Testing a Wavelet-based Variability Model (WVM) for Solar PV Power Plants

M. Lave and J. Kleissl, *Member, IEEE*

***Abstract* — A wavelet variability model (WVM) for simulating photovoltaic (PV) power plant output given a single irradiance point sensor as input is tested at the 48MW Copper Mountain solar PV plant. 4 days with different amounts of variability are chosen for validation of the model. Comparisons of wavelet fluctuation power index (fpi) and power output ramp rates (RRs) between the input point sensor, WVM simulated power output, and actual power output are presented for the 4 test days. At all timescales, the WVM simulated power output is found to match the variability of the actual power output well, and to be a strong improvement over the input point sensor.**

***Index Terms* – photovoltaic, variability, solar energy, grid integration, wavelet transforms**

I. Introduction

Grid-connected solar photovoltaic (PV) power generation has grown exponentially over the past few years. Continuing this trend to achieve a high penetration of solar power is highly desirable from an environmental point of view, but the variability of solar PV power is considered a major obstacle. Solar PV power production is variable due to the rising and setting of the sun, cloud formations, changes in atmospheric composition, and panel specific variables such as temperature and soiling. Cloud-induced fluctuations have the highest potential to affect the electric grid since they introduce changes in power production on short timescales. The other causes of variability typically change over longer timescales and are often more predictable than clouds. Fortunately, though, since clouds are not homogeneous, geographic smoothing will reduce variability.

Here we explore a wavelet based variability model (WVM) for simulating the variability of a solar PV power plant given limited inputs (the model was first presented in [1]). The WVM will help grid operators understand the impact of solar power variability and the benefits of geographic smoothing. For example, the model can quantify the reduction in relative variability by adding additional PV site locations that increase geographic diversity. In this way, it can be used to estimate the variability of a yet-to-be built solar PV power plant, or the variability reduction benefit of adding more PV to an existing power plant. The model presented here scales up, taking the timeseries of a single irradiance point sensor as input, then using a wavelet transform and estimating variability reductions to simulate the power output of an entire PV power plant, by accounting for the geographic smoothing that will occur over the entire plant.

The main focus of this paper is testing this WVM against an actual power plant on various days. In section II we describe previous works related to this method, section III outlines the procedure for running the WVM, section IV is the validation of the WVM for various days at a test solar PV power plant, and section V is the conclusion.

II. Background

A key factor in solar variability is the number of sites and the size of geographic area that is being considered. While solar power at one site may be highly variable, the relative variability will be reduced when many sites are aggregated. Geographic smoothing increases as more sites are aggregated and as the correlation between sites decreases. While adding more sites always reduces variability, the amount of this reduction depends on the geographic diversity of the site mix and the spatial decorrelation scale of cloud cover. Both affect the correlations between sites: the lower the correlation the more diverse the sites are or the more heterogeneous the cloud field, which leads to more geographic smoothing. Several studies have quantified the correlation between sites, and used this as a metric to simulate power plant variability.

Since the power output from distant PV sites is uncorrelated, the variance of aggregate output of multiple sites will be a factor of N smaller than the variance at just one site, where N is the number of systems. We define the variability ratio (VR) as the ratio of variance at one site to variance at the average of all sites.

$$VR = \frac{\sigma^2_{site\ 1}}{\sigma^2_{avg\ all\ sites}}. \quad (1)$$

Defined this way, an increase in VR will indicate a decrease in the aggregate variability. $VR = 1$ for perfectly correlated sites and $VR = N$ for independent sites. Other values of VR will exist depending on the correlation between sites.

The spatial and temporal scales of decorrelation vary by location and meteorological conditions, leading to studies of VR at various locations and number of sites. A fluctuation factor, defined as the root mean squared (RMS) value of a high-pass filtered 1-min timeseries of solar irradiance was used to demonstrate a 1-month average VR value of 2.8 for 9

This work was supported by DOE High Solar PV Penetration grant 10DE-EE002055 to UC San Diego.

M. Lave is with the Mechanical and Aerospace Engineering Department at the University of California, San Diego, CA 92093 (e-mail: mlave@ucsd.edu).

