P50/P90 Analysis for Solar Energy Systems Using the System Advisor Model

Preprint

A. Dobos and P. Gilman
National Renewable Energy Laboratory

M. Kasberg
Pariveda Solutions

Presented at the 2012 World Renewable Energy Forum
Denver, Colorado
May 13-17, 2012

Conference Paper
NREL/CP-6A20-54488
June 2012

Contract No. DE-AC36-08GO28308

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) PIX 16416, PIX 17423, PIX 16560, PIX 17613, PIX 17436, PIX 17721

Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

WREF 2012: P50/P90 ANALYSIS FOR SOLAR ENERGY SYSTEMS USING THE SYSTEM ADVISOR MODEL

Aron P. Dobos
Strategic Energy Analysis Center
National Renewable Energy Laboratory
Phone: 303.384.7422
Email: aron.dobos@nrel.gov

Michael Kasberg
IT Consultant
Pariveda Solutions
Denver, CO

Paul Gilman
Technical Consultant to the
National Renewable Energy Laboratory
Portland, OR

To secure competitive financing for a solar energy generation project, the economic risk associated with inter-annual solar resource variability must be quantified. One way to quantify this risk is to calculate exceedance probabilities representing the amount of energy expected to be produced by a plant. Many years of solar radiation and metereological data are required to determine these values, often called P50 or P90 values for the level of certainty they represent. This paper describes the two methods implemented in the National Renewable Energy Laboratory's System Advisor Model (SAM) to calculate P50 and P90 exceedance probabilities for solar energy projects. The methodology and supporting data sets are applicable to photovoltaic, solar water heating, and concentrating solar power (CSP) systems.

1 Introduction

The economic value of a solar energy generating facility depends on the availability of the solar resource. The solar radiation, and to a lesser extent, temperature, humidity, atmospheric pressure, and wind speed determine the timing and quantity of energy the facility generates. Weather-related events such as passing clouds or storms introduce resource variability at short time scales, and larger-scale events such as volcanic eruptions and climate cycles introduce variability on larger, inter-annual time scales. In their assessment of the value of a solar energy project, financial institutions use statistical methods to determine the likelihood that a power plant will generate a certain amount of energy in any given year over the plant's 20- to 30-year life.

Exceedance probabilities have been widely used in the wind industry to describe the probability that a particular location will experience sufficient wind speeds for a proposed wind farm to be financially sound. Banks and investment firms working on wind farm projects often require P50 and P90 values of the wind resource at a location to determine the risk associated with a project's ability to service its debt obligations and other operating costs. Although the solar resource is generally more predictable than wind [7], the exceedance probability risk assessment approach can also be applied to solar energy projects.

Statistically robust estimates of energy generation exceedance probabilities require many years of resource data, as well as sufficiently detailed system performance models. The System Advisor Model (SAM) is a free software application produced by the National Renewable Energy Laboratory that performs rigorous solar power plant system performance modeling and calculates detailed financial cashflows. SAM calculates metrics such as annual energy output, capacity factor, levelized cost of electricity (LCOE), internal rate of return (IRR), and others for flat-plate and concentrating photovoltaic generators, as well as various configurations of concentrating solar power (CSP) systems including parabolic trough, power tower, dish stirling, and linear Fresnel, with and without thermal energy storage (as applicable).

This paper describes SAM's P50/P90 analysis capability, starting with the available long-term weather datasets, details of the calculation methods for P50 and P90, and explores two representative analysis scenarios for a utility scale flat-plate PV system and a solar power tower system.

2 Solar Radiation and Weather Data

Some solar energy simulation software use files from the Typical Metereological Year (TMY) datasets [1, 2] as input. TMY files are available for many locations in the United States, making them suitable for use in simulation models like SAM. The TMY2 and TMY3 datasets consist of a file for each location, and each file contains hourly data derived from long term measured data. The data are processed by choosing "typical" months to represent the long term properties of the data. Data representing months during outlier events such as large volcanic eruptions are excluded to ensure that the file represents the long-term climate. For project financial analysis, these outlier events may result in worst-case years that affect the project's financial terms. Because the TMY files do not include data from these potential worst-case years, they may be more appropriate for preliminary analysis of a system design than for financial decisions. Using multi-year historical data to model the long term performance of a system instead of a TMY file ensures that the performance prediction accounts for potential worst case years [6].

