

A Model for Evapotranspiration by Fuzzy Least Squares Method

Sriram A.V¹, Rashmi C.N²

¹(Assistant Professor, Department of Civil Engineering, UVCE, Bangalore University)

²(Assistant Professor, Department of Civil Engineering, PESIT, Bangalore)

Abstract: Prediction of evapotranspiration is important for study, design and management of irrigation systems. In this study, the potential of fuzzy Linear Regression model using least squares is investigated in modelling of mean monthly potential evapotranspiration (PET) obtained using the standard FAO-56 Penman-Monteith equation. Various combinations of daily climatic data, namely solar radiation, air temperature, relative humidity and wind speed, are used as inputs to the Fuzzy regression model, so as to evaluate the degree of effect of each of these variables on Penman-Monteith estimated PET. Residual error, Co-efficient of Determination and Nash-Sutcliffe efficiency test are used as comparison criteria for evaluation of the model performance. The Fuzzy regression technique gave Residual errors of (-0.07)

Co-efficient of determination of (0.994) and Nash Sutcliffe efficiency of (0.998) for Bangalore meteorological station. Based on the results, it was found that the fuzzy regression model could be successfully employed in estimating the daily Potential Evapotranspiration in the study area.

Keywords: Fuzzy Regression, Method of Least squares, Potential Evapotranspiration.

I. Introduction

Efficient use of water resources in agro-ecosystems of the World has become increasingly important because of rapid depletion of water resources, industrial development and population increase, drought conditions, and degradation of ground and surface water quality in many regions. Evapotranspiration (ET) is a major component of the hydrologic cycle. A major proportion, an average of 60% of the total precipitation falling on the land surface is returned to the atmosphere across the globe by ET. Thus, accurate quantification of crop water use (evapotranspiration) is crucial for better management and allocation of water resources.

From the several existing ET₀ equations, the FAO-56 Penman-Monteith equation, [1] is currently used widely and can be considered as a sort of standard. Many studies have been conducted worldwide to find the proper method of accurate PET estimation. Numerous temperature-based models have been proposed and used to estimate PET [2]. Several studies done under diverse climatic conditions have revealed a widely varying performance of alternative equations in which local calibration are required [1]. Several models have been proposed to predict PET; however there is no universal consensus on the suitability of any given model for a given climate.

Because evapotranspiration is an incidental, nonlinear, complex, and uncertain process, it is not an easy task to obtain an accurate formula for PET estimation that is suitable to encompass all environments. Therefore, there is a new avenue to use intelligent soft computing techniques, such as Fuzzy regression, Artificial networks etc., to estimate evapotranspiration.

Fuzzy rule systems have been used successfully in reservoir management [3], rainfall-runoff problems [4]; [5] and in parameters of groundwater flow [6]; [7] and [8].

The Objective of this study is to develop fuzzy logic intelligent model using method of least squares to estimate monthly Potential evapotranspiration with reasonable accuracy from reliable climatic measurements, and subsequently to determine the ability of fuzzy logic for estimating PET values as a function of climatic variables and compare the accuracy with PET measured using standardized FAO 56 Penman-Monteith method in the Semi humid city of Bangalore in South India.

II. Methodology

2.1 Introduction

The monthly potential evapotranspiration values are calculated using the daily values for GKVK campus, Hebbal, Bangalore meteorological station for a data of 5 years (2007-2011) by FAO 56 Penman-Monteith equation and this was compared with a fuzzy regression model for Evapotranspiration developed using Least-squares method.

After establishing the regression equation, it is of interest to measure the quality or reliability of the fitted regression equation. The *i*th residual is the difference between the observed value *Y_i* and the corresponding fitted value \hat{Y}_i . This residual is denoted by *e_i* and is defined in general as

$$e = Y - \hat{Y} \quad \text{----- (1)}$$

The fuzzy coefficient of determination (HR^2) is used to interpret the proportion of the total variation in Y explained by the regression line which is defined by where \bar{Y} is the mean of \hat{Y} . ----- (2)

$$(HR)^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2}{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}$$

Nash-Sutcliffe Efficiency (CE):

The efficiency E proposed by Nash and Sutcliffe (1970) is defined as one minus the sum of the absolute squared differences between the predicted (P_i) and observed (O_i) values normalized by the variance of the observed values during the period under investigation. It is calculated as ----- (3) The

$$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}$$

range of E lies between 1.0 (perfect fit) and -1. An efficiency of lower than zero indicates that the mean value of the observed time series would have been a better predictor than the model.

