Accepted Manuscript

Title: An Artificial Neural Network-based Forecasting Model of Energy-Related Time Series for Electrical Grid Management

Authors: Annalisa Di Piazza, Maria Carmela Di Piazza, Giuseppe La Tona, Massimiliano Luna

PII: S0378475420301695

DOI: http://dx.doi.org/doi:10.1016/j.matcom.2020.05.010

Received date: 29-10-2019
Revised date: 23-04-2020
Accepted date: 06-05-2020

An Artificial Neural Network-based Forecasting Model of Energy-Related Time Series for Electrical Grid Management

A. Di Piazza, M. C. Di Piazza[a,*]*, G. La Tona[a], M. Luna[a]

Consiglio Nazionale delle Ricerche (CNR), Istituto di Ingegneria del Mare (INM), Palermo, Italy

Abstract

Forecasting of energy-related variables is crucial for accurate planning and management of electrical power grids, aiming at improving overall efficiency and performance. In this paper, an artificial neural network (ANN)-based model is investigated for short-term forecasting of the hourly wind speed, solar radiation, and electrical power demand. Specifically, the non-linear autoregressive network with exogenous inputs (NARX) ANN is considered, compared to other models, and then selected to perform multi-step-ahead forecasting. Different time horizons have been considered in the range between 8 and 24 hours ahead. The simulation analysis has put in evidence the main advantage of the proposed method, i.e., its capability to reconcile good forecasting performance in the short-term time horizon with a very simple network structure, which is potentially implementable on a low-cost processing platform.

1 Introduction

Forecasting of solar radiation and wind speed lets energy stakeholders estimate the output power of a PV/wind generator in advance; this can help to reduce the effects of power fluctuations on the power grid, thus allowing the overall efficiency and power quality of the plant to be increased [1]–[4]. Similar considerations apply to power load demand, considered one of the most important tasks in power systems operations since its accuracy has an impact on both operational security and economics of the power grids [5].

Optimal energy management and economic planning of a smart grid (SG), including renewable distributed generation, can then be aided by a robust and reliable energy-related data forecasting to optimize decision making at the grid operator level. Besides, both the renewables and the load demand forecasting are useful to give the input to intelligent Energy Management Systems (EMSs), considered today one of the most significant techniques which can enhance the efficiency, reliability, and economy of the SGs [6]–[9].

Many contributions on both wind speed and solar radiation forecasting are present in the technical literature [9]–[13]. Moreover, several contributions have been proposed on the forecasting of energy from renewables and load profiles associated with smart grid management [9], [14]–[16]. As for load demand forecasting, the deep learning approach, together with specific machine learning based approaches, such as support vector machines (SVM) [17], has been often proposed in recent years to obtain more accurate results using, for example, convolutional and recurrent neural networks [18], [19]. However, despite the high performance exhibited by the deep learning approach, it should be observed that it is well suited to very large and complex datasets [18] and this can be an issue if the available dataset for learning is small, as it usually occurs in the case of electrical power production/consumption datasets.

According to a general classification, forecasting methods can be divided into five families: 1. persistence method, also known as Naïve predictor; 2. physical models; 3. time series linear models; 4. Artificial Intelligence (AI)-based models, including machine learning-based approaches; 5. hybrid structures [9]–[12], [20]. In general, the need for a trade-off between good forecasting results for a given time horizon and a simple implementation of the forecasting technique is recognizable. In [21], the suitability of the different models for performing forecasting at specific time scales has been established. Naïve predictor assumes that the variable to be forecast at time '$t_0+\Delta t$' will be the same as it was at time 't_0'. Despite its simplicity, this method is very accurate but only for short/very-short term forecasting (from few seconds to 30 minutes ahead); physical models, such as Numerical Weather Predictors (NWP) are based on complex mathematical models using several weather-related data; for this reason, they require significant computational resources and usually are run no more than one or two times a day. As a result, physical models are useful only for medium/long-term forecasts (i.e., for time horizons exceeding 6 hours ahead); additionally, they produce best results when weather conditions are stable. Time series linear models are generally easy to implement, inexpensive, and provide timely results; they are based on error minimization between the model-based predictions and historical trends and work well for short term horizons (from 30 minutes to 6 hours); on the other hand, they are unsuitable for modelling strongly nonlinear phenomena.

[*] Corresponding author at: INM-CNR, via Ugo La Malfa 153, 90146, Palermo, Italy.
E-mail address: mariacarmela.dipiazza@cnr.it

AI-based and hybrid algorithms have proved to be more effective than other classical autoregressive predictors, due to the behaviour of electric production/consumption-related data that is very much non-linear and seasonal [19]. Indeed, such methods generally outperform linear time series-based models for different time horizons from very short to medium [9], but their formulation and parameter tuning is a challenging task.

For short/medium term (30 minutes- up to 1 day-ahead), several approaches have given satisfactory results, so the choice of the most suitable approach to use is not straightforward. Among them, in general, artificial neural networks (ANNs) have provided accurate forecasting results. This behavior has been assessed for both solar radiation and wind speed modeling [22], [23],[24], as well as for load power demand [25]. Based on the above considerations, in this paper, the *nonlinear autoregressive network with exogenous inputs* (NARX) ANN has been considered for performing the forecasting of wind speed, solar radiation, and load power demand. This network has been chosen because it is well suited to model nonlinear dynamic systems, such as dependencies among meteorological time series; in addition, its use in time-series-based models has proved to be successful thanks to its capability to suitably reproduce relationships among data even if a small dataset is available for learning [26]. Furthermore, it combines the computational power of fully recurrent ANNs with the advantage of a simpler structure due to the feedback being limited only to the output neuron, rather than extended to the hidden states [27]. In this paper, the NARX network is first compared with some common time series linear models to assess its superior capability to learn and reproduce the behavior of the considered variable. Then, considering a case study involving an extensive use of electrical heating/cooling devices in the power grid, such an ANN is used for forecasting the hourly wind speed, solar radiation, and load demand time series, by using the environmental temperature as the ANN's exogenous input.

