Photovoltaic String Sizing Using Site-Specific Modeling

Todd Karin and Anubhav Jain

***Abstract*—One key design decision for photovoltaic (PV) power plants is to select the number of PV modules connected in series, also called the string size. Longer strings typically lower total system costs, but the string size must still meet relevant electrical standards to ensure that the maximum system voltage remains less than the design voltage. Traditional methods calculate string size using the temperature coefficient of open-circuit voltage assuming that the coldest expected temperature occurs simultaneously with a full-sun irradiance of 1000 W/m^2. Here, we demonstrate that this traditional method is unnecessarily conservative, resulting in a string size ∼10% shorter than necessary to maintain system voltage within limits. Instead, engineers may determine string size by modeling open-circuit voltage over time using historical weather data, a method consistent with the 2017 National Electric Code. For bifacial systems, we derive a simple additive term that predicts the additional voltage rise. We demonstrate that this site-specific modeling procedure predicts open-circuit voltages in close agreement with data from field measurements. We further perform a comprehensive sensitivity analysis to identify an appropriate safety factor. By using site-specific modeling instead of traditional methods, we estimate a ∼1.2% reduction in levelized cost of electricity, a significant improvement to PV power plant economics. The method is provided as an easy-to-use web tool and as an open-source Python package.**

***Index Terms*—Design standards, design tools, photovoltaic systems, solar panels.**

I. Introduction and Economic Impact

MINIMIZING costs for solar photovoltaic (PV) power plants is a multifaceted optimization problem. The benefits of increased system dc design voltage are well known and have caused the industry to transition away from 600 V systems in favor of 1000 and 1500 V systems. By increasing the string size (the number of modules connected in series), more power travels through each wire run, reducing system costs associated with wiring, combiner boxes, and installation labor.

In this article, we develop and validate a method for calculating the maximum string size by simulating the PV system using historical weather data, datasheet module parameters, and the system design (see Fig. 1). This method is consistent with the 2017 National Electric Code (NEC) 690.7(A)(3) standard [1]. While this site-specific modeling approach is currently used in some PV power plant designs [2], the complexity of the calculation and lack of industry-standard techniques limit its use.

The traditional 690.7(A)(1) method for determining the maximum string size uses the lowest expected ambient temperature and an irradiance of 1 sun to compute the maximum open-circuit voltage (V_{oc}) using standard translation equations [3], [4]. This method was designed to be a simple way to determine safe string lengths, but yields unnecessarily conservative results. This is because maximum V_{oc} typically occurs during winter mornings when the plane-of-array (POA) irradiance is significantly reduced and the cell temperature is 5–20 °C higher than ambient. This article shows that traditional methods result in a string size ∼10% shorter than necessary, a significant missed opportunity for system-level cost optimization.

Longer strings can lower the levelized cost of electricity (LCOE) in a number of ways.

1) Electrical balance-of-systems costs are reduced because more modules are connected in each wire run and to each combiner box.
2) Installation labor is reduced because there are fewer trenches to dig and wires to install.
3) Some shutoffs due to low voltage during very warm weather can be avoided. Similarly, system degradation can lead to chronic undervoltage that prevents the inverter from starting; longer strings reduce this risk.
4) Longer strings can enable a better utilization of the available ground area, lowering land costs per watt.
5) Often for single-axis tracking systems, one more module per string can be added without changing the number of piles or controllers, reducing structural balance-of-systems costs.
6) DC ohmic losses are slightly reduced due to the use of higher voltages, improving power output.
7) Inverter costs decrease because an inverter with a higher minimum voltage can be used. These inverters have a higher nominal ac output voltage, resulting in improved inverter power output at similar electronics costs and a reduced cost of the step-up transformer.

A rough estimate of the economic impact of increasing string size can be made by assuming that the same number of modules are reorganized into 10% longer strings. A 50-MW utility-scale single-axis-tracking system has a total installed cost of $1.21/WDC [5]. The electrical balance-of-systems costs make

Manuscript received October 9, 2019; revised December 12, 2019 and January 14, 2020; accepted January 20, 2020. Date of publication February 17, 2020; date of current version April 21, 2020. This work was supported by the Durable Modules Consortium, an Energy Materials Network Consortium funded by the U.S. Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy, Solar Energy Technologies Office. The Lawrence Berkeley National Laboratory was supported by the DOE under Award DE-AC02-05CH11231. *(Corresponding author: Todd Karin.)*

The authors are with the Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: toddkarin@lbl.gov; ajain@lbl.gov).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JPHOTOV.2020.2969788

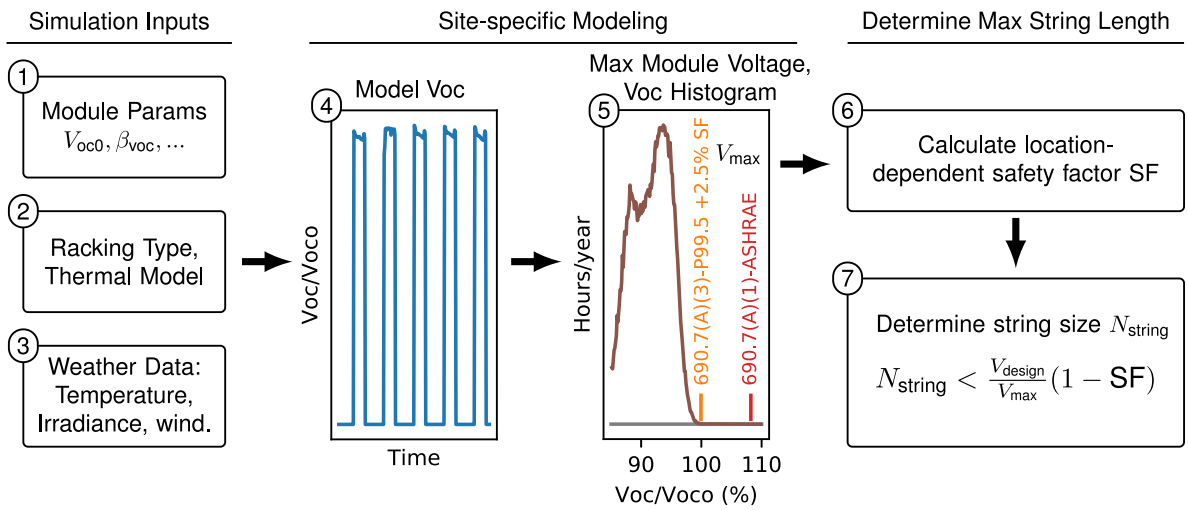


Fig. 1. Method for using site-specific modeling to determine string size. First, simulation inputs are chosen, consisting of datasheet module parameters, the racking type (orientation of the array and whether the system is fixed-tilt or single-axis tracking), and the thermal model for cell temperature as a function of weather data. Next, weather data are acquired for the site of interest. These simulation inputs are used to model V_{oc} over time and, subsequently, produce a histogram of the V_{oc} values. Several standard values for the maximum predicted module voltage V_{max} are calculated based on this histogram. A safety factor is chosen based on local weather data uncertainty and other uncertainties in the calculation. Using the system design voltage V_{design} (typically 600, 1000, or 1500 V), the longest acceptable string size is calculated using the equation shown.

up \$0.14/WDC, the installation labor \$0.11/WDC, and the structural balance of systems \$0.15/WDC. If the system is rearranged into 10% longer strings, the electrical balance-of-systems costs will be reduced by 10%. The installation labor comprises site preparation, pile driving, trench digging, wiring, and more; we estimate a 5% reduction in installation labor. Structural balance-of-systems costs are reduced by 2.3% because of the smaller number of piles and controllers [6]. In total for the 50-MW system, this results in a ∼1.6% reduction in capital costs. Using standard financial models with up-to-date system designs this translates to a ∼1.2% reduction in LCOE [7], demonstrating the significant cost savings possible just by reorganizing strings.

