Assessment of PV system performance with incomplete monitoring data

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Metadata Record: https://dspace.lboro.ac.uk/2134/20641

Version: Published

Publisher: © WIP Wirtschaft und Infrastruktur

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
ASSESSMENT OF PV SYSTEM PERFORMANCE WITH INCOMPLETE MONITORING DATA


Centre for Renewable Energy Systems Technology (CREST), School of Electronic, Electrical and Systems Engineering, Loughborough University, LE11 3TU, United Kingdom
*Tel.: +44 1509 635367, Email: E.Koumli@lboro.ac.uk

ABSTRACT: An analysis of PV system performance requires both meteorological and electrical data for the assessment period. However, actual in-field data acquisition is rarely 100%, often resulting in a significant amount of incomplete data sets for performance assessment. These gaps, if not taken into account, may add noticeable bias in yield assessment and thus estimations of the lacking data need to be made. An approach of back-filling the required data is given and validated here. Three different categories of data loss are identified and case-specific methods of synthesising missing data are developed. The integrity of the performance assessment process is assessed. The three cases of data loss are defined as: missing meteorological data only, missing electrical monitoring data only and missing both electrical and meteorological data. Case-specific methods are proposed and their performance against measured data is evaluated statistically by means of: root mean square error (RMSE), mean absolute error (MAE) and mean bias error (MBE). The inferred monthly performance ratio on two of the selected cases showed accurate agreement against measured data presenting significantly low MBE values, equal or less than -0.01.

Keywords: PV system, performance, monitoring, missing data

1 INTRODUCTION

PV installations in the UK have increased significantly over the last 5 years, reaching a total capacity of 5 GW to date and heading well towards the national 2020 target of 15% of total energy production by renewable energy sources. This means special importance needs to be put on quality assurance and monitoring, to assure high PV energy yield and to avoid system downtime and therefore energy loss [1].

Complete monitoring data are required in order to evaluate the energy yield of a system for a given time period as well as financial performance. Energy yield and performance ratio are essential performance metrics against which contractual guarantees are often verified. Thus, there is need for appropriate monitoring and various studies aim to outline guidelines on how proper monitoring and data analysis should be carried out [2]–[5]. These studies give recommendations on monitoring practices and analysis of the results, but do not sufficiently address back-filling requirements.

Most large PV systems operate independent meteorological and electrical monitoring systems, often purchased through different providers. These systems occasionally may lose data due to communication issues or system malfunctions. These periods of time range from minutes to days, weeks or even months.

System assessment requires an understanding of the resource (irradiance) and yield (energy yield). This can be obtained with confidence only if there are no significant gaps in the datasets. Common back-filling strategies are, e.g., to use data from previous day or same day in the last year, but there are obvious shortcomings with these strategies. Using energy estimates from previous dates is not ideal as weather variability is not taken into account. Due to variable weather as well as potential PV component degradation, it is not ideal to use data from the past year. Using data from co-located systems is another strategy. However, even identical systems often differ in performance due to differences in micro-climates.

2 CASES OF DATA LOSS

This paper presents a strategy to back-fill data with good accuracy for both short and long term periods, while taking into account weather as well as system performance variations. Three cases of data loss are identified. The first case is that of missing meteorological datasets, while electrical readings are available. This case is met in most small systems, either domestic or commercial, where installers reduce the cost by omitting the meteorological sensors. The second case is that of the electrical monitoring system being interrupted. The third case is a failure of both monitoring sub-systems, which could be due to communication or hardware failures. The last two cases are often met in the majority of solar farms. The proposed methods are validated against real measurements using two case studies shown in Table I and the results are assessed by means of i) root mean square error (RMSE), ii) mean absolute error (MAE) and iii) mean bias error (MBE).

The RMSE describes the random error in a distribution and tends to increase with outliers. MAE describes the absolute error and MBE indicates whether the model overestimates or underestimates the measurement value and hence it is probably the most important metric for PR assessment. For the first and third cases, a c-Si module (system A) from CREST PV facilities is used whereas for the second case, a case study from the UK domestic field trials programme (system B) [6] is used.

