Available online at www.sciencedirect.com
ScienceDirect
Solar Energy 84 (2010) 1782–1793
SOLAR ENERGY
www.elsevier.com/locate/solener

Quantifying PV power Output Variability

Thomas E. Hoff[a], Richard Perez[b,*]

[a] Clean Power Research, Napa, CA, USA
[b] ASRC, The University at Albany, Albany, NY, USA

Received 30 September 2009; received in revised form 16 June 2010; accepted 3 July 2010
Available online 11 August 2010

Communicated by: Associate Editor David Renne

Abstract

This paper presents a novel approach to rigorously quantify power Output Variability from a fleet of photovoltaic (PV) systems, ranging from a single central station to a set of distributed PV systems. The approach demonstrates that the relative power Output Variability for a fleet of identical PV systems (same size, orientation, and spacing) can be quantified by identifying the number of PV systems and their Dispersion Factor. The Dispersion Factor is a new variable that captures the relationship between PV Fleet configuration, Cloud Transit Speed, and the Time Interval over which variability is evaluated. Results indicate that Relative Output Variability: (1) equals the inverse of the square root of the number of systems for fully dispersed PV systems; and (2) could be further minimized for optimally-spaced PV systems.

Keywords: Solar resource; Variability; Distributed PV generation; Photovoltaics; Utility

1. Introduction

There is a growing concern about the potential of photovoltaic (PV) power Output Variability having a negative effect on utility grid stability. High levels of high frequency variability during partly cloudy conditions have been reported at some central PV generating stations and have contributed to create an awareness of this issue to the point where some in the utility industry believe it could constrain the penetration of grid-connected PV. The US Department of Energy (2009) recently convened a workshop on the subject of high penetration PV and identified PV variability as a top concern and research priority.

This concern for short-term variability likely originates from utilities' experience with wind generation where "ramping" – i.e., the sudden coming online of a large number of units due to e.g., a frontal passage – is a known issue that requires mitigation, such as reserve, storage, and/or localized operational forecasts (Blatchford, 2009; Noah, 2008). Although fundamentally different, the short-term variability of solar generation identified at central generation plants raises the concern that accelerated on–off ramping from multiple plants may create major problems that will require major mitigation efforts.

The objective of the present work is to provide a general model that quantifies the short-term power Output Variability resulting from an ensemble of arbitrarily configured PV systems. As a first step towards this objective, the paper describes an initial model quantifying the Output Variability resulting from an ensemble of equally-spaced, identical PV systems. This layout, while simplified, facilitates the analysis of most PV deployment scenarios from a single central station to a fully distributed configuration.

There has been a substantial amount of work devoted to understanding the variability associated with the solar resource for a single location. Suehrcke and McCormick (1989) published one of the early papers identifying the nature of high frequency irradiance data (e.g., 1 min) as

* Corresponding author. Address: ASRC, The University at Albany, 251 Fuller Rd., Albany, NY 12203, USA. Tel.: +1 518 437 8751.
E-mail address: perez@asrc.albany.edu (R. Perez).

0038-092X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.solener.2010.07.003

fundamentally different from the lower frequency data (e.g., hourly) that is generally available in archives and typical meteorological year (TMY) files. They found that the frequency distribution of the short-term irradiances is considerably more bi-modal than that of hourly data, expressing the "on or off" nature of solar radiation, particularly for the direct irradiance component. Jurado et al. (1995) confirmed this bi-modal nature with possible implications for the operation of solar systems. Gansler et al. (1995) described the shortcomings of using hourly data for solar system simulations, comparing 1-min and hourly simulations and showing that systems with operational thresholds could not be properly simulated without sub-hourly information. Marwali et al. (1998), however, indicated that adding proper battery buffers and active management of systems would absorb the impact of short-term variability on both the load and supply side.

An influential paper in the modeling of short-term variability is that of Skartveit and Olseth (1992). They demonstrated that the distribution of sub-hourly Global Horizontal Irradiance (GHI) and Direct Normal Irradiance (DNI) could be effectively parameterized and modeled as a function of solar conditions defined by the hourly clearness index as well as the variations of this index from one hour to the next. The Skartveit and Olseth paper supports the thesis that it would be possible to model the absolute variability of a single point based upon the data contained in hourly satellite-derived data sets such as the NSRDB (2007) or Solar Anywhere® (2009). Tovar et al. (2001) proposed a modeling approach similar to that of Skartveit and Olseth, experimentally showing that the frequency distribution of sub-hourly data could be assessed from the insolation conditions defined by the hourly data stream. Woyte et al. (2007) defined operational parameters to quantify power and energy fluctuations, identifying the dimensionless clearness index as the key variable, and corroborating the early findings of Skartveit and Olseth by showing that the probability distribution of the high frequency clearness index was largely independent of season and location, but dependent upon current insolation conditions.

A lesser amount of work has been devoted to understanding the effect of combining multiple locations on irradiance variability. Noteworthy is the paper by Wiemken et al. (2001) that analyzed normalized 5-min output data from 100 PV systems dispersed throughout Germany. The focus of the paper was placed on determining the standard deviation of the combined output of all systems relative to their mean monthly output rather than analyzing short-term fluctuations per se – so as to extract a measure of long-term predictability and stability of country-wide PV generation. The paper, nevertheless, showed that the short-term fluctuations from the ensemble of systems, quantified by the distribution of 5-min normalized power output changes, were drastically reduced compared to single system fluctuations.

Three interesting papers are based on measured data from a network of locations in Japan. Otani et al. (1997) examined variability associated with 1-min irradiance measurements obtained from a nine-site network located in approximately a 4 km by 4 km region. The authors calculated the root mean square of the difference between the instantaneous irradiance and the hourly average irradiance for each site independently versus the combined average irradiance from all nine-sites considered together. They found that during partly cloudy conditions the nine-site average decreased to around 20–50% of each site independently. These authors and their colleagues (Kawasaki et al., 2006) performed further analysis on this data and termed the reduction in variability as the "smoothing effect".

Murata et al. (2009) performed a related analysis using a 1999 data set composed of 52 PV systems. The authors' primary focus was to analyze the ratio of the worst case fluctuation relative to the average fluctuation in the change in output for a given number of PV systems. They found that the number of systems is not a key factor in this ratio. While not the main focus of their paper, they also analyzed the correlation between two systems' short-term change in output, a metric that is similar to the metric retained in the present article to quantify short-term variability. They found that the experimentally derived correlation increased with the inverse of the sites' distance and the length of the considered fluctuation Time Intervals (see Figs. 8 and 9 in that paper). The correlation was found to approach zero (implying that the systems fluctuate independently of each other) for intervals less than a few minutes. They derived an empirical model based on this experimental evidence but the relationship was not fully explained on a mathematical basis and did not result in a general model that could be applicable for any number of PV systems for deployments ranging from central station to distributed generation.

2. Methods

2.1. Definitions

2.1.1. PV Fleet

PV Fleet refers to an ensemble of PV installations. The ensemble considered in this paper consists of a one-dimensional set of N identical, equally-spaced, installations. This ensemble can describe configurations ranging from a centralized power plant (N closely-spaced installations) to regionally dispersed generation (N widely-spaced installations).

2.1.2. Output Variability

Output Variability is a measure of the PV Fleet's power output changes over a selected sampling Time Interval and analysis period relative to PV Fleet capacity. Output Variability is quantified by computing the standard deviation $\sigma_{\Delta t}^{\sum N}$ as follows.