II. Standards for String Size Design

A. National Electric Code

The 2017 NEC section 690.7 includes three acceptable methods for calculating the string size of a PV system [1].

1) Method 690.7(A)(1) specifies that the maximum module voltage should be calculated using the open-circuit voltage corrected by the "lowest expected ambient temperature" and the open-circuit voltage temperature coefficient. One acceptable data source is the extreme annual mean minimum design dry bulb temperature found in the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) handbook [8].
2) Method 690.7(A)(2) is similar to 690.7(A)(1), but provides a V_{oc} correction factor for mono and multicrystalline silicon instead of the datasheet temperature coefficient.
3) Method 690.7(A)(3) is relevant to PV systems 100 kW or larger and states that a documented and stamped design using an industry standard method can be provided by a professional electrical engineer, further suggesting that one standard method for calculating the maximum voltage of a PV system is the Sandia PV array performance model [9].

In this article, we thoroughly explore techniques for calculating string size using the Sandia model that are consistent with 690.7(A)(3). We use weather data from the National Solar Radiation Database (NSRDB) [10], as suggested in the Sandia array performance model [9]. The NEC provides no further guidance on how to select the maximum module voltage; our aim in this article is to develop the industry-standard methodology for this task.

B. Discussion of Standards for String Sizing

Several different standards for the maximum module voltage expected (V_{max}) are compared in this article.

1) Standard 690.7(A)(1)-ASHRAE is the method used traditionally, which calculates the maximum module voltage by correcting V_{oc} at standard test conditions using the extreme annual mean minimum dry bulb temperature from the ASHRAE database and a POA irradiance of 1000 W/m^2.
2) Standard 690.7(A)(1)-NSRDB is similar to 690.7(A)(1)-ASHRAE but uses the extreme annual mean minimum dry bulb temperature calculated from NSRDB.
3) Standard 690.7(A)(1)-DAY uses the extreme annual mean minimum design dry bulb temperature during daytime (defined as when the global horizontal irradiance is larger than 150 W/m^2) and a POA irradiance of 1000 W/m^2 (data source is NSRDB). This standard avoids using temperatures during nighttime when there is little chance of creating a voltage violation.
4) Standard 690.7(A)(3)-P99.5 is a site-specific modeling method that uses a maximum module voltage equal to the 99.5th percentile V_{oc} simulated over 19 years.

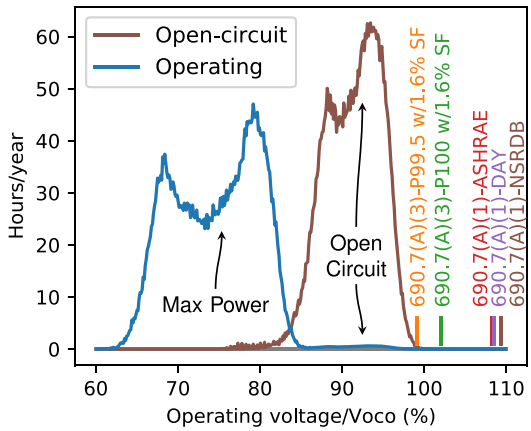


Fig. 2. For a typical PV system, comparison of open-circuit voltage history and operating voltage history assuming the system operates at open-circuit voltage 1% of the time. Also shown are the standard V_{max} values. Simulation is for a single-axis tracking PV system located in Berkeley, CA, USA, with a temperature coefficient of $\beta_{\text{voc}} = -0.37\%/^{\circ}\text{C}$.

5) Standard 690.7(A)(3)-P100 is a site-specific modeling method that uses a maximum module voltage equal to the 100th percentile (historical maximum) of the simulated V_{oc} over 19 years.

When calculating string size, maximum voltage values for 690.7(A)(3)-P99.5 and 690.7(A)(3)-P100 should be increased by a safety factor depending on system design and location, as discussed further in Section V.

Due to unpredictable weather conditions, any choice of string size still allows a statistical possibility that the design voltage could be exceeded. One way to address this risk is to choose an acceptable fraction of time where the system voltage is predicted to remain below the design limit. We first note that under typical operating conditions, the PV array operates at the maximum power point (MPP) voltage (V_{mp}), which is typically 20–25% lower than V_{oc} (see Fig. 2). Thus, for the system to produce a voltage violation, a system shutdown producing open-circuit conditions must occur simultaneously with an unusually cold and clear day.

Typical down-time rates (fraction of time any given array is not exporting power) for solar PV power plants are around 1% [6], and during this time, the array may operate at open-circuit conditions. In Fig. 2, we show two example voltage histograms: open-circuit voltage and a voltage history assuming open-circuit conditions occur 1% of the time. By selecting the 690.7(A)(3)-P99.5 V_{oc} and assuming that the 1% down-time occurs randomly, statistically, the system voltage will remain within voltage limits 99.995% of the time (equivalent to a small voltage violation occurring for less than 30 min per year). While a randomly occurring 1% downtime may be the current norm, curtailment due to high solar penetration may increase the fraction of time a PV system spends at open-circuit voltage in the future. A more conservative option for the maximum module voltage is the 690.7(A)(3)-P100 value.

Typically, no safety hazards or equipment damage occur during a small voltage violation. Inverter manufacturers design inverters with a safety margin so that dc input voltages slightly higher than the specified design voltage do not cause damage. One tier-1 inverter manufacturer suggested that a typical safety margin is 50 V, at which point damage may be expected to the dc link or electromagnetic-interference-reducing capacitors. The integrated gate bipolar junction transistor bridge will typically have a higher margin, e.g., 200 V margin on a 1000 or 1500 V rated inverter, to guard against damage due to cosmic rays [11]. Therefore, exceeding the design voltage by 3–5% would not be expected to cause equipment damage or a safety hazard on some inverters. However, exceeding the inverter's maximum specified voltage by any amount could cause a hazardous failure and/or void the inverter warranty. Therefore, we recommend system design engineers engage with the inverter manufacturer directly to determine the precise overvoltage capability of the specific inverter model used.

C. Impact of Site-Specific Modeling on String Size

Using site-specific modeling improves string size over that from traditional methods. On a grid of locations in the continental US, we simulate an open-rack fixed-tilt system, oriented south at latitude tilt, a temperature coefficient of V_{oc} equal to $-0.35\%/\text{C}$ (a typical value for crystalline Si), and temperature coefficients for open racking and glass/polymer modules. In this calculation, we use a location-dependent safety factor (SF) that includes the weather data uncertainty plus 2.0%, as described in Section V. The improvement in string size by switching from 690.7(A)(1)-ASHRAE to 690.7(A)(3)-P99.5 standards depends on location [see Fig. 3(a)] but on average results in $(9.9 \pm 1.2)\%$ longer strings [see Fig. 3(b)].

III. Site-Specific V_{OC} Modeling Procedure

A. Overview

The site-specific modeling method to determine string size is described schematically in Fig. 1 and summarized as follows.

