Which Models Matter: Uncertainty and Sensitivity Analysis for Photovoltaic Power Systems

Clifford W. Hansen and Andrew Pohl

Sandia National Laboratories, Albuquerque, NM, 87185-1033, USA

Abstract **— Predicting power for a photovoltaic system from measured irradiance requires a sequence of models, e.g.: translation of measured irradiance to the plane-of-array; estimation of cell temperature; and calculation of module electrical output. Uncertainty in predicted power arises from the aggregate uncertainty in the various models used. But which models contribute significantly, or insignificantly, to this aggregate uncertainty? We report an uncertainty and sensitivity analysis that partially addresses this question.**

At each step in the modeling process, we consider commonly-used models and quantify uncertainty in the model and its parameters by analyzing model residuals obtained by comparing model predictions to available measurements. We are not yet able to quantify uncertainty in all relevant modeling steps. We develop stochastic process models for these residuals and use Monte Carlo sampling to propagate uncertainty from step to step. Propagating uncertainty through all modeling steps, we obtain a sample of values for PV system output representing the aggregate uncertainty in this quantity. We use rank correlations to identify the relative contribution to the aggregate uncertainty that can be attributed to each modeling step. For the steps and models we consider, we find that uncertainty in plane-of-array transposition and in effective irradiance models dominates the uncertainty in energy production; uncertainty in cell temperature and in module DC output is significantly less influential.

Index Terms **— uncertainty, photovoltaic system performance, models.**

I. Introduction

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 (e.g., [1]). Uncertainty in these models and their parameters arises from a variety of sources, including measurement errors, inexact model specification, and from the necessarily finite data used to calibrate the models. Because each of the models in the sequence are uncertain, the predictions of system yield are also uncertain.

To understand the degree of confidence that can be placed in energy yield predictions, we first describe the aggregate uncertainty in the predicted energy by *uncertainty analysis*, which quantifies uncertainty in each component model, propagates uncertainty through the sequence of calculations, and obtains a distribution of energy yield predictions. We next determine by *sensitivity analysis* the extent to which uncertainty at each modeling step contributes to the aggregate uncertainty. Our results indicate which models in the sequence are important in the sense of contributing to uncertainty in energy yield predictions, and also how these models could be improved to reduce the aggregate uncertainty.

Previous analyses of the effect of uncertainty on system output (e.g., [2], [3], [4]) are dependent on the assumptions inherent in the Guide to the Expression of Uncertainty in Measurement (GUM) [5], neglect correlations among uncertain quantities (e.g., [6]), or have focused narrowly on individual model steps (e.g., [7]). Ours is the first analysis of which we are aware to quantify uncertainty empirically using time series of measured quantities, and to propagate this uncertainty through a sequence of modeling steps. However, our analysis is limited by excluding uncertainty in the measurements which underlie determination of model parameters. These uncertainties and their effects on measurements of module performance can be substantial (e.g., [8], [9]).

Our paper is organized as follows: Section II identifies the models for which we quantified uncertainty and our method for doing so. Section III outlines the techniques used to propagate uncertainty through the sequence of models to obtain a distribution of system output. Section IV summarizes the sensitivity analysis and our conclusions.

II. Quantifying Uncertainty at each modeling step

Figure 1 depicts a standard sequence of models required to estimate system output, beginning with the necessary meteorological inputs (e.g., irradiance). At each step, options for models may be available: for example, Hay and Davies [10] and Perez [11] are both popular models for estimating plane-of-array (POA) irradiance from global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI).

We quantify uncertainty in models at steps where we have concurrent measurements of the inputs and output of a model. The red boxes on Figure 1 indicate the modeling steps for which uncertainty is quantified in our analysis. Specifically we consider the following modeling steps applied to two PV systems each comprising either a single SunPower 305WHT or a single FirstSolar FS272 module:

1. Translation of GHI, DNI and DHI to POA irradiance G_{POA} by applying one of the following models to

meteorological measurements obtained in Albuquerque, NM and in Golden, CO: Isotropic sky [12]; Sandia [13]; Hay and Davies [10]; and Perez [11].

