Environmental Research Letters

LETTER

Geographic smoothing of solar PV: results from Gujarat

Kelly Klima$^1$ and Jay Apt$^{1,2}$

$^1$ Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
$^2$ Tepper School of Business, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

E-mail: kklima@andrew.cmu.edu

Keywords: solar photovoltaic, geographic smoothing, Gujarat, renewable energy

Supplementary material for this article is available online

Abstract

We examine the potential for geographic smoothing of solar photovoltaic (PV) electricity generation using 13 months of observed power production from utility-scale plants in Gujarat, India. To our knowledge, this is the first published analysis of geographic smoothing of solar PV using actual generation data at high time resolution from utility-scale solar PV plants. We use geographic correlation and Fourier transform estimates of the power spectral density (PSD) to characterize the observed variability of operating solar PV plants as a function of time scale. Most plants show a spectrum that is linear in the log–log domain at high frequencies $f$, ranging from $f^{-1.23}$ to $f^{-1.56}$ (slopes of $-1.23$ and $-1.56$), thus exhibiting more relative variability at high frequencies than exhibited by wind plants. PSDs for large PV plants have a steeper slope than those for small plants, hence more smoothing at short time scales. Interconnecting 20 Gujarat plants yields a $f^{-1.66}$ spectrum, reducing fluctuations at frequencies corresponding to 6 h and 1 h by 23% and 45%, respectively. Half of this smoothing can be obtained through connecting 4–5 plants; reaching marginal improvement of 1% per added plant occurs at 12–14 plants. The largest plant (322 MW) showed an $f^{-1.76}$ spectrum. This suggests that in Gujarat the potential for smoothing is limited to that obtained by one large plant.

1. Introduction

Low-pollution electric power sources, such as solar power, have significant potential to reduce the emissions associated with generating electricity. However, solar photovoltaic (PV) generation is a variable energy source, with large and rapid changes in output $[1, 2]$. This variability of solar PV is sometimes cited as a barrier to its large scale integration into the grid $[3–6]$. Many authors have examined the potential for geographic smoothing of PV in the time domain. That literature is of two broad types. One uses modeled or (less commonly) measured solar illumination. The second uses observed data from power plants. There are few of the second type because data are often proprietary and unavailable to researchers. Here we use 13 months of approximately 1 or 2 minute time resolution data from 30 utility-scale PV plants separated by up to 470 km; we have made these data publically available.

Some irradiance studies suggest that geographic separation may smooth PV variability. The correlation of solar irradiance measured at two locations decreases as the distance between the sites increases $[7–14]$. In addition, cloud models have been used to estimate the smoothing effect of geographic diversity $[15, 16]$, and changes in clear sky index for 23 locations show smoothing is likely for as few as five plants $[2]$. However, studies examining actual generation data provide conflicting results. Five-minute step changes in normalized PV power from one German plant can exceed ±50% but are never larger than ±5% for 100 summed German PV sites $[17]$. Modeled generation at hourly resolution shows smoothing, which is greater on partly-cloudy days $[18]$. Other studies have examined area effects, suggesting that larger capacity plants $[19, 20]$ or plants spread out over a wider area $[21]$ exhibit less variability than smaller or more densely packed solar farms,
respectively. Similarly, geographic smoothing using a large number of smaller plants can reduce variability [22–25], where the maximum variability is theoretically proportional to the square root of the number of plants aggregated [26]. On the other hand, several studies suggest smoothing may not occur. For instance, correlation of real power output for three tracking PV sites in Arizona is high, suggesting smoothing might not be effective there [1]. Similarly, Murata et al find that sites in Japan separated by less than about 200 km are not independent [27], which suggests smoothing might also not be effective there.

Here we examine the potential for geographic smoothing of solar PV in the Indian state of Gujarat using actual generation data from multiple utility-scale solar power installations. We use geographic correlation and Fourier transform techniques to estimate the power spectral density (PSD) [28, 29] and characterize the observed variability of operating solar PV plants as a function of time scale.