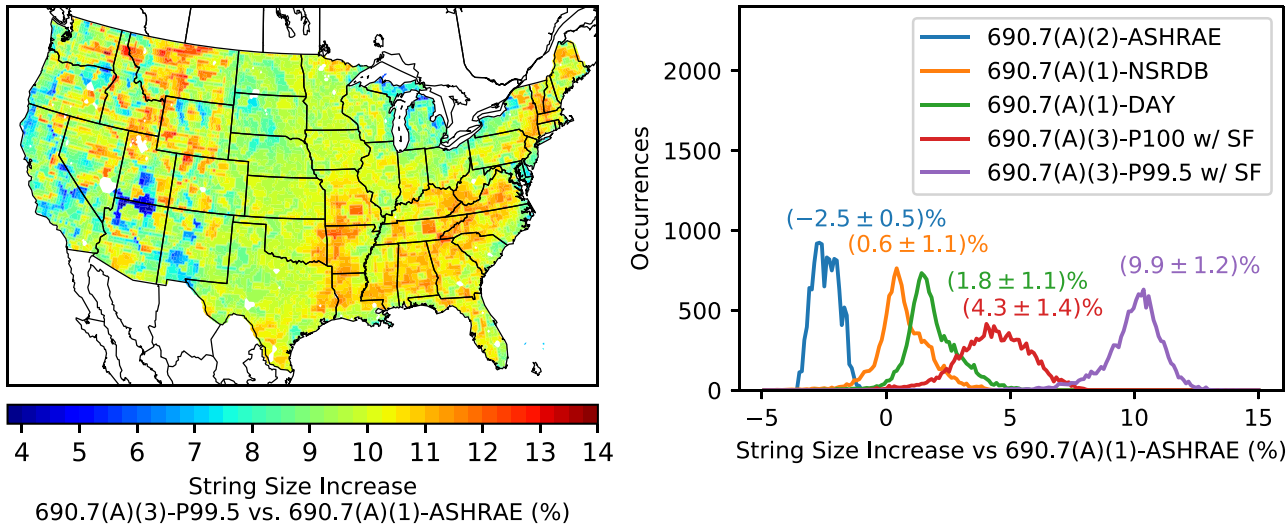
1) Identify module parameters from the module datasheet.
2) Specify racking type: fixed-tilt versus tracking, etc.
3) Acquire weather data from the NSRDB (or other source).
4) Simulate V_{oc} at each time step.
5) Calculate maximum module voltage (V_{max}) expected over system lifetime based on V_{oc} histogram.
6) Determine safety factor SF based on largest uncertainties in the calculation (see Section V).
7) Calculate the maximum string size based on system design voltage, maximum voltage V_{max}, and safety factor.

The maximum string size N_{string} is the largest integer satisfying

$$N_{\text{string}} < \frac{V_{\text{design}}}{V_{\text{max}}}(1 - \text{SF}) \tag{1}$$

where V_{design} is the system design voltage, typically 600, 1000, or 1500 V.

The system design engineer must make multiple practical decisions in order to determine a string size based on this



(a) (b)

Fig. 3. String size increase averaged across the continental US. Simulation considers an open-rack fixed-tilt system, oriented South at latitude tilt, a temperature coefficient of V_{oc} equal to -0.35%/C, and open-rack polymer-back thermal model parameters. (a) Distribution of string length improvements across the continental US from 690.7(A)(1)-ASHRAE to 690.7(A)(3)-P99.5 without any safety factor included. (b) Site-specific modeling standard 690.7(A)(3)-P99.5 results in string sizes $(9.9 \pm 1.2)\%$ longer than allowed using 690.7(A)(1)-ASHRAE. The 690.7(A)(3)-P99.5 and 690.7(A)(3)-P100 values include a safety factor equal to the location-dependent weather data safety factor plus an additional 2.0% to account for V_{oc} manufacturing uncertainty, wind data uncertainty, and diode ideality factor uncertainty.

procedure; we provide explicit guidance in Section VI. We have also developed a convenient freely available web tool [12] and open-source python package [13] for performing the above calculation.

B. Calculation of V_{oc}

Module open-circuit voltage V_{oc} can be modeled by the following equation:

$$V_{\text{oc}} = V_{\text{oc0}} + N_s \frac{n_{\text{diode}} k_B T}{q} \log\left(\frac{E}{E_0}\right) + \beta_{\text{voc}}(T - T_0) \quad (2)$$

where V_{oc0} is the open-circuit voltage at reference conditions, n_{diode} is the diode ideality factor, E is the effective plane-of-array (POA) irradiance, N_s is the number of cells in series in the module, $E_0 = 1000$ W/m^2 is the reference irradiance, $T_0 = 25$°C is the reference temperature, β_{voc} is the open-circuit voltage temperature coefficient, k_B is the Boltzmann constant, and q is the elementary charge (see Appendix A for derivation) [9]. The temperature coefficient of V_{oc} is

$$\beta_{\text{voc}} = \beta_{\text{voc0}} + M_{\text{bvoc}}(1 - E/E_0) \quad (3)$$

where M_{bvoc} is the coefficient of the irradiance dependence of the open-circuit voltage temperature coefficient, typically assumed to be zero. Equation (2) is an excellent approximation over a wide variety of operating conditions, but does not apply for extremely low ($<10\,\mu$ W/m^2) irradiances. Physically, V_{oc} cannot be less than zero, so if the effective irradiance is low enough to produce a negative V_{oc} in (2), V_{oc} is simply replaced by zero (see Appendix A for explanation) [14].

The effective irradiance is the POA irradiance modified by the spectral loss due to changes in air mass and reflection losses from the front surface of the module:

$$E = f_1(\text{AM}_\text{a})\left[E_{\text{dni}} \cos(\text{AOI}) f_2(\text{AOI}) + f_D E_{\text{poa,diffuse}}\right] \quad (4)$$

where $f_1(\text{AM}_\text{a})$ is the spectral loss coefficient depending on absolute (dimensionless) air mass AM_a, E_{dni} is the direct normal irradiance, $f_2(\text{AOI})$ is the reflection loss coefficient, f_D is the fraction of diffuse irradiance arriving at the cell, $E_{\text{poa,diffuse}}$ is the diffuse component of the irradiance in the plane of the array, and AOI is the angle of incidence. In this article, the functions f_1 and f_D are conservatively assumed equal to unity (i.e., no loss) for simplicity. The ASHRAE AOI loss model is used for f_2 [15], [16]. In our calculation of the POA diffuse irradiance, we use the Perez model for the sky component and an albedo of 0.25 for the ground component [14].

In modeling V_{oc}, the most complex part of the process is determining the AOI of radiation onto the PV module. This problem has been explored in detail and is fully implemented in PVLIB [14]. For tracking systems, we utilize the built-in algorithms for determining tracker orientation in PVLIB.

