A RELIABILITY AND AVAILABILITY SENSITIVITY STUDY OF A LARGE PHOTOVOLTAIC SYSTEM

E. Collins, S. Miller, M. Mundt, J. Stein, R. Sorensen, J. Granata, and M. Quintana
Sandia National Laboratories
PO Box 5800, Albuquerque, New Mexico, 87185-1033

ABSTRACT: A reliability and availability model has been developed for a portion of the 4.6 megawatt (MWdc) photovoltaic system operated by Tucson Electric Power (TEP) at Springerville, Arizona using a commercially available software tool, GoldSim™. This reliability model has been populated with life distributions and repair distributions derived from data accumulated during five years of operation of this system. This reliability and availability model was incorporated into another model that simulated daily and seasonal solar irradiance and photovoltaic module performance. The resulting combined model allows prediction of kilowatt hour (kWh) energy output of the system based on availability of components of the system, solar irradiance, and module and inverter performance. This model was then used to study the sensitivity of energy output as a function of photovoltaic (PV) module degradation at different rates and the effect of location (solar irradiance). Plots of cumulative energy output versus time for a 30 year period are provided for each of these cases.
Keywords: effective availability, photovoltaic reliability, sensitivity, solar irradiance, system modeling, accelerated aging.

1 INTRODUCTION

Sandia National Laboratories' (SNL) Department of Energy (DOE) Photovoltaic Reliability Program is developing a suite of software tools to facilitate system-level prediction capabilities. One of these tools is a simulation that will predict cumulative energy output of a system versus time with respect to solar irradiance, PV module performance, and equipment availability. The purpose of this tool is to provide system designers with "realistic" output energy generation projections over the life of a system for a specific system configuration and location.

2 OVERVIEW OF MODELING APPROACH

To estimate the output energy from the PV plant as a function of time, a simulation model was developed that encompasses weather variability, PV module performance, and system design topology. The PV plant's output is simulated assuming all components are functioning and the PV modules have no output degradation over time. This ideal performance includes the effects of system design and weather variability (within the year and between years). Actual plant output is then estimated by multiplying the ideal power for each hour with the plant availability, which is calculated from the reliability model. See Figure 1.

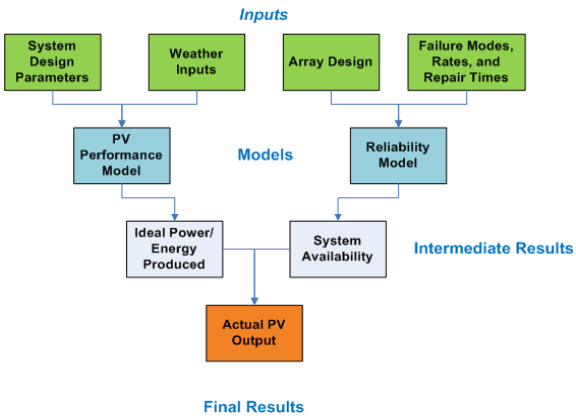

Figure 1. Integration of PV Performance and Reliability Models.

2.1 Weather (Solar Irradiance) Model

Typical Meteorological Year (TMY2) data are used to represent hourly weather and irradiance for a typical year at the test locations. The TMY2 dataset is available for 239 locations in the U.S. and was created by stringing together typical months of measured and simulated data from a 30 year period of observations (1961-1990) at each site. To create 30 years of weather for our simulations, the TMY2 sequence was repeated 30 times and each irradiance value was multiplied by a scaling factor to account for the effect of the inter-annual variability on the mean irradiance for a 30 year record. The inter-annual relative uncertainty of the TMY2 irradiance data is reported to be +/- 9% for the 95% confidence interval. To represent the effect of this uncertainty on the total energy from the PV system over its 30 year life, we calculated the standard relative error on the mean based on 30 yearly samples:

$$\frac{0.09*0.5}{\sqrt{30}} = 0.82\% \qquad (1)$$

This uncertainty is included in the analysis by multiplying the irradiance values from each realization by an irradiance scaling factor based on the standard error.

2.2 PV Performance Model

Baseline output power from the PV plant is estimated with the Sandia Photovoltaic Array Performance Model [1]. This empirically based model calculates the maximum power point for the array IV curve from hourly irradiance, weather data, and PV array design parameters (e.g., module type, mounting orientation, cell temperature, etc.). The module parameters are derived from outdoor testing performed at Sandia National Laboratories.

