Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 126 (201709) 651–658

Energy Procedia

www.elsevier.com/locate/procedia

72nd Conference of the Italian Thermal Machines Engineering Association, ATI2017, 6-8 September 2017, Lecce, Italy

Forecasting of PV Power Generation using weather input data-preprocessing techniques

Maria Malvoni*, Maria Grazia De Giorgi, Paolo Maria Congedo

Dipartimento di Ingegneria dell'Innovazione, Università del Salento, via per Monteroni - 73100 Lecce , Italy

Abstract

Stochastic nature of weather conditions influences the photovoltaic power forecasts. The present work investigates the accuracy performance of data-driven methods for PV power ahead prediction when different data preprocessing techniques are applied to input datasets. The Wavelet Decomposition and the Principal Component Analysis were proposed to decompose meteorological data used as inputs for the forecasts. A time series forecasting method as the GLSSVM (Group Least Square Support Vector Machine) that combines the Least Square Support Vector Machines (LS-SVM) and Group Method of Data Handling (GMDH) was applied to the measured weather data and implemented for day-ahead PV generation forecast.

Peer-review under responsibility of the scientific committee of the 72nd Conference of the Italian Thermal Machines Engineering Association

Keywords: data-driven forecast; data preprocessing, GLSSVM, wavelet, PCA, day-ahead forecast; photovoltaic; solar power.

1. Introduction

The European policies impose strict targets to reach a significant integration of renewable energy sources (RES) by year 2030 [1]. The development of forecasting models represents an essential requirement to support high sharing of renewable such as solar and wind, allowing to deal with the stochastic nature and the variability of the renewable

* Corresponding author. Tel.: +39 0832 297782.
E-mail address: maria.malvoni@unisalento.it

Peer-review under responsibility of the scientific committee of the 72nd Conference of the Italian Thermal Machines Engineering Association
10.1016/j.egypro.2017.08.293

sources in order to ensure the safe and reliability of the electric grid. In the small scale grid integration as microgrid, the prediction of PV power generation leads also to challenges for innovative energy management systems by the development of an integrated control of energy flows between loads and sources [2].

The PV power generation forecasting is a current research topic and several approaches have been investigated in the literature [3]. The data-driven forecast methods also known as time series forecasting methods need historical power measurements [4]. They include statistical and stochastic machine learning techniques and have already demonstrated high performance for such purpose [5]. On the other hand, the performance of a grid-connected PV system depends on the climate in which it is located [6]-[7], therefore the weather variables have a significant impact on the accuracy of the forecasting systems.

The weather parameters as the solar irradiance on the plane of the array and the ambient temperature can improve the forecast accuracy when actual measurements are used as input data to predict the PV output power [8]-[9]. In addition the wind speed has also a relevant influence on the predictions of PV power, leading more accurate estimations [10]-[11].

The data-driven forecast methods have to deal with a huge amount of data. In the PV power forecasts based on the historical data, redundant information can led more complex modelling. The principal component analysis (PCA) is one of the most popular techniques to remove such redundancy. It has already demonstrated its efficiency in the reduction of the dataset size without loss of information, improving the forecast accuracy with a sustainable computational time [12]-[13]. The weather variables are stochastic and introduce random fluctuations in the predicted PV power when they are used as input data in the forecast systems [8]. The wavelet technique represents an efficient solution to reduce the noise in input datasets before to implement a prediction model. The wavelet decomposition (WD) applied to the input data of a forecasting model allows to deal with the solar irradiance fluctuations, leading to the accuracy improvement [14].

The novel supervised learning algorithm GLSSVM was developed by the authors in [5], combining the group method of data handling (GMDH) and the least squares support vector machine (LS-SVM) and was implemented to predict the hourly PV output power in one-day ahead time frame. Results demonstrated that the GLSSVM outperforms the mother models, allowing to obtain lower forecast errors of the PV output power up to 24 hours ahead than the LS-SVM and the GMDH algorithms.

