Quantifying PV Power Output Variability

Thomas E. Hoff and Richard Perez

Clean Power Research

Abstract

This paper presents a novel approach to rigorously quantify power output variability from a fleet of 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 the Dispersion Factor. The Dispersion Factor is a 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 dispersed PV systems; and (2) can be minimized for optimally-spaced PV systems. That is, relative output variability decreases as the number of dispersed PV systems increases. It eventually reaches the point where output variability is negligible relative to the total fleet capacity. For example, the output variability from 1,000 MW of dispersed 4-kW residential PV systems corresponds to the output variability associated with 0.2 percent of what it would be if the capacity was concentrated in a single location. The approach can be used to analyze an existing fleet's output variability or to control output variability when designing or providing incentives for the construction of a new fleet.

Acknowledgements

Thanks to Jeff Peterson at the New York State Energy and Research Development Authority (NYSERDA) for his support of this work. Thanks to Christy Herig (Solar Electric Power Association), Ben Norris (Clean Power Research), Jeff Ressler (Clean Power Research), and Antoine Zelenka (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.

Key Words

photovoltaics, variability, intermittency, distributed generation, central station, high penetration

Executive Summary

Introduction
There is a growing concern about the effects of photovoltaic (PV) power output variability on utility grid stability. High levels of minute-by-minute output variability during partly cloudy conditions reported at some central station PV facilities have created an awareness of this issue. Some industry professionals believe that this issue could constrain the penetration of grid-connected PV.

These and other concerns prompted the US Department of Energy to hold a workshop on "High Penetration of Photovoltaic (PV) Systems into the Distribution Grid" in February 2009. Many participants identified PV output variability as a top research priority.

Objective
The long-term objective of this work is to provide a model that quantifies the absolute power output variability of a fleet of arbitrarily-configured PV systems. This paper describes, validates, and applies a model that quantifies relative power output variability for a fleet of identical PV systems (same size, orientation, and spacing) distributed in the direction of prevailing cloud motion. This simplified layout facilitates the analysis of most PV deployment scenarios, from single central station to fully distributed configurations.

Key Definitions and New Concepts
The analysis begins with the definition of key variables and new concepts.

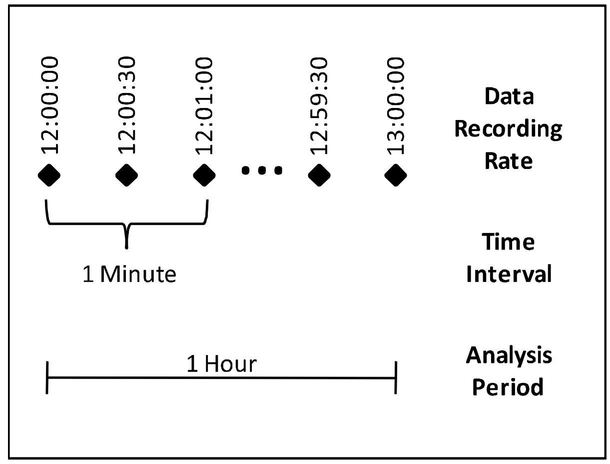
PV Fleet
A *PV Fleet* is composed of a given number of individual PV systems spread out over a geographical area.

PV Power Output and Variability
Four aspects of PV output and variability are defined here. *PV Power Output* is the instantaneous amount of power (kW) produced by a *PV Fleet* (the combined output of all individual systems, regardless of placement). *Change in PV Power Output* is the difference between two *PV Power Output* measurements separated by one *Time Interval*. If a *PV Fleet* has a total *PV Power Output* of 800 kW at 12:00 PM and 1000 kW at 12:01 PM, then the *Change in PV Power Output* over a 1-minute *Time Interval* is 200 kW.

Output Variability summarizes the series of *Changes in PV Power Output* over the *Analysis Period*. It is a measure of the power fluctuations from the *PV Fleet*. It is quantified, in this paper, as the standard deviation of this series. *Relative Output Variability* is the key calculated measure of *PV Fleet* variability, and is used for comparative purposes between alternative fleet designs. It equals the *Output Variability* of a given *PV Fleet* divided by the *Output Variability* of a PV installation with the same capacity as the fleet concentrated at one single location. *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 percent.

Time Measurements

Three time measurements are defined here: *Data Recording Rate*, *Time Interval*, and *Analysis Period*. *Data Recording Rate* is the frequency at which power measurements are recorded. *Time Interval* is the time over which the change in fleet power output is calculated. The analysis of 1-minute plant variability, for example, would have a 1-minute *Time Interval*. *Analysis Period* is the period over which the analysis is performed. The figure to the right illustrates the relationship between data that has a 30-second *Data Recording Rate*, a 1-minute *Time Interval*, and a 1-hour *Analysis Period*.



Dispersion Factor

Dispersion Factor is a new concept defined in this paper. It captures the relationship between *PV Fleet* configuration (i.e., number and orientation of systems and their geographic density), cloud transit speed (the primary source of short-term output variability), and the *Time Interval* of interest. *Dispersion Factor* is defined to be the number of *Time Intervals* required for a cloud to pass across the entire *PV Fleet*.

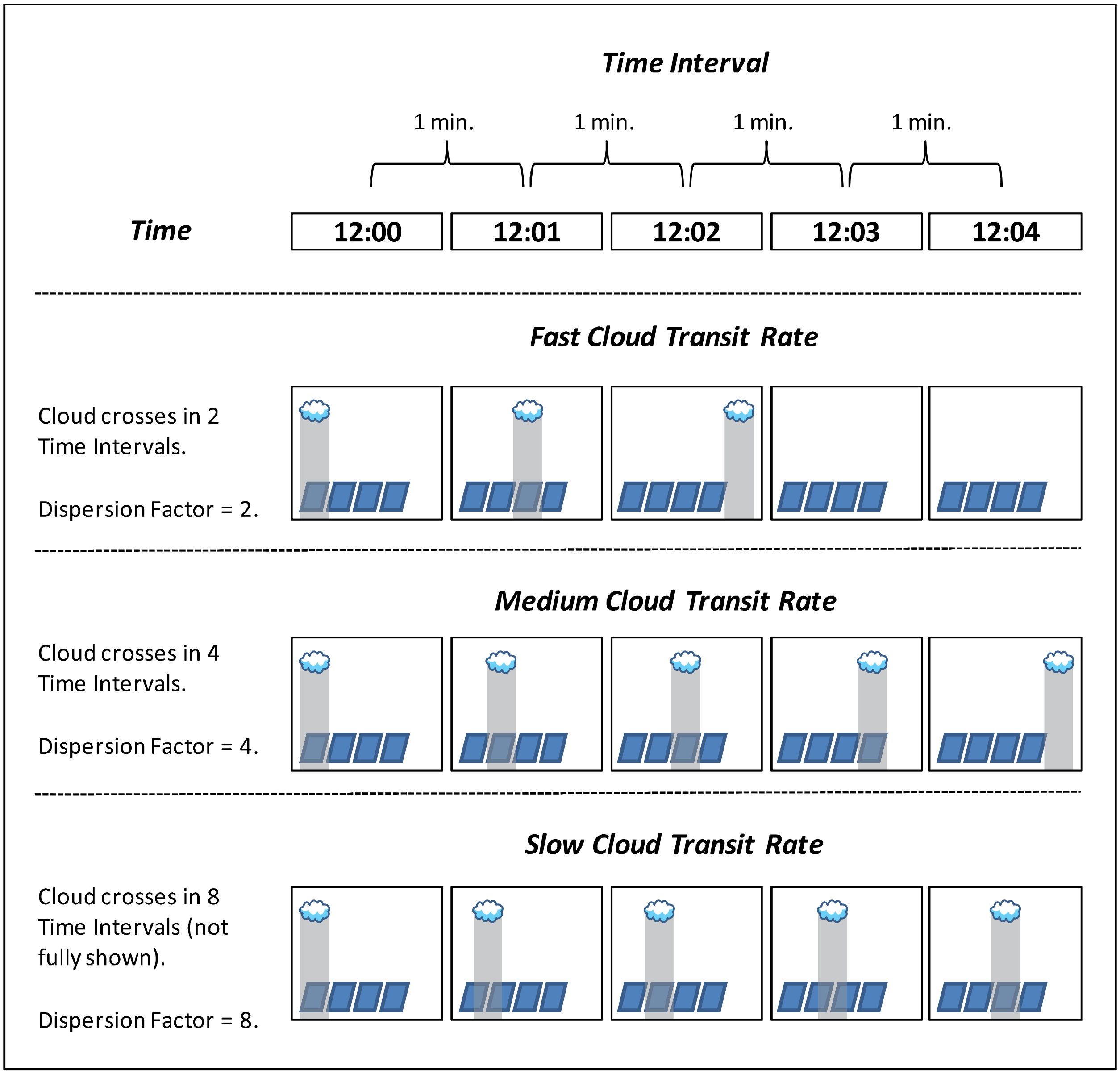
Figure 1 presents the *Dispersion Factor* for three cases: a fast, medium, and slow moving cloud across a *PV Fleet* with 4 PV systems. The *Time Interval* is assumed to be 1 minute. The fast-moving cloud in the top section of the figure crosses the *PV Fleet* in 2 minutes (2 *Time Intervals*) and thus has a *Dispersion Factor* of 2. The medium-speed cloud in the middle section of the figure crosses the *PV Fleet* in 4 minutes (4 *Time Intervals*) and thus has a *Dispersion Factor* of 4. The slow-moving cloud in the bottom section of the figure crosses the *PV Fleet* in 8 minutes (8 *Time Intervals*) and thus has a *Dispersion Factor* of 8. The *Dispersion Factor* increases as the wind speed decreases.

Virtual Network

The ideal data to use in validating the model in this paper would be either high-frequency irradiance data from high-density, large-area, gridded-networks of pyranometers or high-frequency power output data from networks of identical PV systems. Networks specifically designed for this purpose do not currently exist.

An alternative set of data can be developed by constructing a *Virtual Network* based on data measured at a single, actual location. Irradiance data at a virtual location is specified by assuming that the cloud-induced patterns measured at the actual location move at a constant speed across the *Virtual Network*. Data for a virtual location equals data measured at the actual location at the point in time that corresponds to the amount of time required for a cloud to travel from the actual location to the virtual location. That is, the data for the virtual location is obtained by measuring the data at the actual location at a given time stamp and then adding the amount of time required to move from the actual location to the virtual location to the time stamp. Such a *Virtual Network* facilitates the analysis of scenarios that vary the number of systems and the *Dispersion Factor*.

Figure 1. *Dispersion Factor* for a *PV Fleet* with 4 PV systems using a 1-minute *Time Interval* when the *Cloud Transit Rate* is fast, medium, or slow.



Model Description

The model characterizes *Relative Output Variability* over four distinct *Dispersion Factor* regions.

<u>CROWDED REGION</u> The number of PV systems is greater than the *Dispersion Factor*. As illustrated in the top section of Figure 1, a cloud disturbance affects more than one PV system within the *PV Fleet* in one *Time Interval*.

<u>OPTIMAL POINT</u> The number of PV systems equals the *Dispersion Factor*. As illustrated in the middle section of Figure 1, a cloud disturbance affecting one system within the *PV Fleet* will affect the next one in exactly one *Time Interval*.

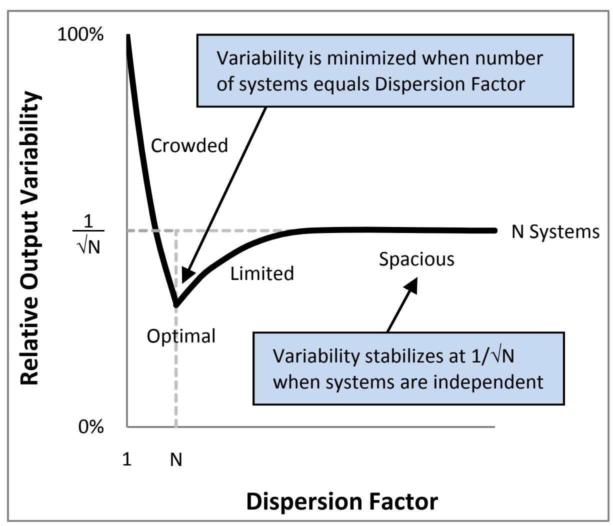
<u>LIMITED REGION</u> The number of PV systems is less than the *Dispersion Factor*. As illustrated in the bottom section of Figure 1, a cloud disturbance does not reach the next system before the next *Time Interval*.

<u>SPACIOUS REGION</u> 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.

Figure 2 presents the model structure for a *PV Fleet* with N *PV Systems*. *Relative Output Variability* declines in the *Crowded* region and reaches a minimum at the *Optimal* point (where the number of systems and *Dispersion Factor* are equal). It increases somewhat in the *Limited* region and then stabilizes in the *Spacious* region.[1]

The figure suggests that *Relative Output Variability* is approximately equal to the inverse of the Dispersion Factor raised to the ¾ power in the *Crowded* region[2] and equals the inverse of the square root of the number of PV systems in the *Spacious* region when systems are independent.

