Energy Conversion and Management 100 (2015) 117–130

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

# Error analysis of hybrid photovoltaic power forecasting models: A case study of mediterranean climate

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

*Dipartimento di Ingegneria dell'Innovazione, Università del Salento, via per Arnesano, I-73100, Italy*

ARTICLE INFO

*Article history:*
Received 30 May 2014
Accepted 27 April 2015
Available online 15 May 2015

*Keywords:*
Photovoltaic power forecast
Least square support vector machine
Artificial neural network
Wavelet decomposition
Forecasting errors
Imbalance penalties
Solar irradiance
Weather variations

ABSTRACT

The advancement of photovoltaic (PV) energy into electricity market requires efficient photovoltaic power prediction systems. Furthermore the analysis of PV power forecasting errors is essential for optimal unit commitment and economic dispatch of power systems with significant PV power penetrations. This study is focused on the forecasting of the power output of a photovoltaic system located in Apulia – South East of Italy at different forecasting horizons, using historical output power data and performed by hybrid statistical models based on Least Square Support Vector Machines (LS-SVM) with Wavelet Decomposition (WD). Five forecasting horizons, from 1 h up to 24 h, were considered. A detailed error analysis, by mean error and statistical distributions was carried out to compare the performance with the traditional Artificial Neural Network (ANN) and LS-SVM without the WD. The decomposition of the RMSE into three contributions (bias, standard deviation bias and dispersion) and the estimation of the skewness and kurtosis statistical metrics provide a better understanding of the differences between prediction and measurement values. The hybrid method based on LS-SVM and WD out-performs other methods in the majority of cases. It is also evaluated the impact of the accuracy of the forecasting method on the imbalance penalties. The most accurate forecasts permit to reduce such penalties and thus maximize revenue.



## 1. Introduction

Productivity forecasting has always been a key issue in power system operation. In particular, with the rise of deregulation and free competition of the electric power industry, loads and productivity forecasting has become more important than ever before. Since renewable energy power plants were used, such as PV systems and wind farm, the productivity forecast for the national energy system becomes difficult due to the high variability of the electricity production of this new system.

The present study is a part of the funded research project "7th Framework Programme Building Energy Advanced Management Systems (BEAMS)". The project aims to develop an advanced, integrated management system for many buildings, in particular for the public ones; this system has to be able to control and improve the energy efficiency of infrastructures in term of using public lighting, ventilation, air conditioning, electric vehicles and other types of energy from renewable sources. Furthermore, part of the BEAMS research program concerns the study of the benefits of installation of PV systems 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 [1]. The short term PV power prediction is very important for the planning and management of electric system, but the critical aspects have to be considered. The forecasting accuracy depends also on the weather conditions of installation site and the randomness of solar source is the main limitation of photovoltaic system, which influences the quality of the connected electrical system. The possibility to predict the solar irradiation or PV power (up to 24 h or even more) [2–4] and the development of real time prediction model [5] help to optimize the integration of PV generator in the electric grids.

The forecasting methods applied in the field of renewable energy can be classified into different categories: the physical model, the conventional statistical model, the spatial correlation model, and the artificial intelligence [6,7]. Some of these prediction models are more accurate at short-term prediction while others are better in long-term prediction [6].

Electric load time series are usually nonlinear functions of exogenous variables. To incorporate non-linearity, Artificial Neural Networks (ANNs) received great attention in solving

* Corresponding author. Tel.: +39 0832297759; fax: +39 0832297777.
*E-mail address:* mariagrazia.degiorgi@unisalento.it (M.G. De Giorgi).

http://dx.doi.org/10.1016/j.enconman.2015.04.078
0196-8904/

problems of electricity price [8], electrical energy consumption [9] or productivity forecasting [10–12].

In [13–15] methods based on artificial neural networks were implemented for estimating the energy provided by a PV generator in the next hours. In particular in [15] four different methods were compared: three of them are classical methods and the fourth one is based on an artificial neural network developed by the R&D Group for Solar and Automatic Energy at the University of Jaen.

In the literature different methods based on artificial intelligence techniques have been implemented, including the Artificial Neural Network (ANN) of Multi-Layer Perceptrons (MLP) [16], Radial Basis Function [17] and Recurrent Neural Networks [18] and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [19].

Studies dealing with the applications of ANNs for PV and wind generation forecasting can be found in [20–22]. In [23] ANNs have been applied for annual energy harvesting calculation of grid-connected PV systems.

Fadare [24,25] applied ANN model to predict wind speed variation [24] and to forecast solar radiation in Nigeria [25].

Artificial neural network models provide better short-term productivity forecasts with respect to standard linear Autoregressive Integrated Moving Average (ARIMA) models [18] and persistent model [26].

