A Data-Driven String Sizing Methodology for Utility-Scale Photovoltaic Systems

Filipe G. Souza[1], Marília Braga[2], Maria Fernanda N. Silva[3] and Ricardo Rüther[2]

[1] *University of Freiburg, Freiburg (Germany)*

[2] *Universidade Federal de Santa Catarina, Florianópolis (Brazil)*

[3] *Kroma Energia, Recife (Brazil)*

Abstract

This paper presents the Data-Driven String Sizing (DDSS) method for optimizing PV module string length in utility-scale systems. Unlike conventional conservative approaches, DDSS utilizes sub-hourly, site-specific meteorological data to refine cell temperature and open-circuit voltage estimates, enabling increased string size. The method incorporates thermal inertia modeling to capture worst-case irradiance scenarios, applies statistical analysis, and integrates safety factors to address uncertainties. Two case studies in distinct Brazilian climates (Jaguaruana and Florianópolis) assessed technical and economic outcomes based on one year of data. DDSS allowed for increased string length by a single module (~3.2%), leaving energy yield essentially unchanged but yielding measurable CAPEX savings – up to 0.84% – mainly from reduced racking and construction costs. An extended string-length scenario, considering a less conservative scenario with additional modules in series, improved CAPEX savings of up to 1.96%. While benefits depend on site conditions and require high-quality on-site data, DDSS offers a replicable, data-driven approach with potential to enhance the economic performance of utility-scale PV plants.

Keywords: Photovoltaic systems, String sizing, Data-driven design, Thermal modeling, Module temperature, Utility-scale solar, Economic assessment, Levelized cost of energy, Voltage optimization.

1. Introduction

Solar photovoltaic (PV) technology has become the leading contributor to global electricity generation capacity additions. In 2024, PV accounted for approximately 81% of new installations worldwide (REN21, 2025), driven by declining costs, rising demand for clean energy, and ambitious decarbonization targets. Utility-scale PV plants, typically ranging from tens to hundreds of megawatts, require careful design optimization to maximize performance and minimize costs. One promising strategy is increasing the DC operating voltage. By enabling longer strings, this approach reduces ohmic losses and lowers material and installation expenses (Gkoutioudi et al, 2013) – ultimately improving the Levelized Cost of Energy (LCOE).

The industry has progressively adopted higher voltage standards: from 600 Vdc in the early 2000s to 1,000 Vdc, and later 1,500 Vdc systems (Jiménez and Bkayrat, 2015). In 2024, the first 2,000 Vdc PV system was commissioned, and the IEC is currently developing specifications for systems up to 3,000 Vdc (Sungrow, 2024). These developments reflect a clear trend toward higher voltages to support more cost-effective plant architectures.

However, increasing string voltage must be carefully managed. Exceeding the system's maximum voltage can compromise safety, reduce equipment lifespan, and affect operational reliability (David et al, 2022). PV module voltage is sensitive to environmental conditions – decreasing with temperature and increasing slightly with irradiance (Skoplaki and Palyvos. 2009). Accurate modeling of string voltage therefore requires precise estimation of both irradiance and module temperature.

Worst-case voltage scenarios typically occur under cold ambient conditions combined with sudden irradiance surges, such as those caused by fast-moving clouds in the early morning. These events are especially critical

10.18086/swc.2025.03.15 Available at http://proceedings.ises.org

when strings operate near open-circuit conditions (e.g., during curtailment or inverter startup), where module voltage peaks. Capturing such transients demands high-resolution meteorological data (sub-30-minute intervals), as overirradiance events may last only seconds and are often missed in hourly datasets (Nascimento et al, 2020).

Moreover, PV modules exhibit thermal inertia: temperature does not respond instantaneously to irradiance changes due to the module's thermal mass and heat transfer properties (Fuentes, 1985). This lag, illustrated in Figure 1 using data from the Fotovoltaica/UFSC laboratory, is often neglected in conventional models, leading to inaccurate voltage predictions during rapid irradiance transitions.

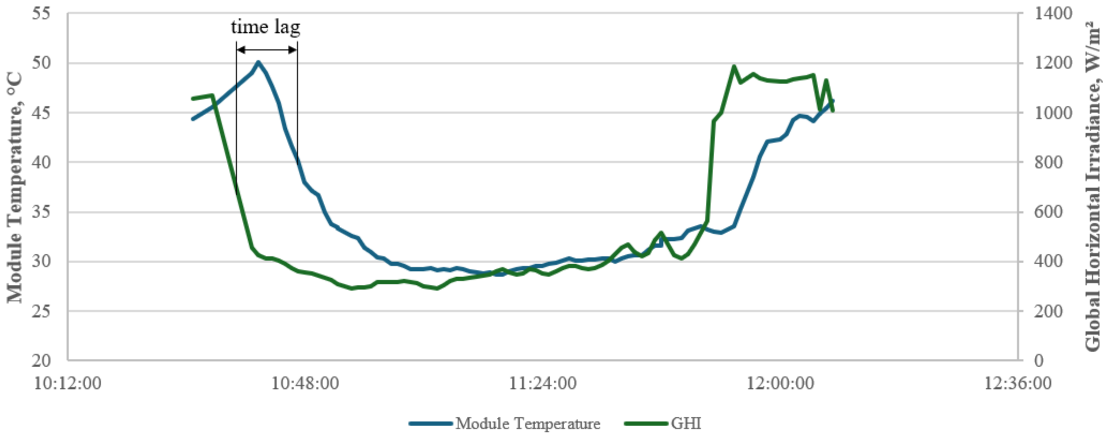


Figure 1: PV module's temperature lag following irradiance variation.