2.2 Penman-Monteith equation

This equation is used to predict evapotranspiration from standard climatological records of sunshine, temperature, humidity and wind speed. This approach includes all parameters that govern energy exchange and corresponding latent heat flux. The Penman-Monteith form of the combination equation is given as

$$ET_o = \left[\frac{0.408 \Delta (R_n - G) + \gamma \left(\frac{900}{T + 273} \right) u_2 (VPD)}{\Delta + \gamma (1 + 0.34 u_2)} \right] \text{-----(4)}$$

- Where:
- ET_o = daily reference ET [mm/d], for longer periods 900 becomes 37
 - T = mean air temperature at 2m high [°C]
 - VPD = vapor pressure deficit [kPa]
 - u₂ = wind speed at 2 m high [m/s] = 2 m/s
 - R_n = net radiation at the crop surface [MJ m⁻² d⁻¹]
 - Δ = slope vapour pressure curve [kPa °C⁻¹]
 - γ = psychometric constant [kPa °C⁻¹]
 - G = soil heat flux density [MJ m⁻² d⁻¹]

2.3

Fuzzy Regression Methodology

For a crisp set A, belonging to an universe X, the characteristic function $\mu_A(x):X \rightarrow \{0, 1\}$ is given by

$$\mu_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} \text{----- (5)}$$

In the fuzzy case, the characteristic function, dubbed as the membership function is simply every function $\mu_A(x):X \rightarrow \{0, 1\}$

The estimated Y_i at a particular μ value is given by

$$\mu_{\hat{Y},L} = (a_0 + a_1 + \dots + a_k) - (1 - \mu)(c_{0,R} + c_{1,R} X_1 + c_{2,R} X_2 + \dots + c_{k,R} X_k) \text{----- (6)}$$

where a_0, a_1, \dots, a_k are the coefficients of Y_i at $\mu = 1$; $(C_{o,L} + C_{1,L} X_1)$ and $(C_{o,R} + C_{1,R} X_1)$ are the left fuzzy width and right fuzzy width for X_1 ; $(C_{o,L} + C_{1,L} X_2)$ and $(C_{o,R} + C_{1,R} X_2)$ are the left fuzzy width and right fuzzy width of X_2 ; $(1-\mu)(C_{o,L} + C_{2,L} X_1)$ and $(1-\mu)(C_{o,R} + C_{2,R} X_1)$ are the left fuzzy width and right fuzzy width for X_1 at a given μ value.

Observed data represented by symmetrical fuzzy numbers can be expressed as

$$Y_i = (Y_i, (1-\mu) e) \text{----- (7)}$$

The method of least-squares is used to find that particular regression line (Yi) where the sum of squared deviations of the data points (Yi) above or below it is minimized. To facilitate fuzzy regression analysis matrix algebra is employed. The general fuzzy linear model can be expressed in the matrix form as:

$$\tilde{Y} = X\tilde{\beta} \tag{8}$$

where,

$$\tilde{Y} = \begin{bmatrix} (y_1, (1-\mu)e_{1,L}, (1-\mu)e_{1,R}) \\ (y_2, (1-\mu)e_{2,L}, (1-\mu)e_{2,R}) \\ \vdots \\ (y_n, (1-\mu)e_{n,L}, (1-\mu)e_{n,R}) \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix} \quad \hat{\beta} = [\hat{\beta}_0 \quad \hat{\beta}_1 \quad \dots \quad \hat{\beta}_k] = \begin{bmatrix} (a_0, (1-\mu)c_{0,L}, (1-\mu)c_{0,R}) \\ (a_1, (1-\mu)c_{1,L}, (1-\mu)c_{1,R}) \\ \vdots \\ (a_n, (1-\mu)c_{n,L}, (1-\mu)c_{n,R}) \end{bmatrix}$$

----- (9)

In the above equations, Matrices Y and X are the data matrices associated with response variable and predicted variables, respectively. Matrix β contains the least-squares estimates of regression coefficients. To obtain regression parameters, equation (9) can be transformed as

----- (10)

$$(X'X)\hat{\beta} = X'\tilde{Y}$$

where X' is the transpose of X.

