NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Sub-Hour Solar Data for Power System Modeling From Static Spatial Variability Analysis

Preprint

Marissa Hummon, Eduardo Ibanez, Gregory Brinkman, and Debra Lew

Presented at the 2nd International Workshop on Integration of Solar Power in Power Systems
Lisbon, Portugal
November 12–13, 2012

Conference Paper
NREL/CP-6A20-56204
December 2012

Contract No. DE-AC36-08GO28308

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) PIX 16416, PIX 17423, PIX 16560, PIX 17613, PIX 17436, PIX 17721

Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

Sub-Hour Solar Data for Power System Modeling from Static Spatial Variability Analysis

Marissa R. Hummon, Eduardo Ibanez, Gregory Brinkman, Debra Lew
National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO, USA
Correspondence: marissa.hummon@nrel.gov

***Abstract*—High penetration renewable integration studies need high quality solar power data with spatial-temporal correlations that are representative of a real system. For instance, as additional solar power sites are added, the relative amount of variability should decrease due to spatial averaging of localized irradiance fluctuations. This presentation will summarize the research relating sequential point-source sub-hour global horizontal irradiance (GHI) values to static, spatially distributed GHI values. This research led to the development of an algorithm for generating coherent sub-hour datasets that span distances ranging from 10 km to 4,000 km. The algorithm, in brief, generates synthetic GHI values at an interval of one minute, for a specific location, using SUNY/Clean Power Research, satellite-derived, hourly irradiance values for the nearest grid cell to that location and grid cells within 40 km. During each hour, the observed GHI value for the grid cell of interest and the surrounding grid cells is related, via probability distributions, to one of five temporal cloud coverage classifications (class I, II, III, IV, V). Synthesis algorithms are used to select one-minute time step GHI values based on the classification of the grid cell of interest in a particular hour. Three primary statistical measures of the dataset are demonstrated: reduction in ramps as a function of aggregation; coherence of GHI values across sites ranging from 6 to 400 km apart over time scales from one minute to three hours; and ramp magnitude and duration distributions as a function of time of day and day of year.**

I. Introduction

Changes in photovoltaic (PV) power or concentrating solar power (CSP) production can exceed 50% of the clear sky generation potential over two to five minutes, as seen in Fig. 1. Rapid changes in solar power output can impact markets with sub-hour intervals, reserve requirements, net load variability, regulation requirements, and the operation of other generators [1]–[4]. To capture these potentially challenging events in renewable energy integration studies, high temporal and spatial resolution solar power production data is needed. The qualities of the dataset, most relevant to an integration study, include:

1) Solar power data must be time synchronized to the weather conditions during each time step and at each geographic location.
2) Solar power data must have sufficient temporal resolution to capture site-specific solar power output ramps.
3) Solar power data must have appropriate spatial-temporal correlations to capture intra-plant and plant-to-plant ramping correlations.
4) Solar power data must have sufficient geographic resolution to represent the relative solar power injection into the power system at each location.

Mills et al. explored the concept of coherence in their analysis of the spatial-temporal irradiance dataset in the South Great Plains network [5]. Several studies show that the aggregation of solar power output, or irradiance, shows a reduction in the variability proportional to the geographic area of aggregation [5]–[10].

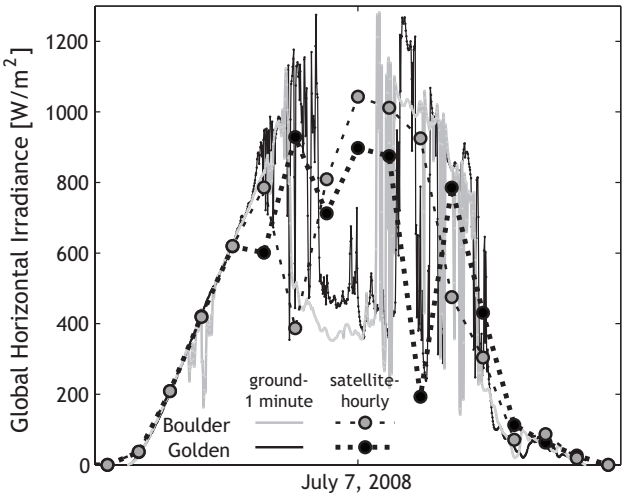

Fig. 1. Simultaneous irradiance measurements from two sites 70.4 km apart at two time scales, $t_{step} = 1$ minute (line) and satellite irradiance values, $t_{step} = 60$ minutes (dotted line with circles). The site near Boulder, CO is located at 39.911°N -105.235°W (gray) and the site near Golden, CO is located at 39.742°N -105.180°W (black).

This paper describes the statistical measures of variability in both temporal and spatial irradiance datasets, the probabilistic relationship between the datasets, and the algorithms used to model sub-hour irradiance data for the Western Wind and Solar Integration Study II (WWSIS II). Lastly, some examples from the WWSIS II are shown, as well as several measures of the quality of the dataset.

II. Irradiance Variability in Spatial and Temporal Statistical Analyses

Sub-hour solar irradiance data is available in select locations throughout the United States [11], while hourly irradiance values from satellite images is available for years 1998-2009 gridded at 0.1°, roughly 10 km. The sub-hour irradiance data used in this study was collected through the National Renewable Energy Laboratory Measurement and Instrumentation Data Center (MIDC). Instrumentation, calibration, and data collection methods for ground-based irradiance data are documented by Stoffel et al. [12]. The locations of ground-measured, sub-hour irradiance used to derive the variability patterns in the dataset for the WWSIS II are in Table I. The hourly irradiance data from satellite images was made available by Clean Power Research (CPR).

