Renewable Energy 85 (2016) 631–641

Contents lists available at ScienceDirect

Renewable Energy

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

An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in Utrecht, the Netherlands

A.G.R. Vaz [a, b], B. Elsinga [b], W.G.J.H.M. van Sark [b], M.C. Brito [c, *]

[a] Faculdade de Ciências da Universidade de Lisboa, Department of Geographic Engineering, Geophysics and Energy, 1749-016 Lisboa, Portugal
[b] Utrecht University, Copernicus Institute of Sustainable Development, Energy & Resources, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands
[c] Instituto Dom Luiz, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 13 March 2015
Received in revised form
4 June 2015
Accepted 25 June 2015
Available online xxx

Keywords:
Photovoltaics
Artificial neural network
NARX model
Time series
Forecasting

ABSTRACT

In order to perform predictions of a photovoltaic (PV) system power production, a neural network architecture system using the Nonlinear Autoregressive with eXogenous inputs (NARX) model is implemented using not only local meteorological data but also measurements of neighbouring PV systems as inputs. Input configurations are compared to assess the effects of the different inputs. The added value of the information of the neighbouring PV systems has demonstrated to further improve the accuracy of predictions for both winter and summer seasons. Additionally, forecasts up to 1 month are tested and compared with a persistence model. Normalized root mean square errors (nRMSE) ranged between 9% and 25%, with the NARX model clearly outperforming the persistence model for forecast horizons greater than 15 min.

1. Introduction

The challenge for electrical grid operators is to continuously synchronize electrical energy demand and supply. As global demand for renewable energy is increasing, the economic and technical issues of photovoltaic (PV) solar power penetrations into the power grid must be addressed. Especially since natural variability of the solar resource, seasonal deviations in production and the high cost of energy storage raises concerns regarding reliability and feasibility of solar power systems. This is due to the fact that solar energy is highly dependent on weather conditions including cloud structure and day/night cycles. Clouds can cause significant ramps in solar insolation and PV output, which may be difficult to handle by the grid operator. Therefore, integration of electricity produced by solar power systems requires accurate solar energy forecasts.

Solar energy forecasts allow grid operators to adapt the load in order to optimize the energy transport, allocate the needed balance energy from other sources if no solar energy is available and plan maintenance activities at the production sites. Accurate solar forecasting methods improve the quality of the energy delivered to the grid and reduce the additional cost associated with weather dependency. The combination of these two factors has been the main motivation behind considerable research activities in solar forecasting.

Linear models, such as Box–Jenkins and autoregressive integrated moving average (ARIMA) type models are regularly used to generate forecasts. They assume linear correlation structures among the time series values and thus no nonlinear patterns can be captured Zhang [30]. Subsequently, Reikard [22], Paoli et al. [25], and Pedro & Coimbra [23] used nonlinear models that show more flexibility in capturing the data underlying characteristics and those nonlinear models outperformed linear models. Moreover, at shorter time interval (less than 1 h), short-term patterns dominate and Artificial Neural Networks (ANN, see Section 3 for definitions and properties) demonstrated good results in solar forecasting Diagne et al. [5].

In Ref. [26] ANN were used to perform one-step ahead forecasting of hourly values of global irradiance and they revealed that those results outperform linear models results. They also compared various models in terms of error and training time and found that the Levenberg–Marquardt algorithm achieved the best performance.

In Ref. [29] a comparative study between different ANN models was conducted to predict insolation 1-day ahead, in which the recurrent neural network outperformed the *feedforward* neural

* Corresponding author.
E-mail address: mcbrito@fc.ul.pt (M.C. Brito).

http://dx.doi.org/10.1016/j.renene.2015.06.061
0960-1481/

network. Additionally, in Ref. [13] the researchers implemented a multilayer neural network for half hour cloudiness forecasting and considered it an important tool for the estimation of cloudiness affecting solar radiation.

ANN forecasting models for hourly solar irradiation for times of up to 6 days ahead were tested in Refs. [14]; concluding that the developed intelligent models outperformed satellite-based models. Moreover, an input selection scheme was used and results revealed that models with slightly larger sets of inputs generally perform better for same-day and 1-day ahead forecasts.

In Ref. [4] forecasting the daily solar radiation with two dynamic artificial neural networks (*Feedforward* Time Delay Neural Network and Nonlinear Autoregressive with eXogenous inputs (NARX)) was proposed. Both models had a satisfactory performance, facilitating energy management of solar systems when storage systems are adopted.

Several studies that compare ANN with other simpler time series techniques (AR, ARIMA, etc.) have been conducted in the past. In Ref. [10] different solar forecasting techniques' performances are assessed based upon a forecasting skill given as

$$s = 1 - \frac{RMSE}{RMSE_p} \quad (1)$$

where RMSEp denotes the root mean square error of the persistence model. Thus, a forecasting skill closer to 1 shows that the model being assessed has significantly improved the accuracy relatively to the persistence model.