2. Data

Real time generation data from the State Load Dispatch Centre of Gujarat Energy Transmission Corporation website are available [30] for 50 solar PV plants in Gujarat, India. These measured power output values are updated at uneven time intervals, generally between 1 and 2 min. We captured website data at 1 min intervals from 17 February 2014 to 16 March 2015. The data capture process, link to our archived data, and power plant characteristics are in the supplementary data.

We used four tests to clean the data. In the first two tests, the full datasets from 13 sites were discarded either because (1) peak generation exceeded the inverter’s capacity (resulting in a flat generation during peak hours), or (2) the resolution of the instruments measuring generation was too coarse (resulting in reported generation at increments of 0.1 MW or larger). For the remaining data, we conducted two tests at each timestep. Reported generation values less than −0.1 MW occurred almost entirely during nighttime hours, but the individual points (as opposed to the full day) were discarded. Finally, we used visual inspection to confirm the ‘goodness’ of the data, resulting in discarding one additional day of a small MW plant that our data cleansing algorithm had not captured.

Figure 1 shows generation data for a 15 MW plant for a year, a week, a clear day, and a partly-cloudy day.

3. Methods

Given an improved understanding of type of variability exhibited by different power sources, a power system operator can understand what combination might be needed to match demand. We use Fourier decomposition to examine the generation data in the frequency domain, where the PSD at a particular frequency indicates the relative amount of variability at the corresponding timestep.

3.1. Calculating the PSD

To handle the observed uneven time steps, we used the Lomb periodogram [31] as coded in Press [29]. An attribute of the Fourier or Lomb methods of estimating the PSD is that increasing the temporal length of the dataset does not reduce the standard deviation of the PSD at any frequency. To increase the signal-to-noise ratio, we used the standard technique of partitioning the dataset into time segments with an oversampling frequency of 4 (such that most data points are in four time segments), resulting in time segments of approximately 1.5 months. Since most time steps were less than 2 min resolution, the highest frequency the data can represent without aliasing (the Nyquist frequency) corresponded to 4 min.

3.2. Scaling plants for comparison

To understand the potential for smoothing plants over thousands of plant combinations, we needed a simplifying process to compare plants. A linear line of best fit would not work due to the unusual shape of the PSDs. Therefore, we make the simplifying assumption that the PSD of a single plant has a flat spectrum (constant PSD) in the log–log domain at low frequencies and an $f^{-m}$ spectrum at high frequencies, such that the PSD can be approximated using coefficients $A$ (the PSD value at low frequencies), $m$ (the slope of the PSD at high frequencies), and $\beta$ (relates to $y$-intercept in log–log domain) via equation (1):

$$\text{PSD}(f) = \frac{A}{1 + \beta f^m} \quad (1)$$

Since the day–night cycle causes solar PSDs to exhibit a peak at 24 h (and its harmonics), we fit this equation in Matlab in the log–log domain to frequencies corresponding to times slower than 48 h and faster than 12 h. Figure 2 shows the PSD of a 5 MW and a 25 MW plant with their respective fitted curves.

To compare the PSD of a single plant to the PSD of interconnected plants in a way that controls for plant capacity, we scale the PSDs using equation (1)’s $A$ values. First we fit equation (1) in the log–log domain to both the PSD of a single plant and the PSD of the interconnected plants to determine the respective $A$ coefficients, $A_{\text{single}}$ and $A_{\text{interconnected}}$. We then multiply the interconnected PSD by $A_{\text{single}}/A_{\text{interconnected}}$ so that $y$-intercept at low frequencies is identical to the single plant PSD. Finally, we refit the PSD of the interconnected plants with $A$, $\beta$, and $m$ such that the lines
of best fit for the single and the interconnected plants cross at $f = 1/24$ h. After scaling, a spectrum with a steeper negative slope (e.g. $f^{-1.76}$) has smaller high-frequency fluctuations than a spectrum with a less-steep slope (e.g. $f^{-1.23}$); in other words, a steeper negative slope represents more high-frequency smoothing. This is the procedure used by Katzenstein [39] for wind plants.

