Effect of Environmental Conditions and Training Algorithms on the Efficiency of a NARX Based Approach to Predict PV Panel Power Output

Toufik Sebbagh[(1*)], **Ridha Kelaiaia**[(1)], **Adlen Kerboua**[(1)], **Abderrazak Metatla**[(1)], **Abdelouahab Zaatri**[(2)]

[(1)]   *LGMM Laboratory, University of 20 Août 1955 - Skikda, Po B 26, Skikda, 21000, ALGERIA*
        e-mail: t.sebbagh@univ-skikda.dz

[(2)]   *University of Constantine 1, Constantine, 25000, ALGERIA*

SUMMARY

Photovoltaic energy is volatile in nature since it depends on weather conditions. It is important to have an idea about the reliability and the economic feasibility of any new project to decide whether it is right to proceed with the installation of such a project. Hence, it is becoming fundamental to know renewable energy state and production that can be combined with other less variable and more predictable sources to justify the choice of regions for the new photovoltaic projects installation. The current research investigates the forecasting abilities of a NARX based approach. The influence of the meteorological data, such as irradiance, ambient temperature, and wind speed, and the impact of training algorithms on the performance of the NARX-based forecaster is studied. For this purpose, four models are discussed, each model is trained based on three training algorithms. The NARX model using a Bayesian Regularization algorithm, trained by the three meteorological data as inputs and the converted power output as output, outperforms the other models. It consists of a simple architecture with one input layer, a hidden layer containing 10 neurons, and an output layer, with a mean square error of 0.0085 W2 for the training phase and 0.0043 W2 testing phase, and the overall regression of 95.48%. This simplified architecture and low values of the mean square error and the regression coefficient suggest that they are promising photovoltaic output prediction tools, particularly in locations where few meteorological parameters are monitored.

KEY WORDS: *Forecast; PV power output; NARX; Training Algorithms; meteorological characteristics.*

1. INTRODUCTION AND RELATED WORKS

Energy is a key factor in the development of humanity. Electrical energy is the most useful form of energy, it was generated mainly from fossil fuels, and is currently mainly related to them. However, this issue has a harmful environmental effect. In this context, renewable energy is an alternative solution to this issue mainly in isolated areas where it is hard or impossible to connect to the grid. Among the large variety of renewable energy sources, solar energy is the most promising and the most acceptable source. Therefore, solar energy is gaining the attention of scientists, industrials, and governments. However, the installation of any new project requires to know its reliability and its economic feasibility.

The basic component of a photovoltaic (PV) converter is the PV cell. A PV cell converts directly the sunlight into low voltage unstable electric power. According to the required current and voltage, several cells are connected in series or in parallel to produce panel.

Since the generation of PV power fully depends on uncertain and uncontrollable meteorological factors, the power output of the PV system changes dynamically. Meanwhile, during a day, the converted power can reach the maximum value in clear sky conditions, but it suddenly drops to low values under the effect of partial shading created by clouds, trees, and buildings. Therefore, a pinpoint prediction of PV power conversion is very difficult. The accurate forecasting of PV power production is a great challenge, that can improve the reliability of installed systems and maintain the power quality [1].

Several papers have been conducted proposing methods to forecast the PV power production. These techniques can be classified as direct and indirect forecasting techniques [1, 2].

Nowadays the most popular techniques in direct forecasting the power output of PV installations are the soft computing techniques based on Artificial Neural Networks (ANN) [3]. It is noteworthy that ANN is a statistical computer-based method that can to explore the relationships between variables with high accuracy [4, 5]. Otherwise, ANN has an inherent ability to model non-linear, dynamic, noisy data, and complex systems [6, 7]. Thus, Mellit and Kalogirou [8] outlined a comprehensive review of the use of ANNs to forecast solar energy. This review cited several papers published between 1994 and 2008. However, all the articles focused on the forecasting of solar radiation but not the power output of the panels. Ding et al. [9] proposed an approach based on ANN to forecast the power output of a PV system. In this approach, the improved backpropagation learning Algorithm was adopted to overcome the shortcomings of the standard backpropagation learning algorithm. Even so, the paper discussed the power output forecasting for a rainy day and a sunny day, the influence of meteorological parameters on the results of the proposed model cannot be proved since sunny day may have lower temperature and wind speed values than the rainy day. Moreover, the impact of the training algorithm on the results of the proposed ANN predictor did not discuss. Another approach based on different topologies of ANN to forecast the power output of PV modules has been proposed by Lo Brano et al. [10]. The authors investigated the influence of some meteorological data, such as air temperature, wind speed and solar irradiance, and some physical parameters of the panel, without taking into account the influence of the training algorithm. Shi et al. [11] proposed algorithms to predict the power output of a PV system based on weather classification and support vector machine (SVM). This research investigated the influence of the weather conditions, but not the training algorithm, and applied it to a grid-connected system. Xiao et al. [12] introduced an ANN forecaster of the output power of three technologies of PV cells. The authors studied the influence of the number of hidden neurons on

the prediction of power. Abedinia et al. [13] proposed a solar output forecaster based on a multilayer perceptron neural network, trained using a hybrid modified Levenberg-Marquardt (LM) and an evolutionary algorithm. The proposed model used irradiation, air temperature, and cell temperature as inputs. However, air temperature and cell temperature are two faces of the same coin, and the effect of the change of the training algorithms on the performance of the forecaster engine is not illustrated. Saberian et al. [14] used a generalized regression neural network (GRNN) and feedforward neural network (FFNN) to predict the power output of a PV model using temperature and irradiance as inputs. It was found that, while both FFNN and GRNN performed satisfactorily, FFNN produced more accurate results. Vaz et al. [15] implemented a NARX predictor to forecast the power production of a PV system using as input meteorological data and measurements of other neighbouring PV systems. They used the Normalized root mean square error (nRMSE) to measure the performance of the network and demonstrate that the proposed model outperforms the persistence model for a forecast horizon greater than *15 min*. Aningo et al. [16] presented an ANN model power prediction. They compared the performance of a nonlinear autoregressive neural network (NARX) with a nonlinear autoregressive neural network (NAR) using a time series modelling approach. The study showed that the NARX model outperformed the NAR model. Recently, Louzazni et al. used a (NARX) model to predict the power output of a monocrystalline PV module installed in Belbis, Egypt [17]. The prediction model considered only the effect of temperature and solar irradiation. Nevertheless, the effect of the training algorithm was not shown since the LM is solely applied to train the proposed NARX model.