String sizing is traditionally based on conservative assumptions to ensure safety (adjusting module's open-circuit voltage from STC to the lowest expected ambient temperature), but this approach may underutilize system voltage capabilities. Some manufacturers have proposed alternative sizing methods using site-specific data, yet these lack standardized procedures and peer-reviewed validation and may pose excessive risk. Advanced methodologies have emerged in the literature (Karin and Kain, 2020; Sanchís-Gómez et al, 2025) but none fully take into consideration the temporal resolution and thermal dynamics needed for comprehensive risk assessment. Although international standards permit alternative approaches (IEC, 2023), there is no consensus on best practices for data-driven string sizing.

This work introduces a novel methodology for determining maximum string length in utility-scale PV systems, integrating historical meteorological data, thermal inertia, and system-specific parameters. The method is preliminary validated using field measurements from the Fotovoltaica/UFSC laboratory in Florianópolis, Brazil, and its economic impact is assessed. The proposed approach aims to enhance plant profitability while ensuring safe and reliable operation.

2. Data-driven string sizing methodology

2.1. Overview

The Data-Driven String Sizing (DDSS) methodology determines the maximum safe number of PV modules in series (N_s) for utility-scale plants by combining site-specific sub-hourly meteorological records, module/inverter/racking parameters, dynamic cell-temperature modeling (including thermal inertia), and a probabilistic treatment of uncertainties. DDSS uses synchronized meteorological time series (with global horizontal irradiance, ambient temeprature, and wind speed) as main input, and produces a time series of per-module open-circuit voltages (V_{oc}), selects a design non-exceedance V_{oc} value, applies an aggregated safety factor, and computes N_s from the system maximum DC voltage. Figure 2 shows the workflow of the DDSS methodology.

The methodology was validated against measurements from an instrumented module at the Fotovoltaica/UFSC facilit. Validation details are provided in Chapter 3. Implementation uses Python coding with pvlib (Anderson et al, 2023), is suited for various system architectures, and is freely available online at https://github.com/godoyf/DDSS-Methodology.

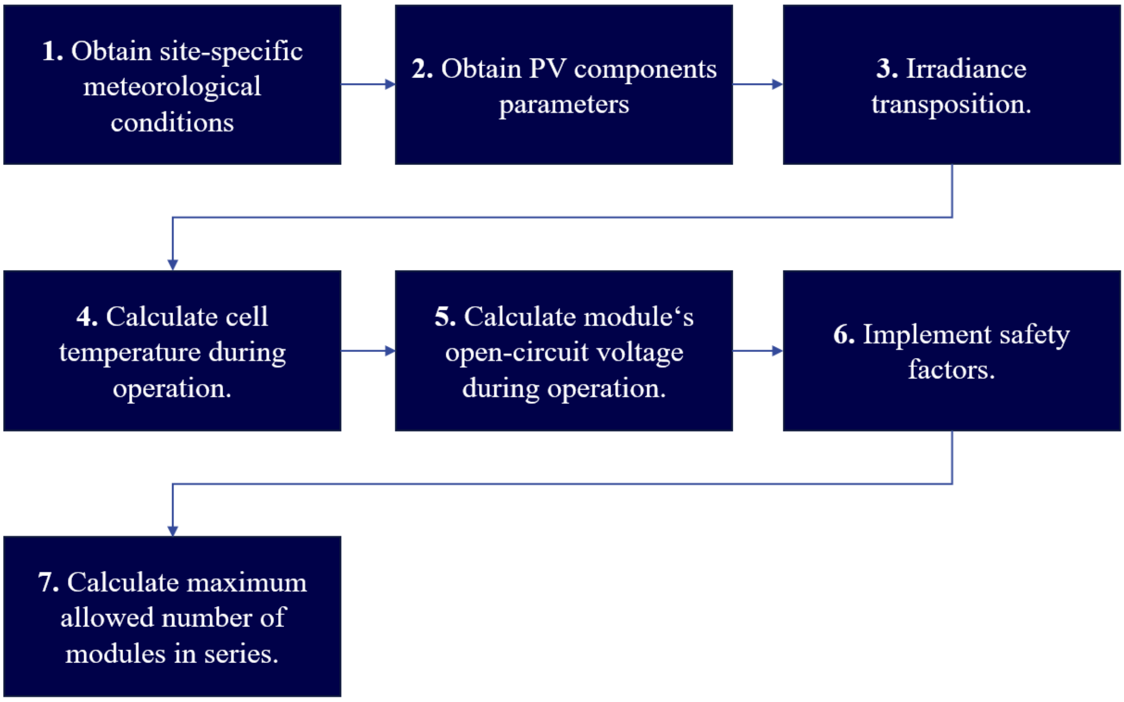


Figure 2: Workflow of the DDSS method.

2.2. Effective irradiance calculation

This step of the DDSS method involves translating the measured global horizontal irradiance (GHI) into the irradiance actually received by the photovoltaic modules – known as effective irradiance (G_{eff}). G_{eff} serves as a main input for subsequent modeling of PV cell temperature and open-circuit voltage behavior. All calculations are performed for each time step in the dataset, ensuring that the dynamic effects of solar geometry and system configuration are captured. The following computational sequence is implemented:

- **Solar Position Calculation**: The solar zenith and azimuth angles for each timestamp are computed using the *get_solarposition* function from pvlib (Reda and Andreas, 2004).

- **Decomposition of GHI**: The measured GHI is decomposed into its two main components – direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) – using the Erbs correlation model (Erbs et al, 1982), applied through the *erbs* function in *pvlib*.

