How much power is lost in a hot-spot? A case study quantifying the effect of thermal anomalies in two utility scale PV power plants

Åsmund F. Skomedal [a,b,*], Bjørn L. Aarseth [b,a], Halvard Haug [a,b], Josefine Selj [a,b], Erik S. Marstein [a,b]

[a] Institutt for Energiteknikk (IFE), Instituttveien 18, 2007 Kjeller, Norway
[b] University of Oslo, Department of Technology Systems, Gunnar Randers vei 19, 2007 Kjeller, Norway

ARTICLE INFO

ABSTRACT

Methods for quick and accurate detection and diagnosis of defects in PV systems are increasingly important as the global photovoltaic (PV) capacity continues to grow at a rapid pace. Two of the most used methods for defect detection involve aerial infrared thermography and data analysis of production data. In this work, we combine the two methods to analyze two utility scale PV plants, providing new understanding about the two methods. We report on the percentage and distribution of thermal anomalies of different categories and quantify their relationship with performance on string level. We find that the most important parameter for determining production losses on string level is the number of module substrings containing thermal anomalies. Due to the large variability of the effect of different thermal signatures, as well as uncertainties in the estimate of the string performance, we find no clear correlation between performance and thermal signature category or temperature. However, a whole hot cell is the thermal signature that on average has the smallest impact on the power output on string level. Finally, in our data, the performance of a string of 20 modules with 3 bypass diodes is *on average* reduced by 1.16 ± 0.12% per module substring containing thermal anomalies.

1. Introduction

The annual global installed PV power capacity has sky-rocketed in recent years, reaching almost 100 GW$_p$ in both 2017 and 2018 (Masson et al., 2019). This amounted to an increase of around 1/3 and 1/4 of the cumulative capacity in these years, respectively. Since 2013, most of the new PV capacity has been installed in utility scale power plants (Masson et al., 2019). For the owners, once a plant has been commissioned, it is important that the energy output is as expected, and preferably higher. This demands adequate *operation and maintenance* (O&M) routines, which includes cleaning, fault detection, and quick replacement of faulty components. Several approaches exist for fault detection, including real-time monitoring of power output on inverter or string level, visual inspection, electroluminescence and *Infrared Thermography* (IRT) (Daliento et al., 2017; Hoyer et al., 2009). Further, much research has been done involving distributed power electronic architectures and dynamical reconfiguration techniques for fault detection and hot-spot mitigation (Balato et al., 2015; Costanzo and Vitelli, 2018; Femia et al., 2008; Olalla et al., 2015). In this study, however, we are interested

in gaining an increased understanding of IRT and power output monitoring by combining the two methods.

IRT is a contactless imaging technique relying only on an infrared camera, as well as stable weather conditions of relatively high irradiation. This allows for detection and classification of thermal anomalies on module, string and inverter level during operation (Buerhop et al., 2012a; Jahn et al., 2018; Tsanakas et al., 2016). Classification of thermal anomalies is normally done by ascribing them into one of several sub-categories, as explained in Section 2.2 and in the literature (Jahn et al., 2018; Tsanakas et al., 2016). To increase throughput of IRT inspections *Unmanned Aerial Systems* (UAS), usually in the form of multicopters, have been adopted (Aghaei et al., 2015; Buerhop et al., 2012b; Gallardo-Saavedra et al., 2018; Zefri et al., 2018). This allows for significantly lower inspection times, especially for larger systems. Interestingly, despite the level of implementation on commercial scale, it is still not well-known how a given thermal signature will affect the power loss of a PV system with a given string design.

In the literature, there are many examples of research on data-based performance analysis. Some of the authors focus on the models/metrics

---

that are used (Bizzarri et al., 2015; Marion et al., 2005; Ventura and Tina, 2016), some focus on the procedure for detecting outliers/faults (Platon et al., 2015; Woyte et al., 2014), while some explore the possibilities of employing machine learning on the problem (Rodrigues et al., 2016). A rich variety of approaches exist, all with their different strengths and weaknesses (Livera et al., 2019). However, many of the proposed methods are only tested on lab-scale systems, and not validated for utility scale scenarios. Furthermore, there is little feedback from the commercial PV system owners about which methods have been successfully employed, and which are not considered relevant. There is therefore a need for scientific contributions based on utility scale systems in order to increase the usefulness of these models in commercial PV plant operation.

Both IRT and performance monitoring are important tools by themselves. However, these two methods are based on different physical principles, and each has the capacity to detect only a limited range of defects. The combination of the two methods might provide information that is lost when each method is applied independently. In the literature, some steps have been taken to illuminate this topic. One approach has been to model thermal signatures observed in the field, with support of field measurements, to better understand the power loss at module level, and how this translates to string level (Buerhop et al., 2015, 2012b). The conclusion was that system power loss is very dependent on defect type and severity, as well as the numbers of modules in the string. Others have focused on comparing IRT to power production data from measurements on module scale, either with power optimizers on 1–3 modules, or module resolved voltage measurements (Buerhop et al., 2018; Teubner et al., 2017). The high granularity of measurements in these systems makes it possible to trace defects back to their origin, and determine when a specific defect has occurred (Buerhop et al., 2018). Another study that combines IRT with production-based performance analysis is based on module resolved measurements (Stegner et al., 2018). This is an advantage for understanding the losses that happen on module level, but the conclusions do not necessarily generalize to losses on string or inverter level, since the electrical operating conditions change when modules are connected in series and parallel. The first study, to our knowledge, that combines IRT analysis with string-level production data from commercial/utility scale PV plants is (Dalsass et al., 2015). The plants they study have a combined capacity of about 9 MWp, giving a decent basis for defect statistics. Dalsass et al., like us, also employ simulations of the PV systems, which demonstrably gives a deeper insight into the mechanisms of the thermal anomalies. However, the ambient conditions during the IRT were not as stable as in our case, leading to higher uncertainties in the thermal signature analysis. Also, even though Dalsass et al. has data from about 40,000 PV modules, this does not provide enough data on thermal anomalies to quantify any correlation between average production losses and hot-spot temperature, defect category, or the number of defects.

In all the cases above, the number of thermal signature categories discussed is limited. This is largely due to the limited availability of thermal signature data from the field. Since PV modules are relatively reliable components, a large number of inspected modules are required in order to get significant statistics on thermal signatures. This begs the need for more data from utility scale systems. However, acquiring production data from such systems is often difficult for independent researches, since the data are considered proprietary or business sensitive.