Among several strengths of ANN-based forecaster, a key strength is the ability of the network designer to select several inputs to improve the forecast accuracy. However, using too many input variables can lead to some negative consequences as the increase in computation time or in the redundant variables that can yield complications to the training stage by increasing the number of local optima [18].

The majority of the above works deal with the problem of the forecast of the photovoltaic power output without taking into account the influence of the training algorithm. However, few studies have been conducted to determine the most influential climatic parameters that can enhance the predictive capabilities of ANN-based PV power output forecasters. The same is with training algorithms that can provide the best performance of the ANN-based forecast models. Hence, the current research investigates the use of ANN for medium-term direct forecasting of PV power output based on time series historical data. The main contribution of this study is to present the influence of the type of inputs (meteorological parameters) and the training algorithms on the performance of the proposed NARX model.

2. ARTIFICAL NEURAL NETWORK PREDICTORS

An Artificial Neural Networks (ANN) is considered as a machine that models a task or a function of interest, performing useful computation through a learning process. In fact, to drive its computing power, the ANN employs its massively parallel distributed structure and its ability to learn and generalize. It can find reasonable outputs for the inputs that are not encountered during the training process [15, 19].

ANNs are the most used machine learning techniques to forecast solar power [20] spurred by the observations of the neuron operation, the interconnection of a group of neurons form a

neural network (NN). The connections have numeric weights, whose final value is given in the training phase, and all together predict the output.

Figure 1 illustrates a schematic of ANN architecture, while the mathematical model of an ANN is expressed as (1):
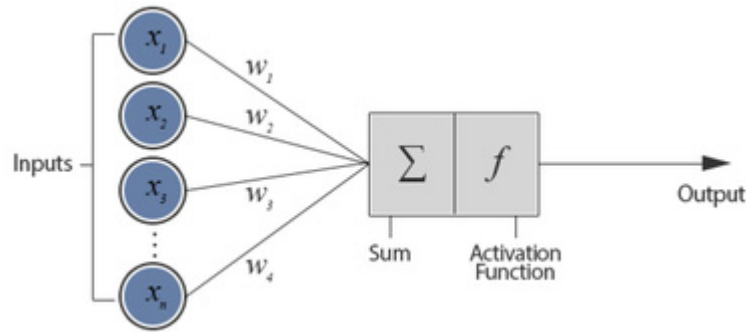

Fig. 1 *Schematic of ANN*

$$U_N = b + \sum_{j=1}^{N}(W_j \times I_j) \tag{1}$$

Where: U_N is the network output; W_j is the connection weight; I_j is the input number; b is the bias weight and N is the number of inputs.

There is a large number of ANN forecasters topologies proposed in the literature. Some specific examples include Multi-layer feedforward neural networks (MLFFNN), Multi-layer perceptron NN (MLPNN), Radial basis function NN (RBFNN), and recurrent neural networks (RNN). The description of these NN types with examples of application to the PV power prediction field is presented in [21].

NARX is a recurrent dynamic NN with feedback connected to several layers of the NN [22]. The architecture of the NARX relies on the actual value of an input time series P(t) to previous values of the same series (*P(t-1)*, *P(t-2)* ...), and the current and the previous of the exogenous series [23]. The simplest architecture of the NARX model is presented in Figure 2, and can be expressed mathematically (2) as follows:

$$y(t) = f(P(t - n_P), ..., P(t - 1), P(t), y(t - n_y), ..., y(t - 1)) \tag{2}$$

Where: f is a nonlinear function; $P(t)$ and $y(t)$ are, respectively, the input and output of the network at time t; n_P and n_y are the order of the input and output, respectively.

The NARX network is adopted due to its good ability to handle problems involving the modelling of nonlinear dynamic systems, such as dependencies among meteorological time series [24, 25]. The benefits of using NARX NNs include the often-superior forecast caused by the extra inputs as well, they are also relatively easy to understand and use, compared with some other statistical or parametric methods that require prior knowledge of statistics and they can use Input-Output (IO) mapping, i.e. the ability to learn from previous samples by creating IO mapping to solve problems [16, 19].
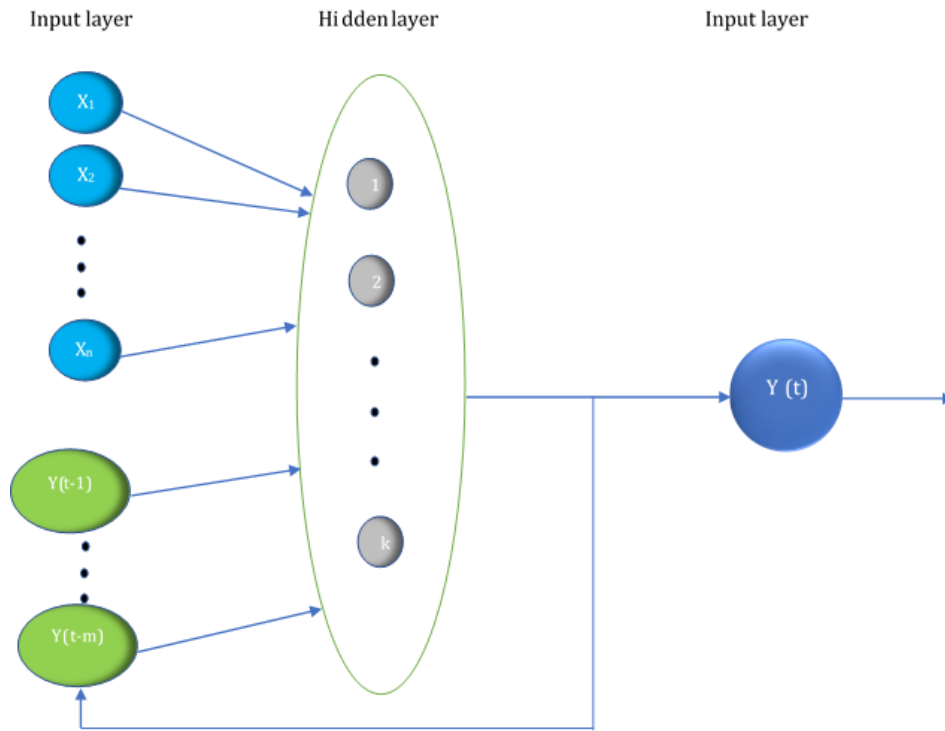
Fig. 2 *General structure of NARX Network*

3. METHODOLOGY

A Multivariate time series model based on a NARX model is proposed to forecast the power generated by the PV panel using three sets of meteorological data.