3.3. Understanding the potential for smoothing plants
To understand the potential for smoothing variability in plants, we summed plants, calculated the PSD, and compared the slopes $m$. Averaging over all possible combinations for 2 through 20 plants, we investigate how interconnecting plants can potentially provide smoothing.
The PSD of some plants was noisy due to frequent data dropouts. For some others, the PSD exhibited low-pass filtering at frequencies above $10^{-3}$ Hz (corresponding to approximately 15 min). In what follows, we used 20 plants with spectra that had neither of these features (figure 3). For each period when good generation data existed for all 20 plants, we calculated the PSDs for all possible initial plants and combinations of 2 through 20 of the plants. For each combination, we normalized the interconnected PSD to the single plant PSD using the process described above. We then compared the line of best fit for the two PSDs at particular frequencies by taking the ratio of the single plant value to the interconnected plants value in the $x$–$y$ domain. If no smoothing occurs when solar plants are interconnected, the result should be close to 1 for all frequencies. If there is a reduction in variability then there will be frequencies for which the fraction is less than 1.

### 4. Results

#### 4.1. Calculating individual PSDs

The supplementary data contains a PSD for each of the plants. Most plants exhibit a spectrum at high frequencies ranging from $f^{-1.23}$ to $f^{-1.56}$, with the very large 322 MW plant displaying a $f^{-1.76}$ slope. We find that larger plant size is correlated with a steeper slope with a correlation coefficient of 0.57 at $p < 0.001$ (figure 3). These results agree with the approximately $f^{-1.3}$ spectrum identified by previous research using generation data [1, 32, 33], (as well as the $f^{-0.7}$ spectrum identified when the $y$-axis is the square of the power [34]). This implies that there is still a large need for fast ramping power or demand response to compensate for PV fluctuations at high frequencies. Our results also validate for a real plant the conjecture based on irradiance data that as the capacity of the plant increases, it is likely that the plant will cover a larger horizontal area and thus be able to naturally filter out some of the variability [19, 20].

#### 4.2. Understanding the potential for smoothing plants

To understand the potential for smoothing plants, we summed generation of 20 plants as described above. The time domain generation data are given in the supplementary data, and figure 4 shows the PSDs and resulting lines of best fit for 1, 5, 10, and 20 plants. The amplitude of variability of 20 interconnected PV plants at a frequency corresponding to 1 h has $\sim$45% of variability than that of a single PV plant (figure 5). Just as for wind’s geographic smoothing, the reduction is very dependent on the time scale examined; at 6 h the variability of 20 plants is 80% of that of a single plant. We find a steepening of the $f^{-1.3}$ spectrum toward $f^{-1.66}$ as the number of plants increases.

For reference, we calculated the Bird and Hulstrom Clear Sky Index for direct normal irradiance upon a horizontal surface and global horizontal irradiance upon a horizontal surface [35] at the location of Plant 50 (Charanka) using 1 min resolution for 1 yr. The slope of both of these models is $-1.84$ at high frequencies, and the full PSDs are given in the supplementary data.

Figure 5 shows the fraction of the spectrum of a single plant retained versus the number of interconnected solar plants ($N$) at different timescales. For reference, we show the reduction that would occur if cloud activity in all locations was independent per Hoff and Perez’s $N^{-0.5}$ calculation [26] (which is very similar to empirical results showing a $N^{-0.46}$ relationship [24]). Interconnecting approximately 20 plants
Figure 4. Sample summed power output for ten power plants. (a) The normalized PSD of 1, 5, 10, and 20 plants (in black, red, green, and magenta, respectively) using only the data for the days all 20 plants are available. (b) The corresponding line of best fit for these PSDs.

Figure 5. Fraction of the spectrum of a single plant retained versus the number of interconnected solar plants normalized at $f = 1/24$ h at different timescales. Interconnecting as few as 12 PV plants achieves the majority of the reduction of variability where additional plants reduce less than 1% of variability.
yields a 25%–45% reduction in variability depending on frequency examined. Approximately half the geographic smoothing occurs by interconnecting 4–5 plants, with marginal returns of less than a 1% change per plant after 12–14 plants have been connected. This observed smoothing is not only much less than that suggested by the theoretical $N^{-0.5}$ [26], it also appears to asymptote to a nonzero value at high $N$.

