Italian Journal of Agrometeorology
Rivista Italiana di Agrometeorologia

Firenze University Press
www.fupress.com/ijam

OPEN ACCESS

Citation: A. Laaboudi, A. Slama (2020) Using Neuro-fuzzy and linear models to estimate reference Evapotranspiration in South region of Algeria (A comparative study). *Italian Journal of Agrometeorology* (2): 55-64. doi: 10.13128/ijam-971

Received: June 18, 2020

Accepted: July 26, 2020

Published: January 25, 2021

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Using Neuro-fuzzy and linear models to estimate reference Evapotranspiration in South region of Algeria (A comparative study)

Abdelkader Laaboudi[1,*], Abdeldjalil Slama[2]

[1] *National Institute of research in agronomy of Algeria. Experimental station of Adrar, Algeria*
[2] *Laboratory of Mathematics, Modeling and Applications (LAMMA), University of Adrar, Algeria*
*Corresponding author. E-mail: Laaboudiaek@yahoo.fr

Abstract. In order to estimate daily reference evapotranspiration (ETo) in arid region of Algeria, Adaptive Neuro-Fuzzy Inference System (ANFIS) and regression methods as Robust Regression (RR), Bayesian Regression (BR) and Multiple Linear Regression (MLR) techniques were used to develop models based on four explanatory climatic factors: temperature, relative humidity, wind speed and sunshine duration. These factors have been used as inputs, and ETo values computed by the Penman-Monteith formula have been used as outputs. Determination coefficient (R^2), root mean square error (RMSE), Mean absolute error (MAE), mean absolute relative error (MARE) and Nash-Sutcliffe efficiency coefficient (NSE) were used to evaluate the performance of models developed with different input configurations. We concluded that RR, BR and MLR models were able to successfully estimate ETo, but ANFIS technique seems to be more powerful. Thus, the obtained results by the best ANFIS model, during the test phase are: 0.98, 0.27 (mm/day)^2, 0.36 (mm/day) and 5.52 % respectively for R, MAE, RMSE and MARE.

Keywords. Reference evapotranspiration, arid regions, Adaptive Neuro-Fuzzy Inference System, robust regression, Bayesian regression, Penman-Monteith formula.

1. INTRODUCTION

Evapotranspiration (ET) is an important factor in climatological, hydrological and agricultural management. So its estimate is of vital importance for irrigation scheduling, water resources planning and management, and for drought forecasting (e.g. Abyaneh et al., 2010, Meng et al., 2018, Lee et al. 2012). Thus, the evapotranspiration is used to compute many Drought Indices as Reconnaissance Drought Index (RDI) (Tsakiris et al., 2007), the water surplus variability index (WSVI) (Gocic and Trajkovic, 2014) and standardized precipitation evapotranspiration index (SPEI) (Vicente-Serrano et al., 2010). To estimate reference evapotranspiration (ETo), the Penman-Monteith (PM) equation has been recommended as the standardized equation, but it has high requirements of climatic data (Peng et al. 2017, Wable et al. 2019). However, in developing countries, application of this equation for ETo estimation has certain limitations due to unavailability of specific data requirements (Naidu and Majhi, 2019).

According to Tabari et al. (2012), Practitioners and researchers need to be provided with guidance on the choice of the most appropriate ETo method to be adopted when weather data are insufficient to apply the Penman-Monteith method. In this context, some researchers was evaluated the reliability of simplified pan-based approaches for estimating ETo directly that do not require the data of meteorological parameters (Trajkovic et al., 2010). Others have used the modeling approaches (Keshtegar et al., 2018). In this direction we have decided to use modeling techniques to estimate ETo based on a daily time step. Thus, Adaptive neuro-fuzzy inference system (ANFIS), robust regression (RR), Bayesian regression (BR) and multiple linear regression (MLR) models were developed and compared to each other.

Multiple Linear Regression method is one of the most widely known modeling techniques. It was used for reference evapotranspiration modeling either alone (Yirga, 2019) or for a comparison (Khoshravesh et al., 2015; 2016; Ozgur et al., 2017) and to examine the relationship between weather parameters and Carbone monoxide (CO) concentration (Ve and Jo, 2016). However, this method is extremely sensitive to deviations from the model assumptions as a normal distribution assumed for the error terms (Stahel, 1997). Consequently, robust regression estimators can be a powerful tool for outlier detection in complicated data sets. For this reason, robust regression model can be the best alternative for multiple linear regression model (Marona et al., 2006).

Bayesian linear regression is an extension of linear regression for modeling and predicting some complex phenomena. It has numerous advantages over classical methods. One of the main advantage of Bayesian predictions over maximum likelihood methods of estimation is an overall increase in accuracy with high levels of reliability on a fraction of the test sample (Braga et al., 2005).

Adaptive Neuro-Fuzzy Inference System (ANFIS) technique which is introduced first by Jange (1993) is a multilayer feed-forward network. It uses neural network training algorithms and fuzzy logic to create an input-output correlation for fuzzy decision rules that perform well on any given task. According to Karimaldini et al. (2012), while neural networks are good at recognizing patterns, they are not good at explaining how they reach to decisions because this technique is, in fact, a black-box for its user. Fuzzy logic systems are good at explaining their decisions, but they cannot learn and adjust themselves to a new environment. These limitations have been solved with ANFIS technique.