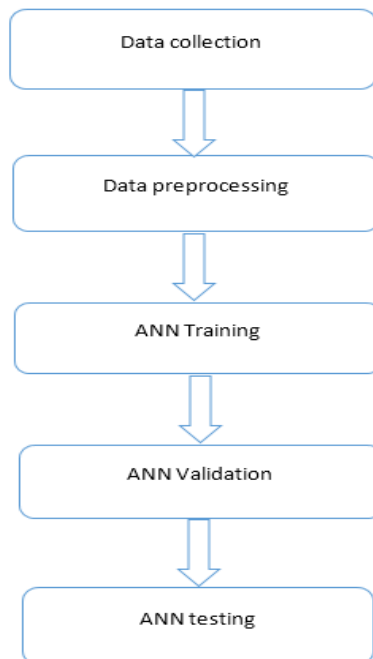The different phases carried out for the current procedure are presented in Figure 3.

Fig. 3 *Different phases of the proposed model*

After collecting the time series dataset, a pre-processing is applied to these data. The dataset is decomposed into three sets, the first one is used for the training and represents *70%* of the total data, while the other sets, composed of *15%* each, are used for the test and validation of the proposed NN.

To study the influence of the inputs on the performance of the model, different sets of data are used to train the NN.

To study the influence of the learning method on the performance of the network, three different algorithms are used for each set of input data.

3.1 SELECTED PV PANEL DESCRIPTION AND CHARACTERISTICS

The PV panel used for this research is a "CEM 235-P" polycrystalline composed of 60 series-connected cells. The panel is covered using hardened-glass glazing covered in transparent ethylene–vinyl acetate. The PV energy converted is measured for a south-oriented panel and a tilt angle of *32°*. Table 1 highlights the electrical characteristics of this panel. The maximum power that can be converted by this panel is 235 W under Standard Test Condition (STC).

Table 1 *The name of table*

Parameters	Values
Maximum power (P_m)	235 Wc
Maximum power current (I_{mpp})	7.78 A
Short circuit current (I_{SC})	8.4 A
Maximum power voltage (V_{mpp})	30.2 V
Open circuit voltage (V_{OC})	37.4 V
Normal operating cell temperature (NOCT)	45∓ 2°C
Temperature coefficient of power	-0.39 % / °C
Temperature coefficient of current	+0.06 % / °C
Temperature coefficient of voltage	-0.33 % / °C

3.2 WEATHER CHARACTERISTICS OF THE SELECTED LOCATION AND DATA COLLECTION

The selected location (Latitude= 36.848739, Longitude=6.889475) is located at the University of Skikda. The panels are installed close to the laboratory of renewable energy of the Faculty of technology (Figure 4) and serve as a platform for practical pedagogic and research studies. Skikda is characterized by its Mediterranean climate marked by wet, mild winter and dry and clear summer.

Fig. 4 *PV panels under study*

In this research, it is assumed to use a set of time series data composed of global solar radiation, ambient temperature, and wind speed and the corresponding direct current power delivered by the PV panel for training, validation, and test of the proposed neural network.

Lack of measured data and the objective of building a database that can be used by students and researchers, the meteorological data used in this research are collected from [26], within the date range from 01/01/2019 to 10/01/2019. During this date range, the PV power output is consistently zero between 4 pm and 6 am. Thus, only the data between 6 am and 4 pm are considered in the simulations. The data are taken on a time scale of 1 hour. According to several works, the time series prediction of PV power on an hourly average basis is more accurate for medium-term forecasters [27, 28]. However, averaging data may obscure some information like peaks occurring in the instantaneous irradiance curve as a result of partial or total shading. Such an effect is not taken into consideration in the current study.

3.2.1. IRRADIANCE LEVEL

The irradiance is defined as the density of the solar radiation power received on a given surface [29]. In the present work, the global irradiance incident on a *32°* inclined and south-oriented panel is considered.

The global irradiance ($G_{g,t}$) depends on the horizontal radiation and the surface orientation. It is composed of three different contributions; beam radiation ($G_{b,t}$), diffuse radiation ($G_{d,t}$) and reflected radiations ($G_{e,t}$):

$$G_{g,t} = G_{b,t} + G_{d,t} + G_{r,t} \tag{3}$$

The three components of the global irradiance were computed according to the procedure proposed in [30]:

$$G_{b,t} = \frac{G_{g,h} - G_{d,h}}{\cos(90 - \alpha)} \cos(\theta) \tag{4}$$

$$G_{d,t} = \frac{1 + \cos(\beta)}{2} \tag{5}$$

$$G_{r,t} = \rho_g G_{d,h} \frac{1-cos(\beta)}{2} \qquad \textbf{(6)}$$

where: $G_{g,h}$ is the global horizontal irradiance (W/m^2); $G_{d,h}$ is the diffuse horizontal irradiance (W/m^2); α is the solar altitude in (°); θ is the incidence angle in (°); β is the tilt angle in (°) and ρ_g is the ground reflectance.

3.2.2. AMBIENT TEMPERATURE

The surface temperature of the panel has a negative impact on the PV panel performance. The power converted by the panel is reduced with the surface temperature increase. The ambient temperature has a direct impact on the panel, it is related to the PV panel surface temperature according to the formula (7) as follows [31, 32]:

$$T_C = T_a + \frac{NOCT-20}{800} G_{g,t} \qquad \textbf{(7)}$$

Where: T_C is the cell temperature (°C); T_a is the ambient temperature (°C); $NOCT$ is the nominal operating cell temperature (°C) and $G_{g,t}$ is the global irradiance on a tilted surface (W/m^2).

