Conceptual Reference Evapotranspiration Models for Different Time Steps

Laaboudi A[1]*, Mouhouche B[2] and Draoui B[3]

[1]Algerian Institute for Research in Agronomy, Experimental Station of Adrar, Algeria
[2]National High School of Agronomic El Harrach, Algeria
[3]Bechar University, Algeria

Abstract

The evapotranspiration is one of the basic components of the hydrologic cycle and is essential for estimating irrigation water requirements. The use of Artificial Neural Networks (ANNs) in estimation of reference evapotranspiration has received enormous interest in the present decade. This paper describes the results obtained using neural network techniques to improve the accuracy of reference evapotranspiration estimation in different situations. Because the Neural networks are proved to be parsimonious universal approximators of nonlinear function, we have exploited this property to build various models in situation of lack of meteorological parameters and in different time steps. The FAO-56 Penman–Monteith equation (PM) was used to compute the reference evapotranspiration values.

The study showed that the neural network technique performed the best models even when it is feared the risk of co linearity and provided the best results by choosing appropriate architecture. They were able to reduce both Root Mean Squared Error and Mean Absolute Relative Error values and at the same time maximize the Nash-Sutcliffe efficiency and coefficient determination values.

Keywords: Water requirements; Neural network; Reference evapotranspiration; Meteorological parameters; Nonlinear function; Appropriate architecture

Introduction

The multiplication and worsening water deficiency states are taking across the world a dimension of the first order. The groundwater level is falling and threatens 1.5 billion people on the planet. It is therefore possible that water will become an international 33 strategic issue, which can lead to serious regional conflicts. In Algeria, the deficit of this blue gold has become disturbing confirming the hypothesis of diverse expertise and thereby making use of different methods which have all concluded that between 2010 and 2025, our country will be facing the most endemic shortage of water [1].

Furthermore, Algeria is an agricultural country where different climatic zones exist and almost 84 percent of the area is represented by arid region. Only a narrow belt of northern regions show humid climate. Most of the areas in the central and southern Algeria are highly arid, while the northern part of the country is humid. In theory, the current water availability per capita in Algeria is 500 cubic meters, down from 1.500 cubic meters in 1962. It is projected that it will further reduce to 450 cubic meters in 2020.

Despite its successes, investments to mobilize additional supplies of potable water and industrial water, irrigation have failed to match the growing demand. Recent droughts have exposed the vulnerability of large-scale irrigation systems and the pressure on groundwater resources. At the same time, new demands are emerging for major investment in wastewater treatment to counter the continuing threat that untreated sewage poses for health and long term sustainability of the country's water resources [2].

Available water appeared as the most important factor limiting wheat crop yields under the semi arid highland of eastern Algeria. The amount of grain yield produced per water use increased with the increase of availability of soil water and consequently water use efficiency increased [3]. Development of the irrigation sector and improvement of its planning system as part of the small-scale irrigation project activities are a big challenge for the government of Algeria.

According Liu et al. [4] for the sustainable development of oasis ecosystem, it is important to clarify an exact distribution of oasis land-use type where the excessive waste of the flood irrigation method has broken the balance between the water supply and requirement.

Moreover, accurate estimation of crop water requirements in the arid and semi-arid regions is crucial and important for sound water-use efficiency [5]. Indeed, semi arid regions are characterized by a water scarcity that is amplified by inefficient irrigation practices such as flood irrigation system.

Furthermore, according Hajare et al. [6], knowledge of exact amount of water required by different crops in a given set of climatological condition of a region is great help in planning of irrigation scheme, irrigation scheduling, effective design and management of irrigation system and also for midterm planning in case of mid season drought.

In this context, it should be noted that reference evapotranspiration (67 ET0) is one of the major components of the hydrologic cycle, and its accurate estimation is of utmost importance for many studies such as hydrologic water balance, irrigation system design and management, crop yield simulation, and water resources planning and management [7,8]. Many equations are used to estimate ET0 [9]. They can be divided into two main groups, i) those that are empirical and have lower data requirements, and ii) those that are physically-based and require proportionately more data, but the comparison between the results of these methods reveals a wide divergence.

*Corresponding author: Laaboudi A, Algerian Institute for Research in Agronomy, Experimental Station of Adrar, Algeria, E-mail: Laaboudiaek@yahoo.fr

Indeed, Alkaeed et al. [10] have compared six reference Evapotranspiration (ET0) methods, which are based on their daily performances under the given climatic condition in the western region of Fukuoka City. Their conclusions are; when considering the availability and reliability of the input data, the use of all these methods are suggested as practical methods for estimating ET0 if the standard FAO56-PM equation is not applicable due to the complexity of its input parameters. But, Wang et al. [11] have noted that four temperature based models including BP, RMBF, Hargreaves (HRG) and Blaney–Criddle (BCR) were employed and compared with the true PM. Based on the statistical evaluation, RMBF, HRG and BCR consistently overestimated the ET0 and showed poor performance.

Because it shows more realistic results [12,13], FAO has recommended the Penman-Monteith method [14], yet this approach was highly criticized due to its high number of meteorological parameters that are usually non available in most meteorological stations. That is why; we have decided to use Artificial Neural Networks (ANN) to model the reference evapotranspiration, based on a monthly, decade and daily time step.

It should be noted that an ANN is a computer model composed of individual processing elements called units or neurons. They are highly interconnected and operate in parallel. These elements are inspired by biological nervous systems. Within the human brain, individual cells, referred to as neurons, undertake discrete computation in massively parallel system. Neurons are responsible for the human capacity to learn and this significant property is used in machine learning in artificial network. An ANN emulates this computational capacity by distributing computations to several interconnected layers of simple processing units known as artificial neurons.

