THE EFFECT OF UNCERTAINTY IN MODELING COEFFICIENTS USED TO PREDICT ENERGY PRODUCTION USING THE SANDIA ARRAY PERFORMANCE MODEL

Larry Pratt (lpratt@sandia.gov) and David L. King (davidlking@aol.com)
Sandia National Laboratories, Albuquerque, NM 87185-0752

ABSTRACT

Predicting photovoltaic array performance is an important part of system design and monitoring, so it's important to quantify the uncertainty associated with the predictions. The Sandia Array Performance Model [1] is one of many tools used to predict annual energy production, but the effect of the uncertainty in model coefficients has not been fully investigated. This paper quantifies the relative importance of voltage and current temperature coefficients, as well as the coefficients relating voltage and current to solar irradiance, for crystalline silicon modules. Using the coefficient variation observed in the Sandia module database and computer simulation, the effect of the uncertainty was quantified in terms of the range in predicted annual energy production relative to actual energy production by three small grid-connected PV systems. The relative importance of each coefficient by month of the year was also determined in order to understand the seasonal behavior of the performance model.

INTRODUCTION

System design and field monitoring depend on accurate photovoltaic array performance predictions. Performance predictions are useful in comparing different technologies prior to installation and for monitoring systems in the field. Understanding the precision in performance predictions is essential to making decisions with high confidence. This paper documents the impact of uncertainty associated with five model coefficients used in the Sandia Photovoltaic Array Performance Model [1]. Previous work has already put some bounds on the effect of the temperature coefficients relative to other system level factors in determining ac-energy available from PV systems [3]. We use the model to calculate the range of predicted annual energy production from three small systems installed at Sandia National Laboratory's outdoor test site as a way of estimating the impact, or sensitivity, of each coefficient on the model residuals.

Sandia National Laboratories (SNL) has for many years conducted outdoor tests on photovoltaic modules to estimate performance coefficients from linear regression for use in the Sandia Photovoltaic Array Performance Model. Sandia also maintains a database of performance coefficients for hundreds of photovoltaic modules for modeling purposes. The Sandia performance model has been implemented in the U.S. Department of Energy's Solar Advisor Model (SAM) [2]. The uncertainty estimates documented in this analysis will also support a separate modeling and simulation effort underway at SNL to understand differences among various competing performance models, including the Solar Advisor Model (SAM).

PROCEDURE

Sandia Array Performance Model

The Sandia Array Performance Model was used to predict annual energy production from three small crystalline silicon arrays of less than 2 kW each located at the Sandia outdoor test facility. The nominal coefficients taken from the module database for the appropriate array provided the baseline prediction for annual energy production. We then applied 32 unique sets of model coefficients generated by a 2^5 experimental design plus one center point in order to generate a range of predicted annual energy production. The uncertainty in the model coefficients was represented by the variation in the 32 unique sets of coefficients. The range in predicted energy production from the simulation was used to quantify the effect of the uncertainty in the model coefficients on performance predictions. A more realistic estimate of the uncertainty in some of the model coefficients was also quantified through analysis of test results on multiple modules from the same manufacturer.

The Sandia Array Performance Model uses three key equations to predict power at any given time, shown in equations 1-3. The uncertainty analysis will address the impact of five parameters in these equations: β_{Vmp}, α_{Imp}, C_0, C_2, and the diode factor 'n'.

$$Imp = Imp_0 \cdot \{C_0 \cdot E_e + C_1 \cdot E_e{}^2\} \cdot \{1 + \alpha_{Imp} \cdot (T_C - T_0)\} \quad (1)$$

$$Vmp = Vmp_0 + C_2 \cdot N_s \cdot \delta(T_C) \cdot \ln(E_e) + C_3 \cdot Ns \cdot \{\delta(T_C) \cdot ln(E_e)\}^2 + \beta_{Vmp}(T_C - T_0) \quad (2)$$

$$Pmp = Vmp * Imp \quad (3)$$

Imp = Predicted maximum-power current
Vmp = Predicted maximum-power voltage
Pmp = Predicted maximum power
Imp_0 = Current at maximum power under standard test conditions.

