Analysis of Photovoltaic System Energy Performance Evaluation Method

Sarah Kurtz
National Renewable Energy Laboratory

Evan Riley
Black & Veatch

Jeff Newmiller
DNV KEMA Renewables

Timothy Dierauf
SunPower Corporation

Adrianne Kimber
Incident Power

Jacob McKee
GCL Solar Energy, Inc.

Robert Flottemesch
Constellation

Pramod Krishnani
Belectric

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Technical Report
NREL/TP-5200-60628
November 2013

Contract No. DE-AC36-08GO28308

Analysis of Photovoltaic System Energy Performance Evaluation Method

Sarah Kurtz
National Renewable Energy Laboratory

Evan Riley
Black & Veatch

Jeff Newmiller
DNV KEMA Renewables

Timothy Dierauf
SunPower Corporation

Adrianne Kimber
Incident Power

Jacob McKee
GCL Solar Energy, Inc.

Robert Flottemesch
Constellation

Pramod Krishnani
Belectric

Prepared under Task No. SS13.4510

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-5200-60628
November 2013

Contract No. DE-AC36-08GO28308

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

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) photo by Pat Corkery, NREL 16416, photo from SunEdison, NREL 17423, photo by Pat Corkery, NREL 16560, photo by Dennis Schroeder, NREL 17613, photo by Dean Armstrong, NREL 17436, photo by Pat Corkery, NREL 17721.

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

Acknowledgments

This work was completed with participation from Gina Binnard (Belectric), Forrest Collins and Owen Westbrook (Juwi Solar); Matthew Deal, Eric Giosa, and Ronald Melchior (Constellation); Matt Donovan and Jenya Meydbray (PV Evolution Labs); Jarom Fariente (Chevron); Tassos Golnas and Joseph Philip (SunEdison); Michael Gostein (Atonometrics); Mary Jane Hale (McHale); Dirk Jordan, Byron Stafford, and John Wohlgemuth, (NREL); James Mokri (San Jose State University); Bob Rutemiller (Automation Consulting & Education); Josh Stein (Sandia National Laboratories); Raji Sundaramoorthy (SUNY College of Nanoscale Science and Engineering); Richard Weinburg (Miasole); and David Williams (Dissigno). This support from the broader community was essential to directing and completing this work. The portion of this work completed at the National Renewable Energy Laboratory was supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308.

Acronyms

AC	Alternating current
ANSI	American National Standards Institute
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
DC	Direct current
EPC	Engineering, procurement, and construction
GHI	Global horizontal irradiance
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
ISO GUM	Guide to the Expression of Uncertainty in Measurement
NREL	National Renewable Energy Laboratory
O&M	Operations and maintenance
OTF	Outdoor test facility
PR	Performance Ratio
POA	Plane of array
PV	Photovoltaic
RMIS	Reference meteorological irradiance system
RSF	Research Support Facility
SRRL	Solar Radiation Research Laboratory

Executive Summary

Documentation of the energy yield of a large photovoltaic (PV) system over a substantial period can be useful to measure a performance guarantee, as an assessment of the health of the system, for verification of a performance model to then be applied to a new system, or for a variety of other purposes. Although the measurement of this performance metric might appear to be straightforward, there are a number of subtleties associated with variations in weather and imperfect data collection that complicate the determination and data analysis. A performance assessment is most valuable when it is completed with a very low uncertainty and when the subtleties are systematically addressed, yet currently no standard exists to guide this process.

This report summarizes a draft methodology for an Energy Performance Evaluation Method, the philosophy behind the draft method, and the lessons that were learned by implementing the method. The general philosophy behind the methodology includes the following features:

- The method is performance-model agnostic.
- The performance model must not be inadvertently modified, when being implemented on the measured meteorological data sets, relative to the model that was used on the historical data set.
- The parties to the test must intentionally define the test boundary—differentiating what is being tested from what is not being tested.
- When correctly implemented, the test result should be independent of the weather and other parameters found outside of the test boundary.

Lessons learned included:

- It is important to collect an accurate, uninterrupted data set.
- It is critical to clearly define and document every step in the process, regardless of how small, especially when multiple parties are involved. A party completely unfamiliar with the process should be able to read the documentation and perform the evaluation with virtually zero deviation from the verified results.
- Strategies for dealing with missing and erroneous data may vary with the data set, but establishing accepted guidelines can facilitate making consistent choices.
- Understanding the subtleties of the meteorological data and the resulting implications of the definition of the test boundary is critical to the meaning and implementation of the test.

The report also summarizes questions requiring additional research and useful modifications to the test procedure, based on the results of the Case Study. These questions and conclusions are summarized in the Conclusions section.

Table of Contents

List of Figures

List of Tables

Introduction: Motivation

The power generation of a photovoltaic (PV) system may be documented by a capacity test [1, 2] that quantifies the power output of the system at set conditions, such as an irradiance of 1000 W/m^2, an ambient temperature of 20°C, and a wind speed of 1 m/s. A longer test must be used to verify the system performance under a range of conditions. A year-long test samples weather and shading associated with all seasons. Shorter tests require less time, but may introduce seasonal bias, especially if the model is of inconsistent accuracy through the year (for example, if the shading is incorrectly estimated). Currently, no comprehensive standard exists for this type of test. Although the documentation of energy yield might appear to be straight forward, there are a number of subtleties [3-7] that complicate the data analysis associated with variations in weather and imperfect data collection. These subtleties can complicate completion of an agreement associated with a performance guarantee, or completion of any test intended to quantify performance of a plant.

Completion of the energy yield evaluation with a very low uncertainty adds value to the project. Performance during a year may vary depending on such factors as:

- Seasonal shading issues
- Seasonal soiling
- Sensitivity to high temperatures
- Sensitivity of the model to weather – for example, if performance ratio is used as the metric, the measured performance will vary strongly with temperature
- Early system degradation
- Clipping or intentional curtailment.

Thus, there is general agreement that an energy test completed over a full year provides greater confidence that a PV system was correctly designed and installed, compared with a shorter test.

A standard for an Energy Performance Evaluation Method is challenging to write because the details of the test and test implementation depend on such things as:

- Size of the project
- Choice of the model
- Choice of the test boundary
- Responsibility for system operation and maintenance (O&M).

The purpose of this report is to communicate a draft of a standard for an Energy Evaluation Test Method (see Appendix B) along with a description of the philosophy that underlies that draft and associated issues. The Test Method may be useful anytime there is a desire to document the long-term performance of a PV system.

Case studies were completed to test the draft test method and to elucidate the issues that were or were not resolved. A particularly complicated data set was chosen for the case studies so as to identify issues that might not arise with a cleaner data set. Because the purpose of the case studies was to test the draft rather than to test the PV system, this report highlights the data that identified issues with the draft, rather than providing an exemplary test report as would be done in a formal application of the standard.

Methodology

Differentiation of Energies

An accurate prediction of system output [3] typically requires a model [8-13] that includes the effects of irradiance, temperature, shading, system design and the local environment. In this report, we apply the draft standard using a sophisticated model (PVsyst) [8] and a simple model (performance ratio) [14-16].

The terms used for describing the energy are defined here.

Predicted energy: The energy generation that is predicted from *historical* weather data that is considered to be representative for the site using a model chosen by the parties to the test.

Expected energy: The energy generation predicted from the same model but using the weather data that is collected during the test.

Measured energy: The electricity that was measured to have been generated by the PV system during the test.

These definitions provide an unambiguous way to differentiate the prediction made, based on historical weather data from the modified prediction of the energy (that would be expected, based on the actual weather data for the time of interest) as shown in Figure 1.

A performance guarantee, assessment of system degradation, or other test result is based on a comparison of the expected and measured values. Emphasis is placed on applying the same model to both the historical and measured weather data. Inadvertent variations from this philosophy may occur, e.g., when the historical weather data are hourly data and the measured data are collected on a shorter sampling interval. Similarly, the historical data usually include global horizontal irradiance or a combination of direct normal and diffuse horizontal irradiance, whereas the measured data may be collected from a sensor in the plane of the array. If the model is modified subtly in one of these ways, the changes may appear to be small, but even a small difference in the model implementation could mean the difference between passing or failing a performance guarantee.

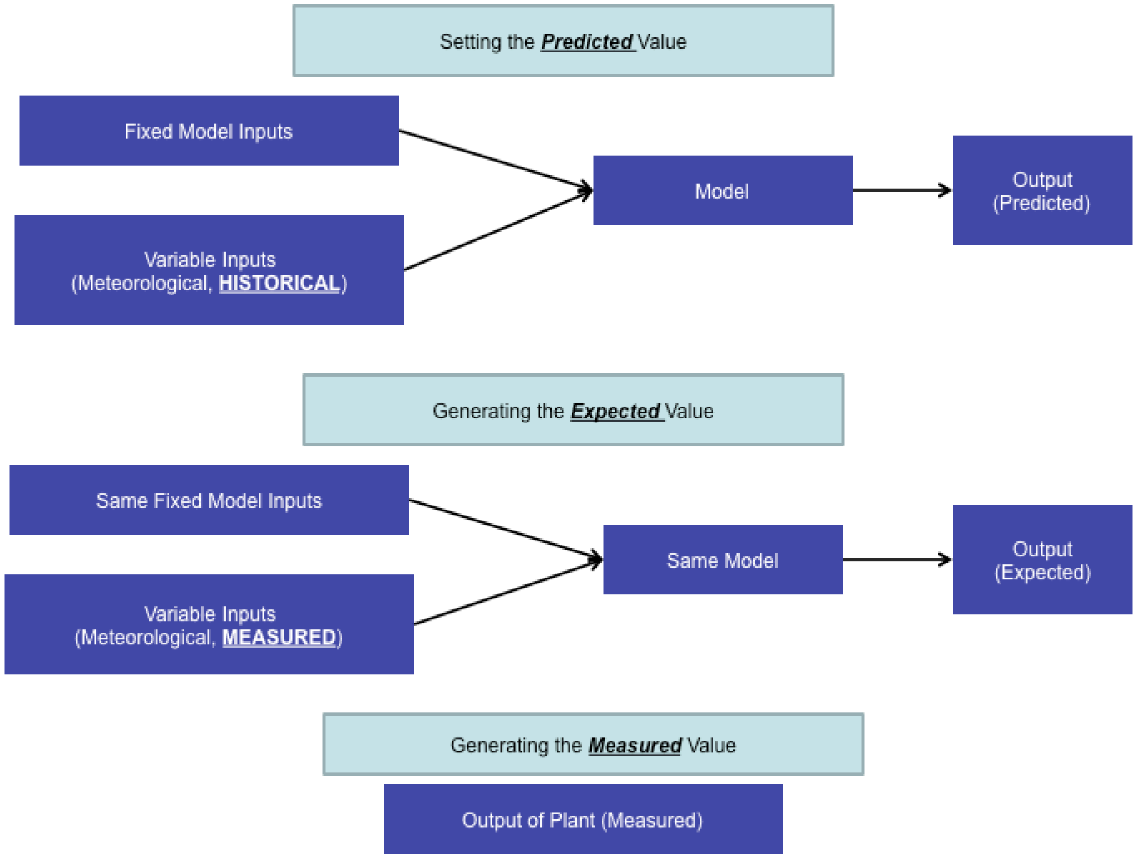

Figure 1. Schematic showing relationship of predicted, expected, and measured energies.

The schematic shows how fixed model inputs are used with both historical and measured data to estimate the predicted output (before the start of the test) and expected output (based on observed weather) of the PV system, which is then compared with the measured output.

Defining the Test Boundary

The Energy Test is meant to define whether a system was installed and performs according to the model that generated the *predicted* energy. If the generation of electricity is lower than the *expected* energy, the system is considered to be deficient. The system output is generally defined by identifying the production meter, while the fuel into the test and the conditions external to the test are defined by the sensors measuring the meteorological conditions. Such a description defines the test boundary; the test is designed to test the performance of everything inside of the boundary. If the *expected* energy differs from the *predicted* energy, this difference is considered to be outside of the test boundary and is not relevant to the test result (which compares the *expected* and *measured* energies.)

Uncertainties and Risks

When comparing the energy yield at the end of the test period with the expected energy yield, there is rarely complete agreement. It is useful to define four different reasons or risks that could cause the predicted energy yield to differ from the measured energy yield.

1. **Variations in weather**. The calculated *expected* energy is intended to correct for the differences between the historical weather file and the weather during the test. The difference between the *predicted* and *expected* energy should exactly align with the weather difference, so this difference should not affect the test result, but will affect the revenue generated by the plant.
2. **Ability of the installed system to perform as represented by the model**. The test is designed to measure this ability directly. In the event that the system does not perform as represented in the model, it may be because the model was flawed or because the installation was flawed.
3. **Uncertainties associated with the test**. There are uncertainties associated with the test that may be impossible to correct. For example, it may be necessary to exclude data or use incomplete data to conduct the test; both of these data anomalies can cause errors in the results. Another example of this risk is the inability of the model to correct for differences between the measured meteorological data and the historical data. This difference (e.g., between an extremely hot year or extremely cloudy year) can amplify biases in the model and cause erroneous test results. Such uncertainty is difficult to predict at the beginning of the test and can be difficult to assess at the end of the test, so may be difficult to address in a contract.
4. **Uncertainties in measurements**. Measurements of meteorological data and energy output data have inherent uncertainty associated with them that should be quantified and reported. The treatment of this uncertainty is more straightforward than the treatment of the uncertainty described previously in item 3. Measurement uncertainties must be defined and agreed upon in writing by all parties. This definition includes how the uncertainty will be determined and propagated.

These four reasons for why the measured energy yield may differ from the predicted energy yield represent different types of risks to the parties involved (see Table 1). For example, considering item 1, if the resource is poor during the measurement period, the measured electricity yield will be lower than the predicted energy yield, causing a risk to the party who is selling the electricity in the event this variation in resource is not appropriately corrected. We refer to this aforementioned risk as "Weather Risk."

Likewise, if the system converts the resource less efficiently into energy than as predicted (assuming all weather variations are properly corrected), the installer will usually be held responsible. Although this concern may be considered a responsibility rather than a risk, we will refer to the possibility that the system converts the resource into electricity less efficiently than modeled as "System Risk."

The third risk is associated with the uncertainty in the test itself, caused by errors implicit in the test methodology. An error in the test could lead to an incorrect conclusion, creating risk for both parties of the test. If the model has been calibrated to be accurate (to give the correct amount of electricity at the end of the test) for a specific set of weather conditions, but is imprecise (gives incorrect results from day to day), the test result will become dependent on the weather during the test period. To distinguish the risk associated with the test's ability to discriminate from the other types of risk, we refer to this third type of risk as "Test Risk."

The final risk is likely the best understood and easiest to quantify—the risk associated with uncertainty in measurements, that is, "Measurement Risk." This risk exists in any performance test and, therefore, is well understood and has been well studied. For more information see the American Society of Mechanical Engineer's Performance Test Code Section 19.1: Test Uncertainty. See Table 1 for further information on these four risks.