They have been successfully implemented to model evapotranspiration in several studies. These studies indicated that the ANN models can be used as an alternative method to estimate ET0. The performance of ANN models reported in these studies was superior to respective conventional methods of ET0 estimation [15]. However, according Xing et al. [16], this method overestimates the daily ET0 by 10% compared to the hourly ET0.

Furthermore, Shirgure and Rajput [17] have indicated that neural network is a new tool which can solve the more complex modelling problems like estimating evaporation from pan, which may be difficult to solve by conventional mathematical equations and multiple linear regression. It is observed from this review that the prediction model for evaporation is superior with neural networks. They have emerged as one of the useful artificial intelligence concepts used in the various engineering applications. Due to their massively parallel structure and ability to learn by example, ANNs can deal with nonlinear modelling for which an accurate analytical solution is difficult to obtain [18,19]. They are highly simplified mathematical models of their biological counterparts and include the ability to learn and generalize from examples to produce meaningful solutions to problems even when input data contain errors or are incomplete, and to adapt solutions over time to compensate for changing circumstances and to process information rapidly [20].

It is important to notice that Liu et al. [21] have found that the GNN model performed better than M-slat and BPNN models for modelling both runoff and evapotranspiration of Chinese fir plantations in China. Also, Laio et al. [22] had compared two nonlinear models, Nonlinear Prediction (NLP) and Artificial Neural Networks (ANN) for multivariate flood forecasting. They had found that, for NLP the calibration of the locally linear model is quite simple, while for ANN the validation and identification of the model can be cumbersome, mainly because of over fitting. Very good results are obtained with the two methods: NLP performs slightly better at short forecast times while the situation is reversed for longer times. Besides, Behzadi et al. [23] had employed about 6 nonlinear regression forms as counterparts to ANN, they had found that ANN generated a slightly better descriptive sheep growth curve than the best one which generated from nonlinear models and made the most accurate prediction.

This study contains three parts, the first discusses the main works performed in modelling field especially which have used neural network approach. The second presents description of study area, experimental conditions and model performance criteria. The third contains the obtained results and discusses them with a comparison at certain results obtained by researchers in the same context.

Materials

Our study was realized in the region of Adrar, located in the south-west of Algeria. Latitude: 27° 49' N and Longitude: 00° 18' E (Figure 1). Adrar is characterized by its extreme meteorological parameters.

Climate characteristic

Adrar's climate is dry throughout the year. The climate is characterized by the extended thermal amplitudes during the year, the month and even the day. The absolute maximum temperature reaches 49.5°C in summer (July and August). On the contrary, ice and frosts are rare in this region. Nevertheless, icy days can cause catastrophic damages, especially to traditional farming. Furthermore, it was recorded:

- A negligible pluviometry (<25 mm / year).

- A relative humidity often below 50%. The dew is very rare.

- A North-East wind blows almost constantly.

- A fluently clear sky with intense brightness.

Estimation of reference evapotranspiration

Next is the Penman-Montheith equation that is used for calculating

Figure 1: Sketch of the investigation area.
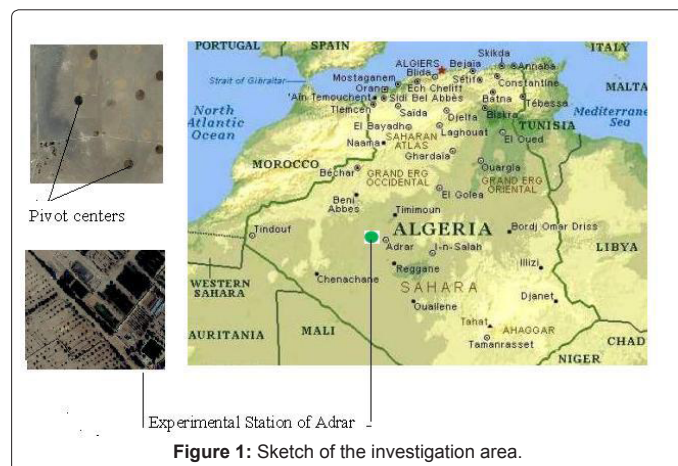
the reference evapotranspiration; it was proposed by Allen et al. [24]:

$$ETO = \frac{0.408\Delta\left(R_n - G\right) + \gamma\dfrac{900}{T+273}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)}$$

where ET0 is the reference evapotranspiration [mm day^{-1}], R_n is the net radiation at the crop surface [MJ m^{-2} day^{-1}], G is the soil heat flux density [MJ m^{-2} day^{-1}], T is the mean of daily air temperature at 2 m height [°C], u_2 is the wind speed at 2 m height [m s^{-1}], e_s is the saturation vapour pressure [kPa], e_a is the actual vapour pressure [kPa], e_s-e_a is the saturation vapour pressure deficit [kPa], Δ is the slope vapour pressure curve [kPa °C^{-1}], γ is the psychometric constant [kPa °C^{-1}].

The parameters: air temperature, sunshine duration, wind speed and relative humidity are taken directly from the meteorological station and are used to estimate other parameters, such as the net radiation, slope vapour pressure, psychometric constant and so on.

Neural network and models evaluation

The neural network is trained with a series of inputs and desired outputs from the training data set. The ANN used in this study is the feed forward network with the back propagation training algorithm. It is a supervised learning technique used for training artificial neural networks. Basically, it is a gradient descent technique to minimize the squared error between the calculated and desired outputs.