This study is aimed at filling in this gap. This paper presents data from two utility scale PV systems with a combined capacity of 115 MWp, giving valuable insight into both the statistics of thermal signatures, and the statistics of the effect on these signatures on system performance. This is important not only to the owner of the particular PV systems being studied, but to PV researchers and asset owners in general. In short, the goal of this work is to explore the cross-section between thermal imaging and production-based performance analysis. The research questions that are addressed in this paper are:

(a) What parameters determined by thermal imaging are most correlated with performance losses on system level?
(b) Are there any systematic differences in the average string-level losses induced by the different thermal signature categories?
(c) What effect does different hot spot temperatures typically have on power losses on string level?

2. Method

2.1. Description of PV systems

The data used in this paper come from 2 utility scale power plants in Sub-Saharan Africa with a combined DC capacity of 115 MWp. According to the Köppen-Geiger climate classification system (Beck et al., 2018), both plants are in a cold, arid steppe climate (BSk). Both the systems were commissioned in 2014 and have single-axis trackers Both sites have a similar array setup, with 43 and 82 central inverters in location 1 and 2 respectively. Each inverter has one MPPT, and about 160 strings connected in parallel, with 20 modules in each string. The current measurements used are for two and two strings connected in parallel. One such unit of two strings is the smallest unit where currents are measured, and this is done with a maximum uncertainty of 1%. We will refer to these units as *double-strings* in the remainder of the paper. The modules have 72 cells and a peak capacity of 290 Wp and 295 Wp in location 1 and 2. All the modules in these plants have three bypass diodes, defining three *module substrings* of 24 cells connected in series.

2.2. Infrared thermography

The thermal images involved in the analysis in this paper were captured with aerial IRT. The imaging and the image analysis were done by Hawk62 Aviation and GeoSUN Africa. We have access to these results through a report, with detailed information about each module with a thermal anomaly (see snippet in Fig. 1). For each module, the number of defective cells (or diodes), hot spot temperatures, thermal signature categories (see Section 2.2.1), visual defects, and ambient conditions (temperature and irradiance) at the time of imaging are recorded. To avoid false positives, temperature anomalies smaller than 5 °C are ignored. The overall findings are summarized in table format, with one row for each defective module, which makes it easy to import the data to an analysis platform. It is the results from these reports that are used in the analysis in this paper.

For imaging, a FLIR Tau2 640 × 512 IR camera, with an absolute temperature uncertainty of 5 K and a sensitivity of 50 mK, and a 16 MP Micro Four Thirds optical camera was utilized. All IR images were taken at an altitude corresponding to a ground sampling distance of <1.3 cm/pixel. For geotagging of images, a combination of real time and post processed Kinematic GNSS receivers were used. According to Hawk62 Aviation and GeoSUN Africa, this leads to an uncertainty of ±1 module
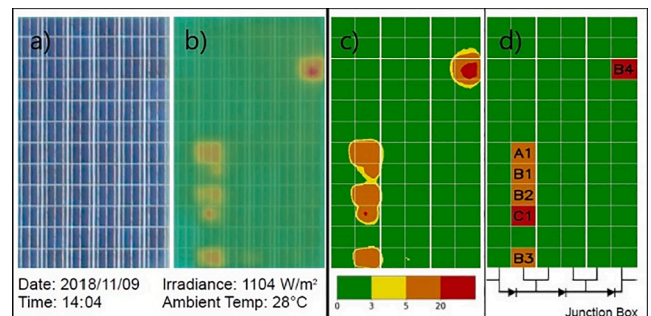

Fig. 1. Example from the IR-report, showing (a) an optical image of a module, (b) an overlay of the optical image and the IR-image, (c) a discretized IR-image, and (d) the categorization of the individual cells according to the categories described in Section 2.2.1.

to the left or right.

The thermal signature of a defect depends on ambient conditions. Since we are comparing thermal images which are necessarily taken at different times, it is important that the air temperature and the in-plane irradiance is as constant as possible during these times, and that the guidelines of IEC TS 62446-3 are upheld (IEC, 2017). In Fig. 2 we show the distribution of ambient temperatures and in-plane irradiances during the imaging process. The temperature varied with about 5 °C, and the irradiance varied with less than 40 W/m^2, during the imaging of the plants.

2.2.1. The thermal defect categories

The IRT images containing thermal anomalies are divided into 6 thermal signature categories (listed in Table 1), adapted from the IEA-PVPS report "Review of Failures of Photovoltaic Modules" (Köntges et al., 2014). We emphasize that this categorization is based on the shape and area of the thermal signatures, not the underlying defects. However, we have listed some possible causes of each category in Table 1. Based on the data that was available to us, it is apparent that the different categories overlap, and that the distinction between them is not always obvious. This means there may be instances of misclassification. This form of categorization, however, is easy to implement, and to our knowledge there is no better alternative in the literature. It is worth to note that, even if category D is listed with the possible cause "fully active bypass diode", all the thermal signatures from A to E may cause the bypass diode to be activated.

2.3. Performance analysis

The granularity in which data-based performance analysis can be done is determined by the granularity of electrical measurements in the power plant. As mentioned in Section 2.1, the smallest unit on which electric measurements are done in the plants analyzed in this paper are two strings connected in parallel (double-strings), and these are thus also the units that are used for performance analysis.

The performance analysis is carried out in three steps: (1) Time series data from the plant are acquired. In this case, these data include only currents and voltages from the double strings. (2) These data are filtered, in order to remove corrupt data and minimize noise, and aggregated into a daily array yield. (3) Finally, the average yield in the period when the thermographic images were taken is compared on string level, giving a relative deviation for each double-string. This relative yield is used as the metric for comparing the string performances. The following sections will explain this procedure in more detail.
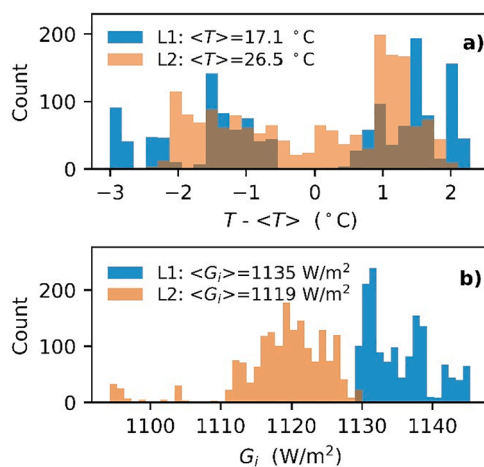

Fig. 2. Distribution of (a) temperatures T and (b) in-plane irradiance G_i during the imaging of the plants. $<T>$ and $<G_i>$ are the average temperature and in-plane irradiance during the imaging of each plant.