The National Solar Radiation Database (NSRDB) is a long term hourly dataset of measured and modeled solar radiation for hundreds of locations in the United States. Each location in the database may have hourly records from 1961-1990, or 1991-2005, or both, meaning that some locations have a full 45 years of data. While solar radiation is the primary driver of a PV or CSP plant's energy projection, metereological weather data (temperature, pressure, humidity, etc) also have secondary but not insignificant impacts on a system's performance. The National Climatic Data Center (NCDC) maintains a corresponding database of the meterological inputs required by the SAM models, and can be obtained with the solar radiation data for a nominal fee. To make robust P50/P90 analysis available to a wider audience, NREL partnered with the NCDC to make the dataset available without charge in an encrypted form specially designed for use in SAM. While the raw data is not accessible to the user, simulations can be performed using the dataset directly in SAM.

The long term NCDC/NSRDB dataset includes the impact of large volcanic activity and other phenomena that occur on timescales larger than one year. In particular relevance to solar plants, the eruption of Mt. Pinatubo introduced large quantities of aerosols into the atmosphere that reduced incident irradiance levels between 1991 and 1993. Other variations include the cyclic El Niño and La Niña phenomena, as well as the 11 and 22 year sun spot cycles.

3 Methodology

An exceedance probability is the probability that a certain value will be exceeded. For example, a P50 value of 10,000 kWh for the annual output of a solar power system means that there is a 50 % likelihood that the system's output will be greater than 10,000 kWh. Similarly, a P90 value of 10,000 kWh would mean that the system is likely to generate over 10,000 kWh 90 % of the time. [6,7]. For analysis of solar energy power plants, the P50 and P90 values of annual annual electricity generation and of the LCOE can both provide useful information for financial analysis of a proposed project.

One method of calculating exceedance probabilities is to fit the dataset to a standard probability distribution, and to calculate the P50 or P90 value from the distribution's cumulative distribution function (CDF). This method is used by [3, 8] and works well when the data is normally distributed. In this case, the P50 value is by definition the mean value μ. The P90 value can be calculated from the CDF of a normal distribution, which is defined by the function in Eqn. 1.

$$\Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x-\mu}{\sigma}\frac{1}{\sqrt{2}}\right)\right] \tag{1}$$

The P90 value occurs when $\Phi((x-\mu)/\sigma) = 0.1$. Defining $\gamma = (x-\mu)/\sigma$, Eqn. 2 can be solved numerically.

$$\Phi(\gamma) = 0.1 \rightarrow \gamma = -1.282 = \frac{x-\mu}{\sigma} \tag{2}$$

Rearranging terms results in Eqn. 3, which is an expression for the P90 value given the mean (μ) and standard deviation (σ) of a dataset that is assumed to fit a normal distribution.

$$x = \mu - 1.282\sigma \tag{3}$$

A second method works best when the data are not normally distributed, as is the case with solar resource data over many years [7], where outlier events such as volcanic eruptions and cyclic solar patterns can skew the data. In this case, no particular statistical probability distribution is assumed to fit the data, and rather an empirical CDF of the data is used to calculate the P50 and P90 values. The empirical CDF is determined by sorting the data in ascending order, and assigning each data point an equal fraction of the total probability, which is equal to one. The estimation procedure is shown in Table 1 for a very small 5 point data set.

Data Value	Estimated CDF
32,457	0.2
34,330	0.4
37,302	0.6
38,451	0.8
39,307	1.0

Table 1. Example Empirical CDF Calculation

It is clear from Table 1 that many more data points are required to establish a representative empirical CDF curve from which to interpolate exceedance probabilities. Supposing that more that 10 values are provided, the P90 value is calculated by linearly interpolating the table to obtain a data value for which the empirical CDF would equal 0.1. Figure 1 shows both the estimated CDF of the irradiance components in Phoenix, AZ, as well as the normal CDF calculated from μ and σ. The plot exposes many deviations from the normal distribution, and shows why interpolating exceedance probabilities directly from the empirical CDF is a more reliable approach, provided the dataset is sufficiently large. The P50 and P90 points using each calculation method are plotted as well.