The present study aims to: (1) Investigate the potential of using robust regression, Bayesian regression and ANFIS models to estimate reference evapotranspiration, (2) choose the best approach for users in arid region conditions and (3) to adapt the best models to the climatic conditions in south region of Algeria.

2. MATERIALS AND METHODS

Our study was carried out in the region of Adrar, located in the south-west of Algeria. Latitude: 27°49'N and Longitude: 00°18'E (Fig. 1).

2.1 Climate characteristics

Adrar region is characterized by its extreme meteorological parameters. Its climate is dry throughout the year and is characterized by the extended thermal amplitudes during the year, the month and even the day. The absolute

Fig. 1- Sketch of the investigation area.

maximum temperature reaches 49.5°C in summer (July and August), while, ice and frost are the rare phenomena.

2.2 Description of data and availability

For estimating reference evapotranspiration, the ANFIS, RR BR and MLR models were trained. The entire database (the overall size is, n = 1825) was splitting into two datasets, 80% were used in training phase and 20 % remaining were used in test phase.

In the present investigation, daily data (temperature, relative humidity, wind speed and sunshine duration) consist of daily series values recorded throughout the period of 1825 days (From January 2013 to December 2017). The registration of these meteorological statements was performed by the meteorological station located within the experimental site. Using these observed climatic data, daily values of ETo were computed initially using the Penman-Monteith (Eq. 1). These computed ETo values were used to train the ANFIS models.

2.3 Estimation of reference evapotranspiration

The Penman-Montheith equation used for estimating reference evapotranspiration is written as bellow (Allen et al., 1998):

$$ETo = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T+273}u_2(e_s - e_a)}{\Delta + \gamma(1+0.34u_2)} \quad (1)$$

Where ETo is the reference evapotranspiration (mm day^{-1}), Rn 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}), es is the saturation vapor pressure (kPa), ea is the actual vapor pressure (kPa), (es−ea) is the saturation vapor pressure deficit (kPa), Δ is the slope vapor pressure curve (kPa $°C^{-1}$), γ is the psychometric constant (kPa $°C^{-1}$).

The parameters air temperature, relative humidity, wind speed and sunshine duration: were taken directly from the meteorological station and were used to estimate other parameters. According to Doorenbos and Pruitt (1977), net radiation and the saturation vapor pressure deficit can be estimated by air temperature and sunshine duration. Net radiation is the difference between the net short wave radiation (R_{ns}), and the outgoing net long wave radiation (R_{nl}).

$$R_n = R_{ns} - R_{nl} \quad (2)$$

if

$$R_{ns} = (1-\alpha)(0.25+0.50\frac{n}{N})Ra \quad (3)$$

and

$$R_{nl} = f(t).f(ea).f(n/N) \quad (4)$$

$$R_n = (1-\alpha)(0.25+0.50\frac{n}{N})Ra - f(t).f(ea).f(n/N) \quad (5)$$

2.4 Statistics of meteorological variables

The statistical features of meteorological variables and evapotranspiration in training and test subsets (Tab. 1) indicate that the data structures of these subsets have the same characteristics.

It can be noted that the variability range of meteorological parameters in the study area was very large. For instance, the daily values of temperature ranged between 7.10 °C and 42 °C: relative humidity ranged between 17.50 % and 95 %, wind speed ranged between 0.00 and 4.27 ms^{-1} and sunshine duration ranged between 0.00 to 12.35 hours per day. Hence, any model developed on this data set should have a wide application in all regions that have meteorological parameters in the range the of the study area. The standard deviation values indicated that the variability of meteorological variable values is very important.

The correlations of all input variables are presented in (Tab. 2). This table shows that the linear correlations between ETo and two independent variables: temperature, and relative humidity are very high. Their values are 0.85 and -0.86 respectively Hence, any model that uses these explanatory climatic parameters should be able to estimate the ETo satisfactorily.

Tab. 1. Statistics of meteorological variables in training, test and validation data subsets. T temperature, RH Relative humidity, U_2 Wind speed and n sunshine duration.

Phases	Statistic parameters	T (°C)	RH (%)	U_2 (ms^{-1})	n ($h.day^{-1}$)
Training	Min	7.10	17.50	0.00	0.00
	Max	42.00	95.00	4.27	12.35
	Mean	25.92	41.60	1.62	9.14
	Std	9.02	14.03	0.72	2.74
Testing	Min	10.75	23.50	0.28	0.00
	Max	42.25	90.00	4.18	12.35
	Mean	26.34	46.71	1.48	8.81
	Std	8.97	16.38	0.61	2.26

Tab. 2. Correlation matrix between input and output variables.

	Temperature	Humidity	Wind speed	Sunshine duration	ET0
Temperature	1.00				
Humidity	-0.81	1.00			
Wind speed	0.07	0.03	1.00		
Sunshine duration	0.28	-0.39	0.05	1.00	
ETo	0.86	-0.85	0.31	0.57	1.00

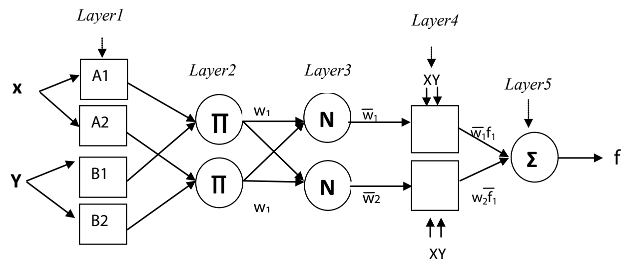

Fig. 2. ANFIS architecture with 5 layers. and are the membership functions (MFs) for inputs x and y respectively.