The purpose of this study is to investigate the performance of the data-driven forecast methods coupled with data pre-processing techniques. The forecast performance of the GLSSVM was evaluated when different preprocessing data techniques are applied to the weather historical data series to forecast the PV power output at several ahead time horizons. The wavelet decomposition and the Principal Component Analysis are the main preprocessing methods adopted in this work to decompose the time series of weather data. The input data decomposition results were applied to train the hybrid forecasting model GLSSVM at different horizons, from 1 h to 24 h ahead. This paper is organized as follows. Section 2 describes the data-preprocessing techniques and the forecasting model GLSSVM which were implemented. The input data description, the error metrics to assess the accuracy performance and the investigation methodology are presented in section 3. Results and discussions are described in section 4, and conclusions of this study are given in section 5.

2. Methods

2.1. PCA and WD for data-preprocessing techniques

2.1.1. Principal Component Analysis

The PCA is a statistical technique that allows resizing a dataset, taking into account uncorrelated and redundant information. The PCA technique finds the most significant variations of the variables. The covariance method is one of the approaches to implement the PCA [15]. Starting from a dataset $\mathcal{S}$ of dimensions Q × M where M is the observations number of Q variables, it is possible to obtain a new subset of L variables with L< Q.

Given the matrix B of dimension M × Q:

$$\mathrm{B} = \mathcal{S} - \mathrm{h\,u^T} \qquad (1)$$

where $\mathrm{u[j]} = \frac{1}{\mathrm{M}}\sum_{\mathrm{i=1}}^{\mathrm{M}} \mathcal{S}[\mathrm{i,j}]$ and $\mathrm{h[i]} = 1$ with j=1…Q and i=1…M, the Q × Q covariance matrix C of matrix B is defined as:

$$\mathrm{C} = \frac{1}{\mathrm{M-1}}\,\mathrm{B^TB} \qquad (2)$$

The D diagonal matrix Q x Q of eigenvalues of C is given by:

$$D[k,l] = \lambda_k \ \text{ for } k = l \qquad D[k,l] = 0 \ \text{ for } k \neq l \qquad (3)$$

where λ_k is the kth eigenvalues of the covariance matrix C.
The Q x Q eigenvectors matrix V that diagonalizes the covariance matrix C is:

$$V^{-1}CV = D \qquad (4)$$

Furthermore the V matrix contains Q column vectors, each of length Q, which represent the Q eigenvectors of the covariance matrix C. It is need sort the columns of the matrix V in order of decreasing eigenvalue of the matrix D. The columns of the matrix W of Q × L dimension are the first L columns of the V matrix as follows:

$$W[k,l] = V[k,l] \qquad (5)$$

with k=1,..Q and l=1,…L where L<Q. Hence the dataset $\mathcal{S}$ can be transformed into a new M-by-L space by:

$$\bar{\mathcal{S}} = BW \qquad (6)$$

In the PCA it is necessary to choose an appropriate number of the components in order to not worsen the accuracy of the predicting model. Generally, the first principal component defines the highest percentage of the variance of the samples and the second one refers to the next highest variance.

2.1.2. Wavelet Decomposition

The wavelet transform provides a mathematical tool for time-scale signal analysis at the same way as the Fourier transform in the time-frequency domain. The main difference is in the use of different time scale resolutions. Wavelet transform uses only two parameters, scale and translation and separates a time signal into different time scale components [16].
The continuous wavelet transform X_ψ is defined as the convolution of a signal x(t) with a wavelet function $\psi(t)$, known as mother wavelet, shifted in time by a translation parameter *b* and a dilation parameter *a:*

$$X_\psi(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} \psi\left(\frac{t-b}{a}\right) x(t)dt \quad (7)$$

The discrete form of the continuous wavelet transform is based on the discretization of parameters a, b that leads to a discrete set of continuous basis functions. This can be achieved as follows:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{a_j}}\psi\left(\frac{t-b_k}{a_j}\right) \qquad (8)$$

where $a_j = a_0^j$ and $b_k = kb_0a_0^j$ with j, k ∈ Z, $a_0 > 1$, $b_0 \neq 0$; the index j controls the observation resolution (dilation) and the index k controls the observation location (translation).
In order to overcome the computational complexity and the redundancy of the continuous wavelet transform, the Discrete Wavelet Transform (DWT) is introduced. Wavelet analysis is simply the process of decomposing of a signal into shifted and scaled versions the mother wavelet $\psi_{j,k}(t)$, using the scaling function $\varphi_{j,k}(t)$. The scaling function $\varphi_{j,k}(t)$, called also father function, allows to obtain different representations of the signal x(t) at different level of resolution when j varies between 0 to n. Then a signal x(t) can be written as:

$$x(t) = \sum_{j=1}^{n} c_{j,n}\varphi_{j,n}(t) \qquad (9)$$

where $c_{j,n}$ are the wavelet coefficients at level n respectively. A fast DWT algorithm based on decomposition and reconstruction low-pass and high-pass filters can be used. This algorithm permits to obtain "approximations" and "details" from a given signal. An approximation is a low-frequency representation of the original signal, whereas a detail is the difference between two successive approximations and depicts high-frequency components of the signal. Then Eq. 9 becomes:

$$x(t) = \sum_{j=1}^{n} c_{j-1,n}\varphi_{j-1,n}(t) + \sum_{j=1}^{n} d_{j-1,n}\psi_{j-1,n}(t) \qquad (10)$$

where $c_{j,n}$ and $d_{j,n}$ are the detail and the approximation coefficients at level n respectively, determined by using a filter bank.

2.2. The GLLSVM for the data-driven forecasting model

The method GLSSVM combines the GMDH and LSSVM. The input variables are selected on the base of the results of the GMDH and the LSSVM is then used as the time series forecasting models. In the hybrid model algorithm the data are divided into two subsets, one for training and the second one for testing at the first. Subsequently, all combinations of two input variables (xi, xj) are considered and the coefficients of the polynomial regression are computed. To determine new input variables for the next layer, the output variable is chosen by the least square approach, which gives the smallest of root mean square error. Then the train dataset is combined with the input variables, which are used as input for the LSSVM model. More detail can be found in [5].

3. Methodology

3.1. Description of input data

In the present work the data used to carry the investigation are referred to the photovoltaic system of nominal power of 960 kW_P, located in the campus of the University of Salento, in Monteroni di Lecce (LE), Puglia (40° 19'32"'16 N, 18° 5'52"'44 E) [17]. The panels are installed on shelters used as car parking. The PV system consists of two southeast oriented subfields with an azimuth angle of 10°. One of subfield has a nominal power of 353.3kW_P, inclined at a tilt angle of 3°. The second one of 606.7kW_P has a tilt angle of 15°. In order to monitor the main parameters of PV plant, an integrated data acquisition system is configured to measure the solar irradiation on tilted planes, ambient temperature and output power.
Generally, the input for prediction methods can consist of different parameters, called forecasting factors, which compound an input vector. In the present work the input vector is given by hourly measurements of ambient temperature, solar irradiance on tilted of array, wind speed and PV power. The monitored data are related to the historical data from 01/01/2013 to 31/12/2013, supplied in [18]. The wind speed data were provided by the laboratory of micrometeorology of the University of Salento, considering that the monitoring system is not equipped with an anemometer [11]. The solar irradiance was considered as the average between the measured irradiance on the module plane with tilt 3° and the measured irradiance on the plane of array tilted 15°, weighted by the field area ratio [11]. In order to implement the forecasting models, the collected data were divided in two sets: the training data set included 65% of the time series data and the 35% remaining of data represented the testing data.

3.2. Metrics for forecast accuracy

Suitable metrics are provided for the assessment of the accuracy of PV power forecasts as follows:

- Normalized error E(i) is the ratio of the difference between the actual and the predicted PV power and the maximum power generation at the i-th time step, as follows:

$$E(i) = \frac{P^p(i) - \mathrm{P}^m(\mathrm{i})}{Max_1^M (P^m(i))} \qquad (12)$$

where P^m is the measured PV power at the i-th time step, P^p represents the corresponding PV power predicted by a forecasting model and M is the number of values forecasted over a given period. The normalized process enables the results comparison with other PV system of different capacity.

- Normalized Mean Bias Error (NMBE) is the average of the normalized errors over the forecast period, given by:

$$NMBE = \frac{1}{M} \cdot \sum_1^M E(i) \qquad (13)$$

The NMBE gives the average forecast bias. A positive NMBE indicates over-forecasting. This means that the predicted PV power is higher than the measured data. A negative NMBE indicates under-forecasting that entails a predicted PV power lower than the measured data.