Table 1

Description of the thermal signature categories provided by the aerial IRT report.

Thermal sign. category	Description	Possible cause
A	Whole cell is hot	Shunted or partially shaded cell
B	Part of cell is hot	Cracked cell or disconnected interconnect
C	Single hot point	Cracked cell or damaged busbar/interconnect
D	Uniformly hot substring	Fully active bypass diode[1]
E	Patchwork pattern of hot cells	Short-circuited diode/cracked front glass
J	Hot junction box	Poor connection causing ohmic heating

[1] By *fully* active, we mean that the diode is conducting almost all the current, presumably because the module substring is unable to carry any significant current. This could be caused by a very high series resistance or a cell with very low current generation.

2.3.1. Filtering

Before the data is processed, it is very important to ensure good data quality. All measurements that are corrupted, either by communication loss during logging, or insufficient maintenance of the sensors, are removed. Subsequently, filtering for noise reduction is employed. The goal of such filtering is to maximize the variations in the performance of the double-strings that is caused by real defects, relative to variations in the performance that is not caused by defects, and that we are unable to account for or model. For instance, we may account for both variations in irradiance and temperature by assuming that they have an approximately linear correlation with the energy output of a PV module. On the other hand, it is difficult to account for varying spectral effects in the incident light, varying irradiance and temperatures across the power plant, or varying degrees of soiling and self-shading. Instead, we try to minimize such effects by filtering away time stamps with low irradiance, low solar elevation angle, high angle of incidence, and cloudy conditions. We achieve this by applying filtering cutoff thresholds, following the procedure described in (Skomedal et al., 2019). The applied cutoff thresholds are summarized in Table 2. Further, clear sky filtering (removing cloudy conditions) is done using the clear sky detection algorithm implemented in Python PVLib (F. Holmgren et al., 2018).

2.3.2. The performance metric

In order to assess the effect of the thermal defects on the performance of the PV system on double-string level, it is necessary to quantify this performance. In the literature, there are many suggested metrics that can be used for this purpose, including the performance ratio (*PR*) (Guerriero et al., 2017; Marion et al., 2005; Ventura and Tina, 2016; Woyte et al., 2014), the closely related temperature-corrected *PR* (*PR_{STC}*), and the weighted relative energy error (Bizzarri et al., 2015). Further, empirical models like the PVUSA model (Whitaker et al., 1997), or the models proposed in (Platon et al., 2015) and (Ventura and Tina, 2016), may be employed to model the system output. What these methods have in common is that they account for the irradiance and possibly the temperature conditions, correcting for these effects in a more or less physical way. There are also many examples in the literature where machine learning (ML) models have been employed for performance analysis, e.g. in (Rodrigues et al., 2016).

Table 2

Filtering cutoff thresholds for minimum solar elevation angle, maximum angle of incidence, and minimum in-plane irradiance for the plant locations.

Location	L1	L2
Min. solar elevation angle (°)	35	20
Max. angle of incidence (°)	55	55
Min. G$_i$ (W/m^2)	430	640

Both physical, empirical and ML models have tuning parameters that need to be adjusted to the system that is to be modeled. For physical models, these are physical parameters (e.g. array orientations, topology, and shading conditions) found in the design drawings of the system. For empirical and ML models, these parameters are set by fitting/training the model to/on historical data. In this case, all defects need to be labeled in advance, so that they can be excluded from the model fitting/ training. If this is not done adequately, one may end up with a model that perfectly models the system, including the effect of the defects, thus defeating the purpose of the study. Doing this manually on thousands of time series is an enormous task, and any automatic labeling would need extensive validation. As far as we know, no method for doing this robustly exists to date. Developing our own method for doing this falls outside the scope of this paper.

For these reasons, our approach to performance assessment does not require any kind of modeling. Instead we rely on the statistics of the utility scale power plant. This approach is applicable if the current (or power) is measured on near string-level (e.g. two strings in parallel, as in this work). It is also required that the strings can be grouped into sections where the physical conditions of the strings (orientation, topology, shading conditions, soiling, etc.) are similar, and that there are many modules in each group.

Our approach is simply to use the relative differences in production between the strings in periods where the weather conditions are most uniform across the plant, and to use this relative difference as a measure of the performance of each string. The method for ensuring uniform weather conditions is to use filtering, as discussed in Section 2.3.1. The performance metric we use is the filtered *relative yield*, which is defined as

$$y^*_{rel} = \frac{Y^*_s - Y^*_{s,med}}{Y^*_{s,med}} \cdot 100\%,$$ (1)

where $Y^*_s = \sum E^*_{i,out}/P_{nom}$ is the filtered specific yield on a given day, and $Y^*_{s,med}$ is the median Y^*_s of the double-strings of the same group on the same day. $\sum E^*_{i,out}$ is the total energy output over a given period and P_{nom} is the nominal power of the sub-array for which the specific yield is calculated. The asterisk * means that the data is filtered prior to aggregation. In this work, the y^*_{rel} is calculated daily for each double-string. Finally, it is the average relative yield $<y^*_{rel}>$ over the one-month period around the time when the thermographic images are taken that is used as a measure of the performance of each double-string. Note that this metric is identical to what is used for performance assessment in (Dalsass et al., 2015), with the exception of filtering.

It is also worth to note that we are comparing $<y^*_{rel}>$, which is based on a whole month worth of data, with an instantaneous scan of our plants (the IRT images). The reason we do this is that, even after filtering, there is a day-to-day variability in y^*_{rel}. When we do a time-average, we take advantage of the law of large numbers, assuming that the impact of faults on performance is stable in this period, and that the variability can be treated as symmetric noise that will average to zero over time. Thus, by doing this we remove the random day-to-day variability in performance. Since we expect a correlation between $<y^*_{rel}>$ and the number of thermal anomalies, the length of the averaging period was chosen in order to maximize this correlation, measured by the R^2 of the regression fit shown in Section 3.2.