An important input to the model is the plane of array (POA) irradiance. The model requires that POA irradiance be divided into direct and diffuse components. The direct component is calculated as the product of the direct normal irradiance (DNI) and the angle of incidence, calculated analytically from the sun position [2]. The diffuse component is estimated using the Perez radiation model [3]. Baseline PV output is strongly correlated with POA irradiance, but the model also includes terms to account for (1) cell temperature effects on electrical performance and efficiency, (2) light

reflected off the cover glass of the modules at high incident angles, and (3) spectral effects due to variations in air mass.

2.3 Reliability and Availability Model

The reliability and availability model for the Tucson Electric Power (TEP) Springerville Generating Station was previously developed [4]. That reliability block diagram (RBD) had 6 levels with 26 inverter-arrays. Only one inverter-array from this system is modeled in this sensitivity study, but the approach is scalable to larger systems if adequate computing resources are available. Figures 2-5 are the RBD levels that describe the inverter-array that is modeled in the GoldSim™ software.

In reference [4] a simple reliability definition was used. Success was defined as a single PV string delivering power to the grid. Availability of the system was calculated for this definition of system success using ReliaSoft's BlockSim 7™ tool.

For the analysis in this paper, the commercial software tool, GoldSim™, enables sensitivity analyses for more complex definitions of reliability, as well as the modeling of environmental influences such as geographical location, seasons, and weather on the kilowatt hour output of the inverter-array over time.

The reliability modeling in GoldSim™ is not configured in reliability block diagram format, but instead as a parent-child hierarchy that follows the flow of kilowatt hour production through critical components in the system. The granularity of the model was based on the major components of the system and the level of identification of field failures in the reporting process. Figures 2-5 represent a hierarchical model of the inverter array, which is the basic power generation unit.

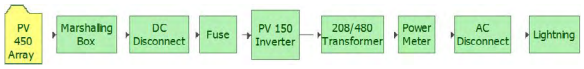

Figure 2. RBD of Inverter-Array, Level 3.

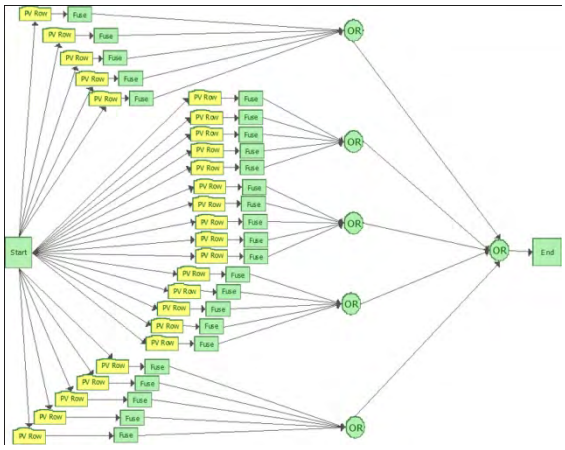

Figure 3. RBD of PV 450 Array, Level 4.

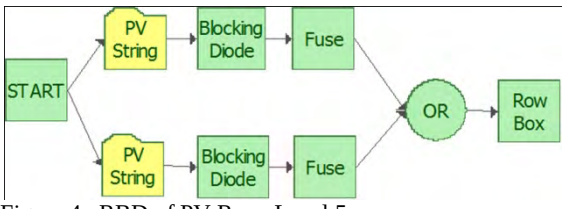

Figure 4. RBD of PV Row, Level 5.

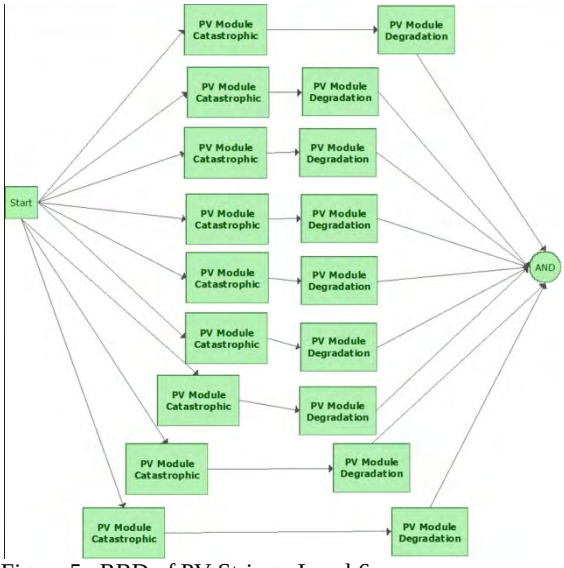

Figure 5. RBD of PV Strings, Level 6.