- Normalized Root Mean Square Error (RMSE) measures the overall performance of the predictions over the entire forecasting period, as follows:

$$NRMSE = \sqrt{\frac{1}{M} \cdot \sum_1^M \left(E(i)\right)^2} \qquad (14)$$

The NRMSE penalizes high forecast errors because of the squaring of each error that gives higher weight to high errors than low errors.

- Skewness (SKEW) is a measure of the asymmetry of the error probability distribution, given by:

$$SKEW = \frac{M}{(M-1)(M-2)} \cdot \sum_{i=1}^{M} \left(\frac{E - NMBE}{\sigma} \right)^3 \quad (15)$$

where σ is the error standard deviation.

The skewness indicates the overall tendency of a forecasting model to over-forecast or under-forecast. If the skewness is negative, the probability distribution is right - skewed and the model under-forecasts. For positive skewness, the probability distribution is left - skewed and the model over-forecasts. If the skewness is near zero, the distribution is symmetric.

- Kurtosis describes the magnitude of the peak or the width of the distribution of forecast errors, as follows:

$$KURT = \left\{ \frac{M(M-1)}{(M-1)(M-2)(M-3)} \cdot \sum_{i=1}^{M} \left(\frac{E_i - NMBE}{\sigma} \right)^4 \right\} * \frac{3(M-1)^2}{(M-2)(M-3)} \quad (16)$$

High kurtosis value indicates a peaked (narrow) distribution. Low kurtosis indicates a flat (wide) distribution. High peak of the distribution means a large number of very small forecast errors and indicates a more accurate forecasting method.

3.3. Implementing

During the training phase, the input data set was defined by measurements of ambient temperature, solar irradiance averaged on two tilted planes and wind speed for each hour of the monitored records. The output data set (target set) was built considering the hourly PV output power in the next 24 h with ahead forecast step of 1 h.

In order to investigate the performance of different data pre-processing techniques, the GLSSVM forecast model was implemented to predict the PV output at different time horizons (from 1 h to 24 h ahead) in five different approaches on the base of the performed data handling as follows:

1. Implementation Baseline: the input datasets were directly used as input for the GLSSVM forecast model without pre-processing;
2. Implementation WD: the input datasets were decomposed by wavelet technique and then two sets of coefficients (approximation coefficients and detail coefficients) were used as input for the GLSSVM forecast model;
3. Implementation PCA: the input datasets were analyzed by PC and then the representation of input data in the principal component space was used as input for the GLSSVM forecast model;
4. Implementation WDPCA: the input datasets were decomposed by wavelet technique, then the PCA was applied to the coefficients of the wavelet transform to obtain their representation by the principal components and successively they were used as input for the GLSSVM forecast model;
5. Implementation PCAWD: the input datasets were analyzed by PCA, the representation of input data in the principal components were decomposed by wavelet decomposition and then the approximation coefficients and detail coefficients were used as input for the GLSSVM forecast model;

All simulations are carried out by a code developed in Matlab® software. The wavelet decomposition was performed at level 8 of each row of matrix of input data, using the Wavelet Daubechies 4 (‘db4’). The principal component analysis was performed on the input data matrix by 3 principal components to obtain the representation of input data in the principal component space.

4. Results and discussion

The results of the GLSSVM forecast method with different data pre-processing techniques were investigated by the comparison of error metrics and the analysis of error statistical distributions. Fig. 1 shows the normalized mean bias errors for the five implementations. The PV power forecasts without the data pre-processing (Baseline) present high NMBE errors up to 1,1% at 6 h ahead and up to -0.8% at 20 h ahead. A positive NMBE during the first 12 h indicates the GLSSVM without input data handling over-forecasts the PV power; meanwhile the predicted PV power is lower than the measured data (under-forecast) for time horizon more than 12 h.

