southern pair of pixels, in which the majority of the pixel area is ocean rather than land. Other possible explanations for the differing levels of agreement among pixels include mis-registration of the NSRDB or ground sensor coordinates (i.e., locating the ground sensors in an incorrect pixel), or a misalignment in time of the time-averaged ground measurements relative to the instantaneous NSRDB value. Mis-registration seems very unlikely given that the pixel coordinates are taken from the GOES image data, and that mis-registration of ground instruments would imply a significant and systematic error in all ground locations. Other authors [20] have observed better correlation between time-averaged ground measurements and instantaneous, spatially-averaged GHI derived from satellite observation when the prevailing winds are taken into account. Intuitively, the time average should be taken along a path running from downwind to upwind of the sensor location, corresponding to the GHI that occurs at the sensor location from earlier to later than the time of the satellite observation. However, neither of these explanations (mis-registration or temporal misalignment) seems capable of explaining the variation observed in Figure 15 or Figure 16, thus it seems more likely that the explanation lies in the PSM algorithm.

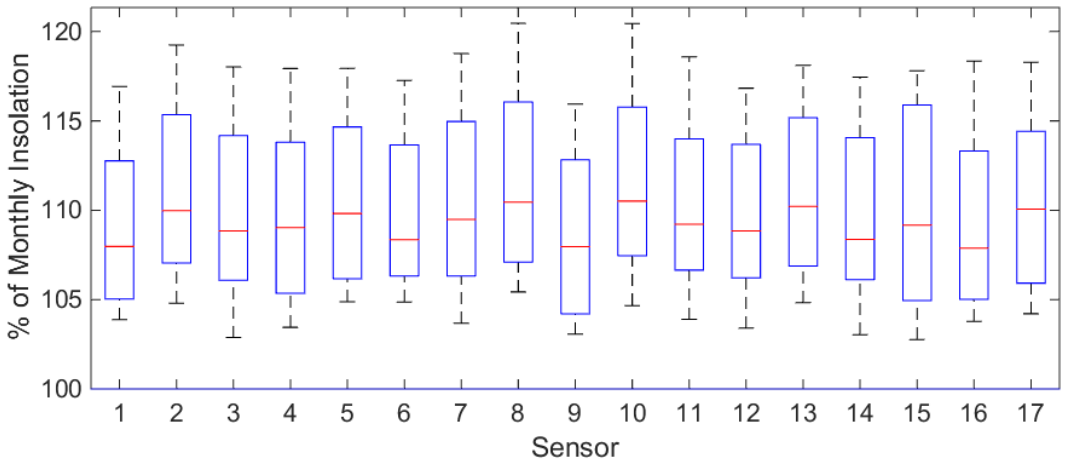



**Figure 17. Comparison of central satellite pixel (13) with all ground sensors**

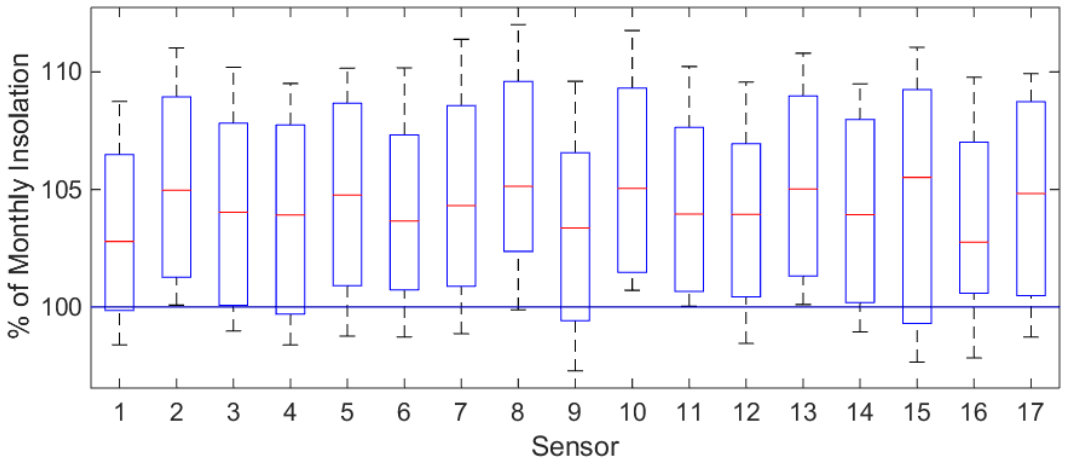



**Figure 18. Comparison of satellite pixel north of center (14) with all ground sensors**

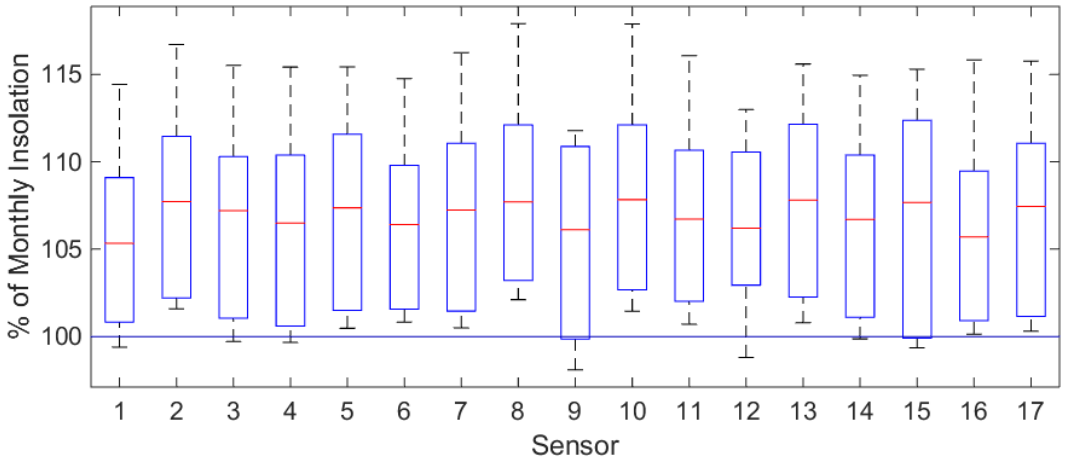



**Figure 19. Comparison of satellite pixel east of center (18) with all ground sensors**

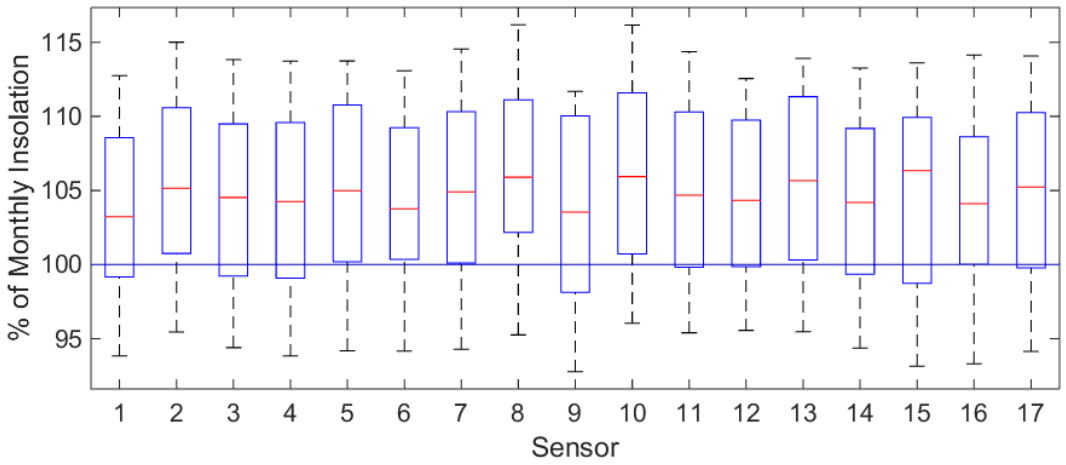



**Figure 20. Comparison of satellite pixel northeast of center (19) with all ground sensors**

### *4.3.2. First Solar Plant 1*

In our initial comparison of NSRDB GHI to ground data, we noted significant, non-random discrepancies between NSRDB GHI and ground data at the location of First Solar Plant 1. These discrepancies apparently resulted from imprecise coordinates for the ground sensor and thus to an incorrect matching between the ground sensor and the NSRDB image pixel, as discussed here. Investigation of this specific location provides insight into the potential variation of NSRDB GHI over short distances.