Table I: Table of PV systems

<table>
<thead>
<tr>
<th>Name</th>
<th>Module Type</th>
<th>Nominal power (W)</th>
<th>N° of Modules</th>
<th>Data Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>System A</td>
<td>Crystalline silicon (c-Si)</td>
<td>245.0</td>
<td>1</td>
<td>CREST outdoor monitoring system</td>
</tr>
<tr>
<td>System B</td>
<td>Poly-crystalline silicon (p-cSi)</td>
<td>960.0</td>
<td>8</td>
<td>UK Domestic Field Trials</td>
</tr>
</tbody>
</table>
2 ESTIMATING IRRADIANCE

Where irradiance data are lost, namely in cases 1, 3, system performance can be assessed by utilising synthetic climatic data. The method to acquire horizontal irradiance and ambient temperature is based on meteorological data collected from more than 80 ground meteorological stations on a national scale through MIDAS database [7]. Horizontal irradiance is interpolated to the nearest point of the PV system and then it is translated into tilt irradiance using separation [8] and translation algorithms [9] given that the location, orientation and tilt of the system are known. Both horizontal irradiation and ambient temperature data are interpolated using Kriging spatial interpolation method, which is described in [10].

3 ESTIMATING ELECTRICAL PERFORMANCE

Module temperature is calculated from in-plane irradiance and ambient temperature using a simple linear thermal model, such as the one presented by Ross [11]:

\[ T_m = T_a + k \cdot G \] (1)

Where \( T_m \), \( T_a \), and \( G \), are module temperature (°C), ambient temperature (°C), and in-plane irradiance (W/m²) respectively. \( k \) is known as Ross coefficient and it takes different values according to the mounting configuration of the module. In this work \( k \) was obtained by linear fitting of (\( T_m-T_a \)) against \( G \) hourly data.

The criteria for the electrical model were i) the available input data and ii) its training capability. The chosen electrical model is based on simplified King model for the maximum power point [12] and the formula is given in the following form [13]:

\[ P'(G', T') = G'((1+k_1\ln(G') + k_2\ln^2(G'))^2 + k_3T'_m+k_4T'_m\ln(G') + k_5T'_m\ln^2(G') + k_6T'_m^2) \] (2)

Where \( P' = P/P_{STC} \), \( G' = G/G_{STC} \) and \( T'_m = T_m - T_{STC} \) (STC = Standard Testing Conditions with \( G_{STC} = 1000 \text{W/m}^2 \), \( T_{STC} = 25^\circ \text{C} \) and \( P \) is maximum power (W)). The model yields a “3D power surface” as the one depicted in Figure 1. For the training process hourly data of \( P,G,T_m \) around the missing period (i.e. the validation set) are fed into Eq.(2) and the coefficients (\( k_i-k_o \)) are determined via a curve fitting algorithm.

Data quality checks were applied prior to feeding the model, as invalid input data could corrupt the training process. Here, an optimisation algorithm was used to detect the best training set for a period of one missing month, which was used as the validation set. It was found that the best training set is approximately 15 and 25 days before and after the missing period respectively, which results in 40 days of hourly data, in total. System performance is, however, installation specific. Thus the agreement is improved by ongoing training to keep performance descriptors recent.

Finally, performance ratio is calculated using Eq. (3):

\[ PR = \frac{E \cdot G_{STC}}{H \cdot P_{STC}} \] (3)

Where \( E \) is the energy output (Wh) and \( H \) is the in-plane irradiation (Wh/m²).

4 VALIDATION RESULTS ON CASES 1, 2, 3

Validation results for the three cases of data loss are discussed in the following sections.

4.1 Missing irradiance and temperature

System A was used for case 1 and monthly PR values were calculated for one year (2014). Total horizontal and in-plane irradiation and PR are compared in Figure 2 (a) and (b). The statistical results are given in Table II.
Global horizontal irradiance is estimated with a very good agreement to measured data with a MBE of approximately 1.2 (kWh) for the whole year. PR shows an average overestimation of about 10% throughout the year as in-plane irradiation is slightly underestimated with MBE of about -8.1 (kWh), which is primarily due to the models involved in the process of separation and translation of horizontal irradiance to plane of the array [14].

4.2 Missing electrical data

In case 2, System B was considered and it has been assumed that a whole month of energy readings is missing, which is a very realistic case. Eq.(1) is used to calculate the energy output while meteorological data are available for this period. Data from dates around the missing period i.e. the past and later days around this gap have been used to extract the coefficients for the electrical model, using the method described in Section 3. The results are shown in Figure 2 for daily analysis including the monthly result. This method gives a very good agreement with measurements with daily and monthly RMSE of about 0.06 and 0.002 kWh respectively.

Moreover, using the replenished period to acquire the monthly PR, it was found that this can be calculated with a very small monthly RMSE of about 0.001 (and MBE of about -0.001). This points out the efficiency of the proposed method to acquire the missing performance data when meteorological data are available from the monitoring system.