The P90 value calculated from the normal distribution is greater than the empirical value for all three of the solar radiation components for the 30 year Phoenix, AZ dataset. Because the data do not fit a normal distribution very closely, the empirical method likely yields a better estimate of the true P90 value. It is just by chance that the P90 values for the diffuse irradiance are nearly identical, since by inspection a P80 value would yield quite different results from the two methods.

SAM calculates P50 and P90 values of the annual energy output using both methods. It performs an hourly simulation for each year in the dataset to calculate the system's output for each year. For the first method, it calculates the CDF of the normal distribution from the mean and standard deviation of the values, and for the second method, it calculates the empirical CDF. For both methods, SAM then determines the P50 and P90 exceedance probabilities either directly from the normal CDF equation or by linearly interpolating the empirical CDF table.

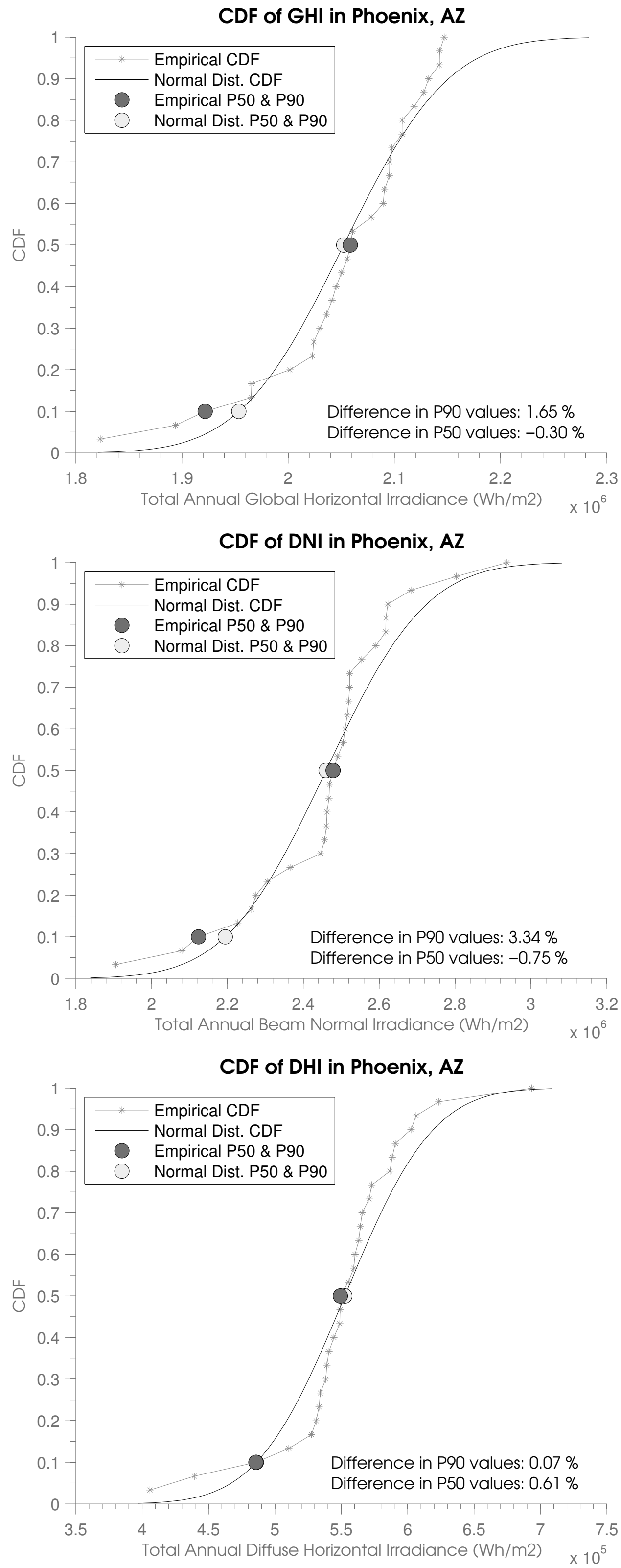

Fig. 1. CDF of Irradiance in Phoenix, AZ, 30 years

4 Example Scenarios

In this section, two hypothetical solar power plant scenarios are investigated using weather data for Phoenix, Arizona. The simulations were performed with SAM version 2011.12.2.