By Matrix operations, the regression coefficients can be derived as follows.

$$\hat{\beta} = (X'X)^{-1} X'\tilde{Y} \tag{11}$$

Where $(X'X)^{-1}$ is the inverse of the matrix $(X'X)$

2.3.1 Fuzzy Estimators

Now we may find the confidence intervals for X1, X2, X3.... Xk. Let $(X'X)^{-1} = A[a_{ij}]$. A $(1-\beta)$ 100% Confidence interval of a is given as

----- (12)

$$[\hat{a} - t_{\beta/2}\hat{\sigma}\sqrt{a_{11}}, \hat{a} + t_{\beta/2}\hat{\sigma}\sqrt{a_{11}}],$$

Where, a11 is the first element along the main diagonal of matrix A.

Similarly the $(1-\beta)$ 100% Confidence intervals of the regression Co-efficients are found.

In the t distribution, to find the critical value $t_{\beta/2}$, we use $n-(k+1)$ degrees of freedom. Now put these confidence intervals together, one on top of another, to get the fuzzy estimators $X1^\wedge, X2^\wedge \dots$ of X1, X2,... respectively. By placing Confidence Intervals one on top of another, the- cuts are obtained and hence the membership functions of the regression coefficients are obtained. The Vertex method is used for propagation of fuzziness and Centroid method is used for defuzzification.

III. Study Area And Data Used

3.1 Study Area: The study area is GKVK campus, Hebbal of Bangalore in Karnataka State lying in tropical climates of India having sub humid climatic conditions. Fig.3 shows the location of the meteorological station. In the present study, a fuzzy linear regression model for evapotranspiration is developed for the study area. The general details of the meteorological station are summarized in Table 1.

TABLE 1: General details of the study area

Location	Climate Type	Latitude	Longitude	Altitude in m
Bangalore	Semi humid	13°05'N	77°34 'E	923

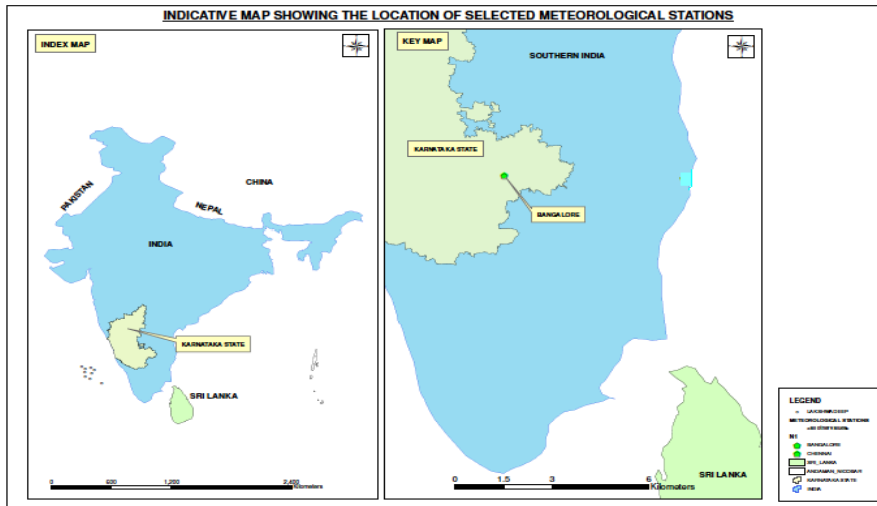


Figure 3: Location of the selected meteorological station

3.2 Data used: The data for Bangalore was obtained from the weather station located at University of Agricultural sciences, GKVK, Hebbal for a span of 5 years (2007-2011). Fig 4 gives the Average annual values of mean temperature, Minimum Humidity, Maximum Humidity and wind speed data of the meteorological station. It is seen that as mean temperature increases, relative humidity decreases and vice-versa for the data obtained for the meteorological station.

