Bayesian Based Neural Network Model for Solar Photovoltaic Power Forecasting

Angelo Ciaramella[1], Antonino Staiano*[1], Guido Cervone[2], and Stefano Alessandrini[3]

1 Dept. of Science and Technology, University of Naples "Parthenope", Isola C4, Centro Direzionale, I-80143, Napoli (NA), Italy
{angelo.ciaramella, antonino.staiano}@uniparthenope.it
2 Penn State University, USA
3 NCAR, National Centre for Atmospheric Research, Boulder, CO, USA

Abstract. Solar photovoltaic power (PV) generation has increased constantly in several countries in the last ten years becoming an important component of a sustainable solution of the energy problem. In this paper, a methodology to 24-hour or 48-hour photovoltaic power forecasting based on a Neural Network, trained in a Bayesian framework, is proposed. More specifically, an multi-ahead prediction Multi-Layer Perceptron Neural Network is used whose parameters are estimated by a probabilistic Bayesian learning technique. The Bayesian framework allows obtaining the confidence intervals and to estimate the error bars of the Neural Network predictions. In order to build an effective model for PV forecasting, the time series of Global Horizontal Irradiance, Cloud Cover, Direct Normal Irradiance, 2-m Temperature, azimuth angle and solar Elevation Angle are used and preprocessed by a Linear Predictive Coding technique. The experimental results show a low percentage of forecasting error on test data, which is encouraging if compared to state-of-the-art methods in literature.

1 Introduction

Solar photovoltaic technology has become one of several renewable energy [1]. It has receiving global research attention due to its natural abundance, noise pollution free, non-emission of greenhouse gases unlike the fossil powered generation sources which affect the climatic conditions and causing global warming. The technology has gained popularity in the establishment of large solar farms in major countries such as the United States of America, Spain, Italy, Japan, China, Australia and in other countries of the world. Mainly for its suitability for power generation in urban and remote isolated rural areas for small scale applications such as water pumping systems and domestic electricity supply for lighting and other uses [6]. In Italy, thanks also to substantial government subsides over the past five years, the annual generation by solar photovoltaic power

* Corresponding author

(PV) has notably increased, from 200 GWh in 2008 to 19418 GWh in 2013, that corresponds to 7% of the total Italian energy demand [19]. The energy produced by photovoltaic farms has a variable nature depending on astronomical and meteorological factors. The former are the solar elevation and the solar azimuth, which are easily predictable without any uncertainty. The latter, instead, deeply impact on solar photovoltaic predictability. Since the power produced by a PV system depends critically on the variability of solar irradiance and environmental factors, unexpected variations of a PV system output may increase operating costs for the electricity system by increasing requirements of primary reserves, as well as placing potential risks to the reliability of electricity supply.

A priority of a grid operator is to predict changes of the PV system power production, mainly using persistence-type methods, in order to schedule the spinning reserve capacity and to manage the grid operations. In addition to transmission system operators, online power prediction of the PV system is also required by various end-users such as energy traders, energy service providers and independent power producers, to provide inputs for different functions like economic scheduling, energy trading, and security assessment.

In this last years, several researches for forecasting the solar irradiance in different scale times have been made by using Machine Learning and Soft Computing methodologies [13]. In particular based on Artificial Neural Networks [12, 7], Fuzzy Logic [16] and hybrid system such as ANFIS [17], Recurrent Neural Networks [4] and Support Vector Machines [18]. Most of the works concentrate only on few impact parameters (e.g., only temperature) and they not consider a confidence interval for the estimated prediction.

In this paper a methodology to PV forecasting based on a Neural Network (Multi-Layer Perceptron), trained in a Bayesian framework, is proposed. The Bayesian framework allows obtaining the confidence intervals and to estimate the error bars of the model prediction. In order to build an effective model for PV forecasting, the time series of Global Horizontal Irradiance, Cloud Cover, Direct Normal Irradiance, 2-m Temperature, azimuth angle and solar elevation angle are used. The features of the observed time series are extracted by a Linear Predictive Coding mechanism.