To calculate the cell temperature, we use the Sandia array performance model

$$T_{\text{cell,MPP}} = T_{\text{ambient}} + \frac{E}{E_0}\Delta T + E\exp(a + bv) \quad (5)$$

where $T_{\text{cell,MPP}}$ is the cell temperature at the MPP, E is the irradiance in W/m^2, v is the wind speed measured at 10-m height, and a, b, and ΔT are model parameters [9]. At open circuit, the module is warmer because no power is exported as electricity. In the steady state, module heating from the sun $Q_{\text{in}} \propto (1 - \eta)$ balances module thermal loss to the environment $Q_{\text{out}} \propto (T_{\text{cell}} - T_{\text{ambient}})$, where η is the module efficiency. Simplifying, we can estimate the temperature of the module at open

TABLE I
SUGGESTED BACKSIDE IRRADIANCE FRACTION f_{BACK} FOR CALCULATING V_{OC} RISE IN BIFACIAL MODULES

Ground type example	Albedo	Backside irradiance fraction (f_{back})
Grass, concrete	0.25	0.15
Dry grass	0.5	0.3
White sand	0.8	0.4
Fresh snow	1.0	0.5

Values are found by simulating backside irradiance using view-factor models at three different locations and selecting the backside irradiance fraction that most closely predicts the rise in max module voltage. Ground type examples are taken from [22].

circuit

$$T_{\text{cell,OC}} = T_{\text{ambient}} + \frac{T_{\text{cell,MPP}} - T_{\text{ambient}}}{1-\eta} \tag{6}$$

from the temperature at the MPP. We note that for the first several minutes of any system shutdown, the voltage will still be determined by the MPP temperature due to the thermal mass of the modules.

Bifacial modules may be treated similarly, except that the irradiance E is increased due to backside illumination [17]. Because of the log dependence of V_{oc} on irradiance, the backside irradiance only increases V_{oc} slightly (typically less than 1%). The cells in bifacial modules may be warmer than those in monofacial modules due to the absence of the reflective white backsheet and the increased thermal insulation present in a glass–glass module [18]. While this temperature rise might completely compensate for the increase in V_{oc} due to backside irradiance, we ignore this effect due to a lack of empirical evidence in the literature.

While backside irradiance can be calculated more precisely using view-factor or ray-tracing models [19], [20], a simple approximation is accurate enough for string sizing calculations. Since open-circuit voltage goes as the log of irradiance, by assuming that the backside irradiance is proportional to the frontside irradiance by a factor f_{back}, the increase in open-circuit voltage becomes an additive constant. From (2), the additive increase in open-circuit voltage due to backside irradiance is

$$\Delta V_{\text{oc}} = N_s \frac{n_{\text{diode}} k_B T}{q} \log\left(1 + f_{\text{back}} \eta_{\text{bifacial}}\right) \tag{7}$$

where η_{bifacial} is the efficiency of the backside of the module relative to the frontside (typically around 0.7). We emphasize that (7) applies either to an instantaneous V_{oc} or to calculated V_{max} values such as 690.7(A)(1)-ASHRAE or 690.7(A)(3)-P99.5.

Prior work has found that ground albedo is one of the strongest drivers of backside irradiance [21]. We simulate bifacial PV systems in three locations using a view-factor model [19] for different ground albedos. Simulations use a bifaciality of 0.7, a row height of 1 m, a row width of 1 m, a ground coverage ratio of 0.28, a horizon band angle of 15°, a backside PV row reflectivity of 3%, and a frontside PV row reflectivity of 1%. For each albedo, we find the f_{back} that closely predicts but still overestimates V_{oc}; see Table I for results. In locations with high wintertime albedos (snowy locations), we suggest to use $f_{\text{back}} = 0.5$, and in locations with wintertime ground albedos less than 0.25, we suggest $f_{\text{back}} = 0.15$. This typically results in a ∼1.2% increase in V_{oc} for snowy locations and a ∼0.3% increase in V_{oc} for non-snowy locations.

IV. COMPARISON WITH MEASURED DATA

To validate the V_{oc} estimation tool, we compare the simulation predictions with field-measured data for multiple deployed PV systems. The mobile performance and energy rating testbed dataset contains a total of 33 different modules deployed at three different locations: Eugene, OR; Golden, CO; and Cocoa, FL [23]. Each module is connected to a current/voltage (I–V) tracer that acquires a full I–V curve every 5 min for one to two years. Additional details are given in Appendix B.

We simulated V_{oc} using the PV system parameters and weather data from the NSRDB and compared the results to using colocated weather stations. Fig. 4(a) shows a histogram of V_{oc} values over the time period for a single module, finding that simulations based on the NSRDB slightly overestimate V_{oc}, while those based on local weather are more accurate. In comparing simulated versus measured 690.7(A)(3)-P99.5 V_{oc} across the 33 different modules [see Fig. 4(b)], we find an average fractional difference

$$f = (\text{Simulated P99.5}V_{\text{oc}})/(\text{Measured P99.5 } V_{\text{oc}}) - 1 \tag{8}$$

of $(0.5 \pm 0.5)\%$ for multicrystalline Si (mSi) and monocrystalline Si (xSi) modules when using NSRDB data. The direction of this error is such that a design based on site-specific modeling would be less likely to see voltage violations.

In order to determine the origin of this V_{oc} overestimate, we also simulated V_{oc} using the local weather station data [see Fig. 4], resulting in a more accurate estimation of V_{oc} with a fractional difference of $(-0.3 \pm 0.4)\%$ for mSi and xSi. This empirical comparison demonstrates the excellent predictive power of site-specific modeling.

V. SAFETY FACTOR AND SENSITIVITY ANALYSIS

With 690.7(A)(3) methods, it is important to add a safety factor in order to account for the uncertainties in the calculation. This is in contrast to 690.7(A)(1,2) where the standard practice is to neglect a safety factor since the calculation is inherently conservative.

A summary of the sensitivity analysis is shown graphically in Fig. 5. We provide explicit guidance on how to combine this information into a safety factor in Section VI. The largest uncertainty in the modeled maximum voltage is the quality of the NSRDB weather data, which leads to a variable safety factor depending on location. We further note that due to the standard overbuild of inverters (see Section I), there can be a built-in additional safety factor not accounted for here.

A. Quality of NSRDB Weather Data

To determine the accuracy of the NSRDB temperatures used in the V_{oc} calculation, we compare the mean yearly minimum dry bulb temperatures between NSRDB and ASHRAE data

