PV POWER OUTPUT VARIABILITY: CALCULATION OF CORRELATION COEFFICIENTS USING SATELLITE INSOLATION DATA

Thomas E. Hoff
Clean Power Research
Napa, CA
www.cleanpower.com

Richard Perez
ASRC, University at Albany
Albany, NY

ABSTRACT

Utility planners and operators are concerned about how short-term PV system output changes may affect utility system stability. Short-term output changes are driven by changes in the clearness index. This paper analyzes the correlation coefficient of the change in the clearness index between two locations as a function of distance, time interval, and other parameters. The paper presents a method to estimate correlation coefficients that uses location-specific input parameters. The method is derived empirically and validated using 12 years of hourly satellite-derived data from SolarAnywhere® in three geographic regions in the United States (Southwest, Southern Great Plains, and Hawaii). Results from 70,000 station pair combinations suggest that: (1) correlation coefficients decrease predictably with increasing distance; (2) correlation coefficients decrease at a similar rate when evaluated versus distance divided by the considered variability time interval; and (3) the accuracy of results is improved by including an implied cloud speed term. The approach has potential financial benefits to systems that are concerned about PV power output variability, ranging from individual distribution feeders to state-wide balancing regions.

1. INTRODUCTION

PV capacity is increasing on utility systems. As a result, utility planners and operators are growing more concerned about potential impacts of power supply variability caused by transient clouds. Utilities and control system operators need to adapt their planning, scheduling, and operating strategies to accommodate this variability while at the same time maintaining existing standards of reliability.

It is impossible to effectively manage these systems, however, without a clear understanding of PV output variability or the methods to quantify it. Whether forecasting loads and scheduling capacity several hours ahead or planning for reserve resources years into the future, the industry needs to be able to quantify expected output variability for fleets of up to hundreds of thousands of PV systems spread across large geographical territories. Underestimating reserve requirements may result in a failure to meet reliability standards and an unstable power system. Overestimating reserve requirements may result in an unnecessary expenditure of capital and higher operating costs.

The present objective is to develop analytical methods and tools to quantify PV fleet output variability. Variability in time intervals ranging from a few seconds to a few minutes is of primary interest since control area reserves are dispatched over these time intervals. For example, regulation reserves might be dispatched at an ISO every five seconds through a broadcast signal. Knowledge about PV fleet variability in five-second intervals could be used to determine the resources necessary to provide frequency regulation service in response to power fluctuations.

Variability of a PV fleet is thus a measure of the magnitude of changes in its aggregate power output corresponding to the defined time interval and taken over a representative study period. Note that it is the *change* in output, rather than the output itself, that is desired. Also note that, for each time interval the change in output may vary in both magnitude and sign (positive and negative). The statistical metric that is employed to quantify variability is the *standard deviation of the change in fleet power output*.

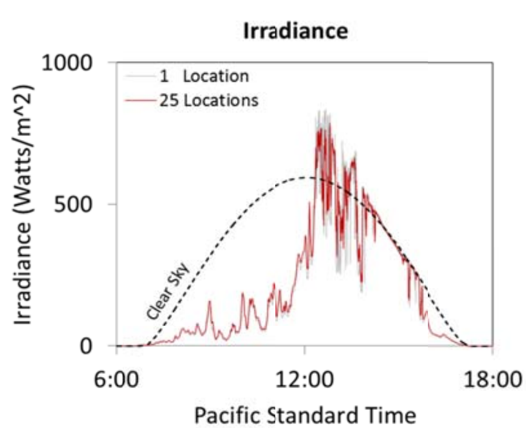
It is helpful to graphically illustrate what is meant by output variability. The left side of Figure 1 presents measured 10-second irradiance data (PV power output is almost directly proportional to irradiance) and the right side of the figure presents the change in irradiance using a 10-second time interval for a network of 25 weather monitoring stations in a 400-meter by 400-meter grid located at Cordelia Junction, CA on November 7, 2010 [1]. The light gray lines correspond to irradiance and variability for a single location and the dark red lines correspond to average irradiance distributed across 25 locations. Results suggest that spreading capacity across 25 locations rather than concentrating it at a single location reduces variability by more than 70 percent in this particular instance.

