International Journal of Agriculture Sciences
ISSN: 0975-3710&E-ISSN: 0975-9107, Volume 8, Issue 62, 2016, pp.-3529-3532.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000217

Research Article

# MODELING OF REFERENCE EVAPOTRANSPIRATION USING REGRESSION TECHNIQUES

**KHEDKAR D.D.[1*] SINGH P.K.[1], BHAKAR S.R.[1], KOTHARI MAHESH[1], JAIN H.K.[2] AND MUDGAL V.D.[3]**

[1]Department of Soil and Water Engineering, Maharana Pratap University of Agriculture & Technology, Udaipur, Rajasthan
[2]Department of Statistics, Maharana Pratap University of Agriculture & Technology, Udaipur, Rajasthan
[3]Department of Process and Food Engineering, Maharana Pratap University of Agriculture & Technology, Udaipur, Rajasthan
*Corresponding Author: Email-ddkhedkar1@gmail.com

Received: December 01, 2016; Revised: December 19, 2016; Accepted: December 20, 2016; Published: December 24, 2016

**Abstract-** Prediction of evapotranspiration is important for design and management of irrigation systems, water resources management and climatological studies. The ASCE had recommended Penman-Monteith model (FAO-56) as the sole standard method for determining ETo over the wide variety of climatic situations over the world and it requires all types of data. At many locations, there is either lack of meteorological data or availability of meteorological parameters is limited. It is necessary to find alternative to Penman-Moneith method with limited data availability. In order to carry out study, average weekly meteorological data, *viz.,* maximum temperature, minimum temperature, maximum relative humidity, minimum relative humidity, bright sun shine hours, wind speed and pan evaporation were collected from whether station located at Dhule (Mahaashtra, India) for period of 1980 to 2014. In this study, the potential of Linear Regression is investigated in modeling of reference evapotranspiration (ETo) using the standard FAO-56 Penman–Monteith equation. The four types of linear regression models were developed by varying the independent variables, these are; LR1(pan evaporation); LR2 (maximum temperature and minimum temperature); LR3 (maximum temperature, minimum temperature and bright sun shine hours); LR4 (maximum temperature, minimum temperature, maximum relative humidity, minimum relative humidity and bright sun shine hours). The results of all performance measures for all LR models during development stage varies in the range as R (0.902 to 0.933), d(IA) (0.946 to 0.964), RMSE (0.701 to 0.841), MAE (0.532 to 0.646), MAPE (11.609 to 14.274) and CE (0.813 to 0.870) and showed the performance in sequence of LR4, LR3, LR1 and LR2. It indicates that all LR models performed satisfactorily and showed marginal difference of performance measures among them in development stage. Similar kind of close difference for each performance measure occurred during validation stage of all LR models. It indicates that all LR models were validated satisfactorily and generalized for estimation of ETo. Overall, the performance suggest that all LR models can be an acceptable approach to predict ETo values for Dhule station as per data availability.

**Keywords**- Reference evapotranspiration, Meteorological data, Linear regression.

**Citation:** Khedkar D.D., et al., (2016) Modeling of Reference Evapotranspiration Using Regression Techniques. International Journal of Agriculture Sciences, ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 8, Issue 62, pp.-3529-3532.



## Introduction

Evapotranspiration denotes the quantity of water transpired by plants during their growth, or retained in plant tissue, plus the moisture evaporated from the surface of the soil and the vegetation [7]. The quantification of evapotranspiration is preceded by reference crop evapotranspiration (ETo) and it is defined as the evapotranspiration rate from a reference surface, not short of water. The reference surface is a hypothetical grass reference crop with an assumed crop height of 0.12 m, a fixed surface resistance of 70 s/m and an albedo of 0.23 [1]. The ETo can be measured directly by lysimeter or water balance approach or indirectly using meteorological parameters. Thus, it depends on meteorological factors such as temperature, relative humidity, wind speed and solar radiation [15]. The various indirect (climate based) methods were established for estimation of ETo at various climatic conditions, however Penman-Monteith method was recommended as sole standard by the Food and Agriculture Organization (FAO 56) and can be applied in wide variety of climatic situations[1]. The main limitation of this method is the difficulty in obtaining all necessary meteorological data, which are not always available in some weather stations or available in questionable quality. This lack of meteorological data leads to the development of simpler approaches to estimate ETo that require only few parameters.

