4th Imeko TC19 Symposium on Environmental Instrumentation and Measurements
Protecting Environment, Climate Changes and Pollution Control
June 3-4, 2013, Lecce, Italy

SHORT-TERM POWER FORECASTING BY STATISTICAL METHODS FOR PHOTOVOLTAIC PLANTS IN SOUTH ITALY

Maria Grazia De Giorgi, Paolo Maria Congedo, Maria Malvoni, Marco Tarantino

Department of Engineering for Innovation, University of Salento, Via per Monteroni, 73100 Lecce, Italy, mariagrazia.degiorgi@unisalento.it

Abstract: Statistical methods based on Multiregression Analysis and Artificial Neural Networks (ANNs) have been developed in order to predict power production of a 960 kWp grid-connected photovoltaic (PV) plant in the campus of the University of Salento, Italy.
The neural network has been used only as a statistic model based on time series of PV power and meteorological variables, as module temperature, ambient temperature and irradiance on module's plain. In particular, a sensitivity analysis has been carried out in order to find those weather parameters with the best impact on the forecasting.
Keywords:
Forecasting, Photovoltaic power, Artificial neural networks, Prediction, Multiregression Analysis

1. INTRODUCTION

An important issue for the growth of PV sector and grid-connected photovoltaic (PV) systems, is the forecasting of energy output throughout its operation. An optimal use of the renewable energy needs its characterization and prediction in order to size detectors or to estimate the potential of power plants. In terms of prediction, electricity suppliers are interested in various horizons to estimate the fossil fuel saving, to manage and dispatch the power plants installed [1]

The uncertainty of power from the sun is a limitation of PV system, influencing the quality of the electrical system that connected. So, the possibility to predict the solar irradiance (up to 24 h or even more) can became a very important role for an efficient planning of the the Grid Connected photovoltaic systems.

In literature different forecasting methods have been developed to evaluate the performance of PV systems. In [2] C. Chupong and B. Plangklang presented the power forecasting of a PV system by calculating the solar radiation, collecting data from weather forecasting, and using Elman neural network to forecast by using data from PV system.

In [3] a MLP network for forecasting of 24 h ahead of solar irradiance was developed. The proposed model used as input parameters the mean daily irradiance and the mean daily air temperature. A good results were obtained from comparison between the measured and the forecasted PV power. Statistical prediction methods are based on models that establish the relation between historical values of the power and the meteorological variables. So, it's important to choose the right ambient data.

ANNs are useful tools to understand the complex and nonlinear relationships among data, without any previous assumption concerning the nature of these correlations. [4, 5]. The training is one of the most critical phase. In this step the choice of input data and of the neural connections have to be properly set in order to have an appropriate simulation of the performance of a PV plant.
An aim of the present study is to underline the influence of several weather parameters with respect to the accuracy of PV power predictions.
This paper presents an artificial neural network (ANN) approach for forecasting the performance of electric energy generated output from a 960 kWp grid-connected photovoltaic (PV) plant installed in the campus of the University of Salento, Italy.

The present study is a part of the funded research project ‘‘7th Framework Programme Building Energy Advanced Management Systems (BEAMS)’’. Part of the BEAMS research program is concerning the study on the benefits of installing PV systems and charging stations for electrical vehicles (EV) and the development of tools to improve/optimize the distribution of loads in the grid composed by the public facility services.
The University of Salento is one of the two pilot sites in which this project is being developed. In the last 2 years, the university has significantly promoted the use of energy from renewable sources by the installation of solar PV roofs on parking areas and charging stations for electric cars.

The ANN interpolates among the solar PV generation output and relevant parameters such as solar radiation, module temperature and ambient temperature

Utilizing the regression analysis, the influence of measured meteorological data on PV power generation has been analyzed.