3.2.3. WIND SPEED

The impact of wind speed on the power output of the PV system is given by the following equation (8) modified from [33]:

$$P_{PV} = \eta_{PV,STC} \left[1 + \frac{\mu}{\eta_{PV,STC}}(T_a - T_{STC}) + \frac{\mu}{\eta_{PV,STC}} \frac{9.5}{5.7+3.8v} \frac{(NOCT-20)}{800}(1 - \eta_{PV,STC})G_{g,t}\right] A_{PV} G_{g,t} \quad \textbf{(8)}$$

Where: P_{PV} is the power output from the PV panel (W); $\eta_{PV,STC}$ is the efficiency of the panel at STC conditions (%); μ is the temperature coefficient (%/°C); v is the wind speed (m/s); A_{PV} is the area of the PV panel (m^2).

The mean values of ambient temperature, global irradiance, wind speed, and the corresponding power output are plotted in Figure 5.

The mean values of temperature vary between $2°C$ and $20°C$. However, the average mean value of the tilted irradiance is about $409 W/m^2$.

The values of wind speed used in the current study are taken for the specified area at the height of $10 m$. The average wind speed varies between $2 m/s$ and $5 m/s$, but rarely, it falls under $1 m/s$ and rarely exceeds $6 m/s$. The average mean value of wind speed is about $3 m/s$.

The above time series are used as inputs and outputs of the neural networks as shown in Figure 3.

3.3 DATA PRE-PROCESSING

The purpose of the data pre-processing stage is to facilitate the training of the network. Since there is a great difference in the values of the input data, higher valued inputs may suppress the influence of smaller ones while training the neural network. Hence, data must be well processed and properly scaled before being inputted to the ANN [33]. In this step, all input and output data are normalized linearly in the range [0, 1].

The selected data are pre-processed as follows.

Firstly, all the input data that leads to a target equal to zero, are deleted from the dataset.

The other data are normalized between *0* and *1* using the following formula:

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

(9)

Where: $x_{normalized}$ is the normalized value, x is the actual value, x_{min} is the minimum value of each time series dataset, and x_{max} is the maximum value from each time series vector.
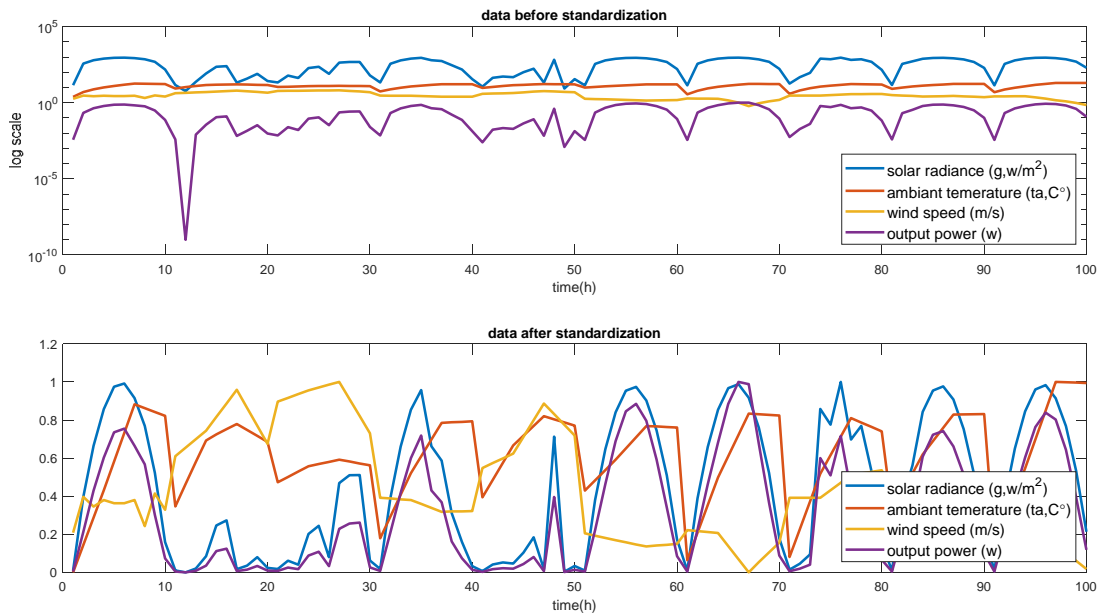
The normalized data are illustrated in Figure 5.

Fig. 5 *Input and output data before and after normalization*

3.4 ANN ARCHITECTURE

For the proposed approach, we use the nonlinear autoregressive network with exogenous inputs (NARX) model to predict the power output converted by the PV panel.

In the absence of a formal mathematical approach to determine the optimal network structure, the "trial and error" procedure is chosen. Several NARX networks are generated by variation of the following internal parameters: activation functions, number of hidden layers and neurons, the number of delays. The optimal network structure is identified among the different configurations, based on the performance plot, error histogram plot, training state plot, and the regression plot. If the performance of the network is unsatisfactory, the weights and biases are initialized and the network is retrained until adequate results are achieved.

The NARX model used is composed of an input layer, one hidden layer, and the output layer with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The sigmoid transfer function enables the network to learn the non-linear correlations between the weather variables and PV output. The linear output layer ensures that the range of the output is not limited.

The number of neurons in the hidden layer is optimized (by trial and error) to 10 neurons, and the number of delays is set to 2.

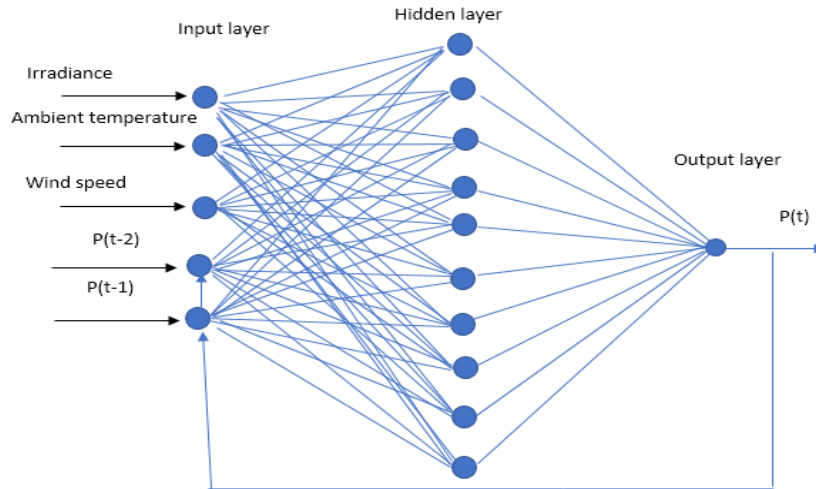A schematic of the proposed NARX model is depicted in Figure 6.