Table 1. Comparison of Different Types of Uncertainties and Associated Risks Involved in Executing This Test Method

Type of Risk	Description	Test boundary	Risks for performance guarantee	Role in test or model development
Weather Risk	Associated with the natural variability of weather	Outside	Risk for reduced revenue is usually taken by buyer	Model corrects for weather variability
System Risk	Associated with the chance that the system is not able to convert the resource into energy as efficiently as modeled	Inside	Responsibility is usually taken by installer	The goal is to measure this
Test Risk	Associated with flaws in the test procedure (e.g., imprecise model)	-	Risk is shared or commercially allocated	A well-designed test can eliminate this. The purpose of this report is to better understand this risk.
Measurement Risk	Associated with the measurement equipment used	-	Risk is shared or commercially allocated	Quantified with industry standard practices, reduced with quality sensors

In general, the goal of the test design and application is to

- Eliminate Test Risk
- Minimize Measurement Risk
- Cleanly separate the System Risk (predicted system performance) from the Weather Risk (variations in the weather).

For example, using a comprehensive model that can accurately represent the effects of variable weather reduces the Test Risk for both parties. One could define a risk associated with an inaccurate model, but Table 1 distributes the effects of an inaccurate model between System Risk and Test Risk. If the model is inaccurate, then the system won't perform as expected; if the model is imprecise[1] then the test risk can be amplified by weather variations. Similarly, high data quality and high-accuracy sensors reduce Measurement Risk. A more subtle point that is crucial to understand is that the definition of the *test boundary* can increase or decrease Test Risk as well as blur the difference between Weather Risk and System Risk. The next section will describe this concept in greater detail.

Selecting an Appropriate Test Boundary to Minimize Test Risk

As previously mentioned, the choices of Test Boundary and how the meteorological and energy data are collected determine what is considered to be part of the system performance. The choice of how to monitor each parameter affects the four types of "Risks." For example, moving the module temperature outside of the test boundary moves the System Risk associated with proper cooling of the modules to Weather Risk. At the same time, this can affect the Test Risk. For example, if the model does not accurately calculate the module temperature, then the Test Risk may be decreased when the module temperature is moved outside of the test boundary.

This test can be used most effectively when based on complete understanding of how the choice of the test boundary relates to the items in Table 1. The subtle effects of these choices can be quite confusing; to elucidate these, we contrast the effects of two common choices of measurement parameters and how these choices affect the location of the system boundary, moving some details between System Risk and Weather Risk. We do not attempt to quantify the effects of these choices on the Test Risk, but note that the Test Risk will also change, depending on the model that is being used. When the test is applied for purposes of verifying a model rather than a contract, the "Risk" terminology may be inappropriate, but the general concepts remain the same.

Table 2 summarizes the parameters that lie inside the test boundaries for two cases. Case #1 measures global horizontal irradiance (GHI) and ambient temperature. Case #2 uses a reference cell to measure the irradiance in the plane of the array (POA) and measures module temperature. In the draft standard for the Energy Performance Evaluation Method, Case #1's definition of the test boundary is the default value because it places all aspects of system performance inside of the test boundary. Table 2 shows that Case #2 moves spectral variations and the difference between ambient temperature and module temperature outside of the test boundary, affecting the distinction between Weather Risk and System Risk. Case #2 reduces the Test Risk by removing the transposition of GHI to POA irradiance. Figures 2a and 2b depict the two choices of test boundary visually.

[1] Note that a model that has a bias (e.g., an incorrect temperature coefficient) is considered "imprecise", but might still be "accurate" if it has been adjusted to give the correct answer at the end of the test period for a specific set of weather data.

Table 2. Test Boundary Definition for Two Cases

Case #1: PVsyst is used as the model with measurement of ambient temperature, global horizontal irradiance (GHI), and wind speed.

Case #2: Temperature-corrected performance ratio is calculated based on POA irradiance measured by matched reference cell and module temperature measurements. In both cases, the installer takes responsibility for O&M, cleaning of modules, etc.

Parameter	Inside test boundary – Case #1	Inside test boundary – Case #2
Global Horizontal Radiation (kWh/m^2/y)		
Transposition Model (GHI to POA irradiance)	X	
Transposition Factor	X	
Spectral Variations	X	
Internal Shadings	X	X
External Shadings	X	X
Incident Angle Modifier Loss	X	
Efficiency Variation with Irradiance Level	X	X
Efficiency Variation Due to Temperature Difference between Ambient and Module Temperature	X	
Ambient Temperature and Wind Speed Effects		
Soiling	X	X
Module Quality	X	X
Module Mismatch	X	X
DC Wire Loss	X	X
Inverter Efficiency Loss	X	X
Inverter Clipping	X	X
Transformer	X	X
AC Wire Losses	X	X
Auxiliary Loads (MWh/y)	X	X
Availability	X	X

For Case #1, the choice of measuring global horizontal irradiance (GHI) aligns the measured irradiance data type with the irradiance in most historical weather data files. Both the historical and measured GHI data are converted to POA irradiance using the same model. (In Case #1, PVsyst is used, consistently using the same internal model for the conversion as long as both weather files provide GHI rather than a combination of direct and diffuse irradiance.) The use of GHI data avoids effects of the local albedo (reflection from the nearby surroundings) on the measured irradiance, but introduces uncertainties associated with the GHI-to-POA transposition. Similarly, the horizontal positioning of the GHI sensor avoids the need to measure the azimuth alignment and often improves the accuracy of the sensor alignment because the horizontal configuration is the easiest to align.

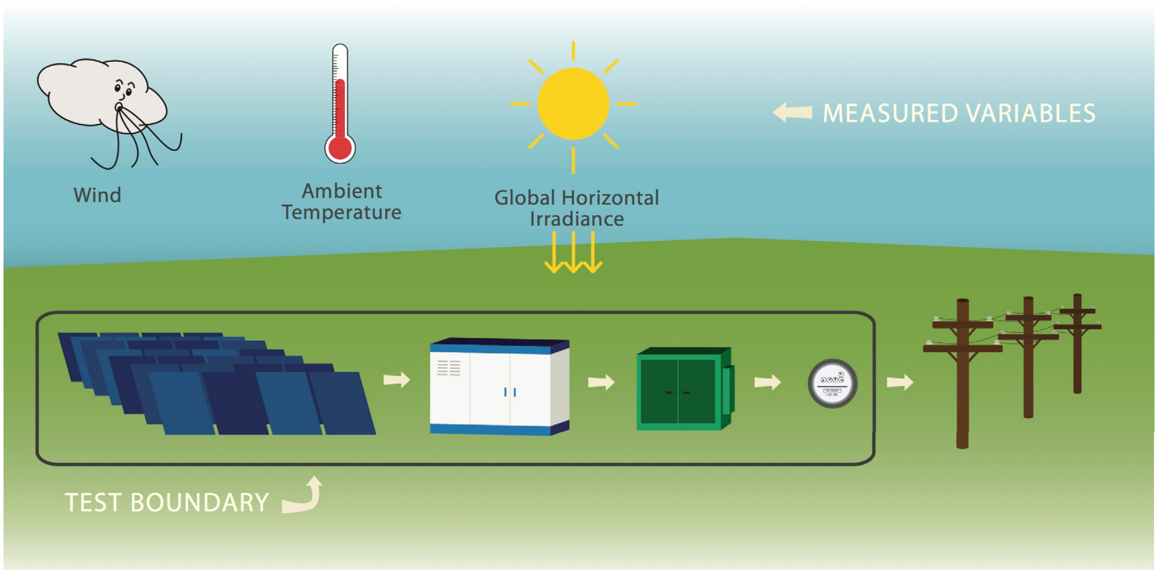

Figure 2a. Test boundary for Case #1 in which the global horizontal irradiance, ambient temperature, and wind speed are measured

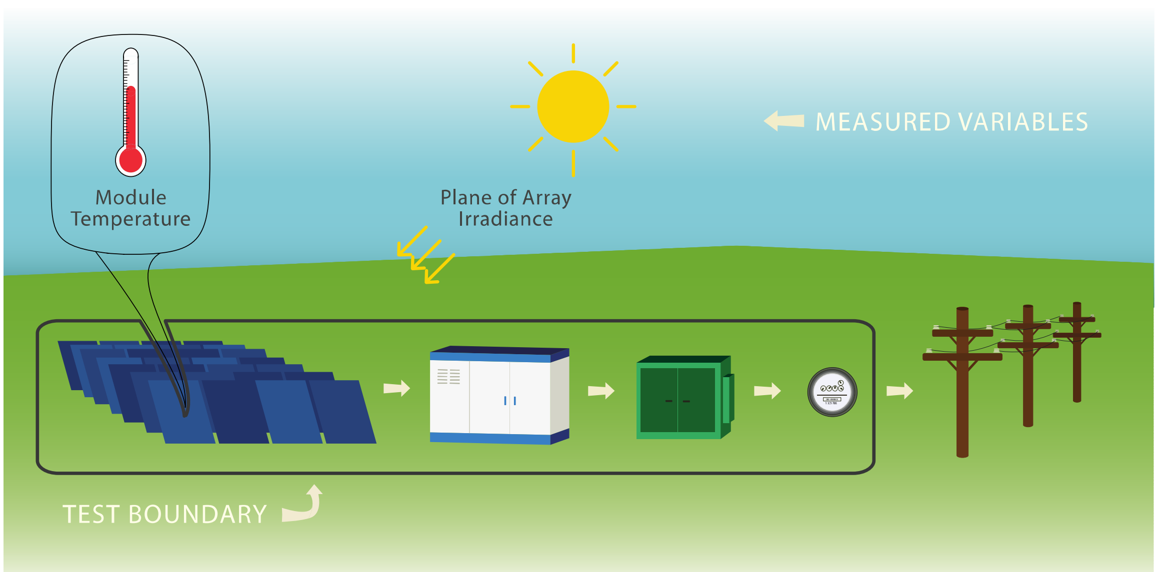

Figure 2b. Test boundary for Case #2 in which the plane-of-array irradiance and module temperature are measured

In Case #1, all aspects of system performance are inside of the test boundary, with only the GHI, ambient temperature, and wind speed outside of the test boundary. In Case #1, the installer, thus, is responsible for all aspects of system installation and performance. The buyer, or whoever is receiving payment for the electricity, uses the test to determine that the system is working correctly, but takes a gamble on the actual electricity produced depending on the weather conditions.

For Case #2 the irradiance is measured in the POA with a matched reference cell and the module temperature is recorded. In this case, some of the uncertainties discussed for

Case #1 disappear. For example, the translation from horizontal irradiance to the irradiance in the tilted plane [17] no longer needs to be completed, reducing the Test Risk. Also, the use of a matched reference cell moves uncertainty associated with spectral variations [18] from System Risk to Weather Risk. Similarly, if ground reflections or module alignment differ from what was designed, these will appear as a different POA irradiance, again moving these uncertainties from System Risk to Weather Risk.

Measuring module temperature instead of ambient temperature also moves the system boundary. If module temperature is measured and the system installation did not provide the modeled air circulation, then the higher module temperatures would indicate a hot year rather than incorrect installation. Thus, Case #2 moves some aspects of workmanship (e.g., array alignment and module operating temperature) outside of the test boundary, causing them to be part of the Weather Risk instead of the System Risk. In choosing the methodology for data collection and model application, the parties should be aware of these subtleties. Tables 1 and 2 are intended to aid in understanding the ramifications of the choices that are made and may serve as a useful tool in evaluating other choices of measurement configurations.

A performance guarantee derived from historical weather data (validated against ground-based thermopile measurements) would logically also use broadband irradiance data from thermopiles. Use of this data allows for a parallel method to set the guarantee, and then to measure this guarantee, even though it may not be the measurement technique with the lowest uncertainty for predicting performance of the system.

If historical weather data sets intended to estimate reference cell data in the POA were more available, there would be more interest in using reference cells. The draft standard requests use of thermopiles for irradiance measurements, but use of reference cells was left as an option. In the end, all parties to the test should agree on which primary instruments are used to measure the guarantee.

Advantages of Possible Choices

The draft standard recommends using Case #1 as the default test boundary because of the clarity it provides and the ease of access to historical weather data that are compatible with that definition of the test boundary. However, other choices are also commonly used. Here is a brief summary of some considerations associated with the choices made in these two cases and other choices applicable to the test method.

Advantages of Monitoring GHI

- Most historical data are represented in terms of GHI. This allows for consistency in how the guarantee was determined and how the plant is measured. However, note that some models calculate POA irradiance from the direct and diffuse irradiance data rather than from GHI when a historical data file provides all three.
- Sensor alignment is simplified (no azimuthal alignment is needed).
- Variations of albedo (in both time and space) do not affect the measurement. Note that snow can enhance the irradiance on systems with substantial tilt and that the

practice of placing an irradiance sensor to the south of the system or off to one side to avoid shading can expose the sensor to a different albedo and sky collection than the system experiences.

Advantages of Monitoring POA Irradiance

- No model is needed to convert the measured GHI to the POA irradiance. The choice of model used for the translation can make a difference of a few percent [17].
- The performance of the system can be more accurately tracked relative to POA irradiance than relative to GHI. Using POA irradiance is useful to tighten the prediction of PV system performance as well as to more quickly detect when the system needs maintenance [19]. For this reason, tracking POA for purposes of O&M is recommended even if GHI is selected for the implementation of this test.

Advantages of Measuring Ambient Temperature

- A guarantee is usually based on measured ambient temperature.
- Modules seldom operate at a uniform temperature. Placement of one sensor in the middle of the module may not accurately reflect the temperature of the entire module or the entire array. (However, note that the ambient temperature may also vary across the array.)
- Sensors on the backs of modules can become loose enough to poorly reflect the true module temperature. A sensor measuring the ambient temperature is less sensitive to the attachment.
- It is difficult to decide how to combine or average sensors for the multiple panels measured.

Advantages of Measuring Module Temperature

- Direct measurement of the module temperature avoids the need to define a model for the expected module temperature.
- For small systems, direct measurement of the module temperature more accurately quantifies transients associated with passing clouds and gusts of wind, compared with using a steady-state model. However, accurate documentation of transients would require many sensors spaced across a system.

Advantages of Measuring Soiling

- Soiling is often a stronger function of the weather than of the quality of the system installation, so soiling is often considered to be outside of the test boundary, and may be considered part of the characterization of the weather.
- Measuring the soiling enables soiling to be placed outside the test boundary.
- Measuring soiling supports the option of separating engineering, procurement, and construction (EPC) and O&M contracts.

Advantages of Not Measuring Soiling

- Placing the effects of soiling inside of the test boundary suggests that parties to the test consider proper operation of a system to include cleaning when the modules become soiled. If cleaning is expected, there is no need to measure soiling levels.
- Addition of an instrument for measuring soiling adds cost to the test.
- Soiling may be nonuniform across a large system, so quantifying the soiling and correcting for it may have a relatively large uncertainty.

Advantages of a Year-Long Test

- Such a test may detect seasonal problems with shading, soiling, snow, frost, model accuracy, clipping, or curtailment.
- A year-long test may detect intermittent problems with system performance such as inverter outages or problems with component degradation.

