# Comparison of High-Frequency Solar Irradiance: Ground Measured vs. Satellite-Derived

Matthew Lave[1] and Andrew Weekley[2]

[1]Sandia National Laboratories, Livermore, CA, 94550, USA
[2]National Renewable Energy Laboratory, Golden, CO, 80401, USA

***Abstract* — High-frequency solar variability is an important to grid integration studies, but ground measurements are scarce. The high resolution irradiance algorithm (HRIA) has the ability to produce 4-sceond resolution global horizontal irradiance (GHI) samples, at locations across North America. However, the HRIA has not been extensively validated. In this work, we evaluate the HRIA against a database of 10 high-frequency ground-based measurements of irradiance. The evaluation focuses on variability-based metrics. This results in a greater understanding of the errors in the HRIA as well as suggestions for improvement to the HRIA.**

***Index Terms* — solar energy, solar power generation, power grids.**

## I. Introduction

High-frequency solar variability is an important input to accurate distribution grid integration studies. Using low-frequency solar variability results in underestimation of the impact of solar photovoltaics (PV) to distribution grid operations. Underestimations of voltage regulator tap change operations of up 20-70% were found when using 15-minute solar variability instead of 30-second solar variability [1].

However, measurements of high-frequency solar variability are scarce. Sandia has collected a database of 10 high-frequency (30-seconds or better) irradiance measurements, mostly in the western United States. Separately, NREL has developed satellite-derived irradiance variability samples with resolution up to 4-seconds [2] and availability across the United States.

In this paper, we present initial results from ongoing work comparing these two datasets to determine the relative accuracy of the satellite-derived high-frequency irradiance, and to suggest improvements to the satellite downscaling (30-minute to 4-second) methods.

### A. *Importance of High-Frequency Solar Variability*

Central to the work presented in this paper is the assumption that low-frequency irradiance data (e.g., satellite-derived 30-minute data) is at insufficient resolution for distribution grid integration studies. The importance of high-resolution measurements for distribution grid simulations is discussed in detail in a parallel work [3]. In that work, errors in simulated number of voltage regulator tap change operations (a measure of the impact of PV to distribution grid operations) of up to 27% were found simply by using low-resolution data. This shows the importance of high-frequency irradiance samples and is motivation for the work presented in this paper.

### B. *Previous Works to Create High-Frequency Variability*

There have been several previous works that have attempted to create high-frequency solar variability samples from low-frequency inputs. Methods have included using a reference library of high-frequency samples [4] and downscaling using wavelet-based methods [5]. However, there is no widely used method that has been extensively validated: individual studies have typically done local validation at only one specific location.

## II. Ground Measured Solar Variability

Sandia's database consists of 10 locations with high-frequency, ground-based measurements of solar irradiance. The locations are shown in Figure 1. Most locations (8 of 10) have irradiance measurements at 3-sceond resolution or better, allowing for comparison against even the highest resolution of satellite-derived data (4-seconds). Additionally, all locations except Mayaguez have a full year of data, allowing for validation of season trends in solar variability.

## III. Satellite Derived Solar Variability

The NREL-developed high-resolution irradiance algorithm (HRIA) is capable of producing irradiance samples at up to 4-second resolution [2]. 4-second samples are produced based on

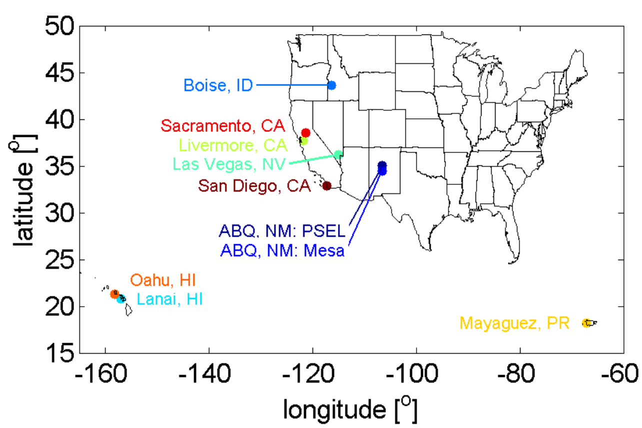


Figure 1: Locations with high-frequency irradiance ground-based measurements. .