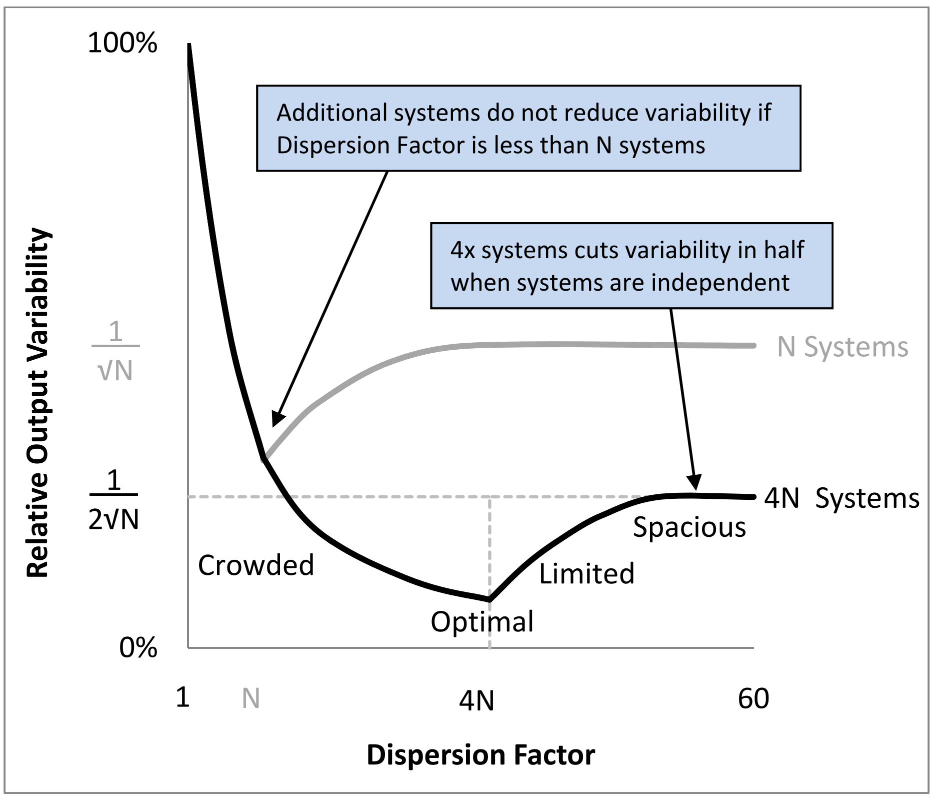
Figure 2. *Variability* for *PV Fleet* with N systems.



[1] A rigorous demonstration of this model is provided in the paper. The paper presents a solution for all regions except the *Limited* region.
[2] This is an empirical result and requires further validation.

Figure 3 presents the structure of the model for four times as many systems (i.e., 4N PV systems). The figure includes the results for N PV systems for comparison purposes. The figure indicates that quadrupling the number of PV systems cuts output variability in half in the *Spacious* region, following the $1/\sqrt{N}$ trend. It also indicates that there is a zone within the very *Crowded* region (*Dispersion Factor* < N) where the additional PV systems provide no added benefit compared to N PV systems. That is, *Relative Output Variability* is the same whether there are N systems or 4N systems when the *Dispersion Factor* is less than N.

Figure 3. *Relative Output Variability* is a function of the number of PV systems and the *Dispersion Factor*.



Model Validation

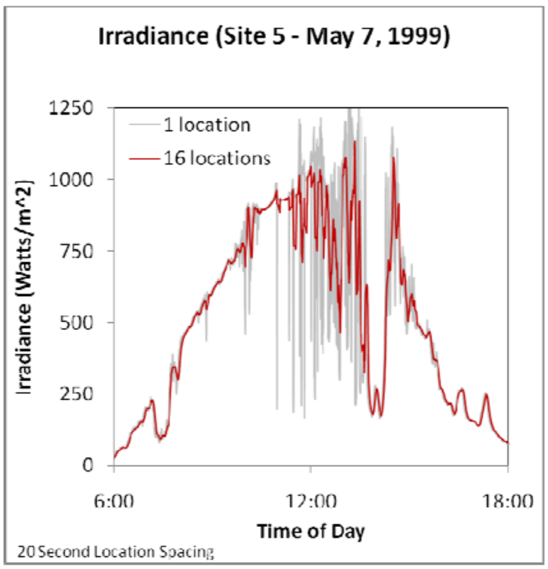
In order to validate the model, 12 *Virtual Networks* are constructed around stations from the Atmospheric Radiation Measurement (ARM) Southern Great Plains Site's extended facility. The ARM facility recorded irradiance data using a 20-second *Data Recording Rate*. The present validation is based on one highly variable day worth of data from the 12 *Virtual Networks*.

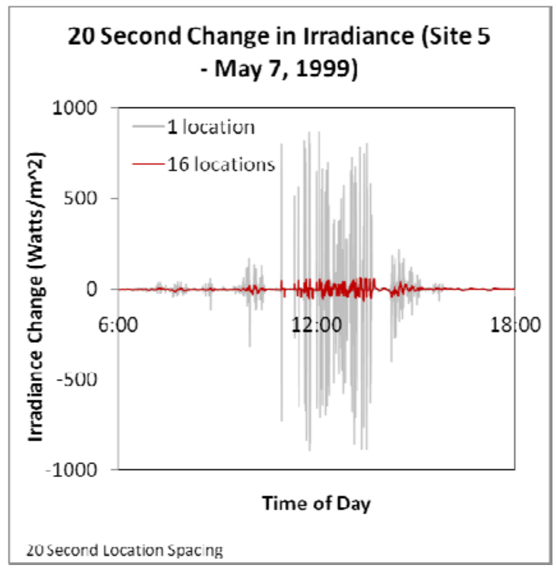
For each network, the model is evaluated using different scenarios, by varying the selected *Time Interval* and/or the number of locations and/or system spacing.

Results for Single Scenario, Single Virtual Network

This subsection presents validation results for one single scenario using one *Virtual Network*. The left side of Figure 4 presents irradiance data (*PV Power Output* is almost directly proportional to irradiance) and the right side of the figure presents the C*hange in Irradiance* using a 20-second *Time Interval* for one of the 12 *Virtual Networks*. The light gray lines correspond to irradiance and variability for a single location and the dark red lines correspond to irradiance and variability for a fleet distributed across 16 locations at the *Optimal* point. Results suggest that, as expected, capacity spread across a fleet of 16 systems greatly reduces variability.

Figure 4. Irradiance and 20-second change in irradiance.

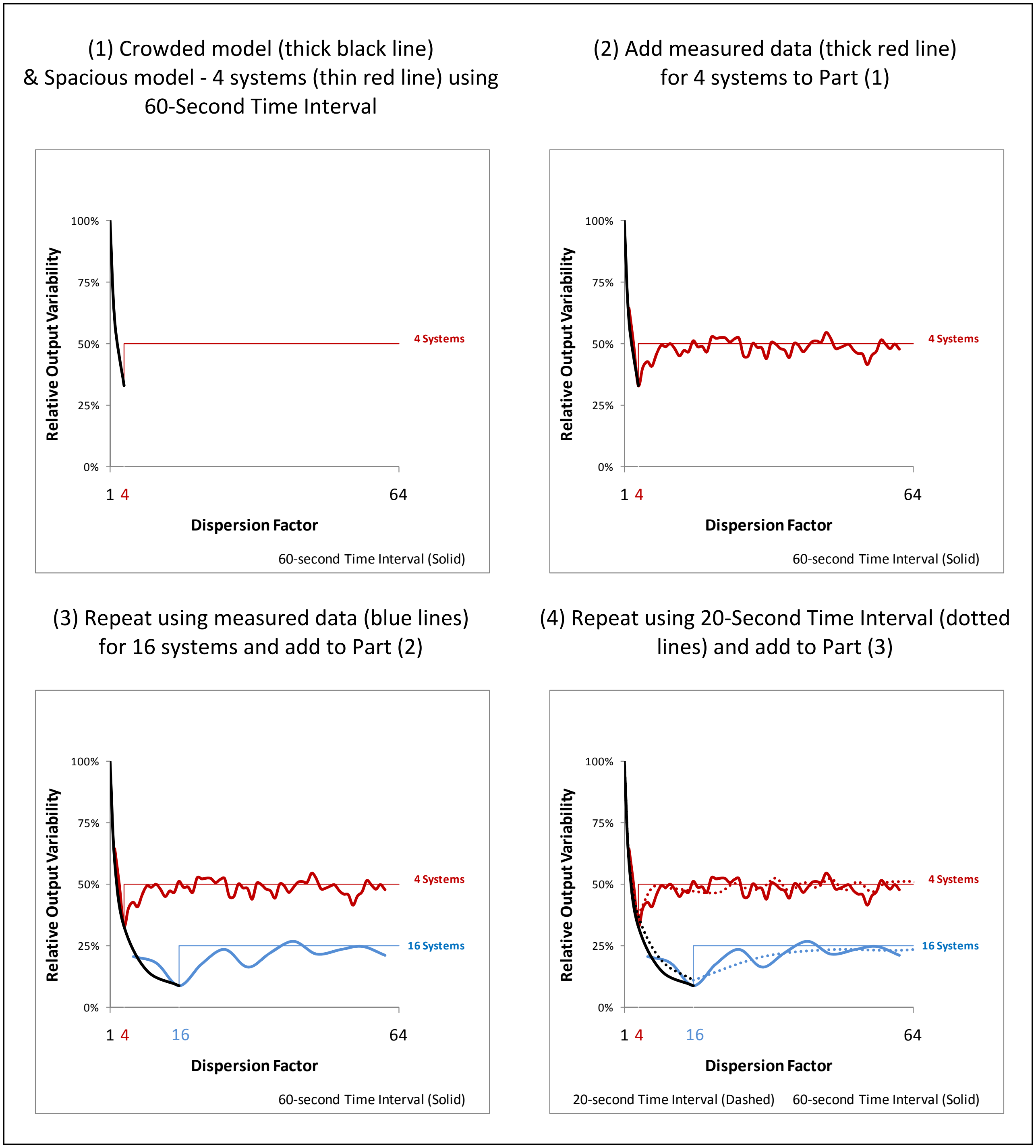




Results for Multiple Scenarios, Single Virtual Network

This subsection presents validation results for scenarios that vary the number of PV systems and the *Dispersion Factor* for a single *Virtual Network*. Figure 5 presents the *Relative Output Variability* resulting from multiple scenarios for one of the 12 *Virtual Networks*. Part (1) of the figure presents the *Crowded* and *Spacious* models for 4 systems using a 60-second *Time Interval*. Part (2) superimposes the ARM experimental data where the PV systems are spaced so as to result in a range of *Dispersion Factors*. Part (3) repeats Parts (1) and (2) and adds model and experimental results for 16 systems. Part (4) repeats Parts (1) through (3) and adds results for scenarios using a 20-second *Time Interval* (dotted lines). The figure suggests that experimental results are closely aligned with the proposed model for all scenarios.

Figure 5. Validation results for One *Virtual Network* (May 7, 1999).



Results for Multiple Scenarios, Multiple Virtual Networks

This subsection presents validation results for multiple scenarios at all 12 *Virtual Networks*.

A key finding of the paper is that *Relative Output Variability* in the *Spacious* region is solely based on the number of systems. That is, results are the same across all networks.

As illustrated in Figure 6, *Relative Output Variability* for a distributed fleet of PV systems (i.e., distant enough to be in the *Spacious* region) equals output variability for capacity concentrated in a single location divided by the square root of the number of systems.

Figure 6. *Variability* for *Spacious* region.

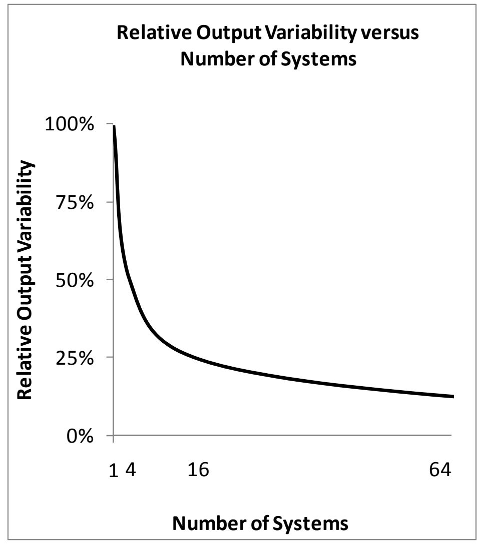


Figure 7. *Variability* for *Crowded* region (includes results for 12 *Virtual Networks*).

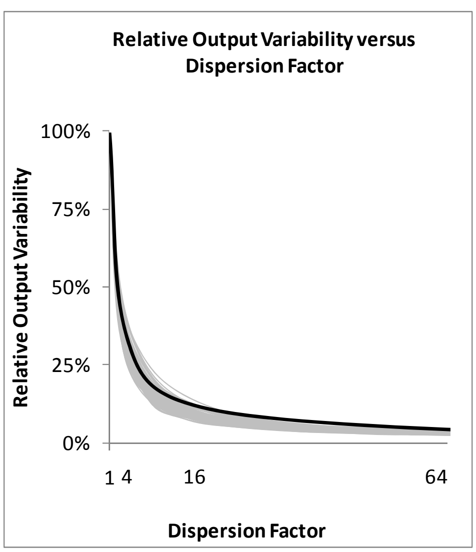


Figure 7 presents an analysis of the results of the *Crowded* region for all 12 *Virtual Networks* across a range of number of systems and a range of *Dispersion Factors* using *Time Intervals* of 20, 40, 60, 80, 100, and 120 seconds. The light gray lines correspond to all scenario results for all 12 *Virtual Networks*.

The figure suggests that results are similar for 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.

These results suggest that both the *Crowded* model and the *Spacious* model may be network-independent. Thus it may be feasible to develop a generally applicable model for all locations. Further work is needed to evaluate this potential.

