International Research Journal of Computer Science (IRJCS) ISSN: 2393-9842
Issue 11, Volume 2 (November 2015) www.irjcs.com

Development of an Artificial Neural Network model to predict The monthly air temperature in the region of Meknes (Morocco)

Mustapha BEN EL HOUARI
Research Team "Analytical Chemistry and Environment"
University Moulay Ismail
Faculty of Sciences

Omar ZEGAOUI
Research Team "Applied Materials and Catalysis"
University Moulay Ismail
Faculty of Sciences

Abdelaziz ABDALLAOUI
Research Team "Analytical Chemistry and Environment"
University Moulay Ismail
Faculty of Sciences

Abstract*— In order to establish an empirical mathematical model to predict the monthly air temperature in the region of Meknes in Morocco, three types of artificial neural networks (ANN): Multi layer perceptron (MLP), cascade feed forward and Elman recurrent network were studied. The performances of the developed models with these types of ANN are discussed and compared with those of the model developed by the multiple linear regressions. The database used contains a monthly historical of meteorological parameters recorded in the region of Meknes between 1996 and 2013. These parameters are atmospheric pressure, humidity, precipitation, visibility, wind speed, maximum wind speed and dew point as input parameters and air temperature as an output parameter. The obtained results, based on the correlation coefficients (R), the mean squared error (MSE) and the sums of squared errors (SSE) demonstrate that the air temperature prediction is optimal and more efficient with the MLP model, Levenberg-Marquard algorithm with architecture [7-4-1] and the Tansig function in the hidden layer and the Purelin function in the output layer.

***Keywords*— ANN, MLP, Levenberg-Marquard algorithm, MSE, R, SSE, Air temperature.**

I. INTRODUCTION

Artificial neural networks are improved techniques of data processing able to model relationships between particularly complex functions. Indeed, artificial neural networks (ANN) have been successfully used in various domains of science and engineering because of its ability to model both linear and non-linear systems without the need to make assumptions as are implicit in conventional statistical approaches. The ANN predictive technique has been used in weather events **[1;2;3]**, stock market **[4]**, cloud classification and identification **[5;6]**, Several authors have been developed ANN models for air temperature prediction in many countries such as Dombayc *et al.* **[7]** in Denizli (Turkey), Almonacid *et al.* **[8]** in several regions in Spain and Tasadduq *et al.* **[9]** in Jeddah in Saudi Arabia. . Also, Behrang *et al.* **[10]** developed the multilayer Perceptron (MLP) and radial basis function (RBF) neural networks to predict the daily sunlight in Dezful in Iran. These authors considered different weather variables, including average air temperature, relative humidity, sunshine hours, evaporation, and values of the wind speed between 2002 and 2006. The objective of this study is to establish an artificial neural network model to predict efficiently the monthly air temperature in the region of Meknes, Morocco. In this context, three ANN models were developed including two static types (Multilayer Perceptron (MLP), cascading network (cascading feed forward)) and the third is the recurrent Elman network. Their performances were compared to those of the multiple linear regressions.

II. MATERIAL AND METHODS

A. *Normalization of the database*

In general, the database must be pretreated so that its values can be adapted to the inputs and outputs of the network. A common pre-processing comprises an appropriate normalization, which is applied in order to consider the amplitude of the values accepted by the network. The learning base consists of seven vectors X_1, X_2, ..., X_7 that are the independent variables who have been normalized between 0.1 and 0.9 according the following formulation **[11]**:

$$X_{ij} = 0.8 \times \frac{X_{ij} - X_{i\min}}{X_{i\max} - X_{i\min}} + 0.1 \quad (1)$$

X_{ij} : Standard of variable X_i for observation values j,
X_{ij} : Actual values of variable X_i for observation j,
$X_{i\,min}$: Minimum value of the variable X_i,
$X_{i\,max}$: Maximum value of the variable X_i.