The paper is organized as following. In Section 2 the Solar Photovoltaic Power Data are described. In Section 3 we introduce the Multi-Layer Perceptron and the Probabilistic Bayesian learning and in Section 4 the experimental results are described. Finally, in Section **??** we focus on conclusions and future remarks.

2 Solar Photovoltaic Power Data

To maintain grid stability at an effective cost, it has now become crucial to being able to predict with accuracy the renewable energy production which is combined with other more predictable sources (e.g., coal, natural gas) to satisfy the energy demand [10][11]. In this work, data collected from three PV farms are considered. They are located in Italy, namely Lombardy and Sicily regions, with a nominal power (NP) of 5.21 kW and in Calabria region with a nominal power around 5 MW [15].

3 Multi-Layer Perceptron and the Probabilistic Bayesian learning

A Neural Network (NN) is usually structured into an input layer of neurons, one or more hidden layers and one output layer. Neurons belonging to adjacent layers are usually fully connected and the various types and architectures are identified both by the different topologies adopted for the connections and by the choice of the activation function. Such networks are generally called Multi-Layer Perceptron (MLP) [2] when the activation functions are sigmoidal or linear. The output of the j-th hidden unit is obtained first by forming a weighted linear combination of the d input values, and then by adding a bias to give:

$$z_j = f\left(\sum_{i=0}^{d} w_{ji}^{(1)} x_i\right) \tag{1}$$

where d is the number of the input, $w_{ji}^{(1)}$ denotes a weight in the first layer (from input i to hidden unit j). Note that $w_{j0}^{(1)}$ denotes the bias for the hidden unit j; and f is an activation function such as the continuous sigmoidal function.

The outputs of the network are obtained by transforming the activation of the hidden units using a second layer of processing elements

$$y_k = g\left(\sum_{j=0}^{M} w_{kj}^{(2)} z_j\right) \tag{2}$$

where M is the number of hidden unit, $w_{kj}^{(2)}$ denotes a weight in the second layer (from hidden unit j to output unit k). Note that $w_{k0}^{(2)}$ denotes the bias for the output unit k; and g is an activation function of the output units which does not need to be the same function as for the hidden units. The learning procedure is the so called back propagation [2].

Due to its interpolation capabilities, the MLP is one of the most widely used neural architectures. The MLP can be trained also using probabilistic techniques. The Bayesian learning framework offers several advantages over classical ones [2]: (i) it cannot overfit the data; (ii) it is automatically regularized; (iii) the uncertainty in the prediction can be estimated.

In the conventional maximum likelihood approach to training, a single weight vector is found which minimizes the error function; in contrast, the Bayesian scheme considers a probability distribution over the weights. This is described by a prior distribution $p(\mathbf{w})$ which is modified when we observe the data $\mathbf{D}$. This process can be expressed with Bayes theorem:

$$p(\mathbf{w}|\mathbf{D}) = \frac{p(\mathbf{D}|\mathbf{w})p(\mathbf{w})}{p(\mathbf{D})} \tag{3}$$

To evaluate the posterior distribution, we need expressions for the prior $p(\mathbf{w})$ and for the likelihood $p(\mathbf{D}|\mathbf{w})$. The prior over weights should reflect the knowledge, if any, about the mapping to be built.
We consider k different sets of weights by using different regularization parameters α_k for each group. To preserve the scaling properties of the network mapping the prior equation can be written as [14]

$$p(\mathbf{w}) = \frac{1}{Z_W(\{\alpha_k\}_k)} e^{-\sum_k \alpha_k E_{W_k}} \tag{4}$$

where k runs over the different weight groups, $Z_W(\{\alpha_k\}_k)$ is a normalization factor, E_{W_k} weight regularizer, α_k is the hyperparameter.