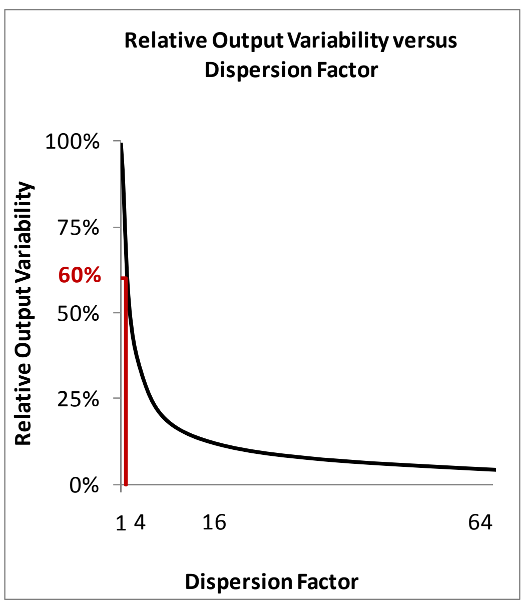
Model Application

The paper includes several examples of how to apply the model. One of the examples is a cursory analysis of the reported wintertime variability of Tucson Electric's 5 MW_{DC} Springerville plant. Assume that the plant is 420 meters across, the *Time Interval* of interest to the utility is 1 minute, and there are at least 3 separate systems. When clouds are moving at a rate of 3.5 meters per second, the plant has a *Dispersion Factor* of 2 because it takes 2 *Time Intervals* for a cloud to pass across the plant.[3]

Figure 8 overlays a *Dispersion Factor* of 2 on Figure 7. According to the model, this *PV Fleet* would be expected to have output variability equal to about 60 percent of a single location under such conditions. This may provide an explanation for why output variability at the Springerville plant has been reported to be high during certain times of the year.

Had the Springerville plant been designed as a more distributed fleet, its variability could have been dramatically reduced. For example, the same 5 MW capacity could have been installed as 1,000 five-kW systems. This fleet, while providing the same capacity, would have had about one-twentieth of the output variability as the central plant. Another approach would have been to build a PV plant that encompassed a larger geographic area to increase the *Dispersion Factor* and thus reduce the *Relative Output Variability*.

Figure 8. Results for Springerville plant when *Dispersion Factor* is 2.



[3] (2 *Time Intervals*) x (60 seconds/*Time Interval*) x (3.5 meters/second) = 420 meters.

Conclusion and Next Steps

Conclusions

This paper presents a novel approach to rigorously quantify power output variability from a fleet of PV systems, ranging from a single central station to a set of distributed PV systems. The approach demonstrates that the *Relative 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 the *Dispersion Factor*. The *Dispersion Factor* is a 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 dispersed PV systems; and (2) can be minimized for optimally-spaced PV systems. The approach can be used to analyze an existing fleet's output variability or to control output variability when designing or providing incentives for the construction of a new fleet.

Next Steps

The paper concludes by outlining suggested research directions. The broad research directions include:

- Validate the model further using high-frequency irradiance or PV output data measured across a fleet of locations or systems.
- Extend the model to calculate results for an arbitrary *PV Fleet* configuration.
- Model absolute, rather than relative, output variability as a function of location, climate, and weather conditions.

Background

There is a sustained and growing interest in grid-connected photovoltaic (PV) systems in the U.S. Incentives from the New York State Energy and Research Development Authority (NYSERDA) over the past several years have resulted in the installation of several thousand PV systems in the state of New York. The increase in federal support for solar, particularly in the form of the 30 percent federal tax credit, is likely to increase the number of PV systems. Other states are having similar experiences.

Residential, commercial, and utility sectors are all likely to see an increase in the number of systems as a result of changes to the federal tax credit. First, more residential customers will install systems because they will receive a larger tax credit as a result of the transition from a capped to an uncapped residential tax credit (it was previously limited to $2,000). Second, more commercial customers will install systems because of the year-to-year tax credit continuity. Third, investor-owned utilities will increase their investment in PV systems because they are now eligible to receive the tax credit.

The result is that there is going to be an increase in the number of grid-connected PV systems. Accompanying this increase is likely to be a greater concern about the potential impact of a high penetration of PV systems on the utility grid.

Introduction

The concern about the potential effects of a high penetration of PV is sufficiently strong that the U.S Department of Energy conducted a workshop on the topic in February 2009. The workshop was designed to assist the Department of Energy in constructing a research agenda that will address core issues associated with high-penetration PV. More than 100 industry professionals attended this workshop, thus signifying the depth of interest in the topic.

A recurring concern at the workshop was that short-term PV output variability, particularly due to cloud movements, could negatively affect electric utility grid stability. Workshop participants assumed that there will be a lot of variability and many moved directly to the question about how to address the issue.

The concern may be triggered by high levels of PV output variability reported at some central station PV facilities, particularly during cloudy winter conditions. Minimal information, however, is available about the extent of short-term variability issues. Even less information is available about how the intelligent design/location of a fleet of PV systems within a geographical area might be used to address the effect of short-term variability.

Or the concern may be triggered by an underlying assumption that some sort of mitigation effort is required to protect against short-term variability because PV, like wind, is powered by a non-controllable renewable resource. It is important to recognize, however, that there are fundamental differences between PV and wind. The most important difference is that PV power is proportional to irradiance while wind power varies as the cube of the wind speed. For example, if both irradiance and wind speed double over a very short interval of time, PV output would increase by 100 percent while wind generation would increase by 700 percent.

Whatever the cause, the general perception is that this issue could adversely impact the adoption of grid-connected PV.

Literature Review

There has been a fair amount of work devoted to understanding the variability associated with the solar resource for a single location. Suehrke and McCormick (1989) published one of the early papers identifying the nature of high frequency irradiance data (e.g., one minute) as fundamentally different from lower frequency data (e.g., hourly) generally available in archives and typical meteorological year (TMY) files, noting 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 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 one-minute and hourly simulations and showing that systems with operational thresholds could not be properly simulated if sub-hourly information was not available. 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 the load and supply side.

An influential paper in the modeling of short-term variability is that of Skartveit & Olseth (1992). They showed that the distribution of sub-hourly GHI and DNI within a given hour 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 an hour to the next. This 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 (2005) or SolarAnywhere® (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.

Minimal work, however, has been devoted to understanding the effect on irradiance variability of combining multiple locations. Three of the most interesting papers are based on measured data from a network of locations in Japan.

Otani et. al. (1997) examined variability associated with one-minute irradiance measurements obtained from a nine-site network concentrated 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 the nine-site average decreased to around 20 to 50 percent of each site independently during partly cloudy conditions. 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 focus of the paper, they presented the correlation between two-system combinations using a 1 minute time interval. They found that the correlation was close to zero (implying that the locations are independent) unless the distance between the two systems was very short (see Figure 9 in that paper) or the time interval was long.

There are four notable limitations in the work of Otani et. al. (1997), Kawasaki et. al., (2006), and Murata et. al. (2009) when considered for PV output variability. First, Otani et. al. (1997) focused on the evaluation of the difference between instantaneous irradiance and average irradiance. A more relevant evaluation from the operational perspective of a utility is how much instantaneous irradiance changes over some given time period. Second, their work did not provide a sufficient mathematical explanation as to why the "smoothing effect" occurs. Third, their work lacked a general model that could be applicable for any number of PV systems in applications that range from central station to distributed generation. Fourth, utility system operators are likely to be very interested in the reduction in output variability that results from having a distributed fleet of PV systems versus having all capacity concentrated in a single location.

Objective

The objective of this work is to provide a general model that quantifies absolute power output variability from a fleet of arbitrarily-configured PV systems. As a first step toward that objective, this paper describes, validates, and applies a model that quantifies relative power output variability for a fleet of identical PV systems (same size, orientation, and spacing) distributed in the direction of prevailing cloud motion. This simplified layout facilitates the analysis of most PV deployment scenarios, from single central station to fully distributed configurations.

Definitions

It is useful to begin the analysis with a set of definitions.

Time Measurements

Three different time measurements are defined here: *Data Recording Rate*, *Time Interval*, and *Analysis Period*.

Data Recording Rate

Data Recording Rate is the frequency at which data observations are recorded.

Time Interval

Time Interval (also referred to as Δt) is the duration over which one data measurement is compared to another data measurement. The difference in power generation from one time interval to the next is the underlying quantity that defines short-term variability.

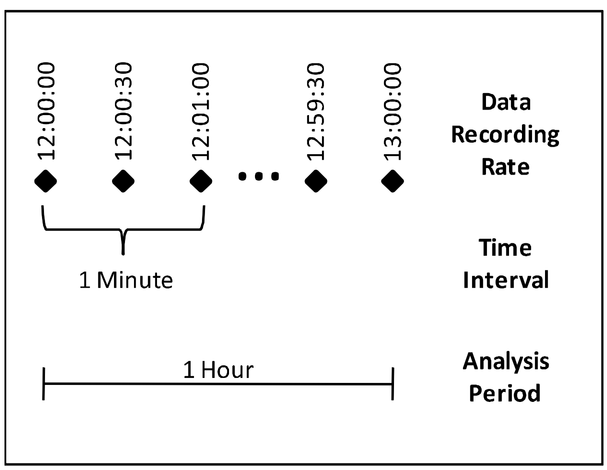
Analysis Period

Analysis Period is the period over which the analysis is performed.

Time Measurement Relationships

Figure 9 illustrates the relationship between data that have a 30-second *Data Recording Rate*, a 1-minute *Time Interval*, and a 1-hour *Analysis Period*.

Figure 9. Relationship between *Data Recording Rate*, *Time Interval*, and *Analysis Period*.



Irradiance

Let I_t^n represent the irradiance (Watts/m^2) that is incident on a PV system at time t in location n.

Change in Irradiance

Let $\Delta I_{t,\Delta t}^n$ represent the change in irradiance between time *t* and t+Δt at location n. That is,

$$\Delta I_{t,\Delta t}^n = I_t^n - I_{t+\Delta t}^n \qquad (1)$$

PV System Capacity

Individual System Capacity

Let C^n represent the capacity of a PV system at location n. For purposes of this paper, it is assumed that there is only one PV system per location.

Fleet Capacity

Let C^{Fleet} represent the capacity of all PV systems at all locations. The total capacity of the fleet of PV systems equals the sum of capacity across all systems at all locations. That is, fleet capacity equals:

$$C^{Fleet} = \sum_{n=1}^{N} C^n \qquad (2)$$

PV Power

Let P_t^n represent the amount of power produced by the PV system at time t in location n. Since PV power production is approximately[4] directly proportional to the product of capacity and incident irradiance, PV power production at time t in location n is equal to some constant (a) times PV capacity (C^n) times incident irradiance.

$$P_t^n = aC^n I_t^n \qquad (3)$$

Change in Power

This paper is concerned with the *Change in Power* over a given *Time Interval* (Δt). Let $\Delta P_{t,\Delta t}^n$ represent the *Change in Power* between time t and time t+Δt at location n. The *Change in Power* at location n equals the power at time t minus the power at time t+Δt.

$$\Delta P_{t,\Delta t}^n = P_t^n - P_{t+\Delta t}^n \qquad (4)$$

Substituting Equation (3) into Equation (4) and referring to Equation (1) results in the expression of the change in power as PV capacity times the change in irradiance.

[4] It is approximate because there is a temperature effect. The temperature effect will be neglected in this paper.

$$\Delta P_{t,\Delta t}^{n} = aC^{n}\Delta I_{t,\Delta t}^{n} \tag{5}$$

Change in Irradiance Random Variable

Let $\Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}}$ represent a random variable[5] that summarizes the change in irradiance over a set of times. In particular, $\Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}}$ is a random variable that represents the set of changes over a given *Analysis Period* (from time 1 to time T) using a *Time Interval* of Δt. That is, $\Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}}$ equals the set of ordered pairs where:

$$\Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}} = \left\{ \left(t_1, \Delta I_{1,\Delta t}^{n}\right), \left(t_2, \Delta I_{2,2+\Delta t}^{n}\right), \cdots, \left(t_T, \Delta I_{T,T+\Delta t}^{n}\right) \right\} \tag{6}$$

Cloud Transit

Three aspects of Cloud Motion or *Cloud Transit* are important: *Cloud Transit Rate*, *Cloud Transit Time*, and *Cloud Transit Distance*.

Cloud Transit Rate

Cloud Transit Rate equals the speed at which clouds are moving.

Cloud Transit Time

Cloud Transit Time is the amount of time required for a cloud to fully pass across the *PV Fleet*. When the PV systems are treated as points, then the measurement equals the amount of time from the first to the last system plus 1 *Time Interval* (so as to move past the last system).

Cloud Transit Distance

Cloud Transit Distance is the distance that a cloud moves to fully pass across the *PV Fleet*.

Relationship to Distance

Since time equals distance divided by rate, *Cloud Transit Time* can be expressed as

$$Cloud\ Transit\ Time\ = \frac{Cloud\ Transit\ Distance}{Cloud\ Transit\ Rate} \tag{7}$$

When PV systems are treated as single points, the amount of time required for a cloud to move completely across the *PV Fleet* equals the distance between the first system and the last system divided by the *Cloud Transit Rate* plus one *Time Interval*.

$$Cloud\ Transit\ Time\ = \Delta t + \frac{Distance\ Between\ First\ and\ Last\ System}{Cloud\ Transit\ Rate} \tag{8}$$

[5] S. Ross (1988) presents a good discussion of random variables.

Dispersion Factor

Dispersion Factor (*D*) is a novel concept that is defined in this paper. It captures the relationship between *PV Fleet* configuration (i.e., number and orientation of systems and their geographic density), *Cloud Transit Time,* and *Time Interval*.

More specifically, *Dispersion Factor* equals the number of *Time Intervals* required for a cloud to move across the *PV Fleet*. Stated another way, *Dispersion Factor* times *Time Interval* (Δt) equals *Cloud Transit Time*.

