Quantifying the Variability in Solar PV Production Forecasts

ASES National Solar Conference
May 17-22, 2010

Steve R. Dean, ASA, P.E.
DAI Management Consultants, Inc.
1370 Washington Pike
Bridgeville, PA 15017

(412) 220-8920
sdean@daimc.com
www.daimc.com

Introduction to DAI

Who is DAI?

- Founded in 1987 as a full-service consulting firm dedicated to the power and energy infrastructure industries
- Over the past two decades, DAI has worked in virtually every area of the power sector, advising lenders, investors, developers, utilities, regulators, legal experts, and government organizations.

- Renewable Industry Experts
  - Advisors to institutional investors in renewable assets
  - Advisors to firms developing new renewable power plants
- Energy Market Experts
  - Industry-leading clients
  - University-affiliated experts at the Carnegie Mellon University Electricity Industry Center
  - Published, peer-reviewed research
- Appraisal & Valuation Specialists
  - ASA-accredited senior appraisers
- Power & Energy Market Engineers
- Electric Market Economists
- Plant Managers & Operators
  - Gas Turbine Combined Cycle (DAI Oildale)
  - Hydroelectric (DAI Great Falls)

Recognized Expertise

- DAI's staff includes economists, engineers, financial analysts, statisticians, and operations experts

- American Society of Appraisers Certified
- Licensed Professional Engineer by National Council of Examiners for Engineering
- Published, peer-reviewed research
  - *The Appraisal Journal*
  - *Journal of Structured and Project Finance*
  - *Journal of Economic Behavior and Organization*
  - *Public Utilities Fortnightly*
  - *The Electricity Journal*
  - *Nuclear Engineering International*
  - *Renewable Energy World*
- 2008 Pittsburgh 100: Fastest-growing Companies

Decision Analysis

- Quantitative Risk Analysis (“QRA”)
- Electric and Fuel Market Studies
- Electric Market Forecasts
- REC Market Forecasts
- Fuel Market Forecasts
- Statistical Analysis of Asset Performance
- Hedging Strategy Analysis
- Analysis of Capital Cost Uncertainty
- Collateral Analysis for Loan Guarantees
- Acquisition and Divestiture Advisory
- Valuation Litigation Support

Appraisal & Valuation

- Equipment Fair Market Value Appraisal
- Residual Value Determination
- Liquidation Value Determination
- Tax Analysis/Support
  - Alternative Energy Property Allocations
  - Business Combinations (SFAS 141)
  - Goodwill and Intangible Assets (SFAS 142)
  - Gain or Loss from Acquisition (IRC 1060)

Engineering Consulting

- Independent Engineer Analysis
- New Technology Commercialization
- Power Plant Operation and Maintenance
- Portfolio Management
- Feasibility Studies
- Expert Witness Testimony/Consulting

Asset Management

- Acquisition and Divestiture Support
- Operations Efficiency Analysis
- Outage Management
- On-Site Operations and Maintenance
- Environmental Compliance Review
- Permitting, Licensing, and Accounting

