SANDIA REPORT

SAND2015-6700
Unlimited Release
Printed August 2015

Photovoltaic System Modeling: Uncertainty and Sensitivity Analyses

Clifford W. Hansen and Curtis M. Martin

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

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

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online order:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2015-6700
Unlimited Release
Printed August 2015

Photovoltaic System Modeling: Uncertainty and Sensitivity Analysis

Clifford Hansen and Curtis Martin
Photovoltaic and Distributed Systems Department, Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-1033

Abstract

We report an uncertainty and sensitivity analysis for modeling AC energy from photovoltaic systems. Output from a PV system is predicted by a sequence of models. We quantify uncertainty in the output of each model using empirical distributions of each model's residuals. We propagate uncertainty through the sequence of models by sampling these distributions to obtain an empirical distribution of a PV system's output. We consider models that: (1) translate measured global horizontal, direct and global diffuse irradiance to plane-of-array irradiance; (2) estimate effective irradiance; (3) predict cell temperature; (4) estimate DC voltage, current and power; (5) reduce DC power for losses due to inefficient maximum power point tracking or mismatch among modules; and (6) convert DC to AC power. Our analysis considers a notional PV system comprising an array of FirstSolar FS-387 modules and a 250 kW AC inverter; we use measured irradiance and weather at Albuquerque, NM.

We found the uncertainty in PV system output to be relatively small, on the order of 1% for daily energy. We found that uncertainty in the models for POA irradiance and effective irradiance to be the dominant contributors to uncertainty in predicted daily energy. Our analysis indicates that efforts to reduce the uncertainty in PV system output predictions may yield the greatest improvements by focusing on the POA and effective irradiance models.

CONTENTS

FIGURES

TABLES

NOMENCLATURE

CDF	cumulative distribution function
DHI	diffuse horizontal irradiance
DOE	Department of Energy
DNI	direct normal irradiance
GHI	global horizontal irradiance
GUM	Guide to the Expression of Uncertainty in Measurement
POA	plane-of-array
PSEL	Photovoltaic System Evaluation Laboratory
PV	photovoltaic
SNL	Sandia National Laboratories

1. INTRODUCTION

As the photovoltaic (PV) industry continues to mature and incentives are reduced, investment in PV increasingly depends on the confidence that can be placed in predictions of the energy yield. Predicting energy yield requires use of a sequence of models, e.g., to translate measured irradiance to the system's plane-of-array, to estimate cell temperature, and to predict DC power for given conditions. Uncertainty in these models and their inputs arises from a variety of sources, including measurement errors, inexact model specification, and from the necessarily finite data used to calibrate models. In aggregate, these uncertainties contribute to uncertainty in predicted energy yield. Therefore, to understand what confidence can be placed in energy yield predictions, to identify how to improve model accuracy and to reduce prediction uncertainty, we must quantify the uncertainty introduced by each model and the effect of each model's uncertainty on energy yield predictions.

This report builds upon earlier work [1] which developed a methodology to quantify uncertainty in PV system output predictions and applied the methodology to PV systems comprising a single module. Here, we extend the analysis in [1] by considering additional steps in the modeling process: specifically, we improve the uncertainty characterization for the models for effective irradiance, and we address uncertainty in models involved in the translation of DC output from a module to AC power. As noted in Section 5 our findings in this analysis are consistent with those reported in [1]. Here we also present the methods for uncertainty quantification and consequently this report may be viewed as superseding the earlier work.

Other analyses of uncertainty in module or system performance have centered on the effects of measurement uncertainties, or on calibration of models to data assumed to be exact. For example, a detailed investigation of module performance uncertainties under natural sunlight including correction for irradiance and temperature was given by Whitfield et al. [2]. The methodology was based on the analytical propagation of respective uncertainties in two dimensions using the Guide to the Expression of Uncertainty in Measurement (GUM) [3]. This methodology was used and expanded for long-term outdoor IV measurements for data of Northern latitude [4]. Dirnberger and Kraling [5] provide a detailed analysis of uncertainty deriving from indoor measurements to determine module rating at standard test conditions (STC). Müller et al. [6] compared measured and predicted performance of operating PV power plants over several years to quantify the uncertainty in predicted annual yield; they identified the solar resource and power reduction due to module degradation and/or soiling as the primary causes of differences between predicted and measured output. Hansen et al. [7] examined the influence of uncertainty in calibrated parameters on performance predictions, where the data used for calibration was assumed to be error-free.

Our report is organized as follows:

- Section 2 describes the methodology we employ to quantify uncertainty in each modeling step, and identifies the models we considered;
- Section 3 describes and illustrates the uncertainty quantification for each modeling step;
- Section 4 presents the results of the uncertainty analysis, i.e., the distribution of results predicted by propagating uncertainty through each modeling step;

- Section 5 summarizes the results of the sensitivity analysis, relating the uncertainty associated to each modeling step to the uncertainty observed in the predicted system output.

2. METHODOLOGY

Here we describe the approach taken to quantify uncertainty in PV system modeling. We outline the conceptual approach to our uncertainty analysis in Section 2.1. In Section 2.2, we describe the process of PV system modeling and identify the scenarios selected for analysis. We then outline the methods used to quantify uncertainty at each modeling step in Section 2.3.

2.1. Conceptual Approach to Uncertainty Analysis

Uncertainty analysis is a systematic process to propagate uncertainty in a model or its inputs to uncertainty in the model's output. An uncertainty analysis involves two primary steps: quantification and propagation. First, we quantify uncertainty in a model and in the model's inputs. Here we use probability distributions to quantify uncertainty, noting that other expressions of uncertainty are available [8]. Second, we propagate uncertainty to a model's output through a set of model calculations using a Monte Carlo technique to sample distributions for uncertainty.

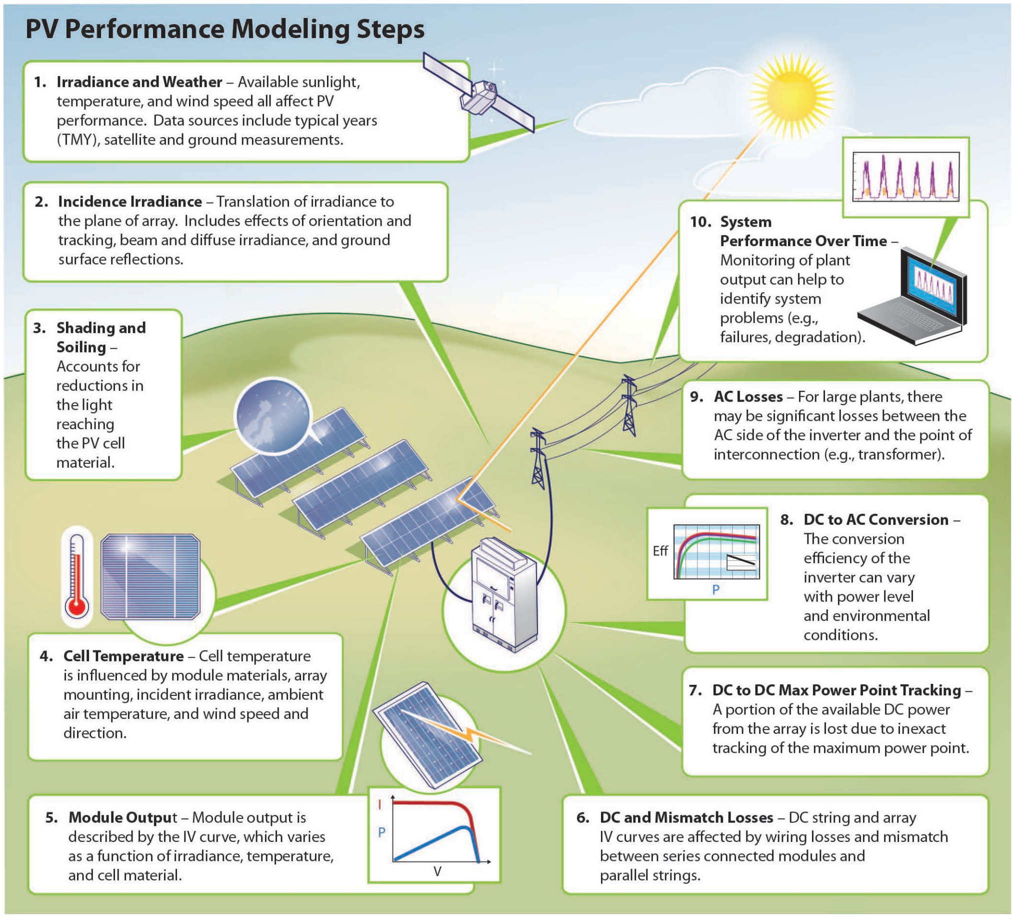
In concept, uncertainty can be categorized as either *parameter* uncertainty or *model uncertainty*. Parameter uncertainty refers to uncertainty in a particular model input, whereas model uncertainty refers to lack of knowledge regarding the model itself. In practice, these two categories tend to overlap, for example, a model for extraterrestrial irradiance may consist of a single, constant (but uncertain) value, or may comprise an equation involving several parameters that accounts for observed systematic variation in extraterrestrial irradiance over time.

In our analysis, we explicitly address model uncertainty by considering several credible alternative models, when such models are available. However, we do not represent parameter uncertainty in the traditional manner, which would be to specify a distribution of possible values for each individual parameter. Instead, we adopt an approach where we characterize the uncertainty in a model's output by quantifying the distribution of each model's *residual*, i.e., the difference between the model's prediction and the true value. This method of representing uncertainty in a model's residuals effectively aggregates the uncertainty resulting from all of the model's parameters into a single quantity.

We adopt this approach because nearly all models involved in estimating PV system output are *calibrated*, i.e., their parameter values are determined by fitting the model's equations to data that are considered representative. When a parameter is determined by fitting an equation to a set of data, uncertainty in the parameter arises from a number of sources, including uncertainty in the data used for the fitting, model uncertainty in the equation used, and numerical error arising from the finite sample of data and the fitting procedure. Moreover, when the parameter is jointly determined with other parameters, uncertainty in each parameter is likely correlated with uncertainty in all other parameters determined in the same fitting procedure. These factors complicate the effort to separately describe uncertainty in each fitted parameter.

2.2. PV System Modeling

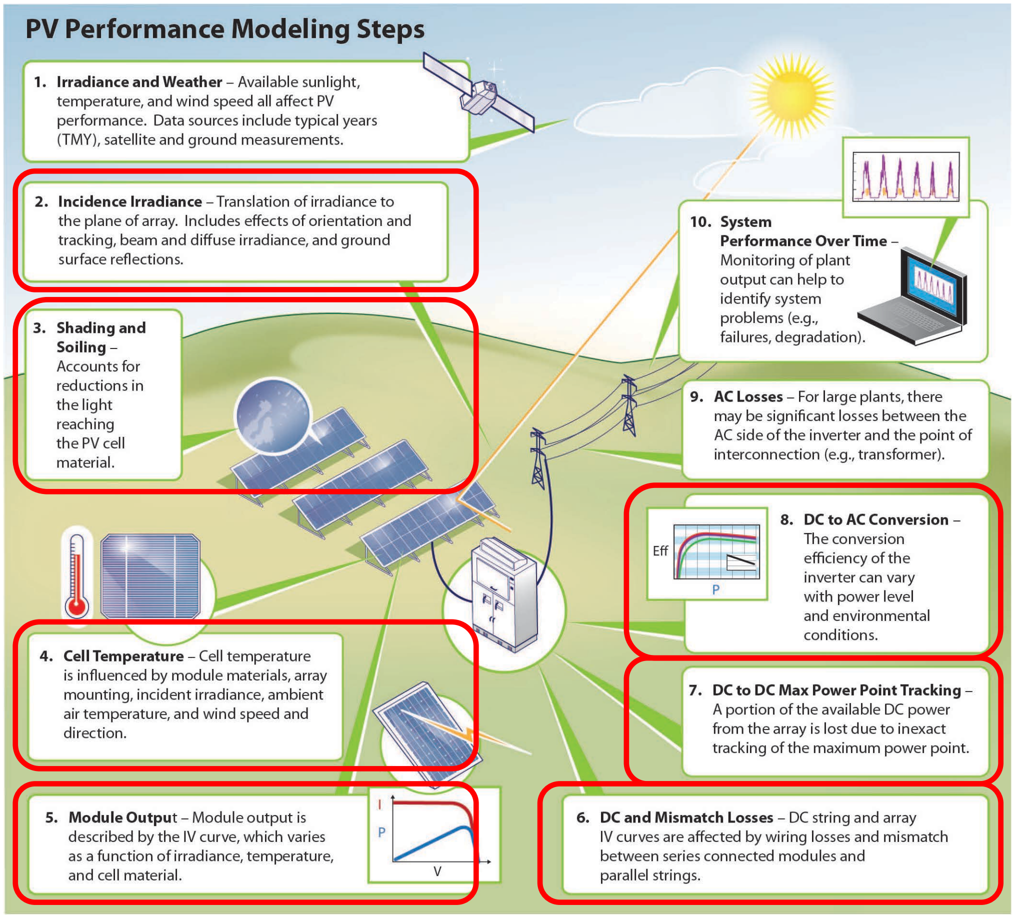
The process of modeling DC power output from a PV system involves nine notional steps, as illustrated in Figure 1. Uncertainty in the outcome of each step may arise from uncertainty in the models employed or from the parameters required by those models. For example, Step 2, “Incident Irradiance,” estimates plane-of-array irradiance (POA) from typical irradiance measurements (global horizontal irradiance (GHI), direct normal irradiance (DNI), and/or diffuse horizontal irradiance (DHI)). This translation involves a choice from among a number of models for sky diffuse irradiance, e.g., the isotropic sky diffuse model [9] or Hay and Davies’ diffuse model [10].


PV Performance Modeling Steps

1. **Irradiance and Weather** – Available sunlight, temperature, and wind speed all affect PV performance. Data sources include typical years (TMY), satellite and ground measurements.
2. **Incidence Irradiance** – Translation of irradiance to the plane of array. Includes effects of orientation and tracking, beam and diffuse irradiance, and ground surface reflections.
3. **Shading and Soiling** – Accounts for reductions in the light reaching the PV cell material.
4. **Cell Temperature** – Cell temperature is influenced by module materials, array mounting, incident irradiance, ambient air temperature, and wind speed and direction.
5. **Module Output** – Module output is described by the IV curve, which varies as a function of irradiance, temperature, and cell material.
6. **DC and Mismatch Losses** – DC string and array IV curves are affected by wiring losses and mismatch between series connected modules and parallel strings.
7. **DC to DC Max Power Point Tracking** – A portion of the available DC power from the array is lost due to inexact tracking of the maximum power point.
8. **DC to AC Conversion** – The conversion efficiency of the inverter can vary with power level and environmental conditions.
9. **AC Losses** – For large plants, there may be significant losses between the AC side of the inverter and the point of interconnection (e.g., transformer).
10. **System Performance Over Time** – Monitoring of plant output can help to identify system problems (e.g., failures, degradation).

Figure 1. PV System Modeling Process

Figure 2 indicates the sequence of models considered in the quantification of uncertainty for PV system modeling in this analysis. Most, but not all modeling steps are addressed in the uncertainty quantification illustrated by Figure 2. We intentionally did not consider uncertainty represented by the models and measurements in Step 1, because that uncertainty is dominated by uncertainty in measured irradiance which will directly (and proportionally) affect predicted power. There are several ongoing research programs to better understand and improve measurement uncertainty. Our focus is to understand the relative contribution of all other modeling steps.



PV Performance Modeling Steps

1. **Irradiance and Weather** – Available sunlight, temperature, and wind speed all affect PV performance. Data sources include typical years (TMY), satellite and ground measurements.
2. **Incidence Irradiance** – Translation of irradiance to the plane of array. Includes effects of orientation and tracking, beam and diffuse irradiance, and ground surface reflections.
3. **Shading and Soiling** – Accounts for reductions in the light reaching the PV cell material.
4. **Cell Temperature** – Cell temperature is influenced by module materials, array mounting, incident irradiance, ambient air temperature, and wind speed and direction.
5. **Module Output** – Module output is described by the IV curve, which varies as a function of irradiance, temperature, and cell material.
6. **DC and Mismatch Losses** – DC string and array IV curves are affected by wiring losses and mismatch between series connected modules and parallel strings.
7. **DC to DC Max Power Point Tracking** – A portion of the available DC power from the array is lost due to inexact tracking of the maximum power point.
8. **DC to AC Conversion** – The conversion efficiency of the inverter can vary with power level and environmental conditions.
9. **AC Losses** – For large plants, there may be significant losses between the AC side of the inverter and the point of interconnection (e.g., transformer).
10. **System Performance Over Time** – Monitoring of plant output can help to identify system problems (e.g., failures, degradation).

Figure 2. Sequence of Models Considered in Uncertainty Quantification.

To quantify uncertainty in the outcome of any particular modeling step, we needed concurrent measurements of model inputs and outputs with sufficient resolution and/or quality to have confidence that the data would fairly represent the uncertainty present. Except for soiling rates (part of Step 3), we found data for each of the steps indicated by Figure 2.

Following the approach of the previous study [1], we characterize the uncertainty of each modeling step using measurements of inputs and outputs of that step. We then use measured irradiance (GHI, DNI, and DHI) and weather (ambient temperature and wind speed) data and the uncertainty characterizations to propagate uncertainty through the sequence of models to get an estimate of uncertainty in the overall modeling chain. Finally, we perform sensitivity analyses to identify which modeling steps contribute most to the overall uncertainty. In contrast to the previous study, which considered multiple locations and solar module technologies, here we consider only a single scenario: a notional fixed-tilt PV system comprising an array of 2,493 First Solar 387 CdTe modules connected in 277 parallel strings (9 modules per string) to a 250 kW SMA SC250U inverter.

Except the POA irradiance models, all models for Steps 2 through 5 illustrated in Figure 2 require component-specific coefficients, many of which we determine from measurements made in Albuquerque, NM. In the uncertainty propagation, we use solar and weather measurements collected during 2013 and 2014, also at Albuquerque, NM. Because our analysis focuses on the relative influence of various uncertainties on the overall uncertainty in the modeling chain, rather than on the absolute values of the modeling outcomes, we do not believe that our use of site-specific data greatly affects our study's conclusions.

2.3. Experimental Set-up

Multiple experiments were performed in order to collect the data necessary for this study. Where possible, the experiments were conducted concurrently.

In order to characterize the uncertainty in the plane-of-array irradiance (G_{POA}, or POA) model, solar irradiance measurements were collected at Sandia National Laboratories' Photovoltaic Systems Evaluation Laboratory (PSEL) in Albuquerque, NM between October 2013 and September 2014. GHI was measured with a Kipp and Zonen CM11 pyranometer, DNI was measured with a CHP-1 pyrheliometer, and a shaded Epply PSP pyranometer was used to measure DHI. POA for a southward-facing surface tilted 35.05° from horizontal was measured using an additional Epply PSP pyranometer installed at the same orientation. Data from these instruments were captured at 3-second time intervals and time-stamped with a GPS time reference. The 3-second data were subsequently time-averaged to produce one-minute samples with sample times matching those of the PV data described below. Together with ambient air temperature and wind speed collected concurrently at the PSEL, these measurements of GHI, DNI, DHI, and POA also provided the data inputs for the uncertainty propagation phase of the study.

To characterize the uncertainty in the model for effective irradiance (the amount of solar energy available to be converted by the solar module), a First Solar FS-390 module was installed at the PSEL facing south with a tilt of 35.05° (the latitude of the PSEL). Automatic data loggers captured short-circuit current from the First Solar module, along with back-surface temperature measured with thermocouples at approximately 30 second intervals, time-stamped with a GPS time reference. The 30-second data were time-averaged to produce one-minute samples.

We attempted to characterize uncertainty in soiling at the PSEL by simultaneously measuring output from two FS-390 modules, one of which was cleaned daily and the other left to accumulate soiling naturally. However, the soiling rates at the PSEL are low enough that any systematic effects from soiling were obscured by daily variation in effective irradiance and by noise inherent in measuring module short-circuit current. Consequently, we only use data from the clean FS-390 module and examine the effect of soiling rates parametrically.

Module-specific parameters for FS-390 modules were obtained at the PSEL using a separate module mounted on a two-axis tracker. These parameters include reflection and spectral mismatch losses expressed in the effective irradiance model, as well as temperature coefficients and electrical performance parameters used in Steps 4 and 5 of Figure 2. All these parameters were estimated from measured I-V curves recorded from the module under a wide range of illumination and environmental conditions achieved by positioning the module with the tracker. The characterization tests and parameter estimation methods are summarized in [7].

To estimate power losses due to module mismatch and MPPT tracking inaccuracies (Steps 6 and 7 of Figure 2), we monitored operating voltage and DC power for an array of 48 First Solar FS-387 modules connected to a Fronius IG-TS 5.0 inverter in four parallel strings of 12 modules each. We placed an additional single FS-387 module adjacent to the array so that it would be under the same illumination conditions as the array, and recorded I-V curves of this reference module at regular intervals. As described in Step 5 of the following section and Section 3.6, these data provide the basis for a new model for the aggregate power loss due to module mismatch and MPPT tracking inefficiencies.

2.4. Methods for Uncertainty Quantification

Let $\hat{f}\left(x|p\right)$ represent a model $\hat{f}$ applied to inputs x with a fixed set of parameter values p. Denote the true value at x by $f\left(x\right)$; then the residual is given by

$$\varepsilon_f\left(x|p\right)=\hat{f}\left(x|p\right)-f\left(x\right) \tag{1}$$

We regard $\varepsilon_f\left(x|p\right)$ as a random variable and develop distributions for $\varepsilon_f\left(x|p\right)$ for each selected model $\hat{f}$ using the representative data.

The distribution for $\varepsilon_f\left(x|p\right)$ characterizes the aggregate uncertainty in the model $\hat{f}$ and the inputs x conditional on the parameters p. Different distributions for $\varepsilon_f\left(x|p\right)$ can result when the parameters p are varied. Because parameters are generally obtained by calibration of a model to data, for a given set of data there are ‘best’ values for these parameters, but other values may arise from different data sets. Thus parameter variation can arise from alternate data sets, which in turn results if the analysis is done using measurements from a different location or time of year.

Parameter values used here are those regarded as default values for each model. We did not attempt to quantify uncertainty in the parameters by finding alternate data sources and recalibrating models to obtain alternate parameter values.

As indicated in Figure 2, calculation of AC power from a PV-based system involves a sequence of models. At each step in the process, uncertainty is quantified for the models used in that step. Results (with uncertainty) from each step are then used as input to the next step.

Step 1: Estimation of POA irradiance

Plane-of-array (POA) irradiance G_{POA} is defined as the total broadband irradiance incident on the face of a module. It is typically represented as the sum of beam and diffuse components, E_b and E_{diff}, respectively:

$$G_{POA} = E_b + E_{diff} . \tag{2}$$

In turn, the diffuse component is expressed as the sum of ground-reflected diffuse irradiance $E_{diff,g}$ and diffuse irradiance from the sky, $E_{diff,sky}$:

$$E_{diff} = E_{diff,g} + E_{diff,sky} . \tag{3}$$

For reasons presented later, uncertainty in estimating POA irradiance arises from uncertainties in models for $E_{diff,sky}$. Thus, we estimate POA irradiance using a model for the sky diffuse component $\hat{f}_{sky}$ that operates on GHI, DNI and DHI:

$$\widehat{G}_{POA}(t) = E_b + E_{diff,g} + \hat{f}_{sky}\left(GHI(t), DNI(t), DHI(t) \middle| p_{POA}\right) \tag{4}$$

We consider the following models for $\hat{f}_{sky}$: isotropic sky diffuse [9]; Sandia simple sky diffuse [11], Hay and Davies [9]; and Perez [12]. Expressions for $\hat{f}_{POA}$ can be complex and are found in the listed references, along with each model's parameter values.

For POA irradiance models, the residual $\delta_{POA}\left(t \middle| p_{POA}\right)$ is expressed a fraction of the measured POA irradiance, denoted by $POA(t)$:

$$\delta_{POA}\left(t \middle| p_{POA}\right) = \frac{\widehat{G}_{POA}(t) - G_{POA}(t)}{G_{POA}(t)} \tag{5}$$

We expressed the residual as a fraction because it allowed for a simpler detrending of the residuals as functions of the solar angle of incidence (AOI) (see Section 3.1 for details).

Analysis of POA residuals revealed systematic trends in $\delta_{POA}\left(t \middle| p_{POA}\right)$ that changed with time of day, season, location and POA irradiance model (see Section 3.1). To facilitate random sampling from these results, we fit empirical expressions $y_{POA}\left(t \middle| p_{POA}\right)$ to the residuals to separate trends from random effects:

$$\delta_{POA}\left(t \middle| p_{POA}\right) = y_{POA}\left(t \middle| p_{POA}\right) + \varepsilon_{POA}\left(t \middle| p_{POA}\right) \tag{6}$$

Illustrative results are provided in Section 3.1 for each POA irradiance model.

Step 2: Estimation of effective irradiance

Effective irradiance E_e represents the irradiance converted to electrical current within the module. E_e differs from G_{POA} due to several mechanisms: optical losses at the module's face (primarily specular reflection); mismatch between the solar spectrum and the module's quantum efficiency; losses due to shading and soiling; and parasitic losses due to electrical resistance internal to the module. Models are available to separately quantify reflection losses and spectral mismatch. Parasitic losses are normally implicitly included in the module performance model (see Step 4). Here, we explicitly consider uncertainty in spectral mismatch. Although uncertainty exists regarding models for reflection losses we judge that uncertainty in reflection loss models is small compared to uncertainty in spectral mismatch models. We attempted without success to locate and analyze data describing soiling losses. We do not consider shading losses in this analysis because such losses are highly dependent on local features and on a PV system's configuration. Thus, all uncertainty represented for the effective irradiance model is attributed to the spectral mismatch terms.

The effective irradiance is modeled using an empirical expression common to many module performance models (e.g., [13], [14]):

$$\widehat{E_e}\left(t \middle| p_E, q_E\right) = \hat{f}_1\left(AM\left(t\right) \middle| p_E\right)\left[\hat{f}_2\left(AOI\left(t\right) \middle| q_E\right)E_b(t) + E_{diff}(t)\right]SF \tag{7}$$

where E_{diff} represents the diffuse irradiance incident on the solar module, including circumsolar, sky and ground-reflected diffuse irradiance. The functions $\hat{f}_1$ and $\hat{f}_2$ model spectral mismatch and reflection losses, respectively. The scalar multiplier SF represents the losses due to soiling on the face of the solar module, with a value of 1.0 corresponding to a clean module.