Vmp_0 = Voltage at maximum power under standard test conditions.
E_e = Effective irradiance on a scale from 0 to 1+ suns.
C_0 and C_1 are regression coefficients estimated under clear and cloudy sky conditions by regressing Imp at 50 °C on $E_e + E_e^2$ and intercept = 0. $C_0 + C_1 = 1$.
T_C = Cell temperature
T_0 = Reference cell temperature, 25 °C
C_2 and C_3 are regression coefficients estimated by regressing Vmp at 50° C on nkT/q*ln(Ee).
α_{Imp} = Temperature coefficient for maximum-power current
β_{Vmp} = Temperature coefficient for maximum-power voltage
$\delta(T_C)$ = nkT/q = "Thermal voltage" where k is Boltzmann's constant, q is elementary charge, T is cell temperature in Kelvin, and n is the measured 'diode factor' of a cell n a module

The Sandia report provides a more detailed explanation of these equations and how they are generated for use in the model [1].

Sandia Module Database

The Sandia module database contains model coefficients for over 50 tested crystalline silicon modules and 300 crystalline silicon modules with estimated coefficients. The 5th and 95th percentiles from the distribution of model coefficients were used for the high and low levels for the experimental design and simulation, Figure 1. This plot includes both measured and estimated coefficients, since the distributions are similar.

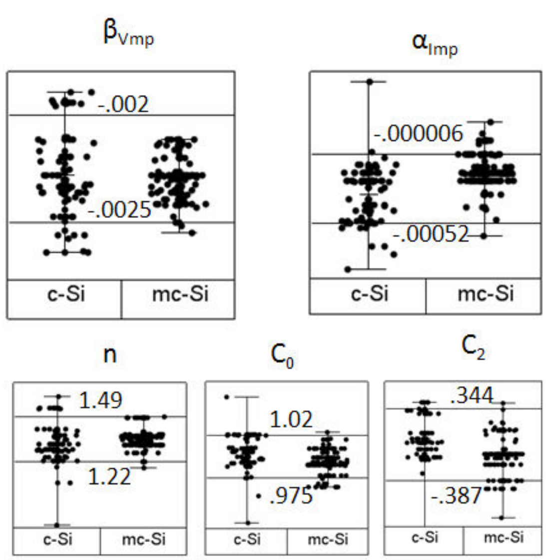

Figure 1: Distribution of model coefficients for crystalline silicon modules from the Sandia module database with reference lines at the 5th and 95th percentiles.

C_1 and C_3 are not included explicitly in the simulation because they do not vary independently from C_0 and C_2, respectively. $C_0 + C_1 = 1$. The equation $C_3 = -8.5 + 10.5 * C_2$ was used to estimate C_3 from C_2, based on the correlation between the two terms observed in the historical database.

Computer Simulation

A computer simulation was used to quantify the effect of uncertainty in the model coefficients on predicted energy production. For each of the three arrays, the Sandia Array Performance model was run with 32 unique sets of model coefficients from the factorial design plus the nominal set. The nominal set comes from the module database for the given technology.

DC power was measured every two minutes between March 2007 and March 2008 on three small grid-connected systems installed at the Sandia outdoor test facility. DC power was predicted at each observation using the Sandia Array Performance Model. Measured and predicted energy were calculated as the weighted sum of power by the hour, the month, and the year for each system, as shown in Equations 4 and 5. The weight of each sum was determined by the time interval between the current measurement and the previous measurement. If the interval was greater than one hour, that observation was ignored because the irradiance conditions may have changed a great deal in that interval of time. The number of observations was constant for each simulation so the comparison across the experimental design is valid, but not necessarily an accurate representation of a full year in production.

$$Energy_{Pred}\ [kWh] = \frac{Imp_{Pred}\ [A] * Vmp_{Pred}[V]}{1000} * \Delta Time\ [hours] \qquad (4)$$

$$Energy_{Meas}\ [kWh] = \frac{Imp_{Meas}\ [A] * Vmp_{Meas}[V]}{1000} * \Delta Time\ [hours] \qquad (5)$$

The residual was calculated as DC predicted energy minus DC measured energy, such that a positive residual indicated the model over-predicted the actual energy production and a negative residual meant the model under-predicted actual energy. Effective irradiance was collected from a calibrated crystalline reference cell in the plane of array. Cell temperature was estimated from thermocouples on the back of the modules and equation 12 from Sandia report [1] using a 3 °C temperature offset.