CPR uses a semi-empirical model developed by Perez et al. [13], [14] to derive the irradiance at the surface of the earth from the visible channel of satellite images. A cloud cover index, based on an empirical relationship between the top of the atmosphere and surface radiation, is used to modify clear sky global horizontal irradiance (GHI) and estimate GHI at the ground consistent with the cloud cover in the visible satellite image. The CPR SolarAnywhere data product used in this research is the unshifted, snapshot dataset, which includes GHI, direct normal irradiance (DNI), diffuse horizontal irradiance (DFI), surface temperature, and surface wind speed [15].

TABLE I
LOCATIONS WITH MEASURED IRRADIANCE DATA AT A TEMPORAL RESOLUTION OF ONE-MINUTE, MAINTAINED BY THE MEASUREMENT AND INSTRUMENTATION DATA CENTER AT THE NATIONAL RENEWABLE ENERGY LABORATORY.

Location	Latitude	Longitude	Years
Arcata, CA	40.880	-124.080	2009
Los Angeles, CA	33.967	-118.423	2010
Las Vegas, NV	36.060	-115.080	2006-2010
Las Vegas, NV	36.086	-115.052	2006-2009
Alamosa, CO	37.561	-106.086	2009
Boulder, CO	39.911	-105.235	2005-2010
Golden, CO	39.742	-105.180	2005-2010

A. Clearness Index

All irradiance values are converted to a fraction of the clear sky irradiance value called the clearness index (ci) to remove the diurnal effects of the solar zenith angle. The result is a number between 0 and 1, where 1 is clear sky conditions and 0 is no visible irradiance. To produce an accurate clear sky GHI profile for each site at each moment in time, we modified the Bird Clear Sky model [16] to fit the broadband aerosol optical depth, τ_A, to surrounding days that appear to be clear. τ_A is modified because it has the greatest affect on irradiance transmission. The aerosol density varies over the course of the day and by season. For instance, aerosol density decreases as the sun rises because of scattering and increases again as the sun sets. It is lowest in the summer months and highest in the winter months. In order to get a closer fit between the Bird Clear Sky model and actual clear days, we iteratively fit τ_A, hourly, to designated clear days and then transfer those calculated τ_A values to surrounding days. The steps of the algorithm are as follows:

1) Using a wide range of τ_A: $[0 : 0.001 : 1.1]$, the clear sky GHI value is calculated for a specific location at noon for each day of the year.
2) Find the "best noon value" of τ_A by comparing the clear sky GHI to the actual GHI at noon.
3) Using the "best noon value" of τ_A per day, we calculate the hourly clear sky value for all hours of the year.
4) We evaluate the net error between the clear sky and actual GHI values per day; select the days with a mean percentage difference between actual and clear hourly GHI values less than 10%.
5) These selected days are then fit, hour-by-hour for the τ_A value that best fits the clear sky data to the measured GHI values.
6) We propagate the hourly τ_A values for clear days to surrounding days.
7) Finally, we run two smoothing functions: robust local regression using weighted linear least squares and a second degree polynomial on (i) time of day, seasonal variation with a multi-week span (i.e., smooth 8 am values over the entire year) and on (ii) sequential daytime values with an eight-hour span.

For the CPR dataset, the first clear sky data point after sunrise and the last data point before sunset are modified to match the measured data, resulting in a ci equal to 1 for those two data points.

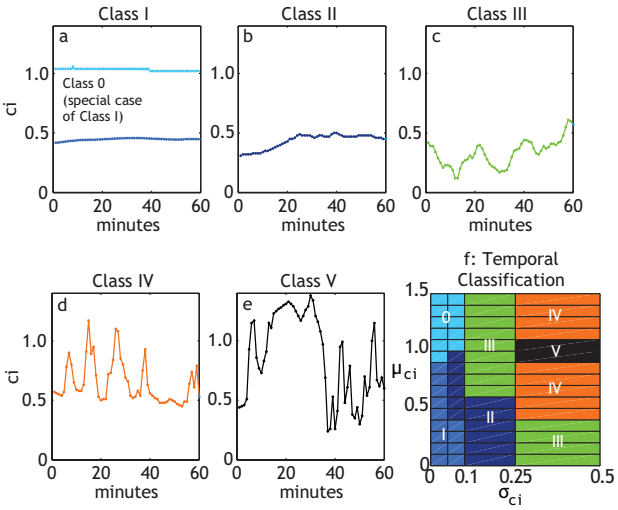

Fig. 2. Examples of the five classes of temporal variability are shown in plots (a) through (e). Classes I-III (a-c) are based on the width of the distribution of ramps. Classes IV-V (d,e) are characterized by a rapid change between two or more different cloud cover densities (e.g., clear sky with small, dense clouds moving at a high altitude). Panel (f) shows how the temporal classes are defined in terms of the mean (μ_{ci}) and standard deviation (σ_{ci}) of the clearness index for 60 consecutive minutes.