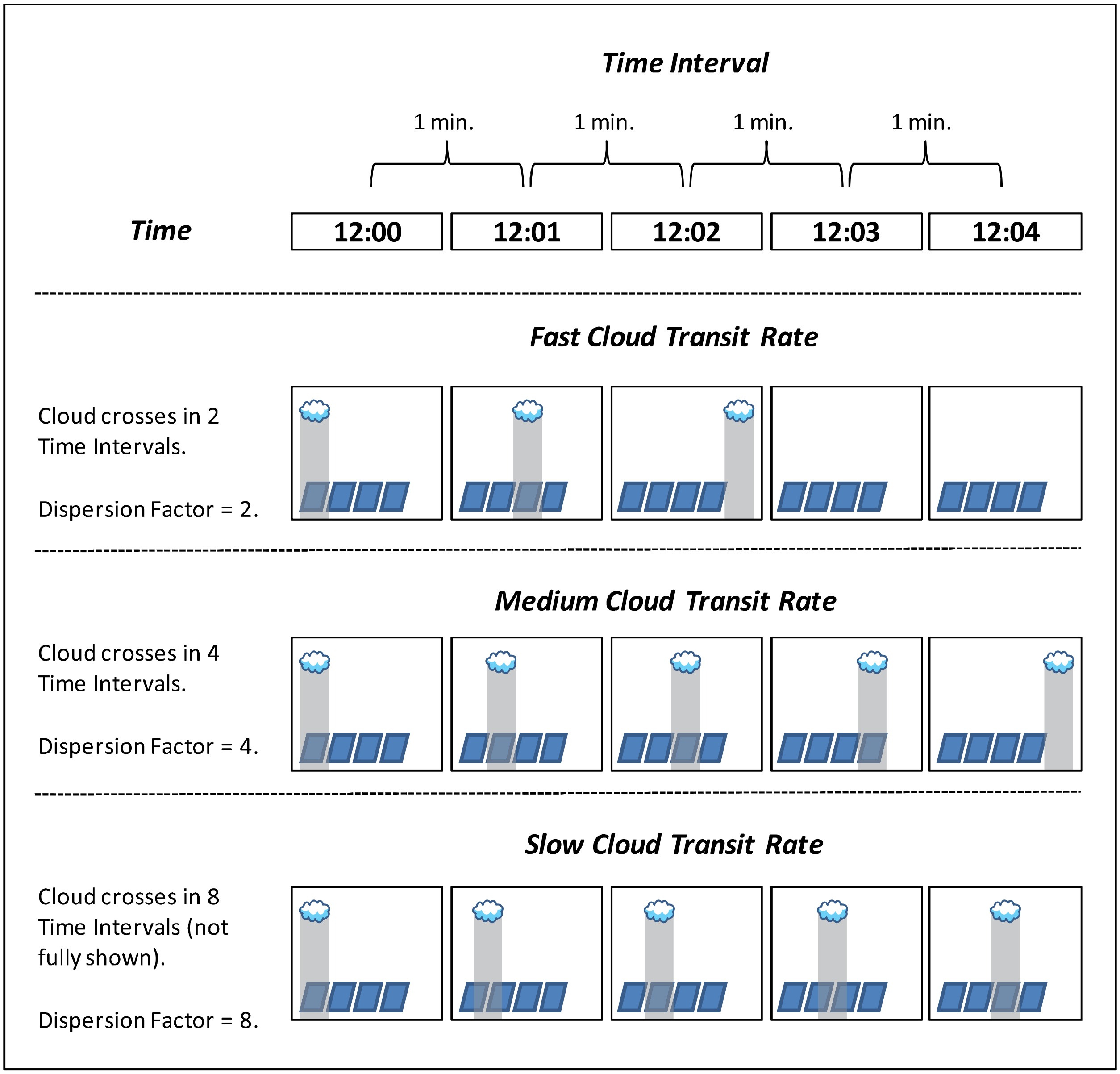
$$D\Delta t = Cloud\ Transit\ Time \qquad (9)$$

Substituting Equation (8) into Equation (9) and dividing by Δt results in

$$D = 1 + \frac{Distance\ Between\ First\ and\ Last\ System}{Cloud\ Transit\ Rate\ \ x\ \Delta t} \qquad (10)$$

Figure 10 illustrates how the *Dispersion Factor* is calculated for a *PV Fleet* with 4 PV systems using a 1-minute *Time Interval* under three *Cloud Transit Rate* conditions: fast, medium, or slow. The fast-moving cloud in the top section of figure crosses the *PV Fleet* in 2 minutes (2 *Time Intervals*) and thus has a *Dispersion Factor* of 2. The medium-speed cloud in the middle section of the figure crosses the *PV Fleet* in 4 minutes (4 *Time Intervals*) and thus has a *Dispersion Factor* of 4. The slow-moving cloud in the bottom section of the figure crosses the *PV Fleet* in 8 minutes (8 *Time Intervals* – not fully shown) and thus has a *Dispersion Factor* of 8. The *Dispersion Factor* increases as the wind speed decreases.

Figure 10. *Dispersion Factor* for a *PV Fleet* with 4 PV systems using a 1-minute *Time Interval* when the *Cloud Transit Rate* is fast, medium, or slow.



Variance and Standard Deviation

Two statistics that summarize the behavior of a random variable are the mean (μ) and the standard deviation (σ). The mean is the average value of the random variable. A limitation of the mean is that is does not describe the random variable's level of variability. The standard deviation is a measure of the random variable's variability. A more mathematically convenient form of standard deviation is the square of the standard deviation, also known as the variance (σ^2).

Problem Formulation

The objective of this paper is to develop a model to quantify plant output variability based on weather conditions and PV fleet configuration. Two simplifications will facilitate initiating such an analysis. First, since PV output is proportional to irradiance, the paper will focus on irradiance variability. Second, it will be assumed that PV plant dispersion is uniform across N locations. Thus, the objective is to develop a model that quantifies PV output variability for a fleet of N identical, evenly-spaced PV systems.

The magnitude of the change in power output is measured by the standard deviation. For purposes of convenience throughout the paper, rather than solving for the standard deviation of the change in power, the solution will be for *a* times total PV capacity times the effective standard deviation across N locations. That is, PV fleet output variability equals:

$$aC^{Fleet}\sigma_{\Delta t}^{\sum N} = \sqrt{Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{P}_{\Delta t}^{n}\right]} \tag{11}$$

Or

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{aC^{Fleet}}\right)\sqrt{Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{P}_{\Delta t}^{n}\right]} \tag{12}$$

Substituting Equation (5) into Equation (12) results in:

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{aC^{Fleet}}\right)\sqrt{Var\left[\sum_{n=1}^{N} aC^{n}\Delta \boldsymbol{I}_{\Delta t}^{n}\right]} \tag{13}$$

It is assumed throughout the paper that the PV plant capacity is equally distributed among N locations. That is, for any location n, $C^n = C^{Fleet}/N$. In addition, it is assumed that all plants have the same orientation.

As a result, Equation (13) simplifies to:

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{N}\right)\sqrt{Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{I}_{\Delta t}^{n}\right]} \tag{14}$$

Problem Solution

The focus of this paper is on solving Equation (14). The model characterizes *Relative Output Variability* over 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 Figure 10, 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 Figure 10, 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 Figure 10, 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.

Spacious Region: Number of PV Systems Much Less Than Dispersion Factor

First, consider the *Spacious* region. This applies when the number of PV systems is much less than the *Dispersion Factor*.[6] In this region, systems are sufficiently far apart such that the change in irradiance at any system is independent of the change in irradiance at any other system. When each random variable $\Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}}$ is independent, the variance of their sum equals the sum of the variances. Equation (14) can be simplified by moving the variance inside the summation.

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{N}\right) \sqrt{\sum_{n=1}^{N} Var[\Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}}]} \qquad (15)$$

If the change in irradiance at all systems has the same standard deviation, the standard deviation for any system can be selected. The system at location 1 is arbitrarily selected.

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{N}\right) \sqrt{\sum_{n=1}^{N} (\sigma_{\Delta t}^{1})^2} \qquad (16)$$

[6] For example, suppose that 4 PV systems are each located 2,000 meters apart so that *Cloud Transit Distance* equals 6,000 meters from the first to last PV system. At a *Cloud Transit Rate* of 5 meters per second, it takes a cloud 1,200 seconds to move across the *PV Fleet*. The *Dispersion Factor* is 21 if the *Time Interval* is 60 seconds.

The result is that the standard deviation of the change in irradiance across all N systems equals the standard deviation in the change in irradiance at the first system divided by the square root of the number of systems.

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

Optimal Point: Number of PV Systems Equal Dispersion Factor

Next, consider the *Optimal* point. This applies when the number of PV systems equals the *Dispersion Factor* (N = D). An analysis of power output variability at the *Optimal* point requires a novel analytical approach.

Virtual Network

The ideal data to use in validating the model in this paper would be either high-frequency irradiance data from high-density, large-area, gridded-networks of pyranometers or high-frequency power output data from networks of identical PV systems. Networks specifically designed for this purpose do not currently exist.

An alternative set of data can be developed by constructing a *Virtual Network* based on data measured at a single, actual location. Irradiance data at a virtual location is specified by assuming that the cloud-induced patterns measured at the actual location move at a constant speed across the *Virtual Network*. Data for a virtual location equals data measured at the actual location at the point in time that corresponds to the amount of time required for a cloud to travel from the actual location to the virtual location. That is, the data for the virtual location is obtained by measuring the data at the actual location at a given time stamp and then adding the amount of time required to move from the actual location to the virtual location to the time stamp. Such a *Virtual Network* facilitates the analysis of scenarios that vary the number of systems and the *Dispersion Factor*.

Table 1 illustrates how to construct a *Virtual Network*. Measured irradiance data is presented in the first column of Table 1. Assume that the *Cloud Transit Rate* is 2.5 meters per second and that irradiance is translated to the next location after a *Time Interval* (Δt) of 20 seconds. The first yellow stair-step in Table 1 illustrates that irradiance that is 852 Watts/m^2 at 12:00:00 will be 852 Watts/m^2 after 1 *Time Interval* at 12:00:20 in a virtual location that is 50 meters away, 852 Watts/m^2 after 2 *Time Intervals* at 12:00:40 in a virtual location that is 100 meters away, and 852 Watts/m^2 after 3 *Time Intervals* at 12:01:00 in a virtual location that is 150 meters away.

If one-quarter of the PV fleet is located at each of the four locations, the irradiance of the fleet will be 479 Watts/m^2 at 12:01:00. Since the combined irradiance after one *Time Interval* (20 seconds later) is 366 Watts/m^2, the change in irradiance is 113 Watts/m^2 over a 20-second *Time Interval*.

Table 1. Projection of measured irradiance (Watts/m^2) to virtual network.

	IRRADIANCE					CHANGE IN IRRADIANCE
	Actual Location	Virtual Locations listed by Cloud Transit Time (& Cloud Transit Distance) 20 Sec. (50 meters)	40 Sec. (100 meters)	60 Sec. (150 meters)	**Fleet Distributed Across 4 Locations**	Watts per m^2 per 20 seconds
12:00:00	852					
12:00:20	352	852				
12:00:40	112	352	852			
12:01:00	600	112	352	852	**479**	
12:01:20	400	600	112	352	**366**	***113***
12:01:40	952	400	600	112	**516**	***-150***
12:02:00	88	952	400	600	**510**	***6***
12:02:20	112	88	952	400	**388**	***122***
12:02:40	300	112	88	952	**363**	***25***
12:03:00	600	300	112	88	**275**	***88***

The Appendix uses this approach to define a *Virtual Network* and then solves the equation. The result is that the standard deviation for N systems each separated by one *Time Interval* equals the standard deviation between output separated by a time of N x Δt divided by the number of systems. That is,

$$\sigma_{\Delta t}^{\sum N} = \frac{\sigma_{N\Delta t}^{1}}{N} \qquad (18)$$

According to Equation (9) and since D = N, *NΔt* corresponds to the *Cloud Transit Time* (i.e., the amount of time required for a cloud to move across the *PV Fleet*)

Limited Region: Number of PV Systems Less Than Dispersion Factor

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, Equation (17) will apply as the upper bound in this range, because the independence condition would be violated if the locations were any closer than that.

Crowded Region: Number of PV Systems Greater Than Dispersion Factor

Finally, consider the *Crowded* region. This applies when the number of PV systems is greater than the *Dispersion Factor*. 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. In this range, Equation (18) will apply. However, it will use *Dispersion Factor* (D) in place of the number of systems (N).

Solution

The result is that the solution to this problem depends upon the region that is applicable. In more concrete terms, it depends upon whether the *PV Fleet* is central station (high concentration) or if it is a set of geographically diverse distributed generation system.

The four regions can be summarized in the following equation. The top part of the solution covers the *Spacious* region. The middle part covers the *Limited* region. The bottom part covers the *Optimal* point and *Crowded* region.

$$\sigma_{\Delta t}^{\sum N} \begin{array}{l} = \dfrac{\sigma_{\Delta t}^{1}}{\sqrt{N}} \; for\ N \ll D \text{ (Spacious)} \\ < \dfrac{\sigma_{\Delta t}^{1}}{\sqrt{N}} \; for\ N < D \text{ (Limited)} \\ = \dfrac{\sigma_{D\Delta t}^{1}}{D} \; for\ N \geq D (\text{Optimal } N = D, \text{Crowded } N > D) \end{array} \quad (19)$$

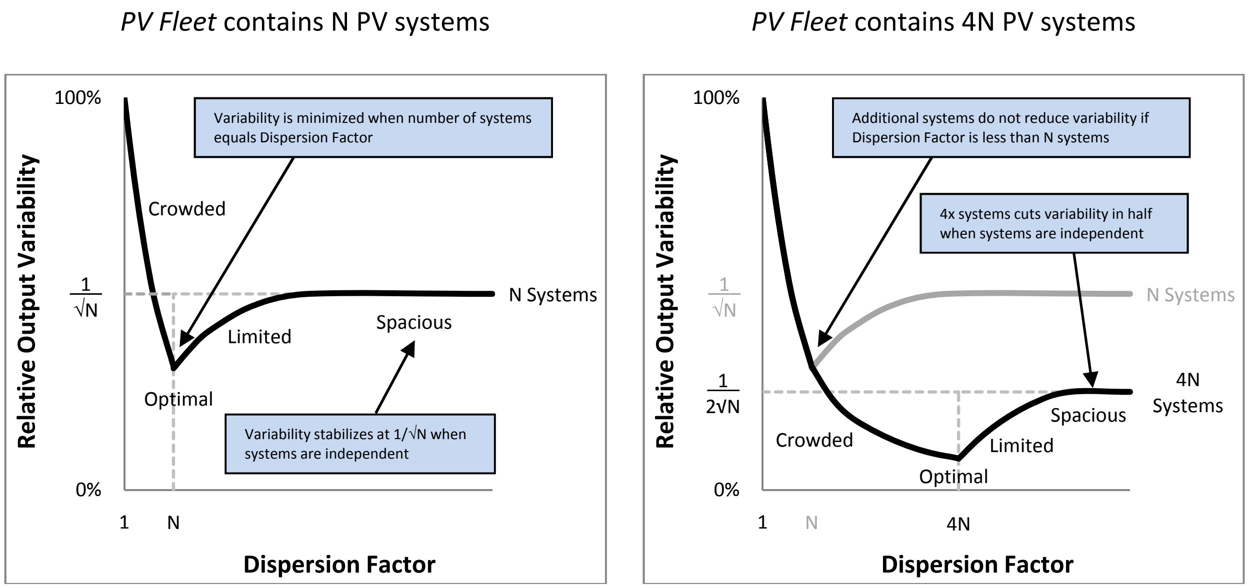
An inspection of Equation (19) suggests that solutions are related. First consider the *Spacious* region. The standard deviation of a fleet that is composed of N identical systems equals the standard deviation of the change in output over a *Time Interval* of Δt for one system divided by the square root of the number of systems. Next consider the *Optimal* point and *Crowded* region. $D\Delta t$ equals *Cloud Transit Time*. This means that, for the *Optimal* point and *Crowded* region, the standard deviation of the combined locations equals the standard deviation of the change in output over a *Time Interval* of "Cloud Transit Time" seconds (rather than the *Time Interval* of Δt seconds as in the *Spacious* region) divided by the *Dispersion Factor* (rather that the square root of the number of *PV Systems* as in the *Spacious* region).