One potential shortcoming of our approach is that we do not differentiate between the performance of arrays connected to different inverters. Since we are only comparing strings connected to the same inverter, and since the MPPTs are located at the inverters, any differences in the MPP voltage (V_{mpp}) are neglected. Thus, if one MPPT enforces a reduced V_{mpp} due to a large number of defects, this is not accounted for. Further, a lowered V_{mpp} caused by active bypass will give a smaller, or even zero current reduction in strings with defects. In general, the more the V_{mpp} is lowered by active bypass diodes, the smaller the difference in y^*_{rel} will be between defective and non-defective strings. This effect is neglected by our approach. However, in our data there is no significant correlation between V_{mpp} and the number of thermal anomalies in each inverter. We interpret this to mean there are not enough defects to significantly alter the V_{mpp}, which gives us confidence that we can rightly neglect the effect of defects on the operating voltage.

2.4. Defect power loss simulations

In a previous publication it was shown that the IV-curves of modules containing different thermal signatures can be recreated using a relative simple MATLAB Simulink circuit model based on the single diode equation (Aarseth and Marstein, 2019). In this study we use a similar model, based on 72 cell modules with 3 bypass diodes, but since the exact effect of the different thermal signatures on the IV characteristics of the modules in this study are unknown, we limit the possible defect parameters to series resistance. This is of course an oversimplification, that assumes the entire loss from the module substring defect can be approximated as ohmic losses, as long as the bypass diode is not activated. In the simulations, the module power is 290 W, and the module IV characteristics are fitted to the module datasheet. The resulting cell parameterization in the one diode model is a diode saturation current $I_0 = 2.9 \times 10^{-10}$ A, photo generated current $I_{ph} = 8.6$ A, quality factor of $N = 1$, and a cell series resistance $R_s = 6 \times 10^{-2}$ Ω. A piecewise linear diode is used for the module bypass diode, with an on-voltage of $V_f = 0.1$ V, on-resistance $R_{on} = 1 \times 10^{-2}$ Ω and an off-conductance $G_{off} = 1 \times 10^{-8}$ S.

To simulate how the different thermal signature categories affect measurable string performance, we simulate 1 to 4 module substrings in a double-string of otherwise fault-free modules, using the different resistances for thermal signature categories A-C listed in Table 3. The results of this are shown in Section 3.3. The magnitude of the resistances do not lead to the activation of the bypass diode. For category D, however, the module substring is removed so that the current can only flow through the bypass diode. Lastly, for category E, the bypass diode is removed, causing the module substring to short-circuit.

The two systems reviewed here both have central inverters into which between 75 and 80 double-strings connected in parallel are fed, all controlled by one MPPT. Simulations of full systems of this size require substantial simulation times. For this reason, only two double strings have been simulated at a time, one fault-free and the other containing the defects under review. The system V_{mpp} is then set to the V_{mpp} of the fault-free string, thus assuming that the MPP voltage of the entire 80 parallel double-string setup will only change negligibly due to one defective double-string, as discussed in Section 2.3.2. The resulting power of both double-strings are used to calculate the relative yield, using Eq. (1) from Section 2.3.2.

3. Results and discussion

3.1. Statistics of thermal anomalies

A summary of the number of thermal anomalies relative to module number (left) and differentiated by defect category relative to the

Table 3
The simulation parameters of the thermal signature categories. Category J has not been simulated because of the low number of signatures in this category.

Thermal signature	Simulated by
A	$R_s = 0.2$ Ω
B	$R_s = 1.2$ Ω
C	$R_s = 1.4$ Ω
D	Removing module substring
E	Short circuiting module substring

number of defects (right), found in each location is shown in Fig. 3. The most numerous thermal signature is category B, which is a part of a cell that is hot. There is some difference in both the fraction of defective modules and the relative distributions of the thermal signature categories between the sites. The reason for this is probably not related to climatic conditions, as both plants are located in the same climatic zone. The largest relative difference between the locations is in category E (short-circuited diode), which constitutes about 12% of defects observed in location 2, while only 2% in location 1. The reason for this can be a higher frequency of reported thunder events at location 2.

3.2. Number of thermal anomalies and string performance

There is a clear relationship between the performance on double-string level and the number of thermal anomalies in the double-string. In Fig. 4, the average relative yield $<y^*_{rel}>$ is shown as it varies with the number of module substrings containing thermal anomalies (*affected substrings*). The variation in $<y^*_{rel}>$ within each x-value, and especially at zero affected substrings, illustrates the implicit uncertainty in our estimate of the performance. Despite this, there is a clear tendency that the performance decreases with increasing number of affected module substrings. In fact, there is a near linear decrease in the median $<y^*_{rel}>$ in all locations, at least up to five affected substrings. The trend lines in Fig. 4 are obtained by ordinary least squares regression based on all data points in each location, and, in the case of the aggregated value, on all data points from all locations combined. The R^2-value is based on the mean $<y^*_{rel}>$ at each x-value. Note that we are not saying that the performance loss is expected to vary linearly with the number of affected substrings. On the contrary, in Section 3.3 we argue, based on simulations, that the response is probably non-linear. However, the linear regression is useful to quantify the typical relationship between performance loss and number of affected substrings.

Because the data from the two locations exhibit a similar trend, and since the locations have virtually identical configurations, we assume that aggregating the data from the two plants is meaningful, and that further analysis based on the whole dataset, including both locations, is valid.

Interpreting the trend line of the aggregated data to represent the average effect of the number of defects, we conclude that there is a 0.58 ± 0.06% decrease in the performance of a double-string per affected module substring in the systems we consider. Let us call this the *performance loss rate*. The uncertainty in the performance loss rate is two times the standard error of the regression estimate. If we extrapolate linearly from the result at double-string level, the performance loss rate is on average 1.16 ± 0.12% per affected substring in a 20-module string, and $1.16 \cdot 20 \cdot 3 \approx 70\%$ (±7%) in each substring that is affected.

A defect that limits the maximum current of a module substring will

often lead to the activation of a bypass diode. With an active diode, the voltage contribution of the module substring is lost. This loss of voltage leads to a current reduction in the whole string (because the string is operating at a voltage set by the MPPT, and the string current is adjusted so as to accommodate this), which constitutes the loss caused by the module substring. As long as the diode is active, the voltage loss is independent of how many defects there are in the module substring. If the performance loss in each string has a stronger correlation with the number of module substrings containing thermal anomalies (affected substrings) than the number of cells with thermal anomalies, this could mean that most of the defects lead to activated diodes. Although there is a correlation between $<y^*_{rel}>$ and the number of cells with thermal anomalies in a string, in our data the correlation is stronger with number of affected module substrings. This is confirmed by the fact that $R^2 = 0.73$ for the linear regression on number of *cells* with thermal anomalies, while $R^2 = 0.88$ for the linear regression on number of *affected substrings*. This is the reason we choose to report the performance loss per affected substring, rather than per affected cell.

