Photovoltaic Array Condition Monitoring Based on Online Regression of Performance Model

Spataru, Sergiu; Sera, Dezso; Kerekes, Tamas; Teodorescu, Remus

Published in:
Proceedings of the 39th IEEE Photovoltaic Specialists Conference, PVSC 2013

DOI (link to publication from Publisher):
10.1109/PVSC.2013.6744271

Publication date:
2013

Document Version
Early version, also known as pre-print

Photovoltaic Array Condition Monitoring Based on Online Regression of Performance Model

Sergiu Spataru, Dezso Sera, Tamas Kerekes and Remus Teodorescu
Aalborg University, Aalborg, 9220, Denmark

Abstract — Photovoltaic (PV) system performance can be degraded by a series of factors affecting the PV generator, such as partial shadows, soiling, increased series resistance and shunting of the cells. This concern has led to greater interest in improving PV system operation and availability through automatic supervision and condition monitoring of the PV system components, especially for small PV installations, where no specialized personnel is present at the site.

This work proposes a PV array condition monitoring system based on a PV array performance model. The system is parameterized online, using regression modeling, from PV array production, plane-of-array irradiance, and module temperature measurements, acquired during an initial learning phase of the system. After the model has been parameterized automatically, the condition monitoring system enters the normal operation phase, where the performance model is used to predict the power output of the PV array. Utilizing the predicted and measured PV array output power values, the condition monitoring system is able to detect power losses above 5%, occurring in the PV array.

Index Terms —condition monitoring, fault detection, performance model, photovoltaic systems, regression analysis.

I. INTRODUCTION

Photovoltaic (PV) system performance can be degraded by factors such as partial shadows, soiling, increased series resistance and shunting [1] as well as inverter or other balance-of-system component faults. These factors can lead to a significant amount of power loss, of up to 18% per year, as reported in [2]. This concern has led to greater interest in improving PV system operation and availability through automatic supervision and condition monitoring of the PV system components [3], especially for small PV installations, where no specialized personnel is present at the site.

Several methodologies for offline or online PV plant condition monitoring have been reported in the literature. In [4], statistical analysis methods are used to analyze energy production data from each inverter of the plant, and to determine if there are significant discrepancies between inverter energy production measurements. Such an approach can provide valuable insight in the operation of the plant, but does require a representative measurement data set for each inverter and extensive statistical analysis skills.

Some condition monitoring methods make use of empirical performance models of the PV system, together with sensor measurements, such as plane-of-array (POA) irradiance, module temperature [2], or satellite observations [5], and energy production data, to detect faults in the PV system. Usually these methods are employed to analyze large data sets of measurements, and are often time and resource intensive, thus being more suitable for offline condition monitoring of PV plants.

Other condition monitoring techniques are based on electrical models of the PV generator [3], or intelligent decision systems [6]. Some of these methods can achieve very good performance prediction and fault detection results, but often require extensive modeling and/or detailed tests of the PV modules/array, prior to system installation, which may not always be feasible for practical application.

Considering the existing and proposed PV system condition monitoring solutions, and the specific requirements of the PV installation (PV plant or residential PV system), a condition monitoring system should provide a good tradeoff between: costs; installation and operation complexity requirements; applicability to a wide range of PV technologies and configurations; and good fault detection/monitoring performance.

This work proposes a condition monitoring system based on the Sandia Array Performance Model (SAPM) [7]. The condition monitoring system is parameterized automatically/online during the operation of the PV system; using POA irradiance, module temperature, and PV array production/maximum power (MPP) measurements. The parameterization and optimization of the system is realized automatically using regression modeling methods.

The novelty of the proposed method in comparison with other model based condition monitoring systems is that the proposed system can makes use of the current-voltage (I-V) scanning capabilities of a new generation of commercial PV inverters [8]. Using the real MPP extracted from the I-V curve, together with ambient conditions sensors, the system is able to calculate online an accurate performance model of the PV array in question, during the field operation of the PV array/inverter.