The latitude and longitude coordinates provided for Plant 1 in the First Solar data set were apparently rounded to the second decimal place of minutes. These low-precision coordinates led to misregistration of the 5x5 grid of satellite pixels requested for this location, as illustrated in Figure 21. Pixel centers are indicated in Figure 21 by the intersection of the blue lines, while red lines midway between each center mark the pixel boundaries. As indicated, the red arrow shows the location of the low-precision coordinates for the ground station toward the bottom of the pixel marked B. Along the northern edge of the pixel marked C, a solar installation is clearly evident. Closer inspection of the area reveals a number of support buildings on the north side of this installation. Assuming that the GHI instrument for this site is actually located near these support buildings , we estimated coordinates for the ground sensor and found them to be at the southern edge of the pixel marked A, approximately 4 km away from the provided coordinates.

Comparing the ground data with the NSRDB GHI in each of these two pixels, we found distinct differences between the ground sensor data and that of different satellite pixels. Figure 22 shows time series for the ground sensor and the pixels marked A and B during a representative three-day period. Ground sensor data is shown in blue; open circles indicate daylight samples and red dots within circles indentify days determined to be clear (28 February and 1 March). The red curve represents NSRDB GHI from pixel B, while the green curve shows the data from pixel A (just west of pixel B). From these curves, it is immediately apparent that the NSRDB GHI data from pixel A is much more consistent with the ground data. In contrast pixel B reports significantly lower GHI than either pixel A or the ground sensor for much of each day.