B. Temporal Classification and Spatial Statistics

The one-minute ci data is categorized into one of five classifications of cloud cover (Classes I-IV). Visual inspection of 60 consecutive one-minute values of ci, grouped by the mean (μ_{ci}) and standard deviation (σ_{ci}) of the clearness index over that period, yielded classification of irradiance variability into one of five classes, as shown in Fig. 2. Classes I, II, and III show relatively low variability, with less than 0.02%, 0.87%, and 4.53% of one-minute ramps exceeding 0.1 ci, respectively. Classes IV and V demonstrate distinct/sharp shifts in ci with greater than 20% of one-minute ramps exceeding 0.1 ci. Class V irradiance variability is distinct from Class IV not by the magnitude of ramps (in fact, the portion of Class IV ramps between 0.05 and 0.2 exceeds Class V), but by the physical interpretation of the temporal data: Class V represents clear sky conditions with intermittent clouds, while Class IV irradiance variability is characterized by multiple ci states. The temporal classification scheme is based on inferential statistics. By grouping the 60 minute periods by mean and variance, we were able to make an inductive inference from a subset of observations to

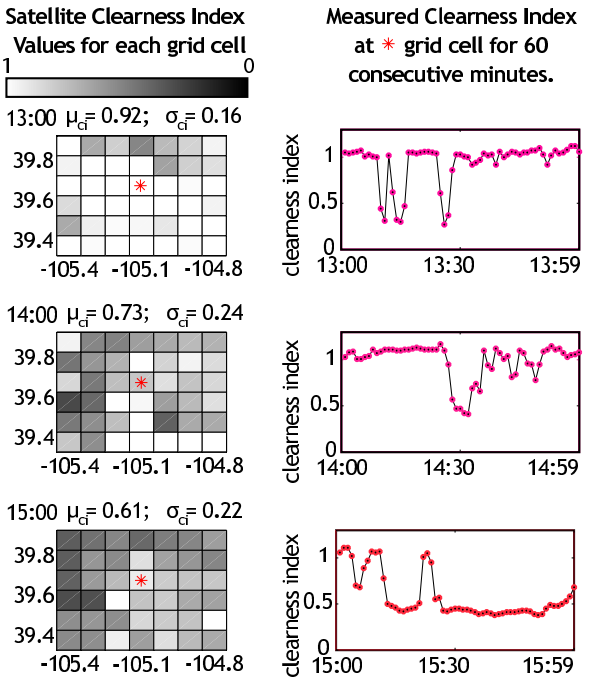

Fig. 3. Three consecutive hours (from 1 pm to 3 pm) on Aug 10, 2005 at SRRL. The spatial pattern shows a trend towards increasing cloudiness over the afternoon.

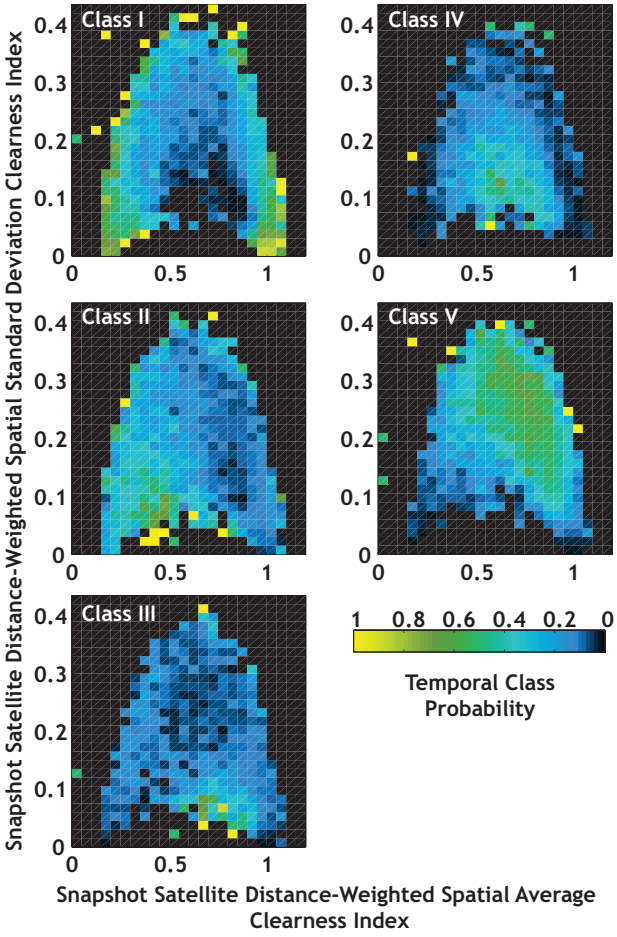

Fig. 4. Probability density distribution of temporal classes across snapshot satellite patch measures: μ_{ci} and σ_{ci}. These are a visualization of the lookup tables used to select a temporal variability class given a patch of satellite data.

the parameters of the whole group. The classification matrix based on μ_{ci} and σ_{ci}, as shown in Fig 2(f).

Similar to the one-minute data, we converted satellite irradiance values into ci values. The satellite hourly data provides a snapshot of the sky at that location. We find that the single snapshot of irradiance has a correlation coefficient of 0.669 with the mean ci of the next 60 minutes, across all temporal classes, while the distance-weighted [17] mean of satellite ci data from approximately 40 nearby locations has a correlation coefficient of 0.736. This set of time-synchronized ci satellite data representing an area of approximately 4,500 km^2 is called a patch. The middle of the satellite data area is called the point of interest, the location of the simulated data. Fig. 3 shows the patch and point of interest (left) associated with each ground-based irradiance measurement (right) for the following 60 minutes. Measures of the variance, mean, and statistical distribution of ci values in each patch, characterized at each hour, form the basis for the probability statistics relating patch data and temporal classification data.