Discussion

Before proceeding to a validation of the results, it is useful to discuss the fundamental shape of the results based on Equation (19). A sample curve is presented in the left side of Figure 11 for N Physical Locations. The figure suggests that output variability reduces in the *Crowded* region and reaches a minimum at the *Optimal* point. It increases somewhat in the *Limited* region and then stabilizes in the *Spacious* region where each location is independent.

One of the findings of this work is that the output variability is proportional to the inverse of the square root of the number of *Physical Locations*. The right side of Figure 11 presents the structure of the model for four times as many locations (i.e., 4N *Physical Locations*). The figure suggests that quadrupling the number of *PV Systems* cuts output variability in half in the *Spacious* region. It also indicates that output variability is the same as that of output variability of N *PV Systems* when *Dispersion Factor* is less than or equal to N (i.e., the first part of the *Crowded* region). This implies that more *PV Systems* do not translate into reduced output variability in this region, a situation that can apply to central generation plants.

Figure 11. *Relative Output Variability* is function of the number of PV systems and the *Dispersion Factor*.



The following observations can be made:

- Output variability general decreases with an increasing number of systems.
- The *Dispersion Factor* relative to the number of locations can limit the output variability reduction. For example, locating too many individual plants too close together reduces the value of the larger number of locations.
- Maximizing location spacing for a fixed number of locations does not minimize output variability. While this is initially a surprising result, it makes sense upon reflection, because the locations benefit from being close together up to a point due to negative covariance.
- Output variability goes to a fixed number (1 divided by the square root of N) given large enough location spacing.

Validation

The results in Equation (19) can be validated as follows.

1. Obtain measured high-frequency irradiance data for a single location
2. Translate irradiance to a *Virtual Network* of locations
3. Combine the output from the locations in the *Virtual Network* to create a *PV Fleet*
4. Calculate the rate of irradiance change
5. Summarize results in the standard deviation

Single Virtual Network

The analysis was performed using data from a single actual location (labeled as Site 5).

1. Figure 12 presents the irradiance data measured every 20 seconds at Site 5 for May 7, 1999. The figure illustrates that this site had a high degree of irradiance variability on that particular day.
2. Translate the measured irradiance from the single location to virtual locations as described above. In this example, the measured data is translated to 15 virtual locations located 20 seconds apart to result in a total of 16 locations. That is, it requires 5 minutes for the measured irradiance to translate from the measured location to the 15th virtual location.
3. Take 1/16 of the output from a plant based on the irradiance from each location to create a fleet that is the combined output from 16 systems. The resulting irradiance is overlaid on top of the measured irradiance from the single location in Figure 13.
4. Calculate the 20-second rate of irradiance change by subtracting the irradiance 20 seconds later from the irradiance at the current time. Repeat this for all times. The results are presented in Figure 14. As expected, results indicate that the rate of irradiance change from the combination of 16 locations is greatly reduced compared to the rate of irradiance change from a single location.
5. The standard deviation of the rate of change for the one location is 117 Watts/m^2 over the 20-second interval. The standard deviation of the rate of change for the combined 16 locations is 13 Watts/m^2 over the 20-second interval. Thus, the 16-location network has 11 percent the standard deviation as the single location.

Figure 12. Irradiance measured every 20 seconds.

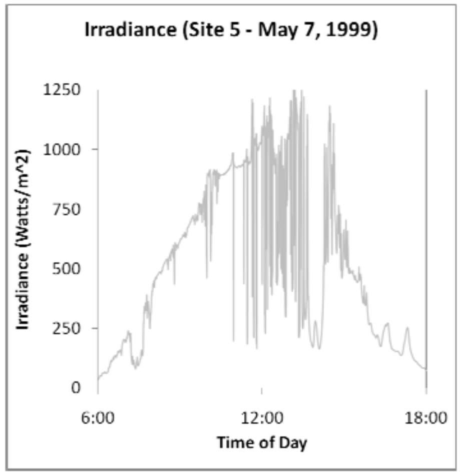


Figure 13. Combined irradiance at 16 locations (1 measured, 15 virtual).

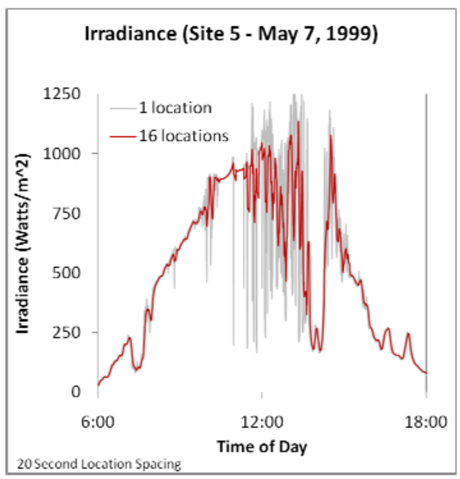
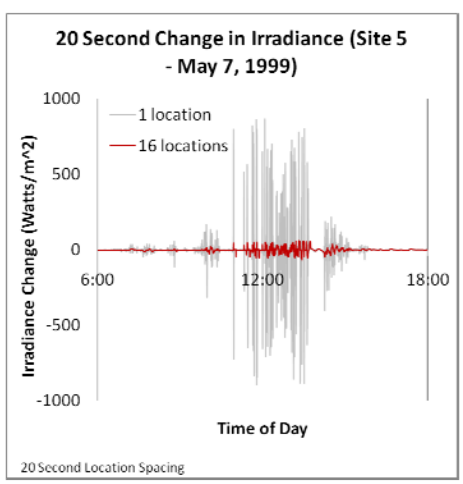
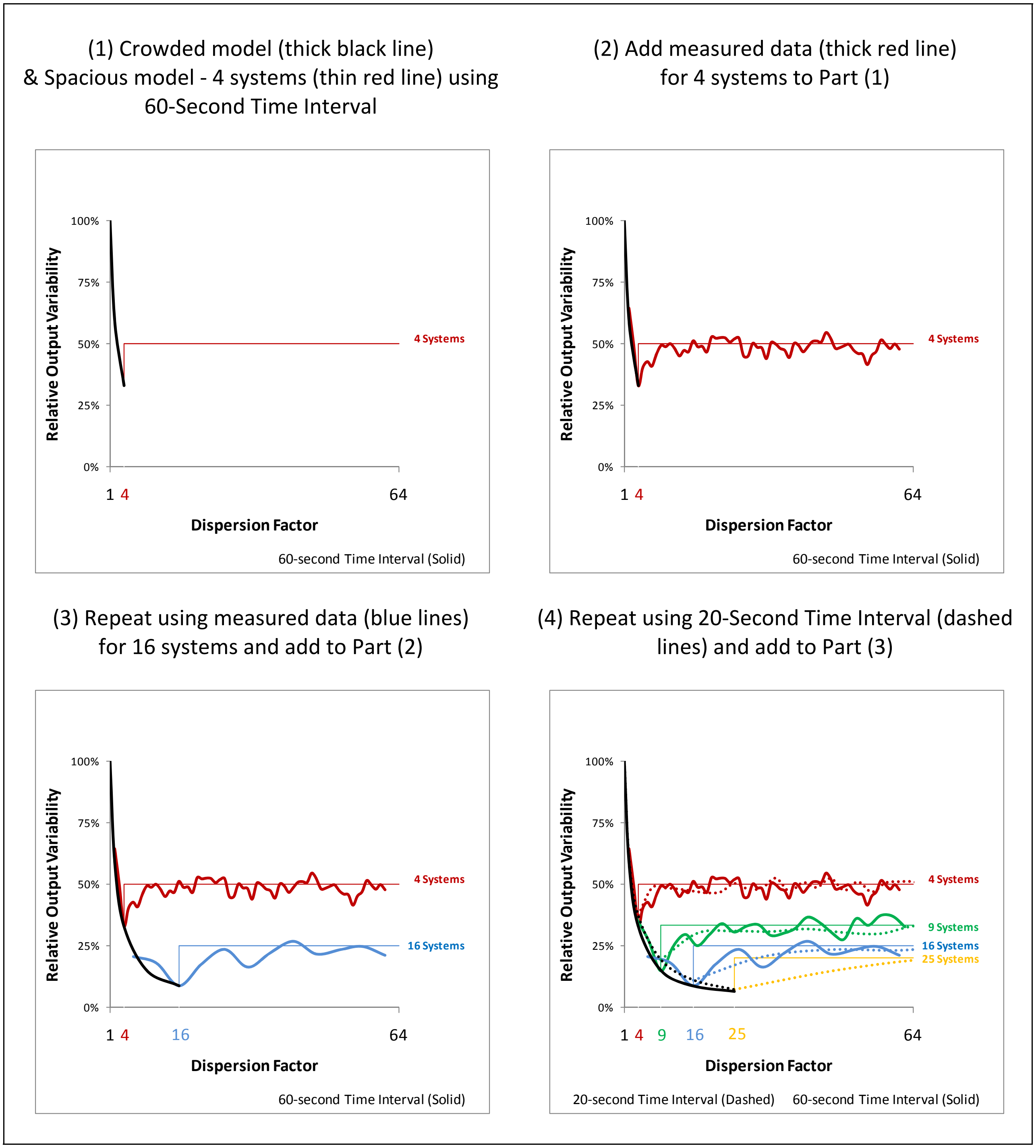


Figure 14. 20-second change in irradiance.



While the results for this one scenario are promising, it is useful to examine the results for a wider range of scenarios by varying the number of PV systems and the *Dispersion Factor*. Figure 15 presents the *Relative Output Variability* resulting from multiple scenarios for one of the 12 *Virtual Networks*. Part (1) of the figure presents the *Crowded* and *Spacious* models for 4 systems using a 60-second *Time Interval*. Part (2) superimposes the ARM experimental data where the PV systems are spaced so as to result in a range of *Dispersion Factors*. Part (3) repeats Parts (1) and (2) and adds model and experimental results for 16 systems. Part (4) repeats Parts (1) through (3) and adds results for scenarios for 9 and 25 systems and uses a 20-second *Time Interval* (dotted lines). The figure suggests that experimental results are closely aligned with the proposed model for all scenarios.

Figure 15. Model validation results for Site 5 (May 7, 1999).



Replication of Results to 12 Virtual Networks

While these single-network results are promising, it is beneficial to determine how results compare across multiple independent sites. *Virtual Networks* were created based on the experimental ARM data for 12 locations that had different weather characteristics. The mean and standard deviation of the single-site variability are presented in Figure 16. The figure illustrates that the 12 sites represent a variety of weather scenarios.

Figure 17 presents the irradiance and the 20-second change in output for a single system and for 16 systems for all 12 virtual networks. In each site, the 16-system fleet greatly reduces output variability.

Figure 18 summarizes the results for a 20-second and 60-second *Time Interval* for 4, 9, 16, and 25 systems with a wide variety of location spacing. The results are consistent across all sites and are similar to the results presented above.

Figure 16. Summary statistics for 12 geographically distinct sites.

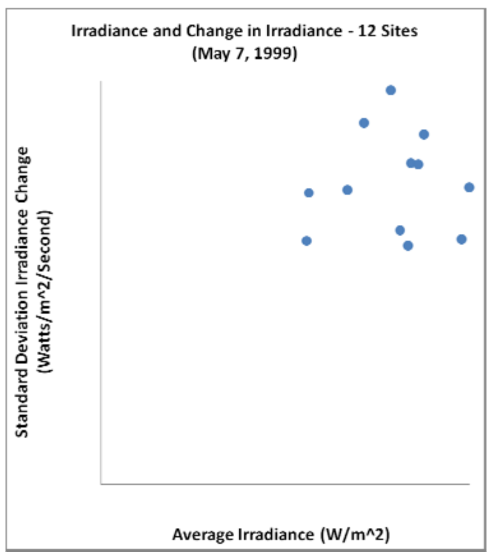


Figure 17. Irradiance and change in irradiance.