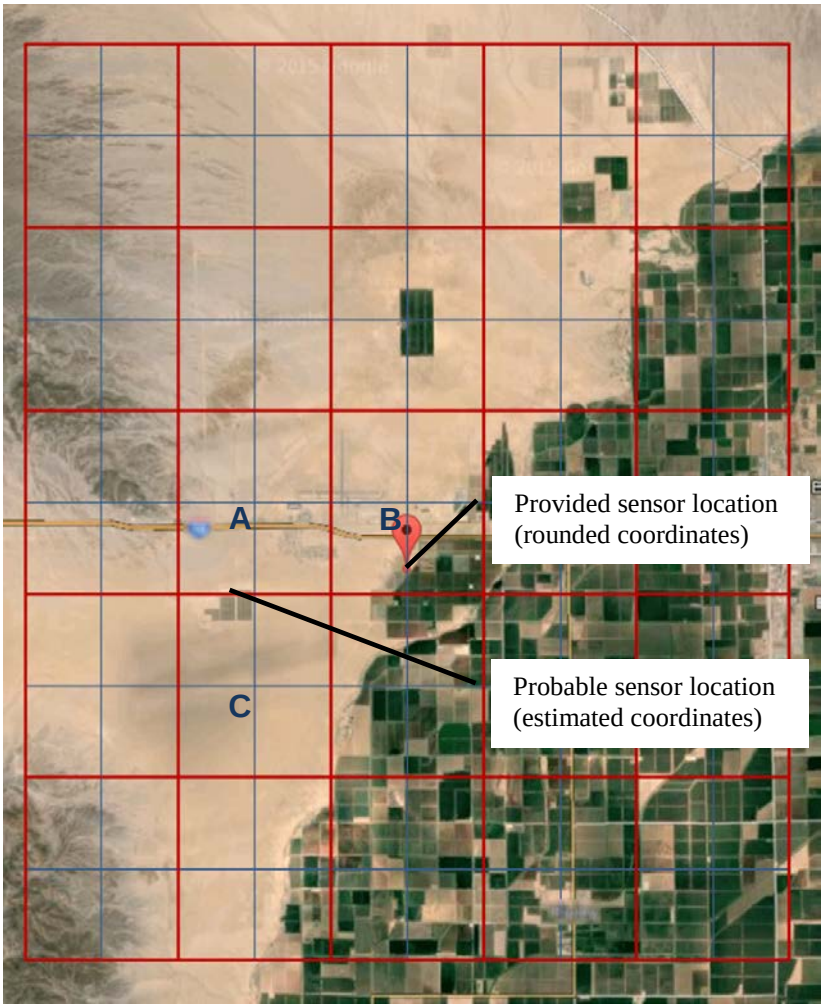



**Figure 21. First Solar Plant 1 Location**

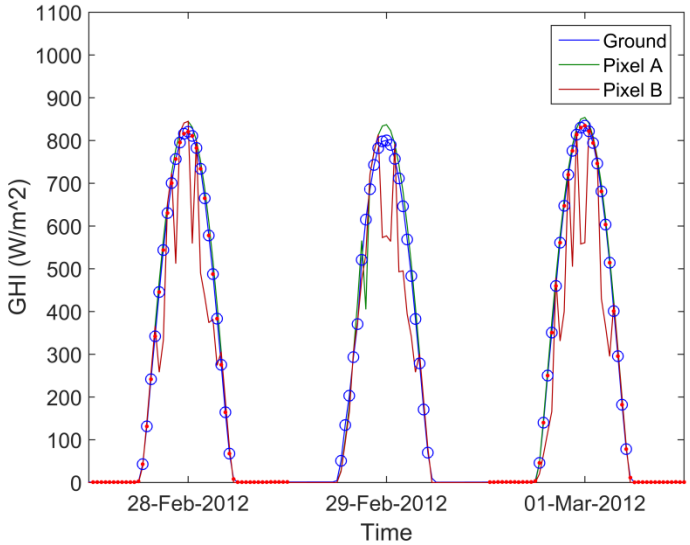



**Figure 22. GHI Time series of First Solar Plant 1 and NSRDB pixels A and B**

The differences between these two pixels persist over multiple days as shown by Figure 23. The data for pixel A are more tightly clustered about the 1:1 line, and there is a dramatic difference in all three comparative statistics. Unsurprisingly, using the data from pixel A also has a sizable effect on the monthly insolation ratios, as shown in Figure 24. Figure 24 also corroborates the observation noted in the discussion of the Hawaii data, that the PSM algorithm appears to underestimate GHI in the winter months, and overestimate it in the summer. Though significantly reduced, this effect persists when comparing data from pixel believed to contain the ground sensor.

