M. J. Reno and C. W. Hansen, "Identification of Periods of Clear Sky Irradiance in Time Series of GHI Measurements," *Renewable Energy*, 2016. Preprint. DOI: 10.1016/j.renene.2015.12.031

Identification of Periods of Clear Sky Irradiance in Time Series of GHI Measurements

Matthew J. Reno * and Clifford W. Hansen

Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185-1033, USA

*Corresponding author. *E-mail address:* mjreno@sandia.gov
TEL.: +1 505 844 3087; Fax: +1 505 844 7231

Abstract

We present a simple algorithm for identifying periods of time with broadband GHI similar to that occurring during clear sky conditions from a time series of global horizontal irradiance (GHI) measurements. Other available methods to identify these periods do so by identifying periods with clear sky conditions using additional measurements, such as direct or diffuse irradiance. Our algorithm compares characteristics of the time series of measured GHI with the output of a clear sky model without requiring additional measurements. We validate our algorithm using data from several locations by comparing our results with those obtained from a clear sky detection algorithm, and with satellite and ground-based sky imagery.

Keywords: Solar resource, global horizontal irradiance, clear sky

1. Introduction

We present an algorithm to identify periods of time with broadband irradiance similar to that occurring during clear sky conditions in a time series of measured global horizontal irradiance (GHI). We developed this algorithm to support automated performance evaluation of photovoltaic (PV) systems, which are typically monitored by recording the power produced concurrently with global horizontal irradiance, temperature and wind measurements. Examining system performance during periods of time with irradiance similar to clear sky conditions facilitates detection of system degradation and faults.

A clear sky condition is defined generally as the absence of visible clouds across the entire sky dome, and clear sky irradiance is the irradiance (e.g., GHI) occurring during these conditions. A more precise definition of clear sky irradiance relies to some extent on judgment. Even in the absence of visible clouds, aerosols, precipitable water, and other atmospheric conditions can vary in time and space and affect the transmittance of solar radiation through the atmosphere. Accordingly, there are several alternatives for precise definitions of the clear sky condition and clear sky irradiance (Younes and Muneer, 2007). Methods are available which can identify clear sky conditions from ground measurements; we survey these methods in the Background section. These methods are intended to identify only periods with clear sky irradiance (distinguishing periods with irradiance similar to clear sky irradiance) and often rely on measurements such as diffuse horizontal irradiance (DHI) which are not typically available in conjunction with PV system monitoring data. The proposed method is valuable to identify both clear sky periods and periods with irradiance similar to clear sky irradiance relying only on GHI measurements.

We use the term 'clear sky equivalent GHI' to refer to either GHI during clear sky conditions or GHI during all-sky conditions that is similar to clear sky irradiance. When clouds are present in the sky but are sufficiently separated from the direct path between the sun and a point of interest on the ground, the GHI observed at that point may be indistinguishable from what the GHI would be under clear skies. For our purpose both 'clear sky irradiance' and 'irradiance similar to clear sky irradiance' are meaningful because a PV system will produce similar levels of power under either condition. When monitoring a PV system, it is desirable to make assessments in as timely a manner as possible to avoid delays in identifying and

correcting system faults. For this reason, we wish to identify both time periods. Throughout the paper, we use the terms 'period with clear sky equivalent irradiance' and 'clear sky period' to refer to a period of time with the previously defined 'clear sky equivalent GHI', to avoid cumbersome repetition of the distinction between these irradiance terms.

Our algorithm identifies periods with clear sky equivalent GHI by comparing characteristics of the time series of GHI measurements with corresponding output from a clear sky model. The algorithm is designed for high time resolution irradiance data, and all results presented here are for 1-minute resolution data. The algorithm is not sensitive to the choice of clear sky model (many are available) because an iterative adaptive method can be used to reconcile the clear sky model to the measured GHI. Our method could also be applied to a time series of measured plane-of-array (POA) irradiance if partnered with an appropriate model for clear sky POA irradiance.

2. Background

2.1 Previous Work in Identifying Clear Sky Conditions

Many clear sky detection methods rely on measured irradiance quantities. Younes and Muneer presented an overview of several clear sky detection algorithms and evaluated the accuracy of nine methods (Younes and Muneer, 2007). The simplest method for identifying clear sky conditions is to compare the clearness index k_t, defined as the ratio of global horizontal irradiance to the extraterrestrial irradiance on a horizontal plane, with a threshold value. Authors have used different thresholds for the value that defines clear sky: 0.6 (Muneer et al., 2000), 0.65 (Cucumo et al., 2008; Cucumo et al., 2010), or 0.7 (Lam and Li, 1996; Li and Lam, 2001). The ratio of diffuse to global solar radiation K can also be used to define clear skies (Rahim et al., 2004), with a proposed value of less than 0.27 being clear (Lam and Li, 1996; Li and Lam, 2001). The ratio of zenith luminance to horizontal diffuse illuminance can also be used to determine clear skies (Li and Tang, 2008). Other authors have proposed using more refined measures of sky clearness; for example, Perez et al. (Perez et al., 1990) have proposed a formulation of sky clearness ε using the diffuse irradiance I_D and beam irradiance I_B that takes into account the zenith angle z.

$$\varepsilon = \frac{\left(I_D + I_B/\cos z\right) I_D + 1.041 \times z^3}{1 + 1.041 \times z^3}. \quad (1)$$

Perez et al. (Perez et al., 1990) suggested a value $\varepsilon = 6.2$ as a definition of a clear sky.

In addition to distinguishing clear and non-clear periods, some authors have also classified cloud type using measured GHI and diffuse. Calbó, Gonzalez, and Pages used clearness index, diffuse fraction, and the short-term variability of GHI and diffuse to classify the cloudiness of the sky into categories using supervised classification techniques (Calbó et al., 2001; Pages et al., 2003). Harrison, Chalmers, and Hogan used GHI and diffuse to determine cloud amount and discriminate between stratiform and convective cloud types (Harrison et al., 2008). DeFelice and Wylie use a four-band, ground-based, sun photometer to detect and classify clouds (DeFelice and Wylie, 2001). Other methods use the variability in a time-series of irradiance measurements to classify the period as clear using metrics such as variability index (Stein et al., 2012) or probability of persistence (Kang and Tam, 2013). Duchon and O'Malley used the mean clearness index and standard deviation of irradiance over a 21 minute window to categorize seven types of clouds (Duchon and O'Malley, 1999).

Long and Ackerman (Long and Ackerman, 2000) outline a detailed approach comprising four tests which use GHI and diffuse irradiance that together detect all cloud scenarios. The tests are: (1) the normalized total shortwave (i.e., GHI) limits test; (2) the maximum diffuse shortwave test; (3) the change in shortwave magnitude with time test; and (4) the normalized diffuse ratio variability test. Their method is designed for high frequency (sub-15 minute resolution) data and uses an iterative method for

determining threshold values. Their method has become a de facto standard in the solar measurement community, for example, for data quality control.

Clear sky conditions can also be defined using instrumentation other than irradiance measurements. Markou classified the sky condition based on sky luminance scan data (Markou et al., 2007). Dupont, Haeffelin, and Long compared the results of cloud detection using shortwave and longwave radiation and Lidar backscatter measurements (Dupont et al., 2008). Hogan, Jakov, and Illingworth estimated cloud cover using ground-based radar (Hogan et al., 2001). Marty and Philipona use longwave downwelling radiation along with air temperature and humidity to detect clear skies during the day and night (Marty and Philipona, 2000). Orsini et al. use downwelling shortwave radiation and downwelling longwave radiation to detect cloud type, cloudiness, and cloud height (Orsini et al., 2002). More recently, sky imagers have been used to detect clouds and classify their opacity (Ghonima et al., 2012).

2.2 Clear Sky Models

The algorithm proposed in this paper is dependent on the use of a clear sky model, which estimates the irradiance at the earth's surface at a given location and time under clear sky conditions. The clearness index k_t can be directly calculated from the extraterrestrial radiation, but clear sky irradiance is also dependent on atmospheric parameters. Much research has gone into creating models for calculating surface irradiance under a clear sky. While none of these clear sky models are perfect, they provide a much more accurate representation of surface irradiance during clear skies than does extraterrestrial horizontal irradiance.