In this study, two ANN models are implemented and validated with reasonable accuracy on real electric energy generation output data. In the first model, the PV power output for the next 1 hour (t+1) is calculated, using a time series of measured hourly data, included the main parameter at time t, module temperature, ambient temperature, irradiance and instant PV power. In the second approach, the PV power measure at the instant t is implemented to prevision PV power at t+1, without weather parameters.

2. HYSTORICAL DATA AND SITE DESCRIPTION

The site under study is the PV park, located in the campus of the University of Salento, in Monteroni di Lecce, Apulia (40_190320016N, 18_50520044E). It is characterized by a warm Mediterranean climate with a dry summer. In order to define a prediction model for PV power, the most significant problem remains the selection of the best parameter to use from among the several variables of the system. A detailed description of this PV system is in [6].

The data acquisition system consists of three inverters, the solar irradiance sensors and the PV module/ambient temperature sensors. The data from the inverters and the sensors are characterized by protocols Modbus, Profibus, clean contacts or digital inputs, and they are collected by a PLC Siemens with a scada WINCC for processing and storage. In particular, an analysis of the time series represented by the following daily data (collected every 1 hour) has been carried out: module temperature (°C), ambient temperature (°C), irradiance on plain inclined at a tilt angle of 3° and irradiance for a tilt angle of 15° (W/m^2), PV power(W). The time series data used included 365 days (from 05/03/2012 to 05/03/2013).

3.MULTIREGRESSION ANALYSIS

Multiple regression is a data analysis technique that permits to measure of how well a given parameter variable can be predicted using a linear function of a set of other variables.
The aim of the multi-regression analysis was to obtain a relationship between PV power, module temperature and the ambient conditions (ambient temperature and irradiance on plain of modules). The first effort made was to develop a model to predict PV power based on four inputs: ambient temperature, module temperature, irradiance on two plains inclinated. The general form of the model equation obtained is:

$$P = b_1*T_{Amb} + b_2*T_{Mod} + b_3*I_3 + b_4*I_{15}$$

The regression coefficients have been calculated by an iteratively reweighted least squares algorithm, with the weights at each iteration calculated by applying the bi-square function to the residuals from the previous iteration..
First of all, a detailed sensitivity analysis has been carried out in order to find those weather parameters with the highest impact on the forecast by a linear regression between each weather parameter and the PV power.
The best regression for the inputs selection could be evaluated in terms of squared correlation coefficient R^2.
Figures 1a-1d show the hourly PV power versus, respectively, hourly ambient temperature, module temperature, irradiance 3° and 15° on the basis of one year collected data.

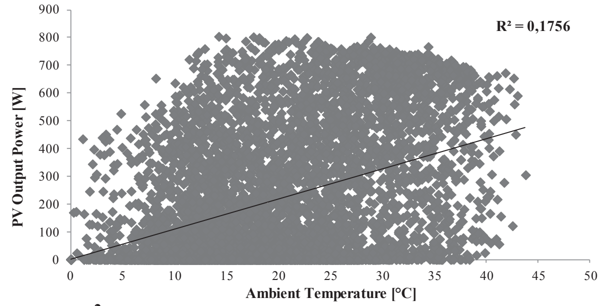

Fig. 1.a R^2 coefficient for linear regression between Ambient Temperature and PV Power

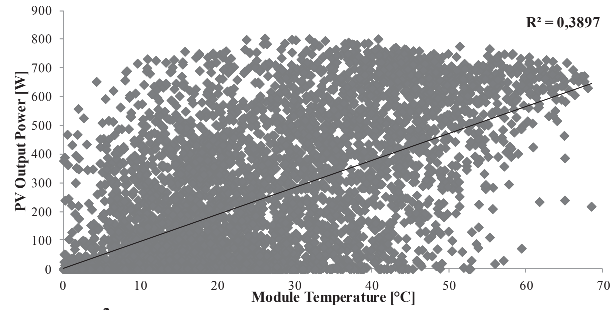

Fig. 1.b R^2 coefficient for linear regression between Module Temperature and PV Power