Consistent with the approach of [13] and [14], the spectral mismatch function, $\hat{f}_1\left(AM\left(t\right) \middle| p_E\right)$, is modeled here as a fourth-order polynomial:

$$\hat{f}_1(AM(t) \mid p_E) = p_{E,4}AM^4(t) + p_{E,3}AM^3(t) + p_{E,2}AM^2(t) + p_{E,1}AM(t) + p_{E,0}\,, \tag{8}$$

where AM is the absolute air mass calculated from solar position and site altitude using tools coded in the PV_Lib toolbox [11]. Solar position (i.e., zenith Z and azimuth Az, both measured in degrees) is computed using a legacy algorithm originating with G. Hughes [15]. Relative air mass AMr is calculated from the zenith angle using the empirical model described in [16]:

$$AMr = \frac{1}{\cos\left(Z\right) + 0.50572 \times \left(6.07995 + \left(90 - Z\right)\right)^{-1.6364}} \tag{9}$$

Finally, using site pressure P (Pa) estimated from site elevation H (m) by [17]

$$P = 100 \times \left(\frac{44331.514 - H}{11880.516} \right)^{1/0.1902632}, \tag{10}$$

absolute air mass AM is obtained as

$$AM = AMr \times \frac{P}{101325}. \tag{11}$$

With absolute air mass determined for a set of measurement conditions, the parameter vector p_E for a given module is determined by fitting the expression in Eq. (8) to the measured short-circuit current of the module [7]. No uncertainty is ascribed to the models represented by the calculations in Eqs. (9) through (11).

Reflection losses are quantified by calculating the complement, i.e., the fraction f_2 of direct irradiance that is captured by the module. The factor f_2 is modeled as a polynomial function of the angle of incidence AOI:

$$\hat{f}_2(AOI(t) \mid q_E) = q_{E,5}AOI^5(t) + q_{E,4}AOI^4(t) + q_{E,3}AOI^3(t) + q_{E,2}AOI^2(t) + q_{E,1}AOI(t) + q_{E,0}. \tag{12}$$

Given solar position (azimuth and zenith as defined above) and module orientation (azimuth Az_A and tilt θ_T), AOI is determined by geometry:

$$AOI = \cos^{-1}[\cos(Z)\cos(\theta_T) + \sin(Z)\sin(\theta_T)\cos(Az - Az_A)] \tag{13}$$

Because reflections are generally specular in nature, $\hat{f}_2$ is applied only to the beam component of the incident irradiance. The parameter vector q_E is determined by fitting the expression in Eq. (12) using measurements of short-circuit current taken over a suitable range of module-sun geometries obtained using a two-axis tracker (see [18] for details). In our analysis we regard the effect on uncertainty in $\widehat{E_e}$ due to uncertainty in $\hat{f}_2$ as small in comparison with the effect of uncertainty in $\hat{f}_1$. Alternatives to Eq. (12) are available (e.g., [19], [14]).

Effective irradiance is not measured directly. Rather, effective irradiance is calculated from measured short-circuit current I_{SC} and cell temperature T_C [13]:

$$E = \frac{I_{SC}}{I_{SC0}\left(1 + \alpha_{Isc}\left(T_C - 25\right)\right)}, \tag{14}$$

where α_{Isc} is the pre-determined value for the temperature coefficient of I_{SC} and I_{SC0} is the pre-determined value for I_{SC} at standard test conditions (STC), i.e., broadband POA irradiance of

1000 W/m^2 and cell temperature of 25°C. For the modules in our experiments, α_{Isc} and I_{SC0} were estimated from testing of a similar model at Sandia National Laboratories using techniques described in [7]. In computing effective irradiance, we estimated the cell temperature T_C from the module back-plane temperature measurements using Eq. (17).

As with POA irradiance we quantified the residual for effective irradiance as the relative difference between modeled values $\hat{E}\left(t|p_E, q_E\right)$ calculated from measured irradiance and measured values $E\left(t\right)$ obtained from corresponding short-circuit current measurements:

$$\delta_E\left(t|p_E, q_E\right) = \frac{\widehat{E}\left(t|p_E, q_E\right) - E\left(t\right)}{E\left(t\right)} \tag{15}$$

Analysis of these residuals revealed systematic trends in $\delta_E(t \mid p_E, q_E)$ that changed with season, time of day, and sky condition (clear vs. cloudy). To facilitate random sampling from these results, we fit empirical expressions $y_E\left(t|p_E\right)$ to the residuals to separate trends from random effects:

$$\delta_E\left(t \mid p_E, q_E\right) = y_E\left(t \mid p_E, q_E\right) + \varepsilon_E\left(t \mid p_E, q_E\right) \tag{16}$$

Illustrative results for the residuals are shown in Section 3.2.

Step 3: Estimation of cell temperature

We model cell temperature $T_C\left(t\right)$ using the empirical approach proposed by [13]. Cell temperature is modeled by adding to module back surface temperature $T_M\left(t\right)$ a temperature difference that is proportional to POA irradiance G_{POA}:

$$T_C\left(t\right) = T_M\left(t\right) + \frac{G_{POA}\left(t\right)}{1000\ \text{W/m}^2}\Delta T\ . \tag{17}$$

$T_M\left(t\right)$ is modeled as a function of ambient temperature $T_{amb}\left(t\right)$ and wind speed $WS\left(t\right)$:

$$T_M\left(t\right) = G_{POA}\left(t\right)\exp\left(a + bWS\left(t\right)\right) + T_{amb}\left(t\right). \tag{18}$$

Combining Eq. (17) and Eq. (18) obtains a model for cell temperature:

$$\widehat{T_C}\left(t\,|\,p_{TC}\right)=\hat{f}_{TC}\left(\widehat{G}_{POA}\left(t\right),T_{amb}\left(t\right),WS\left(t\right)|\,p_{TC}\right) \\ =\widehat{G}_{POA}\left(t\right)\exp\left(a+bWS\left(t\right)\right)+T_{amb}\left(t\right)+\frac{\widehat{G}_{POA}\left(t\right)}{1000\ \mathrm{W/m^2}}\Delta T \tag{19}$$

The parameter vector $p_{TC}=\left(a,b,\Delta T\right)$ is determined by fitting the expression in Eq. (19) to module temperatures measured over a range of irradiance, air temperature and wind conditions.

Cell temperature T_C is not measured directly. Instead, T_C can be calculated from measured short-circuit current and open-circuit voltage using a technique similar to [20]. The residual for cell temperature is quantified by the difference between modeled values $\widehat{T_C}\left(t\right)$ and values $T_C\left(t\right)$ calculated from IV curve measurements for a module of the same model:

$$\varepsilon_{TC}\left(t\,|\,p_{TC}\right)=\widehat{T_C}\left(t\,|\,p_{TC}\right)-T_C\left(t\right) \tag{20}$$

Illustrative results are provided in Section 3.3.

Step 4: Calculation of DC power

We obtain DC power $\hat{P}_{DC}\left(t\right)$ by separately predicting DC voltage $\hat{V}_{DC}\left(t\right)$ and current $\hat{I}_{DC}\left(t\right)$ from the PV modules by using the Sandia Array Performance Model (SAPM) $\hat{f}_{DC}$ [13]:

$$\left[\hat{V}_{DC}\left(t\,|\,p_{DC}\right),\,\hat{I}_{DC}\left(t\,|\,p_{DC}\right)\right]=\hat{f}_{DC}\left(\widehat{E}\left(t\right),\,\widehat{T_C}\left(t\right)|\,p_{DC}\right)+\left[\varepsilon_{VDC}\left(t\,|\,p_{DC}\right),\,\varepsilon_{IDC}\left(t\,|\,p_{DC}\right)\right] \tag{21}$$

In Eq. (21) the residuals $\varepsilon_{VDC}\left(t\,|\,p_{DC}\right)$ and $\varepsilon_{IDC}\left(t\,|\,p_{DC}\right)$ represent the residuals for DC voltage $\hat{V}_{DC}\left(t\right)$ and current $\hat{I}_{DC}\left(t\right)$, respectively. Module DC power (with uncertainty) is then determined by multiplying:

$$\hat{P}_{DC}\left(t\,|\,p_{DC}\right)=\hat{V}_{DC}\left(t\,|\,p_{DC}\right)\times\hat{I}_{DC}\left(t\,|\,p_{DC}\right) \tag{22}$$

The parameter vector p_{DC} contains 13 module-specific values that are determined from IV curve measurements during a sequence of performance tests [7].

The residuals $\varepsilon_{VDC}\left(t\,|\,p_{DC}\right)$ and $\varepsilon_{IDC}\left(t\,|\,p_{DC}\right)$ are quantified by comparing modeled values with measured I-V curves. Illustrative results are provided in Section 3.4.

Step 5: Estimation of array DC loss

With the estimate of module output power given illumination and temperature conditions, we now consider an array of similar modules connected to a DC-to-AC inverter in N_P parallel strings, each with N_S series modules. Ideally, maximum power from such an array is generated by operating each module at its maximum power point, i.e., the voltage at which the module produces maximum power output. Hence, array performance models (e.g., [13]) typically use effective irradiance and cell temperature to estimate the maximum power point of a single representative module and then assume that every module in the array can be operated identically so that the array output can be computed by multiplying the representative module's output:

$$P_{array}(t) = \left[N_s V_{DC}(t) \right] \times \left[N_P I_{DC}(t) \right] = N_P \, N_s \, P_{DC}(t) . \tag{23}$$

This approach generally overestimates the amount of DC power an actual array will produce for two reasons: first, differences between modules in a string prevent operating all models at their individual maximum power points, and second, even if all the modules in an array were identical, modern control systems cannot perfectly maintain the maximum power point at each instant. Power decreases due to these two mechanisms are referred to as mismatch and maximum power point tracking (MPPT) losses, respectively.

Mismatch loss arises from the fact that the modules in an array are not identical. Some differences between modules are systemic, such as variations in module current or voltage due to manufacturing or in operating conditions due to a module's installation. For example, the location of a module in an array may affect its operating temperature. Other variations are ephemeral, such as non-uniform illumination across an array due to passing clouds or differences in temperature due to changing air currents. Because of these variations, each module will have a slightly different maximum power point at each instant in time. However, the DC current in a string must be the same for each module causing individual modules to vary from their maximum power points. Consequently the string will produce less than the maximum DC power that would be predicted by the array performance model.

MPPT loss arises from the inability of control systems in modern inverters to precisely and instantly find the operating voltage that will yield maximum power output from the array. In practice, the control system adjusts the string voltage by set increments and at fixed time intervals based on present and previous measurements of array output. Imprecise or delayed adjustments to string voltages result in the array delivering to the inverter less than the maximum DC power that would be predicted by array performance model.

While these two loss mechanisms are reasonably well-understood, few widely-accepted models exist for accurately estimating these losses. Analyses have shown that the string mismatch losses are typically quite small even when the modules comprising the string have substantial differences (e.g., [21]). MPPT losses are dependent on the tracking algorithm in use but are also small in magnitude ([22], [23]) and are acknowledged as difficult to measure precisely [24].

Rather than attempt to separately quantify the uncertainty in a string mismatch loss model and an MPPT loss model, we construct a data-driven model of the aggregate of these two losses which we term *array DC loss*, and quantify the uncertainty in this aggregate model. Given a nonuniform array driven at a sub-optimal control voltage, we define effective module power, $P_{DC,eff}(t)$, as the power delivered to the inverter by the 'average' module, i.e., a module considered representative of the modules comprising the array. Denoting the maximum power this module could produce as $P_{DC}(t)$, we write $P_{DC,eff}(t)$ as

$$P_{DC,eff}(t) = P_{DC}(t) - \Delta P(t), \tag{24}$$

where $\Delta P(t)$ represents the DC power loss *per module* due to string mismatch and MPPT inefficiency. Based on experimental results described in Section 3.5, we use a constant value of $\Delta P(t)$ for each day, which we select depending on the variability of the day's irradiance.

We use the model described above to estimate DC output power from an array as

$$\hat{P}_{array}(t) = N_P \, N_S \left(\hat{P}_{DC}(t) - \Delta P(t) \right) \tag{25}$$

where $\hat{P}_{DC}(t)$ is obtained from Eq. (22). Uncertainty in $\hat{P}_{array}(t)$ arises from uncertainty in $\Delta P(t)$, which is characterized by comparing the maximum power of a reference module with the power output of an array driven by a maximum power point tracking algorithm. Illustrative results are shown in Section 3.5.

Step 6: Estimation of AC power

A final loss mechanism considered is the imperfect efficiency of the inverter. Some of the DC power that is provided to an inverter is lost in the conversion to AC power. Much of this lost power is converted to heat by the physical mechanisms involved in the conversion process. We model inverter output AC power $P_{AC}(t)$ as

$$\hat{P}_{AC}(t) = \hat{f}_{inv}\left(P_{DC}(t) \mid p_{inv} \right), \tag{26}$$

where $\hat{f}_{inv}\left(P_{DC}(t) \mid p_{inv} \right)$ represents the generally-used model proposed in [25]:

$$\hat{f}_{inv}\left(P_{DC}(t) \mid p_{inv} \right) = \frac{P_{AC0} - C}{A - B}\left(P_{DC}(t) - B \right) + C\left(P_{DC}(t) - B \right)^2 . \tag{27}$$

In this model, P_{AC0} represents the maximum ac-power "rating" for the inverter at reference or nominal operating conditions. The intermediate variables *A*, *B*, and *C* are defined in terms of the operating voltage $V_{DC}(t)$ and several other parameters:

$$\begin{aligned} A &= P_{DC0}\left[1 + C_1 \cdot \left(V_{DC}(t) - V_{DC0}\right)\right] \\ B &= P_{s0}\left[1 + C_2 \cdot \left(V_{DC}(t) - V_{DC0}\right)\right] \\ C &= C_0\left[1 + C_3 \cdot \left(V_{DC}(t) - V_{DC0}\right)\right] \end{aligned} \tag{28}$$

The parameter vector $p_{inv} = \left(P_{AC0}, P_{DC0}, V_{DC0}, P_{s0}, C_0, C_1, C_2, C_3\right)$ is determined by fitting the expression in Equations (27) and (28) to inverter output power measured over a range of input power and voltage conditions, as specified, for example, in a test protocol jointly developed by Sandia National Laboratories and BEW and promulgated by the California Energy Commission (CEC) [26].

Because the inverter tests are replicated a number of times (10 is typical) the parameter vector p_{inv} is regarded as uncertain. A set of parameter vectors $\left\{p_{inv,k} \,\middle|\, k = 1,\ldots,N\right\}$ is obtained by fitting the model indicated in Eq. (27) and Eq. (28) to the measurements for each individual test. Uncertainty in $\hat{f}_{inv}\left(P_{DC,eff}(t) \mid p_{inv}\right)$ is thus quantified by elements of $\left\{p_{inv,k} \,\middle|\, k = 1,\ldots,N\right\}$. Illustrative results are shown in Section 3.6.

3. QUANTIFYING UNCERTAINTY

Here we describe the quantification of uncertainty for each modeling step: POA irradiance (Section 3.1), effective irradiance (Section 3.2), cell temperature (Section 3.3), DC power (Section 3.4), array DC loss (Section 3.5), and DC-to-AC conversion (Section 3.6).

3.1. POA Irradiance

As noted earlier, estimating POA irradiance from GHI, DNI and DHI requires estimating the beam and diffuse components:

$$G_{POA} = E_b + E_{diff} \tag{29}$$

Beam irradiance E_b is determined from DNI by accounting for the sun's angle of incidence AOI on the module:

$$E_b = DNI \times \cos\left(AOI\right) \tag{30}$$

Angle of incidence AOI is computed using the module's assumed fixed orientation (latitude tilt and 180° azimuth) and the sun position computed using a legacy algorithm originating with G. Hughes [15], which has sufficiently high accuracy that we did not consider uncertainty in AOI.

The diffuse component is divided into ground reflected irradiance $E_{diff,g}$ and sky diffuse irradiance $E_{diff,sky}$:

$$E_{diff} = E_{diff,g} + E_{diff,sky} \tag{31}$$

Ground reflected irradiance $E_{diff,g}$ is estimated using an uncertain value for ground albedo a:

$$E_{diff,g} = GHI \times a \times \frac{1-\cos\left(A\right)}{2} \tag{32}$$

where A is the tilt angle of the module towards the equator (assumed constant and precisely known). The model for $E_{diff,g}$ in Eq. (32) assumes horizontal surrounding terrain and isotropic reflection of GHI from the ground [27] that is independent of solar zenith angle, solar azimuth, or time of day. Thus, the albedo parameter a represents a spatially and temporally averaged fraction of GHI reflected from the ground. No precise quantification of albedo is available. It is typical to assume a value $a = 0.2$ [27]. We use 0.2 as a base value; our earlier work [1] concluded that the value had little influence in our analysis.

Sky diffuse irradiance $E_{diff,sky}$ is calculated using one of four alternative models (listed in order of increasing model complexity):

- Isotropic sky diffuse model [9];
- Sandia simple sky diffuse model [11];
- Hay and Davies diffuse model [9];
- Perez sky diffuse model [9].

For models that require parameter values we used values generally regarded as typical, as follows:

- Sandia simple sky diffuse model – empirical coefficients 0.012 and 0.004, from [11]. These values were determined by calibrating the model to measurements of GHI, DHI, and ambient temperature at Sandia in Albuquerque, NM, prior to 2010.
- Hay and Davies diffuse model – annual average extraterrestrial radiation E_a taken equal to 1367 W/m^2.
- Perez sky diffuse model – the Perez model requires a large number of empirical coefficients. We used values from Table 6 in [12] that are regarded as typical and recommended by the model's originators.

For each model we determined an empirical distribution for the model residuals using concurrent measurements of GHI, DNI, DHI and POA irradiance at Sandia's Photovoltaic Systems Evaluation Laboratory during 2013-2014. We used measured GHI, DNI and DHI to predict POA irradiance then computed model residuals by comparing predicted to measured POA irradiance. Using scatterplots and other techniques, we identified systematic trends in the residuals that are reflected in the characterization of uncertainty in each model's residuals.

To avoid including residuals primarily resulting from measurement artifacts we exclude data measured with sun elevation angles less than 10°. At the PSEL site, local shadowing occasionally affects measurements at low sun elevation angles. In addition, the instrument used to measured POA irradiance shows measurement aberrations due to internal reflections at high incident angles. For these reasons our analysis precludes low sun elevation angles.

3.1.1. Isotropic sky diffuse model

Figure 3 displays the residuals for the Isotropic sky diffuse model as a function of angle of incidence AOI. Dependence of the residuals on AOI is evident. Figure 4 shows residuals for two different months and demonstrates dependence of residuals on time of year. Figure 5 illustrates that residuals in a given month are different for clear sky conditions as opposed to cloudy conditions. Finally, Figure 6 and Figure 7 show that residuals can also depend on time of day. In Figure 6 and Figure 7, we separate residuals into two subsets (before and after noon), and fit each subset with a second order polynomial in AOI to quantify the different trends.

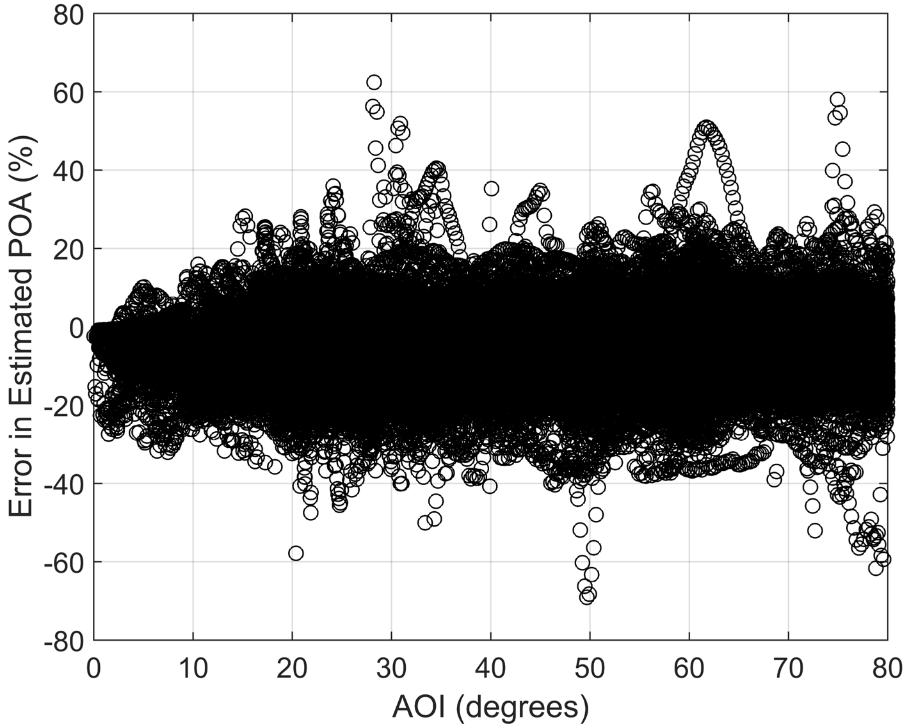


Figure 3. Residuals for predicted POA irradiance using Isotropic sky diffuse model.

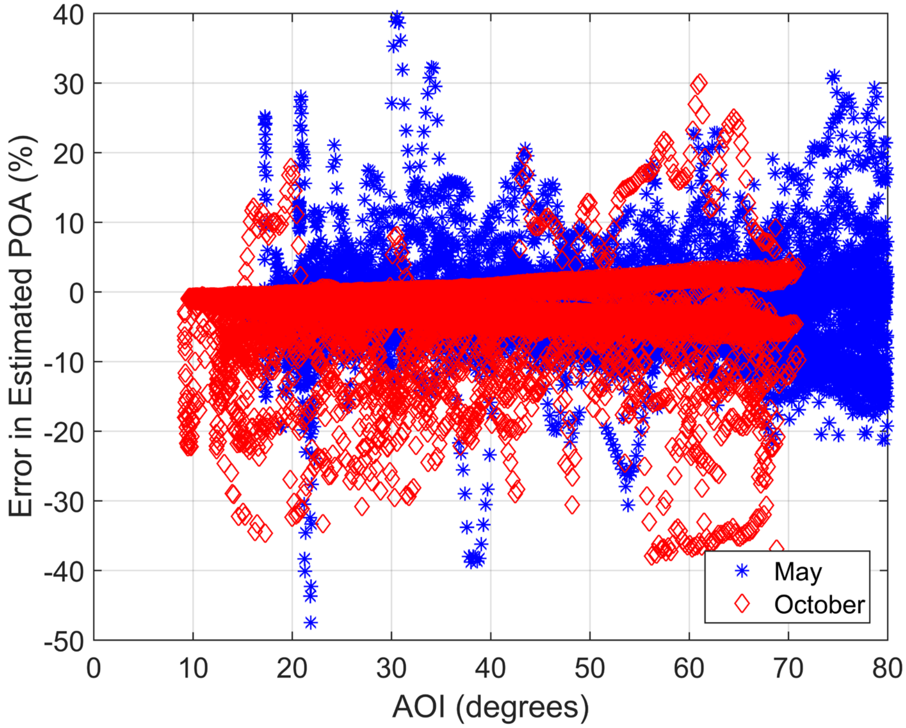


Figure 4. Dependence of residuals on time of year for Isotropic sky diffuse model.

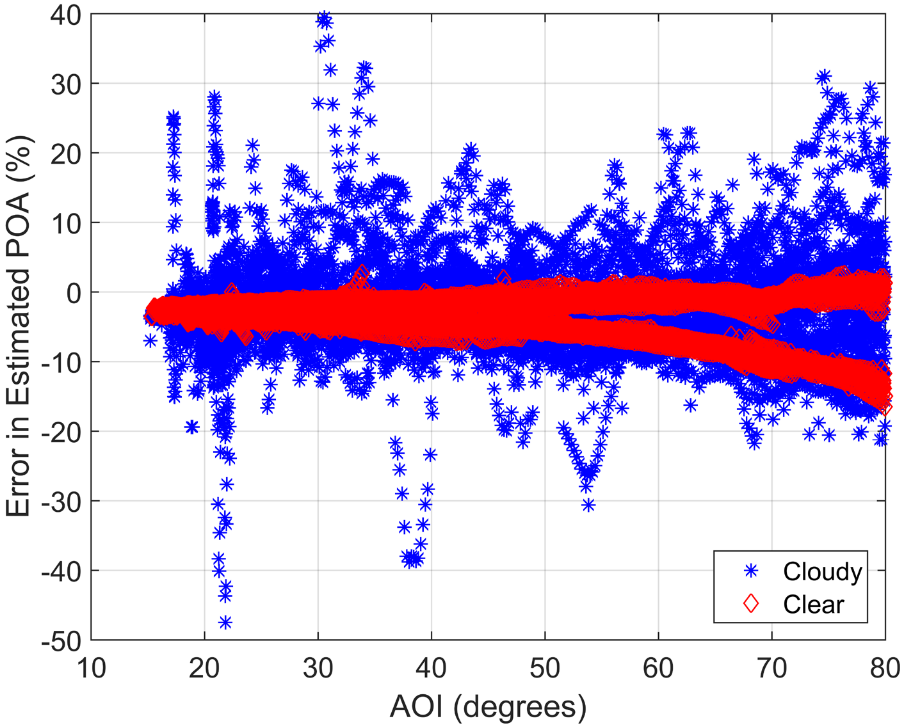


Figure 5. Dependence of residuals on sky condition for Isotropic sky diffuse model (May 2014 shown).

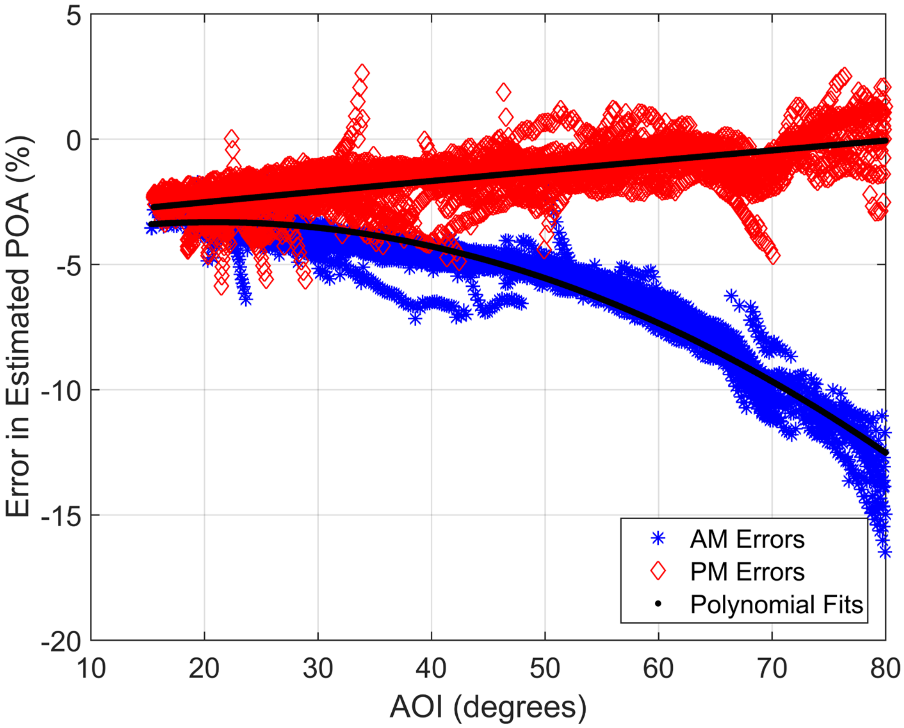


Figure 6. Dependence on time of day of residuals for Isotropic sky diffuse model (clear periods during May 2014 shown).