The back propagation algorithm, as noted by Parizeau [25], allows training the multilayer networks. To be useful, this network must have a non linear transfer function on hidden layers and the output layer according to the application type, either linear function or non linear function.

In this study, we have used three architectures according the time step modelling. In monthly time step, we have used two - layered Feed Forward Neural Network (FFNN) (Figure 2).

In decade time step, we have used also two - layered Feed Forward Neural Network (FFNN), but with 15 neurons in hidden layer and in daily time step we have used a structure of three - layered Feed Forward Neural Network (FFNN) (Figure 3). Its topology uses two sigmoid functions in hidden layers and one linear function in output layer as it is depicted in Figure 3.

Where IW {1,1} is the weight matrix in the first hidden layer; b {1} is the bias vector in the first hidden layer; LW {2,1} is the weight matrix in the second hidden layer; b {2} is the bias vector in the second hidden layer; LW {3,1} is the weight matrix in the output layer; b {3} is the bias vector in the output layer.

Description of data and availability

In the present investigation, daily data (temperature, sunshine duration, wind speed and humidity) consist of a series of daily values registered throughout the period of 1464 days. The registration of these meteorological statements was performed by the meteorological station within the experimental site and was used for the estimation of ET0.

Using these observed climatic data, daily values of ET0 were initially computed using the Penman-Monteith Equation (1). These computed ET0 values were used to train the ANN models. The database was divided into three subsets: 70% of data are used in the training phase, 15% in the testing phase; the remaining is reserved for validation. The idea behind this division is to:
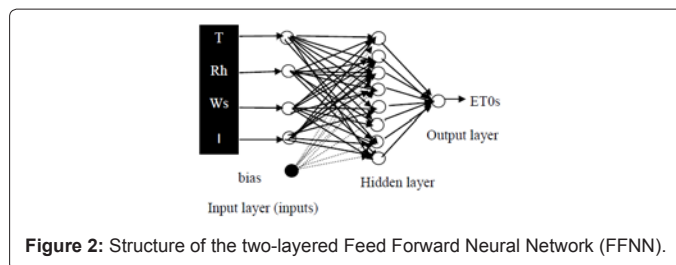

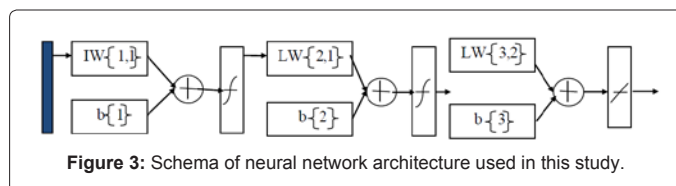**Figure 2:** Structure of the two-layered Feed Forward Neural Network (FFNN).

Figure 3: Schema of neural network architecture used in this study.

1- Take into consideration the seasonal tendency in ANN model.

2- Overcome the over fitting problem.

This also ensures the statistical properties of the training and testing data to be of similar order. As the climatic characteristics of the arid zones are important in assessing the applicability of the models in general, the statistical data of meteorological parameters in the study area are presented in Table 1.

We noticed that the variability range of meteorological parameters in the study area was large. For instance, the daily values of temperature ranged between 7.5°C and 41.6°C; relative humidity between 13% and 95%; duration of insolation between 0.00 and 12.30 hours/day; and wind speed was between 0.00 and 5.09 ms^{-1}. Hence, any model developed on this data set should have a wide application.

Selection of input variables

The correlations of all input variables are presented in Table 2. This table shows that the linear correlations between temperature and ET0 and also between relative humidity and ET0 are very high. Hence, any model that uses temperature and relative humidity should be able to estimate the ET0 satisfactorily. The model's accuracy can be improved by considering other variables that possess the aerodynamic effects on ET0, such as insolation duration and wind speed.

The temperature and humidity are also highly correlated. Therefore, a combination of these two factors may provide a good estimate. Wind speed and insolation are not well correlated with ET0. Nevertheless, these parameters are included in our model for better accuracy of the ET0 estimation. It should be noted that all these correlations between variables are linear type but the ET0 process is considered to be highly nonlinear. However, the high correlation conditions between input variables may pose problem of the co linearity in MLR modelling.

One can assess it by examining tolerance and the Variance Inflation Factor (VIF) which are collinear. The variable's tolerance is 1-R^2. A small tolerance value indicates that the variable under consideration is almost a perfect linear combination of the independent variables already in the equation.

The Variance Inflation Factor (VIF) is 1/tolerance, it measures the impact of collinearity among the variable in a regression model.

At the first glance (Table 2), there is a strong correlation between

	Meteorological parameters	Min	max	Mean	median	STDEV
Monthly averages data	T	11.58	38.86	25.39	25.75	8.69
	Rh	17.32	64.72	39.8	38.69	11.91
	Ws	1.04	2.26	1.67	1.66	0.27
	I	6.49	10.71	8.37	8.13	1.09
	ET0	2.733	10.214	6.364	6.322	2.376
decadal averages data	T	39.93	11.05	25.39	25.38	8.83
	Hr	70.15	16.6	39.8	37.4	37.4
	Ws	2.71	0.76	1.67	1.68	0.39
	I	12.25	4.54	8.37	8.37	1.46
	ET0	11.24	2.59	6.36	6.24	2.42
Daily data	T	7.5	41.6	25.34	26.05	8.84
	Hr	13	95	41.88	38.5	14.38
	Ws	0	5.09	1.7	1.63	0.75
	I	0	12.3	8.39	9	2.81
	ET0	1.43	13.52	6.37	6.31	2.6

Table 1: Statistical data of meteorological 1 parameters.