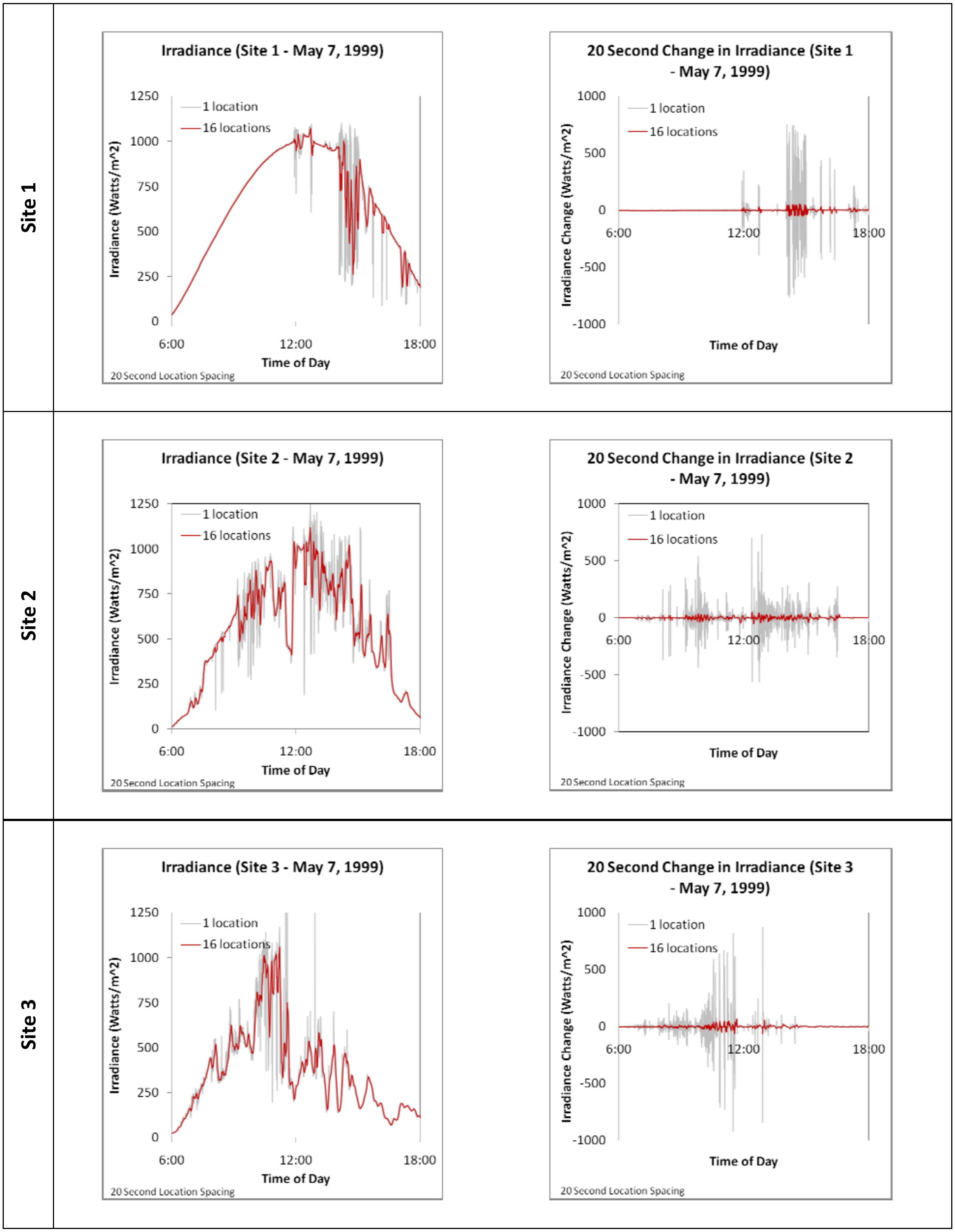


Figure 17. Irradiance and change in irradiance (cont.).

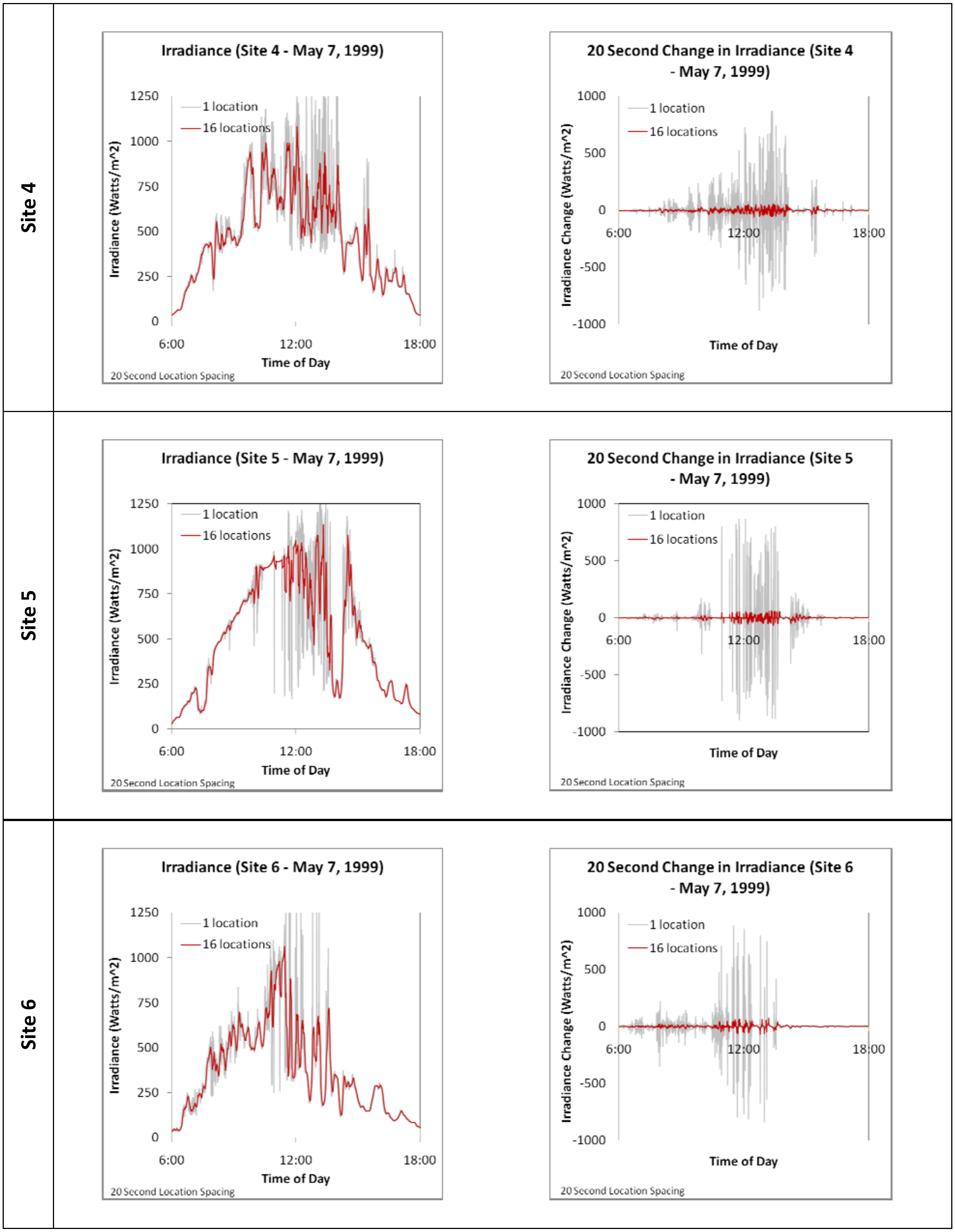


Figure 17. Irradiance and change in irradiance (cont.).

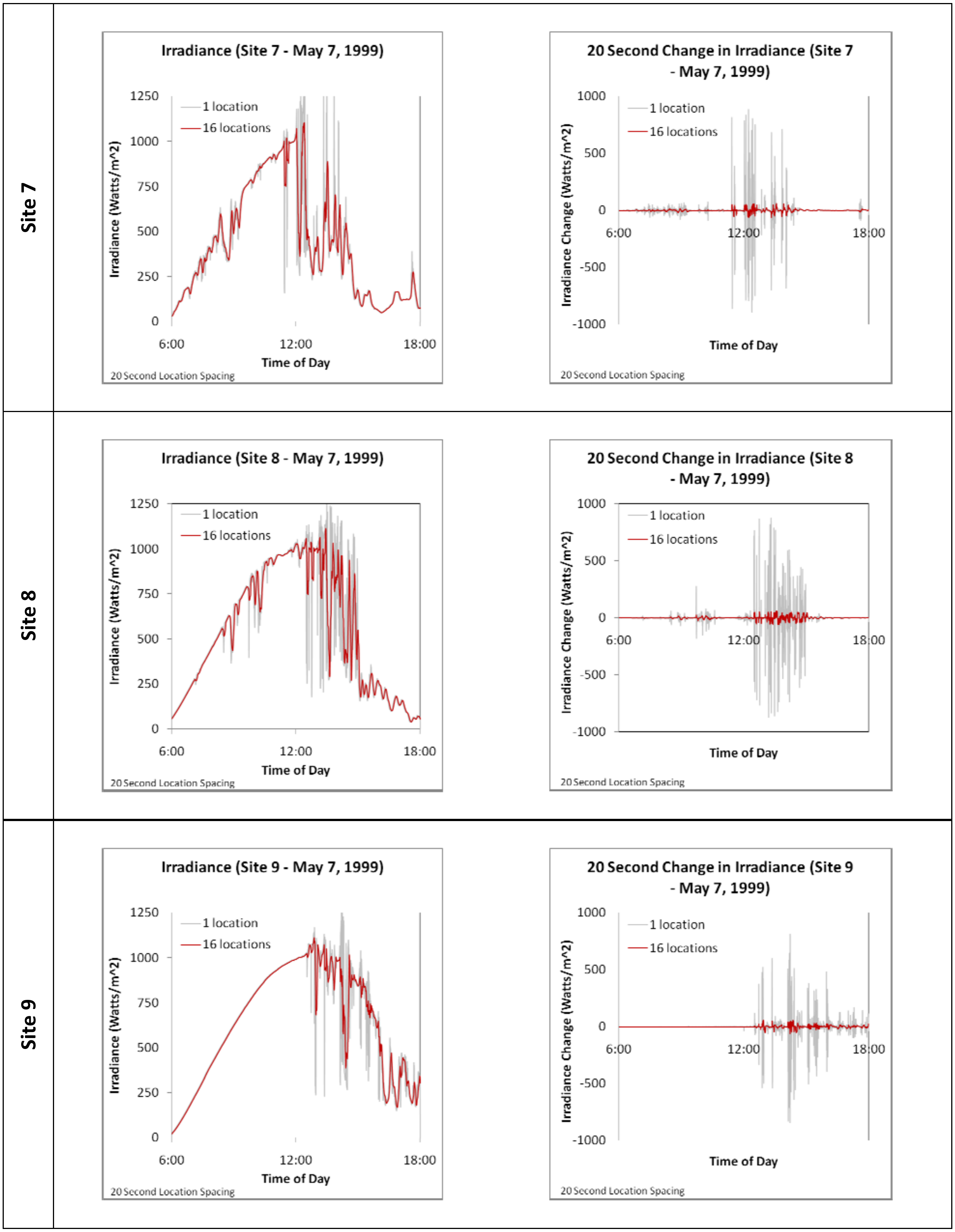


Figure 17. Irradiance and change in irradiance (cont.).

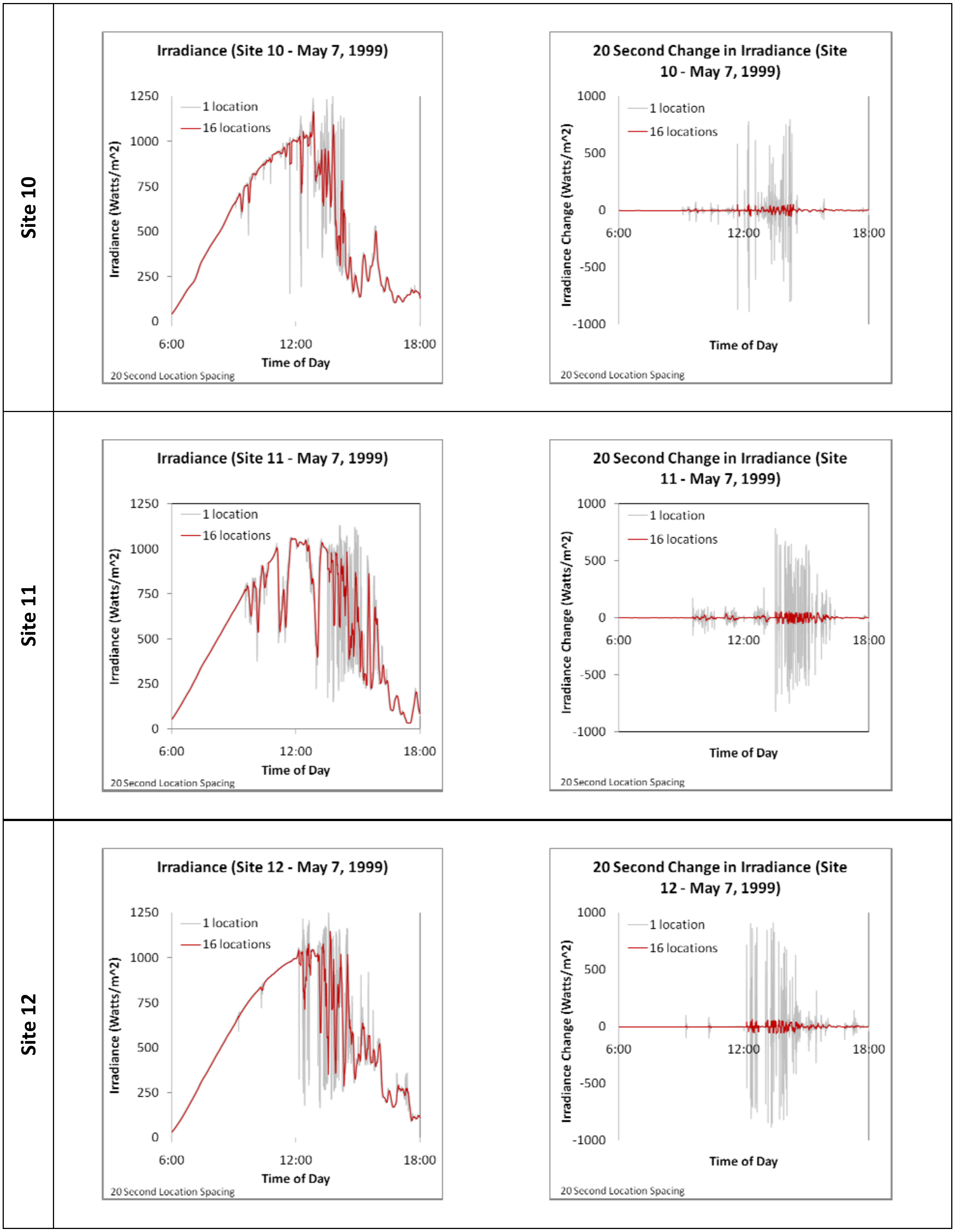


Figure 18. Summary Results.