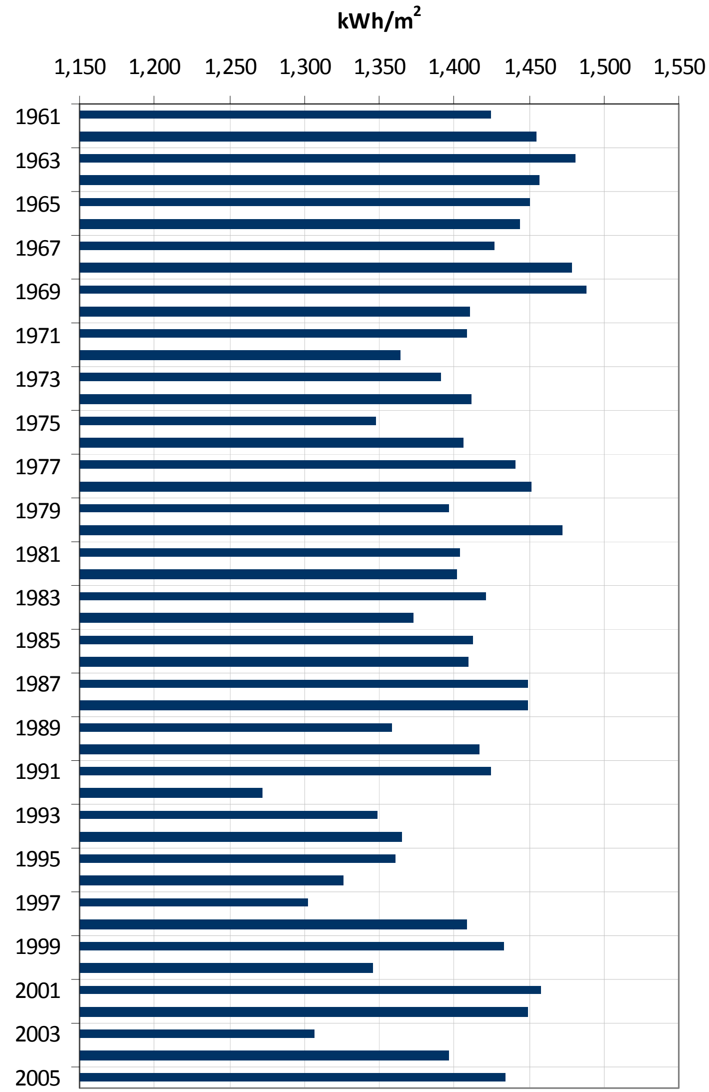
Quantifying the Variability in Photovoltaic Power Production

Annual Solar Insolation Varies Significantly

- We start with a rather basic premise: solar variability matters.

- However constant we think the sun is, solar insolation actually varies significantly over time.
- Our concern is purely short- and intermediate-term variation; we don't address long-term structural factors that may be caused by climate change and global dimming.
- The primary causes of this short-term variation are almost purely random: weather, cloud cover, and infrequent extreme events (*e.g.*, volcanic eruptions).
- Even as statistical "noise," this variation presents problems for PV investors and lenders who may have cash flow constraints for debt service or lease payments.