The main contribution of the proposed forecasting technique is the simplicity of the chosen network and its high performance, which gives comparable or better results with respect to time series-based models, in particular the vastly used linear models. This result is achieved without resorting to more complex methods such as ANNs with multiple recurrent layers, which require very large datasets, or complex pre-processing procedures [14], [18], [20], [28], [29]. Therefore, the implementation of the proposed network is also feasible on low-cost hardware platforms, such as small single-board computers (e.g., Raspberry Pi boards). As such, it can ease the use of forecasting processes in energy management schemes that optimize the efficiency and performance of the smart grids. Moreover, in addition to the forecasting method and its assessment, this paper provides an evaluation of the NARX capability to handle different forecasting horizons for the two considered variables. The obtained results allow finding the optimal time horizon for each variable and assessing whether its forecasting by NARX is suitable to the decision processes and actions on which smart grid management is based (e.g., regulation, load dispatch, maintenance, etc.).

2 Case Studies and Data Description

The meteorological dataset used in this study for solar radiation forecasting comes from a station placed in northwestern Sicily (Italy) and has been provided by SIAS (*Servizio Informativo Agrometeorologico Siciliano*). The used dataset is complete and contains hourly global solar radiation (MJ/m^2) and hourly maximum/minimum temperature recorded for seven consecutive years; for this study, the hourly mean temperature has been calculated and used.

The meteorological dataset used for wind speed forecasting comes from a publicly available database published by the U.S. National Renewable Energy Laboratory (NREL) [30]. Hourly wind speed (m/s) and outdoor temperature values have been extracted from this database considering a period of seven consecutive years. Specifically, the data used for simulations refer to the city of Amarillo, Texas, USA.

As for the load demand dataset, the public dataset for the Global Energy Forecasting Competition 2012 [31] is used. This dataset consists of 4.5 years of hourly load power demand (in kW) for a US utility and temperature history data. The load power demand profiles are reported for 20 zones, and one of those is considered for the case study.

3 Performance Indices

The Normalized Root Mean Square Error (NRMSE) and the Coefficient of Variation of the Root Mean Squared Error CV(RMSE) are used in this study to define the deviation between observed and estimated values. They are defined respectively as:

$$NRMSE = N^{-1}\left(Y_{max} - Y_{min}\right)^{-1}\sqrt{\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2} \tag{1}$$

$$CV(RMSE) = N^{-1}\overline{Y}^{-1}\sqrt{\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2} \tag{2}$$

where Y_i is the observed sample of the time series, $\hat{Y}_i$ is the predicted sample of the time series, Y_{max}, Y_{min} are the maximum and minimum observed values, and $\bar{Y}$ is the mean of the observed values. In this paper, the tables that show the performance indices report the above described indices with subscripts *e, v, t, r,* and *f* that indicate, respectively, estimation set, and validation set for the time series linear models, training phase, and recalling phase (application of models to validation set) for the neural models, and forecasting phase for both time series linear models and neural models. With specific reference to the forecasting results, they are evaluated considering an additional index, i.e., the *mean absolute percentage error* (MAPE), which provides information on the short-term performance and represents a measure of the variation of the forecast values around the measured data. The MAPE is defined according to equation (3):

$$MAPE = 100\, N^{-1} \sum_{i=1}^{N} \left| Y_i - \hat{Y}_i \right| Y_i^{-1} \tag{3}$$

4 The NARX Model

The NARX neural network is derived by a class of discrete-time nonlinear systems, i.e., the nonlinear autoregressive with exogenous input (NARX) models. It is well suited to model nonlinear dynamic systems. In a dynamic system, the values of the output signals depend on both the instantaneous values of its input signals and the past behavior of the system [32]. The NARX network has been used in time-series modeling thanks to its adaptive learning process which works very well also with small-scale data, e.g., collected in less than one year [26].

The mathematical formulation of the NARX model is expressed in (4):

$$y(t) = F\left(u(t-n_k), u(t-n_k-1), \ldots, u(t-n_k-n_u), y(t-1), \ldots, y(t-n_y)\right) \tag{4}$$

where $y(t)$ and $u(t)$ are the output and the input of the model at a discrete time step t, $n_y \geq 1$, $n_u \geq 1$, $n_u \leq n_y$ are the input and output memory orders (delay), $n_k \geq 0$ is the number of input samples after which the output is affected by the input, and F is a nonlinear mapping function. When the function F is approximated by a multilayer perceptron (MLP), the resulting neural network is called a NARX network. In other words, a NARX network consists of a MLP that takes as input a window of past independent (exogenous) inputs and past outputs and then calculates the current output. It is possible to introduce x as the vector of the state variables, so that $x_i(t)$ is the *i-th* state variable in the NARX network. Then, it is possible to say that the states of the NARX, given by a set of two tapped-delay lines (i.e., n_u taps on the input values and n_y taps on the output values), are updated according to the law described in (5):

$$x_i(t+1) = \begin{cases} u(t-n_k) & i = n_u \\ y(t) & i = n_u + n_y \\ x_{i+1}(t) & 1 \leq i < n_u \ or \ n_u < i < n_u + n_y \end{cases} \tag{5}$$

Therefore, at the time t, the taps correspond to the values given by (6):

$$\boldsymbol{x}(t) = [u(t-n_k-1), \ldots, u(t-n_k-n_u),\ y(t-1), \ldots, y(t-n_y)] \tag{6}$$

The MLP is a network organized into two layers (hidden and output layers of the NARX). The N_k nodes of the hidden layer perform the function (7)

$$z_i(t) = x_i(t+1) = \sigma\left[\sum_{j=1}^{N} a_{i,j} x_j(t) + b_i u(t) + c_i\right] \quad i = 1, \ldots, N_k \tag{7}$$

where $a_{i,j}$, b_i and c_i are fixed real-valued weights, σ is the sigmoid function, and N is the number of state variables.