Here, we use the clear sky model proposed by Ineichen (Ineichen and Perez, 2002), which is based on the Kasten model (Kasten, 1980) and has been shown to be generally accurate and fairly easy to implement (Badescu et al., 2012; Gueymard, 2012; Ineichen, 2006; Reno et al., 2012). The inputs to this model are normal incidence extraterrestrial irradiance I_0, solar zenith z (radians), air mass AM (unitless), Linke turbidity factor TL (unitless), and elevation h (m):

$$\mathrm{GHI} = c_{g1} \times I_0 \times \cos(z) \times \exp(-c_{g2} \times \mathrm{AM} \times (f_{h1} + f_{h2}(\mathrm{TL}-1))) \times \exp(0.01 \times \mathrm{AM}^{1.8}); \quad (2)$$

$$\text{with } f_{h1} = \exp(-h/8000) \text{ and } f_{h2} = \exp(-h/1250), \quad (3)$$

$$\text{where } c_{g1} = 5.09\text{e-}5 \times h + 0.868 \text{ and } c_{g2} = 3.92\text{e-}5 \times h + 0.0387. \quad (4)$$

Kasten and Young's formula (5) is used to calculate the air mass (Kasten and Young, 1989):

$$\mathrm{AM} = \frac{1}{\cos(z) + 0.50572(96.07995 - z)^{-1.6364}} \quad (5)$$

The Linke turbidity factor, proposed by Linke in 1922, is a measure of the absorption by the water vapor and the absorption and scattering by the aerosol particles compared to a dry and clean atmosphere (Linke, 1922). Ineichen proposed an air mass independent formulation for the Linke turbidity (Ineichen and Perez, 2002) that has the advantages of being solar altitude independent and matches the original Linke turbidity factor at air mass 2. Remund calculated and produced Linke turbidity maps for the world for each month using a combination of ground measurement and satellite data (Remund J., 2003). Monthly images of Linke turbidity for the world that can be downloaded from either the HelioClim website (HelioClim, 2013) or Solar Radiation Data (SoDa) website (SoDa, 2011).

As long as the clear sky model is reasonably accurate, we do not believe that choice of the specific model is critical. Our analysis has shown that the largest errors in a clear sky model for a given location result from improper scaling rather than from an incorrect shape of the clear sky irradiance profile (Reno et al., 2012). Our algorithm includes an iterative method for adjusting the scaling of the selected clear sky model to adapt the chosen clear sky model to local data (Section 4).

3. Algorithm

Intuitively, a plot of GHI vs. time during clear sky conditions is a smoothly-varying curve, with values that are a relatively constant fraction of extraterrestrial irradiance on a plane parallel to the earth's surface. Our proposed method for identifying periods with clear sky equivalent GHI calculates five quantities that characterize the smoothness, shape, and magnitude of the GHI vs. time curve and compares these quantities with those derived from a clear sky model. Unlike many of the methods previously discussed in Section 2.1, our method only requires GHI.

The method can be applied to an entire day or to shorter time periods. Because the method examines the shape of the measured irradiance profile, it classifies time periods as either clear or cloudy. However, by using a sliding time window individual measurements can be classified as clear or cloudy (Section 3.3).

Our algorithm uses five criteria to classify a period of time as clear, i.e., with GHI equivalent to GHI during clear sky conditions:

1. Mean value of GHI;
2. Maximum value of GHI;
3. Line length of irradiance vs. time curve;
4. Standard deviation of rate of change in GHI;
5. Maximum difference between changes in GHI and clear sky time series.

The time series of GHI and corresponding clear sky model output are divided into intervals each containing *n* values, and each of the five quantities are calculated for each time period. In the discussion that follows, values from the clear sky model are indicated by a subscript $\sim_{CS}$. Then, for each time period the calculated quantities are compared with threshold values, which depend on *n*, to classify the time period as clear or not. An application of the method to 10-minute periods and corresponding threshold values is discussed in Section 3.2.

3.1 Criteria

Mean value of GHI. The first criterion examines the mean value of irradiance $\bar{G}$ during a time period. Specifically, we compute

$$\bar{G} = \frac{1}{n}\sum_{i=1}^{n} GHI_i \tag{6}$$

As illustrated in Figure 1, the mean irradiance $\bar{G}$ will be significantly lower for cloudy conditions than under a clear sky. The mean value of GHI criterion is met if $\bar{G}$ is close enough to $\bar{G}_{CS}$.

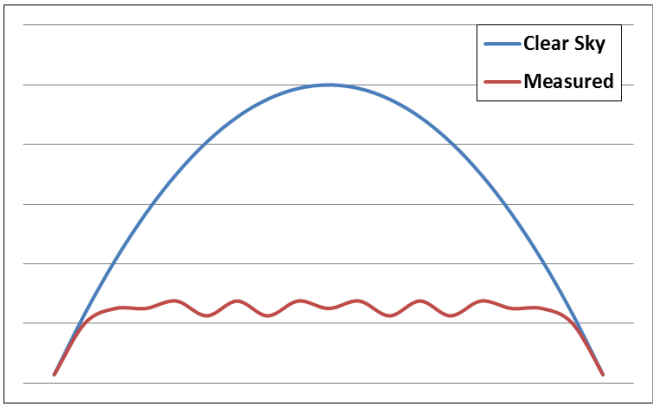

Figure 1. Motivation for mean GHI criterion.

Maximum value of GHI. When the sky is partially cloudy during the time period, the average measured irradiance could be reasonably close to the average irradiance from the clear sky model even though GHI varies above and below the clear sky values. Figure 2 illustrates examples where the mean value of the measured time series is the same as the mean value of the clear sky time series. Figure 2a is an example where the sky gradually and uniformly brightens, and Figure 2b shows a more realistic case where cloud enhancement originally increases the irradiance before the cloud occludes the sun. The simplest way to distinguish these conditions from clear sky conditions is to also consider the maximum irradiance value M in the time period:

$$M = \max[GHI_i] \quad \forall i \in \{1, 2, ..., n\} \tag{7}$$

The second criterion is satisfied if M is close enough to M_{CS}.

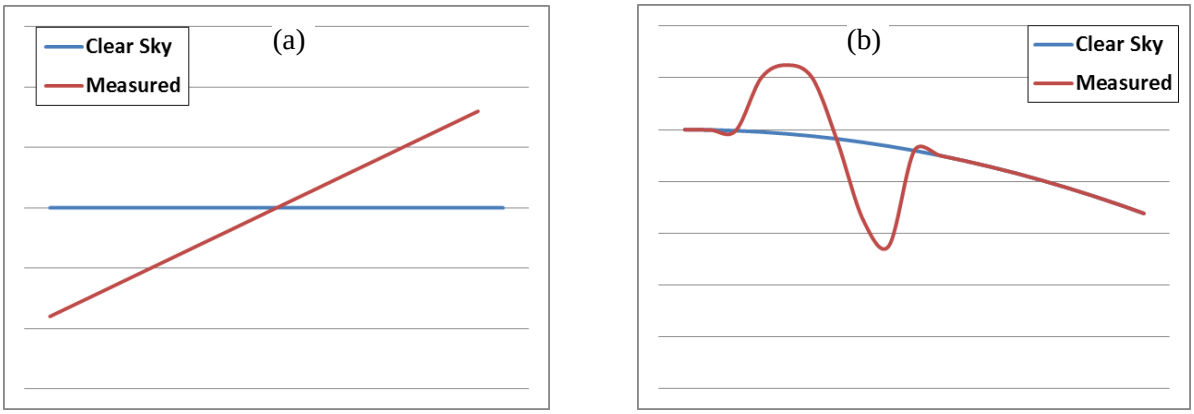

Figure 2. Motivation for maximum value criterion.

By themselves, mean and maximum GHI are not sufficient to distinguish clear from cloudy periods. Figure 3 shows two examples of time series with variability but with the same mean and maximum values: Figure 3a represents GHI during passage of a small thick cumulus cloud, and Figure 3b represents the effects on GHI of persistent high wispy cirrus clouds. Accordingly, we use three additional criteria intended to distinguish a variable time series, such as would result if clouds were present, from a smoothly changing time series. Variability in GHI resulting from the effects of clouds, both large and small, will change the shape of the GHI curve relative to a clear sky profile. However, because cloud size, transparency and speed vary, the effects on the time series of GHI also vary. The third, fourth and fifth criteria quantify different aspects of variability to distinguish clear sky periods from periods with different types of clouds.

