Jestr

Journal of Engineering Science and Technology Review 13 (1) (2020) 67 - 71

JOURNAL OF
**Engineering Science
and Technology Review**

www.jestr.org

Research Article

Levenberg-Marquardt Training Function using on MLP, RNN and Elman Neural Network to Optimize Hourly Forecasting in Tetouan City (Northern Morocco)

Yousra Amellas*, A. Djebli and A. Echchelh

Ibn Tofail University of Kenitra, Laboratory of Energy and Materials, B.P. 242, Kenitra, Morocco

Received 17 December 2019; Accepted 21 January 2020

Abstract

The study aims to find the most appropriate training function for three types of neural networks, MLP, RNN and Elman Neural Network from which we have chosen seven types of training functions. Levenberg-Marquardt algorithm has shown good results. The study also showed that Elman Neural Network has a good result for the short-term forecast for the Tetouan region, in a particular case.

Keywords: Levenberg-Marquardt, MLP, RNN, Elman Neural Network, Hourly Forecasting.

1. Introduction

In order to preserve the environment and produce green electricity renewable and ecological energies [1][2] have become an essential solution [3]. Our study focuses on wind energy [4] in the first place, which considers a predominant source of electricity production. This technology is completely dependent on wind. Not only does the wind behave randomly with respect to its characteristics and weather changes but also its continuous intermittency creates fluctuations in the production process that directly influence the stability of the power produced.

To properly manage the electrical system wind forecasting at different time scales becomes essential. To do this, our study focuses on predicting short-term wind speed [5]. Several so-called statistical methods [6] can be applied to carry out this type of forecast. Some of these methods are based on time series. They are generally based on historical data in order to predict the future activity, the most famous models are ARMA (Auto-Regressive Moving Average) [7], ARIMA (Auto-Regressive Integrated Moving Average) is the most used model in the field of wind energy which was established by box and Jenkins , ARX(Autoregressive with Exogenous Inputs)… which are generally linear models. On the other hand, wind is characterized by its non-linear behavior for this reason; several studies propose artificial intelligence methods such as fuzzy models, Random Forest (RF), k-Neighbors (kNN), Adaptive Neural Fuzzy Inference Systems (ANFIS) [8] ... and Artificial Neural Networks, the latter were the case of our study of which we have chosen three types: FNN (Feed-forward Neural Network), NARX (Non-Linear Autoregressive Exogenous) [9] et Elman Neural Network [10].

Our Study aims to find the most suitable training function for the three neural network used which are: FNN, NARX and Elman. We have chosen to compare seven of these functions: LM (Levenberg-Marquardt) [11], RB(Resilient back-propagation), SCG (Scaled Conjugate gradient), CGB(Conjugate Gradient with Powell Beale Restarts), CGF (Fletcher- Powell Conjugate Gradient), CGP(Polak-Ribiére Conjugate Gradient and OSS(One Step Secant).

According to the calculations, we have deduced that Levenberg-Marquardt training function [12] has the best results R^2=0,778, R^2=0,775, R^2=0,772, for the three networks respectively. One Step Secant has the lowest results with R^2=0.666 for FNN, R^2=0.739 for NARX and R^2= 0.755 for Elman neural network.

2. Research Method

In order to find the best model to predict wind speed on a short time scale we rely on meteorological data from the city of Tetouan [13][14] taken every hour throughout 2018 from the site www.wunderground.com. Our data is identified by wind speed (WS) en m/s, wind direction (DIR) en deg, the temperature (TMP) en °C, the dew temperature (DEW) en °C, atmospheric pressure (ATM) en hPa, and humidity (H) en %. We used our weather data in MATLAB software by performing the simulation by the NNTOOL application for the three types of neural networks.

2.1. Neural Network

An artificial neural network is part of the artificial intelligence technology [15] its functioning mimics the activities of the human cognitive system. Our artificial neural networks can be divided into two main classes FNN belongs to the machine learning class when NARX and Elman are part of the deep learning class named by Recurrent Neural Networks (RNN) [16]. Although, the three networks maintain the same three-layer architecture, an input layer with 5 inputs, an intermediate layer and an output layer holding a single output.