2. Transfer of G_{POA} to effective irradiance Ee (i.e., the irradiance converted to electrical current, being reduced by reflection and spectral mismatch losses) using $Ee = f_1(AM)G_{POA}$ where $f_1(AM)$ is an empirically-determined polynomial in air mass AM [14] that is determined for each module using measurements in Albuquerque, NM (see [7] for model calibration details).
3. Estimate cell temperature T_C using the expression $T_C = T_A + G_{POA}\exp(a + bWS) + \frac{G_{POA}}{1000}\Delta T$ ([14], Eq. 13) in which T_A and WS are ambient temperature and wind speed, respectively. The empirical coefficients a, b and ΔT are determined for each module using measurements in Albuquerque, NM (see [7] for model calibration details).
4. Calculate DC voltage and current using the Sandia Array Performance Model (SAPM) [14] which comprises a system of equations to predict short-circuit current I_{SC}, open circuit voltage V_{OC}, and maximum power current I_{MP} and voltage V_{MP} from Ee and T_C. The equations require a set of empirical coefficients, which are determined from measurements in Albuquerque, NM [7].

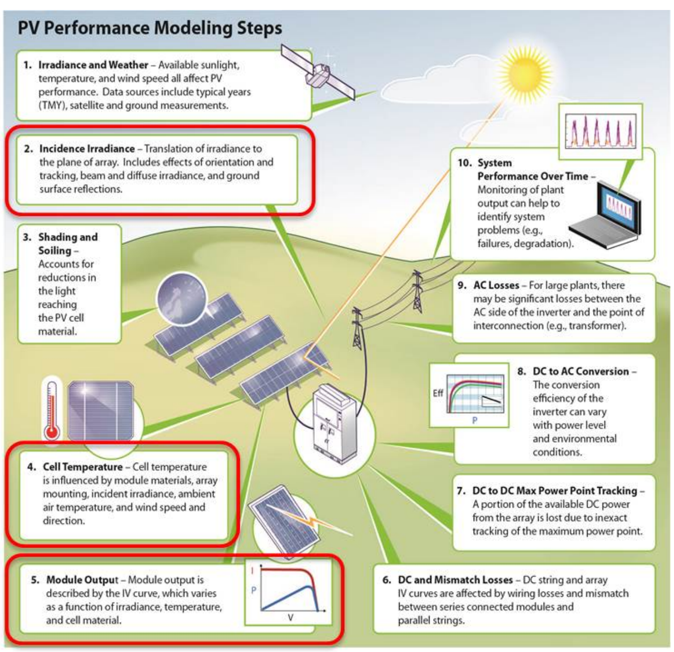

Figure 1. Standard sequence of models used to estimate PV system output (www.pvpmc.org).

We intentionally exclude measurement uncertainty from our analysis. In particular, because uncertainty in irradiance measurements will directly and proportionally transfer to uncertainty in system power output including irradiance measurement error in our analysis would likely obscure the relative importance of uncertainty in all remaining modeling steps. Our intent here is to determine the extent to which uncertainty in these other modeling steps affects uncertainty in system output predictions.

We quantify uncertainty in each modeling step by computing time series of model residuals, i.e., the difference between the model's prediction and the measured value of the predicted quantity. For example, we obtained measured DNI and DHI for 2011 in Albuquerque, NM, concurrent with measured total irradiance on a pyranometer fixed at 35° tilt, oriented southwards. We use measured DNI and DHI as input to several models for predicting plane of array (POA) irradiance (e.g., isotropic sky [12], Hay and Davies [10], Sandia [13], and Perez [11]) and computed the model residuals by subtracting the measured POA irradiance from the model predictions. Figure 2 illustrates the model residuals for POA irradiance that we obtained using the isotropic sky model; for this model, we note a strong, systematic dependence of error on both angle-of-incidence (AOI) and on season. We also found systematic dependencies on sky condition (clear or non-clear) and on time of day (AM vs. PM) (Figures 2 and 3). When these trends are removed the remaining residuals are distributed relatively uniformly across the range of AOI and are amenable to statistical modeling. Thus, we partition the data by month, time of day (AM or PM) and sky condition (clear or non-clear), empirically detrend each of these 48 data sets, and then obtain an empirical distribution for the detrended residuals. We repeat this analysis for each of the four POA models [13].