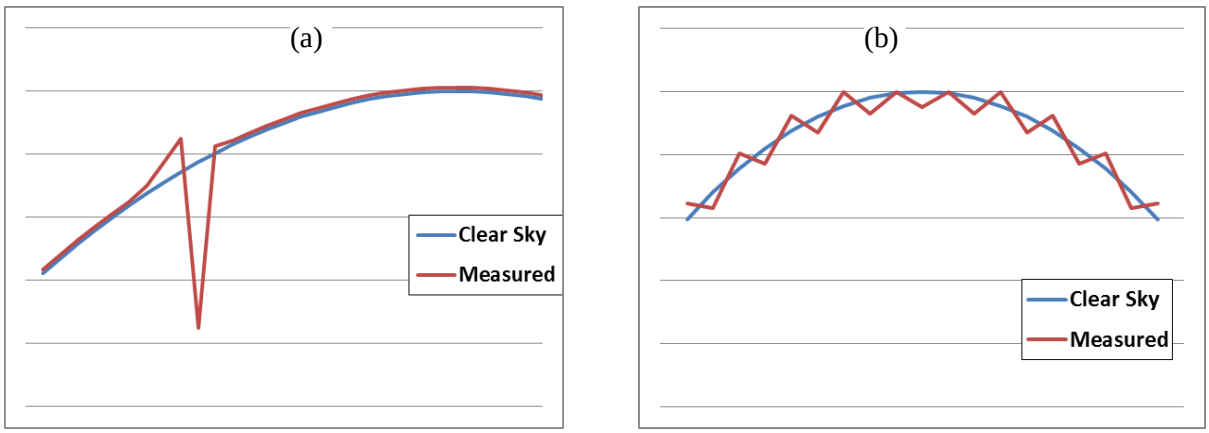

Figure 3. Variable GHI time series with same mean and maximum values.

Line length. The length L of the sequence of line segments connecting the points of the GHI time series is calculated by Eq. (8):

$$L = \sum_{i=1}^{n-1} \sqrt{\left(GHI_{i+1} - GHI_i\right)^2 + \left(t_{i+1} - t_i\right)^2} \tag{8}$$

Any variability in the measured GHI will increase L relative to the line length of the clear sky model L_{CS}. Line length is similar to the "variability index" unit proposed to quantify variability in (Stein et al., 2012). The third criterion is met if L is near enough to L_{CS}. However, the line length for a time series with a few large changes, or with many small changes, may not be significantly greater than that of a smooth curve when a long interval of time is considered. Consequently, we add two additional criteria.

Standard deviation of rate of change in irradiance. We calculate the standard deviation σ of the slope (s) between sequential points in the time series, normalized by the average GHI during the time interval:

$$s_i = \frac{GHI_{i+1} - GHI_i}{t_{i+1} - t_i}, \quad \forall i \in \{1, 2, ..., n-1\} \tag{9}$$

$$\bar{s} = \frac{1}{n-1} \sum_{i=1}^{n-1} s_i \tag{10}$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} \left(s_i - \bar{s}\right)^2} \Big/ \bar{G} \tag{11}$$

Although σ cannot easily distinguish between situations such as Figure 3b where the slope between measurements varies slightly, it readily distinguishes between the two time series illustrated in Figure 4 which both have the same line length. Because we expect σ_{cs} to be small relative to the value of σ during a cloudy period, the fourth criterion is satisfied when σ is less than a threshold value.

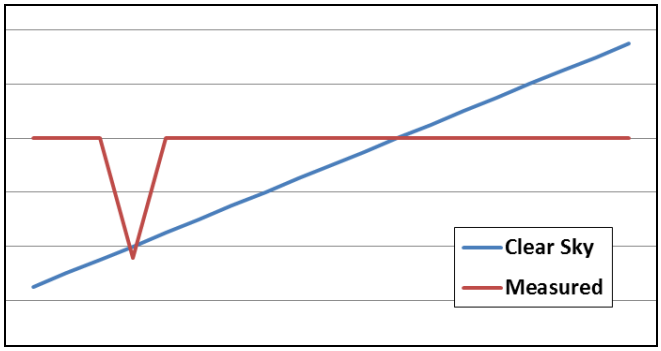

Figure 4. Motivation for the standard deviation of slope criterion.

Maximum difference between changes in GHI and clear sky time series. The final quantity considered is the maximum difference X between the change in measured irradiance and the change in clear sky irradiance over each measurement interval:

$$x_i = GHI_{i+1} - GHI_i, \quad \forall i \in \{1, 2, ..., n-1\} \tag{12}$$

$$x_{i,CS} = GHI_{i+1,CS} - GHI_{i,CS} \tag{13}$$

$$\begin{aligned} X &= \max\left\{\left|x_i - x_{i,CS}\right|\right\} \\ &= \max\left\{\left|\left(GHI_{i+1} - GHI_{i+1,CS}\right) - \left(GHI_i - GHI_{i,CS}\right)\right|\right\} \quad \forall i \in \{1, 2, ..., n\} \end{aligned} \tag{14}$$

During a clear sky period, the change between successive measurements should be similar to the change in irradiance predicted by the clear sky model; the fifth criterion is met when X is less than a threshold value. Figure 5 shows an example of measured GHI during and after a cloud passes over an irradiance sensor. At measurement 1, the cloud is between the sun and the sensor, but by measurement 2, the cloud's shadow has moved away from the sensor. Because the change of measured GHI between observations 1 and 2 is much larger than the change in the corresponding clear GHI values, the fifth criterion identifies the period including measurement 1 as cloudy. This last criterion also ensures that successive periods with cloud shadows occurring on the period boundaries are classified as cloudy rather than clear. The change between measurements should not vary more than the measurement error for the given sensor.

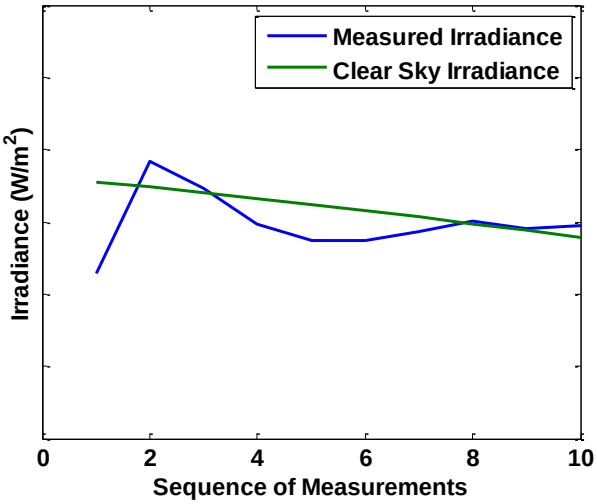

Figure 5. Motivation for maximum difference between change in GHI and clear sky irradiance.

3.2 Threshold Values for Criteria

To apply the algorithm, a threshold value must be established for each criterion. For a 10-minute window, we combined evaluations of clear sky models with analysis of measured GHI to determine appropriate threshold values for each criterion. If a different length time period was chosen, the thresholds for each criterion would need to be adjusted.

For the mean and maximum GHI criteria, a threshold of ±75 W/m^2 within the mean and maximum of the clear sky model was chosen. Most evaluations of clear sky models find that the average bias error of the model is less than 10%, often around 7% (Alam, 2006; Badescu, 1997). Setting the threshold to be similar to these bias errors avoids misclassifying a clear time period as cloudy due to bias error in the clear sky model.

To determine threshold values for the remaining criteria, measured GHI for 2011 from NREL's Solar Radiation Research Laboratory (SRRL) at Golden, CO, was analyzed (NREL, 2013). We visually identified clear and cloudy periods by comparing the shape of the GHI vs. time curve and its magnitude to a simple clear sky model. Each quantity is calculated for both clear and cloudy periods, and thresholds are selected that distinguish between clear and cloudy conditions.

The calculated line length for the 10-minute moving window on a clear day is illustrated in Figure 6. Using the data shown in Figure 6 and data for other clear days (not shown), thresholds of within +10 (calculated with GHI in W/m^2 and time in minutes) and −5 of the clear sky line length were chosen. Figure 6 shows how the line length increases when variability from clouds is present.