The values of the dependent variable were normalized in the range [0, 1] following the equation (2):

$$Y_j = \frac{Y_j - Y_{min}}{Y_{max} - Y_{min}} \quad (2)$$

Y_j : Normalized value of the variable Y for observation j,
Y_j : Actual value of the variable Y for observation j,
Y_{max} : maximum value of the variable Y,
Y_{max} : Minimum value of the variable Y,

Table 1 : Meteorological variables

Parameter	Designation	Min	Average	Max	Type of parameter
Atmospheric pressure (hpa)	Pr	1011.8	1017.1	1027.8	Independent variables
Humidity (%)	H	39.5	64.2	83.4	
Precipitation (mm)	P	0	39.1	331.7	
Visibility (km)	Vis	6.8	8.8	10.2	
Wind speed (km/h)	V	3.0	10.6	17.6	
Maximum wind speed (km/h)	Vmax	11.0	20.4	27.8	
dew point (°C)	Tr	2.0	11.1	17.0	
Air temperature (°C)	Tair	1.8	17.4	29.2	Dependent variable

B. Multiple Linear Regression

Multiple linear regression (MLR) was used to predict the values of the dependent variable from independent variables. It is used to find the best linear model to predict the dependent value that produces the minimum error. In such model, each independent variable is weighted so that the value of the correlation coefficients maximizes the influence of each variable in the final equation **[1]**.

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n \quad (3)$$

a_i : The regression coefficients i = 1, 2, …; n.

C. Artificial neural networks

The neural network is defined as an assembly of identical structural elements called interconnected cells (or neurons) like the ones of the vertebrates' nervous system (Biological neurons) **[12]**. The authors as in **[13]** were inspired to develop formal or artificial neurons. Thus, similarities were established between the elements of biological neurons and components of Artificial Neural Formal (Figure 1). An artificial neuron is defined as a non-linear algebraic function set, to bounded values, actual variables called "input", with three basic elements: a set of connection weights, a threshold, and an activation function **[14]**.

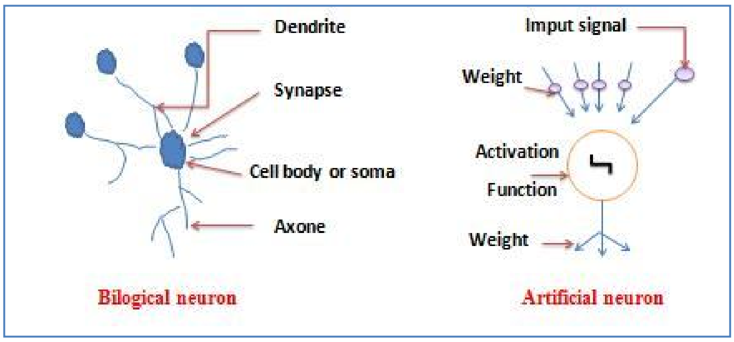

Figure 1 : Analogy between biological and artificial neuron **[15].**

1) Multilayer perceptrons: The Multilayer Perceptrons are networks having structure follows a logic of information processing through successive layers of artificial neurons, from input to output, without return upstream information **[16].** For this type of network, each neuron in a layer is connected to each neuron in the preceding layer and in the next layer (except for the input and output layers) and no connection exists between the cells of the same layer. Figure 2 shows an example of Multilayer Perceptron type network with single hidden layer neurons.

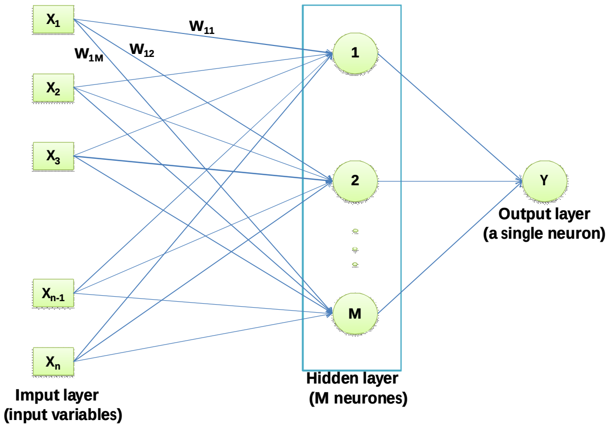

Figure 2 : Example of Multilayer Perceptron neural network simplified with a single hidden layer.

2) Cascade feed forward: These artificial neural networks are similar to MLP networks. They have unidirectional connections forward (feed forward) and each neuron fully connected to the next layer units. The remaining fact is that the neurons of the input layer are connected to the neurons of the output layer. Figure 3 shows a network in cascade with a single hidden layer.