Note that the performance loss rate of 1.16 ± 0.12% per affected substring (in a 20-module string) is based on two particular PV systems where 150–160 strings in parallel feed into each MPPT, and where each string has 20 modules and each module has three substrings. It does not necessarily generalize to systems with different MPPT, string, and module configurations. This is discussed further in the next section.

3.3. Thermal signature category and string performance

To study the effect of the different thermal signature categories, we have isolated double-strings that contain exclusively one kind of thermal signature. In Fig. 5 we show how the average relative yield depends on the number of affected module substrings for double-strings containing *exclusively* defects of category A, B, C, D and E respectively. In this way we hope to isolate the performance loss rate associated with each category. There were not enough defects of category J to make this kind of plot. Note that, for simplicity of visualization, in this plot we show the median along with the 25th and 75th percentile, defined by the boxes.

Judging by the median relative yield, it seems that thermal signatures of category A (whole cell hot) are not correlated with as large a performance loss as the other categories. There is a measured increase in performance at two affected substrings, and a drop in performance at three and four affected substrings. There are only three data-points at four affected substrings, and this is not enough data to average out the uncertainty in our performance estimate. On the other hand, category B, C, D and E correlate with larger performance losses, except at three affected substrings for category D. However, there are only 5 double-strings at this position, and we consider it an outlier.

Simulations of the relative reduction in P_{mpp} on double-string level for the thermal signature categories are also shown in Fig. 5. Assuming the relative reduction in P_{mpp} to be proportional to the change in the performance, and hence to the relative yield, we can compare the measurements to the simulations. Of course, the exact effect of a given thermal signature on the IV-curve of a module is not predictable. Small variations in the underlying defect can give large variations on cell and module IV characteristics, but the thermal signature category may remain the same. Therefore, series resistance has been chosen as the simulation parameter for all categories, and all defects of the same category have been simulated with the same module substring series resistance. With this approach we are not able to correlate each individual thermal defect with a corresponding IV-curve. However, it makes us able to estimate "typical" losses in P_{mpp} of strings containing each defect category, and how of the number of affected substrings relates to mismatch between the V_{mpp} of the string and the central inverter.

Although there is a large variation in the measured relative yield of the double-strings at each x-value, there is a clear tendency that the median decrease in relative yield is similar to the simulated decrease in
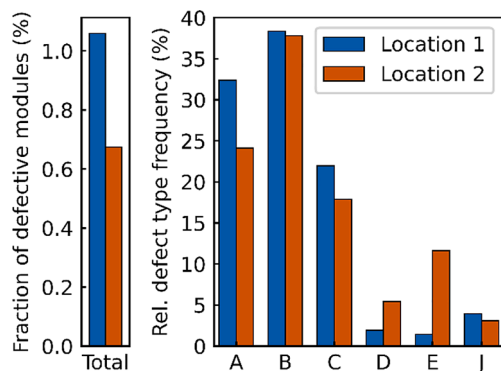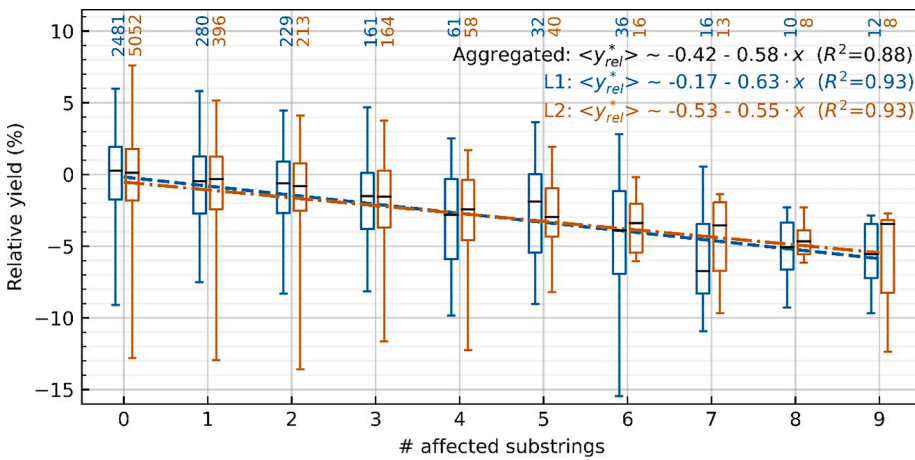

Fig. 3. Number of modules with thermal anomalies relative to the total number of modules (left) and the number of anomalies of each category relative to the total number of defects (right) at the two locations.

Fig. 4. Average relative yield $<y_{rel}^*>$ of the double-strings as it varies with number of module substrings with thermal anomalies in the two locations. The boxplots encompass the 25th and 75th percentiles, the horizontal lines show the median values, and the whiskers show the 2.5th and 97.5th percentiles. The numbers close to the top of the plot show how many data-points (each point representing one double-string) are represented in each respective x-value. The trend lines are obtained by ordinary least squares regression based on all data points in each location, and, in the case of the aggregated value, on all data points from all locations combined.
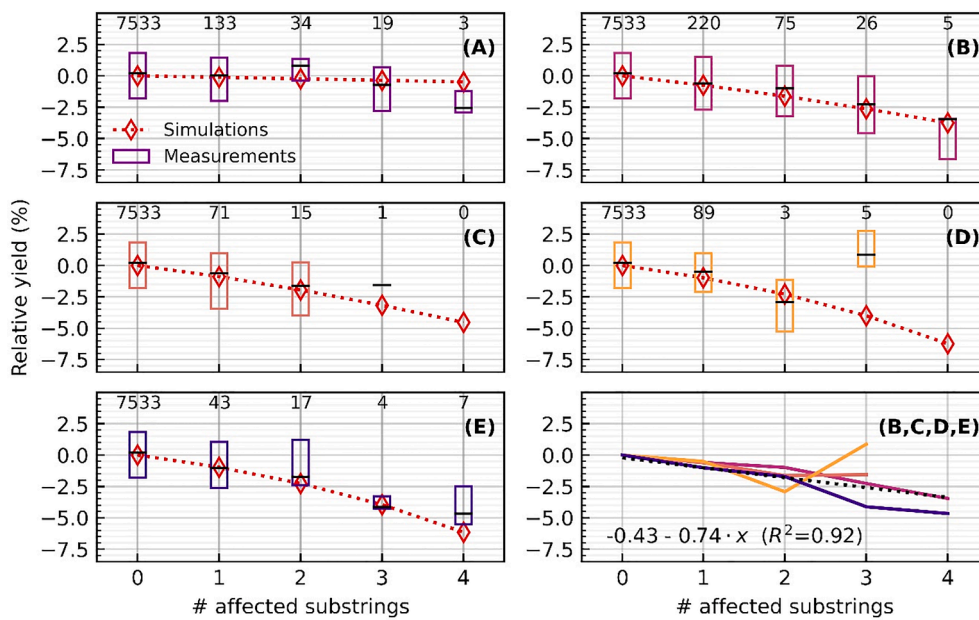
Fig. 5. The measured average relative yield $<y_{rel}^*>$ of double-strings containing only defects category A (whole cell hot), B (part of cell hot), C (single hot point), D (uniformly hot substring), E (patchwork of hot-spots), respectively, as well as B, C, D and E combined. The boxes encompass the 25th and 75th percentiles, and the black line marks the median of $<y_{rel}^*>$. Simulations of the effect the different defects on the relative reduction in P_{mpp} are also shown. A linear regression on the data with B, C, D and E defects is shown in the subplot marked by (B, C, D, E).

