Impact of Climatological Parameters on Reference Crop Evapotranspiration using Multiple Linear Regression Analysis

H R Mahida[1], V N Patel[2]

[1]Civil Engineering Department, A D Patel Institute of Technology, Gujarat Technological University, India

[21]Civil Engineering Department, A D Patel Institute of Technology, Gujarat Technological University, India

ABSTRACT : *The reference crop evapotranspiration (ETo) of Bhaniyara station have been estimated using multiple linear regression (MLR) technique in XLSTAT tool. The meteorological data such as maximum temperatures, sunshine hours, relative humidity and wind speed were collected for the Bhaniyara station of Vadodara district, Gujarat state, India for the period of nine years and the missing value of that data series was also determine using SPSS20 software. The observed ETo values have been estimated using the equation of evapotranspiration (FAO-56).MLR is carried out using ETo as predictor variable and maximum temperatures, sunshine hours, relative humidity and wind speed as independent variable to find out predominant factor on ETo. This whole procedure is done for five different Models. In model 1, Maximum temperature, relative humidity, sunshine hours and wind speed are correlated with ETo. In model 2, Maximum temperature, relative humidity and sunshine hours are correlated with ETo. In model 3, Maximum temperature, sunshine hours and wind speed are correlated with ETo. In model 4, Maximum temperature, relative humidity and wind speed are correlated with ETo. In model 5, wind speed, sunshine hours and relative humidity are correlated with ETo. In case of model 1 the value of R, R^2 and RMSE for 70% dataset is 0.911, 0.830 and 0.341 respectively and for 30% dataset it is 0.954, 0.910 and 0.325 respectively. As the value of R and R^2 are nearer to 1 and the value of RMSE is low, which is good. As the model 1 gives the best correlation values as compared to model 2 model 3, model 4 and model 5 it can be accepted as the best fit model for prediction of ETo. Considering maximum temperature the model gives good correlation values hence maximum temperature is accepted as predominant factor and the presence of relative humidity does not play an important role in prediction of ETo for this study area.*

Keywords *- Climate change, multiple linear regression, performance evaluation, reference crop evapotranspiration*

I. INTRODUCTION

India is more susceptible to effect of climate change due to its high dependence on climate sensitive sectors like agriculture and forestry. According to the IPCC [1], the average global surface temperature increased by 0.74°C over the last 100 years. General agreement have revealed that global warming and related changes to the hydrological cycle are likely to enhance the frequency and severity of extreme climate events, causing more severe floods and droughts. Global warming due to the enhanced greenhouse effect is expected to cause major changes in various climatic variables, such as precipitation, relative humidity, solar radiation and temperature. Atmospheric temperature is the most widely used indicator of climatic changes as global and regional scales, and global land-surface air temperatures have increased in the Northern Hemisphere by 0.3°C/ decade from 1979 to 2005.

The combination of two separate processes, where water is lost from the soil surface by evaporation and from the crop by transpiration, is called as evapotranspiration. Hydrological parameters such as precipitation, evapotranspiration, soil moisture and ground water are likely to change with climate [2], and the impact of climate change on evapotranspiration rate is important for hydrologic processes. Crop water requirements depend upon several climatic variables like rainfall, radiation, temperature, humidity and wind speed. Therefore, any change in climatic parameters due to global warming will also affect evapotranspiration, [3]. An indirect way to obtain estimates of evapotranspiration is the evaporation rate from pans filled with water, known as pan evaporation (Epan). Trends in Epan have been reported with different conclusions depending on the region studied. Jhajharia [4] found both decreasing and increasing tendencies in Epan in northeast India, depending on the location of the station.

Decreases in Epan have been attributed to decreasing surface solar radiation and wind speed [5]., and increases in cloud cover, greater air pollution and higher concentrations of atmospheric aerosols. Pan evaporation depends on the water surface temperature and energy balance between the evaporation pan, water and the atmosphere. If the humidity does not change, increasing water temperature should increase evaporation. If the humidity increases, it will partially offset the impact of higher temperature on the evaporation. Small changes in evapotranspiration can have important consequences in arid climates. 1% temperature increase could increase evapotranspiration by 12.69% in arid regions of Rajasthan, India, where the annual rainfall varies from 100 to 400 mm and mean temperature varies by about 25°C, [3].

