Title

High-frequency irradiance fluctuations and geographic smoothing

Journal

Solar Energy, 86(8)

ISSN

0038092X

Authors

Lave, Matthew
Kleissl, Jan
Arias-Castro, Ery

Publication Date

2012-08-01

DOI

10.1016/j.solener.2011.06.031

Peer reviewed

High-frequency irradiance fluctuations and geographic smoothing

Matthew Lave[1], Jan Kleissl[1], Ery Arias-Castro[2]

[1]Dept. of Mechanical & Aerospace Eng., University of California, San Diego

[2]Dept. of Mathematics, University of California, San Diego

Abstract

Using six San Diego solar resource stations, clear-sky indices at 1-sec resolution were computed for one site and for the average of six sites separated by less than 3 km to estimate the smoothing of aggregated power output due to geographic dispersion in a distribution feeder. Ramp rate (RR) analysis was conducted on the 1-sec timeseries, including moving averages to simulate a large PV plant with energy storage. Annual maximum RRs of up to 60% per second were observed, and the largest 1-sec ramp rates were enhanced over 40% by cloud reflection. However, 5% per second ramps never occurred for a simulated 10 MW power plant. Applying a wavelet transform to both the clear-sky index at one site and the average of six sites showed a strong reduction in variability at timescales shorter than 5-min, with a lesser decrease at longer timescales. Comparing these variability reductions to the Hoff and Perez (2010) model, good agreement was observed at high dispersion factors (short timescales), but our analysis shows larger reductions in variability than the model at smaller dispersion factors (long timescales).

1. Introduction

The variable nature of solar radiation is a concern in realizing high penetrations of solar photovoltaics (PV) into an electric grid. High frequency fluctuations of irradiance caused by fast moving clouds can lead to unpredictable variations in power output on short timescales. Short-term irradiance fluctuations can cause voltage flicker and voltage fluctuations that can trigger automated line equipment (e.g. tap changers) on distribution feeders leading to larger maintenance costs for utilities. Given constant load, counteracting such fluctuations would require dynamic inverter VAR control or a secondary power source (e.g. energy storage) that could ramp up or down at high frequencies to provide load following services. Such ancillary services are costly to operate, so reducing short-term variation is essential. Longer scale variations caused by cloud groups or weather fronts are also problematic as they lead to a large reduction in power generation over a large area. These long-term fluctuations are easier to forecast and can be mitigated by slower ramping (but larger) supplementary power sources, but the ramping and scheduling of power plants also adds costs to the operation of the electric grid. Grid operators are often concerned with worst-case scenarios, and it is important to understand the behavior of PV power output fluctuations over various timescales.

Many previous studies have shown the benefit of high-frequency irradiance data. Suehrcke and McCormick (1989) and Gansler et al. (1995) found 1-min data to have different statistics from lower-frequency data, including a much more bi-modal distribution than 1-hour or 1-day data. Gansler et al. (1995) mention that while using 1-hour data may be acceptable for space and water heating systems, where the thermal capacitance effects dampen out short-term variations, the time response of PV systems is much faster and using 1-hour data will likely lead to errors.

Understanding that high-frequency fluctuations are important, further studies have looked to characterize these fluctuations, often by comparing fluctuations at one site to fluctuations at the average of multiple sites. Otani et al. (1997) use a fluctuation factor defined as the root mean squared (RMS) value of a high-pass filtered 1-min time series of solar irradiance to demonstrate a 2-5 times reduction in

variability when considering 9 sites located within a 4 km by 4 km grid. Curtright and Apt (2008) and Lave and Kleissl (2010) used 1-min timeseries to show reductions in the mean, maximum, and standard deviation of ramp rates (RRs) when considering the average of three or four sites versus only one site. Power spectral densities (PSDs) presented in Otani et al. (1997), Curtright and Apt (2008), and Lave and Kleissl (2010) all show strong reductions in power content of fluctuations of the average of multiple sites versus the power content of fluctuations at one site. Lave and Kleissl (2010) also present coherence spectra which show that the sites in Colorado which were 60 km or more apart were uncorrelated on timescales shorter than 12-hours. Two sites that were only 19 km apart were uncorrelated on timescales shorter than 3-hours.

Wiemken et al. (2001) used 5-min normalized output from 100 PV sites spread throughout Germany. They found the standard deviation of the average of 100 sites to be 0.61 that of 1 site for the month of June, and that 5-min fluctuations of ±5% of power output at nameplate capacity are virtually nonexistent in the average, yet single sites have fluctuations larger than ±50%. Also included in that paper is a figure from Beyer et al. (1991), which shows exponential decay of cross-correlation as a function of distance for hourly irradiance data from six sites in Germany which can be used to estimate the reduction in standard deviation when averaging sites. Murata et al. (2009) analyzed 1-min data from 52 PV systems spread across Japan to determine the "smoothing effect" of aggregating multiple systems. The authors introduce a fluctuation index, which is the maximum difference in aggregated power output over a given time interval. They found that over 1-min, sites more than about 50-100 km apart were uncorrelated and thus that there was a limit reached whereby adding more PV sites had no effect on reducing variability, since the variability introduced by the diurnal cycle eventually becomes larger than the cloud-induced variability. For times greater than 10-min, however, they reject the hypothesis that sites within 1000 km are independent, though some of the dependence may be due to diurnal solar cycles and could be eliminated by using a normalized solar radiation.

Hoff and Perez (2010, hereafter HP10) present a framework to estimate the decrease in standard deviation of irradiance achieved by aggregating PV sites. The reduction in standard deviation is a function of the number of PV sites and a dispersion factor, D, defined as the number of time intervals it takes for a cloud to pass over all PV sites across the region being considered. The dispersion factor is useful in determining when the transition from PV sites being uncorrelated to correlated occurs. They predict a factor of $\sqrt{N}$ reduction in standard deviation of the average of N sites compared to the standard deviation of one site for the "spacious region," where the number of sites is much less than the dispersion factor, $N << D$. This corresponds to the sites being fully independent of one another, and is a known result from statistics on independent random variables. At an "optimal point" where the number of sites equals the dispersion factor, $N = D$, they derive a factor of N reduction in standard deviation. At this point, sites would be perfectly correlated when shifted by an appropriate timestep (the dispersion factor divided by the number of sites), and the standard deviation of the average of all sites will be reduced more than would be expected if the sites were entirely independent. HP10 also define a "limited region," between the "spacious region" and the "optimal point," $N < D$. In the "crowded region" where the number of sites is larger than the dispersion factor, $N > D$, they propose that the standard deviation will be reduced by a factor of D, since the sites are at least partially dependent, and adding additional sites will not reduce the standard deviation of the average since the reduction is only a function of the dispersion factor, and not the number of sites. HP10 perform a limited model validation by simulating a fleet of PV systems based on measured irradiance at only one site. In the simulated system, the irradiance at the non-measured sites was found by shifting the timestamp but otherwise maintaining the measured data, thus simulating an equally spaced system in which frozen clouds move at a constant speed along a line containing all sites.