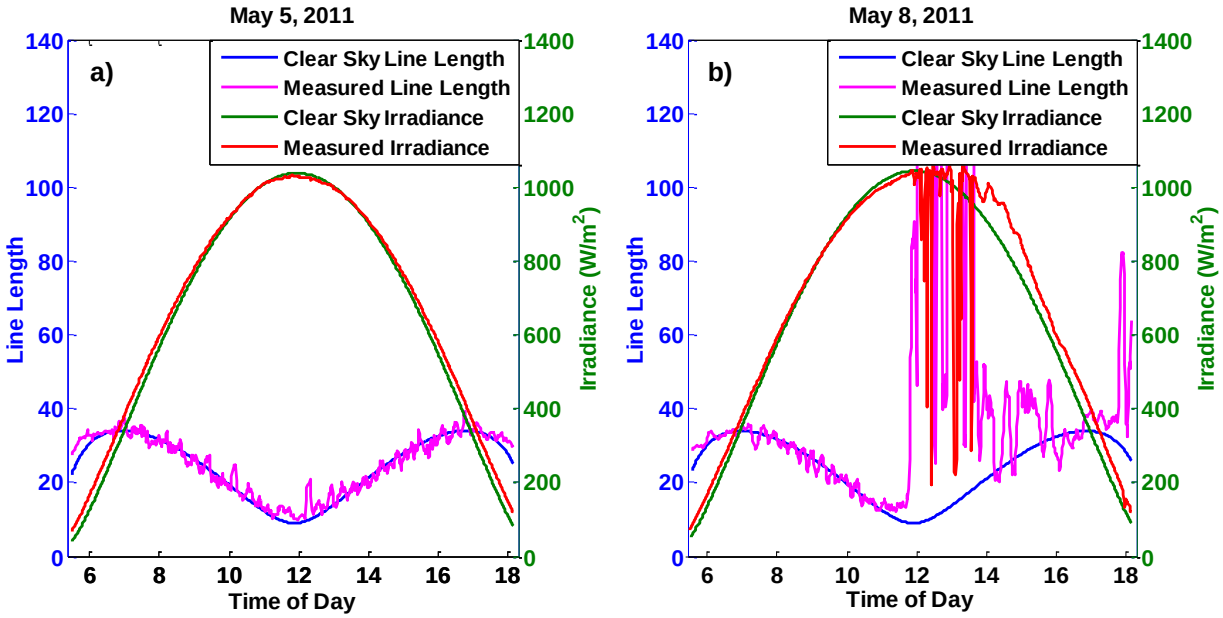

Figure 6. Line length of GHI (W/m^2) in a 10-minute moving window on a clear day (a) and a cloudy day (b) in Golden, CO.

Figure 7 shows that on a clear day the standard deviation of changes in GHI measurements in the 10-minute moving window is consistently low throughout the day. Based on these data, a fixed value of 0.005 is chosen as the threshold for the standard deviation of rate of change in irradiance criterion. Under cloudy conditions as shown in Figure 7, the standard deviation of changes in GHI clearly exceeds the threshold.

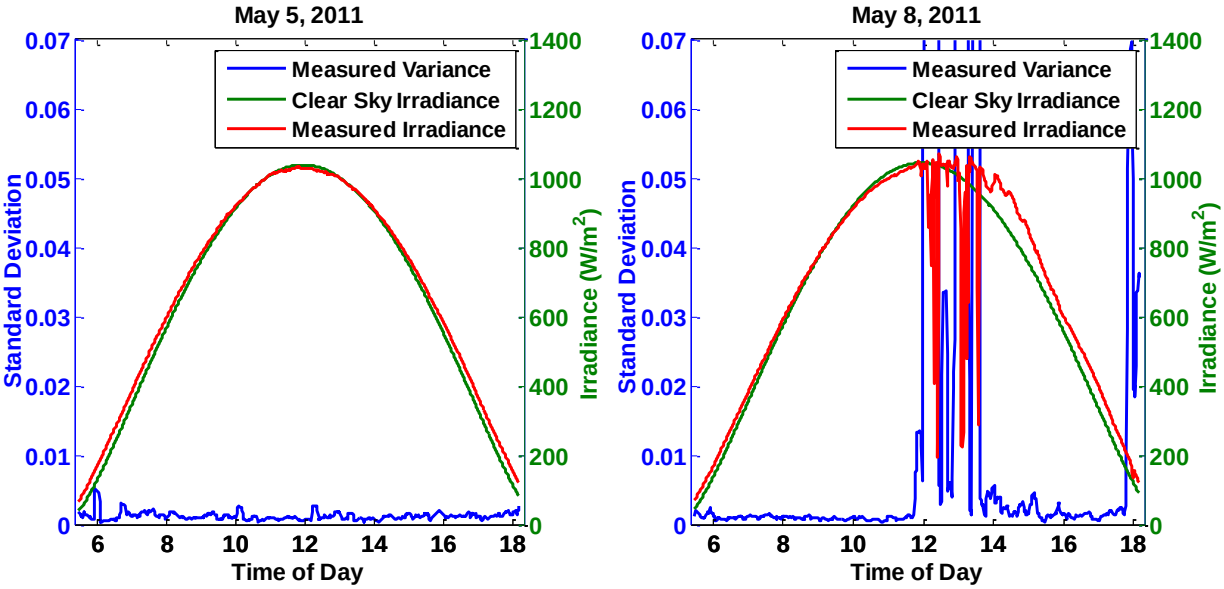

Figure 7. Standard deviation of the changes in GHI in a 10-minute moving window on a clear day (a) and a cloudy day (b) in Golden, CO.

The last threshold is for the maximum difference between change in consecutive GHI measurements and the change in the clear sky time series. We select a threshold of 8 W/m^2 per time period for the fifth criterion, using the clear periods we visually identified in the SRRL data. A certain degree of measurement error is present in the GHI data. For most GHI sensors, both thermopiles and silicon-based sensors, the sensor error is about 5% and at best it is 2% (Gueymard and Myers, 2009; Reda, 1998). Because we observed a significant level of autocorrelation in the time series of differences between measured GHI and GHI from the clear sky model, we set the threshold value much less than these estimates of sensor error.

3.3 Moving Window Method for Classifying Individual Measurements

Our method identifies periods of time as clear or cloudy by examining the overall shape of the GHI vs. time curve. However individual measurements may also be classified by using a sliding window of time. We illustrate this application using a moving 10-minute window with irradiance measurements at 1-minute intervals. Each GHI time-series window that meets all criteria is flagged as clear. Individual measurements will be evaluated several times as the window passes, and if an individual measurement is within at least one window declared as clear, the measurement is identified as clear because all 10 measurements in that window must be clear.

Classifying individual measurements has the benefit of not classifying an entire day or time period as cloudy due to temporary shading from a building, a pole, or power lines. This approach is illustrated using measured GHI data for two clear days at the Nevada Power Clark Station in Las Vegas, Nevada (NREL, 2013), which is subject to intermittent shading by nearby power lines and utility poles. Figure 8a shows shadows in the morning and evening as the sun crosses nearby power lines; Figure 8b shows a longer period in the evening when the sun crosses behind a nearby utility pole. The figures show that the rest of each day has been correctly classified as clear (marked with red x's).

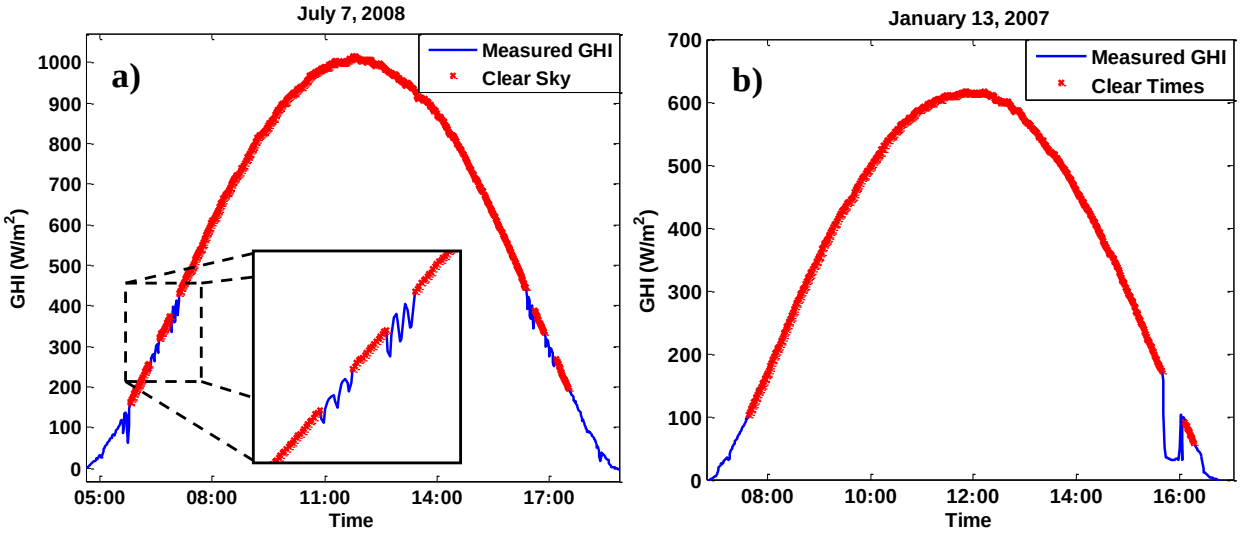

Figure 8. Measured GHI at Clark Station showing a) shadows from nearby power lines and b) the effect of a pole.