The temperature and humidity are also highly correlated. Therefore, a combination of these two factors may provide a good estimate of reference evapotranspiration. It should be noted that all these correlations between variables are linear type but the ETo is universally considered a nonlinear process dependent on interacting meteorological variables (Laaboudi et al., 2012; Fang et al., 2018).

2.5 Adaptive neuro-fuzzy inference system (ANFIS)

An adaptive network, as its name implies, is a network structure consisting of nodes and directional links through which the nodes are connected. Moreover, parts or all of the nodes are adaptive, which means each output of these nodes depends on the parameters pertaining to this node and the learning rule specifies how these parameters should be changed to minimize a prescribed error measure (Jange, 1993). In ANFIS, fuzzy rule bases are combined with neural networks to train the system using experimental data and obtain appropriate membership functions for process prediction and control. The inference system has two input variables x and y as each variable has two fuzzy subsets. A typical rule set with two fuzzy if then rule set for a first order Sugeno fuzzy model can be defined as Eq. 6 and 7:

Rule 1: If x is A_1 and y is B_1 Then $f_1 = p_1x + q_1y + r_1$ (6)

Rule 2: If x is A_2 and y is B_2 Then $f_2 = p_2x + q_2y + r_2$ (7)

Where A_1, A_2 and B_1, B_2 are the membership functions (MFs) for inputs x and y respectively, p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output function. f_1 and f_2 are constant output respectively for rule 1 and rule 2 in ANFIS for the first-order Sugeno inference system.

The general architecture of ANFIS consists of five layers, namely, a fuzzy layer, a product layer, a normalized layer, a defuzzy layer and a total output layer.

The membership function (MF) of each input was tuned using the hybrid method consisting of back propagation for the parameters associated with the input membership function and the least square estimation for the parameters associated with the output membership functions. The architecture of the ANFIS is shown in Fig. 2.

2.6 Multiple Linear Regression (MLR)

Multiple linear regression (MLR) is a statistical approach to modeling the linear relationship between a response (dependent) variable and one or more explanatory (independent) variables.

Given an independent and identically distributed (i.i.d) observations (x_i,y_i), i = 1, . . . , n, in order to understand how the response y_i's are related to the covariates x_i's, we traditionally assume the assume the following linear regression model:

$$y_i = x_i^T\theta + \varepsilon_i, \tag{8}$$

Where θ is an unknown p × 1 vector, and the ε_i's are i.i.d and independent of x_i with $E(\varepsilon_i|x_i)=0$.

The most commonly used estimate for θ is the ordinary least-square (OLS) estimate that minimizes the sum of squared residuals

$$\sum_{i=1}^{n}(y_i - x_i^T\theta)^2. \tag{9}$$

2.7 Robust regression (RR)

It is well known that the OLS estimate is extremely sensitive to the outliers. A single outlier can have large effect on the OLS estimate (Yu and Yao, 2017). Thus robust regression analysis provides an alternative to a least squares regression model when fundamental assumptions are unfulfilled by the nature of the data, such as if the distribution of errors is asymmetric or prone to outliers.

The Statistics Toolbox function "robustfit" is useful in these cases. The function implements a robust fitting method that is less sensitive than OLS to large changes in small parts of the data (Matlab Statistics Toolbox 2010).

Robust regression works by assigning a weight to each data point. Weighting is done automatically and iteratively using a process called *iteratively reweighted least squares*. In the first iteration, each point is assigned equal weight and model coefficients are estimated using ordinary least squares. At subsequent iterations, weights are recomputed so that points farther from model predictions in the previous iteration are given lower weight. Model coefficients are then recomputed using weighted least squares. The process continues until the values of the coefficient estimates converge within a specified tolerance (Matlab Statistics Toolbox 2010). For more details on robust regression, see Fox and Weisberg (2002) and Yu and Yao (2017).

2.8 Bayesian Regression (BR)

Suppose that we are interested in estimating a parameter θ from the data $y=(y_1,y_2,\dots,y_n)$ using a statistical model described by a density $l(y|\theta)$, called the likelihood function or likelihood. Bayesian philosophy states that θ can be considered as random variable with probability distribution $\pi(\theta)$, which is known as the prior distribution, or just the prior. The prior distribution expresses our beliefs about the parameter before examining the data. Given the observed data y, update of beliefs about θ by combining information from the prior distribution and the data by the use of Bayes' theorem, and so the calculation of the posterior distribution, $\pi(\theta|y)$. For the prior distribution we have considered a Jeffrey non-informative prior based on the Fisher information (see Bernard et al. 2000 and Ghosh et al. 2007).

Consider a standard linear regression problem given in (12) and consider ε_i, $i=1,\dots,n$ are independent and identically normally distributed random variables $N(0,\sigma^2)$, $\sigma>0$.