- **Obtaining Tracking Angles (for tracking systems)**: For systems utilizing single-axis trackers, the surface tilt angle for each step is computed using *pvlib*'s *singleaxis* function. This calculation uses the site location, time, and tracking configuration to determine the tilt angle of the PV array.

- **Plane-of-Array Front and Rear Irradiance**: With the solar position, DNI, and DHI known, the total irradiance incident on the PV module surface (G_{front}) is computed using *pvfactors's (*Huawei Digital Power, 2025*) timeseries* function, available in pvlib. If the system employs bifacial PV modules, the contributions of rear-side irradiance (G_{back}) is also estimated using the same function.

Once G_{front} and G_{back} are obtained, G_{eff} can be calculated using the following equation:

$$G_{eff} = G_{front} + (G_{back}\ x\ Bif) \qquad \text{(eq. 1)}$$

where *Bif* is the module bifaciality coefficient of the PV module.

2.3. Cell temperature calculation

To reflect real operating conditions, accounting not only for ambient temperature and irradiance but also for the transient thermal behavior of the module, the cell temperature (T_{cell}) is estimated using the Fuentes thermal model

$$mc\frac{dT_{cell}}{dt} = \alpha G_{eff} - h_c(T_{cell} - T_{amb}) - \epsilon\sigma(T_{cell}^4 - T_{sky}^4) - \epsilon\sigma(T_{cell}^4 - T_{grnd}^4) \quad \text{(eq. 2)}$$

where m is the module mass per unit area, c the specific heat, α the absorptance, h_a the convective heat transfer coefficient, ε the emissivity, σ the Stefan-Boltzmann constant, and T_{sky} and T_{grnd} the effective sky and ground temperatures (Fuentes, 1985). The model is derived from an energy balance solution, treating the module as a uniform solid body that absorbs irradiance and dissipates heat through convection to ambient air and radiation to both sky and ground.

Thermal inertia is especially relevant during fast irradiance transitions – such as passing cloud edges – where the irradiance will rise faster than the module temperature. In these scenarios, cell temperature lags behind irradiance, potentially leading to worst-case voltage scenarios..

The temperature time series is computed using an adaptation of the *temperature.fuentes* function from the pvlib library. Required parameters of the Fuentes model can be obtained from the engineering design, module's specification and relevant literature.

2.4. V_{oc} calculation

The open-circuit voltage (V_{oc}) calculation is central to the DDSS methodology, as string sizing must ensure that the sum of module voltages under extreme conditions does not exceed system's maximum voltage.

V_{oc} is computed using a custom python routine adapted from the Sandia Array Performance Model (SAPM)

$$V_{oc} = V_{oc,STC} + N_{cs}\frac{n_d k(T_{cell}+273.15)}{q}\ln\left(\frac{G_{eff}}{G_{STC}}\right) + \beta_{Voc}(T_{cell} - T_{cell,STC}) \quad \text{(eq. 3)}$$

where $V_{oc,STC}$ is the module's open-circuit voltage at Standard Test Conditions (STC), N_{cs} is the number of cells in series in the module, n_d is the diode ideality factor, k is the Boltzmann constant, q is the electron charge, G_{STC} = 1000 W/m², β_{Voc} is temperature coefficient of V_{oc} and $T_{cell,STC}$ = 25 °C. This formulation accurately reproduces the dependence of V_{oc} on both temperature and irradiance (King et al, 2004). Besides the cell temperature obtained in the previous step, the inputs include module-specific parameters, available in manufacturer's datasheet or PVsyst's .pan file.

2.5. Safety factors implementation

The V_{oc} obtained in the previous step is influenced by several sources of uncertainty, both intrinsic to the modeling process and external to data acquisition. Given the long operational lifetime of utility-scale PV systems (often exceeding 25 years), it is essential to incorporate safety factors that account for variability and uncertainty. The amount and quality of available data can significantly reduce uncertainty.

From the V_{oc} timeseries, the engineer may select a non-exceedance threshold (e.g., P100, P99.9, P99.5, or P99), depending on the risk tolerance. Choosing a higher threshold (e.g., P99.9) results in more conservative designs (fewer modules in series), while lower thresholds (e.g., P99) can lead to slightly higher risk but better use of system voltage potential. In most practical applications, a P99.5 threshold ($V_{oc,p99.5}$) is recommended due to the limited open-circuit operation. High downtime (such as heavy curtailment scenarios) can shift the operating voltage towards V_{oc} and should be further assessed.

Uncertainty in V_{oc} estimation arises from multiple sources and must be accounted for in string sizing:

- **System Specification**. Flash test results from over 600,000 TOPCON provided by Kroma Energia showed that 90% of the modules exhibit up to +1.0% higher V_{oc} than specified in datasheet. while β_{Voc} values generally match manufacturer specs. A +1.0% uncertainty margin is recommended for specification variance.
- **Meteorological**. Long-term GHI and temperature variability contribute to V_{oc} uncertainty; wind speed variation across large PV sites adds ~0.6% uncertainty (Karin and Jain, 2020). A total of 1.5 – 2.5% combined meteorological uncertainty is recommended depending on duration of measurements and data availability.
- **Instrumentation Accuracy**. Class A and B pyranometers introduce ±2% and ±5% measurements uncertainty with rigorous quality control, respectively (Kipp & Zonen, 2025). These irradiance uncertainties can translate into 0.2% and 0.5% V_{oc} uncertainty under high irradiance.
- **Temporal Resolution**. Lower-resolution datasets miss short-term irradiance peaks (Braga et al, 2020; Nascimento et al, 2020), reducing maximum recorded GHI and underestimating V_{oc}; preliminary assessment by the authors recommend the following safety margins for different time-resolution of measurements: +0.3% (1-min), +0.8% (10-min), and +1.2% (60-min).
- **Modeling Error**. Errors from irradiance transposition, temperature modeling, and electrical conversion contribute an estimated +1.0% uncertainty, based on validation results presented in the following chapter.