The long term datasets available in the public version of SAM 2011.12.2 contained several years of invalid metereological data and prohibited some simulations from completing successfully or gave erroneous results. For these examples, files for years with problematic data were removed from the Phoenix 45-year dataset, so the results are based on 30 years of data. The modified weather dataset is available for download on the SAM website [10]. Efforts are underway at NREL to better characterize potential errors and missing data in long term historical solar and metereological datasets.

4.1 20 MW Utility-scale PV System

The annual solar energy generation of a 20 MW utility-scale PV system is considered in Phoenix. Three system configurations are investigated. Two flat-plate systems with fixed and 1-axis tracking are modeled using SAM's California Energy Commission (CEC) 5 parameter PV module model, and a concentrating PV system is modeled using the concentrating PV simple efficiency model. The system parameters are shown in Table 2. For all three systems, the soiling derate is assumed to be a constant 0.95, and total system derate is 0.89.

Figure 2 shows CDFs calculated using both exceedance probability methods for each system configuration along

Parameter	Fixed	1 Axis	CPV
Nameplate size	20 MW	20 MW	20 MW
Tracking	Fixed	1 Axis	2 Axis
Tilt	33 °	0 °	n/a
Azimuth	South	South	n/a
Module efficiency	16 %	16 %	27 %

Table 2. Example PV System Specifications

with the P50 and P90 values. For each configuration, the normal distribution method yields a P90 value that is a few percent more optimistic than the empirical CDF method. The P50 values calculated using each method are in better agreement. Table 3 lists the P50 and P90 annual energy values from both methods and the annual energy value calculated using the TMY2 weather file for Phoenix. For all of the systems, the annual energy value from the TMY2 file is slightly greater than the P50 value. The TMY2 and P50 values would perhaps be in better agreement if the TMY2 files were weighted only for solar radiation rather than for both solar radiation and meteorological conditions [1].

Metric	Fixed (MWh)	1 Axis (MWh)	CPV (MWh)
P50*n*	36221	45205	46507
P90*n*	34182	42264	41549
P50*e*	36533	45844	46695
P90*e*	33604	41070	40348
Min	31658	39430	36145
Max	38442	49487	55465
TMY2	37323	46573	47629

Table 3. PV System Annual Energy Metrics. *n* denotes P-value from normal distribution, *e* denotes empirical P-value

4.2 100 MW Solar Power Tower

Two solar power tower configurations are considered, also in Phoenix. The first plant has a solar multiple of 2 with 6 hours of thermal energy storage, and the second one has a solar multiple of 1.2 with no thermal energy storage. Key parameters are listed in Table 4.

Table 5 lists P50 and P90 annual output values and the output value from the TMY2 file for the two power tower plants in Phoenix. As with the PV scenario, the TMY2 weather file gives a higher annual output than the P50 value of the long-term dataset. From figure 3, it appears that the addition of thermal energy storage does not significantly affect the shape of the empirical distribution. For both systems, the normal distribution P90 exceeds the empirical value on the order of 2 %, similar to the PV examples.