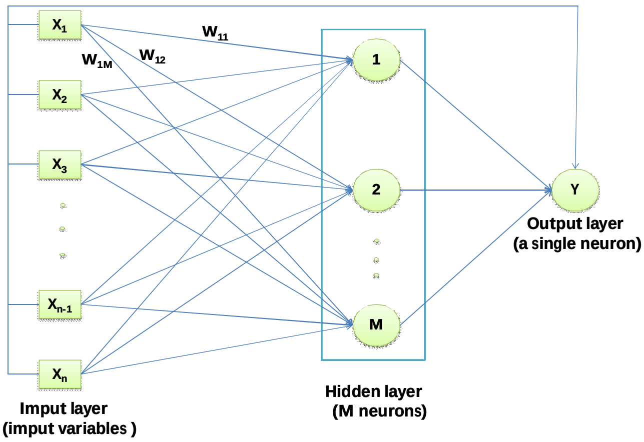

Figure 3 : Example of cascade feed forward neural network with a single hidden layer

3) Recurring networks "Elman": These networks' structure can comprise recurrences (Figure 4). These recurrences can dramatically change the dynamics that will be established in a network of neurons and make it self-perpetuating. The use of induction will be included in the context of multi-layer perceptrons with the Elman network. In this type of network, the activation of the hidden layer is duplicated back into the input layer. Recurrent networks using leaky integrators may be designated as the recurrent neural network for a continuous time. They are known to be theoretically capable of replicating any dynamic systems.

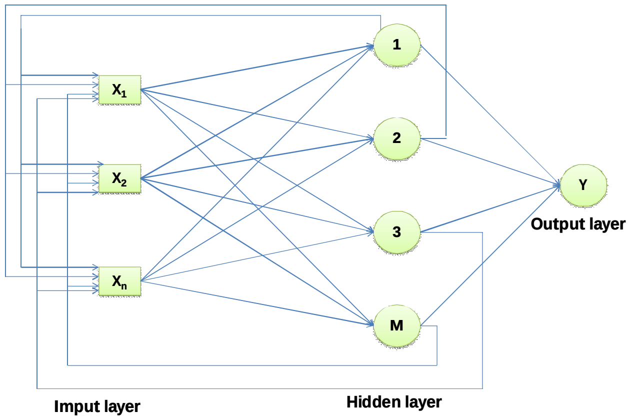

Figure 4 : Example of Elman network with one hidden layer.

D. Model Performance Evaluation

The correlation coefficient (R), the mean square error (MSE) and the sum of squared errors (SSE) were used to evaluate the predictive quality and performance of artificial neural network models developed for predicting monthly air temperature in the city of Meknes.

1) *Correlation coefficient*: The correlation coefficient (R) between the desired values and those estimated for each neuron of the output layer is an additional parameter to estimate network performance model. It is calculated according to **[17]**:

$$R = \sqrt{1 - \frac{\sum_{j=1}^{N} (Y_{Pj} - Y_{Oj})^2}{\sum_{j=1}^{N} (Y_{Oj} - Y_m)^2}} \tag{6}$$

With $\mathbf{Y_m}$ as the average of the observed values

The correlation coefficient is between -1 and 1, reflecting a good performance of the network when its value is close to 1.

2) *Mean square error*: This allows the combined statistical index assessing variance and bias. It is used as the measure of the overall performance of the model. The model is well optimized if the value of MSE is close to zero, which tends towards a better performance and a perfect forecast. Its mathematical formulation is given by the following equation **[18]**:

$$MSE = \frac{1}{N} \sum_{j=1}^{N} (Y_{Pj} - Y_{Oj})^2 \tag{5}$$

With: $\mathbf{Y_{Pj}}$ and $\mathbf{Y_{Oj}}$ are respectively the predicted and observed values on the observation j;
N is the number of observations.

3) Sums of squared errors: This statistical index also allows a combined assessment of the variance and bias. The model is well optimized if the SSE value is close to zero. Its mathematical formulation is given by the following equation **[9]**:

$$SSE = \sum_{j=1}^{N} (Y_{Pj} - Y_{Oj})^2 \tag{7}$$

With Y_{Pj} et Y_{Oj} are respectively the observed and predicted values for the observation j.