Previous work [9, 10] proposed PV system fault detection methods based on I-V curve analysis, which can yield good detection results, but does require the inverter to have I-V scanning capabilities. This requirement is not mandatory for the current method; PV array production data, estimated by the maximum power point tracking (MPPT) of the inverter, can be used instead of the real MPP, which naturally leads to somewhat lower accuracy of the monitoring system.

The main advantages of the proposed condition monitoring system are: simple commissioning and operation requirements from the point of view of the installer/system operator, as well as a potential applicability to a wide range of PV technologies and systems configurations. The main requirements being

POA irradiance and module temperature sensors; optionally, I-V scan capable commercial inverters [8] can lead to increased accuracy of the condition monitoring system.

In the following sections, the main concept and theory of operation of the conditions monitoring system will be presented. The influence of measured data quality on the accuracy of the system will be analyzed, along with suggestions on how to improve the prediction.

Experimental results are presented for two study cases: one in which lower quality MPP data is used, estimated by the solar inverter tracking; and a second where higher quality data is used, extracted from the I-V characteristic of the PV array. In both cases the fault detection capabilities of the condition monitoring system are tested on a PV array affected by increased series resistance, or by partial shadows.

II. Condition Monitoring System

The proposed condition monitoring system is based on the well-known Sandia Array Performance Model [7], widely used in PV system design, analysis and troubleshooting. The SAPM utilizes a database of empirically derived PV module parameters developed by testing modules from a variety of manufacturers [7], and can predict PV array power and energy production with high accuracy [11, 12].

The inputs of the condition monitoring system are the *POA irradiance, PV module temperature* and the *maximum power point of the PV generator*. These system variables are monitored periodically using ambient sensor measurements, and I-V scans (to find the real MPP of the PV generator), or the MPPT of the solar inverter (which estimates the MPP).

Having these inputs available, the proposed model does not need to be parameterized by using a parameter database, or by performing tests on PV modules. The model parameters are obtained online through regression modeling, after the system has been installed, during a so-called *commissioning* or *learning phase*.

After the learning phase has been completed, and the condition monitoring model has been parameterized, the system enters the *normal operation phase*, when the power output of the PV array is monitored and compared with the predicted/expected output of the condition monitoring model.

A. *Inputs of the condition monitoring system*

The condition monitoring system periodically measures and monitors the POA irradiance, module temperature, and the MPP of the PV generator. The quality of the measured data has a significant influence on the prediction accuracy of the performance model [13], and thus on the condition monitoring system. Measurement errors and sensor calibration procedures represent an important issue for monitoring systems, and efforts to solve them are ongoing [13].

Another important factor affecting prediction accuracy is the collection procedure of the measurements, especially the PV array's MPP. Normally the PV output power is measured and it is assumed the MPPT has correctly located the MPP, as depicted in Fig. 1a. However, the MPPT may not track the real MPP precisely, especially in cloudy sky conditions, therefore the quality of the reported data can be low [2]. Furthermore the PV output power/production data is averaged over a period of time (one, five or a ten minutes period). This averaging of the process introduces further nonlinearities and inaccuracy in the performance model, as the averaging period increases.

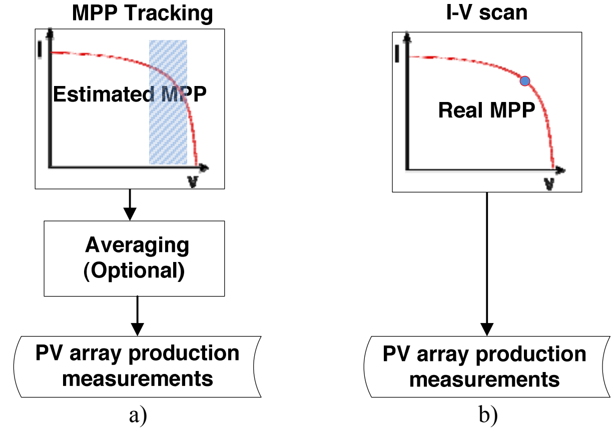

Fig. 1 PV array production acquisition: a) Maximum power point estimated by the inverter MPPT, and optionally averaged by the inverter or data logging system; b) Maximum power point extracted from the I-V characteristic.