Finally, the output layer is formed of a single linear node defined as in (8).

$$y(t) = \sum_{j=1}^{N_k} w_{i,j} z_j(t) + c_0 \tag{8}$$

The activation function of the hidden neurons is the usual sigmoid function σ, which approximates the Heaviside step function to assess whether its input is above or below a threshold. On the other hand, the activation function of the output neuron is linear because a continuous output is desired.

Unlike a conventional recurrent neural network, the NARX network has limited feedback coming only by the output neuron rather than by the hidden states. Only the output of the NARX is fed back to the input of the feed-forward neural network. Nevertheless, it has been demonstrated that such a neural network is as computationally powerful as a fully-connected recurrent neural network [33]. The NARX block scheme is illustrated in Fig. 1.

The NARX network is trained in *open loop* removing the feedback from the output neuron and presenting it with observed data of the considered time series (wind speed, solar radiation, and load power demand) and of the exogenous variable (temperature). This means that the network can be trained like a classic feed-forward network using the backpropagation algorithm instead of the more complicated backpropagation through time algorithm [33], which considers the recursion of the output. Furthermore, another advantage is that with the *open loop* configuration the training is more accurate since it only considers observed values.

In order to exploit these two advantages, the following approach is used in the application under study. Firstly, the network has been implemented as an open loop network, as shown in Fig. 2, which is used for learning the wind speed/solar radiation/power demand versus temperature behavior and for evaluating its performance for the prediction of independent variables (wind speed, solar radiation, power demand) in the recalling phase. Then, the NARX structure, shown in Fig. 1, is used to obtain the forecasting, according to the multi-step-ahead approach by iterating the execution in closed loop for a number of times corresponding to the desired forecasting horizon.

The use of temperature as exogenous input for load power demand forecasting deserves a further explanation. This study considers an electrical power consumption profile framed in a decarbonization scenario where massive use of electrical heating/cooling devices such as heat pumps is envisaged [34] [35]. In such a scenario, that is going to be established worldwide to pursue the goals of Paris climate agreement, the electrical power load has a strong dependence on environmental temperature.

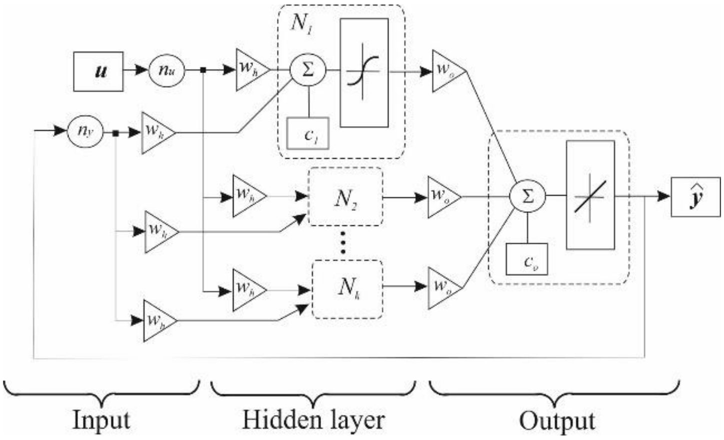

Fig. 1 Structure of the NARX network

4.1. The training algorithm

The training method is the conjugate gradient backpropagation with Polak-Ribiére algorithm [27]. This algorithm has been chosen because it converges faster than other more widespread algorithms, such as gradient descent and gradient descent with momentum. In the case under study, the value of the learning rate, initially equal to 0.01, decreases according to an exponential law. The epochs used for the training phase are determined using the early stopping approach [36], and the best model is defined evaluating the performance indices defined in Section 3, calculated both in training and recalling phases. In order to make the ANN training more efficient, a normalization step is applied to both the input vectors and the output vectors in the datasets; then, a post-processing procedure is applied to the network outputs transforming them back into the units of the original data.

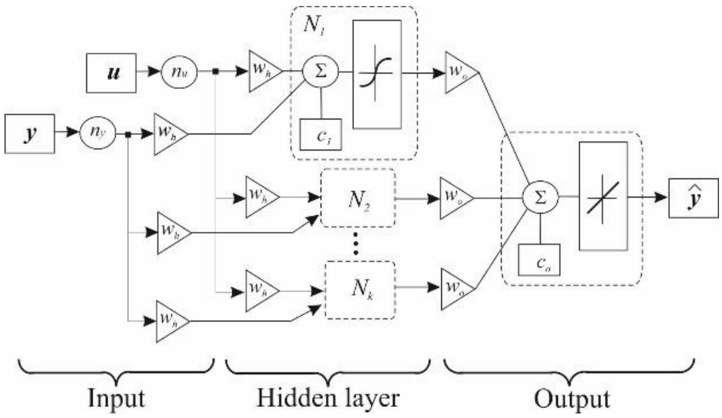

Fig. 2 Structure of the “open loop” NARX network, used for training

It should be observed that the parameter n_k is usually set to zero in the literature. However, this would be a severe constraint, because it would imply the need to know future values of temperature in the forecasting horizon. To avoid this problem, n_k was set equal to the considered forecasting horizon.