Woyte et al. (2007) present a unique study in that they use very high frequency data (1-sec, 5-sec, or 1-min depending on the site) collected for up to 2-years, instantaneous clearness index, and a wavelet transform to analyze fluctuations of all scales in time, from very short to very long. The Haar wavelet was applied to each clearness index dataset to detect fluctuations over various timescales. They introduce a fluctuation power index, which is the sum of the square of the wavelet mode at each timescale, and is

used to quantify the amplitude and frequency of occurrence of fluctuations on a specific timescale. In other wavelet studies, Kawasaki et. al (2006) applied the Daubechies 4 wavelet to 1-min irradiance 2-year timeseries from nine sites in a 4x4 km grid, and Perpinan and Lorenzo (2010) applied the MODWT wavelet to 1-sec solar irradiance timeseries from a few days in October, 2009.

This paper builds on these previous works by using 1-sec clear-sky index (K_c) data from 6 sites on a microgrid similar to urban distribution feeders (Section 2) to quantify extreme ramp rates (RRs). Methods are described in Section 3. RRs were analyzed by computing statistics at different time steps and by using varying moving average intervals to represent large PV plants or storage (Section 4.1). Coherence spectra are employed in Section 4.2 to analyze the correlation between six sites at different time scales. We apply a wavelet to detect variability over various timescales relevant to the operation of a power grid (Section 4.3). Wavelet analysis allowed for a localized study of the power content of variations over various timescales. The power content of variations at one site was compared to the power content of variations at the average of six sites in close proximity to study the reduction in variability over various timescales achieved by using multiple site locations (Section 4.4) and to test the model of HP10 (Section 4.5). Conclusions are presented in Section 5.

2. Data

Global Horizontal Irradiance (GHI) was recorded once per second at sites throughout the University of California, San Diego (UCSD) campus as part of the UCSD Decision Making using Real-time Observations for Environmental Sustainability (DEMROES) network of sensors (Kleissl et al., 2010, Fig. 1). All sites employ a LICOR Li-200SZ silicon pyranometer sampling at 1Hz. The collection of 1-sec data proved to be a challenge of both data storage on the datalogger and sensor reliability, and so data availability is inconsistent. While there are 8 sites maintained as part of the DEMROES network, at any given time a maximum of 6 sites recorded 1-sec data.

The main site used in this paper was the Engineering Building II (EBU2, 32.8813^{o}N, 117.2329^{0}W), for which data was available for all of 2009 except for May 23 through June 4. We do not expect for the 13 days of missing data to lead to a strong seasonal or other bias, and therefore will refer to this as 1-year of data. Five other sites also recorded data from July 31 to August 25, 2009, and are used to study the benefits of aggregating sites. These five sites ranged from Hubbs Hall (HUBB, 32.8670^{0}N, -117.2533^{0}W) which is 0.1 km from the Pacific Ocean to Moores Cancer Center (MOCC, 32.8782^{0}N,-117.2229^{0}W) 3km to the east. The Biomedical Science Building (BMSB, 32.8758^{0}N, 117.2362^{0}W), RIMAC Arena (RIMC, 32.8852^{0}N, 117.2402^{0}W), and Tioga Residence Hall (TIOG, 32.8790^{0}N 117.2434^{0}W) are more centrally located (Fig. 1). Since these data covers nearly the entire month of August, we will refer to this as a 1-month dataset.

After applying the factory calibration, clear days were used (assuming identical atmospheric composition) to create linear fits against RIMC, and each site was cross-calibrated by this linear fit. In addition, careful quality control was carried out by visually examining each site for shading and other errors. We noticed 5 occurrences at EBU2 and 15 at RIMC of greater than 400 W m^{-2} decreases lasting less than 3-sec during otherwise clear periods. We suspect that these were due to birds or airplanes instantaneously shading our sensors, and not cloud effects which were the focus of this study. Therefore, such dips were removed from the data using a linear interpolation plus characteristic variance to maintain statistics, which should be appropriate over such a short time. It is possible that such events also occurred during cloudy conditions, but is not clear how to detect such events then without possibly removing real variability.

To eliminate the deterministic effect of diurnal cycles, GHI measurements (in W m^{-2}) were converted into a dimensionless clear-sky index by dividing the measured GHI by the clear-sky irradiance. We used the Sunny Days model (Gonzales and Wilcox, 2004) based on Long and Ackerman (2000), which uses input GHI and diffuse horizontal irradiance (DHI, measured by a Dynamax SPN1 pyranometer at EBU2) to calculate clear-sky irradiance. Since Sunny Days is locally calibrated day-by-day it was found to be more accurate than the Ineichen and Perez (2002) climatological clear-sky model especially for mornings

and evenings. Times when the solar altitude angle was less than 10° were removed to eliminate both nighttime values when the clearness index is expected to always be zero and early morning and late evening periods when the pyranometer is subject to errors in cosine response.

Figure 1: Map of the UCSD solar resource sites, showing proximity to the Pacific Ocean (left), and Interstate 5 (center). From EBU2, distances and headings are: BMSB(0.69km, 205°), HUBB (2.47km, 230°), MOCC (0.95km, 105°), RIMC (0.80km, 300°), and TIOG (1.00km, 255°). Map © 2010 Google – Image © 2010 TerraMetrics.

3. Methods

3.1 1-year analysis at one site: Ramp Rate analysis

The frequency of occurrence and magnitude of RRs of solar PV are of critical interest to power system operators. From the 1-sec clear-sky indices, we can extract two different averages which have different practical relevance.

First, block averages were taken on time intervals varying from 1-sec to 1-hour, which shows the difference in statistics over various data averaging intervals. Typically irradiance or power output data are averaged over longer periods and our analysis allows comparison to such data. The block average method produces fewer data points as the block size increases.

Second, *moving* averages over intervals of $T = 2^j$ sec ($j = 1,2, \ldots 12$ corresponding to $T = 2,4, \ldots 4096$ sec) were computed at time steps of 1-sec such that the average at any given time, *t*, is the average of values $(2^j - 1)$ seconds before and $2j$ seconds after *t*. Intervals of 2^j seconds were chosen to be consistent with wavelet analysis presented later. No moving average was computed when $TOD < 2^j - 1$ *or* $\max(TOD) - TOD < 2^j$, where TOD is the time of day, such that the moving average would not be complete. Since this clipping removes more data at the beginning and end of the day for larger j reducing the available data, no analysis was performed for timescales larger than 4096-sec. Moving averages at different T are representative of power sampled every second, but averaged spatially over the dimensions of a solar power plant or by using energy storage.