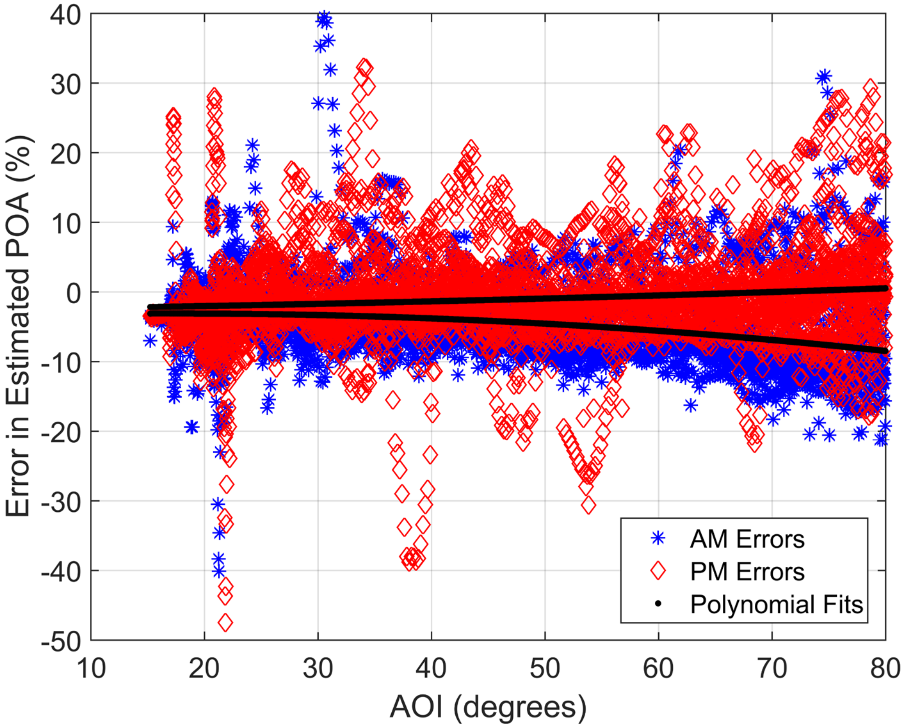


Figure 7. Dependence on time of day of residuals for Isotropic sky diffuse model (cloudy periods during May 2014 shown).

Using this partition of the data, we assembled 48 empirical distributions of model residuals for the Isotropic clear sky model. For each month, we distinguished clear and cloudy conditions by the ratio of measured DNI to GNI (Global Normal Irradiance—measured with a pyranometer bore-sighted with the DNI pyrheliometer): if the ratio was greater than 0.85, we considered the sky to be clear. We further separated each day into morning (AM) and afternoon (PM) time periods, resulting in 48 data subsets. Within each subset we quantified the systematic dependence of the residuals on AOI by fitting a second order polynomial (as illustrated in Figure 6 and Figure 7). The polynomial fit is generally successful at removing the systematic trends (see, for example, Figure 8 and Figure 9). We then estimated one or more empirical cumulative distribution functions (CDFs) from the difference between each residual and the fitted polynomial (e.g., Figure 10 and Figure 11). To accommodate changes in variance of the de-trended residuals across AOI, we partitioned AOI into several bins and obtained separate CDFs for each bin.

We note that the de-trended residuals exhibit a relatively consistent daily pattern during clear sky conditions (e.g., Figure 8) but a rather random pattern during cloudy conditions (e.g., Figure 9). Accordingly, to propagate uncertainty we sample one quantile value for all times with clear conditions for each day, but randomly sample a quantile value for each time with cloudy conditions.

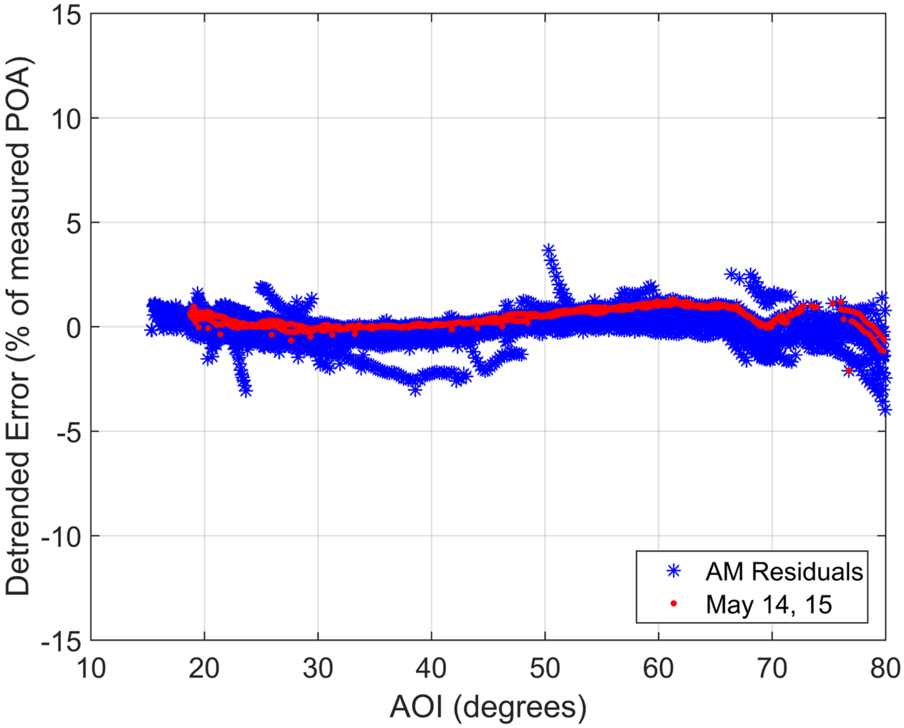


Figure 8. Residuals for clear sky conditions, May 2014, before noon, after removing trend.

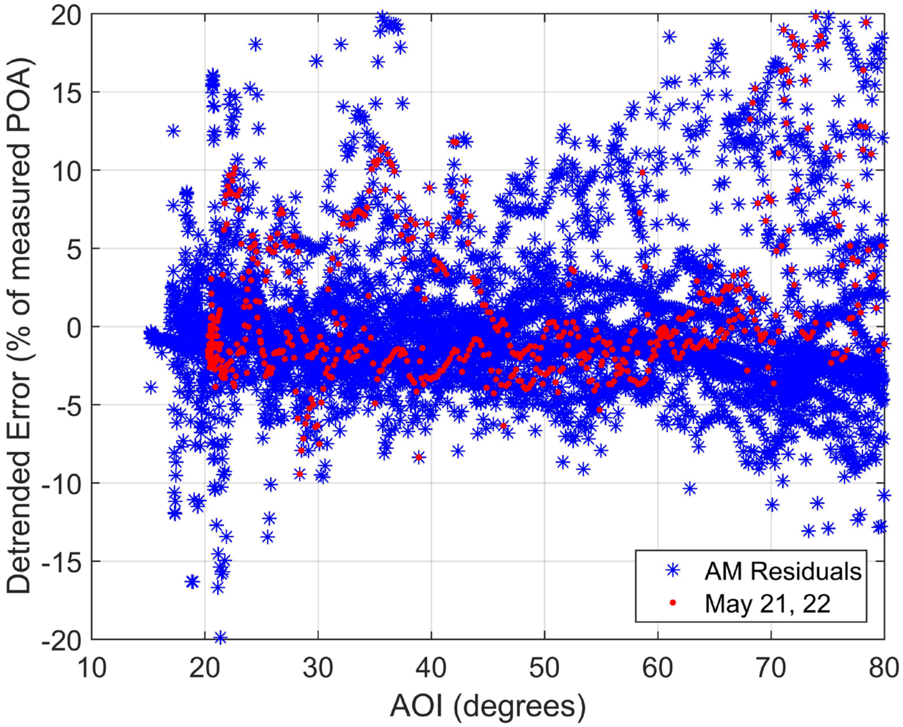


Figure 9. Residuals for cloudy conditions, May 2014, before noon, after removing trend.

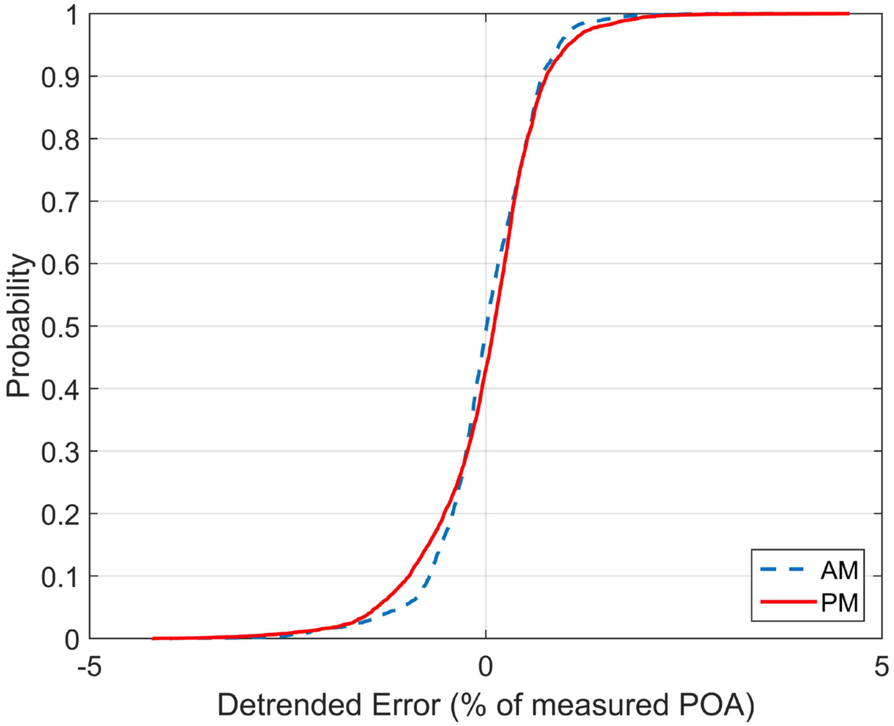


Figure 10. Empirical CDFs for de-trended residuals during clear sky conditions (May 2014 shown).

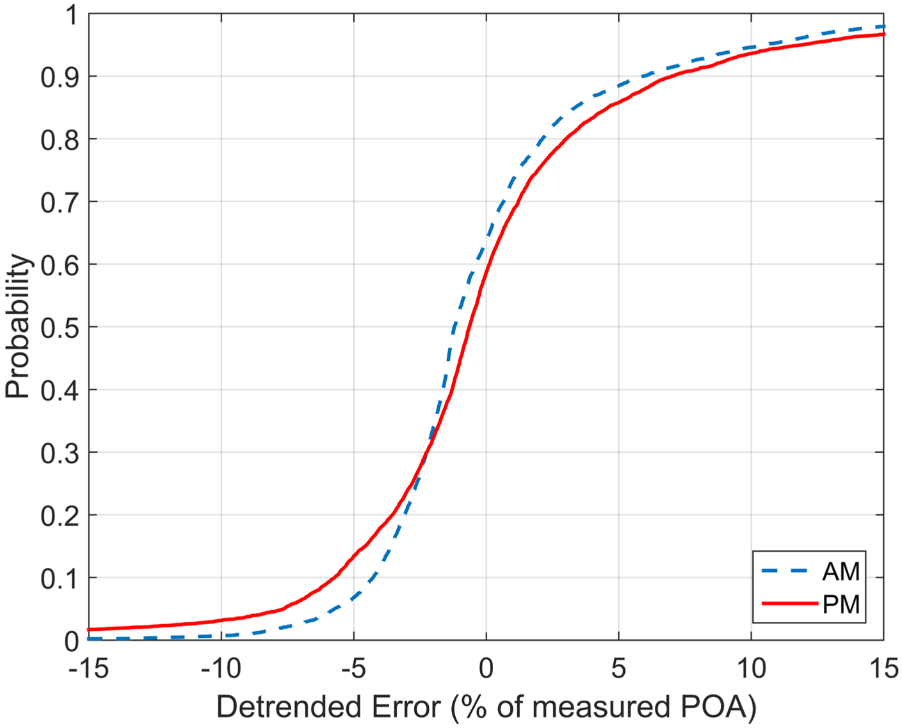


Figure 11. Empirical CDFs for de-trended residuals during cloudy sky conditions (May 2014 shown).

3.1.2. Sandia simple sky diffuse model

For the Sandia simple sky diffuse model we adopted the same general approach for quantifying uncertainty in model residuals as is used for the Isotropic sky diffuse model. We observed a range of residuals generally comparable to the Isotropic sky diffuse model (compare Figure 12 and Figure 5).

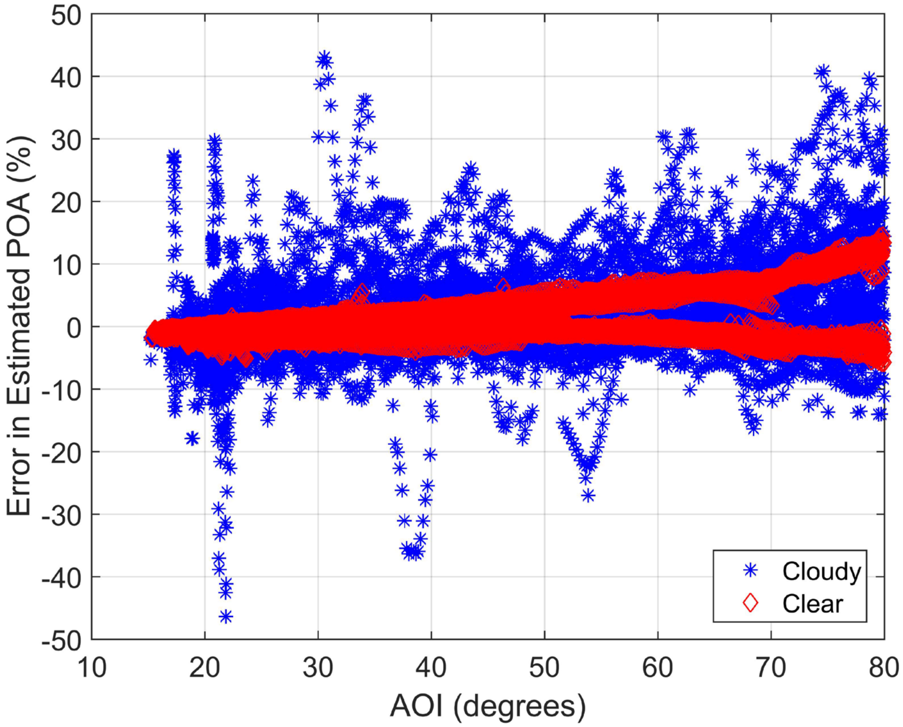


Figure 12. Dependence on sky condition of residuals for Sandia simple sky diffuse model (May 2014 shown).

3.1.3. Hay and Davies diffuse model

For the Hay and Davies diffuse model we adopted the same general approach for quantifying uncertainty in model residuals as is used for the Isotropic sky diffuse model, and observed similar ranges of residuals (compare Figure 13 and Figure 5).

3.1.4. Perez sky diffuse model

For the Perez sky diffuse model we adopted the same general approach for quantifying uncertainty in model residuals as is used for the Isotropic sky diffuse model. Similar ranges of residuals are observed in these models (compare Figure 14 and Figure 5).

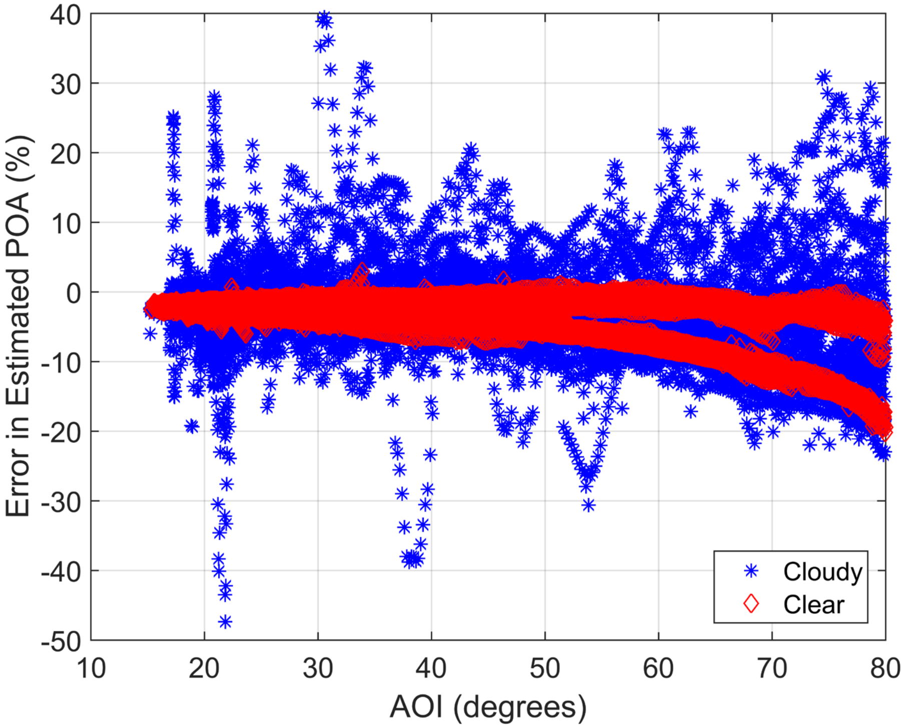


Figure 13. Dependence on sky condition of residuals for Hay and Davies diffuse model (May 2014 shown).

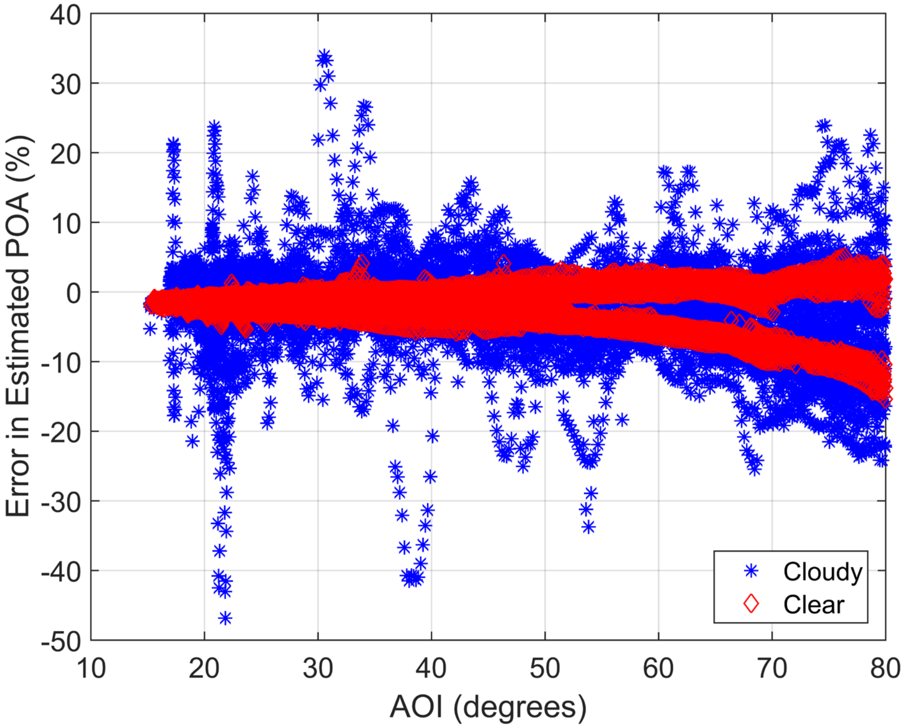


Figure 14. Dependence on sky condition of residuals for Perez sky diffuse model (May 2014 shown).

Finally, comparing the systematic trends in the residuals of the four POA irradiance models, we observe that the trends are similar among the models, with more or less the same general shape in each month, and the magnitude of each trend varying depending on the model (see Figure 15). In a given month, the systematic trends in the Hay and Davies model residuals have the most variance; the Sandia model seems to have the least-varying trends.

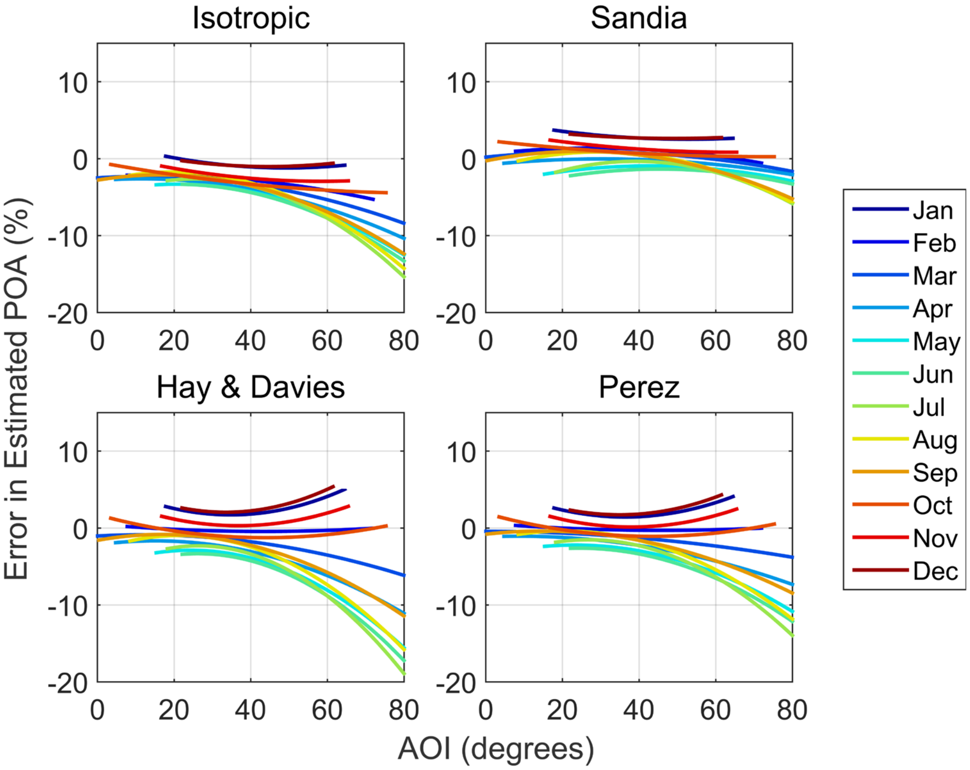


Figure 15. Trends in POA model residual by model and month (clear morning conditions shown).

3.2. Effective irradiance

Estimating effective irradiance involves beam irradiance and diffuse irradiance in the plane of the array (E_b and E_{diff} respectively), two polynomial functions $f_1(AM)$ and $f_2(AOI)$, and a derate factor SF to account for surface soiling:

$$E_e = f_1(AM)\left[f_2(AOI)E_b + E_{diff}\right]SF \tag{33}$$

Beam irradiance E_b is computed from measured DNI using Eq. (30). Because we have more reliable measurements of POA than E_{diff}, we obtain E_{diff} from POA irradiance G_{POA} and E_b by

$$E_{diff} = G_{POA} - E_b \tag{34}$$

We use Eq. (34) to compute diffuse POA irradiance at this stage in order to avoid compounding effective irradiance model residuals with error from the sky diffuse models discussed in Section 3.1. Assuming a value of 1.0 for the soiling factor SF of a clean module, our model for effective irradiance could be stated as

$$E_e = f_1(AM)\left[G_{POA} - \left(1 - f_2(AOI)\right)E_b\right]. \quad (35)$$

We determined coefficients of the polynomials $f_1(AM)$ and $f_2(AOI)$ for the First Solar modules by measuring electrical performance of one module over a range of *AOI* and *AM* values using a two-axis tracker. We also determined effective irradiance from concurrent measurements of module short-circuit current and cell temperature (Eq. (14)). We then compiled an empirical distribution for the effective irradiance residuals (see Eq. (15)).

We found systematic trends in the effective irradiance residuals similar to those found in the POA model residuals. Figure 16 displays the residuals for the First Solar effective irradiance model as a function of *AOI*. The dependence of the residuals on *AOI* is immediately apparent. Successive subdivision of the residuals by time of year, sky condition and time of day further highlights this dependence, as depicted in Figure 17 through Figure 20. Figure 17 shows the residuals for two different months, while Figure 18 shows how the residuals for a single month differ between clear and cloudy sky conditions. Figure 19 illustrates the difference between morning and afternoon residuals for clear conditions, and Figure 20 shows the difference between morning and afternoon residuals for cloudy conditions.

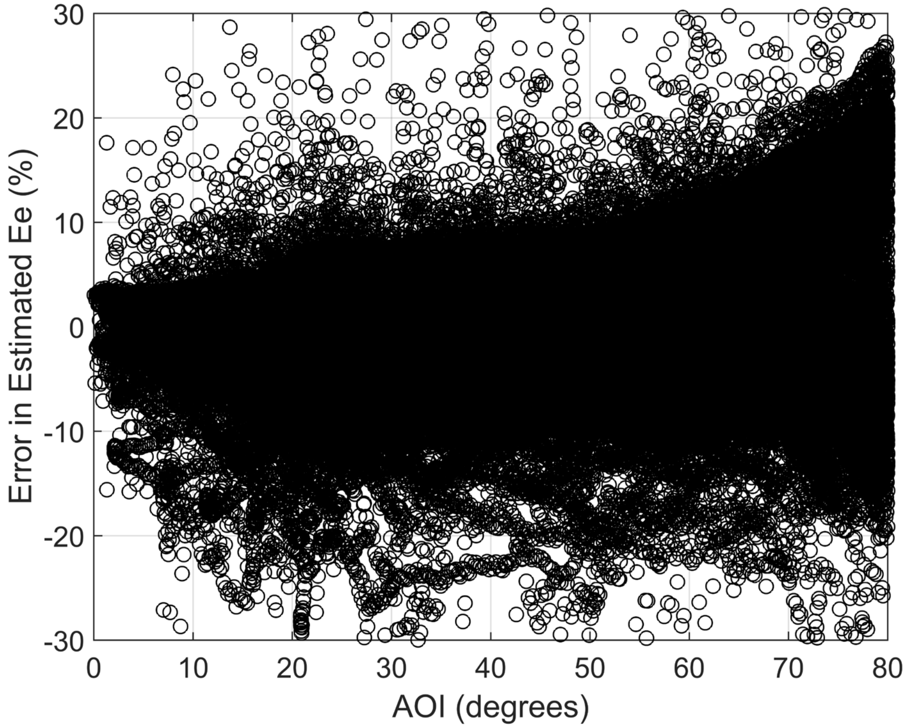


Figure 16. Residuals for predicted effective irradiance.