In order to define the best set of NARX parameters (delay, number of neurons in the hidden layer), several simulations have been run. In particular, simulations have been performed with the delay ranging between 1 and 7 and the number of neurons in the hidden layers ranging between 3 and 10. Each dataset has been split into two parts: a training set using 80% of data and a validation set using the remaining 20%. A pre-processing procedure (i.e., detrending) has been applied to the variables to focus their dynamics better. The flow chart in Fig. 3 describes the approach followed to tune the parameters of the NARX network.

4.2. ANN structure and parameters

The open loop structure is used to perform the time series prediction of wind speed, solar radiation, and load power demand hourly profiles. It is worth observing that the ability to predict a given variable is intended as the ability to correctly learn and reproduce the behavior of the considered variable during the validation (open loop execution) phase.

As for the solar radiation prediction, the best configuration is 2-5-1 (where the three digits are the number of neurons in the input, hidden, and output layer, respectively) with a time delay of 7, and it exhibits the following values of the performance indices: $NRMSE_t$=0.062, $NRMSE_r$=0.063, $CV(RMSE)_t$=0.34, and $CV(RMSE)_r$=0.34. The performance of the NARX for this case is defined by the indices given in Table 1.

Table 1 Performance indices on training and validation datasets for the NARX in predicting solar radiation

Neural network structure	Delay	$NRMSE_t$	$NRMSE_r$	$CV(RMSE)_t$	$CV(RMSE)_r$
2-3-1	1	0.096	0.096	0.522	0.528
2-3-1	3	0.071	0.071	0.383	0.391
2-3-1	5	0.068	0.069	0.370	0.381
2-3-1	7	0.067	0.068	0.366	0.374
2-5-1	1	0.095	0.095	0.517	0.518
2-5-1	3	0.069	0.069	0.377	0.378
2-5-1	5	0.064	0.064	0.348	0.351
2-5-1	7	**0.062**	**0.063**	**0.340**	**0.340**
2-10-1	1	0.095	0.095	0.516	0.518
2-10-1	3	0.066	0.067	0.361	0.365
2-10-1	5	0.062	0.063	0.342	0.342
2-10-1	7	0.064	0.064	0.345	0.350

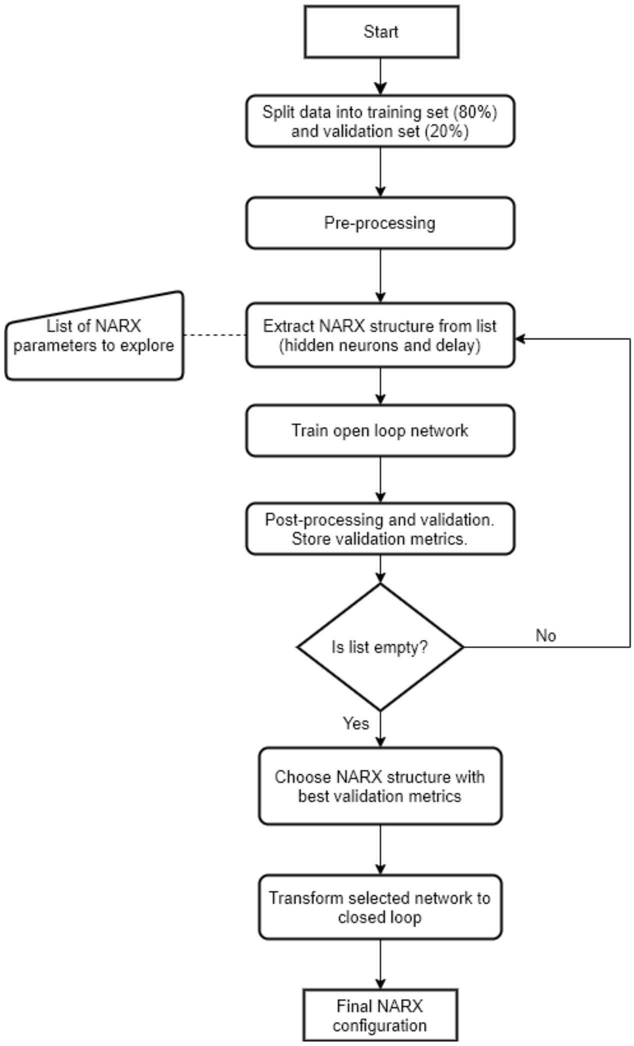

Fig. 3 Approach followed to tune the parameters of the NARX network

Table 2 shows that, for the wind speed case, the best configuration is 2-5-1 with a delay of 7, exhibiting the following values of the performance indices: $NRMSE_t$=0.0236, $NRMSE_r$=0.0297, $CV(RMSE)_t$=0.0943, and $CV(RMSE)_r$=0.0911.

As far as the load demand is concerned, the best performing configuration is 2-10-1 with a time delay of 7. As shown in Table 3, the performance indices for this case are the following: $NRMSE_t$=0.0217, $NRMSE_r$=0.024, $CV(RMSE)_t$=0.0345, and $CV(RMSE)_r$=0.0368.