In order to enable comparison with earlier work [2, 9, 18], we also investigated the potential for smoothing by examining the distance dependence of the correlation of generation ramps. We first calculated ramp data (difference in generation between timesteps) between 10:00 and 17:59 local time. We then interpolated these data to even 1 min timesteps, then decimated the values to 5, 15, 30, and 60 min timesteps (thus accounting for selection bias). As shown in the supplementary data, we found a decrease in correlation as distance between plants increases toward 50 km. Near 100 km, the correlation becomes almost constant as a function of distance, as one might expect with the correlation in the solar cycle (e.g., $p = 0.55, 0.25, 0.1a$, and 0.01 for 1 h, 30 min, 15 min, and 5 min timesteps, respectively). These findings agree with the previous results examining distance versus correlation of irradiance data. In addition, to enable comparison with work on wind power [36, 37], we calculated the coherence. As shown in the supplementary data, we found that coherence is high at the 24 h and aliased frequencies. As the distance between plants increases, coherence at higher frequency decreases. However, even up to 242 km distance, there is some coherence at low frequencies.

5. Discussion

Prior to wide-spread solar PV adoption, the power sector will need to address solar generators’ intermittency and variability. Here we study the potential for geographic smoothing of PV using 13 months of observed power production from utility-scale plants in Gujarat, India.

All of the plants examined displayed similar power spectra to those in previously published literature [1, 32, 33], albeit with slightly different slopes. The expected diurnal peaks at 24 h and harmonics are present. At high frequencies, the plants exhibit a spectrum similar to cloud processes [38, 39]. These processes may be a function of the $f^{-5/3}$ spectrum displayed by wind [32] and the $f^{-3}$ spectrum displayed by hydrologic processes [40], or the PV plant may act as a low pass filter [34].

When interconnecting Gujarat PV plants, approximately half of the smoothing can be obtained through connecting 4–5 plants. Reduction in variability exhibits diminishing marginal returns, with interconnecting 12–14 plants yielding a $f^{-1.66}$ spectrum. However, the largest plant, the 322 MW Charanka Solar Park (at 23°54′N, 71°12′E), has an $f^{-1.76}$ spectrum. This suggests that in Gujarat, the potential for smoothing may be limited to that obtained by one large plant.

We further note that the PSD of the clear sky index at the Charanka Solar Park showed a $f^{-1.84}$ slope, a steeper slope than all of our observed individual and combined plants. While this suggests more smoothing to reduce the noise from clouds may be possible, we do not observe this smoothing, suggesting that other limiting factors may be in play. Examining the distance dependence of the correlation of generation ramps, we find in agreement with prior studies (7–14) limits that may be associated with the clear sky index results. This is a physical characteristic that would be common to many locations throughout the world. However, at the higher frequencies where we are measuring the slope, we suggest that different physics dominate, such as those associated with cloud processes (which are in turn a function of local geography, weather, and climate) and mechanical processes associated with power plant machinery. Of course, this work was limited to one region, Gujarat, and if high time resolution data become available from other regions it would be of great interest to determine if these results are general.

The power sector may also wish to compare the potential for geographic smoothing of solar PV to that from other renewable energy types. In comparison to the geographic smoothing of distributed wind plants in Texas [41], the PV plants we examined show substantially less geographic smoothing. The two areas are of comparable size (roughly 400 km × 400 km). Interconnecting 20 wind plants in Texas was found to reduce fluctuations at frequencies corresponding to 6 h and 1 h by 65% and 95% respectively, substantially more than the 23% and 45% observed for PV plants in Gujarat. We also find that when interconnecting observed PV plants, reaching marginal returns of less than a 1% change per plant requires two or three times the number of interconnected plants than for wind (12–14 for PV, 3–6 for wind). Since the area examined is comparable in size to many balancing areas, the relatively small amount of smoothing is likely to be relevant to practical application of solar PV generation at grid scale.