To yield useful reliability and availability metrics, the basic blocks in the model are populated in GoldSim™ with life distributions and repair distributions for each block. The failure modes for all of the blocks except the inverter in the model are populated with the same failure distributions as those found in reference [4]. The repair distributions available in GoldSim™ are limited to exponential, lognormal, and gamma. Thus, the same field repair data from reference [4] was used to choose the best fit from among those 3 available distributions for corrective maintenance and for off-time descriptions of all the blocks in the performance model. Although the available repair distributions in GoldSim™ were not the best fit distributions to the field data, visual comparison of both indicated that they were representative, so their usage should not unduly affect the predictions from the simulation.

There are several improvements not included in reference [4] that are implemented in the performance model. Fidelity of both inverter and photovoltaic module modeling is improved. In GoldSim™ the scale parameters of the irradiance, failure, degradation, and inverter disturbance distributions are varied during the simulation. A sensitivity analysis supporting an improved definition of reliability is included, as well as the modeling of environmental influences such as geographical location, seasons, and weather on the cumulative kilowatt hour production of the inverter-array over time. Thus, the implementation in GoldSim™ should represent significant improvement in the fidelity of the effective availability metric, which can be directly compared to what is used by the utility operators. Effective availability is defined in terms of actual kilowatt hour production compared to what could have been produced if the inverter array was perfectly available (no hardware failures or downtime) for the weather conditions at a particular geographical location.

For the inverter, a preventive maintenance distribution and a grid perturbation distribution were modeled in GoldSim™ as an additional failure mode and as an interruption of power production, respectively, along with their associated off-time distributions. In reference [4] these interruptions were not modeled, so their associated off times were triggered upon a system downing event, instead of the specific interruptive event.

Lightning was modeled in GoldSim™ as a triggering external disturbance resulting in failure of the inverter, instead of a series reliability element representing a common failure mode [4].

In the field the inverters were repaired rather than replaced upon failure. Reliability growth analysis in ReliaSoft's reliability growth analysis RGA 7™ tool was used to estimate the cumulative MTBF (360 days) at 5 years, the end of the period in which data was taken. This value is used in the exponential distribution as a catastrophic failure distribution to trigger corrective maintenance in the simulation. Cumulative MTBF represents a conservative estimate of inverter reliability. The instantaneous MTBF (482 days) after 5 years of reliability growth represents the achievable reliability, assuming prudent design that anticipates known field failure modes.

Table I summarizes the inverter failure mode and other triggering descriptions of inverter off-times that were modeled in the simulation. Scale parameters are in days or its transformation.

Table I. Inverter Failure Distributions Summary.

Failure Mode	Distribution	Log Mean or Lambda	Log Std Dev
Catastrophic	Exponential	0.00278	
Preventive Maintenance	Exponential	2.62	
Grid Perturbation	Lognormal	3.62	1.70

For each of the photovoltaic modules, in addition to the catastrophic failure mode characterized by field data, a degradation mode was implemented to represent linear degradation of the module's output over time. An additional simulation was run to include tape joint degradation from module packaging.

3 SENSITIVITY STUDIES

To evaluate the sensitivity of total lifetime energy production to various failure modes a Monte Carlo simulation was executed. Eight scenarios were run that explored the sensitivity of module degradation rates and mechanisms, as well as weather on lifetime energy production for the PV system. Each scenario consisted of 100 stochastic realizations that are intended to represent the epistemic uncertainty in system performance. Epistemic uncertainty includes uncertainty in the rates of failure and repair. For each realization parameter values were sampled using Latin Hypercube Sampling (LHS). In many cases, these parameter values define distributions used in the reliability model for determining the timing and severity of component failures. These uncertainties are considered to be aleatory, since they occur simply by chance as a function of time.

3.1 Comparison of Model to TEP Performance Data

The annual kilowatt hour (kWh) production for the 5 calendar years from 2003 to 2007 at the TEP Springerville Photovoltaic Generating Station is shown in Figure 6 in a histogram. The histogram displays counts of inverter-arrays in categories of annual kWh energy production. All of the 26 inverter-arrays were not up and running for an entire calendar year until 2005, so only 106 inverter-years were used for the comparison with the model in Table 2.