In order to underline the effect that fast changing irradiance conditions can have on the MPPT estimated maximum power of the array, MPPT measurement data is analyzed for two representative days, one cloudy sky day and clear sky day, as presented in Fig. 2. The PV array dc power data was collected from a commercial inverter [8], connected to a PV array (eight BPMSX120 series connected multi-crystalline Si modules), averaged and sampled every minute.

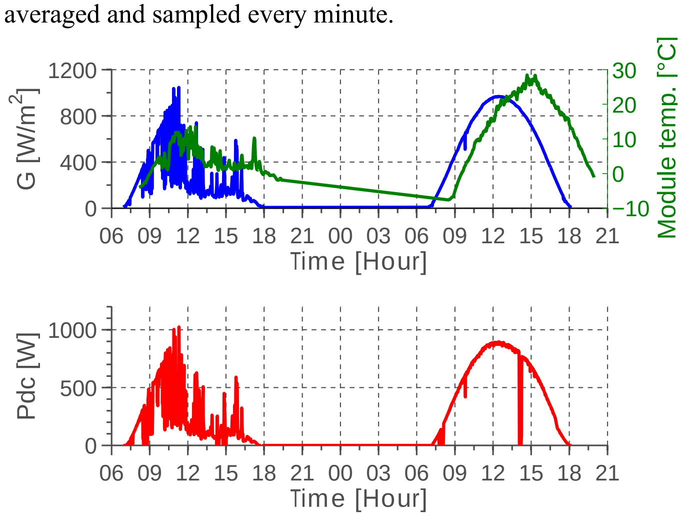

Fig. 2 DC power data of a mc-Si (8xBPMSX120) PV array, measured over two reference days (one cloudy sky day and one clear sky day). The MPP of the PV array is estimated by the MPPT of a commercial inverter, and averaged over a period of one minute.

The PV array dc power P_{dc} is plotted in Fig. 3 and Fig. 4, as a function of irradiance and module temperature, for the two reference days analyzed.

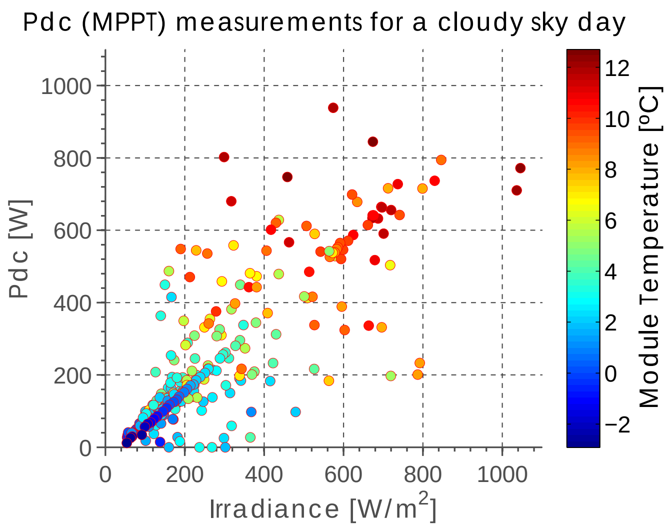

Fig. 3 PV array DC power estimated by the inverter MPPT during a cloudy sky day, averaged over a 1 minute period.

As can be observed from Fig. 3, due to the fast changing conditions, the MPPT is not able to track the MPP accurately. This type of MPPT production data (P_{dc}), if unfiltered, can cause the condition monitoring system to yield poor results.