A "fleet computation" approach can be taken to calculate output variability for a fleet of PV systems as follows: identify the PV systems that constitute the fleet to be studied; select the time interval and time period of concern (e.g., one-minute changes evaluated over a one-year period); obtain time-synchronized solar irradiance data for each location where a PV system is to be sited; simulate output for each PV system using standard modeling tools; sum the output from each individual system to obtain the combined fleet output; calculate the change in fleet output for each time interval; and finally calculate the resulting statistical output variability from the stream of values.

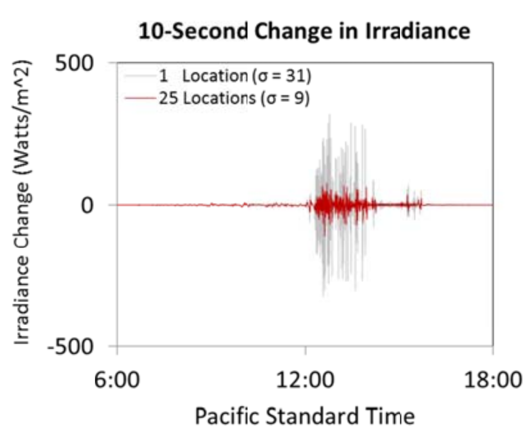
This "fleet computation" approach, while technically valid, is difficult to implement in practice for several reasons. First, insolation data is not available in sufficient resolution (either time resolution or geographical resolution). For example, SolarAnywhere [2], which provides operational real-time insolation data for the continental U.S. and Hawaii, is currently based on a 10 km x 10 km grid and a one-hour time interval. It has a practical real-time limit of one-half hour and 1 km based on current satellite technology. Fleet computation could not be performed for, say, systems spaced 0.5 km apart with a four-minute time interval. Second, PV variability determined using the fleet computation approach is only applicable to studies having a matching time interval of interest and a fixed fleet selection. The study would have to be re-commissioned whenever additional PV systems came on-line. Finally, calculations are highly computation intensive, and thus are not suitable for real-time operations.

A more viable approach is to streamline the calculations through the use of a general-purpose PV output variability methodology. The method needs to quantify short-term fleet power output variability using the observations that sky clearness and sun position drive the changes in the short-term output for individual PV systems and that physical parameters (i.e., dimensions, plant spacing, number of plants, etc.) determine overall fleet variability.

Figure 1. Twenty-five location network reduces 10-second variability by more than 70 percent in a 400 meter x 400 meter grid at Cordelia Junction, CA on November 7, 2010.



10 sec. data from 400m x 400m grid at Cordelia Junction, CA (Nov. 7, 2010)



10 sec. data from 400m x 400m grid at Cordelia Junction, CA (Nov. 7, 2010)

Hoff and Perez [3] developed a simplified model as a first step towards a general method to quantify the output variability resulting from an ensemble of equally-spaced, identical PV systems. They defined output variability to be the standard deviation of the change in output over some time interval (such as one minute) using data taken from some time period (such as one year).

The simplified model covered the special case when the change in output between locations is uncorrelated (i.e., cloud impacts at one site are too distant to have predictable effects at another for the considered time scale), fleet capacity is equally distributed, and the variance at each location is the same. Under these conditions, Hoff and Perez showed that fleet output variability equals the output variability at any one location divided by the square root of the number of locations. Mills and Wiser [4] have derived a similar result that relates variability to the square root of the number of systems when the locations are uncorrelated.

2. CORRELATION VERSUS DISTANCE

The critical factors that affect output variability are the clearness of the sky, sun position, and PV fleet orientation (i.e., dimensions, plant spacing, number of plants, etc.). To improve accuracy, Hoff and Perez introduced a parameter called the Dispersion Factor. The Dispersion Factor is a parameter that incorporates the layout of a fleet of PV systems, the time scales of concern, and the motion of cloud interferences over the PV fleet. Hoff and Perez demonstrated that relative output variability resulting from the deployment of multiple plants decreased quasi-exponentially as a function of the generating resource's Dispersion Factor. Their results demonstrated that relative output variability (1) decreases as the distance between sites increases; (2) decreases more slowly as the time interval increases; and (3) decreases more slowly as the cloud transit speed increases.

Mills and Wiser analyzed measured one-minute insolation data over an extended period of time for 23 time-synchronized sites in the Southern Great Plains network of the Atmospheric Radiation Measurement (ARM) program. Their results demonstrated that the correlation of the change in the global clear sky index (1) decreases as the distance between sites increases and (2) decreases more slowly as the time interval increases.