J. Kleissl is with the Mechanical and Aerospace Engineering Department at the University of California, San Diego, CA 92093 (e-mail: jkleissl@ucsd.edu).

sites located within a 4 km by 4 km grid in Tsukuba, Japan [2]. 1-min steps of power output timeseries in Arizona show VR values of 1.7 to 3.3 for 3 sites 100s of km apart [3]. 5-min clear-sky index timeseries from 4 sites up to 100s of km apart in Colorado had VR values of 2.4 to 4.1 [4]. Wiemken et al. 5-min normalized output from 100 PV sites spread throughout Germany were found to have a VR of 2.7 [5].

Analyzing 1-min data from 52 PV systems spread across Japan, it was 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 [6]. 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.

Step changes (deltas) in block averages of the clear-sky index for 23 Southern Great Plains (SGP) GHI stations, with sites separated by 20 to 440km showed that 1- and 5-min fluctuations had nearly zero correlation between all sites, even at 20km distances, but deltas for times longer than 5-min increased in correlation with decreasing distance [7]. Through simulation, the authors determine that six times less reserve resources are required to mitigate fluctuations for a distributed plant over 20 x 20 km than would be required for a central plant of the same power capacity.

Another study used 24 irradiance sensors – 17 stations in the ARM network and 7 stations in the SURFRAD network – to create virtual networks of irradiance sensors by displacing and time-shifting the sensor measurements [8]. 20-sec, 1-min, 5-min, and 15-min fluctuations become uncorrelated at 500m, 1km, 4km, and 10km, respectively. They extrapolate the correlation relationships to model a homogeneously dispersed solar resource over a 40x40 km grid, and find variability to be reduced by a factor of 80, 40 10, or 4 over the variability of a single site. Further work [9] has shown that the correlation values collapse onto a line when the distance is divided by timescale. Accounting for cloud speed as determined from satellite further decreased the scatter suggesting a universal correlation law.

With a similar objective to the WVM, [10] used a solar irradiance point sensor timeseries to simulate variability of a larger power plant. A cut-off frequency was defined as the intersection of short-timescale and long-timescale linear fits of the irradiance Fourier power spectrum. A smaller cut-off frequency indicated smoothing up to a longer timescale, and cut-off frequency was found to exponentially decay with increasing PV plant area. To simulate a power plant form a single irradiance sensor, a transfer function based on a low pass filter which is scaled by the power plant area is used. Validation against actual PV power output showed good agreement between maximum power fluctuations of simulated and actual data.

III. WVM Procedure

In this section, we give a brief outline of the WVM. A full description of the WVM process is given in [1]. The WVM simulates power plant output given measurements from only a single irradiance point sensor by determining the geographic smoothing that will occur over the entire plant. The simulated power plant may be made up of either distributed generation (i.e., a neighborhood with rooftop PV), centrally located PV as in a utility-scale power plant, or a combination of both. In the WVM, we assume a statistically invariant irradiance field both spatially and in time over the day, and we assume that correlations between sites are isotropic: they depend only on distance, not direction. The main steps to this procedure are:

1) Apply a wavelet transform to the clear-sky index of the original irradiance timeseries, decomposing the clear-sky index into wavelet modes $w_{\bar{t}}(t)$ at various timescales, $\bar{t}$, which represent cloud-caused fluctuations at each timescale.
2) Determine the distances, $d_{m,n}$, between all pairs of sites in the PV power plant; $m = 1, \dots, N$, $n = 1, \dots, N$.
3) Determine the correlations, $\rho(d_{m,n}, \bar{t})$, between the irradiances all sites in the plant at timescales corresponding to wavelet modes. We assume that correlations are isotropic; they do not depend on direction but merely the magnitude of distance between sites. Central to determining these correlations is a correlation scaling coefficient (A). A small A value (1-3) results in lower correlations between sites and indicates small, variable cloud formations, while a large A (>4) means higher correlations and indicates larger, more continuous clouds. The A value varies day-by-day as cloud types and patterns change.
4) Use the correlations to find the variability reduction, $\mathrm{VR}(\bar{t})$, at each timescale.
5) Scale each mode of the wavelet transform by the VR corresponding to that timescale to create simulated wavelet modes of the entire power plant. Apply an inverse wavelet transform to create a simulated clear sky index of areal-averaged irradiance over the whole power plant, $< GHI_{norm}^{sim} >_{pp} (t)$.
6) Convert this area-averaged irradiance into power output, $P(t)^{sim}$ using a clear-sky power model. Clear-sky power models range from simple linear models to more complicated, temperature dependent non-linear models.