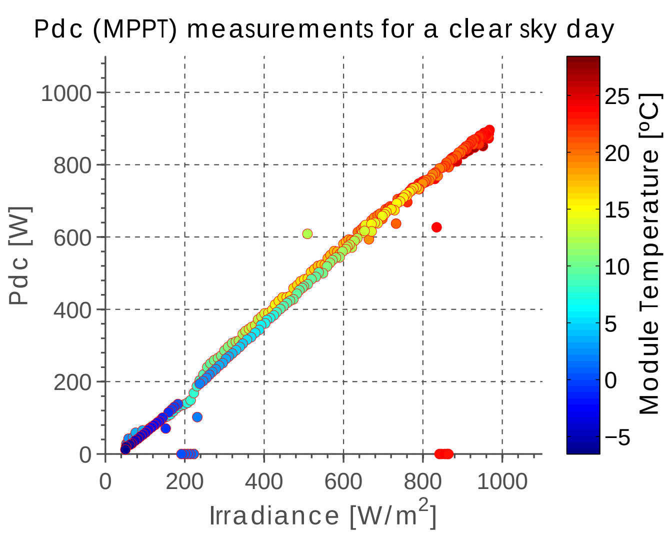

Fig. 4 PV array DC power estimated by the inverter MPPT during a clear sky day, averaged over a 1 minute period.

In clear sky conditions, the irradiance is changing much slower in comparison with the MPPT dynamics, such that the MPP is much better tracked by the MPPT, as can be observed from the P_{dc} measurements in Fig. 4, which have a quasi-linear dependence with the irradiance. The outliers in Fig. 4 are caused by starting/restarting of the solar inverter/tracking

A method to improve the prediction result is to filter out the low quality data, by filtering out the low irradiance (below 200 W/m^2) and low module temperature data (below 20 °C), which usually correlates with cloudy sky days.

Better quality MPP data can be obtained by extracting the real MPP from the I-V curve of the PV array, as depicted in Fig. 1b. New generation solar inverters [8] have already the capabilities to measure the I-V curve of each PV string [14].

B. *Learning phase*

The learning phase consists of the acquisition of *training data*, in the form of PV array MPP measurements for a range of operating conditions (irradiance and module temperature), provided the PV system is operating normally and is unaffected by faults. The condition can be met by pre-filtering the training data, and eliminating poor MPP estimates due to fast changing conditions, inverter/MPPT restarting, temporary partial shading, or even faulty PV system installation.

The pre-filtering of the training data can be achieved by roughly estimating the MPP of the array, using PV module datasheet values, as in (1), and eliminating measurements which differ by more than ±10% of the P_{dc_valid} rough estimate.

$$P_{dc\_valid} = P_{mp\_STC}\left(1 + k_p \Delta T\right) G \tag{1}$$

Where: P_{mp_STC} is the maximum power of the PV module in *Standard Testing Conditions* (STC - 1000 W/m^2, 25°C and air mass 1.5), as stated in the datasheet; T_c-cell temperature [K]; *G=plane-of-array irradiance/1000*; $\Delta T = T_c - 298.15$ [K], k_p MPP temperature coefficient [%/K].

Once sufficient data has been collected, a *PV array performance model candidate*, based on the SAPM, is parameterized according to the steps in Fig. 5, using regression modeling.

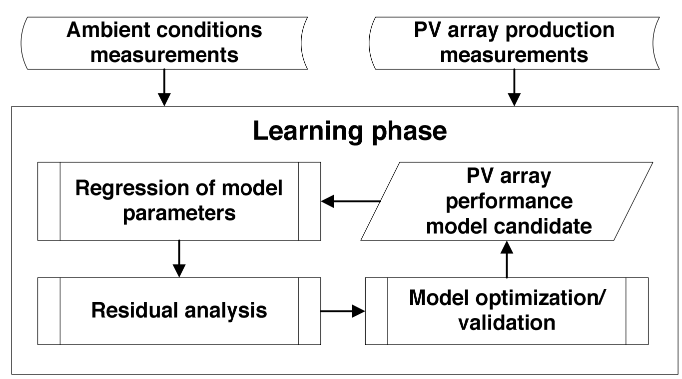

Fig. 5: Learning/commissioning phase of the condition monitoring system, where the PV permance model paremeters are optimized using regression modelling and nominal operation data.

