14th IMEKO TC10 Workshop Technical Diagnostics
New Perspectives in Measurements, Tools and Techniques
for system's reliability, maintainability and safety
Milan, Italy, June 27-28, 2016

A Nonlinear Predictor for the Supervision of Photovoltaic Strings Performances

Giacomo Leone[1], Loredana Cristaldi[1]

[1] *Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy, {giacomo.leone, loredana.cristaldi}@polimi.it*

***Abstract* –In this paper, a nonlinear predictor of the electrical power produced by a PV string is proposed. The first phase of the approach is the training of the predictor, during which four characteristic parameters are determined. Such coefficients are representative of the string under study and define its electrical signature (identikit). Once trained the model, when new monitoring data are available, the mismatch between the forecasted and measured electrical power can be assumed as a reliable marker of the performances of the string, since the greater the mismatch, the worse the string efficiency. The analysis of the forecasting error, therefore, enables the detection of losses of energy production. In particular, a strength of the proposed approach is the possibility to distinguish the losses due to aging phenomena from the losses due to the dust or dirt accumulation. The method has been tested and validated for a real case study and the obtained results are presented in the paper.**

I. INTRODUCTION

Photovoltaic (PV) solar technology constitutes an important source of electrical energy, covering an increasing percentage of the global electricity demand, so that the supervision of the efficiency of PV plants has become a hot topic in the research community. The factors that can lead to a reduction of the PV modules are manifold. A well know problem is the dust accumulation on the modules surface [1]. Dust, indeed, blocks incident photons, which are not able to reach PV cells, and might determine a non-negligible effect on the incomes. Obviously, this issue is more relevant in arid, semi-arid and desert regions. The dust, or in general dirt, accumulation, indeed, is highly influenced by the environmental and weather conditions of the area in which the panels are installed, as demonstrated in [2]. The authors, in fact, among a wide range of climatic parameters have found that high humidity, rain and snow have significant effects to the efficiencies of the PV installations. Another leading factor in the performances degradation of PV panels is the aging [3-4]. In [3], the variation of PV modules performances after 11 years of exposure in a cool and marine environment is taken into account. The authors found a little change in the average value of open circuit voltage V_{oc} and an average decrease of 6.38% for the short-circuit current I_{sc}, leading to an average 4.39% decrease in the maximum power point P_{max}.

The most relevant effect of all the cited problematics that might occur in a PV plant is the reduction of its performances in terms of energy production. It follows that the evaluation of such losses is a crucial task for a correct management and maintenance schedule of the plant. In order to achieve such objectives the definition of models or electrical signatures that describe the plant production in optimal conditions and at the beginning of its operative life is required. The differences between the actual production and the reference value provided by such models, indeed, is an effective marker of the eventual performance decrease. In the literature, many proposals are available. The authors in [5-6], for instance, propose methodologies for the monitoring of the energy performances of the PV plant. In particular, the work in [6] deals with the monitoring of the energy performance of PV fields by means of low cost hardware. It proposes a methodology based on inferential tools, which return information about the correct operation of the PV field. The methodology needs an initial training that allows defining one or more reference strings, which are used as benchmarks for future comparisons. Other previous works focus on the forecast of the maximum power generated by PV plants. The authors in [7] have implemented a model-based MPPT tracker that allows forecasting the producible energy in condition by a clean and not degraded panel. In [8], a forecasting approach based on artificial intelligence techniques, namely a Neural Network architecture, is proposed. On the other hand, other authors have proposed circuits for simulating PV arrays [9-10]. Their main drawbacks, however, is that they rely on simplified equations and/or require many computational efforts.

In this paper a nonlinear predictor for the supervision of fielded photovoltaic strings performances is proposed. The independent variables used for the forecast of the producible power in optimal conditions are functions of only environmental parameters, namely solar radiation and module temperature, so that the proposed methodology is suitable also for medium-sized PV plant equipped with low cost and standard condition monitoring system, usually not designed for diagnostic purposes. The paper is structured in the following way: in Section II, the proposed method is introduced, whereas Section III describes the validation procedure adopted for a real case application. In Section IV, the obtained results are reported, finally the contribution is closed with the conclusions.