IV. Validation of the WVM

To compare the results of the WVM to actual power plant power output, we used the Sempra Generation Copper Mountain (CM) 48MW$_p$ utility scale PV power plant in Boulder City, NV. The footprint of the CM plant is shown in Fig. 1. The plant contains ground-mounted cadmium telluride (CdTe) thin-film PV at a fixed tilt of 25°. Plane of array (POA) irradiance at 1-sec resolution from a Kipp&Zonen CMP11 was used for input to the WVM model, and the WVM output was compared to power output of the entire plant also measured at 1-sec. Total power output was the sum of all inverter power outputs, so ignores AC transmission losses.

At the time of writing, complete data (all necessary measurements for validating the WVM) is only available from CM for August 2011 through November 2011. From this time range, 4 test days with varying daily irradiance profiles were chosen: August 24th, October 1st, October 10th, and November 7th. POA irradiance profiles for each of the days are shown in Fig. 2. Since CM is in Southern Nevada, many of the days in August to November were mostly clear. Clear days are of little interest in validating the WVM, since there is no variability

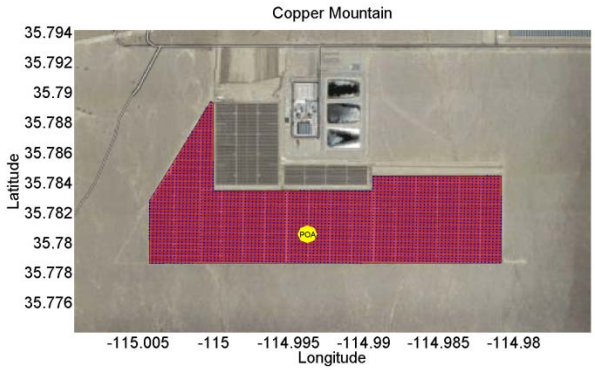

Fig. 1. Polygon showing the footprint of the Copper Mountain power plant. The red shading shows the polygon footprints, while the blue dots show the simulation containers representing small groups of PV panels. The large yellow dot indicates the location of the POA point sensor used as input. The map is approximately 3 × 2 km. Maps © Google Maps.

and hence no effect due to geographic smoothing between sites. On a completely clear day, the power plant output can be determined by using a linear multiplier of the irradiance. However, the WVM is useful and must be validated on variable days, and so the 4 test days chosen have at least some variability. The mostly clear November 7th day will also serve as a test to make sure that the WVM does not introduce any issues during clear periods.

For each of the days, a different correlation scaling coefficient (A) value was found by using 15 plane-of-array reference cells at CM (note the A values can also be determined from a network of irradiance sensors or estimated based on nearby or historical values). The A values for the 4 test days are shown in Table 1. Based on these values, we see that August 24th ($A = 4.16$) and October 1st ($A = 5.16$) had similar typical correlations and hence typical cloud sizes. October 10th ($A = 11.16$), on the other hand, had a much larger A values and so much larger typical clouds leading to larger correlations between sites. This tells us that on August 24th and October 1st, the benefits of geographic diversity across the CM plant will be larger than on October 10th.

Table 1: Correlation scaling coefficients.

	A value
August 24th	4.16
October 1st	5.16
October 10th	11.16
November 7th	2.87

Having determined the correlation scaling coefficients, we can now apply the WVM to estimate the power plant power output variability statistics at CM on the test days. For the purposes of validating the WVM, we will be comparing statistic in units of both clear-sky indices and power. In all plots, we present the input (POA point sensor), WVM output of power plant simulation, and actual power plant data. Table 2 shows the nomenclature that will be used in comparing these three values. In all cases, the goal will be to match the simulated power plant output to the actual power plant output. The POA point sensor is included to show the improvement of the simulation over the input to the WVM.