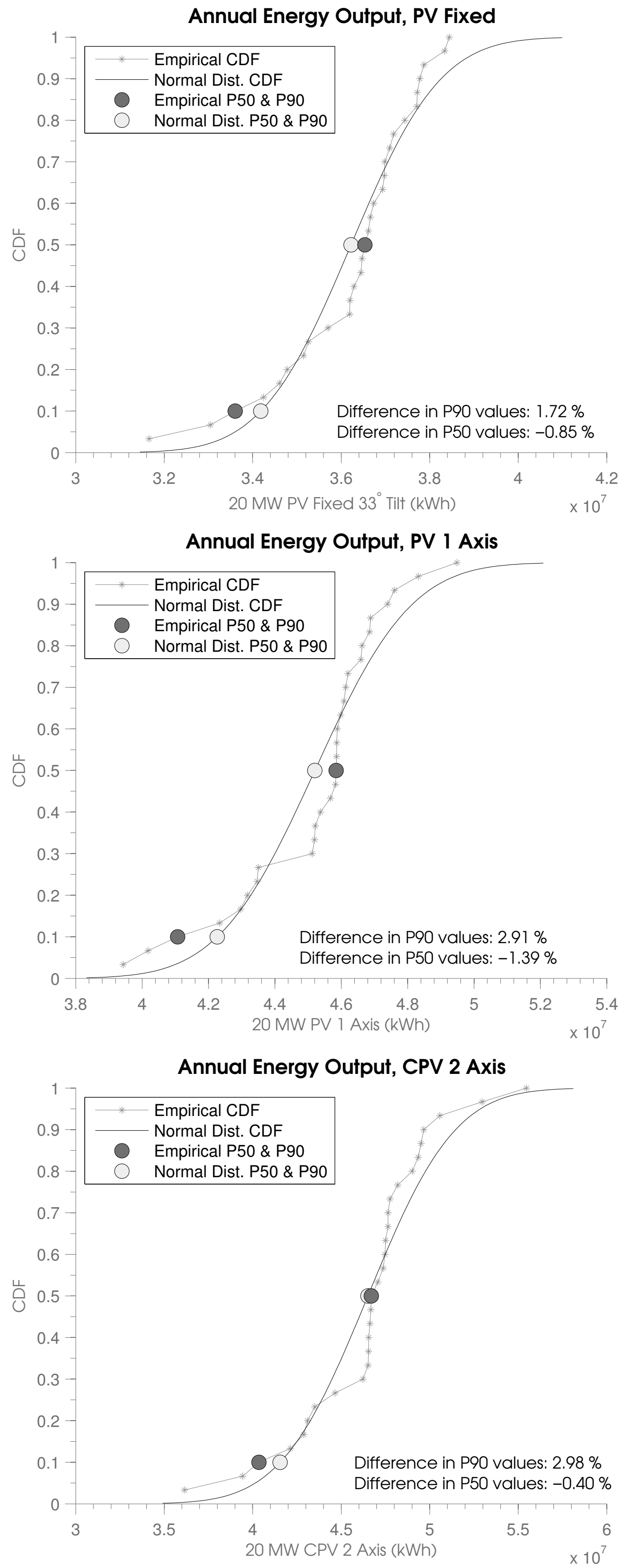

Fig. 2. PV System CDF

Parameter	Value
Nameplate capacity	100 MWe
Solar multiple	2 or 1.2
Tower height	184 m
Thermal storage	Two tank, 6 hrs or none
Cooling	Wet

Table 4. Example Power Tower System Specifications

Metric	6 hrs TES (MWh)	No Storage (MWh)
P50*n*	398956	224688
P90*n*	351651	196678
P50*e*	404301	226673
P90*e*	343603	193209
Min	297142	164569
Max	479520	273566
TMY2	410484	232358

Table 5. Solar Power Tower Annual Energy Metrics. n denotes P-value from normal distribution, e denotes empirical P-value