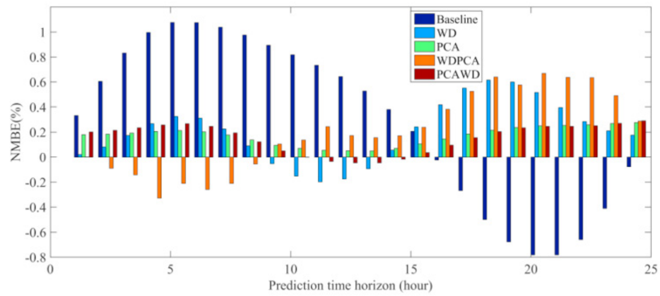

Fig. 1 NMBE for the GLSSVM forecasting model coupled with different data processing techniques

When the data handling techniques are applied, mean errors are drastically reduced. The approach based on the PCA shows higher performance than the one with the WD, leading to NMBE up to 0.3% at long time horizons, errors lower than 0.1% between 10 h and 14 h of the time horizon and with mean errors always positive, which means an over-estimation of the PV power. The WD presents mean error up to 0.6% at long time horizons, with underestimations of the power between 9 h and 13 h of the time horizon. Substantial different results are obtained when the PCA and the WD are combined and implemented on the input data. Implementation WDPCA returns higher error than implementation PCAWD, up to 0.7% at 20 h ahead. During the first 6 h ahead the WDPCA underestimates the PV power and it overestimates at long horizons. High performance is found for the implementation PCAWD with error lower than 0.3% and a tendency to underestimate the PV power mainly.

Results for the implementation PCAWD and PCA are quite similar. This means that the wavelet decomposition of the principal components doesn't improve significantly the accuracy. Unlike accuracy improvements have been found when the PCA is applied to the wavelet decomposition coefficients (WDPCA).

A further analysis can be carried out in terms of NRMSE, considering that it penalizes high forecast errors. All implementations show the same trend of the RMSE (Fig. 2). It increases up to 6.6% at short time horizons and it decreases up to 4% at 12 h ahead. In the second half of the one-day time frame it increases again up to 5.1% and then it decreases with the lowest value at 12 h. This indicates that the predicted power values are very spread from the measured data.

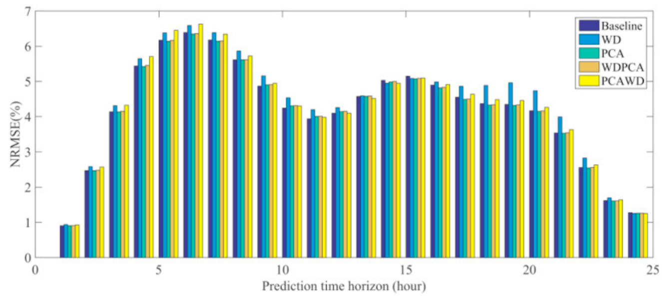

Fig. 2 NRMSE for the GLSSVM forecasting model coupled with different data processing techniques

In Fig. 3 the probability distributions of the normalized error E is plotted at different time horizons. For all implementations the trend is quite similar. High probability more than 60% to find low errors in the range [-0.1, 0.1] is given for the 1 and 24 ahead hour forecasts. When the time horizon increases up to 6 h, the probability decreases up to 40% and the normalized error rises between 0.1 and 0.3. Probabilities to find errors in the range [-0.1, 0.1] are lower than 20% for time horizons higher than 6 h. For short time horizons up to 6 h, the histograms are peaked and became flat in the next hours except at 24 h.