A generic model candidate $\hat{P}_{mp}(G,T_c)$, derived from the SAPM, is proposed in (2), where the nonlinearities of the model have been separated in eight linear terms x_1 to x_7 denoted as *predictor variables* or *regressors* [15].

$$\begin{cases} x_1 = T_c \ln(G), x_2 = (T_c \ln(G))^2 \\ x_3 = \Delta T, x_4 = G, x_5 = G^2 \\ x_6 = G\Delta T, x_7 = G^2 \Delta T \end{cases} \tag{2}$$

Based on these variables, a *multiple linear regression model* can be formulated as in (3), where y is the *respons*e variable, in this case the PV array maximum power, p_i are the *regression coefficients* which have to be calculated, and ε represents random model error [15].

$$\hat{P}_{mp}(G,T_c) = y = \sum_{i=0}^{k} p_i x_i + \varepsilon, \text{where } x_0 = 1 \tag{3}$$

Using (2), (3) and the training data acquired during the learning phase, the regression coefficients can be calculated using the *least squares normal equation* presented in [15], or by a gradient descent algorithm. Once these parameters have been calculated, the PV array maximum power can be predicted using the proposed regression model.

In order to validate the new regression parameters of the model, several indicators are used in the learning phase, such as Root-Mean-Square Error (RMSE) and coefficient of determination R^2 [15]. The last step in the model validation process is to analyze the residuals resulted from the regression, and to check for trends that would suggest modeling or measurement errors. A detailed procedure for PV performance model validation is presented in [13].

If the model is validated, the condition monitoring system enters the *normal operation phase*; otherwise the model is optimized using a *stepwise regression* procedure, where terms from model candidate are added or removed successively [13].

The iterative optimization and the modification of the PV performance model candidate structure may be necessary for different types of PV technologies and system configurations, where other nonlinearities are more accentuated and are not characterized in the initial model candidate.

C. Normal operation phase

During the normal operation phase of the condition monitoring system, depicted in Fig. 6, PV array MPP, as well as the POA irradiance and module temperature measurements are acquired. The MPP measurements are either tracker by the MPPT of the inverter (Fig. 1a), or the real MPP is calculated from the I-V characteristic curve of the PV array (Fig. 1b), if the inverter has I-V tracing capabilities.

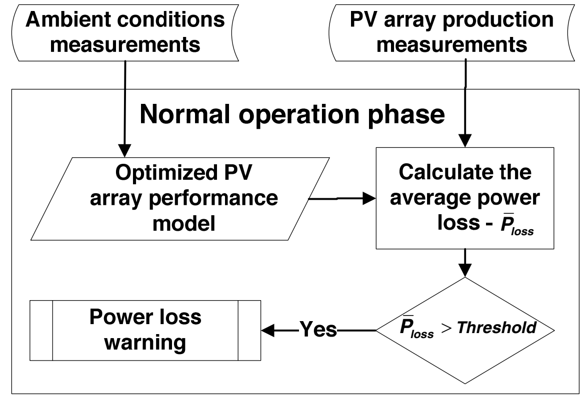

Fig. 6 Normal operation phase, when the PV array production data is monitored, together with the ambient conditions. The average power loss factor is calculated, and if it is larger than a certain threshold, a warning is issued to the plant operator to check the PV system.

The irradiance and temperature measurements are used as inputs to the optimized PV array performance model, to predict/estimate the maximum power output of the PV array, under these conditions.

The condition monitoring systems uses the predicted and measured MPP values to calculate the average power loss $\bar{P}_{loss}$ of the PV array, according to (4).

$$\bar{P}_{loss} = \frac{1}{N} \sum_{i=0}^{N-1} \frac{\hat{P}_{mp}[\mathrm{i}] - P_{mp}[\mathrm{i}]}{P_{mp}[\mathrm{i}]} \cdot 100 \tag{4}$$

If the average power loss is above a preset threshold value, a warning is issued to the plant operator. The threshold value is directly related to the prediction accuracy of the performance model, the quality of the model input data, and the uncertainty in the sensor measurements.

In a comprehensive study done by [12], five minute averaged MPP values, POA irradiance and module temperature, are used in conjunction with the SAPM, to predict the annual energy yield of eight modules of different technologies. For six of the eight modules (one c-Si, three mc-Si and two CIS modules), the annual energy production, predicted by the SAPM, agreed to within 5% of the measured values. The last two modules (tandem junction amorphous silicon) suffered significant degradation during this period, which led to poor predictions of the performance model.