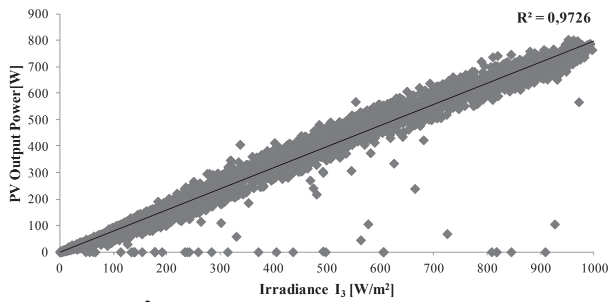

Fig. 1.c R^2 coefficient for linear regression between Irradiance on plain 3° and PV Power

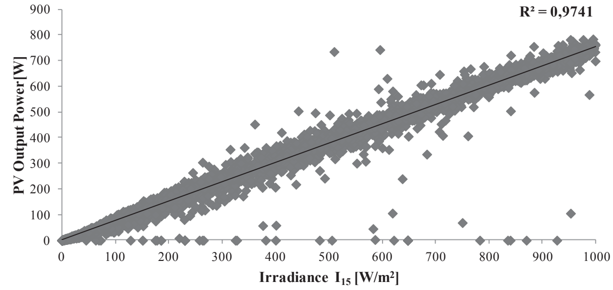

Fig.1.d R^2 coefficient for linear regression between Irradiance on plain 15° and PV Power

It is evident that the most correlated parameter with PV power is given by irradiance.

Table 1 Coefficients for the four input parameter

T_{Mod} Module Temperature	T_{Amb} Ambient Temperature	I_3 Irradiance on plane of module with tilt 3°	I_{15} Irradiance on plain of module with tilt 15°
b_1= 0.07	b_1=-0.3	b_3=0.39	b_4=0.39

In view of the R^2 values obtained, all parameters have been taken into consideration to implement the multi-regression analysis. The coefficients b_1, b_2, b_3, b_4 are presented in table 1. Figure 2 shows coefficient R^2 in the case of the multi-regression analysis, underlining good correlation.

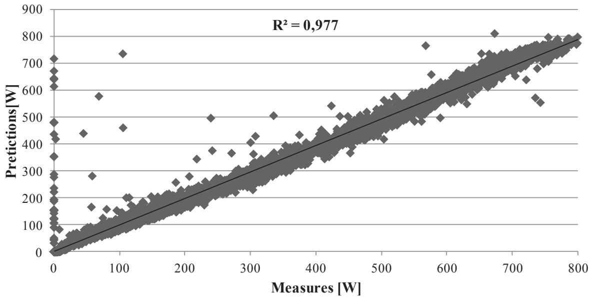

Fig. 2. PV power measured versus values forecasted by multi-regression analysis.

4. ARTIFICIAL NEURAL NETWORKS (ANNS)

Neural networks are composed of simple elements operating in parallel, inspired by biological nervous systems. The network function is given by the connections between elements. A neural network is trained to perform a particular function by adjusting the values of the connections (weights) between elements (Fig.3).

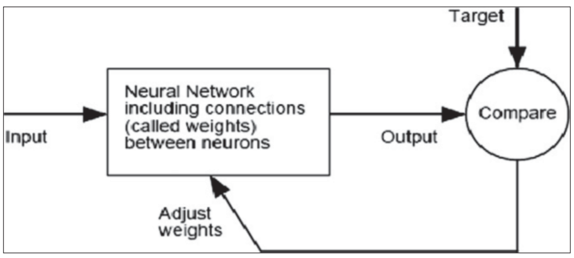

Fig.3 Basic Block Diagram of Neural Network

The basic component of such a system is a neuron. When active, electrochemical signals are received through synapses to the neuron cell. Each synapse has its own weight that determines the contribution and extent to which the respective input affects the output of the neuron. The weighted sum of the input electrochemical signals is fed to the nucleus that sends electrical impulses in response, being transmitted to other neurons or to other biological units as actuation signals. Neurons are interconnected through synapses. The synaptic weights modify continuously during learning. Groups of neurons are organized into subsystems and integrate to form the brain.