To develop a probabilistic relationship between a patch of snapshot ci and a class of sequential ci values, we evaluate co-located satellite snapshot irradiance values and ground-based irradiance values. The pixel selected by the satellite is within 10 km of the ground-based measurement apparatus. Fig. 4 demonstrates the probability densities for the patch μ_{ci} and σ_{ci} for each of the five classes. The color scale, which ranges from 0 to 1, represents the probability that the combination of patch μ_{ci} and σ_{ci} are indicative of a particular variability class. This information is a visual representation of the lookup tables used to select a variability class prior to synthesizing one-minute time step ci values from hourly patch ci values.

III. Sub-Hour Irradiance Algorithm

The sub-hour irradiance algorithm used to generate the WWSIS II dataset starts with associating a patch of snapshot satellite data with a temporal variability class via the lookup tables shown in Fig. 4. Patches with $\mu_{ci} > 0.9$ and a three-hour rolling average of ci, also greater than 0.9, are considered a special type of Class I, shown as Class 0 in Fig. 2(a). The algorithms for Classes I-III are similar in method, only differing in the distribution of ramps used to generate the sequential values of $ci_{modeled}$. The algorithms for Classes IV-V are similar in concept, though they are based on different qualities of the patch dataset. This section describes the algorithms used to generate modeled ci values for a location. Lastly, we show the modeled and measured data at both the Boulder (Fig. 5) and Golden (Fig. 6) locations concurrently to demonstrate how two sites, with overlapping spatial statistics, result in correlated modeled data.

A. Synthesizing Classes I-III

Classes I-III, and Class 0, are based on the distribution of ramps derived from the six locations, throughout the year, as described in Table I. These classes have a slowly changing

ci, where the ramps, the change in ci from one minute to the next, are independent of the μ_{ci}. The algorithm procedure is:

1) Randomly select 60 ramps from the appropriate distribution for the temporal class.
2) Sum the first ramp of the series with the satellite ci at the top of the hour.
3) Cumulatively sum the first ramp with the second ramp to get the second ci value; continue through the 60^{th} ramp to get the 60^{th} ci.
4) Sum the slope between the satellite ci values at the beginning of the hour and the next hour.

In Fig. 5, the ci patch (left) at 12 pm and 4 pm were classified as Class I and Class III, respectively. The modeled ci, converted to GHI, is shown on the right. The Class I modeled irradiance at 12 pm has less variability than the Class III modeled irradiance at 4 pm.

B. Synthesizing Classes IV-V

Classes IV and V are modeled on the basis that the cloud cover in the sky, and the resulting ci classification, can be characterized as a system undergoing transitions between states. The states represent different opacities of cloud cover and the transition probabilities represent cloud movement over the area. Class IV is modeled as a system with the probability of transitioning between six ci states: 0.01 – 0.03, 0.31 – 0.45, 0.46 – 0.6, 0.61 – 0.75, 0.76 – 0.9, and 0.91 – 1.1. The probability of the state is determined by the distance-weighted [17] proportion of the ci values from surrounding sites in that ci state. The state duration is determined by a random draw from a distribution of state durations, which were derived through empirical analysis of state durations in measured data. In Fig. 5 at 3 pm, the ci patch is classified as Class IV, yet the modeled irradiance is very smooth, because the distance-weighted sites are all roughly the same ci. This means that, while the patch was classified as a highly variable temporal class, the actual ci data generated irradiance data consistent with the patch characteristics.

Class V is a simpler system, consisting of just two states: clear ($ci > 0.9$) and cloudy ($ci < 0.75$). The cloudy, clear state ci value is determined by finding the distance-weighted mean of ci values, such that $ci < 0.9$, $ci > 0.9$, respectively. The transition probability of moving from a clear to a cloudy state is determined by the distance-weighted proportion of sites that are cloudy compared to the total number of distance-weighted sites. The transition probability of returning to a clear state is determined by drawing a random cloudy state duration from empirical analysis, similar to Class IV. Fig.s 5 and 6 have multiple hours with ci patch classified as Class V.

In Fig. 6 (SRRL, Golden) at 1 pm, the Class V algorithm resulted in modeled data that oscillates frequently between a clear sky condition and occasional cloudiness, with a $\mu_{ci} \approx$ 800 W/m^2. At the same time, in Fig. 5 (NWTC, Boulder), the Class V algorithm resulted in a near inversion, with the $\mu_{ci} \approx$ 450 W/m^2. The two locations share a substantial fraction of the spatial irradiance data, yet the statistics are dominated by the local conditions. In the Boulder case, at 1 pm, the surrounding sites are mostly cloudy, while in Golden, the surrounding sites are still mostly clear.