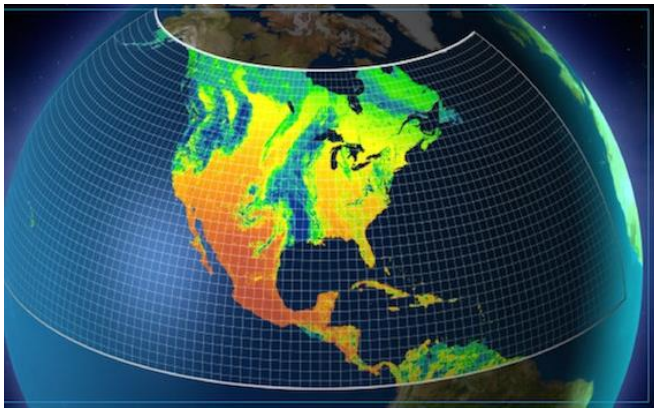

Figure 2: NSRDB coverage [6].

low-frequency, satellite derived irradiance available through the National Solar Radiation Database (NSRDB) [6]. As can be seen in Figure **2**, NSRDB measurements, and, hence HRIA 4-second samples, are available for most of North America. NSRDB irradiance is resolved on a 4 by 4km grid.

The HRIA predicts the temporal variability for a satellite-derived irradiance pixel using two steps.

### *A. SIND Method (30-minute to 1-minute)*

First, the method used for the Solar Integration National Database (termed the "SIND" method) [7] is used to downscale satellite-derived irradiance to 1-minute resolution. This is done using a spatial "patch" of satellite data points: both the pixel containing the location of interest plus several surrounding pixels are used to determine the "Class" of solar variability. Classes roughly range from low variability to high variability, but can also have features related to changing sky conditions (e.g., clear to cloudy). Once a class is selected, it is used to model the 1-minute solar variability. Figure 3 gives an overview of the process used to create 1-minute resolution samples.

### *B. Further Downscaling (1-minute to 4-second)*

Second, the 1-minute samples are further downscaled to 4-seconds using an extension of the Fourier transform [8], as shown in Figure 4. For each class of sky conditions (as defined in the SIND method), a library was assembled of 1-second ground measurements from Oahu, Hawaii [9]. The average Fourier power content was found for each class, and then used to fill in the Fourier spectrum in the 1-minute to 4-second range. 4-second HRIA samples were then created using an inverse Fourier transform.

## IV. Comparison Metrics

Due to the method used to create them, the satellite-derived high-frequency irradiance samples are not expected to exactly match the ground measured irradiances. Specifically, the timing of clouds may not match between satellite-downscaled and ground-measured irradiance variability. Instead, it is important that the overall irradiance variability statistics are captured by the HRIA model. Thus, a direct comparisons using traditional evaluation metrics (e.g., RMSE) which compare measurements at the exact same timestamp are not appropriate.

Since the variability samples are most likely to be used to understand the relative impact of solar variability to electric grid operations, comparison metrics which evaluate the variability over a longer period of time such as a day, month, or year, are more appropriate.