The likelihood function $l(y|\theta,\sigma)$ is given by

$$l(y|\theta,\sigma)\propto\sigma^{-n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-x_i^T\theta)^2\right). \quad (10)$$

With Jeffrey's non-informative prior for (θ,σ) given by

$$\pi(\theta,\sigma)\propto\frac{1}{\sigma}. \quad (11)$$

The posterior distribution of θ, obtained by combination of (10) and (11) is given by:

$$\pi((\theta,\sigma)|y)\propto\sigma^{-n-1}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-x_i^T\theta)^2\right). \quad (12)$$

The posterior distribution (12) is used to estimate the vector parameter (θ,σ).

The performances of linear regression and ANFIS models were evaluated to compare their predictive accuracies based on the following statistical criteria:

- The coefficient of determination (R^2) is the square of correlation coefficient (r) between the observed and estimated data values of the dependent variable. The coefficient r is expressed as:

$$r=\frac{\sum_{i=1}^{n}(Y_{obs}-\overline{Y}_{obs})(Y_{sim}-\overline{Y}_{sim})}{\sqrt{\sum_{i=1}^{n}(Y_{obs}-\overline{Y}_{obs})^2\sum(Y_{sim}-\overline{Y}_{sim})^2}} \quad (13)$$

- The root mean squared error (RMSE):

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(Y_{obs}-Y_{sim})^2}{n}} \quad (14)$$

- The mean absolute error (MAE):

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_{obs}-y_{sim}| \quad (15)$$

- The mean absolute relative error (MARE):

$$MARE=\frac{1}{n}\sum_{i=1}^{n}\frac{|Y_{obs}-Y_{sim}|}{|Y_{obs}|}\text{ x }100 \quad (16)$$

- The Nash- Sutcliffe efficiency (NSE) coefficient, It is calculated as follows:

$$NSE=1-\frac{\sum_{i=1}^{n}(y_{sim}-y_{obs})^2}{\sum_{i=1}^{n}(y_{sim}-\bar{Y}_{obs})^2} \quad (17)$$

Where,

NSE: Nash-Sutcliffe efficiency; $\overline{Y}_{obs}$: Average of observations; Y_{sim} is a simulated variable, Y_{obs} the observed variable, $\overline{Y}_{sim}$ the average of simulated variable, $\overline{Y}_{obs}$ the average of observed variable, n a number of observations.

The NSE coefficient is the Nash-Sutcliffe efficiency coefficient, proposed by Nash and Sutcliffe (1970), It is used to evaluate the predictive power of the model (Latrech et al., 2019).

The Matlab software was used for the implementation and application of ANFIS approach and regression methods (BR, RR MLR).Two phases were employed in ANFIS modeling: training and testing phases.

3. RESULTS AND DISCUSSIONS

For estimating reference evapotranspiration, the ANFIS, RR, BR and MLR models were trained. The resultsob-

Tab. 3. Comparison of the models in terms of R, E, RMSE, MAE and MARE

Modeling methods	Datasets	R^2	NSE	RMSE (mm/day)	MAE (mm/day)	MARE (%)
Robust regression	Dataset1	0.951	0.950	0.6045	0.4584	9.6774
	Dataset2	0.966	0.966	0.4622	0.3744	9.0191
Linear Multiple Regression	Dataset1	0.951	0.949	0.5967	0.4665	9.5345
	Dataset2	0.966	0.965	0.4631	0.3756	8.9746
Bayesian regression	Dataset1	0.951	0.949	0.5967	0.4668	11.7285
	Dataset2	0.966	0.965	0.4617	0.3757	8.970
ANFIS (Tr. P) (T.P)	Dataset1	0.982	0.981	0.3660	0.2506	4.7574
	Dataset2	0.984	0.980	0.3579	0.2709	5.5199

Tr.P: Training phase, T.P: test phase.

tained from evaluating performance of these models: in terms of statistical criteria are given in Tab. 3.

As it can be seen in (Tab. 3), the ANFIS model represented more consistent estimates. The ANFIS model has the smallest MAE, RMSE and the highest R^2 and NSE in the Training and testing phase. In testing phase, the ANFIS model has the smallest RMSE (0.3579), MAE (0.2709) and MARE (5.5199) and the highest R^2 (0.984) and NSE (0.981). The performance of the ANFIS model on the testing dataset showed that the ANFIS model can be used to provide accurate and reliable reference evapotranspiration (ETo) prediction. The models RR, BR and MLR are almost similar one to each other. Moreover, from this table, it is evident that all performance criteria illustrate a reasonably good performance for all models. This is meant that all models could provide a good estimation of reference evapotranspiration. Nevertheless, if we deal with each method separately, we find slight differences between them: according to the size of the samples, its dispersion and the number of inputs.

3.1 Adaptive Neuro-Fuzzy Inference System

For the ANFIS approach the model developed in this study provided consistent RMSE, MAE and MARE values during training and testing phases when compared to other models. To find out the best model in among the all ANFIS models 50 and 100 epochs, 2 and 3 number of membership functions were tried for each model (Tab. 4).

While in training phase there were slight differences in terms of RMSE according to number of membership functions (MF), in testing phase, all models with the same number of MF are similar one to each other in terms of RMSE. Thus, their values are 0.358 mm/day for 2 MF and 0.475 mm/day for 3 MF respectively.