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{C^{\text{Fleet}}}\right)\sqrt{\text{Var}\left[\Sigma_{n=1}^{N}\Delta P_{\Delta t}^{n}\right]} \quad (1)$$

where C^{Fleet} is the total installed peak power of the fleet and $\Delta P^n_{\Delta t}$ is a random variable that represents the time series of changes in power at the nth PV installation using a sampling Time Interval of Δt defined over an analysis period extending time t_1 to time t_T.

$$\Delta P^n_{\Delta t} = \left\{(t_1, \Delta P^n_{t_1,\Delta t})(t_2, \Delta P^n_{t_2,\Delta t}), \ldots, (t_T, \Delta P^n_{t_T,\Delta t})\right\} \quad (2)$$

Each $\Delta P^n_{t,\Delta t}$ represents the change in power output at the nth PV installation between times t and $t + \Delta t$. More specifically, $\Delta P^n_{t,\Delta t} = P^n_t - P^n_{t+\Delta t}$.

Power output at the nth installation is approximately equal to the product of the plane-of-array irradiance at the nth PV system (I^n), the capacity of the nth PV system (C^n), and a constant sizing factor α that is the same across all systems. Thus, Eq. (1) can be written as:

$$\sigma^{\sum N}_{\Delta t} = \left(\frac{1}{C^{\text{Fleet}}}\right)\sqrt{\text{Var}\left[\Sigma^N_{n=1}\alpha C^n \Delta I^n_{\Delta t}\right]} \quad (3)$$

Furthermore, C^{Fleet} equals N times C^n because all systems are assumed to be identical. The result is that Eq. (3) can be simplified and Output Variability is given by:

$$\sigma^{\sum N}_{\Delta t} = \left(\frac{\alpha}{N}\right)\sqrt{\text{Var}\left[\Sigma^N_{n=1}\Delta I^n_{\Delta t}\right]} \quad (4)$$

2.1.3. Relative Output Variability

Relative Output Variability is the central variable of the proposed model. It is defined as the ratio of the Output Variability for the PV Fleet (equal to $\sigma^{\sum N}_{\Delta t}$) to Output Variability of the same PV Fleet concentrated in one single location (equal to $\sigma^1_{\Delta t}$). The Relative Output Variability quantifies the noise reduction associated with the dispersion of the fleet over a region. Relative Output Variability ranges between 0% and 100%.

2.1.4. Dispersion Factor

Dispersion Factor (D) captures the relationship between PV Fleet configuration (i.e., the number of systems and their geographic density), Cloud Transit Speed (the primary source of short-term Output Variability), and Time Interval (Δt). The Dispersion Factor is a dimensionless variable defined as the number of Time Intervals required for a cloud disturbance to pass across the entire PV Fleet.

$$D = \frac{L}{V\Delta t} \quad (5)$$

where L is the length of the considered PV Fleet in the direction of the clouds motion and V is the transit rate. D is dimensionless because L is in meters, V is in meters per second, and Δt is in seconds.

The Dispersion Factor increases as the cloud speed decreases and/or as the distance between installations increases.

Fig. 1 illustrates the Dispersion Factor for three cases: a fast, medium, and slow Cloud Transit Speed across a PV Fleet with four PV systems. The fast-moving cloud in the top section of the figure crosses the PV Fleet in $2\Delta t$, and thus D equals 2. The medium transit speed requires $4\Delta t$ for a cloud to cross the PV Fleet, and thus D equals 4. The slow transit speed in the bottom would result in a Dispersion Factor of 8.

2.1.5. Independence

In the context of this article, independence between any pair of system in a fleet implies that the correlation between the time series of the two systems, sampled at the considered Time Interval Δt, approaches zero.

2.2. Output Variability Model Formulation

The model consists of a solution to Eq. (4) in four distinct Dispersion Factor regions:

Crowded Region	The number of PV systems is greater than the Dispersion Factor. As illustrated in the top section of Fig. 1, a cloud disturbance affects more than one PV system within the PV Fleet in one Time Interval.
Optimal Point	The number of PV systems equals the Dispersion Factor. As illustrated in the middle section of Fig. 1, a cloud disturbance affecting one system within the PV Fleet will affect the next one in exactly one Time Interval.
Limited Region	The number of PV systems is less than the Dispersion Factor. As illustrated in the bottom section of Fig. 1, a cloud disturbance does not reach the next system before the next Time Interval.
Spacious Region	The number of PV systems is much less than the Dispersion Factor. This is an extension of the Limited Region such that the short-term fluctuations of each PV system become independent of each other.

2.2.1. Spacious Region ($N \ll D$)

Beginning with the Spacious Region, systems are sufficiently far apart such that Output Variability for any system is independent of the Output Variability for any other system. When each random variable $\Delta I^n_{\Delta t}$ is independent, the variance of their sum equals the sum of the variances. Eq. (4) can thus be simplified by moving the variance inside the summation.[1]

$$\sigma^{\sum N}_{\Delta t} = \left(\frac{\alpha}{N}\right)\sqrt{\Sigma^N_{n=1}\text{Var}\left[\Delta I^n_{\Delta t}\right]} \quad (6)$$

[1] This is of course only valid for short-term sub-hourly fluctuations where the deterministic solar geometry effects that are not independent from one system to the next can be ignored. However these effects are well known and do not need a new modelization.