4.3 Missing meteorological and electrical data

Case 3 is a combination of the two cases described above. It has been described in previous work [15], here, performance ratio is also included in the analysis. A missing month of meteorological and electrical data (April 2014) is inferred. The comparison of real and modelled energy output and performance ratio are shown in Figure 4.

The following Tables III-V describe the statistical results for irradiation, ambient and module temperature as well as energy output and performance ratio.

Table II: Statistical results for monthly and annual analysis of measured and interpolated climatic data.

<table>
<thead>
<tr>
<th></th>
<th>Total Horizontal irradiation (kWh)</th>
<th>In-plane irradiation (kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monthly RMSE</td>
<td>Monthly MAE</td>
</tr>
<tr>
<td></td>
<td>2.13</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Figure 3: Comparison of modelled and measured energy output and PR for the missing month (March 2003). The last column represents the monthly energy output.

Figure 4: Comparison of modelled and measured energy output and PR for the missing month

The results for ambient temperature show that it can be interpolated to the location of interest with a very small bias and RMSE. This is expected as temperature is temporally and spatially more homogeneous than irradiance considering the same distance and the UK climate. This bias increases for module temperature as it propagates from both in-plane irradiation (inherent underestimation) and ambient temperature, but the effect is generally very small (a couple of degrees less than the measured value). The statistical results for energy output are affected by in-plane irradiation and module

Table III: Statistical results for ambient and module temperature comparisons

<table>
<thead>
<tr>
<th></th>
<th>Ambient temperature (K)</th>
<th>Module temperature (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily</td>
<td>Monthly</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.54</td>
<td>0.45</td>
</tr>
<tr>
<td>MAE</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>MBE</td>
<td>-0.45</td>
<td>-0.45</td>
</tr>
</tbody>
</table>

Table IV: Statistical results for in-plane irradiation

<table>
<thead>
<tr>
<th></th>
<th>Irradiation (kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.44</td>
</tr>
<tr>
<td>MAE</td>
<td>0.34</td>
</tr>
<tr>
<td>MBE</td>
<td>-0.21</td>
</tr>
</tbody>
</table>

Table V: Statistical results for energy output and PR

<table>
<thead>
<tr>
<th></th>
<th>Energy output (kWh)</th>
<th>PR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily</td>
<td>Monthly</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.11</td>
<td>1.52</td>
</tr>
<tr>
<td>MAE</td>
<td>0.09</td>
<td>1.52</td>
</tr>
<tr>
<td>MBE</td>
<td>-0.05</td>
<td>-1.52</td>
</tr>
</tbody>
</table>

The results for ambient temperature show that it can be interpolated to the location of interest with a very small bias and RMSE. This is expected as temperature is temporally and spatially more homogeneous than irradiance considering the same distance and the UK climate. This bias increases for module temperature as it propagates from both in-plane irradiation (inherent underestimation) and ambient temperature, but the effect is generally very small (a couple of degrees less than the measured value). The statistical results for energy output are affected by in-plane irradiation and module
temperature resulting in slight underestimation of the final result with an MBE of about 0.1. However, the result shows very good agreement for the PR. Daily and monthly statistical metrics appear to be very small (MBE is -0.02 and -0.01 respectively). This is due to the fact that the inherent underestimation of in-plane irradiation in energy output is diminished (see Eq.(3)) in the PR.

5 CONCLUSIONS

This paper examined typical data loss and identified three distinct cases requiring different back-filling solutions. The results were validated against measured data from two PV systems. An interpolation technique which exploits climatic data from ground based stations was used to acquire the missing meteorological data at a given system location and an empirical model was used to calculate the energy output using as input data in-plane irradiation and module temperature.

Global horizontal irradiance showed very good agreement with measured data with a MBE of 1.2 kWh for the annual result. For the first case of data loss, namely missing irradiance data, PR was overestimated by about 10%. Future work will focus on reducing this bias mainly resulting from the underestimation of in-plane irradiation.

For the second case a missing month of electrical data was considered. The analysis gave an average monthly MBE in energy output of about 0.04% and performance ratio of approximately 0.001 for a whole missing month, which points out the efficiency of the applied back-filling method.

For the third case, a missing month of both meteorological and electrical data was considered. The results were satisfactory for the back-filled monthly energy output and for performance ratio with the latter presenting a significantly small MBE of about -0.01, due to the elimination of the inherent underestimation which derives from in-plane irradiation. This is an important outcome given that PR is a very significant performance metric and it was estimated for a period of complete lack of monitored data.

6 REFERENCES