Finally, a multiple linear regression model was fit to the residuals versus the model coefficients to determine relative importance of each coefficient, Equation 6.

$$Y_i = \beta_0 + \beta_1 \cdot \beta_{Vmpi} + \beta_2 \cdot \alpha_{Impi} + \beta_3 \cdot C_{0i} + \beta_4 \cdot C_{2i} + \beta_5 \cdot ni + e_i \qquad (6)$$

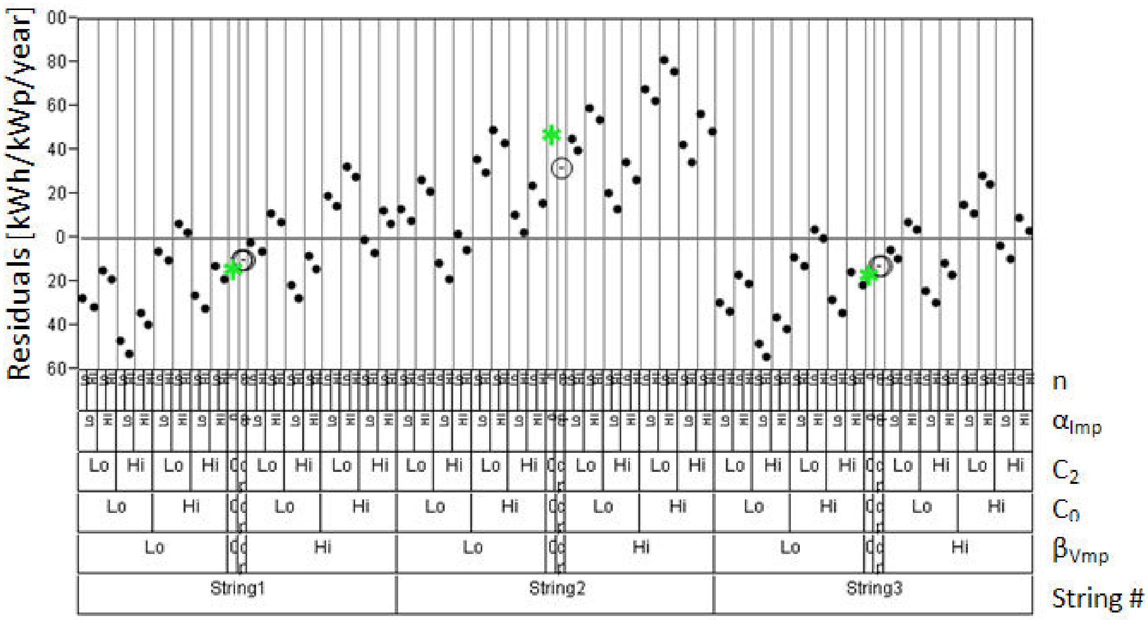

Figure 2: Residuals for three systems over the 34 simulations (32 factorial points + 1 center point + 1 nominal setting) for which actual DC energy was approximately 2000 kWh/kWp/yr. The residual using the nominal coefficients is shown with the green asterisk and the center point is shown with 'o'.

RESULTS

Annual Energy Residuals

Table 1 shows the sensitivity of annual energy yield on the uncertainty in model coefficients. Predicted yield varied by as much as 5% from the lowest to highest prediction for a given array of modules.

Energy Residuals [kWh/kWp/year]

String	Min	Max
1	-52 (-2.6%)	32 (1.6%)
2	-19 (-0.9%)	82 (4.1%)
3	-54 (-2.7%)	27 (1.4%)

Table 1: Minimum and maximum of annual energy residuals for three systems over the simulation space.

The residuals for all 34 sets of model coefficients are shown in Figure 2. The residual divided by the power rating of the system is plotted on the y-axis, so the interpretation of the residual is straightforward in terms of annual yield (kWh/kWp/yr). The 5th and 95th percentiles of the model coefficients are plotted along the x-axis as 'lo' and 'hi', respectively.

Several general patterns emerge from the plot. First, the Sandia model using the nominal coefficients does a good job of predicting the annual DC energy production. Predicted yield for String 1 and String 3 are low by 15 kWh/kWp/yr, or 0.75%, relative to the 2000 kWh produced. The prediction for String 2 was high by about 40 kWh/kWp/yr, or 2%. The residuals for the center points of the design space, shown as circles in the graph, appear close to the nominal coefficients, as well. This result suggests that 'generic' model coefficients for this specific technology did just as well as the nominal coefficients in predicting energy yield, for these three system examples.

Several additional observations can be taken from the graph, some of which should be expected from a careful examination of equations 1 and 2, some of which are less obvious.