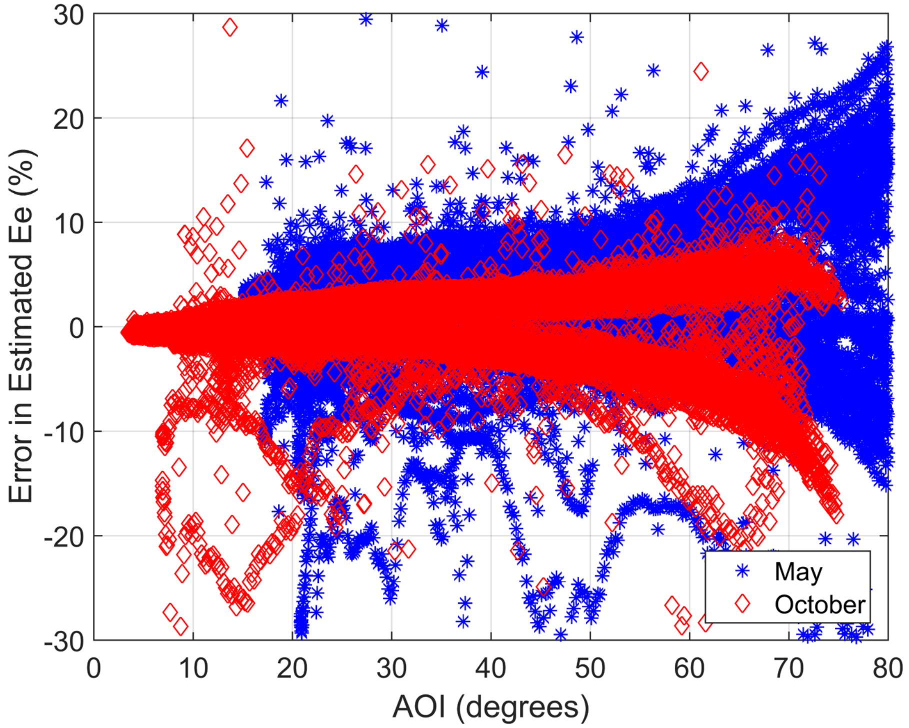


Figure 17. Dependence of effective irradiance residuals on time of year.

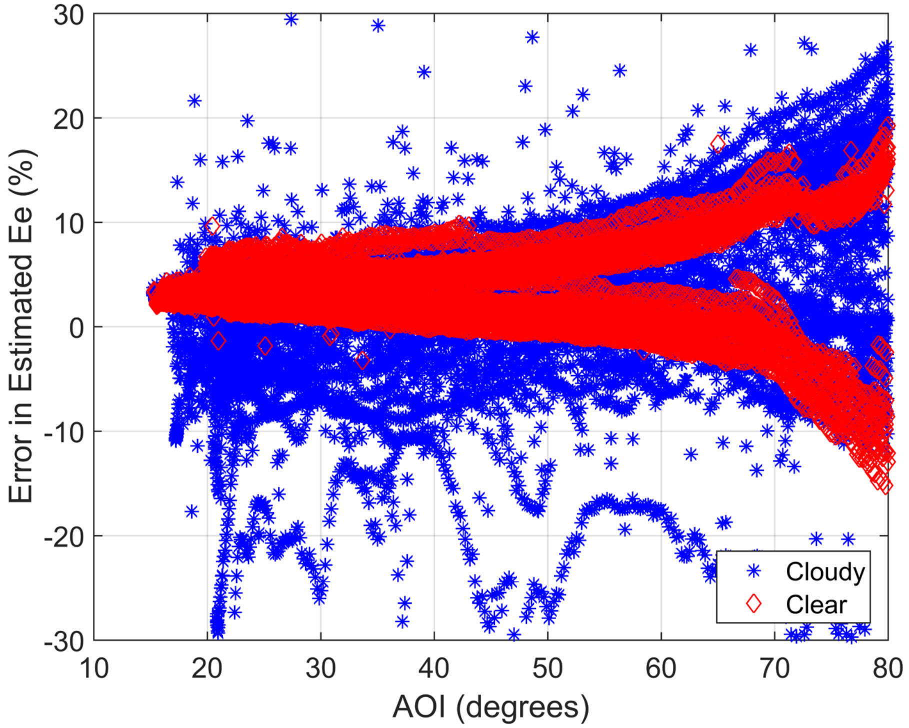


Figure 18. Dependence of effective irradiance residuals on sky condition (May 2014 shown)

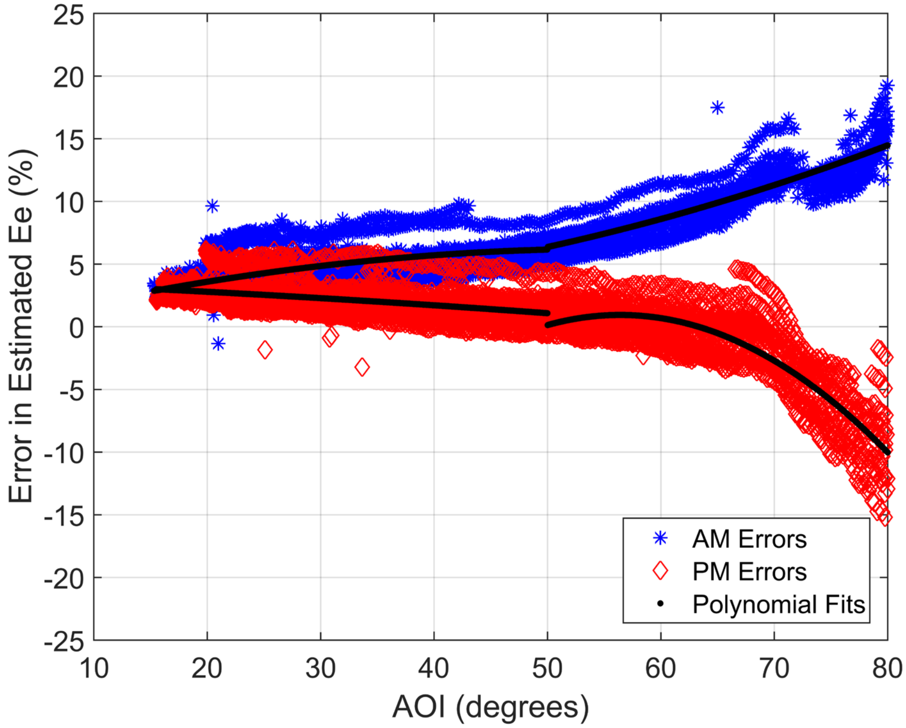


Figure 19. Dependence of effective irradiance residuals on time of day (clear periods during May 2014 shown)

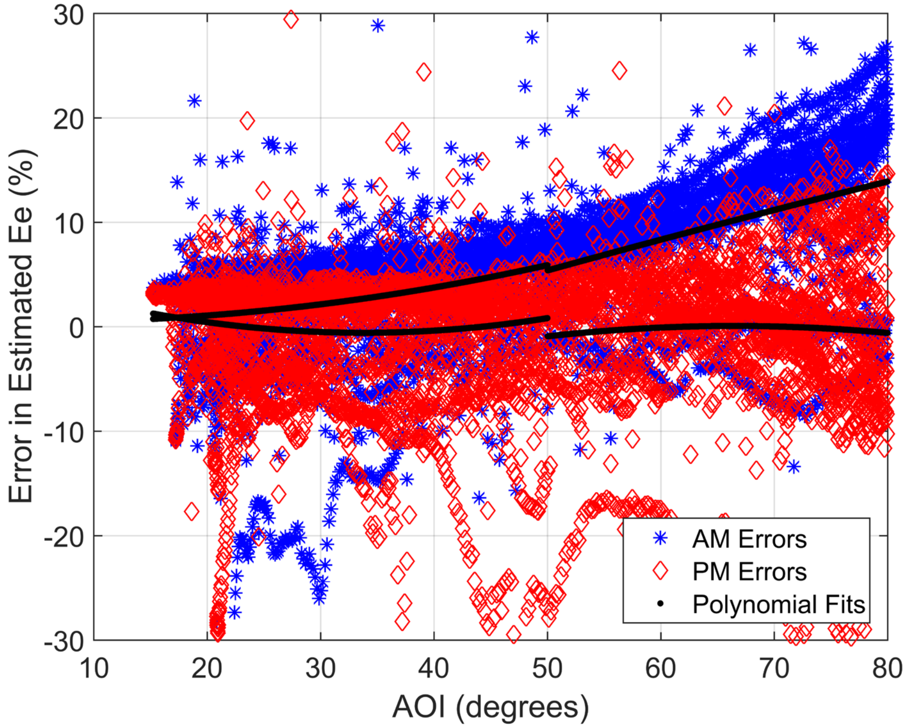


Figure 20. Dependence of effective irradiance residuals on time of day (cloudy periods during May 2014 shown).

We subdivided the effective irradiance residuals by month, sky condition, and time of day, and we further subdivided each month/sky condition/time of day subset into two additional subsets ($AOI \leq 50$ and $AOI > 50$), resulting in a total of 96 subsets. We de-trended the residuals by fitting a second-order polynomial in AOI to the data in each subset. No attempt was made to ensure continuity of the polynomials across the $AOI = 50$ boundary. Examples of the de-trending polynomials are depicted in Figure 19 and Figure 20.

Figure 21displays monthly trends fit to effective irradiance residuals for each time-of-day and sky condition combination. The collection of trends reveals that the model generally overestimates effective irradiance in the morning and underestimates it in the afternoon, and the error is generally greater the closer the sun is to the horizon.

Subtracting the fitted polynomial from the residuals in each subset generally removes the systematic trend (see typical results in Figure 22 and Figure 23). To characterize the remaining variation in the residuals, we estimated an empirical CDF of the de-trended residuals in each of the 96 subsets (see Figure 24 and Figure 25). The polynomial fit and CDF for each subset thus comprise an empirical distribution of the First Solar effective irradiance residuals for each combination of time of year (month), time of day, angle of incidence range, and sky condition. Figure 22 and Figure 23 show that the de-trended effective irradiance residuals have a fairly constant daily pattern during clear sky conditions, but vary randomly during cloudy sky conditions. Accordingly, when we propagate uncertainty, we use a single random quantile value for all times with clear sky conditions in a day, but we sample a different random quantile value for each time with cloudy sky conditions.

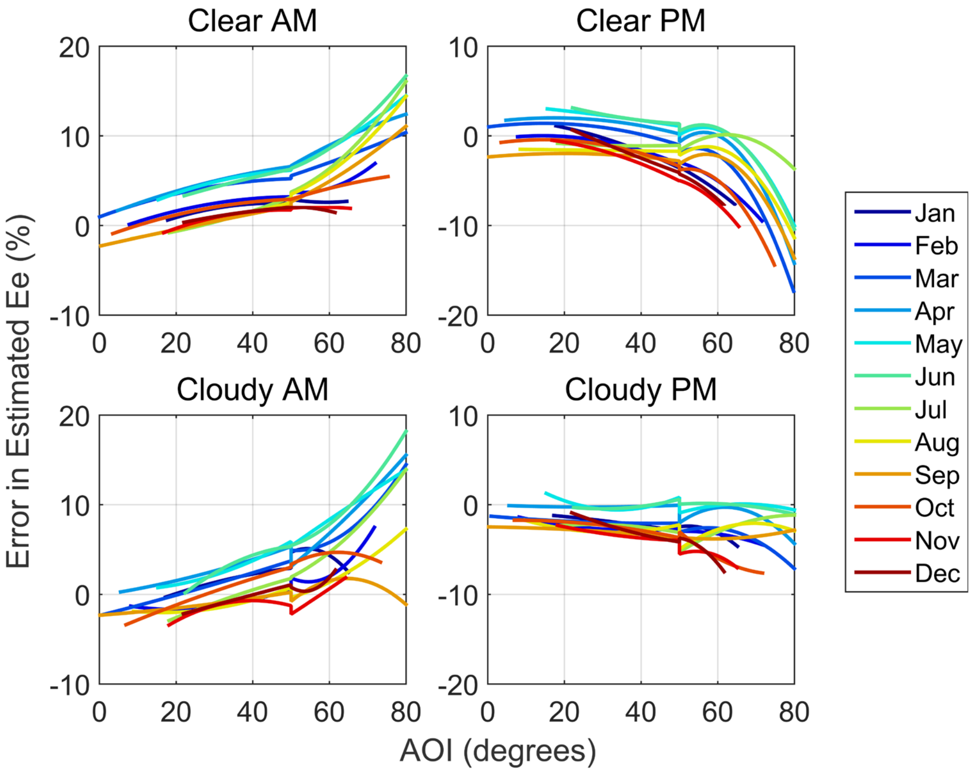


Figure 21. Monthly trends fit to effective irradiance residuals

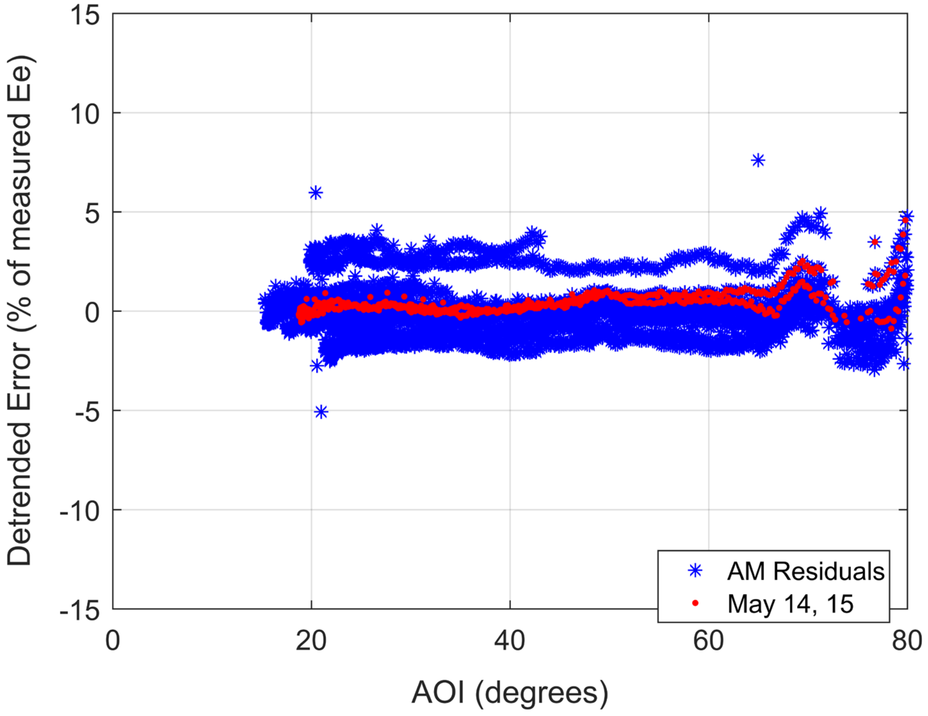


Figure 22. Effective irradiance residuals for May 2014 morning clear sky conditions after removing trends.

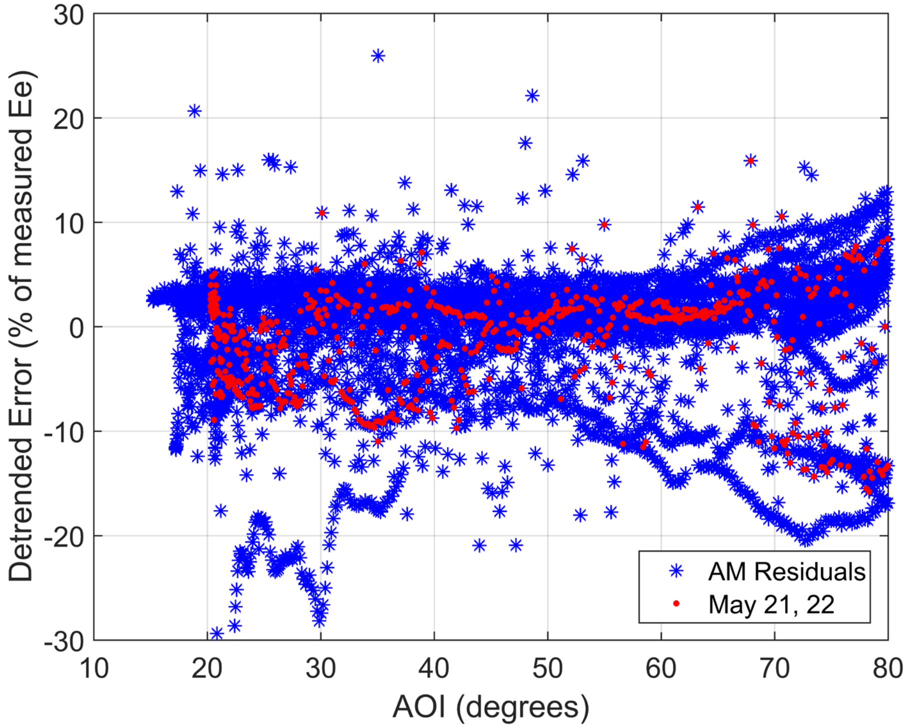


Figure 23. Effective irradiance residuals for May 2014 morning cloudy sky conditions after removing trends.

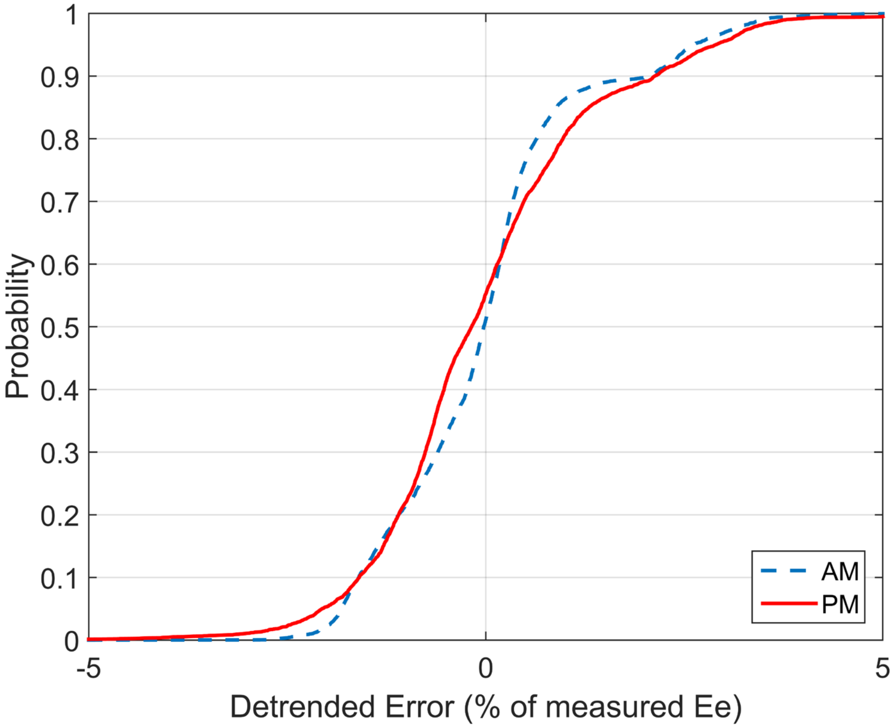


Figure 24. Empirical CDFs for de-trended effective irradiance residuals during clear sky conditions (May 2014 shown)

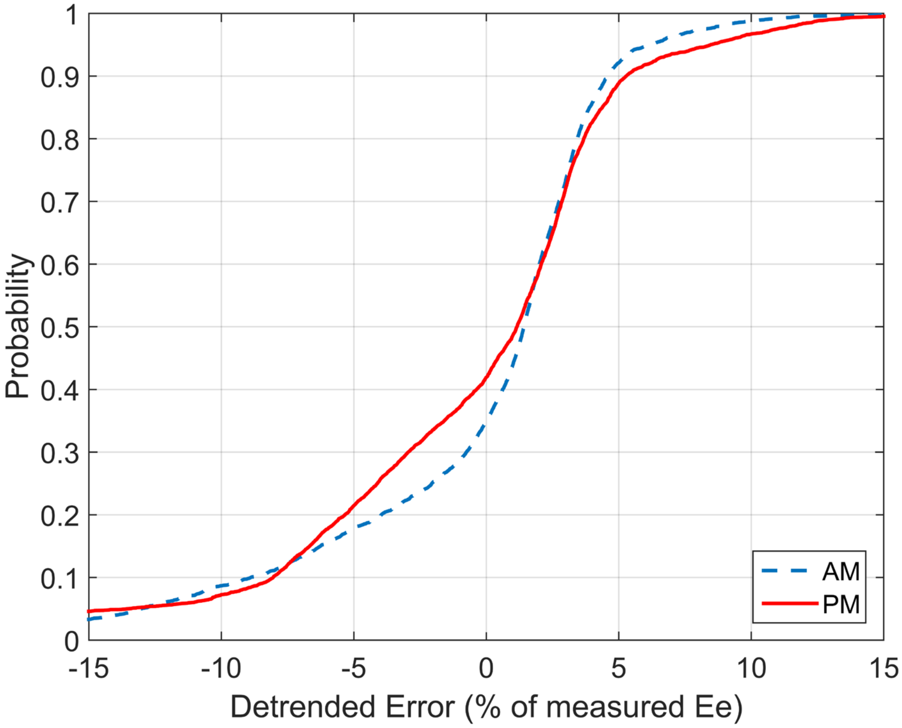


Figure 25. Empirical CDFs for de-trended effective irradiance residuals during cloudy sky conditions (May 2014 shown)

Because the partitions for the effective irradiance distributions are not based on POA irradiance, the same distributions for effective irradiance are used with each of the four different POA irradiance models. Also, because calculated effective irradiance was obtained from the same measurements as were used to estimate the effective irradiance model in Eq. (7), the CDFs for residuals for effective irradiance are generally unbiased.

3.3. Cell Temperature

Cell temperature is modeled using Eq. (19). We determined an empirical distribution of the cell temperature model residuals using measurements of POA irradiance, wind speed and ambient temperature at Sandia's Photovoltaic Systems Evaluation Laboratory (PSEL) during January 2013, along with I-V curves recorded for a First Solar module during this same time. We estimated cell temperature from the I-V curves using a technique similar to [20], and computed a set of residuals by comparing modeled and estimated cell temperatures. Examining the residuals, we found different behavior for different sky conditions and ranges of wind speed WS (see Figure 26). Accordingly, we created three empirical distributions for the model residuals. Conditional on clear or cloudy conditions, we constructed three distribution models for the residual in cell temperature by partitioning wind speed into two intervals for cloudy periods:.

$$\text{Clear:}\quad \mathrm{E1} = \{0 < \mathrm{WS} < \infty\} \tag{36}$$

$$\text{Cloudy:}\quad \mathrm{E2} = \{0 < \mathrm{WS} \le 5\} \tag{37}$$

$$\mathrm{E3} = \{5 < \mathrm{WS}\} \tag{38}$$

We constructed an empirical CDF for the residuals in cell temperature for each interval, as displayed in Figure 27. Because the partitions for the distributions are not based on POA irradiance, the same distributions are used for each of the four POA models created above.

3.4. PV Module Output

Figure 28 shows measured DC power and DC voltage from testing of a First Solar FS-387 module at PSEL during February 2013. Using these measurements we determined coefficients for the Sandia Array Performance Model (SAPM) [13], which we then used to produce model predictions of DC power and DC voltage, also depicted in Figure 28.

We find the residuals in predicted DC current and predicted DC voltage to be uncorrelated (Figure 29; correlation coefficient 0.0067). We see a relatively strong correlation between effective irradiance and cell temperature (Figure 30), as should be expected.

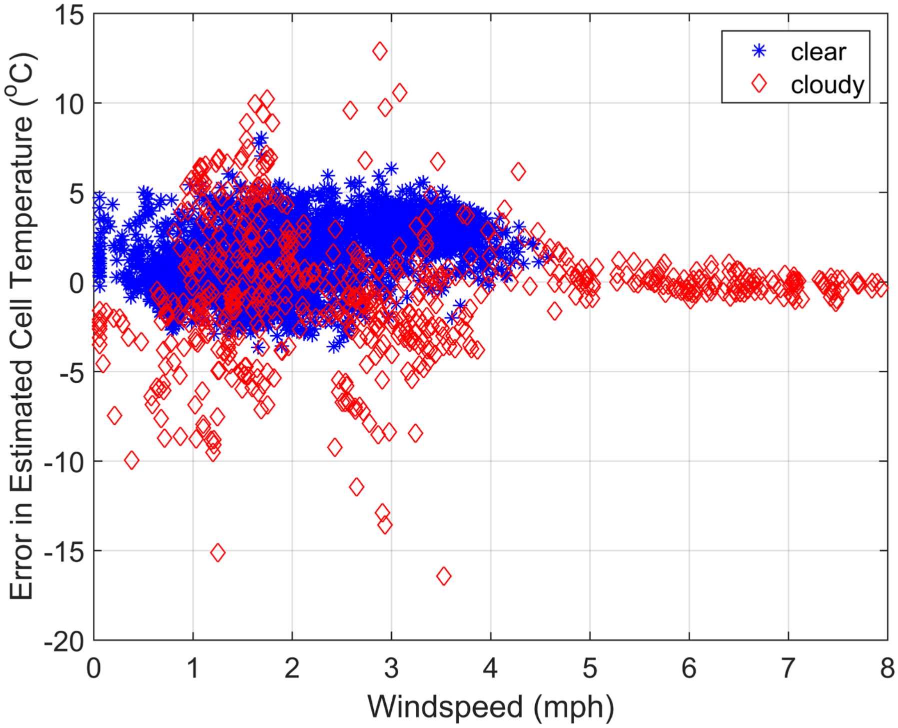


Figure 26. Residuals for Cell Temperature Model.

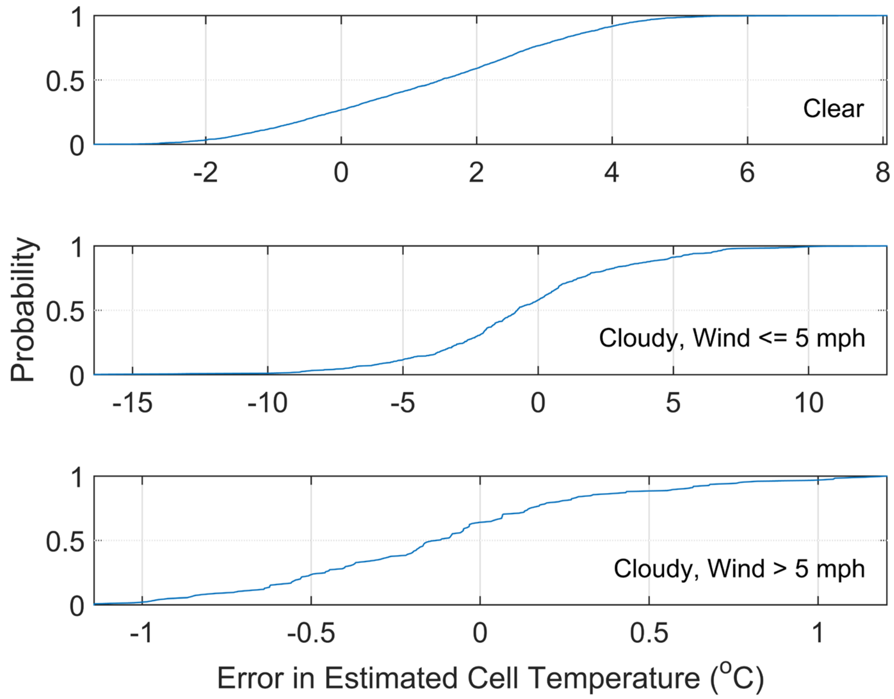


Figure 27. Empirical CDFs for Cell Temperature Model Residual.