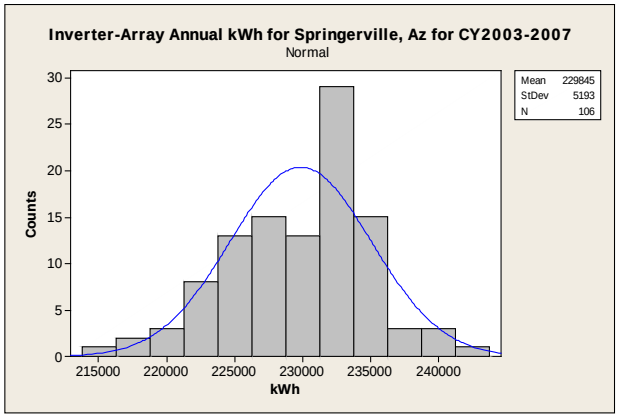

Figure 6. Histogram of Field Data Variation in Annual kWh

The simulation model predicted results in kWh for 30 years for the TEP Springerville Photovoltaic Generating Station. These are shown in Figure 7 for an assumed module linear degradation of 0.5% per year. Since weather data is not available for Springerville, Arizona, weather data from Flagstaff, which is approximately the same elevation, was used instead [5].

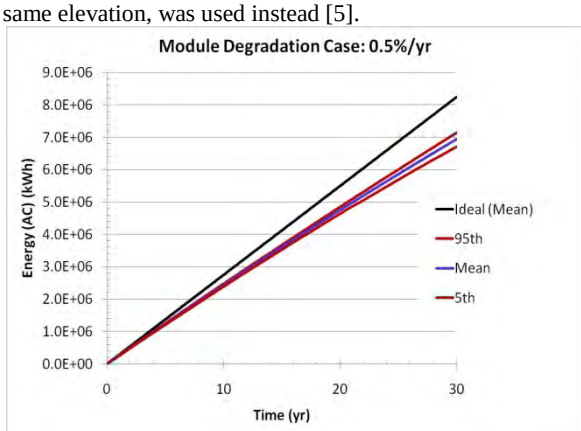

Figure 7. Predicted Cumulative Energy for 30 Years for Springerville Generating Station.

Table II compares the uncertainty in the model prediction to the field data at 5 years at the 5th, 50th, and 95th percentiles. These 5% to 95% intervals for the prediction and the field data overlap with 8% difference between the medians.

Table II. Comparison of Model with Field Data.

Percentile	Field Data (kWh)	Prediction (kWh)
5	221,300	229,900
50	229,800	247,700
95	238,400	263,600

Although 30 year predictions are given for cumulative energy plots in this paper, it should be recognized that only five-year field data and hypothesized accelerating aging models for the photovoltaic modules are used in the 30 year simulations. Accelerated stress testing is needed to determine aging and degradation models for the inverter and other

components in the system in order to make realistic projections for cumulative energy output.

The difference between the prediction and the field data can be explained by the error sources which include:

- The irradiance data used in the reliability is for Flagstaff, Arizona, because weather data is not available for the location where field data was collected on the system;
- The weather model assumptions from the TMY2 source have significant influence on the predicted power output;
- Year-to-year weather variation at the sites is not well characterized;
- Even with 5 years of operating experience, there was not enough field data to define precise failure, repair, and other off-time distributions for the system.

Stepwise linear regression analysis was performed with kilowatt hour production as the response variable at time equals 30 years. The independent variables were the parameters that were varied during the simulation, 11 in all. These were irradiance scaling factor, catastrophic failure distribution scale parameters for the six system hardware components in the simulation, module degradation rate, and inverter disturbance distribution scale parameters for grid disturbance, preventive maintenance, and lightning. This resulted in a five variable model with a final correlation coefficient squared of 96%, which is the proportion of variability in the response attributed to the model. Independent variables were added to the regression model based on their contribution to the overall model variability in the regression prediction of the response variable.

To investigate the sensitivity of the predictive model to weather variations, ideal energy production was simulated for 30 years with zero failures, no degradation, and no maintenance using the Flagstaff weather model. The uncertainty in irradiance due to weather variation given in reference [5] was simulated. Figure 8 depicts 30-year cumulative ideal energy production along with its 90% confidence interval. Irradiance variation due to weather appears to affect energy production predictions slightly.