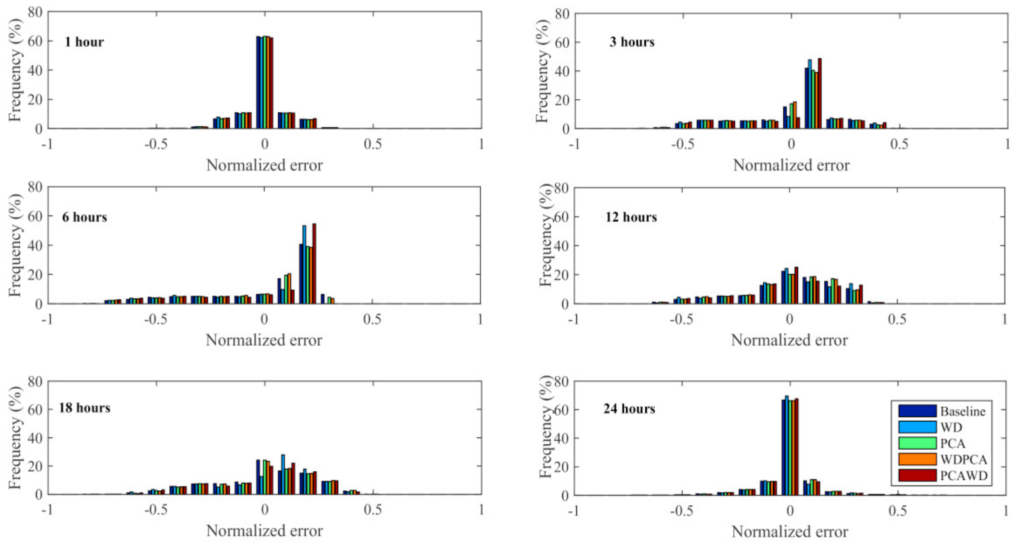

Fig. 3 Error probability distributions for the GLSSVM forecasting model coupled with different data processing techniques at a different time horizon.

Furthermore, when the time horizon increases, the histograms are shifted to the right. This is in agreement with the skewness and kurtosis parameters (Table 1). In fact the skewness is negative for all implementations, which means that the errors are spread out more to the right of the mean error than to the left, demonstrating the investigated techniques mainly trend to under-forecast the PV power generation. Furthermore the implementations with skewness closer to zero present the lowest mean errors, which is the case of the PCAWD. High kurtosis value indicates a peaked and narrow probability distribution that leads a large number of very small forecast errors, implying more accurate forecasts. The PCAWD technique returns the highest Kurtosis such as the highest probability to perform the lowest forecast error.

Time horizon	1 h	3 h	6 h	12 h	18 h	24 h
	SKEWNESS					
Baseline	-0,58	-0,81	-1,18	-0,68	-0,55	-0,36
WD	-0,52	-0,80	-1,17	-0,61	-0,82	-0,27
PCA	-0,63	-0,85	-1,18	-0,73	-0,55	-0,38
WDPCA	-0,60	-0,84	-1,18	-0,73	-0,55	-0,36
PCAWD	-0,49	-0,81	-1,22	-0,62	-0,71	-0,31
	KURTOSIS					
Baseline	6,27	3,34	3,15	3,06	2,77	11,12
WD	5,98	3,21	2,99	2,90	2,88	11,27
PCA	6,29	3,35	3,16	3,01	2,81	11,10
WDPCA	6,30	3,34	3,15	3,00	2,79	11,18
PCAWD	6,06	3,22	3,12	2,99	2,80	11,00

Table 1 Skewness and Kurtosis for the GLSSVM forecasting model coupled with different data processing techniques at a different time horizon.

Conclusions

This paper investigates the accuracy performance of the data-processing techniques applied to the hybrid forecasting model GLSSVM that combines the GMDH and LS-SVM methods. Measured data of solar irradiance, ambient temperature and wind speed were used for the predictions in a day-ahead time frame of the PV power of the grid-connected system, located in the campus of the University of Salento- Italy. The principal component analysis and wavelet decomposition techniques were chosen and combined to perform input data handling methods. Five different implementations were performed and their performances were analyzed in terms of mean error and probability distribution in order to identify the most accurate predictions. Results show the data-driven forecast method combined with the data preprocessing techniques can improve the forecast accuracy. The PCAWD outperforms the PCA, WD and WDPCA investigated techniques in most of the hour-ahead horizons. Therefore the wavelet decomposition of the principal components of input data can be applied to generate new input data set for the time series forecast models, allowing higher forecast performance with a decreasing of mean error bias up to 1%.

References

[1] Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions a policy framework for climate and energy in the period from 2020 to 2030 [COM/2014/015]

[2] Donateo, T., Congedo, P.M., Malvoni, M., Ingrosso, F., Laforgia, D., Ciancarelli, F. An integrated tool to monitor renewable energy flows and optimize the recharge of a fleet of plug-in electric vehicles in the campus of the University of Salento: Preliminary results (2014) IFAC Proceedings Volumes (IFAC-PapersOnline), 19, pp. 7861-7866. DOI: 10.3182/20140824-6-ZA-1003.01184

[3] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de-Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting (2016), Solar Energy 136, pp 78-111, DOI:10.1016/j.solener.2016.06.069.