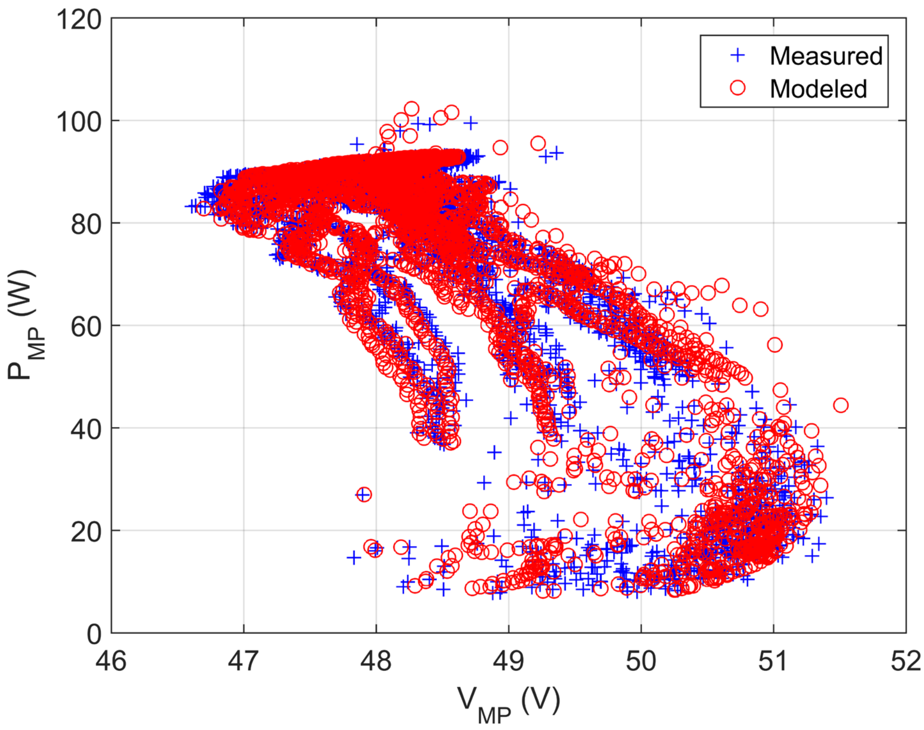


Figure 28. Measured and modeled DC power and DC voltage for First Solar FS-387 module.

We note a small systematic dependence of the residual for DC current on effective irradiance, and an increase in the variation of the residual with increasing effective irradiance (Figure 31, top panel). We also observe different ranges of the DC voltage residual for different effective irradiance values (Figure 31, bottom panel). Due to the correlation between effective irradiance and cell temperature, the residual in DC voltage similarly changes with cell temperature. Consequently, we construct an empirical model for the de-trended DC current and voltage residuals by partitioning effective irradiance (expressed in unit of ‘suns’) into three intervals:

$$E_1 = \{0 \le E \le 0.5\} \tag{39}$$

$$E_2 = \{0.5 < E \le 0.9\} \tag{40}$$

$$E_3 = \{0.9 < E\} \tag{41}$$

For each interval we de-trended the DC current residuals by subtracting a line fit to the data as depicted in Figure 31. We then constructed an empirical distribution for the de-trended residuals. We also constructed an empirical distribution for the DC voltage residuals for each interval without applying any de-trending. CDFs for the resulting current and voltage distributions are displayed in Figure 32 and Figure 33, respectively. When propagating uncertainty, we sample each distribution independently and assume no temporal correlation between sampled values.

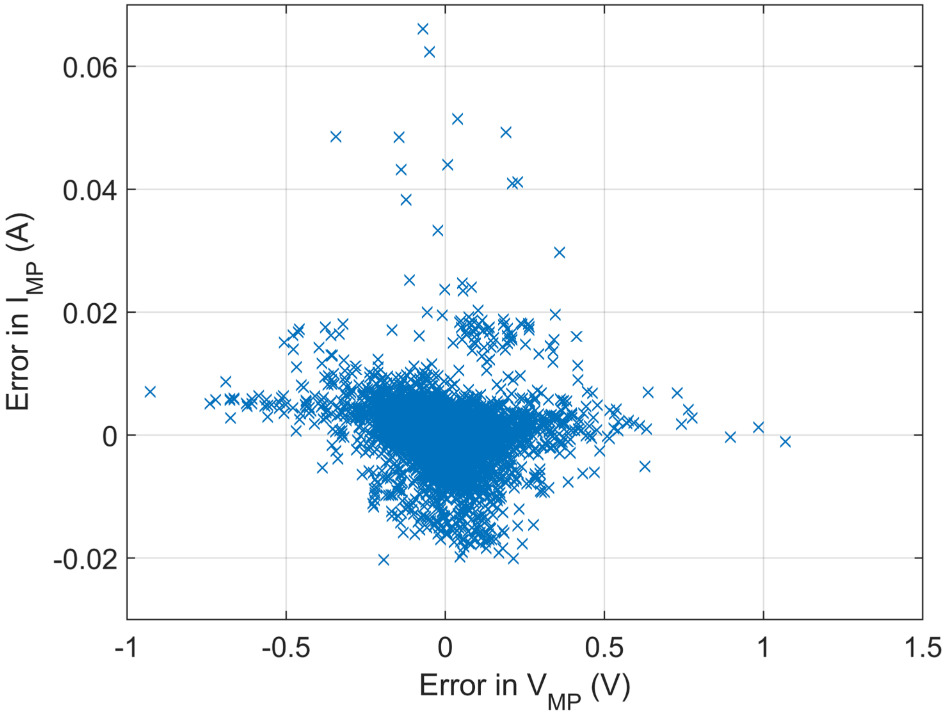


Figure 29. Residual for predicted DC current and DC voltage for First Solar FS-387 module.

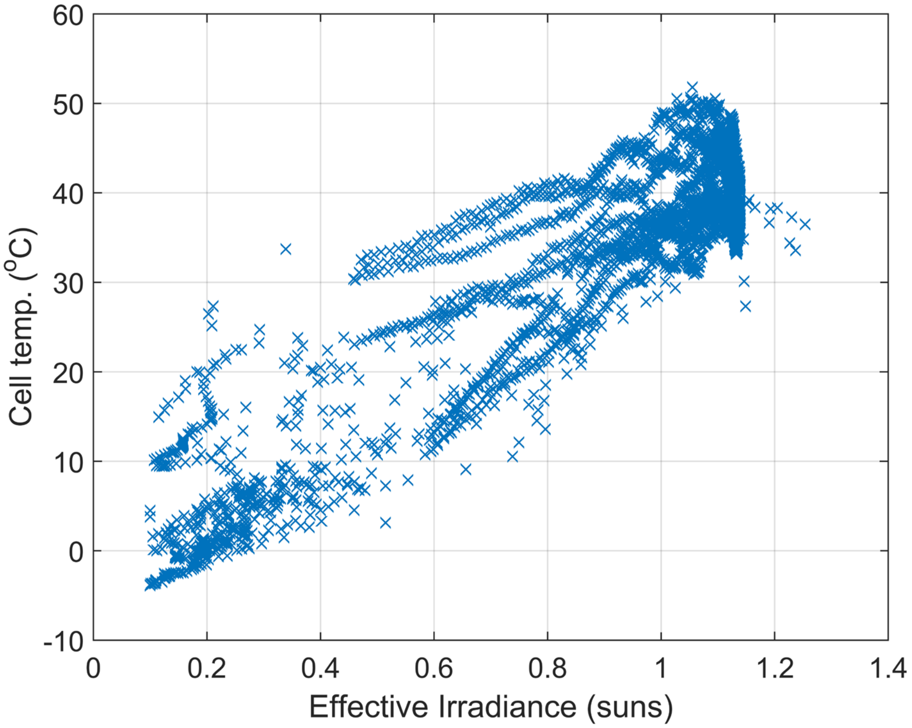


Figure 30. Correlation between effective irradiance and cell temperature for First Solar FS-387 module.

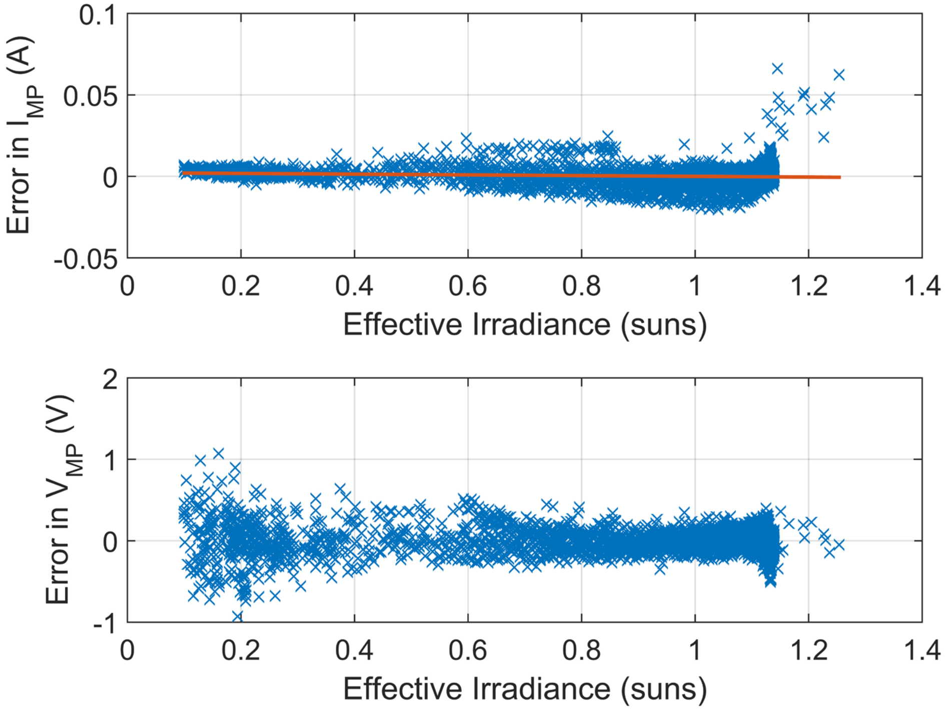


Figure 31. Correlation between effective irradiance and residual for DC voltage and current for First Solar FS-387 module.

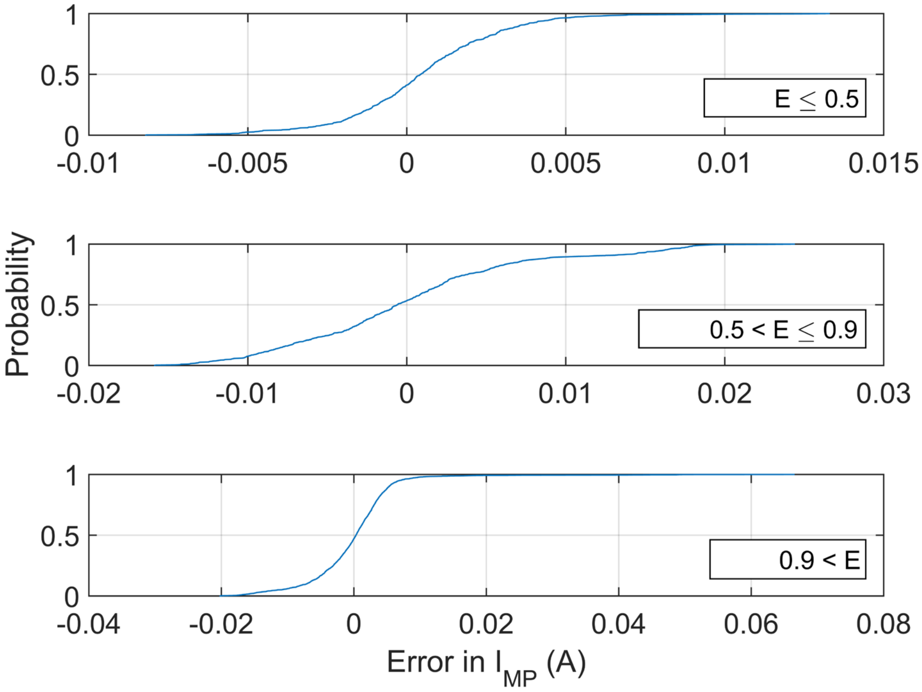


Figure 32. Distributions of de-trended DC current residual for First Solar FS-387 module.

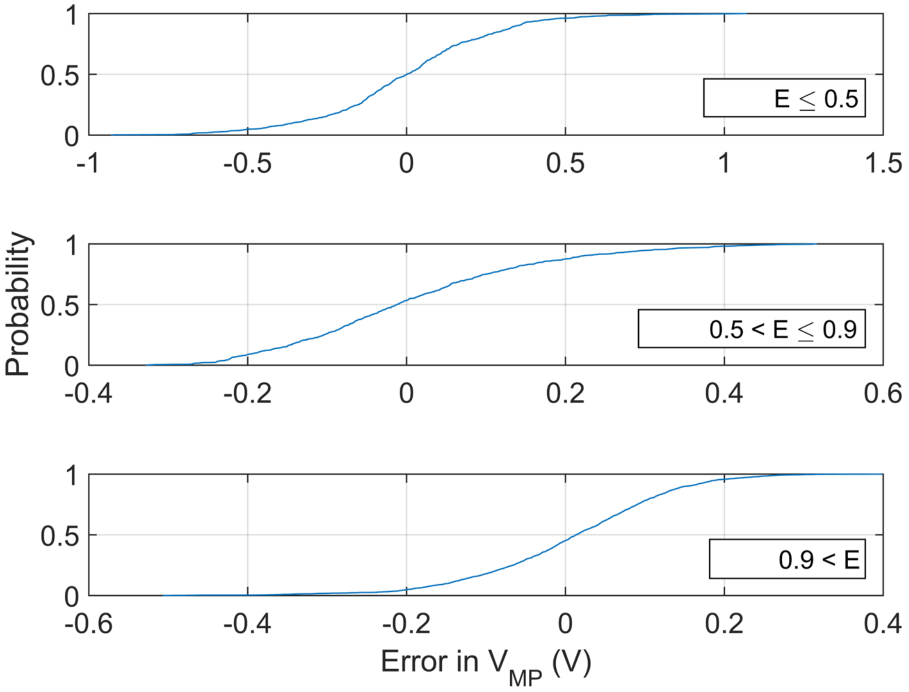


Figure 33. Distributions of DC voltage residual for First Solar FS-387 module.

3.5. Array DC Loss

The model for array DC loss (i.e., the aggregate of mismatch and MPPT loss) requires an estimate of the average loss per module $\Delta P(t)$ (Eq. (24)). To quantify $\Delta P(t)$ we followed a technique proposed by Jantsch et al. [24]. We use measured performance of a PV array at the PSEL with 48 First Solar FS-387 modules in four parallel strings of 12 modules each, connected to (and controlled by) a Fronius IG-TS 5.0 inverter. A single First Solar FS-387 reference module is operated separately and adjacent to the array. Both the array and the reference module are held at maximum power by the inverter. The I-V curve of the reference module, is measured at one-minute intervals simultaneous with measurement of the inverter operating voltage and input current. These data were collected for 20 February through 15 March 2014, a period which included a wide range of sky conditions from clear to overcast.

By dividing the inverter operating voltage by twelve (the number of modules in each string), we compute the voltage of an "average" array module. For the same instant in time, we obtain the maximum power voltage and current of the reference module by finding the maximum power point in the corresponding I-V curve.

Figure 34 shows the average array module voltage plotted against the reference module maximum power voltage for a clear day. In Figure 34 the quantities μ and σ are respectively the average and standard deviation of the difference between the array average and reference module voltages. If the reference module and all the array modules were completely identical and there were no error in the MPPT algorithm of the inverter, these data would be expected to lie along the 1:1 line, which is shown in the figure for reference. We note that the data do lie generally along a line of unity slope, but there is a constant difference of nearly one volt between the average array module voltage and the reference module maximum power voltage. Data from other days show a similar pattern, generally following a line with unity slope offset from the 1:1 line by about one volt. The offset results primarily from the intrinsic difference between the reference module and the array's 'average' module. Between days the primary difference is in the variance of the data, which is much higher on cloudy days than on clear days resulting from the inverter's less efficient tracking of the array MPPT during rapidly changing irradiance conditions.

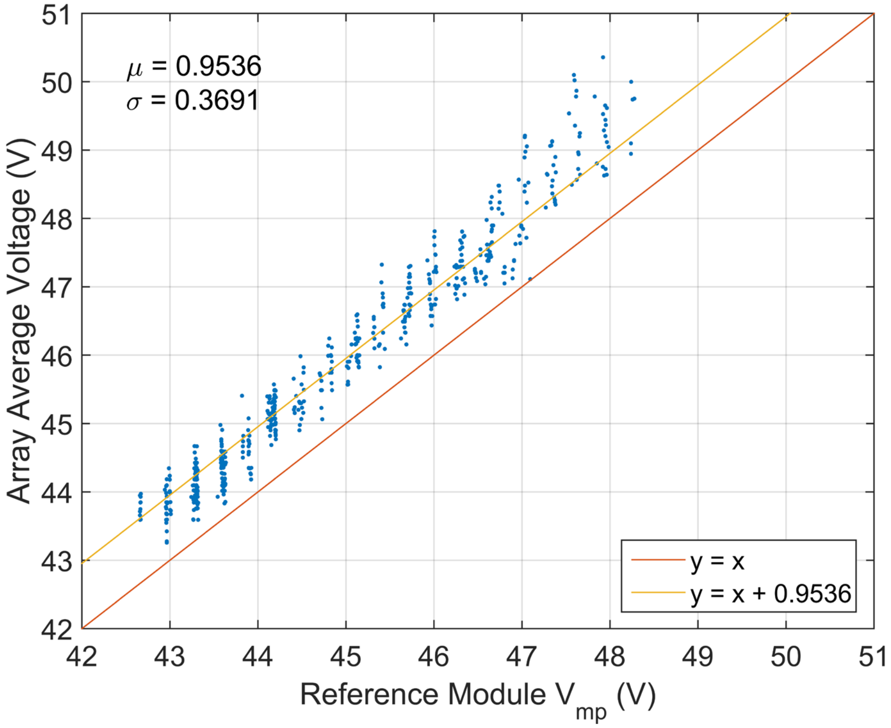


Figure 34. Comparison of array average module voltage to reference module maximum power voltage.

Assuming that the array and the reference module are uniformly illuminated on clear days, we attribute the voltage offset illustrated in Figure 34 to systemic differences between the average array module and the reference module. We average this voltage difference over all the clear days, denoting the result as ΔV_{offset}, and use ΔV_{offset} to estimate the effective module DC power $P_{DC,eff}(t)$ for each array measurement as follows. We first compute the effective reference module voltage $V_{DC,eff}(t)$ corresponding to the measured array voltage:

$$V_{DC,eff}(t) = \frac{V_{array}(t)}{N_S} - \Delta V_{offset}\,. \tag{42}$$

We then use the reference module I-V curve at time t to find the effective module current $I_{DC,eff}(t)$ corresponding to $V_{DC,eff}(t)$, as depicted in Figure 35. We then compute $P_{DC,eff}(t)$ as the product of $V_{DC,eff}(t)$ and $I_{DC,eff}(t)$, and the instantaneous power loss $\Delta P(t)$ as the difference between the reference module maximum power and $P_{DC,eff}(t)$, as specified in Eq. (24).

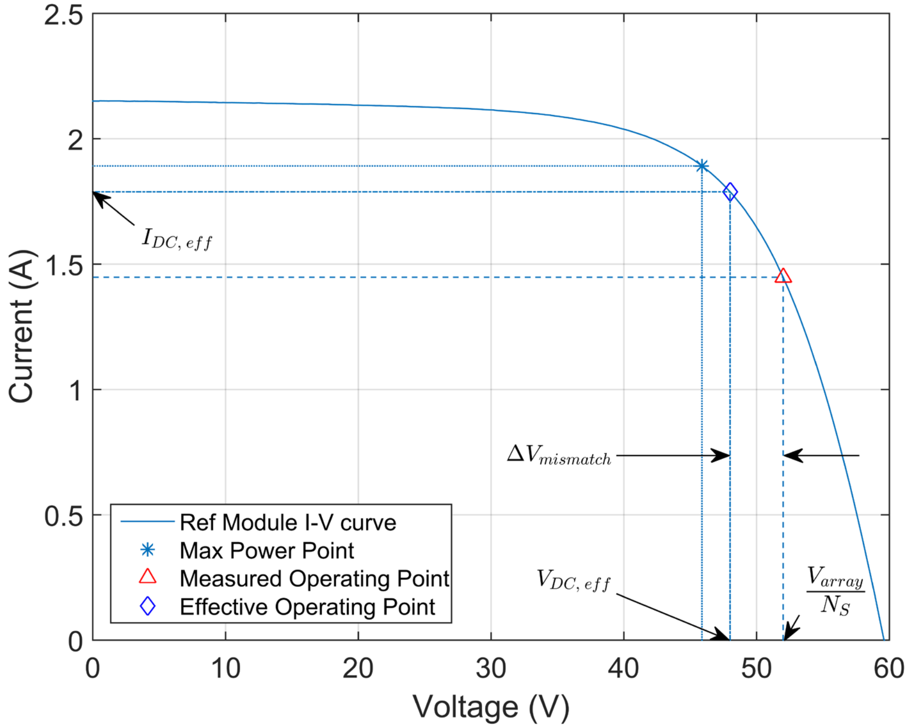


Figure 35. Calculation of module operating current and output power from array operating voltage. The array operating voltage and ΔV_{offset} are exaggerated here for clarity.

Analysis of array DC losses observed in our measured data showed a tendency towards roughly constant daily loss rates, with rates that vary between days. Figure 36 illustrates the cumulative array DC loss for several days with different degrees of variability in irradiance. On inspection of the irradiance profiles for the days represented in Figure 36, we found that each day's loss rate was generally correlated with the degree of variability in the irradiance that day. Based on this finding, we decided to represent array DC loss using a daily constant rate indexed by sky condition as determined by the variability of irradiance.

We quantify variability in irradiance by the method in [28], which characterizes a day's irradiance in terms of two indices: clearness and variability. Clearness index *CI* is the ratio of total measured daily insolation to total clear-sky insolation. This index is near zero on an overcast or rainy day (little light received), and approaches a value of 1.0 on a clear day. Variability index *VI* is the ratio of the path length of the measured GHI curve to that of the theoretical clear-sky GHI curve [28]. This index approaches a minimum value of 1.0 on clear days and increases as the magnitude and frequency of changes in irradiance increase. However, *VI* may also approach 1.0 on overcast days, when the increase in path length due to frequent but small amplitude variations in the irradiance curve is offset due to receiving less light in the day. For highly variable days, *VI* is typically much larger than 1.0. Figure 37 plots *VI* against *CI* values for the days during which the array and representative module were monitored. Using this graph, together with time-series plots of measured GHI for these days, we assigned each day to one of four categories based on their *VI* and *CI* index values (see Table 1).

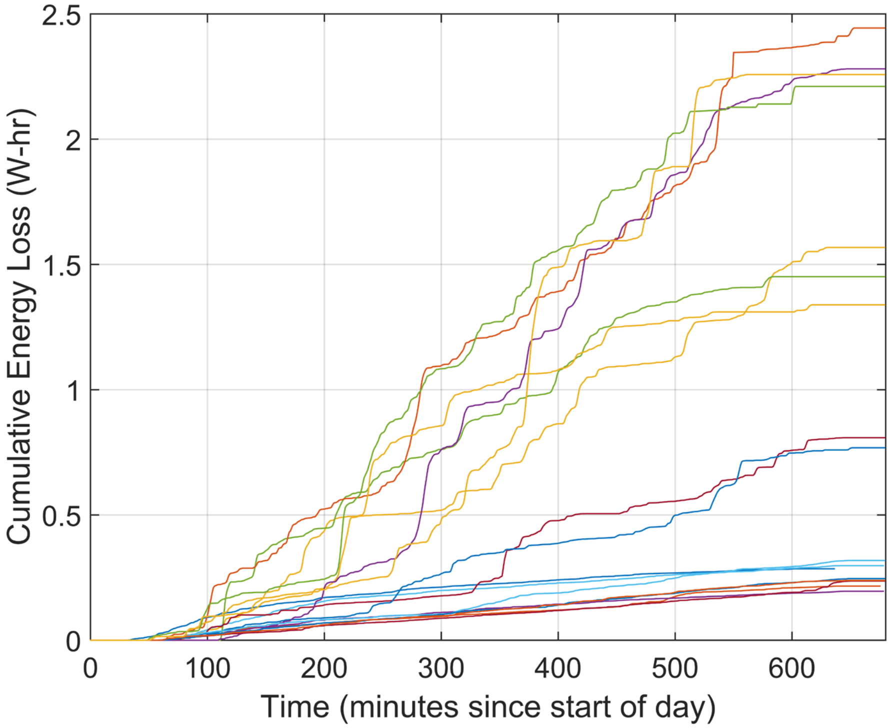


Figure 36. Single-module cumulative array DC loss during single days.

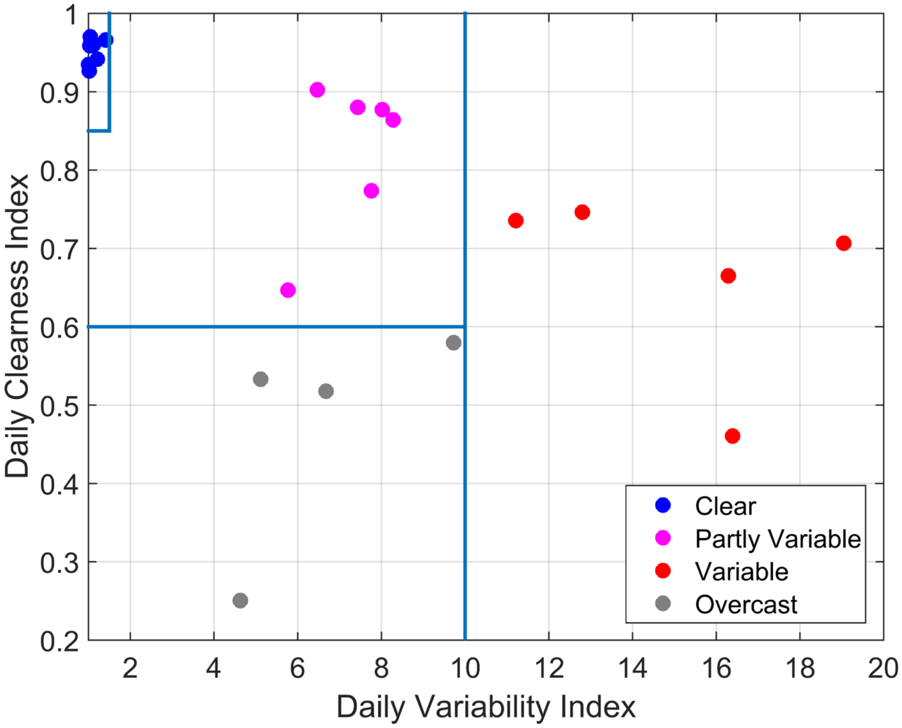


Figure 37. Clearness and Variability index values 20 Feb – 15 Mar 2014, Albuquerque, NM

Table 1. Rules for grouping days based on Variability Index *VI* and Clearness Index *CI*

Day Type	*VI/CI Range*
Clear	$VI < 1.5$ and $CI > 0.85$
Partly Variable	$VI < 10$ and $CI > 0.6$ and not Clear
Variable	$VI \geq 10$
Overcast	$VI < 10$ and $CI \leq 0.6$

For each day d, we estimate the daily average array DC loss rate $\overline{\Delta P}(d)$ as

$$\overline{\Delta P}(d) = \frac{1}{t_N - t_0} \int_{t_0}^{t_N} \Delta P(t)\,dt \cong \frac{1}{t_N - t_0} \sum_{k=0}^{N-1} \frac{\Delta P(t_k) + \Delta P(t_k)}{2} (t_{k+1} - t_k). \tag{43}$$

where $t_0, \ldots, t_k, \ldots t_N$ are the times at which values of $\Delta P(t_k)$ are available. Grouping the days into the categories indicated in Table 1, we arrived at a set of values $\left\{\overline{\Delta P}(d)\right\}$ for each type of day, shown in Figure 38. These results are consistent with observed behavior of MPPT algorithms: tracking errors are generally minimal on clear days, but increase as variability in irradiance increases. String mismatch loss results in a positive minimum value for each type of day. Loss rates may be lower on overcast days due to lower net irradiance received.

