U.S. DEPARTMENT OF
ENERGY

Office of
ENERGY EFFICIENCY &
RENEWABLE ENERGY

Modeling of Photovoltaic Systems: Basic Challenges and DOE-Funded Tools

Introduction

Photovoltaic (PV) systems are expected to operate in varying conditions for at least 20 to 30 years, and the U.S. Department of Energy (DOE) supports research and development (R&D) to extend the useful PV system life to 50 years. System performance directly affects project cash flows, which largely determine the value of those systems. It also affects operation and planning activities for the electric grid. Therefore, to reduce financial risk (relevant to the system owners) and reliability risk (relevant to the electric power system), it is important to accurately model the operation of PV systems before they are constructed. Such a model will use meteorological inputs and a mathematical representation of the system to calculate the energy that will be generated over any time interval of interest—from minutes to decades.

However, PV systems involve components with complex electrical, thermal, and mechanical behavior. This means it is not computationally efficient to simulate the operation of systems with models that only use physical laws. For this reason, all platforms that simulate the behavior of PV systems make use of empirical and semi-empirical models to describe the performance of various components.

The importance of accurate modeling is hard to overstate given the rapid deployment of PV systems in the United States and around the world. According to the Solar Energy Industries Association, the installed capacity of solar energy in the United States at the end of 2021 exceeded 100 gigawatts direct current (GWdc), distributed across nearly 3 million systems, while statistics from the International Renewable Energy Association show the world has been adding more than 100 GWdc every year since 2019. Given the goals for a decarbonized power sector by 2035 and a net-zero economy by 2050, DOE expects the installation rate of PV systems to reach new records, further increasing the significance of having accurate models.

Model Inputs

Models of actual or proposed PV systems generally need two types of inputs: design specifications or actual design parameters, and environmental data. Specifications (often referred to, somewhat misleadingly, as metadata) include electrical characteristics of the PV modules, electrical connection topology, specifications of the inverters, geographic coordinates, orientation and spacing of the modules, tracking algorithms of the trackers, and shading conditions. These parameters mostly stay constant or change in a predetermined way.

Environmental data are necessary to simulate operation under realistic conditions and over extended periods of time. Such inputs are solar radiation, ambient temperature, wind speed and direction, volume of rain and snow, reflectance of the ground (albedo), and air quality levels. These parameters change constantly in ways that cannot be predetermined with sufficient accuracy or precision. To allow for the simulation of realistic performance by a PV system, modelers make assumptions for these environmental variables. The most frequent assumption is that over long timelines (e.g., 30 or more years), average environmental conditions stay the same. This allows for the use of historical records of these variables as inputs to models that will simulate the performance of a PV system. With climate change, the assumption of long-term environmental invariance becomes questionable, meaning the outputs of models using historical records will be subject to greater uncertainty when it comes to estimates of long-term performance.

Furthermore, modeling the realistic, long-term performance of a PV system requires additional data; specifically, information about the reliability of certain components whose malfunction or degradation reduces or interrupts the power generated by the system. This type of reliability, capacity, or availability information can be approximated using historical records of similar systems or calculated using engineering models combined with the modeled behavior of a system that is unencumbered by component failures or degradation.

Modeling Uncertainty

Even with all the information mentioned above, models cannot perfectly simulate the performance of the system. First, the models themselves have some built-in uncertainty due to imperfect observations during model development and the approximating nature of empirical models. As alluded to earlier, a perfect, from-first-principles (ab initio) representation of complex systems is computationally inefficient if not impossible. Additionally, the specification information comes with its own uncertainty, errors, and discrepancies from the as-built reality. Finally, if historical environmental data are used as inputs to estimate future performance, the fidelity of the observations or their proximity to the real environment of the system can exacerbate uncertainty.

Then, there is the assumption that the climatic parameters (the "average weather") will remain the same over the next decades. Of course, even without anthropogenic climate change, real weather can deviate from the historical average, sometimes significantly so. In addition to this inherent variability, the extreme difficulty to predict such deviations caused by climate change over the course of 20-30 years inserts additional, irreducible uncertainty in our models.

An accurate assessment of risks associated with the technical performance of PV systems needs to account for these uncertainties in models, specifications, input data, and epistemic assumptions. A full risk assessment, including financial performance, should extend to additional parameters, such as the cost of maintenance, the financial outlook of key component vendors, and the macroeconomic and policy factors affecting the value of electricity generated by the power plant, among others.

DOE-Funded Tools

System and Component Modeling

The Solar Energy Technologies Office (SETO) has provided sustained funding for projects that have delivered results across the full spectrum of elements necessary for simulating a PV system.

For example, the System Advisor Model (SAM) allows performance simulation of a PV system with one-minute resolution and an arbitrary length of time. SAM is powered by component-simulating models developed by national labs and, recently, the broader community. Many of these models are empirical—they describe relationships between physical quantities without explicit reference to a basic theory of the underlying physics. Other models use explicitly physical laws, such as those that calculate the ohmic losses in the wires connecting the DC components in a PV system. SAM can be used to enter system location and shading conditions as well as specifications and connectivity of components.

For applications that do not need the full feature-set of SAM, the PVWatts Calculator can be used instead. PVWatts is a simple, empirical model that allows a user to enter the location of a PV system along with a few key inputs related to the size and type of the system. The calculator models the behavior of a typical system with the help of two numbers—the first is the conversion efficiency of the solar inverter (the power electronics equipment that converts the direct current (DC) output of the PV modules to grid-compatible alternating current (AC)), and the second is for all the other effects that result in losses from incident sunlight to electric power, such as the effects of temperature, shading, soiling, and conductor resistance. The calculator uses the National Solar Radiation DataBase (NSRDB) to retrieve the environmental data related to the desired location and the other user-defined input to calculate the expected monthly or annual generation of energy from the system.