Once the expressions for the prior and the noise model is given, we can evaluate the posterior

$$p(\mathbf{w}|\mathbf{D}, \{\alpha_k\}_k, \beta) = \frac{1}{Z(\{\alpha_k\}_k, \beta)} e^{-\beta E_D - \sum_k \alpha_k E_{W_k}} \tag{5}$$

where β is an hyper parameter, $Z(\{\alpha_k\}_k, \beta)$ is a normalization factor and E_D is an appropriate error function [2].
This distribution is usually very complex and multi-modal, and the determination of the normalization factor is very difficult. Also, the hyperparameters must be integrated out, since they are only used to determine the form of the distributions.
The approach followed is the one introduced by [9], which integrates the parameters separately from the hyperparameters by means a Gaussian approximation and then finds the mode with respect to the hyperparameters. This procedure gives a good estimation of the probability mass attached to the posterior, in particular way for distributions over high-dimensional spaces [9].
Using a Gaussian approximation the mode of the resulting distribution can be evaluated

$$\{\{\hat{\alpha}_k\}_k, \hat{\beta}\} = \mathrm{argmax}_{\{\alpha_k\}_k, \beta} p(\{\alpha_k\}_k, \beta|\mathbf{D}) = \int p(\{\alpha_k\}_k, \beta, \mathbf{w}|\mathbf{D}) d\mathbf{w} \tag{6}$$

and the hyperparameters thus found can be used to evaluate:

$$\hat{\mathbf{w}} = \mathrm{argmax}_{\mathbf{w}} p(\mathbf{w}|\mathbf{D}, \{\alpha_k\}_k, \beta) \tag{7}$$

The above outlined scheme must be repeated until a self-consistent solution $(\mathbf{w}|\{\alpha_k\}_k, \beta)$ is found.

4 Experimental results

Data have been collected from the three PV farms previously mentioned. In particular, the data used in the experiments are a combination of model forecasts and observational measurements relative to three stations, called Lombardy, Calabria and Sicily (approximate locations: (45.5N, 9.2E, 39.7N, 16.2E,

37.4N, 15.0E). The data were available for the following periods: Sicily: January 2010-December 2011, Lombardy: July 2010-December 2011, Calabria: April 2011-March 2013.
The forecast data is composed of atmospheric forecasts generated using the Regional Atmospheric Modeling System (RAMS), initialized with boundary conditions from ECMWF deterministic forecast fields starting at 00 UTC. The forecasted parameters used in the experiments include Cloud Cover (CC), the Global Horizontal Irradiance (GHI) and air Temperature at 2 m above the ground (T2M). Additionally, the data were augmented with computed measurements for solar AZimuth angle (AZ) and solar ELevation angle (EL) have been included. The observations include the real quantity of electrical power generated from each of the three solar farms. The data was averaged into hourly values to be consistent with the forecast data.
These predictors are then employed to train an MLP for a 24-hour or 48-hour ahead prediction. In particular, we consider 24 hours of observations to forecast 24 or 48 hours of PV concentrations. For each parameter, previously mentioned, a Linear Predictive Coding (LPC) technique is used to extract the main features. LPC is for representing the spectral envelope of a digital signal in compressed form, using the information of a linear predictive model [5]. More in detail, if we consider k LPC coefficients for each parameter (GHI, CC, T2M, AZ, EL, PV), at time t, the NN has $6 \times k$ inputs, $x_i, i = 1, \ldots, 6 \times k$ corresponding to the observed 24 hours $\mathbf{x} = [x_1, \ldots x_{6\times k}]$. The outputs are 24 or 48 corresponding with the PV observations to forecast. In Figure 1 the architecture of the NN with 24 outputs is shown. The data set has been divided in training and test sets (70% and 30%, respectively). A K-fold cross-validation mechanism is used to select the optimal number of hidden nodes [3]. In the first experiment we concentrate on the prediction of PV for 24-hour. The NN is composed by 10 hidden nodes. We obtained a cross-correlation coefficient of 97% between source and predicted PV sequences on the training set and of 91% on the test set. Moreover we use a Root Mean Square Error (RMSE) to calculate the percentage of error and we obtain 6% and 8% on the training and test sets, respectively. In Figures 2 and 3 we report some predictions obtained by the model on test data. In particular, in Figure 2 a regular PV concentration is considered and in Figure 3 an anomaly sequence. In both cases the model permits to predict the concentrations with high accuracy. Successively, we concentrate on the prediction of PV for 48-hour. We obtained a cross-correlation coefficient of 94% between source and predicted PV sequences on the training set and of 89% on the test set. The RMSEs are of 7% and 9% on the training and test sets, respectively. In Figures 4 and 5 we report some predictions obtained by the model on test data. Also in this case, in both cases the model permits to predict the concentrations with high accuracy.