From either the block average or the moving average, RRs were computed as the difference between successive clear-sky indices. Cumulative distribution functions (cdf) of RRs show the statistical distributions and extreme values. Additionally, we computed the averages of 1-sec ramps with magnitude greater than 0.25 to show the typical behavior before large positive or negative ramps. RRs of clear sky indices give the percent change (as a fraction of clear-sky irradiance) over one timestep, regardless of the

TOD when that change occurred. The clear sky index provides the best measure to compare cloud induced solar variability analyses between different sites. If the occurrence of clouds is independent of TOD, it also provides the most relevant measure to characterize solar energy variability at a site, especially for 2D tracking power plants (whose output fluctuates less over a clear day). However, if clouds occur preferentially over certain TODs and a fixed-tilt plant is considered, then clear sky index variability does not translate directly to power output variability of a PV plant.

3.2 Geographic smoothing at six sites over one month

3.2.1 Coherence spectrum

As a measure of spatial correlation of the clear-sky index over various time scales, we calculated the coherence spectrum between EBU2 and the other 5 sites. The coherence spectrum provides normalized covariance at each frequency, allowing for visualization of correlation over various timescales. The coherence is expected to be large at long timescales as large weather systems will lead to similar clear-sky indices for all the sites. Note, however, that solar cycles have been removed by using the clear-sky index and thus the coherence will not be as large as if irradiances had been used. The timescale at which sites become weakly correlated is an indication of the longest timescale on which the sites are nearly independent and will dampen aggregate variability. Although negative correlation would reduce variability more than zero correlation, negative correlation is not expected physically.

3.2.2 Wavelet analysis

The stationary or dyadic wavelet transform, *W*, of a signal $x(t)$ is (Mallat, 2009):

$$W_{2^j}^{\tau} = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) dt \tag{1}$$

where t is time, τ is the time offset from the beginning of the day, ψ is the wave used to produced the wavelet transform, and s is the scaling factor. Since we used a real wavelet and a discrete transform, we required that j be a positive integer. Although Woyte et al. (2007) used the Haar wavelet to detect dips and compute power content at each timescale, we found the Haar wavelet to be lacking in that large wavelet coefficients exist only at sharp signal transitions. This means that changes from one state to another (e.g. a step from cloudy to clear) are detected by the Haar wavelet rather than the duration of an up or down fluctuation (a top hat).

We instead chose to employ the top hat wavelet as the basis function of our analysis of clear-sky index timeseries. The top hat wavelet is defined as

$$\psi(t) = \begin{cases} 1, & \frac{1}{4} < t < 3/4 \\ -1, & 0 < t < \frac{1}{4} \quad || \quad \frac{3}{4} < t < 1 \\ 0, & else \end{cases} \tag{2}$$

and is shown in Figure 2. Although typically in wavelet analysis $s = 2^j$ is used in Eq. 1, we chose to define $s \equiv 2^{j+1}$ instead, so Equation 1 becomes

$$W_{2^j}^{\tau} = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^{j+1}}} \psi\left(\frac{t-\tau}{2^{j+1}}\right) dt. \tag{3}$$

This altered definition allows for the timescale, 2^j seconds, to describe the duration of the clear or cloudy period of interest rather than the duration of the entire wavelet. Substituting the clear-sky index, $x(t) = K_c(t)$, into Eq. 3 will result in a separate timeseries $w_j(\tau)$ for each j value (mode), where $w_j(\tau)$ is defined such that

$$w_j(\tau) = \int_{-\infty}^{\infty} K_c(t) \frac{1}{\sqrt{2^{j+1}}} \psi\left(\frac{t-\tau}{2^{j+1}}\right) dt \,. \tag{4}$$

Just as for the moving averages, we chose to limit our analysis to $j \leq 12$ (corresponding to 1.1 hours or less), and only τ values for which data were available over the entire interval of size 2^{j+1} around τ were retained. As such, early morning and late evening periods are not resolved at the longer modes. It is common in wavelet transforms to extend the original timeseries using either a periodic extension or zero padding, but we feel that neither is appropriate in this situation as they will both introduce effects that were not present in the original timeseries.

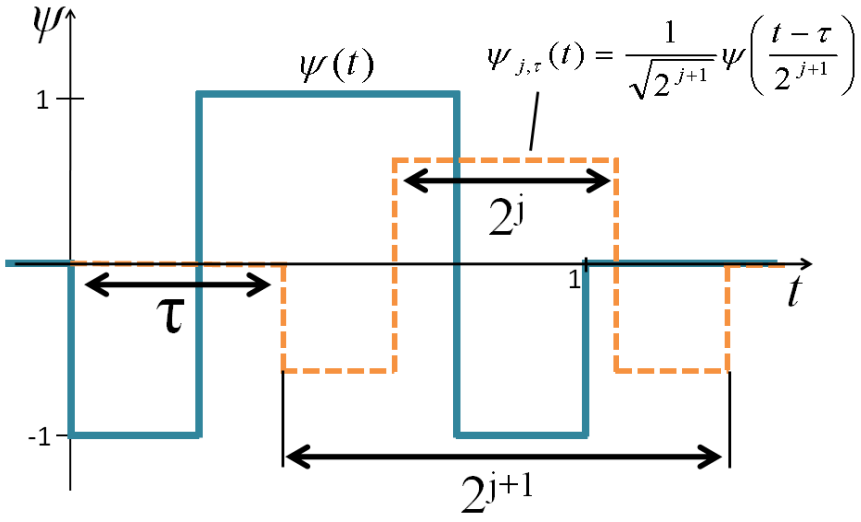

Figure 2: Top hat wavelet $\psi(t)$ (solid line) and the scaled and translated wavelet $\psi_{j,\tau}(t)$ (dashed line). This scaled wavelet would capture a clear period of duration 2^j bordered by cloudy periods.

The power content of each timeseries $w_j(\tau)$, can be found by calculating the wavelet periodogram *I*. Following the definition of the Fourier periodogram, the wavelet periodogram is the square of the coefficients of the wavelet transform, normalized by the length over which the wavelet was applied, which in this case is 2^{j+1}:

$$I_j(\tau) = \frac{1}{2^{j+1}} |w_j(\tau)|^2 \tag{5}$$

3.2.3 Application of wavelet analysis to determine reduction in variability from averaging 6 sites

The wavelet periodograms are still timeseries, and are difficult to examine visually for periods longer than one day. Therefore, we use the 'fluctuation power index,' as described by Woyte et al. (2007) to quantify the power contained in fluctuations at each timescale. The fluctuation power index, fpi, is:

$$\text{fpi}(j) = \frac{1}{T_j} \int_0^{T_j} I_j(\tau) d\tau, \tag{6}$$

where T_j is the length of the timeseries $I_j(\tau)$, which decreases as j increases due to unresolved periods of the higher modes. Using T_j instead of a constant value based on the length of the original $K_c(t)$ timeseries means that fpi(j) is an average value, which allows for comparison of fpi at different j values.