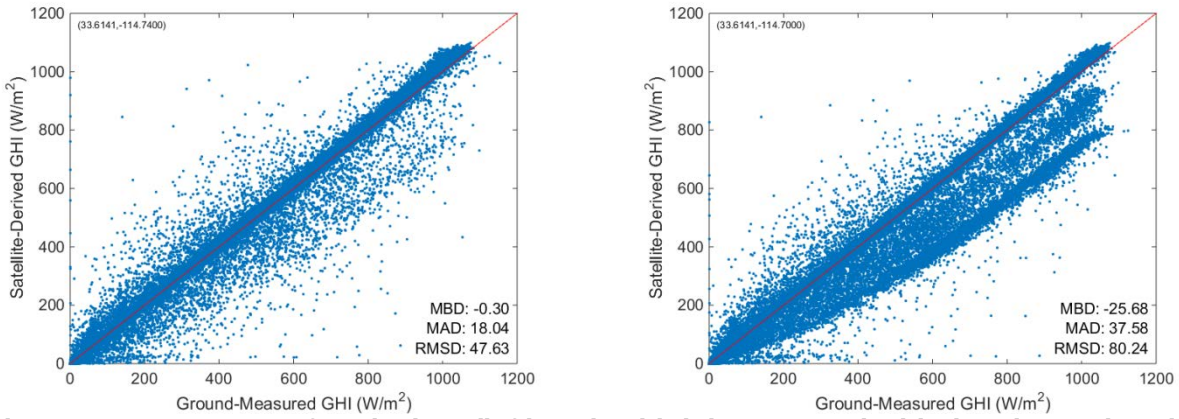



**Figure 23. NSRDB GHI for pixels A (left) and B (right) compared with the First Solar Plant 1 ground sensor**

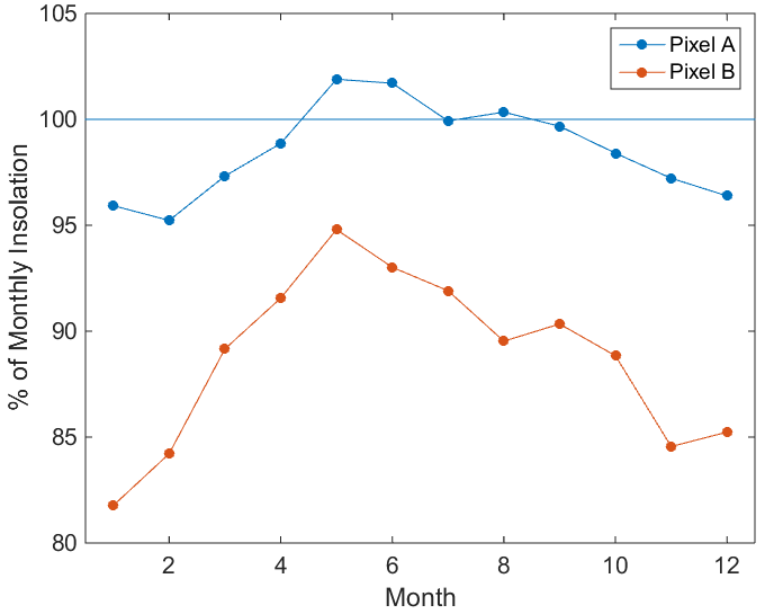



**Figure 24. Monthly insolation ratios of two NSRDB pixels compared with the First Solar Plant 1 ground sensor**

The results of the analysis of the NREL Hawaii and the First Solar sites highlight the potential for significant variation in NSRDB GHI agreement with ground data over the range of a few

pixels. We note that, in both cases, the underlying terrain varied significantly (land/ocean in the case of Hawaii and dry/irrigated land in the case of the First Solar site). These short-range discrepancies might provide an opportunity to improve PSM by focusing on the algorithms that recognize and separate cloud structures from ground.

## 4.4. Systematic Underestimation of GHI in the NSRDB

Figure 25 illustrates an interesting phenomenon that we observed while comparing time series of the NSRDB GHI to ground data at the First Solar Plant 1 site. The investigation was prompted by the horizontal line of data in Figure 4 where the NSRDB GHI is nearly zero but the ground data indicate more than 200 W/m$^2$ of GHI. It is possible that these times indicate when the data fueling the PSM algorithm are faulty, e.g., satellite images are missing or poorly captured. We looked in detail at the time series of NSRDB GHI and ground data to see if systematic patterns were present.

During the month of July 2012, a pronounced spike appears at 10:00 a.m. each day in the time series from pixel C (Figure 21). During the first few days of August, the spikes diminish in magnitude until they completely disappear. Similar spikes occur at the same time of day in July of 2010 and 2011 although to a lesser extent. These spikes appear to be unique to the pixel C time series. Other than the spike evident on 22 July in the top of Figure 25, no similar spike is apparent in the pixel A time series. Because the spikes occur every day at the same time of day for an extended period of days, with no corresponding perturbations in the data from the neighboring pixel, the spikes in the pixel C time series are almost certainly indicative of some kind of systematic failure of the PSM algorithm as applied to pixel C. We believe the spikes may be caused by sunlight reflected from the solar array apparent in Figure 21 that are captured by the GOES satellite and interpreted as a dense cloud. Additional analysis of satellite imagery and the solar geometry would be required to confirm or refute this hypothesis.