An aggregated safety factor (SF) is applied to account for meteorological, instrumentation, temporal-resolution of measurements and modelling uncertainties. Total SF will typically range from 2.5% to 6.0%. This safety factor is applied to the selected non-exceedance value of open-circuit voltage to obtain the module's maximum open-circuit voltage ($V_{oc,max}$):

$$V_{oc,max} = V_{oc,p99.5}\,(1 + SF) \qquad \text{(eq. 4)}$$

The final maximum number of modules (N_s) in series is then

$$N_s = \left\lfloor {}^{V_{max}}\!\big/_{V_{oc,max}} \right\rfloor \qquad \text{(eq. 5)}$$

where V_{max} is the system's rated maximum DC voltage.

3. Results and discussion

3.1. Method validation

The DDSS validation used high-quality measurements from the Fotovoltaica/UFSC research laboratory in Florianópolis (27.4°S, 48.4°W). The instrumented testbed comprises a single-row single-axis tracker (±55° with backtracking) with 24 bifacial modules; one module (CSI's CS7N-645MB-AG, c-Si, monoPERC, bifacial module) was monitored in detail. Class-A meteorological and electrical sensors logged plane-of-array irradiance (front and back), ambient temperature, wind speed, back-of-module temperature, and V_{oc}. I–V traces were captured every minute and meteorological signals were synchronized.

Measurements covered 2023-10-01 to 2024-09-30 and were split into a 5-month calibration period (2023-10-01 to 2024-02-29) and a 7-month testing period (2024-03-01 to 2024-09-30). Rigorous quality control removed low-irradiance noise (GHI < 10 W/m²), zero and repeated sensor records, and days with >10% bad data; 22 calibration-period days and 54 test-period days were excluded. Data were resampled from 1-min to 10-min to reduce processing burden and to match common measurement campaigns resolutions.

Five thermal models were compared: Faiman (2008), PVsyst (2025), SAPM (King et al, 2004), Fuentes (1985), and Prilliman (Prilliman et al, 2020). Only Fuentes and Prilliman models consider thermal inertia, while the others assume steady-state operation. Where necessary, cell temperature (T_{cell}) to module temperautre (T_{mod}) conversions were applied, since some of the models reproduce T_{mod} and other T_{cell}. Thermal model parameters were optimized using SciPy's *minimize* function (Virtanen et al, 2020) to reduce the Root Mean Square Error (RMSE) over the calibration period.

As expected, calibrated models outperformed defaults. The Fuentes model gave the best calibrated performance (RMSE = 1.82 °C, MAE = 1.45 °C, MBE = -0.51 °C, R^2 = 0.98) and captured thermal inertia during rapid irradiance ramps. Prilliman, which also incorporates thermal inertia, also performed well. Steady-state models tracked daily profiles closely but tended to misrepresent fast transient ramps, as can be seen in Figure 3. Table 1 summarizes the calibrated models results.

Table 1: Evaluation metrics for each thermal model utilizing optimized parameters.

Model	**MAE** *[°C]*	**MBE** *[°C]*	**RMSE** *[°C]*	**R²**	**Max.** *[°C]*	**Min.** *[°C]*	**Mean** *[°C]*
Measured	-	-	-	-	69.26	7.81	40.62
Faiman	1.76	0.74	2.41	0.96	67.54	11.69	40.84
PVsyst	1.76	0.74	2.41	0.96	64.03	11.23	39.08
SAPM	1.83	0.01	2.56	0.96	63.61	11.07	39.11
Fuentes	1.45	-0.51	1.82	0.98	67.55	6.34	40.88
Prilliman	1.72	0.00	2.37	0.97	61.52	6.78	38.87