		Temperature	Humidity	Wind speed	Insolation	ET0
Monthly averages data	Temperature	1.00				
	Humidity	-0.82	1.00			
	Wind speed	0.20	-0.06	1.00		
	Insolation	0.36	-0.50	0.24	1.00	
	ET0	0.93	-0.86	0.37	0.57	1.00
decade averages data	Temperature	100				
	Humidity	-0.79	1.0 0			
	Wind speed	0.10	0.03	1.00		
	Insolation	0.24	-0.41	0.21	1.00	
	ET0	0.91	-0.82	0.32	0.49	1.00
Daily data	Temperature	1.00				
	Humidity	-0.78	1.00			
	Wind speed	0.01	0.07	1.00		
	Insolation	0.16	-0.27	0.02	1.00	
	ET0	0.86	-0.79	0.31	0.42	1.00

Table 2: Correlation matrix between input 3 and output variables.

T (temperature) and Rh (relative humidity) with R = -0.82. This gives VIF higher than 3.

Although some authors [26] accept a VIF lower than 5, otherwise, |R| lower than 0.90. It is best to observe a VIF lower than 3 or |R| <0.80.

If a low tolerance value is accompanied by large standard errors and non significance, multicollinearity may be an issue. That is why we prefer to not use this approach in daily time step.

Data normalization

We have used in this study the statistical or Z-score normalisation technique which uses mean ($\overline{x}$) and standard deviation (σ_x) of the original data in normalization process. The normalization scheme given was

$$Z_i = \quad Z_i = \frac{x_i - \overline{x}}{\sigma_x} \quad (2)$$

where Z_i, x_i, $\overline{x}$ and σ_x, are normalized value, real value, mean, and standard deviation respectively.

Criteria of evaluation

The performances of ANNs and MLR models were evaluated to compare their predictive accuracies based on the following statistical criteria:

The Nash-Sutcliffe efficiency (E) was proposed by Nash and Sutcliffe. It is calculated by formula (2) according to Krause et al. [27], the square value of the correlation coefficient (R), the Root Mean Squared Error (RMSE), the Mean-Squared Error (MSE) and the Mean Absolute Relative Error (MARE) were calculated as follows:

$$E = 1 - \frac{\sum_{i=1}^{n}(Y_{sim} - Y_{obs})^2}{\sum_{i=1}^{n}(Y_{sim} - Y_{obs})^2} \quad (3)$$

$$R = \frac{\sum_{i=1}^{n}(Y_{obs} - \overline{Y}_{obs})(Y_{Sim} - \overline{Y}_{sim})}{\sqrt{\sum_{i=1}^{n}(Y_{obs} - \overline{Y}_{obs})}\sqrt{\sum_{i=1}^{n}(Y_{Sim} - \overline{Y}_{sim})}} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{obs} - Y_{sim})^2}{n}} \quad (5)$$

$$MARE = \frac{1}{n} \sum_{i=1}^{n} \frac{|Y_{obs} - Y_{sim}|}{|Y_{obs}|} \times 100 \qquad (6)$$

with E Nash-Sutcliffe efficiency, Y_{sim} Simulated variable, Y_{obs} observed variable, $\overline{Y}_{sim}$ Average of Simulated variable, $\overline{Y}_{obs}$ Average of observed variable, n Number of observations.

For multiple regressions, we added the test of non colinearity parameters using the matrix of covariance, VIF (Variance Inflation Factor), the F statistical and T statistical.

We have used the neural network toolbox in Matlab (version 7), where we can find all necessary functions, already set, and we have programmed all the needed equations.

Results and Discussions

Monthly time step

The results presented in Table 3 showed that the ANNs can model the ET0 process by using only minimum climatic variables. Also they showed that the model performance is influenced by the number of inputs (Table 3). However, it turned out that the presence of temperature or relative humidity in the model gives satisfactory results. Farther, they showed the ability of the ANNs to estimate the ET0 by using air temperature, relative humidity and light hours.

By removing the daily sunshine duration variable, the result presented (Figure 4) showed a high performance. So, there is no doubt about the computational ability of the ANN to estimate the ET0 where only few climatic variables are available.

Entrées	R²	E	RMSE (mm/j)	MARE (%)
T+Rh	0.94	0.93	0.58	5.62
T+Wv	0.96	0.95	0.49	6.81
T+I	0.95	0.94	0.55	6.58
Wv+I	0.41	0.42	1.44	19.47
T +Wv+I	0.97	0.96	0.43	4.81
Rh+ Wv+I	0.97	0.94	0.53	6.72
T+Rh+Wv	0.99	0.98	0.26	3.02
T+Rh+Wv+I	0.99	0.98	0.25	2.71

Where T is the air temperature, Rh is the relative humidity, I is the sunshine duration, Wv the wind speed, R² is the determination coefficient, E is the Nash-Sutcliffe efficiency, RMSE is the root mean squared error (mm/day) and MARE is the mean absolute relative error (%).

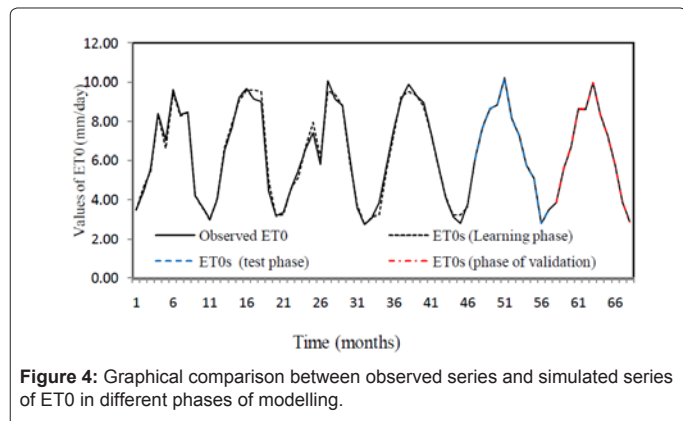**Table 3:** Performance criteria according 7 to number of inputs.