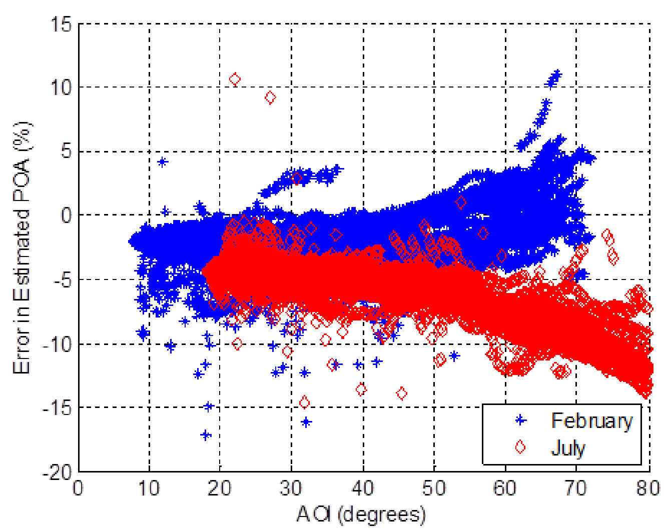

Figure 2. Dependence on season of error in POA irradiance prediction for the isotropic sky model.

As indicated in Figure 1, we similarly quantified uncertainty in the other models we consider. For the effective irradiance model, the residual is modeled with distributions conditional on sky condition and on ranges of values for air mass. The distributions that model the residual of modeled cell temperature depend on sky condition and ranges of values for wind speed. The residual for modeled DC voltage is conditional on ranges of effective irradiance but no systematic dependencies are evident for the residual of modeled DC current. These dependencies and the statistical models for the residuals are described and illustrated in [13].

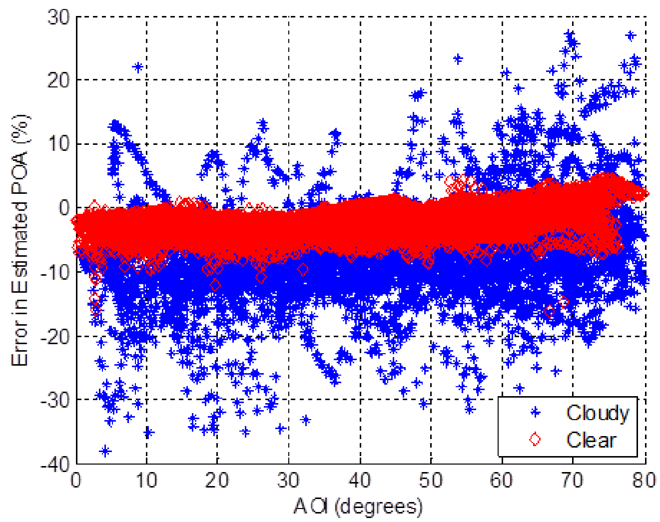

Figure 3. Dependence on sky condition of POA irradiance error for the isotropic sky model (March 2011 shown).