Our first type of FNN (Feed-forward Neural Network) neuron belongs to the family of Perceptrons based on the gradient back-propagation algorithm. A multi-layer Perceptron or Perceptron (depending on the number of hidden layers) [17] is a neural network of a classic hierarchy containing inputs or a single input, weights and Biases

*E-mail address: yousra.ame@gmail.com
ISSN: 1791-2377
doi:10.25103/jestr.131.09

corresponding to the inputs, a linear summation function, an activation function and one or more outputs, the network is defined by the following relationship:

$$y = a\,(\sum_{n=1}^{p} w_j\, x_j + w_0) \qquad (1)$$

There are several types of classical neural networks such as Adaline (Adaline (Adaptive Linear Neuron), Madaline (Multi-Adaline) also RBF (Radial Basis Function) [18].

The second class of networks is recurrent networks (RNN). This is a type of loop network, for which we have chosen closed loop architecture for NARX [19] and Elman neural network [20]. NARX feedback algorithm is similar to a multi-layer Perceptron only it relies on the system's input and output regressors to train the network from which it reports a new input from the output. The function that represents the operation of the NARX model in parallel mode or in other words close loop is:

$$\hat{y}\,(t+1) = \\ = \hat{f}(\hat{y}_{t-1}\,, \hat{y}_{t-2}\,, \dots\,, \hat{y}_{t-n}\,, \hat{x}_{t-1}\,, \hat{x}_{t-2}\,, \dots\,, \hat{x}_{t-n}\,) \qquad (2)$$

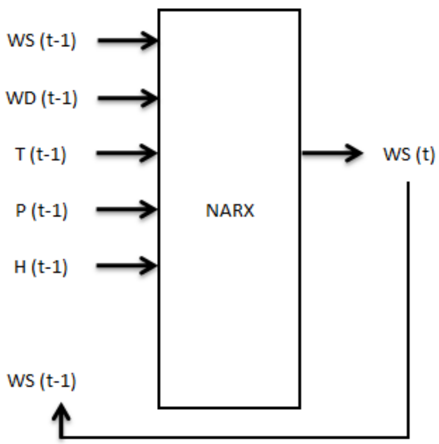

Fig. 1. Example of NARX-NN structure (Close loop structure)

Elman Neural Network (ENN) is a traditional neural network that has an additional input from the hidden layer named the context layer it is defined as partially recurring. ENN are generally used to study dynamic systems such as wind. There are many other networks similar to ENN such as the Jordan Neural Network, although there are other types of recurring neural networks such as LSTM (Long Short-Term Memory) [21] and GRU (Gated Recurrent Unit).

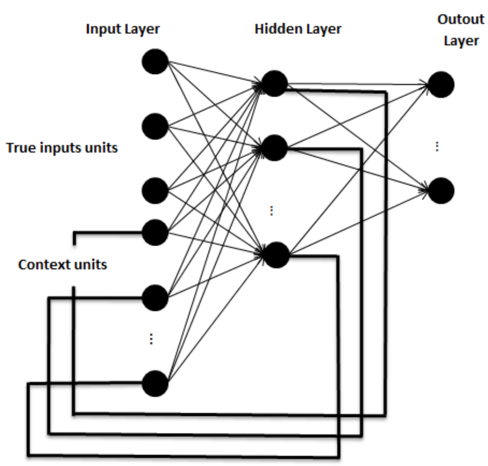

Fig. 2. Example of Elman structure (Close loop structure)

The Elman network model can expressed as:

The expression of output layer at *t* time is written:

$$O_q{}^{(t)} = f(\sum_{j=1}^{L} h_j{}^{(t)}\, w_{j,q}) \qquad (3)$$

The expression of output of hidden layer at t time is written:

$$h_j{}^{(t)} = f(\sum_{i=1}^{L} u_i{}^{(t)}\, w_{i,j} + \sum_{K=1}^{L} C_k{}^{(t-1)}\, w_{j,k}) \qquad (4)$$

For the input function at *t* time is written:

$$c_k{}^{(t-1)} = h_j{}^{(t)} \qquad (5)$$

Some case studies have been processed by NARX and Elman for the prediction of wind speed, hourly solar radiation as well as other case studies such as network traffic forecasting, pollution.