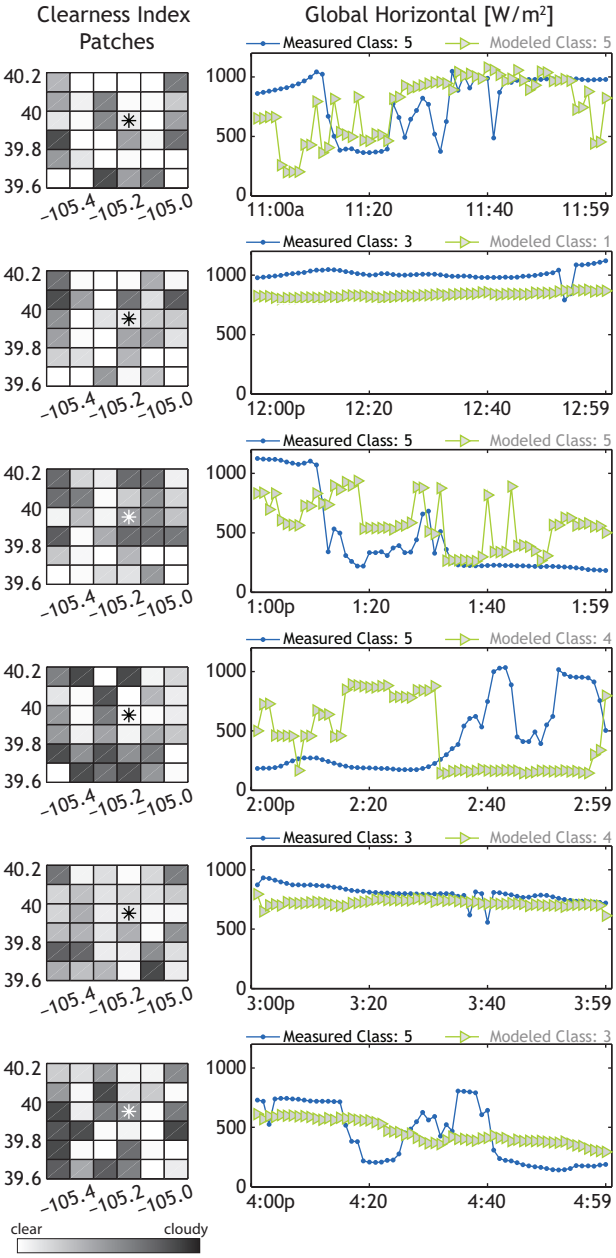

Fig. 5. Snapshot satellite ci (left) and temporal GHI (right) from the National Wind Technology Center at the National Renewable Energy Laboratory on August 22, 2005. Satellite ci patches represent the snapshot at the end of the hour. Measured (black circle) and modeled (right, green triangle) time series GHI, located at 39.911°N, -105.235°W, with a time step of one minute are shown on the right. The temporal variability class of the measured data (in black) and modeled data (in gray) are located above each time series plot.

Sandia National Laboratory [18] conducted an independent validation of the sub-hour algorithm described here and found that the modeled data could reasonably be used for the WWSIS II. Most of the validation sites showed that the modeled data agreed reasonably with the measured data. The modeled data showed significant disagreement with two sites that may have microclimate conditions.