II. PROPOSED METHOD

As already mentioned in Section I, the energy production of a PV plant is not always optimal over time since many different phenomena determining a decrease of its performance may occur. In particular, dust accumulation, aging process and degradation of the cells have been highlighted as the most impacting factors. These considerations justify the need for a forecasting model able to estimate the producible power by a fielded well-operating plant in each environmental conditions, namely for each given solar radiation and modules temperature. The proposed algorithm gets ideas from the method described in [11]. In that work, the authors propose a model for the estimation of the produced power P_{mmp} by a PV module for a given solar irradiance value G and cells temperature T_c, when operating at its Maximum Power Point (MPP). In particular, the relation that links the environmental quantities with the electrical power production is the following:

$$P_{mpp} = C_0 G + C_1 G T_c + C_2 G \ln(G) + C_3 G \ln^2(G) \quad (1)$$

The methodology allows the definition of a PV reference PV panel, by means of the four parameters C_0, C_1, C_2 and C_3. Their estimation is performed at the PV plant installation time, which corresponds to the optimal reference condition characterized by no aging and clean surface. The comparison of the actual energy production with the estimated one allows the evaluation of the production decrease and the planning of the most adequate maintenance activity.

The methodology presented in this contribution is based on the same aforementioned equation (1), but with the substantial difference that the target output is the power produced by a PV string and not by only one module. What makes interesting and challenging our proposal are the different operational constraints that characterize the case study analyzed in [11], which is a single PV module, and the one here considered, a PV string. The first important difference regards the uncertainty on the environmental conditions data. In experiments like the one described in [11] the measurement setup adopted for the acquisition of temperature and solar radiation values is physically close to the equipment under test, so that the confidence in assuming that the measured data correspond to the real conditions is high. Conversely, this is not usually the case for fielded PV plants, where the measurements of temperature and solar radiation are often performed in a single point of the plant, so that the assumption that the acquired data can be taken as reference for the entire plant is weaker. This is even more evident in extended plants, where phenomena like passenger clouds can cause the shading of part of the plant, so that the acquisition of environmental data in only one point gives a poor contribution. Another critical difference between the two considered application cases regards the tracking of the MPP. In the reference work [11], indeed, for each sample of temperature and solar radiation, the complete V-I curve of the panel have been acquired thanks to a proper control of the electronic load, so that the actual related maximum power value could be stored. This process is quite hard to repeat for fielded PV plants, especially for those equipped with low cost monitoring systems. In such situations, indeed, the data of delivered electrical power for given solar radiation value and modules temperature not always represents the maximum power producible by the strings in those conditions. In the usual configuration of medium-sized PV plants, the central inverter architecture, in fact, several strings are connected to a unique converter that sets a unique MPP for the entire plant. It is evident that such solution represents a trade-off between plant costs and plant productivity, since the inverter imposes a global optimal operating point, that, however, may differ from the best operating point of the single strings. Furthermore, in grid-connected plants, the grid itself sometimes limits the plant productivity according to the energy demand. An ulterior source of power losses to take into account when considering PV strings is related to the mismatch of the I-V characteristics of the different modules, due to the manufacturing processes. Such losses increases with the number of modules connected to the same Maximum Power Point Tracking (MPPT) converter, so that passing from a central inverter architecture to a string inverter solution may help to limit the efficiency decrease related to such issues [12].

It is evident from the previous reasoning that the application of the model described by the Equation (1) for a case study like a PV string implies the introduction of some approximations that do not take place when the model is applied for the performances prediction of a single module. It follows that the objective of this paper is to investigate about the possibility to apply the described model for PV strings. The discriminating factor will be the error introduced by the model in the forecast of the producible power by the strings. An estimation error restrained in a small range, indeed, would allow the use of the predictor for the identification of performance losses due to different processes such as dust or dirt accumulation, aging and cells degradation.

III. PREDICTOR VALIDATION

The presented algorithm is based on the equation (1) that correlates the maximum delivered power by a string in determined conditions of solar radiation G and cell temperature T_c. The model is very simple to be applied and the main computational effort is required for the estimation of the four coefficients C_0, C_1, C_2 and C_3. Such task, however, is not resources demanding as it can be carried out through an ordinary least square estimation, being the model linear in the coefficients. This distinguishes our proposal with respect to other models, as [9-10], that are much more complicated and require a not negligible computational effort.