Following the steps of the WVM outlined in section III, a simulated normalized irradiance timeseries of the entire CM power plant was created, $< POA_{norm}^{sim} >_{pp} (t)$. This was turned into a simulated power output, P^{sim}, using a simple power clear-sky model derived by using a linear multiplier

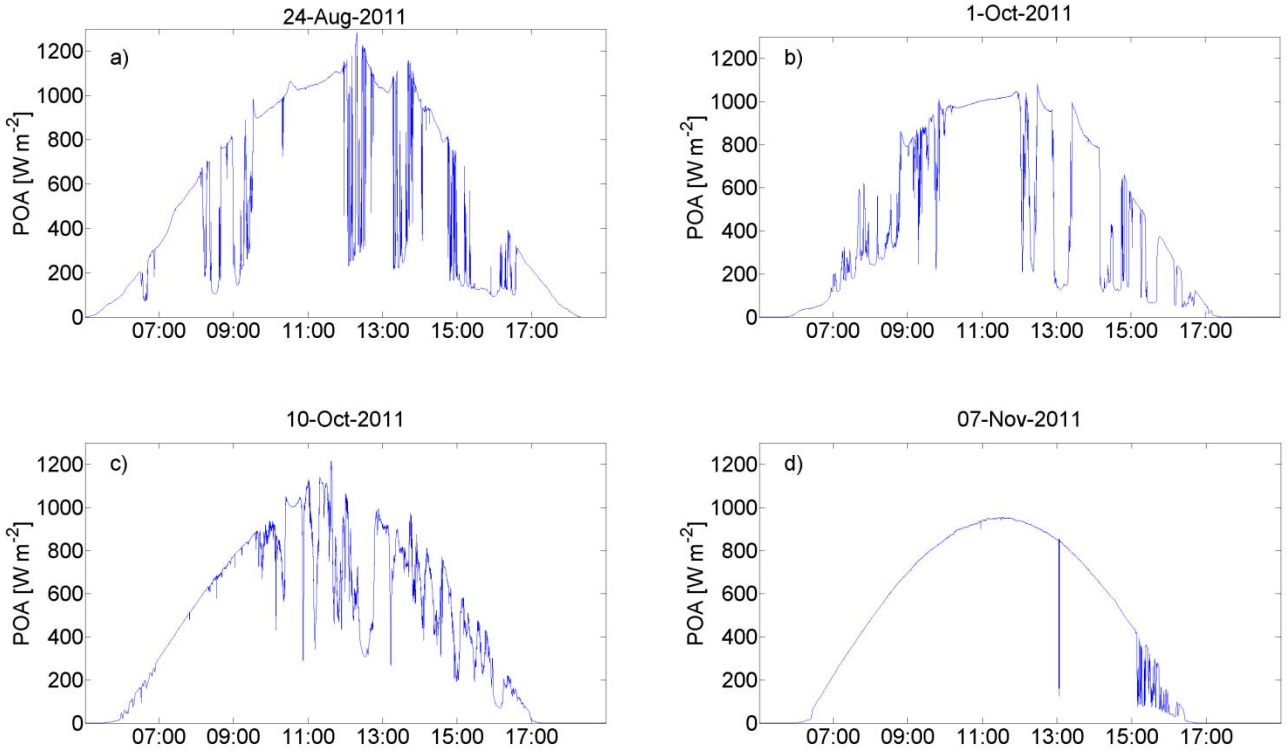

Fig. 2. Daily POA irradiance profiles at Copper Mountain. (a) August 24th, 2011, a variable day with a few midday clear periods. (b) October 1st , 2011, a highly variable day with one midday clear period. (c) October 10th, 2011, a variable day with a mostly clear morning. (d) November 7th, 2011, a mostly clear day with a little variability.

Table 2: Nomenclature for POA, simulated power output, and actual power output in different units. Angle brackets denote averaging. POA_{norm} is the clear-sky index, while $< POA_{norm} >_{pp}$ is the ratio of actual to clear-sky output power. Since the later does not align with the definition of the clear-sky index, the POA_{norm} notation was chosen.