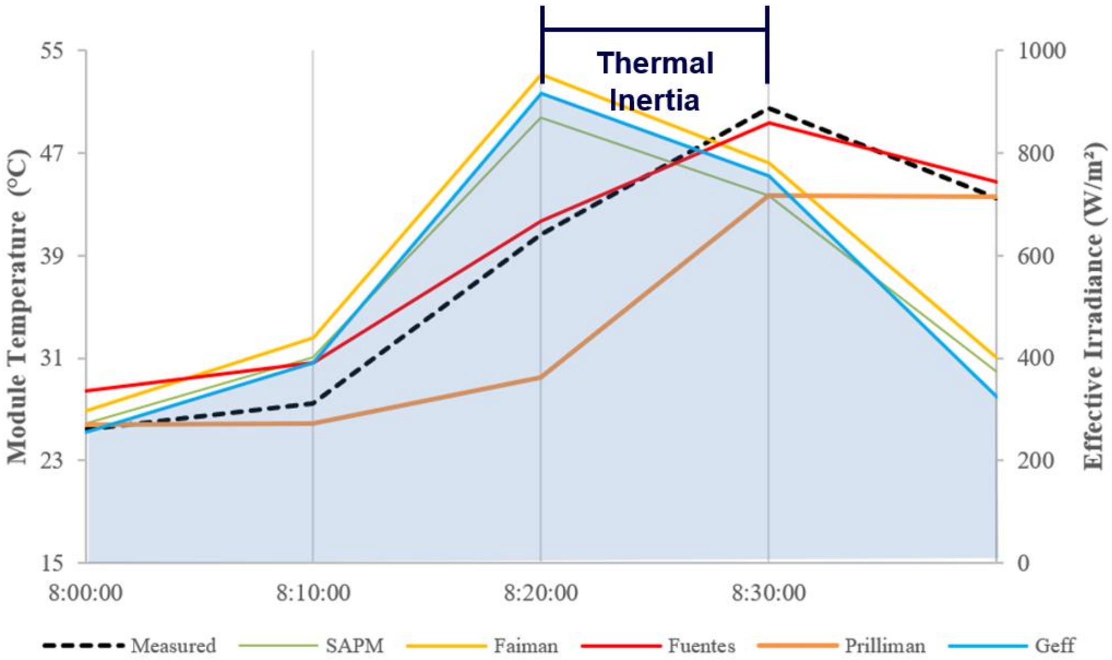


Figure 3: Zoomed-in view of a 50-minute interval with rapid variation in effective irradiance (G_{eff}, blue shaded curve). Only displaying four models (solid lines named according to the models) and measured temperature (black dashed line) to ease visualization.

V_{oc} was computed with the SAPM electrical model using each modelled T_{cell} (or converted T_{mod}) and the measured back-of module measurements. Metrics over operating conditions (GHI > 100 W/m²) show small errors: MAE values ranging from 0.3 – 0.5 V and RMSE from 0.8 – 0.9 V. Fuentes (and Prilliman) yielded the best percentile alignment (P99.9, P99.5, P99) when paired with calibrated parameters. The clear improvement captured from SAPM to Prilliman (which uses SAPM's T_{cell} as input) highlights the importance of thermal inertia when assessing maximum V_{oc}. Most models tend to underestimate V_{oc} (implied from negative MBE values), which could be explained by T_{cell} overestimation. Assuming $T_{cell} = T_{mod}$ reduced bias and improved percentile agreement for the tested double-glass modules, indicating measured module temperature is an acceptable proxy in this context in allignment with recent literature (Oliveira et al, 2025). Table 2 summarizes the calculated metrics.

Table 2: Evaluation metrics for V_{oc} estimation, utilizing each thermal model with optimized parameters when $T_{mod} = T_{cell}$. Red

and blue colored cells indicate under and overestimation of V_{oc}, respectively.

Model	MAE [V]	MBE [V]	RMSE [V]	Max [V]	P99.9 [V]	P99.5 [V]	P99 [V]
Measured	-	-	-	44.68	44.54	44.10	43.70
Estimated (from Temp. Measurement)	0.23	-0.13	0.78	44.78	44.61	44.16	43.68
Faiman	0.31	0.045	0.83	44.13	43.67	43.25	43.05
PVsyst	0.32	-0.16	0.83	44.24	43.68	43.26	43.05
SAPM	0.32	-0.16	0.83	44.35	43.62	43.18	42.86
Fuentes	0.27	-0.13	0.79	44.48	44.20	43.58	43.27
Prilliman	0.31	-0.13	0.81	44.81	44.51	44.02	43.61

SAPM electrical model coupled with a thermal-inertia model (particularly if calibration can be performed) provide accurate V_{oc} time series for DDSS application in utility-scale contexts. Calibration materially improves thermal estimates and high-percentile voltage alignment; however, extending validation to longer, higher-resolution, and multi-site datasets is necessary to safely reduce conservatism and generalize the method.

3.2. Case study

This subsection applies the DDSS workflow to two ~100 MWp utility-scale projects with contrasting climates and system architectures to illustrate technical and economic impacts. Case 1 (Jaguaruana, Ceará, Brazil) is a high-irradiance (2110 kWh/m² per year), semi-arid site, with single-axis-tracking (GCR = 36.6%) and central inverters; Case 2 (Florianópolis, Santa Catarina, Brazil) is a lower-irradiance (1615 kWh/m² per year), subtropical site, with fixed-tilt (GCR = 59.6%) and string inverters. For comparability both use the same c-Si, PERC, bifacial module (CSI's CS7N-645MB-AG, 645 Wp) and only one year of on-site measurements. Even though 10+ years of measurements are recommended (and likely to be available for utility-scale projects), 1 year of measurements is considered to assess the impact of DDSS methodology even under severe safety factors. Table 3 summarizes the main characteristics of each case.

Table 3: Key parameters of each case study.

	Case 1 (JAG)	Case 2 (FLN)
Location	Jaguaruana, Ceará, Brazil (-4.946°, -37.728°)	Florianópolis, Santa Catarina, Brazil (-27.431°, -48.441°)
Average GHI (kWh/m²)	2110	1615
Average T_{amb} (°C)	26.9	20.9
Installed Capacity	~ 100 MWp / 80 MWac (DC/AC ratio: ~1.25)	~ 100 MWp / 80 MWac (DC/AC ratio: ~1.25)
PV Module	CSI/Canadian Solar CS7N-645MB-AG	CSI/Canadian Solar CS7N-645MB-AG
Inverter topology	Central	String
Racking	Single-axis tracker	Fixed tilt
GCR	36.6%	59.6%

Three sizing approaches are compared:

- Traditional ($V_{oc,STC}$ corrected to annual minimum ambient temperature).