There are many equations available to estimate ETo, but simpler equations give inconsistent values [5,8,13-15] due to their different weather data requirements or because they were developed for specific climatic regions. As a result many climate based methods performed better for a particular climatic condition and needs local adjustments or calibrations, hence there is need to assess advanced statistical techniques than existing methods for better accuracy in prediction of ETo. Most of researchers [12,9,3,6] found that linear regression models perform better than climate based models with same data requirement for alternatives to Penman-Monteith method. The practitioners and researches need to be provided guidance on the choice of the most appropriate ETo equation to be adopted when weather data are insufficient to apply the FAO-56 PM equation. Thus, it is important to apply statistical techniques in order to develop simple empirical equations to estimate ETo with limited data availability. Predictions of ETo rates are central to many planning applications in water resources and agriculture. We therefore propose an alternate method of ETo prediction using regression equations. This approach maintains the structural simplicity and minimal data requirements needed for many practical applications but also accounts for the key variables that affect ETo in different climates. The linear models provided here to calculate ETo can be used for location-specific or for regional scale evaluations and for estimations of ETo in areas where limited meteorological measurements are available. The models have the advantage of a simple structure enabling them to be used with existing weather datasets or with generated datasets from climate models. Thus, in this paper the linear regression models were developed for prediction of ETo with different combinations of meteorological parameters under limited data availability.

## Materials and Methods

### Study area and Data Collection

The study was carried out using meteorological data collected from whether station located at Dhule (Mahaashtra, India) for period of 1980 to 2014. Its geographical location is 20º54'N and 74º46'E with elevation above mean sea level is about 263 m. In order to carry out study, daily/weekly meteorological data, viz., maximum temperature, minimum temperature, maximum relative humidity, minimum relative humidity, bright sun shine hours, wind speed and pan evaporation were collected.

### FAO 56 Penman-Monteith method

The Penman-Monteith method is recommended by FAO-56 as the sole method for determining ETo, hence this method is used in this study. For calculating the reference evapotranspiration the detailed procedure given in FAO-56 is adopted in the present study. The FAO Penman-Monteith model to estimate ETo is given by[1];

$$ETo = \frac{0.408 * \Delta * (R_n - G) + \gamma * \left(\frac{900}{T+273}\right) * U_2 * (e_s - e_a)}{\Delta + \gamma * (1 + 0.34 * U_2)} \quad [3.1]$$

Where,
ET= Reference evapotranspiration (mm/day),
λ= Latent heat of vaporization (KJ/Kg),
Δ= Slope of saturation vapour pressure temperature curve (kPa/$^0$C),
γ = Psychometric constant (kPa/$^0$C),
T= Mean air temperature ($^0$C),
$e_s$= Saturated vapour pressure (kPa),
$e_a$= Actual vapour pressure (kPa),
$R_n$= Net radiation (MJ/m$^2$/day),
G = Soil heat flux density (KJ/m$^2$s),
$U_2$= Wind speed at 2m height (m/s),

### Linear Regression Modeling

In this study, average weekly data for the period 1980 to 2014 was used for development and validation of model for estimation of ETo under limited data availability. Out of total data period 80 percent data i. e. from 1980 to 2007 (1456 data sets) were used for development of model and 20 percent data i.e. 2008 to 2014 (364 sets) were used for validation of model. The SPSS 21.0 software was used to develop statistically optimal models of simple and multiple linear regression for prediction of ETo values.