Reference crop evapotranspiration (ETo) refers to crop evapotranspiration in the open short grass land where the soil moisture is adequate, ground is completely covered, grass grew normally with the similar height (grass height is about 8 ~ 15 cm) [6]. ETo is the most important parameter while predicting the crop water requirement. In the context of climate change, changes of temperature, wind speed, rainfall, solar radiation and other factors will lead to the change of ETo, thus affecting the crop water demand and agricultural water usage. In the context of climate change, changes of temperature, wind speed, rainfall, solar radiation and other factors will lead to the change of ETo, thus affecting the plan of crop water demand and agricultural water usage.

The objective of the present study is;

1. To estimate the reference crop evapotranspiration using Epan data of the Bhaniyara station.
2. To explore quantitative relationship between ETo and the influencing climatic variable.
3. To determine the predominant factor for prediction of ETo.

II. MATERIALS AND METHODS

A. *Study Area and Data Collection*

Bhaniyara is a village in Waghodia Taluka in vadodara district of Gujarat state, India. It is located 17 km towards east from district head quarters vadodara.132 km from state capital Gandhinagar. Bhaniyara is located at 22˚23’ N latitude, 73˚16’ E longitude at an altitude of 33 m above mean sea level.

The daily meteorological data such as maximum temperature, sunshine hours, relative humidity and wind speed for the Bhaniyara station were collected and used in the data analysis and model development. The monthly average metrological data for the nine years has been used for this study.

B. *Methodology*

In this study, daily metrological data for the Bhaniyara station were collected. These data has been converted into the monthly average data. Some data gaps or missing values are indentified in data. The missing values were found out with the SPSS 20 software. The linear trend at a point method is used to find the missing values of the data. To study the impact of metrological data on reference crop evapotranspiration (ETo) multiple linear regression analysis was performed using XLSTAT software. The observed ETo values have been estimated using the general equation of evapotranspiration (FAO-56).The equation is;

$$\mathbf{ETo = Epan \times Kp.} \qquad [1]$$

Where,

ETo = Reference Crop Evapotranspiration (mm/day).

Kp = Pan Co-Efficient for class A Pan placed in short green cropped area.(FAO-56), assume windward side distance of green crop is 100 m.

Epan = Pan Evaporation (mm/day).

Relative humidity and wind speed data has been used from the given metrological data to find out the Kp value using TABLE 5 from FAO-56 (Allen et al., 1998).

MULTIPLE LINEAR REGRESSION (MLR) MODEL:

The objective of the model is the transfer of information among several variables observed simultaneously and the estimation of the dependent variable from the several other observed independent variables. The monthly reference crop evapotranspiration (ETo) at a meteorological center is expressed as a simple linear model as

$$\mathbf{ET_O = C + a_1 X_1 + a_2 X_2 + a_3X_3 + a_4X_4 + a_5X_5} \qquad [2]$$

Where, $a_1, a_2, \ldots$ and C are empirical constants and $X_1, X_2, \ldots$ are the meteorological parameters influencing the region.

Now, MLR is carried out using ETo as dependent variable and maximum temperature, sunshine hours, relative humidity and wind speed as independent variable to find out predominant factor on ETo and obtain best model.

This whole procedure is repeated for five different Models.

In Model 1, Maximum temperature, relative humidity, sunshine hours and wind speed are correlated with ETo.

In Model 2, Maximum temperature, relative humidity and sunshine hours are correlated with ETo.

In Model 3, Maximum temperature, sunshine hours and wind speed are correlated with ETo.

In Model 4, Maximum temperature, relative humidity and wind speed are correlated with ETo.

In Model 5, wind speed, sunshine hours and relative humidity are correlated with ETo.

This entire data set is divided into two datasets i.e. 70% data for training and 30% data for validation. Then, the corresponding co-efficient of correlation (R) for training and testing data set is evaluated.