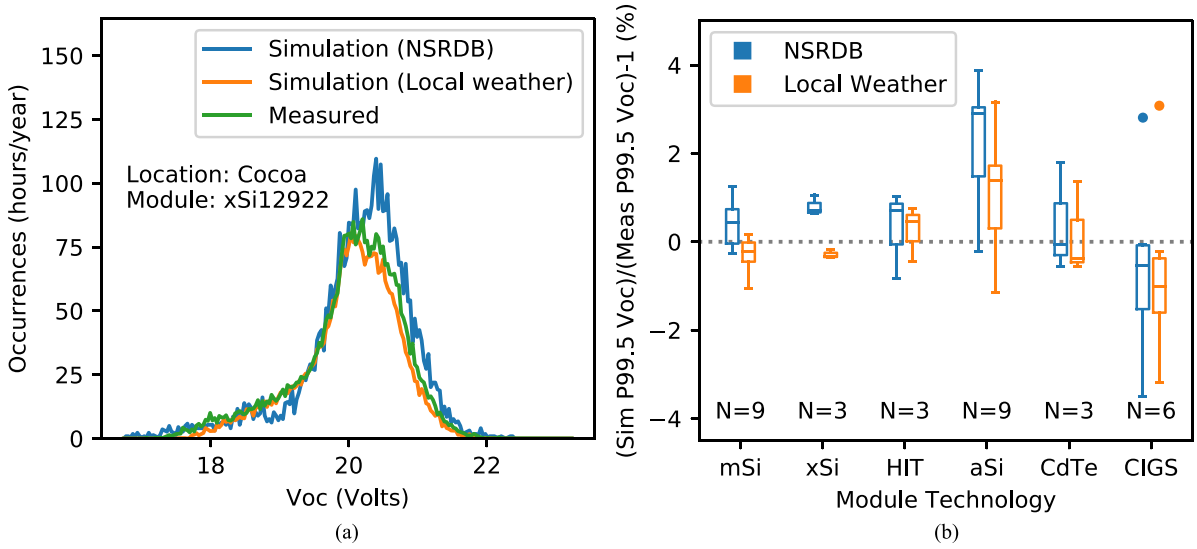


Fig. 4. Comparison of measured V_{oc} with simulation. (a) Histogram of V_{oc} for a single module type in Cocoa, FL, USA (other locations/modules not shown). The simulated V_{oc} using NSRDB weather data slightly overestimates V_{oc}, where the simulation using the local weather data is more accurate at reproducing observed V_{oc}. (b) Comparison of the simulated and measured 690.7(A)(3)-P99.5 values using two different weather data sources using measured data from three locations. Simulations using NSRDB tend to overestimate V_{oc} by $(0.5 \pm 0.5)\%$ for mSi and xSi (as compared to measured values), while simulations using the local weather data underestimate V_{oc} by $(-0.3 \pm 0.4)\%$.

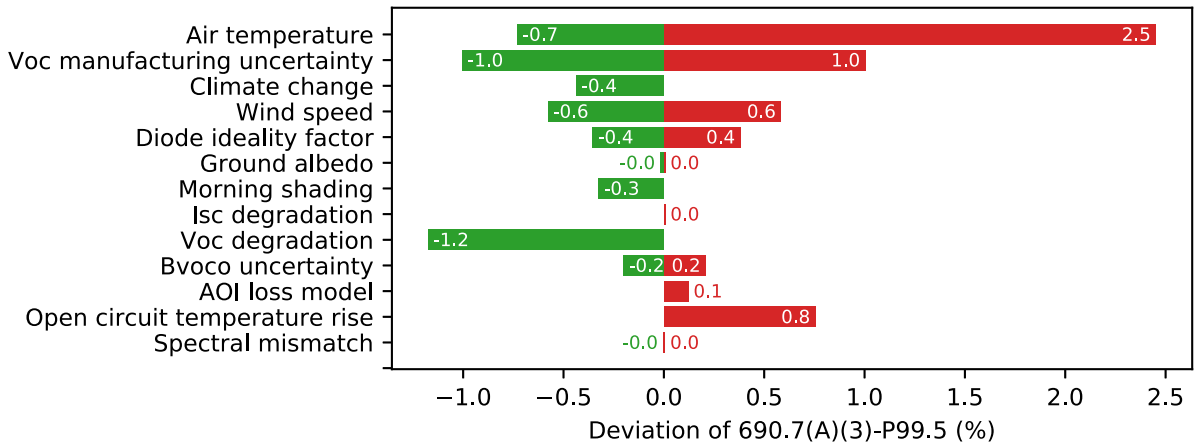


Fig. 5. Sensitivity of 609.7(A)(3)-P99.5 value with respect to the major uncertainties in the calculation. The length of the bars represents the average deviation for a monofacial single-axis-tracker Si PV system in three locations: Pheonix, Seattle, and Chicago. Positive deviations (shown in red) imply that the true P99.5 V_{oc} will be higher than the simulation predicts with using standard assumptions and necessitate the inclusion of a safety factor. The air temperature sensitivity arises from the error in the NSRDB database of -2 to 6.7 °C. V_{oc} manufacturing uncertainty occurs because not all modules in a production line will have V_{oc} with exact specification. The wind speed limits show the impact of changing the wind speed to 50% or 150% of its value in the dataset. The diode ideality factor represents the difference between using an ideality factor of 1.1 or 1.3 (1.2 is reference). Ground albedo is studied from 0 to 1, with 0.25 being the reference. Shading modifies the irradiance so that the first and last hours of the day have zero irradiance. I_{sc} degradation shows the impact of a 0.75%/year degradation in I_{sc} due to, e.g., encapsulant discoloration or soiling (simulation time is 19 years). V_{oc} degradation shows the impact of a -0.15%/year V_{oc} degradation. β_{voc0} uncertainty shows the impact of increasing or decreasing the V_{oc} temperature coefficient by 5% relative (reference is -0.35%/°C). The AOI loss model shows the impact of assuming no loss from the top glass layer, as opposed to the reference case using the ASHRAE angular reflection model with an AIM parameter of 0.05. Open-circuit temperature rise shows the impact of ignoring the increased temperature at open-circuit conditions; default is to include this effect. Spectral mismatch shows the effect of a $\pm 2\%$ spectral mismatch occurring uniformly over time.

sources [8]. For each location in the ASHRAE database, we find the closest data point in the NSRDB database and calculate the extreme annual mean minimum dry bulb temperature. Of the 8118 points in the global ASHRAE dataset, we kept 1756 points that were within 5 km of an NSRDB data point and in the USA. The temperature difference between the two datasets for ASHRAE points in the continental US is shown in Fig. 6(a). The weather data are more accurate in the less mountainous regions of the country. For systems in these locations, a lower safety factor may be used due to the greater certainty of the temperature data.

A histogram of the temperature difference between these two data sources [see Fig. 6(b)] shows the degree to which the NSRDB temperature data are consistent with ASHRAE values. On average, the NSRDB overestimates the minimum temperature by 1.9 °C, with 95% of locations having an

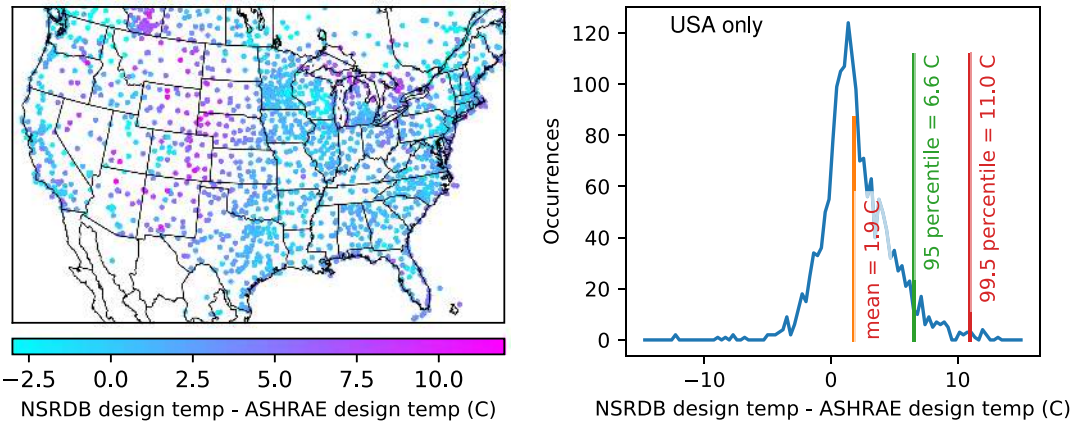


Fig. 6. (a) Difference between the NSRDB and ASHRAE mean yearly minimum design dry bulb temperature at locations of ASHRAE weather stations. On average, the temperatures agree to within $(1.9 \pm 2.8)°$C. (b) Histogram showing design temperature differences between the NSRDB and ASHRAE. For 95% of ASHRAE locations in the USA, the NSRDB design temperature exceeds the ASHRAE temperature by less than 6.6 °C. This temperature uncertainty is used to find a safety factor using the module's temperature coefficient of V_{oc}.

overestimate below 6.7 °C. Using a typical temperature coefficient for silicon solar cells of $\beta_{voc}/V_{oc0} = -0.35\%/°$C, a safety factor of 2.3% would correct weather data quality inaccuracy for 95% of locations in the US, and we suggest this safety factor if a location-dependent safety factor is not used. The location-dependent temperature error in Fig. 6(a) can be used to justify a lower safety factor for locations where the NSRDB design temperature agrees closely with the ASHRAE design temperature (these data are also available in the python package [13] and web tool [12]).