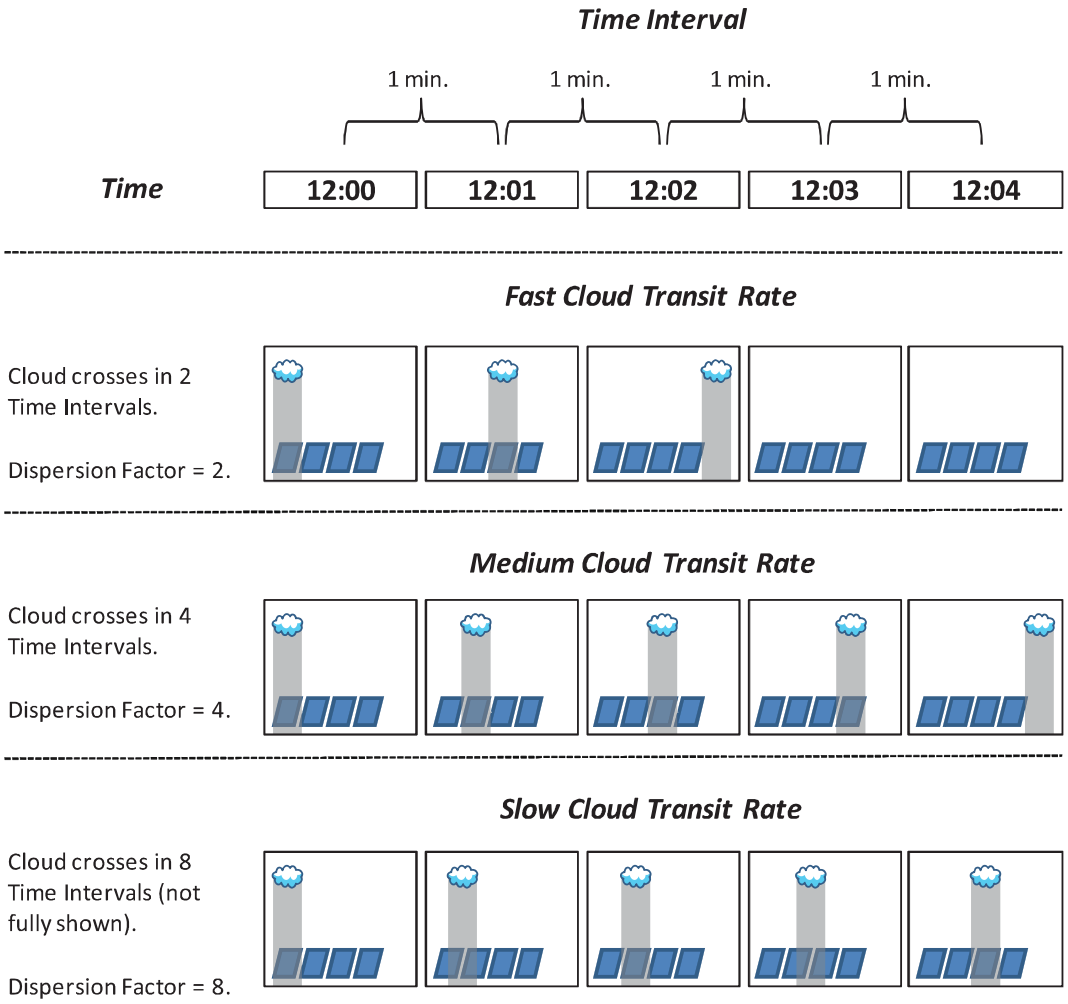

Fig. 1. Dispersion Factor for a PV Fleet with 4 PV systems using a 1-min Time Interval when the cloud transit rate is fast, medium, or slow.

Assuming that irradiance changes at all systems have the same standard deviation, the standard deviation for any system can be selected. The system at location 1 is arbitrarily selected and is substituted for all $\Delta^n_{\Delta t}$ with the result being:

$$\sigma^{\sum N}_{\Delta t} = \left(\frac{\alpha}{N}\right)\sqrt{N\left(\frac{\sigma^1_{\Delta t}}{\alpha}\right)^2} \tag{7}$$

When simplified, the Output Variability across all N systems equals the Output Variability at any one system divided by the square root of the number of systems.

$$\sigma^{\sum N}_{\Delta t} = \frac{\sigma^1_{\Delta t}}{\sqrt{N}} \tag{8}$$

This result ought to be expected since Eq. (8) represents the mean squared error of a random sample of size N. This relationship is traditionally known as the Bienayme formula (e.g., see Loeve, 1977).

2.2.2. Optimum Point ($N = D$)

By definition of the Dispersion Factor and the Optimum Point, and assuming as a first order approximation that the cloud patterns remain largely unchanged as they cross over the PV deployment zone, a cloud disturbance reaching one system will affect the next system during the next Time Interval, the following one in two Time Intervals, and so on.

Thus starting from Eq. (4) and squaring it to extract the variance,

$$\left(\sigma^{\sum N}_{\Delta t}\right)^2 = \left(\frac{\alpha}{N}\right)^2 \mathrm{Var}\left[\Sigma^N_{n=1}\Delta I^n_{\Delta t}\right] \tag{9}$$

and expanding the variance term as follows,

$$\left(\sigma^{\sum N}_{\Delta t}\right)^2 = \left(\frac{\alpha}{N}\right)^2\left(\frac{1}{T}\right)\Sigma^T_{t=1}\left[\left(\Sigma^N_{n=1}I^n_t - \Sigma^N_{n=1}I^n_{t+\Delta t}\right)\right]^2 \tag{10}$$

Eq. (10) can be rewritten as:

$$\left(\sigma^{\sum N}_{\Delta t}\right)^2 = \left(\frac{\alpha}{N}\right)^2\left(\frac{1}{T}\right)\Sigma^T_{t=1}\left[\left(\Sigma^N_{n=1}I^n_t - \Sigma^{N-1}_{n=0}I^{n+1}_{t+\Delta t}\right)\right]^2 \tag{11}$$

The last term in the first summation and the first term in the second summation can be extracted and the summations combined to result in

$$\left(\sigma_{\Delta t}^{\sum N}\right)^2 = \left(\frac{\alpha}{N}\right)^2 \left(\frac{1}{T}\right) \Sigma_{t=1}^{T} \left[\left(I_t^N + \Sigma_{n=1}^{N-1}\left(I_t^n - I_{t+\Delta t}^{n+1}\right) - I_{t+\Delta t}^1\right)\right]^2 \tag{12}$$

The Optimum Point assumption facilitates interchanging the PV system position and the Time Interval. More specifically, $I_t^n = I_{t+\Delta t}^{n+1}$ for all values of n. The result is that the terms in the interior summation cancel and Eq. (12) simplifies to:

$$\left(\sigma_{\Delta t}^{\sum N}\right)^2 = \left(\frac{\alpha}{N}\right)^2 \left(\frac{1}{T}\right) \Sigma_{t=1}^{T} \left[\left(I_t^N - I_{t+\Delta t}^1\right)\right]^2 \tag{13}$$

At the Optimum Point, $I_t^N = I_{t+(1-N)\Delta t}^1$. Thus, Eq. (13) can be rewritten as

$$\left(\sigma_{\Delta t}^{\sum N}\right)^2 = \left(\frac{\alpha}{N}\right)^2 \left(\frac{1}{T}\right) \Sigma_{t=1+(1-N)\Delta t}^{T+(1-N)\Delta t} \left[\left(I_t^N - I_{t+N\Delta t}^1\right)\right]^2 \tag{14}$$

where t starts at $1 + (1 - N)\Delta t$ rather than at 1.

Finally taking the square root of Eq. (14), it follows that the Output Variability of the PV Fleet equals $\frac{1}{N}$ times that of a single system at the Optimum Point using a Time Interval of $N\Delta t$.

$$\sigma_{\Delta t}^{\sum N} = \frac{\sigma_{N\Delta t}^1}{N} \tag{15}$$

2.2.3. Limited Region ($N < D$)

Next, consider the Limited Region. This applies when the number of PV systems is less than the Dispersion Factor. That is, the plants are located farther apart than the Optimal Point but not as far apart as the Spacious Region. Unfortunately, there is not a specific solution to this region. Rather, Eq. (8) applies as the upper bound in this range, because the independence condition would be violated if the locations were any closer than that.

$$\sigma_{\Delta t}^{\sum N} < \frac{\sigma_{\Delta t}^1}{\sqrt{N}} \tag{16}$$

2.2.4. Crowded Region ($N > D$)

Finally, consider the Crowded Region. In this range, one might think of the Crowded Region effectively being an increased concentration of PV in each location that is greater than $1/N$. Therefore, because the premise behind Eqs. (11)–(15) remains true if N is replaced by D, the Output Variability in the Crowded Region can be expressed as

$$\sigma_{\Delta t}^{\sum N} = \frac{\sigma_{D\Delta t}^1}{D} \tag{17}$$

2.2.5. Model visualization

The proposed model amounts to the combination of Eqs. (8, 15, 16, and 17), each applying in its respective spatial domain. The fundamental shape of this model is presented in the left side of Fig. 2 for N PV systems. Relative Output Variability reduces in the Crowded Region and reaches a minimum of $1/N$ at the Optimal Point. It increases somewhat in the Limited Region and then stabilizes in the Spacious Region where each PV location is independent reaching a value of $1/\sqrt{N}$. The right side of Fig. 2 presents the structure of the model for four times as many locations (i.e., 4 N physical locations). Quadrupling the number of PV systems cuts Relative Output Variability in half in the Spacious Region.