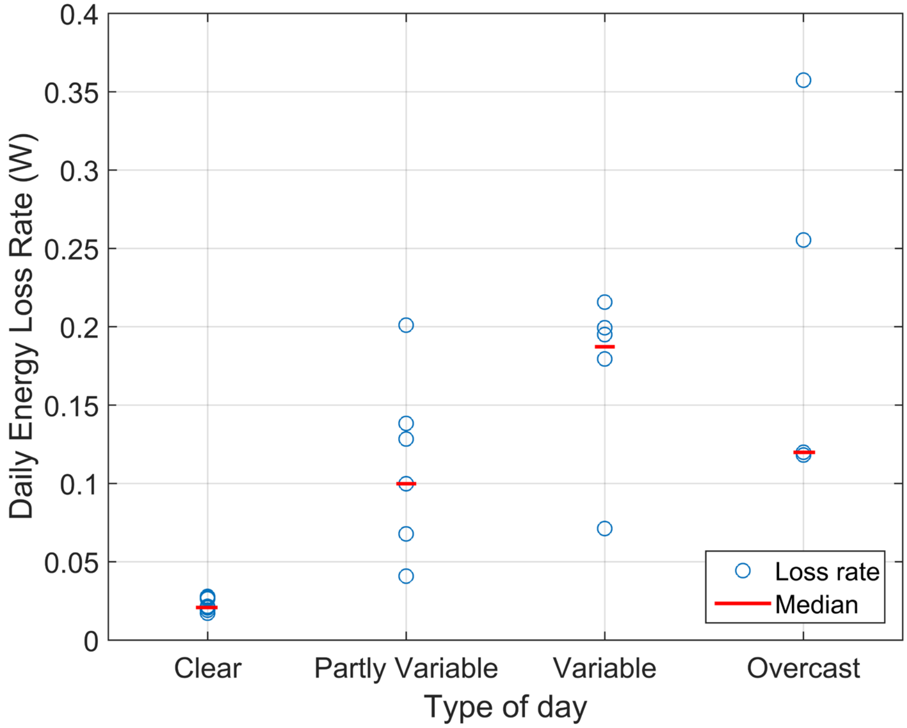


Figure 38. Daily array DC loss rates due to string mismatch and MPPT losses.

Array DC power loss rate, denoted by $\Delta P_{DC}(t) = \Delta P_{DC}(t, CI(t), VI(t))$, is modeled by $\overline{\Delta P_{50}}(t)$, i.e., the median of $\left\{\overline{\Delta P}(d)\right\}$, for each type of day $d(t)$, where the type of day is determined by $CI(t), VI(t)$ as described above. Uncertainty in $\Delta P_{DC}(t)$ is quantified by an empirical distribution for each type of day constructed from the data underlying Figure 38. A time series of $\Delta P_{DC}(t)$ is independently determined for each realization of the system model by randomly sampling the empirical distribution once for each day.

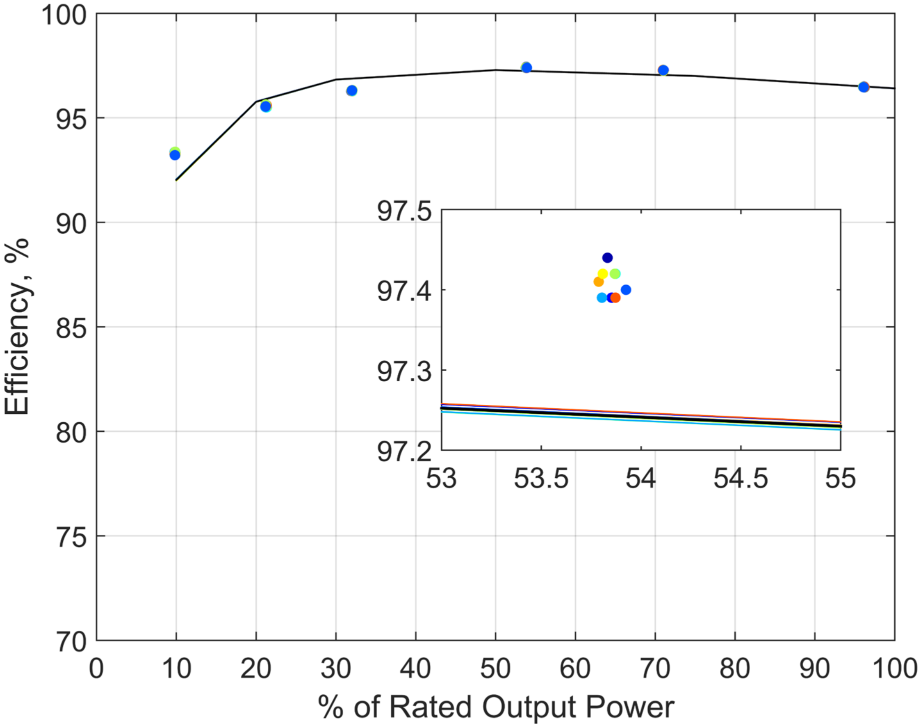
3.6. AC Power

For DC to AC conversion we use the inverter model described by Eq. (27) and Eq. (28). We quantify uncertainty using test data reported by the California Energy Commission (CEC) for a 250 kW SC250U inverter manufactured by SMA Technologies. The specification sheet provides data from ten replicated tests, each of which consists of measuring conversion efficiency while the operating the inverter to produce six different output power levels at each of three input voltage settings, yielding a total of eighteen measurements of conversion efficiency per test replicate.

We estimate a base inverter model by fitting the model equations (Eq. (27) and Eq. (28)) to the aggregate data from all ten test replicates. We characterize the uncertainty in this base model by a set of ten alternative models which result from independently fitting the model equations to the data for each test replicate.

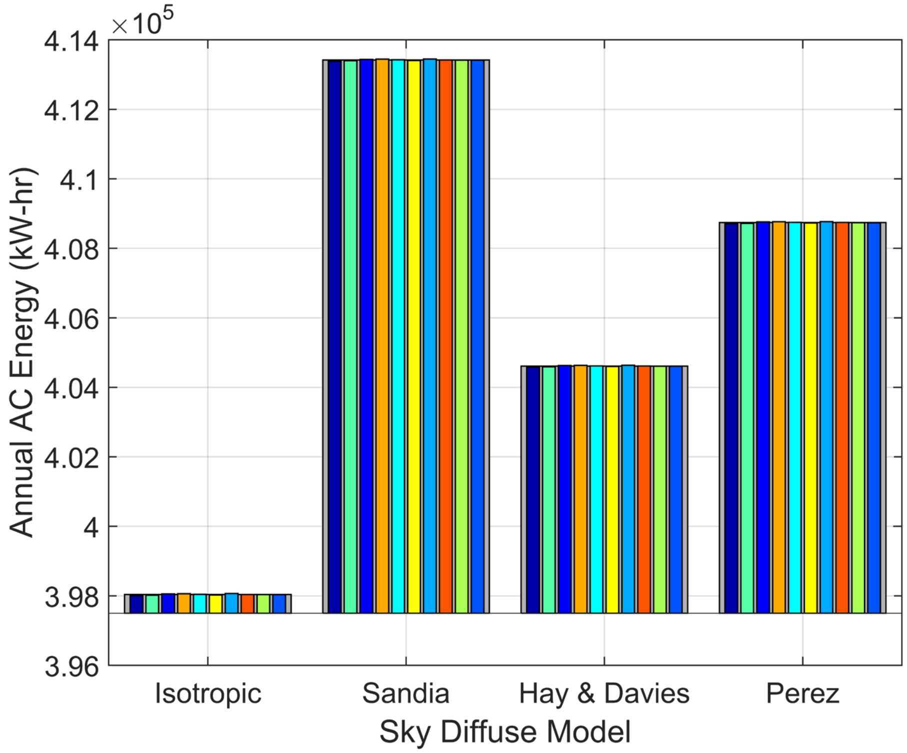
Figure 39 depicts a subset of the CEC data along with efficiency predictions for the base and alternative models. The figure shows that the uncertainty in the inverter model is small relative to the bias in the model. Measured efficiency (dots) shows very low variability which leads to extremely small differences between the models (lines). The differences between the replicated tests and the resulting differences between the models are only apparent at the scale of the figure inset. However, both the base and alternative models depart from the measured data indicating a bias in the model, perhaps arising from a structural deficiency which prevents greater accuracy in the representation of measured inverter performance.

The extremely low uncertainty in the inverter model is further illustrated in Figure 40. We estimate the total annual AC energy output for each alternative inverter model by propagating measured irradiance (GHI, DNI, DHI) through the chain of base models up to and including the array DC loss model, and use the results as input to each of the alternative inverter models (narrow colored bars) and the base inverter model (broad grey bar). The greatest difference among inverter models is extremely small—around 0.015%—compared with the greatest difference in annual POA insolation due to the sky diffuse models (about 4%).



Dots represent test data; lines represent base model (black) and alternative models (colored).

Figure 39. Comparison of Inverter Models (nominal voltage case).



Narrow bars represent models fit to data from a single test. Broad bars represent model fit to data from all tests.

Figure 40. Uncertainty in Annual AC Energy for Inverter Models.

4. UNCERTAINTY ANALYSIS

4.1. Uncertainty Propagation

Conceptually, uncertainty about a model or the model's inputs implies that the model's outputs are also uncertain. Here we employ a Monte Carlo technique to generate a sample representing the distribution of uncertainty in each model's output, using the characterization of uncertainty for each model derived from the model's residuals.

The true value of the quantity being predicted by a model is not known. The model's output provides a baseline estimate of this unknown value. We also have information about the possible errors in each model's prediction in the form of a distribution for each model's residuals. By sampling the residuals and combining the sample with each model's baseline estimate, we generate a sample of the true (unknown) quantity predicted by each model. For example, consider Step 1, Estimation of POA irradiance. The baseline estimate of POA irradiance, $\widehat{G}_{POA}(t)$, is the result from a POA model. The true value of POA irradiance, denoted by $G_{POA}(t)$, is unknown, but we have a distribution for the POA model's residual $\delta_{POA}(t|p_{POA})$ and an equation relating these quantities:

$$\delta_{POA}(t|p_{POA}) = \frac{\widehat{G}_{POA}(t) - G_{POA}(t)}{G_{POA}(t)} \tag{44}$$

We then regard $\delta_{POA}(t|p_{POA})$ as a random variable, sample a value for $\delta_{POA}(t|p_{POA})$, and estimate the true (unknown) value for POA irradiance as:

$$G_{POA}(t) = \frac{\widehat{G}_{POA}(t)}{1 + \delta_{POA}(t|p_{POA})} \tag{45}$$

Thus, we obtain a sample for the true value of $G_{POA}(t)$. Except as noted below, analogous methods are used to generate samples for the true values for other modeled quantities. These samples are passed from one modeling step to the next to compute a sample of AC output from the PV system.

Sampling the distributions of model residuals must account for any observed correlations between the residuals and the model's inputs. For example, the residual in the Isotropic sky diffuse model for estimating POA irradiance from GHI, DNI and DHI exhibits systematic variation over a range of solar zenith angles (see Figure 3). Accordingly, we construct uncertainty distributions that account for these correlations.

Calculating PV system output over time inherently involves models with time series inputs, e.g., GHI. Consequently, the sampled model residuals are themselves time series and must reflect appropriate temporal correlations. We address temporal correlations in a rather simplified manner that will tend to overstate the influence of an uncertain time series input on the model's output. For a small number of distinct conditions, e.g., clear or cloudy skies, morning or afternoon hours, we first judge whether the time series of model residuals exhibits any significant temporal correlation. When a model residual exhibits temporal correlation for a given condition, we assume that time series values remain perfectly correlated until the condition changes. When a model residual shows little or no temporal correlations during a given condition, we randomly and independently sample the model residual at each time step.

In the uncertainty propagation, Step 1, Estimation of POA irradiance, produces a sample for the true (unknown) value of $G_{POA}(t)$ from measured GHI, DNI and DHI. The subsequent Step 2, Estimation of Effective Irradiance, requires as input beam and diffuse POA irradiance (E_b and E_{diff} respectively). We translate the sample of $G_{POA}(t)$ to samples of $E_b(t)$ and $E_{diff}(t)$ by using Eq. (29), Eq. (30) and measured DNI. In other words, we regard $E_b(t)$ as known exactly without error and regard all of the error in $G_{POA}(t)$ as arising from the diffuse POA irradiance model.

Steps 3 through 5 proceed as described in Section 2, with each step accepting as input the sample of output from earlier steps, and the ouput from each step being modified by a random sample from the step's error distribution (as described in Section 3). At Step 6, Estimation of AC Power, rather than an empirical distribution of error in the model output, we have instead a set of ten equally likely alternative models for the inverter's performance. We propagate uncertainty in the inverter model by randomly selecting one of these alternative models for each realization of the system model.

4.2. Uncertainty Analysis Results

Using measured weather data (i.e., GHI, DNI, DHI, wind speed and ambient temperature) collected at Albuquerque NM from October 2013 through September 2014, we applied the methods outlined above to generate a total of four samples of possible output from a hypothetical PV system consisting of 277 parallel strings, each string comprised of nine First Solar FS-387 modules, connected to an SMA SC250U inverter. Each sample uses a different POA irradiance model, and comprises 100 time series, one year in length, of one-minute array AC power, which are reduced to time series of daily AC energy.

Figure 41 shows the results obtained when using the isotropic sky model. In Figure 41, a single red curve shows the CDF resulting from the baseline estimate of daily AC energy. The range of variation in daily energy in this curve is primarily determined by the variation in daily insolation over the year. The group of blue curves comprises 100 CDFs for daily AC energy that result from the uncertainty propagation. The separation between the red curve and the group of blue curves indicates a bias on the order of 3% of daily energy that is incurred by using the isotropic sky model in combination with the other component models. The variation among the blue curves represents the range of uncertainty in daily energy that results when uncertainty in each component model is propagated through the complete modeling chain.

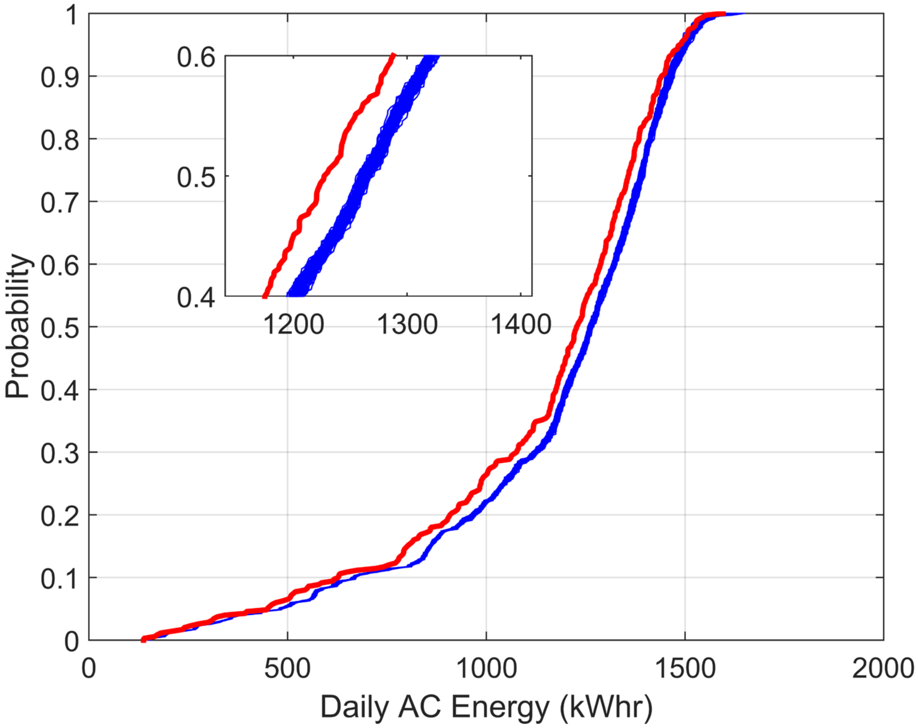


Figure 41. Distribution of Daily AC Energy using the Isotropic Sky Model

Figure 42 shows the graph from Figure 41 alongside similar plots of daily AC energy using the three other POA irradiance models (the subsequent component models are not varied). Figure 42 shows that:

- For each POA irradiance model, the blue CDFs are tightly grouped, indicating that the overall variation in predicted daily energy is relative small (on the order of 1%).
- The greatest offset between the red and blue CDFs is observed for the isotropic sky and the Hay and Davies models, indicating that the default parameter values for these models result in POA irradiance predictions that are systematically biased when compared with onsite measurements.

These observations are similar to those noted at the module DC power level in the previous study [1], which suggested that bias in predicted daily energy output can be attributed to bias in the POA irradiance models. This suggestion is supported by comparison of the daily AC power distributions in Figure 42 with similar plots of predicted daily POA insolation shown in Figure 43, in which similar, model-related biases are apparent.

Figure 44 displays the annual POA insolation measured at Albuquerque, NM during the period simulated in the uncertainty propagation, along with the baseline estimate and the histogram of values from each sample for annual POA insolation for each POA irradiance model. As in the previous study, this figure confirms that model baseline estimates can be substantially different from measured values, that biases tend to be systematic for each POA model, and that applying residuals to the model baseline estimate generally obtains a value similar to the measured annual insolation.

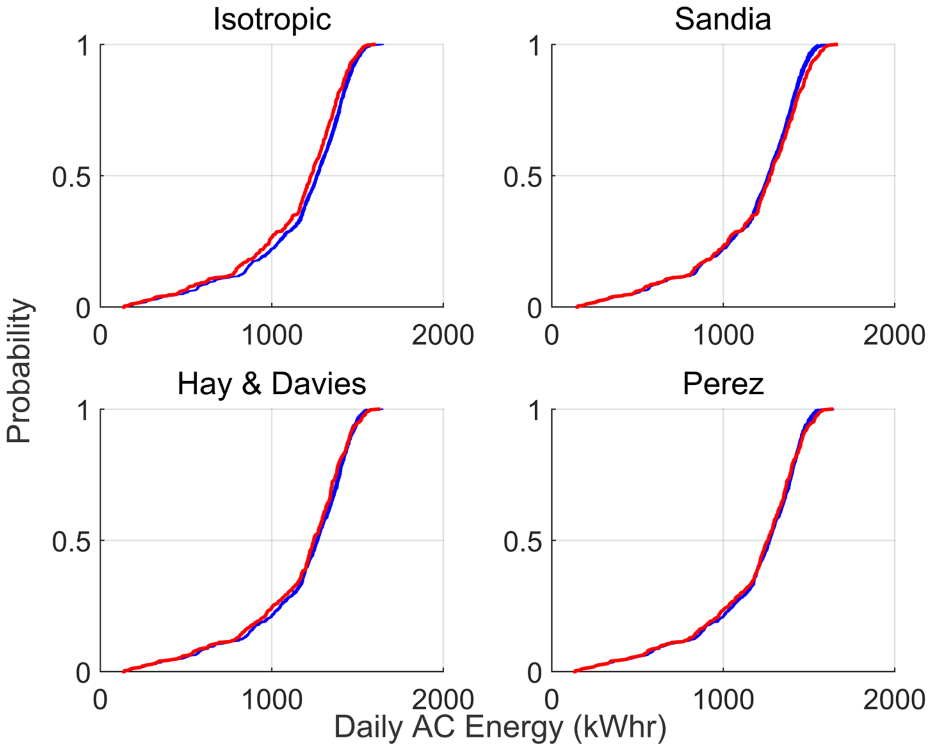


Figure 42. Distributions of Daily AC Energy

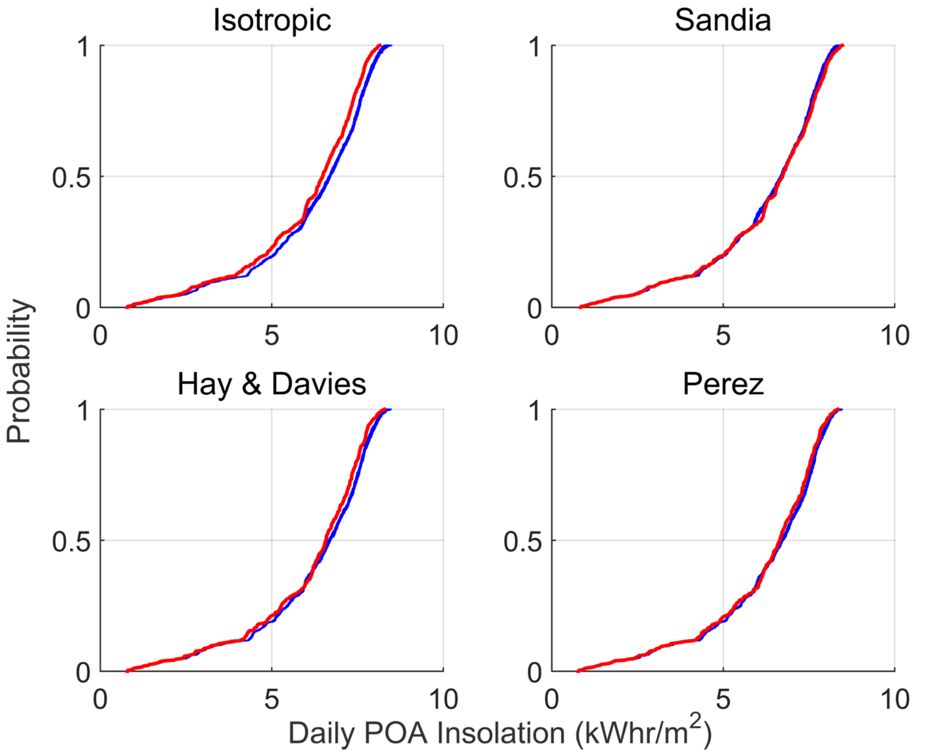


Figure 43. Distributions of Daily POA Insolation

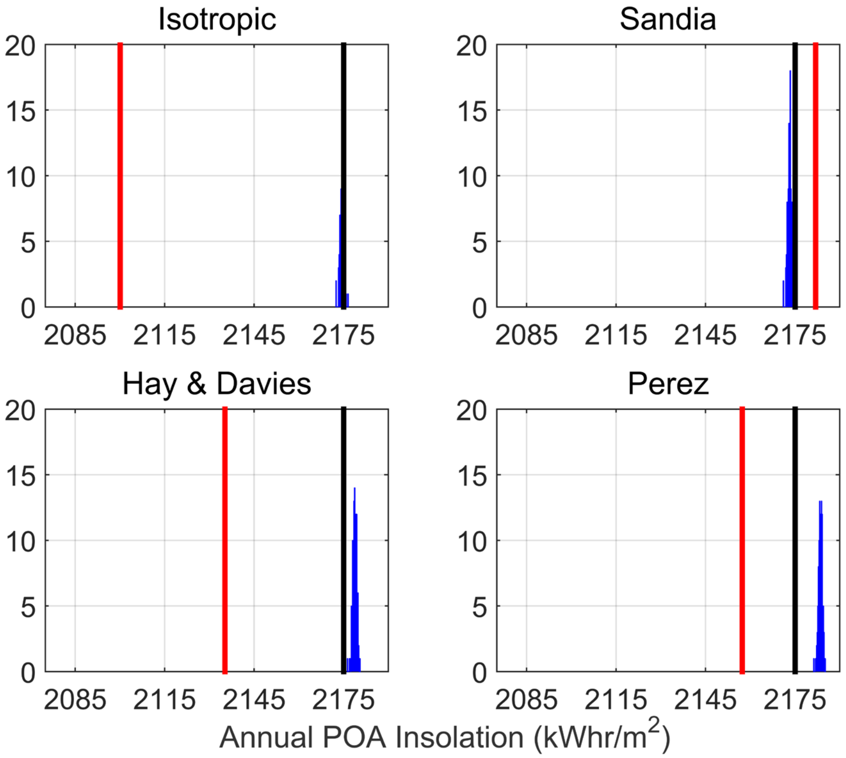


Figure 44. Distributions of Annual POA Insolation

Note: black = measured, red = model baseline estimate; blue = sample histogram.

Besides looking at distributions of annual POA insolation estimates, the previous study also examined the residual in monthly POA insolation predicted by each model. Results of repeating this analysis with the weather data used in the present study are shown in Figure 45, where we note biases consistent with those in the previous report. In addition, we also observe two fairly distinct seasonal patterns in these biases: the shapes of the isotropic sky and Sandia model curves are almost identical, as are those of the Hay and Davies and Perez models. This similarity was also evident, but not noted in the previous report.

The next stage in the modeling chain after POA is effective irradiance. Figure 46 and Figure 47 show the distributions of daily and annual effective insolation, respectively. We observe that the amount of uncertainty after the effective irradiance model is not much different than after the POA model, as the group of the family of blue curves is still very tight. Furthermore, the distance between each group and the red curve representing the baseline model appears to change only very slightly from the POA to the effective irradiance model steps. Close comparison of Figure 47 and Figure 44 confirms that the effective irradiance model step alters the overall bias in the results slightly: in all four models, the baseline estimate for effective irradiance is closer to the measured value than in the POA stage.