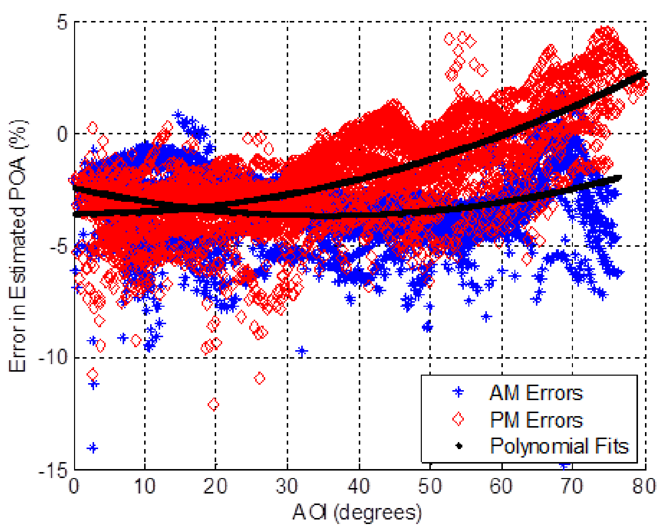

Figure 4. Dependence on time of day of POA irradiance error for the isotropic sky model showing empirical detrending (clear periods of March 2011 shown).

III. Uncertainty Propagation

Because the inputs to models for a PV system (e.g., irradiance) naturally are time series, output from the models employed are also time series. Moreover, each model in the sequence is deterministic: for a fixed set of time series inputs, the model's output is also fixed. We first computed a "baseline" time series of output for each model by assuming that the input and output from each model are exactly correct. The baseline time series serves as a reference against which to determine the direction (increasing or decreasing) of influence of a model's uncertainty on system output.

To obtain an ensemble of time series of possible system output we propagate uncertainty from each modeling step to the next. Conceptually, we randomly sample errors for each model consistent with the distribution of model residuals, apply the random errors to the model's output, and pass the modified output as input to the next model in the sequence. To illustrate symbolically, denote the model for estimating G_{POA} as $f_{POA}(GHI,...)$ and the model for estimating Ee as $f_E(G_{POA},...)$. Then an element of the ensemble of effective irradiance at any particular time can be represented symbolically as

$$Ee_i = f_E\left(G_{POA,i}\right) + \varepsilon_{E,i} = f_E\left(f_{POA}\left(GHI,...\right) + \varepsilon_{POA,i}\right) + \varepsilon_{E,i}$$

where $\varepsilon_{E,i}$ and $\varepsilon_{POA,i}$ are the randomly sampled values for the errors in $f_E(G_{POA},...)$ and $f_{POA}(GHI,...)$, respectively.

Because a model's error may be correlated between time steps we created stochastic process models for each model's error that in our judgment represent the expected correlations [13]. For example, the residual for modeled POA irradiance is temporally correlated during clear-sky conditions for each day, reflecting our judgment that the POA model's error would tend to be in the same direction throughout a day's clear periods. By contrast, the residuals for modeled POA irradiance are considered uncorrelated during non-clear conditions because we judge that during these times the model errors are derived from random effects (e.g., cloud movements, ground reflections, etc). We sample these process models (i.e., use a Monte Carlo approach) to generate an ensemble of time series of possible model errors. Where multiple models are available for a particular step, we perform separate analyses with each model.

We apply each model to the initial data or to output of the previous model and then apply the sampled time series of model errors to obtain an ensemble of each model's output. Because the model errors are quantified by comparing a model's output with measurements, any deviation between model and measurement is reflected in the distribution of model errors. Thus, each element of the ensemble of model realizations (obtained by applying errors to the baseline model output) represents an estimate of measured quantity the model should predict. By contrast the baseline time series is shifted away from the measured quantity if any bias is present in the model.

Figure 5 illustrates the distribution of daily energy represented in the baseline time series (red curve) and in each of time series in the ensemble (blue curves), when using the isotropic sky model for POA irradiance, the SunPower module, and meteorological data for Albuquerque, NM. We observe that the uncertainty in the predicted daily energy is relatively small (i.e., the blue curves are tightly grouped). However, the ensemble is also shifted from the baseline realization, resulting from a bias in the isotropic sky model causes a systematic over-estimate in daily energy. Similar figures were obtained for the First Solar module and when using the data from Golden, CO [13].