The following observations can be made:

- Relative Output Variability generally decreases with an increasing number of systems.
- The Dispersion Factor relative to the number of locations can limit the Relative Output Variability reduction. For example, locating too many individual plants too close together reduces the value of the larger number of locations.
- Relative Output Variability decreases with $\frac{1}{\sqrt{N}}$ for large location spacing ($N \ll D$).[2]
- Maximizing location spacing (or increasing D) for a fixed number of locations does not minimize Relative Output Variability. While this is initially a surprising result, it makes sense upon reflection, because the locations may benefit from being close together up to a point due to negative covariance.

It is interesting to note that in the left side of the Crowded Region, near the y axis, the Relative Output Variability is independent of the considered number of systems, when D is less than or equal to N. This implies that more PV systems do not automatically reduce Relative Output Variability within a very Crowded Region – a situation that can apply to central generation plants.

3. Validation results

3.1. Experimental data

The ideal data to use in validating the proposed model would consist of either high frequency irradiance data from high-density, large-area, regularly gridded networks of pyranometers, or identical PV installations. Unfortunately, such comprehensive networks specifically designed for this purpose have not yet been fully deployed and/or made available to the scientific community. However, an alternative set of data can be developed by constructing a *virtual network* based on high frequency data measured at a single, actual location. Irradiance data at a virtual location in the virtual network are obtained by assuming that the cloud-induced patterns measured at the actual location move at a constant velocity across the virtual network. For example, if it takes 4 min for a cloud disturbance to travel from the actual location to a virtual location, then the irradiance data for the virtual location at 12:13 pm would be the same

[2] In practice, large spacing would represent several 100s of meters between stations (see below).

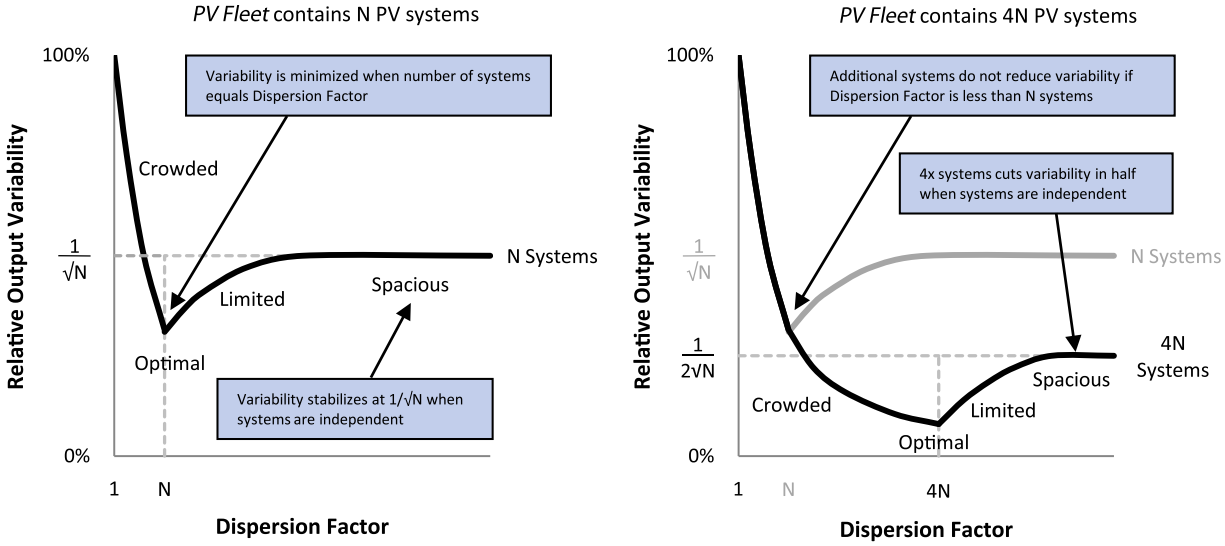

Fig. 2. Relative Output Variability is a function of the number of PV systems and the Dispersion Factor.

as the irradiance data for the actual location at 12:09 pm. In effect, the virtual network translates the time dimension at a single measured location into a spatial dimension.

Virtual networks were thus constructed from twelve stations of the Atmospheric Radiation Measurement (ARM) Southern Great Plains Site's extended facility, where irradiance is recorded at a 20-s rate (Stokes and Schwartz, 1994). Global Horizontal Irradiance (GHI) is used as a proxy for plane-of-array irradiance (an assumption equivalent to considering horizontal PV arrays, i.e., without fundamental implications on Relative Output Variability).

3.2. Limited model validation

The limited validation presented below is based on data from one highly variable day's worth of data from the 12 virtual networks. This validation is undertaken to substantiate the model's framework and underlying assumptions with a sample of measured high frequency irradiance data. For each network, the model is evaluated using different *scenarios*, by varying the Time Interval (Δt), the number of locations (N), and/or their one-dimensional spacing. One scenario at a single network is considered first. Results are then extended to include multiple scenarios at all 12 networks.

3.2.1. Single network/single scenario

This first illustrative scenario consists of a fleet of 16 installations distributed so that $N = D$ (Optimal Point). Fig. 3 presents observed GHI and the change in irradiance ($\Delta I_{\Delta t}$) with $\Delta t = 20$ s. The light gray lines correspond to irradiance and variability for a single location in the selected network and the dark lines correspond to irradiance and variability for the fleet. Variability is greatly reduced compared to a single station.

3.2.2. Single network/multiple scenarios

A series of scenarios at the same network varying the number of systems and the Dispersion Factor are considered next. Fig. 4 presents the resulting Relative Output Variability. Part (1) of the figure presents the *Crowded* and *Spacious* models for four systems using a Time Interval (Δt) of 60 s. Part (2) superimposes the virtual network's experimental data where the spacing for four PV systems is varied so as to result in a range of Dispersion Factors. Part (3) repeats Parts (1) and (2) and adds modeled and experimental results for 16 systems. Part (4) repeats Parts (1) through (3) and adds results for scenarios using a Time Interval of 20 s (dotted lines). The figure suggests that experimental results are closely aligned with the proposed model for all scenarios.

3.2.3. Multiple networks/multiple scenarios

While the results presented above for one *virtual* network substantiate the proposed model, it is useful to see how results compare across multiple actual sites. The mean Output Variability of each 12 independent sites for the selected day is plotted against the mean irradiance observed for each site in Fig. 5. This figure suggests that the 12 sites represent a diverse sample of partly cloudy conditions where short-term variability is expected to be significant.

Fig. 6 repeats the analysis presented in Fig. 3. It presents GHI and the 20-s change in GHI for a single system and for 16 virtual systems in an optimum point configuration for each of the 12 virtual networks based on 12 actual sites. The 16-system fleet consistently reduces variability for each network.