Advantages of a Test Shorter Than A Year

- Such a test can be completed in a shorter time.
- A test that is shorter provides earlier results regarding system performance that can be used to improve system performance and inform financial models.
- A shorter test may be of adequate accuracy if the model gives consistent predictive quality at all times during the year and if there are no intermittent problems.

Advantages of Using Reference Cells

- Reference cells may be less expensive to purchase and maintain compared with thermopile pyranometers.
- Reference cells typically have an angular and spectral response that is closer to that of PV modules. Matched reference cells should minimize these differences, leading to lower uncertainty in characterizing the irradiance available for solar electric conversion.

Advantages of Using Thermopiles

- Historical irradiance data files are based on comparisons to broadband thermopile measurements, so thermopiles are required to provide a fair comparison.

In the end, the choice of how to model the weather and how to characterize the weather during the measurement period is complex and is likely to depend on the size of the project, the availability of onsite support to help with cleaning and maintaining sensors, etc. All parties should come to agreement on these issues and document the agreement during planning of the test.

Conclusions

A case study was completed, as described in Appendix A. The lessons learned from this case study and from other studies are summarized here for brevity.

Lessons Learned from Evaluating the Data Set, Especially Regarding Data Handling

Importance of Data Quality

- The importance (and challenge) of obtaining high-quality data must be emphasized. Examples of unexpected problems include shadowing of sensors caused by a handrail or safety/security lighting, dataloggers that frequently miss a data point for no obvious reason, and inaccurate zero calibration of sensors.
- Monitoring data quality throughout the year may identify and resolve data quality issues, resulting in a higher quality data set. The test plan should define who is responsible for monitoring data quality and how data issues are corrected.
- Attempts to rectify the glitches in data collection may be only partially successful, so the test method should give some guidance about reasonable strategies for dealing with missing data while encouraging utmost attention to high data quality.
- Comparison of multiple data streams can help in the analysis of the uncertainty of the measured data. In some cases, having multiple data streams may allow reduction of the uncertainty.
- The evaluation of the seasonality may require finer granularity than monthly evaluations if outages reduce performance in many months.

Early and Periodic Data Quality Checks

Such checks should include:

- Values out of range and/or missing data.
- Night time measurements differing from zero may indicate incorrect zero calibration.
- Problems may be identified by comparison of outputs of similar subarrays on sunny days or of the integrated outputs of similar subarrays on cloudy days.

Routine Data Handling

Attention should be paid to the following details:

- The data must be carefully examined for missing or erroneous data and each of these appropriately dealt with (see the section on Missing Data).
- Similarly, inconsistencies in the frequency of data collection and/or duplicate records should be identified and addressed.
- Nighttime data should be removed from the analysis if nonzero irradiance or electrical measurements might affect the expected or measured energy. However, if parasitic power losses are inside of the system boundary, then these must be measured and quantified through the night.

- Conventions of time documentation vary greatly including:
    - Format
    - Whether the time stamp indicates the beginning, middle, or end of the period
    - Whether standard time is used throughout the year or converted to local "summer" or "daylight savings time" during some parts of the year
    - Handling of leap years (Feb. 29th)
    - Indication of midnight as 0:00 or 24:00.

Handling Out-of-Range or Missing Data

- Identify values that are out of range or missing.
- For each hour, the data for each of these four parameters was evaluated according to the available data.
    - If < **10%** of the **electrical** data were missing/rejected, then the remaining values for the electrical data were averaged; or else the hour was recorded as null for all four parameters.
    - If < **10**% of the **irradiance** data were missing/rejected, then the remaining irradiance values were averaged; or else the hour was recorded as null for all four parameters.
    - If < **20**% of the **temperature** data were missing/rejected, then the remaining temperature values were averaged; or else the hour was recorded as null for all four parameters.
    - If < **50**% of the **wind speed** data were missing/rejected, then the remaining values were averaged; or else the wind speed for that hour was replaced with the irradiance-weighted average of the wind speed that was found for the entire year of historical weather data.

Irradiance Data

- Variations of the outputs of thermopiles can vary by several percent as a function of time of year, time of day, etc.
- For the case study, a sensor placed immediately adjacent to the system would have received shading during the late afternoon. Evaluation of the potential size of this effect was not completed, but it is a clear concern, especially when a system may be installed in a location with no shade-free access.
- The uncertainty associated with sensor alignment for the GHI sensors was estimated to be small enough that it could be neglected, compared with other uncertainties. Because of the ease of sensor alignment for GHI, this is a smaller issue than it would be for POA irradiance measurements.
- The choice of irradiance datasets would affect the output by ~3%. In most cases, only a single data set is available, but the uncertainty would still be present; using high-accuracy irradiance sensors is key to reducing uncertainty of the test.

Lessons Learned About Designing the Test

The test plan should specify all primary sources of data that will be used in the analysis. Specifically, the following should be identified:

- The electrical meter to which the performance guarantee applies
- Exactly which sensors will be used to measure the raw data along with any redundancy requirements
- How soiling will be considered
- How measurement uncertainty will be considered when comparing test results with guarantee values
- The effects of shade and reflections on the irradiance measurement
- How the placement of the wind sensor will be consistent with the model.

While converting this draft into a standard, it will be useful to consider the following modifications:

- Evaluation of alignment of the irradiance sensor. When GHI is used, the alignment is unlikely to be a major source of error, but for models that use POA irradiance data, the alignment of the irradiance sensor could cause a bias error.
- Evaluation of placement of the wind sensor. The sensor should be placed as indicated in the model, but if it's not ideally positioned to represent the entire system, is there value in applying a correction?
- In the case of more than one week total missing data, the draft requests estimation of the associated error by using a monthly comparison of the performance relative to the model. For the data set, the monthly comparison was ineffective because of the small number of months with "normal" (no outages) performance. Consideration should be given to using seasonal corrections based on a shorter time analysis. This treatment should be agreed upon as part of the initial test plan.
- The current draft recommends omitting hours with missing data from the analysis. Another strategy would be to substitute simulated data for those hours. For example, the weather data could be used to estimate the electricity production for times when that data is missing.
- Record the fraction of data that was found for each hour as part of the report (or note that that data weren't available if the datalogger did not record this information).
- Add more detail about how the equipment calibration and alignment should be verified at appropriate times.

Remaining Issues

We expect that ongoing research will strengthen the implementation of this draft standard. Some valuable research topics may include:

- Uncertainty assessments related to missing data, imperfect irradiance measurement, shading effects, handling of missing wind data, etc., which could address the following:
    - What useful guidance could be given about uncertainty analysis?
    - The analysis used for the case study discarded data for time periods with less than 90% of irradiance or electrical data or less than 80% of temperature data. Is this an optimal choice?
    - When there are missing data at one time of year or another, what is the best way to identify the associated uncertainty or bias?
    - Is it preferable to omit data for a time period with missing electrical data or to replace the missing electrical data with the modeled data for that hour?
    - If there are missing wind data, is it better to use historical wind data, or data taken from a similar time period?
    - How can one quantify the uncertainty associated with application of the test in less than one year?
    - Is a checklist needed to assess the potential bias errors associated with shading, soiling, snow, etc.?
- Quantifying the circumstances under which using reference cells instead of pyranometers could reduce uncertainty.
- Alignment of the irradiance and wind sensors including:
    - What alignment requirement is needed for irradiance sensors? (If irradiance is measured in the plane of the array, what accuracy is required for the tilt and azimuth alignments?)
    - If the choice is made to monitor irradiance in the plane of the array, where should that sensor be placed (e.g., front row, height from ground) and what checks are needed to confirm accurate placement?
    - If the choice is made to monitor irradiance in the plane of the array, what checks should be completed to confirm that the local albedo is representative of the albedo used in the design and experienced by the entire system?
    - If wind sensors are not positioned at the same height as was used to collect the historical weather data, should a correction be applied?
- Quantifying effects of using GHI vs. POA irradiance and/or the algorithm used for conversion between these.

- When estimating the error associated with missing data in a particular season, evaluating how soiling and clipping should be included in the seasonal analysis.
- How to address a possibility of substantial erroneous data as a result of a datalogger that is averaging only a few points per time interval.

Because this draft is debated in a formal standards committee, the following changes to the current draft should be discussed:

- Adding more guidelines for the handling of uncertainty.
- The length of the test should be revisited; it could be useful to provide a checklist of factors that could cause large uncertainty for shorter tests even when a model shows consistent accuracy through the seasons.
- The alignment (tilt and azimuth) requirement for the irradiance sensor could be determined at the time the predicted energy is calculated.
- The methods for handling missing data and assessing uncertainty associated with seasonal bias may be agreed upon as part of the test plan.
- Specifically, methods for assessing uncertainty associated with missing data in one season may be better defined, including quantifying effects of clipping, soiling, etc.
- The current draft recommends omitting hours with missing electrical production data from the analysis. Another strategy would be to substitute simulated data for those hours. For example, the weather data could be used to estimate the electricity production for times when that data is missing.
- Record the fraction of data that was found for each hour as part of the report (or note that that data weren't available if the datalogger did not record this information).
- Add more detail about how the equipment calibration and alignment should be verified at both the beginning and the end of the test.
- Consider an appendix or other guidelines if reference cells are chosen instead of pyranometers.

References

[1] C.M. Whitaker, T. Townsend, J. Newmiller, D. King, W. Boyson, J. Kratochvil, D. Collier, and D. Osborn. "Application and validation of a new PV performance characterization method." Proc. Photovoltaic Specialists Conference, 1997. Conference Record of the Twenty-Sixth IEEE, pp. 1253-1256.

[2] A. Kimber, T. Dierauf, L. Mitchell, C. Whitaker, T. Townsend, J. Newmiller, D. King, J. Granata, K. Emery, C. Osterwald, D. Myers, B. Marion, A. Pligavko, A. Panchula, T. Levitsky, J. Forbess, and F. Talmud. "Improved test method to verify the power rating of a photovoltaic (PV) project." Proc. 35th Photovoltaic Specialists' Conference, June 20-25, 2010, Philadelphia, PA.

[3] E. Riley. "Photovoltaic system model validation and performance testing: What project developers and technology providers must do to support energy production estimates." Proc. 38th IEEE Photovoltaic Specialists Conference, 2012, pp. 003053-003055.

[4] O.W. Westbrook, B.H. Copanas, and F.D. Collins "New approaches for characterizing photovoltaic system performance." Proc. 38th IEEE Photovoltaic Specialists Conference, June 2012, pp. 001529-001534.

[5] S. Ransome and P. Funtan "Why hourly averaged measurement data is insufficient to model PV system performance accurately." Proc. 20th European Photovoltaic Solar Energy Conference, June 2005, Barcelona.

[6] S.J. Ransome and J.H. Wohlgemuth. "kWh/kWp dependency on PV technology and balance of systems performance." Proc. 29th IEEE Photovoltaic Specialists Conference, 2002, pp. 1420-1423.

[7] S. Ransome "4EP. 1.1 How Well Do PV Modelling Algorithms Really Predict Performance?" 22nd European PVSEC, Milan, 2007.

[8] A. Mermoud. "Use and validation of PVSYST, a user-friendly software for PV-system design." Proc. 13th European Photovoltaic Solar Energy Conference, Nice, France, April 1995, pp. 736-739.

[9] D.L. King, J.A. Kratochvil, and W.E. Boyson. "Photovoltaic Array Performance Model." Sandia National Laboratories, 2004, Report no. SAND2004-3535.

[10] N. Blair, M. Mehos, C. Christensen, and C. Cameron. "Modeling photovoltaic and concentrating solar power trough performance, cost, and financing with the solar advisor model." Solar 2008, American Solar Energy Society, 2008.

[11] B. Marion, and M. Anderberg. "PVWATTS An Online Performance Calculator for Grid-Connected PV Systems." Proc. ASES Solar 2000 Conference, Madison, Wisconsin, 2000.

[12] B. Marion. "Comparison of predictive models for photovoltaic module performance." Proc. 33rd IEEE Photovoltaic Specialists Conference, 2008, pp. 1-6.

[13] B. Kroposki, D. Myers, K. Emery, L. Mrig, C. Whitaker, and J. Newmiller. "Photovoltaic module energy rating methodology development." Proc. 25th IEEE Photovoltaic Specialists Conference, 1996, pp. 1311-1314.

[14] B. Marion, J. Adelstein, K. Boyle, H. Hayden, B. Hammond, T. Fletcher, B. Canada, D. Narang, D. Shugar, H. Wenger, A. Kimber, L. Mitchell, G. Rich, and T. Townsend. "Performance Parameters for Grid-Connected PV Systems." Proc. 31st Photovoltaic Specialists' Conference, Lake Buena Vista, FL 2005.

[15] IEC 61724, "Photovoltaic system performance monitoring - Guidelines for measurement, data exchange and analysis" IEC Central Office: Geneva, Switzerland 1998.

[16] S.J. Ransome, J.H. Wohlgemuth, S. Poropat, and R. Morgan. "Can Grid Tied PV Systems Be Characterised with only Monthly Average Values of PR?" 19th PVSEC, Paris, France 2004, pp. 5BV-1-56.

[17] R. Perez, R. Seals, P. Ineichen, R. Stewart, and D. Menicucci. "A new simplified version of the Perez diffuse irradiance model for tilted surfaces", Solar Energy, 39, (3), 1987, pp. 221-231.

[18] B. Marion. "Influence of Atmospheric Variations on Photovoltaic Performance and Modeling Their Effects for Days with Clear Skies," Proc. 38th Photovoltaic Specialists Conference, 2012, Austin, Texas.

[19] S.K. Firth, K.J. Lomas, and S.J. Rees. "A simple model of PV system performance and its use in fault detection." Solar Energy, 84, (4), 2010, pp. 624-635.

[20] T. Dierauf, A. Growitz, S. Kurtz, J.L.B. Cruz, E. Riley, and C. Hansen. "Weather-Corrected Performance Ratio." National Renewable Energy Laboratory, NREL/TP-5200-57991, (2013).

Appendix A: Description of System Chosen for Evaluation

A complex and simple model, much as described in Cases #1 and #2, were applied to a data set that was available at NREL. Considerations in the choice of this system included:

- A commercial size that would be representative of the types of systems relevant to the Energy Performance Evaluation Method.
- Availability of the types of data needed for applying both PVsyst and performance ratio modeling.
- A preference to have a data set with a variety of interesting features to identify challenges with the data analysis that might be addressed by the draft standard.

A 524-kW system installed over the Visitor's Parking lot at the National Renewable Energy Laboratory in Golden, Colorado, was chosen for the study based on these criteria and data availability. The system is briefly summarized in Table A-1.

Table A-1. System Rating Summary

System part	Power (kWp)		Property (for all parts of system)	Value
Inverter A	262.08		Tilt	10°
Inverter B	262.08		Azimuth	165°
Total system	**524.16**		Analysis period	Jan 10, 2012 – Jan 9, 2013

Evaluation of Electricity Production Data

The evaluation of the electricity production data is common to both case studies.

Production data is available from three meters. In Figure A-1 it can be seen that it could be possible to either use the data from the first two meters or from the single meter connected directly to the grid. Normally, the test plan must specify the exact meter that will be used to document the performance of the system.