P_{mpp}. Except at four affected substrings in category A and E, and three affected substrings in category D, the simulations fall within the 25th and 75th percentiles of the measurements. However, there are only five or less datapoints at these positions. This means that the empirical performance assessment agrees well with the simulations.

Considering the simulations in Fig. 5, it is evident that the relationship between performance loss and number of affected substrings is not linear: The more affected substrings, the larger the effect of each added fault. The reason for this is that the IV-curve, and hence the MPP, of a string is altered by faults, and the higher the number of affected substrings is, the further away the string MPP will be from the system MPP enforced by the central inverter. In other words, according to our simulations, faults typically lead to a reduction in string current that is larger, relatively speaking, than the voltage loss in the isolated MPP of the string, for large parallel connected systems like L1 and L2.

The categories associated with the biggest loss is category D and E, which we have simulated as a fully activated bypass diode and a short circuited bypass diode, respectively. In a 40-module double-string, both give a loss of ~1% in P_{mpp} with one affected substring. This is higher than the 0.83% loss one might expect (if one assumes the loss is limited to 100%/40/3 = 0.83%, and neglects the bypass diode on-voltage and -resistance), and this is due to the mismatch between the MPP of the

defective double-string and the system MPP. To simulate category E as a perfect short-circuit may be an oversimplification; in some cases there will be some resistance in the shunt. Furthermore, a patchwork pattern of hot cells caused by a short circuited bypass diode can be hard to distinguish from several A and B signatures within a substring. Thus, it is also possible that there is a number of mis-categorizations between signature A, B and E. This is important, because robustly distinguishing a short-circuited bypass diode from other categories will potentially limit the O&M measure to replacing the bypass diode instead of the entire module. However, in Fig. 5E we see a very good agreement between simulations and measured performance loss, indicating that in our data most of the E-anomalies are indeed short circuits.

We judge that the variation in the relative yield is too high to distinguish the relative magnitude of the effect of the different categories on performance, except that category A is less detrimental than the other categories. Because of this, we have gathered category B-E in the last subfigure of Fig. 5 and applied an ordinary least squares linear regression on all data points from the four categories. The result is a reduction in relative yield by 0.74% per affected substring in a double-string, or equivalently 1.5% for a 20-module string. This means that the linear decrease in relative yield per affected substring is increased by near 30% by excluding category A, underlining the lower significance of

this category in our data.

Although it is tempting to assume that the performance loss shown in Figs. 4 and 5 is *caused* by the observed thermal anomalies, we would like to emphasize that we don't have sufficient grounds to make claims about causality. Granted, we observe clear correlations, but there may be other factors involved, such as a correlation between thermal anomalies and other, non-thermal loss factors, that we have not controlled for. Still, through the simulations, we have shown that the average performance loss rates may be caused by module-level defects, and hence it is *possible* that all the observed performance loss is associated with thermal anomalies.

While the average performance loss per affected substring (Fig. 4) does not necessarily generalize to other PV system configurations, we would argue that the performance loss rate of each thermal signature category (Fig. 5) may be similar for other systems. Variations in climatic conditions and c-Si cell technology (i.e. cell thickness, etc.) may lead to different numbers and distributions of thermal signatures (ref. Fig. 3), but the performance loss will be limited by the bypass diodes. In systems similar to the ones we have studied, with central inverters, 20 c-Si modules in each string, and 72 cells and 3 bypass diodes in each module, we expect that categories B-E will have a performance loss per affected substring similar to what we have found. This, however, needs to be confirmed by further studies.

3.4. Hot spot temperature and string performance

In order to study the effect of hot spot temperature we have isolated double-strings with exactly one thermal signature, and plotted the average relative yield versus the difference between the maximum hot spot temperature in the double-string and the average temperature of non-defective modules, ΔT. This is shown in Fig. 6, with the different categories shown in the different subplots.

There is no statistically significant correlation between performance and temperature in any category. This was tested statistically with the Pearson correlation coefficient at a significance level of 5%. We note that (Dalsass et al., 2015) seemingly find a correlation, although they don't quantify it.

In our data, the variation in the performance between non-defective double-strings is around 20%. The performance loss per affected substring is around 0.58%. We expect the effect of temperature to be of a smaller magnitude than this. This means that we are looking for a correlation that is much smaller than the noise in the data. Also, we do not

necessarily expect two thermal signatures of the same category that have the same temperature to have the same effect on the IV curve of their modules, for instance if they have different sizes. Given these facts, it is not unlikely that any correlation between performance and temperature is so small that it drowns in noise in our data. This makes us unable to say anything quantitative about the effect of hot-spot temperature on the performance on double-string level. This is not to say that performance is not negatively correlated with hot spot temperature, for instance on module or cell level. However, in our data, any correlation between hot-spot temperature and performance is overshadowed by variations that are not related to temperature.

We note that, from a thermodynamics perspective, it is not necessarily the maximum temperature that determines the thermal losses of a defect. As is noted in (Teubner et al., 2019), a more relevant parameter to correlate with performance losses would be the difference in the average temperature between the defective areas and the non-defective areas. In this way, the *area* of the thermal anomaly, as well as its temperature, is taken into account. Unfortunately, as we did not have access to the raw IR images, we were not able to perform this analysis.