Fig. 7 repeats the last part of the analysis presented in Fig. 4. It summarizes the results obtained for all networks for a 20-s and 60-s Time Interval and for 4 and 16 systems with a variety of location spacing. The experimental results

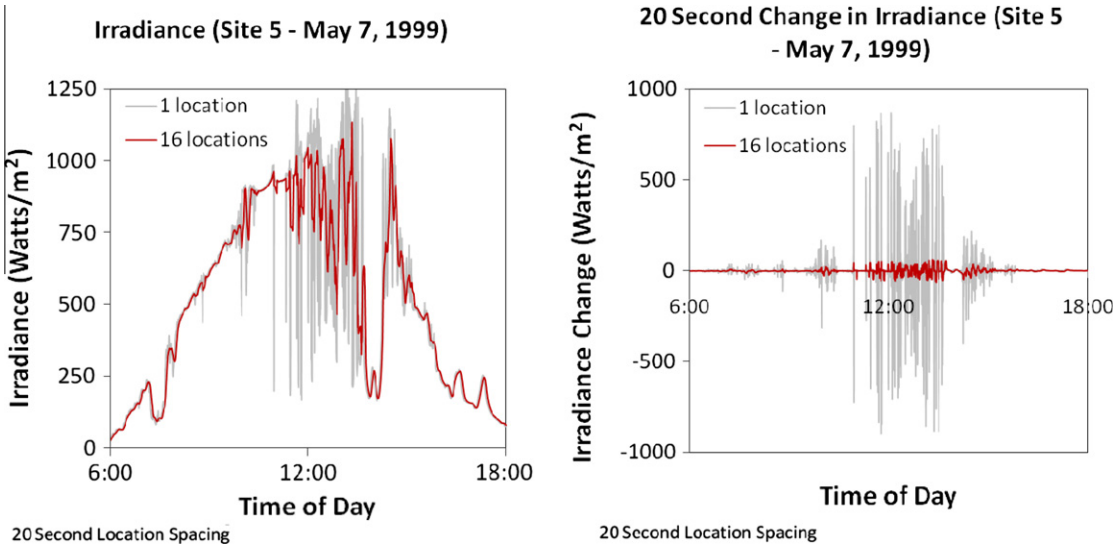

Fig. 3. Irradiance for one location and averaged over 16 locations at the Optimal Point (left), and change in irradiance using a Time Interval of 20 s (right).

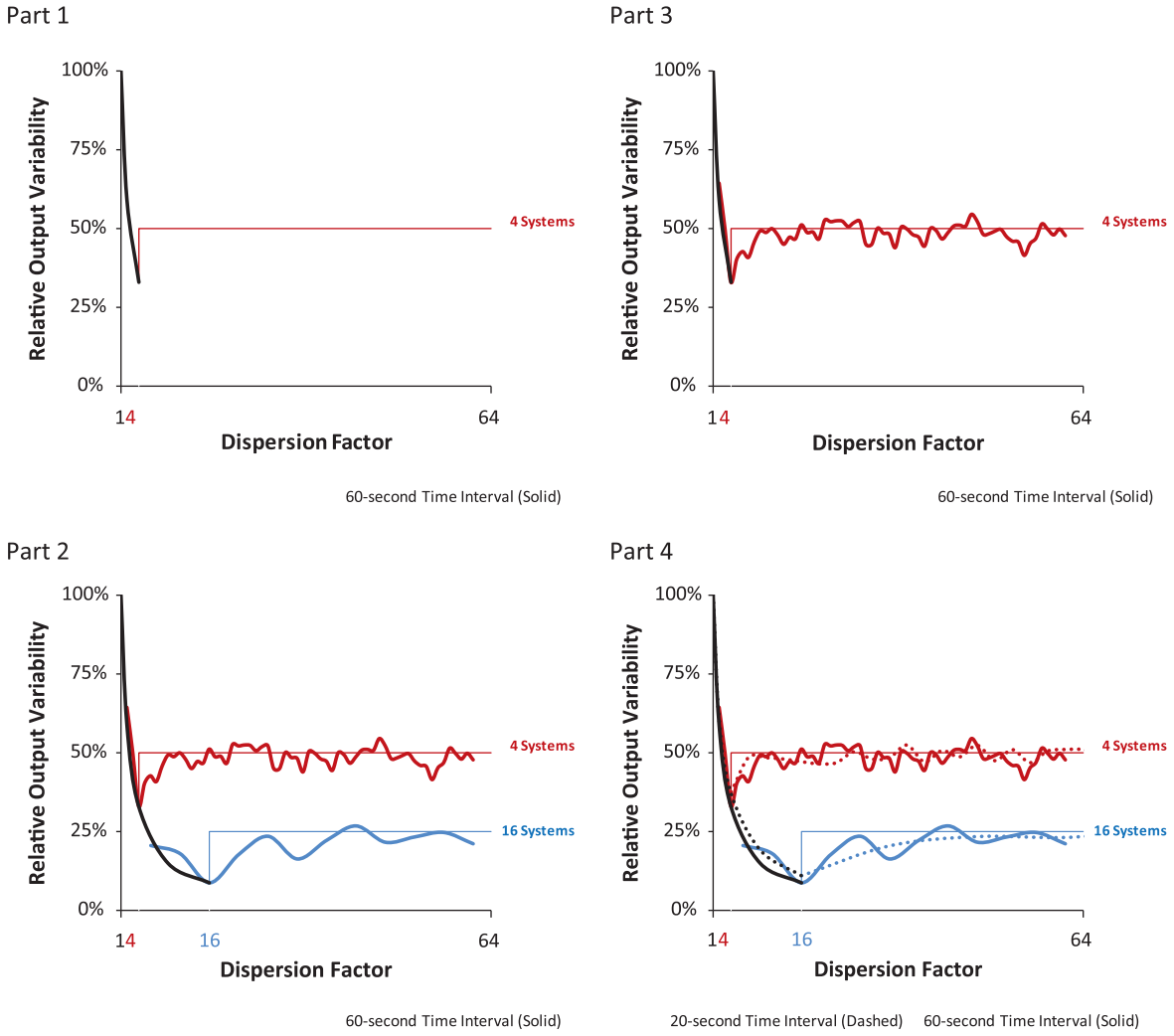

Fig. 4. Validation results for one virtual network and multiple N, D and ΔT scenarios. Part 1 shows the crowded model (thick black line) & Spacious model – four systems (thin red line) using 60-s Time Interval. Part 2 adds measured data (thick red line) for four systems. Part 3 adds measured data (blue lines) for 16 systems. Part 1 repeats Parts 1–3 using a time interval of 20 s (dotted lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

are consistent across all 12 networks and agree well with the proposed model. In order to investigate this further, the validation was expanded to include a large number of scenarios at all 12 virtual networks.

Looking first at the spacious system distribution, Fig. 8 reports the mean Relative Output Variability experimentally observed across all networks/scenarios when the number of spaciously distributed installations is increased from 1 to 64. This experimental validation shows that the model closely follows a $1/\sqrt{N}$ Relative Output Variability as the number of independent systems increases.

Fig. 9 presents an analysis of the results of the Crowded Region for all 12 virtual networks across a range of Dispersion Factors using Time Intervals of 20, 40, 60, 80, 100, and 120 s. The light gray lines correspond to all the scenarios simulated for all 12 networks. The figure indicates that results are similar under all scenarios at all 12 virtual networks. *The dark solid line suggests that an empirical result (one that requires further validation), is that the Crowded model is approximately equal to the inverse of the Dispersion Factor raised to the ¾ power.*

3.3. Indirect complementary validation – the Otani et al. article

As mentioned above, Otani et al. (1997) performed an analysis of 1-min irradiance measurements obtained from a nine-site network concentrated in an area of 12 km^2. The authors calculated the root mean square (RMS) of the difference between the instantaneous irradiance and the hourly average irradiance for each site independently versus the combined average irradiance from all nine-sites considered together. The present paper calculates the standard deviation (this is similar to their RMS calculation) of the difference between two consecutive irradiance measurements. Thus, while the results in the Otani paper tend to understate the minute-to-minute changes because some of the variability is eliminated when the hourly average is taken, the study offers a valuable opportunity to partially validate results from this paper. The proposed model predicts that Output Variability for a system with capacity spread out across nine spacious locations relative to single site-variability should be 33% ($1/\sqrt{9}$). Otani et al. report that the nine-site irradiance variability decreased to around 20–50% relative to each representative site during cloudy conditions. While not an exact comparison due to methodological differences, it does provide a complementary validation of the approach documented herein.