Annalisa Di Piazza et al used NARX to perform the hourly forecast of solar irradiation and wind speed, considering temperature as an exogenous variable. NARX was based on two techniques, the optimization technique based on a genetic algorithm (GA) and method that determine the optimal network architecture by pruning (optimal brain surgeon (OBS) strategy) [22].

Hong Thom Pham et al, chose to work with a high-powered hybrid NARX-ARMA model in order to predict the long-term condition of a machine. This forecast was based on vibration data. NARX is used to predict the deterministic component and ARMA to predict the error component [23].

As the NARX model is a type of dynamic neural network. Ines Sansa et al used a simulation to prove NARX's predictive effectiveness against static network models [24].

N. Mohana Sundaram et al, used the Elman type close loop architecture network to predict air pollution mortality, specifying the type of mortality (respiratory mortality, cardiovascular mortality and mortality from meteorological data) and finding the link between the two phenomena. As Elman is characterized by a good prediction of time series, they were able to satisfy the study [25].

Vlastimil Clupek et al studied the prediction of Network Traffic by several typologies of Elman neuron network on four types of Network Traffic. The results also showed Elman's good adaptation to dynamic systems [26].

2.2. Input variables

The table below shows the correlation between the meteorological variables, this allows us to eliminate some variables in order to avoid costly calculations during the simulation.

Table 1. Presentation of correlation coefficients between variables .

Variables	Wind Speed	Pressure	Temperature	Humidity	Dew point temperature
Wind speed	1	-0,1675	0,0508	-0,4094	-0,2233
Pressure		1	-0,2186	0,0718	-0,1265
Temperature			1	-0,3127	0,6850
Humidity				1	0,3581
Dew point temperature					1

After having calculated the correlation coefficient between the variables, we notice that the correlation coefficient between the temperature and the dew point temperature equal to 0.68507, this means that the two variables have the same climatological characteristics, and then we can eliminate the dew point temperature to facilitate the calculation.

3. Results and Discussion

The goal of our work is to find the most appropriate training function (or training algorithm) for the three types of neural networks; this will allow us to more quickly choose the short-term wind speed forecast model for the Tetouan region. As mentioned above, we have chosen three types of RNA, a feed-forward neural network model, NARX and Elman close loop Neurons Network. These were trained by seven types of training algorithms: Levenberg-Marquardt (LM), Resilient Back-propagation (RB), Scaled Conjugate Gradient (SCG), Conjugate Gradient with Powell/Beale Restarts (CGB), Fletcher-Powell Conjugate Gradient (CGF), Polak-Ribiére Conjugate Gradient (CGP) and One Step Secant (OSS). Choosing the right training function reduces the complexity of the problem in terms of calculation even more, choosing the right number of neurons in the hidden layer of each type of network has a positive influence on the accuracy of the choice of parameters. We have 5 inputs at the moment (t-1): WS(t-1), DIR(t-1), ATM(t-1), H(t-1), TMP(t-1). For the hidden layer, the numbers of neurons have been varied, FNN shows good results by integrating 200 Neurons, with a sigmoid activation function, the most frequently used, it allows the reduction of the input value to reduce it between 0 and 1.

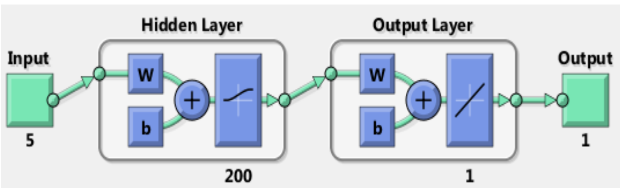

Fig. 3. FNN Neural Network – 200 Neurones – Logsig fonction activation

As for the NARX close loop type, it has 70 hidden layer neurons, and a Tanh function (hyperbolic tangent), similar to the sigmoid function except that this first one produces results between -1 and 1 (centered in zero).