Another model validation study [11] performed by National Institute for Standards and Technology and Sandia National Laboratory shows the SAPM model can predict power output to within 1% of measured power. In this study all the model parameters are accurately determined for the tested module samples, and high quality measurements are collected using precision instruments.

In the current paper, a power loss threshold of 5% has been determined to be optimal for the operation of the condition monitoring system, which is also in agreement with the reported SAPM prediction accuracy.

Furthermore, the power loss factor is averaged over a window of several samples (N=5 to 10 samples), in order to reduce the possibility of false alarms.

III. EXPERIMENTAL RESULTS

Two study cases have been considered for analyzing the proposed condition monitoring system operation, on a PV array consisting of eight multi-crystalline Si (BPMSX120) modules. The first study case considers MPP estimated by the MPPT (simply measuring the dc output power, assuming the MPPT keeps the operating point at MPP). The second case considers the MPP extracted from the I-V curve of the PV array, measured every ten minutes (for several days) by the same inverter.

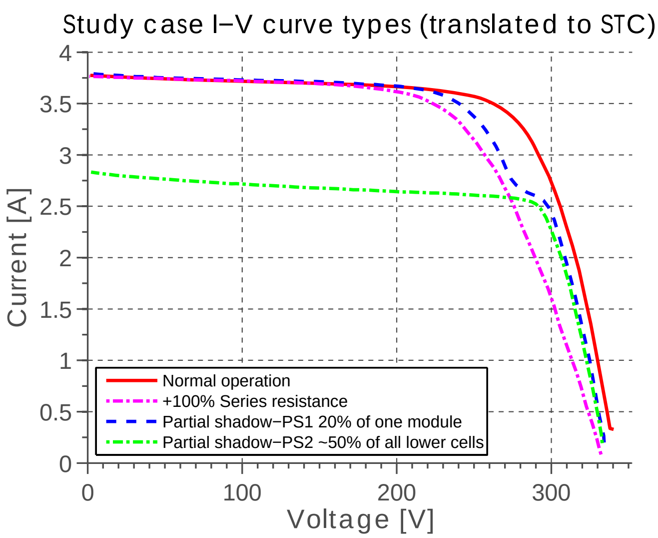

Fig. 7 I-V curves of a mc-Si PV array used in the two study cases to test the fault detection capability of the condition monitoring system. The PV array was either affected either by: partial shadows affecting the lower cells of the array (similar to dirt accumulation) – PS 2 or partial shadows affecting a few cells within one module – PS1; or by increased series resistance (+ 100%).

In order to evaluate the accuracy of the regression model, as well as to test the capability of the condition monitoring system to detect power losses in the PV array, several scenarios of partial shading and increased series resistance of the PV array, have been designed and measured, for the two study cases. Representative I-V curves of each scenario, translated to STC, are presented in Fig. 7.

A. Condition monitoring based on inverter MPPT/PV array dc output power

In the first study case, several days of one-minute averaged PV array dc output power measurements are collected by a commercial inverter [8], as presented in Fig. 1. The dc output power is acquired using the inverter's data logging capabilities, whilst the POA irradiance and module temperature are measured with a digital silicon irradiance sensor (Si-RS485-TC-T).

Since the measurements contain both clear sky day and cloudy sky day data, they are filtered using, (1) before parameterizing the condition monitoring system, as described in the previous section.