III. Results and discussion

In this study, the database was divided in three phases: learning, testing and validation phases. The obtained values of R, MSE and SSE (results not presented in the text) indicate that the best distribution of the global database is 70%, 15% and 15% for the learning phase, the testing phase, and the validation phase respectively.

A. Multiple Linear Regression

Statistical analysis by the method of multiple linear regression was performed using Xlstat software on all of the database to predict the monthly air temperature. The obtained regression equation is:

 ISSN: 2393-9842
www.irjcs.com

YMLR = 62,1572 - [0,0450 x Pr] - [0,1277 x H] - [0,0123 x P[- [0,2201 x Vis[+ [0,1331 x V] - [0,0502 x Vmax] + [1,0209 x Tr]. (8)

The coefficient of correlation obtained by the model MLR (R = 0.77) and the mean squared error (MSE = 0.01) indicates that the air temperature does not correlate linearly with the other meteorological parameters.
Figure 5 shows the relationship between the observed and the estimated of the monthly values obtained for air temperature MLR method.

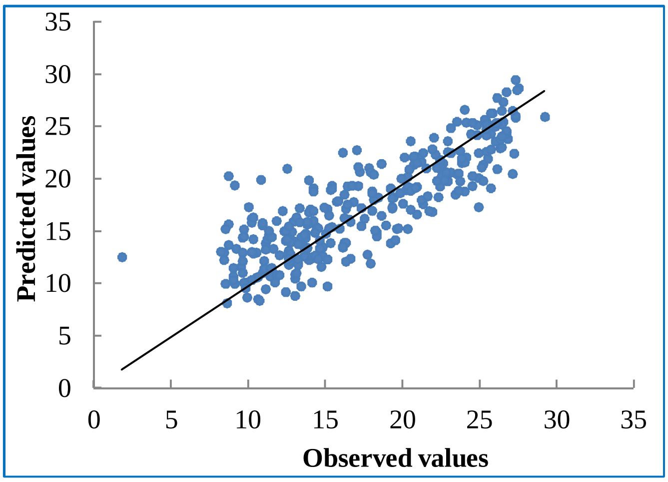

Figure 5 : Relationship between the observed and the estimated of the monthly values obtained for air temperature and estimated air by the MLR method.

B. Multilayer Perceptron

Because of the robustness of the Levenberg-Marquardt algorithm **[1;19;20]**, this algorithm was used with one hidden layer while changing the number of hidden neurons and the transfer functions couples. Table 2 presents the obtained values for the correlation coefficient, the mean squared error and the sum of errors. These results indicates that the MLP model, with the Levenberg-Marquardt as learning algorithm, the Tansig function in the hidden layer, and the Purelin function in the output layer, with the configuration [7-4-1] are most optimal architecture to predict the monthly air temperature.

Table 2 : Performance model developed by MLP type of neural network based on the transfer functions couples

Hidden layer	Output layer	Designation	R	MSE x 10^{+4}	SSE	Architecture
Tansig	Tansig	TT	0.982	62	1.93	[7-8-1]
Tansig	Logsig	TL	0.936	421	13.13	[7-9-1]
Tansig	**Purelin**	**TP**	**0.992**	**29**	**0.93**	**[7-4-1]**
Logsig	Tansig	LT	0.986	50	1.56	[7-10-1]
Logsig	Logsig	LL	0.941	427	13.32	[7-2-1]
Logsig	Purelin	LP	0.987	48	1.49	[7-7-1]
Purelin	Tansig	PT	0.977	85	2.65	[7-3-1]
Purelin	Logsig	PL	0.937	444	13.85	[7-8-1]
Purelin	Purelin	PP	0.981	66	2.06	[7-9-1]

Figure 6 shows the relationship between the predicted and observed values obtained for the monthly air temperature by MLP network. It shows a good correlation between observed and predicted air temperature values.