Increasing number of MF more than 2 MF, enhances the model performance in training phase but contrary in testing phase. Thus, it can be seen (Tab.4) that the model 4 (2, gebellmf, linear) and model 5 (2, gauss2mf, linear) perform well, as they have high coefficients of determination R^2 (0.964 and 0.984) and the lowest values of RMSE =0.366 and 0.358 in training and testing phase respectively. These models perform better than model 7 (3, gauss2mf, linear) and model 8 (3, gauss2mf, linear) in testing phase. The given coefficient of determination values are higher than values of R^2 = 0.67 obtained by Areerachakul (2012) and R^2 0.943 obtained by Pour-Ali Baba et al. (2015) but they are much closer to R^2 = 0,986 obtained by Kumar et al. (2012).

Regarding RMSE values obtained by this study, they were higher than RMSE = 0.265 mmday^{-1} obtained by Shamshirband et al. (2016) but they were lower than RMSE = 0.753 and 0.821 mmday^{-1} obtained by Patil et al. (2017) in training and test periods respectively.

3.2 Linear regression models

To study the performance of linear regression models: RR, BR and MLR models were evaluated together and compared one to each other.

Tab. 4. R and RMSE (mm.day^{-1}) values of the ANFIS models in training and testing phases.

Model N°	MF	number of MF	Training phase		Testing phase	
			R^2	RMSE	R^2	RMSE
Model 1	Trimf	2	0.953	0.414	0.984	0.358
Model 2	Trapmf	2	0.964	0.371	0.984	0.358
Model 3	Psigmf	2	0.964	0.367	0.984	0.358
Model 4	Gbellmf	2	0.964	0.366	0.984	0.358
Model 5	gauss2mf	2	0.964	0.366	0.984	0.358
Model 6	Gbellmf	3	0.972	0.323	0.974	0.475
Model 7	Trapmf	3	0.968	0.341	0.974	0.475
Model 8	gauss2mf	3	0.972	0.328	0.974	0.475

Although limitations of MLR, this technique is widely used at present, it has been used by Tabari et al. (2012) who have reported that MLR model provided good agreement with the ETo obtained by the PM method. They have got a $R^2 = 0.96$ with the best MLR model which was much closer to $R^2 = 0.966$ obtained by the present study. These values were higher than the best value ($R^2 = 0.82$) obtained by Saylan et al. (2019).

Khoshravesh et al. (2016) who have used the multivariate fractional polynomial (MFP), robust regression and Bayesian regression to estimate the monthly ETo, their results showed that the accuracy of MFP model was greater than the other models. RR and BR models gave the same results in terms of R^2 and RMSE in different locations. The higher value of R^2, which was 0.97, it was closer to $R^2 = 0.9662$ obtained in this study.

3.2.1 Coefficients of regression estimation

The regression coefficients are the least squares estimates of the parameters. Their values indicate how much change in Y occurs for a one-unit change in X when the remaining X's are held constant. These coefficients are the values of β_0, $\beta_{1\ldots}$ β_p. They are illustrated in Tab. 5.

It is clear that there were differences between these coefficients from one method to another. Robust regression has provided much better regression coefficient estimates when outliers are present in the data. Thus, as can be seen from Tab. 6, models 9, 10, and 11, MARE values given by RR were slightly less than those given by MLR and BR methods. Contrary, models 13, 14 and 15, MARE values were slightly high.

This indicates that robust regression models were affected by the sample size. Consequently, with small samples, RR method performed more accurate models than MLR and BR techniques. However, BR models become more accurate than RR and MLR methods in case of larger samples. Similar result was shown by Grzenda (2015).

Contrary, models 13, 14 and 15 performed by LMR were slightly better that those performed by RR. They were very closer to models 13, 14 and 15 performed by BR. Consequently, RR method was more effectiveness for the small sample sizes and BR method was effectiveness with larger sample sizes.

Another parameter that could affect the model accuracies was the number and nature of the inputs. Thus, according to the MARE values (Tab. 7), there were differences between the different methods, Sometimes RR models are better than LMR models (16, 17, 18, 19 and 21). May be these models were affected by outlier effect and RR has overcome this problem.

With non-informative prior, BR method was always much closer to LMR method.

Fig. 3 shows Scatter plots of observed versus simulated values of reference evapotranspiration (ETo) for the ANFIS, RR, BR and MLR models in the testing phase database. This figure has confirmed that the used regression models were closer each other but ANFIS model was closest to ETo. In this context, Ladlani et al. (2014) have proved that ANFIS model was more accurate than MLR model.

Fig. 4 demonstrates the observed values of ETo compared with the estimated values from different approach-

Tab. 5. Regression coefficients according to each linear modeling method.

Inputs	β_i	LMR	RR	BR
Contant	β_0	0.957	0.823	1.028
Temperature	β_1	0.131	0.131	0.130
Relative humidity	β_2	-0.057	-0.056	-0.058
Wind speed	β_3	0.013	0.014	0.013
Sunshine duration	β_4	0.277	0.282	0.275

Tab. 6. Comparison between MLR, RR and BR in terms of MARE according to the sample size (number of observations).