The fpi is essentially variance at each timescale (is variance if $\text{mean}(w_j(\tau)) = 0$), so we used fpi to evaluate the reduction in variability achieved by averaging six sites versus the variability at EBU2 alone and to compare our results to the HP10 model. HP10 define a dispersion factor $D = \frac{L}{V\,T}$, where L is the length of the region with the PV sites, V is the cloud velocity, and T is the relevant timescale. Although L and V remain constant for a given area and time, varying the timescale changes D. Since variability at multiple timescales was calculated through the fpi, we were able to test the HP10 model over various dispersion factors for the 1-month data.

4. Results

4.1 Ramp Rate analysis

The cdf of the absolute value of step sizes (SS) for K_c averaged over blocks of 1-sec, 10-sec, 1-min, 10-min, and 1-hr simulating data averaged over and sampled at those intervals are shown in Figure 3. The probability of occurrence of SSs greater than 5%, 10%, and 25% are shown in Table 1. Both Fig. 3 and Table 1 show SS statistics vary significantly over all timescales, which is consistent with previous findings that 1-min and 1-hr data have different statistics (i.e., Suehrcke and McCormick, 1989 and Gansler et al., 1995). These variations in statistics of SSs down to 1-sec show the importance of sampling data as frequently as possible when studying irradiance fluctuations. Large step sizes have a much greater probability of occurring when using 1-hr averages than when using 1-sec averages. However, due to the nature of block averaging, at longer time intervals, the sample size is small and events with high probabilities of occurrence do not happen very often in a day (Table 1). Still, the cdf of SSs shows a trend toward SS magnitude decreasing as the averaging time decreases – short-time steps will not be as extreme as long-time steps.

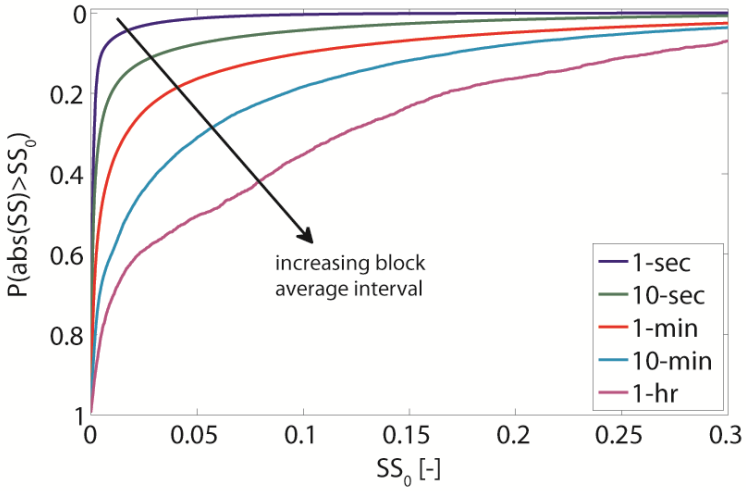

Figure 3: Cumulative distribution function of SSs for block averages over 1-sec to 1-h at EBU2 for 2009. The probability of occurrence of a certain SS (or larger SSs) can be determined by locating the SS on the *x*-axis and going up to intercept the line of the desired block averages. The *y*-value at that point provides the probability. For example, a 25% SS for 1-h block averages occurs 11% of the time or about once per day, on average. The 1-hr curve is based on a smaller sample (the number of 1-hr blocks contained in the 1-year of data) compared to the other curves. For example, the 10-min curve is based on a sample that is 6x larger. This explains why the 1-hr curve is slightly different in shape and slightly more irregular than the other curves.

Table 1: Probabilities of SSs larger than 10%, 25% or 50% at each timescale of block averages along with approximate number of occurrences per day. Occurrences per day were found using an estimated annual average of 10-hours per day when solar altitude angle is greater than 10°.

Block average interval	abs(SS)>0.10		abs(SS)>0.25		abs(SS)>0.50	
	P(abs(SS)>0.10)	#/day	P(abs(SS)>0.25)	#/day	P(abs(SS)>0.50)	#/day
1-sec	0.37%	132	0.02%	6.3	0.0002%	0.1
10-sec	4.29%	155	1.07%	38.4	0.10%	3.5
1-min	9.96%	59.8	3.48%	20.9	0.63%	3.8
10-min	18.39%	11.0	5.26%	3.2	0.85%	0.5
1-hr	35.22%	3.5	11.23%	1.1	0.91%	0.1

While block averages represent sampling data at certain periods where the actual variability is unaffected, moving averages can be used to simulate the effects of fast-ramping energy storage (e.g. flywheels). If the length of the moving average is equal to the time over which energy storage has the capacity to eliminate fluctuations through charging or discharging, then a moving average timeseries will be representative of the PV + storage output timeseries. Moving averages are also relevant to simulating power output of a large PV array or a fleet of PV sites that all sit along the cloud motion vector and are spaced evenly, as used in HP10. In this case, a longer moving average interval will simulate the output of a larger PV plant, since large systems will ideally average over a timescale of $A^{1/2} / V$, where $A^{1/2}$ is the square root of the area of the array and V is the cloud velocity. Moving averages at various timescales are shown in Figure 4 for August 22, 2009.

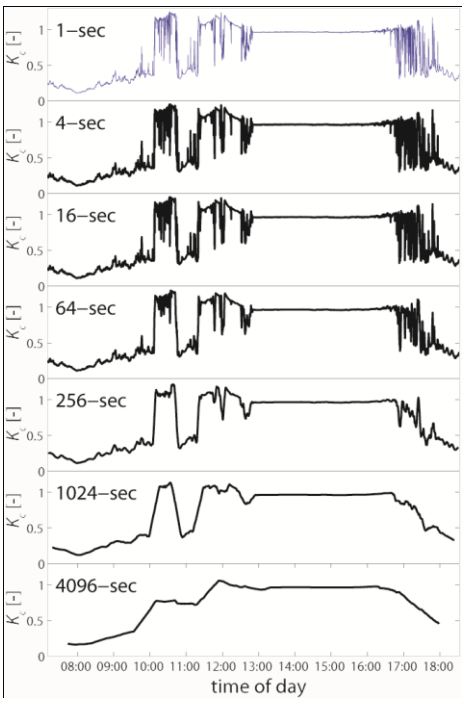

Figure 4: Moving averages of the clear-sky index, K_c, over various averaging intervals for EBU2 on August 22, 2009.