PERFORMANCE EVALUATION PARAMETERS:

The performance evaluation parameters used in the present study are the Pearson correlation coefficient (R), coefficient of determination (R^2) and root mean square error (RMSE).

A. Pearson correlation coefficient (R): Correlation – often measured as a correlation coefficient which indicates the strength and direction of a linear relationship between two variables (for example model output and observed values). A number of different coefficients are used for different situations. The formula to find out the correlation coefficient is;

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad [3]$$

The correlation is +1 in the case of a perfect increasing linear relationship and 1 in case of a decreasing linear relationship, and the values in between indicates the degree of linear relationship between model and observations. A correlation coefficient of 0 means the there is no linear relationship between the variables.

The square of the Correlation coefficient (R^2), known as the ***coefficient of determination***,

The coefficient of determination, R^2, is useful because it gives the proportion of the fluctuation of one variable that is predictable from the other variable. It is a measure that allows determining how certain one can be in making predictions from a certain model/graph. The coefficient of determination is such that $0 < R^2 < 1$, and denotes the strength of the linear association between x and y.

The coefficient of determination represents the percent of the data that is the closest to the line of best fit.

B. Root mean square error (RMSE): The Root Mean Square Error (RMSE) is used to measure the difference between predicted values by a model and the actually observed values from the location. These individual differences are also called **residuals**. The RMSE of a model prediction with respect to the estimated variable X_j is defined as the square root of the mean squared error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - X_j)^2}{n}}$$

Where, X_i is observed values and X_j is modeled values at time/place *i*.

The RMSE values can be used to distinguish model performance in a calibration period with that of a validation period as well as to compare the individual model performance to that of other predictive models.

III. Results and Discussion

In this study the performance evaluation parameters such as R, R^2 and RMSE has been calculated for the two stages i.e. training and validation for each model and are given in Table I.

Table 1: Performance Evaluation Parameters for Each Model

MODEL	RMSE		R		R^2	
	70%	30%	70%	30%	70%	30%
1	0.341	0.325	0.911	0.954	0.830	0.910
2	0.345	0.328	0.909	0.948	0.826	0.899
3	0.341	0.324	0.911	0.954	0.829	0.910
4	0.355	0.357	0.904	0.949	0.818	0.900
5	0.566	0.538	0.727	0.750	0.528	0.563

Referring to the table 1, in case of model 1 the value of R, R^2 and RMSE for 70% dataset is 0.911, 0.830 and 0.341 respectively and for 30% dataset it is 0.954, 0.910 and 0.325 respectively, which is very good. In case of model 2 the value of R, R^2 and RMSE for 70% dataset is 0.909, 0.826 and 0.345 respectively and for 30 % dataset it is 0.948, 0.899 and 0.328 respectively, which is comparatively good. In case of model 3 the value of R, R^2 and RMSE for 70% dataset is 0.911, 0.829 and 0.341 respectively and for 30% dataset it is 0.954, 0.910 and 0.324 respectively, which is also very good. In case of model 4 the value of R, R^2 and RMSE for 70% dataset is 0.904, 0.818 and 0.355 respectively and for 30% dataset it is 0.949, 0.900 and 0.357 respectively, which is also good. In case of model 5 the value of R, R^2 and RMSE for 70% dataset is 0.727, 0.528 and 0.566 respectively and for 30% dataset it is 0.750, 0.563 and 0.538 respectively, which is comparatively low. Here, it is easily noticeable that while, considering maximum temperature co-efficient of correlation achieved as the best. From the above exercise one can observed that value of maximum temperature is significantly affecting the value of ETo as compared to the value of sunshine hours, relative humidity and wind speed. Addition of Relative Humidity does not perform important task. Considering the all models, model 1 gives the best correlation values Hence, It is accepted as best fit model. The Fig 1 and Fig 2 show the plot of observed ETo vs. predicted ETo for training and validation respectively for model 1.