	Clear-sky index [-]	Power units [MW]
POA point sensor (input to WVM)	POA_{norm}	P^{POA}
WVM simulated power plant output	$< POA_{norm}^{sim} >_{pp}$	P^{sim}
actual power plant output	$< POA_{norm} >_{pp}$	P

from a plane of array irradiance clear-sky model (here derived from the Ineichen [11], Boland model [12], and Page Model [13]). Note that while the WVM produces output that is just as resolved as the input (i.e., a 1-sec input timeseries produces a 1-sec output timeseries). However, the CMP11 pyranometer used for POA irradiance measurements has a response time of 5-sec, and so times shorter than 5-sec are not presented here for validation.

For validation, as a variability metric we use the fpi, which is the power content of fluctuations in the wavelet modes at each timescale (for a further definition, see [1, 14]), analogous to the power content of a Fourier transform. The fpi is a valuable metric for measuring the effect of geographic smoothing at each timescale. Geographic diversity will lead to noticeable reductions in the short timescale fpis, with little effect on the long timescales. The goal of the WVM is to create simulated power output that statistically has the same variability distribution across all timescales as the actual power output, and this will be achieved if the fpis of $< POA_{norm}^{sim} >_{pp}$ and $< POA_{norm} >_{pp}$ match. POA and total power output can be slightly offset in time based on the direction of cloud movement and the location of the POA sensor versus the centroid of the power plant. Since the fpi describes the variability content (and total variance) rather than the time of occurrence, it allows measuring the accuracy of the WVM independent of these limitations.

Fig. 3 shows the fpis of the POA sensor, simulated power, and actual power for the 4 test days. The simulated power $< POA_{norm}^{sim} >_{pp}$ matches the actual power $< POA_{norm} >_{pp}$ fpi well at all timescales for these 4 test days. The simulation is a strong improvement over the input $< POA_{norm} >_{pp}$ fpi, especially at short timescales. At long timescales, all three fpis approach the same value, because the effect of geographic smoothing reduces as the timescale increases. This means that if concerned with long timescale (i.e., > 30-min) power plant variability, a simple linear scaling of the variability of the POA point sensor may be appropriate. At shorter timescales, though, a simple scaling of the POA point sensor variability will not be appropriate, and instead the WVM works well to simulate the total power plant output variability.

While the WVM has matched the fpis well on these 4 test days (Fig. 3), ramp rate (RR) statistics are often of greater interest to power plant and grid operators. Fig. 4 shows the cumulative distribution of RRs at 1-sec, 10-sec, 30-sec, and 60-sec for the 4 test days. At all of these timescales, but