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 general form of statistically optimal simple and multiple linear regression is given by;

$$Y = a + bx \quad ..[3.2]$$

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 \quad ......[3.3]$$

Where, Y is the dependent variable and $x_1, x_2, x_3, x_4, x_5$ are the independent variables, $a$ is interceptor and $b_1, b_2, b_3, b_4, b_5$ are the partial regression coefficients.

The ETo estimated from Penman-Monteith method were considered as dependent variable while meteorological parameters were assumed as independent ones for development of LR models. The four types of LR models were developed by varying the independent variables, these are;

Model 1 - LR1 - with single independent variable i.e. pan evaporation (Epan).

Model 2 - LR2 -with two independent variables as maximum temperature (Tmax) and minimum temperature (Tmin).

Model 3 - LR3 - with three independent variables as maximum temperature (Tmax), minimum temperature (Tmin) and bright sun shine hours (SSH).

Model 4 - LR4 - with five independent variables as maximum temperature (Tmax), minimum temperature (Tmin), maximum relative humidity (RHmax), minimum relative humidity (RHmin) and bright sun shine hours (SSH).

In this study, Model 1 i.e. represented by LR1 was in the form of simple linear regression, while Model 2, Model 3, Model 4 represented by LR2, LR3, LR4 respectively and were in the form of multiple linear regressions.

The performance evaluation of models during development and validation period were carried out with comparison of predicted and observed values of ETo. The performance criteria adopted here is the highest correlation coefficient (R), index of agreement d(IA), coefficient efficiency (CE) and the lowest values of root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE).

## Results and Discussion

Using the SPSS 21.0 software, the statistically optimal models of simple and multiple linear regression were developed for prediction of ETo values and results presented in [Table-1].

**Table-1** *Statistical Criteria for best fit LR models with mathematical expressions for prediction of ETo at Dhule station*

| Sr. No. | Statistical Coefficients | LR1 | LR2 | LR3 | LR 4 |
|---|---|---|---|---|---|
| 1 | Standard error (S) | 0.756 | 0.841 | 0.453 | 0.701 |
| 2 | Multiple Correlation Coeff. (R) | 0.921 | 0.902 | 0.924 | 0.933 |
| 3 | Adjusted $R^2$ | 0.849 | 0.813 | 0.853 | 0.870 |
| 4 | F - test value (F) | 8173.97 | 3161.48 | 2822.37 | 1952.37 |
| **Models** | **Mathematical expressions** | | | | |
| **LR1** | ETo = 1.707+0.440Epan | | | | |
| **LR2** | ETo = -9.203+0.393Tmax+0.029Tmin | | | | |
| **LR3** | ETo = -8.227+0.235Tmax+0.149Tmin+0.288SSH | | | | |
| **LR4** | ETo = -3.892+0.135Tmax+0.198Tmin -0.028RHmax+0.02RHmin+0.293SSH | | | | |

[Table-1] shows the statistical criteria for defining LR models with their mathematical expressions. It was observed that the values of multiple correlation coefficient (R), coefficient of correlation (adjusted $R^2$) and F test value for all models were more than 0.90, 0.81 and 1952 respectively, whereas the values of standard error for all models were less than 0.85. It indicates that all LR models satisfies statistical criteria well so they were considered for prediction of ETo values at Dhule station.

### Performance evaluation of LR models

**Table-2** *Performance evaluation of LR models with limited data for Dhule Station*

| Model | Statistical Criterias | | | | | |
|---|---|---|---|---|---|---|
| | R | d(IA) | RMSE | MAE | MAPE | CE |
| Development Period (1980-2007) | | | | | | |
| LR1 | 0.921 | 0.958 | 0.756 | 0.544 | 12.499 | 0.849 |
| LR2 | 0.902 | 0.946 | 0.841 | 0.646 | 14.274 | 0.813 |
| LR3 | 0.924 | 0.959 | 0.744 | 0.574 | 12.823 | 0.854 |
| LR4 | 0.933 | 0.964 | 0.701 | 0.532 | 11.609 | 0.870 |
| Validation Period (2008-2014) | | | | | | |
| LR1 | 0.951 | 0.969 | 0.543 | 0.417 | 11.291 | 0.895 |
| LR2 | 0.919 | 0.943 | 0.832 | 0.692 | 18.018 | 0.752 |
| LR3 | 0.938 | 0.958 | 0.720 | 0.579 | 15.160 | 0.815 |
| LR4 | 0.946 | 0.960 | 0.709 | 0.569 | 14.377 | 0.820 |