ISSN: 2393-9842
www.irjcs.com

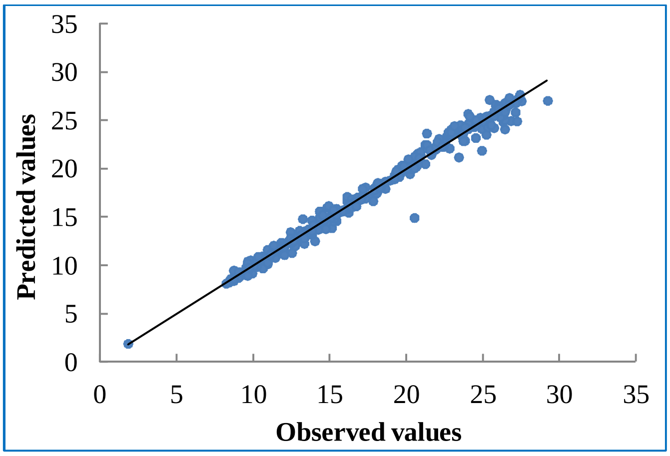

Figure 6 : Relationship between the predicted and observed values of the monthly air temperature obtained by MLP neural network.

C. Cascade feed forward

The Cascade feed forward is a type of neural network architectures that are similar to MLP networks. It also has forward unidirectional connections (feed forward); the neurons of the input layer are also connected to the output layer.

ISSN: 2393-9842
www.irjcs.com

Table 3 : Performance of the model developed by the Cascade neural network based on the transfer functions couples.

Hidden layer	Output layer	Designation	R	MSE x 10^{+4}	SSE	Architecture
Tansig	Tansig	TT	0.987	45	1.4	[7-4-1]
Tansig	Logsig	TL	0.941	420	13.1	[7-7-1]
Tansig	**Purelin**	**TP**	**0.988**	**40**	**1.24**	**[7-10-1]**
Logsig	Tansig	LT	0.989	36	1.12	[7-3-1]
Logsig	Logsig	LL	0.941	430	13.4	[7-3-1]
Logsig	Purelin	LP	0.987	44	1.37	[7-6-1]
Purelin	Tansig	PT	0.976	83	2.58	[7-8-1]
Purelin	Logsig	PL	0.930	440	13.72	[7-6-1]
Purelin	Purelin	PP	0.981	67	2.09	[7-7-1]

Figure 7 shows a strong correlation between the observed and estimated values of the air temperature obtained by the cascade neural network.

To determine the optimal network architecture, we varied the pair of transfer functions and number of neurons in the hidden layer. Table 3 shows the obtained values of R, MSE, and SSE for different pairs of transfer functions and different number of hidden neurons. For LM learning algorithm, the best performance is obtained with neural network architecture [7-10-1], the Tansig function as transfer function for the hidden layer and the Purelin function for the output layer. With 10 hidden neurons,

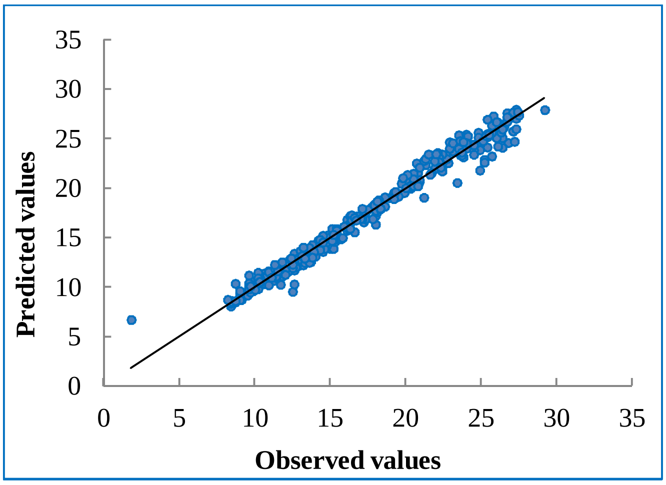

Figure 7 : Relationship between monthly observed air temperature and those predicted by the model developed by the cascade neural network.

D. Recurring networks "Elman"

For Elman model, the activation of the hidden layer is duplicated back into the input layer and the optimum structure of the network has found while performing a variation of the pairs of transfer functions and neuron numbers of the hidden layer. Table 4 shows the calculation of R, MSE and SSE for different pairs of transfer functions and a different number of hidden layer neurons. For LM learning algorithm, the best performance(R=0.982, MSE = 6.2×10^{-3} and SSE = 1.93) is obtained with neural network architecture [7-7-1], the Tansig function as transfer functions for the hidden layer and the Purelin function for the layer output with 7 hidden neurons.