Model N°	Sample size (n)	Robust Regression	Multiple linear Regression	Bayesian Regression
Model 9	30	2.214	2.239	2.220
Model 10	50	4.072	4.544	4.520
Model 11	100	8.798	8.829	8.802
Model 12	150	8.827	8.851	8.822
Model 13	200	8.042	7.996	7.979
Model 14	300	7.108	7.076	7.047
Model 15	375	9.019	8.975	8.970

Tab. 7. Comparison between MLR, RR and BR in terms of MARE according to the number of inputs. T temperature, Hr: Relative humidity, U_2 Wind speed and n sunshine duration.

Model N°	Input combinaisons	RR	LMR	BR
Model 16	T	19.35	19.52	18.34
Model 17	T+n	15.73	16.08	23.35
Model 18	T+ U_2	15.73	16.08	18.03
Model 19	T+Hr	15.74	16.08	15.99
Model 20	T+Hr + U_2	13.91	13.96	13.75
Model 21	T+Hr + n	12.71	13.08	13.11
Model 22	T+ U_2+ n	12.27	12.46	12.48
Model 23	Hr + U_2+ n	15.36	15.19	15.19
Model 24	T+Hr + U_2+n	9.02	8.97	8.97

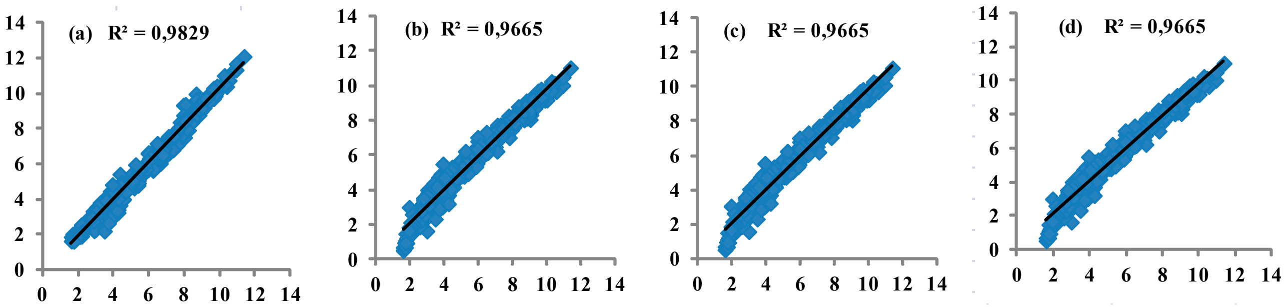

Fig. 3. Scatters plots of observed ET0 (mm/day) and simulated ET0 (mm/day) obtained by different approaches in (test phase database) a: ANFIS; b Bayesian regression; c: Robust regression ; d: Multiple regression. Observed ETo (ET0o), Simulated ET by ANFIS (ET0a), Simulated ET by RR (ET0r), Simulated ET by MLR (ET0m), ET0 Simulated ET by BR (ET0b).

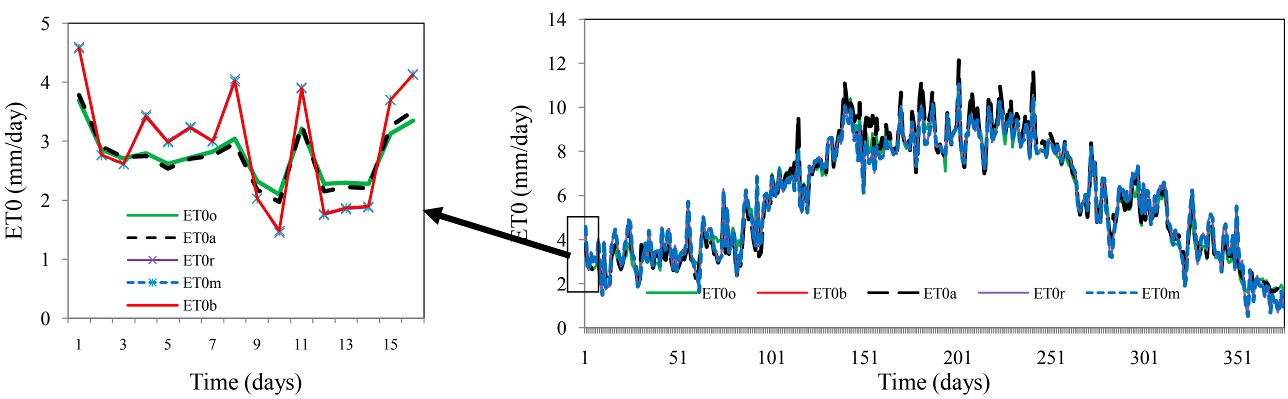

Fig. 4. Graphical comparison of observed ET0 and simulated ET0 obtained by the two approaches in testing phase ; Observed ETo (ET0o), Simulated ET by ANFIS (ET0a), Simulated ET by RR (ET0r), Simulated ET by MLR (ET0m), ET0 Simulated ET by BR (ET0b).

es. The graph has illustrated that predicted values from the ANFIS were closer to the observed values than those obtained from linear regression techniques.