De Giorgi et al. [27] compared ARMA models, which perform a linear mapping between inputs and outputs with Artificial Neural Network (ANNs) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS), which perform a non-linear mapping, underlining that, at long time horizon, ANNs presents higher accuracy in wind power forecasting. This was also confirmed in [28] for PV power predictions.

In [29] Radial Basis Functions and Multilayer Perceptron ANNs were compared to predict solar radiation by estimating the clearness index. To forecast the hourly global horizontal solar radiation, a method, based on the combination of the k-means algorithm and NAR (nonlinear autoregressive) network, was proposed in [30]. In [31] a regression neural network was implemented to predict the solar radiation on tilted surfaces.

In [32] the power forecasting of a PV system was performed by Elman neural network, which was based on solar radiation and weather forecasting data as inputs. However, a major risk in the use of ANN models is the possibility of excessive training data approximation, i.e., over-fitting, which usually increases the out-of-sample forecasting errors.

Recently, new methods for time series forecasting that are based on Learning Machines were developed, using Support Data Machines (SVMs) [33,34]. Several studies underlined that SVMs are more resistant to the over-fitting problem, by achieving high generalization performance in solving forecasting problems of various time series. SVM can model complex problems with datasets given by several variables and a reduced training dataset. In [35] the SVM was used to model the battery nonlinear dynamics. The feasibility of using SVMs to forecast electricity load was discussed in [36]. An advantage in the use of SVM is that it is less computational expensive than traditional ANN models based on back-propagation algorithms [37]. Mohandes et al. [38] compared favorably the performance of SVMs with the multilayer perceptron (MLP) neural networks for the prediction of the wind speed in Madina city, Saudi Arabia.

In [39] the SVM was applied to estimate daily solar radiation using sunshine duration. In [40] an estimation of the monthly solar radiation was obtained by SVM methods that were trained on air temperature data. In [41] the impact of different prediction horizons was evaluated for photovoltaic power forecasting methods, that were based on support vector regression and numerically predicted weather variables.

In the literature various hybrid SVM methods were also developed [42]. An adaptive two-stage hybrid network with self-organized map (SOM) and support vector machine (SVM) was developed for short-term load forecasting in [43].

Beyond the hybridization of the SVM, in the recent literature a variant of the standard SVM has been introduced that is the Least Square Support Data Machine (LS-SVM), which uses a simplified linear model, simpler and computationally easier but with the same advantages of the ANNs and SVMs models [44]. LS-SVM models were already applied for wind power forecasting [45–47].

Regarding the hybrid methods, prediction forecast models that are based on wavelet decompositions WD, could be used to improve the prediction performance of short-term load forecast, as shown in [48,49]. Least Square Support Vector Machine (LS-SVM) with Wavelet Transform were used in [50] to predict day-ahead electricity prices.

The PV power time series generally include low and high frequency components. WD decomposes the PV power time series into its components, which could be used separately as input in the prediction model. In [51] a hybrid approach based on WD and ANNs and evolutionary algorithm was successfully proposed for accurate short-term load forecasting of power systems.

Forecasting the produced energy with high accuracy is a key issue in microgrid control, where the photovoltaic (PV) energy sources are dominating the market.

The integration of energy sources into micro-grid operation, as PV generators or wind turbines, needs the consideration of power generation uncertainty. Hence, for optimal operation of PVs and wind turbines, the capacity of solar and wind generation must be considered in the scheduling of the micro-grids. The dependable capacity of PVs and wind turbine is an important factor that is related to the accuracy of photovoltaic and wind power forecast [52].

Finally forecast errors can have substantial economic consequences, if they are large enough that they cause a different commitment than would have been performed with an optimal forecast.

Furthermore, in the liberalized markets, e.g. in Italy [53], if there is a mismatch between the injections of a photovoltaic power plant and the day-ahead market power, the energy injections out of a tolerance band are charged of imbalance penalties [54]. For these reasons, very important is the analysis of the accuracy of the forecasting method by the evaluation of several statistical metrics and of the forecast errors distribution, e.g. the tails of the forecast error distribution have the greatest economic impact and there is more uncertainty in the forecasts.

Despite the importance of a deep analysis of the accuracy of the forecasting methods, several works in the literature performed the evaluation of the different forecasting methods by the estimation of conventional metrics, as the root mean square error (RMSE), mean bias error (MBE), and mean absolute error (MAE). In the present study PV power output forecasting are performed by LS-SVM with Wavelet Decomposition of the input data. Two different input datasets are implemented. The first one is based on the measured power output, the second one uses also the module temperature, the ambient temperature, and the irradiance on plain inclined at the tilt angle. The results in term of accuracy are compared with those of ANN. The performance evaluation is performed by a detailed error analysis [55].

In the literature few works focused on the comparison of ANN and hybrid LS-SVM forecasting models for PV power based on the evaluation of several error metrics.

In the present work a deep study of the statistical error distribution, a decomposition of the standard deviation by amplitude and phase error and the evaluation of the skewness and kurtosis