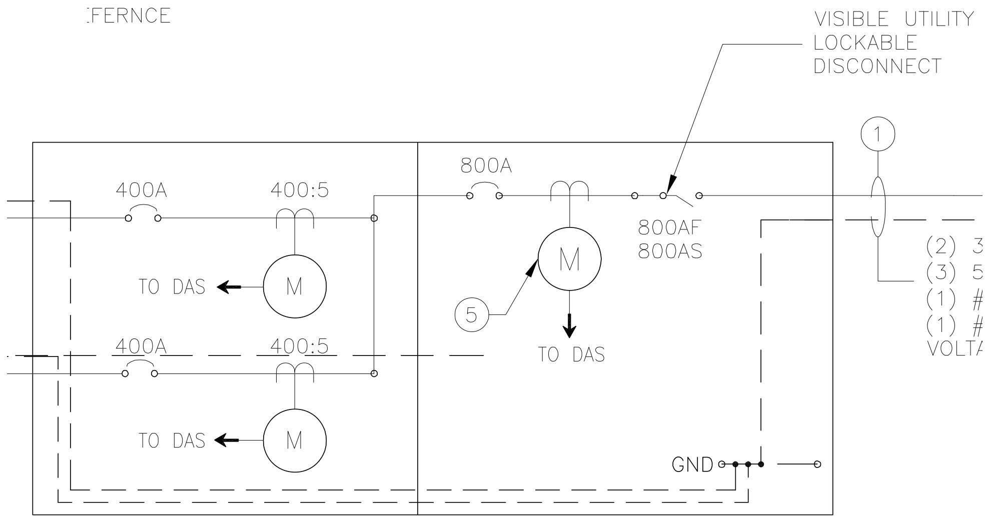

Figure A-1. Schematic indicating wiring of three meters.

The two meters on the left (AC meter #1 and AC meter #2) are connected to the two inverters; the meter on the right (Grid-tie meter) gives the total for electricity flowing into the grid. In a contract, the choice of measurement location needs to be clearly defined because the actual values will differ slightly.

Data were recorded for all three of the meters with a 15-s frequency, as shown in Figure A-2.

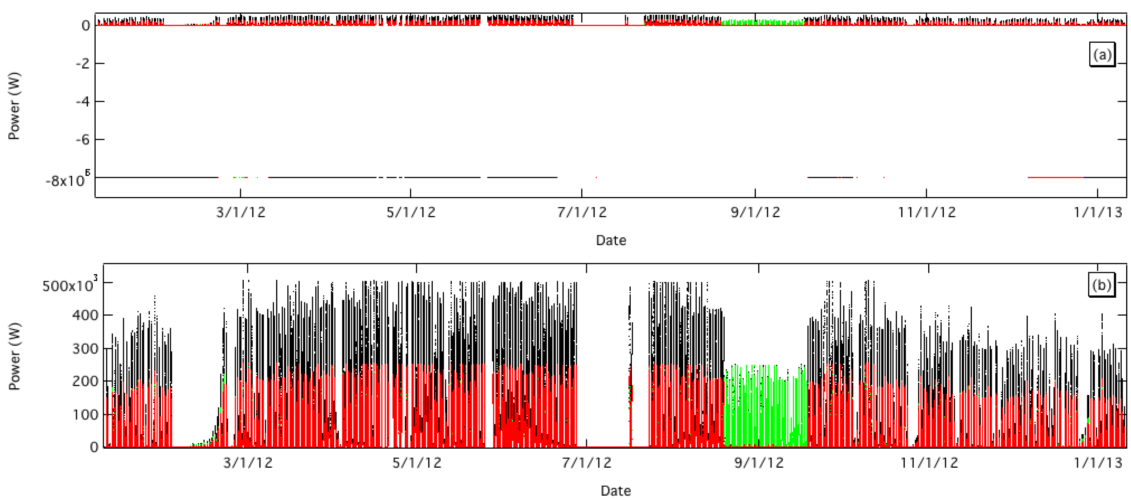

Figure A-2. Power recorded as a function of time/date for the three meters with the grid-tied meter indicated in black, meter #1 in red, and meter #2 in green.

The top graph (a) shows how -7999000 appeared many times in the data set. The lower graph (b) is scaled to hide the negative data. Obvious in the data is evidence of 1) loss from snow coverage for a prolonged time in February and for short times in the spring and following fall/winter, 2) a prolonged outage in July that was implemented by the owner because of lack of adequate access to the local grid, 3) an outage for inverter #1 in August and September, and 4) throughout the year, but especially in April and May, there were hours or days of missing data, probably the result of failed communication with the data loggers. The red data for meter #1 hide the green data for meter #2, except for the period when inverter #1 was not operating. Also evident is clipping of the power at ~ 500 kW for the entire system (black) and ~ 250 kW for each individual inverter (red and green).

Table A-2 summarizes statistics for the data sets. The -7999000 points (indicating missing data) were spaced throughout the data, often losing one or two data points out of an hour. The cause of the missing data was unclear, but the > 30,000 missing data points for the Grid-tie meter (see Table A-2) represented >1% of the overall data, so replacing them with zeroes would introduce unacceptable error. The data processing was discussed and we decided to follow this procedure:

- Identify values that are out of range or missing.
- For each hour, the data for each of these four parameters is evaluated according to the available data.
- If < **10%** of the **electrical** data is missing/rejected, then the remaining values for the electrical data are averaged; or else the hour is recorded as null for all four parameters.
- If < **10**% of the **irradiance** data is missing/rejected, then the remaining irradiance values are averaged; or else the hour is recorded as null for all four parameters.
- If < **20**% of the **temperature** data is missing/rejected, then the remaining temperature values are averaged; or else the hour is recorded as null for all four parameters.
- If < **50**% of the **wind speed** data is missing/rejected, then the remaining values are averaged; or else the wind speed for that hour is replaced with the irradiance-weighted average of the wind speed that was found for the entire year of historical weather data.

Using this approach, it was determined that 224 h lacked 90% of the data for AC meters #1 and #2, while the grid-tied meter was deficient for 229 h. This might be considered surprising, (the grid-tie meter showed 31,633 incidences of -7999000) but, as seen in Figure A-1, these were dispersed through the year, and almost never represented more than 10% of the data in one hour. Table A-2 also compares evidence of slight differences in the calibrations of the three meters.

Table A-2. Summary of Data Quality Assessment Criteria

Meter	AC meter #1	AC meter #2	Sum of #1 + #2	Grid-tie meter
Occurrences of -7999000 (missing data)	692	818		31,633
Hours with > 10% of data missing	224	224		229
Nighttime typical reading	150 ±25 W	0 W		-200 W
Comparison of peak outputs on sunny day	232.8 kW	232.7 kW	465.5 kW	466.0 kW
Average after removing bad data	34497	38871	73368	73086

The draft standard used for this case study specified:

> The test shall last 365 days unless otherwise specified in the test plan. In the event that more than 7 days (or 80 daylight hours) of data are excluded from the analysis after filtering the data per section “Data Filtering”, then the seasonal bias must be estimated (by applying this test procedure on a

month-to-month basis) and the effect of the loss of data on the final result be estimated and included in the final report.

For this case study, the 224 h of missing data represent 224/24= 9.3 days. The 224 h of missing data spanned both daytime and nighttime hours, but were concentrated in April and May. The draft test specifies that the seasonal bias must be estimated in this case by using a month-to-month analysis. However, as will be seen directly, the seasonality of the accuracy of the model prediction could not be quantitatively assessed using a month-by-month analysis because essentially every month of the year was affected by one of the outages described previously. Thus, the standard may need to use a more granular method of assessing the effect of the seasonality for a situation with missing data, such as is seen in this case.

It can be difficult to collect one year's worth of data without having some missing data. In this case, we have documented thousands of missing data points, either as a result of no data appearing in the file or as a result of an error code, giving us the opportunity to assess the impact of these missing data. Commonly, dataloggers may be programmed to average data for some number of minutes. It is not always clear how a data logger handles missing data. If the data logger averages one-second data, then reports the average after 5 minutes, the analyst may not know whether a value was successfully captured every second for the entire 5 minutes. For this case study, the large number of missing data points led to several lessons.

Additionally, the evaluation of this data set and other data sets considered in this study found that the data evaluation was complex and exhausting. Some of the details that contribute to the challenge are summarized in the following bullets.

Lessons learned from this part of the evaluation:

- Monitoring data quality throughout the year may identify and resolve data quality issues, resulting in a higher quality data set. The test plan should define who is responsible for monitoring data quality and how data issues are corrected.
- Attempts to rectify the glitches in data collection may be only partially successful, so the test method should give some guidance about reasonable strategies for dealing with missing data, while encouraging utmost attention to high-quality data.
- Comparison of multiple data streams can help in the analysis of the uncertainty of the measured data. In some cases, having multiple data streams may allow reduction of the uncertainty.
- The evaluation of the seasonality may require finer granularity than monthly evaluations if outages reduce performance in many months.

Early and periodic data quality checks should include:

- Values out of range and/or missing data
- Night time measurements differing from zero—these may indicate incorrect zero calibration

- Comparison of outputs of similar subarrays on sunny days or of the integrated outputs of similar subarrays on cloudy days to identify problems

Attention should be paid to the following details as a part of routine data handling:

- The data must be carefully examined for missing or erroneous data and each of these appropriately dealt with.
- Similarly, inconsistencies in the frequency of data collection and/or duplicate records should be identified and addressed.
- Nighttime data should be removed from the analysis if nonzero irradiance or electrical measurements might affect the expected or measured energy. However, if parasitic power losses are inside of the system boundary, then these must be measured and quantified through the night.
- Conventions of time documentation vary greatly including:
    - Format
    - Whether the time stamp indicates the beginning, middle, or end of the period
    - Whether standard time is used throughout the year or converted to local "summer" or "daylight savings time" during some parts of the year
    - Handling of leap years (Feb. 29th)
    - Indication of midnight as 0:00 or 24:00.

Evaluation of Irradiance Data for Case #1

For Case #1, GHI data are needed as inputs into the PVsyst model. In a test plan, the source of the irradiance data may be defined, but in cases of missing data, other sources of data may also be evaluated for substitution. Four potential data sets were explored: 1) data from a silicon pyranometer (LI-200) on the nearby roof of the Research Support Facility (RSF), 2) data from the Reference Meteorological Irradiance System (RMIS) near the Outdoor Test Facility, 3) global horizontal data measured at the Solar Radiation Research Laboratory (SRRL), and 4) the combination of direct normal and diffuse horizontal irradiance measured at the SRRL and combined mathematically to obtain GHI. The data quality identified for these 4 options (#3 and #4 are summarized together) are summarized in Table A-3. At the SRRL, multiple thermopile pyranometers are deployed side-by-side, so the rightmost column of TableA-3 actually summarizes many more options. The SRRL data are the most complete, the cleanest, and of the highest accuracy, we expect, so these were used in the assessment. However, the availability of so many irradiance data sets provides some opportunity to consider the consistency; Figures A-3 through A-5 show comparisons of the GHI measured on different days during the year by the various sensors, including additional sensors at SRRL.

Table A-3. Options for Global Horizontal Irradiance Data

Identity	RSF	RMIS	SRRL
Type of sensor	Silicon pyranometer	Thermopile	Thermopile
Distance from system (km)	0.2	0.7	0.6
Missing data (h)	232	3	1*
Cleaning	None	During week	During week
Frequency (s)	15	5	60

*Nov. 9th between 16:00 and 17:00 – data are missing.

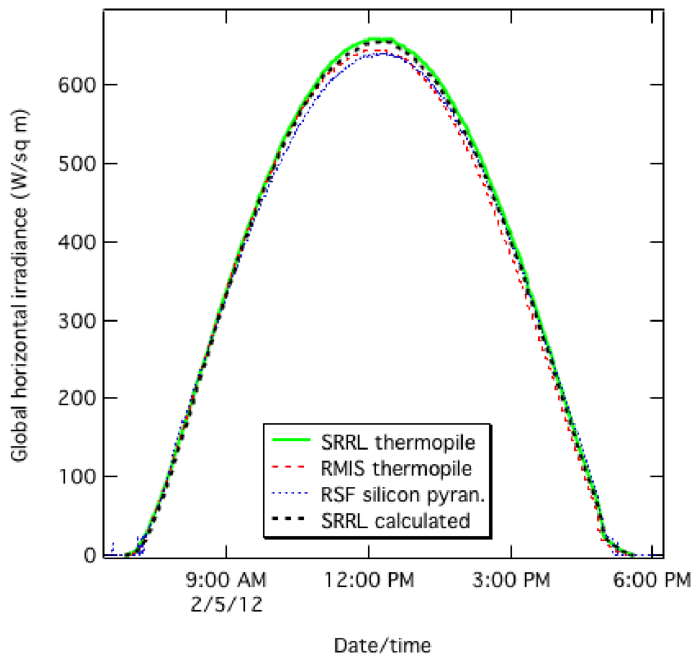

Figure A-3a. Comparison of GHI data on a sunny day in February.

The calibration of the SRRL thermopile results in a larger irradiance measurement than the RMIS thermopile by 2%-3%. The RSF sensor shows a result closer to RMIS during the winter. If SRRL direct and diffuse data is used to calculate global, there is slightly better alignment, as indicated by the dashed black line.

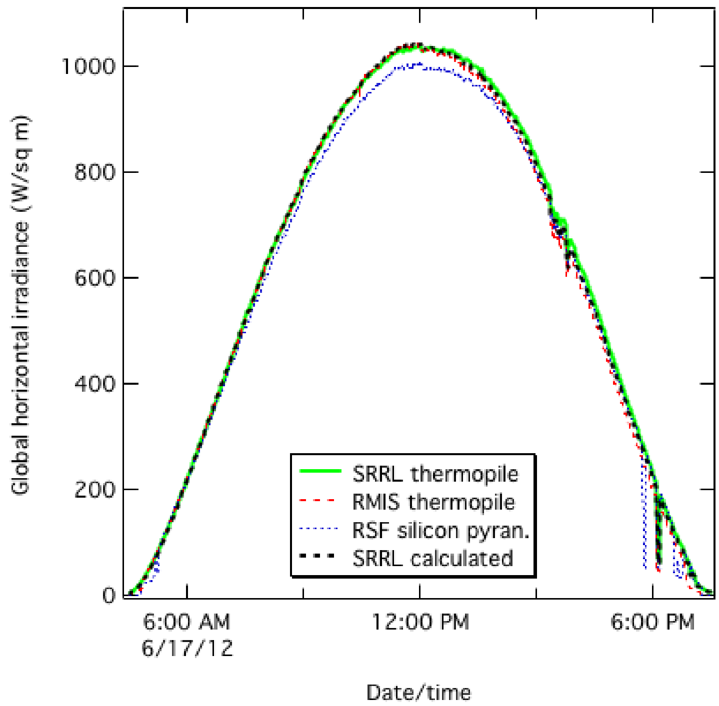

Figure A-3b. Same comparison as Figure A-3a except for a sunny summer day.

The RMIS and SRRL data are better aligned, though there is still a difference in alignment or clock synchronization. The global data calculated from the SRRL direct and diffuse data are in excellent agreement.