One such metric that will be used for comparison is the daily variability score [1]. The variability score is a way to quantify

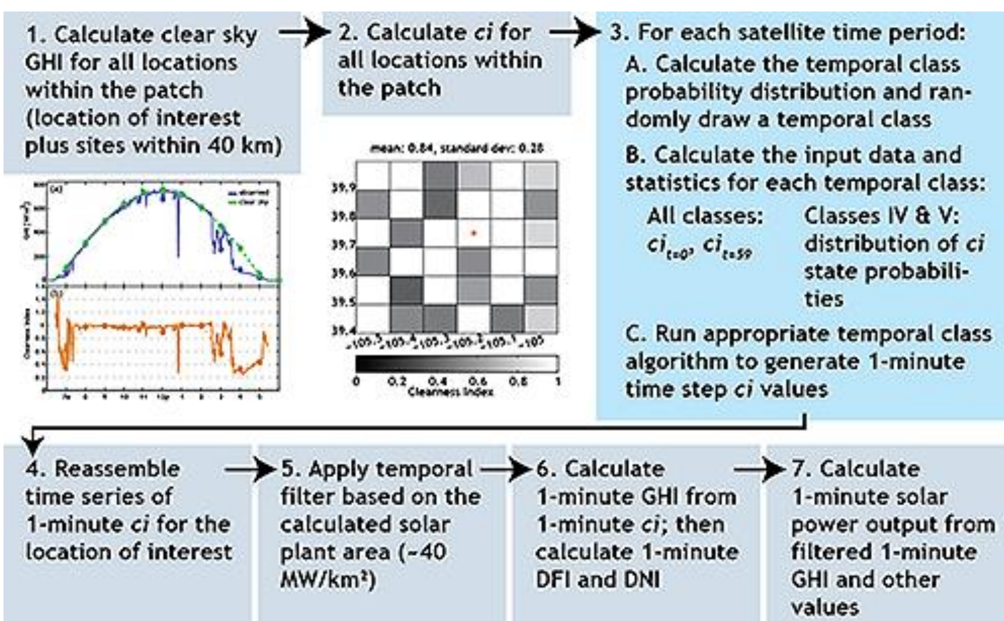


Figure 3: SIND method flowchart, showing how 30-minute satellite data is downscaled to 1-minute resolution [7]. The method uses the cloud index (ci) to classify the irradiance in each satellite pixel.

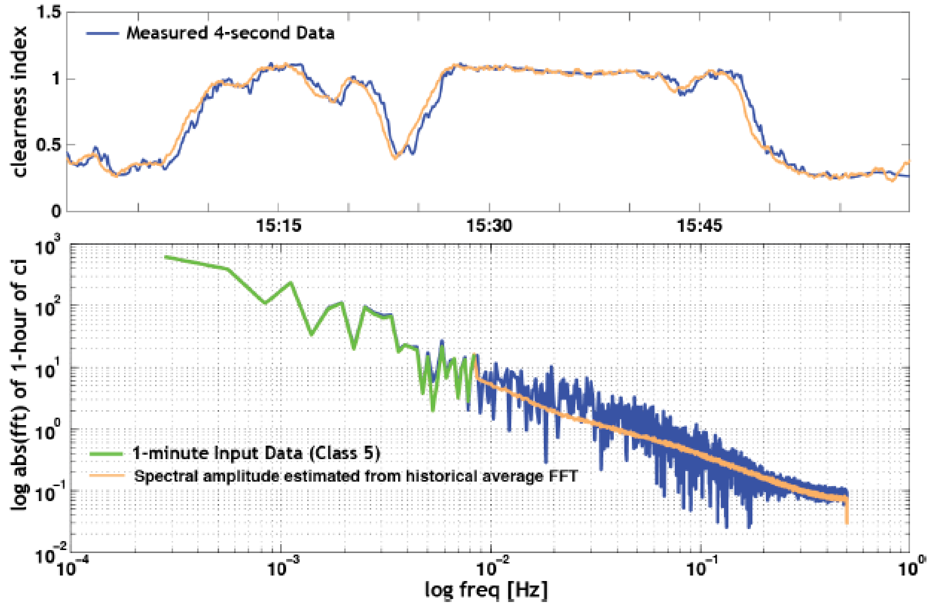


Figure 4: [Top] Clearness index samples: blue measured, orange modeled with HRIA. [Bottom] Fourier transforms: green 1-minute SIND data, blue 1-second measured, orange modeled with HRIA based on average of library samples [8].