The application case considered in this paper regards a PV plant located in South Africa. As for the technical characteristics of the considered plant, each single string is composed by 24 modules, characterized by an open circuit voltage V_{oc}=36.9V, a short circuit current I_{sc}=8.47A and a peak rated power of 235 W (5.64 kW of peak power for the string). The strings are put in parallel in group of 8, by means of a string box. Finally, 17 groups of 8 paralleled strings are connected to a unique inverter with maximum DC input power equal to 802 kW. The entire plant is constituted by the repetition of the described configuration. Since the data are confidential, however, the ratings of the overall plant are not reported. For what concerns the information about the environmental conditions of the plant, they are measured in correspondence of a single module. In particular, the solar radiation is measured through a pyrometer at the same orientation of the module. Considering the huge extent of the PV plant under study, it is evident, as aforementioned in Section II, how the knowledge of the environmental conditions for application cases as the one here presented is approximate with respect to the experiment carried out in [11]. The acquired data refer to the period between the 1st January and the 31th December 2105. The data gathering period is much more extent than in [11], where it was of 1 month (sampling period of 60s), so that a larger set of environmental conditions can be considered in the analysis. In particular, the small sampling time of the monitoring system, that is 60s, has allowed the acquisition of 34000 data of solar radiation, module temperature and string delivered power. More in detail, the gathered solar radiation values data are distributed in an interval between 400 W/m^2 and 1550 W/m^2.

The model proposed in (1) is heavily dependent on the solar radiation G, as it appears in all the four predictor variables. It follows that in the validation of the model, the analysis was focused in the composition of the training and test sets. In particular, it has been chosen to investigate about the effect that the different distribution of the values of G in the training set has on the prediction error, in order to check if there is evidence of correlation between such quantities. A further relevant quantity of interest is the variance of the coefficients in consequence of a change of the training set.

The possible combinations of the radiation values in the training set are infinite, so that some constraints have been imposed in order to limit the combination space.

First, the radiation values have been grouped in N_c=4 different classes. The upper limit values of the classes have been selected in correspondence of the 25%, 50%, 75% and 100% quantile of the radiation dataset in order to make the classes composed by the same number of samples. The final result of such classification is reported in Table 1:

Table 1. Solar radiation values classification

Class Index	Lower solar radiation limit [W/m^2]	Upper solar radiation limit [W/m^2]
Class 1	400.00	691.10
Class 2	691.10	875.40
Class 3	875.40	1054.60
Class 4	1054.60	1560.00

At this point, the next steps are carried out as follows:

1. Set the desired number of samples N_{tr} composing the training set.
2. Define the vector $p=\{p_1,p_2,p_3,p_4\}$, where the generic value p_i corresponds to the percentage of radiation values in the training set that belong to the *i-th* class of Table 1. In other words, p describes one of the possible strategies according to which extracting values of G from each class. In the definition of the vector p, the following constraints are applied:
   - p_i can range from p_{min}=1/12 to p_{max}=3/4, with an admissible incremental step equal to p_{step}=1/12.
   - The sum of the values composing p is equal to 1.

   The values p_{min}, p_{max}, p_{step} and N_c have been chosen in order to limit the otherwise infinite number of possible combinations of the solar radiation values in the training set. In particular, applying the reported values, N_{comb}=165 possible different combinations for the definition of p exist.
3. Define a training set, taking care that the percentage of radiation samples belonging to each class corresponds to the values defined in step 2.
4. Estimate the coefficients C_0, C_1, C_2 and C_3.
5. Define a test set of length N_{test}=9200 (corresponding to about 100 days of acquisition). The only constraint in the definition of the set is that the radiation samples are uniformly distributed among the different classes, that is N_{test}/4 radiation samples for each class are selected. In this way, it is ensured that the evaluation of the prediction error is not biased by the values of G in the test set.
6. Apply the equation (1) to the test set and compute the percentage prediction error, defined as:

$$\varepsilon_j = 100 \cdot \frac{P_j - P_{j,est}}{P_j} \qquad (2)$$

being P_j the j-th observation of power in the test set and $P_{j,est}$ the corresponding estimation.