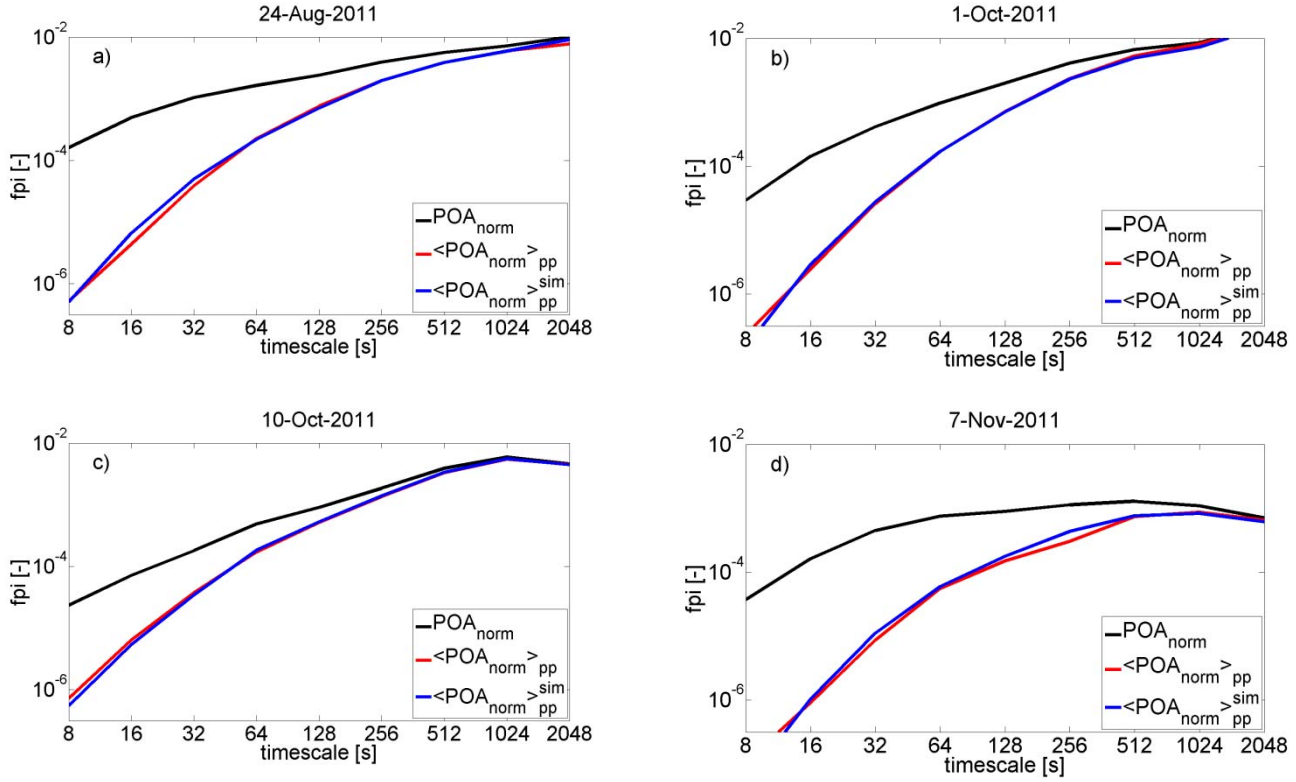

Fig. 3. Fluctuation power index (***fpi***) for the POA point sensor (black), actual power output of Copper Mountain (red), and simulated power output (blue line), for the 4 test days: (a) August 24th, (b) October 1st, (c) October 10th, and (d) November 7th. The simulation matches the actual power output well on these 4 days, and is a strong improvement over the POA point sensor.

especially at short timescales, we see the same trend as in the short-timescale fpis: the RRs are much reduced in P versus P^{POA}, and P and P^{sim} show similar RR distributions. Even on the mostly clear day (November 7th), the RRs of P and P^{sim} match well, even though the average magnitude of extreme ramp rates has been greatly reduced (i.e., the 99% RRs on November 7th are much smaller than the 99% RRs on the other days). This consistent matching of RRs over both highly variable and mostly clear days shows the versatility and broad application of the WVM.

V. Conclusion

The wavelet-based variability model (WVM) for simulating the power output of a solar photovoltaic (PV) plant was briefly described and then tested on 4 days at the Copper Mountain (CM) power plant. The 4 test days represented a range of variable days, from highly variable to nearly clear.

For each of the test days, the correlation scaling coefficient (A) was determined using reference cells at CM, and then used to run the WVM. The results of the WVM simulation were then compared to the actual power output in two ways: by comparing a wavelet fluctuation power index (fpi), and by comparing the ramp rates of power output. Both of these test showed strong agreement in the statistics of simulated and actual power output at all timescales.

The fpi comparisons showed the strong difference between in the power content of short-timescale, cloud-caused fluctuations between the single plane of array (POA) point sensor and the total power plant output. Geographic smoothing was found to greatly reduce the variability at short timescales. The WVM simulated this geographic smoothing well, and fpis of simulated and actual power output matched at all timescales.

Ramp rate (RR) distributions were also matched well by the WVM simulation, and showed a distinct improvement over the POA point sensor RRs, especially at short timescales. It was found that for CM on these test days, linearly scaling the POA point sensor to estimate the variability of the total plant will heavily overestimate the variability.

Future work will focus on characterizing the A values needed to determine correlations as a function of distance, and on validating the model at more locations and over more days in different seasons of the year.

Acknowledgement

We appreciate the help of David Jeon, Leslie Padilla, and Shiva Bahuman from Sempra Energy, Darryl Lopez from Eldorado Energy, and Bryan Urquhart from UCSD for providing and supporting the Copper Mountain data.