1. As β_{Vmp} varies from -0.0025 (Lo) to -0.002 (Hi), the coefficient is approaching 0, so predicted power increases because the voltage loss due to temperature diminishes.
2. As C_0 varies from 0.975 (Lo) to 1.021 (Hi), it moves from a number less than one to a number greater than one, so predicted power increases because $E_e > 0$.
3. As C_2 varies from -0.387 (Lo) to 0.344 (Hi), the predicted power decreases for all three systems,

which is not obvious from the equations. As C_2 varies from Lo to Hi, it changes from a negative number to a positive number. For $C_2 < 0$, predicted Vmp increases when $E_e < 1$ and decreases when $E_e > 1$, because $\ln(E_e) < 0$ when $E_e < 1$. When C_2 is positive, predicted Vmp decreases when $E_e < 1$ and increases when $E_e > 1$. The effect of C_2 has to be understood in terms of the distribution of E_e for a given area or time period. (See Figure 7.)

4. As α_{Imp} varies from -0.00052 (Lo) to -0.000006 (Hi), the coefficient is actually getting closer to 0, so predicted power increases because the current loss due to temperature decreases .
5. As n varies from 1.216 (lo) to 1.49 (Hi), predicted power decreases. This effect is also not obvious from the equations, as it depends on the sign of the C_2 and the distribution of irradiance.

The relative importance of each coefficient in explaining the variation in predicted annual energy yield is shown in Figure 3. The points represent the difference between the high and low annual energy predictions for each term averaged across the other factors in the experimental design.

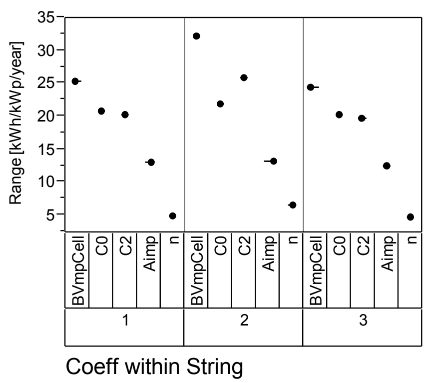

Figure 3. Relative importance of model coefficients on variation in annual energy predictions for systems that produced 2000 kWh/kWp/yr.

B_{Vmp} has the biggest impact on energy prediction, but the effect is limited to $\pm$ 0.75% of actual energy. C_0 and C_2 follow as the second and third most important coefficients for predicting energy production. The α_{Imp} coefficient effects the prediction by $\pm$ 0.4%, while the diode factor is limited to less than $\pm$ 0.2%. The sum of the individual effects is equal to the overall variation shown in Figure 2.

The relative rank is likely to vary for systems in different climates, depending on the distribution of temperature and irradiance, as well as for different technology types. As previously reported, the effect of temperature coefficients on predicted annual energy production for mc-Si is different for Albuquerque as compared to Buffalo, NY, where it's much colder and less sunny [3].

Monthly Energy Residuals

The relative importance of each model coefficient also changes by month of the year (Figure 4). There is no significant difference from one array versus another in terms of relative importance by month, so the three arrays were aggregated to the same panel in this graph. The importance of C_0, C_2, and n remain constant from month to month, but the influence of temperature coefficients α_{Imp} and β_{Vmp} changes over the year. β_{Vmp} can impact predictions by as much as $\pm$ 2.5 kWh/KWp/month in June and July when actual energy production was around 200 kWh, but it showed virtually no impact in the cooler winter months.

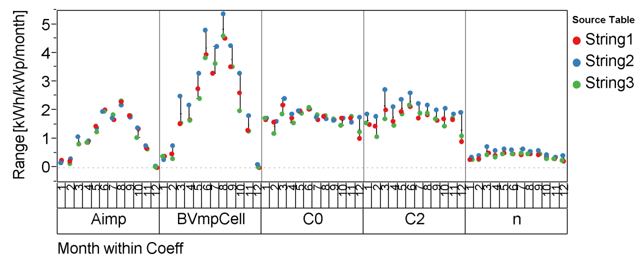

Figure 4. Relative importance of each model term by month of the year for three small crystalline systems in Albuquerque, NM.

DISCUSSION

Uncertainty of Model Coefficients

The estimates of uncertainty in the model coefficients used for the simulation represent a worst case scenario because the analysis assumes all the variation in the module database represents random variation about the true coefficient. Physics based models should explain at least some of the variation in the database: cell technology, cell production variability, cell mismatch in modules, module sampling uncertainty, test procedures including outdoor environmental conditions or laboratory equipment used to vary temperature, data analysis or regression algorithms, and measurement error [4].