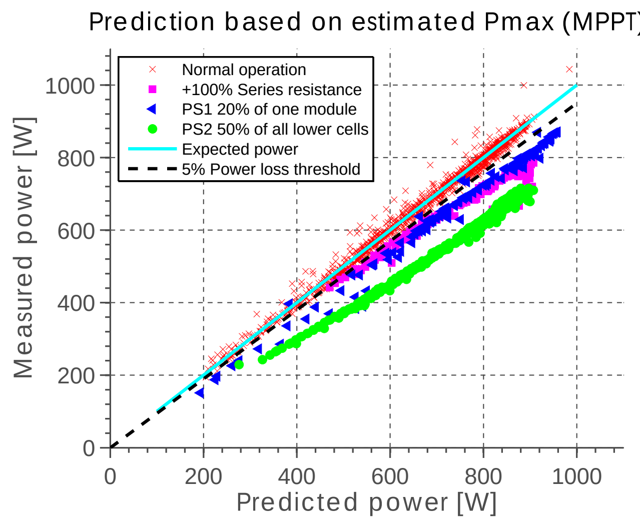

Fig. 8 Predicted vs. measured power for the case when the performance model is trained with MPPT 1 minute averaged data, measured with a commercial inverter. Regression of performance model resulted in $RMSE = 6.21W$ and a coefficient of determination $R^2 = 0.998$.

As can be observed from the results in Fig. 8, the normal operation predicted dc output power/ $\hat{P}_{mp}(G,T_c)$ is below the 5% power loss threshold. Some dc output power estimates (outliers) are above the 5% threshold, which is to be expected when using MPPT estimated and averaged MPP data. False alarms due to outliers can be avoided by averaging the power loss factor, over several samples, as in (4). Still, power losses of over 5%, due to increased series resistance of the PV system (+100%), or due to shading (such as PS1 or PS2B), can be detected by the condition monitoring system.

B. Condition monitoring based on inverter - measured I-V characteristics

In the second study case the MPP data is extracted from the I-V characteristic of the same PV array, measured periodically (every 10 minutes) by the same PV inverter, during the same period as in the previous case. The irradiance and temperature are measured using the same sensors as in the previous case.

In the learning phase, the performance model candidate is parameterized using I-V curves of the array recorded at various irradiances and temperatures, in normal operation conditions, over a period of several days. A very accurate prediction of the maximum power can be achieved in this case, as can be observed from the normal operation data in Fig. 9.

Since the I-V curves were acquired in identical conditions with the previous case (same PV array, inverter, ambient sensors, same partial shading, and during the same period), a comparison can be made between MPPT/dc output power based condition monitoring (Fig. 8) and real MPP/I-V curve based condition monitoring (Fig. 9). As can be observed in Fig. 9, the real MPP extracted from the I-V curve makes it possible to detect even relatively small power losses.

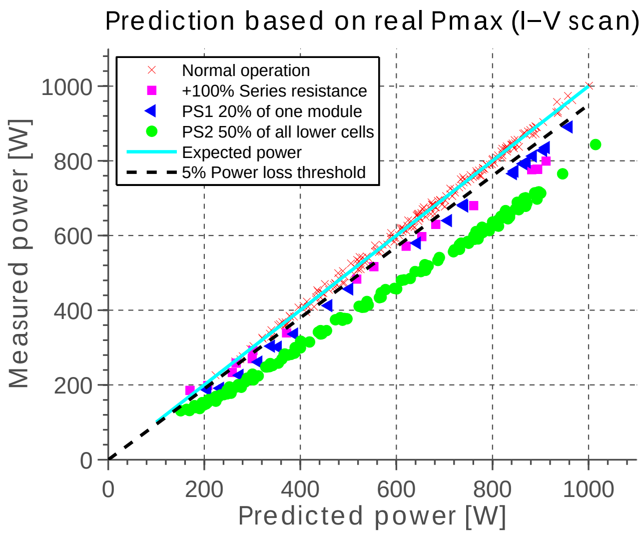

Fig. 9 Predicted vs. measured power for the case when the performance model is trained with data extracted from I-V curves, measured with a commercial inverter. Regression of performance model resulted in $RMSE = 4.73W$ and a coefficient of determination $R^2 = 1$.

IV. Conclusions

A photovoltaic array condition monitoring system based on SAPM has been proposed, which can be parameterized online, using regression methods. The main advantage of the proposed system is that it does not require modeling and testing of the PV modules/arrays before installing the system. But it has an initial commissioning phase, when ambient conditions and PV array MPP measurements are used to automatically parameterize the condition monitoring systems by regression modeling.