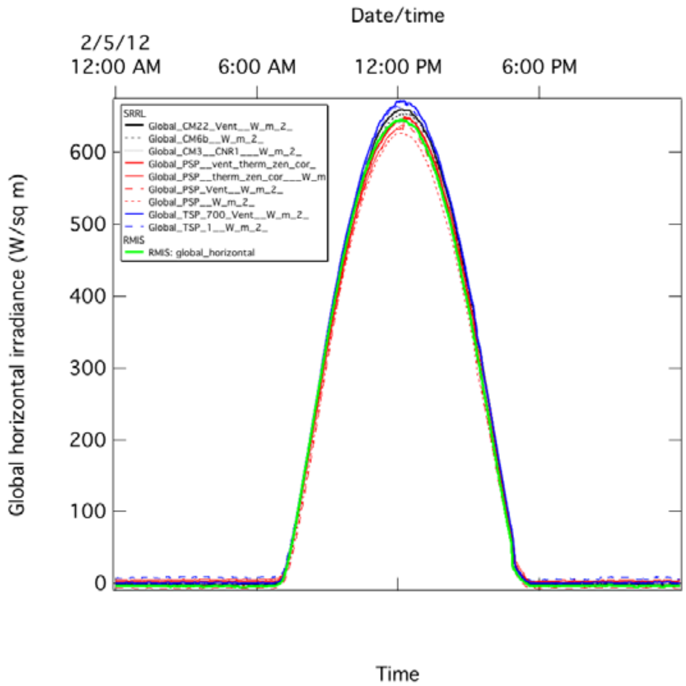

Figure A-4a. Comparison of measured GHI for a set of thermopiles on Feb. 5, 2012.

The legend uses the same naming convention as the SRRL website: http://www.nrel.gov/midc/srrl_bms/.

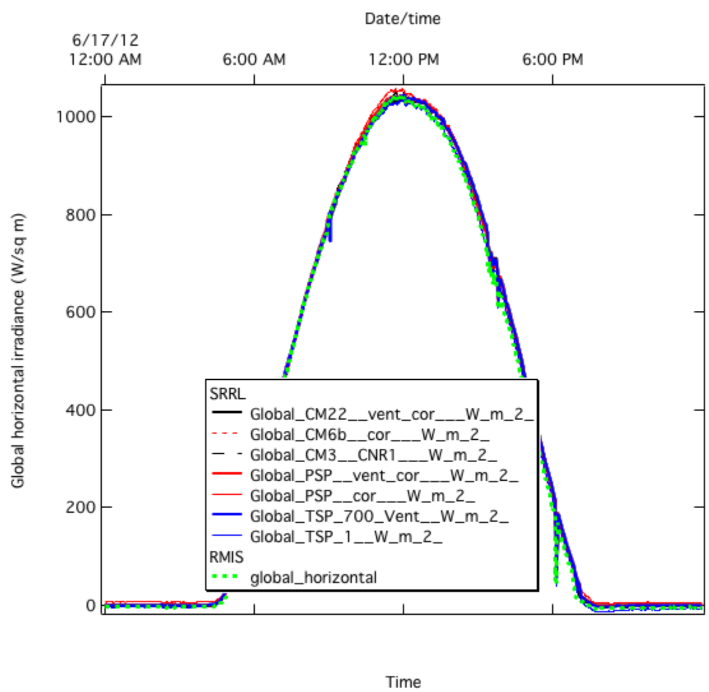

Figure A-4b. Comparison of measured GHI for a set of thermopiles on June 17, 2012.

Note that the configuration at SRRL was changed between these dates.

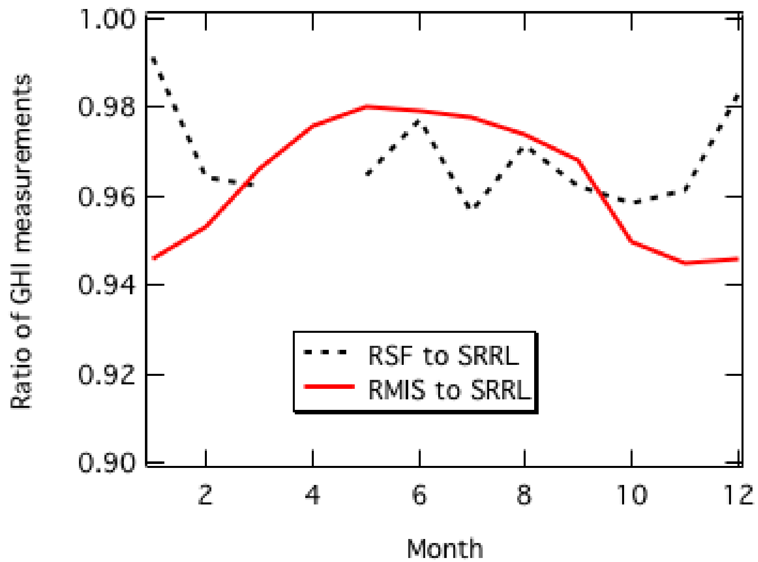

Figure A-5. Comparison between RMIS, RSF, and SRRL data.

The data have been adjusted for missing data; the data point for April is omitted because of the high uncertainty associated with missing data.

Using the 90% data requirement described previously, the GHI data are summarized in Table A-4 for each month of the year. The values for the partial months in January of 2012 and 2013 are combined.

Table A-4. One-Hour Aggregated GHI Data (kWh/m^2/month)

January includes 1/10-1/31 in 2012 and 1/1-1/9 in 2013. Data are first screened for hours with > 90% of the data present, then those data are averaged for that hour.

Month	PVSyst Met file	SRRL	SRRL cal*	SRRL miss hours	RSF	RSF miss hours	RMIS	RMIS miss hours	RMIS to SRRL (%)	RMIS to SRRL* (%)
1	71.8	81.5	80.4	0	80.8	0	77.1	0	94.6	95.9
2	85.4	106.5	104.9	0	102.7	0	101.5	0	95.3	96.8
3	145.7	164.7	162.2	0	156.5	10	159.1	0	96.6	98.1
4	170.6	185.5	183.0	0	153.6	130	181.1	0	97.6	99
5	189.5	200.9	198.6	0	175.1	70	196.8	2	98.0	99.1
6	187.9	210.2	208.1	0	205.4	0	205.7	0	97.9	98.8
7	191.4	201.4	199.8	0	192.7	0	197.0	0	97.8	98.6
8	182.7	182.5	180.8	0	177.3	0	177.7	0	97.4	98.3
9	147.9	152.9	151.3	0	146.1	4	148.0	1	96.8	97.8
10	107.9	110.6	109.7	0	104.8	6	105.1	0	95.0	95.8
11	86.4	81.9	80.7	1	78.8	0	77.4	0	94.5	95.9
12	52.7	70.1	69.1	0	67.8	12	66.3	0	94.6	95.9
Total	1619.9	1748.7	1728.6	1	1641.6	232	1692.8	3	96.8	97.9

*This column used the direct CH1 and the diffuse CM22 irradiance sensors.

The measured data are substantially greater when compared to the historical data (for every month of the year except November and August). As an accuracy check, the measured data are also compared with a clear-sky model in PVsyst in Figure A-6. The data are consistent with the clear-sky model, though the data fall short during July and August when there were fires and smoky skies in Colorado.

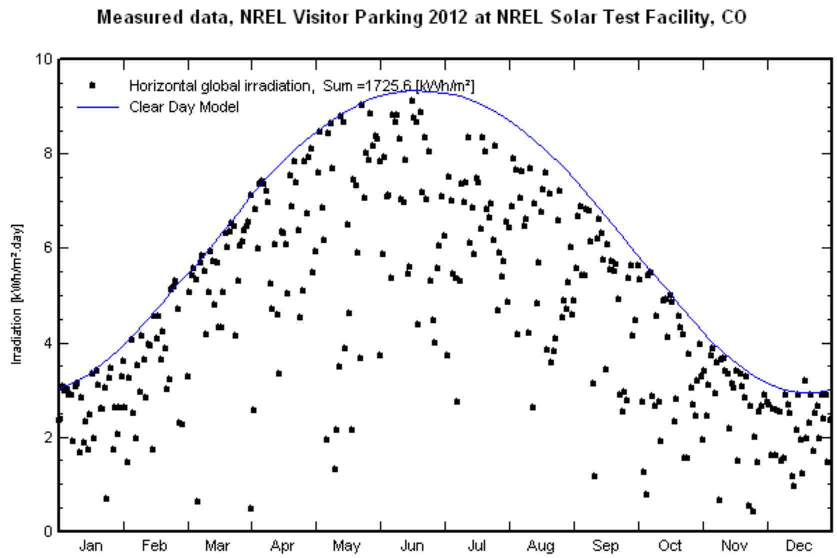

Figure A-6. PVsyst summary of measured irradiance data compared with clear-day model

Lessons learned from the evaluation of the irradiance data:

- The measured data deviated from the historical data by 6%-7%. This deviation was within the 10% window that the draft standard suggests is reasonable. If the difference had been larger than 10%, the test method requires a comment about the expected accuracy of the data.
- Variations of the outputs of thermopiles can vary by several percent as a function of time of year, time of day, etc.
- For this system, a sensor placed immediately adjacent to the system would have received shading during the late afternoon. Evaluation of the potential size of this effect was not completed, but it is a clear concern, especially when a system may be installed in a location with no shade-free access.
- The uncertainty associated with sensor alignment for the GHI sensors was estimated to be small enough that it could be neglected, compared with other uncertainties. Because of the ease of sensor alignment for GHI, this is a smaller issue than it would be for POA irradiance measurements.
- The choice of datasets would affect the output by ~3%. In most cases, only a single data set is available, but the uncertainty would still be present.

Evaluation of Ambient Temperature Data for Case #1

For Case #1, ambient temperature data are needed as inputs into the PVsyst model. Three potential data sets were explored: 1) data from a nearby roof (RSF), 2) data from the Reference Meteorological Irradiance System (RMIS) near the Outdoor Test Facility, and 3) data measured at the Solar Radiation Research Laboratory (SRRL). These data options are summarized in Tables A-5 and A6.

Table A-5. Options for Ambient Temperature Data

The temperature was averaged over the entire year. The number of hours of missing data were defined as hours with > 20% of the data missing.

Identity	RSF	RMIS	SRRL
Distance from system (km)	0.2	0.7	0.6
Missing data (h)	231	0	0
Frequency (s)	15	5	60
Average temperature (°C)	13.1	12.5	12.5

Given the small difference between the average temperatures and the much smaller number of missing data points for SRRL, the SRRL data set is used for the rest of the analysis. Note that the Visitor's Parking structure might operate at a slightly lower average temperature than the others because it will receive some shading in the late afternoon.

Table A-6. One-Hour Aggregated Temperature Data (°C)

January includes 1/10-1/31 in 2012 and 1/1-1/9 in 2013 for all but the PVsyst Met file. Data are screened for hours with > 80% of the data present, and averaged for each hour.

Month	PVsyst Met file	SRRL	SRRL miss hours	RSF	RSF miss hours	RMIS	RMIS miss hours	RMIS - SRRL	RSF - SRRL
1	1.13	2.92	0	3.27	0	2.85	0	0.35	-0.07
2	-1.29	-0.58	0	0.02	0	-0.54	0	0.6	0.04
3	3.16	10.33	0	11.11	9	10.26	0	0.78	-0.07
4	8.39	12.09	0	12.49	130	12.18	0	0.4	0.09
5	13.11	15.40	0	16.33	70	15.51	0	0.93	0.11
6	18.83	23.84	0	24.38	0	23.74	0	0.54	-0.1
7	20.79	24.20	0	25.48	0	24.25	0	1.28	0.05
8	21.52	23.24	0	24.14	0	23.23	0	0.9	-0.01
9	16.99	18.85	0	19.48	4	18.77	0	0.63	-0.08
10	9.35	10.29	0	10.64	6	10.26	0	0.35	-0.03
11	2.57	8.21	0	8.32	0	7.99	0	0.11	-0.22
12	1.76	1.12	0	1.35	12	1.01	0	0.23	-0.11
Total	9.76	12.53	0	13.12	231	12.50	0	0.59	-0.03

The average temperatures for the year were consistent within one degree for the three sets of data. The uncertainty associated with this variation is less than the uncertainty associated with the variability of the irradiance data.

Evaluation of Wind Data for Case #1

For Case #1, wind data are needed as inputs into the PVsyst model. Three potential data sets were explored: 1) data from a nearby roof (RSF), 2) data from the Reference Meteorological Irradiance System (RMIS) near the Outdoor Test Facility, and 3) data measured at the Solar Radiation Research Laboratory (SRRL). These data options are summarized in Table A-7.

Table A-7. Options for Wind Speed Data

The wind speed was averaged for the entire year of data. Missing data were flagged for hours in which > 50% of the data were missing.

Identity	RSF	RMIS	SRRL
Distance from system (km)	0.2	0.7	0.6
Location of sensor	~1 m above roof	~1 m above ground	7 m off ground at higher altitude than the other locations
Missing data (h)	223	4	0
Frequency (s)	15	5	60
Average (m/s)	3.7	1.7	3.2

The RSF data were rejected because of poor calibration as shown in Figure A-7. The wind speed that is relevant to the Visitors' Parking structure may be higher than for a ground mount system. The SRRL data for 22 ft off the ground might or might not be appropriate, but was used in this analysis because it is similar to the height of the structure. The placement of a wind sensor should be defined in the test plan.

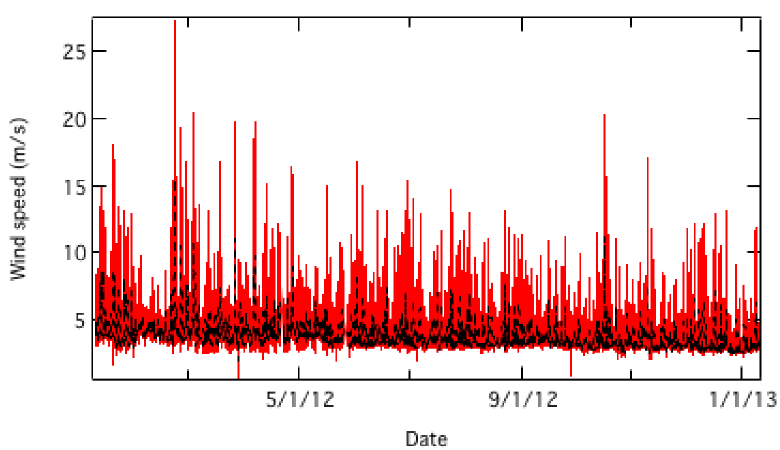

Figure A-7. Wind speed measured by RSF anemometer.

The calibration appears to be in error, so this data was not selected for use in the analysis.

Comparison of Expected and Measured Production for Case #1

The data summarized in Tables A-4 and A-6 and from a companion table for wind were prepared as inputs into PVsyst. The model was run side-by-side to compare the calculations with the historical weather file and the measured weather file and the results were summarized in Table A-8 and in Figure A-8. The performance was lowest relative to that expected for the month of July when the system was intentionally turned off (~70% loss). The second lowest month was February, when snow covered the system, causing a loss of ~60% relative to the expected electricity. The month of September showed ~25% loss because of the inverter outage. Months that had no obvious reason for performance loss delivered between 15% and 16% more electricity than was expected.