Using historical data from Sandia outdoor module testing, we can estimate the variation associated with different modules which are nominally equivalent in terms of cell technology and module construction. The within group variation of nominally identical modules is about one-half to one-third as much as the entire distribution for β_{Vmp} and α_{Imp} , Figure 5.

The range of predicted annual energy yield for Array 2 was cut in half when using uncertainty estimates equal to half of the original estimates. A Gage Repeatability and Reproducibility (Gage R and R) study would be necessary to further refine the estimate of uncertainty for a broad range of module products and test conditions, and to identify opportunities for improving the measurement process.

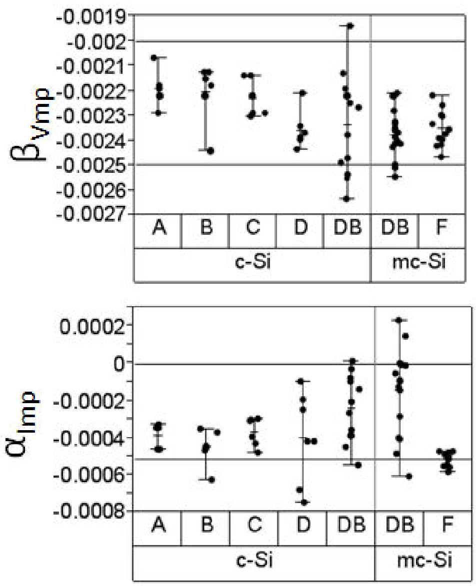

Figure 5. β_{Vmp} [V/°C/cell] and α_{Imp} [1/°C] estimates for crystalline silicon modules. A, B, C, D, and F show groups of nominally equivalent modules. The references lines are the 5th and 95 the percentiles from the module database. DB represents the remaining records in the module database for silicon modules.

Figure 6 shows the β_{Vmp} estimates of several different modules from three different manufacturers of crystalline silicon modules. β_{Vmp} for manufacturer 1 is significantly higher than the other two by 0.1 mV to 0.27 mV/ C/ cell, based on a 95% confidence interval for the difference in the means. This analysis was based on measured modules only. By contrast, there is no significant difference in α_{Imp} for these same three manufactures.

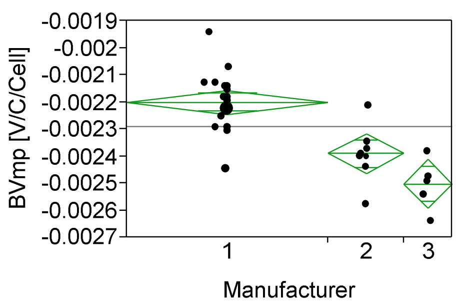

Figure 6. Comparison of β_{Vmp} (V/C/cell) for three different module manufacturers of crystalline silicon modules.

The estimates of within group variation and the significant difference in β_{Vmp} for at least one manufacturer suggest that the uncertainty estimates used in the simulation were over-stated by a factor of 2 or 3. The combined sensitivity of annual energy yield to model coefficients drops from $\pm$ 2.5% to $\pm$ 1.5% using these more realistic estimates of uncertainty in model coefficients. The sensitivity of any single model coefficient is considerably less than the combined effect of all five terms.

Relative Importance by Month

The relative importance of the two temperature coefficients changes over the course of the year. Predicted energy yield is more sensitive to β_{Vmp} and α_{Imp} in the summer months because the average module temperature is higher and further away from standard test conditions. The average monthly cell temperature ranged from 18 °C in January to 45 °C in July and August.

The relative importance of C_0 and C_2 remain constant over the year because they are the coefficients for effective irradiance effects. Effective irradiance generally varies from 0 to full sun over the course of each day all year round, as opposed to operating temperature which is more seasonal.

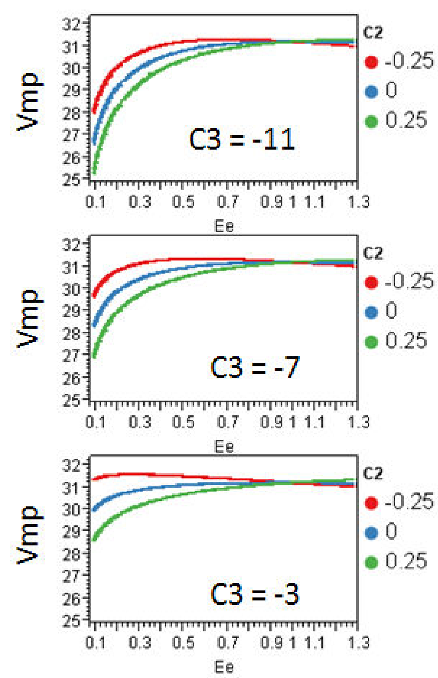

Figure 7. Predicted Vmp vs effective irradiance for a single module over a range of C_2 and C_3 values taken from the module database.