The cdf of RRs for various moving averages is shown in Figure 5, and specific values are shown in Table 2. For the moving averages, increasing the averaging time decreases the probability of a large ramp. For example, for a 4096-sec (about 1-hour) moving average, the probability of a ramp larger than 0.1% s^{-1} is zero. This is intuitive, since the change in the moving average is the change in the step size divided by the averaging interval. Since a 1-sec average under both the block and moving averages simply represents the original timeseries, the 1-sec cdf which appears in both Figures 3 and 5 and Tables 1 and 2 serves as a reference for comparison between the two averaging methods. To create the power plant size column Table 2, we assumed that the 1-sec data was representative of the fluctuations of a typical household PV installation of 2.5kW. Then, using the $A^{1/2} \,/\, V$ relation mentioned earlier, we determined the relationship between moving average intervals and PV plant sizes. This assumed a frozen cloud field traveling at a constant speed, V, over the entire PV plant. While this is unlikely physically, it givens an indication of the best-case scenario and allows for a comparison of fluctuations over various PV plant sizes.

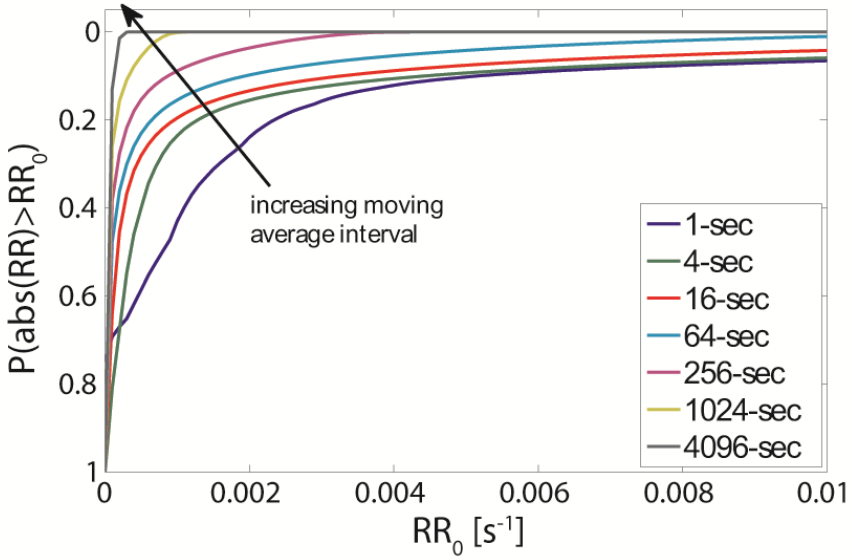

Figure 5: Cumulative distribution function of 1-sec RRs and RRs of moving averages over various timescales (representing large PV plants or plants with energy storage) at EBU2 for 2009. The 1-sec value at $RR_0 = 0$ is 0.75 and not 1.0 due to the very small changes that can occur over 1-sec resulting in $RR < 0.0001$. For all other timescales, $RRs < 0.0001$ never occur.

Table 2: Probabilities of RRs exceeding 0.1%, 1%, or 5% s^{-1} at moving average timescales along with approximate number of occurrences per day. Occurrences per day are based on a 10 sunlight-hour day.

Moving average interval	Power plant size	RR>0.001 s^{-1}		RR>0.01 s^{-1}		RR>0.05 s^{-1}	
		P(RR>0.001)	#/day	P(RR>0.01)	#/day	P(RR>0.05)	#/day
1-sec	2.5 kW	42.98%	15,472	6.55%	2,359	1.35%	486
4-sec	40 kW	23.57%	8,486	5.90%	2,125	0.81%	292
16-sec	640 kW	19.53%	7,031	42.98%	1,511	0.04%	15
64-sec	10.2 MW	15.46%	5,564	1.03%	370	0%	0
256-sec	164 MW	8.84%	3,181	0%	0	0%	0

In order to examine the typical behavior leading up to and after the largest 1-sec ramps, Figure 6 displays the mean (or conditional average) of all 1-sec ramp events greater than 25%, separated into positive and negative ramps. An 'ideal' ramp would simply be a step function from a small K_c to a large K_c or vice versa. However, in practice K_c is variable before or after large ramps as the clear or cloudy period before or after the ramp is often shorter than one minute. For the negative (or clear to cloudy) ramp, there is successive enhancement in clear-sky index in the 1-min before the ramp. This is a manifestation of short clear periods but also of cloud edge enhancement; as a cloud nears the path between the sun and the sensor, some sunlight is reflected off the near edge of the cloud and down to the sensor, while the sun-sensor path is mostly unobstructed. Cloud enhancement leads to irradiances larger than the clear-sky model due to additional diffuse irradiance, resulting in a clear-sky index greater than 1 (Fig. 6). A similar but opposite behavior is observed for the up-ramp. The change in mean clear-sky index from one minute before a large negative ramp to one minute after is about 10%, which indicates a change of state from clear to cloudy. For large positive ramps, this change is only about 3%, and so represents a much smaller change in average state of the sky.

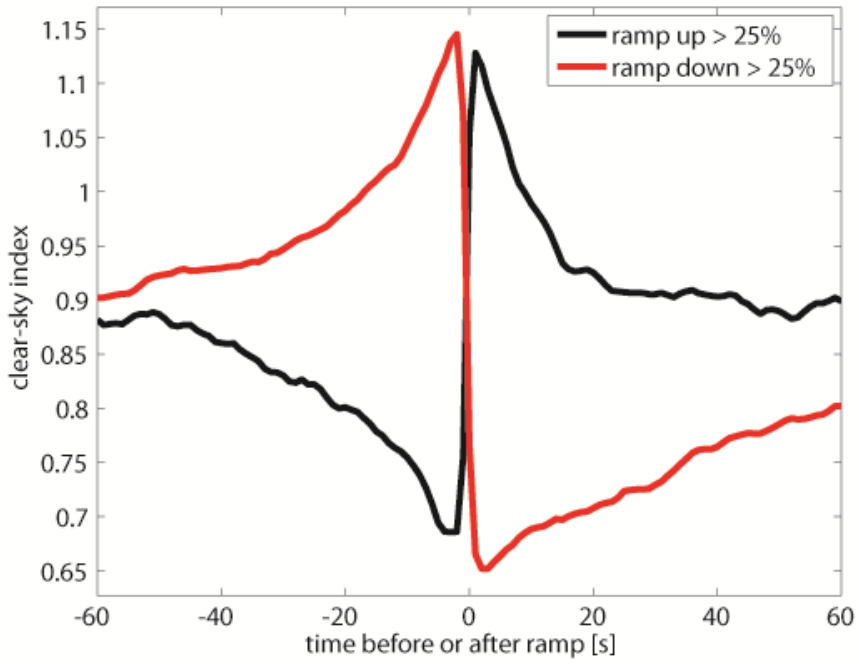

Figure 6: Means of all ramps at EBU2 in 2009 that were greater than 25% s^{-1}, separated into positive and negative ramps. The red line shows the mean of 1006 timeseries starting 1-min before and ending 1-min after a ramp that was more than a 25% s^{-1} decrease in clear-sky index. The black line shows the mean of 511 such timeseries that were centered around a greater than 25% s^{-1} increase in clear-sky index.