solar variability: variability scores are low (0 to 10) for clear conditions which have low variability and large (>100) for highly variable conditions. Two data samples that have the same variability score have similar solar variability. The variability score is the maximum value of the quantity ramp rate magnitude ($RR_0$, expressed in % of 1000 Wm$^{-2}$) times ramp rate probability, multiplied by 100 to give an easier to interpret number:

$$VS_{RRdist} = 100 \times \max[RR_0 \times P(|RR| > RR_0)] \qquad (1)$$

Here, all comparisons are done at 30-second resolution. Since some of the samples in Sandia's database were collected at 3-second resolution and so do not have identical timestamps as the HRIA data, 30-second averages are a fairer comparison. Temporal sensitivity analysis [3] shows at most around 3% errors in distribution grid simulations when using 30-second irradiance data instead of higher-frequency, so 30-second comparisons are sufficient for this analysis.

## V. Initial Results

In the initial comparison, days were separated into clear and cloudy with separate analysis of each. Based on previous experience with the VS [1], values VS<10 are typically clear days. Thus, day when VS<10 were classified as clear and days when VS>10 were classified as cloudy.

### A. *Clear Days*

In general, on clear days the HRIA variability score (VS) is similar to the ground VS. As seen in Figure 5, on clear days the HRIA produces VS values that are close to the ground VS values. In other words, the HRIA does not produce extremely variable days when the ground measurements indicated clear days.

However, a trend is seen in Figure 5 whereby the HRIA VS almost always exceeds the ground VS (i.e., there are more points above the 1:1 line than below).

Figure 6 shows ground measurements and HRIA simulations on a clear day. Even though the HRIA follows the general clear-sky shape, it has some variability that is not reflected in the ground measurements. This is the reason why HRIA VS values are slightly higher on clear days than ground VS values: the HRIA is adding a small amount of variability, even on fully clear days. This is likely caused by the use of the average Fourier transform for each class, as described in Section III. B.

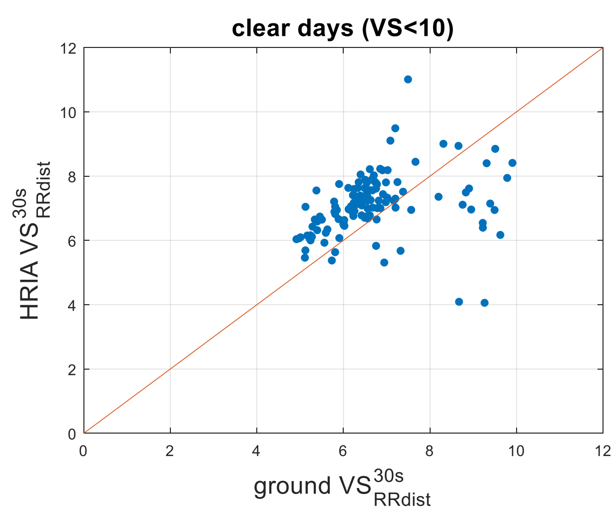


Figure 5: Scatter plot of HRIA 30-second variability score (y-axis) versus ground 30-second variability score (x-axis) on clear days during the year 2013 in Albuquerque, NM..

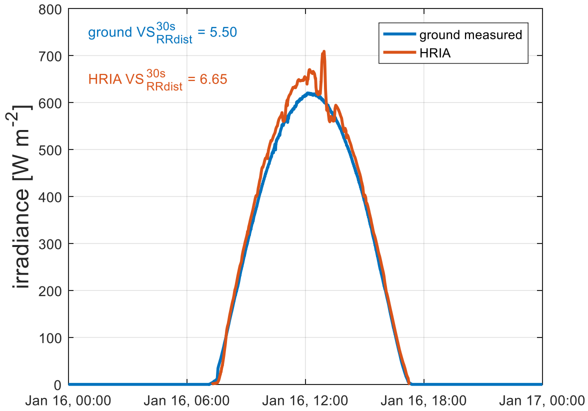


Figure 6: Plot of ground measurements (blue) and HRIA simulated (red) timeseries on a clear day: January 16th, 2013 in Albuquerque, NM. 30-second variability scores are also included in the top left.