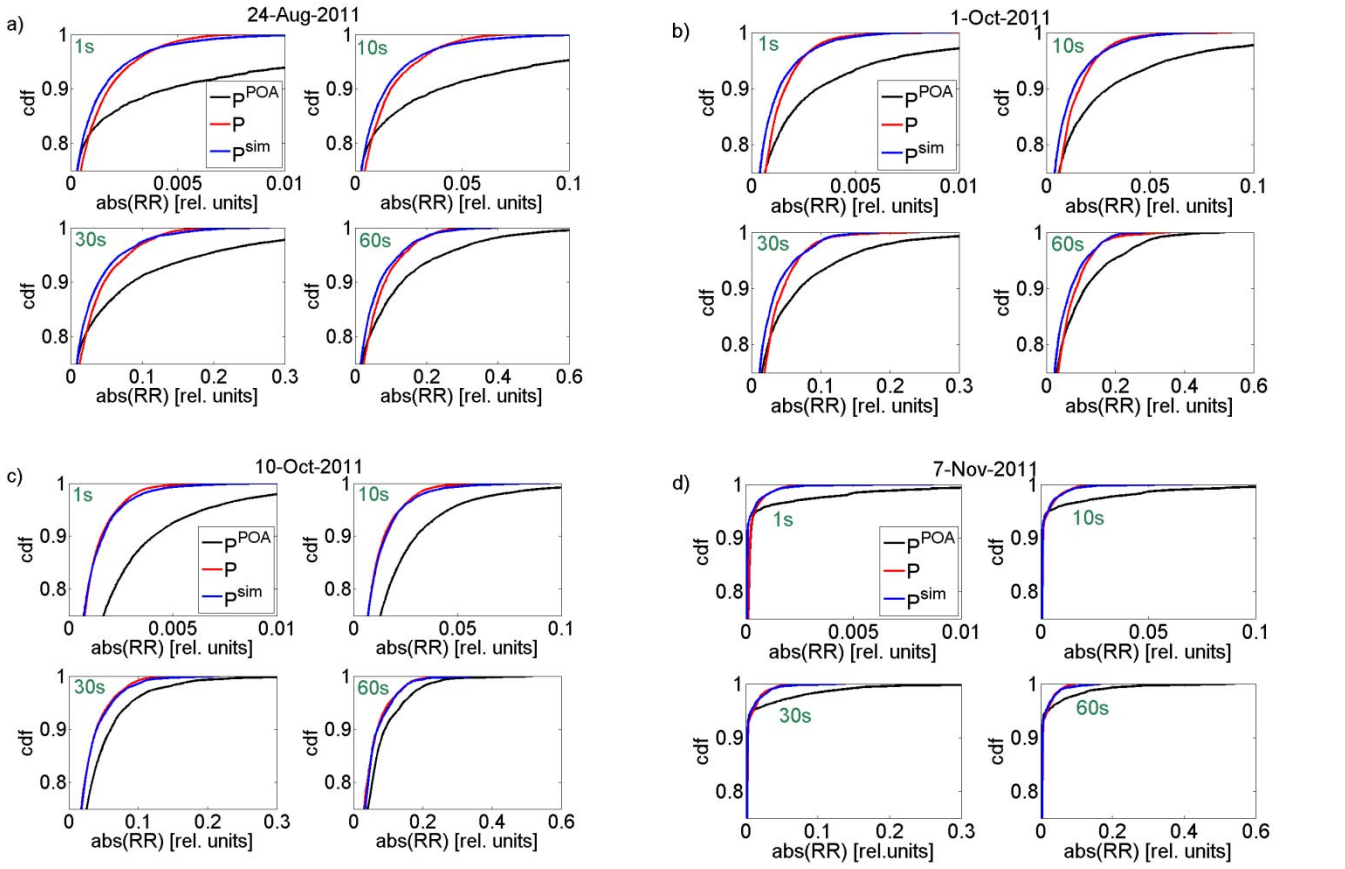

Fig. 4. Extreme (>75 percentile) ramp rate distributions at (clockwise from top left) 1-sec, 10-sec, 30-sec, and 60-sec for P^{POA} (black), P (red), and P^{sim} (blue) for the 4 test days: (a) August 24th, (b) October 1st, (c) October 10th, (d) November 7th. The match between the simulated and actual power ramp rates (red and blue lines) is strong for all days and all timescales.