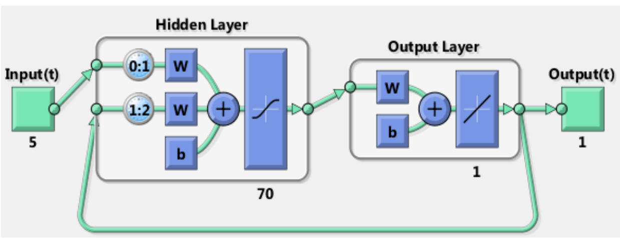

Fig. 4. NARX Neural Network – 70 Neurones – Tansig fonction activation

For Elman Neural Network, it has 30 neurons in a hidden layer, keeping the Tanh function. The output layer for all networks contains the wind speed in time t.

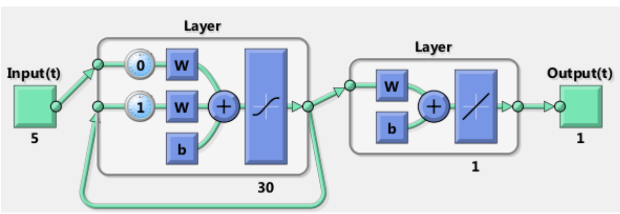

Fig. 5. Elman Neural Network – 70 Neurons – Tansig function activation

Figures 6, 7 and 8 present the results of the three networks, FNN, NARX and Elman according to the seven training algorithms. In order to compare the results between the desired output and the output calculated by the networks we have chosen the windiest day of the year, which is January 29, 2018. We found that the output calculated by the Levenberg-Marquardt algorithm is similar to the actual output .The Levenberg-Marquardt algorithm or in other words damped least-squares (DLS)known for its good memory and fast calculation On the other hand, the One Step Secant algorithm shows low similarity.

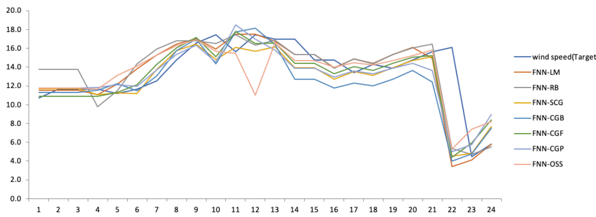

Fig. 6. FNN (with all training functions)

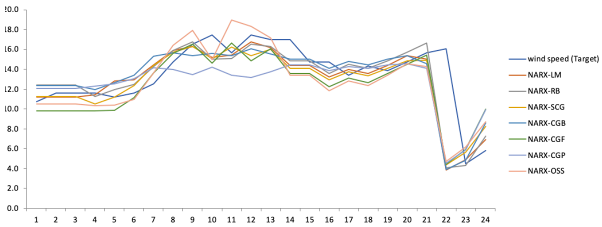

Fig. 7. NARX (with all training functions)

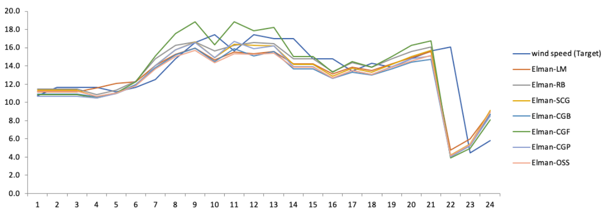

Fig 8. Elman (with all training functions)

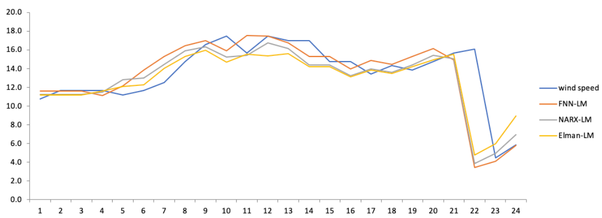

Fig. 9. Comparison of LM

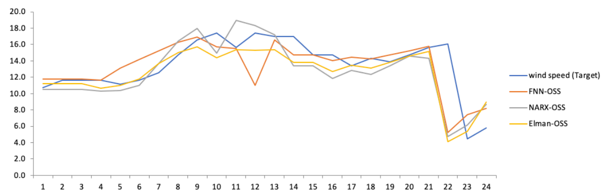

Fig. 10. Comparison of OSS

After making sure that LM is generally the best algorithm for training dynamic and static networks. The figures above 9 and 10 allow us to conclude that Elman is the best network for processing our non-linear time series.
As we have a problem predicting quantitative variables, it is automatically a regression problem. To analyze this

regression, we used four types of forecast error measurements as shown in the table below. MSE, RMSE, R^2 and MAE [27]. With:

MSE (*Mean Square Error*) = $\frac{1}{n}\sum_1^n \ (O-d)^2$

RMSE (*Root Mean Square Error*) = $\sqrt{\frac{1}{n}\sum_1^n \ (O-d)^2}$

R^2 (*Coefficient of determination*) = $1-\frac{\sum_{i=1}^n (O-d)^2}{\sum_{i=1}^n (O-\bar{d})^2}$

MAE (*Mean absolute error*) = $\frac{\sum_{i=1}^n |O-d|}{n}$

Table 2. Statistical errors generated by different algorithms

		LM	**RB**	**SCG**	**SCB**	**CGF**	**CGP**	**OSS**
FNN	MSE	1,7954	2,2464	1,98004	2,0262	2,098	2,1455	2,708
	RMSE	1,3399	1,4988	1,4071	1,4234	1,4487	1,4647	1,645
	R²	0,7787	0,7231	0,7559	0,7502	0,7413	0,735	0,666
	MAE	0,99	1,129	1,043	1,0598	1,0763	1,0876	1,2289
NARX	MSE	1,82122636	2,101210513	1,96408045	1,9513738	2,05772554	2,03122551	2,113257586
	RMSE	1,3495282	1,449555281	1,40145654	1,39691582	1,43447745	1,42521081	1,45370478
	R²	0,7755485	0,741042704	0,75794288	0,75950887	0,74640188	0,74966775	0,739558
	MAE	0,99840004	1,080492228	1,03139979	1,03173873	1,06420477	1,06386072	1,073848202
Elman	MSE	1,8449411	1,941150921	2,02079432	1,97844128	2,01660377	2,00397398	1,983718146
	RMSE	1,35828609	1,393251923	1,42154645	1,40657075	1,42007175	1,41561788	1,408445294
	R²	0,77262585	0,767076876	0,75095335	0,75617303	0,75146981	0,75302633	0,755522694
	MAE	1,00364669	1,024598747	1,04702693	1,03570849	1,044074	1,04261823	1,03333355

4. Conclusion

The article tries to find the most suitable training algorithm for the three types of neural networks in order to find the best wind speed prediction model. The study included FNN, NARX and Elman Neural Network while keeping the same architecture, and by modifying only the parameters of each network. Training algorithms showed good overall similarity but the output variables found by the Levenberg-Marquardt (LM) algorithm were very similar to the real variables. At the same time, we were able to conclude that Elman provides a good model for processing our weather compared to other networks.