Fig. 6 *Proposed NARX model structure*

2.4.1. INPUTS AND OUTPUTS

For the present study, four different sets of data are used as input data to illustrate the impact of the input data on the performance of the chosen neural network model. These sets are defined as:

- Irradiance.
- Irradiance + ambient temperature.
- Irradiance + wind speed.
- Irradiance + ambient temperature + wind speed.

The target data is the hourly values of the PV panel power output.

2.4.2. ANN TRAINING

To study the influence of the training technique on the performance of the NARX model, three training algorithms are applied to each of the four cases. These are the Levenberg-Marquardt (LM), the Bayesian Regularization (BR), and the Scaled Conjugate Gradient (SCG).

LEVENBERG-MARQUART ALGORITHM

The LM is an iterative algorithm that determines the minimum of a multivariable function expressed as the sum of squares of nonlinear real-valued functions [34]. The LM is characterized by its quick learning feature [33].

Similarly to quasi-Newton methods, the Levenberg Marquardt algorithm is designed to approach second order training speed without having to compute the Hessian matrix. Under the assumption that the error function is some kind of squared sum, then the Hessian matrix can be approximated as (10) [35]:

$$H = J^T J \qquad \qquad \textbf{(10)}$$

and the gradient can be computed as (11):

$$H = J^T e \tag{11}$$

Where: J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases. The Jacobian matrix determination is less computationally expensive than the Hessian matrix; e is a vector of network errors. Then the update can be adjusted as (12):

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{12}$$

The parameter μ is a scalar controlling the behavior of the algorithm. For $\mu = 0$, the algorithm follows Newton's method, using the approximate Hessian matrix. When μ is high, this becomes gradient descent with a small step size.

BAYESIAN REGULATION ALGORITHM

The BR minimizes the combination of squared errors and weights, and is more likely to generalize the performance better and is thus more reliable out-sample [27]. A detailed description of the BR Algorithm can be found in [36].

SCALED CONJUGATE GRADIENT ALGORITHM

The SCG algorithm is an iterative technique that converges to the minimum of a quadratic function in a finite number of iterations [33]. SCG is well suited to handle large-scale problems [37].

The datasets are randomly divided into three parts (*70%* for training; *15%* for validation and *15%* for testing) when invoking the training process. The advantage of a random data selection is that the network is exposed to data corresponding to different days during the training session. Otherwise, it memorizes the order in which data is fed to the network and may fail to generalize [7]. However, this classification is not followed when training the different models with the BR algorithm, since BR uses the whole dataset for training.

2.4.3. EVALUATION CRITERIA:

After the training and validation of the forecasting model, the accuracy of the prediction is measured by calculating the difference between the forecasted values and the target.

In this paper, the Mean Squared Error (MSE) is used to quantitatively evaluate the performance of the proposed forecasting model.

The MSE is defined as the average squared difference between the outputs of the forecasting model and the targets. It can be calculated using (13):

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (P_{forcasted_i} - P_i)^2 \tag{13}$$

Where: $P_{forcasted_i}$, is the power forecasted by the NARX model (W); P_i is the target power (W).

4.   RESULTS AND DISCUSSIONS

The prediction model is defined and trained to obtain the best forecasting results. The time series meteorological data consist of irradiance on an inclined surface (*X1*), ambient

temperature (*X2*), and wind speed (*X3*), taken with an hourly interval and regrouped in four datasets, and used in four NARX models as follows:

- Case 1 (*M1*): *X1*;
- Case 2 (*M2*): *X1 + X2*;
- Case 3 (*M3*): *X1 + X3*;
- Case 4 (*M4*): *X1 + X2 + X3*.

Each of the four models constructed using the previous datasets is trained by the three training algorithms.

To validate the proposed approach, the MSE performance function was examined during the training and the testing phases. If the validation vectors failed to improve or remained the same they were able to stop the training process early. This is explained by the increase in the MSE value during the validation. Test vectors were used to check whether the network was generalizing well but did not have any effect on training.

Table 2 presents the values of MSE calculated by the four NARX during training, validation, and testing phases. In this part, the models are trained using LM Algorithm.

Table 2 *MSE for models trained by LM*

Model		M1	M2	M3	M4
MSE	*Training*	8.17061 e-3	9.87257 e-3	7.49080e-3	7.31118e-3
	Validation	2.70227 e-2	7.71166 e-3	7.04412e-3	2.24171 e-2
	Test	9.10208 e-3	1.30422 e-2	8.76837e-3	8.72177 e-3

From Table 2, it can be seen that the NARX model (*M4*) gives the best performance when trained using the LM algorithm, and the MSE calculated with this model has the lowest value compared to other models with an error value of *0.0073 W²* and *0.0087 W²* at epoch 9 with three input variables, for the training and testing phases, respectively.

Table 3 presents the MSE values of the different phases of the four NARX models trained using the SCG Algorithm. According to this scenario, the model (*M3*) is the most efficient in the training phase, since the result of the MSE calculation was *0.0092 W²* at epoch 17.

Table 3 *MSE for models trained by SCG*

Model		M1	M2	M3	M4
MSE	*Training*	1.24595 e-2	1.13236 e-2	9.22995 e-3	1.15250e-2
	Validation	3.18113 e-2	1.24963 e-2	4.38230 e-2	1.87916 e-2
	Test	7.25040 e-3	8.16040 e-3	9.36666 e-3	6.52453 e-3

Table 4 shows the MSE calculated during the training, validation, and testing of the four networks when trained by the BR training algorithm. Accordingly, *M4* was the performant model for training and testing phases. The values of MSE are *0.0085 W²* and *0.0043 W²* at epoch 278, respectively.