7. Repeat steps 3-6 N_t=1000 times, defining each time a new training set (the values in the vector *p* do not change) and test set, in order to check the sensitivity of the model with respect to their composition. Once defined *p*, indeed, the possible sets of G values extractable from each class are infinite. Setting a high value of N_t allows to explore a statistically significant region of the combinations space.
8. Repeat steps 2-7 defining a new vector *p*, selecting one of the possible N_{comb} combinations.
9. Repeat steps 1-8 varying N_{tr} from 1200 (about 13 days of acquisition) to 6000 (approximately 60 days), with incremental steps of 1200 (i.e. N_l=5 different possible lengths of the training set are taken into account) in order to investigate about the model sensitivity with respect to the length of the training set.

At the end of the described procedure, for each possible combination of the training set c_j, with j=1,2,…, N_{comb}, and desired length of the training set N_{tr}, N_t different estimations for the four coefficient C_0, C_1, C_2 and C_3 and for the prediction error ε are available.

The final objective is to check if the performances of the model are whether influenced or not by the distribution of the G in the training set, as well as from the number of samples composing it.

In particular, the performances of the model are evaluated taking into account the following parameters:

- ε_m, the absolute value of the median percentage prediction error. Such quantity is considered more robust than the mean value to the outliers;
- *w*, the width of the 95% confidence interval (CI) for the prediction error: it is a good indicator of the extent of the error range;
- σ_k, the variation of the coefficient C_k, k=0,1,2,3, in consequence of a change of the training set.

In the next Section, the results obtained for the considered case study are reported and commented.

IV. RESULTS

This Section is divided in two parts: in sub-section A, the description and the comment of the result obtained applying the procedure proposed in Section III are reported. Sub-section B, instead, proposes a methodology for an effective definition of the training set in order to maximize the prediction performances.

A. Results description

The first result of interest deriving from the application of the procedure described in Section III is that it is possible to associate to each of the possible combinations c_j, j=1,..,N_{comb}, of the solar radiation training set two parameters, R_{med} and R_{skew} , corresponding to the median value and skewness value, respectively. Such values are peculiar parameters of the various trainings sets and, as expected, are independent on the number of samples composing it.

In the following Fig.1, the results obtained about the median value ε_m of the percentage prediction error in correspondence of different training combinations and training lengths are shown.

The values in the abscissa correspond to the values of R_{med} related to each combination. It is worth to remind that each combination has been tested, for a given training set length, N_t=1000 times, so that the values reported in the y-axis represents the median value of the obtained results. In particular, each curve is related to a different length N_{tr} of the training set.

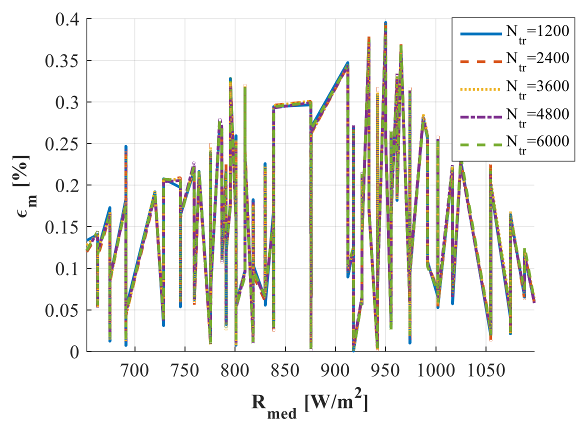

Fig. 1. Absolute value of the median percentage prediction error ε_m as function of the median solar radiation value R_{med} in the training set.

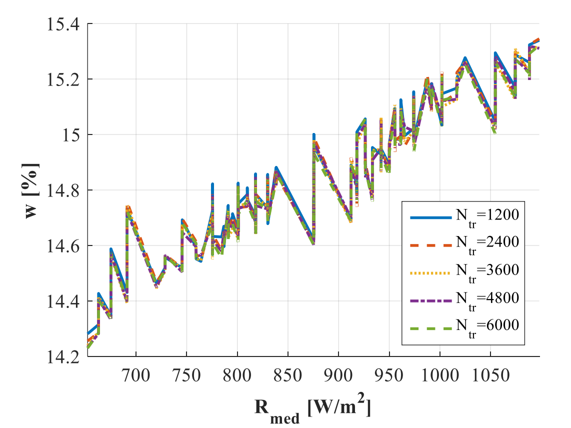

Fig. 2. Width of the 95% Confidence Interval for the percentage prediction error as function of the median solar radiation value R_{med} in the training set.

From the observation of Fig.1 it is possible to conclude that there is no evidence about correlation between ε_m and R_{med}. The median value, however, is always restrained in the range [0, 0.4]%. Further, the prediction results seem to be not influenced by the length of the training set, as the different curves almost completely overlap.