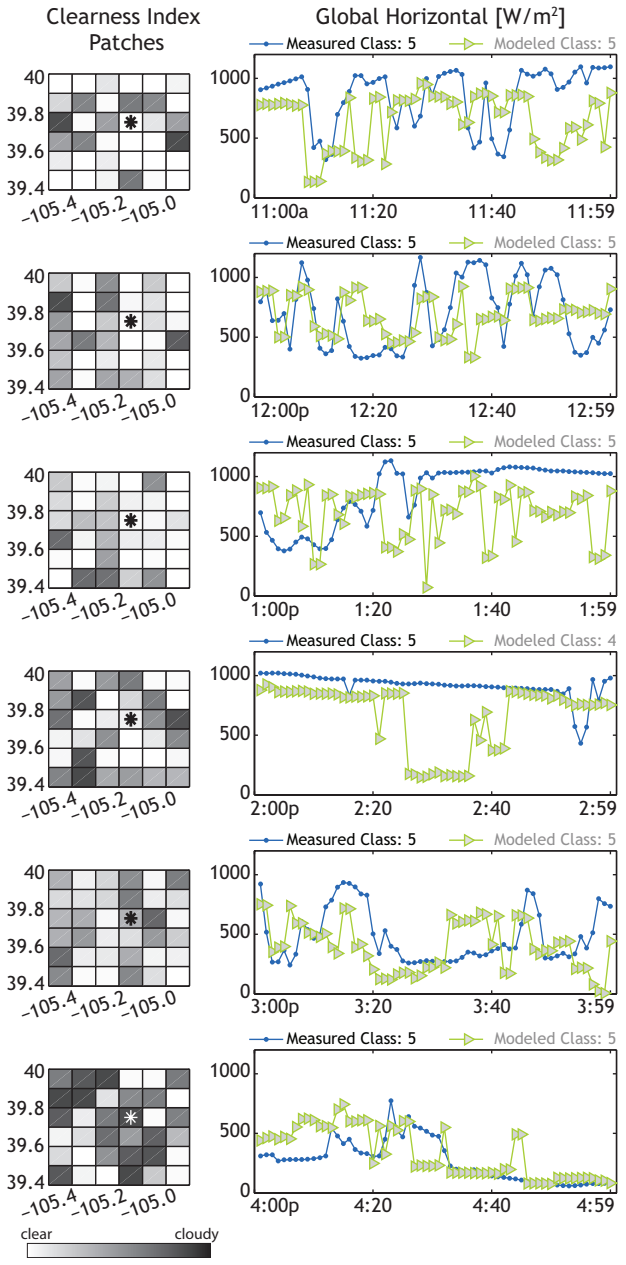

Fig. 6. Snapshot satellite *ci* (left) and temporal GHI (right) from the Solar Radiation Research Laboratory at the National Renewable Energy Laboratory on August 22, 2005. Satellite *ci* patches represent the snapshot at the end of the hour. Measured (black circle) and modeled (right, green triangle) time series GHI, located at 39.742°N, -105.180°W, with a time step of one minute are shown on the right. The temporal variability class of the measured data (in black) and modeled data (in gray) are located above each time series plot.

IV. Western Wind and Solar Integration Study Dataset

The WWSIS II dataset was designed to be incorporated into a security constrained unit commitment (SCUC) and security constrained economic dispatch (SCED) model. The unit commitment (UC) model optimizes, over the defined horizon, the units of generation committed to be "on" based on balancing generation and demand, while meeting reserve requirements and operational constraints of the generators and transmission system. The UC problem is typically solved multiple times at different scales – planning horizons of one or more years, day-ahead UC horizons of 24 hours plus some lookahead to take into account the state of the following day, and hour-ahead UC horizons of an hour or more with some lookahead. Economic dispatch is limited to the real time operation of the electric system; however, most UC models also perform a mock economic dispatch to calculate the cost of operating the generators (variable operation and maintenance and fuel costs). In the past, integration studies [2], [3] have modeled the effects of a high penetration of solar and wind generation on a system at a resolution of one hour, with some statistical analysis of sub-hour variability. Recently, studies underway at the National Renewable Energy Laboratory, including the WWSIS II, Eastern Renewable Generation Integration Study, and the Demand Response and Storage Integration Study, are based on simulating the SCED at a resolution of five minutes, consistent with the shortest time step real time energy market in the United States. Thus, solar power production data needed to have at least a five minute resolution, coherence over a very large area (e.g. the western United States), and exhibit appropriate ramp distributions at individual sites and a reduction in variability as the solar power production data is aggregated.

A. Solar Plant Characteristics

Fig. 7 shows the locations and installed capacity of the solar power plants in the WWSIS II. The study footprint covers two time zones and 3,101,476 km^2. The study modeled both CSP with six hours of thermal energy storage and photovoltaic PV power plants. The PV plants were modeled in one of two ways: (i) utility-scale PV plants were modeled at locations with high resource quality and tied into the nearest bus location in the model, and (ii) rooftop PV plants were modeled in population centers as distributed PV. The characteristics of the modeled plants, including installed capacity and the total number of plants, are recorded in Table II.

TABLE II
Summary of WWSIS Solar Plants

Solar Plants	Total	Capacity Range [MW]	Tracking	Storage Hours
Rooftop PV	474	0.03 - 123	0%	
Utility PV	533	0.1 - 200	45%	
CSP	182	64 - 200		6

The algorithm described in Section III is applied to each hour of snapshot *ci* patch data, to develop modeled data at each solar power plant sites, with a time step of one minute. The result is a time series of GHI values representative of a point source. Both utility-scale and rooftop PV is expected to show some reduction in variability if the footprint of the plant is greater than the shadow of intermittent clouds moving over the plant. Marcos et al. [9] compared the point source irradiance measurements and the power output from several PV plants in Spain and found that the PV plant acts like a low-pass filter. Below the cut-off frequency, the power of the higher frequency modes is lower. The cut-off frequency scales inversely with the footprint of the plant; some example values are given in Table III.

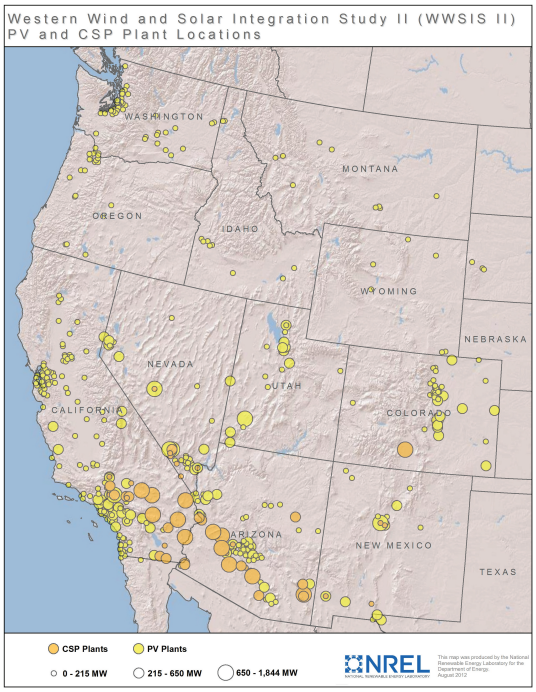

Fig. 7. Western Wind and Solar Integration Study sites for PV and CSP power generation.

TABLE III
TEMPORAL FILTER PROPERTIES APPLIED TO UTILITY-SCALE PV PLANTS WITH A PACKING DENSITY OF 38 MW/KM^2.