IV. Sensitivity Analysis

Propagation of uncertainty through the models as described earlier produces an ensemble of time series of system output that represents the aggregate uncertainty resulting from the models used in the calculation. The marginal distributions over each time series in this ensemble comprise the blue curves in Figure 5. We subtract the baseline time series from each member of the ensemble to obtain an ensemble of time series of deviations in DC power. For the sensitivity analysis, at each time step we correlate the deviations in system output with each model's error ensemble to determine the degree of influence of each model on the uncertainty in system output. The sensitivity analysis identifies the models that cause the spread among the blue curves in Figure 5.

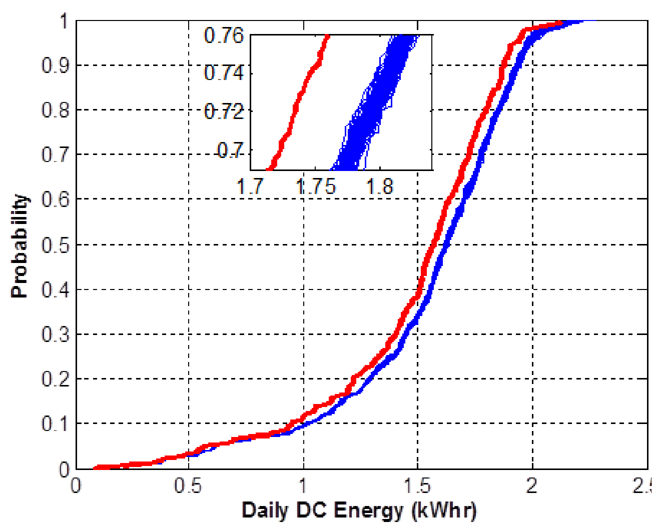

Figure 5. Distributions of daily energy: baseline (red) and ensemble (blue) for the SunPower module using meteorological data from Albuquerque, NM and the Isotropic sky model.

We reduce each time series of deviations in system output to deviations in daily, monthly and annual DC energy, and correspondingly reduce the time series of sampled errors for each modeling step comparably (i.e.., to error in daily POA insolation, error in daily degree hours, error in daily effective insolation, error in daily amp-hours). At each day (or month) we perform a stepwise rank regression between the deviation in DC energy and the daily summation of error for each model. The results of these correlations indicate the relative influence of each model on the deviations in DC power from the baseline values, thereby answering the question: which models matter to the uncertainty in predicted system output?

Figure 6 illustrates the results of the stepwise rank regression when using the isotropic sky model, the SunPower module, and meteorological data from Albuquerque, NM. For each day of the year, we plot the regression coefficient associated with the daily error in the modeled POA irradiance (POA), effective irradiance (*Ee*), cell temperature (T_C), DC voltage at maximum power (V_{mp}), and DC current at maximum power (I_{mp}) (coefficients range between –1 and +1 due to the use of rank regression). We readily observe that uncertainty associated with the POA irradiance model is dominant, with a secondary contribution from uncertainty in the effective irradiance model. Uncertainty in the other models has little effect on the uncertainty in daily energy. Similar conclusions are obtained for monthly and annual energy.

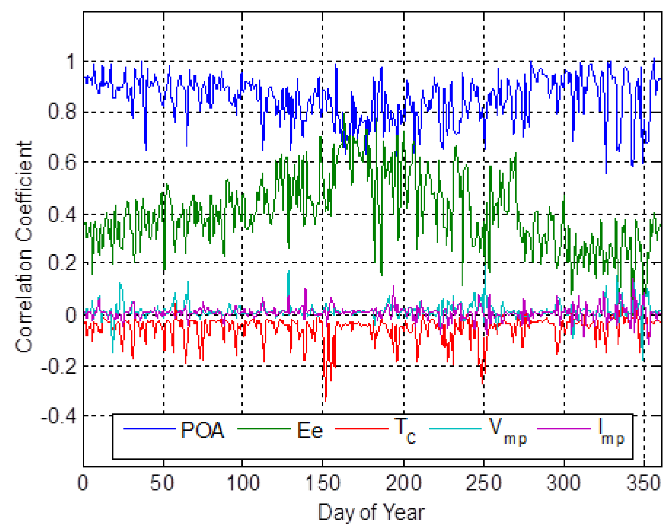

Figure 6. Stepwise rank regression results showing the significance of error in the Isotropic sky model.