Fig.2, instead, reports in an analogous way the results achieved for what regards the width w of the 95% CI of the prediction error. In this case, a very interesting correlation between the CI width and R_{med} is highlighted. Moreover, also in this case an independence of the prediction results with respect to N_{tr} is evident

As already said, another interesting parameter to consider for the evaluation of the proposed model performances is the standard deviation of the coefficients C_k, k=0,1,2,3, obtained testing N_t times a specific combination for a given training set length N_{tr}. This information, indeed, is an indicator of the stability of the model with respect to the training set. We refer to this quantity as σ_k, where k is the index of the coefficient. In particular, in the following Fig.3, the values obtained for the quantity σ_0, related to the coefficient C_0, are reported in per unit with respect to the coefficient mean value (absolute value of such quantity if the coefficient is negative).

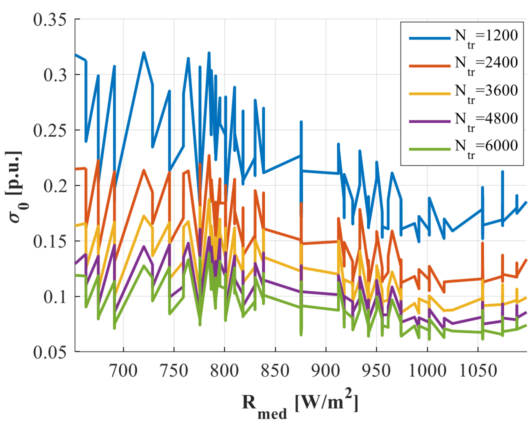

Fig. 3. Standard deviation of the model coefficient C_0 as function of the median solar radiation value R_{med} in the training set.

The proposed results show a slight dependence of σ_0 on R_{med}. In particular, the values of σ_0 decreases as R_{med} increases. The same trend is replicated for each value of training length N_{tr}, even if in correspondence of different values as the dispersion of the coefficient estimation decreases as the training sample length increases. As for the values of σ_k related to the other coefficients (C_1, C_2 and C_3), they are not reported because of the limited space of the paper. They show, however, a trend similar to the one of σ_0, except for the coefficient C_1 that seems to be more stable with respect to the training set composition.

The following Table 2, reports the correlation coefficients between the median radiation value in the training set R_{med} and the quantities ε_m, w and σ_k obtained for N_{tr}=1200. The value for the remaining values of N_{tr} are not reported for the sake of the space, as they are quite similar:

Table 2. Correlation coefficient between R_{med} and ε_m, w and σ_k (k ranging from 0 to 3)

ε_m	w	σ_0	σ_1	σ_2	σ_3
-0.3936	0.9305	-0.6853	-0.3708	-0.6076	-0.6115

The required best composition of the training set is the one that guarantees the minimization of these three quantities of interests. In this regard, the results of Fig.2 and Fig.3 are contrasting since as R_{med} increases the CI width w of the error diminishes. Conversely, the deviation of the coefficients increases. Therefore, a selection criterion that takes into account all these aspects is required. In the next sub-section the criteria adopted in this paper is described.

B. Results elaboration

In the previous sub-section, it has been observed that the trend of the quantities ε_m, w and σ_k as function of R_{med} is independent on the training length N_{tr}. Let us denote with $\varepsilon_m^{j,N_{tr,n}}$, $w^{j,N_{tr,n}}$ and $\sigma_k^{j,N_{tr,n}}$ (k from 0 to 3) the values of the quantities ε_m, w and σ_k, respectively, obtained for the combination j (j=1,2,…,N_{comb}=165) in correspondence of a given training set length $N_{tr,n}$ (n=1,2,..,N_l). Then, a synthetic estimator of the performances of the combination j for the median prediction error ε_m can be defined as:

$$\varepsilon_m^j = \frac{1}{N_l}\sum_{n=1}^{N_l}\varepsilon_m^{j,N_{tr,n}} \qquad (3)$$

that is the mean value of the values obtained taking into account different training lengths. The same equation can be applied in analogous way to the other two quantities of interest, so that w_j and σ_k^j (k=0,1,2,3) are obtained. At this point, for each combination j it is possible to associate a score coefficient s_j defined as:

$$s_j = \sqrt[6]{\left(\frac{\varepsilon_m^j}{\max_j(\varepsilon_m^j)}\right)^2 + \left(\frac{w^j}{\max_j(w^j)}\right)^2 + \frac{1}{4}\cdot\sum_{k=0}^{3}\left(\frac{\sigma_k^j}{\max_j(\sigma_k^j)}\right)^2} \qquad (4)$$