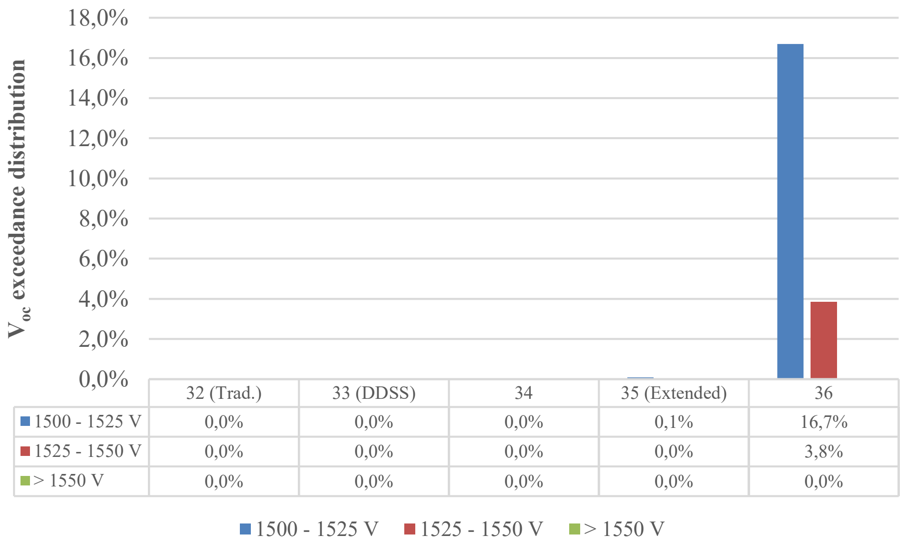
- DDSS (P99.5 Voc from the full year of site measurements, inflated by a high SF of 6.0% to produce $V_{oc,max}$).
- Extended DDSS (upper limit of DDSS, considering P99.5 Voc, inflated by a SF of 2.5% to produce $V_{oc,max}$ and considering that the system can handle 1525 Vdc).

In Case 1 (JAG), the conventional method resulted in 32 modules per string, whereas the DDSS and Extended DDSS allowed 33 and 35 modules in series, respectively. In Case 2 (FLN), the traditional limited string sizeto 31 modules and it was increased to 32 and 34 on DDSS and Extended DDSS, respectively. Table 4 summarizes the string sizing results.

Table 4: String sizing summary for both cases using traditional, DDSS, and Extended DDSS methods.

Method	Case 1 (JAG)	Case 2 (FLN)
Traditional	32	31
DDSS	33 (+3.12%)	32 (+3.23%)
Extended DDSS	35 (+9.38%)	34 (+9.68%)

To further evaluate evaluate voltage exceedance, the V_{oc} timeseries obtained from measurements was multiplied by the number of modules in series to simulate system-level voltages. Exceendances are divided into three intervals, according to voltage tolerances: i) 1500–1525 V, still within the tolerance margin of most manufacturers; ii) 1525–1550 V, caution as operation approaches typical safety limits; and iii) values above 1550 V, regarded as severe. Figure 4 and Figure 5 present the exceedance distributions in each cases.



	32 (Trad.)	33 (DDSS)	34	35 (Extended)	36
1500 - 1525 V	0,0%	0,0%	0,0%	0,1%	16,7%
1525 - 1550 V	0,0%	0,0%	0,0%	0,0%	3,8%
> 1550 V	0,0%	0,0%	0,0%	0,0%	0,0%

Figure 4: V_{oc} exceedance for various string configurations for Case 1 (JAG), based on one year of measured data.

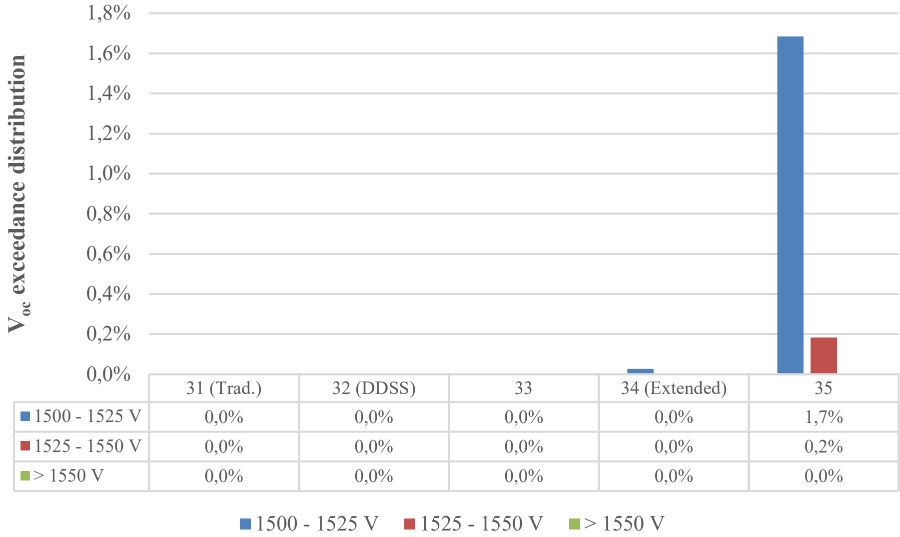


Figure 5: V_{oc} exceedance for various string configurations for Case 2 (FLN), based on one year of measured data.

For Case 1, although DDSS recommends 33 modules, voltages exceed 1500 V only with 35 modules (≈0.1% probability in the one-year dataset) and reach 1550 V only at 36 modules; for Case 2 the DDSS limit is 32 modules, but the 1500 V threshold is likewise exceeded only from 35 modules upward and 1550 V is reached only at 36 modules. These outcomes indicate that DDSS, together with the applied safety factor, is conservative and maintains compliance even for rare high-voltage events. However, it should be taken into consideration that only one year of measurements (with 10-minute logging interval) were assessed – thus the high SF –, meaning that those results do not justify increasing string sizes beyond the calculated limits.

Am energy yield assessment was carried out through simulations in PVsyst v8.0.14. The increase in string length (N_s) derived from DDSS did not produce any measurable improvement in energy yield and Extended scenario introduced minor improvement. The results confirm that longer strings slightly reduce DC wiring losses (which works best for central inverters since those systems have longer DC cabling) while maintaining inverter efficiency nearly constant, leading to a combined loss reduction of up to 0.24% in Case 1 (JAG) and 0.13% in Case 2 (FLN). These performance improvements translate directly to Performance Ratio (PR) improvement. Table 5 summarizes the DC ohmic and inverter efficiency losses.

Table 5: DC ohmic losses, inverter efficiency losses, and their combined sum for different Ns configurations in Cases 1 and 2.

Case 1 (JAG)	N_s	31	32 (Traditional)	33 (DDSS)	34	35 (Extended)
	DC Ohmic Losses	1.51%	1.43%	1.35%	1.27%	1.20%
	Inverter Loss (Effic.)	1.55%	1.57%	1.59%	1.61%	1.62%
	Combined:	**3.04%**	**2.98%**	**2.92%**	**2.86%**	**2.80%**
Case 2 (FLN)	N_s	30	31 (Traditional)	32 (DDSS)	33	34 (Extended)
	DC Ohmic Losses	0.65%	0.60%	0.56%	0.54%	0.51%
	Inverter Loss (Effic.)	0.99%	1.00%	1.01%	0.99%	1.00%
	Combined:	**1.63%**	**1.59%**	**1.56%**	**1.52%**	**1.50%**

Economically, longer string configuration translates into fewer strings, reducing several cost items. According to unpublished results obtained from Kroma Energia:

- **Mounting structures**. More modules per string reduces linearly the total number of structures required.

- **DC electrical balance of system**. A 5% increase in string length typically results in a 6% reduction in DC electrical balance of system costs, mainly due to decreased cabling and combiner box requirements.
- **Labor for mounting structure assembly and pile driving**. Labor and materials are expected to decrease proportionally with fewer structures and less civil work needs.

These benefits should be verified case-by-case, since longer strings can introduce additional gains but also trade-offs such as the need for stronger structural components, additional earthworks, or layout changes that may offset some savings.

Considering a typical capital expentidure scenario for utility-scale projects in Brazil, for Case 1 (JAG), the DDSS design with one additional module per string led to CAPEX savings of 0.84%, a LCOE reduction of 0.10 USD/MWh. The Extended scenario with two additional modules produced CAPEX savings of up to 1.96% and a LCOE reduction of 0.20 USD/MWh. To provide a technical perspective, these improvements are equivalent to an effective increase in module efficiency from 20.84% to 21.00% (DDSS scenario) and 21.17% (Extended scenario). For Case 2 (FLN), the DDSS methodology led to CAPEX savings of 0.42%, a LCOE reduction of 0.08 USD/MWh, while the Extended scenario produced CAPEX savings of up to 0.95% and a LCOE reduction of 0.19 USD/MWh. Since Case 2 (FLN) utilizes stirng inverter and fixed-tilt racking architecture, the economic benefits of increasing string size are less pronouced.

The results demonstrate that the DDSS method effectively balances safety and optimization. By relying on measured data and physical modeling rather than conservative assumptions, it provides a realistic assessment of the maximum achievable voltage and thus allows project engineers to safely exploit the system's voltage range. The method is especially relevant for large-scale power plants, where marginal cost reductions per megawatt can translate into hundreds of thousands of dollars in savings.

Nevertheless, the method requires the availability of high-resolution on-site meteorological measurements and can benefit from precise calibration of thermal. It is not recommended for distributed generation systems, where such data are usually unavailable. Future improvements could include integrating machine-learning-based estimators for temperature and irradiance using satellite data, as well as extending the DDSS framework to upcoming 2000 Vdc system architectures specified in IEC TS 63543-1.

4. Final remarks

This study proposed and evaluated a DDSS methodology that determines the maximum safe number of PV modules in series (N_s) by combining sub-hourly site meteorology (GHI, ambient temperature, wind speed), module/inverter/racking specifications, dynamic cell-temperature modeling (including thermal-inertia effects), and a probabilistic analysis of uncertainty. The approach generates a V_{oc} time series, selects a design non-exceedance percentile, applies an aggregated safety factor, and derives N_s from the system maximum DC voltage. Implementation uses Python with pvlib and is validated against measurements.

Thermal-model validation demonstrated that models incorporating thermal inertia deliver superior transient fidelity; among the models tested the Fuentes formulation provided the best match to measured module temperatures after calibration and is therefore recommended for DDSS when high-resolution local data are available. Differences in temperature model choice translated into relatively small absolute V_{oc} differences (generally <1.0 V), but inertia-aware models reduced short-term bias during rapid irradiance ramps and thus yield more robust high-percentile voltage estimates critical for string sizing safety margins.

Application to two ~100 MWp utility-scale cases with contrasting climates, different architectures and only one year of measurements, showed that DDSS safely increased allowable string length by one module in each case using a P99.5 design threshold and an aggregated SF of 5.4%. Energy-yield simulations found negligible impacts on performance metrics (PR differences of 0.1% or less) because reductions in DC ohmic losses were offset by marginal inverter-efficiency effects at higher DC voltages.

Economically, modest but measurable CAPEX savings accrue from longer strings due to fewer racking rows, reduced DC EBOS and lower labor requirements. In the reported DDSS scenarios CAPEX fell by 0.4 – 0.8%, producing small improvements in LCOE. Extended, exploratory N_s scenarios (larger increases in string length) amplified these savings (CAPEX reductions up to 1.9%), indicating larger projects or markets with high labor/racking costs could profit more substantially from DDSS.

Key limitations temper these conclusions: the analysis relied primarily on one year of 10-min on-site data and validation on a single instrumented module, which necessitated conservative safety factors; irradiance transposition, long-term variability and broader device/geometry validation remain to be completed. Future work should prioritize longer and higher-resolution measurement records (including 1 – 10 s), expanded model validation across climates and technologies, refined uncertainty decomposition to reduce conservatism, and exploration of advanced electrical models and tracking algorithms to further increase confidence and potential gains from DDSS.

5. References

Anderson, K.S., Hansen, C.W., Holmgren, W.F., Jensen, A.R., Mikofski, M.A., Driesse, A., 2023. pvlib python: 2023 project update. JOSS 8(92), 5994. https://doi.org/10.21105/joss.05994

Braga, M., et al., 2020. Solar Over-Irradiance Events: Preliminary Results from a Global Study. In: 47th IEEE Photovoltaic Specialists Conference (PVSC), Calgary, Canada, 15 Jun–21 Aug 2020, pp. 2764–2770.

David, J., Ciufo, P., Elphick, S., Robinson, D., 2022. Preliminary evaluation of the impact of sustained overvoltage on low voltage electronics-based equipment. Energies 15(4), 1536. https://doi.org/10.3390/en15041536

Dhandale, K.K., 2021. A review of design, manufacturing of grid tied PV inverter and its impact on site performance, reliability and safety. International Research Journal of Engineering and Technology (IRJET) 08.

Erbs, D.G., Klein, S.A., Duffie, J.A., 1982. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. Solar Energy 28(4), 293–302. https://doi.org/10.1016/0038-092X(82)90302-4

Faiman, D., 2008. Assessing the outdoor operating temperature of photovoltaic modules. Progress in Photovoltaics 16(4), 307–315. https://doi.org/10.1002/pip.813

Fuentes, M.K., 1985. A simplified thermal model for flat-plate photovoltaic arrays. SAND85-0330.

Huawei Digital Power, Webinar: A technical guide for string configuration of Fusion Solar solutions. [Online]. Available: https://www.youtube.com/watch?v=CUCu4T5urao&t=889s (accessed: Sep. 14, 2025).

International Electrotechnical Commission (IEC), 2023. IEC 62548-1: Photovoltaic (PV) arrays – Part 1: Design requirements.

Jiménez, A.N., Bkayrat, R., 2015. Utility scale 1,500 VDC PV power plant architecture evolution: advantages and challenges. Integration of Renewable Energy into High and Medium Voltage Systems Conference & Exhibition, 2015.

Karin, T., Jain, A., 2020. Photovoltaic string sizing using site-specific modeling. IEEE J. Photovoltaics 10(3), 888–897. https://doi.org/10.1109/JPHOTOV.2020.2969788

King, D., Boyson, W.E., Kratochvil, J.A., 2004. Photovoltaic array performance model, 2004.

Kipp & Zonen, 2025. Pyranometer Site Confidence Check V3.0. Kipp & Zonen Documentation. [Online]. Available: https://www.kippzonen.com/Download/1081/Pyranometer-Site-Confidence-Check-V3 (accessed: Aug. 6, 2025)

Nascimento, L.R. do, Braga, M., Campos, R.A., Naspolini, H.F., Rüther, R., 2020. Performance assessment of solar photovoltaic technologies under different climatic conditions in Brazil. Renewable Energy 146, 1070–1082. https://doi.org/10.1016/j.renene.2019.06.160

Nascimento, L.R. do, Viana, T.S., Campos, R.A., Rüther, R., 2019. Extreme solar overirradiance events: Occurrence and impacts on utility-scale photovoltaic power plants in Brazil. Solar Energy 186, 370–381.

https://doi.org/10.1016/j.solener.2019.05.008

Oliveira, A., Braga, M., Naspolini, H. F., & Rüther, R., 2025. Validation of thermal models for bifacial photovoltaic systems under various albedo conditions. Progress in Photovoltaics. https://doi.org/10.1002/pip.3892

Prilliman, J., Stein, J.S., Riley, D., Tamizhmani, G., 2020. Transient Weighted Moving-Average Model of Photovoltaic Module Back-Surface Temperature. IEEE Journal of Photovoltaics 10(4), 1053–1060. https://doi.org/10.1109/JPHOTOV.2020.2992351

PVsyst, 2025. Array Thermal losses. PVsyst Documentation. https://www.pvsyst.com/help/project-design/array-and-system-losses/array-thermal-losses/index.html?h=thermal (accessed: Aug. 3, 2025)

REN21, 2025. Renewables 2025 Global Status Report. ISBN: 978-3-948393-19-9. [Online]. Available: https://www.ren21.net/gsr-2025/global_overview/. Last accessed: 28-11-2025.

Reda, I., Andreas, A., 2004. Solar position algorithm for solar radiation applications. Solar Energy 76(5), 577–589. https://doi.org/10.1016/j.solener.2003.12.003

Sanchís-Gómez, C., Aleix-Moreno, J., Vargas-Salgado, C., Alfonso-Solar, D., 2025. The novel evaluation method for PV module temperature and string size risk in utility-scale solar projects. Solar Energy 295, 113520. https://doi.org/10.1016/j.solener.2025.113520

Skoplaki, E., Palyvos, J.A., 2009. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. Solar Energy 83(5), 614–624. https://doi.org/10.1016/j.solener.2008.10.008

Sungrow Power Supply, 2024. 2023 Sustainability Report. [Online]. Available: https://info-support.sungrowpower.com/application/pdf/2024/05/21/Sungrow%202023%20Sustainability%20Report.pdf (accessed: Oct. 28, 2025).

Virtanen, P., et al., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2