Annual Global Insolation (kWh/m²), Newark, New Jersey, 1961-2005

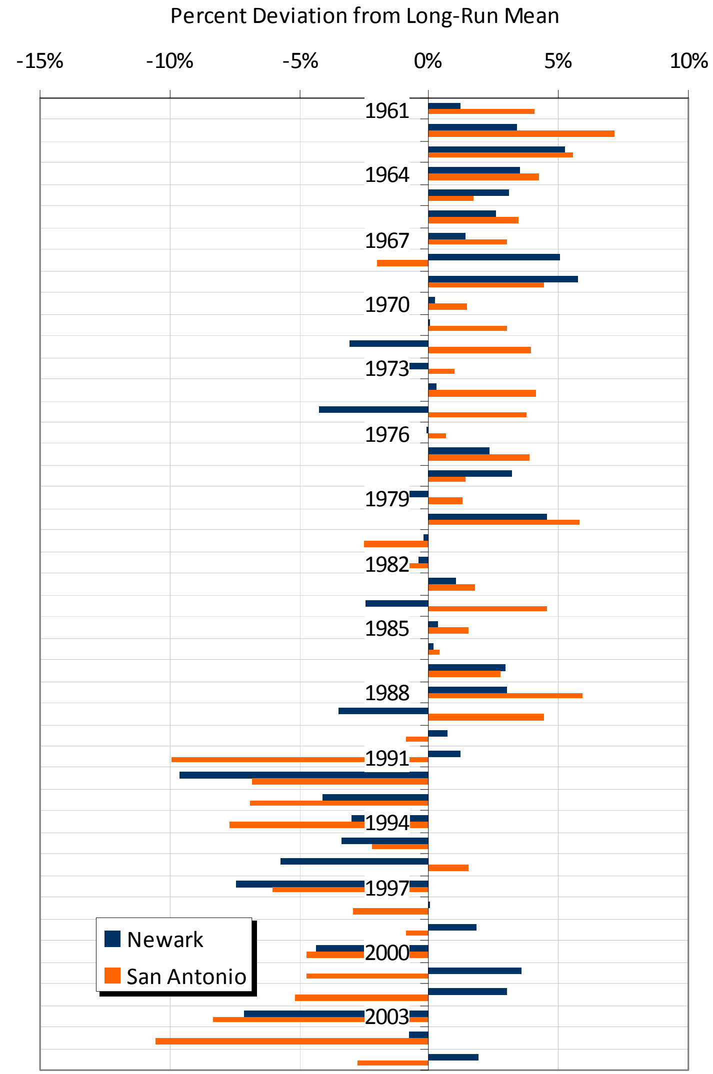


DAI MANAGEMENT CONSULTANTS

Geographically-Varying Variability

- Not only does insolation vary over time, it also varies by geographic location. There is no one-size-fits-all level of variability that one can assume to reflect this "noise"
- The variations tend not to be strongly correlated (here, the correlation is 0.44), and relative variability can be quite different.
  - San Antonio's standard deviation of deviations from the long-run mean is 30% higher than Newark's
- Models like PVWatts that advocate "blanket" error levels (*e.g.*, 20% for individual years) miss this geographic variability.
- More importantly, they lack the ability to ascribe probabilities to errors of any given magnitude. For example, is a 5% error just as likely as a 15% error? Is it twice as likely? Are positive errors more likely than negative errors?

Annual Deviations in Global Insolation, Newark vs San Antonio



DAI Management Consultants

Financing Implications of Variation in Insolation

• Having established the variability of solar insolation, we turn to its implications for investors and lenders in solar power projects.

• Insolation variability drives power production, and thereby revenue variability. With costs largely fixed, this variability drops directly to the bottom line: cash available for debt service and cash flows to equity.

• Traditionally, accounting for variability has involved setting up reserve accounts for debt service to help smooth out cash flow fluctuations. Additionally, many lenders are only willing to offer financing based on revenue modeled at the 90^{th} percentile (or 95^{th}, 99^{th}, etc) of projected production.

• There is a constant tension, however, between lenders and investors and over- and under-mitigating cash flow variability.

- If risks are over-mitigated, returns suffer and good projects may be unjustly rejected.
- If risks are under-mitigated, projects may unnecessarily experience financial distress.

• The traditional model of financing projects assuming n^{th} percentile production is eminently reasonable – provided that the probabilistic model of production is itself reasonable. Here, we encounter the core of our analysis: ***how can this production variability be quantified probabilistically? Is the n^{th} percentile actually the "true" n^{th} percentile?***

Probable vs Typical

• Most locations that are targeted for solar power projects do not have extensive historical records showing actual production. Instead, project stakeholders must rely on models (and their underlying data) to develop probabilistic forecasts of future production.

• The question, then, is: is the 90th percentile of the estimated probability distribution sufficiently close to the underlying true (but unobserved) distribution?

- Is the production model well-calibrated?
- Is the underlying historical data an accurate representation of historical experience?

• There are a variety of production models (PVWatts, PVsyst, Polysun, etc), but generally there is one common source of underlying historical data: The National Solar Radiation Database, as expressed in the TMY2 (or TMY3) files.

• The TMY databases refer to the “Typical Meteorological Year” and are the common measure of historical experience for solar power models.

• TMY2 is a dataset of hourly values of solar radiation and meteorological elements for a one-year period that defines a “typical” year for a particular location. It includes global horizontal radiation, direct normal radiation, dry bulb temperature, dew point temperature, and wind speed for more than 230 locations across the country from 1961 to 1990. A similar TMY3 database covers the period 1976 to 2005 and covers more than 1,000 locations.