In the ANN technique, a simulation of a small part of the central nervous system is done which is a rather basic mathematical model of the biological nervous system. Inputs are fed into the corresponding neurons, and the electrochemical signals are altered by weights.

The weighted sum is operated upon by an activation function, and outputs are fed to other neurons in the network. All these neurons are highly interconnected and the activation values constitute final output or may be fed to the next model. These connection weights are continuously modified during training to obtain desired accuracy and generalization capabilities.

Elman ANN network

In this work, ELMAN ANN network have been used to forecasting and evaluating the PV power of the park.

This kind of network is characterized by feedback from the first layer output to the first-layer input.

This recurrent connection allows the Elman network to detect and generate time-varying patterns (Fig. 4).

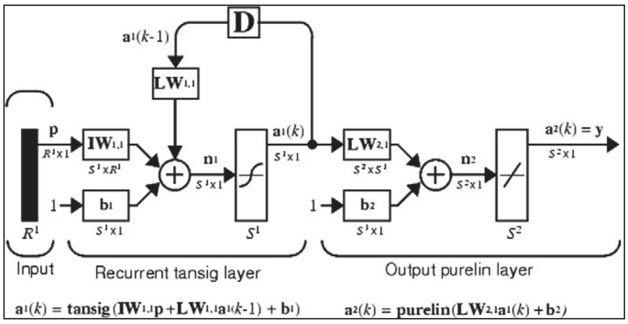

Fig.4 Typical architecture of an Elman Back Propagation network.

Results and discussion

In this study, the ANN has been compiled with the Matlab© software and its Neural Network toolbox.

Firstly, an accurate elaboration of the measured values was necessary in order to check, in each month, the days in which the parameters were either unavailable or incorrect. Subsequently, the real values of all data were normalized in a range [-1, 1].

The neural network has been used only as a statistic model based on time series of on-line measured PV power.

For each time instant t, the input value is given by the average hourly power at that time, while the target is given by the average hourly powers along the forecast horizon h=1.

Table 2 shows the network parameters used in the training.

As said two different ANN forecasting systems were implemented. Table 3 describes the numerical parameters included in each of the forecast systems.

Table 2 Elman network parameters used in the training for the forecast system I and II

Training function TRAINGDX
Adapt learning function LEARNGD
Performance function MSE
Number layers 3
Neurons (layer 1) 5
Neurons (layer 2) 5
Neurons (layer 3) 1
Activation function hidden layer TANSIG
Activation function output layer PURELIN
Epochs 500

Table 3 Numerical parameters included in each of the forecast systems

Forecast system	Numerical parameters included in the forecast system
I	P PV output power at instant t
II	T_{Mod} Module Temperature T_{Amb} Ambient Temperature I_3 Irradiance on plane of module with tilt 3° I_{15} Irradiance on plain of module with tilt 15° P PV output power at instant t

Model I is based on one inputs: the hourly average data of PV power and applied on a training period of 1 years for a forecasting horizon at the time t + 1 (1 h).
The performance of the ANN is evaluated using a data set of input variables (testing data set) different from that used in the training process.
The testing data set is given by the data collected in eight months, while the training data is given by the data collected in 3 months.
All the collected data time series data (365 days/6297 hourly records) were divided in two sets: training and testing data sets. The training data set included 65% of the time series data, the testing data set 35%.
These forecast values are compared with the actual values recorded at site (Fig.5).
The second model is based on five inputs: the hourly average data of the weather parameters and PV power and applied on a training period of 1 years for a forecasting horizon at the time t + 1 (1 h). 4