Figure 4: Graphical comparison between observed series and simulated series of ET0 in different phases of modelling.

Choosing of the appropriate architecture

One can choose the number of hidden layers and number of neurons in each layer; it should be considered that, as the number of neurons increase, the network is capable of identifying complex phenomena. But it should be taken into consideration that too many neurons lead to the over fitting.

However, in our case, the phenomenon studied is not very complex. That is why it turned out that using a single neuron in the hidden layer for a model with 4 inputs, gives in test phase, 0.99, 1, 0.13 mm /day and 1.97% respectably for R², E, RMSE and MARE.

The figure below indicates that the observed series and the simulated series have almost the same speed, although they diverge up and down several times especially in learning phase. Nevertheless, the two series are close to each other in test phase and in validation phase (Figure 5). Despite these good performance criteria, it is useful to seek a more appropriate architecture. Each time we add a neuron to hidden layer, the values of R² and MSE changed. But addition of neurons to hidden layer does not necessarily improve the model performance (Figure 6).

The performance criteria indicate that there is perfect agreement between the observed series and the simulated series of ET0. For the three modelling phases, R² and E values reached 1. MSE, RMSE and MARE are decreased to 0. The R² values found in different phases are better than R²=0.98 found by Archana and Shrivastava [28] in monthly time step. The graphical comparison between the observed series and the simulated series shows a very high performance (Figure 7).

So, the model selected in monthly time step was characterized by architecture containing a single hidden layer of 8 neurons and an output layer of single neuron.

The parameters retained by the model developed are presented in Table 4.

Decade time step

The statistical analysis of data shows a close relationship between the observed and the simulated series; the determination coefficient R² reached 97%. Generally, all parameters that were used, contributed significantly in estimating ET0. Results showed at a confidence level of 0.05 the marginal contribution of each variable is significant. But these results are not better than those obtained by using ANN with a single neuron in the hidden layer (Table 5).

We applied the trial-and-error technique by increasing the number of neurons in the hidden layer until we found a smaller value of error and the higher value of R² (Figure 8). In fact, at architecture of 15 neurons in hidden layer, all performance criteria are best.

The comparison of the performance criteria obtained during the different phases of the neural network modelling with those obtained by MLR for the various sets of the data, shows the importance of the neural network modelling. The MARE (%), i.e., the percentage of recorded errors between real and simulated values of ET0, indicates a higher performance of the neural networks over MLR.

It should be mentioned that the number of inputs influenced the model performance. Indeed, despite sunshine duration being not an important parameter, removing it has decreased the value of MARE from 2.62% to 3.79% (learning phase).

The graphical comparison between the observed ET0 series and the simulated ET0 series shows a merger between them especially in test phase and in validation phase (Figure 9).

Daily time step

Using MLR, the statistical analysis of data shows a close relationship between the observed and the simulated series; the determination coefficient R^2 and Nash efficiency reached 0.97.

The MSE, RMSE and MARE values are respectably; 0.20 (mm/day)^2, 0.45 (mm/day) and 5.19%. Generally, all parameters used contributed significantly in estimating ET0. Results showed that at a confidence level of 0.05, the marginal contribution of each variable is significant. But these results are not better than those obtained by using ANN with 4 neurons in the hidden layer where R^2 and E are 0.99 and

Figure 5: Graphical comparison between observed ET0 and simulated ET0 in different phases (single neuron).
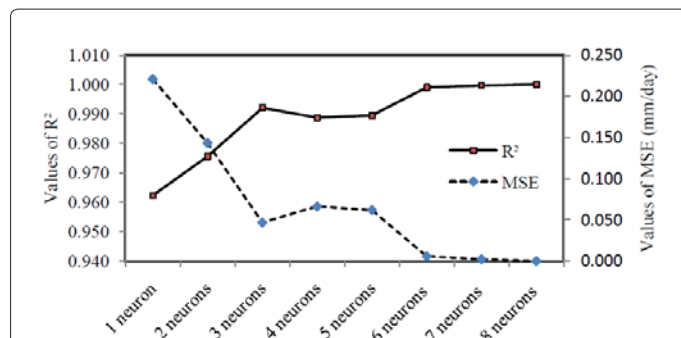
Figure 6: Evolution of R^2 and MSE in term of number of neurons in the hidden layer (learning phase).

Figure 7: Graphical comparison between observed series and simulated series of ET0 in different phases of modelling.

N° of Neurons	Weight matrix of hidden layer					output	
	T	Rh	Ws	I	biais	weights	bias
1	-1.23	-2.46	-0.02	1.62	-5.93	-0,89	
2	-1.42	7.01	-4.17	1.88	3.15	0.18	
3	6.91	10.58	-13.25	-7.26	-1.50	0.56	-7.9
4	1.77	1.41	1.36	-4.37	-1.61	6.06	
5	-1.34	-0.90	-0.21	-1.98	6.88	16.50	
6	5.67	3.73	1.84	-7.42	-2.15	-4.65	
7	6.95	-0.37	7.80	-4.09	-17.27	0.61	
8	-0.48	-1.07	2.16	-1.69	12.60	-9.56	

Where T is the average of temperature, Rh is relative humidity Ws is the average of wind speed and I is the unshine duration

Table 4: Parameters retained for the monthly time step model.