• A typical meteorological year is constructed via an algorithm that selects the most typical month of the thirty years in each database. The algorithm minimizes the difference between the year in question and the long-run average for each parameter. The parameters are then weighted according to their importance to the determination of locational solar resource availability. Certain data is then excluded on the basis of persistence or unusual occurrence.

• Having identified the “most typical” month for each of the twelve months, they are concatenated to produce the typical year.

• This concatenation of data is <u>not</u> a probabilistic representation of historical experience.

Is "Typical" Central?

• The use of the "typical year" as the production forecast implies a degree of probabilistic meaning that does not actually exist. There is no statistical reason to believe that the "typical" forecast level is the mean, median, or modal production level.

• In short, "typical" is not central (in the statistical sense). A "TMY" is not like an actual year.

• As a result, probabilistic assessments based on "typical" data (like TMY2 and TMY3) are not accurate.

• Consider some basic concerns that reveal problems with the "typical year" model:

- The data selection process in TMY2 is asymmetric with regard to outliers. Only years in which solar output is *adversely* affected by unique circumstances are excluded; unusually "good" years are not excluded.
- Months are constructed by a weighting of parameters related to solar resource availability, rather than actual experience.
- Years are constructed in a manner that neglects the potential for sequential dependence. One especially "cold" month, for example, may more likely be followed by another "cold" month.
- The likelihood of an entire year of "typical" months may be overstated.

• Because years are constructed as the concatenation of twelve typical months, no unusual periods can *ever* be incorporated into the forecast. Moreover, since it is a construction of months culled from different years, this "typical year" also reflects a year that has *never* actually occurred – a seemingly strange outcome for something deemed "typical."

Can Historical Data Be Used to Produce a More Probabilistically-Appropriate Forecast?

• If TMY data is not suitable for probabilistic analysis, can the actual historical data be used to develop a more representative distribution of production experience?

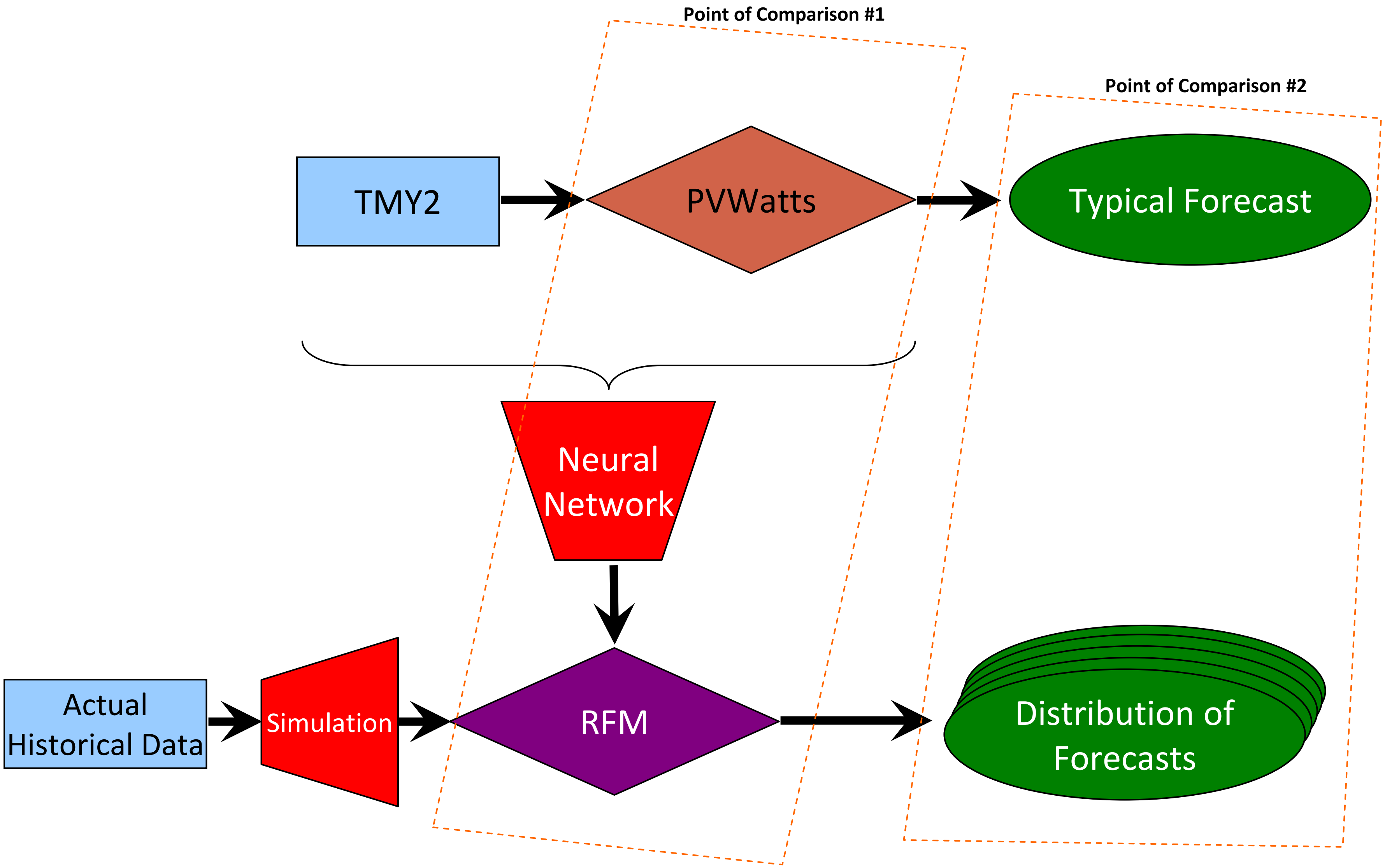
• Focusing exclusively on the underlying data, we use a quantitative risk analysis approach to explore the impact of parametric uncertainty on solar production.