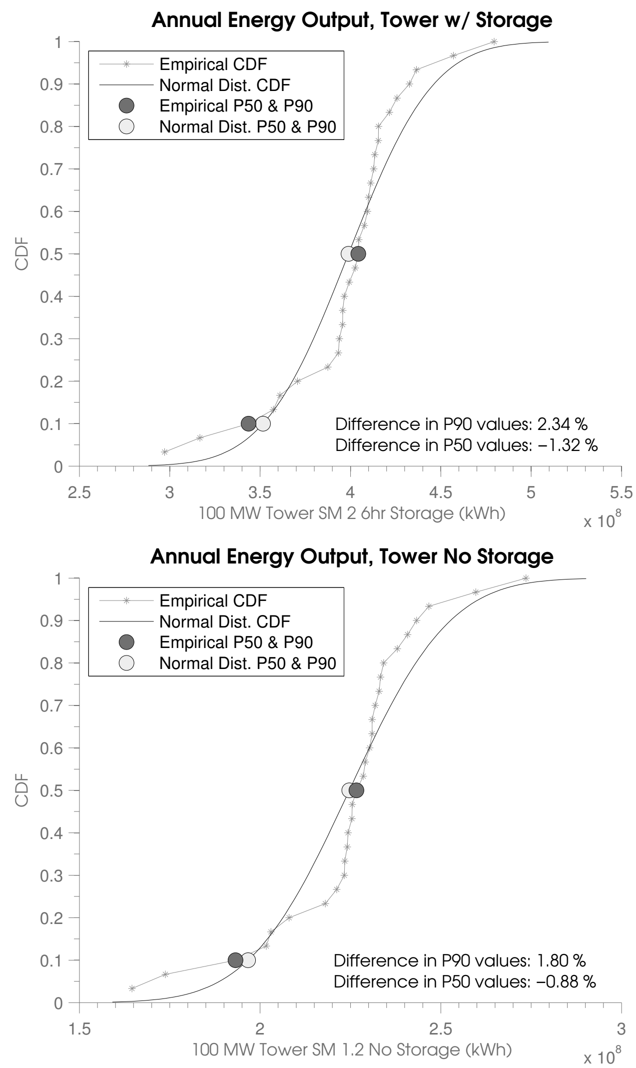

Fig. 3. Power Tower CDFs

4.3 Case Analysis

Figure 4 shows the percentage difference between each year's calculated annual energy output and the 30 year minimum value. Because the fixed PV system uses both direct and diffuse radiation, it shows the least variation, while the CPV and power tower systems show the most variation. The worst case years show the effect of large volcanic eruptions in 1982 (El Chichon, Mexico), 1991 (Pinatubo, Philippines), and of El Nino in 1997-1998. The statistical distributions used to create the TMY datasets do not include these worst case years.

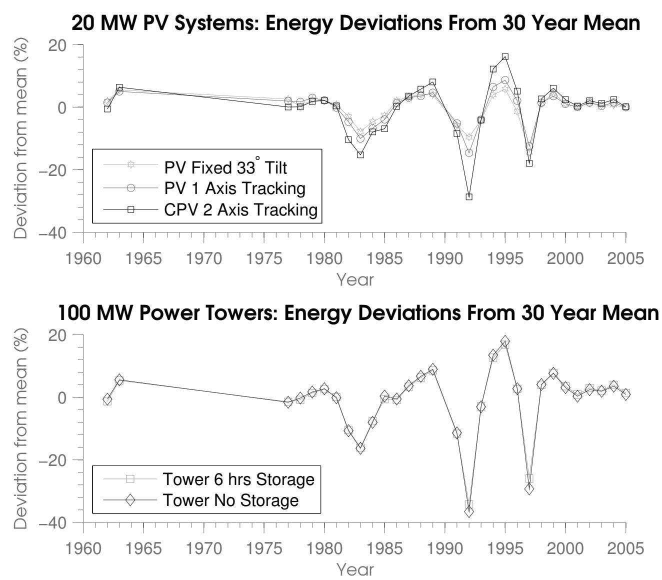

Fig. 4. Percentage deviations from long term mean values for both technologies.

Two metrics, Δ_{EX} and Δ_P, are defined for the purposes of this paper in Eqn. 4. Δ_{EX} represents the percentage difference between the mean and worst case scenarios (i.e. extreme minimum value), while Δ_P looks at the deviation between the P50 and P90 values.

$$\Delta_{EX} = 100 \cdot \frac{E_{ann,mean} - E_{ann,min}}{E_{ann,min}} \quad \Delta_P = 100 \cdot \frac{P50_e - P90_e}{P90_e} \tag{4}$$

Both of these metrics are plotted in Figure 5 for the solar radiation components and all system configurations considered.

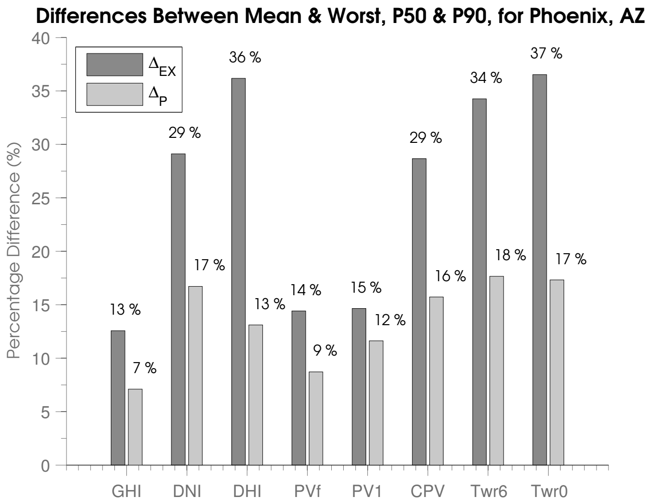

Fig. 5. Percentage difference between mean and worst case scenarios.