Table 4 *MSE for models trained by BR*

Model		M1	M2	M3	M4
MSE	*Training*	8.8249 e-3	8.76857 e-3	9.60012 e-3	8.56229e-3
	Validation	-	-	-	-
	Test	1.41387 e-2	8.07837 e-3	5.11074 e-3	4.30502 e-3

The above results show that the MSE lies in the range of *0.0073* and *0.012* for the training phases and between *0.0043* and *0.014* for the testing phases. The ranges of this performance indicator demonstrate that all the models have good predictive abilities. Meanwhile, competitive MSE results are generated when training the model *M4*. For the training phase, *M4* performs well when trained using LM, and for the test phase, it gives the best result when trained using BR.

The scatter plot for the NARX model *M4* trained by BR is given in Figure 7, and the scatter plot for *M4* trained using LM in Figure 8.

The values of the regression (R) are equal to *0.95* for *M4* for both training algorithms (BR and LM), and for the training phase are equal to *0.97* and *0.95* for BR and LM respectively. The values of the overall regression reveal that the model *M4* trained using BR ($R=0.95$) outperforms the same model when trained by LM.
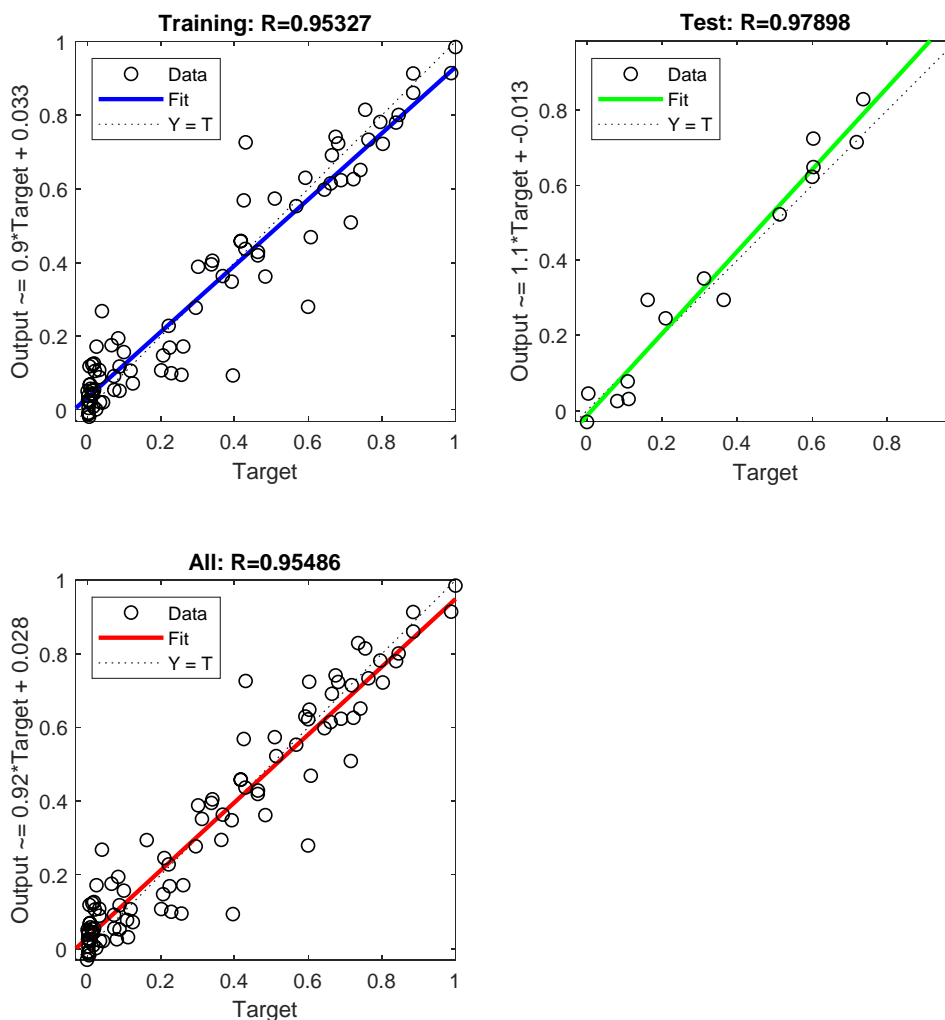

Fig. 7 *Regression analysis of the network outputs with respect to targets for training, validation and test sets for the model M4 trained by BR*
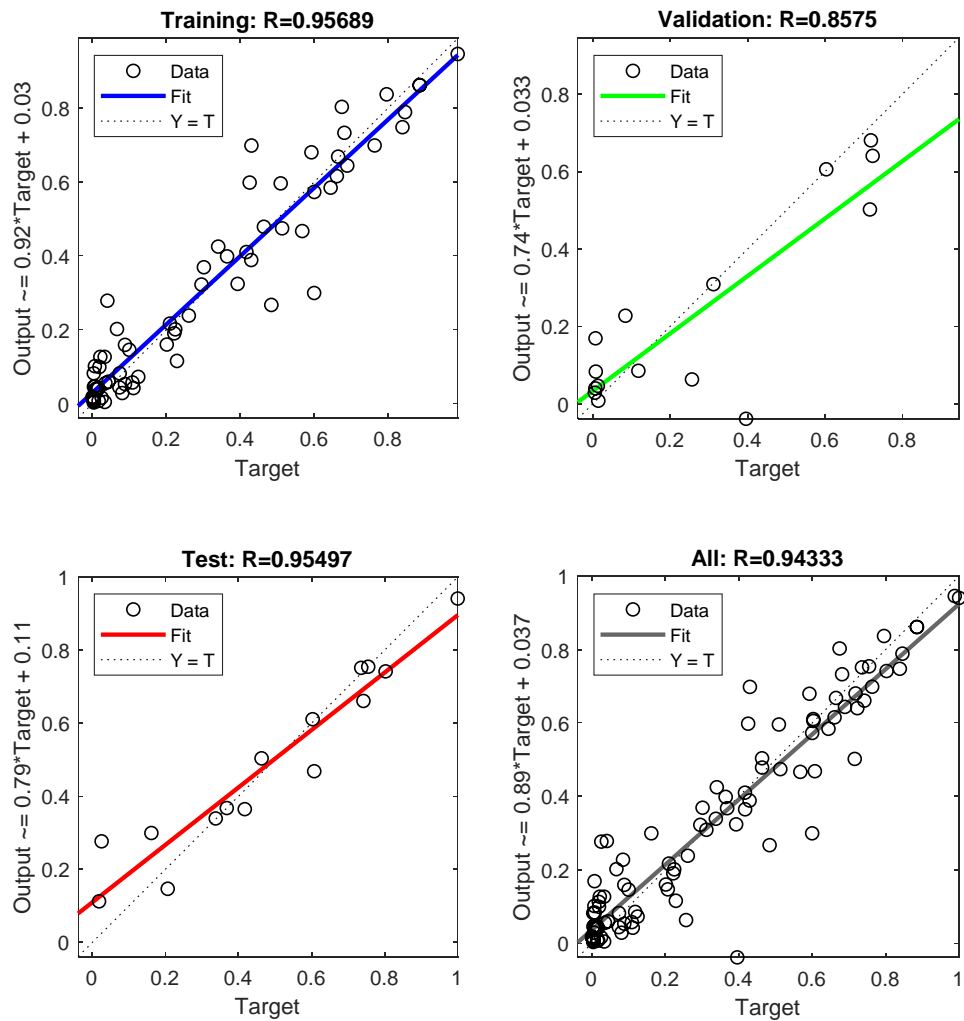
Fig. 8 *Regression analysis of the network outputs with respect to targets for training, validation and test sets for the model M4 trained by LM*