4. Conclusion

By combining IRT with production-based performance analysis, we have gained an increased understanding of how thermal anomalies relate to performance on string and double-string level. We find that the most important parameter is the number of module substrings containing thermal anomalies. Although a single defect may reduce the power output of a string of 20 modules with 3 bypass diodes each by as much as 2% (according to our simulations), *on average*, in the plants in this study, the power output is reduced by $1.16 \pm 0.12\%$ per module substring with a thermal anomaly. Of the different defect categories, cells that are uniformly hot (category A) are the least detrimental for the power output. We are not able to determine the relative magnitude of the effect of the other categories independently, but by excluding the effect of the thermal signature category A, the power output is on average reduced by 1.5% per affected substring. We argue that this result, applying to thermal anomalies of category B, C, D or E, might generalize to other systems with c-Si modules and similar configuration to the plants in this study. Lastly, we do not observe any significant correlation between maximum hot spot temperature and performance on double-string level.

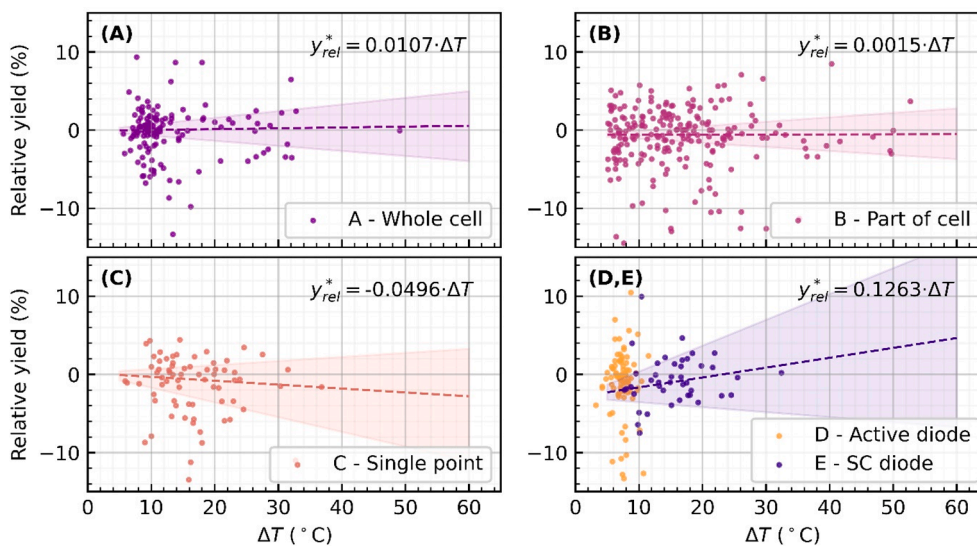Previously, much research on how different defects affect the

Fig. 6. The measured average relative yield $<y^*_{rel}>$ as it varies with temperature for the different defect categories. Only double-strings with 1 affected module substring have been included, in order to isolate the defect categories. Theil-Sen regression lines with 95% confidence intervals (marked by shaded region) are shown.

performance of PV modules has been published. However, little research exists quantifying how common different defects actually are in the field, and how much production loss one can expect from the different defects *on system level*. This case study is, to this date, the study where IRT analysis has been combined with production-based loss analysis that is based on the largest number of PV modules. Thus, we contribute to the state of the art by quantifying the distribution of thermal signatures in three utility scale PV power plants, quantifying the relationship between performance loss and thermal anomalies, and identifying correlations between parameters found by IRT and performance loss on string level. This is valuable information for PV stakeholders, as well as PV system operators, IRT service providers, policy makers, and other researchers in the field.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aarseth, B.L., Marstein, E.S., 2019. Defect recognition and power loss estimation of PV systems using infrared thermography. In: 36th European Photovoltaic Solar Energy Conference and Exhibition.

Aghaei, M., Grimaccia, F., Gonano, C.A., Leva, S., 2015. Innovative automated control system for PV fields inspection and remote control. IEEE Trans. Ind. Electron. 62, 7287–7296. https://doi.org/10.1109/TIE.2015.2475235.

Balato, M., Costanzo, L., Vitelli, M., 2015. Series-parallel PV array re-configuration: Maximization of the extraction of energy and much more. Appl. Energy 159, 145–160. https://doi.org/10.1016/j.apenergy.2015.08.073.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future köppen-geiger climate classification maps at 1-km resolution. Sci. Data. https://doi.org/10.1038/sdata.2018.214.

Bizzarri, F., Brambilla, A., Caretta, L., Guardiani, C., 2015. Monitoring performance and efficiency of photovoltaic parks. Renew. Energy 78, 314–321. https://doi.org/10.1016/j.renene.2015.01.002.

Buerhop, C., Fecher, F.W., Pickel, T., Häring, A., Adamski, T., Camus, C., Hauch, J., Brabec, C.J., 2018. Verifying defective PV-modules by IR-imaging and controlling with module optimizers. Prog. Photovolt. Res. Appl. 26, 622–630. https://doi.org/10.1002/pip.2985.

Buerhop, C., Scheuerpflug, H., Pickel, T., 2015. Defect analysis of installed PV-modules - IR-thermography and in-string power measurement. 31st Eur. Photovolt. Sol. Energy Conf. Exhib. Defect 49, 1692–1697. https://doi.org/10.4229/EUPVSEC20152015-5BO.12.6.

Buerhop, C., Schlegel, D., Niess, M., Vodermayer, C., Weißmann, R., Brabec, C.J., 2012a. Reliability of IR-imaging of PV-plants under operating conditions. Sol. Energy Mater. Sol. Cells 107, 154–164. https://doi.org/10.1016/j.solmat.2012.07.011.

Buerhop, C., Weißmann, R., Scheuerpflug, H., Auer, R., Brabec, C.J., 2012b. Quality control of PV-modules in the field using a remote-controlled drone with an infrared camera. In: Proc. 27th Eur. Photovolt. Sol. energy Conf. Exhib. (EU PVSEC) 3370–3373. https://doi.org/10.4229/27thEUPVSEC2012-4BV.2.43.

Costanzo, L., Vitelli, M., 2018. On the link between the operating point and the temperature distribution in PV arrays working under mismatching conditions. Int. J. Photoenergy. https://doi.org/10.1155/2018/6359543.

Daliento, S., Chouder, A., Guerriero, P., Pavan, A.M., Mellit, A., Moeini, R., Tricoli, P., 2017. Monitoring, diagnosis, and power forecasting for photovoltaic fields: A review. Int. J. Photoenergy. https://doi.org/10.1155/2017/1356851.