4. Iterative Process for Removing Bias in Measurements or in the Clear Sky Model

Detection of clear times depends on the clear sky model to which the measured data are compared. An inaccurate clear sky model with a large bias could cause clear periods to be misidentified as cloudy, by failing to meet the criteria for mean or maximum irradiance. To address this problem, we propose an iterative method to locally calibrate the clear sky model, as needed, to available measurements.

Iteration (Figure 9) comprises detecting clear periods, fitting the clear sky model to the measurements during clear sky periods, and then detecting additional clear periods using the improved clear sky model. Fitting of the clear sky model depends on variables and parameters considered in the clear sky model itself. One simple approach would be to estimate a scaling factor (α) by minimizing the RMSE (Eq. (15)) between measured (GHI) and clear sky irradiance (CSI). More refined calibration may be achieved by fitting a scaling factor which depends on zenith angle, time of year, or other exogenous quantities. In our validation results, we calibrated a simple scaling factor for each site using only clear periods on days when greater than 50% of the day was classified as clear.

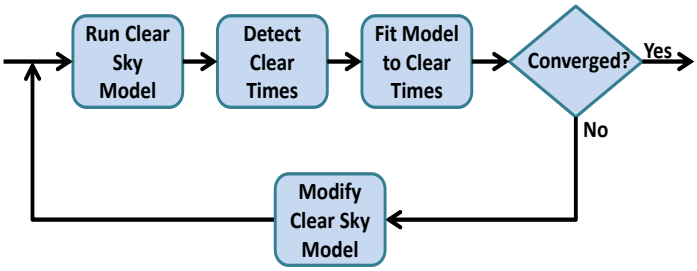

Figure 9. Flow chart for adaptive clear time detection algorithm

$$\text{Minimize } f(\alpha) = \sqrt{\frac{\sum_{i=1}^{n} (\alpha \times GHI_i - CSI_i)^2}{n}} \qquad (15)$$

subject to $\alpha > 0$

5. Validation

We validated our algorithm using several approaches, including visual confirmation, comparison with Long and Ackerman's method (Long and Ackerman, 2000), comparison with satellite imagery and with ground-based sky imagers.

5.1 Visual Verification

First, we verified the detection algorithm by visually examining one month of measured GHI in Las Vegas, NV; results are shown in Figure 10 and Figure 11. Measured GHI was obtained from the Las Vegas Valley Water District at the Fort Apache facility. Red markers indicate minutes classified as clear, using a sliding 10-minute window and the threshold values determined from the SRRL data as described earlier. For clear sky conditions where the sun is not obstructed by cloud, each day should look similar to April 10 in Figure 10 with the smooth diurnal parabola shape. For example, on April 8th the morning and middle part of the afternoon show large variability and decreases in measured irradiance typical of clouds blocking the sun. However the same date also includes several hours of clear sky in the early and late afternoon. Even on very cloudy days like April 15th, there can still be brief periods where there is a clear path to the sun. Conversely, during mostly clear days like April 18th clouds can block the sun occasionally.

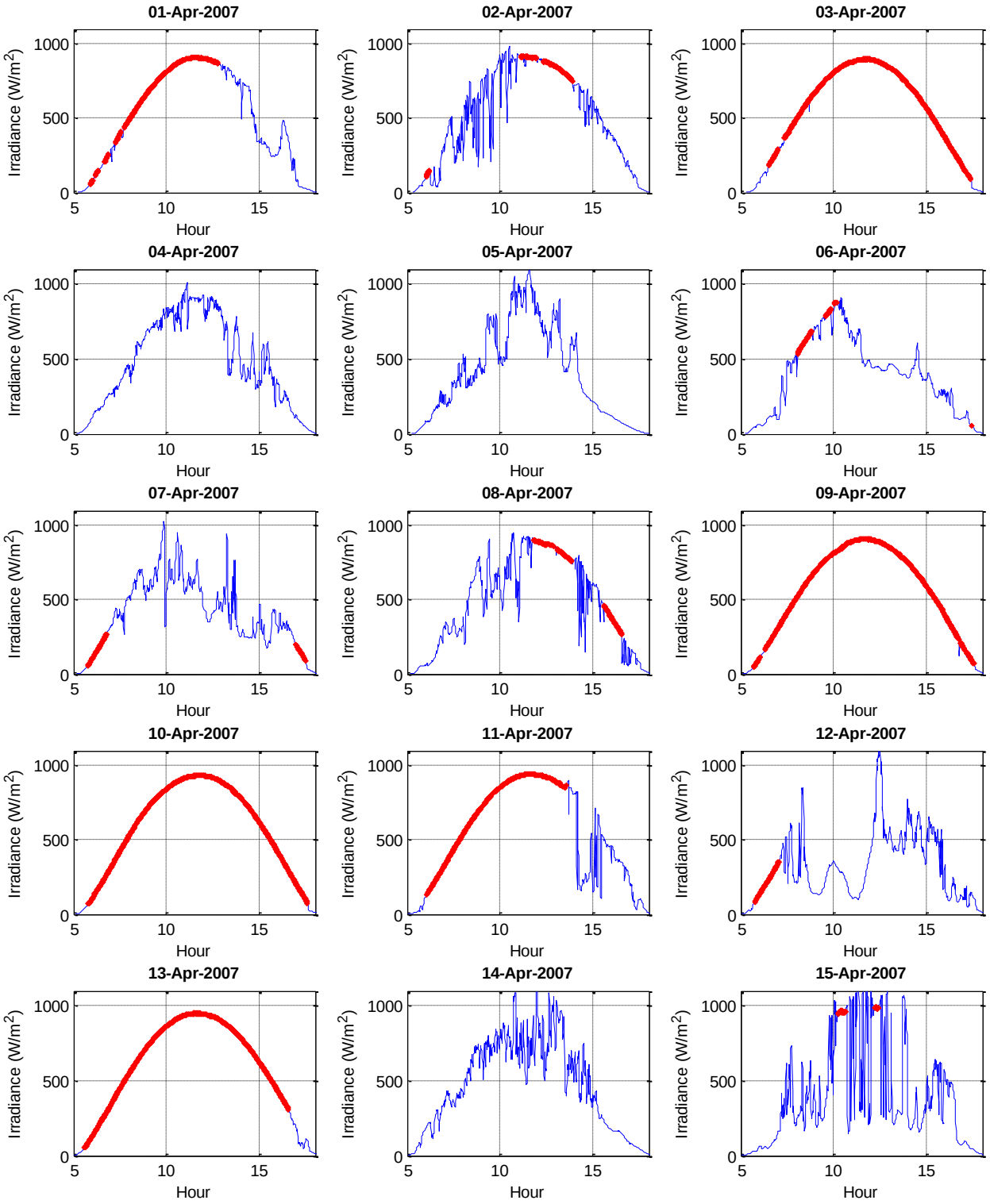

Figure 10. Visual representation of clear sky detection for Las Vegas, NV. Measured irradiance is in blue, with red markers signifying minutes identified as clear.