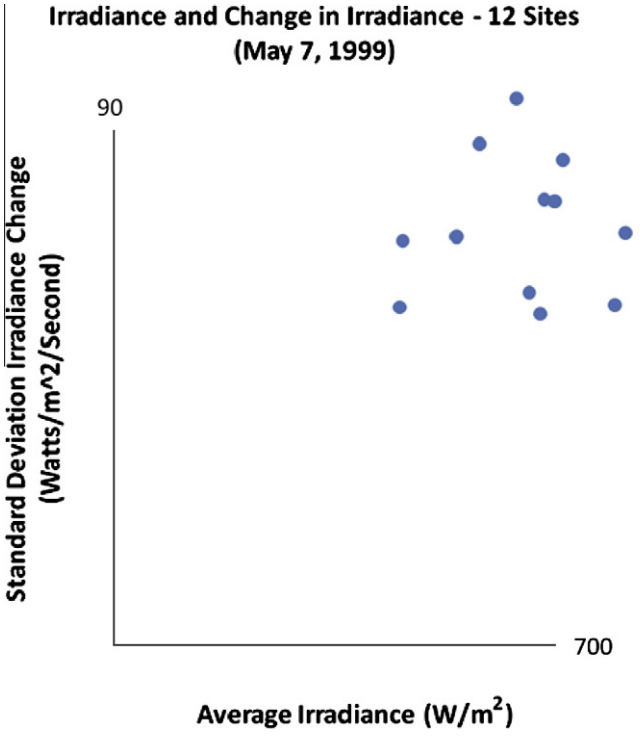

Fig. 5. Summary statistics for 12 geographically distinct sites.

4. Discussion

4.1. Prospective model application examples

The proposed modeling approach is simple and produces convincing results. This section presents two examples of how the model might be applied. The first example is designed to provide insights on the performance of a medium size (5 MW) PV plant operated by Tucson Electric in Springerville, Arizona. The second example is for a hypothetical 100 MW PV Fleet implemented as either a central station plant or as distributed generation.

4.1.1. 5 MW Springerville, Arizona PV plant

Tucson Electric reported that the utility experienced changes in output up to 50% over Time Intervals of 60 s (Hansen, 2007). This reported finding has become a cause of concern for some in the industry while others have presented information that suggests that the concern is unfounded (Hoff et al., 2008; Perez et al., 2009). The results from this paper can be used to shed additional light on the subject.

The Springerville plant covers 17.8 hectares (Moore et al., 2005). While not perfectly square, it can be assumed that the plant is about 420 m by 420 m. As stated above, the Time Interval (Δt) of concern to the utility is 60 s. Consider a case when the cloud transit rate equals 3.5 m/s. A disturbance impingent upon the first portion of the system will be completely transitioned off the plant after $2\Delta t$'s, implying $D = 2$. As shown in Fig. 9, a Dispersion Factor of 2 translates to a standard deviation that is a Relative Output Variability of 60%. The low Dispersion Factor provides for a good understanding of why there is a high degree of Relative Output Variability at the Springerville plant under certain conditions.

One question that arises, however, is how this situation could have been avoided. Suppose that a distributed generation approach had been taken instead of concentrating all 5 MW in a single location. Rather than having the plant concentrated in a single location, the 5 MW plant could

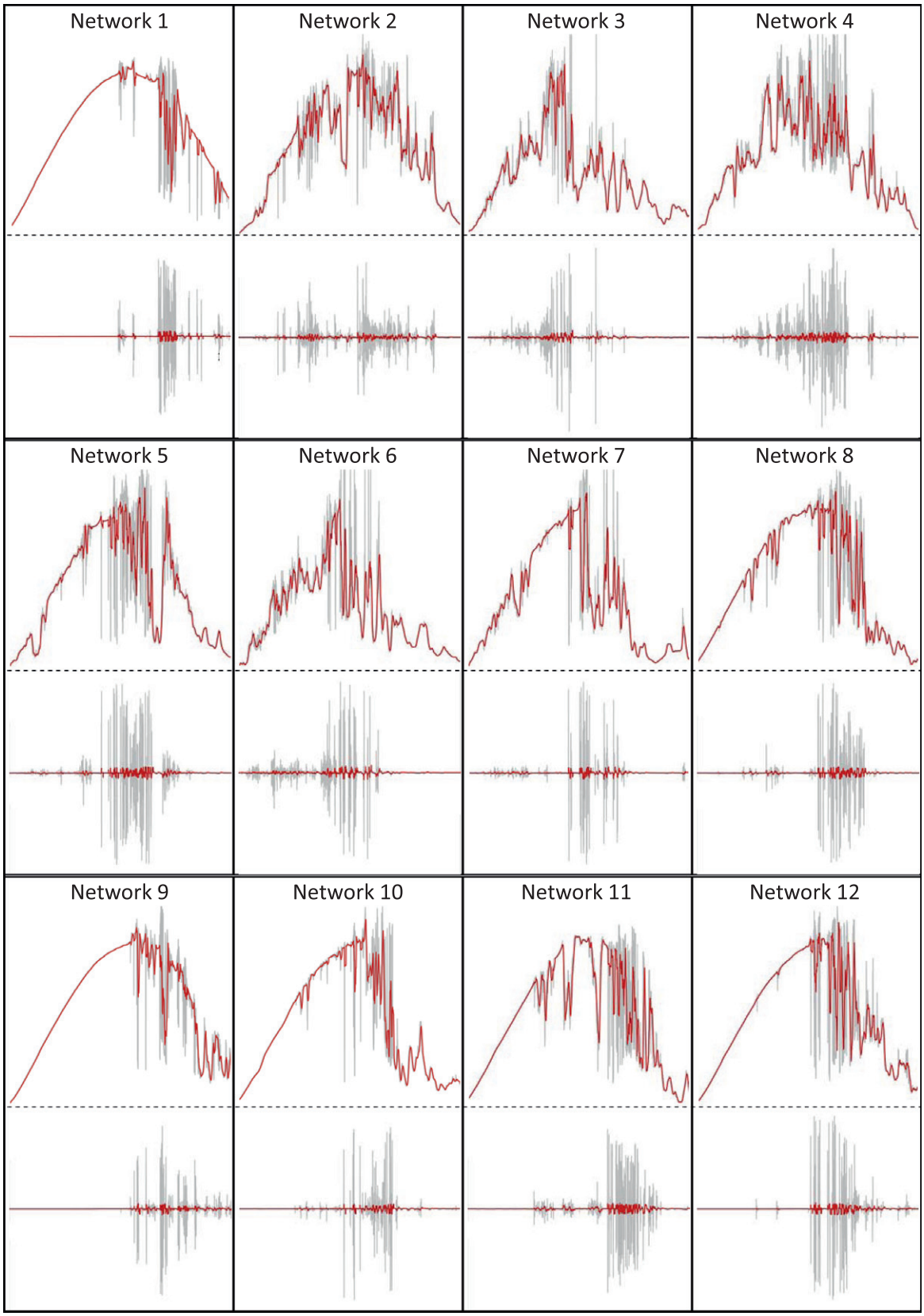

Fig. 6. Same as Fig. 3 for all 12 virtual networks ($N = 16$, $\Delta t = 20$ s).

