Le Séminaire International sur
L'Hydrogéologie et l'Environnement
5 - 7 Novembre 2013, Ouargla (Algérie)

Modelling Monthly Potential Evapotranspiration (ETP) Using Generalized Regression Neural Networks (GRNN): Case Study of the Semi-Arid Region of Guelma Northeast of Algeria.

HEDDAM Salim [1], LADLANI Ibtissem [2], HOUICHI Larbi [3], DJEMILI Lakhdar [4]

[1]Maître de conférences (MCB), Faculté des Science, Département des Sciences Agronomiques, Université 20 Août 1955, SKIKDA, *E-mail: heddamsalim@yahoo.fr*.
[2]Doctorant(e) en Sciences, Université Badji Mokhtar Annaba, Faculté des Sciences de l'Ingénieur, Département d'Hydraulique. *E-mail:* ladlaniibtissem@yahoo.fr
[3]Maître de conférences (MCA). Institut d'Architecture de Génie Civil et d'Hydraulique, Université Hadj Lakhdar - Batna. *E-mail:* houichilarbi@yahoo.fr
[4]Maître de conférences (MCA), Université Badji Mokhtar Annaba, Faculté des Sciences de l'Ingénieur, Département d'Hydraulique. *E-mail:* l_djemili@hotmail.com

***Abstract*— The aim of this study is to estimate the monthly potential evapotranspiration (ETP) based on class pan evaporation (E_P), using climatic data, in the agro meteorological conditions of the semi-arid region of Guelma, Northeast of Algeria country, using Generalized Regression Neural Networks (GRNN) based approach and multiple linear regression model (MLR). For the purpose of this paper, the generalized regression neural networks model (GRNN) and multiple linear regression models are developed and compared in order to estimate ETP. Various monthly climatic data, that is, monthly sunshine duration, maximum, minimum and mean air temperature, and wind speed from Guelma, Algeria, are used as inputs to the GRNN and MLR models. The performances of the models are evaluated using root mean square errors (RMSE), mean absolute error (MAE), Willmott index of agreement (d) and correlation coefficient (CC) statistics. Based on the comparisons, the GRNN was found to perform better than the MLR model.**

Key-Words— Potential Evapotranspiration (ETP), Modelling, Artificial Neural Network, GRNN, MLR

I. INTRODUCTION

Evapotranspiration (ET) is the simultaneous process of transfer of water to the atmosphere by transpiration and evaporation in a soil–plant system **[1]**. Accurate estimation of evapotranspiration is required for efficient irrigation management. Evapotranspiration is a complex process because it depends on several weather factors, such as temperature, radiation, humidity, wind speed and type and growth stage of the crop. Evapotranspiration (ET) being the major component of hydrological cycle will affect crop water requirement and future planning and management of water resources **[2]**. Evaporation pans (class A pan, US Weather Bureau) are used extensively throughout the world to estimate ETP. Evaporation pan (Ep) provides a measurement of the combined effect of temperature, humidity, wind speed and solar radiation on the reference crop evapotranspiration. This measurement can successfully be used to estimate ETP with a reasonable accuracy **[3]**. In this study, the potential of the generalized regression neural network (GRNN) is investigated for modeling monthly ETP based on class pan evaporation (E_p) using climatic data, in Guelma northeast of Algeria and to assess its performance relative to multiple linear regression (MLR).

II. METHODS

II.1. Artificial neural networks (ANNs)

Artificial Neural Networks (ANNs) are computational modeling tools that have recently emerged and found extensive acceptance in many disciplines for modeling complex real-world problems. ANNs modelling is basically a non-linear statistical analysis technique. It is essentially a black box linking input data to output data using a particular set of non-linear functions. An artificial ANNs consists of some basic elements called neurons. Each neuron (see **Fig. 2**) includes a set of inputs (called x_i), weight coefficients (called synaptic weights: w_i), and an activation function (f) **[4-5]**. Input variables are processed through successive layers of neurons. There is always an input layer, with a number of neurons equal to the number of variables of the problem (in this study five climatic variables), and an output layer, where the response is made available, with a number of neurons equal to the desired number of quantities (in this study only one output variable correspond to the reference evapotranspiration ET_P) computed from the inputs. Layers between the input and output layers are called hidden layers and may contain a large number of hidden processing units.