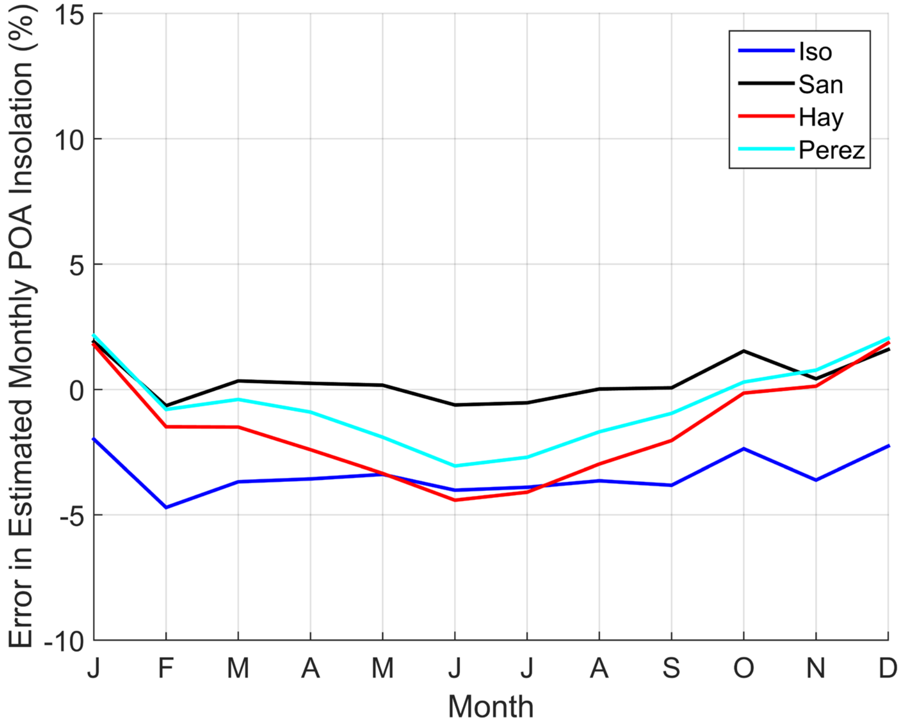


Figure 45. Residuals for monthly POA insolation for Albuquerque, NM.

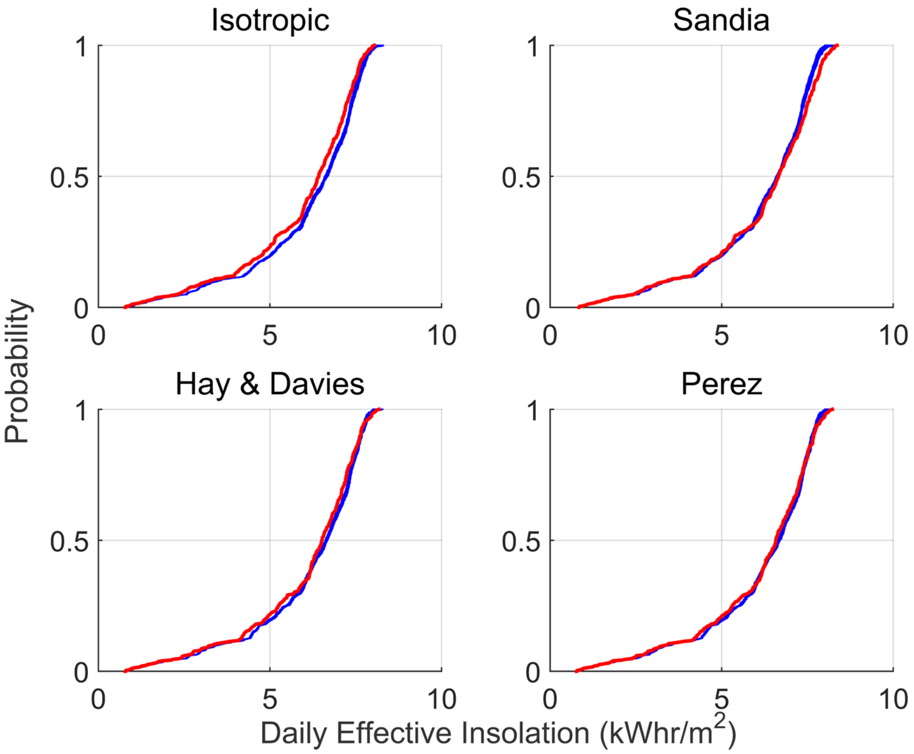


Figure 46. Distributions of daily effective insolation

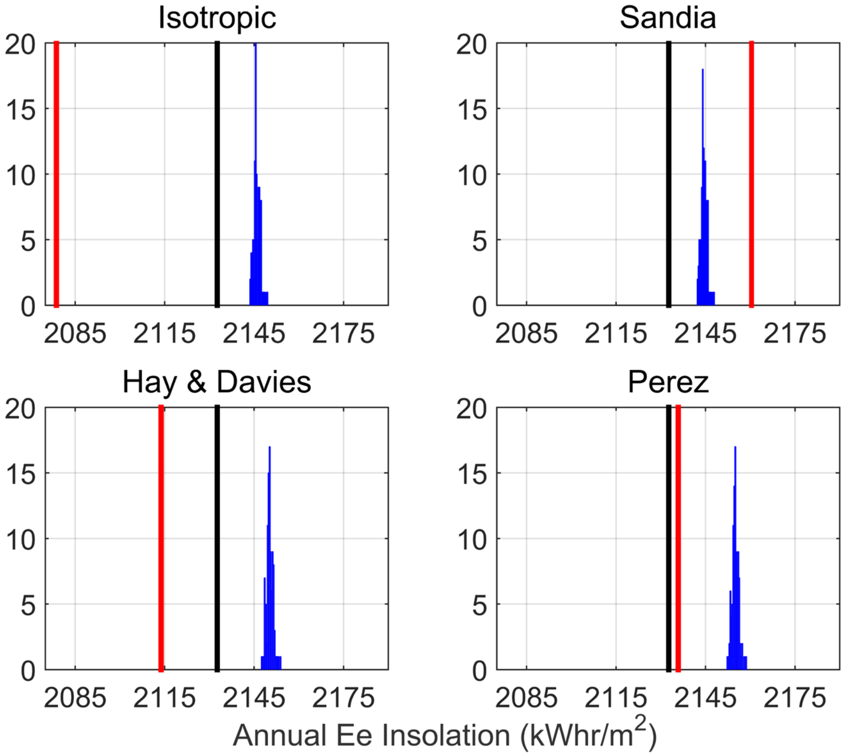


Figure 47. Distributions of annual effective insolation

Figure 48 and Figure 49 show the distributions of daily module DC power output before and after the application of array DC losses, respectively. These figures follow the same general patterns as in the POA and effective irradiance figures discussed previously. Overall, the difference between the baseline estimate and the random realizations appears to be somewhat smaller after these modeling steps than after the effective irradiance modeling step.

Finally, Figure 50 shows the annual AC output energy predictions for the hypothetical array. Above each sky diffuse model label, the wide gray bar represents the baseline estimate of inverter output energy, obtained by propagating the weather data through the chain of models. The narrow colored bars represent the 100 sample elements obtained by applying model uncertainty at each step of the modeling chain. These narrow bars have the same pattern within each sky diffuse group because we used the same set of random numbers for a given realization in each sky diffuse model. The colors of the bars in each group indicate the ten realizations of the inverter model that represent uncertainty in the DC-to-AC conversion modeling step.

Several observations can be noted from inspection of Figure 50:

- For each sky diffuse model, the maximum difference between the realizations is on the order of 1MW-hr, or 0.25% of the average annual energy output of approximately 400 MW-hr. This is consistent with the very close grouping in the daily AC energy output distributions among the random realizations noted in Figure 42.

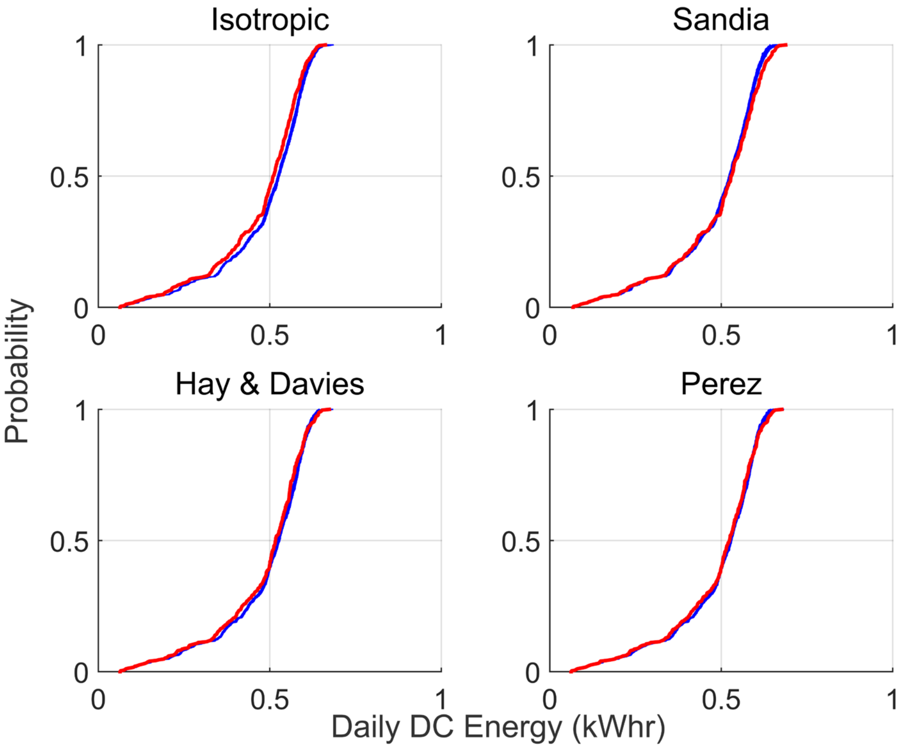


Figure 48. Distributions of daily module DC output

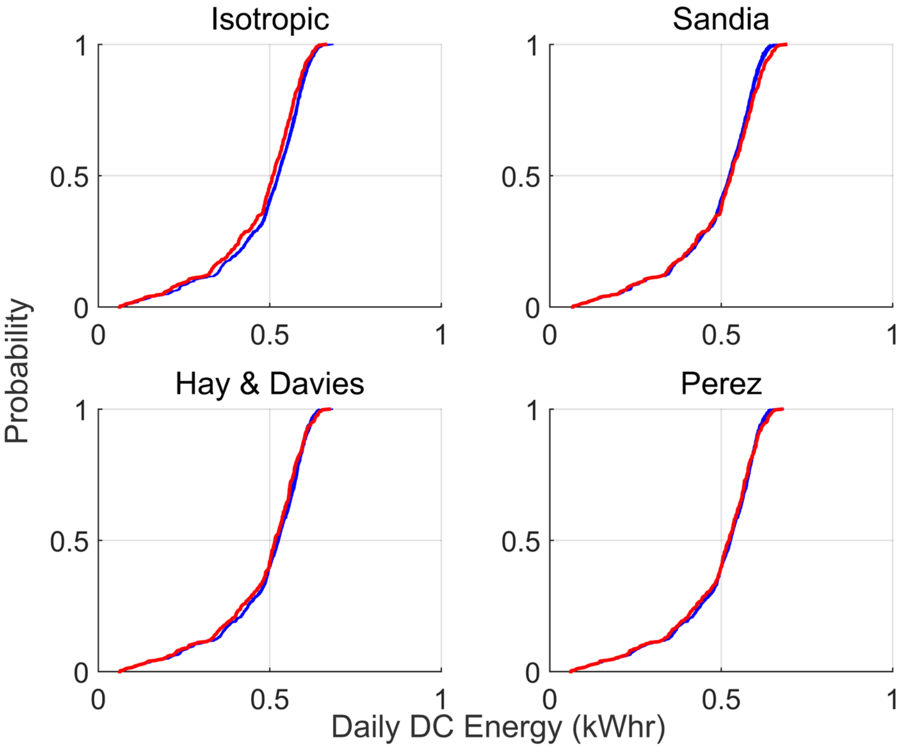
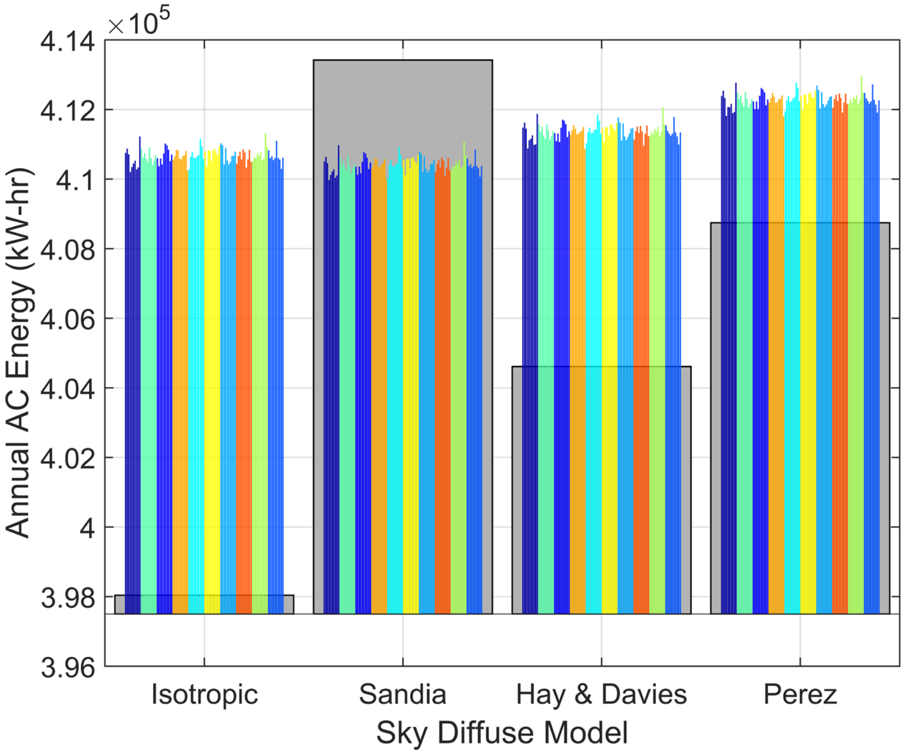


Figure 49. Distributions of daily module DC output with array DC loss applied.



Broad bars = baseline estimates; narrow bars = realizations; colors identify inverter models.

Figure 50. Annual AC output for the baseline model and the sample elements.

- Comparison of the various colored groups within each sky diffuse model indicates that the variation between the ten alternative inverter models appears to be minimal. This is consistent with the close grouping of the baseline and alternative inverter models noted previously in Figure 39.
- The maximum difference between the four baseline estimates is approximately 15.4 MW-hr, approximately 3.8% of the average of the four estimates. This difference is the same as the relationship between the maximum difference and average of the four baseline annual POA insolation estimates (red lines in Figure 44). In fact, all the relative differences between the four baseline estimates are about the same at the inverter stage as they are at the POA irradiance stage.
- Within each sky diffuse model group, the relationships between the sample elements and the baseline estimate of the annual AC energy appear similar to the same relationships between the baseline and sample elements for annual POA insolation.
- The differences between the four sets of realizations are much smaller than the differences between the four baseline estimates. These relationships are also evident in the distributions of annual POA and effective insolation (the blue histograms in (Figure 44 and Figure 47, respectively).

These observations support our method of ‘correcting’ model error by applying sampled residuals, in that we observe closer agreement among the four sets of realization than among the four baseline estimates. Moreover, consistency from one modeling step to the next also indicates that the primary sources of uncertainty, i.e., the residuals with the greatest effect, are in the early modeling steps: translation to POA irradiance and conversion from POA irradiance to effective irradiance.

5. SENSITIVITY ANALYSIS

5.1. Sensitivity Analysis Methods

Given the results of an uncertainty analysis, a *sensitivity analysis* examines the relationship between uncertainty in models or model inputs and the resulting uncertainty in the modeling outcomes. The basic model output at each time step of our analysis is AC power. We quantify uncertainty in AC power in a manner similar to that used to quantify model output uncertainty at each modeling step. We first compute a baseline estimate of AC power by using the baseline models at each step without applying values for each model's residual. We obtain a sample of residuals for AC power by subtracting this baseline estimate from each of the 100 realizations for AC power obtained by applying the sampled residual at each modeling step. Finally, we examine the relationship between the AC power residuals and the samples for model residuals by means of correlations and scatterplots.

Formally, we calculate a difference for AC output by $\Delta P_{AC}(t)$:

$$\Delta P_{AC}(t) = \hat{P}_{AC}(t) - P_{AC}(t) \tag{46}$$

where $P_{AC}(t)$ is the baseline estimate of AC power, obtained from the chain of models without applying residuals or using the alternate inverter models. Specifically, the baseline estimate $P_{AC}(t)$ is obtained by setting $\delta_{POA}(t|p_{POA}) = 1$, $\delta_E(t|p_E, q_E) = 1$, $\varepsilon_{TC}(t|p_{TC}) = 0$, $\varepsilon_{DC}(t|p_{VDC}) = 0$, $\varepsilon_{DC}(t|p_{IDC}) = 0$, $\Delta P_{DC}(t) = \overline{\Delta P_{50}}(t)$ and using the base inverter model.

We then perform correlations at each time step between ranked values for $\Delta P_{AC}(t)$ and ranked values for each of $\delta_{POA}(t|p_{POA})$, $\delta_E(t|p_E, q_E)$, $\varepsilon_{TC}(t|p_{TC})$, $\varepsilon_{DC}(t|p_{VDC})$, $\varepsilon_{DC}(t|p_{IDC})$, $\Delta P_{DC}(t)$ and an index g_{Inv} which ranks the alternate inverter models in order of increasing annual energy. The index g_{Inv} is computed by applying the alternate inverter models to the output of the array DC loss step for the base models, summing and then ranking the annual energy.

We use stepwise regression to build a sequence of regression models for $\Delta P_{AC}(t)$ using $\delta_{POA}(t|p_{POA})$, $\delta_E(t|p_E, q_E)$, $\varepsilon_{TC}(t|p_{TC})$, $\varepsilon_{DC}(t|p_{VDC})$, $\varepsilon_{DC}(t|p_{IDC})$, $\Delta P_{DC}(t)$ and g_{Inv} as predictors. The first model uses the single predictor that provides the best predictions (among all single predictor models) for $\Delta P_{AC}(t)$; the second model uses the first predictor plus one additional predictor; and so forth. The order in which the predictors are selected for the sequence of regression models indicates the strength of correlation between a predictor and $\Delta P_{AC}(t)$.

Intuitively, one expects to see AC power to increase with increasing POA irradiance, increasing effective irradiance, decreasing temperature, increasing DC voltage or DC current, decreasing array DC loss and increasing AC power conversion. However, for residuals expressed as ratios (e.g., $\delta_{POA}\left(t \middle| p_{POA}\right)$) we can instead find a negative correlation between AC power and the sampled values for residuals because of the manner in which the residuals are quantified. For example, we observe negative correlation between the POA model residual $\delta_{POA}\left(t \middle| p_{POA}\right)$ and the deviation in $\Delta P_{AC}\left(t\right)$ because there is an inverse relationship between $\delta_{POA}\left(t \middle| p_{POA}\right)$ and $G_{POA}\left(t\right)$ (see Eq. (45)). To permit an intuitive interpretation of our results, we switched the signs of several correlation coefficients.

5.2. Sensitivity Analysis Results

Figure 51 displays the stepwise rank regression coefficients for daily AC energy for our hypothetical array in Albuquerque, NM, using the isotropic sky diffuse model. We use rank regression to remove the effects of the widely varying magnitudes among predictor variables. The predicted variable $\Delta P_{AC}\left(t\right)$ is the time series of the difference between each realization's daily AC energy and the daily AC energy for the base realization. Because daily AC energy for the base realization is fixed, greater values of $\Delta P_{AC}\left(t\right)$ correspond to increased AC energy output. The predictor variables must be summarized to daily quantities which reflect any monotonic relationships with $\Delta P_{AC}\left(t\right)$. In general, we summed the model residuals for each realization, reasoning as follows:

- For POA irradiance, we summed the sampled model residuals (i.e., $\delta_{POA}\left(t \middle| p_{POA}\right)$ in Eq. (6)) for each day, reasoning that the sum of residuals is monotonically related with AC energy.
- For effective irradiance, we summed the sampled model residuals (i.e., $\delta_{E}\left(t \mid p_{E}, q_{E}\right)$ in Eq. (16)).
- For cell temperature, we summed the sampled model residuals (i.e. $\varepsilon_{TC}\left(t \middle| p_{TC}\right)$ in Eq. (20)), reasoning that the sum is monotonically related with total daily cell temperature, which in turn is monotonically related to AC energy.
- For DC voltage and DC current, we summed the sampled model residuals $\varepsilon_{VDC}\left(t \middle| p_{DC}\right)$ and $\varepsilon_{IDC}\left(t \middle| p_{DC}\right)$ reasoning that these sums are monotonically related to AC energy.
- For array DC output (i.e., after array DC losses), the uncertainty propagation (see Section 3.5) produces a daily DC loss rate which we use directly in the regression, reasoning that the DC loss rate is monotonically related with AC energy.
- Uncertainty in AC output is characterized by a set of alternate inverter performance models (see Section 3.6). We created an index an index g_{Inv} which ranks the alternate inverter models in order of increasing annual energy and thus is monotonically related to $\Delta P_{AC}\left(t\right)$.

After performing rank regressions at a daily time scale we summed over months and repeated the regression (Figure 55).

5.2.1. Daily AC Energy

Figure 51 clearly shows that uncertainty arising from the POA irradiance and effective irradiance modeling steps dominate the uncertainty in the deviation of daily AC energy from its baseline value. Generally, uncertainty in effective irradiance shows the greatest influence on uncertainty in daily AC energy, although uncertainty in POA irradiance shows a stronger correlation during certain periods (July/August and November/December). Uncertainty arising from the array DC loss model appears to have a secondary but still significant effect. Uncertainty arising from each of the other models (cell temperature, DC voltage and DC current, and inverter) is relatively insignificant.

Figure 52 compares the regression analysis results for daily AC energy for all four POA irradiance models. This figure confirms that the residuals arising from the POA and effective irradiance models are dominant regardless of which sky diffuse irradiance model is used. Figure 53 and Figure 54 present the stepwise regression results ordered by the variability and clearness indices, respectively. These figures confirm that the alternating order of the coefficients associated with the effective irradiance (Ee) and POA irradiance models in Figure 52 result from seasonal dependencies. The daily variation in the magnitude of these coefficients likely results from our stochastic models for the model residuals, which imperfectly preserve the autocorrelations observed in the daily data.

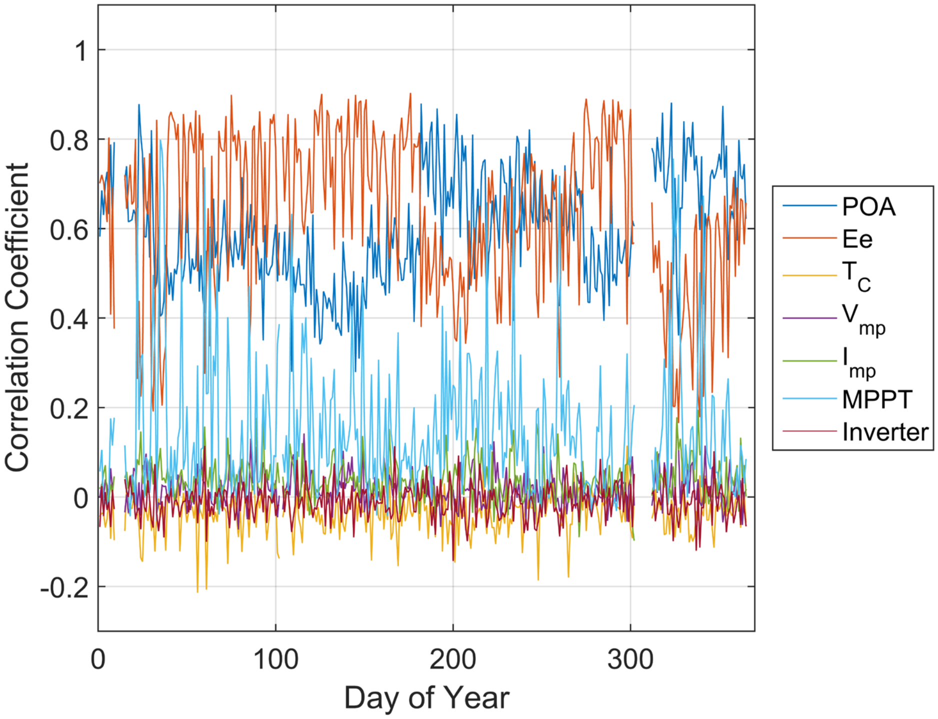


Figure 51. Stepwise rank regression coefficients for daily energy (isotropic sky diffuse model).

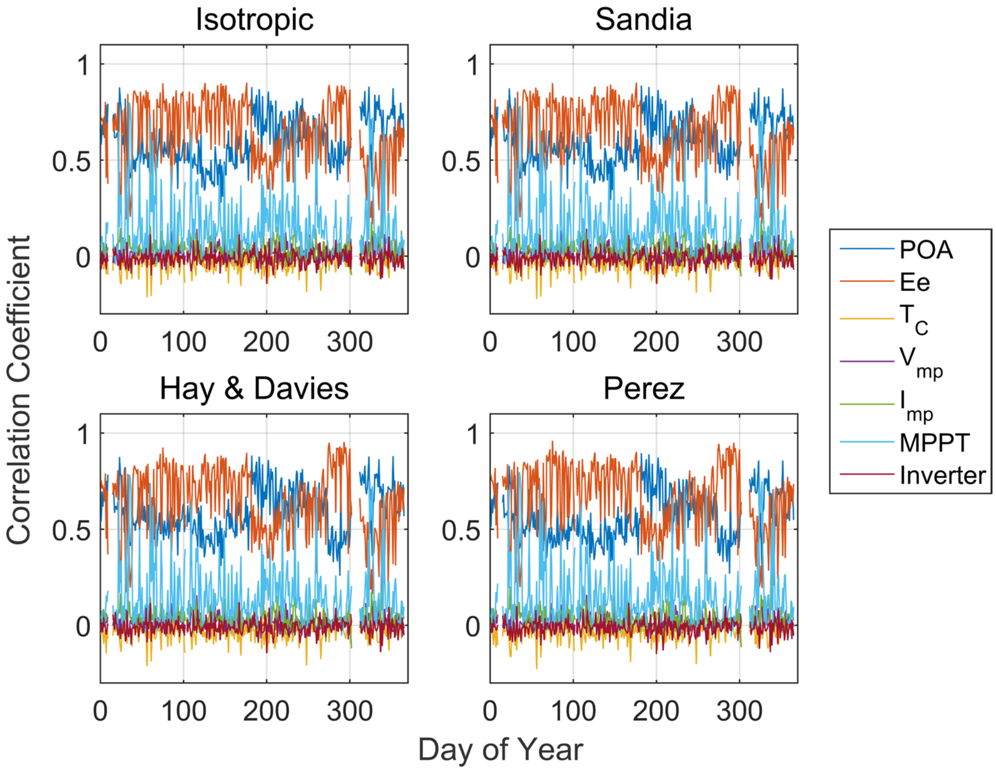


Figure 52. Stepwise rank regression coefficients for daily energy (all 4 POA models).