Table 2 Performance indices on training and validation datasets for the NARX in predicting wind speed

Neural network structure	Delay	$NRMSE_t$	$NRMSE_r$	$CV(RMSE)_t$	$CV(RMSE)_r$
2-3-1	1	0.0344	0.0439	0.1375	0.1345
2-3-1	3	0.0247	0.0312	0.0989	0.0956
2-3-1	5	0.0243	0.0306	0.0972	0.0938
2-3-1	7	0.0243	0.0306	0.0973	0.0938
2-5-1	1	0.0344	0.0439	0.1376	0.1347
2-5-1	3	0.0245	0.0307	0.0978	0.0942
2-5-1	5	0.0243	0.0306	0.0970	0.0937
2-5-1	7	**0.0236**	**0.0297**	**0.0943**	**0.0911**
2-10-1	1	0.0347	0.0445	0.1388	0.1364
2-10-1	3	0.0242	0.0304	0.0967	0.0933
2-10-1	5	0.0241	0.0304	0.0966	0.0932
2-10-1	7	0.0237	0.0298	0.0950	0.0914

Table 3 Performance indices on training and validation datasets for the NARX in predicting power load demand

Neural network structure	delay	$NRMSE_t$	$NRMSE_r$	$CV(RMSE)_t$	$CV(RMSE)_r$
2-3-1	1	0.0375	0.0389	0.0595	0.0597
2-3-1	3	0.0257	0.0283	0.0408	0.0434
2-3-1	5	0.0245	0.0266	0.0389	0.0409
2-3-1	7	0.0242	0.0272	0.0384	0.0416
2-5-1	1	0.0347	0.0381	0.0550	0.0584
2-5-1	3	0.0255	0.0283	0.0404	0.0434
2-5-1	5	0.0250	0.0277	0.0396	0.0425
2-5-1	7	0.0228	0.0248	0.0362	0.0380
2-10-1	1	0.0341	0.0373	0.0541	0.0572
2-10-1	3	0.0247	0.0272	0.0392	0.0417
2-10-1	5	0.0229	0.0251	0.0363	0.0385
2-10-1	7	**0.0217**	**0.0240**	**0.0345**	**0.0368**

5 Prediction by Linear Approaches

In order to benchmark the proposed approach in terms of performance indices, predictions of wind speed, solar radiation, and load power demand for the case study, have also been performed by some common time series linear models.

The most commonly used models for time series data are the autoregressive processes, also known as linear time series models. In particular, in this study, the three couples of related variables (wind speed vs. temperature, solar radiation vs. temperature, and load power demand vs. temperature) are taken into account and studied as dynamic systems.

The linear models considered for this study are autoregressive exogenous models. In particular, the following processes have been used: ARX model, ARMAX model, Box-Jenkins (BJ) model, and Output-Error (OE) model [37], [38]. The parameters of the models have been estimated using the corresponding training set used for the NARX (80% of each dataset). Then, validation of the models has been performed using the corresponding validation set used for the NARX (the remaining 20%). Table 4 summarizes the obtained performance indices for each of the considered models, for solar radiation, wind speed, and load power demand prediction. From the comparison of the performance indices obtained by different methods, it is possible to observe that the NARX network, in its best performing configurations, outperforms all the considered time series linear models at predicting the one-step-ahead evolution of all the considered time series. As far as the NARX performance is concerned, it is possible to note that an improvement of about 8.8% on NRMSE and of about 49.4% on CV(RMSE) is achieved in solar radiation prediction in comparison with the best performing linear model (ARX). Moreover, a slight reduction of about 0.1% on NRMSE and of about 0.2% on CV(RMSE) is obtained in wind speed prediction in comparison with the best performing linear model (BJ). Finally, with reference to load demand prediction, an improvement of about 0.4% on NRMSE and of about 0.6% on CV(RMSE) is observed with respect to the results obtained by the best performing linear model (BJ).

Table 4 Performance indices on estimation and validation datasets for the time series linear models

	Linear Models	$NRMSE_e$	$NRMSE_v$	$CV(RMSE)_e$	$CV(RMSE)_v$
Radiation prediction	ARMAX	0.174	0.186	0.943	1.033
	BJ	0.149	0.167	0.808	0.926
	ARX	**0.144**	**0.151**	**0.790**	**0.834**
	OE	0.149	0.167	0.808	0.840
Wind speed prediction	ARMAX	0.0243	0.306	0.0971	0.0940
	BJ	**0.0241**	**0.0305**	**0.0965**	**0.0937**
	ARX	0.0260	0.0326	0.1038	0.0999
	OE	0.117	0.154	0.470	0.474
Load demand prediction	ARMAX	0.0263	0.0296	0.0429	0.0453
	BJ	**0.0252**	**0.0279**	**0.0410**	**0.0427**
	ARX	0.0264	0.0298	0.0429	0.0456
	OE	0.1680	0.1924	0.2738	0.2948

6 Forecasting Results

Once the effectiveness of the NARX network in predicting the one-step ahead evolution of all the considered variables has been demonstrated, this ANN has been used for performing the forecasting (closed loop execution) of the same variables using the validation datasets (data not seen during training). For each considered time series, the NARX network with the best performing structure (delays and numbers of neurons) as defined in Section 4.1, is then used for this purpose, and compared to best performing linear model. Table 5 shows the selected linear models and NARX structures.

Table 5 Selected linear models and NARX structures for assessing forecasting results

	Selected Linear Model	Selected NARX structure
Radiation	ARX	2-5-1 delay 7
Wind Speed	BJ	2-5-1 delay 7
Load Demand	BJ	2-10-1 delay 7

Several time horizons ranging from 8 to 24 hours ahead have been explored. In the case of forecasting results, the indices selected to measure the performance of the network are the NRMSE, the CV(RMSE), and the MAPE. Table 6 and Table 7 show the forecasting results obtained with the selected NARX structure and linear model respectively for each time series at different time horizons. It is possible to observe that the best forecasting for solar radiation is obtained by NARX for 12-hours-ahead with the following values: {$NRMSE_f$=0.087, $CV(RMSE)_f$= 0.627} @12h. Furthermore, the best forecasting results for wind speed are found by NARX for 8-hours-ahead time horizon with the following values: {$NRMSE_f$=0.136, $CV(RMSE)_f$= 0.288} @8h. Finally, for the load power demand, the best forecasting results are found by NARX for 8-hours-ahead time horizon with the following values: {$NRMSE_f$ = 0.177, $CV(RMSE)_f$ = 0.113} @8h.