Performance criteria	ANN Learning	Test	MLR Learning	Test
R^2	0.95	0.97	0.94	0.96
E	0.94	0.97	0.94	0.96
MSE (mm/d)^2	0.33	0.16	0.34	0.23
RMSE (mm/d)	0.57	0.40	0.58	0.48
MARE (%)	7.03	5.81	7.57	6.97

Where R^2 is the determination coefficient, E is the Nash-Sutcliffe efficiency, MSE is the mean-square error (mm/day)^2, RMSE is the root mean squared error (mm/day) and MARE is the mean absolute relative error (%)

Table 5: Comparison between performance criteria of the ANN model with 4 inputs and single neuron and the MLR model.
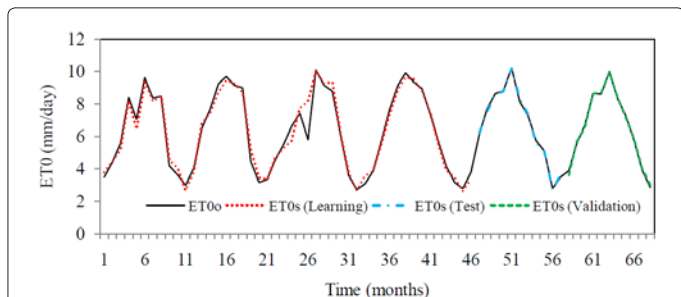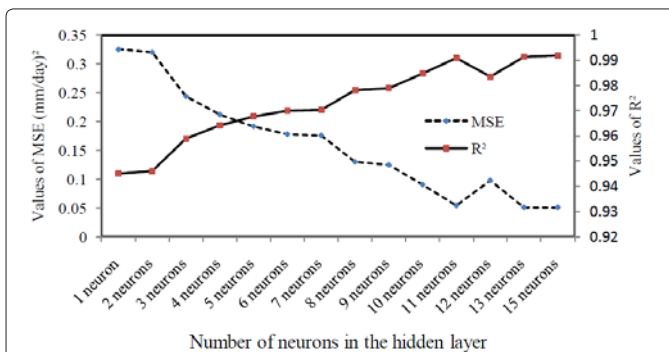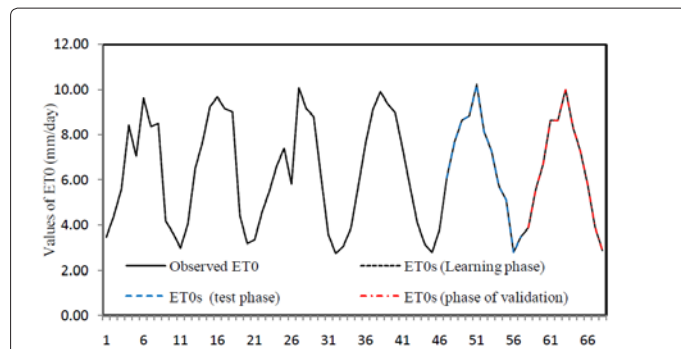

Figure 8: Evolution of R^2 and MSE in term of number of neurons in the hidden layer (learning phase).

the MSE, RMSE and MARE values are respectably 0.07 (mm/day)^2, 0.27 (mm/day) and 2.22% in test phase.

These results are very satisfactory and we can stop addition of neurons in hidden layers at this simple architecture. In this context, Tabari et al. [29] have noted that among several tested architectures, a single hidden layer with 5 neurons was the best architecture. So we can say that an ANN with only a hidden layer is enough to represent the nonlinear relationship between the climatic elements and the corresponding ET0. But, it should be noted that the advantage of the neural method lies in the possibility of having improvements in the performance criteria by modifying the network architecture. Koleyni [30] believed that the neural networks' performance is very often related to its architecture. This performance is usually determined

simply through experiments because of the lack of theory which leads to determine the adequate architecture easily. The choice of the neural network capacity primordially reflects its ability of learning and generalizing. If the network model is proportionally small, it will be unable to obtain the desired function. However, if it is too complex, it will be unable to generalize the model.

Extensive test experiments were conducted in order to select the optimal network architecture (Figure 10). Consequently, these tests led to a network of 2 hidden layers, each of 8 neurons.

Also, the neural network requires setting up the learning rate and the number of iterations. Hence, after different combinations, we chose a learning rate = 0.2 and a number of iterations = 1000.

The obtained values of R^2 and RMSE are better than those obtained by Diamantopoulou et al. [31], which ranged from 0.939 to 0.956 and 0.549 to 0.609 mm/day respectively. They are better than the highest R^2 (0.907) and lowest RMSE (0.356) found by Traore et al. [32].

The error (MSE) at the testing phase is always lower than this at learning phase; this shows the absence of over fitting.