A second observation from Figure 6 is that the HRIA simulated irradiance exceeds the clear-sky values at certain times (e.g., around 13:00). Irradiance should only exceed clear-sky values when nearby clouds provide reflections, termed cloud enhancement. On fully clear days such as the one shown in Figure 6, there will be no clouds nearby and so clear-sky values should not be exceeded by as much as they are in the HRIA simulation.

Because of both the slight overestimation of variability during clear conditions and the exceedance of clear-sky values, a possible modification to the HRIA would be to simply assume a clear-sky model when the HRIA predicts a fully clear day. However, care should be exercised to make sure this method does not then underestimate the variability. Mostly clear days with short variable periods might be predicted to be fully clear, and hence the variability underestimated. A mixed statistical approach whereby e.g., 90% of clear days are fully clear and assigned clear-sky values while 10% of clear days are created using the current HRIA method with small amounts of variability added could also be investigated.

### *B. Cloudy Days*

On some cloudy days, the HRIA was found to underestimate the high-frequency irradiance variability. A highly variable day is shown in the top plot of Figure 7. On this day, while the HRIA captured the basic trends in the ground data (e.g., the reduced irradiance around 08:00), but it did not match the many high-frequency up and down ramps seen in the ground data. Thus, the HRIA VS (10) did not match the large ground VS (136).

A partial explanation for this variability underestimation may be the spatial averaging implicit in the HRIA samples. The training library used for the 4-second algorithms is based on the average of 18 point sensors in Oahu. Thus, there was inherent spatial variability smoothing. The SIND method may also suffer from this inherent smoothing,

To show the impact of spatial smoothing, the bottom plot in Figure 7 compares the HRIA sample to a smoothed version of the ground sample that was smoothed using the wavelet variability model [10] to represent the spatial average over the area covered by the 18 point sensors in Oahu. The VS of this smoothed sample (31) is closer to that of the HRIA sample, but the HRIA sample still underestimates the variability when compared to this smoothed sample. Thus, spatial smoothing is likely only a partial explanation of the underestimation.

Since the HRIA is based on two methods of downscaling the satellite irradiance – the SIND method from 30-minutes to 1-minute and the further Fourier downscaling to 4-seconds – the variability underestimation could be caused by one or both methods. For example, if the 1-minute data does not have sufficient variability, the 4-second data will also not have sufficient variability, regardless of the ability of the 4-second algorithm to accurately downscale from 1-minute to 4-seconds. Conversely, if the 1-minute data does have sufficient variability, errors may be in the 4-second algorithm.