Over the entire year, there were five 1-sec ramps up (probability of 2.1×10^{-7} s^{-1}) and 17 1-sec ramps down (probability of 7.3×10^{-7} s^{-1}) with magnitudes greater than 50%. The maximum up ramp was 58% s^{-1} and maximum down ramp was 59% s^{-1}. Thus, as an absolute worst case scenario, a maximum change of 60% over 1-sec can be assumed. The worst irradiance fluctuations were 432 W m^{-2} for an up ramp (June 5, 14:01:42) and 516 W m^{-2} for a down ramp (April 15, 13:33:42), which corresponded to 45% and 54% clear-sky index ramps, respectively. We emphasize, however, that this applies only for one point sensor, and when sites are averaged or PV arrays are considered, these maximum ramps are expected to be strongly reduced.

4.2 Coherence spectra

The coherence spectra over 1-month showing the coherence between EBU2 and the other 5 sites are shown in Fig. 7. At long timescales, the coherence spectra all approach 1. This is expected since hourly and longer weather phenomena such as changes in synoptic cloudiness and atmospheric composition changes affect all sites. Since the coherence spectra were calculated using clear-sky indices, the spectra do not approach 1 as quickly as would be expected with irradiances since the daily cycle of the sun rising and setting is (mostly) removed. The sites are uncorrelated for time scales shorter than 10 min. BMSB, RIMC, and TIOG have the highest coherence values against EBU2 at long timescales. HUBB and MOCC have consistently lower coherence values for timescales longer than 10-min. While it is expected that HUBB will have lower coherence due to it being at the coast and more than twice as far away from EBU2 than the other sites, it is somewhat surprising that MOCC also has such low coherence. MOCC (~1km ESE) and TIOG (~1km WSW) are almost the same distance away from EBU2, albeit in nearly opposite directions, and yet the coherence spectra for each is markedly different. This indicates different weather patterns to the west of EBU2 as to the east. Anecdotal sky observations have confirmed that clouds often evaporate as they move eastward which would result in a smaller coherence.

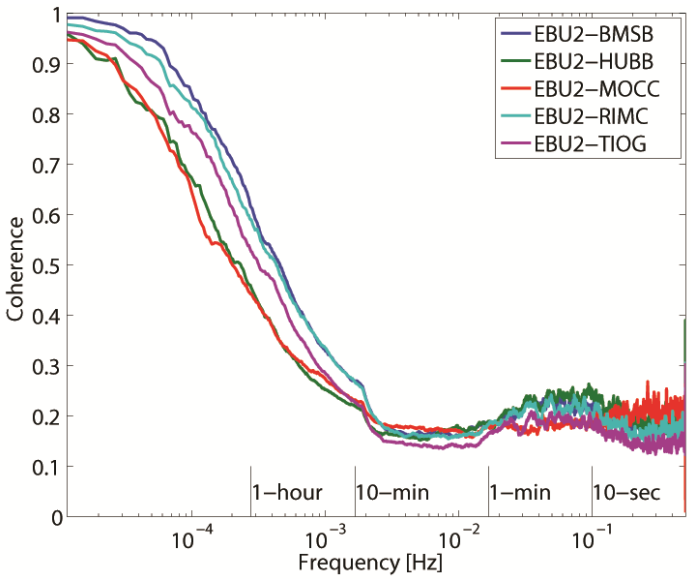

Figure 7: Coherence spectra for EBU2 and each of the other 5 sites for July 31 through August 25, 2009. Each spectrum is smoothed by a moving average smoothing filter for clarity. The apparent vertical lines at the far right of the plot are fluctuations that were not dampened by the smoothing filter since they are too close to the boundary. Different time scales are marked through vertical lines.

4.3 Wavelet decomposition

Wavelet periodograms were computed from the clear-sky index for EBU2 as well as from the clear-sky index for the average of 6 sites for each timescale, $j = 1$ to 12 for the month when 6 sites were simultaneously available. The periodograms from August 22, 2009 over modes $j = 6$ (about 1-min) to $j = 12$ (about 1-hr) are shown in Fig. 8. August 22 was chosen because it has both cloudy and clear periods and because it has a distinct clear period followed by a distinct overcast period both lasting about 30-min. This serves as a validation of our application of wavelets, as we expect this period to produce two peaks at the $j = 11$ mode (34-min). Indeed, the most distinct peaks in the wavelet periodogram shown in Fig. 8 are on the $j = 11$ mode, and occur at about 10:30 and 11:00. We can also see from the periodogram that the dominant timescale of fluctuations between 16:30 and 18:00 was 256-sec ($j = 8$).

This was not obvious by inspecting the original timeseries, but rather is a useful result found through wavelet decomposition.

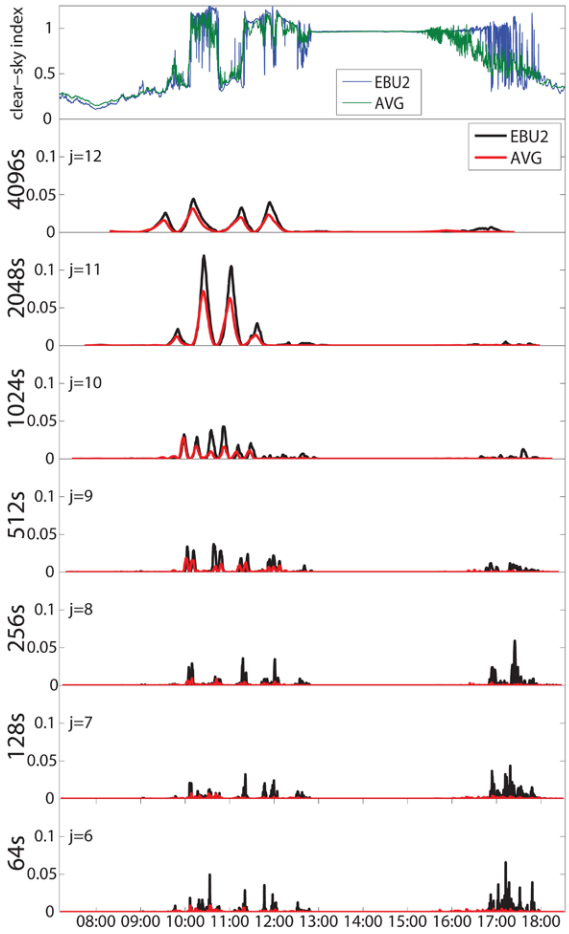

Figure 8: Clear-sky index (blue and green thin lines) and wavelet periodogram (black and red thick lines) of modes $j = 6$ through $j = 11$ for EBU2 and the average of all 6 sites on August 22, 2009.