In order to evaluate the correlation between the observed values of the ET0 and the simulated values, we plotted them in a graph as shown in Figure 11. The result is: scattered points statistically distributed around the line y=x. This shows a very good resemblance that explains a high correlation coefficient at test phase. We mentioned that most of the predicted values with ANN are lying near the y=x line. Further, this study concludes that a combination of: mean air temperature,

Figure 11: Scatter plots of observed and computed ET0 (test phase).
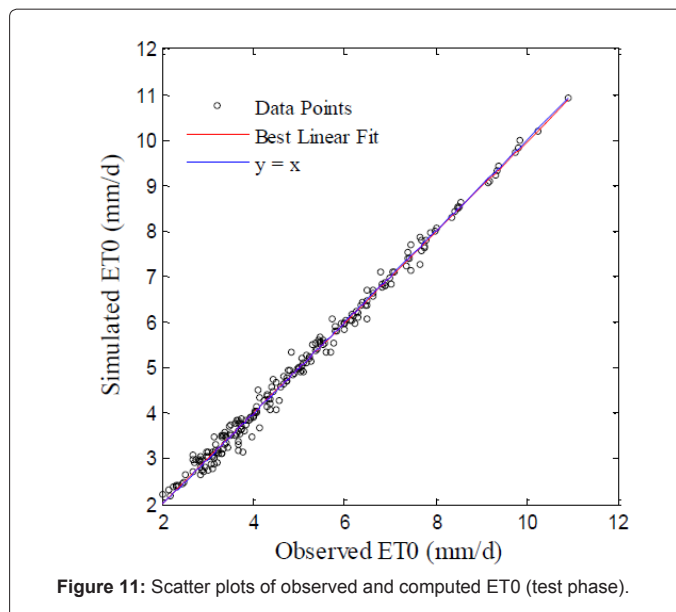
wind speed, sunshine hour and mean relative humidity provides better performance in predicting the reference evapotranspiration.

At the end, we must confess that the performance of the models varies according to the number of inputs as well as of the predicted time step. Hence, Wang et al. [33] have noted that wind velocity and relative humidity were found to improve the temperature-based back propagation accuracy when incorporated into the network input sets.

Indeed, this performance will be even better when we are interested in modelling a more extensive time step. With a simple architecture, we can obtain a very strong correlation, i.e., R^2 close to 1. Yet, this performance decreases when the number of the inputs is reduced.

Furthermore, other factors may intervene and may affect model performance; these include the standard deviation between the values of input parameters. It is found that standard deviation values of inputs reduce from a daily time step to monthly time step. The opposite occur for the correlations between input parameters and output. Indeed, when the standard deviation of the input parameter values is reduced and the correlation between them and the output is high, the model performance will be good without a complex architecture.

It should be mentioned that, the number of neurons in the input layer depends on the number of climatic variables used in estimating ET0. The individual node in the input layer corresponds to respective variables. Thus, the number of neurons in the input layer varies according to the climatic data requirement of the model.

Conclusions

The present study discusses the application and usefulness of the artificial neural network modelling approach in predicting the evapotranspiration reference. The results are quite encouraging and suggest the usefulness of a neural network based modelling technique for an accurate prediction of the evapotranspiration as an alternative to the multiple linear regression approach, because the advantage of the neural method lies in the possibility of having improvements in the performance criteria by modifying the network architecture.
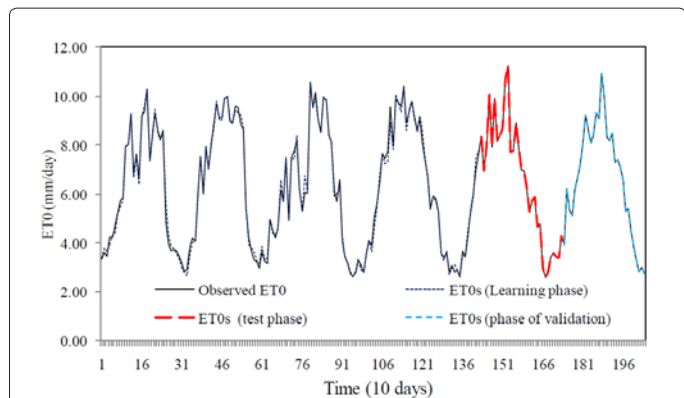

Figure 9: Graphical comparison between observed series and simulated series of ET0 in different phases of modelling.
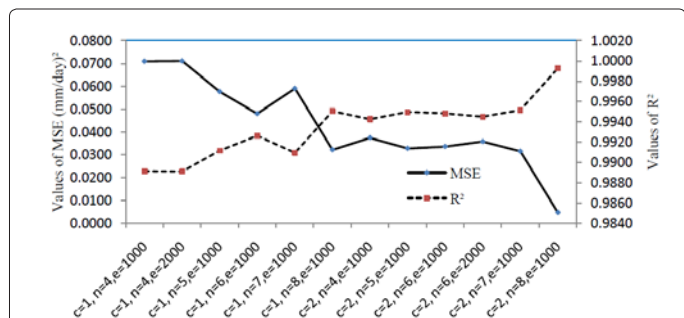

Figure 10: Evolution of MSE and R^2 in term of network architecture and number of epochs. c number of hidden layers, n number of neurons per each hidden layer, e number of epochs.

ANN modelling now has been added a new dimension in computational science. The application of ANN models in ET0 estimation is now being widely discussed.

In order to build nonlinear models in different time steps we should multiply the elements in hidden layer. Thus, we have obtained the satisfactory models with 8, 15 neurons in hidden layer respectively at the monthly and decade time steps. For daily time step it required two hidden layers.

This work provides users different options according data availability and the required accuracy. Furthermore, results may be used for dam dimensions, small scale irrigation or large scale irrigation and also for responding to important farmer questions; how much water will my crop need before I can turn the water off? How much water to apply?

References

1. Hadef R, Hadef A (2001) Le déficit d'eau en Algérie: une situation alarmante. Desalination 137: 215–218.

2. World Bank (2007) People's Democratic Republic of Algeria. Public expenditure review assuring high quality public investment. Document of the World Bank (in two volumes) volume I: Main Text. Report no 36270-dz.