[Table-2] shows the statistical performance of different LR models with limited data during development and validation period for Dhule Station. During development stage, it was observed that LR4 model showed the best values of all performance measures as higher values of R (0.933), d(IA) (0.964) CE (0.870) and lower values of RMSE (0.701), MAE (0.532), MAPE (11.609) followed by the performance measures of LR3 and LR1 model while the LR2 model showed lower performance among them having R (0.902), d(IA) (0.946), CE (0.813), RMSE (0.841), MAE (0.646) and MAPE (14.274). However, it was also observed that the results of all performance measures for all LR models varies in the range as R (0.902 to 0.933), d(IA) (0.946 to 0.964), RMSE (0.701 to 0.841), MAE (0.532 to 0.646), MAPE (11.609 to 14.274) and CE (0.813 to 0.870) and showed the performance in sequence of LR4, LR3, LR1 and LR2. It indicates that all LR models performed satisfactorily and showed marginal difference of performance measures among them in development stage. It also indicates from LR2 to LR4 model that as the number of independent variables (those required in Penman-Monteith method) increases the performance of models increases.

The results for the validation of all LR models. It reveals that all models showed numerically at par results for each performance measures in validation stage, however sequence for performance of models slightly change as LR1, LR4, LR3 and LR2 than that of development stage. In LR1 model the value of the R in development stage is 0.921 and it enhances to 0.951 in validation stage. Similar kind of enhancement also occurred in d(IA) (0.958 to 0.969) and CE (0.849 to 0.895) during validation of LR1 model. It was also observed that there were reduction in the values of RMSE, MAE, MAPE for LR1 model during validation as 0.756 to 0.543, 0.544 to 417, 12.499 to 11.291 respectively. It was observed that LR1 model showed better performance in validation stage than development stage. However, it has shown close difference in enhancement and reduction of each performance measure. Similar kind of close difference for each performance measure occurred during development and validation stage of remaining LR models. It indicates that all LR models were validated satisfactorily and generalized for estimation of ETo. Overall, the performance suggest that all LR models can be an acceptable approach to predict ETo values for Dhule station as per data availability. The most of researchers [2,10-12] also found that linear regression models with varying independent variable can be adopted for prediction of reference evapotranspiration.

The [Fig-1] represents comparison of ETo values of Penman-Monteith method with those of LR1 model which has pan evaporation as independent variable. The [Fig-2] to 4 represents comparison of ETo values of Penman-Monteith method and those of predicted by LR2, LR3 and LR4 which have independent parameters as those required in Penman-Monteith method. It was observed that as the number of independent variables increases the difference between predicted and observed values of ETo were decreases [Fig-2 to 4]. However, the difference shown by [Fig-1] has irrespective of the trend due to the pan evaporation as independent variable. Evaporation from pan provides a measurement of a combined effect of temperature, humidity, sunshine hours, and wind speed on the reference crop evapotranspiration [4].