5 Conclusions

In this work, a methodology to PV forecasting based on a neural network, trained in a Bayesian framework, has been introduced. More specifically, a Multi-Layer

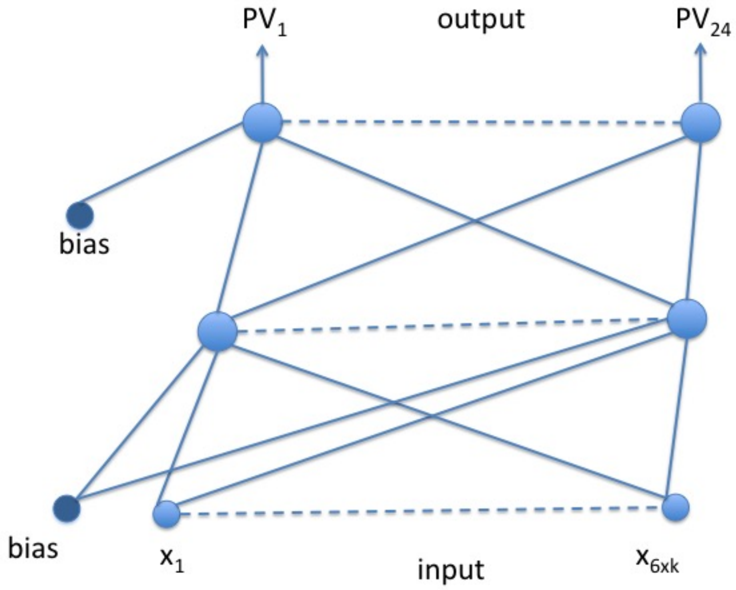

Fig. 1. MLP for 24 hours ahead prediction.

Perceptron Neural Network for 24 or 48 hours ahead prediction is used whose parameters are estimated by a probabilistic Bayesian learning technique. The experimental results show a low percentage of error on test data. In the next future the work will be focused on 72-hours prediction and the use of other Neural Network based methodologies such as Radial Basis Functions NNS and Recurrent NNs [8].

References

1. E. B. Balogun, X. Huang, D. Tran, Comparative Study of Different Artificial Neural Networks Methodologies on Static Solar Photovoltaic Module, International Journal of Emerging Technology and Advanced Engineering, Volume 4, Issue 10, 2014.
2. C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
3. C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
4. S. Cao, J. Cao, Forecast of solar irradiance using recurrent neural networks combined with wavelet analysis, Applied Thermal Engineering 25, 161-172.
5. Li Deng, O. S. Douglas, Speech processing: a dynamic and optimization-oriented approach. Marcel Dekker. pp. 4148. ISBN 0-8247-4040-8, 2003.

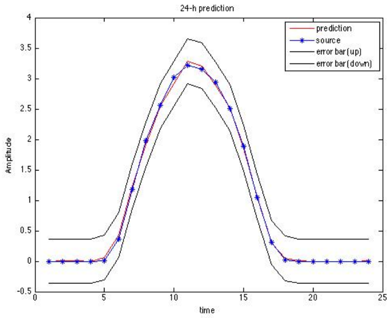

Fig. 2. 24-hour PV prediction obtained by MLP NN and Bayesian framework.