Figure 7 demonstrates that similar conclusions are obtained for any of the POA irradiance models considered. Detailed analysis of POA irradiance models [15] concludes that these models differ primarily in terms of bias. Figure 8 illustrates that similar sensitivity analysis results are obtained for the First Solar module, and when meteorological data from Golden, CO are used.

IV. Conclusion

We have presented an uncertainty and sensitivity analysis for the daily DC output of two representative PV modules that considered the uncertainty in models for POA irradiance, effective irradiance, cell temperature and module DC output. We found that bias in the POA models is primarily responsible bias in the predicted energy output, and that uncertainty in the POA irradiance models is the dominant cause of uncertainty in the predicted output. Uncertainty in the effective irradiance model makes a secondary contribution to uncertainty in system output; the effects of uncertainty in cell temperature and module DC output are negligible. These conclusions hold for different module technologies and different sources of meteorological data, but are conditional on having well-calibrated base models. Poorly calibrated models will cause a greater bias and/or uncertainty in predicted system output.

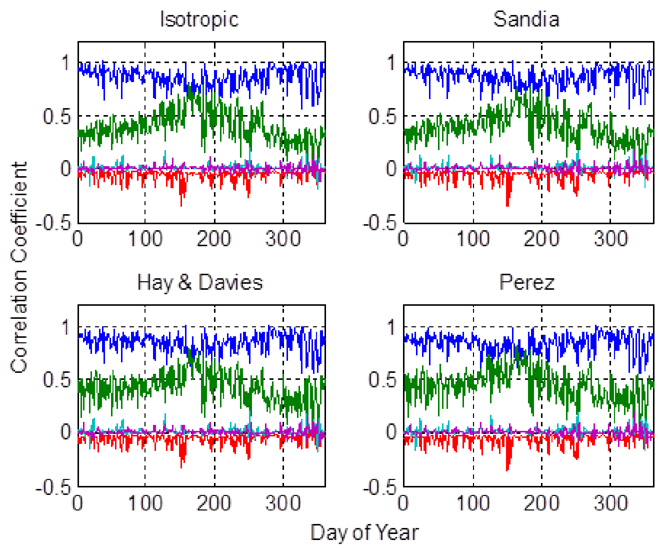

Figure 7. Stepwise rank regression results for each POA irradiance model for the SunPower module using meteorological data in Albuquerque, NM.

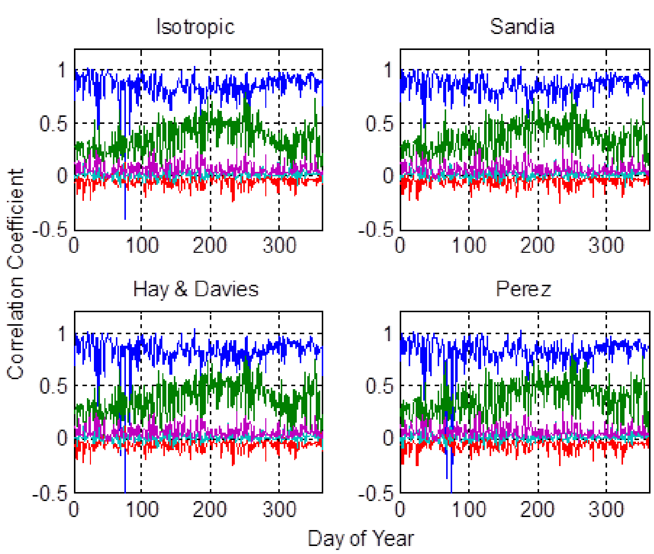

Figure 8. Stepwise rank regression results for each POA irradiance model for the First Solar module using meteorological data in Golden, CO.