have been regionally distributed across 1000 5-kW independent plants. In that case, it would be the Spacious Region and the standard deviation relative to a single location would have been 3% ($1/\sqrt{1000} = 3\%$). That is, the dis-

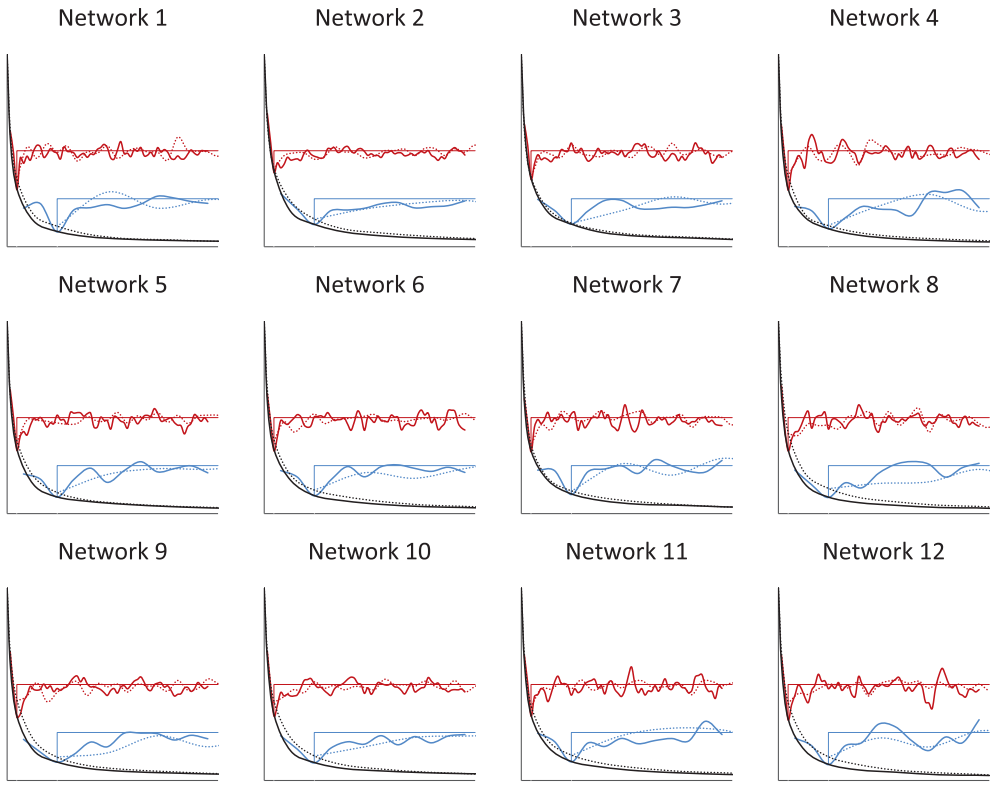

Fig. 7. Same as Fig. 4, Part 4 for all 12 virtual networks.

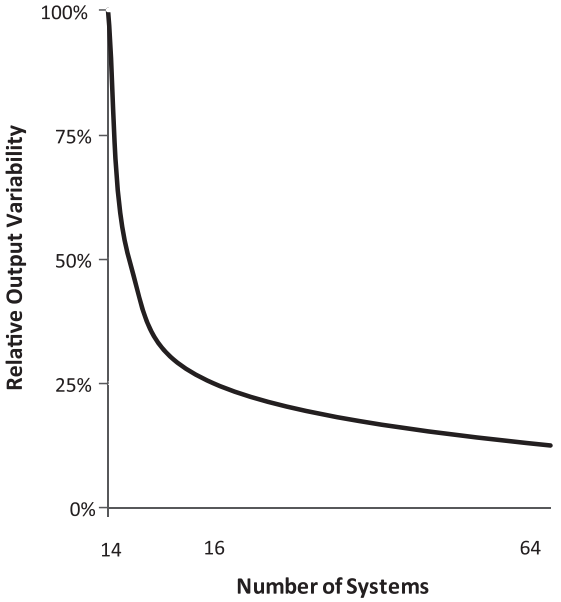

Fig. 8. Experimentally observed Relative Output Variability for Spacious Region.

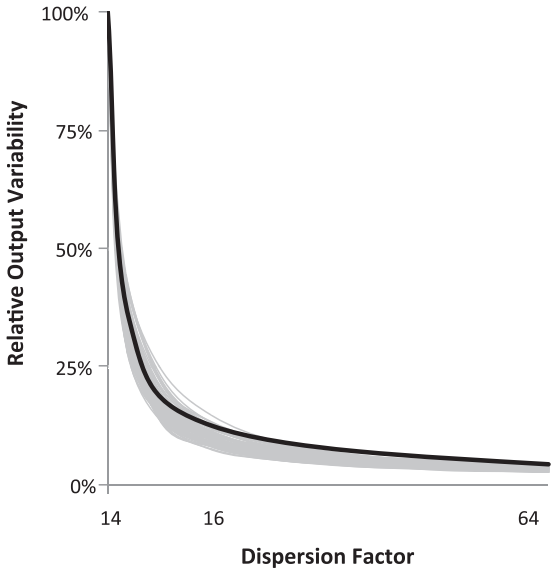

Fig. 9. Experimentally observed Relative Output Variability for Crowded Region.

tributed generation scenario has 3% relative variability versus the central plant's 60%.

4.1.2. Hypothetical 100 MW PV

This second example is for a utility that intends to install 100 MW of PV (assume that it is all in the same orientation and configuration) and is contemplating a variety of installation scenarios. The options under consideration include:

- Central scenario: one 100 MW central station facility.
- Distributed scenario 1: 100 moderately-sized (1 MW) distributed plants distributed throughout the utility system.

- Distributed scenario 2: 20,000 residential PV systems that are 5 kW each distributed throughout the utility system.

The utility wants to gain a better understanding of the potential impact of the Relative Output Variability from a utility systems operation perspective before selecting an installation scenario.

The two distributed scenarios are the most simple to deal with because the assumption can be made that the systems are independent. Eq. (6) (spacious model) can be applied directly if it is assumed that weather conditions are similar throughout the utility system and stations are sufficiently far apart for their number to be much less than the Dispersion Factor applicable to the considered variability Time Interval.[3] The 100 1-MW plants will result in 10% Relative Output Variability ($1/\sqrt{100}$). The 20,000 5-kW plants will result in less than 1% Relative Output Variability ($1/\sqrt{20,000}$).

The central station plant can be viewed as being composed of a large number of closely-spaced systems. Thus, the limiting factor is not likely to be the number of systems. Rather, it will be the Dispersion Factor as defined by the available distance and Cloud Transit Speed.

The following assumptions are made to perform the analysis:

- The 100 MW occupies a 1 km^2 (i.e., output density of 100 W/m^2).
- One side of the square is perpendicular to the direction of the wind.
- Maximum cloud transit rate is 10 m per second.
- The utility is concerned with the Output Variability over a Time Interval of 10 s.