Fig. 4 shows the 12-hours-ahead forecasting result for solar radiation; Fig. 5 shows the 8-hours-ahead forecasting result for wind speed, and the 8-hour-ahead forecasting result for power load demand is illustrated in Fig. 6.

Table 6 Performance indices for the selected NARX in forecasting for several time horizons

	Hours ahead forecast (h)	$NRMSE_f$	$CV(RMSE)_f$	MAPE (%)
Radiation prediction	8	0.113	0.816	21.0
	12	**0.087**	**0.627**	**15.2**
	24	0.136	0.976	30.7
Wind speed prediction	8	**0.136**	**0.288**	**36.99**
	12	0.140	0.296	38.34
	24	0.168	0.356	47.92
Load demand prediction	8	**0.177**	**0.113**	**7.9**
	12	0.262	0.168	11.7
	24	0.250	0.160	11.2

Table 7 Performance indices for the selected linear model in forecasting for several time horizons

	Hours ahead forecast (h)	$NRMSE_f$	$CV(RMSE)_f$	MAPE (%)
Radiation prediction	8	0.205	0.769	24.00
	12	**0.185**	**0.693**	**20.51**
	24	0.197	0.738	22.14
Wind speed prediction	8	**0.139**	**0.296**	**38.003**
	12	0.149	0.316	43.88
	24	0.184	0.390	58.83
Load demand prediction	8	**0.181**	**0.203**	**14.25**
	12	0.284	0.240	15.92
	24	0.293	0.261	18.44

It should be observed that, for PV plants under uniform radiation, the forecast of solar radiation can be used to forecast their output power, once the main geometric and electric characteristics of the plant are known. Similarly, the power from a wind generator can be easily deduced by wind speed information, once the wind plant parameters are set. Therefore, the proposed approach is particularly suitable to optimize decision making at the grid manager level. In particular, the good forecasting obtained in this study in the range between 8 and 12 hours ahead is useful for electricity market and characterization of the viability of solar/wind production in a given location [39].