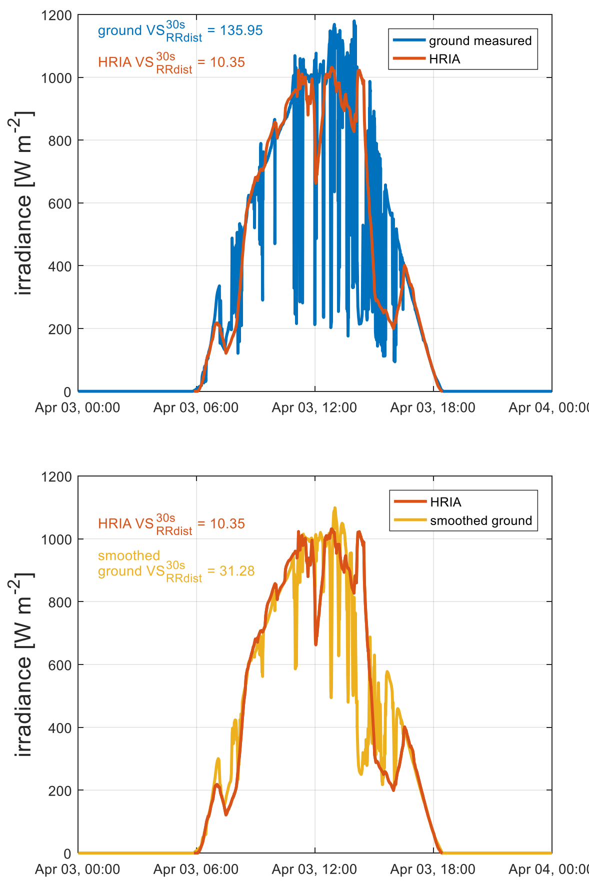


Figure 7: [Top] Plot of ground measurements (blue) and HRIA simulated (red) timeseries on a cloudy day: April 3rd, 2013 in Albuquerque, NM. 30-second variability scores are also included in the top left. [Bottom] Same HRIA sample (red), compared to ground measurements smoothed over the area covered by the 18 pyranometers in the NREL Oahu sensor network (yellow).

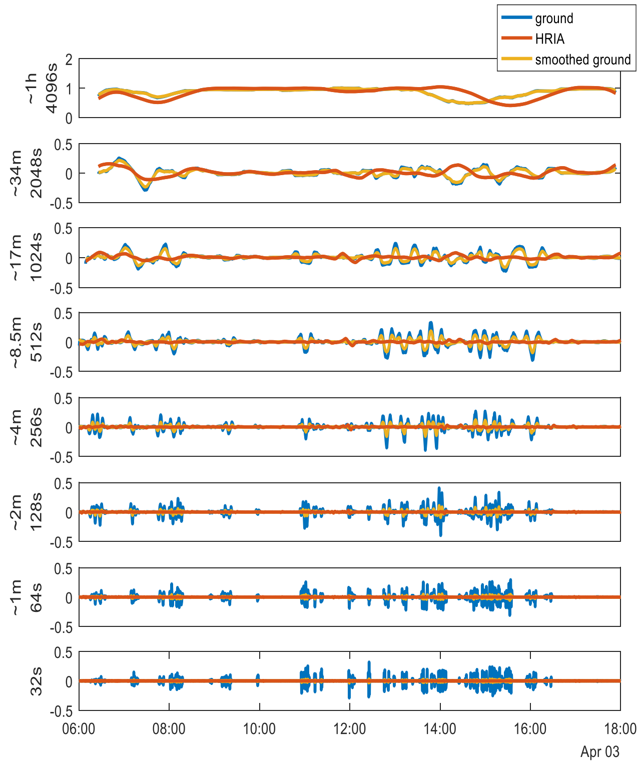


Figure 8: Wavelet decomposition showing variability performance at timescales ranging from 32s to ~1h. Irradiance samples were divided by a clear-sky model to create clear-sky index values before performing the wavelet decomposition.

To help understand the performance at each timescale on this highly variable day, we used a wavelet decomposition. The wavelet decomposition allows for resolution of variability at a variety of timescales [11]. For example, small clouds may cause variability at short timescales (e.g., 30-seconds), while longer-term weather trends will lead to long-term variability (e.g., 1-hour): these are resolved as fluctuations short or long wavelet timescales. Figure 8 shows the wavelet decomposition of the clear-sky index for both the ground, HRIA, and smoothed ground samples on April 3rd, 2013 (the day shown in Figure 7). The smoothing applied to create the smoothed ground sample can be seen to reduce the ground variability at shorter timescales (i.e., the 32s wavelet timescale).