The ability to effectively approximate non-linear systems is due to the presence of one or more hidden layers and non-linear transfer functions in the hidden layer's neurons. The output of each neuron is determined by using an activation function, usually nonlinear activation functions are used such as Sigmoid or Gaussian. To obtain the desired output for any given input, the coefficients should be determined by training the network where sets of inputs with the corresponding outputs are given to the network through a training algorithm. This process should be repeated several times in order to minimize the output error. Each run of a complete set is called an epoch **[4-5].**

II.2. Generalized Regression Neural Networks (GRNN)

The generalized regression neural network **(**GRNN) is a neural network architecture that can solve any function approximation problems in the sense of estimating a probability distribution function. The network was firstly developed by **[6]**. GRNN is a universal approximator for smooth functions, so it should be able to solve any smooth function-approximation problem given enough data **[6]**. The GRNN consists of four layers, including the input layer, pattern layer, summation layer, and output layer. The first layer is fully connected to the second, pattern layer, where each unit represents a training input pattern and its output is a measure of the distance of the input from the stored patterns. Each pattern layer unit is connected to the two neurons in the summation layer: S-summation neuron and D-summation neuron. The S-summation neuron computes the sum of the weighted outputs of the pattern layer while the D-summation neuron calculates the unweighted outputs of the pattern neurons. The output layer merely divides the output of each S-summation neuron by that of each D-summation neuron, yielding the predicted value to an unknown input vector *x* as

$$Y_i = \frac{\sum_{i=1}^{n} y_i . \exp\left[-D\left(x \quad , x_i\right)\right]}{\sum_{i=1}^{n} \exp\left[-D\left(x \quad , x_i\right)\right]} \qquad (1)$$

Where n indicates the number of training patterns and the Gaussian D function in (1) is defined as

$$D\left(x \quad , x_i\right) = \sum_{k=1}^{m} \left(\frac{x_k - x_{ik}}{\pm}\right)^2 \qquad (2)$$

y_i is the weight connection between the *i*th neuron in the pattern layer and the S-summation neuron, *n* is the number of the training patterns, *D* is the Gaussian function, *m* is the number of elements of an input vector, x_k and x_{ik} are the *j*th element of *x* and x_i, respectively. The σ notation, known as the spread (or width), determines the generalization performance of the GRNN.

II.2. Multiple Linear Regression Model (MLR)

Multiple linear regression (MLR) is a well-known method of mathematically modeling the relationship between a dependent variable and one or more independent variables. In general, response variable Y may be related to *n regressor* variables. The following model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_n x_n \qquad (3)$$

Is called a multiple linear regression model with *n* regressor variables. Where β_0 is a constant and β_i, i = 1… n are regression coefficients.

Table 1. Statistical parameters of data set for the period of 1990-1997, Guelma station

Parameters	**Units**	X_{mean}	X_{max}	X_{min}	S_x	**Cv** (S_x/X_{mean})	**CC/ETP**
T_{max}	(°C)	24.51	39.80	13.60	7.38	0.30	**0.922**
T_{min}	(°C)	11.19	21.40	2.20	5.45	0.48	**0.848**
T_{mean}	(°C)	17.42	30.10	7.80	6.33	0.36	**0.913**
SN	h	7.44	12.20	3.50	2.14	0.28	**0.847**
U_2	m/s	2.10	9.90	1.30	0.93	0.44	**0.092**
ETP	mm	84.37	176.00	14.80	51.08	0.60	**1.000**

III. STUDY AREA AND DATA COLLECTION

Monthly measured climatic data for an 8 year period were obtained from Guelma weather station (latitude 36° 28‟ N, longitude 7° 28‟E, Altitude: 228 m).The climate is semi-arid. **Table 1** shows monthly maximum, minimum and mean air temperature (T_{max}, T_{min} and T_{mean}), Sunshine duration (SN), and wind speed (U_2) for Guelma station, used in this study. In this study, ETP (mm /month) was calculated by

$$ETP = Kp \times EP \tag{4}$$

In which EP is measured evaporation from a Class A Evaporation Pan (mm d-1) and Kp is a pan coefficient, which depends on the exposure of the pan, wind speed, humidity and distance of the pan from homogeneous materials [7]. Under ideal condition of the pan, *Kp* was reported to be 0.85 [8]. The monthly statistical parameters of each data are given in Table 1. In the table, the X_{mean}, X_{max}, X_{min}, S_x and CV denote the mean, maximum, minimum, standard deviation and variation coefficient, respectively.