3. Chennaf H, Aidaoui A, Bouzerzour H, Saci A (2006) Yield response of durum wheat triticum dirum Desf) Cultivar Waha to deficit irrigation under semi arid growth conditions. Asian J Plant Sci 5: 854-860.

4. Liu B, Zhao W, Chang X, Li S, Zhang Z, et al. (2010) Water requirements and stability of oasis ecosystem in arid region, China. Environ Earth Sci 59: 1235–1244.

5. Er-Raki S, Chehbouni A, Ezzahar J, Khabba S, Lakhal EK, et al. (2011) Derived Crop coefficients for winter wheat using different reference evapotranspiration estimates methods. J Agr Sci Tech 13: 209-221.

6. Hajare HV, Raman NS and Dharkar ERJ (2008) New technique for evaluation of crop water requirement. WSEAS Trans Environ Dev 4: 436-446.

7. Kumar M, Raghuwanshi NS, Singh R (2011) Artificial neural networks approach in evapotranspiration modeling: a review. Irrig Sci 29: 11-25.

8. Parasuraman K, Elshorbagy A, Carey SK (2007) Modeling the dynamics of the evapotranspiration process using genetic programming. Hydrol Sci J 52: 563-578.

9. López-Moreno JI, Hess TM, White SM (2009) Estimation of reference evapotranspiration in a mountainous mediterranean site using the Penman-Monteith equation with limited meteorological Data. Pirineos 164.

10. Alkaeed O, Flores C, Jinno K, Tsutsumi A (2006) Comparison of several reference Evapotranspiration methods for Itoshima Peninsula Area, Fukuoka, Japan. Memoirs of the Faculty of Engineering, Kyushu University 66: 1-14.

11. Wang YM, Traore S, Kerh T and Leu JM (2011) Modeling reference evapotranspiration using feed forward backpropagation algorithm in arid regions of Africa. Irrig drainage 60: 404-417.

12. Saidati B, Samuel P (2006) Evapotranspiration de référence dans la région aride de Tafilalet au sud-est du Maroc. AJEAM-RAGEE 11: 1-16.

13. Hazrat MA, Lee TS (2009) Potential Evapotranspiration Model for Muda irrigation Project, Malaysia. Water Resour Manage 23:57-69.

14. Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration - Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Food and Agriculture Organization: Rome.

15. Laaboudi A, Mouhouche B, Draoui B (2011) Neural network approach to reference evapotranspiration modeling from limited climatic data in arid regions. Int J Biometeorol.

16. Xing ZS, Chow TL, Meng F, Rees HW, Monteith JO, et al. (2008) Testing Reference evapotranspiration estimation methods using evaporation pan and modeling in maritime region of Canada. J Irrig Drainage Eng 134: 417-424.

17. Shirgure PS, Rajput GS (2011) Evaporation modeling with neural networks-A Research review Int J Res Rev Soft & Intelli Comp (IJRRSIC) 1.

18. Abedi-Koupai J, Amiri MJ, Eslamian SS (2009) Comparison of Artificial Neural Network and Physically Based Models for Estimating of Reference Evapotranspiration in Greenhouse. Aust J Basic Appl Sci 3: 2528-2535.

19. Johannet A,Vayssade B, and Bertin D (2007) Neural networks: from black box towards transparent box Application to evapotranspiration modeling. Proc world Acad Sci, Eng Technol vlume 24.

20. Jain SK, Nayak PC, Sudheer KP (2008) Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. Hydrol Process 22: 2225-2234.

21. Liu Z, Peng C, Xiang W, Deng X, Tian D, et al. (2012) Simulations of runoff and evapotranspiration in Chinese fir plantation ecosystems using artificial neural networks. Ecological Modeling 226: 71-76.

22. Laio F, Porporato A, Revelli R and. Ridolfi L (2003) A comparison of nonlinear flood forecasting methods. Water Resour Res 39: 1129.

23. Behzadi MRB, Aslaminejad AA (2010) A Comparison of Neural Network and Nonlinear Regression Predictions of Sheep Growth. J Ani Vet Adv 9: 2128-2131.

24. Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration - Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Food and Agriculture Organization: Rome.

25. Parizeau M (2004) Réseaux de neurones. GIF-21140 et GIF-64326 124.

26. Université Laval, Rakotomalala R (2011) Colinéarité et section des variables régression linéaire multiple. Université Lyon 2. Laboratoire ERIC.

27. Krause P, Boyle DP, Base F (2005) Comparison of different efficiency criteria for hydrological model assessment. Adv Geosci 5: 89-97.

28. Archana C, Shrivastava RK (2010) Reference Crop Evapotranspirationestimation using Artificial Neural Networks. Int J Eng Sci Technol 2: 4205-4212.

29. Tabari H, Marofi S and Sabziparvar AA (2009) Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. Irrig Sci 28: 399-406.

30. Koleyni K (2010) Using Artificial neural Networks for Income Convergence. Global J Busi Res 3: 141-152.

31. Diamantopoulou MJ, Georgiou PE, Papamichail DM (2011) Performance evaluation of artificial neural networks in estimating reference evapotranspiration with minimal meteorological data**.** Global NEST J 13: 18-27.

32. Traore S, Wang YM, Kerh T (2008) Modeling Reference evapotranspiration by generalized regression neural network in semiarid zone of Africa. WSEAS trans info Sci App 5.

33. Wang YM, Traore S, Kerh T and Leu JM (2011) Modeling reference evapotranspiration using feed forward backpropagation algorithm in arid regions of Africa. Irrig drainage 60: 404-417.