Perez et. al. [6] analyzed the correlation between the variability observed at two neighboring sites as a function of their distance and of the considered variability time scale. The authors used 20-second to one-minute data to construct virtual networks at 24 US locations from the ARM program [7] and the SURFRAD Network and cloud speed derived from SolarAnywhere to calculate the station pair correlations for distances ranging from 10 meters to 100 km and from variability time scales ranging from 20 seconds to 15 minutes. Their results demonstrated that the correlation of the change in global clear sky index (1) decreases as the distance between sites increases and (2) decreases more slowly as the time interval increases.

The consistent conclusions of these studies are that correlation: (1) decreases as the distance between sites increases and (2) decreases more slowly as the time interval increases. Hoff and Perez add that the correlation decreases more slowly as the speed of the clouds increases.

3. OBJECTIVE

Utility planners clearly require a tool that can reliably quantify the maximum output variability of PV fleets using a manageable amount of data and analysis. The methods referred to above would potentially meet this requirement if the changes in output between locations were uncorrelated (i.e., correlation coefficient is zero). In actual fleets, however, PV systems will generally have some degree of correlation, so any planning tool will have to incorporate correlation effects in calculating actual fleet variability.

This paper takes a step towards a general method by analyzing the correlation coefficient of the change in clearness index between two locations as a function of distance, time interval, and other parameters. It uses hourly global horizontal insolation data from SolarAnywhere to calculate correlation coefficients for 70,000 station pair combinations across three separate geographic regions in the United States (Southwest, Southern Great Plains, and Hawaii). The correlation coefficients taken from these combinations are then compared to a method that could prove useful when integrated into utility planning and operations tool. Further details are presented in [5].

4. APPROACH

Hoff and Perez defined PV fleet variability as the standard deviation of its power output changes using a selected sampling time interval (e.g., such as one minute or one hour) and analysis period (such as one year), as expressed relative to the fleet capacity. To simplify the work, they formulated it in terms of the change in insolation rather than the change in PV power.

As stated earlier, sky clearness and sun position drive the changes in short-term output for individual PV systems. Mills and Wiser [4] and Perez, et. al. [8] subsequently isolated the random component of output change and examined changes attributable only to changes in global clear sky (or clearness) index. The global clearness index equals the measured global horizontal insolation divided by the clear sky insolation. This paper continues in the direction of Mills and Wiser and Perez, et. al. and focuses on changes in the global clearness index.

The global clearness index at a specific point in time is typically referred to as Kt^*. It equals the measured global horizontal insolation (GHI) divided by the clear-sky insolation. This paper refers to the change in the index between two points in time as ΔKt^*. Since the change occurs over some specified time interval, Δt, at some specific location n, the variable is fully qualified as $\Delta Kt^{*n}_{t,\Delta t}$. This only represents one pair of points in time. A set of values is identified by convention by bolding the variable. Thus, $\boldsymbol{\Delta Kt^{*n}_{\Delta t}}$ is the set of changes in the clearness indices at a specific location using a specific time interval over a specific time period.

Correlation and dependence in statistics are any of a broad class of statistical relationships between two or more random variables or observed data values. Let $\boldsymbol{\Delta Kt^{*1}_{\Delta t}}$ and $\boldsymbol{\Delta Kt^{*2}_{\Delta t}}$ represent sets of observed data values for the change in the clearness index at location 1 and location2. Pearson's product-moment correlation coefficient (typically referred to simply as the correlation coefficient) is calculated for each pair of locations.

The analysis is performed as follows:

1. Select a geographic region for analysis
2. Select a location for the first part of the pair
3. Select a location for the second part of the pair
4. Select a time interval for the analysis
5. Select a clear sky irradiance level bin
6. Obtain detailed insolation data
7. Calculate the change in the clearness index
8. Calculate the correlation coefficient
9. Repeat the calculation for all sets of location pairs, time intervals, and clear sky irradiance bins.

The focus of this paper is on trying to determine if patterns existing that help to better quantify correlation coefficients. As part of the objective, a method is tested that produces the desired output parameter of the correlation coefficient of the change in the clearness index between two separate locations. The inputs into this method include the distance between the two locations, time interval, and location-specific parameters based on empirical weather data, in particular, cloud speed.