Applications that need more granular simulation of PV systems or relevant parameters may use PVlib, an open-source library of empirical and semi-empirical functions, written in MATLAB and python, useful in modeling numerous aspects of system and component behavior, including the incident irradiance according to various clear-sky models, and the PV cell temperature according to a number of environmental inputs and the support structure and type of the PV module, among others.

The industry has created modeling tools for use by solar energy practitioners over the last two decades. While this document cannot offer an exhaustive list of commercially available software, among the more established products are PVSYST, HelioScope, PlantPredict, Aurora, PV*SOL, and CASSYS (Canadian Solar System Simulator). It's worth noting that the design of a PV system is critical in the realization of projects small and large; therefore, many of these applications place significant focus on the features, capabilities, and interface associated with system design.

Environmental Variables

For meteorological data, the National Solar Radiation DataBase (NSRDB) can provide more than 20 years of irradiance, surface temperature, and other ancillary data with a minimum resolution of 4x4 km and 30 minutes. NSRDB data are generated from satellite imagery, a large number of environmental data, and a physics-based model of radiation transfer in the atmosphere. The data are not post-processed using ground measurements and consequently carry a level of uncertainty that does not meet the requirements of project financing. However, the database is used for feasibility studies or large-scale analyses by the industry and the national laboratories.

Performance Loss Factors

Summary field data that can inform empirical models of PV performance can be found in two reports created by the PV Fleet Performance Initiative: the PV Fleet Performance Data Initiative: March 2020 Methodology Report and the PV Fleet Performance Data Initiative: Performance Index–Based Analysis. The PV Fleet Performance Initiative analyzes the multi-year performance of hundreds of PV systems and derives information about the major performance loss factors, such as soiling, shading, electrical mismatch, wiring resistance, degradation rate, and availability. While the reports contain significantly more detail, it is worth mentioning that the analysis of the more than 1,000 systems in the study points to an average inverter availability of 97.7% compared to a commonly assumed 99%.

More reliability-related data has been collected by projects such as the PV Reliability Operations and Maintenance (PVROM) project, led by Sandia National Labs and the Electric Power Research Institute. PVROM refers collectively to the tool developed for the collection of data on system specification, performance, availability, and maintenance activities; the database constructed to house those datasets; and the project that created the tools, collected the data, and performed the analysis. The project has been able to provide data-driven PV system reliability and operation and maintenance findings that can be utilized to notify more strategic long-term thinking around solar plant operation and value.

Modeling Uncertainty

SETO has funded a team dedicated to defining—through stakeholder consensus—a methodology that can be used to estimate the impact of uncertainty in the inputs on the estimation of system performance. The project, led by NREL and Sandia National Labs, will also provide benchmark estimates of uncertainty for specific inputs and parameters. Preliminary analysis indicates the uncertainty in the energy output is generally smaller than the uncertainty in the irradiance sensor measurements. More detailed results will be presented in workshops and conferences in 2022 and an implementation of the methodology will be included in a future version of SAM.

Assessment of uncertainty and risk for solar projects is central to project development and operation. The full treatment for such risk analysis is beyond the scope of DOE-funded projects; however, industry stakeholders have long attempted to understand and quantify risk factors from technology to commercial and macroeconomic risk. A recent example is a 2019 report that features multiple contributions from members of the solar industry.

Performance Analysis

PV system models can be used for more than future performance estimates (and, therefore, valuation). They can also be populated with real-time observations of environmental variables, such as solar radiation and ambient temperature, as well as design information that reflects the as-built system. The model output can then be compared with the observed output of the system to detect anomalies in performance.

This analysis can occur either in real time or retrospectively. Additionally, identification of anomalies in performance is possible and more efficient without the use of an explicit modeled representation of the PV system. To facilitate these types of analyses, SETO has supported the development of toolkits like RdTools and Solar Data Tools that can detect anomalies using very little information about the specifications of the system.

Future Directions

The ability to model PV system behavior is important in a wide range of applications from project development to power plant monitoring, to electric grid planning. Each application has its own specific requirements in terms of temporal resolution, component granularity, and output precision. The introduction and adoption of new technologies such as new solar cell materials, bifacial modules, and grid-forming inverters requires not just incremental improvement of existing models, but the introduction of new models, too.

SETO is committed to supporting the PV modeling community through funding opportunities relevant to the varied challenges and through the dissemination of the models and algorithms developed by national labs in conferences, publications, and open-source software repositories. Important forums that have fostered information exchange and collaborations among professionals in the industry, academia, and the national laboratories are the workshops and symposia organized since 2010 by the PV Performance Modeling Collaborative, led by Sandia National Laboratories.

At the same time, SETO encourages the greater community to actively engage in open-source software projects that tackle PV system or component modeling challenges. The permissive licensing for those projects is meant to promote low-overhead use by non-profit and for-profit entities alike.

Finally, SETO understands that both model development and validation need high-quality datasets. While the solar R&D community has long lamented the lack of such datasets, the office is exploring solutions that can address this challenge. Plausible paths include clearinghouses for historical data from selected privately owned and many government-owned systems, and marketplaces for historical data from privately owned systems.

U.S. DEPARTMENT OF
ENERGY

Office of
ENERGY EFFICIENCY &
RENEWABLE ENERGY

For more information, visit:
energy.gov/eere/solar

DOE/EE-2603 • May 2022