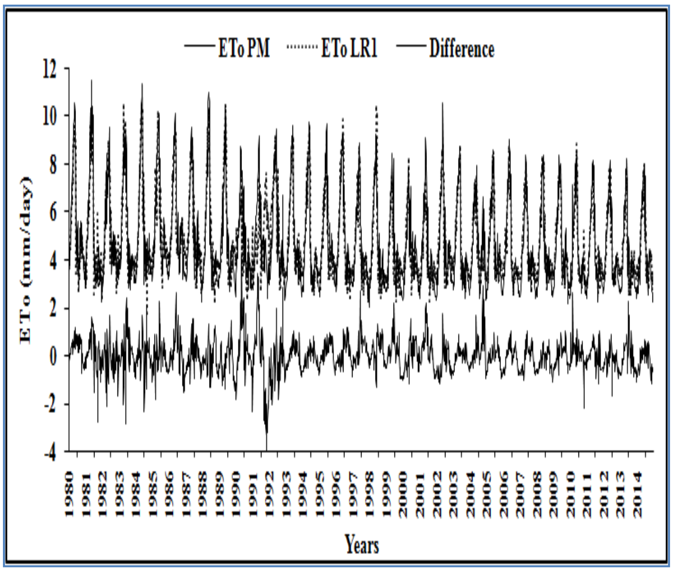


**Fig-1 Comparison of the values of ETo by Penman Monteith and LR1 model (Epan) for Dhule station**

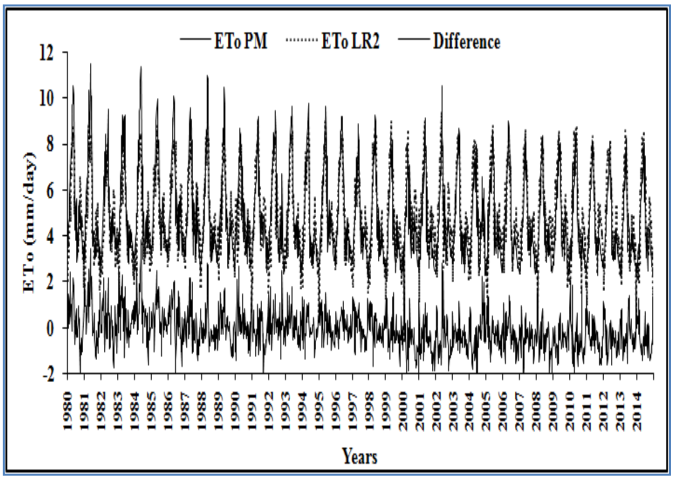


**Fig-2 Comparison of the values of ETo by Penman Monteith and LR2 model(Tmax and Tmin) for Dhule station**

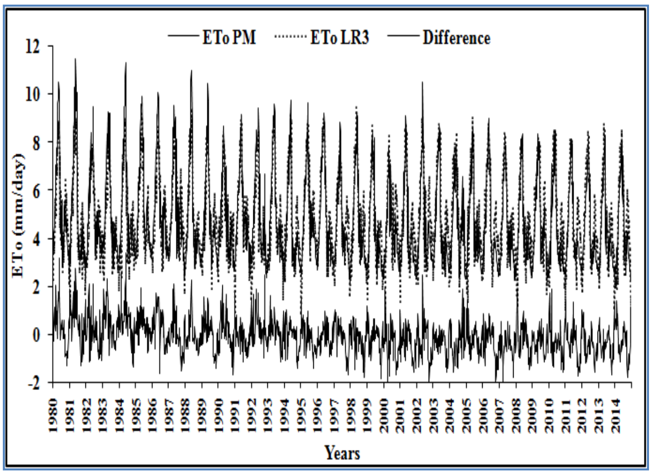


**Fig-3 Comparison of the values of ETo by Penman Monteith and LR3 model (Tmax, Tmin and SSH) for Dhule station**

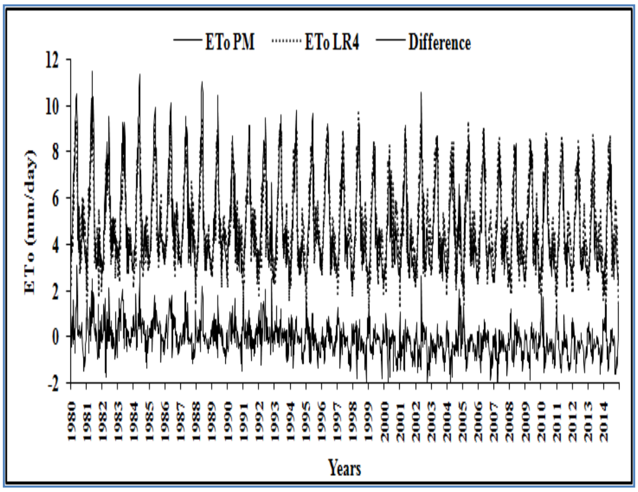


**Fig-4 Comparison of the values of ETo by Penman Monteith and LR4 model (Tmax, Tmin, RHmax, RHmin and SSH) for Dhule station.**