[4] Chatfield, C. (2000). Time-series forecasting. CRC Press.

[5] De Giorgi, M.G., Malvoni, M., Congedo, P.M. Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine (2016) Energy, 107, pp. 360-373. DOI: 10.1016/j.energy.2016.04.020

[6] M. Malvoni, A. Leggieri, G. Maggiotto, P.M. Congedo, M.G. De Giorgi, Long term performance, losses and efficiency analysis of a 960 kWP photovoltaic system in the Mediterranean climate, (2017) Energy Conversion and Management, 145, pp 169-181, doi.org/10.1016/j.enconman.2017.04.075

[7] Khatib, T., Sopian, K., & Kazem, H. A. Actual performance and characteristic of a grid connected photovoltaic power system in the tropics: A short term evaluation (2013): Energy Conversion and Management 71, 115-119.

[8] De Giorgi, M.G., Congedo, P.M., Malvoni, M. Photovoltaic power forecasting using statistical methods: Impact of weather data (2014) IET Science, Measurement and Technology, 8 (3), pp. 90-97. DOI: 10.1049/iet-smt.2013.0135

[9] De Giorgi, M.G., Congedo, P.M., Malvoni, M., Tarantino, M. Short-term power forecasting by statistical methods for photovoltaic plants in south Italy (2013) 4th IMEKO TC19 Symposium on Environmental Instrumentation and Measurements 2013: Protection Environment, Climate Changes and Pollution Control, pp. 171-175.

[10] Du Y., J. Fell C., Duck B., Chen D, Liffman K., Zhang Y., Gu M, Zhu Y, Evaluation of photovoltaic panel temperature in realistic scenarios, (2016) Energy Conversion and Management, 108, pp 60-67, http://dx.doi.org/10.1016/j.enconman.2015.10.065.

[11] Malvoni, M., Fiore, M.C., Maggiotto, G., Mancarella, L., Quarta, R., Radice, V., Congedo, P.M., De Giorgi, M.G. Improvements in the predictions for the photovoltaic system performance of the Mediterranean regions (2016) Energy Conversion and Management, 128, pp. 191-202. DOI: 10.1016/j.enconman.2016.09.069

[12] Malvoni, M., De Giorgi, M.G., Congedo, P.M.Photovoltaic forecast based on hybrid PCA–LSSVM using dimensionality reducted data(2016) Neurocomputing, 211, pp. 72-83. DOI: 10.1016/j.neucom.2016.01.104

[13] M. Malvoni, M.G. De Giorgi, P.M. Congedo, Data on Support Vector Machines (SVM) model to forecast photovoltaic power, (2016) Data in Brief, 9 pp 13-16, DOI: 10.1016/j.dib.2016.08.024

[14] De Giorgi, M.G., Congedo, P.M., Malvoni, M., Laforgia, D. Error analysis of hybrid photovoltaic power forecasting models: A case study of mediterranean climate (2015) Energy Conversion and Management, 100, pp. 117-130. DOI: 10.1016/j.enconman.2015.04.078

[15] Jolliffe, Ian. Principal component analysis. John Wiley & Sons, Ltd, 2002.

[16] Mallat, Stephane G. "A theory for multiresolution signal decomposition: the wavelet representation." IEEE transactions on pattern analysis and machine intelligence 11.7 (1989): 674-693.

[17] Congedo, P.M., Malvoni, M., Mele, M., De Giorgi, M.G. Performance measurements of monocrystalline silicon PV modules in South-eastern Italy (2013) Energy Conversion and Management, 68, pp. 1-10. DOI: 10.1016/j.enconman.2012.12.017

[18] Malvoni, M., De Giorgi, M.G., Congedo, P.M. Data on photovoltaic power forecasting models for Mediterranean climate (2016) Data in Brief, 7, pp. 1639-1642. DOI: 10.1016/j.dib.2016.04.063