• We use a neural network to create a reduced-form model ("RFM") of PV production as a function of certain key input parameters. Then, we use stochastic simulation to estimate the probability distribution of PV production based on distributions of input parameters estimated from actual historical data (rather than a "typical" year).

• The technical details are in our paper, but the basic process is as follows:

1. Create a neural network using six inputs and one hidden layer to estimate power production.
2. Train the network based on data from TMY2. The fully-trained network is then able to replicate PVWatts accurately within a single (albeit complex) analytic function. This function is the RFM.
3. Using the RFM, we can replace the static input parameters with probability distributions to facilitate simulation of PV production.
4. Instead of relying on TMY data, we develop probability distributions of the input parameters directly from the underlying data without reference to a "typical" year. These probability distributions also preserve the covariance structure present between the input parameters.
5. Using these historically-derived inputs, we then use Monte Carlo simulation to estimate the distribution of power production.

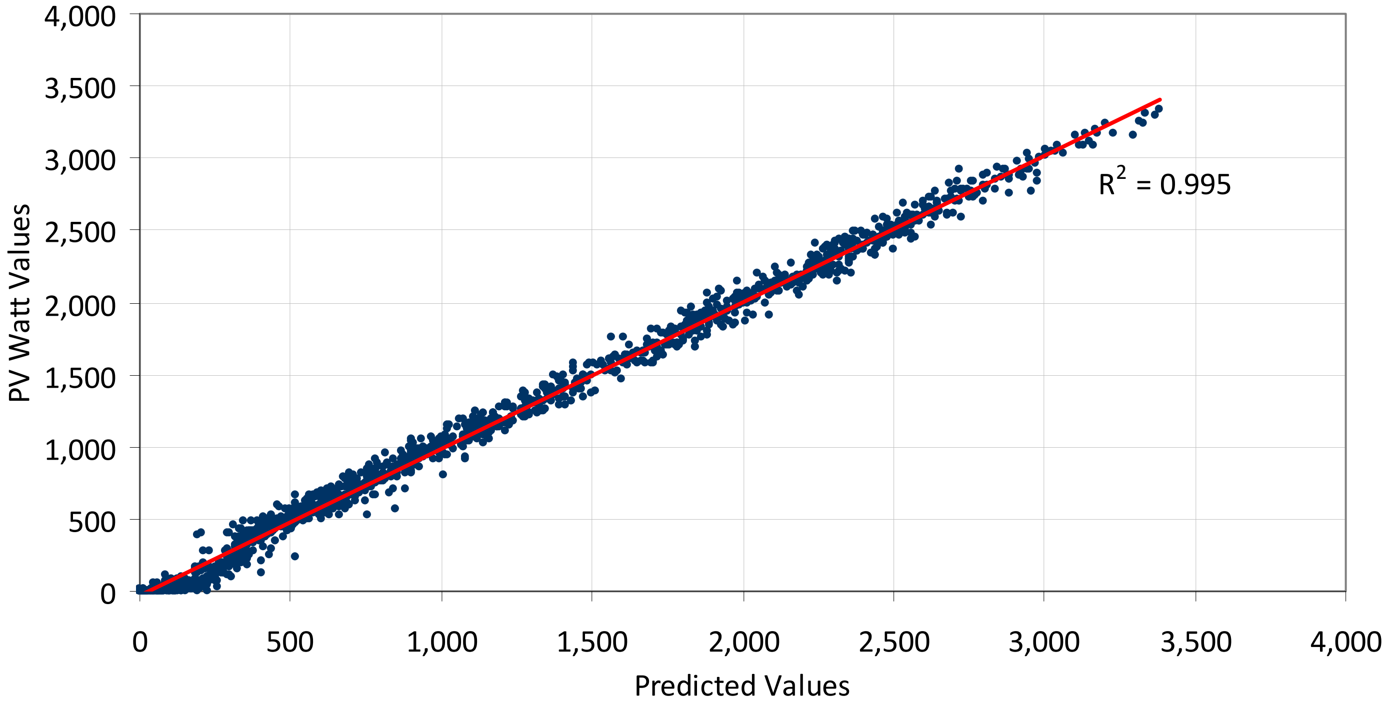
• The remaining examples are based on a 4 kW reference system located in Newark, NJ, with a fixed tilt of 40.7° (latitude). We compare our results to those produced by PVWatts for a system with the same characteristics.



DAI Management Consultants

Point of Comparison #1: Does the RFM Faithfully Reproduce the Underlying Model?

- Since our simulation model works through the production model (the RFM), it is important that the neural network has allowed us to capture the essence of PVWatts – a faithful, unbiased replication of the underlying model is essential.
- Fortunately, the performance of the neural network is exceptional in allowing us to replicate PVWatts, as the graph below illustrates. The RFM captures 99.5% of the variation produced by PVWatts.