The flat plate PV systems exhibit the smallest Δ_{EX} values, because they capture both diffuse and beam radiation. One would therefore expect the variation in the output of the PV systems to be similar to the variation in the global solar resource. As a result, the fixed and 1 axis tracking systems' Δ_{EX} is closest to that of GHI. For the CPV system, which depends solely on beam irradiance, on the other hand, the Δ_{EX} is identical to the Δ_{EX} for DNI. The Δ_P values follow the same pattern, but are lower than the extreme case, because they effectively filter out the extreme years, which are only 3 of the 30 years.

For the power tower, Figure 5 shows that the system with storage shows less variation between the mean and worst case scenarios (i.e. smaller Δ_{EX}). In both tower configurations, Δ_{EX} is slightly higher than Δ_{EX} for DNI, likely due to the delay between changes in DNI and the system's output caused by thermal inertia, and startup and shutdown times. Because the concentrating solar thermal plants cannot utilize diffuse radiation, their Δ_{EX} degree of variation is consistent with that of the CPV system and DNI resource, but greater than the flat plate PV system and GHI resource.

This analysis of exceedance probabilities of the solar resource and simulated energy output provides insight into uncertainty in the expected energy production of a system. Further analysis on other metrics could account for other important factors not considered in this paper, for example, the time-dependent value of energy that is particularly relevant to systems with energy storage [11]. For example, SAM could be used to generate exceedance probability values of other annual metrics such as capacity factor, revenue, LCOE, NPV, IRR, and others.

5 Conclusion

Estimates of the amount of energy a solar power plant will generate in the future are necessary to determine a project's financial risk. P50 and P90 exceedance probabilities based on many years of historical weather data are useful for this determination. The case studies in this paper show that computer simulations based on TMY datasets representing the solar resource over a multi-year period can overpredict the P50 value over the same period. For locations where sufficient data are available, SAM can be used to compare predictions using TMY files and P50 or P90 values calculated from multi-year data sets.

Achnowledgements

This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08-GO28308 with the National Renewable Energy Laboratory.

References

[1] Marion, W., Urban, K. *Users's Manual for TMY2s*, NREL, 1995. *http://rredc.nrel.gov/solar/pubs/tmy2/*

[2] Wilcox, S., Marion, W., *Users Manual for TMY3 Data Sets*, NREL Technical Report NREL/TP-581-43156, 2008.

[3] 3Tier, *Site Climate Variability Analysis (Seattle, WA)*, 3Tier, Seattle, 29 Nov 2010

[4] 3Tier, *Site Resource Assessment (Desert Rock, NV)*, 3Tier, Seattle, 2011

[5] 3Tier, *3Tier Global Solar Dataset: Methodology and Validation*, 3Tier, Seattle, 2010

[6] Vignola, F., Grover, C., Lemon, N., *Building a Bankable Solar Radiation Dataset* Proc. Solar 2011, American Solar Energy Society, Raleigh, NC, 2011

[7] Zempleni, M., Davidson, A., *PV Solar Power Generation Projects*, Moody's Investor Service, 2010

[8] Crawley, N. *Modelling P50 versus P90 wind profiles*, Online, *http://www.navigatorpf.com/blog/modelling-p50-versus-p90-wind-profiles*, 2009

[9] Thevenard, D., *Uncertainty in Long-Term Photovoltaic Yield Predictions*, Natural Resources Canada, Varennes, 2010

[10] National Renewable Energy Laboratory, *System Advisor Model*, http://www.nrel.gov/analysis/sam

[11] Sioshansi, R., Denholm, P., *The Value of Concentrating Solar Power and Thermal Energy Storage*, IEEE Transactions on Sustainable Energy, vol.1, no.3, pp.173-183, Oct. 2010 doi: 10.1109/TSTE.2010.2052078