Plant Capacity [MW]	Plant Footprint [km^2]	Cut-Off Freq. [Hz]	Filter Limit [min]
25	0.66	0.00252	6.6
50	1.32	0.00178	9.4
100	2.63	0.00126	13.3
200	5.26	0.00089	18.7
380	10.00	0.00065	25.8

To convert irradiance to power output the power conversion model needs several other data streams, including DNI and DFI, as well as meteorological data. The main concern with sub-hour irradiance is the conversion of GHI to DNI and DFI. It is common to see GHI values exceed clear sky conditions, due to scattering off of cloud edges. However, such scattered light would not appear in the DNI spectrum. The GHI is converted to DNI by first transferring GHI values greater than clear sky to linearly interpolated hourly satellite DFI, and then calculating the DNI using the zenith angle (ZA) of the sun, $DNI_t = (GHI_t - DFI_t) * \cos ZA_t$, at each time t. The complete set of visible irradiance measurements is combined with linearly interpolated wind speed and temperature to form the input dataset for PV. For CSP, a complete set of psychrometric is needed in addition to the irradiance and wind speed values. The WWSIS II dataset uses the System Advisor Model [19] to convert solar irradiance into solar power output, using separate modules for PV and CSP modeling.

B. Analysis of Sub-Hour Western Wind and Solar Integration Study Solar Data

In Section I we enumerated several qualities of a sub-hour solar power dataset that are essential for renewable integration studies: individual sites have a representative number, magnitude, and duration of solar power output ramps; aggregates of sites show a reduction in the relative ramp magnitude, duration, and frequency; the correlation of ramps across sites decreases with increasing distance or longer time periods; and individual sites have a representative capacity factor. Fig. 8 shows the power output from a single 200 MW PV plant in Southern Nevada, and the simultaneous power output from all solar plants in the Western Interconnect. In July 2020, the 95^{th} and 99^{th} percentile ramps for the single site in southern Nevada is 10.7 MW and 32.2 MW, respectively, while the 95^{th} and 99^{th} percentile ramps for the entire Western Interconnect is 638.7 MW, 990.7 MW, respectively. When those percentile are scaled relative to the installed capacity, the single site has a ramp equal to 5.35% and 16.1% of the installed capacity for the 95^{th} and 99^{th} percentile ramp, while the aggregated solar power output has a ramp equal to 1.1% and 1.8% of the installed capacity across the region. The reduction in variability is nearly an order of magnitude for the largest ramps (99^{th} percentile).

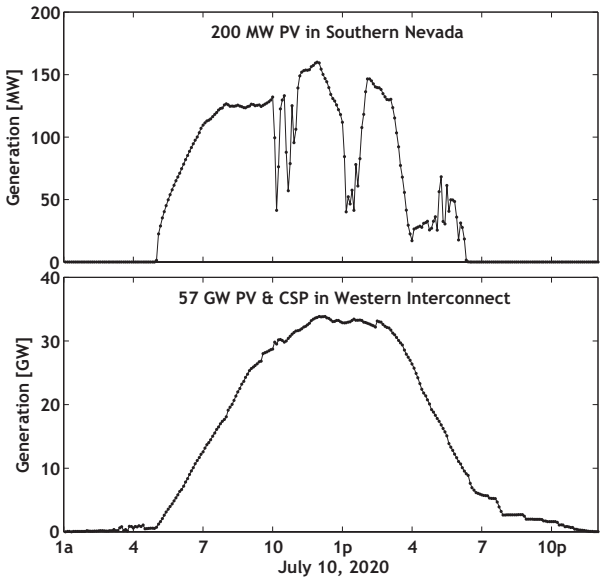

Fig. 8. Solar power output from a single site in Southern Nevada and from all sites in the Western Interconnect, with installed capacity of 200 MW and 57,000 MW, respectively.

Plant-to-plant ramp correlations are performed on ci to remove the correlation of the diurnal cycle of the sun through sky. Fig. 9 shows the correlation coefficient for a group of sites in Arizona spanning inter-plant distances of 10 km to 400 km. The correlation of five time scales of changes in ci are shown: 3 hours, 1 hour, 30 minutes, 10 minutes, and 1 minute. We expect the correlation to decrease as the distance between the sites increases or the time scale of the ci ramp is decreased. In other words, we expect little correlation of solar power output over one minute between two sites; however, long-term weather phenomena, such as afternoon storm clouds, should show up as correlated events in the longer ci ramps. Fig. 9 was inspired by a similar analysis of measured data by Mills et al. [5]

V. CONCLUSION

High penetration renewable integration studies need solar power data at specific location with representative local

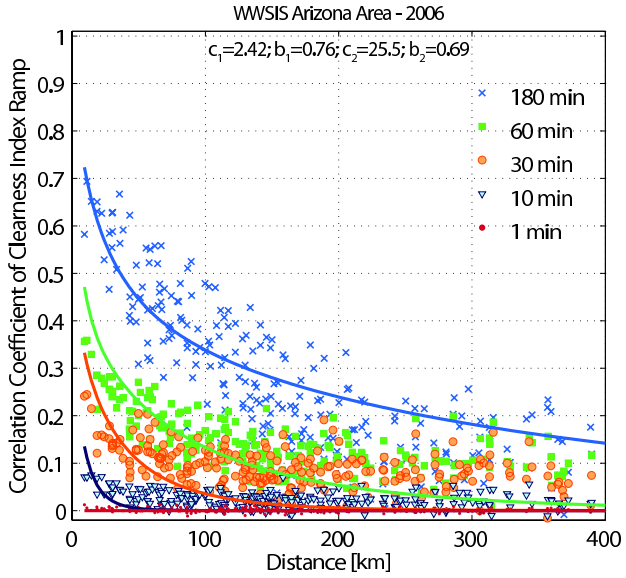

Fig. 9. Site-to-site correlation of clearness index ramps. Ramps are measured by the instantaneous change in ci over a the time, t. The data were fit with $c_1 \exp -d^{-b_1/\bar{t}} + c_2 \exp -d^{-b_2/\bar{t}}$, where d is the distance between sites in kilometers.