In fact, according to the works of Yaseen et al. (2016), in some cases, the ANFIS models were the best predictors, in other cases, Bayesian models were the best. Thus, the ANFIS spatial model structure was the best predictor of flow and Bayesian temporal model structures performed better than the ANFIS spatial model structure. In our case ANFIS model performances were always better than Bayesian model performances this is probably due to the consideration of the non-informative prior.

If we focus on the 16 first values of different series obtained by the different methods we clearly see that the series of simulated values by ANFIS was very close to the values of the observed series of ETo. The other series of values simulated by the other methods were somewhat distant from the observed series of the ETo but they were very close one to each other and they hided one behind the other.

Overall, the ANFIS models provided the best ETo estimates than statistical models. For practical uses, the ANFIS model with the RMSE values less than 0.3 mm/day and MARE values less than 6 % had good accuracy in ETo modeling and can be used where climatic data are limited. This is especially true in some regions of developing countries where reliable weather data sets required for FAO -56 formula are always not available.

As a whole, the findings of this study revealed that the ANFIS models can be employed successfully in reference evapotranspiration estimation. The main disadvantage of this approach is the complexity of implementation with additional of inputs or membership functions, this task required more time and the results could be very poor.

4. CONCLUSION

In this study we have analyzed and compared adaptive neuro fuzzy inference system (ANFIS) and linear regression models to well estimate reference evapotranspiration when climatic data sets are not enough available. Results showed that the best ANFIS model compared with linear regression models were approximately similar and they were very satisfactory. But in terms of accuracy, ANFIS model seems to be the most reliable. The results were quite encouraging and suggest the usefulness of ANFIS-based modeling techniques for accurate prediction of evapotranspiration as an alternative to statistical approaches. Because the advantage of the ANFIS method lies in the possibility of having improvements in the performance criteria by modifying the membership functions, ANFIS have become powerful tools for modeling in many varied fields of research.

Another advantage may be behind its powerful in modeling is its nonlinear feature, because evapotranspiration process is a nonlinear phenomena. Using linear regression methods in the modeling could result in satisfactory findings, however the ANFIS method has justified its superiority in the power of prediction. Indeed, the designed ANFIS model showed higher performance than other models because the simulated series matches the observed series perfectly.

For the other class of extended linear models (splines, thin-plate, additive,) and the Bayesian regression models with informative prior distribution can be considered in future work.

ACKNOWLEDGEMENT

The authors would like to extend their gratitude to the many people who helped to bring this research project to fruition. Thanks are also to the anonymous reviewers for their careful reading, their useful comments and suggestions which have helped to improve the quality of the paper.

REFERENCES

Abyaneh H. Z., Nia A. M., Varkeshi M. B., Marofi S., Kisi O. 2011. Performance evaluation of ANN and ANFIS models for estimating garlic crop evapotranspiration. Journal of irrigation and drainage engineering, 137(5): 280-286.

Allen R.G., Pereira L.S., Raes D., Smith M. 1998. Crop evapotranspiration (guidelines focomputing irrigation crop water requirement).Irrigation and drainage paper no. 56. Food and Agriculture Organization. Rome.

Areerachakul S. 2012. Comparison of ANFIS and ANN for Estimation of Biochemical Oxygen Demand Parameter in Surface Water. International Journal of Chemical and Biological Engineering, 6: 286-290.

Bernard M. and Smith A. F. 2000. Bayesian Theory. Hoboken, NJ, USA: Wiley.

Braga J., Heuze Y., Chabadel O., Sonan. N. K., Gueramy A. 2005. Non-adult dental age assessment: correspondence analysis and linear regression versus Bayesian predictions. Int J Legal Med. 119: 260-274 DOI 10.1007/s00414-004-0494-8

Demelash N. 2013. Deficit irrigation scheduling for potato production in North Gondar, Ethiopia. African Journal of Agricultural Research, 8(11): 1144-1154.

Doorenbos J., Pruitt W.O. 1977. Guidelines for predicting crop water requirements. Irrigation and drainage.pap. 24. Food and Agriculture Organization, Rome.

Fang W., Huang S., Huang Q., Huang G., Meng E., Luan J. 2018. Reference evapotranspiration forecasting based on local meteorological and global climate information screened by partial mutual information. *Journal of Hydrology*, 561: 764-779.

Fox J., and Weisberg S. 2002. Robust regression. *An R and S-Plus companion to applied regression*, 91.

Ghosh J. K., Delampady M., Samanta T. 2007. An introduction to Bayesian analysis: theory and methods. Springer Science & Business Media.

Gocic M. and Trajkovic S. 2014. Drought characterisation based on water surplus variability index. Water resources management, 28(10), 3179-3191.

Grzenda W. 2015. The advantages of Bayesian methods over Classical methods in the context of credible Intervals. Information systems in management, 4(1): 53-63.

Hunsaker D.J., Fitzgerald G.J, French A.N., Clarke T.R., Ottman M., Jand P.J. 2007. Wheat irrigation management using multispectral crop coefficients: II. Irrigation scheduling performance. grain yield and water use efficiency. Am Soc Agric Biol Eng, 50(6): 2035-2050.

Jange J. 1993. ANFIS: Adaptive-network-based fuzzy inference system." IEEE transactions on systems. man. and cybernetics, 23(3): 665-685.