Dalsass, M., Scheuerpflug, H., Maier, M., Brabec, C.J., 2015. Correlation between the monitoring data of a photovoltaic power plant and module defects detected by drone-mounted thermography. 31st Eur. Photovolt. Sol. Energy Conf., 1793–1798.

F. Holmgren, W., W. Hansen, C., A. Mikofski, M., 2018. pvlib python: a python package for modeling solar energy systems. J. Open Source Softw. 3, 884. https://doi.org/10.21105/joss.00884.

Femia, N., Lisi, G., Petrone, G., Spagnuolo, G., Vitelli, M., 2008. Distributed maximum power point tracking of photovoltaic arrays: Novel approach and system analysis. IEEE Trans. Ind. Electron. 55, 2610–2621. https://doi.org/10.1109/TIE.2008.924035.

Gallardo-Saavedra, S., Hernández-Callejo, L., Duque-Perez, O., 2018. Technological review of the instrumentation used in aerial thermographic inspection of photovoltaic plants. Renew. Sustain. Energy Rev. https://doi.org/10.1016/j.rser.2018.05.027.

Guerriero, P., Piegari, L., Rizzo, R., Daliento, S., 2017. Mismatch based diagnosis of PV fields relying on monitored string currents. Int. J. Photoenergy 2017, 1–10. https://doi.org/10.1155/2017/2834685.

Hoyer, U., Burkertand, A., Auer, R., Buerhop-Lutz, C., Vodermayer, C., Mayer, M., Wotruba, G., 2009. Analysis of PV modules by Electroluminescence and IR Thermography. In: 24th Eur. Photovolt. Sol. Energy Conf., 3262–3266.

IEC, 2017. IEC TS 62446-3:2017.

Jahn, U., Herz, M., Köntgens, M., Parlevliet, D., Paggi, M., Stein, J.S., Tsanakas, I., Berger, K.A., Ranta, S., French, R.H., Richter, M., Tanahashi, T., Friesen, G., Hottenrott, D., Ndiro, E., Mathiak, G., Weinreich, B., Fada, J.S., Kubicek, B., Buerhop-Lutz, C., 2018. Review on infrared and electroluminecence imaging for PV field applications. IEA-PVPS.

Köntges, M., Kurtz, S., Packard, C.E., Jahn, U., Berger, K., Kato, K., Friesen, T., Liu, H., Van Iseghem, M., 2014. Review of failures of photovoltaic modules. IEA-PVPS. https://doi.org/978-3-906042-16-9.

Livera, A., Theristis, M., Makrides, G., Georghiou, G.E., 2019. Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems. Renew. Energy 133, 126–143. https://doi.org/10.1016/j.renene.2018.09.101.

Marion, B., Adelstein, J., Boyle, K., Hayden, H., Hammond, B., Fletcher, Canada, B., Narang, D., Kimber, A., Mitchell, L., Rich, G., Townsend, T., 2005. Performance parameters for grid-connected PV systems. Conf. Rec. IEEE Photovolt. Spec. Conf. 1601–1606. https://doi.org/10.1109/PVSC.2005.1488451.

Masson, G., Kaizuka, I., Lindahl, J., Jäger-Waldau, A., 2019. 2019 Snapshot of Global Photovoltaic Markets.

Olalla, C., Deline, C., Clement, D., Levron, Y., Rodriguez, M., Maksimovic, D., 2015. Performance of power-limited differential power processing architectures in mismatched PV systems. IEEE Trans. Power Electron. 30, 618–631. https://doi.org/10.1109/TPEL.2014.2312590.

Platon, R., Martel, J., Woodruff, N., Chau, T.Y., 2015. Online Fault detection in PV systems. IEEE Trans. Sustain. ENERGY 6.

Rodrigues, S., Ramos, H.G., Morgado-Dias, F., 2016. Machine learning PV system performance analyzer. 33rd Eur. Photovolt. Sol. Energy Conf. Exhib., 2892–2897.

Skomedal, Å., Øgaard, M.B., Selj, J., Haug, H., Marstein, E.S., 2019. General, robust, and scalable methods for string level monitoring in utility scale PV systems. In: 36th European Photovoltaic Solar Energy Conference and Exhibition.

Stegner, C., Dalsass, M., Luchscheider, P., Brabec, C.J., 2018. Monitoring and assessment of PV generation based on a combination of smart metering and thermographic measurement. Sol. Energy 163, 16–24. https://doi.org/10.1016/j.solener.2018.01.070.

Teubner, J., Buerhop, C., Pickel, T., Hauch, J., Camus, C., Brabec, C.J., 2019. Quantitative assessment of the power loss of silicon PV modules by IR thermography and its dependence on data-filtering criteria. Prog. Photovolt. Res. Appl. 27, 856–868. https://doi.org/10.1002/pip.3175.

Teubner, J., Kruse, I., Scheuerpflug, H., Buerhop-Lutz, C., Hauch, J., Camus, C., Brabec, C.J., 2017. Comparison of drone-based IR-imaging with module resolved monitoring power data. Energy Procedia 124, 560–566. https://doi.org/10.1016/j.egypro.2017.09.094.

Tsanakas, J.A., Ha, L., Buerhop, C., 2016. Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: A review of research and future challenges. Renew. Sustain. Energy Rev. 62, 695–709. https://doi.org/10.1016/j.rser.2016.04.079.

Ventura, C., Tina, G.M., 2016. Utility scale photovoltaic plant indices and models for on-line monitoring and fault detection purposes. Electr. Power Syst. Res. 136, 43–56. https://doi.org/10.1016/j.epsr.2016.02.006.

Whitaker, C.M., Townsend, T.U., Newmiller, J.D., Engineering, E., Ramon, S., King, D.L., Boyson, W.E., Kratochvil, J.A., Laboratories, S.N., 1997. Application and validation of a new Pv performance. In: Conf. Rec. Twenty Sixth IEEE Photovolt. Spec. Conf. – 1997, 1253–1256. https://doi.org/10.1109/PVSC.1997.654315.

Woyte, A., Richter, M., Reich, N., Green, M., Mau, S., Hassan, G.G., Beyer, H.G., 2014. Analytical monitoring of grid-connected photovoltaic systems. IEA-PVPS. https://doi.org/10.13140/2.1.1133.6481.

Zefri, Y., ElKettani, A., Sebari, I., Lamallam, S.A., 2018. Thermal infrared and visual inspection of photovoltaic installations by UAV photogrammetry—application case: Morocco. Drones 2, 41. https://doi.org/10.3390/drones2040041.