B. Climate Change

Most climate models predict that extreme cold temperatures will gradually disappear. By mid-century, current 1-in-20-year temperature minima are not projected to occur at all in a higher emission scenario [24], making voltage violations less likely. The projected increase in the yearly extreme minimum temperature ranges from 2 to 5.5 °C per 55 years across the continental US [25]. Over a 35-year lifetime of a solar power plant, this corresponds on average to an increase of 0.6–1.8° C in minimum temperatures and, therefore, to a decrease in the maximum module voltage of of 0.2–0.6%.

C. Degradation

Multiple degradation mechanisms can cause V_{oc} to degrade as the module ages. However, in a review of over 1000 studies, which measured V_{oc} degradation, V_{oc} degradation was not found to be significantly different from zero even though individual studies are on the order of -0.15% per year [26]. For the purpose of string size estimates, we do not suggest using a V_{oc} degradation rate in the calculation.

D. Other Uncertainties

We compare the sensitivity of the 690.7(A)(3)-P99.5 value to different assumptions and uncertainties in the model. We calculate the fractional change in 690.7(A)(3)-P99.5 due to different uncertainties for a single-axis-tracking monofacial xSi PV system with a temperature coefficient of V_{oc} of 0.35%/°C at three locations (Pheonix, Seattle, and Chicago), summarized in Fig. 5.

1) The V_{oc} manufacturing uncertainty was estimated to be on the order of 1% based on a literature search: finding standard deviations of 0.7% [27] and 0.5% [28].
2) Wind speed is notoriously difficult to predict from gridded weather data. The impact of wind speed errors was found by changing the wind speed in the input data to 50% or 150% of its original value, leading to a $\pm 0.5\%$ change in V_{max}.
3) The diode ideality factor is an important parameter of solar cells that is not provided directly on the datasheet. Suggested values of diode ideality factor are 1.1 for mono-c-Si, 1.2 for multi-c-Si, and 1.4 for CdTe. An additive error of 0.1 in the diode ideality factor leads to a $\pm 0.4\%$ change in V_{max}.
4) Errors in ground albedo do not significantly change V_{max} for a monofacial system.
5) The impact of morning shading was simulated by assuming that the first and last hours of the day have no irradiance, finding a change of -0.3%.
6) Some locations have larger extreme minimum temperature fluctuations than other locations. This can be quantified by comparing the extreme annual mean minimum dry bulb temperature with the 20-year return period extreme minimum dry bulb temperature. Gulf-coast states and California have extreme minimum temperatures that are only moderately colder than average, while parts of the desert southwest and midwest have infrequent but anomalously cold periods (see Fig. 7). When using 690.7(A)(3)-P99.5, the propensity for some locations to experience anomalously cold weather can be accounted for as an additional safety factor by multiplying the cold temperature fluctuation magnitude by $|\beta_{voc}|/V_{oc0}$.

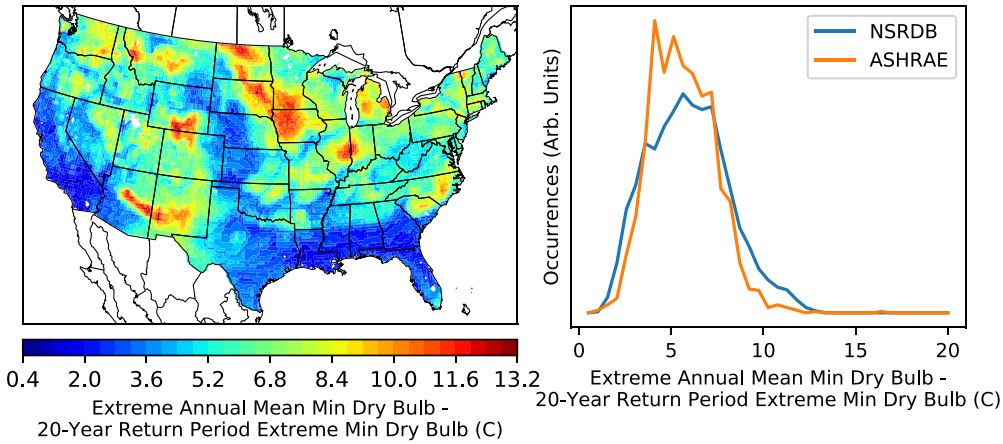


Fig. 7. (a) Distribution of extreme cold temperature fluctuations in the continental US. Data are calculated from the NSRDB. The extreme annual mean minimum dry bulb temperature is the average of the coldest temperature each year. (b) Histogram of cold temperature fluctuations in the continental US. The histograms for NSRDB and ASHRAE datasets are similar even though the NSRDB is uniformly distributed, while ASHRAE stations are nonuniformly distributed across the USA. Using NSRDB data, 99% of locations have an extreme temperature fluctuation under 11.2 °C.

VI. Guidance for PV System Design

In this section, we provide explicit guidance for designing strings according to NEC 690.7(A)(3).

1) *Choice of standard:* Use the 690.7(A)(3)-P99.5 voltage instead of the 690.7(A)(3)-P100 voltage unless one of the following conditions is met.
   a) Open-circuit conditions are expected to occur more than 1% of the time ($>$ 88 h per year).
   b) Equipment (e.g., inverter) damage or safety hazards are expected if system voltage exceeds the design voltage by 3%.
   c) A more conservative design is desired.
2) *Safety factor:* An appropriate safety factor can be determined by summing the following uncertainties.
   a) Determine the air temperature uncertainty for the location of interest from the data available in Fig. 6, the web tool [12] or the Python library [13]. If the location is highly mountainous, ASHRAE data points are too sparse or a simpler method desired, use a standard weather temperature uncertainty of 6.6 °C. Multiply this temperature uncertainty by the absolute value of the module temperature coefficient of open-circuit voltage (in %/°C) to find the safety factor.
   b) Add 1.0% due to V_{oc} manufacturing uncertainty.
   c) Add 0.6% due to wind speed uncertainty.
   d) If the diode ideality factor is unknown, add 0.4%.
   e) If the 690.7(A)(3)-P99.5 standard is used and extra protection against overvoltage is desired, add a safety factor due to cold weather variability. Find the difference between the mean yearly minimum dry bulb temperature and the 20-year return period extreme minimum dry bulb temperature (NSRDB data available in Fig. 7, web tool, Python library, or ASHRAE data from [8]). Multiply this difference by the absolute value of the module temperature coefficient of open-circuit voltage (in %/°C) to find the additional safety factor.
3) *Temperature rise at open circuit:* The module temperature rise due to open-circuit conditions should typically be included in order to accurately predict the voltage during system shutdowns. However, if an extra degree of caution is desired for the first several minutes of a shutdown where temperatures have not risen yet, do not include temperature rise in the modeling.
4) *Bifacial modules:* The extra voltage rise due to a module's bifaciality can be calculated using (7). Use a backside irradiance fraction from Table I.
5) *Diode ideality factor:* If the diode ideality factor is unknown, suggested values are 1.1 for mono-Si, 1.2 for multi-Si, and 1.4 for CdTe.
6) *Nonstandard technologies:* The methods in this article should only be applied to crystalline Si and CdTe modules mounted with fixed-tilt or single-axis tracking. Application to niche technologies such as concentrating PV, perovskites, organic PV, and others is not recommended.