Acknowledgments

This research is based upon work supported by the Solar Energy Research Institute for India and the U.S. (SERIUS) funded jointly by the U.S. Department of Energy subcontact DE AC36-08GO28308 (Office of Science, Office of Basic Energy Sciences, and Energy Efficiency and Renewable Energy, Solar Energy Technology Program, with support from the Office of International Affairs) and the Government of India subcontact IUSSTF/ICERDC-SERIUS/2012 dated 22nd November 2012. We thank Rangan Banerjee and Rhythm Singh (IIT Bombay) for sharing the location of
the dataset (the website of the State Load Dispatch Center for Gujarat, India). We also thank our download and conversion advisors Eric Brunwick, Terrence Wong, Joe Jewell, Jeremy Boulton and Eric Morganson.

References

[1] Curtright A E and Apt J I 2008 The character of power output from utility-scale photovoltaic systems Prog. Photovolt., Res. Appl. 16 241–7
[2] Mills A D and Wiser R H 2010 Implications of wide-area geographic diversity for short-term variability of solar power Technical Report LBNL-3884E Lawrence Berkeley National Laboratory
[3] Edenhofer O et al (ed) 2011 IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation, Prepared by Working Group III of the Intergovernmental Panel on Climate Change (Cambridge: Cambridge University Press) p 1075 (http://srren.ipcc-wg3.de/report)
[4] Widen J, Carppman N, Castellucci V, Lingfors D, Olsson J, Remouf T, Bergkvist M, Grabbé M and Waters R 2015 Variability assessment and forecasting of renewables: a review for solar, wind, wave and tidal resources Renew. Sustainable Energy Rev. 44 356–75
[5] McConnell B W, Hadley S W and Xu Y T 2011 A review of barriers to and opportunities for the integration of renewable energy ORNL/TM-2010/138 Oak Ridge National Laboratory
[6] Denholm P and Margolis R M 2007 Evaluating the limits of solar photovoltaic (PV) in electric power systems utilizing energy storage and other enabling technologies Energy Policy 35 4624–33
[7] KleisI J 2013 Solar Energy Forecasting and Resource Assessment (Amsterdam: Academic)
[8] Long C N and Ackerman T P 1995 Surface measurements of solar irradiance: a study of the spatial correlation between simultaneous measurements at separated sites J. Appl. Meteorol. 34 1039–46
[9] Barnett T P, Ritchie J, Fost J and Stokes G 1998 On the space–time scales of the surface solar radiation field J. Clim. 11 88–96
[10] Lave M and KleisI J 2010 Solar variability of four sites across the state of Colorado Renew. Energy 35 2687–93
[11] Hinkelmann L 2013 Differences between along-wind and cross-wind solar irradiance variability on small spatial scales Sol. Energy 88 192–203
[12] Otani K, Minowa J and Kurokawa K 1997 Study on area solar irradiance for analyzing areally–totalized PV systems Sol. Energy Mater. Sol. Cells 47 281–8
[13] Lave M, KleisI J and Arias-Castro E 2012 High-frequency irradiance fluctuations and geographic smoothing Solar Energy 86 2190–9
[14] Kawaski N, Oozeki T, Otani K and Kurokawa K 2006 An evaluation method of the fluctuation characteristics of photovoltaic systems by using frequency analysis Sol. Energy Mater. Sol. Cells 90 13556–63
[15] Jewell W T and Unruh T D 1990 Limits on cloud-induced fluctuation in photovoltaic generation IEEE Trans. Energy Convers. 5 8–14
[16] Kern E J and Russell M 1988 Spatial and temporal irradiance variations over large array fields Conf. Record of the 26th IEEE Photovoltaic Specialists Conf., (1) vol 2 pp 1043–50
[17] Wiemken E, Beyer H G, Heydenreich W and 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 Sol. Energy 70 513–8
[18] Rowlands I H, Kemery B P and Beausoleil-Morrison I 2014 Managing solar–PV variability with geographical dispersion: an Ontario (Canada) case-study Renew. Energy 68 171–80