C_2 is also the only parameter to change signs from the low to high setting. C_2 and the corresponding C_3 coefficients are used to predict the Vmp as a function of E_e, and they are especially important in lower light conditions, as shown in Figure 7. As Ee approaches 1, Vmp approaches Vmp0, and C_2 and C_3 have little effect on predicted Vmp.

However, C_2 and C_3 are important for modeling the relationship between Vmp and effective irradiance below 0.7 suns.

Finally, the importance of the diode factor on energy predictions remains fairly constant and relatively small over the year. The diode factor enters in to the equation as part of the thermal voltage, which is roughly 26 mV per cell at 25 °C and n = 1. With five modules in series and 72 cells per module, the predicted Vmp would increase (or decrease depending on the sign of C_2) from 2.4 V to 3.5 V between n=1 and n=1.5 when C_2 =0.25. This represents a 0.5% relative change for a 200 V system.

Overall Modeling Accuracy

It is important to keep the results presented in this document in context with the overall uncertainty associated with modeling PV system energy production. This study specifically addresses only variation in temperature coefficients and coefficients relating the behavior of Vmp and Imp to solar irradiance. As discussed elsewhere, there are many other factors influencing both the DC performance characteristics of a PV array as well as the AC performance of the overall system [3]. All of these factors must be considered in modeling or monitoring PV system energy production.

This study quantifies the influence and the relative importance of specific coefficients in the Sandia performance model for crystalline silicon systems operating in Albuquerque. Knowing the relative influence of these coefficients will assist in guiding development of module testing procedures and in improving the Sandia performance model. Similar analyses will need to be conducted for other PV technology types (thin-film, concentrator) at different geographic locations in order to better understand the overall uncertainty in annual energy calculations using the Sandia performance model.

CONCLUSION

The effect of uncertainty in the model coefficients for predicting annual energy yield using the Sandia Array Performance Model ranged from 2% to 5% relative to actual energy production. Using the worst case estimates of uncertainty, the predicted annual energy yield varied by 100 kWh/kWp/yr on a system that produced 2000 kWh/kWp/yr of DC power (5%). Using slightly more realistic estimates of coefficient uncertainty, the predicted annual DC energy yield ranged only 2% relative to actual energy yield. While the relative importance of each parameter can change over the course of the year, β_{Vmp} was most important in terms of predicting annual energy yield. Under the worst case estimates of uncertainty however, the sensitivity to β_{Vmp} was roughly $\pm$ 0.75%. C_0 and C_2 each had slightly less influence on the predicted annual yield compared to β_{Vmp}. α_{Imp} was about half as important and the diode factor had virtually no impact on predicted annual energy production.

These conclusions were reached for crystalline silicon systems operating in environmental conditions typical of Albuquerque, New Mexico. The relative importance of the modeling coefficients considered will likely differ for other PV technology types and for sites with different environmental conditions. However, relatively liberal bounds on coefficient variability, due to either measurement error or module production variability, were found to introduce relatively small errors (<5%) when comparing predicted to measured annual energy production.

ACKNOWLEDGEMENT

Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Sandia acknowledges the support of the DOE Solar Energy Technologies Program in particular for the work presented here.

The authors also thank Gary Galbraith for his design, installation, and maintenance of the small systems test facility at Sandia National Labs. He also provided valuable input to the implementation of the simulation used in this study.

REFERENCES

[1] D. King, W. Boyson, J. Kratochvil, " Photovoltaic Array Performance Model," Sandia National Laboratories Report, SAND2004-3535, August 2004.

[2] Solar Advisor Model, National Renewable Energy abs, https://www.nrel.gov/analysis/sam/

[3] D. King, J. Kratochvil, W. Boyson, "Analysis of Factors Influencing the Annual Energy production of Photovoltaic Systems," 29th IEEE PVSC, 2002.

[4] D. King, J. Kratochvil, W. Boyson, "Temperature Coefficients for PV Modules and Arrays: Measurement Methods, Difficulties, and Results," 26th IEEE PVSC, Anaheim, CA, 1997.

[5] Lasnier et al. 1990, Photovoltaic Engineering Handbook, IOP Publishing Ltd