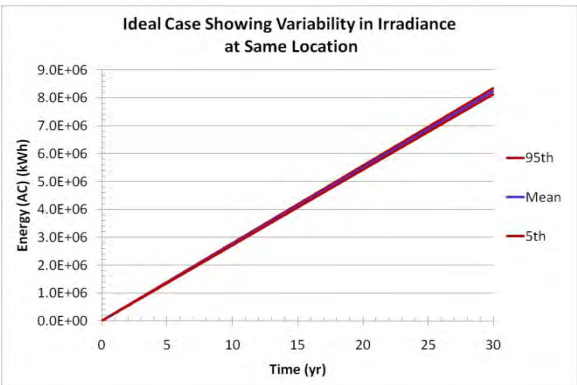

Figure 8. Effect of Weather Variation on Predicted Energy.

3.2 Photovoltaic (PV) Module Linear Degradation

The assumed degradation pattern due to aging for crystalline silicon (cSi) modules is linear. This rate for the modules at the Springerville Generating Station was estimated by Tom Hansen, the project engineer, to be approximately 0.5%, based on his measurements for the first 5 years after installation. Module linear degradation rates are separately simulated for cases of 0.5%, 1.0%, and 1.5% using the normal distribution in this paper. A standard deviation of 0.2% was used for all three distributions.

The mean parameter was allowed to vary randomly according to a triangular distribution of width plus or minus 3 normal standard deviations. Thus, for a given degradation case for a given module, a mean was randomly drawn from the triangular distribution and then this value, along with the appropriate constant standard deviation, was used to randomly sample a mean and standard deviation for the module from the normal distribution. This sample mean and sample standard deviation were then used throughout the entire 30 year simulation for that particular module. Figure 9 shows the degradation variation among modules from the 5th to the 95th percentiles for the 1% module linear degradation case over 30 years. Figure 10 shows the mean degradation curves over 30 years for the 0.5%, 1%, and 1.5% degradation cases.

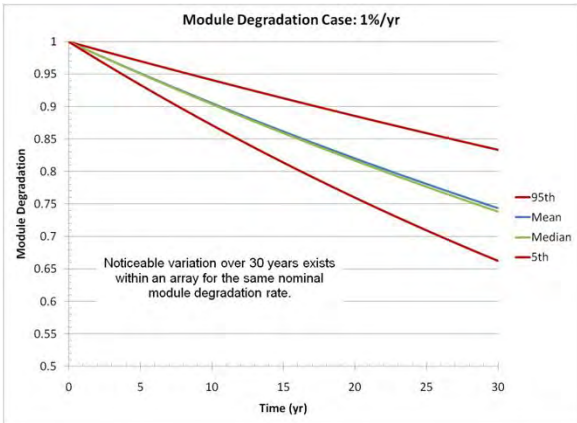

Figure 9. Module to Module Variation for 1% Degradation.

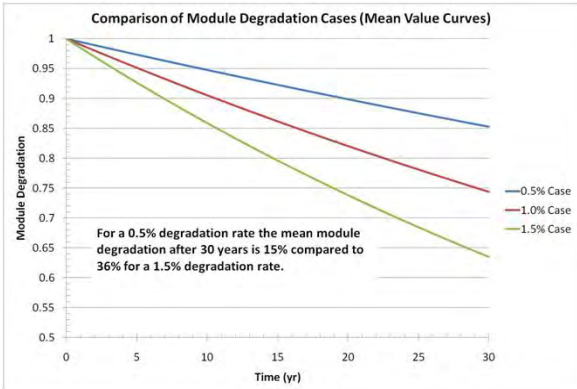

Figure 10. Comparison of 0.5%, 1%, and 1.5% Module Degradation Rates for 30 Years

Each of the module degradation cases of 0.5%, 1%, and 1.5% was simulated for 30 years. Figure 11 depicts median energy output for each of those cases. The differences among relative energy outputs increase dramatically after about 15 years. The 30 year cumulative energy production for the Flagstaff scenario of 1.5% degradation is 15% lower than the 0.5% degradation rate for the Flagstaff location. This suggests

that modules with high degradation rates should be replaced much earlier than modules with low degradation rates.