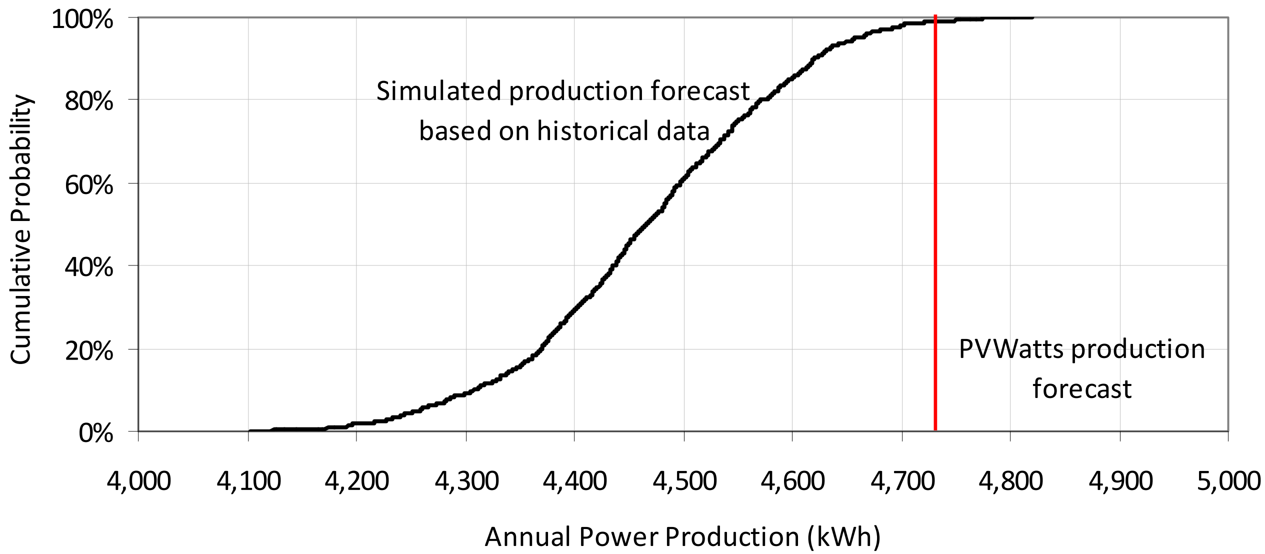


DAI Management Consultants

Point of Comparison #2: Do the Production Forecasts Compare?

• At the second point of comparison, we actually expect a different result. Our claim is that use of TMY data (instead of "raw" historical data) produces biased estimates of the likelihood of certain production outcomes. As a result, we would expect there to be a meaningful difference between the TMY-based production forecast and the center of our probabilistic production forecast that uses actual historical data.

• Since we have verified that our model functions nearly identically to PVWatts, by using raw historical data instead of TMY-adjusted data, we can examine whether or not the production forecasts are probabilistically-equivalent.



• They are not.

• PVWatts produces a "typical" production forecast of approximately 4,725 kWh – a level that corresponds to the 99th percentile of the underlying distribution and 6% higher than the median production level calculated from raw historical data.

DAI Management Consultants

Explaining the Origin of the Variability Bias

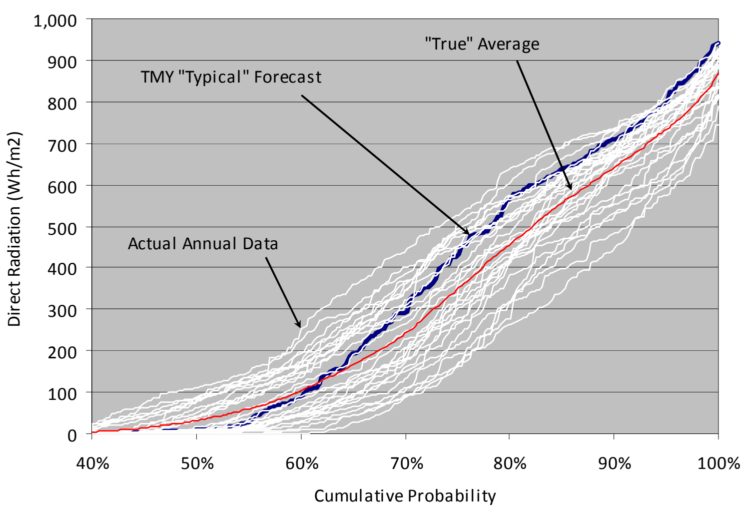
• We have previously suggested several reasons why the TMY data may bias probabilistic estimates of production. By examining the distribution of TMY production relative to the actual historical data, we can observe this bias directly.