The first term under the radical sign is the square of the value of ε_m for the j-th combination, normalized with respect to the maximum value of ε_m among all the combinations. The same approach is applied for the rest of the quantities; it follows that Equation (4) is very similar to the equation that describes the distance from the origin of a point in a six-dimensional space. In this case, the six variables range from 0 to 1 and provide information of the performance of the model about the median error, the width of the error CI and the dispersion of the four characteristic coefficients. In particular, the values related to the coefficient dispersion σ_k are divided by 4 (the number of coefficients) in order to avoid making the model performance evaluation too much biased from the information about their distribution, as they represent 4 out of 6 terms in (4). Finally, the best combination is the one characterized by the smallest score coefficient s_j as it is the combination that guarantees the best trade-off in the minimization of the quantities ε_m, w and σ_k (k ranging from 0 to 4).

A for the considered application case, the best configuration found applying the proposed approach is characterized by the following vector p_{opt} of training combination:

$$p_{opt} = \{0.1667, 0.2500, 0.2500, 0.3333\} \qquad (5)$$

It can be observed that the composition of p_{opt} is such that the values of G in the resulting training sets are almost uniformly distributed among the different classes defined in Table 1, except for a slight bias towards the high radiation values. In particular, the resulting solar radiation median value is equal to 941.82 W/m^2. Such result is reasonable since, as observable in Fig. 1, Fig. 2 and Fig. 3 , this radiation value guarantees a good trade-off represented by a slightly wider 95% CI of the prediction error but lower value of median error and especially a limited dispersion of the characteristic coefficients. In Table 3, the values of the coefficients and the related standard deviation obtained for different values of N_{tr}, by means of extractions of radiation values as indicated by (5), are reported.

Table 3. Mean value and standard deviation for the model coefficients obtained with the selected training set

N_{tr}	C_0 [m^2]		C_1 [10^{-2}·m^2/°C]		C_2 [m^2]		C_3 [m^2]	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
1200	-33.52	5.71	1.56	0.07	12.33	1.69	-0.97	0.13
2400	-33.25	4.18	1.56	0.05	12.24	1.24	-0.97	0.09
3600	-33.25	3.27	1.55	0.04	12.26	0.97	-0.97	0.07
4800	-33.30	2.85	1.55	0.03	12.27	0.84	-0.97	0.06
6000	-33.39	2.45	1.55	0.03	12.30	0.73	-0.97	0.05

The results highlight how the mean value of the coefficients is independent on the training set length. At the same time, larger training set guarantee a lower standard deviation.

In Fig. 4 the related prediction error is shown. It can be observed that the largest error values lie in correspondence of lower radiation values. A valid motivation is the not negligible error affecting the pyrometers measurements al low solar irradiance conditions. It is obvious that a relevant error in the solar radiation measurement has a catastrophic effect in the power prediction, as it is the most important parameter of the model.

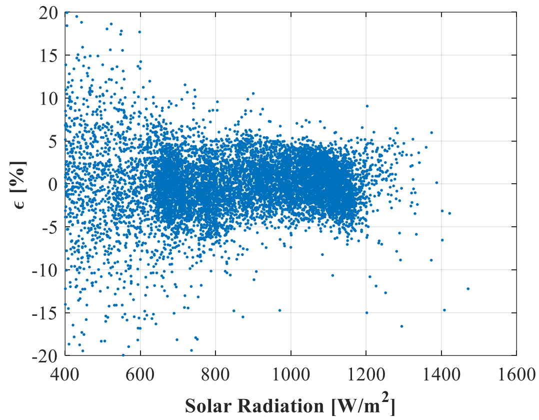

Fig. 4. Percentage prediction error associated to the selected training set as function of the solar radiation values in the test set

Table 4 below reports the 95% CI for the prediction error when in the test set radiation values smaller than 600 W/m2 are whether considered or not. In particular, it can be observed that the exclusion of such values allows to reduce efficiently the error range. As for the median error, it is practically the same since a variation of the 0.04% is found. In both cases, it can be stated that the prediction error is centered at the zero value so that it is plausible that a performance decrease of the string would polarize the error towards the negative values (see Equation (2)).