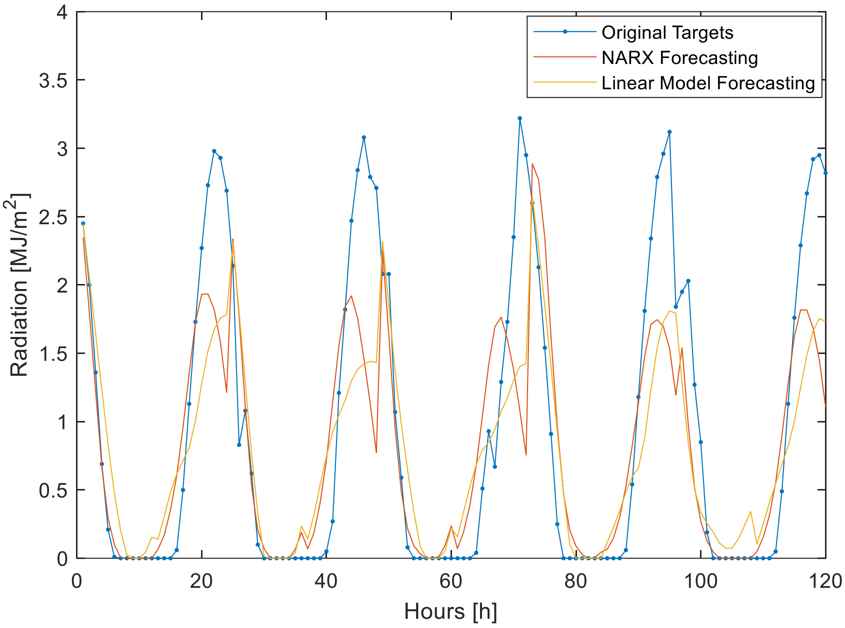

Fig. 4 Twelve-hours-ahead solar radiation forecasting result

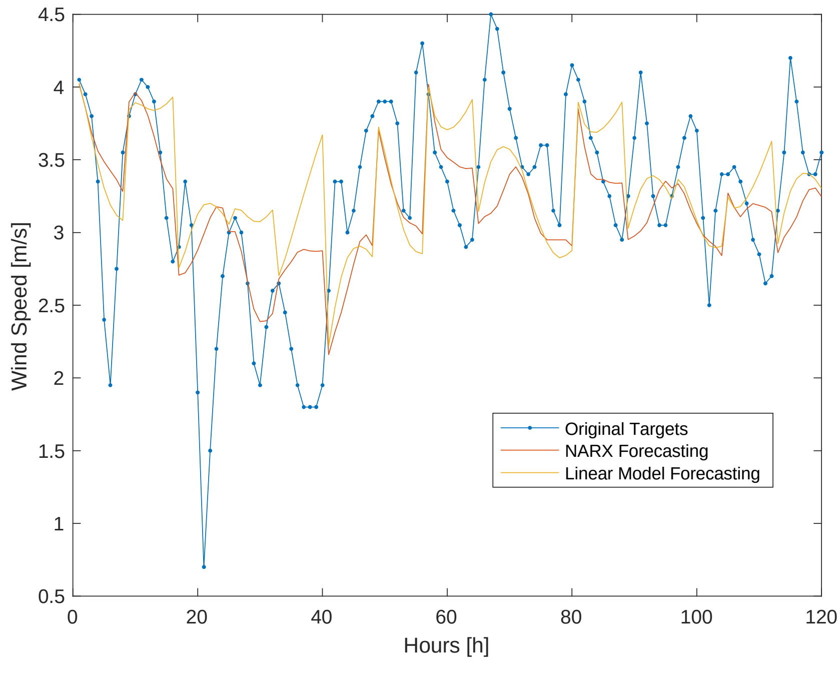

Fig. 5 Eight-hour-ahead wind speed forecasting result

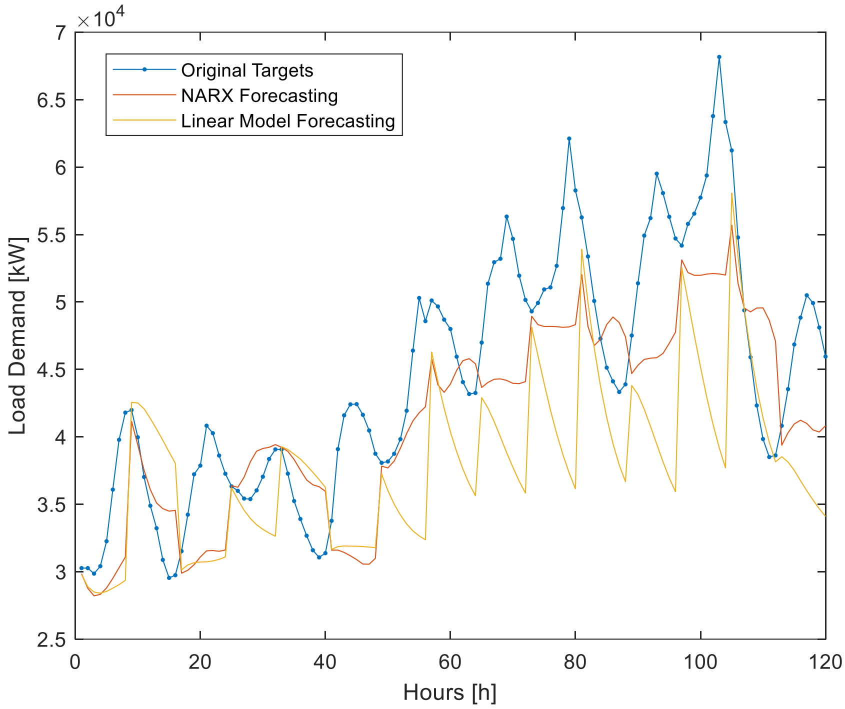

Fig. 6 Eight-hours-ahead power load demand forecasting result

7 Conclusions

In this paper, the use of a NARX ANN-based model for solar radiation, wind speed, and load power demand forecasting is presented. This forecasting technique is aimed at knowing in advance, with a time horizon that is within 24 hours ahead, the output power of renewable sources and the power demand from users of the electrical grid. This result is useful to reduce the impact of power fluctuations on the power grid, and therefore to increase its overall efficiency and power quality. Meteorological datasets, spanning seven years have been used for this study, together with an hourly load power demand dataset spanning 4.5 years. The performance obtained by classic time series linear models and by the proposed ANN is compared in terms of both prediction and forecasting results. The NARX network exhibits the best behavior in terms of both prediction and multi-step-ahead forecasting of the considered variables. The best forecasting results are obtained for the 12-hours-ahead time horizon for solar radiation and for 8-hour-ahead for both wind speed and load power demand. The obtained results show that the NARX model is well suited to perform energy-related time series forecasting in the short-term time horizon. Due to this optimal forecasting time horizon, the proposed method is particularly useful for electricity market management, for characterization of solar/wind production viability, as well as for supporting intelligent EMS in optimal scheduling of electrical power flows. Furthermore, despite the good forecasting performance, the ANN structure is very simple, with a limited number of neurons in a single hidden layer; this is an additional advantage of the chosen approach in comparison to other more complex methods proposed in the literature, where ANNs with multiple recurrent layers are employed.

References

[1] S. Haben, J. Ward, D. Vukadinovic Greetham, C. Singleton, and P. Grindrod, “A new error measure for forecasts of household-level, high resolution electrical energy consumption,” *Int. J. Forecast.*, vol. 30, no. 2, pp. 246–256, Apr. 2014.

[2] M. N. I. Maruf, L. A. Hurtado Munoz, P. H. Nguyen, H. M. Lopes Ferreira, and W. L. Kling, “An enhancement of agent-based power supply-demand matching by using ANN-based forecaster,” in *2013 4th IEEE/PES Innovative Smart Grid Technologies Europe, ISGT Europe 2013*, 2013.

[3] A. Di Piazza, M. C. Di Piazza, A. Ragusa, and G. Vitale, “Environmental data processing by clustering methods for energy forecast and planning,” *Renew. Energy*, vol. 36, no. 3, pp. 1063–1074, Mar. 2011.

[4] A. Di Piazza, M. C. Di Piazza, and G. Vitale, “Estimation and Forecast of Wind Power Generation by FTDNN and NARX-net based models for Energy Management Purpose in Smart Grids,” *Renew. Energies Power Qual. J.*, vol. 12, 2014.

[5] Q. Liu, Y. Shen, L. Wu, J. Li, L. Zhuang, and S. Wang, "A hybrid FCW-EMD and KF-BA-SVM based model for short-term load forecasting," *CSEE J. Power Energy Syst.*, vol. 4, no. 2, pp. 226–237, Jun. 2018.

[6] R. Palma-Behnke *et al.*, "A Microgrid Energy Management System Based on the Rolling Horizon Strategy," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 996–1006, Jun. 2013.

[7] E. Y. Bitar, R. Rajagopal, P. P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1225–1235, 2012.

[8] M. C. Di Piazza, G. La Tona, M. Luna, and A. Di Piazza, "A two-stage Energy Management System for smart buildings reducing the impact of demand uncertainty," *Energy Build.*, vol. 139, pp. 1–9, Mar. 2017.