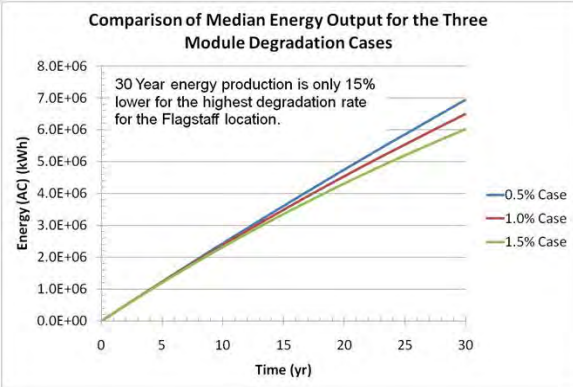

Figure 11. Comparison of Predicted Energy for 0.5%, 1%, and 1.5% Module Degradation.

3.3 PV Module Tape Joint Degradation

Packaging of a thin film PV technology was selected to demonstrate how materials degradation phenomena can be included in the system performance model. One high likelihood failure process identified through a Failure Modes and Effects Analysis (FMEA) of the thin-film technology was degradation of the metal foil tape joints, which are depicted in Figure 12. To generate degradation data, accelerated tests were run on samples with overlapping tape joints. The degradation data obtained in these experiments are being used in the system model [7]. Although a thin film technology example is used here to illustrate how the system simulation can use accelerated aging models, there is no intent to limit this analysis technique to any particular technology. The technique can be used with any failure mode, for which accelerated aging models are available, on any component in the system.

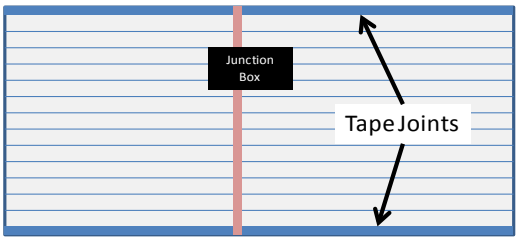

Figure 12. Schematic representation of a thin film module showing location of tape joints being analyzed.

To calculate the effect of the parasitic resistance on module performance, a simple electrical system analysis was performed. In general, the electrical properties of a module can be used to calculate the effect of resistance on performance. To illustrate this, we can consider a "standard" module with a 70 watt output, and a max power point voltage of 40V. The current, based on these values, is 1.75 A. Connected through an external load and running at peak power, the resistance of the external load (R_{load}) is 22.86Ω (R = E/I). Parasitic resistances (R_{tape}) were then introduced into the circuit (one for each tape joint on the module). Combining these with R_{load} provides a total resistance. Assuming that the voltage remains constant across the module, a new current can be calculated, and then used to calculate the power dissipated through the external load and through the two tape joints. The available output power then becomes that consumed by the external load. A decrease in power can be calculated as the ratio P_{load} to $P_{load\text{-}initial}$. We then fit the curve of relative power vs. resistance to obtain an analytical solution that can be programmed into the performance model (the equation used to fit this curve has no physical significance – it is merely a mathematical description of the generated data). For the "standard" module, the description is:

$$Rel.Power = 0.054 + 0.896 * EXP(-R/7.39) \quad (2)$$

Using accelerated aging data generated during thermal cycling of the tape joints, a description of the resistance as a function of thermal cycles can be obtained. If we assume a diurnal cycle and no shading due to clouds, a description of resistance as a function of time is derived.

$$R = 10^{(a+b\sqrt{t}+e)} \quad (3)$$

Where *a* is the initial resistance at t=0, *b* is the slope (in this case 0.028) obtained from the thermal cycling tests, and *e* is a "unit-specific" random normal variable with mean of zero and a standard deviation of 0.3. In this example, we set *a* to zero as we are interested in the resistance increases, and the initial resistance is already included in the module specifications. For this example, the equation then becomes:

$$R = 10^{(0.028\sqrt{t}+e)} \quad (4)$$

These calculations include the assumption of two parasitic resistors for each module (top and bottom). The tape interface on the test sample is smaller than on the module (factor of 20), so the calculations need to be corrected for area. The longer tape interface on the module can be treated as 20 parallel resistors so the actual parasitic resistance for the module is reduced by a factor of 20.

Thus, two equations are needed for the simulation. The first is the one that describes module power as a function of resistance. It is combined with the equation that gives resistance as a function of time. For the simulation, each module is assigned a specific equation for R = *f*(t). That is, a value for "*e*" is assigned based on the distribution parameters for *e*. That value remains constant (for that module) throughout the realization, and is used to determine the decrease in output power for the module.