The ~1h timescale HRIA variability matches well (at least in magnitude) with both the ground and the smoothed ground variability. Matches vary at other timescales, but in general the HRIA appears to underestimate the variability on this day at all timescales less than 30-minutes, even when compared to the smoothed ground sample. Specifically, the variability underestimation in the ~1m to ~17m range suggests that the SIND method is largely responsible for the underestimation of variability on this day.

Possible improvements to better match high-frequency variability on a cloudy day include adding more data to the library of lookup samples for both the SIND and the 4-second algorithms, and making sure that the library measurements match the spatial diversity of the ground measurements they are meant to represent.

## Conclusions and Expected Future Improvements

Out of the initial analysis of the HRIA dataset, we have two directed suggestions to improve the method:

1. Use a clear sky model to represent clear days to counter the slight overestimation of variability on clear days. The clear-sky model may be coupled with a probabilistic method of determining occasional partly cloudy periods in otherwise clear days.
2. Ensure that appropriate spatial scaling is used. A possible reason for HRIA underestimation of variability when compared to ground point sensor measurements may be that the HRIA relies of data libraries that already include spatial smoothing, such as libraries based on the average of many sensors.

Further evaluation comparing HRIA datasets to ground-measured data will look at which model, the SIND or the 4-second algorithm, is most responsible for errors in the HRIA dataset (i.e., using wavelet decomposition). For example, variability underestimation may be caused by the SIND method identifying too many periods as mostly clear instead of highly variable.

Through this evaluation and suggested improvements, we hope to leverage the ubiquitous availability of HRIA derived high-frequency irradiance samples to drastically increase the number of high-frequency solar inputs available to grid integration studies. With careful validation and modification of the methodology to ensure accurate simulated variability, HRIA datasets would be very valuable to integration studies in areas with no ground-based measurements.

## Acknowledgment

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Report number SAND2016-5481 C.

## References

[1] M. Lave, M. J. Reno, and R. J. Broderick, "Characterizing local high-frequency solar variability and its impact to distribution studies," *Solar Energy*, vol. 118, pp. 327-337, 8// 2015.

[2] National Renewable Energy Laboratory. (Jan. 22, 2016). *Solar Power Data for Integration Studies*. Available: http://www.nrel.gov/electricity/transmission/solar_integration_methodology.html#algorithm

[3] M. Lave, J. Quiroz, M. Reno, and R. Broderick, "High Temporal Resolution Load Variability Compared to PV Variability," presented at the Photovoltaic Specialists Conference, Portland, OR, 2016.

[4] C. W. Hansen, J. S. Stein, A. Ellis, and C. Lenox, "Assessment of a new simulation approach for estimating PV output variability from satellite imagery," 2011, pp. 002600-002604.
[5] J. Wegener, M. Lave, J. Luoma, and J. Kleissl, "Temporal downscaling of irradiance data via Hidden Markov Models on Wavelet coefficients: Application to California Solar Initiative data," ed, 2012.
[6] NREL. (2015). *National Solar Radiation Data Base*. Available: http://rredc.nrel.gov/solar/old_data/nsrdb/
[7] M. Hummon, E. Ibanez, G. Brinkman, and D. Lew, "Sub-Hour Solar Data for Power System Modeling from Static Spatial Variability Analysis," in *2nd International Workshop on Integration of Solar Power in Power Systems*, Lisbon, Portugal, 2012.
[8] M. Hummon, A. Weekley, K. Searight, and K. Clark, "Downscaling Solar Power Output to 4-Seconds for Use in Integration Studies," 2013.
[9] *NREL: Measurement and Instrumentation Data Center*. Available: www.nrel.gov/midc
[10] M. Lave, J. Kleissl, and J. S. Stein, "A Wavelet-Based Variability Model (WVM) for Solar PV Power Plants," *IEEE Transactions on Sustainable Energy*, 2013.
[11] M. Lave, J. Kleissl, and E. Arias-Castro, "High-frequency irradiance fluctuations and geographic smoothing," *Solar Energy*, vol. 86, pp. 2190-2199, 8// 2012.