Table 4. 95% Confidence Interval for the percentage prediction error associated to the selected training set

R_{min} [W/m^2]	Lower 95% CI Limit [%]	Upper 95% CI Limit [%]	ε_m [%]	w [%]
400	-8.13	6.74	-0.09	14.86
600	-6.31	5.15	-0.13	11.46

The results reported in Table 3 and Table 4 are indicator of good performances of the model and justify its exploitation for application case like PV strings. Once estimated the model parameters at the PV plant installation time (reference condition), the comparison of the actual energy with the estimated one allows the evaluation of the production reduction and the planning of the most adequate maintenance activities. After having cleaned the string, a comparison between the actual and estimated production for a period of time short enough to consider negligible the presence of dust on the surface allows to evaluate the performance reduction due exclusively to the aging process. At this point, the model parameters redefinition represents the new reference point for the evaluation of the dust effects. The trend of the model parameters along time is a clear indicator of the panel degradation due to the aging.

V. CONCLUSIONS

In this paper, a nonlinear predictor for the performances supervision of PV strings has been proposed. The approach is based on the estimation of four coefficients that provide a signature of the string under study. Their value, indeed, strictly depends on the string performance and its capability to deliver power in the different environmental conditions. The dependency of the model effectiveness on the statistical distribution of the solar radiation values in the training set has been investigated. In particular, a methodology for the identification of the training set that provides the best trade-off between prediction error limitation and model stability has been proposed and tested in a real application case. The result obtained for the PV strings have highlighted a high effectiveness of the model, as its good prediction accuracy permits the estimation of energy production decrease of PV string in consequence of aging process or dust deposition on the modules surface, enabling the planning of the most adequate maintenance activities.

REFERENCES

[1] S. Ghazi, A. Sayigh, and K. Ip, "Dust effect on flat surfaces – A review paper," *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 742–751, 2014.

[2] S. Ghazi and K. Ip, "The effect of weather conditions on the efficiency of PV panels in the southeast of UK," *Renewable Energy*, vol. 69, pp. 50–59, 2014.

[3] A. M. Reis, N. T. Coleman, M. W. Marshall, P. A. Lehman, and C. E. Chamberlin, "Comparison of PV module performance before and after 11-years of field exposure," in *Conference Record of the Twenty-Ninth IEEE Photovoltaic Specialists Conference 2002*, pp. 1432–1435.

[4] M. Catelani et al, "FMECA technique on photovoltaic module," in *2011 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–6.

[5] CEI-IEC International Standard 61724-Photovoltaic system performance monitoring- Guidelines for measurement, data exchange and analysis, Ed. April 1998.

[6] L. Cristaldi, M. Faifer, G. Leone, and S. Vergura, "Reference strings for statistical monitoring of the energy performance of photovoltaic fields," in *2015 International Conference on Clean Electrical Power (ICCEP)*, pp. 591–596.

[7] L. Cristaldi, M. Faifer, M. Rossi, and S. Toscani, "An Improved Model-Based Maximum Power Point Tracker for Photovoltaic Panels," *IEEE Trans. Instrum. Meas*, vol. 63, no. 1, pp. 63–71, 2014.

[8] L. Cristaldi, G. Leone, S. Vergura, "Neural Network-Based Diagnostics for PV plant," in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, [to be published].

[9] R. C. Campbell, "A Circuit-based Photovoltaic Array Model for Power System Studies," in *2007 39th North American Power Symposium*, pp. 97–101.

[10] I. H. Altas and A. M. Sharaf, "A Photovoltaic Array Simulation Model for Matlab-Simulink GUI Environment," in *2007 International Conference on Clean Electrical Power*, pp. 341–345.

[11] L. Cristaldi *et al*, "Simplified method for evaluating the effects of dust and aging on photovoltaic panels," *Measurement*, vol. 54, pp. 207–214, 2014.

[12] E. Roman, R. Alonso, P. Ibanez, S. Elorduizapatarietxe, and D. Goitia, "Intelligent PV Module for Grid-Connected PV Systems," *IEEE Trans. Ind. Electron*, vol. 53, no. 4, pp. 1066–1073, 2006.