These assumptions translate to a Dispersion Factor of 10, and thus a Relative Output Variability of about 18% per results presented in Fig. 9.

This example illustrates how Relative Output Variability of a central plant is primarily dependent upon the Cloud Transit Speed (upon which the Dispersion Factor depends) while that of the distributed systems is primarily a function of the number of sites.

5. Concluding remarks

This paper presented a novel approach to rigorously quantify the variability in power output from a fleet of PV systems, ranging from a single central station to a set of distributed PV systems. The approach demonstrated that the *Relative Output Variability* for a fleet of identical PV systems (same size, orientation, and spacing) is a function of the number of PV systems and the *Dispersion Factor*. The *Dispersion Factor* captures the relationship between PV Fleet layout, the *Time Interval* over which variability is evaluated, and *Cloud Transit Speed*. Results indicated that *Relative Output Variability* for widely-spaced PV systems equals the inverse of the square root of the number of systems. Results also indicated that optimally-spaced PV systems can minimize *Relative Output Variability*.

Model results were compared to measured 20-s irradiance data during high variability conditions. The measured output data were translated to virtual networks. Model results were compared to multiple locations and system spacings. The model agreed well with measured results across all configurations for all considered virtual networks.

The applicability of the model was demonstrated using two examples. One example was a cursory analysis of Tucson Electric's 5 MW PV plant, where the model provided a rational explanation for the high observed variability. The second example was for a hypothetical 100 MW PV Fleet constructed in either a central station or distributed application, and contrasting the short-term variability of both options.

It is of course very important that the development of the proposed model, applied and tested under a set of limited assumptions, including uni-dimensionality, homogeneousness of PV installations and reliance on virtual networks for initial testing be followed by a comprehensive "real-life" implementation and testing. Therefore future work will proceed in the two following directions:

- Validating the basic model framework on actual, rather than virtual networks – although fully comprehensive high frequency regularly gridded networks are not available yet, smaller networks are now being deployed and operated that could be used to expand the limits of the current tests, e.g., the UC-San Diego Network (Kleissl, 2009) to address both microclimatic and variability effects.
- Expanding the current simplified one-dimensional, homogeneous model to account for arbitrary system sizes, system specifications and geographical distribution. Such a task should be a straightforward application of the current basic model, because it already can provide the key underlying information relating the variability of arbitrarily spaced points.

In addition, future activities will also address the question of quantifying absolute single site-variability as a function of insolation conditions, likely along the lines proposed by Skartveit and Olseth (1992). This would allow the model to exploit hourly data bases such as the NSRDB (Myers et al., 2005; Wilcox et al., 2007) or Solar Anywhere® (2009) via the generation of physically representative sub-hourly data time series.

[3] For instance with a prevailing cloud speed of 3.5 m/s and a applicable Δt of 20 s, a 50 km linear region would have a dispersion factor of ~700 a density of one PV installation per km would lead to $N < D/10$ satisfying the spacious region criterion under the assumptions of this study.

Acknowledgements

Thanks to Jeff Peterson at the New York State Energy and Research Development Authority (NYSERDA) for his support of this work. Thanks to Ben Norris, Jim Haugen, and Jeff Ressler (Clean Power Research), Christy Herig (Solar Electric Power Association), and Antoine Zelenka (Emeritus, MeteoSchweitz) for their comments. All opinions in this paper are those of the authors and do not necessarily reflect the views, positions or opinions of NYSERDA or its staff.

References

Blatchford, J., 2009. Personal Communication. California Independent System Operator. <http://www.caiso.com/>.

Gansler, R.A., Klein, S.A., Beckman, W.A., 1995. Investigation of minute solar radiation data. Solar Energy 5 (1), 21–27.

Hansen, T., 2007. Utility Solar Generation Valuation Methods, USDOE Solar America Initiative Progress Report, Tucson Electric Power, Tucson, AZ.

Hoff, T.E., Perez, R., Ross, J.P., Taylor, M., 2008. Photovoltaic Capacity Valuation Methods. SEPA REPORT # 02-08.

Jurado, M., Caridad, J.M., Ruiz, V., 1995. Statistical distribution of the clearness index with radiation data integrated over 5 min intervals. Solar Energy 55 (1), 469–473.

Kawasaki, N., Oozeki, T., Otani, K., Kurokawa, K., 2006. An evaluation method of the fluctuation characteristics of photovoltaic systems by using frequency analysis. Solar Energy Materials and Solar Cells 90 (18–19), 3356–3363.

Kleissl, J., 2009. <http://maeresearch.ucsd.edu/kleissl/demroes/map.html>.

Loeve, M., 1977. Probability Theory, Graduate Texts in Mathematics, vol. 45, fourth ed. Springer-Verlag, p. 12.

Marwali, M.K.C., Haili, M., Shahidehpour, S.M., Abdul-Rahman, K.H., 1998. Short term generation scheduling in photovoltaic-utility grid with battery storage. IEEE Transactions on Power Systems 13 (3), 1057–1062.

Moore, L., Post, H., Hansen, T., Mysak, T., 2005. Photovoltaic power plant experience at Tucson electric power. Progress in Photovoltaics: Research and Applications 13 (4), 353–363.

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

Myers, D., Wilcox, S., Marion, W., George, R., Anderberg, M., 2005. Broadband model performance for an updated national solar radiation data base in the United States Of America. Proc. Solar World Congress, International Solar Energy Society.

Noah, F., 2008. Predicting Sudden Changes in Wind Power Generation. North American Wind Power, October 2008 Issue. www.nawindpower.com>.

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

Perez, R., Taylor, M., Hoff, T.E., Ross, J.P., 2009. Redefining PV capacity. Public Utilities Fortnightly, 147(2).

Skartveit, A., Olseth, J.A., 1992. The probability density of autocorrelation of short-term global and beam irradiance. Solar Energy 46 (9), 477–488.

Solar Anywhere®, 2009. Web-Based Service that Provides Hourly, Satellite-Derived Solar Irradiance Data Forecasted 7 days Ahead and Archival Data back to January 1, 1998. <www.SolarAnywhere.com>.

Stokes, G.M., Schwartz, S.E., 1994. The atmospheric radiation measurement (ARM) program: programmatic background and design of the cloud and radiation test bed. Bulletin of American Meteorological Society 75, 1201–1221.

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

Tovar, J., Olmo, F.J., Batlles, F.J., Alados-Arboledas, L., 2001. Dependence of 1-min global irradiance probability density distributions on hourly irradiation. Energy an International Journal 26 (7), 659–668.

US Department of Energy, 2009. Workshop Report on High Penetration of Photovoltaic (PV) Systems into the Distribution Grid, Ontario, CA, February 24–25.

Wiemken, E., Beyer, H.G., Heydenreich, W., Kiefer, K., 2001. Power characteristics of PV ensembles: experience from the combined power productivity of 100 grid-connected systems distributed over Germany. Solar Energy 70, 513–519.

Wilcox, S., Anderberg, M., Beckman, W., DeGaetano, A., George, R., Gueymard, C., Lott, N., Marion, W., Myers, D., Perez, R., Renné, D., Stackhouse, P., Vignola, F., Whitehurst, T., 2007. National Solar Radiation Database 1991–2005 Update: User's Manual Technical Report NREL/TP-581-41364. National Renewable Energy Laboratory, Golden, Co.

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