Figure 13 illustrates a pronounced effect due to tape degradation over 30 years, which is approximately 36%, compared to the Flagstaff scenario with 0.5% degradation rate without tape degradation. A hypothesized new tape joint equation using less conservative assumptions for temperature cycling is also shown.

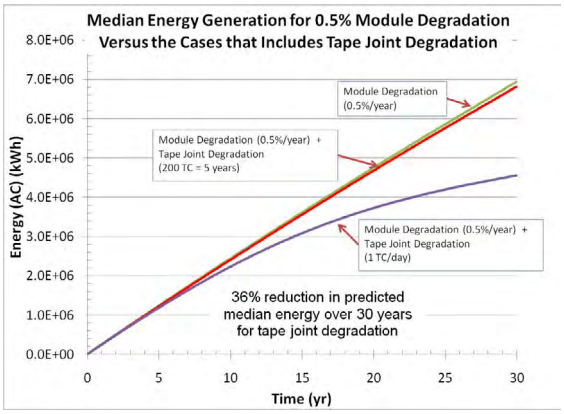

Figure 13. Effect of Tape Joint Degradation on Energy Production

3.4 Effects of Solar Irradiance Variability

We compared energy generation over 30 years assuming 0.5% degradation for 4 different locations: Flagstaff, Las Vegas, Miami, and New York City. Figure 14 illustrates cumulative energy production for those 4 locations for 30 years assuming a module degradation rate of 0.5% per year with no tape joint degradation.

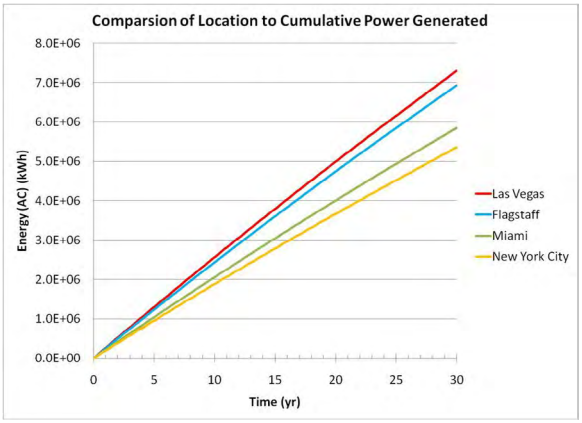

Figure 14. Effect of Solar Irradiance Variability Due to Location

Additionally, we compared energy generation over 30 years for the Flagstaff location using different degradation rates and degradation mechanisms. We also investigated the effect of holding irradiance variation constant for the scenario of the Flagstaff location with 0.5% per year module degradation and with no tape joint degradation. The sensitivity studies consisted of eight scenarios:

- Flagstaff, AZ location at 0.5% degradation (*This is the baseline scenario.*)
- Baseline scenario without weather variation
- Baseline scenario at 1% degradation
- Baseline scenario at 1.5% degradation
- Baseline scenario with tape joint degradation
- Las Vegas, NV location at 0.5% degradation
- Miami, FL location at 0.5% degradation
- New York City, NY location at 0.5% degradation

These scenarios are illustrated in figure 15. It is apparent that while irradiance variability is less significant for a particular location with respect to energy production, irradiance mean value at a particular location significantly affects energy production. Similarly, with all other factors held constant, either the presence of tape joint degradation or differences in module percent degradation significantly affect energy production.

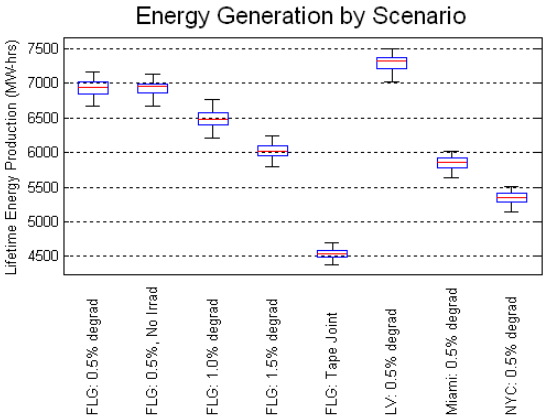

Figure 15. Boxplot Comparison of Energy Generation Scenarios

4 CONCLUSIONS

We demonstrated a simulation model that predicts energy output as a function of the parameters that were varied during the simulation, 11 in all. These were irradiance scaling factor, catastrophic failure distribution scale parameters for the six system hardware components in the simulation, module degradation rate, and inverter disturbance distribution scale parameters including grid disturbance, preventive maintenance, and lightning. This resulted in a five factor model with a correlation coefficient squared of 96%, which is the proportion of variability of the predicted energy explained by the model. Two of these factors (location and degradation rate) explained over 90% of the variation in the data.