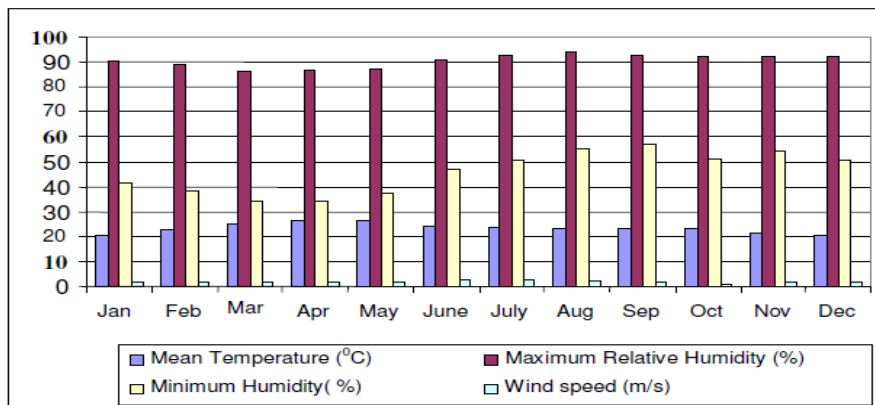


Figure 4: Average Annual values of Mean Temperature, Maximum Humidity, Minimum Humidity and Wind speed for Bangalore (2007 - 2011)

IV. Results And Discussions

Table 2: PET values in mm/day at Bangalore Meteorological station by FAO 56 PM Method.

Month	Bangalore				
	2007	2008	2009	2010	2011
January	4.24	4.06	3.71	3.87	3.69
February	4.92	4.48	4.73	4.94	4.18
March	5.97	5.02	4.95	5.61	5.14
April	6.09	5.36	5.38	5.54	4.69
May	6.06	6.14	4.81	5.45	4.80
June	4.86	4.93	4.53	4.73	4.60
July	4.54	4.35	4.58	3.74	3.77
August	4.05	3.67	3.97	3.36	3.39
September	3.97	4.13	2.97	3.29	3.85
October	3.60	3.38	3.81	3.30	3.22
November	3.46	3.37	3.19	2.68	3.16
December	3.40	3.33	3.37	2.80	3.09

Table 2 shows the PET values obtained at Bangalore station for a period of 5 years. It can be seen that high values of PET are obtained in the months of March, April and May which may be due to the fact that these months have higher temperature, lesser humidity and high wind speeds compared to other months. It can also be seen that amongst all the years of data considered the highest PET value was obtained in May 2008 and the lowest value in November 2010. From the dataset it can be seen that May 2008 has recorded the highest temperature and wind speed and a low value of Relative humidity and November 2010 has recorded low temperature and wind speed and high relative humidity.

4.1 Fuzzy regression model for Bangalore

The Fuzzy Regression Model obtained for Bangalore is $Y = (4.386, 3.712) + (-0.072, -0.079) X_1 + (0.089, 0.067) X_2 + (0.542, 0.498) X_3 + (0.2598, 0.218) X_4$

The Variance-Covariance matrix form by Computation is obtained as

$$\begin{pmatrix} 13.5058 & -0.12551 & 0.001813 & 0.041220674 & 0.475834319 \\ -0.12551 & 0.001437 & -0.00158 & -0.00057216 & 0.006202886 \\ 0.001813 & -0.00158 & 0.01475 & -0.0077651 & 0.021352462 \\ 0.041221 & -0.00057 & -0.00777 & 0.06009479 & 0.006361901 \\ -0.47583 & 0.006203 & -0.02135 & 0.006361901 & 0.051227451 \end{pmatrix}$$

The β matrix for regression Co efficient is obtained as

$$\begin{pmatrix} 4.033379 \\ -0.07546 \\ 0.078561 \\ 0.518592 \\ 0.238776 \end{pmatrix}$$

The Fuzzy estimators for $(1-\beta)*100\%$ confidence intervals is got by the equation 12. The $(1-\beta)$ 100% Confidence interval for each of the regression co-efficients A0, A1, A2, A3 and A4 was estimated. Table 3 shows the $(1-\beta)$ 100% confidence intervals for the four regression co-efficients.

Table 3: $(1-\beta)$ 100% Confidence intervals for regression coefficients.