5. RESULTS

Three separate geographic regions in the United States were selected for analysis: Southwest, Southern Great Plains, and Hawaii (see Table 1). The first location was selected using a grid size of 2.0°, 1.0°, or 0.5° for the Southwest, Southern Great Plains, and Hawaii, correspondingly. The second location was selected between 0.1° and 2.9° (about 10 to 300 km) from the first location (other map coordinates were available but the selected points provided sufficient data for the analysis). Hourly insolation data was obtained for each of the two locations covering the period January 1, 1998 through September 30, 2010 from SolarAnywhere [2]. The analysis was then performed as described above for time intervals of 1, 2, 3, and 4 hours and for 10 separate clear sky irradiance bins. This analysis resulted in more than 70,000 correlation coefficients.

Figure 2 presents a randomly selected set of correlation coefficients for the Southwest. The figures in the columns summarize the results for each time interval and the figures in the rows present the measured correlation coefficients versus several alternative candidate sets of variables. The first column summarizes results for a time interval of 1 hour. The second, third, and fourth columns plot the same results using time intervals of 2, 3, and 4 hours. Results in the top row present correlation coefficients versus the distance between the two locations. Results in the middle row present correlation coefficients versus distance divided by time interval. Results in the bottom row present correlation coefficients versus distance divided by time interval multiplied by relative speed; this term is related to the Dispersion Factor introduced by Hoff and Perez [3]. The dashed line in the bottom figures represents the results of a generalized method, proposed in this paper for use in future tools, that will be validated in the present analysis. Results are calculated using parameters obtained from SolarAnywhere.

Figure 3 and Figure 4 present comparative results for the Great Plains and Hawaii. The patterns presented in the figures are similar across all time intervals in the three geographic locations.

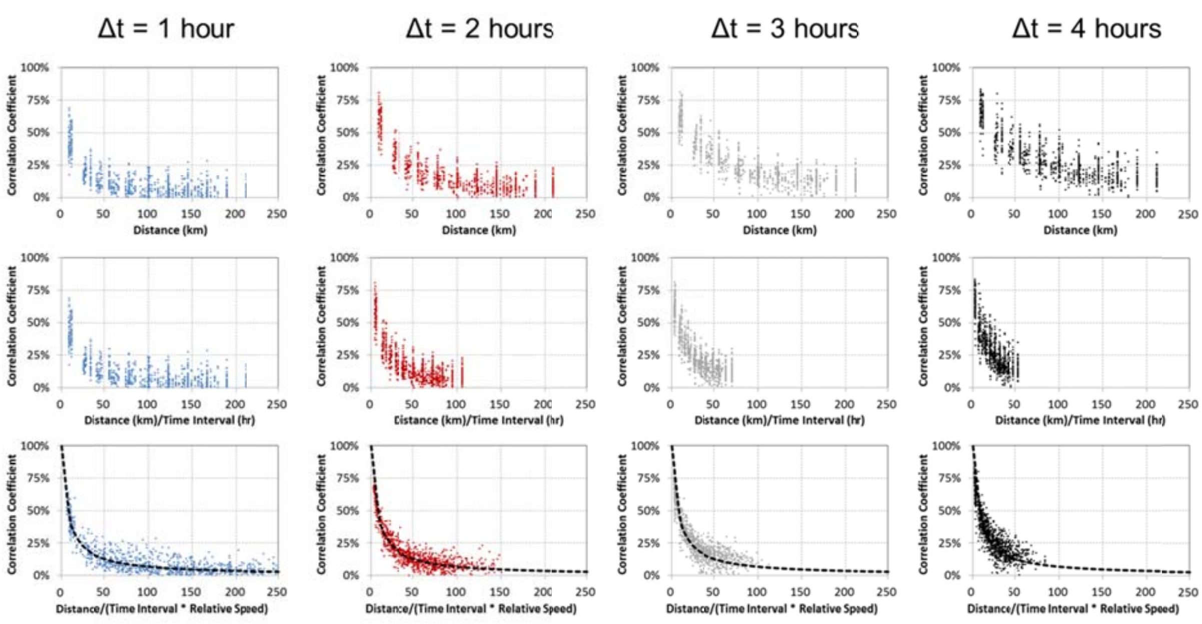
6. CONCLUSIONS

The analysis provides several key findings. First, consistent with previous studies, the correlation coefficients decrease with increasing distance. Second, also consistent with previous studies, this decrease occurs more slowly with longer time intervals. An alternative way of viewing this result is that correlation coefficients decrease at a similar rate when plotted versus distance divided by time interval. Third, the scatter in results is further decreased when a relative speed is introduced for the first location in the pair of locations. Finally, the generalized method (shown by the dashed black lines) fits the empirical data quite well when calibrated using the location-specific derived input parameters.