Table A-8. Comparison of Measured And Predicted Electricity Production Using Case #1

Month	Sum of two meters adjusted (see Table A-2) (MWh)	PVsyst predicted energy (from historical data) (MWh)	PVsyst expected energy (from measured data) (MWh)	Ratio measured to expected	Comment explaining primary difference	Missing hours
1	39.943	30.96	34.82	**1.15**		0
2	19.284	36.61	44.44	0.434	Snow	0
3	76.404	65.35	73.42	1.04	Missing data	10
4	69.785	76.32	80.38	0.868	Missing data; snow	122
5	79.544	81.43	85.75	0.928	Missing data	70
6	83.012	79.21	85.11	0.975	Outage starts on 28th	0
7	26.306	79.68	82.22	0.32	Outage most of month	0
8	66.299	77.43	76.57	0.866	Inverter 1 out starting Aug. 19	0
9	48.366	64.06	65.29	0.741	Inverter 1 fixed Sept. 18	4
10	49.293	45.28	45.41	1.09	Snow	6
11	39.156	36.48	33.78	**1.16**		0
12	29.328	22.97	30.18	0.972	Snow	12
Total	**626.7**	**695.78**	**737.37**	**0.850**		**224**

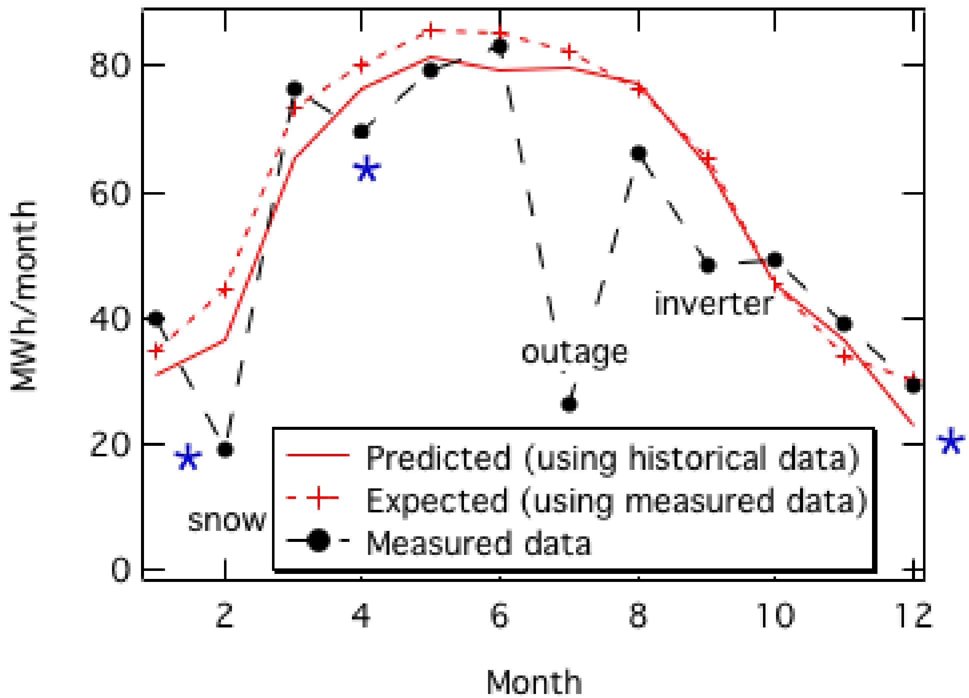

Figure A-8. Comparison of predicted, expected, and measured electricity production

For this comparison, snow coverage affected the production measurably in 3 or 4 of the months. A quantitative model for the snow loss would have improved the comparison significantly. Outages associated with intentional disconnection from the grid and with a failed inverter also affected the output. For the two months of data that were not affected by these, the measured yield exceeded the expected yield by a fairly constant amount.

Performance Ratio Evaluation (Case #2)

Performance ratio (PR) [14-16] is a simple model that allows trivial prediction of the yield of a plant as in equation 1:

Electrical yield = POA irradiance * PR * Nameplate rating/Ref. irradiance (1)

where the electrical yield and POA irradiance are integrated over the same time period and the Ref. irradiance is the irradiance used to assign the nameplate rating (usually 1000 W/m^2).

Because the PR varies with temperature, shading, and many other details, the PR shows substantial seasonal variation. Thus, when a performance guarantee is based on a performance ratio metric, the value is usually derived either from experience at other plants or through application of a more sophisticated model, such as the PVsyst model used in the previous analysis. One way to use the PR but reduce the dependence on weather is to use Dierauf's method in which the annual PR can be derived from data collected at any time of the year by correcting for weather [20].

In this case, we have included application of the PR as an example of a simple model using the draft standard. Although Table 2,addresses measurement of module temperature and correction of that, the module temperature was not recorded in this case, so no temperature correction was made. The evaluation is quite similar to the analysis done previously, but in this case only two data streams need to be considered: 1) the electrical generation that is already discussed, and 2) the POA irradiance. There is only one set of POA irradiance data available because no other sensors in the known data sets were positioned at 10° tilt.

The 10° irradiance data were analyzed as described previously, averaging data measured during an hour, rejecting hours for which more than 10% of the data were missing. The data could also be analyzed by rejecting 15-second periods for which the electrical production or the irradiance data were missing, instead of averaging for the hour and rejecting entire hours when >10% of the data were missing.

The annual summary is given in the next lines; Table A-9 summarizes the data for each month of the year.

Predicted PR from PVsyst:
Plane-of-array irradiance for year = 1761.6 kWh/m^2
Array energy for year = 762.1 MWh
Predicted PR = 762100 kWh/[(524.16 kW/1 kWm^{-2})*1761.6 kWh/m^2]
= 0.825

Measured POA irradiance for year = 1656.9 kWh/m^2
Array energy for year = 626.7 MWh
Measured PR = 626700 kWh/[(524.16 kW/1 kW/m^2)*1656.9 kWh/m^2]
= 0.722

Table A-9. Summary of Monthly Production for Performance Ratio Evaluation

Month	POA irradiance PVsyst from historical data (kWh/m^2)	Predicted energy PVsyst (MWh)	PR predicted by PVsyst	RSF 10° tilt reference cell irradiance (kWh/m^2)	Expected energy (MWh)	Sum of two meters adjusted (MWh)	Measured performance Ratio	Comment	Missing hours
1	90.2	30.96	0.655	85.73	29.426	39.943	**0.889**		0
2	100	36.61	0.698	106.17	38.869	19.284	0.347	Snow	0
3	161.6	65.35	0.772	157.12	63.538	76.404	0.928	Missing data	10
4	179.3	76.32	0.812	151.54	64.504	69.785	0.879	Snow	130
5	192.3	81.43	0.808	168.07	71.17	79.544	0.903	Missing data	70
6	189	79.21	0.8	196.44	82.328	83.012	0.806	Outage from 28th	0
7	194.3	79.68	0.782	192.07	78.766	26.306	0.261	Outage most of month	0
8	190	77.43	0.777	177.41	72.299	66.299	0.713	Inverter out from 19th	0
9	160.4	64.06	0.762	150.59	60.142	48.366	0.613	Inverter fixed on 18th	4
10	123.2	45.28	0.701	111.92	41.134	49.293	0.840	Snow	6
11	108.2	36.48	0.643	86.59	29.194	39.156	**0.863**		0
12	63.5	22.97	0.69	73.22	26.486	29.328	0.764	Snow	12
Total	**1751.9**	**695.78**	**0.758**	**1656.9***	**657.86**	**626.7**	**0.722**		**232**

*Note that the measured POA irradiance is less than the 1875.1kWh/m^2 POA irradiance calculated from the GHI measured at SRRL. It is unclear the extent to which this is a difference in the calibration/alignment of the RSF sensor or an error in the translation from GHI to POA irradiance.

Examination of Table A-9 shows how the measured PR is highly variable. The model is designed to predict the variability of the PR through the year.

Application of the complex and simple models reached the same general conclusions:

- The system did not deliver the predicted nor the expected energy because of a series of outages. Whether the system would have met the performance guarantee would have depended on whether the test plan considered the outages to be allowable or not.
- During the months when no significant outage was observed, the system outperformed the model by 15% or more.

Also, though not studied carefully here, the POA irradiance that was measured did not align well with the POA irradiance calculated from the measured GHI.

Lessons Learned Summary

The ramifications of *how the choice of model and implementation of the model affect the system boundary* (who takes responsibility for which aspects of the system performance) may be the most significant outcome of this study. Thus, although the type of model may be decided by those wishing to install low-cost sensors or by those creating the model, *the choice of the model could easily affect the outcome of the guarantee and should be considered as a business decision*.

The importance (and challenge) of obtaining high-quality data must be emphasized. Examples of unexpected problems include installation of a handrail or lighting for safety or security considerations that cause shadows on sensors, dataloggers that frequently miss a data point for no obvious reason, and inaccurate zero calibration of sensors.

The test plan should specify all primary sources of data that will be used in the analysis. Specifically, the following should be identified:

- The electrical meter to which the performance guarantee applies must be specified.
- Exactly which sensors will be used to measure the raw data along with any redundancy requirements.
- How soiling will be considered.
- How measurement uncertainty will be considered when comparing test results with guarantee values.
- How shade and reflections may affect the irradiance measurement.
- How the placement of the wind sensor will be made consistent with the model.

While converting this draft into a standard, it will be useful to consider the following modifications:

- Evaluation of alignment of the irradiance sensor. When GHI is used, the alignment is unlikely to be a major source of error, but for models that use POA irradiance data, the alignment of the irradiance sensor could cause a bias error.
- Evaluation of placement of the wind sensor. The sensor should be placed as indicated in the model, but if it's not ideally positioned to represent the entire system, is there value in applying a correction?
- In the case of more than one week total missing data, the draft requests estimation of the error that this might introduce by using a monthly comparison of the performance relative to the model. For the data set, the monthly comparison was ineffective because of the small number of months with "normal" (no outages) performance. Consideration should be given to using seasonal corrections based on a shorter time analysis. This treatment should be agreed upon as part of the initial test plan.
- The current draft recommends omitting hours with missing data from the analysis. Another strategy would be to substitute simulated data for those hours. For example, the weather data could be used to estimate the electricity production for times when that data is missing.
- Record the fraction of data that was found for each hour as part of the report (or note that that data weren't available if the datalogger did not record this information).
- Add more detail about how the equipment calibration and alignment should be verified at appropriate times.

In addition, a question remains about whether there is a need to address the possibility of substantial erroneous data as a result of a datalogger that is averaging only a few points per time interval.

We expect that ongoing research will strengthen the implementation of this draft standard. Some valuable research topics may include:

- Uncertainty assessments related to missing data, imperfect irradiance measurement, shading effects, handling of missing wind data, etc.
- Understanding the benefits and challenges of using reference cells as the irradiance sensor.
- Assessing the requirement for alignment of the irradiance sensor.
- Quantifying effects of using GHI vs. POA irradiance and/or the algorithm used for conversion between these.
- When estimating the error associated with missing data in a particular season, evaluating how soiling and clipping should be included in the seasonal analysis.

Appendix B: Draft of Procedure—Photovoltaic System Energy Performance Evaluation

1 Background

The performance of a PV system is dependent on the weather, seasonal effects, and other intermittent issues, so demonstrating that a PV system is performing as predicted requires determining that the system functions correctly under the full range of conditions relevant to the deployment site. The standard, IEC 62446 "Grid connected photovoltaic systems – Minimum requirements for system documentation, commissioning tests and inspection," describes a procedure for ensuring that the plant is wired correctly, but does not attempt to verify that the output of the plant meets the design specification. ASTM E2848-11 "Standard Test Method for Reporting Photovoltaic Non-Concentrator System Performance" describes a method for determining the power output of a photovoltaic system, and is intended to document completion or subsequent operation of a PV system, but is not intended for quantifying performance over all ranges of weather or times of year.

Multiple aspects of PV system performance are dependent on both the weather and the system quality, so it is essential to have a clear understanding of the system being tested. For example, the module temperature is primarily a function of irradiance, ambient temperature, and wind speed. However, the mounting configuration can affect the module temperature and is one aspect of the system that is being tested. The PV industry and PV customers will benefit from understanding how measurement choices can affect the outcome of the test. It is the purpose of this standard to streamline the definition of this test and encourage the adoption of consistent definitions, while still allowing flexibility in the application of the test so as to provide broader value.

2 Scope

This procedure measures and analyzes the energy production of a photovoltaic system relative to production expected from weather data—horizontal irradiance, ambient temperature and wind speed, if using the default test boundary (see definitions section)—or other data, as defined by the parties of the test. The method for predicting the production is outside of the scope of this document.

For best results, this procedure should be used for long-term performance (energy production) testing of photovoltaic systems to evaluate performance over the entire range of operating conditions encountered through the year. Such an evaluation provides evidence that long-term expectations of system energy production are accurate.

The procedure evaluates the quality of the PV system performance, reflecting both the quality of the initial installation and the quality of the ongoing operation of the plant, with the assumption that the model accurately describes the system performance.

In this procedure, actual photovoltaic system energy produced is measured and compared to the expected amount of energy produced for the observed weather. Expectations of energy production are developed using a model of the PV system that will serve as the guarantee or evaluation basis. All parties must agree to this model.

The intent of this standard is to specify a framework procedure for comparing the measured energy produced to the expected energy from a PV system. The framework procedure focuses on items such as test duration, data filtering methods, data acquisition, and sensor choice. To reiterate, the procedure does not proscribe a method for generating predictions of expected energy. The prediction method is left to the user of the test. The end result is documentation of how the PV system performed relative to the performance predicted by the chosen model for the measured weather.

This test procedure is intended for application to grid-connected photovoltaic systems that include at least one inverter and the associated hardware, and is expected to be applied to medium- or large-scale projects.

This procedure is not designed to be applied to concentrator (> 3X) photovoltaic systems, but such application will be explored in the future.

This test procedure was created with a primary goal of facilitating the documentation of a performance guarantee, but may also be used to verify accuracy of a model, track performance (e.g., degradation) of a system over the course of multiple years, or to document system performance for any other purpose.

3 Definitions

Predicted energy: The energy generation that is predicted from *historical* weather data that is considered to be representative for the site using a model chosen by the parties to the test.

Expected energy: The energy generation calculated from the same model but using the weather data that is collected during the test. Exclusions and details are discussed in Section 6.

Measured energy: The electricity that was measured to have been generated by the PV system during the test. Exclusions are discussed in Section 6.

System Performance: Attributes of the system performance that can be traced to the quality of the system components and quality of installation. Generally, the installer is held responsible for the quality of the system performance.

System Operation: Attributes of the system performance that can be traced to the quality of service provided. For example, low availability of the system may be a result of slow response to a disruption. If different entities are responsible for the installation and the operations, then it is useful to distinguish between aspects of the performance that are traced to the initial installation and to the operation.