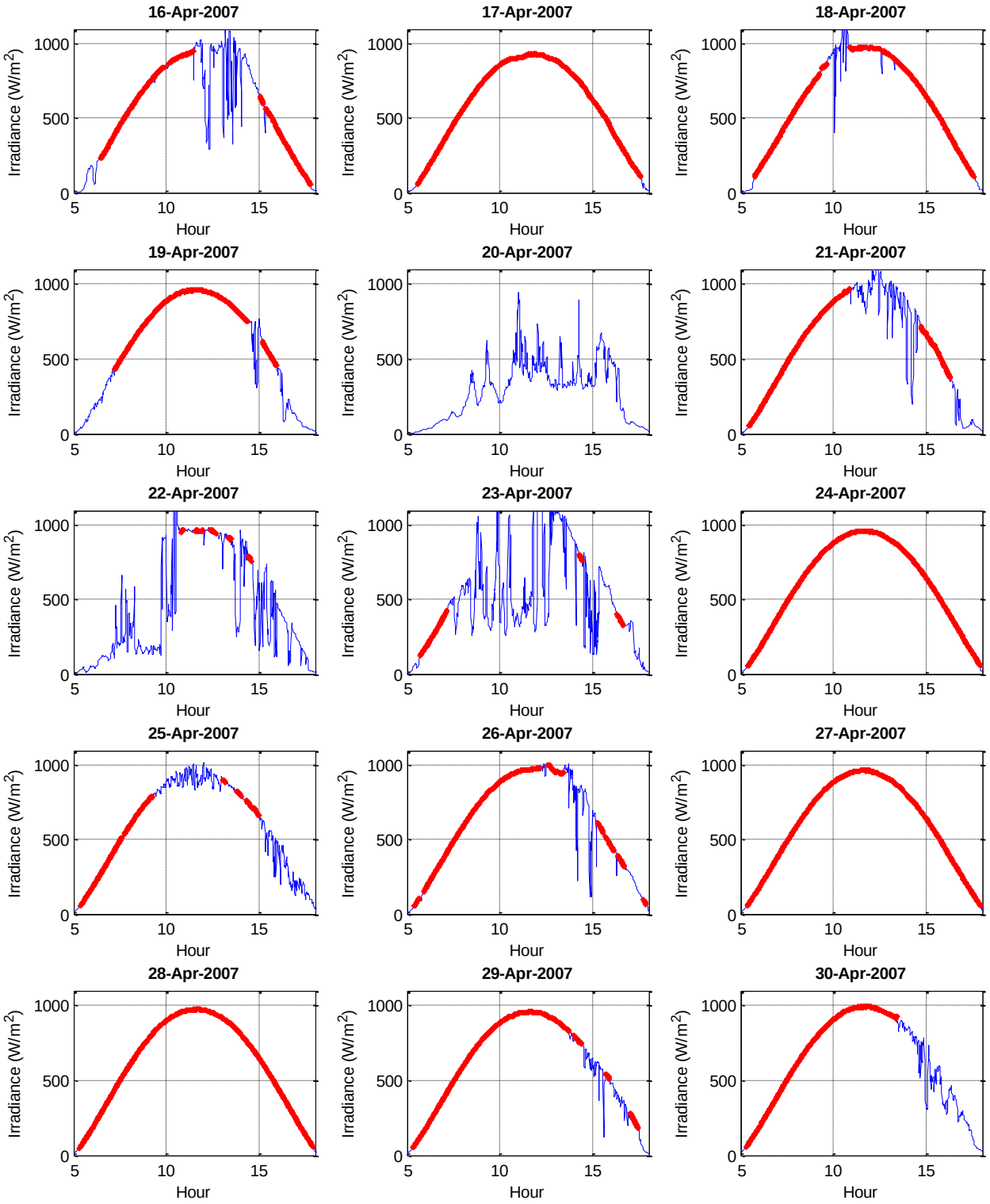

Figure 11. Visual representation of clear sky detection for Las Vegas, NV. Measured irradiance is in blue, with red markers signifying instants identified as clear.

5.2 Comparison with Long and Ackerman method

We compared the results of the clear sky detection method presented here with the results of the algorithm proposed by Long and Ackerman (Long and Ackerman, 2000). The Long and Ackerman algorithm uses GHI, DNI and diffuse measurements to detect clear sky conditions within a 160° field of view, although as applied by its authors the algorithm classifies each point in time as clear or cloudy (as does our algorithm). We used data from the MIDC (NREL, 2013) Anatolia site in Rancho Cordova, California, for the year 2011 at 1-minute resolution with both GHI and diffuse irradiance measurements available from a LICOR LI-200 Pyranometer mounted on a Rotating Shadowband Radiometer (RSR). The 10-minute moving window method and the threshold values presented in Section 3 were used for the analysis. We used the Ineichen (Ineichen and Perez, 2002) clear sky model and the iterative process described in Section 4.

The Long and Ackerman algorithm has four tests and contains many constants that need to be calibrated for the specific site and measurement sensors. The first test is the normalized total shortwave magnitude test. The constant *b* depends on both site and instrument, and for Anatolia it was found to be equal to 1.18 (compared to Long and Ackerman's value of 1.31 for the ARM Southern Great Plains Central Facility). The second test is the maximum diffuse shortwave test and all constants were used as published in (Long and Ackerman, 2000). The third test is the change in magnitude with time test. This formula contains a constant *C* that is dependent on the intrinsic sensor and measurement noise, and it was tuned to be equal to 8. The fourth test is the normalized diffuse ratio and variability test. It also contains the site dependent *b* constant along with a subjective threshold limit based on sensor noise. Long and Ackerman used a threshold of 0.0012 for the ARM Southern Great Plains Central Facility, and a threshold of 0.003 was found to be appropriate for Anatolia. Other than these noted differences, the Long and Ackerman algorithm was implemented exactly as described in (Long and Ackerman, 2000).

The proposed method and the Long and Ackerman method are applied to classify each minute in 2011 as clear or cloudy. This application uses the Long and Ackerman method to identify not only periods with clear sky conditions but also periods with GHI similar to clear sky values. The results are shown in Table 1 as a percent of the total daytime minutes; results for four days are illustrated in Figure 12. The two algorithms agreed for more than 94% of times, and of the clear times identified by the Long and Ackerman algorithm, 96.08% were also identified as clear by our algorithm. For the entire year, the correlation coefficient between indicator variables distinguishing clear from cloudy times for the two methods is 0.899. In general the disagreements between the two algorithms occur on partially cloudy days. For example, a cloud near the sun may not affect the GHI measurement significantly, but the Long and Ackerman algorithm may classify this time as cloudy because the diffuse ratio has changed. Also, the proposed algorithm will declare the entirety of highly variable periods as cloudy, whereas the Long and Ackerman algorithm can identify single minutes in the period as clear if the ratios and criteria are correct.

Table 1. Percent of 1-minute daytime measurements identified by each algorithm as cloudy or clear.

		Long and Ackerman	
		Clear	**Cloudy**
Reno and Hansen	**Clear**	51.1345 %	3.1914 %
	Cloudy	2.0868 %	43.5874 %

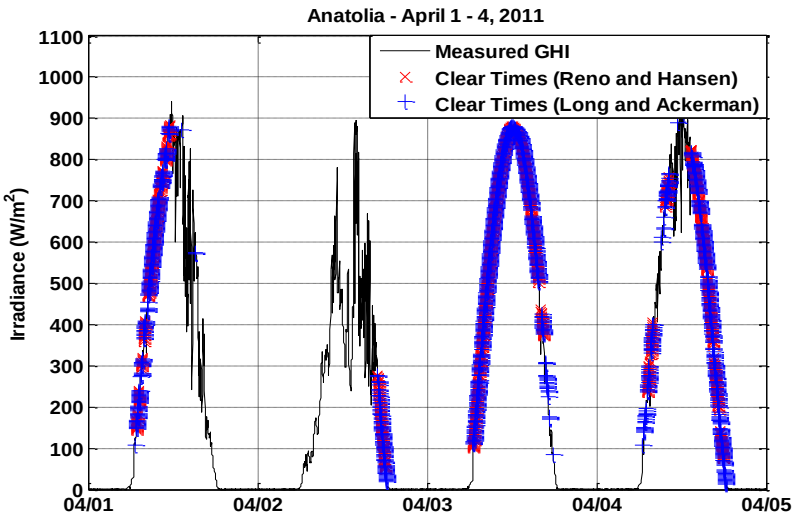

Figure 12. Comparison of our algorithm with Long and Ackerman (Long and Ackerman, 2000). Measured data for Rancho Cordova, CA.