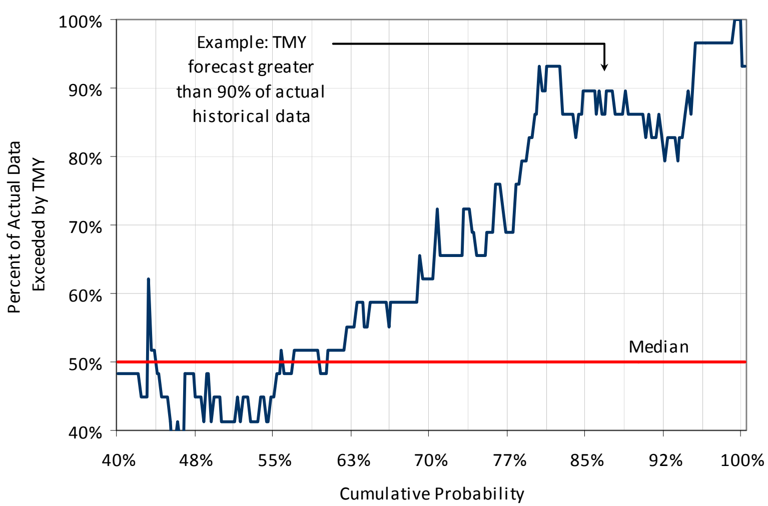
• In the first panel below, the white lines represent the actual data underlying each year from 1976 to 2005 (the period covered by TMY3 data). The data are presented as a cumulative distribution function, or the proportion of hours during the year with direct radiation less than or equal to a given level (the function starts at ≈40% because the remaining data points represent evening hours).

• The thin red line represents the true statistical average. In contrast, the thick blue line represents the "typical" year used by TMY3. The fact that these lines do not overlap makes clear that "typical" is most certainly not average or central. More importantly, the TMY line is almost uniformly above the "true average" line, suggesting a degree of persistent bias.

• The second panel below highlights this bias by plotting which percentile in the actual data is equivalent to the typical year. Instead of remaining close to the median, the typical year if most often significantly higher than the majority of actual experience.

• Modeling approaches that make direct inferences based on the TMY data are at risk of making misleading or inaccurate probabilistic representations of overall production.





DAI MANAGEMENT CONSULTANTS

Accurate Measurement of Solar Production Variability is Essential to Lender and Investor Risk Management

• Many aspects of PV investment depend on accurate measurement of production and production uncertainty.

- Lenders may determine debt-sizing based on certain production levels
- Lenders may seek to set debt service reserve account levels based on perceived cash flow variability
- Equity investors may face liquidated damages claim risk based on underdelivery against PPAs
- Equity investors may realize lower returns than would be appropriate given the risks incurred

• Although one-size-fits-all estimates of production variability (*e.g.*, ±20%/year) may provide a quick way to evaluate risk, the "plus-and-minus" approach only works as intended if it is applied around the *true* center of the distribution (in other words, if the TMY was actually a central measure of production).

• Our analysis suggests it is not a reliably central measure of production. As a result, probabilistic inferences obtained via a "typical" year approach may be misleading, inaccurate, and/or biased. Typical is not central.

• This bias may be high or low. In some cases, lower actual production is more likely, but in other cases, *higher* actual production is more likely.

• Solar insolation and energy production vary significantly over time and by geographic location. DAI's quantitative risk analysis methodology facilitates true, unbiased probabilistic analysis of production uncertainty based on actual historical data.

• The advantage of DAI's probabilistic approach is not only a more accurate measure of average production, but also greater insight into how that production varies and the likelihood of observing any particular level of production.

acciona Energía

WestLB

SUNPOWER

TESORO

THE STATE OF TEXAS

SOCIETE GENERALE

SIEMENS

PRAXAIR

ARCLIGHT

Bank of America

Bank of Tokyo-Mitsubishi UFJ

MUFG

THE BANK OF NEW YORK

CAITHNESS ENERGY

Chesapeake ENERGY

NOMAC DRILLING, LLC

citi

Constellation Energy

CYPRESS

EnergySouth, Inc.

DAI

MANAGEMENT CONSULTANTS

DAI Management Consultants, Inc.

1370 Washington Pike

Bridgeville, PA 15017

(412) 220-8920 voice

(412) 220-8925 fax

www.daimc.com

info@daimc.com

PG&E Pacific Gas and Electric Company

ORMAT

NEW YORK LIFE

New York City Law Department

JPMorgan

Invenergy

ICON

Exelon

FPL Energy

FORTRESS

HSH NORDBANK

IBERDROLA RENEWABLES