Test Boundary: A (physical) differentiation of what is considered to be part of the system under test and what is outside of the system. In addition to defining the physical boundaries and which electrical meter is quantifying the electricity production, the test boundary definition must include the locations of all measurements.

Default Test Boundary: To facilitate the description of the test method, this document defines a default test boundary. Global horizontal irradiance, ambient temperature, and wind speed measurements lie outside of this default test boundary. All other aspects of the system are considered to be part of the PV system that is under test, including the module temperature and the plane-of-array irradiance. The parties to the test may define the test boundary however they wish; the default test boundary is defined only as a tool to clarify the application of the test method described here and as an example for how to define the test boundary.

Parties to the test: The individuals or companies that are applying the test. Commonly, these parties may be the PV customer and the PV installer, with the test method applied to define completion of a contract, but the test method may be applied in a variety of situations and the parties to the test may be a single individual or company in some cases.

"The test": Test that compares the measured output of a PV system over a prolonged time period and compares it to the output that was expected for the PV system for the measured set of weather conditions, as defined by this document.

Model: The simulation model used to calculate expected PV generation.

Other definitions are available from the International Electrotechnical Commission Glossary (http://std.iec.ch/glossary) or ASTM G113 – 09 Standard Terminology Relating to Natural and Artificial Weathering Tests of Nonmetallic Materials. Also ASTM E772 – Terminology of Solar Energy Conversion, and the ASME Power Test Codes.

4 Test Scope, Schedule, and Duration

This test may be applied at one of several levels of granularity of a PV plant. The users of the test must agree upon the level(s) at which the test will be applied. The smallest level to which the test may be applied is the smallest AC power generating assembly capable of independent on-grid operation; typically, this would be at a level > 1 MW.

PV plant construction is often divided into phases; phases may have separate or shared interconnection points and may be spread over a period of months or even years. In general, it is recommended that the test be applied at the highest level, that which encompasses the entire PV project. However, for very large plants scheduled for interconnection in parts, with the first and last interconnection separated by a period of more than 6 months, it is recommended that the test be applied to smaller subsets of the plant as they become available for interconnection. In such cases, upon full plant completion the test may be applied again in a way that encompasses the entire plant, but in these cases the expected energy must be modified to include expected plant performance degradation in accordance with the model accepted by the parties to the test.

Some PV modules show measurable performance changes within hours or days after being installed in the field; others do not. The start of the test should be negotiated between the parties using the manufacturer's guidance of the number of days or the irradiance exposure needed for the plant to reach the modeled performance along with the details of the actual installation and interconnection dates. Any degradation assumptions should be agreed to by all parties and documented as part of the model description.

It is recommended that the test last 365 days. The actual test term should be agreed upon in advance. In the event that more than 7 days (or 80 daylight hours) of data are excluded from the analysis after filtering the data according to the "Data Filtering" section, then the seasonal bias must be estimated (by applying this test procedure on a month-to-month basis), and the effect of the loss of data on the final result must be estimated and included in the final report.

If the test is not continued for a full year, seasonal variations (including shading, spectrum, temperature, and wind) may cause the performance to deviate from what would be obtained over a full year.

All parties to the test must agree on a detailed test procedure before the test commences as described in Sections 5 and 6.

5 Equipment and Measurements

Using the default test boundary, the weather is characterized by:

- Global Horizontal Irradiance (direct and diffuse may also be measured)
- Ambient temperature
- Wind speed
- Soiling (if the test agreement assumes a clean system).

Some models use other inputs. Whereas it is encouraged to monitor many aspects of the PV system operation to best understand the status of the system and optimize its performance, the use of data from the system as a characterization of the weather inputs to the model risks compromising the integrity of the test. When data are used for such characterization there is the risk that some aspects of the system performance are then considered to be part of the uncontrolled weather. For example, if modules are mounted without adequate ventilation, they will run hotter than necessary, reducing system output. Similarly, a tracked system that does not track correctly will measure a plane-of-array irradiance that is lower than what it would have been with optimal tracking. Although the frequency of rain and snow will affect system performance, the design of the system may aid in shedding snow and/or being resistant to soiling.

The system output is characterized by:

- AC energy delivered to the grid
- AC power factor, if of interest to the parties to the test.

The definition of the AC energy, including the point of measurement (such as at a utility-grade meter at the point of interconnection) must be documented as part of the Test Boundary definition. If parasitic loads outside the system boundary exist (e.g., trackers and nighttime electricity use by inverters and transformers), the contract must be clear as to whether adjustments are made for these, and, if so, these adjustments must be characterized, as well.

5.1 Global Horizontal Irradiance

- Pyranometers shall be calibrated prior to the start of the test using ASTM Test Method G167 or ASTM Test Method E824, with Test Method G167 preferred, and must be calibrated at least once per year thereafter, or according to manufacturer recommendations. Angle of incidence and temperature corrections to pyranometer measurements should be considered. See ASTM Practice G183. If reference cells are used for measuring horizontal global irradiance, it should be recognized that reference cells have a different spectral response from the broadband sensors that have been used for deriving most historical irradiance data and calibrating most models. In the case that the historical data were derived from reference cell measurements, or if the model has been validated using reference cells, then similar reference cells should be used in place of pyranometers.
- All irradiance sensors shall be mounted in a position that eliminates shading and reflections on the instrument.
- A minimum of two global horizontal irradiance sensors are required for redundancy, unless agreed to by the parties. For large systems, at least one sensor per 5 MW is required.
- The first two sensors should be independently calibrated.
- Sensors should be placed to maximize coverage of the module array footprint, or may be placed side by side to facilitate consistency comparisons.
- Global horizontal irradiance sensors shall be mounted within 1° of the horizontal plane.
- In locations where the average monthly value of daily minimum temperature is below 0˚C the global irradiance sensors shall be heated and ventilated such that they remain free of ice and snow.
- All irradiance sensors shall be cleaned a minimum of once per week during the period of the test, or as agreed to by the parties to the test. Some types of irradiance sensors require more frequent cleaning, as indicated in specific sensor references in the following sections.
- Thermopiles defined as secondary standard by ISO 9060 are preferred for characterizing broadband irradiance. Often the final uncertainty of the measurement is dominated by the uncertainty of the irradiance measurement, so high-accuracy sensors are desired.

5.1.1 Global Irradiance

Global Horizontal Incident irradiation is measured using calibrated hemispherical pyranometers with a field of view of at least 170˚ to reflect historical data.

5.1.2 Diffuse Irradiance

A second strategy for measuring global irradiance is to measure both diffuse and direct irradiance and combine these. Additionally, some models may use the diffuse irradiance as a direct input. Diffuse irradiance may be measured by a shaded pyranometer or diffuse rotating shadow band. Only one diffuse measurement is required, regardless of plant size, if global horizontal and direct irradiance measurements are available as backup.

5.1.3 Direct Normal Irradiance

A second strategy for measuring global irradiance is to measure both diffuse and direct irradiance and combine these. Additionally, some models may use the direct irradiance as an input to the model. Pyrheliometers must be mounted on a 2-axis tracked surface with a tracking accuracy equal to or better than 1°. Calibration must be confirmed at the beginning and end of the test, or annually if the test extends longer than a year (as noted in IEC 60904-2) according to ASTM E816 or another calibration procedure. Pyrheliometers must be cleaned at least once every 2 days during the test period. Only one pyrheliometer is required, regardless of plant size, if a global horizontal irradiance measurement and diffuse irradiance measurement are available to calculate direct normal irradiance as backup.

Other devices may be used, e.g., a rotating shadow band.

5.2 Ambient Temperature

The ambient temperature is measured according to ASTM D6176 with a sensor (with radiation shield) placed at least 1 m from the nearest PV module. Care should also be taken to ensure that the ambient temperature sensor will have similar exposure to the majority of the PV array, and should not be affected by balance-of-system equipment operation; for instance, an ambient temperature sensor should be placed away from the exhaust of inverters and equipment shelters.

Plants less than 5 MW in size must have at least one ambient temperature sensor. Plants larger than 5 MW must have at least two ambient temperature sensors. Ambient temperature sensors must be calibrated prior to the start of the test and should be replaced or recalibrated at least once every two years thereafter.

5.3 Wind speed

The wind speed is measured 10 m above grade to be consistent with historical data or at a different height to be consistent with the model and historical data that are used.

Plants less than 5 MW in size must have at least one wind sensor. Plants larger than 5 MW must have at least two wind sensors. Wind sensors must be calibrated prior to the start of the test and should be replaced or recalibrated at least once every two years thereafter.

5.4 Energy Measurements

All energy measurements will be made with revenue grade meters per ANSI C12.1. Either energy or power may be recorded, with power measurements integrated to derive the energy. Metering also includes electronic power quality (including power factor) and revenue meters, which are IEC 6100-4-30 compliant. Intertie protection and metering devices with instrument-grade resolution are also acceptable. Recording energy instead of power avoids loss of data from communication outages, but frequent readings are still useful in the case of missing irradiance or temperature data.

6 Procedure

6.1 Overview

The terms "predicted" and "expected" energy are defined in the definitions section to provide an unambiguous way to differentiate the prediction based on historical weather data from the modified prediction of the energy that would be expected, based on the measured weather data for the time of interest.

The comparison of measured energy to expected energy is simplified by collecting the new weather data in the same format as the historical data. Alternatively, both parties must agree upon and document a method to bin data into a compatible format.

The comparison of the modeled and test results is documented in detail in the following sections (this list summarizes Sections 6.2-6.8):

- Calculate and document the Predicted Value for energy using the chosen model by listing all inputs including historical weather data, assumptions regarding soiling, shading, outages, etc.; the raw data are included in the final report as an appendix.
- Complete the measurement of data from the operating system over the test period.
- Evaluate the measured data to identify and document anomalies that may require extra treatment. Such anomalies include missing or erroneous data that can be replaced with acceptable error.
- Calculate the Expected Energy using the remaining (not excluded) data.
- Aggregate the Measured Energy.
- Compare the Expected and Measured Values for energy from the plant.
- Compute the Uncertainty of the Test.

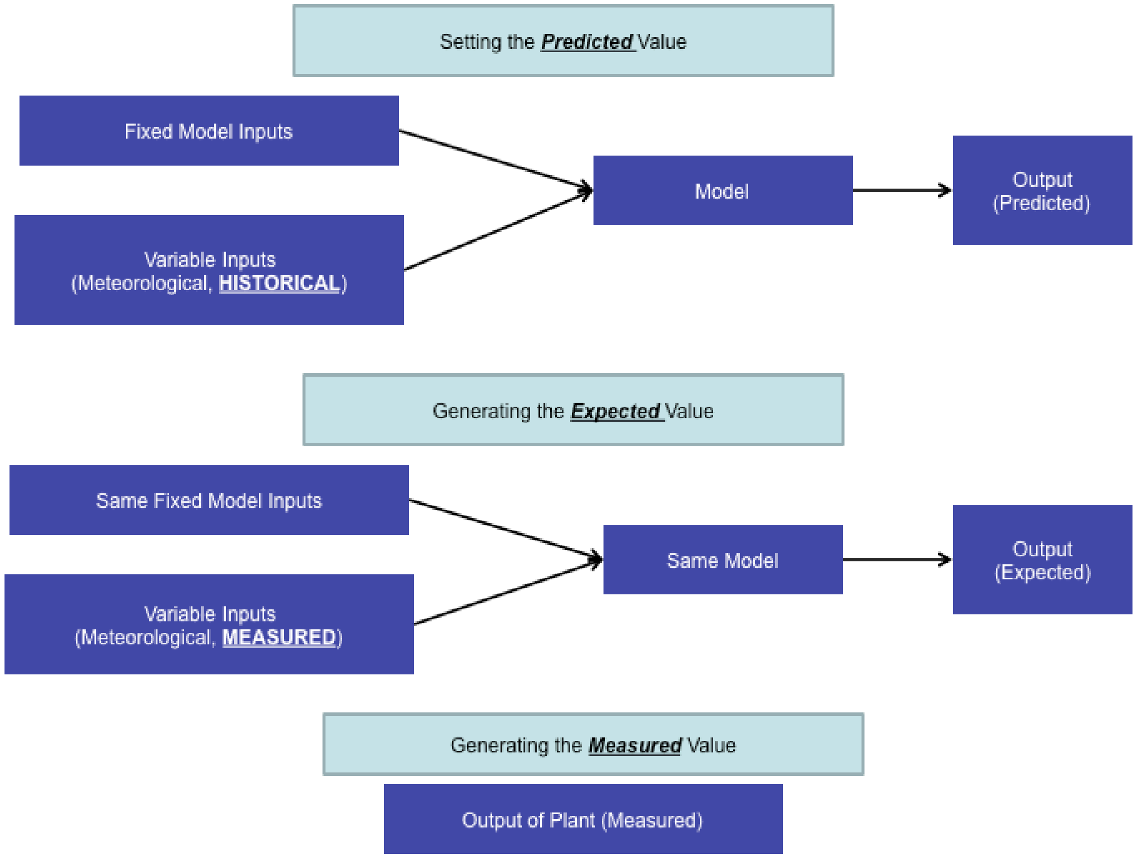

Figure B-1. Schematic showing relationship of predicted, expected, and measured energies.

The schematic shows how fixed model inputs are used with both historical and measured data to estimate the predicted output (before the start of the test) and expected output (based on observed weather) of the PV system, which is then compared with the measured output.

6.2 Calculate and Document Predicted Energy and Method that Will Be Used To Calculate the Expected Energy

As shown in Figure B-1, the first step in the process is to predict the performance of the PV system based on historical weather data using a model that has been agreed to by the parties. The model must be defined in terms of the model inputs, calculation process, and how the measured meteorological data will be input into the model. It is expected that the information required per this Section (6.2) is documented before the beginning of the test; although the final comparison of expected and measured energy does not use the predicted energy directly, the predicted energy is usually required for project planning.

6.2.1 Definition of the Meteorological Inputs Used for the Prediction

The sources of the global horizontal irradiance, ambient temperature, and wind speed data are described and the raw data are included as an appendix in the final report.

6.2.2 Definition of the PV Inputs Used for the Prediction

Table B-1 shows the information requested about each input data type.

Table B-1. PV Performance Input Parameters to the Model for the Initial Prediction

This example table defines the information that is requested about each parameter. Enough information should be given so that the prediction could be duplicated.

Input parameter	Value	Conditions	Source of information
Pmax	205 W	STC = 1000 W/m^2, 25°C cell temperature	Data sheet
All module parameters used in model are enumerated in this table including assumptions made about - Soiling and/or cleaning schedule - Loss factors - O&M availability assumptions - Utility availability and curtailment, other outages - Inverter clipping - Snow losses.			

6.2.3 Definition of Data that Will Be Collected During the Test

The test plan will include documentation of a set of information for each input data type, as shown in Table B-2. The test may identify a primary irradiance/temperature/wind/electricity sensor that will be used as long as data appear to be valid. Alternatively, if multiple sensors of any type are used, the test plan may indicate use of the mean of the sensors. The choice of how to average data from multiple sensors should be defined at the beginning of the test, but some data may be discarded by mutual consent of the parties if there is evidence that the data are in error by more than the expected uncertainty.