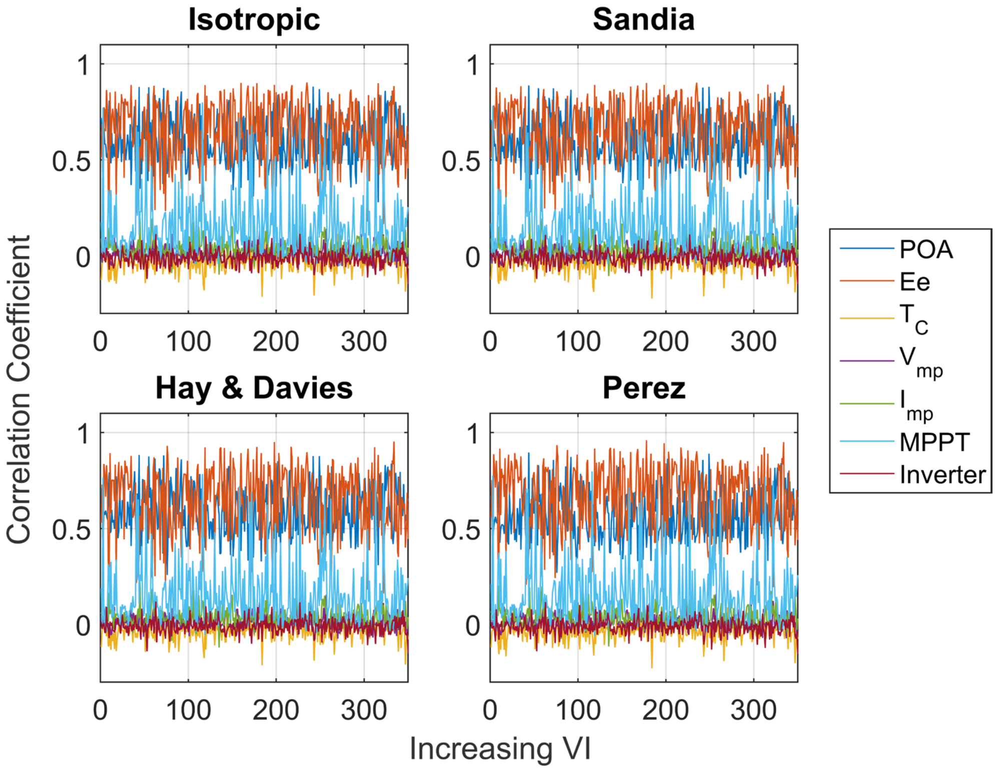


Figure 53. Stepwise rank regression coefficients sorted by variability index.

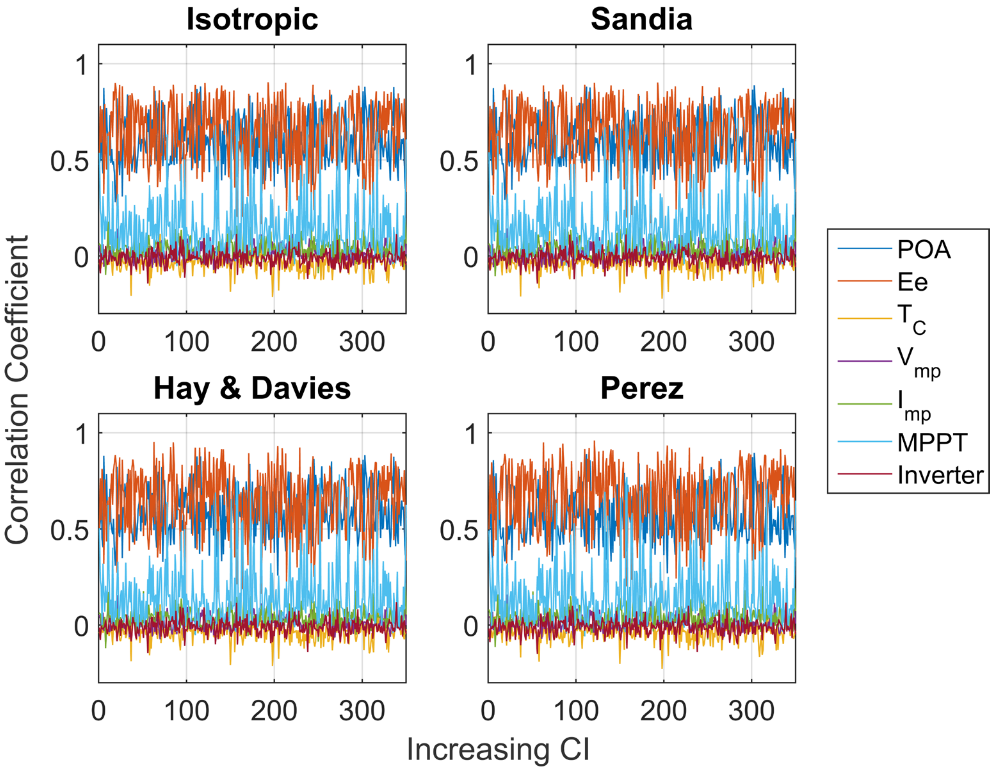


Figure 54. Stepwise rank regression coefficients sorted by clearness index

5.2.2. Monthly AC Energy

Figure 55 shows the stepwise rank regression analysis results for monthly AC energy for each POA irradiance model. Table 2 lists the correlation coefficients found to have statistical significance in the stepwise regression process and the associated R^2 values achieved as each variable is added to the model. POA and effective irradiance are ranked either #1 or #2 in all cases, while array DC loss (indicated as MPPT in Figures 51-53 and Table 2) occurs frequently in the #3 position. Together with the relatively high values of R^2 (the lowest exceeds 0.86), the consistency of these rankings confirms the dominant contribution of POA and effective irradiance models to uncertainty in the end-to-end model and, suggest that array DC loss may warrant additional attention.

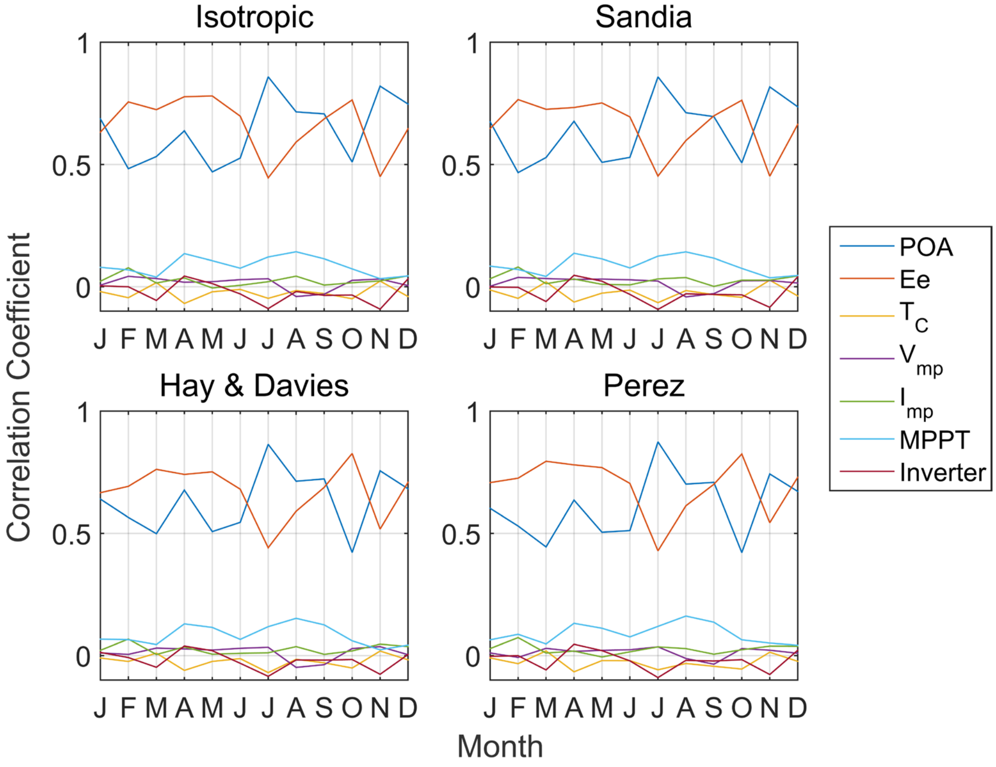


Figure 55. Stepwise rank regression coefficients for monthly energy (all 4 POA models).

Table 2. Stepwise rank regression models for montly AC energy

	Isotropic			Sandia			Hay & Davies			Perez		
	Variable	Coeff	R^2	Variable	Coeff	R^2	Variable	Coeff	R^2	Variable	Coeff	R^2
Jan	POA	0.6866	0.5372	POA	0.6741	0.5169	Ee	0.6661	0.5225	Ee	0.7074	0.5735
	Ee	0.6319	0.9288	Ee	0.6467	0.9272	POA	0.6400	0.9259	POA	0.6021	0.9306
	MPPT	0.0795	0.9351	MPPT	0.0843	0.9342	MPPT	0.0671	0.9303	MPPT	0.0644	0.9347
Feb	Ee	0.7556	0.6907	Ee	0.7651	0.7034	Ee	0.6921	0.6096	Ee	0.7255	0.6448
	POA	0.4828	0.9217	POA	0.4667	0.9185	POA	0.5651	0.9306	POA	0.5301	0.9269
	Imp	0.0776	0.9281	Imp	0.0807	0.9255	Imp	0.0675	0.9356	MPPT	0.0870	0.9351
	MPPT	0.0689	0.9329	MPPT	0.0705	0.9304	MPPT	0.0652	0.9398	Imp	0.0734	0.9404
Mar	Ee	0.7239	0.6402	Ee	0.7254	0.6434	Ee	0.7617	0.6824	Ee	0.7946	0.7295
	POA	0.5325	0.9248	POA	0.5285	0.9232	POA	0.4987	0.9270	POA	0.4445	0.9273
	Inverter	-0.0560	0.9278	Inverter	-0.0604	0.9268				Inverter	-0.0584	0.9306
Apr	Ee	0.7768	0.4696	Ee	0.7327	0.4200	Ee	0.7409	0.4262	Ee	0.7799	0.4733
	POA	0.6375	0.8828	POA	0.6765	0.8822	POA	0.6767	0.8896	POA	0.6358	0.8871
	MPPT	0.1352	0.9003	MPPT	0.1366	0.9000	MPPT	0.1296	0.9057	MPPT	0.1319	0.9037
	Tc	-0.0684	0.9048							Tc	-0.0663	0.9079
May	Ee	0.7799	0.7321	Ee	0.7514	0.6914	Ee	0.7514	0.6936	Ee	0.7689	0.6989
	POA	0.4695	0.9316	POA	0.5088	0.9265	POA	0.5072	0.9262	POA	0.5047	0.9312
	MPPT	0.1064	0.9426	MPPT	0.1137	0.9390	MPPT	0.1150	0.9391	MPPT	0.1118	0.9433
Jun	Ee	0.6973	0.6838	Ee	0.6943	0.6785	Ee	0.6804	0.6658	Ee	0.7041	0.6998
	POA	0.5263	0.9602	POA	0.5291	0.9584	POA	0.5451	0.9596	POA	0.5114	0.9585
	MPPT	0.0759	0.9657	MPPT	0.0769	0.9641	MPPT	0.0662	0.9638	MPPT	0.0763	0.9641
Jul	POA	0.8583	0.6682	POA	0.8581	0.6583	POA	0.8640	0.6705	POA	0.8742	0.6937
	Ee	0.4452	0.8631	Ee	0.4525	0.8581	Ee	0.4412	0.8615	Ee	0.4284	0.8737
	MPPT	0.1215	0.8761	MPPT	0.1248	0.8713	MPPT	0.1181	0.8738	MPPT	0.1194	0.8860
	Inverter	-0.0895	0.8839	Inverter	-0.0928	0.8797	Inverter	-0.0847	0.8808	Inverter	-0.0898	0.8939
Aug	POA	0.7148	0.5142	POA	0.7115	0.5046	POA	0.7130	0.5095	POA	0.7012	0.4876
	Ee	0.5919	0.8777	Ee	0.5980	0.8755	Ee	0.5903	0.8719	Ee	0.6126	0.8782
	MPPT	0.1427	0.8979	MPPT	0.1423	0.8956	MPPT	0.1523	0.8949	MPPT	0.1616	0.9041
Sep	POA	0.7067	0.3839	Ee	0.6987	0.3831	POA	0.7224	0.3889	POA	0.7083	0.3725
	Ee	0.6861	0.8532	POA	0.6955	0.8525	Ee	0.6879	0.8589	Ee	0.7005	0.8564
	MPPT	0.1140	0.8661	MPPT	0.1167	0.8660	MPPT	0.1255	0.8747	MPPT	0.1365	0.8749
Oct	Ee	0.7637	0.6827	Ee	0.7625	0.6876	Ee	0.8265	0.7909	Ee	0.8244	0.7924
	POA	0.5099	0.9293	POA	0.5067	0.9296	POA	0.4218	0.9466	POA	0.4211	0.9450
	MPPT	0.0730	0.9344	MPPT	0.0750	0.9350	MPPT	0.0604	0.9509	MPPT	0.0644	0.9498
							Tc	-0.0507	0.9532	Tc	-0.0549	0.9525
Nov	POA	0.8202	0.7350	POA	0.8170	0.7272	POA	0.7557	0.6459	POA	0.7427	0.6268
	Ee	0.4503	0.9241	Ee	0.4520	0.9187	Ee	0.5175	0.8989	Ee	0.5435	0.9064
	Inverter	-0.0917	0.9320	Inverter	-0.0843	0.9255	Inverter	-0.0763	0.9044	Inverter	-0.0773	0.9121
Dec	POA	0.7469	0.5017	POA	0.7357	0.4780	Ee	0.7088	0.4447	Ee	0.7290	0.4648
	Ee	0.6460	0.9175	Ee	0.6638	0.9166	POA	0.6821	0.9082	POA	0.6715	0.9134

5.2.3 Comparison with previous results

In prior work Hansen and Pohl [1] found that uncertainty in the POA irradiance and effective irradiance models had a dominant influence on the uncertainty in predicted module output DC energy. They repeated the analysis for two different module technologies in each of two locations and obtained similar results for all four cases, suggesting that ranking of important model residuals will not depend greatly on location or module technology. The analysis in [1] included several but not all of the models that are considered here: their work included models for POA irradiance, effective irradiance, cell temperature, module voltage and module current, but not for array DC losses or AC energy. In addition, the uncertainty quantification for the effective irradiance model in [1] did not fully capture the range of model error because the uncertainty was quantified from the data used to estimate the model. In this work uncertainty in the effective irradiance model was quantified using out-of-sample data.

The present study considers one of the cases addressed in [29] (i.e., a CdTe array of modules in Albuquerque, NM), but uses weather/irradiance data from a different time period. The results presented here confirm the previous conclusion that, among the models considered, uncertainty in POA irradiance and effective irradiance models have the greatest influence on the uncertainty in predicted system output. As was done in [1], this analysis excludes the influence on system output of uncertainty in the measurement of irradiance.

5.3 Discussion of omitted modeling steps

5.3.1 Uncertainty in Irradiance Measurement

Our uncertainty analysis does not include characterization of uncertainty in irradiance measurements, nor of its effect on the uncertainty in power and energy predictions. Intuition suggests that error in measured GHI and DNI conveys roughly (but not exactly) to a proportional error in POA irradiance predicted by a model. Quantification of error in measured GHI and DNI depends strongly on the type of instruments used to make these measurements and the errors can exhibit complex patterns that depend on solar angles, temperature and weather conditions (e.g., [30]. Often, error is reported only as a statistic derived from a lengthy data record; errors in the range from 2% to 8% are often cited as general guidelines for a range of irradiance instruments [31]. Due to the known complexity of the uncertainty in irradiance measurements and the intuitive understanding that errors of at least 2% will translate roughly to 2% errors in predicted power, we chose to omit treatment of error in irradiance measurements from our analysis. We note that current research efforts are intended to develop a much deeper understanding of the performance and uncertainty of available irradiance instruments [32].

5.3.2 Albedo

Albedo characterizes the fraction of irradiance incident on the ground that is reflected away. Prediction of POA irradiance includes a term quantifying the ground-reflected irradiance (Eq. (32)). Our uncertainty quantification for POA irradiance (Section 3.1) and uncertainty and sensitivity analyses are conditional on a constant, assumed albedo value $a = 0.2$, which is typical for PV systems [27]. In our prior work [1] we demonstrated that uncertainty in albedo over the range from 0.1 to 0.3 has a minor effect on POA insolation and concluded that uncertainty in albedo would not be a major influence on uncertainty in AC energy.

5.3.3 Soiling

Accumulated soil on the face of a PV module reduces the irradiance reaching the module's cells to less than the POA irradiance. Characterization of power loss due to soiling as a function of time is complex (e.g., [33], [34]); soiling losses rates are dependent on material composition, local weather, PV array construction and solar angles; monthly loss rates exhibit significant variation [35]. Few models are available (e.g., [36]) and the available models generally comprise a constant soiling loss rate without a more detailed accounting for variation in factors that affect soiling.

Measurements of soiling rates exhibit relatively large uncertainties. Soiling rates can be estimated by comparing short-circuit current measured for two modules (or reference cells), one kept clean and one allowed to soil, operated outdoors side-by-side. During our work to characterize uncertainty, we obtained one such data set comprising measured short-circuit current for an appropriate pair of modules over a period of several months. However, the soiling rates in Albuquerque, NM are relatively low, and the systematic trend that would indicate soiling was generally obscured by variability in the measured short-circuit current. The variability could be the result of variation in solar spectrum for which we attempted to account using the effective irradiance model described in Section 2.4, Step 2. Uncertainty in this effective irradiance model may be as great as a few percent of measured current (e.g., Figure 22) even during clear sky conditions. Such variability is likely to obscure the effect of the minor amount of soiling occurring at the test location.

It is generally understood and accepted that the effect of soiling is to reduce the irradiance reaching a module cells. Careful laboratory measurements have shown a linear relationship between soiling areal density and the reduction in irradiance in which the slope is related to the composition of the soil [37]. Thus it stands to reason that uncertainty in soiling will convey proportionally to uncertainty in electrical current and thus to power. Because uncertainty in measurement of soiling losses is relatively high (e.g., [35]) compared with uncertainty in other models we have considered, it is reasonable to conclude that uncertainty in soiling losses is likely to be as influential on uncertainty in annual energy predictions as are the other models identified as influential (e.g., translation to POA irradiance and conversion from POA irradiance to effective irradiance.)

6. CONCLUSIONS

We have completed an uncertainty and sensitivity analysis for modeling AC energy from photovoltaic systems. We considered a single system comprising 2493 FirstSolar modules connected to a 250 kW DC to AC inverter, located at Albuquerque, NM. We quantified uncertainty in the following modeling steps:

- Translation from measured GHI, DNI and DHI to POA irradiance;
- Estimation of effective irradiance (i.e., irradiance converted to electrical current);
- Prediction of cell temperature from measured air temperature and wind speed;
- Production of DC voltage and current from the module;
- Estimation of array DC power loss due to module mismatch and to maximum power point tracking inaccuracy;
- Estimation of DC-to-AC conversion efficiency.

Due to the complexity and correlations among each model's parameters, we adopt an approach where we characterize the uncertainty in a model's output by quantifying the distribution of each model's *residual*, i.e., the difference between the model's prediction and the true value, rather than the traditional approach of quantifying uncertainty in each model's input parameters.

We found that the overall uncertainty in predicted PV system output, i.e., daily energy, to be relatively small, on the order of 1%. We considered four alternative models for the POA irradiance modeling step; and found that variance in predicted PV system output is not greatly dependent on the choice of one of these models. However, we found that all POA irradiance models exhibited a systematic bias of upwards of 4% that depends on location, and that this bias translates proportionally to predicted energy. Thus, choice of a POA irradiance model implies a bias to some degree in the predicted output, but not a greater (or smaller) variance in the predictions.

We performed a sensitivity analysis to relate uncertainty in the PV system output to uncertainty arising from each model. We found that uncertainty in the models for POA irradiance and effective irradiance to be the dominant contributors to uncertainty in predicted daily energy. Our analysis indicates that efforts to reduce the uncertainty in PV system output predictions may yield the greatest improvements by focusing on the POA and effective irradiance models.

Our analysis excludes consideration of the uncertainty in the irradiance measurements and of uncertainty in predictions of soiling losses. It stands to reason that uncertainty in irradiance measurements and in soiling convey proportionally to uncertainty in predicted system output, and thus, these uncertainties are as least as influential, if not more influential, than uncertainty in the models we have considered.

7. REFERENCES

1. Hansen, C., A. Pohl, and D. Jordan, *Uncertainty and Sensitivity Analysis for Photovoltaic System Modeling*, SAND2013-10358, Sandia National Laboratories, Albuquerque, NM, 2013.
2. Whitfield, K. and C.R. Osterwald, *Procedure for determining the uncertainty of photovoltaic module outdoor electrical performance.* Progress in Photovoltaics: Research and Applications, 2001. **9**(2): p. 87-102.
3. ANSI, *American National Standard for Expressing Uncertainty--U.S. Guide to the Expression of Uncertainty in Measurement*, 1997, ANSI.
4. Dominé, D., et al. *Uncertainties of PV Module – Long-term Outdoor Testing*. in *25th PVSEC*, 2012. Valencia, Spain.
5. Dirnberger, D. and U. Kraling, *Uncertainty in PV Module Measurement; Part I: Calibration of Crystalline and Thin-Film Modules.* Photovoltaics, IEEE Journal of, 2013. **3**(3): p. 1016-1026.
6. Müller, B., et al., *Yield predictions for photovoltaic power plants: empirical validation, recent advances and remaining uncertainties.* Progress in Photovoltaics: Research and Applications, 2015: p. n/a-n/a.
7. Hansen, C., et al. *Parameter Uncertainty in the Sandia Array Performance Model for Flat-Plate Crystalline Silicon Modules*. in *37th IEEE Photovoltaics Specialists Conference*, 2011. Seattle, WA.
8. Helton, J.C., J.D. Johnson, and W.L. Oberkampf, *An exploration of alternative approaches to the representation of uncertainty in model predictions.* Reliability Engineering & System Safety, 2004. **85**(1–3): p. 39-71.
9. Loutzenhiser, P.G., et al., *Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation.* Solar Energy, 2007. **81**(2): p. 254-267.
10. Hay, J.E. and J.A. Davies. *Calculations of the solar radiation incident on an inclined surface*. in *First Canadian Solar Radiation Data Workshop, 59*, 1980. Ministry of Supply and Services, Canada.
11. Sandia National Laboratories, www.pvpmc.org, Accessed Nov. 7, 2013.
12. Perez, R., P. Ineichen, and R. Seals, *Modeling Daylight Availability and Irradiance Components from Direct and Global Irradiance.* Solar Energy, 1990. **44**(5): p. 271-289.
13. King, D.L., E.E. Boyson, and J.A. Kratochvil, *Photovoltaic Array Performance Model*, SAND2004-3535, Sandia National Laboratories, Albuquerque, NM, 2004.
14. De Soto, W., S.A. Klein, and W.A. Beckman, *Improvement and validation of a model for photovoltaic array performance.* Solar Energy, 2006. **80**(1): p. 78-88.
15. Hughes, G.W., *Engineering Astronomy*, 1985, Sandia National Laboratories.
16. Kasten, F. and A.T. Young, *Revised optical air mass tables and approximation formula.* Applied Optics, 1989. **28**(22): p. 4735-4738.
17. Portland State Aerospace Society, *A Quick Derivation relating altitude to air pressure*, http://psas.pdx.edu/RocketScience/PressureAltitude_Derived.pdf, Accessed Nov. 7, 2013.
18. Knisely, B., et al. *Validation of IEC 61853-2 standard (Draft): Angle of incidence effect on photovoltaic modules*. in *Photovoltaic Specialists Conference (PVSC), 2013 IEEE 39th*, 2013.

19. Martin, N. and J.M. Ruiz, *Calculation of the PV modules angular losses under field conditions by means of an analytical model.* Solar Energy, 2001. **70**: p. 25-38.
20. (IEC), I.E.C., *60904-5 Ed. 2.0: Photovoltaic devices - Part 5: Determination of the equivalent cell temperature (ECT) of photovoltaic (PV) devices by the open-circuit voltage method*, 2011.
21. MacAlpine, S., *Characterization And Capture Of Photovoltaic System Losses Due To Nonuniform Conditions*, in *Department of Civil, Environmental, and Architectural Engineering*2013, University of Colorado.
22. Hohm, D.P. and M.E. Ropp, *Comparative study of maximum power point tracking algorithms.* Progress in Photovoltaics: Research and Applications, 2003. **11**(1): p. 47-62.
23. Latham, A.M., C.R. Sullivan, and K.M. Odame. *Performance of photovoltaic maximum power point tracking algorithms in the presence of noise*. in *Energy Conversion Congress and Exposition (ECCE), 2010 IEEE*, 2010.
24. Janstch, M., et al. *Measurement of PV Maximum Power Point Tracking Performance*. in *26th European Photovoltaic Energy Conference*, 1997. Barcelona, Spain.
25. King, D.L., et al., *Performance Model for Grid-Connected Photovoltaic Inverters*, SAND2007-5036, Sandia National Laboratories, Albuquerque, NM, 2007.
26. Bower, W., et al. *Performance Test Protocol for Evaluating Inverters Used in Grid-Connected Photovoltaic Systems*. 2004.
27. Lorenzo, E., *Energy Collected and Delivered by PV Modules*, in *Handbook of Photovoltaic Science and Engineering*, A. Luque, Hegedus, S., Editor. 2011, John Wiley and Sons.
28. Stein, J., C. Hansen, and M. Reno. *The Variability Index: A New and Novel Metric for Quantifying Irradiance and PV Output Variability*. in *World Renewable Energy Forum*, 2012. Denver, CO.
29. Cameron, C., J. Stein, and C. Hansen. *Evaluation of PV Performance Models and Their Impact on Project Risk*. in *1st PV Rollout Conference*, 2011. Boston, MA.
30. Gueymard, C.A. and D.R. Myers, *Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling.* Solar Energy, 2009. **83**(2): p. 171-185.
31. Badescu, V., *Modeling Solar Radiation at the Earth's Surface: Recent Advances*, ed. V. Badescu. 2008, Berlin: Springer-Verlag.
32. Driesse, A., P. Jain, and e. al. *Beyond the Curves: Modeling the Electrical Efficiency of Photovoltaic Inverters.* in *33rd IEEE Photovoltaic Specialists Conference*, 2008. San Diego, CA.
33. Sarver, T., A. Al-Qaraghuli, and L.L. Kazmerski, *A comprehensive review of the impact of dust on the use of solar energy: History, investigations, results, literature, and mitigation approaches.* Renewable and Sustainable Energy Reviews, 2013. **22**(0): p. 698-733.
34. Mejia, F.A. and J. Kleissl, *Soiling losses for solar photovoltaic systems in California.* Solar Energy, 2013. **95**(0): p. 357-363.
35. Caron, J.R. and B. Littmann, *Direct Monitoring of Energy Lost Due to Soiling on First Solar Modules in California.* Photovoltaics, IEEE Journal of, 2013. **3**(1): p. 336-340.
36. Kimber, A., et al. *The Effect of Soiling on Large Grid-Connected Photovoltaic Systems in California and the Southwest Reguion of the United States*. in *IEEE Photovoltaics Specialist COnference*, 2006. Waikoloa, Hawaii.

37. Burton, P.D. and B.H. King, *Application and Characterization of an Artificial Grime for Photovoltaic Soiling Studies*. Photovoltaics, IEEE Journal of, 2014. **4**(1): p. 299-303.

DISTRIBUTION

1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories