Estimation of Photovoltaic Array Production Mismatch in Field Conditions

E. Gozani, Y. Massuri, S. Sabag, R. Shirazi, A. Shaton: SolarEdge Technologies, Inc.

Abstract

Differences in illumination and electrical conditions between photovoltaic panels in an array lead to production mismatch losses. An estimation of this loss for 13 different photovoltaic sites is presented here. Based on measurements from each module taken at field conditions, individual I-V curves per module are obtained, and the combined curves are used for an assesments of the mismatch present in the system. Aggregated mismatch values are calculated for each site for a selected set of days during the first year of production, and a combined mismatch value of $\mu_{loss} = 5.4\% \pm 1.9\%$ is observed.

Introduction

The rapidly increasing number of solar sites operating in field conditions worldwide allows us to reconsider long-standing industry assumptions regarding photovoltaic array mismatch power losses. These account for the discrepancy between the sum of the maximum potential power of each individual panel in the array and the actual total array output, and arise when modules with non-identical peak operating conditions are connected together. Numerous effects contribute to the observed mismatch: non-uniformities in manufacturing, light-induced degradation, module aging, soiling, and module temperature; site-specific impacts such as partial shading or non-identical orientations; and pathologies such as potential-induced degradation and burnt module diodes. These individual impacts have been described [1] and studied previously [2][3], with some observing impactful reductions [4] in the output power potential. Meanwhile, other studies found the total impact of mismatch losses to be minor [5][6][7].

Several companies in the photovoltaic industry produce hardware solutions that mitigate mismatch losses. Leading this innovative approach, SolarEdge Technologies, Inc. designed DC-DC converters called optimizers that allow modules to retain their maximal power operating conditions, achieving string and array uniformity along the optimizer outputs. In addition, these devices continuously report the working conditions of each module, allowing for site monitoring at an individual panel level. Data collected from these field devices provides an immense opportunity to study the behavior of solar arrays in field conditions.

This paper aims to utilize real-time optimizer data from several SolarEdge photovoltaic sites to estimate the combined mismatch effect present, and alleviated, in the corresponding arrays. For this initial study, we chose to focus on single-string arrays. A subset of such sites was carefully chosen to exclude the impact of avoidable mismatch causes, partial shading and non-identical orientations, while controlling the impacts of aging by using only first-year data in the estimation.

Estimation Technique

The one-diode model is widely used to describe the electrical properties and response of crystalline silicon photovoltaic modules. Assuming that the cells in the module behave identically, it uses an equivalent electrical circuit comprised of a current source, a single Shockley diode, and two resistors, illustrated in Fig 1. It relies on 6 parameters: the light-induced current I_L, reverse saturation current in the diode I_0, series resistance R_s, shunt resistance R_{sh}, the cell ideality factor γ, and the number of cells connected in series in the module, N_s.

The characteristic relation between module voltage and current is then described by the following equation:

$$I = I_L - I_0 \left[e^{\frac{V + I \cdot R_s}{N_s \cdot V_{th}}} - 1 \right] - \frac{V + I \cdot R_s}{R_{sh}} \tag{1}$$

$$V_{th} := \frac{\gamma k_B T}{q}$$

Where q is the elementary charge, k_B is Boltzmann's constant, and T is the module's temperature, assumed here to be $T = 25^0 C$. Of the six parameters in the model, four are provided by the module's manufacturer following

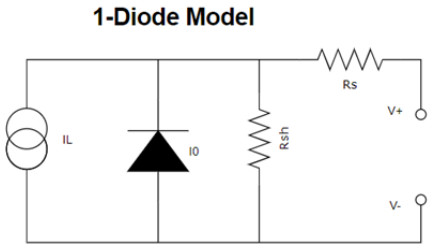

Figure 1: Equivalent electrical circuit representing a crystalline silicon solar module in the one-diode model

laboratory testing under controlled conditions - R_s, R_{sh}, N_s, and γ. As these parameters were only given in STC conditions, they are taken as constants throughout the estimation process. The remaining parameters, I_L and I_0, vary strongly with illumination conditions and thus need to be estimated simultaneously in order to obtain a valid module description. To do so, the optimizer data at each point in time, (V_k, I_k) is used, where k is a temporal index. In standard working conditions, when the solar array is not overproducing power with respect to inverter size or grid limitations, each optimizer ensures that its correponding module operates at maximal power (MP). Thus, the current and voltage measurements are taken to represent the MP conditions of the module - $(V_k, I_k) \sim (V_{MP,k}, I_{MP,k})$. This assumption, used in conjunction with Eq. 1 yields two equations that lead to a unique solution for I_L, I_0 at each point in time:

$$\begin{cases} 1) & I\left(V_{MP}\right) = I_{MP} \\ 2) & \frac{dP}{dV}|_{V=V_{MP}} = 0 \end{cases} \Longrightarrow \begin{cases} I_0 = \frac{I_{MP}\left(1+\frac{R_s}{R_{sh}}\right) - \frac{V_{MP}}{R_{sh}}}{1-\frac{I_{MP}R_s}{V_{MP}}} \frac{N_s V_{th}}{V_{MP}} e^{-\frac{V_{MP}+I_{MP}\cdot R_s}{N_s \cdot V_{th}}} \\ I_L = I_{MP} \cdot \left(1+\frac{R_s}{R_{sh}}\right) + \frac{V_{MP}}{R_{sh}} + I_0 \left(e^{\frac{V_{MP}+I_{MP}R_s}{N_s V_{th}}} - 1\right) \end{cases} \tag{2}$$

Once those parameters are set, an analytic solution to the implicit Eq. 1 is employed, using Lambert W functions, to obtain the full I-V curve of each module at the time of measurement:

$$\begin{aligned} I\left(V\right) &= \frac{\left(I_L + I_0\right) R_{sh}}{R_s + R_{sh}} - \frac{V}{R_s + R_{sh}} \\ &\quad - \frac{N_s V_{th}}{R_s} \cdot W\left(\frac{I_0 R_s R_{sh}}{N_s V_{th}\left(R_s + R_{sh}\right)} \cdot e^{\frac{R_{sh}\cdot\left[V + R_s\left(I_L + I_0\right)\right]}{N_s V_{th}\left(R_s + R_{sh}\right)}}\right) \\ V\left(I\right) &= \left(I_L + I_0\right) R_{sh} - I\left(R_s + R_{sh}\right) \qquad (3.a) \\ &\quad - N_s V_{th} \cdot W\left(\frac{I_0 R_{sh}}{N_s V_{th}} \cdot e^{\frac{\left(I_L + I_0 - I\right) R_{sh}}{N_s V_{th}}}\right) \qquad (3.b) \end{aligned}$$

As the process we just described can be performed for each module in the array simultaneously, the combined I-V curve of the each string can be deduced. In cases where the combined current pushes one of the modules into reverse voltage, ideal bypass diodes employed, assumed to work without dissipation. Focusing on single-string arrays, the mismatch estimation procedure then works as follows, for each time index k:

- Use condition 2 along with equation 3.a to obtain individual module I-V curves in the string
- Combine the I-V curves in series assuming ideal bypass diodes
- Obtain combined P-I curve using equation 3.b
- Find the maximal string power by scanning the combined P-I curve, P_{MP}^{string}
- The mismatch is calculated as the relative difference between the integrated maximal string power and the integrated sum of optimizer input powers, over the selected time period:

$$\mu_{loss} = \frac{\int\left(\sum_{i\in\{optimizers\}} P_i^{in} - P_{MP}^{string}\right) \mathrm{d}t}{\int\sum_{i\in\{optimizers\}} P_i^{in} \,\mathrm{d}t}$$

Figure 2 shows a small-scale example of how this technique works in practice. Focusing on a single day, 2018-06-02, and on a single site, the third site in table 1, it contrasts the estimated string power with the sum of optimizer input powers. It also presents the measured current variation among optimizers, which was found to be a key indicator for mismatch. Overall, the simulation was found to arrive at the maximal observed power when the current mismatch is small. Individual I-V curves and string P-I/V-I curves obtained by the process are also shown in figure 3 for two selected timestamps. Here the specific MPP tracking can be observed, as well as the behavior of the ideal bypass diodes. It is important to note here that we assume that the current and voltage measurements provided by the optimizers convey the maximal power point of the module. This assumption, which is key to the entire estimation technique, generally holds as long as the power output of the inverter connected to the array doesn't exceed its limit. Those cases of overproduction, caused either by hardware limitations of the inverter or by external grid circumstances, are also monitored by the SolarEdge system. For those particular data points, a conservative mismatch value of 0 was taken.

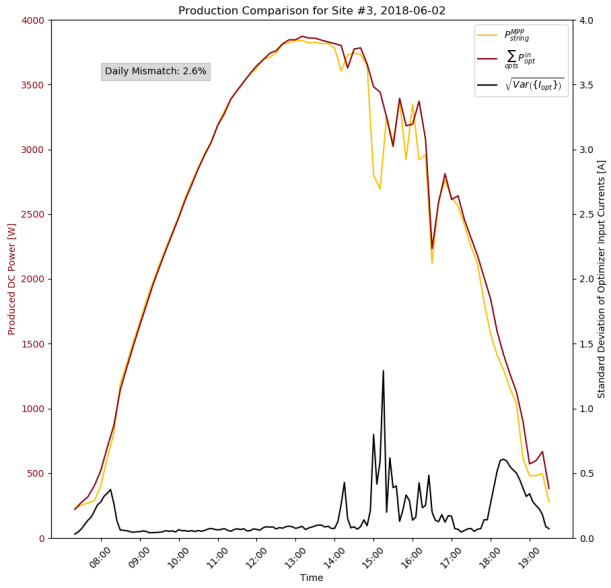

Figure 2: A comparison between the simulated string power (yellow) and the sum of input optimizer powers (red) for a single day (2018-06-02) for the third site on table 1. The difference between the integrals of the two curves sets the daily mismatch, found in this case to be 2.8%. In addition, the standard deviation between optimizer currents throughout the day is shown (black). Current mismatch between the optimizers was found to be a prime indicator of power mismatch.

Sample selection and aggregation

The procedure outlined in the previous section can, in principle, be performed generally on any SolarEdge photovoltaic site. The data collected from sites for monitoring purposes contains current and voltage measurements, allowing us to non-intrusively probe the individual and combined I-V curves in the site at any given point in time. However, as we are interested in a sober estimate of the nominal mismatch loss, additional criteria were applied to sites used in this study to exclude more preventable causes of mismatch. Sites chosen for this study use single-string arrays of identical optimizers, each connected to a single solar panel. In each site, the panels are all set at the same tilt and azimuth, and are of the same manufacturer, age and model. The ratio of total array power and maximal inverter output power was also limited to avoid overproduction as much as possible. In addition, we made sure no artifacts such as trees or chimneys cause partial shading or other transient effects in the sites chosen. A variety of string lengths and geographic locations was encouraged, and sites with strings longer than 25 panels were excluded to avoid mismatch that could be prevented by string rearrangements. Finally, in order to limit the effects of aging mismatch, only the first year of production in the site was used for the analysis, excluding the first two months. Future work may study the onset of aging mismatch by expanding this time frame. Table 1 details the locations, array compositions, and panel model details for the sites selected for the current work.

index	location	# Panels	(tilt, azimuth)	panel STC power	array power	inverter size	inst. date
1	USA (CA)	18	(19, 90)	290W	5220W	5000W	2017/09
2	USA (MA)	11	(30, 235)	260W	2860W	3000W	2017/08
3	Italy	14	(22, 198)	333W	4660W	4000W	2017/11
4	Netherlands	20	(50, 197)	280W	5600W	5000W	2018/01
5	USA (CA)	23	(23, 242)	300W	6900W	7600W	2017/07
6	USA (NY)	14	(27, 265)	290W	4060W	3800W	2017/11
7	USA (CA)	10	(18, 179)	290W	2900W	3000W	2017/09
8	USA (AZ)	19	(23, 178)	285W	5415W	5000W	2017/10
9	Australia (VIC)	16	(22, 335)	325W	5200W	5000W	2017/07
10	USA (CA)	13	(23, 270)	285W	3705W	3800W	2017/07
11	Netherlands	18	(40, 138)	320W	5760W	7000W	2018/08
12	Netherlands	21	(35, 230)	295W	6200W	5000W	2018/04
13	Netherlands	16	(40, 190)	320W	5120W	5000W	2018/06

Table 1: Description of the sites used for the analysis

Since optimizers are set to work only when their input power exceeds a certain threshold, weak production during early morning and late evening makes it possible for optimizers to start and stop reporting at slightly different times. Thus, in order to ensure the validity of the estimate, only time indices for which data from every optimizer was present were used. Data points with only partial data are excluded and considered invalid. We require at least 95% of the daily data to be valid when calculating the daily mismatch. For each site, the daily mismatch is calculated for every valid day. In order to focus on the evaluation of mismatch during nominal production, only days in which the total production exceeded $3kWh/kWp$ were aggregated when calculating the site's mismatch. Thus, the nominal site mismatch loss is defined as the weighted (by daily energy) average of daily mismatch values for selected dates in the period.

Results

The simulation and analysis outlined in the previous sections were performed on every site in table 1. The mismatch was studied both at the daily level and at the aggregate level over the whole period. These two quantities are related, as the aggregated mismatch loss is defined as the energy-weighted average of daily mismatch values. Thus, both were examined simultaneously. Figure 4 shows the distribution of daily mismatch and daily production values for each site. It includes a line where the sample quality cut of $kWh/kWp > 3$ was made. It shows how weaker days in terms of production, while having only a minor contribution to the total mismatch loss, reach more erratic daily mismatch values. The main clusters of daily mismatch values for each site, seen at the lower end of the plot, indicate the variety of weather, physical, and geographical conditions between sites, as the clusters are located around different values of daily production. The final metric used to represent the mismatch loss associated with each site is the weighted average between the daily mismatch values, with the weights being the daily production values:

$$\hat{\mu}_{loss}^{tot} = \frac{\sum_{i=1}^{N_{days}} w_i \cdot \mu_{loss,i}}{\sum_{i=1}^{N_{days}} w_i} \mid w_i = \frac{E_i^{daily}}{P_{inverter}^{max}} \tag{4}$$

Beyond these global features, it is important to compare the difference in statistics between sites. Figure 5 shows normalized distribution of daily mismatch for high-quality days, split into sites by color. Four sites are presented there, for the sake of legibility. While the distribution of mismatch of different sites follows a similar shape, there are still considerable differences between them. Thus, any combination of results between sites must be done carefully.

The analysis we performed attracts two kinds of uncertainties: statistical uncertainties that mainly arise from variable field conditions between sites, and measurement uncertainties related to the optimizer voltage/current data. Since the nominal mismatch is calculated as the weighted average 4, assigning a standard error to it is ambiguous. We chose to calculate the error on the estimate based on the results of bootstrap analysis [8] that studied the empirical distribution of the estimator, finding the best-fitting formula to be:

$$\hat{\sigma}_{\hat{\mu}_{loss}^{tot}} = \sqrt{\frac{N_{days} \cdot \sum_{i=1}^{N_{days}} w_i^2 \left(\mu_{loss,i} - \hat{\mu}_{loss}^{tot}\right)^2}{(N_{days} - 1) \cdot \left(\sum_{i=1}^{N_{days}} w_i\right)^2}} \tag{5}$$

Evaluation of the measurement uncertainty required a more involved approach, as well as a model for inherent optimizer measurement errors. This model was developed in a separate study, where voltage and current readings of a solar array simulator power generator were compared between optimizers and external sensors. Errors on the current and voltage measurements were found to be $\sigma_I = 0.1A, \sigma_V = 0.3V$. The larger error on the voltage measurement is driven by the optimizer's continuous voltage scan around the MPP. These errors were incorporated into the analysis by using a Monte Carlo ensemble method. For each site, 101 different runs were performed - the first using the nominal optimizer data, and the remaining by randomly varying each data point according a gaussian error model with the aforementioned scales. These variations are then propagated along the same estimation technique described previously, each yielding different string powers. Their mean and standard deviation were used to create an envelope of width $\pm 1\sigma$ in possible string array powers, leading to up/down variations in the resulting mismatch values. The measurement uncertainty for the mismatch loss is then calculated as the weighted average of daily uncertainties for each site, throughout the selected period. Finally, the total combined effect is taken as the sum, in quadrature, of individual site errors.

Table 2 shows the aggregated mismatch values per site, along with the statistical errors, calculated using equations 4, 5. The string length and number of valid days per site are also noted. Differences between the sample sizes used by each site, driven mainly by geographic and weather variability and the data quality cuts acting more stringently on some sites, make it challenging to combine site results into a single figure. The simplest approach to this task, which was taken here, is to give each site an equal weight, making the average and standard deviation between sites the primary value for the estimated effect of mismatch in arrays at field conditions. Figure 6 shows the total aggregated mismatch values from each site and the combined mismatch loss value. The combined value of $\mu_{loss}^{combined} = 5.4\% \pm 1.9\%(stat.)^{+0.8\%}_{-0.9\%}(meas.)$ is found to consistently describe the site estimates.

Site Index	String Length	Sample Size [days]	Total Aggregated Mismatch
1	18	120	$4.7\% \pm 0.7\%(stat.)^{+0.2\%}_{-0.3\%}(meas.)$
2	11	102	$8.1\% \pm 1.2\%(stat.)^{+0.3\%}_{-0.3\%}(meas.)$
3	14	159	$4.7\% \pm 0.6\%(stat.)^{+0.3\%}_{-0.3\%}(meas.)$
4	20	163	$6.8\% \pm 0.9\%(stat.)^{+0.2\%}_{-0.3\%}(meas.)$
5	23	139	$4.0\% \pm 0.8\%(stat.)^{+0.2\%}_{-0.1\%}(meas.)$
6	14	147	$7.1\% \pm 0.9\%(stat.)^{+0.3\%}_{-0.3\%}(meas.)$
7	10	210	$3.4\% \pm 0.4\%(stat.)^{+0.2\%}_{-0.3\%}(meas.)$
8	19	243	$2.3\% \pm 0.4\%(stat.)^{+0.2\%}_{-0.1\%}(meas.)$
9	16	93	$7.0\% \pm 1.1\%(stat.)^{+0.2\%}_{-0.2\%}(meas.)$
10	13	141	$3.8\% \pm 0.6\%(stat.)^{+0.2\%}_{-0.3\%}(meas.)$
11	18	74	$4.3\% \pm 1.0\%(stat.)^{+0.2\%}_{-0.2\%}(meas.)$
12	21	110	$8.7\% \pm 1.3\%(stat.)^{+0.2\%}_{-0.3\%}(meas.)$
13	16	79	$5.7\% \pm 1.0\%(stat.)^{+0.3\%}_{-0.3\%}(meas.)$

Table 2: Results of the combined mismatch analysis from each site, along with the corresponding statistical errors

Discussion

Several other groups have studied the topic of mismatch loss in photovoltaic arrays in the past. Chamberlin et al. [6] performed an analysis of mismatch as it unfolds in a 4-module array in configurations taken from 192 48W modules, assumed to be identical. By carefully controlling the irradiance and temperature of each module throughout the experiment to create near-identical conditions for the array, they were able to study mismatch in sterile conditions. By fitting I-V curves measured for each module to a 5-parameter model, which offers a slight simplification to the model used here, they were able to estimate variations in the physical parameters between the modules. Despite finding large variations in the R_s, R_{sh}, γ parameters between modules, when measuring the mismatch loss between any 4-configuration of them no value above $\mu_{loss} = 0.6\%$ was found, leading the group to conclude that the effect is minimal. This conclusion contradicts the results obtained here, but the apparent differences in the experimental premise may explain the difference. Since our study focused on modules operating in field conditions, there are many factors we did not control for. In addition, there is a difference between the array sizes tested. While Chamberlin et al. conducting their study on a system of 4 modules at a time, with a total nominal array power of 192W, the systems tested here were considerably larger, possibly exposing further sources of mismatch that are irrelevant to small systems.

Conclusions

An estimation of the array mismatch was performed on 13 photovoltaic sites based on I_{MP}, V_{MP} measurements taken at field conditions. Novel mathematical techniques were used to reconstruct individual I-V curves per module based on real-time optimizer data, leading to a non-intrusive estimate of the continously-evolving combined I-V curve of the array. Aggregated mismatch values were calculated for each site, and were found to lie in the range $[2.3\%, 8.7\%]$. The combined mismatch observed is $\mu_{loss} = 5.4\% \pm 1.9\%^{+0.8\%}_{-0.9\%}(meas.)$. The quantity and detail of optimizer data provides a unique opportunity for future studies in this field.

References

[1] R. Zilles and E. Lorenzo. "An Analytical Model for Mismatch Losses in PV Arrays". en. In: *International Journal of Solar Energy* 13.2 (Jan. 1992), pp. 121–133. ISSN: 0142-5919. DOI: 10.1080/01425919208909779. URL: http://www.tandfonline.com/doi/abs/10.1080/01425919208909779 (visited on 05/11/2020).

[2] D. Picault et al. "Forecasting photovoltaic array power production subject to mismatch losses". en. In: *Solar Energy* 84.7 (July 2010), pp. 1301–1309. ISSN: 0038092X. DOI: 10.1016/j.solener.2010.04.009. URL: https://linkinghub.elsevier.com/retrieve/pii/S0038092X10001556 (visited on 05/11/2020).

[3] Joao Lucas de Souza Silva et al. "Study of Power Optimizers for Grid-connected Photovoltaic Systems". In: *IEEE Latin America Transactions* 17.01 (Jan. 2019), pp. 127–134. ISSN: 1548-0992. DOI: 10.1109/TLA.2019.8826704. URL: https://ieeexplore.ieee.org/document/8826704/ (visited on 05/11/2020).

[4] D. Nguyen and B. Lehman. "An Adaptive Solar Photovoltaic Array Using Model-Based Reconfiguration Algorithm". In: *IEEE Transactions on Industrial Electronics* 55.7 (2008), pp. 2644–2654.

[5] Thomas S. Wurster and Markus B. Schubert. "Mismatch loss in photovoltaic systems". en. In: *Solar Energy* 105 (July 2014), pp. 505–511. ISSN: 0038092X. DOI: 10.1016/j.solener.2014.04.014. URL: https://linkinghub.elsevier.com/retrieve/pii/S0038092X14001996 (visited on 05/11/2020).

[6] Charles E. Chamberlin et al. "Effects of Mismatch Losses in Photovoltaic Arrays". en. In: *Solar Energy* 54.3 (Mar. 1995), pp. 165–171. ISSN: 0038092X. DOI: 10.1016/0038-092X(94)00120-3. URL: https://linkinghub.elsevier.com/retrieve/pii/0038092X94001203 (visited on 05/11/2020).

[7] F. Spertino and J.S. Akilimali. "Are Manufacturing I–V Mismatch and Reverse Currents Key Factors in Large Photovoltaic Arrays?" In: *IEEE Transactions on Industrial Electronics* 56.11 (Nov. 2009), pp. 4520–4531. ISSN: 0278-0046. DOI: 10.1109/TIE.2009.2025712. URL: http://ieeexplore.ieee.org/document/5129957/ (visited on 05/11/2020).

[8] Donald F. Gatz and Luther Smith. "The standard error of a weighted mean concentration—I. Bootstrapping vs other methods". en. In: *Atmospheric Environment* 29.11 (June 1995), pp. 1185–1193. ISSN: 13522310. DOI: 10.1016/1352-2310(94)00210-C. URL: https://linkinghub.elsevier.com/retrieve/pii/135223109400210C (visited on 05/11/2020).

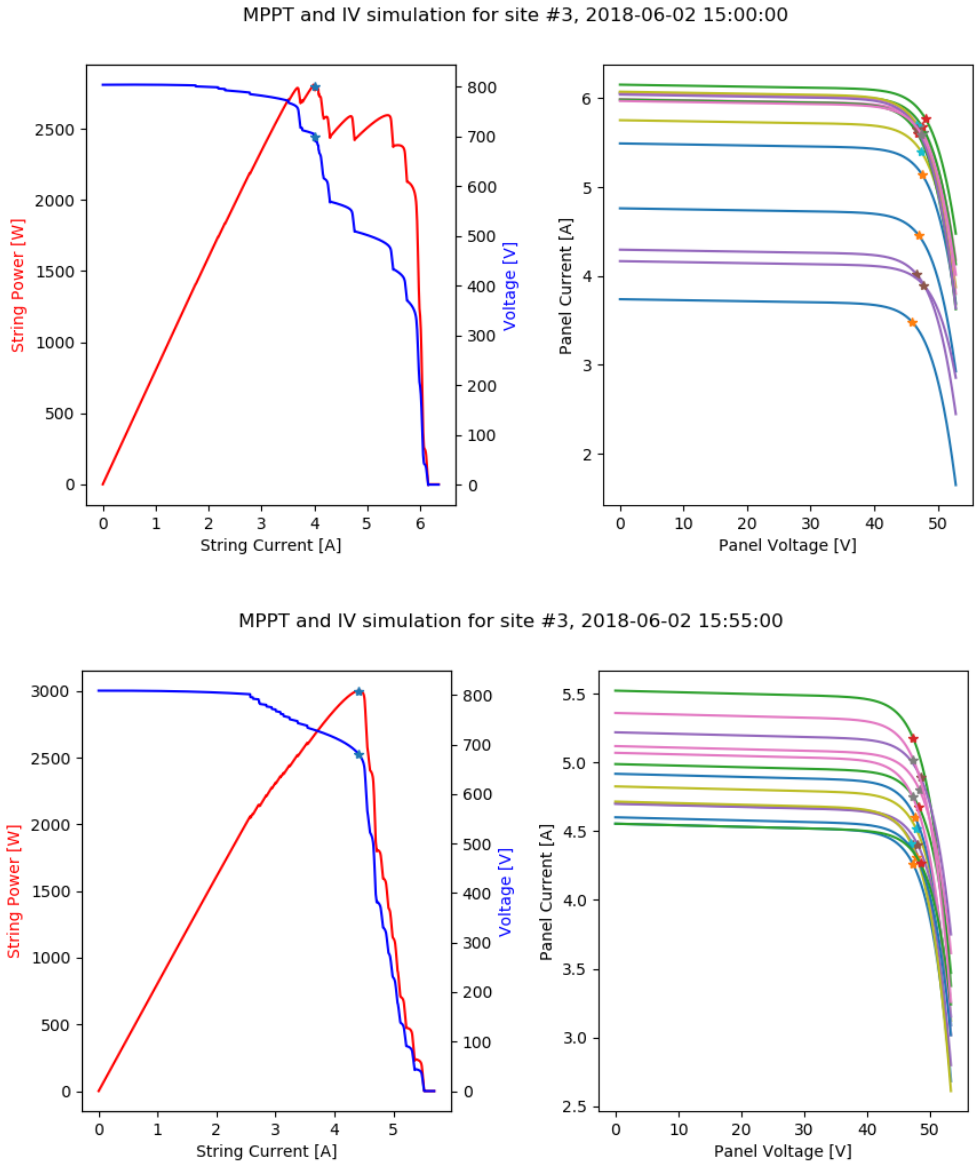

Figure 3: Right: Individual I-V curves obtained by one-diode model, based on the optimizer inputs at a given moment. The measured current, voltage are shown in asterisks. Left: Resulting combined string power (red) and voltage (blue) as a function of string current. The string MPP obtained during the process is shown as an asteirsk. These examples are taken from site #3 on table 1, from two specific times: 2018-06-02 15:00 (top) and 2018-06-02 15:55 (bottom)

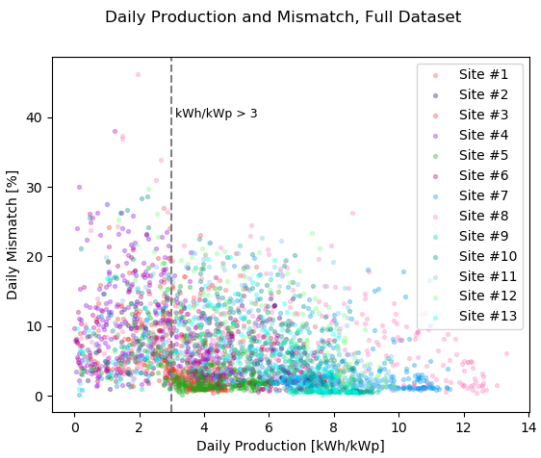

Figure 4: Distribution of daily production and mismatch values for the entire dataset used for the analysis. Each point here describes a single day in the dataset. Different colors indicate data from different sites. The vertical line at daily production $> 3kWh/kWp$ shows the impact of the data quality cut.

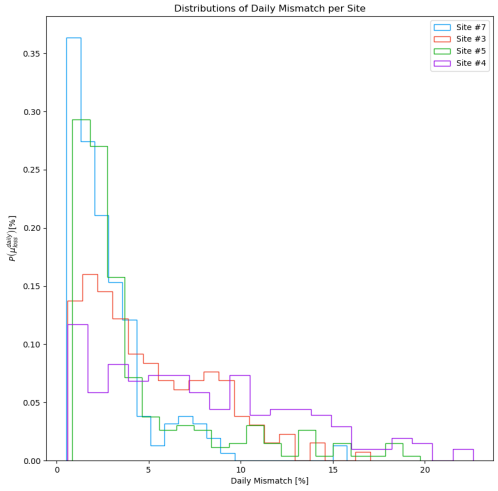

Figure 5: Normalized distribution of daily mismatch values for days with daily production $> 3kWh/kWp$ for 4 selected sites. Different colors indicate different sites.

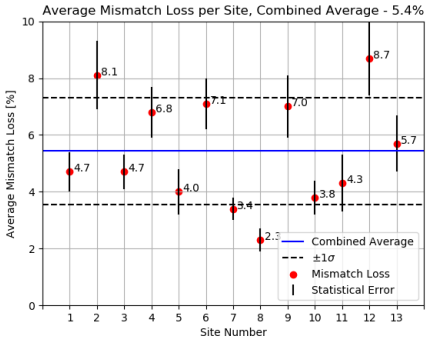

Figure 6: Combined results of the total aggregated mismatch per site. Red markers along with corresponding error bars represent results from each site. The blue solid line shows the combined average value, found to be $\mu_{loss} = 5.4\%$. Bands of ± 1 standard deviations are shown in dashed red lines.