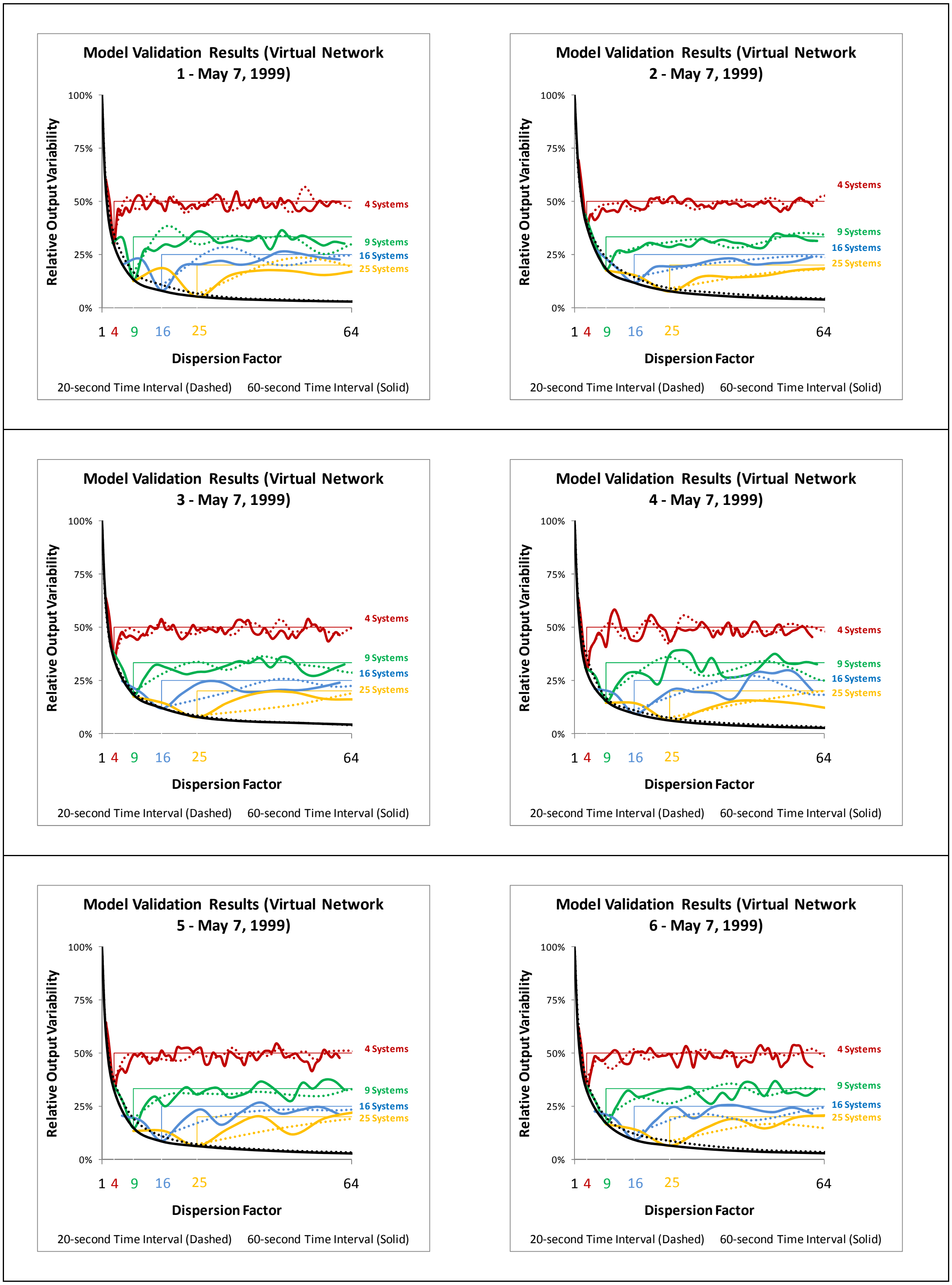
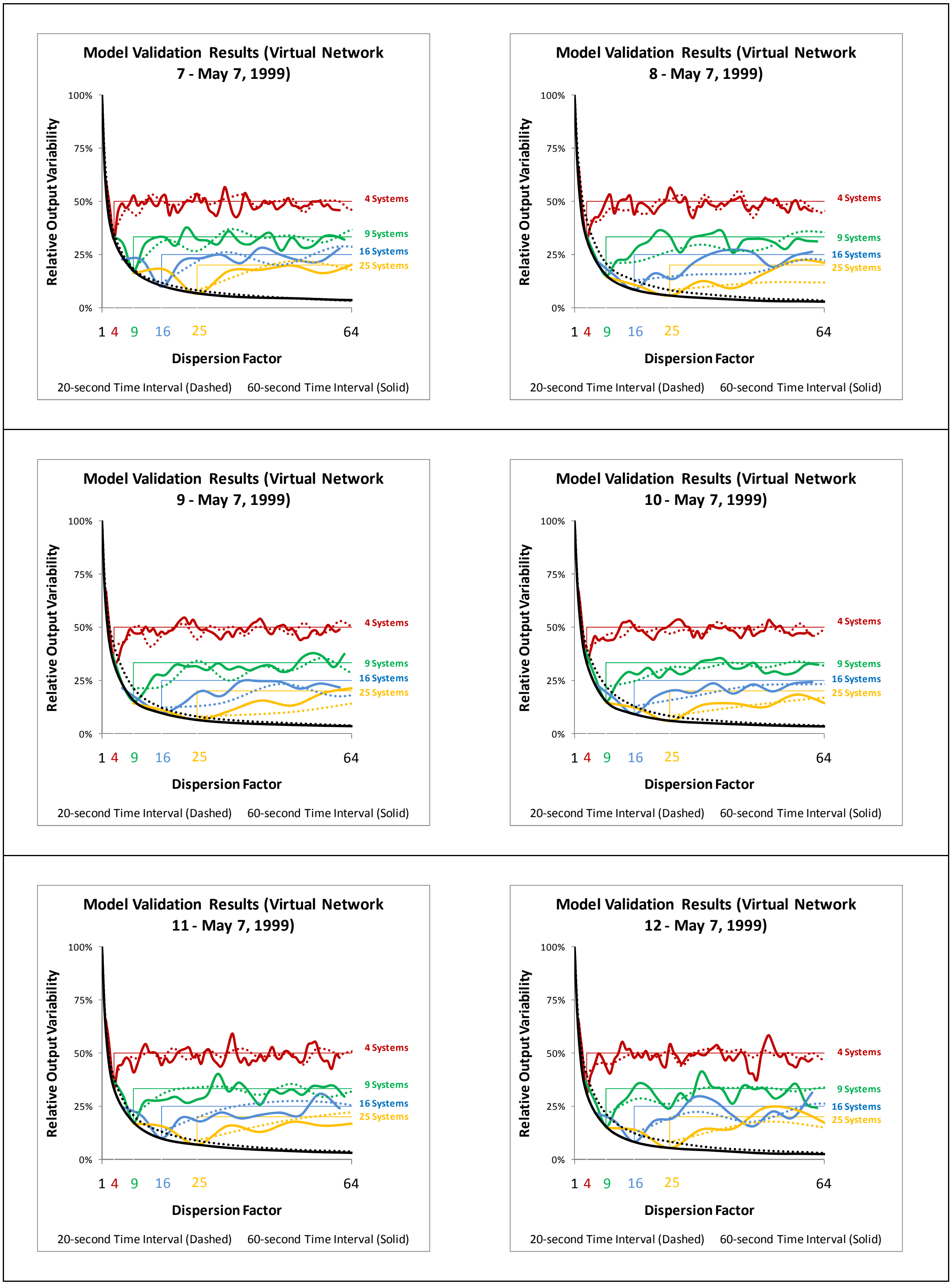


Figure 18. Summary Results (cont.).



General Model

Figure 18 suggests that results are similar for all scenarios for all 12 *Virtual Networks*. In order to investigate this further, the validation was expanded to include multiple scenarios at all 12 *Virtual Networks*.

A key finding of the paper is that *Relative Output Variability* in the *Spacious* region is solely based on the number of systems. That is, results are the same across any network.

As illustrated in Figure 19, *Relative Output Variability* for a distributed fleet of PV systems (i.e., distant enough to be in the *Spacious* region) equals output variability for capacity concentrated in a single location divided by the square root of the number of systems.

Figure 19. *Variability* for *Spacious* region.

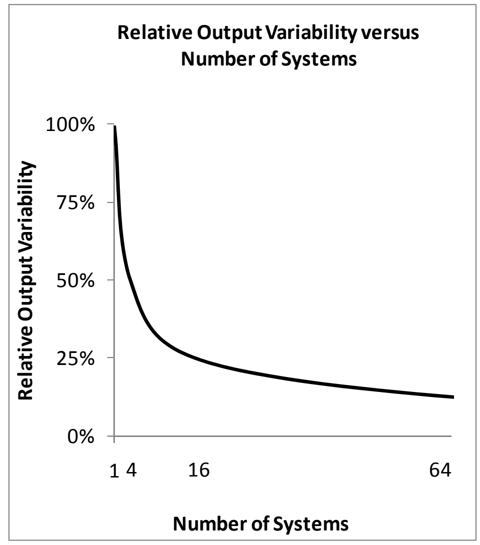


Figure 20. *Variability* for *Crowded* region (includes results for 12 *Virtual Networks*).

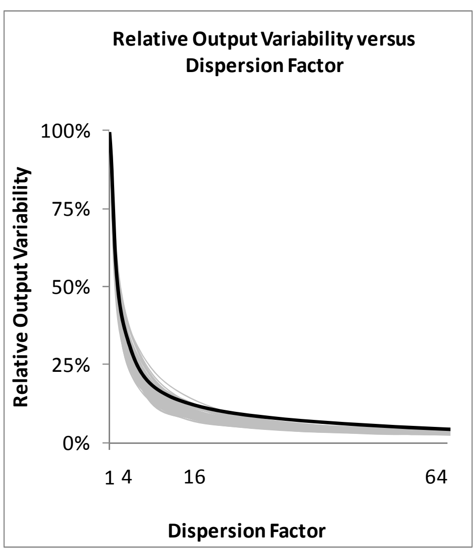


Figure 20 presents an analysis of the results of the *Crowded* region for all 12 *Virtual Networks* across a range of *Dispersion Factors* using *Time Intervals* of 20, 40, 60, 80, 100, and 120 seconds. The light gray lines correspond to all scenario results for all 12 *Virtual Networks*.

The figure suggests 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.

Validation: Nine-Location Actual Network

As mentioned above, Otani et. al. (1997) performed an analysis of one-minute irradiance measurements obtained from a nine-site network concentrated in a 12 square km area. 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. This current paper calculates the standard deviation (this is similar to the root mean square calculation) of the difference between two *instantaneous* irradiance measurements. Thus, while the results in the Otani et. al. (1997) tend to understate the minute-to-minute level of because some of the variability is eliminated when the average is taken, the study offers a valuable opportunity to partially validate results from this current paper.

The current paper predicts that output variability for a system with capacity spread out across 9 locations relative to a single site variability should be 33 percent ($1/\sqrt{9}$). Otani et. al. (1997) report in the conclusions section that the nine-site irradiance variability decreased to around 20 to 50 percent relative to each representative site during cloudy conditions. While not an exact comparison due to methodological differences, it does provide additional validation of the approach documented herein.

Application

The results of the previous two sections are promising: the model 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 Tucson Electric's 5 MW_{DC} Springerville plant. The second example is for a hypothetical 100 MW PV fleet implemented as either a central station plant or as distributed generation.

5 MW_{DC} PV System in Springerville, Arizona

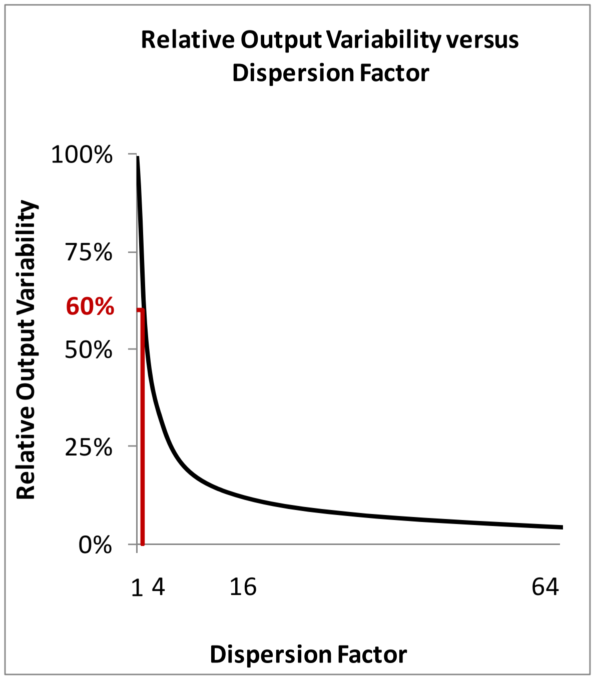
The first example is for the 5 MW_{DC} PV plant located in Springerville, Arizona that is owned by Tucson Electric. Tucson Electric personnel have reported that the utility experienced changes in output up to 50 percent over 60-second *Time Intervals*. This reported finding has become a cause of concern for some in the industry while others have presented information that suggests that the concern is unfounded (Perez et. al., 2009). The results from this paper can be used to shed additional light on the subject.