Table 1. Summary of input data.

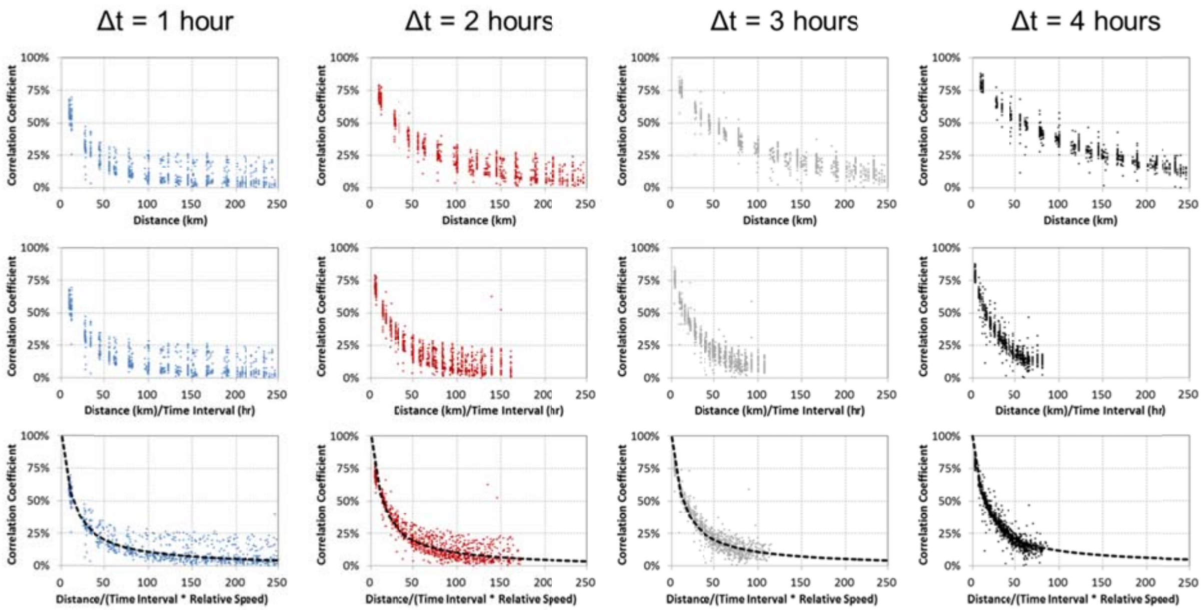
Region	Southwest	Southern Great Plains	Hawaii
Location #1	Latitude: 32° to 42° Longitude: -125° to -109° Grid Size: 2.0°	Latitude: 35° to 38° Longitude: -99° to -96° Grid Size: 1.0°	Latitude: 19° to 20° Longitude: -156° to -155° Grid Size: 0.5°
Location #2	0.1°, 0.3°, …, 1.9° from #1	0.1°, 0.3°, …, 2.9° from #1	0.1°, 0.2°, …, 1.0° from #1
Time Intervals	1, 2, 3, and 4 hours	1, 2, 3, and 4 hours	1, 2, 3, and 4 hours
Clear Sky Irradiance	10 irradiance bins in intervals of 0.1 kW/m^2	10 irradiance bins in increments of 0.1 kW/m^2	10 irradiance bins in increments of 0.1 kW/m^2

Figure 2. Correlation coefficients presented by time interval for Southwest.



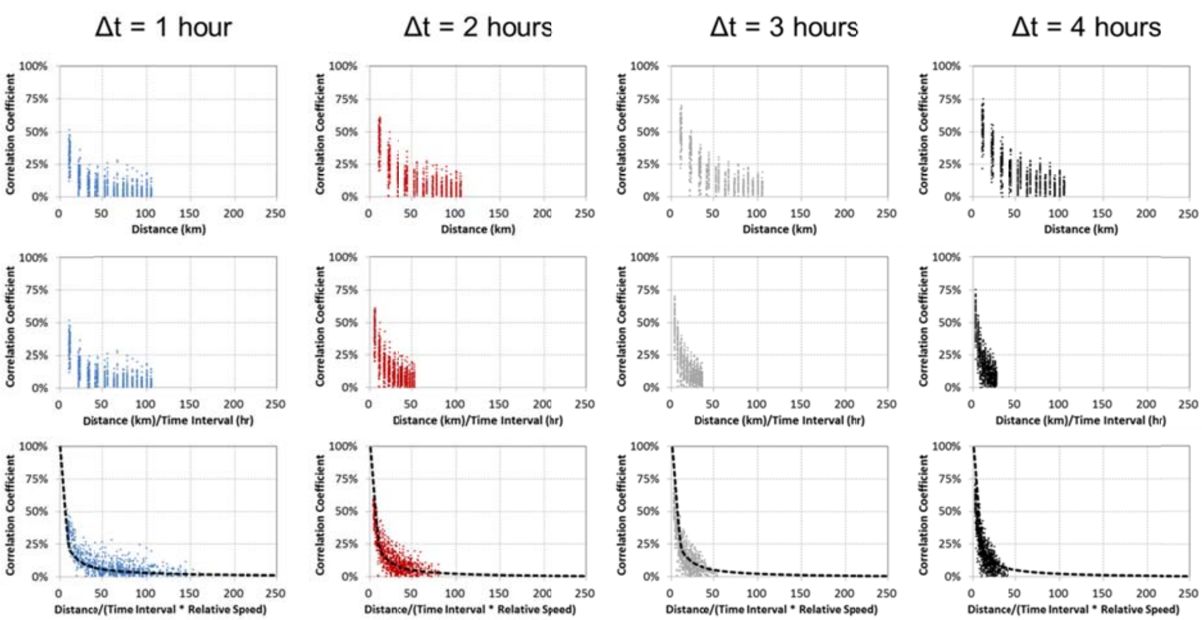
Note: Distance / (Time Interval * Relative Speed) is related to Dispersion Factor

Figure 3. Correlation coefficients presented by time interval for Great Plains.



Note: Distance / (Time Interval * Relative Speed) is related to Dispersion Factor

Figure 4. Correlation coefficients presented by time interval for Hawaii.



Note: Distance / (Time Interval * Relative Speed) is related to Dispersion Factor

7. ACKNOWLEDGEMENTS

Portions of this study were funded under a California Solar Initiative (CSI) Grant Agreement titled "Advanced Modeling and Verification for High Penetration PV." The California Public Utilities Commission is the Funding Approver, Itron is the Program Manager, and the California IOUs are the Funding Distributors. Thanks to Ben Norris who designed, implemented, and operated the mobile irradiance network and provided valuable comments on the paper. Thanks to Jeff Ressler for his comments. Opinions expressed herein are those of the authors only.

8. REFERENCES

[1]. Hoff, T.E., Norris, B. 2010. Mobile High-Density Irradiance Sensor Network: Cordelia Junction Results.

[2]. Solar Anywhere, 2011. Web-Based Service that Provides Hourly, Satellite-Derived Solar Irradiance Data Forecasted 7 days Ahead and Archival Data back to January 1, 1998. www.SolarAnywhere.com.

[3]. Hoff, T. E., Perez, R. 2010. Quantifying PV power Output Variability. Solar Energy 84 (2010) 1782–1793.

[4]. Mills, A., Wiser, R. 2010. Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power. Lawrence Berkeley National Laboratory Technical Report LBNL-3884E.

[5]. Hoff, T. E., Perez, R. 2010. PV Power Output Variability: Correlation Coefficients. www.cleanpower.com.

[6]. Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., C., Hoff, T. E. 2010. Short-term irradiance variability correlation as a function of distance. Submitted to Solar Energy.

[7]. Stokes, G.M., Schwartz, S.E., 1994. The atmospheric radiation measurement (ARM) program: programmatic background and design of the cloud and radiation test bed. Bulletin of American Meteorological Society 75, 1201–1221.

[8]. Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., C., Hoff, T. E. 2010. Parameterization of site-specific short-term irradiance variability. Submitted to Solar Energy.