Table 4 : Performance of the model developed by Elman neural network based on the transfer functions couples.

Hidden layer	Output layer	designation	R	MSE x 10^{+4}	SSE	Architecture
Tansig	Tansig	TT	0.980	68	2.12	[7-4-1]
Tansig	Logsig	TL	0.932	460	14.35	[7-5-1]
Tansig	**Purelin**	**TP**	**0.982**	**62**	**1.93**	**[7-7-1]**
Logsig	Tansig	LT	0.980	67	2.09	[7-9-1]
Logsig	Logsig	LL	0.932	450	14.04	[7-5-1]
Logsig	Purelin	LP	0.981	64	1.99	[7-8-1]
Purelin	Tansig	PT	0.975	84	2.62	[7-3-1]
Purelin	Logsig	PL	0.936	430	13.41	[7-6-1]
Purelin	Purelin	PP	0.982	66	2.05	[7-8-1]

Figure 8 shows a best correlation between the observed results and those estimated for the air temperature during the period of the study.

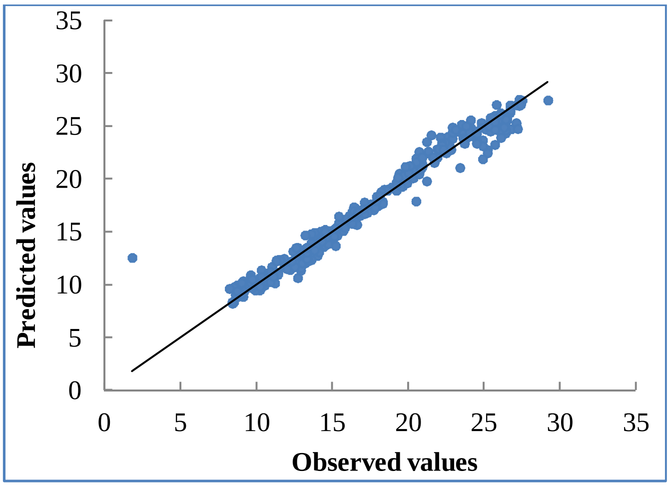

Figure 8 : Relationship between monthly observed and predicted values of air temperature by the developed Elman neural network's model.

E. Performance of the developed ANN models

Based on the R, MSE and SSE obtained values for the three models studied in this work, the best performance of MLP model was obtained with the Levenberg-Marquardt algorithm and pair transfer functions (Tansig-Purelin) with network architecture [7-4-1]. For the cascade network, the best performance is achieved with neural network architecture [7-10-1], the Tansig function as transfer function for the hidden layer and the Purelin function for the output layer with 10 hidden neurons.. For the recurrent model of Elman, the transfer functions (Tansig-Purelin) with LM algorithm, was obtained with network architecture [7-7-1].

Table 5 summarizes the obtained values of correlation coefficient, mean square errors and sums of squared errors for neural networks models of MLP, Cascade feed forward, Elman network, and multiple linear regression.

Table 5 : Correlation coefficients, mean square errors, and sum quadratic errors obtained for the MLP models, Cascade, Elman, and MLR

ANN model	R	MSE	SSE
MLP	**0.992**	**0.0029**	**0.93**
CASCADE	0.988	0.0040	1.24
ELMAN	0.982	0.0062	1.93
MLR	0.778	0.0100	3.12

So, the most effective model to predict the monthly air temperature in Meknes in Morocco is the MLP network with LM learning algorithm, network architecture [7-4-1], while using the office as Tansig transfer function for the hidden layer and the Purelin function for the output layer and four hidden neurons.

IV. Conclusion

In this study, three types of artificial neural networks including Multi layer perceptron, cascade feed forward, and Elman recurrent network were studied to predict efficiently the monthly air temperature in the region of Meknes, Morocco. The obtained results, expressed in terms of R, MSE and SSE, were compared with each other and with those obtained for the model developed by the multiple linear regression. These results indicated that the MLP model using the Levenberg-Marquard algorithm, the network architecture [7-4-1], the sigmoidal transfer function in the hidden layer, linear in the output layer, and 75% of the database chosen randomly for the learning phase was the optimal combination to predict successfully the monthly air temperature in the region of Meknes in Morocco.