Regression Coefficients	$(1-\beta)$ 100% confidence intervals
A0	$4.049 \pm t\alpha/2(0.3225)$
A1	$-0.0755 \pm t\alpha/2(0.0033)$
A2	$0.0779 \pm t\alpha/2(0.0107)$
A3	$0.5198 \pm t\alpha/2(0.0215)$
A4	$0.239 \pm t\alpha/2(0.0119)$

The confidence intervals for each of the α levels were placed one on top of another to obtain the membership function of A0, A1, A2, A3 and A4. Assuming a symmetrical triangular model, figure 5 shows the plot of A0, A1, A2, A3 and A4 for different α -cut values. The confidence intervals are taken in accordance with the T- table value function from 0.005 to 1. The fitted fuzzy regression equation is developed based on the estimated regression coefficients. Table 4 shows the centroidal fuzzy linear regression equations obtained for each of the regression coefficients and also the total error in the model along with fuzzy coefficient of determination.

Table 4: Fuzzy linear regression model for Bangalore meteorological station

Station	Model	HR ²	e = Yi - Ŷi	Nash-Sutcliffe Coefficient E
Bangalore	$Y = (4.386, 3.712) + (-0.072, -0.079) X_1 + (0.089, 0.067) X_2 + (0.542, 0.498) X_3 + (0.2598, 0.218) X_4$	0.994	-0.07	0.998