The most influential factors on energy production are geographical location and photovoltaic module degradation modes. For all combinations of location and degradation, the highest sensitivity for variation for a particular geographical location for a particular photovoltaic module degradation mode was attributed to the module degradation rate, followed by irradiance factor. All other parameters that were varied together accounted for less than an additional 2% of the variation in kilowatt hour production. This suggests that variation in energy production is mainly dependent upon the degradation rate or degradation mode of the modules installed in the system followed by the irradiance factor. Surprisingly, energy production was not strongly sensitive to variations in inverter failure rate or inverter disturbances. Not surprisingly, the mean value of energy production is strongly affected by the geographical location of the system.

5 FUTURE WORK

Available computing resources limited the scope of this simulation to the basic power generating unit, which is the inverter array. This is 1 of 26 power generating units in the 4.6 MW photovoltaic generating facility in Springerville, Arizona. The simulations were performed in GoldSim™ using a PC with 8 cores. Each of the 8 energy generation scenarios took approximately 3 hours

for only 100 realizations per scenario simulation. The replicates show that 100 realizations is an adequate number.

Utilization of supercomputing or development of efficient behavioral models above the inverter-array level is necessary for timely simulations of "what-if" scenarios for utility-scale photovoltaic generating facilities. For example, instead of building a model with 26 identical generating units, it would be more efficient to build a behavioral model with 1 generating unit and simulate it 26 times, or more for larger utility installations. The simulation can be augmented with financial cost models and logistics policies to analyze scenarios for calculating levelized cost of energy metrics.

Actual operations data for failures, repairs, and other off-times is not readily available to better characterize the distributions in the reliability model. A single database like XFRACAS™ could support a life cycle approach for both data collection and managing conceptual development through retirement of fielded PV systems. It is capable of recording and tracking design iterations and upgrades including review boards, failure analyses, corrective actions, field operations, test results, and data for reliability growth and system availability predictions. Sandia has adapted XFRACAS™ for use as a photovoltaic database [6].

Finally, weather modeling improvements are needed to provide more realistic simulation of natural year to year variation.

6 ACKNOWLEDGEMENTS

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 funding the work presented in this paper.

The authors acknowledge Tucson Electric Power for sharing failure and maintenance data for the Springerville, AZ Photovoltaic Generating Facility. Also, we thank Tom Hansen and Kaleb Brimhall who assisted in interpretation of the data logs. This work would not have been possible without this valuable information.

7 REFERENCES

[1] King, D.L., Boyson, W.E., and Kratochvill, J.A. (2004): Photovoltaic Array Performance Model, Sandia Report SAND2004-3535.

[2] Reda, I. and Andreas, A. (2008): Solar Position Algorithm for Solar Radiation Applications, NREL Report TP-560-34302.

[3] Perez, Richard; Ineichen, Pierre; and Seals, Robert (1990): Modeling Daylight Availability and Irradiance Components from Direct and Global Irradiance, Solar Energy 44 (5): 271-289

[4] Elmer Collins, Michael Dvorack, Jeff Mahn, Michael Mundt, and Michael Quintana, "Reliability and Availability Analysis of a Fielded Photovoltaic System," presented at the 34th IEEE Photovoltaic Specialist Conference, June 2009.

[5] National Renewable Energy Laboratory (NREL), Solar Radiation Data Manual for Flat-Plate and Concentrating Collectors, April 1994, NREL / TP-463-5607 DE93018229.

[6] Elmer Collins, Jeff Mahn, Michael Mundt, Jennifer Granata, and Michael Quintana, "Field Data Collection for Quantification of Reliability and Availability for Photovoltaic Systems," presented at the 35th IEEE Photovoltaic Specialist Conference, June 2010.

[7] N. Robert Sorensen, Michael A. Quintana, Michael J. Mundt, Edward V. Thomas, Steven P. Miller, and Samuel J. Lucero "Accelerarted Testing of Metal Foil Tape Joints and Their Effect on Photovoltaic Module Reliability", 25th European Photovoltaic Solar Energy Conference and Exhibition, September 5-11, 2010.