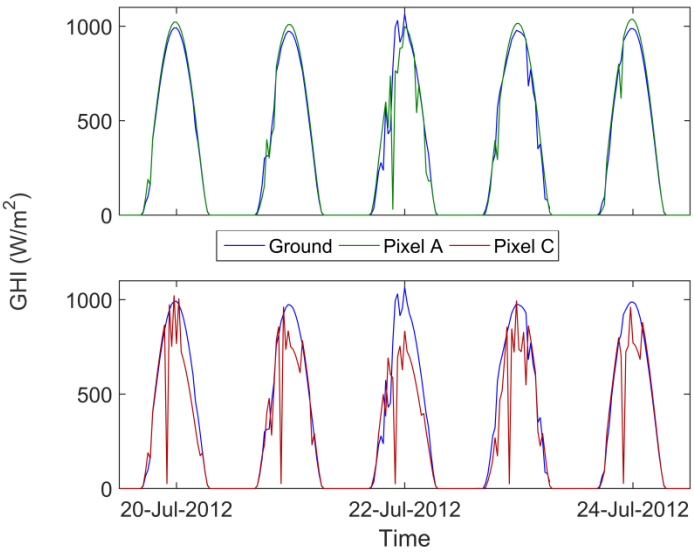



**Figure 25. GHI time series of First Solar Plant 1 and NSRDB pixels A and C**

# 5. CONCLUSIONS

We have compared NSRDB GHI to ground measurements of GHI at a number of locations representing a wide variety of climate classes. Our comparison focuses on ratios of annual and monthly insolation totals because these quantities directly relate to predicted energy from solar power systems.

## 5.1. Conclusions about NSRDB v3 GHI

We find that generally, NSRDB GHI exhibits a bias towards overestimating insolation totals, on the order of 5% when all sky conditions are considered, and somewhat less (~3%) when only clear sky conditions are considered. The biases persist across multiple years and are evident at many locations. In our opinion the bias originates with PSM. We view as less credible the explanation that the bias stems from calibration drift or soiling of ground instruments, because the bias is present at most locations, is persistent across years, and is consistent across a large network of lower quality irradiance instruments (AZMET).

In climates that are subject to snow or ice cover during winter months, we observe that NSRDB GHI may significantly underestimate monthly insolation. We found examples of days where NSRDB GHI apparently misidentified snow cover as clouds, resulting in significant underestimates of GHI during these days leading to substantial understatement of monthly insolation.

Analysis of NSRDB GHI in adjacent pixels shows that the level of agreement between NSRDB GHI and ground data can vary substantially over short distances on the order of 2 km (e.g., Figure 24). Because it is not reasonable that monthly insolation varies by as much as 15% at the location examined we conclude that the variance mostly likely originates with the dramatic contrast in the ground's appearance in the two pixels (Figure 21) where one pixel comprises light brown desert and the other includes irrigated, dark green fields. We observed similar discrepancies (but of less magnitude) among adjacent pixels which contain only land and a mix of land and ocean (Figure 14).

Finally, we observed a small fraction of times (less than 1%) where the NSRDB GHI is nearly zero while the ground data indicate substantial GHI (greater than 200 $W/m^2$). Our investigation did not identify conclusively a cause, but in our opinion the ground data for these times are not wildly incorrect. Possible explanations include faulty satellite data (e.g., missing images) or incorrect operation of the PSM algorithm (e.g., interpreting temporarily bright reflections from the ground as a cloud). We recommend further, detailed investigation of these occurrences because, although rare, the discrepancy is almost certainly with the PSM algorithm, and the very low NSRDB value will affect the accuracy of energy predictions.

## 5.2. Conclusions about comparing NSRDB GHI with ground data

Comparison of NSRDB GHI with ground data necessarily entails equating a spatially averaged, instantaneous observation (NSRDB GHI) with a time series of single point values (ground GHI measurements). These two quantities are inherently different because a time average of the

ground data does not assure that the sun's path across the sky contains clouds at the same proportion as is observed across the spatial extent of the NSRDB pixel containing the ground location. Other authors have investigated the magnitude of this inherent disagreement and the effect of averaging over different time intervals [9], [20]. We explored this issue to some extent but did not arrive at any improvements to methods of comparison used previously.

The comparison of NSRDB GHI with ground data reveals the accuracy of NSRDB GHI only to the extent that the ground data is known to be accurate. We use ground data of varying quality here, and infer evidence about the accuracy of NSRDB GHI from consistency of the comparison across many locations. Other analysts comparing NSRDB GHI to ground data should bear in mind the calibration and cleaning of the instruments involved, and the value of corroborating the comparison across multiple instruments. In our work, we illustrate the comparison across locations to 1) identify uncertainty arising from instrument calibration: e.g., for AZMET data in Figure 5, and for NREL's Oahu data in Figure 15. We also illustrate how the reasonable consistency of NSRDB GHI across years can serve to expose issues with ground data: e.g., with the University of Georgia data in Figure 5.