Table 2. Performances of the GRNN and MLR models in different phases

Model	Training			Validation		
	CC	RMSE	MAE	CC	RMSE	MAE
GRNN	0.974	10.474	7.184	0.922	17.386	12.605
MLR	0.956	13.312	9.625	0.899	22.036	14.476

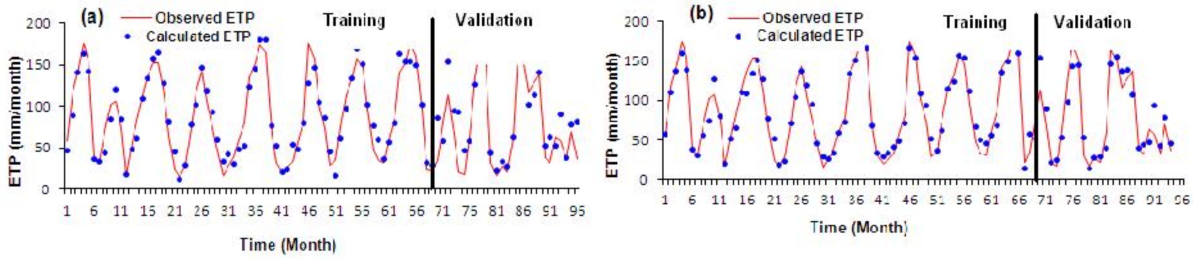

Figure 1. Comparison of observed and simulated series of ETP obtained using: (a) MLR and (b) GRNN models in the Training and validation phase.

IV. RESULTS AND DISCUSSIONS

The comparison of the models was based on both graphical plots and statistical indices (Table 2, Figs. 1 and 2). The GRNN model developed herein was found to yield better agreement with experimental observations for the training, validation, and testing data set compared to data predicted by the multiple linear regression (MLR) model. **Table 2** shows a statistical analysis of the GRNN and MLR models for training, validation and testing performances. According to **Table 2**, for the GRNN model in the training phase, the values of CC, RMSE, and MAE, are 0.974, 10.47, and 7.18, , respectively. In addition, in the validation phase, the values of CC, RMSE, and MAE, are 0.922, 17.38, and 12.60, respectively. The results of the MLR models for ETP prediction are presented also in **Table 2**. It may be seen from **Table 2,** the RMSE, MAE, and CC values for the model in the case of multiple linear regression (MLR) were found to be lower than those for the GRNN models, thereby establishing the superiority of the GRNN model. The prediction accuracy for the regression models was lower when compared to GRNN model for all the two phases. The simulated series of the observed versus calculated values of the ETP for the GRNN and MLR analyzed herein are shown in **Figs. 1** for the training and validation phases, respectively.

V. CONCLUSION

This paper investigate the potential of generalized regression neural network (GRNN) in modelling monthly potential evapotranspiration (ETP) using various climatic data from Guelma , northeast of Algeria country, in comparison to Multiple linear regression (MLR). A major conclusion from this study is that GRNN approach works well in predicting the potential evapotranspiration (ETP) in comparison to multiple linear regression (MLR). The comparison shows that there is better agreement between the results of the GRNN models and monthly potential evapotranspiration values compared to those of Evaporation pans (class ''A'' pan evaporation).

REFERENCES

[1] Tabari H, Grismer ME (2011). Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrig Sci.

[2] Blaney H.F., Criddle W.D. (1950). Determining water requirements in irrigated area from climatological irrigation data. US Department of Agriculture, Soil Conservation Service, Technical Paper no. 96, Washington, DC, USA.

[3] Irmak S, Allen RG, Whitty EB (2003) Daily grass and alfalfa reference evapotranspiration estimates and alfalfa-to-grass evapotranspiration ratios in Florida. J Irrig Drain Eng 129(5):360-370.

[4] Haykin S (1994) Neural Networks a Comprehensive Foundation. Prentice Hall, Upper Saddle River.

[5] Bishop C M (1996) Neural networks for pattern recognition. Oxford: Clarendon Press.

[6] Specht, D.F. (1991). A general regression neural network. IEEE Trans Neural Networks. 420 2(6):568-76.

[7] Jensen ME (1983) Design and operation of farm irrigation system. American Society of Agricultural Engineers, St. Joseph.

[8] Doorenbos J, Pruitt JO (1977) Guidelines for predicting crop water requirement. FAO Irrigation and Drainage Paper No. 27, Food and Agriculture Organization, Rome.