VII. Conclusion

In this article, we have validated a method to determine maximum string size consistent with NEC 690.7(A)(3) using site-specific modeling. We suggest safety factors based on the largest sources of uncertainty in the calculation. On average, this technique allows for a 10% increase in string size over traditional methods and a 1.2% reduction in LCOE, a substantial improvement achieved simply by reorganizing modules into longer strings. This analysis is implemented in a freely available web tool [12] and an open-source Python package [13]. We suggest that this work should be added to future editions of the NEC as a suggested method for designing systems of all sizes.

Appendix A
Dependence of V_{oc} on Irradiance and Temperature

The dependence of V_{oc} on irradiance and temperature [see (2)] is derived in this section from basic principles. While the final derived equation is familiar, it is helpful to demonstrate the

assumptions typically used in calculating V_{oc}. We begin with the Shockley diode equation, modified by adding the standard diode ideality factor n_{diode}

$$I = I_L - I_s \left[\exp\left(\frac{qV}{n_{\text{diode}} k_B T} \right) - 1 \right] \tag{9}$$

where I is the cell current, V is the cell voltage, I_L is the photocurrent, I_s is the reverse saturation current, and T is the absolute temperature in Kelvin [29]. The photocurrent is

$$I_L = qA \int_{E_g}^{\infty} \alpha(\hbar\omega) \frac{d\Phi_\gamma(\hbar\omega)}{d(\hbar\omega)} d(\hbar\omega) \tag{10}$$

where $\alpha(\hbar\omega)$ is the absorption coefficient of the cell, $\Phi_\gamma(\hbar\omega)$ is the number of photons incident with energy $\hbar\omega$ [30], and A is the area of the cell.

Equation (10) shows that the photocurrent is proportional to the incident light intensity. In addition, the photocurrent increases with increasing temperature because as the bandgap decreases, more photons are absorbed [31]. The standard first-order approximation of this effect around a reference condition is

$$I_L = \frac{E}{E_0} \left[I_{L0} + \alpha_{\text{Isc}} (T - T_0) \right] \tag{11}$$

where E is the irradiance reaching the cell, modified by spectral losses and reflection losses, E_0 is the reference irradiance of 1000 W/m^2, T_0 is the reference temperature of 25 °C, α_{Isc} is the temperature coefficient of short-circuit current, and I_{L0} is the photocurrent at reference conditions [9], [32]. In (11), α_{Isc} has units of A/°C to agree with the convention in PVLIB [14], while in [9], normalized 1/°C units are used. Using the AM1.5 spectrum and $E_g = 1.1$ eV, the photocurrent has a value of $I_L/A = 35$ mA/cm^2 [29]. The ideal saturation current I_s is typically provided as an empirically measured parameter.

An expression for the open-circuit voltage can be found by setting $I = 0$ in (9), resulting in

$$V_{\text{oc}} = \frac{n_{\text{diode}} kT}{q} \log\left(\frac{I_L}{I_s} + 1 \right). \tag{12}$$

Note that V_{oc} is zero if $E = 0$. At standard test conditions (25 °C and 1000 W/m^2), the ratio I_{L0}/I_{s0} ranges from 1×10^8 to 1×10^{15} for modules in the California energy commission (CEC) database (with mean value 8×10^9 for crystalline silicon cells) [14]. Thus, for all practical nonzero irradiances (i.e., greater than 10 μ W/m^2), the 1 can be ignored in (12). With this approximation (as is conventional), a correction must be applied in any numerical calculation to ensure that V_{oc} does not go to negative infinity when the irradiance is zero, for example, by replacing all negative V_{oc} by zero.

Combining (11) and (12), we find

$$\begin{aligned} V_{\text{oc}} = {} & \frac{n_{\text{diode}} kT}{q} \log\left[\frac{E}{E_0} \right] + \frac{n_{\text{diode}} kT}{q} \log\left[\frac{I_{L0}}{I_s} \right] \\ & + \frac{n_{\text{diode}} kT}{q} \log\left[1 + \frac{\alpha_{Isc}}{I_{L0}} (T - T_0) \right]. \end{aligned} \tag{13}$$

By evaluating (13) at reference conditions, incorporating (11) and subtracting, we find

$$\begin{aligned} V_{\text{oc}} = {} & V_{\text{oc0}} + \frac{n_{\text{diode}} kT}{q} \log\left[\frac{E}{E_0} \right] \\ & + \frac{n_{\text{diode}} k}{q} \left[\log\left(\frac{I_{L0}}{I_{s0}} \right) - \frac{E_g}{kT_0} \right] (T - T_0) \\ & + \frac{n_{\text{diode}} kT}{q} \log\left[1 + \frac{\alpha_{Isc}}{I_{L0}} (T - T_0) \right] \end{aligned} \tag{14}$$

where I_{s0} is the saturation current at reference conditions. For silicon PV, a typical value of α_{Isc}/I_{L0} is 0.05%/K [33], making the argument of the log in the third term of (14) close to unity for common operating temperatures. Therefore, to an excellent approximation, the log can be expanded and combined with the second term, leading to the well-known equation [9] for open-circuit voltage

$$V_{\text{oc}} = V_{\text{oc0}} + \frac{n_{\text{diode}} kT}{q} \log\left(\frac{E}{E_0} \right) + \beta_{v\text{oc}} (T - T_0) \tag{15}$$

where we have identified the temperature coefficient of open-circuit voltage

$$\beta_{\text{voc}} = \frac{n_{\text{diode}} k}{q} \left[\log\left(\frac{I_{L0}}{I_{s0}} \right) - \frac{E_g}{kT_0} + \frac{\alpha_{Isc} T}{I_{L0}} \right]. \tag{16}$$

Most commonly, the temperature-dependent term in (16) is ignored. For typical silicon PV modules, this leads to an error in predicting V_{oc} of less than 0.01%. We note that for certain devices with higher values of α_{Isc}, the effect can be much larger [34]. For the purpose of calculating V_{oc} for string sizing, we ignore the temperature-dependent term in (16) and assume that β_{voc} is constant, similarly to other treatments [9], [14].

Appendix B
Details on Comparison With Measured Data

The field-measured I–V curve is processed to obtain the open-circuit voltage with a typical uncertainty of 0.4 V (with module V_{oc0} ranging from 21 to 223 V). The module parameters, including open-circuit voltage at standard test conditions, temperature coefficients, and diode ideality factor, were measured on the individual modules after deployment. We note that the precision and accuracy found in this comparison may be due to the fact that the individual modules were characterized, rather than using generic datasheet parameters.

Proper handling of missing data points in the measured data (measurement only occurs during daytime) is important to not bias the 690.7(A)(3)-P99.5 value. First, we linearly interpolate up to two missing data points to account for short lapses. Small changes to the number of interpolated data points do not change the results. The majority of missing data occurs during the night, so missing values are set to a V_{oc} of zero. Some periods of daytime data (8–19%) are also missing from the series; these days were removed before calculating the 690.7(A)(3)-P99.5 value.