Inspection of the wavelet periodogram shows that the amplitude is only slightly reduced for the average versus EBU2 at high modes ($j \geq 10$), but the average amplitude is much smaller at modes corresponding to shorter timescales. Since the amplitude of the periodogram at each scale is the variance at that scale, this allows quantifying how averaging multiple sites will lead to a stronger reduction in variability at shorter timescales.

4.4 Fluctuation power index

The reduction in variability as a function of timescale due to averaging sites for the 1-month period is shown in Figure 9, by plotting the fpi for each timescale. Figure 9 also shows the ratio fpi_{EBU2}/fpi_{AVG}, which we will call the variability ratio (VR), for each timescale. The VR is a measure of the reduction in the power (or variance) of fluctuations. A higher VR means a larger reduction in fluctuations, while a variability ratio of 1 means no reduction in variability compared to a single site. For timescales shorter than 256s (about 4-min), VR was close to 6 for the average of the 6 sites. This is consistent with the factor of 6 reduction in variance that we would expect for 6 sites spread far enough apart such that their clear-sky indices can be considered independent of one another (or uncorrelated). At timescales longer than 128s, the fpi ratio decreased in an exponential fashion as the sites become more and more correlated.

Eventually, at 4096-sec, the VR was nearly one, indicating that on timescales longer than 1-hour, the clear-sky indices at these 6 sites are too correlated to cause significant reductions in variability.

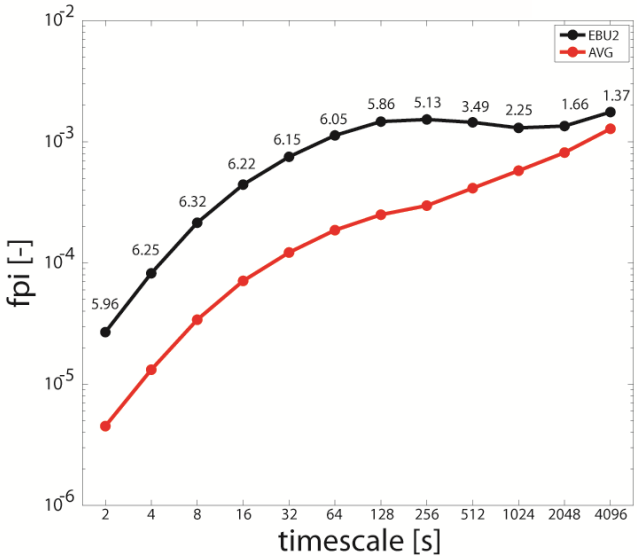

Figure 9: Fluctuation power index for EBU2 and the average of 6 (AVG) sites over 1-month. The numbers above the EBU2 black line are the ratio of fpi_{EBU2}/fpi_{AVG} for each timescale.

4.5 Comparison to Hoff and Perez (2010) model

The VR at each timescale was used to test the Hoff and Perez (2010, hereinafter HP10) theoretical model for the decrease in variability at various dispersion factors D for our six sites. To compute D, cloud speed V and the distance L that clouds must travel in the direction of cloud movement to pass over all sites are required. We estimated distance by assuming that clouds typically travel from west to east over our campus, which was consistent with sky imagery analysis. The east-west distance between HUBB and MOCC, the furthest west and the furthest east sites, respectively, is 2.75 km. We assumed a 5 m s^{-1} cloud velocity given analysis from total sky imagery (Chow et al. 2011), such that $D = \frac{L}{V\,\Delta t} = \frac{550s}{\Delta t}$, with $\Delta t = 2^j$. Since HP10 modeled the ratio of *standard deviation* at the average of all sites to the standard deviation at one site ("relative output variability," ROV), for comparison we take the inverse of the square root of the VR (Fig. 10). In order to plot a single curve for the HP10 ROV, we assumed the lower bound ($ROV = \frac{1}{\sqrt{6}}$) for the reduction in variability in the "limited region." At long timescales (in the "crowded region", expected $ROV = \frac{1}{D}$), D became less than 1 and would have led to an increase in variability at the average versus just one site. This does not make physical sense, and thus the ROV was capped at 1.

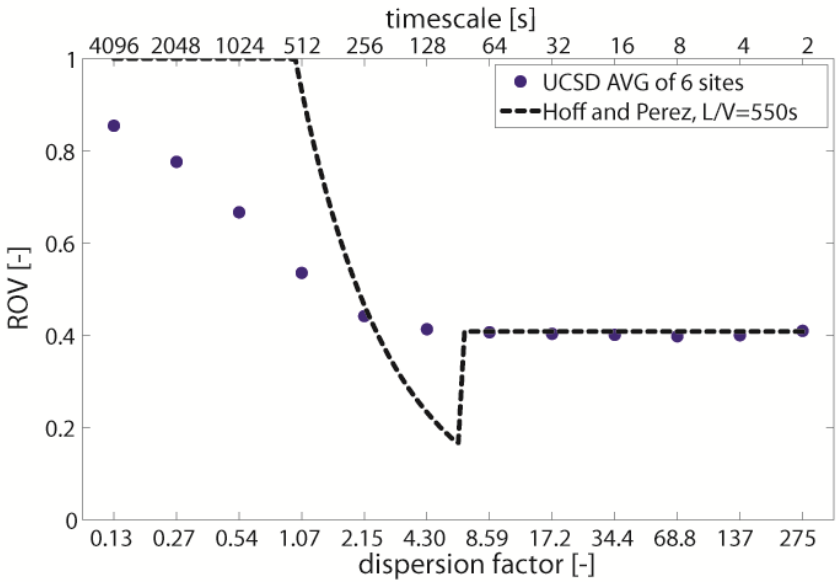

Figure 10: Relative output variability (ROV) for various dispersion factors calculated at each of the timescales used in the wavelet decomposition compared to the Hoff and Perez (2010) model. Hoff and Perez model values greater than 1 were set equal to 1, and the lower limit of $1/\sqrt{6}$ for the "limited region" was used. AVG is the average of 6 sites.

At large D corresponding to short timescales, the sites are indeed independent, and our variability reductions fit the HP10 model. However, the agreement is worse at smaller D corresponding to longer timescales. Our data show no signs of an optimal point in the "limited region". In calculating the optimum point, HP10 assumed that each pair of sites was perfectly correlated (with some timeshift), but we suspect that small lateral dimensions of clouds and cloud condensation / evaporation cause each site to experience a different cloud pattern. Even for two sites only 10 m apart (not shown in Figure 1), on a highly variable day (November 4, 2008) we calculated a maximum time-shifted correlation of only 0.94 (4-sec time shift). In the "crowded region" the ROV increases with increasing time scale, but the HP10 increase is much sharper than the increase found in our data, again presumably due to decorrelation in the cloud fields. Overall, our data indicate a model for the ROV that starts at 1 for very long timescales and then follows an exponential decay that asymptotically approaches $1/\sqrt{N}$ for very short timescales.