While the proposed algorithm matches the clear/cloudy identification from Long and Ackerman more than 94% of the time, one key difference occurs during short intervals. Due to the 10-minute moving window, the proposed algorithm identifies a single minute as clear depending on the values at adjacent times. In contrast, the Long and Ackerman algorithm regularly identifies a single minute as clear because the ratios fall within the correct range. One instance when the Long and Ackerman method identifies a clear minute during a variable period is clearly evident on the afternoon of April 1st in Figure 12. To judge how often these disagreements occur, we counted clear segment lengths, i.e., the number of minutes in a row that are identified as clear. The histogram of segment lengths for each algorithm is shown in Figure 13. We observe general similarity in the two distributions, although Long and Ackerman identifies roughly 25% more short (3 minutes or less) segments as clear than are identified as clear by our method.

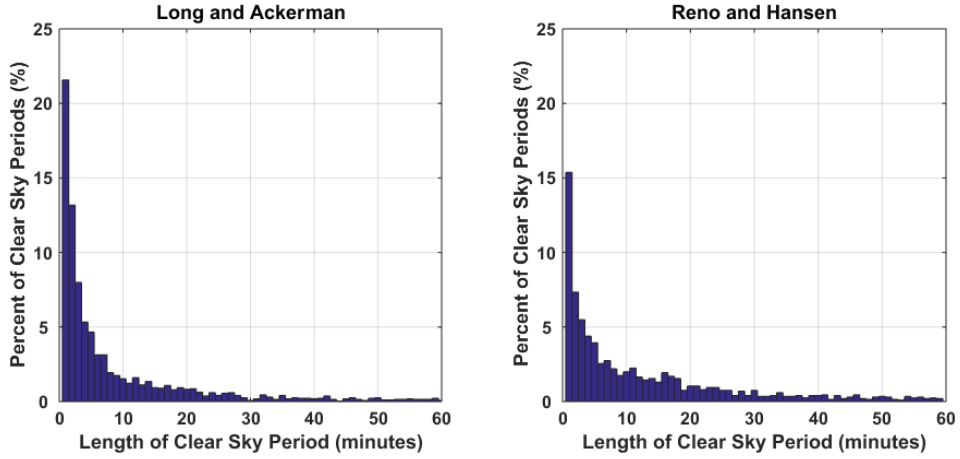

Figure 13. Histogram comparison of the length of clear sky periods for the proposed algorithm with Long and Ackerman.

5.3 Comparison with Cloud Detection in Satellite Images

We compared clear sky periods identified by the proposed algorithm with a cloud classification algorithm that operates on satellite images. The GOES Surface and Insolation Product (GSIP) created every hour by NOAA from processed GOES satellite data includes cloud classification for each image pixel at a spatial resolution of 1/8 x 1/8 degrees (NOAA, 2013), or approximately 10×10 km per pixel for the continental United States. NOAA processes GSIP data using all five imaging channels on GOES together with NCEP GFS forecast model data and IMS daily snow data. The algorithm detects clouds and determines the cloud fraction (i.e., fraction of pixels with clouds), dominant cloud type, and radiative fluxes. The GSIP cloud type parameter is correlated with the solar variability and radiation that reaches the ground (Reno and Stein, 2013).

We applied the proposed algorithm to ground measurements of GHI at the NREL MIDC Anatolia site in Rancho Cordova, California (NREL, 2013). For the period of April 1st, 2009 to June 30th, 2011, once per hour GSIP images are cropped to a 4x4 pixel (roughly 40x40 km square) window around the ground irradiance sensor site. We classified each hour as clear if each of the 16 pixels was identified as clear, otherwise the hour is classified as not clear. From the ground data, we classified a 60-minute period centered on the satellite image time as clear if all 60 minutes were identified as clear by the proposed algorithm. Only daylight hours are considered in this analysis. An example timeframe is shown in Figure 14 with the proposed algorithm classifying each minute and the satellite images classifying whole hours around the time the image was taken.

Fundamentally, satellite imagery provides a once per hour observation over an area of several km^2 per pixel, while a ground measurement provides a point measurement at a higher temporal scale. We expect some differences between these two data sources. For example, clouds visible in the satellite image might never interpose between the sun and sensor during the hour. For this reason, a direct minute by minute comparison between the satellite imagery and proposed algorithm is not attempted.

The results for the 3174 one-hour daylight time periods are shown in Table 2. For every hour during which the GHI clear sky detection algorithm classifies every minute as clear, the satellite image was also completely clear. The contrapositive is also true; if the satellite image was cloudy to any extent, at least some of the irradiance measurements during the hour are identified as not clear by the proposed algorithm. Interestingly for 27% of the hours, the satellite image is clear but the ground measurements indicate some minutes with clouds. This difference can result from the limited spatial and temporal resolution of the satellite data. For example, the satellite image will not show clouds that move through the pixel window or that form and dissipate between image times.

Table 2. Percent of 3174 hourly daylight periods classified by each algorithm as cloudy or clear.

		Satellite Cloud Classification	
		Clear (100% of pixels)	**Cloudy (<100% of pixels)**
Reno and Hansen	**Clear (all 60 minutes are clear)**	42.1 %	0 %
	Cloudy (< 60 minutes are clear)	27.3 %	30.6 %

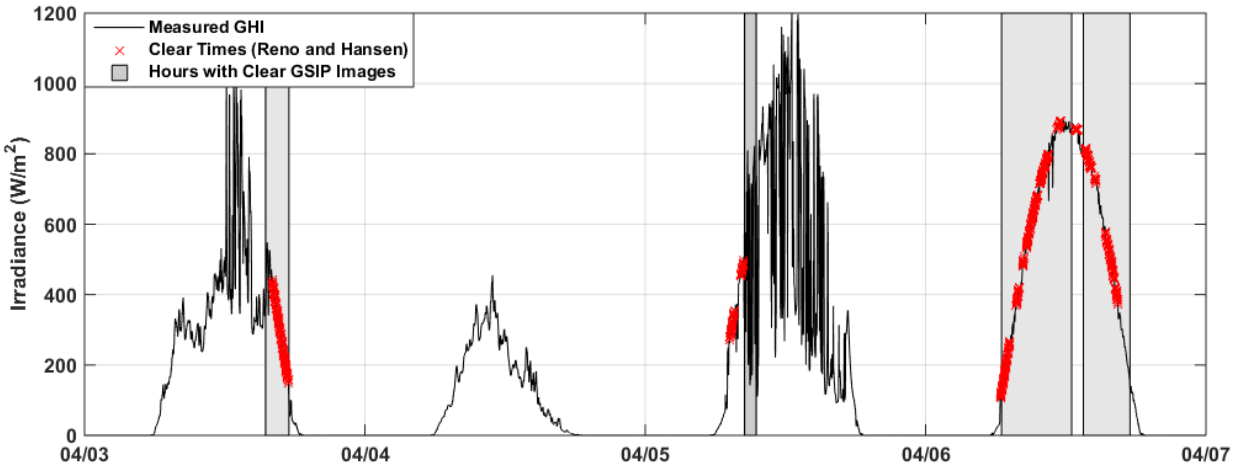

Figure 14. Comparison of clear sky detection algorithm with satellite imagery analysis (GSIP).

5.4 Comparison with Cloud Detection in Ground Imagers

Our algorithm is also compared to cloud detection using ground-based sky imagers. NOAA collects data from Total Sky Imagers (TSI) located at each SURFRAD site (NOAA, 2013). The TSI is a camera suspended over a convex mirror with a rotating black shadow stripe on the mirror. Images are taken every minute and processed to determine cloud cover data. Each pixel is classified as clear, thin, or opaque. The cloud fractions of percent opaque and percent thin are logged for the entire image and for the part of the image around the sun. The number of pixels near the sun varies throughout the day.

The TSI cloud fractions and GHI data are compared for the SURFRAD Table Mountain site in Boulder, Colorado, for daylight hours for the year of 2010. The GHI time series is processed by the clear sky detection algorithm for each minute using the 10-minute moving window.

Our algorithm rarely classifies a minute as clear when the TSI image shows substantial cloudiness, and conversely, essentially always classifies a minute as clear when the TSI image is essentially free of clouds. When the TSI declares the image as cloudy (greater than 95% of pixels having clouds), more than 99.9% of these minutes are classified as not clear by the clear sky detection algorithm. When the TSI declares the image as clear (less than 5% of pixels in the entire image with clouds), more than 98.3% of these minutes are also classified as clear by our algorithm.