irradiance variability and appropriate spatial-temporal correlations across all sites. We have related sequential point-source sub-hour global horizontal irradiance (GHI) values to static, spatially distributed GHI values drawn from satellite images. We have developed and implemented an algorithm for generating coherent sub-hour datasets that span distances ranging from 10 km to 4,000 km. We demonstrate the modeled data at locations with measured data, as well as for a large-scale integration study with over a 1,000 solar power plant locations. We demonstrate two statistical measures of the WWSIS II dataset: comparison of the ci ramps at a single site versus across the entire study region, and coherence of GHI values across sites ranging from 6 to 400 km apart over time scales from one minute to three hours.

ACKNOWLEDGMENT

This work was supported by the U. S. Department of Energy under Contract No. DE-AC36-08-GO28308 with the National Renewable Energy Laboratory. The authors would like to thank the Measurement and Instrumentation Data Center at NREL for collecting, calibrating, and providing quality control/quality assurance for the measured datasets used in this study.

REFERENCES

[1] K. Clark, R. D'Aquila, M. D. MacDonald, N. Miller, and M. Shao, "Impact of dynamic schedules on interfaces," GE, Tech. Rep., 2011.

[2] D. Corbus, M. Schuerger, L. Roose, J. Strickler, T. Surles, D. Manz, D. Burlingame, and D. Woodford, "Hawaii: Oahu wind integration and transmission study," National Renewable Energy Laboratory, Tech. Rep., November 2010.

[3] D. Lew and R. Piwko, "Western wind and solar integration study," National Renewable Energy Laboratory, Tech. Rep., 2010.

[4] S. Venkataraman, G. Jordan, R. Piwko, L. Freeman, U. Helman, C. Loutan, G. Rosenblum, M. Rothleder, J. Xie, H. Zhou, and M. Kuo, "Integration of renewable resources: Operational requirements and generation fleet capability at 20% rps," California ISO, Tech. Rep., August 2010.

[5] A. Mills and R. Wiser, "Implications of wide-area geographic diversity for short-term variability of solar power," Lawerence Berkeley National Laboratory, Tech. Rep. LBNL-3884E, 2010.

[6] C. A. Gueymard and S. M. Wilcox, "Assessment of spatial and temporal variability in the us solar resource from radiometric measurements and predictions from models using ground-based or satellite data," *Solar Energy*, vol. 85, pp. 1068–1084, 2011.

[7] W. T. Jewell and T. D. Unruh, "Limits on cloud-induced fluctuation in photovoltaic generation," *IEEE Transactions on Energy Conversion*, vol. 5, no. 1, pp. 8–14, 1990.

[8] A. Longhetto, G. Elisei, and C. Giraud, "Effect of correlations in time and spatial extent on performance of very large solar conversion systems," *Solar Energy*, vol. 43, no. 2, pp. 77–84, 1989.

[9] 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*, March 2011.

[10] 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, no. 6, pp. 513–518, 2001.

[11] D. Renné, R. George, S. Wilcox, T. Stoffel, D. Myers, and D. Heimiller, "Solar resource assessment," National Renewable Energy Laboratory, Tech. Rep. NREL/TP-581-42301, February 2008.

[12] T. Stoffel, D. Renné, D. Myers, S. Wilcox, M. Sengupta, R. George, and C. Turchi, "Best practices handbook for the collection and use of solar resource data," National Renewable Energy Laboratory, Tech. Rep. NREL/TP-550-47465, September 2010.

[13] R. Perez, P. Ineichen, K. Moore, M. Kmiecik, C. Chain, R. George, and F. Vignola, "A new operational satellite-to-irradiance model," *Solar Energy*, vol. 75, no. 5, pp. 307–317, 2002.

[14] R. Perez, "Time specific irradiances derived from geostationary satellite images," *Journal of Solar Energy Engineering-Transactions of the Asme*, vol. 124, no. 1, pp. 1–1, 2002.

[15] CPR, "Clean power research: Solaranywhere." [Online]. Available: https://www.solaranywhere.com

[16] R. E. Bird and R. L. Hulstrom, "A simplified clear sky model for direct and diffuse insolation on horizontal surfaces," Solar Energy Research Institute, Tech. Rep., 1981.

[17] Weight applied to surrounding sites is determined by: $13 \exp^{-0.1d}$, where d is the distance to the site of interest in km.

[18] C. W. Hansen, "Validation of simulated irradiance and power for the western wind and solar integration study, phase ii," Sandia National Laboratories, Tech. Rep., 2012.

[19] System Advisor Model (SAM), is available for free from NREL at https://sam.nrel.gov/.