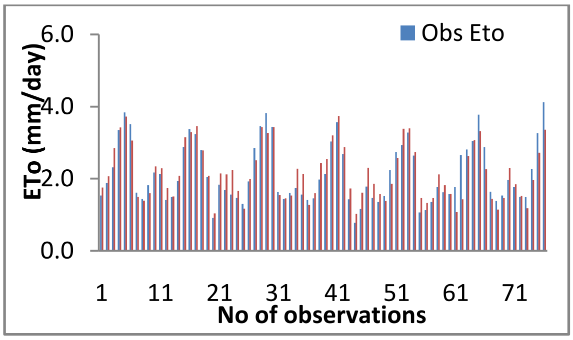

Fig 1: Observed ETo v/s predicted ETo for training of Model 1

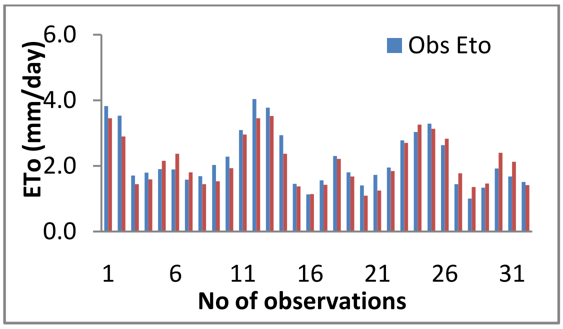

Fig 2: Observed ETo v/s predicted ETo for validation of model 1

IV. CONCLUSION

Comparing results of all five models, following are the conclusions:

In case of model 1 the value of R and R^2 are nearer to 1 and the value of RMSE is low, which is good.

In case of model-2 the value of R and R^2 are nearer to 1 but lower than model-1 and also, the value of RMSE is higher than model-1. Hence the model can't be accepted as best fit model.

In case of model 3 the value of R and R^2 are nearer to 1 and higher than model-2, also the value of RMSE is less compared to model-2.

In case of model 4 the value of R and R^2 are nearer to 1 but, lower than model-1, model-2 and model-3, also the value of RMSE is higher than model-1, model-2 and model-3. Hence the model can't be accepted as best fit model.

In case of model 5 the value of R and R^2 are not nearer to 1 and also RMSE value is very high as compared to other models.

As the model 1 gives the best correlation values as compared to model 2 model 3, model 4 and model 5 it can be accepted as the best fit model for prediction of ETo. When considering maximum temperature the model gives good correlation values hence maximum temperature is accepted as predominant factor and the presence of relative humidity does not play an important role in prediction of ETo for this study area.

REFERENCES

[1] IPCC Summary for policymakers. In: S. Solomon, D. Qin, M. Manning and Z. Chen, "Climate change: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change", Cambridge University Press, Cambridge, 2007

[2] P. H. Gleick, "Methods for evaluating the regional hydrologic impacts of global climatic changes." J Hydrol., 1986, 88, 97–116.

[3] R. K. Goyal, "Sensivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India)." Agric Water Manag., 2004, 69, 1–11.

[4] D. Jhajharia, S. K. Shrivastava, D. Sarkar, S. Sarkar, "Temporal characteristics of pan evaporation trends under the humid conditions of northeast India.", Agric Meteorology,2009, 149, 763–770.

[5] C, Xu, L. Gong, T. Jiang, D. Chen, V. P. Singh, "Analysis of spatial distribution and temporal trend of reference evapotranspiration and pan evaporation in Changjiang (Yangtze River) catchment.", J Hydrol. 2006, 327, 81–93.

[6] R. G, Allen, Pereira LS, Raes D, Smith M, "Crops evapotranspiration. Guidelines for computing crop requirements." Irrigations and Drainage Paper 56. FAO, 1998, Rome.

[7] F .P. Parekh, and T. M. V. Suryanarayana, "Climatic Variability and Crop Weather Model of Wheat Crop for Vallabh Vidyanagar", the Indian Journal of Technical Education, (Special Issue for NCEVT'12), April 2012, 325-330, 2012.

[8] F .P. Parekh, and T. M. V. Suryanarayana, "Impact of Climatological Parameters on Yield of Wheat Using Neural Network Fitting." International Journal of Modern Engineering Research (IJMER), 2012 Vol.2, Issue.5, Sep-Oct. pp-3534-3537.