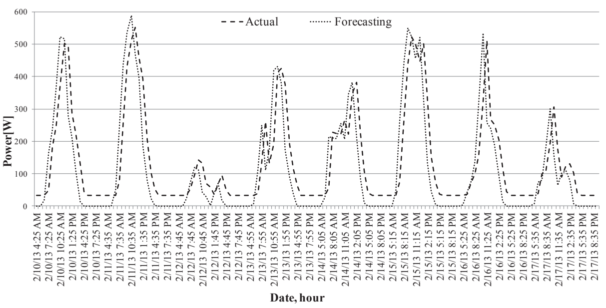

Fig.5 Compare Forecast value and Actual value in Forecast system I

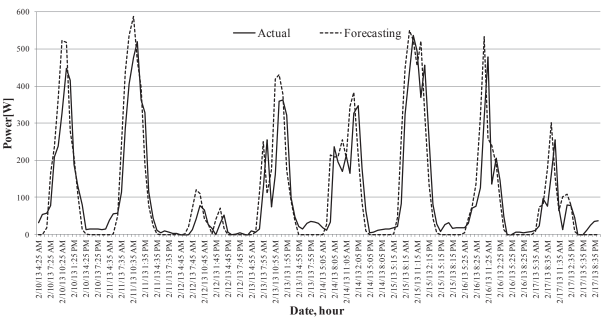

Fig.6 Compare Forecast value and Actual value in Forecast system II

The data included in the forecast systems are: ambient temperature, module temperature, irradiance on plain 3° and irradiance on plain 15°.
Figure 6 shows measured and predicted values.
The comparison of the results obtained with the two different forecasting models was carried out by means of the normalized absolute average error for the forecast method at the time horizon of 1 h, defined as:

$$E_i = \frac{P_i - T_i}{Max_{i=1}^{n}(T_i)} * 100$$

Where
i = generic time instant;
n = number of observations;
P_i = predicted power at instant i;
T_i = real power at instant i.

Then calculated as the mean absolute normalized percentage error, the value is equal to 9.56% for model I and 6,53% in the second model. This confirms the importance of input data based also on weather parameters.

5. CONCLUSIONS

This study is focused on the implementation of a short-term forecasting system for the hourly electrical energy production in a real, grid-connected PV plant
The analyzed forecast systems are based on Elman neural network.
The input variables used for the development of the models were past values of hourly energy production in the PV plant, as well measured weather variables.
A sensitivity analysis has been done to verify the impact of the different parameters to PV power generation. In particular multiple regression analysis has been performed to measure of how well the PV power can be predicted using a linear function of a set of other variables. Results underline the high impact of irradiance on PV power.
Then ANN based on both measured power and meteorological data was revealed as the best forecasting model.

Funds

This work is supported by the Project BEAMS, Project Number 285194, 7th Framework Program.

6. REFERENCES

[1] C. Paoli, C. Voyant M. Muselli, M.L. Nivet,"Forecasting of preprocessed daily solar radiation time series using neural networks", Solar Energy no. 84, pp 2146–2160, 2010

[2] C. Chupong and B. Plangklang "Forecasting power output of PV grid connected system in Thailand without using solar radiation measurement", Energy Procedia no. 9, pp 230 – 237, 2011

[3] A. Mellit , A. Massi Pavan,"A 24-h forecast of solar irradiance using artificial neural network:

application for performance prediction of a grid-connected PV plant at Trieste, Italy" Solar Energy n.84, pp 807–821, 2010
[4] M.G. De Giorgi, A. Ficarella, M. Tarantino, "Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods" Energy no.36, pp 3968-3978, 2011
[5] M.G. De Giorgi, A. Ficarella, M. Tarantino, "Error analysis of short term wind power prediction models" Applied Energy no. 88, pp 1298–1311, 2011
[6] P.M. Congedo , M. Malvoni, M. Mele, M.G. De Giorgi "Performance measurements of monocrystalline silicon PV modules in South-eastern Italy", Energy Conversion and Management no.68 , pp 1–10, 2013