Bacher et al. [1] compare an AR and ARX model for short-term and medium-term solar forecasting and conclude that the most important input is the lagged PV values for short-term horizon (2 h) and, for long horizon NWP models become more important. The authors showed a forecasting skill *s* of 0.27 and 0.34 for a horizon of 1 h–6 h for the AR and ARX model, respectively. Paoli et al. [22] presented a forecasting skill *s* of 0.19 and 0.20 for AR and ANN respectively, for one day ahead. The forecasting skill in Ref. [16] for intra-day forecasts was 0.16 and 0.17 for an AR and ANN model, respectively; for day ahead, the forecasting skill was similar varying between 0.18 and 0.20 for the AR and ANN model respectively. Voyant et al. [28] suggests that an ANN model for daily forecasts on 6 months-cloudy period improves the power production prediction by 9% and 1% relatively to the persistence and ARMA model, respectively. Pedro and Coimbra [23] also presented a comparison between ARIMA and ANN to predict 1 h and 2 h average power output; the forecasting skill for the 1 h forecasts were 0.02 and 0.18, and for the 2 h forecasts were 0.10 and 0.11, for the ARIMA and ANN model respectively; thus showing neural networks have greater potential for short-term forecasting.

The present work improves on A.G.R. Vaz [27] and uses an ANN model to capture the short-term (15 min) ramping patterns caused by cloud formations and to forecast a PV system power output up to 1-month ahead. Moreover, using different input combinations, we assess whether or not solar power forecasts can be improved by knowing beforehand the power output of other neighbouring (few km distance) grid-connected PV systems and meteorological information. Additionally, the forecasting accuracy of the ANN is compared to the persistence model. In Section 3, principles of ANNs are briefly discussed, and in Section 4 the used methodology is presented.

2. Clear sky persistence model

The persistence model is a simple forecasting model that requires knowledge of clear sky irradiance. Usually, this simple forecasting tool is very accurate for very short-time horizons and for low irradiance variability. The model has the clear sky conditions persist for the next time-step and meets the definition of Marquez and Coimbra [15] applied to the power production of a PV system,

$$\hat{k}^*(t + \Delta t) = \hat{k}^*(t) = \frac{PVPP(t)_{measured}}{PVPP(t)_{clr}} \quad (2)$$

$$PVPP(t + \Delta t) = \hat{k}^*(t + \Delta t) \times PVPP(t + \Delta t)_{clr} \quad (3)$$

where $\hat{k}^*$ is defined as clear-sky index, t denotes the time instant, $PVPP(t)$ is the measured photovoltaic power production and $PVPP(t)_{clr}$ is the photovoltaic power production for clear sky conditions, calculated according to Ineichen and Perez [9]. Other clear sky models are possible and Gueymard [7] is suggested for thorough analyses of different models.

3. Artificial neural networks

3.1. Definitions and properties

In its most general form, an ANN is a machine that models a task or function of interest, performing useful computation through a process of *learning*. In fact, the artificial neural network derives its computing power through its massively parallel distributed structure and its ability to learn and generalize, which means finding reasonable outputs whenever inputs are not encountered during training (learning) [8].

The ANN consists of simple processing units, the neuron, and directed, weighted connections between those neurons. The inputs channels have an associated weight, such that the incoming information x_i is multiplied by a corresponding weight w_i. The network input is the result of the so-called propagation function. Here, the strength of a connection between two neurons *i* and *j* is a connecting weight w_{ij}. Experimental knowledge, acquired by the network through a learning process, is stored by massively interconnecting these units (synaptic weights). These connecting weights can be inhibitory or excitatory and by being connected with the neurons, data are transferred.

The output is a function of the particular activation function chosen and a possible bias. The latter is similar to a weight, albeit it has a constant input of 1. This bias term is used by the neuron to generate an output signal in the absence of input signals. Fig. 1 illustrates the nonlinear model of a neuron [8].

The transfer function or activation function controls the amplitude of the output of the neuron and is based on the neuron reactions to the input values and depends on the level of activity of the neurons (activation state). Essentially, neurons are activated

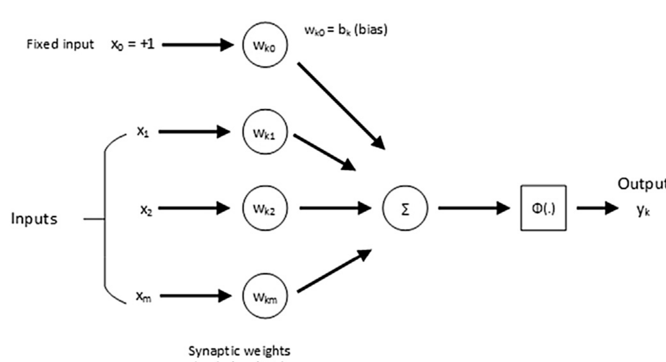

Fig. 1. Nonlinear model of a neuron (Redrawn from Ref. [8]).