References

1. A. Alhamwi, D. Kleinhans, S. Weitemeyer, T. Wogt, Moroccan National Energy Strategy reviewed from a meteorological perspective, Energy Strategy Reviewed 6, p. 39-47 (2015)
2. M. Azeroual, A. El Makrini, H. El Moussaoui, H. El Markhi, Renewable Energy Potential and Available capacity for wind and solar Power in Morocco Towards 2030, Journal of Engineering Science and Technology Review 11(1), p. 189-198 (2018)
3. A. Nouri, M. Ait Bahram, E. Elwanaki, M. Enzili, Moroccan wind farm potential feasibility. Case Study, Energy conversion and Management 112, p. 39-51 (2016)
4. A. Allouhi, Energetic, exergetic, economic and environmental (4E) assessment process of wind power generation, Journal of Cleaner Production 235, p.123-137 (2019)
5. Y. Amellas, O. El Bakkali, A. Djebli, A. Echchelh, Short-term wind speed prediction based on MLP and NARX networks models, Indonesian Journal of Electrical Engineering and Computer Science, Vol 18, N°1, p.150-157 (2020)
6. R. Sharma, D. Singh, A review of wind Power and Wind Speed Forecasting, Journal of Engineering Research and Application 8, Issue 7(Part 3), p.01-09 (2018)
7. E. Erdem, J. Shi, ARMA based approaches for forecasting the tuple of wind speed and direction, Applied Energy 88, p. 1405-1414 (2011)
8. M. Mohandes, S. Rehman, S.M. Rahman, Estimation of wind speed profile using adaptive neuro-fuzzy inference system, Applied Energy 88, p.4024-4032 (2011)
9. O. Bektas, J.A. Jones, NARX Time Series model for Remaining used life Estimation of gas turbine Energies, European Conference of the Prognostics and Helth Management Society (2016)
10. A. Wysocki, M. Lawrynczuk, Elman neural network for modeling and predictive control of delayed dynamic systems, Archives of Control Sciences, Vol 26, N°1, p.117-142 (2016)
11. S. Mammadi, Financial time series prediction using artificial neural network based on Levenberg-Marquardt algorithm, 9th International International Conference on Theory and Application of Soft Computing, Computing with words and Perception, ICSCCW, Procedia Computer Science 120, 602-607 (2017)
12. M. Hock Fun, M.T. Hagan, Levenberg-Marquardt Training for Modular Networks, Proceedings of International Conference on Neural Networks (ICNN'96) (1996)
13. Y. El Khchine, M. Sriti, N. Eddine El Kadri Elyamani, Evaluation of wind energy potential and trends in Morocco, Heliyon 5, e01830 (2019)
14. H. Bidaoui, I. El Abbassi, A. El bouardi, A. Darcherif, Wind Speed Data Analysis Using Weibull and Rayleigh Distribution Function, Case Study: Five cites Northern Morocco, Procedia Manufacturing, Vol 23, p.786-793 (2019)
15. A. El Ibrahimi, A. Baali, Application of several Artificial Intelligence Models for forecasting Meteorological Drought Using the standardized precipitations Index in the Saiss Plain (Northern Morocco), International Journal of Intelligent Engineering ans Systems, Vol 11, N°1, p.267-275 (2018)
16. G. Gelly, Réseaux de neurones récurrents pour le traitement automatique de la parole (Thèse de Doctorat) (2017)
17. A. El Ibrahimi, A. Baali, Application of Neural Modeling and the SPI Index for the prediction of weather Drought in the Saiss Plain (Northern Morocco), Vol 10, N°5, p.1-10 (2017)
18. G.W, H.J. Lu, Y.R. Chang, Y.D.Lee, An improved neural network-based approach for short-term wind speed and power forecast, Renewable Energy 105, p.301-311 (2017)
19. A. A. Hussein, ADerivation and comparison of open loop and closed loop Neural Network Battery, State-of-Charge Estimators, Energy Procedia 75, p.1856-1861 (2015)
20. C. Li, L. Zhu, Z.He, H. Gao, Y.Yang, D.Yao, X.Qu, Runoff Prediction Method Based on adaptive Elman Neural Network, Water, 11(6), 1113 (2019)
21. X. Le , H. Viet Ho, G. Lee, S. Jung, Applications of Long Short-term Memory (LSTM) Neural Network for Flood Forecasting, Water 11, 1387 (2019)

22. A. Di Piazza, M. Carmela Di Piazza, G. Vitale, Solar and Wind forecasting by NARX Neural Networks, Renewable Energy and Environmental Sustainable 1, 39 (2016)
23. T. Van Tung, P, H.Thom,Yang, BoSuk ,An improved hybrid of nonli near autoregressive with exogenous input and autoregressive movig average for longterm machine state forecasting based on vibration signal. Expert Systems With Applications, 37 (4). p. 3310-3317. ISSN 0957-4174 (2010)
24. I. Sansa, N. Mrabet Bellaaj, Solar Radiation using NARX model, Advanced Applications for Artificial Neural Network, Chapter 13 (2018)
25. N. Mohana, S.N. Sivanandam, R. Subla, Elman Neural Network Mortality Predictor for Prediction for Prediction of Mortality due to pollution, International Journal of Applied Engineering Research ISSN 0973-4562 Vol 11, N°3, p. 1835-1840 (2016)
26. V. Clupek, P. Dzurenda, Prediction of Network Traffic by Elman Neural Network, elektrorevue, Vol 5, N°1, ISNN 1213-1539 (2014)
27. M.V. Shcherbakov, A.Brebels, N.L. Shcherbakova, A.P. Tyukov, T.A. Janosky, V.A.Kamaev, Survey of Forecast Error Measures, World Applied Sciences Journal 24 (Information Technologies in Modern Industry, Education and Society) , p. 171-176 (2013)