REFERENCES

[1] M. Lave, J. Kleissl, and j. Stein, "A Wavelet-based Variability Model (WVM) for Solar PV Powerplants," submitted to IEEE Transactions on Sustainable Energy Special Issue on Solar Energy, 2011.

[2] K. Otani, J. Minowa, and K. Kurokawa, "Study on areal solar irradiance for analyzing areally-totalized PV systems," Solar Energy Materials and Solar Cells, *vol. 47, pp. 281-288, Oct 1997.*

[3] A. E. Curtright and J. Apt, "The character of power output from utility-scale photovoltaic systems," Progress in Photovoltaics, *vol. 16, pp. 241-247, May 2008.*

[4] M. Lave and J. Kleissl, "Solar variability of four sites across the state of Colorado," Renewable Energy, *vol. 35, pp. 2867-2873, Dec 2010.*

[5] E. Wiemken, H. G. Beyer, W. Heydenreich, and K. Kiefer, "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, *vol. 70, pp. 513-518, 2001.*

[6] A. Murata, H. Yamaguchi, and K. Otani, "A Method of Estimating the Output Fluctuation of Many Photovoltaic Power Generation Systems Dispersed in a Wide Area," Electrical Engineering in Japan, *vol. 166, pp. 9-19, Mar 2009.*

[7] A. Mills and R. Wiser, "Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power," LBNL Report No. 3884E, *2010.*

[8] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, and T. Hoff, "Short-term irradiance variability: Station pair correlation as a function of distance," Submitted to Solar Energy, *2011.*

[9] T. Hoff and R. Perez, "Modeling PV Fleet Output Variability," Submitted to Solar Energy, *2011.*

[10] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira, and E. Izco, "From irradiance to output power fluctuations: the pv plant as a low pass filter," Progress in Photovoltaics, *vol. 19, pp. 505-510, Aug 2011.*

[11] P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," Solar Energy, *vol. 73, pp. 151-157, 2002.*

[12] J. Boland, B. Ridley, and B. Brown, "Models of diffuse solar radiation," Renewable Energy, *vol. 33, pp. 575-584, Apr 2008.*

[13] J. Page, "The role of solar radiation climatology in the design of photovoltaic systems," in Practical handbook of photovoltaics: fundamentals and applications*, T. Markvart and L. Castaner, Eds., ed Oxford: Elsevier, 2003, pp. 5-66.*

[14] M. Lave, J. Kleissl, and E. Arias-Castro, "High-frequency irradiance fluctuations and geographic smoothing," Solar Energy.

BIOGRAPHIES

Matthew Lave is a PhD student in the Dept. of Mechanical and Aerospace Engineering at the University of California, San Diego (UCSD). Matthew's research focuses on variability analysis of solar radiation and solar power. Projects have included determining ramp rate distributions, the effect of geographic smoothing over large areas, and analysis of solar radiation timeseries using a wavelet transform. His PhD thesis is exploring upscaling and downscaling solar radiation timeseries using wavelets.

Jan Kleissl is an assistant professor at the Dept. of Mechanical and Aerospace Engineering at the University of California, San Diego (UCSD) and Associate Director, UCSD Center for Energy Research. Kleissl received a Ph.D. in 2004 from Johns Hopkins University in Environmental Engineering and joined UC San Diego in 2006. Kleissl supervises 12 PhD students who work on solar power forecasting, solar resource model validation, and solar grid integration work funded by DOE, CPUC, NREL, and CEC. Kleissl teaches classes in Renewable Energy Meteorology, Fluid Mechanics, and Laboratory Techniques at UC San Diego. Kleissl received the 2009 NSF CAREER Award, the 2008 Hellman Fellowship for tenure-track faculty of great promise, and the 2008 UC San Diego Sustainability Award.