References

[1] M. Ben El Houari, O. Zegaoui, A. Abdallaoui, Development of Mathematical Models to Forecasting the Monthly Precipitation. *American Journal of Engineering Research*, 03(11), 2014, 38-45.
[2] C. Marzbam, and G. Stumpf, Multiresolution wavelet transform and neural networks methods for rainfall estimation from meteorological satellite and radar data. *J. App. Meteor.*, *35*, 1996, 617-626.
[3] K. Hsu, H.V. Gupta, X. Gao, and S. Sorooshian, Rainfall Estimation from Satellite Imagery, Chapter 11 of Artificial Neural Networks in Hydrology, *Edited by R.S. Govindaraju and A.R. Rao, Published by Kluwer Academic Publishers*, 2000, 209-234.
[4] E. Collins, S. Ghosh, and C. Scofield, An application of a multiple neural network learning system to emulation of mortgage underwriting judgements. *Proceedings of the IEEE International Conference on Neural Networks*, 1988, 429.
[5] R. L. Bankert, Cloud classification of a advanced very high resolution radiometer imagery in maritime regions using a probabilistic neural network. *J. Appl. Meteor*, *33*, 1994, 909-918.
[6] J.E. Peak, and P.M. Tag, Sesmentation of satellite imagery using hierarchical thresholding and neural networks. *J. Appl. Meteor.*, *33*, 1994, 605-616.
[7] O. A. Dombayc, and M. Golcu. Daily means ambient temperature prediction using artificial neural network method: *A case study of Turkey*. *Renewable Energy*, 34, 2009, 1158-1161.
[8] F. Almonacid, P. Pérez-Higueras, P. Rodrigo, and L. Hontoria. Generation of ambient temperature hourly time series for some Spanish locations by artificial neural networks. *Renewable Energy*, 51, 2013, 285-291.
[9] I. Tasadduq. S. Rehman. K. Bubshait. Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia. *Renewable Energy*, 25, 2002, 545-554.

[10] M.A. Behrang. E. Assareh. A. Ghanbarzadeh. A. R. Noghrehabadi. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. *Solar Energy*, 84, 2010, 1468-1480.
[11] I. A. Basheer, M. Hajmer, Artificiel Neural Networks: Fundamentals, computing, design and application. *Journal of Microbiological Methods*, 43, 2000, 3-31.
[12] P. Coulibaly, F. Anctil, B. Bobee, Prévision hydrologique par réseaux de neurones artificiels : Etat de l'art. Revue canadienne de génie civil, 26, 1999, 293-304
[13] P. J. WERBOS, Applications of advances in nonlinear sensitivity analysis, *System modeling and optimization, New York*,1981, 762-770.
[14] S. Haykin. Neural Networks: *A Comprehensive Foundation, 2nd edition. Prentice Hall PTR*. 1998
[15] Y. B. koffi, K.E.Ahoussi, A.M. Kouassi, O. Kouassi, L.C. Kpangui, J. Biemi, Application des réseaux de neurones formels pour la prévision des débits mensuels du Bandoma blanc à la station de Tortiya (Nord de la cote d'ivoire), *Afrique science*, 10(3), 2014, 134-145.
[16] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323, 1986, 533–536.
[17] H. El Badaoui, A. Abdallaoui, S. Chabaa, Using MLP neural networks for predicting global solar radiation, *The International Journal Of Engineering And Science (IJES)*, 2(12), 2013, 48-56.
[18] N. Cheggaga, F. Youcef Ettoumi. Estimation du potentiel éolien. *Revue des Energies Renouvelables SMEE'10 Bou Ismail Tipaza*, 2010, 99 – 105.
[19] H. El Badaoui, A. Abdallaoui, and S. Chabaa. Perceptron Multicouches et réseau à fonction de base radiale pour la prédiction du taux d'humidité, *International Journal of Innovation and Scientific Research*, 5(1), 2014, 55-67.
[20] J. Moody, C. J Darken., Fast Learning in Network for Locally Tuned Processing Units. *Neural Computation*, 1, 1989, 281-294.