6. Conclusions

We have proposed a new method for identifying periods with GHI similar to clear sky conditions using only high-time resolution GHI data and a clear sky model. The algorithm uses five statistics computed from the GHI data and clear sky model results to identify these periods. Iterative application of the method can be used to compensate for any bias errors in the clear sky model for a location or time period of year. The method and threshold criteria were developed using irradiance data from Golden, Colorado, and are shown to perform well at three other sites (i.e., Las Vegas, NV, Rancho Cordova, CA, and Boulder, CO) with varying GHI sensors, elevations, and management groups.

Results from the method are verified visually and are compared with results from a different, well-regarded method for identifying clear sky conditions that uses measured diffuse irradiance, and with sky images from both satellite and ground-based cameras. The novel algorithm presented here is shown to work as accurately as previously published algorithms while not requiring measurement of the direct or diffuse irradiance. Identification of periods with clear sky equivalent GHI has wide application to design and operation of photovoltaic power systems, including system performance monitoring, data error detection and data cleaning, and quickly identifying the occurrence of partial shading.

Acknowledgement

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.

References

Alam, S., 2006. Prediction of direct and global solar irradiance using broadband models: Validation of REST model. Renewable Energy, 31, 1253-1263.

Badescu, V., 1997. Verification of some very simple clear and cloudy sky models to evaluate global solar irradiance. Solar Energy, 61, 251-264.

Badescu, V., Gueymard, C.A., Cheval, S., Oprea, C., Baciu, M., Dumitrescu, A., Iacobescu, F., Milos, I., Rada, C., 2012. Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models. Renewable and Sustainable Energy Reviews, 16, 1636-1656.

Calbó, J., González, J.-A., Pagès, D., 2001. A Method for Sky-Condition Classification from Ground-Based Solar Radiation Measurements. Journal of Applied Meteorology, 40, 2193–2199.

Cucumo, M., De Rosa, A., Ferraro, V., Kaliakatsos, D., Marinelli, V., 2008. Correlations of global and diffuse solar luminous efficacy for all sky conditions and comparisons with experimental data of five localities. Renewable Energy, 33, 2036-2047.

Cucumo, M., De Rosa, A., Ferraro, V., Kaliakatsos, D., Marinelli, V., 2010. Correlations of direct solar luminous efficacy for all sky, clear sky and intermediate sky conditions and comparisons with experimental data of five localities. Renewable Energy, 35, 2143-2156.

DeFelice, T.P., Wylie, B.K., 2001. Sky type discrimination using a ground-based sun photometer. Atmospheric Research, 59-60, 313-329.

Duchon, C.E., O'Malley, M.S., 1999. Estimating cloud type from pyranometer observations. Journal of Applied Meteorology, 38, 132-141.

Dupont, J.C., Haeffelin, M., Long, C.N., 2008. Evaluation of cloudless-sky periods detected by shortwave and longwave algorithms using lidar measurements. Geophysical Research Letters, 35.

Ghonima, M.S., Urquhart, B., Chow, C.W., Shields, J.E., Cazorla, A., Kleissl, J., 2012. A method for cloud detection and opacity classification based on ground based sky imagery. Atmos. Meas. Tech., 5, 2881-2892.

Gueymard, C.A., Myers, D.R., 2009. Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling. Solar Energy, 83, 171-185.

Gueymard, C.A., 2012. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. Solar Energy, 86, 2145-2169.

Harrison, R.G., Chalmers, N., Hogan, R.J., 2008. Retrospective cloud determinations from surface solar radiation measurements. Atmospheric Research, 90, 54-62.

HelioClim, 2013. HelioClim Solar Radiation. <http://www.helioclim.org/linke/linke_helioserve.html>.

Hogan, R.J., Jakov, C., Illingworth, A.J., 2001. Comparison of ECMWF winter-season cloud fraction with radar-derived values. Journal of Applied Meteorology, 40, 513-525.

Ineichen, P., Perez, R., 2002. A new airmass independent formulation for the Linke turbidity coefficient. Solar Energy, 73, 151-157.

Ineichen, P., 2006. Comparison of eight clear sky broadband models against 16 independent data banks. Solar Energy, 80, 468-478.

Kang, B.O., Tam, K.-S., 2013. A new characterization and classification method for daily sky conditions based on ground-based solar irradiance measurement data. Solar Energy, 94, 102-118.

Kasten, F., 1980. A Simple Parameterization of the Pyrheliometric Formula for Determining the Linke Turbidity Factor. Meteorologische Rundschau, 33, 124-127.

Kasten, F., Young, A.T., 1989. Revised Optical Air-Mass Tables and Approximation Formula. Applied Optics, 28, 4735-4738.

Lam, J.C., Li, D.H.W., 1996. Correlation between global solar radiation and its direct and diffuse components. Building and Environment, 31, 527-535.

Li, D.H.W., Lam, J.C., 2001. An analysis of climatic parameters and sky condition classification. Building and Environment, 36, 435-445.

Li, D.H.W., Tang, H.L., 2008. Standard skies classification in Hong Kong. Journal of Atmospheric and Solar-Terrestrial Physics, 70, 1222-1230.

Linke, F., 1922. Transmissions-Koeffizient und Trubungsfaktor. Beitr. Phys. fr. Atmos, 10, 91-103.

Long, C.N., Ackerman, T.P., 2000. Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. Journal of Geophysical Research-Atmospheres, 105, 15609-15626.

Markou, M.T., Bartzokas, A., Kambezidis, H.D., 2007. A new statistical methodology for classification of sky luminance distributions based on scan data. Atmospheric Research, 86, 261-277.

Marty, C., Philipona, R., 2000. The Clear-Sky Index to separate clear-sky from cloudy-sky situations in climate research. Geophysical Research Letters, 27, 2649-2652.

Muneer, T., Gul, M.S., Kubie, J., 2000. Models for Estimating Solar Radiation and Illuminance From Meteorological Parameters. Journal of Solar Energy Engineering, 122, 146-153.

NOAA, 2013. ESRL Global Monitoring Division. <http://www.esrl.noaa.gov/gmd/grad/surfrad/>.

NOAA, 2013. NOAA's Comprehensive Large Array-data Stewardship System. <http://www.class.ngdc.noaa.gov/saa/products/search?datatype_family=GSIP>.

NREL, 2013. Measurement and Instrumentation Data Center (MIDC). <http://www.nrel.gov/midc/>.

Orsini, A., Tomasi, C., Calzolari, F., Nardino, M., Cacciari, A., Georgiadis, T., 2002. Cloud cover classification through simultaneous ground-based measurements of solar and infrared radiation. Atmospheric Research, 61, 251-275.

Pages, D., Calbo, J., Gonzalez, J.A., 2003. Using routine meteorological data to derive sky conditions. Annales Geophysicae, 21, 649-654.

Perez, R., Ineichen, P., Seals, R., Michalsky, J., Stewart, R., 1990. Modeling daylight availability and irradiance components from direct and global irradiance. Solar Energy, 44, 271-289.

Perez, R., Ineichen, P., Seals, R., Zelenka, A., 1990. Making full use of the clearness index for parameterizing hourly insolation conditions. Solar Energy, 45, 111-114.

Rahim, R., Baharuddin, Mulyadi, R., 2004. Classification of daylight and radiation data into three sky conditions by cloud ratio and sunshine duration. Energy and Buildings, 36, 660-666.

Reda, I., 1998. Improving the Accuracy of Using Pyranometers to Measure the Clear Sky Global Solar Irradiance, National Renewable Energy Laboratory, Golden, CO, NREL/TP-560-24833.

Remund J., W.L., Lefevre M., Ranchin T., Page J., 2003. Worldwide Linke turbidity information, Proceedings of ISES Solar World Congress, Göteborg, Sweden.

Reno, M.J., Hansen, C.W., Stein, J.S., 2012. Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis, Sandia National Laboratories, SAND2012-2389.

Reno, M.J., Stein, J.S., 2013. Using Cloud Classification to Model Solar Variability, ASES National Solar Conference, Baltimore, MD.

SoDa, 2011. Solar Radiation Data Service. <http://www.soda-is.com/eng/index.html>.

Stein, J.S., Hansen, C.W., Reno, M.J., 2012. The Variability Index: A New and Novel Metric for Quantifying Irradiance and PV Output Variability, World Renewable Energy Forum,

Younes, S., Muneer, T., 2007. Clear-sky classification procedures and models using a world-wide data-base. Applied Energy, 84, 623-645.