If cleanliness of the modules is considered to be a part of the system performance then the soiling level does not need to be measured. If module fouling is not part of the energy guarantee (as defined by the parties to the test) additional measurements will be needed to calculate soiling loss that will be credited to the energy measurement. It must be documented whether soiling and snow effects are included in the performance model.

Table B-2. Example Table Documenting the Meteorological Input Parameters to the Model for the Calculation of the Expected Energy

Input parameter	Type of sensor	Location, orientation, and/or positioning of sensor	Number of sensors	Calibration and maintenance (indicate who will provide maintenance if it is not the system operator)	Alignment check (indicate who will check the alignment if it is not the installer)	Data frequency and analysis
Horizontal global irradiance	Pyranometer model # XXX*	Mounted at height of 2 m as located in drawing Y*	3	Once per year using IEC 60904-4; cleaned weekly	Within 1 °	Average data over one hour and use mean value from all functioning sensors
Ambient temperature	Type T thermocouple	As located in drawing Y*	2	Calibration before and after test	None	Average data over one hour and use mean value from all functioning sensors
Wind speed	Anemometer Model X*	As located in drawing Y*	1	Calibration before and after test	None	Average data over one hour
AC energy	Utility-grade meter: model XXX*	Output of entire system as shown on drawing Y, meter ###*	1	Once per year	Not applicable	Integrated energy is read daily
Power factor						Use information from inverter manual
Indication that inverters are MPP tracking correctly	*Table is filled in as in examples above*					
Parasitic energy losses						
Add lines for additional parameters**						

* X, Y, or ## are used as place holders for the actual information.

** (this table is not intended to be an exhaustive list, but to give examples)

6.2.4 Definition of the Model Calculations

The modeling procedure must be defined with as much detail as required so that a technically competent individual can reproduce the calculation of predicted energy. The description may be documented through a reference that is readily available. The model definition is outside the scope of this document.

Some common models neglect to include the effects of snow and soiling. The model must make some assumptions about the cleaning of the array (as well as the cleaning of irradiance sensors, as included in Table B-2) and about snow coverage. These assumptions should be documented as part of the model description. It is recommended that the system operator take responsibility for the cleanliness of the array and that the losses are assumed to be independent of the weather. The decrease in output may be quantified from direct measurement of cleaned and naturally soiled modules, but the parties should recognize that high soiling can be a result of poor system design. If correction is desired for lost production from snow coverage, it is recommended to screen for such days and treat the associated data as missing data if the model does not directly include losses associated with snow.

6.2.5 Predicted Energy for the Specified System and Time Period

Using the inputs and processes described in Sections 6.2.1, through 6.2.4, state the resulting predicted energy for the designated system and how this relates to the system outputs that are defined in Table B-2. The energy may be predicted for DC and/or AC output and additional predictions may be supplied for parasitic losses, such as for operating trackers. If the system is not well described by a separate document, the system must be described in this section including all details that are relevant to the model, such as the number of modules, mounting configuration, etc. If the test may be applied in a phased way, the system description may define each subsystem. If the time period may be long enough to result in degradation of the array and/or if the test will be delayed to allow for light-induced changes, these must be described.

6.2.6 Uncertainty Definition

The predicted energy will be taken as a given with no uncertainty associated with it.

Test uncertainty should be computed following methods presented in the ASME Performance Test Code 19.1, ISO 5725, or ISO GUM. The uncertainty definition and its role in defining the pass/fail test outcome comparing the *expected* and *measured* energy must be agreed upon. It is highly recommended that this agreement is documented in advance of the test. Typically, the uncertainty agreed to by the parties will form a dead band around the guarantee. This dead band hurts the parties of the test, so should be kept as small as possible. A 95% confidence interval is a common industry practice.

The expected energy and measured energy uncertainty will be discussed in Section 6.9.

6.3 Measurement of Data

The data specified in Table B-2 are collected and recorded with every effort made to avoid gaps in data, to maintain sensor function and calibration through early detection of failures, and to strictly adhere to agreed-upon procedures. The cleaning of sensors is documented through a log of the date/time of cleaning and notes on any unusual observations (a photograph is recommended).

6.4 Identification of Excludable Data

Data will be excluded (the expected energy will be decreased for these time periods and energy generated during these exact same time periods will not be included in the final measured energy) in the cases of:

- Utility grid failures or other outages outside of the control of the PV field.
- Any outages that are not within the scope of the measurement and have been agreed upon in writing by all parties. The upfront agreement should include treatment of a full range of situations, including inspections that require taking the plant off line, outages that can't be addressed in a timely way because of lack of access to the site or unavailability of replacement hardware, routine maintenance events, etc.

Alternatively, when electricity production data are missing, the data may be replaced by the expected values calculated from the model from the measured weather data.

6.5 Identification of Erroneous Data and Replacement or Adjustment of such Data/Preparation of Model Input Dataset

Data must be examined for errors; the exact procedure may vary depending on the data that are collected. While it is recommended to document approved methods for filtering data prior to testing, system complexity makes this difficult and a new, mutually agreed-upon process may be needed during the test; the filters that are applied and the data that are removed will be documented in the report. The following are suggestions and may not be applicable in all situations:

- Create preliminary plots and/or perform a visual check, including a time series for every measured value. Inspect for data that are out of range or other variations in the data that can't be explained by seasonal and daily weather variations.
- Document a sequence of time as "missing data" if PV output data is missing because of a communications or metering problem. If the PV system was turned off because of system failure, then the PV output should be noted as zero (even if it wasn't directly measured) and meteorological data reported as measured. Data for this time period should be included in the analysis. The cause for the missing data must be included in the report in all cases.
- If irradiance data is missing and cannot be replaced (see next bullets), then the time period is labeled as "missing data." The report must explain the reason for the missing data.
- Because of the sensitivity of the test to the irradiance data, special attention must be given to the irradiance data. Specifically, irradiance data that may result from accidental shading of a sensor or sensor malfunction should be removed before taking

the average of the data from the remaining sensors. A recommended procedure for identifying such data is:

A. Identify a clear day in each quarter.
B. Compute the average irradiance value for each sensor during each time period and compare each individual value with the average value for all sensors. If this difference is greater than the uncertainty of the sensors, inspect the data to identify a probable cause. (Note that if the data is taken more frequently than once per minute, the data should be averaged over a time period of at least 1 minute.)
C. Look for drift of the calibrations of the sensors.
D. Discard data that can be traced to malfunctioning of the sensor or data acquisition system.
E. Discard individual data points that are compromised by sensor maintenance or cleaning.

If all data for some time periods are removed, this time period is treated as missing data. The missing data, cause for removal of the data, and the impact of the removal of the data are presented in the report.

- In the case where multiple sensors have been used and one sensor has not been identified as the primary sensor, if the visual inspection identifies error in the output of a sensor, that data should be discarded before taking the average of the data pool. This action should be done only with mutual consent of the parties.
    - The irradiance used as input to the model should be the average of the available measurements, except where one measurement is determined to be erroneous, in which case the input to the model should be the average of the remaining measurements, as described previously.
    - The ambient temperature used as input to the model should be the average of the available measurements, except where one measurement is determined to be erroneous, in which case the input to the model should be the average/median of the remaining measurements.
- In the case where primary sensors have been designated, if irradiance, wind, temperature and/or production data are missing from the primary sensors, but are available from another sensor that is representative of the actual data, the data from the other sensor may be substituted. The report must document (a) the rationale for determining that the other measurements are representative and (b) the uncertainty associated with this substitution.
- Out-of-range data and poor data that result from equipment malfunction (e.g., drift out of calibration, tracker dysfunction, etc.) will be treated as described previously. The method for determining equipment malfunction is based on nearby sensor data or clear sky models, rather than by comparison to the modeled output of the PV system.
- When data are available for part of a time period (e.g., if the model is using hourly averages and the data are available only for part of the hour) the data will be

documented as missing data if > 10% of the data are missing. For temperature and wind data, this requirement is reduced to >20% and >50%, respectively. When the fraction of missing data is small enough to use the data for that hour, the existing data are averaged for that hour. In any case, data for the same time period must be removed from both the irradiance and PV performance data.

- In the case of curtailment, it is assumed that the model originally quantified the output assuming curtailment. The expected energy should be calculated in the same way. The parties should agree about whether the measured energy is credited by the expected energy calculated during the curtailment period. If curtailment is inconsistently implemented or the algorithm is modified in any way during the test, this must be documented in the test report.
- Planned Outage or Force Majeure: If a planned outage was documented in the original contract as excludable, then the predicted energy is taken to be zero during this time period. If a planned outage or an unplanned outage does not qualify as excludable as defined in Section 6.4, then the predicted energy for this time period must be included in the performance guarantee comparison. Treatment of Force Majeure events should be agreed to in advance of the test and documented for later reference.
- Grid Support events (Reactive power, power factor adjustment, etc.) Most models assume a power factor of 1.0 at the inverter. Power plants may need to operate away from unity power factors for reasons outside of their control. Deviations from a unity power factor can affect power output and should be considered when developing the model and agreed to by all parties. Measurements of the local power factor may be collected during the planning phase of the project to determine if operation away from a unity power factor may be required. The power factor should be documented during the measurement period if it is included in the model to produce the expected value or is somehow considered in the application of the test.

6.6 Calculation of Expected Energy

The expected energy generated by the facility shall be calculated by inputting the measured variable input data during the test period into the performance model. Following is a step-by-step procedure for calculating the expected energy:

1. Measure all variable inputs, including meteorological data and plant-specific parameters necessary to update the predicted average-year performance model to account for the actual conditions during the test period. These were specified in Table B-2.
2. As necessary, validate the measured variable input data per Section 6.5.
3. Ensure that the time interval of the measured variable input data is consistent with the input requirements of the performance model. For example, if running an hourly simulation program as the performance model, and higher than one hour resolution data is measured, create an hourly data file by averaging the measured variable inputs at the collected time interval. This procedure should have been

defined in Section 6.2.4. See Section, 6.9 Uncertainty Analysis, for further details.

4. Make sure that hourly data, such as hour ending, hour beginning, or average, are at the proper time stamp.
5. Input measured meteorological data into the performance model using the details in Section 6.2 to calculate the expected energy during the test period.
6. Document all excludable events (see Section 6.4) and missing or replaced data (see Section 6.5) during the test period.
7. The expected energy is calculated for the entire test period, excluding time periods for the excludable events and missing data.

If the expected energy deviates from the predicted energy significantly (by more than 10%), then a root cause diagnosis should be completed. For example, such a diagnosis might be that the weather for the year was unexpected, the simulation model is different than the as-built plant, or there was unusual excludable/missing data. The test report must comment on whether the test should still be considered valid.

6.7 Calculation of Measured Energy

The measured energy shall be the result of all energy generated by the facility as measured at the metering location during the test period after subtracting out energy generated during excludable events (see Section 6.4), time periods determined to be missing data (see Section 6.5), and parasitic power losses.

6.8 Comparison of Measured and Expected Energy

The measured energy (section 6.7) and expected energy (section 6.6) can be compared either as a simple difference, percent difference, or ratio calculation.

Difference Calculation: Measured – Expected

Percent Difference Calculation: [Measured – Expected] X 100 / Expected

Ratio (units of %): (Measured X 100) / Expected

Alternatively, the measured data may be adjusted by the ratio of the predicted/expected energy and compared directly with the initial prediction.

The comparison of measured and expected energy must include a consideration of the uncertainties calculated in 6.9, as guided by the initial agreement.

6.9 Uncertainty Analysis

As part of the performance guarantee or test plan, the agreement must state whether the uncertainty of the measurement is considered. Thus, it can be essential to quantify the uncertainty of the measurement as part of determining whether the measured performance meets expectations.

The uncertainty is determined for the test result, not for the original prediction. Uncertainties associated with the model used for the original prediction are neglected because the agreement is based on the original prediction. However, uncertainties associated with the measured weather data will introduce uncertainty in the calculated expected energy.

Both systematic (bias) and random (precision) uncertainties are included in the analysis. The contributions to the uncertainty depend on the model that is used, but generally include uncertainty in the measurements of the irradiance, temperature, and electricity generated.

All measurements and associated uncertainties are tabulated and combined using standard propagation of errors as described in:

- ASME Power Test Codes, 19.1
- ISO 5725
- ISO GUM – Guide 98, Guide to the Expression of Uncertainty in Measurement.

The uncertainties associated with each sensor are taken from the manufacturer's specification and/or from the calibration report provided by the calibration laboratory. As noted previously, if inspection of the data identifies sensor data with drift or other error outside of the manufacturer's specifications, this data may be discarded by mutual consent of the parties. If such data are not discarded, then the uncertainty is increased to be commensurate with the observed discrepancy.

7 Test Procedure Documentation

A detailed test procedure should be published before test commencement. This test procedure includes all specific requirements and agreements for test execution and data reduction. All parties to the test must have a sufficient opportunity to review and approve this test procedure. It is recommended that the test procedure contain the following sections:

- Purpose
- Guarantee values and basis for guarantee or performance prediction
- Test schedule
- Parties to the test and respective roles and responsibilities for details of installation, operation, and data analysis, including responsibility for:
    - Calibrations
    - Ongoing data quality
    - Cleaning of sensors
    - Cleaning of array
    - Detection of system issues

  - Resolution of system issues
  - Determination of curtailment (if applicable)
  - Analysis of data
  - Writing/review of final report
  - Any other relevant roles.
- Plant operating requirements
- Instrumentation
- Pre-Test Uncertainty Analysis
- Detailed data treatment and reduction methods
- Criteria for a successful test
- Instrumentation cut-sheets and calibration certificates
- Historical meteorological data as an appendix.

8 Test Report

The final test report will include both the Test Procedure (either explicitly or by reference) as well as the following items. (For items that are duplicated on both lists, the final report should either duplicate the original information, verify that the project was executed as originally planned, or note modifications that occurred during the test period.)

1. Description of the party doing the test
2. Description of the site being tested, including latitude, longitude, and altitude.
3. Description of the system being tested, specifically including Table B-2, which describes all of the inputs to the model. Specific note should be made of whether there are parasitic loads and how these are documented by the test.
4. Description of the historical meteorological data that were used for the initial prediction as in Table 1.1 and/or inclusion of the raw data as an appendix if the referenced data are not publically available.
5. A summary of the initial performance prediction that was made based on the historical data
6. A summary of the definition of the meteorological data taken during the test as described in Table B-2, including calibration data for all sensors (sensor identification, test laboratory, date of test, and observed changes in calibration)
7. A summary of the definition of the system output data collected during the test as defined in Table B-2, including records of completed calibrations.
8. The raw data that were collected during the test, including note of which data, if any, were flagged for removal.

9. An explanation of why data (if any) were removed.
10. A list of any deviations from the test procedure and why these were taken.
11. Summary comparison of the expected performance calculated from the measured weather data and the actual measured performance.
    A. Breakdown of difference in Expected Energy from Measured Energy into major categories.
12. Description of uncertainty analysis and statement of uncertainty associated with the expected performance, based on the uncertainty of the weather measurements (see section 6.9).
13. Description of uncertainty analysis and statement of the uncertainty associated with the measured performance (see section 6.9).