5. Conclusion

Clear-sky indices at 1-sec resolution from a solar resource station at UC San Diego were used to compute statistics over an entire year, showing the benefits of high-frequency data to obtain accurate irradiance statistics. Two types of averages were applied to the 1-year data: a block average representing different temporal resolutions of averaged solar radiation data, and a moving average. For the block averages, the probability of a large SS increases with increasing averaging time, but the number of occurrences per day decreases. This was consistent with Fig. 3 in Mills and Wiser (2010) where 10-min and 1-hour block averages had increasingly higher probabilities at any given ramp than the 1-min block average. This same trend was also seen down to 1-sec resolution for one day in Fig. 4 in Mills et. al (2009).

With moving averages, the probability and number of occurrences per day of large RRs both decrease with increasing averaging interval. Using storage with a certain capacity solely for RR reduction will result in a PV power output similar to the moving average over an interval equivalent to the storage-time capacity, and will therefore lead to a reduction in both magnitude and occurrence of extreme RRs. Likewise large solar systems will ideally average irradiance over a time scale of $A^{1/2} / V$, where $A^{1/2}$ is the square root of the area of the array (which is a proxy for array length in the direction of the cloud

velocity V) and a frozen cloud field with infinite lateral dimension is assumed (similar as in HP10). This leads to RR reductions at short time scales.

A new framework for quantifying geographic smoothing was presented through coherence spectra and wavelet analysis of 1-sec K_c data over one month at six sites within 3 km distance. Comparisons of clear-sky indices showed the reduction in variability when averaging 6 sites. Coherence spectra indicated strong correlations between sites up to timescales as short as 10-min and the correlation at longer timescales decreased with distance but not isotropically. A top hat wavelet transform was applied to both EBU2 and the average of the 6 sites. At timescales of 256-sec and shorter, there was approximately a factor of 6 reduction in variance for the average versus just the one site. This indicates that all 6 sites were independent at timescales shorter than about 5-min, reducing variability. At longer timescales, however, the reduction in variance became smaller, and there was almost no reduction in variance at 4096-sec. The variability reductions were compared to the theoretical model presented by HP10. While there was strong agreement between the model and our data at high dispersion factors (short timescales), there was good qualitative but poor quantitative agreement at smaller dispersion factors (long timescales). We suspect that the HP10 assumption that sites are perfectly correlated when time-shifted does not apply for real data. Instead, our data suggests an exponential decay of the relative output variability as a function of dispersion factor.

As PV gains higher and higher penetration, it is important to understand the typical fluctuations on various timescales, as well as the potential for storage, PV array size, and geographic dispersion to dampen these fluctuations. The top hat wavelet transform is a novel approach to deconstruct clear-sky indices into separate timescale components, and was very useful in determining the benefits of storage and geographic averaging. However, cloud statistics depend on local meteorological conditions so more tests in different regions are required to ensure applicability of our results.

Acknowledgments

We appreciate funding from the DOE High Solar PV Penetration grant 10DE-EE002055.

References

Addison, P., 2002. The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine, and finance. Institute of Physics Publishing, London.

Beyer, H. G., Luther, J. and Steinberger-Willms, R., 1991. Reduction of fluctuations in lumped power output from distantly spaced PV-arrays, in: 1991 Solar World Congress Proceedings of the Biennial Congress of the ISES. Denver, Colorado, USA, 19–23 August 1991, pp. 321–326. Pergamon Press, Oxford, New York.

Chow, C.W., Urquhart, B., Lave, M., Kleissl, J., Shields, J., 2011. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. submitted to *Solar Energy*.

Curtright, A.E., Apt, J., 2008. The character of power output from utility-scale photovoltaic systems. *Prog. Photovolt* 16, 241-247.

Gansler, R.A., Klein, S.A., Beckman, W.A., 1995. Investigation of minute solar radiation data. *Solar Energy* 55 (1), 21-27.

Gonzales, C., Wilcox, S., 2004. Sunny Days. Developed by the National Renewable energy Laboratory. Algorithm based on methods developed in Long and Ackerman (2000).

Hoff, T.E., Perez, R., 2010. Quantifying PV power output variability. *Solar Energy* 84 , 1782-1793.

Ineichen, P. Perez, R., 2002. A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy* 73 (3), 151-157.

Kawasaki, N., Oozeki, T., Kurokawa, K., Otani, K., 2006. An evaluation method of the fluctuation characteristics of photovoltaic systems by using frequency analysis. *Solar Energy Materials and Solar Cells* 90. doi:10.1016/j.solmat.2006.02.034.

Kleissl, J., Harper, J., Dominguez, A., 2010. A solar resource measurement network at high spatio-temporal resolution, American Solar Energy Society 2010 Conference, Phoenix, AZ.

Lave, M., Kleissl, J., 2010. Solar variability of four sites across the state of Colorado. *Renewable Energy* 35, 2867-2873.

Lave, M., Kleissl, J., 2011. Modeling Solar Variability Effects on Power Plants. Presentation given at the American Solar Energy Society (ASES) SOLAR 2011 Conference, Raleigh, NC.

Long, C.N., Ackerman, T.P., 2000. Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. *J. Geophys. Res.* 105, 15609-15626.

Mallat, S. 2009. A Wavelet Tour of Signal Processing. Academic Press, San Diego.

Mills, A., Ahlstrom, M., Brower, M. Ellis, A., George, R., Hoff, T., Kroposki, B., Lenox, C., Miller, N., Stein, J., Wan, Y., 2009. Understanding variability and uncertainty of photovoltaics for integration with the electric power system.Berkley National Laboratory. LBNL-2855E.

Murata, A., Yamaguchi, H., Otani, K., 2009. A method of estimating the output fluctuation of many photovoltaic power generation systems dispersed in a wide area. *Electrical Engineering in Japan* 166 (4), 9-19.

Otani, K. Minowa, J. Kurokawa, K., 1997. Study of areal solar irradiance for analyzing areally-totalized PV systems. *Solar Energy Materials and Solar Cells* 47, 281-288.

Perpinan, O., Lorenzo, E., 2011. Analysis and synthesis of the variability of irradiance and PV power time series with the wavelet transform. *Solar Energy* 85 (1), 188-197.

Suehrcke, H., McCormick, P.G., 1989. Solar radiation utilizability. *Solar Energy* 43 (6), 339-345.

Wiemken, E. Beyer, H.G., Heydenreich, W. Kiefer, K., 2001. Power characteristics of PV ensembles: experiences from the combined power production of 100 grid connected PV systems distributed over the area of Germany. *Solar Energy* 70 (6), 513-518.

Woyte, A., Belmans, R., Nijsb, J., 2007. Fluctuations in instantaneous clearness index: analysis and statistics. *Solar Energy* 81 (2), 195-206.