Analysis of Existing Plant

The Springerville plant covers 44 acres (Moore) or 178,062 square meters. While not perfectly square, assume that the plant is about 420 meters by 420 meters. As stated above, the rate of change of concern to the utility is the 60-second change in output. Consider a case when the *Cloud Transit Rate* equals 3.5 meters per second. A cloud impingent upon the first portion of the system will be completely transitioned off the entire 5 MW plant after two 60-seconds *Time Intervals*.[7] As illustrated in Figure 21 (Figure 20 with an overlay), a *Dispersion Factor* of 2 translates to a standard deviation that is 60 percent of single location.

[7] (2 Time Intervals) x (60 seconds/Time Interval) x (3.5 meters/second) = 420 meters.

Figure 21. 2 Dispersion Factor translates to 60% standard deviation (relative to one location).



Possible Solutions

The low *Dispersion Factor* provides for a good understanding of why there is a high degree of 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 have been distributed across 1,000 five-kW independent plants. In this region, the standard deviation relative to a single location would have been 3 percent ($\frac{1}{\sqrt{1000}} = 3\%$). That is, the distributed generation scenario has 3 percent variability versus the central plant's 60 percent variability.

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. As presented in Table 2, the options under consideration include:

- A 100 MW central station facility
- 100 moderately-sized (1 MW) distributed plants distributed throughout the utility system
- 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 output variability from a utility systems operation perspective before selecting an installation scenario. As in the body of the paper, all results are presented relative to the uncertainty associated with a single location.

Table 2. Installation scenarios.

	System Size	Quantity	Total Capacity
Central Station			
One Location	100 MW	1	100 MW
Distributed Generation			
Large Systems	1 MW	100	100 MW
Small Systems	5 kW	20,000	100 MW

Distributed Scenarios

The distributed scenarios are the most simple to deal with because the assumption can be made that the systems are independent. If it is further assumed that weather conditions are similar throughout the utility system, the *Spacious* model in Equation (19) can be directly applied. The 100 1-MW plants will result in a standard deviation that is 10 percent of that of a single plant ($\frac{1}{\sqrt{100}}$). The 20,000 5-kW plants will result in a standard deviation that is less than 1 percent of that of a single plant ($\frac{1}{\sqrt{20{,}000}}$).

Central Station Scenario

If properly designed, the central station plant can be viewed as being composed of a large number of 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* that determines the variability for this scenario.

To perform the analysis, make the following assumptions:

- The 100 MW will be located in a 1 km x 1 km square (i.e., output density of 100 Watts/m^2)
- One side of the square is perpendicular to the direction of the wind
- Maximum *Cloud Transit Rate* is 10 meters per second
- The utility is concerned with the output variability over a 10-second time period

It will take 10 *Time Intervals* for a cloud to pass across the fleet under these conditions and thus the *Dispersion Factor* is 10.[8] A *Dispersion Factor* of 10 translates to a standard deviation that is approximately 18 percent of single location.

In summary, the uncertainty for the central station plant is based on area and *Cloud Transit Speed*. It has an 18 percent *Relative Output Variability*. The uncertainty for the distributed systems is based on the number of distinct locations: 100 1-MW distributed systems have a 10 percent *Relative Output Variability*, and 20,000 small distributed systems have a 1 percent *Relative Output Variability*.

[8] (10 Time Intervals) x (10 seconds / Time Interval) x (10 meters per second) = 1,000 meters

Summary and Conclusions

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 minimize *Relative Output Variability*.

Model results were compared to measured irradiance data collected at a 20-second *Data Collection Rate* during high variability conditions. The measured output data were translated to *Virtual Networks*. Model results were compared to a variety of number of locations combined with a variety of system spacing increments. The model correlated well with measured results across all configurations for all 12 *Virtual Networks*.

The paper concluded with two examples to illustrate how to apply the results. One example was a cursory analysis of Tucson Electric's 5 MW_{DC} Springerville plant and the second example was for a hypothetical 100 MW PV fleet constructed in either a central station or distributed application. The first example provided an explanation for why the Springerville output variability was high at certain times of the year.

Research Directions

This paper represents a first step toward quantifying PV output variability. As such, there are a number of research directions on which to focus. They can be broadly categorized as:

- Validate results
- Extend model
- Compute absolute single site output variability

Validate Results

The focus of this paper was on model development. The model validation was performed for 12 *Virtual Networks* for a single day. As such, there is much room to validate the results of the model presented here. There are several directions that can be taken to provide a further level of result validation. They include:

- Validate the *Virtual Network* approach by obtaining high-frequency irradiance data from a high-density, large-area, gridded-network of pyranometers
- Obtain measured irradiance data that has a *Data Collection Rate* that is faster than 20 seconds
- Validate results using measured PV output data from distributed PV plants
- Evaluate results on an hourly basis rather than a daily basis; in particular, an assessment needs to be made relative to the form of the *Relative Output Variability* when performed on an hourly basis rather than a daily basis

Extend Model

There are several directions that can be taken to enhance the model developed in this paper.

Generalize Model

This paper developed a model that quantifies relative power output variability for a fleet of identical PV systems (same size, orientation, and spacing) distributed in the direction of prevailing cloud motion. An important next step is to extend the model to calculate results for an arbitrary PV fleet configuration.

Evaluate Results Using Clear-sky Irradiance and Clearness Index

This paper assumed that the underlying force driving variability is the change in irradiance. It is likely that results could be improved by modifying the model to reflect the fact that irradiance can be expressed as the product of clear-sky irradiance and a clearness index.

Evaluate Independence as a Function of Distance

Results suggested that the variation can be accurately quantified when plants are *Crowded* or *Spacious* but only an upper bound can be given when the plants are in the *Limited* region. Two areas of research are needed here. First, refine the model to provide a transition from the model for adjacent locations to the model for independent systems. Second, evaluate how far apart systems must be in order for them to be independent.

Relax Irradiance Translation Assumption

This analysis assumed that irradiance measured at one location will be the same as the irradiance measured at a time Δt later in the next location. In reality, there may be some other changes that occur to the irradiance of the time Δt. This assumption should be further evaluated and any changes implemented in the model.

Calculate Single Station Noise

This paper focused on the calculation of relative uncertainty. That is, it presented a method to calculate the standard deviation in the output variability for N systems relative to the standard deviation of a single system. A method needs to be implemented to calculate the absolute standard deviation for a single location in order to calculate the absolute standard deviation for N locations. The method would be parameterized as a function of weather conditions noting that this would be minimal during clear sky conditions.

References

Gansler R. A., S. A. Klein and W.A. Beckman (1995): Investigation of minute solar radiation data. Solar Energy, Volume 55, No. 1, pp. 21–27.

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

Jurado M., J. M. Caridad, and V. Ruiz (1995): Statistical distribution of the clearness index with radiation data integrated over five minute intervals. Solar Energy, Volume 55, No. 1, pp. 469–473.

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

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

Moore, L., H. Post, T. Hansen, and T. Mysak: Photovoltaic Power Plant Experience at Tucson Electric Power. Progress in Photovoltaics: Research and Applications, Volume 13, Issue 4, pp. 353 – 363.

Murata, A., H. Yamaguchi, and K. Otani (2009): A Method of Estimating the Output Fluctuation of Many Photovoltaic Power Generation Systems Dispersed in a Wide Area, Electrical Engineering in Japan, Volume 166, No. 4, pp. 9-19.

Myers, D., S. Wilcox, W. Marion, R. George, and M. Anderberg (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.

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

Perez, R., M. Taylor, and T. E. Hoff, J.P. Ross (2009): Redefining PV Capacity. Public Utilities Fortnightly, Volume 147, No. 2.

Ross, S. (1988): A First Course in Probability. Macmillan Publishing Company.

Skartveit A., and J. A. Olseth, (1992): The Probability Density of Autocorrelation of Short-term Global and beam irradiance. Solar Energy, Volume 46, No. 9, pp. 477-488.

Suehrcke H., and P. G. McCormick, (1989): Solar Radiation Utilizability. Solar Energy, Volume 43, No. 6, pp. 339-345.

Solar Anywhere® (2009): Web-based service that provides hourly, satellite-derived solar irradiance data as current as yesterday and archival data back to January 1, 2003. www.SolarAnywhere.com.

Stokes, G. M. and S. E. Schwartz, (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

Tovar J., F. J. Olmo, F. J. Batlles, and L. Alados-Arboledas, (2001): Dependence of one-minute global irradiance probability density distributions on hourly irradiation. Energy, an International Journal, Volume 26, Issue 7, pp. 659-668.

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

Woyte A, R. Belmans, and J. Nijsb (2007): Fluctuations in instantaneous clearness index: Analysis and statistics. Solar Energy, Volume 81, No. 2, pp. 195-206.

Appendix: Optimal Model Region Solution

This appendix solves the standard deviation equation for N systems that are spaced at exactly one *Time Interval* (Δt) seconds between each location.

Based on the standard deviation in Equation (14), the variance equals

$$\left(\sigma_{\Delta t}^{\Sigma N}\right)^2 = \left(\frac{1}{N}\right)^2 Var\left[\sum_{n=1}^{N} \Delta \boldsymbol{I}_{\Delta t}^{\boldsymbol{n}}\right] \qquad (20)$$

The variance term can be expanded and expressed as follows.

$$\left(\sigma_{\Delta t}^{\Sigma N}\right)^2 = \left(\frac{1}{N}\right)^2 \left(\frac{1}{T}\right) \sum_{t=1}^{T} \left[\left(\sum_{n=1}^{N} I_{\boldsymbol{t}}^{\boldsymbol{n}} - \sum_{n=1}^{N} I_{\boldsymbol{t}+\Delta \boldsymbol{t}}^{\boldsymbol{n}}\right)\right]^2 \qquad (21)$$

In the second summation, the summation can be reduced by 1 and the notation of the number of locations increased by 1 and it can be rewritten as

$$\left(\sigma_{\Delta t}^{\Sigma N}\right)^2 = \left(\frac{1}{N}\right)^2 \left(\frac{1}{T}\right) \sum_{t=1}^{T} \left[\left(\sum_{n=1}^{N} I_{\boldsymbol{t}}^{\boldsymbol{n}} - \sum_{n=0}^{N-1} I_{\boldsymbol{t}+\Delta \boldsymbol{t}}^{\boldsymbol{n}+\boldsymbol{1}}\right)\right]^2 \qquad (22)$$

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

$$\left(\sigma_{\Delta t}^{\Sigma N}\right)^2 = \left(\frac{1}{N}\right)^2 \left(\frac{1}{T}\right) \sum_{t=1}^{T} \left[\left(I_{\boldsymbol{t}}^{\boldsymbol{N}} + \sum_{n=1}^{N-1} I_{\boldsymbol{t}}^{\boldsymbol{n}} - \sum_{n=1}^{N-1} I_{\boldsymbol{t}+\Delta \boldsymbol{t}}^{\boldsymbol{n}+\boldsymbol{1}} - I_{\boldsymbol{t}+\Delta \boldsymbol{t}}^{\boldsymbol{1}}\right)\right]^2 \qquad (23)$$

Locations are spaced at a distance of Δt from the next one where Δt is the amount of time it takes the irradiance in one location to move to the next location. Assume that the irradiance is translated unchanged to the next location after a time of Δt. When n is greater than or equal to 1 and less than or equal to N-1, this means that

$$I_{\boldsymbol{t}+\Delta \boldsymbol{t}}^{\boldsymbol{n}+\boldsymbol{1}} = I_{\boldsymbol{t}}^{\boldsymbol{n}} \qquad (24)$$

Equation (22) simplifies to

$$\left(\sigma_{\Delta t}^{\sum N}\right)^2 = \left(\frac{1}{N}\right)^2 \left(\frac{1}{T}\right) \sum_{t=1}^{T} \left[\left(I_{\boldsymbol{t}}^{\boldsymbol{N}} - I_{\boldsymbol{t+\Delta t}}^{\boldsymbol{1}}\right)\right]^2 \tag{25}$$

It is also known that $I_{\boldsymbol{t}}^{\boldsymbol{N}} = I_{\boldsymbol{t+(1-N)\Delta t}}^{\boldsymbol{1}}$ is based on the irradiance movement. Making this substitution and adjusting the summation results in

$$\left(\sigma_{\Delta t}^{\sum N}\right)^2 = \left(\frac{1}{N}\right)^2 \left(\frac{1}{T}\right) \sum_{t=1+(1-N)\Delta t}^{T+(1-N)\Delta t} \left[\left(I_{\boldsymbol{t}}^{\boldsymbol{1}} - I_{\boldsymbol{t+N\Delta t}}^{\boldsymbol{1}}\right)\right]^2 \tag{26}$$

This equals

$$\left(\sigma_{\Delta t}^{\sum N}\right)^2 = \left(\frac{1}{N}\right)^2 Var\left[\Delta \boldsymbol{I}_{\boldsymbol{N\Delta t}}^{\boldsymbol{1}}\right] \tag{27}$$

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

The final result is that the standard deviation equals

$$\sigma_{\Delta t}^{\sum N} = \left(\frac{1}{N}\right) \sigma_{N\Delta t}^{1} \tag{28}$$