[9] Y. Wang, Y. Shen, S. Mao, X. Chen, and H. Zou, "LASSO and LSTM Integrated Temporal Model for Short-Term Solar Intensity Forecasting," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2933–2944, Apr. 2019.

[10] H. Quan, D. Srinivasan, and A. Khosravi, "Short-Term Load and Wind Power Forecasting Using Neural Network-Based Prediction Intervals," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 25, no. 2, pp. 303–315, Feb. 2014.

[11] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, "Solar forecasting methods for renewable energy integration," *Prog. Energy Combust. Sci.*, vol. 39, no. 6, pp. 535–576, 2013.

[12] T. McElroy and M. Wildi, "Multi-step-ahead estimation of time series models," *Int. J. Forecast.*, vol. 29, no. 3, pp. 378–394, Jul. 2013.

[13] Y. Ren, P. N. Suganthan, and N. Srikanth, "A Comparative Study of Empirical Mode Decomposition-Based Short-Term Wind Speed Forecasting Methods," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 236–244, Jan. 2015.

[14] D. Masa-Bote *et al.*, "Improving photovoltaics grid integration through short time forecasting and self-consumption," *Appl. Energy*, vol. 125, pp. 103–113, Jul. 2014.

[15] C. Sandels, J. Widén, and L. Nordström, "Forecasting household consumer electricity load profiles with a combined physical and behavioral approach," *Appl. Energy*, vol. 131, pp. 267–278, Oct. 2014.

[16] N. Liu, Q. Tang, J. Zhang, W. Fan, and J. Liu, "A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids," *Appl. Energy*, vol. 129, pp. 336–345, Sep. 2014.

[17] B. Yildiz, J. I. Bilbao, and A. B. Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting," *Renew. Sustain. Energy Rev.*, vol. 73, no. February, pp. 1104–1122, 2017.

[18] Z. Guo, K. Zhou, X. Zhang, and S. Yang, "A deep learning model for short-term power load and probability density forecasting," *Energy*, vol. 160, pp. 1186–1200, Oct. 2018.

[19] S. Kumar, L. Hussain, S. Banarjee, and M. Reza, "Energy Load Forecasting using Deep Learning Approach-LSTM and GRU in Spark Cluster," in *Proceedings of 5th International Conference on Emerging Applications of Information Technology, EAIT 2018*, 2018.

[20] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renew. Sustain. Energy Rev.*, vol. 27, pp. 65–76, 2013.

[21] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *North American Power Symposium 2010, NAPS 2010*, 2010.

[22] Z. Liu, W. Gao, Y.-H. Wan, and E. Muljadi, "Wind power plant prediction by using neural networks," in *2012 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2012, pp. 3154–3160.

[23] C. Paoli, C. Voyant, M. Muselli, and M. L. Nivet, "Forecasting of preprocessed daily solar radiation time series using neural networks," *Sol. Energy*, vol. 84, no. 12, pp. 2146–2160, 2010.

[24] J. Liu, W. Fang, X. Zhang, and C. Yang, "An Improved Photovoltaic Power Forecasting Model With the Assistance of Aerosol Index Data," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 434–442, Apr. 2015.

[25] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, *Recurrent Neural Networks for Short-Term Load Forecasting*. Cham: Springer International Publishing, 2017.

[26] C. Tao, D. Shanxu, and C. Changsong, "Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement," in *The 2nd International Symposium on Power Electronics for Distributed Generation Systems*, 2010, pp. 773–777.

[27] I. M. Navon and D. M. Legler, “Conjugate-Gradient Methods for Large-Scale Minimization in Meteorology,” *Mon. Weather Rev.*, vol. 115, no. 8, pp. 1479–1502, Aug. 1987.

[28] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, “A review of wind power and wind speed forecasting methods with different time horizons,” in *North American Power Symposium 2010*, 2010, no. June 2014, pp. 1–8.

[29] J. C. Cao and S. H. Cao, “Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis,” *Energy*, vol. 31, no. 15, pp. 3435–3445, Dec. 2006.

[30] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, “The National Solar Radiation Data Base (NSRDB),” *Renewable and Sustainable Energy Reviews*, vol. 89. Elsevier Ltd, pp. 51–60, 01-Jun-2018.

[31] T. Hong, P. Pinson, and S. Fan, “Global Energy Forecasting Competition 2012,” *Int. J. Forecast.*, vol. 30, no. 2, pp. 357–363, Apr. 2014.

[32] H. T. Siegelmann, B. G. Horne, and C. L. Giles, “Computational capabilities of recurrent NARX neural networks,” *IEEE Trans. Syst. Man Cybern. Part B*, vol. 27, no. 2, pp. 208–215, Apr. 1997.

[33] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice-Hall, Inc., 2004.

[34] Eurelectric, “Decarbonization pathways, European economy EU electrification and decarbonization scenario modelling Synthesis of key findings, May 2018.”

[35] B. NEF, “Sector Coupling in Europe: Powering Decarbonisation, Potential and Policy Implications of Electrifying the Economy, February 2020.”

[36] L. Prechelt, “Early Stopping - But When?,” 1998, pp. 55–69.

[37] L. C. Brown and P. M. Berthouex, *Statistics for environmental engineers*. Boca Raton (FL): CRC Press, 2002.

[38] G. E. P. Box and G. M. Jenkins, *Time series analysis: forecasting and control*. Holden Day, 1976.

[39] H. T. C. Pedro and C. F. M. Coimbra, “Assessment of forecasting techniques for solar power production with no exogenous inputs,” *Sol. Energy*, vol. 86, no. 7, pp. 2017–2028, Jul. 2012.