[19] Kankiewicz A, Moon D and Sengupta M 2010 Observed impacts of transient clouds on utility-scale PV fields ASES National Solar Conf. (Phoenix, AZ)
[20] Marcos J, Marrooyo L, Lorenzo E, Alviara D and Icoz E 2011 Power output fluctuations in large scale PV plants: one year observations with one second resolution and a derived analytic model Prog. Photovolt., Res. Appl. 19 218–27
[21] Golnas A and Bryan J M 2011 Output variability in Europe’s largest PV power plant Proc. 26th European Photovoltaic Solar Energy Conf. and Exhibition
[22] Golnas A and Voss S 2010 Power output variability of PV system fleets in three utility service territories in New Jersey and California Proc. 2010 35th IEEE Photovoltaics Specialists Conf. (PVSC) pp 535–9
[23] Golnas A and Bryan J M 2011 11-Month power output variability study of a PV system fleet in New Jersey Proc. 26th European Photovoltaic Solar Energy Conf. and Exhibition
[24] Marcos J, Marrooyo L, Lorenzo E and Garcia M 2012 Smoothing of PV power fluctuations by geographical dispersion Prog. Photovolt., Res. Appl. 20 226–37
[25] Van Haaren R, Morjaria M and Fthenakis V 2014 Empirical Assessment of short-term variability from utility-scale solar PV Plants Prog. Photovolt., Res. Appl. 22 548–59
[26] HofT E and Perez R 2012 Modeling PV fleet output variability Sol. Energy 86 2177–89
[27] Murata A, Yamaguchi H and Otani K 2009 A method of estimating the output fluctuation of many photovoltaic power generation systems dispersed in a wide area Electr. Eng. Japan 166 9–19
[28] Cha P and Molinder J 2006 Fundamentals of Signals and Systems: A Building Block Approach (New York : Cambridge University Press)
[29] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 Numerical Recipes in FORTRAN: The Art of Scientific Computing 2nd edn (Cambridge: Cambridge University Press)
[30] SLDC-Gujarat, ‘State Load Despatch Centre (SLDC)—Real-Time System Data’ (https://sldgcgu.com/RealTimeData/GujSolar.asp) accessed by authors every minute from 17 February 2014 to 16 March 2015
[31] Lomb NR 1976 Least-squares frequency analysis of unevenly spaced data Astrophys. Space Sci. 39 497–50
[32] Apt J and Jaramillo P 2014 Variable Renewable Energy and the Electricity Grid (Oxford: RFF Press)
[33] Lueken C, Cohen G E and Apt J 2012 Costs of solar and wind power variability for reducing CO2 emissions Environ. Sci. Technol. 46 9761–7
[34] Marcos J, Marrooyo L, Lorenzo E, Alviara D and Icoz E 2011 From irradiance to output power fluctuations: the PV plant as a low pass filter Prog. Photovolt., Res. Appl. 19 505–10
[35] Bird R E and Hulstrom R L 1991 A simplified clear sky model for direct and diffuse insolation on horizontal surfaces SERI Technical Report SERI/TR-642-761 Solar Energy Research Institute, Golden, CO
[36] Nanahara T, Asari M, Sato T, Yamaguchi K, Shibata M and Maejima T 2004 Smoothing effects of distributed wind turbines: I, Coherence and smoothing effects at a wind farm Wind Energy 7 61–74
[37] Nanahara T, Asari M, Sato T, Yamaguchi K, Shibata M and Maejima T 2004 Smoothing effects of distributed wind turbines: II, Coherence among power output of distant wind turbines Wind Energy 7 75–83
[38] Fliflet A W and Manheimer W M 2006 Measurement of correlation functions and power spectra in clouds using the NRL WARELOC radar IEEE Trans. Geosci. Remote Sens. 44 3247
[39] von Savigny C, Davis A B, Funk O and Pfeilsticker K 2002 Time-series of zenith radiance and surface flux under cloudy skies: radiative smoothing, optical thickness retrievals and large-scale stationarity Geophys. Res. Lett. 29 1825
[40] Kirchner J W and Neal C 2013 Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection Proc. Natl Acad. Sci. USA 110 12213–8
[41] Katzenstein W, Fertig E and Apt J 2010 The variability of interconnected wind plants Energy Policy 38 4400–10