ACKNOWLEDGMENT

The authors would like to thank J. Meydbray, who brought our attention to this problem. The authors would also like to thank M. Woodhouse for assistance with the technoeconomic analysis. B. Marion provided field data from the mobile performance and energy rating testbed project. The authors would further like to thank a number of experts in solar PV system design who provided feedback on the method: J. Webber, G. Loomis, M. Waters, R. Blumenthal, K. Anderson, R. White, B. Custodio, P. Brucke, K. Lichtenstein, M. Rosi, and T. Tolliver.

REFERENCES

[1] N. N. F. P. Association, *NFPA 70: National Electrical Code 2017*. Delmar Cengage Learning, 2016.

[2] C. Ladd, "Simulating NEC voltage and current values," *Sol. Professional*, vol. 11.4, pp. 12–18, Jul. 2018.

[3] String sizing tool, PV system builder Rev 3.13. [Online]. Available: https://www.solectria.com/technical-support/string-sizing-tool/. Accessed on: Feb. 6, 2020.

[4] B. Brooks, "Array voltage considerations," *Sol. Professional*, vol. 3.6, pp. 74–78, Oct. 2010.

[5] R. Fu, D. Feldman, and R. Margolis, "U.S. solar photovoltaic system cost benchmark: Q1 2018," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-6A20-72399, Nov. 2018. [Online]. Available: https://www.nrel.gov/docs/fy19osti/72399.pdf

[6] "Private communication with cypress creek renewables," Private Communication, Apr. 2019.

[7] N. Blair *et al.*, "System advisor model (SAM) general description (Version 2017.9.5)," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-6A20-70414, 2018.

[8] *ASHRAE Handbook Fundamentals 2017: Inch-Pound Edition*, Amer. Soc. Heating, Refrigerating Air-Conditioning Eng., Atlanta, GA, USA, Jun. 2017.

[9] D. King, W. Boyson, and J. Kratochvill, "Photovoltaic array performance model," Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. SAND2004-3535, 2004.

[10] M. Sengupta *et al.*, "The National Solar Radiation Data Base (NSRDB)," *Renew. Sustain. Energy Rev.*, vol. 89, pp. 51–60, 2018.

[11] U. Schilling, "Cosmic ray failures in power electronics," Semikron, Nuremberg, Germany, Appl. Note AN 17-003, 2017.

[12] T. Karin and A. Jain, String voltage calculator. [Online]. Available: https://pvtools.lbl.gov/string-length-calculator

[13] T. Karin and A. Jain, "vocmax: Calculate the maximum string length for a photovoltaic system," 2019. [Online]. Available: https://github.com/toddkarin/vocmax

[14] W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, "Pvlib Python: A python package for modeling solar energy systems," *J. Open Source Softw.*, vol. 3, no. 29, 2018, Art. no. 884.

[15] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Sol. Energy*, vol. 44, no. 5, pp. 271–289, 1990.

[16] J. E. Hill, J. P. Jenkins, and D. E. Jones, "Testing of solar collectors according to ASHRAE Standard 93–77," *ASHRAE J.*, vol. 84, pp. 107–125, Jan. 1978.

[17] B. Wittmer and A. Mermoud, "Yield simulations for horizontal axis trackers with bifacial PV modules in PVsyst," in *Proc. 35th Eur. Photovolt. Sol. Energy Conf.*, Sep. 2018, pp. 1929–1934.

[18] M. Mittag, L. Vogt, C. Herzog, and H. Neuhaus, "Thermal modelling of photovoltaic modules in operation and production," in *Proc. 36th Eur. Photovolt. Sol. Energy Conf. Exhib.*, Sep. 2019, pp. 892–900.

[19] M. Anoma *et al.*, "View factor model and validation for bifacial PV and diffuse shade on single-axis trackers," in *Proc. 44th IEEE Photovolt. Spec. Conf.*, 2017, pp. 1549–1554.

[20] C. Deline, W. Marion, and S. Ayala Pelaez, "Bifacial radiance," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL SWR-18-03, 2017. [Online]. Available: https://www.osti.gov/doecode/biblio/6869

[21] I. Shoukry, J. Libal, R. Kopecek, E. Wefringhaus, and J. Werner, "Modelling of bifacial gain for stand-alone and in-field installed bifacial PV modules," *Energy Procedia*, vol. 92, pp. 600–608, 2016.

[22] T. C. R. Russell, R. Saive, A. Augusto, S. G. Bowden, and H. A. Atwater, "The influence of spectral albedo on bifacial solar cells: A theoretical and experimental study," *IEEE J. Photovolt.*, vol. 7, no. 6, pp. 1611–1618, Nov. 2017.

[23] W. Marion *et al.*, "User's manual for data for validating models for PV module performance," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-5200-61610, 2014.

[24] D. J. Wuebbles *et al.*, "Climate science special report: Fourth national climate assessment, volume I," U.S. Global Change Res. Program, Washington, DC, USA, Tech. Rep. 21564, 2017, Ch. 6, pp. 185–206.

[25] E. M. Fischer, U. Beyerle, and R. Knutti, "Robust spatially aggregated projections of climate extremes," *Nature Climate Change*, vol. 3, pp. 1033–1038, 2013.

[26] D. C. Jordan, J. H. Wohlgemuth, and S. R. Kurtz, "Technology and climate trends in PV module degradation," in *Proc. 27th Eur. Photovolt. Sol. Energy Conf. Exhib.*, 2012, pp. 3118–3124.

[27] A. M. Reis, N. T. Coleman, M. W. Marshall, P. A. Lehman, and C. E. Chamberlin, "Comparison of PV module performance before and after 11-years of field exposure," in *Proc. Conf. Rec. 29th IEEE Photovolt. Spec. Conf.,* May 2002, pp. 1432–1435.

[28] C. E. Chamberlin *et al.*, "Comparison of PV module performance before and after 11 and 20 years of field exposure," in *Proc. 37th IEEE Photovolt. Spec. Conf.*, Jun. 2011, pp. 000101–000105.

[29] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. New York, NY, USA: Wiley-Interscience, 2006.

[30] P. Wurfel, *Physics of Solar Cells*, 2nd ed. New York, NY, USA: Wiley-VCH, Mar. 2009.

[31] M. A. Green, "General temperature dependence of solar cell performance and implications for device modelling," *Prog. Photovolt., Res. Appl.*, vol. 11, no. 5, pp. 333–340, 2003.

[32] W. D. Soto, S. Klein, and W. Beckman, "Improvement and validation of a model for photovoltaic array performance," *Sol. Energy*, vol. 80, no. 1, pp. 78–88, 2006.

[33] Y. Hishikawa *et al.*, "Temperature dependence of the short circuit current and spectral responsivity of various kinds of crystalline silicon photovoltaic devices," *Jpn. J. Appl. Phys.*, vol. 57, no. 8S3, Jul. 2018, Art. no. 08RG17.

[34] M. Amara, B. Guillo Lohan, A. Kaminski-Cachopo, M. Lemiti, "A new measurement of voc temperature coefficients at very large temperature range," in *Proc. 36th Eur. Photovolt. Sol. Energy Conf. Exihb.*, Sep. 2019. [Online]. Available: https://www.eupvsec-planner.com/presentations/c48173/a_new_measurement_of_voc_temperature_coefficients_at_very_large_temperature_range.htm.