Finally, we note that comparison between NSRDB GHI and ground data involves an assumption which we could not verify: that both data sources are correctly geo-referenced. If coordinates are mistaken, it is possible that the comparison could be quite misleading, as illustrated by Figure 23 comparing NSRDB GHI for two adjacent pixels to the same ground data.

# 6. REFERENCES


1. National Renewable Energy Laboratory, *National Solar Radiation Data Base User's Manual (1961-1990)*, http://rredc.nrel.gov/solar/pubs/NSRDB/, Accessed July 2015.
2. Wilcox, S., *National Solar Radiation Database 1991-2010 Update: User's Manual*, National Renewable Energy Laboratory, Golden, CO, 2012.
3. *National Solar Radiation Data Base*, https://nsrdb.nrel.gov/, Accessed July 2015.
4. Perez, R., et al., *A new operational model for satellite-derived irradiances: description and validation.* Solar Energy, 2002. **73**(5): p. 307-317.
5. Sengupta, M., et al., *Validation of the National Solar Radiation Database (NSRDB) (2005-2012)*, CP-5D00-64981, National Renewable Energy Laboratory, Golden, CO, 2015.
6. *National Solar Radiation Database (NSRDB)*, https://nsrdb.nrel.gov/, Accessed Sept. 2015.
7. *GOES Surface and Insolation Products*, http://www.ospo.noaa.gov/Products/land/gsip/index_v3.html, Accessed Sept. 2015.
8. Sengupta, M., et al. *A Physics-Based GOES Product for Use in NREL's National Solar Radiation Database*. in *29th European Photovoltaic Solar Energy Conference and Exhibition*, 2014. Amsterdam, The Netherlands.
9. Habte, A., M. Sengupta, and S. Wilcox. *Surface Radiation from GOES: A Physical Approach*. in *27th European Photovoltaic Solar Energy Conference*, 2012. Frankfurt, Germany.
10. Kottek, M., et al., *World Map of the Köppen-Geiger climate classification updated.* Meteorologische Zeitschrift, 2006. **15**(3): p. 259-263.
11. The University of Arizona, *AZMET Weather Data*, http://ag.arizona.edu/AZMET/az-data.htm, Accessed May 2015.
12. Lave, M., et al. *Evaluation of GHI to POA Models at Locations across the United States*. in *40th IEEE Photovoltaic Specialists Conference*, 2014. Denver, CO: IEEE.
13. Sengupta, M. and A. Andreas, *Oahu Solar Measurement Grid (1-Year Archive): 1-Second Solar Irradiance; Oahu, Hawaii (Data)*, NREL Report No. DA-5500-56506, National Renewable Energy Laboratory, 2010.
14. *SURFRAD (Surface Radiation) Network*, http://www.esrl.noaa.gov/gmd/grad/surfrad/, Accessed March 2015.
15. Anderberg, M. and M. Sengupta, *Comparison of Data Quality of NOAA's ISIS and SURFRAD Networks to NREL's SRRL-BMS*, NREL/TP-5D00-62526, National Renewable Energy Laboratory, 2014.
16. Prabha, T. and G. Hoogenboom, *Evaluation of solar irradiance at the surface—inferences from in situ and satellite observations and a mesoscale model.* Theoretical and Applied Climatology, 2010. **102**(3-4): p. 455-469.
17. Reno, M.J., C.W. Hansen, and J.S. Stein, *Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis*, SAND2012-2389, Sandia National Laboratories, Albuquerque, NM, 2012.
18. Reno, M.J. and C.W. Hansen, *Detection of clear sky periods in GHI measurements.* Renewable Energy, 2015: p. (submitted).

19. Perez, R., et al. *Improving The Performance Of Satellite-To-Irradiance Models Using The Satellite's Infrared Sensors*. in *Solar 2010 (American Solar Energy Society)*, 2010. Phoenix, AZ.
20. Zelenka, A., et al., *Effective Accuracy of Satellite-Derived Hourly Irradiances.* Theor. Appl. Climatol., 1999. **62**: p. 199-207.

# DISTRIBUTION

| | | | |
|---|---|---|---|
| 1 | MS0899 | Technical Library | 9536 (electronic copy) |


Sandia National Laboratories