5. CONCLUSION AND PERSPECTIVES

Solar PV energy is an alternative to the traditional sources of energy, especially in remote areas. Having an informative idea about the behaviour of PV systems, before its installation is important. An accurate prediction of the power output of a PV panel is a complicated task due to its dependence on uncertain variables.

The objective of this paper is to investigate the influence of some climatological parameters (irradiance, ambient temperature, and wind speed) and the 3 training algorithms (Levenberg Marquardt, Bayesian Regularisation and Scaled Conjugate Gradient) on the ability of the proposed NARX approach to predict the power output of a PV panel.

Four NARX models were developed to predict the power output of a PV panel from hourly meteorological parameters, and the performance of the models is compared in terms of the

MSE and R values. Trained by the Bayesian Regularisation algorithm, the NARX model that used the irradiance, the ambient temperature, and the wind speed as inputs and the corresponding power output as output outperformed other models. This NARX model is composed of an input layer, an output layer, and one hidden layer with 10 hidden Neurons. The sigmoid activation function and the linear activation function are used for hidden and output layers, respectively.

The values of MSE for the most performant NARX forecaster are equal to *0.0085 W* for the training phase and *0.0043 W* for the testing phase, and the overall regression (*R*) is equal to *95.48%.*

Given the simplicity of its architecture, the proposed NARX model is a capable tool for PV power output prediction and can be used by students and research students to build their database. The model can also be applied to estimate the feasibility of new PV projects in new locations.

The future development of this work will consider the following points:

- The use of more meteorological data such as humidity.
- The classification of the time series data into several classes such as clear sky time data, cloudy time data…
- The use of other models and techniques such as the convolutional neural network.

6. REFERENCES

[1] U.K. Das, K.S. Tey, M. Seyedmahmoudian, S. Mekhilef, M.Y.I. Idris, W. Van Deventer, B. Horan, A. Stojcevski, Forecasting of photovoltaic power generation and model optimization: A review, *Renewable and Sustainable Energy Reviews*, Vol. 81, pp. 912-928, 2018. https://doi.org/10.1016/j.rser.2017.08.017

[2] M. Kudo, A. Takeuchi, Y. Nozaki, H. Endo, J. Sumita, Forecasting electric power generation in a photovoltaic power system for an energy network, *Electrical Engineering in Japan*, Vo. 167, No. 4, pp. 16-23, 2009. https://doi.org/10.1002/eej.20755

[3] S. Leva, A. Dolara, F. Grimaccia, M. Mussetta, E. Ogliari, Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power, *Mathematics and Computers in Simulation*, Vol. 131, pp. 88-100, 2017.

https://doi.org/10.1016/j.matcom.2015.05.010

[4] S. Khomfoi, L.M. Tolbert, Fault diagnostic system for a multilevel inverter using a neural network, *IEEE Transactions on Power Electronics*, Vol. 22, No. 3, pp. 1062-1069, 2007.

https://doi.org/ 10.1109/TPEL.2007.897128

[5] M. Kayri, Predictive abilities of bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data, *Mathematical and Computational Applications*, Vol. 21, No. 2, 20, 2016.

https://doi.org/10.3390/mca21020020

[6] H.T. Pedro, C.F. Coimbra, Assessment of forecasting techniques for solar power production with no exogenous inputs, *Solar Energy*, Vol. 86, No. 7, pp. 2017-2028, 2012.

https://doi.org/10.1016/j.solener.2012.04.004

[7]   P. Ramsami, V. Oree, A hybrid method for forecasting the energy output of photovoltaic systems, *Energy Conversion and Management*, Vol. 95, pp. 406-413, 2015.

https://doi.org/10.1016/j.enconman.2015.02.052

[8]   A. Mellit, S.A. Kalogirou, Artificial intelligence techniques for photovoltaic applications: A review, *Progress in energy and combustion science*, Vol. 34, pp. 574-632, 2008.

https://doi.org/10.1016/j.pecs.2008.01.001

[9]   M. Ding, L. Wang, R. Bi, An ANN-based approach for forecasting the power output of photovoltaic system, *Procedia Environmental Sciences*, Vol. 11, pp. 1308-1315, 2011.

https://doi.org/10.1016/j.proenv.2011.12.196

[10] V. Lo Brano, G. Ciulla, M. Di Falco, Artificial neural networks to predict the power output of a PV panel, *International Journal of Photoenergy*, Vol. 2014, 2014.

https://doi.org/10.1155/2014/193083

[11] 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, pp. 1064-1069, 2012.

https://doi.org/10.1109/TIA.2012.2190816

[12] W. Xiao, G. Nazario, H. Wu, H. Zhang, F. Cheng, A neural network based computational model to predict the output power of different types of photovoltaic cells, *PLoS ONE*, Vol. 12, No. 9, e0184561, 2017. https://doi.org/10.1371/journal.pone.0184561