The conclusions presented here do not consider the relative importance of measurement errors in the meteorological data, which we believe will be comparable to or eclipse the uncertainties due to the POA irradiance model, nor do they consider uncertainties introduced by other modeling steps not considered, such as surface reflection losses, module qulity mismatch, or soiling losses. However, the analysis structure presented here can be extended to consider additional models and modeling steps and we plan to present future analysis examining all necessary modeling steps from measured meteorological data to AC power.

ACKNOWLEDGEMENT

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 (SAND2014-4609C).

IV. REFERENCES

[1] J. Stein, "The Photovoltaic Performance Modeling Collaborative " in *IEEE Photovoltaic Specialists Conference*, Austin, TX, 2012.

[2] K. Whitfield and C. R. Osterwald, "Procedure for determining the uncertainty of photovoltaic module outdoor electrical performance," *Progress in Photovoltaics: Research and Applications,* vol. 9, pp. 87-102, 2001.

[3] D. Dominé, A. Jagomägi, A. Guérin de Montgareuil, G. Friesen, and H. D. M. E. Mõttus, D. Stellbogen, T. Betts, R. Gottschalg, T. Zdanowicz, M. Prorok, F. Fabero, D. Faiman, W. Herrmann, "Uncertainties of PV Module – Long-term Outdoor Testing," presented at the 25th PVSEC, Valencia, Spain, 2012.

[4] L. Dunn, M. Gostein, and K. Emery, "Comparison of Pyranometersvs. PV Reference Cells for Evaluation of PV Array Performance," in *38th IEEE Photovoltaic Specialists Conference*, Austin, TX, 2012.

[5] N. N. C. o. S. Laboratories), "American National Standard for Expressing Uncertainty--U.S. Guide to the Expression of Uncertainty in Measurement," ANSI1997.

[6] D. Thevenard and S. Pelland, "Estimating the uncertainty in long-term photovoltaic yield predictions," *Solar Energy,* vol. 91, pp. 432-445, 2013.

[7] C. Hansen, J. Stein, S. Miller, E. E. Boyson, J. A. Kratochvill, and D. L. King, "Parameter Uncertainty in the Sandia Array Performance Model for Flat-Plate Crystalline Silicon Modules," in *37th IEEE*

Photovoltaics Specialists Conference, Seattle, WA, 2011.

[8] D. Dirnberger and U. Kraling, "Uncertainty in PV Module Measurement—Part I: Calibration of Crystalline and Thin-Film Modules," *Photovoltaics, IEEE Journal of,* vol. 3, pp. 1016-1026, 2013.

[9] M. Campanelli and K. Emery, "Device-dependent light-level correction errors in photovoltaic I-V performance measurements," in *39th IEEE Photovoltaic Specialists Conference*, 2013, pp. 0067-0072.

[10] J. E. Hay and J. A. Davies, "Calculations of the solar radiation incident on an inclined surface," presented at the First Canadian Solar Radiation Data Workshop, 59, 1980.

[11] R. Perez, P. Ineichen, and R. Seals, "Modeling Daylight Availability and Irradiance Components from Direct and Global Irradiance," *Solar Energy,* vol. 44, pp. 271-289, 1990.

[12] P. G. Loutzenhiser, H. Manz, C. Felsmann, P. A. Strachan, T. Frank, and G. M. Maxwell, "Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation," *Solar Energy,* vol. 81, pp. 254-267, 2007.

[13] C. Hansen, A. Pohl, and D. Jordan, "Uncertainty and Sensitivity Analysis for Photovoltaic System Modeling," Sandia National Laboratories, Albuquerque, NM SAND2013-10358, 2013.

[14] D. L. King, E. E. Boyson, and J. A. Kratochvil, "Photovoltaic Array Performance Model," Sandia National Laboratories, Albuquerque, NM SAND2004-3535, 2004.

[15] M. Lave, C. Hansen, A. Pohl, W. Hayes, and W. Hobbs, "Evaluation of GHI to POA Models at Locations across the United States," in *40th IEEE Photovoltaic Specialists Conference*, Denver, CO, 2014.