The results show that the method has good sensitivity to deviations from normal operation; it can detect average power losses above 5%. More accurate results can be achieved if real MPP data is extracted from the I-V characteristic of the PV array.

V. References

[1] E. L. Meyer, and E. E. van Dyk, "Assessing the reliability and degradation of photovoltaic module performance parameters," *IEEE Transactions on Reliability,* vol. 53, no. 1, pp. 83-92, Mar, 2004.

[2] S. K. Firth, K. J. Lomas, and S. J. Rees, "A simple model of PV system performance and its use in fault detection," *Solar Energy,* vol. 84, no. 4, pp. 624-635, 2010.

[3] A. Chouder, and S. Silvestre, "Automatic supervision and fault detection of PV systems based on power losses analysis," *Energy Conversion and Management,* vol. 51, no. 10, pp. 1929-1937, 2010.

[4] S. Vergura, G. Acciani, V. Amoruso, G. E. Patrono, and F. Vacca, "Descriptive and Inferential Statistics for Supervising and Monitoring the Operation of PV Plants," *IEEE Transactions on Industrial Electronics,* vol. 56, no. 11, pp. 4456-4464, Nov, 2009.

[5] A. Drews, A. C. de Keizer, H. G. Beyer, E. Lorenz, J. Betcke, W. G. J. H. M. van Sark, W. Heydenreich, E. Wiemken, S. Stettler, P. Toggweiler, S. Bofinger, M. Schneider, G. Heilscher, and D. Heinemann, "Monitoring and remote failure detection of grid-connected PV systems based on satellite observations," *Solar Energy,* vol. 81, no. 4, pp. 548-564, 4//, 2007.

[6] P. Ducange, M. Fazzolari, B. Lazzerini, and F. Marcelloni, "An intelligent system for detecting faults in photovoltaic fields," in Intelligent Systems Design and Applications (ISDA), 2011 11th International Conference on, 2011, pp. 1341-1346.

[7] D. L. King, W. E. Boyson, and J. A. Kratochvil, *Photovoltaic Array Performance Model*, Sandia National Laboratories, Albuquerque, New Mexico 87185-0752, 2004.

[8] Danfoss, *TLX Reference Manual,* L00410320-07_02, Denmark, 2012.

[9] D. Sera, S. Spataru, L. Mathe, T. Kerekes, and R. Teodorescu, "Sensorless PV Array Diagnostic Method for Residential PV Systems," in 26th European Photovoltaic Solar Energy Conference and Exhibition, Hamburg, Germany, 2011, pp. 3776 - 3782.

[10] S. Spataru, D. Sera, T. Kerekes, and R. Teodorescu, "Detection of increased series losses in PV arrays using Fuzzy Inference Systems," in Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE, Austin, Texas, 2012, pp. 000464-000469.

[11] A. H. Fanney, M. W. Davis, B. P. Dougherty, D. L. King, W. E. Boyson, and J. A. Kratochvil, "Comparison of photovoltaic module performance measurements," *Cell,* vol. 125125, no. 150150, pp. 119340, 2006.

[12] A. H. Fanney, B. P. Dougherty, and M. W. Davis, "Comparison of predicted to measured photovoltaic module performance," *Journal of Solar Energy Engineering,* vol. 131, no. 2, pp. 21011, 2009.

[13] J. S. Stein, C. P. Cameron, B. Bourne, A. Kimber, J. Posbic, and T. Jester, "A Standardized Approach to PV System Performance Model Validation," in 35th IEEE Photovoltaic Specialists Conference, Honolulu, Hawaii, 2010, pp. 1079-1084.

[14] S. B. Kjær, O. Oprea, and U. Borup, "Adaptive Sweep for PV Applications," in 26th European Photovoltaic Solar Energy Conference and Exhibition, Hamburg, Germany, 2011, pp. 3708 - 3710.

[15] D. C. Mongomery, *Design and Analysis of Experiments*, 8th Edition ed.: Hoboken, N.J. : John Wiley & Sons 2009.