[13] O. Abedinia, D. Raisz, N. Amjady, Effective prediction model for Hungarian small-scale solar power output, *IET Renewable Power Generation*, Vol. 11, No. 13, pp. 1648-1658, 2017. https://doi.org/10.1049/iet-rpg.2017.0165

[14] A. Saberian, H. Hizam, M.A.M. Radzi, M.Z.A. Ab Kadir, M. Mirzaei, Modelling and prediction of photovoltaic power output using artificial neural networks, *International journal of Photoenergy*, Vol. 2014, 2014. https://doi.org/10.1155/2014/469701

[15] A.G.R. Vaz, B. Elsinga, W.G.J.H.M. van Sark, M.C. Brito, An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in Utrecht, the Netherlands, *Renewable Energy*, Vol. 85, pp. 631-641, 2016.

https://doi.org/10.1016/j.renene.2015.06.061

[16] N.U. Aningo, A. Hardy, D. Glew, Evaluating Solar Prediction Methods to Improve PV Micro-grid Effectiveness Using Nonlinear Autoregressive Exogenous Neural Network (NARX NN), In: L. Scott, M. Dastbaz, C. Gorse (eds), *Sustainable Ecological Engineering Design,* Springer, pp. 363-376, Cham 2020.

https://doi.org/10.1007/978-3-030-44381-8_28

[17] M. Louzazni, H. Mosalam, A. Khouya, A non-linear auto-regressive exogenous method to forecast the photovoltaic power output, *Sustainable Energy Technologies and Assessments*, Vol. 38, 100670, 2020. https://doi.org/10.1016/j.seta.2020.100670

[18] K.L. Priddy, P.E. Keller, *Artificial neural networks: an introduction*, SPIE press, 2005. https://doi.org/10.1117/3.633187

[19] S. Haykin, *Neural Network: A Comprehensive Foundation,* Pearson Education (Singapore) Pte. Ltd., Indian Branch, 482 F.I.E. Patparganj Delhi 110092, India, Pearson Prentice Hall, 2005.

[20] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de-Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, *Solar Energy*, Vol. 136, pp. 78-111, 2016. https://doi.org/10.1016/j.solener.2016.06.069

[21] M.N. Akhter, S. Mekhilef, H. Mokhlis, N.M. Shah, Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques, *IET Renewable Power Generation*, Vol. 13, No. 7, pp. 1009-1023, 2019.

https://doi.org/10.1049/iet-rpg.2018.5649

[22] T. Lin, B.G. Horne, P. Tino, C.L. Giles, Learning long-term dependencies in NARX recurrent neural networks, *IEEE Transactions on Neural Networks*, Vol. 7, pp. 1329-1338, 1996. https://doi.org/10.1109/72.548162

[23] S.A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains,* John Wiley & Sons, 2013.

https://doi.org/10.1002/9781118535561

[24] R.H. Inman, H.T. Pedro, C.F. Coimbra, Solar forecasting methods for renewable energy integration, *Progress in Energy and Combustion Science*, Vol. 39, No. 6, pp. 535-576, 2013.

https://doi.org/10.1016/j.pecs.2013.06.002

[25] S. Gupta, D. Srinivasan, T. Reindl, Forecasting solar and wind data using dynamic neural network architectures for a micro-grid ensemble, In *2013 IEEE Computational Intelligence Applications in Smart Grid (CIASG)*, pp. 87-92, 2013.

https://doi.org/10.1109/CIASG.2013.6611503

[26] https://power.larc.nasa.gov/data-access-viewer/ Checked on August 1st, 2019.

[27] H. Sharadga, S. Hajimirza, R.S. Balog, Time series forecasting of solar power generation for large-scale photovoltaic plants, *Renewable Energy*, Vol. 150, pp. 797-807, 2020.

https://doi.org/10.1016/j.renene.2019.12.131

[28] D.K. Chaturvedi, I. Isha, Solar power forecasting: A review, *International Journal of Computer Applications*, Vol. 145, No. 6, pp. 28-50, 2016.

https://doi.org/10.5120/ijca2016910728

[29] K. Jazayeri, M. Jazayeri, S. Uysal, Comparative analysis of Levenberg-Marquardt and Bayesian regularization backpropagation algorithms in photovoltaic power estimation using artificial neural network, in Industrial Conference on Data Mining, pp. 80-95, 2016.

https://doi.org/10.1007/978-3-319-41561-1_7

[30] J.A. Duffie, W.A. Beckman, *Solar engineering of thermal processes*, 3rd ed. Wiley, 2006.

[31] A. Luque, S. Hegedus, *Handbook of Photovoltaic Science and Engineering,* John Wiley & Sons, 2011. https://doi.org/10.1002/9780470974704

[32] A. Gabour, S. Benzahioul, A. Metatla, T. Sebbagh, Optimization and Control Strategy of Multi-Source System Using Genetic Algorithm, *Carpathian Journal of Electrical Engineering*, Vol. 13, No. 1, pp. 45-60, 2019.

[33] H.B. Demuth, M.H. Beale, O. De Jess, M.T. Hagan, *Neural Network Design*, Martin Hagan, 2014.

[34] B. Goksu, M. Bayraktar, M. Pamik, Prediction of Solar Energy Potential with Artificial Neural Networks. In: Dincer I., Colpan C., Ezan M. (eds) Environmentally-Benign Energy Solutions, *Green Energy and Technology*, Springer, pp. 247-258, Cham, 2020.

https://doi.org/10.1007/978-3-030-20637-6_13

[35] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the Marquardt algorithm, *IEEE transactions on Neural Networks*, Vol. 5, No. 6, pp. 989-993, 1994.

https://doi.org/10.1109/72.329697

[36] F.D. Foresee, M.T. Hagan, Gauss-Newton approximation to Bayesian learning, In *Proceedings of International Conference on Neural Networks (ICNN'97)*, Vol. 3, pp. 1930-1935, IEEE 1997.

[37] M.F. Møller, A scaled conjugate gradient algorithm for fast supervised learning, *DAIMI Report Series*, Vol. 19, No. 339, pp. 1-22, 1990.

https://doi.org/10.7146/dpb.v19i339.6570