Karimaldini F., Teang Shui L., Ahmed Mohamed T., Abdollahi M., Khalili N. 2012. Daily evapotranspiration modeling from limited weather data by using neuro-fuzzy computing technique. Journal of irrigation and drainage engineering, 138(1): 21-34.

Keshtegar B., Kisi O., Arab H. G., Zounemat-Kermani M. 2018. Subset modeling basis ANFIS for prediction of the reference evapotranspiration. Water resources management, 32(3): 1101-1116.

Khoshravesh M., Sefidkouhi M.A.G., Valipour M. 2015. Estimation of reference evapotranspiration using multivariate fractional polynomial, Bayesian regression, and robust regression models in three arid environments. Appl Water Sci., 7(4): 1911-1922.

Kumar P., Kumar D., Jaipaul Tiwari A.K. 2012. Evaporation Estimation Using Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference System Techniques. Pakistan Journal of Meteorology, 8 (16): 81-88.

Laaboudi A., Mouhouche B., Draoui B. 2012. Neural network approach to reference evapotranspiration modeling from limited climatic data in arid regions. International journal of biometeorology, 56(5): 831-841.

Ladlani I., Houichi L., Djemil L., Heddam S., Belouz K. 2016. Estimation of Daily Reference Evapotranspiration (ETO) in the North of Algeria Using Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multiple Linear Regression (MLR) Models:A Comparative Study.Arab J Sci Eng, 39(8).

Latrech B., Ghazouani H., Lasram A., M'hamdi B. D., Mansour M., Boujelben A. 2019. Assessment of different methods for simulating actual evapotranspiration in a semi-arid environment. Italian Journal of Agrometeorology, (2): 21-34.

Lee K. H., Cho H. Y. 2012. Simple method for estimating pan coefficients: Conversion of pan evaporation to reference evapotranspiration. Journal of irrigation and drainage engineering, 138(1): 98-103.

Malamos N., Barouchasa P.E., Tsirogiannisb I.L., Liopa-Tsakalidia A., Koromilasc Th. 2015. Estimation of monthly FAO Penman-Monteith evapotranspiration in GIS environment, through a geometry independent algorithm Agriculture and Agricultural Science Procedia, 4 2015: 290-299

Meng Li., Ronghao Chu., Abu Reza Md Towfiqul Islam, Shuanghe, Shen. 2018. Reference Evapotranspiration Variation Analysis and Its Approaches Evaluation of 13 Empirical Models in Sub-Humid and Humid Regions: A Case Study of the Huai River Basin, Eastern China .Water ,10(4): 493-502

Naidu D. and Majhi B. 2019. Reference evapotranspiration modeling using radial basis function neural network in different agro-climatic zones of Chhattisgarh. Journal of Agrometeorology, 21(3): 316-326.

Patil A. P, and Deka P.C. 2017. Performance evaluation of hybrid Wavelet-ANN and Wavelet-ANFIS models for estimating evapotranspiration in arid regions of India. Neural Computing and Applications, 28(2): 275-285.

Peng L., Li Y., Feng H. 2017. The best alternative for estimating reference crop evapotranspiration in different sub-regions of mainland China. Scientific reports, 7(1): 1-19.

Pour-Ali Baba A., Shiri J., Kisi O., Fard A. F., Kim S., Amini R. 2015. Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). *Hydrology Research*, 44(1): 131-146.

Saylan L., Kimura R., Altınbaş N., Çaldağ B., Bakanoğulları F. 2019. Modeling of Surface Conductance over Sunn Hemp by Artificial Neural Network. Italian Journal of Agrometeorology, (3): 37-48.

Shamshirband S., Amirmojahedi M., Gocić M., Akib S., Petković D., Piri J., Trajkovic S. 2016. Estimation of reference evapotranspiration using neural networks and cuckoo search algorithm. Journal of Irrigation and Drainage Engineering, 142(2): 04015044.

Stahel W. 1997. "Robust alternatives to least squares," *Adv. Math. Tools in Metrol. III, Ser. Adv. Math. for Appl. Sci.*, World Scientific Publishing Company, 45: 118-133.

Tabari H., Kisi O., Ezani A., Talaee, P. H. 2012. SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. Journal of Hydrology, 444: 78-89.

Trajkovic S., and Kolakovic S. 2010. Comparison of simplified pan-based equations for estimating reference evapotranspiration. Journal of irrigation and drainage engineering, 136(2): 137-140.

Tsakiris G., Pangalou D., Vangelis H. 2007. Regional drought assessment based on the Reconnaissance Drought Index (RDI). Water resources management, 21(5): 821-833.

Vicente-Serrano S. M., Beguería S., López-Moreno J. I. 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. Journal of climate, 23(7): 1696-1718.

Wable P.S., Jha M.K., Gorantiwar S.D. 2019. Assessing suitability of temperature-based reference evapotranspiration methods for semi-arid basin of Maharashtra. Journal of Agrometeorology, 21(3): 351-356.

Yaseen A H., Amir P.N., Zhen Z., Subhasis G., Sean A.W. 2016. Bayesian Regression and Neuro-Fuzzy Methods Reliability Assessment for Estimating Streamflow. Water, 8, 287 pp. 15.

Yirga S. A. 2019. Modelling reference evapotranspiration for Megecha catchment by multiple linear regression. Modeling Earth Systems and Environment, 5(2): 471-477.