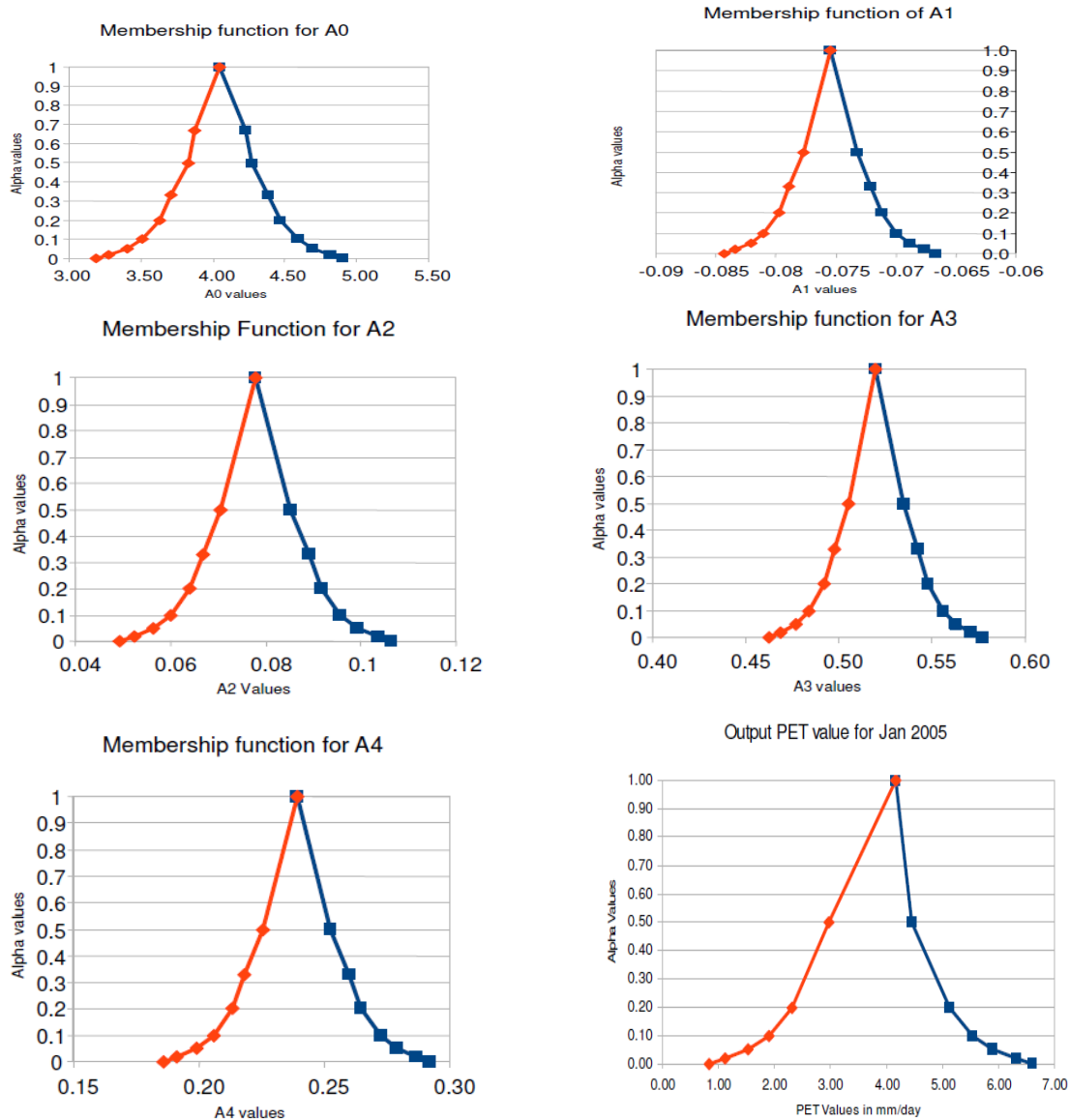


Figure 5: Membership function for Fuzzy Regression Coefficients A0, A1, A2, A3, A4 and Output Values of PET for January 2005 for Bangalore Meteorological station

The Fuzzy regression model is propagated by the vertex method and output variable Y is determined. The membership function for each of the months for 5 years of data is calculated using the fuzzy linear regression model developed. An example output function of January 2005 for Bangalore station is shown in fig 5. It can be seen that the output membership function is an asymmetrical triangular function, while the inputs were assumed to be symmetrical fuzzy numbers which may be probably due to the fact that only four parameters (Temperature, Relative humidity, Wind speed and Solar radiation) were considered as the predictors for the model. The coefficient of determination of model is 0.994 and hence it can be deduced that 99% of the Potential evapotranspiration values can be explained by the four predictors, signifying a strong linear correlation and a Nash Sutcliffe efficiency of 0.998 ascertains the adequacy of the Fuzzy regression model for the study area considered.

V. Conclusions

From this study, it has been found that

- The PET values are obtained by FAO-56 full data Penman-Monteith equation and it was found that highest PET value was obtained in May 2008 (6.14 mm/day) and the lowest value in November 2010 (2.68 mm/day). Higher values of PET values indicate higher temperature, lower Humidity and higher wind speed as observed from the dataset.

- A Fuzzy linear regression model was developed with a fuzzy coefficient of determination HR^2 of 0.994 and a Nash-Sutcliff Coefficient of efficiency of 0.998 for the study area.
- It is found that the fuzzy linear regression model ($e = -0.07$) is a good fit the study area since it has a lower value of residual error and can be effectively used for the prediction of Potential Evapotranspiration in the study area.

References

- [1] Allen, R.G, Pereira, L.S., Raes, D., and Smith, M., 1998: Crop evapotranspiration - Guidelines for computing crop water requirements - *FAO Irrigation and drainage paper 56*, FAO, Rome.
- [2] Wang, Yu-Min, Willy, N., Lenn.,X., Alexander, G., Seydou, T., and Lian-Tsai, D.(2011):Comparative study on estimating reference evapotranspiration under limited climate data condition in Malawi. *International Journal of the Physical Sciences Vol. 6(9)*, pp. 2239-2248.
- [3] Panigrahi D.P. and Majumdar P.P, 2000. Reservoir operation modelling with fuzzy logic, *WaterResour. Manag.*14, 89–109.
- [4] Nayak, P., Sudheer, K., and Ramasastri, K., 2005. Fuzzy Computing based Rainfall-runoff model for real time flood forecasting. *Hydrol. Processes*, 19: 955-968.
- [5] Yu, P.S., and Chen, S.T., 2005. Updating real-time flood forecasting using a fuzzy rule-based model. *Hydrol. Sci. J.*, 50: 265-278.
- [6] Tzimopoulos, C., Mpallas, L., and Papaevangelou, G., 2008. Estimation of Evapotranspiration using fuzzy systems and comparison with the Blaney-Criddle method. *Journal of Environmental Science and Technology*, 1; 181-186.
- [7] Moutsopoulos, K., Chakidis, I., and Tzimopoulos, C., 2005. Fuzzy analysis of ground water resources including incomplete data of aquifer spatial extent. *J. Mech. Behav. Mater.*, 16: 95-101.
- [8] Chalkidis, I. N., Tzimopoulos, C., Evangelides, H., and Yannopoulos, S., 2006. Application of Linear and Fuzzy Programming for water management of an aquifer, *WSEAS Transactions on systems*, 12(5), 2844-2849.