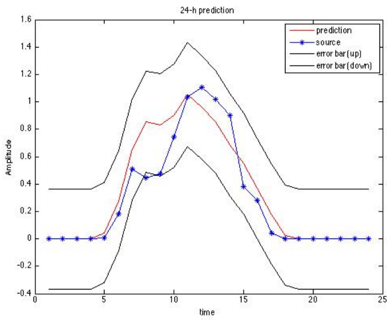

Fig. 3. 24-hour PV prediction obtained by MLP NN and Bayesian framework.

6. R. Deshmukh, R. Bharvirkar, A. Gambhir, A. Phadke, C. Sunshine, Analyzing the dynamics of solar electricity policies in the global context, Renewable and Sustainable Energy Reviews, Volume 16, Issue 7, Pages 5188-5198, ISSN 1364-0321, 2012.
7. D. Elizondo, G. Hoogenboom, R. W. Mcclendon, Development of a neural network model to predict daily solar radiation, Agricultural and Forest Meteorology 71, 115132.
8. S. Haykin, Neural Networks - A Comprehensive Foundation, 2nd Ed., Prentice Hall, 1999.
9. D. J. C., MacKay, Hyperparameters: optimise or integrate out? Maximum entropy and Bayesian methods, Dordrecht, 1993.
10. Mahoney WP, Parks K, Wiener G, Liu Y, Myers B, Sun J, Delle Monache L, Johnson D, Hopson T, Haupt SE, A Wind Power Forecasting System to Optimize Grid Integration, special issue of IEEE Transactions on Sustainable Energy on Applications of Wind Energy to Power Systems, 3 (4), 670-682, 2012.
11. Marquis M, Wilczak J, Ahlstrom M, Sharp J, Stern A, Charles Smith J, Calvert S, Forecasting the Wind to Reach Significant Penetration Levels of Wind Energy, Bull. Amer. Meteor. Soc., 92, 11591171, 2011.
12. A. Mellit, A. M. 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, 84, 807821, 2010.

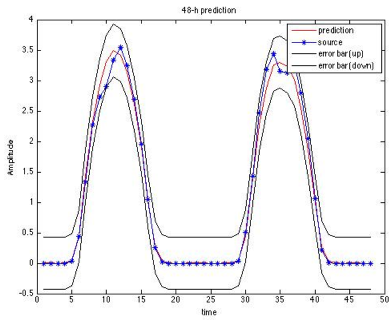

Fig. 4. 48-hour PV prediction obtained by MLP NN and Bayesian framework.

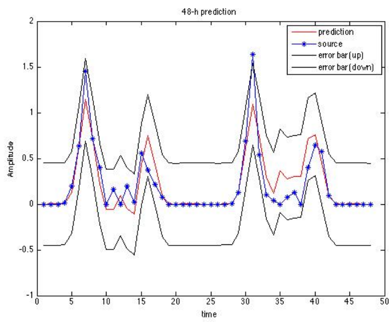

Fig. 5. 48-hour PV prediction obtained by MLP NN and Bayesian framework.

13. A. Mellit, Artificial intelligence techniques for modelling and forecasting of solar radiation data: a review. International journal of Artificial Intelligence and Soft Computing 1, 5276.
14. R. M. Neal, Bayesian learning for neural networks, Berlin: Springer, 1996
15. Pielke RA, Cotton WR, Walko RL et al. A Comprehensive Meteorological Modeling System RAMS, Meteorology and Atmospheric Physics, 49, 69.
16. Z. Sen, Fuzzy algorithm for estimation of solar irradiation from sunshine duration, Solar Energy 63, 3949.
17. A. Sfetsos, A. H Coonick, Univariate and Multivariate forecasting of hourly solar radiation with artificial intelligence techniques, Solar Energy 68, 169178.
18. J. Shi, W.-J. Lee, Y. Liu, Y. Yang, P. Wang, Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines, IEEE Transactions on Industry Applications, vol. 48, n. 3, 2012.
19. Early data on 2013 electricity demand: 317 billion KWh of demand, −3.4% compared to 2012 Terna (company) press release, `http://www.terna.it/LinkClick.aspx?fileticket=GjQzJQkNXhM3d&tabid=901&mid=154`