## Conclusion

The four linear regression ETo models were developed using FAO-56 Penman-Monteith method with varying climatic parameters for the Dhule region. The performance of linear regression models developed was verified based on the evaluation criteria *viz.,* coefficient correlation, index of agreement, root mean square error, mean absolute error, mean absolute percentage error and coefficient efficiency. All linear regression models performed satisfactorily and showed marginal difference of performance measures among them in development and validation stage. It reveals that all linear regression models were performed satisfactorily and generalized for estimation of ETo. Overall, the performance suggest that all LR models can be an acceptable approach to predict ETo values for Dhule region. Therefore, linear regression models are an alternatives to Penman-Monteith method under limited data availability for Dhule region or also the other regions of similar climatic conditions for satisfactory ETo estimation

**Acknowledgement / Funding:** We are thankful and acknowledged to Indian Meteorological Department, Pune and MPKV, Rahuri, Maharashtra (India) for availing meteorological data.

**Author Contributions:** Data collection, analysed, inferred and conclusion

**Abbreviations:** Linear regression (LR),maximum temperature (Tmax), minimum temperature (Tmin), maximum relative humidity (RHmax), minimum relative humidity (RHmin) and bright sun shine hours (SSH), correlation coefficient (R), index of agreement d(IA), coefficient efficiency (CE) and the lowest values of root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE).

**Ethical approval:** This article does not contain any studies with human participants or animals performed by any of the authors.

**Conflict of Interest: None declared**

**References**

[1] Allen R.G., Pereira L.S., Raes D. and Smith M. (1998) Crop Evapotranspiration, Guideline for Computing Crop Water Requirements. FAO Irrigation and Drainage, Paper 56.FAO Rome.
[2] Bogawski Paweł and Ewa Bednorz (2014) *Water Resour Manage,* 28, 5021–5038
[3] Dehbozorgi Fatemeh and Ali Reza Sepaskhah (2012) *Archives of Agronomy and Soil Science,* 58(5), 477-497.
[4] Doorenbos J. and Pruitt W.O. (1977) Guidelines for predicting Crop Water Requirement: FAO Irrigation and Drainage Paper No.24, FAO, Rome, Italy: 144 -156.
[5] George B.A., Reddy B., Raghuwanshi N.S. and WWallender W. (2002) *Journal of Irrigation Drainage Engineering,* 128(1), 1–10.
[6] Huo Z., Feng S., Kang S. and Dai X. (2012) *Journal of Arid Environments,* 82 , 81-90
[7] Michael A.M. (2008) Irrigation Theory and Practices Second edition, Vikas publishing house pvt.ltd., pp 479.
[8] Nandagiri L. and Kovoor G. (2006) *Journal of Irrigation and Drainage Engineering,* 132, 238–249.
[9] Naidu G., Bogayya K.V., Siva Kumar Babu and Srinivasulu V. (2015) *International Journal of Civil Engineering and Technology,* 6(11), 71-75.
[10] Reddy K. ChandraSekhar, Aruna Jyothy S. and Mallikarjuna P. (2010) *Proc. of Int. Conf. on Advances in Civil Engineering,* 83-88.
[11] Sriram A.V. and Rashmi C.N. (2014) *Journal of Mechanical and Civil Engineering*, 11(2 IV), 65-70
[12] Tabari Hossein, Ozgur Kisi, Azadeh Ezani and Hosseinzadeh Talaee P. (2012) *Journal of Hydrology,* 444-445, 78-89.
[13] Temesgen B., Eching S., Davidoff B. and Frame K. (2005) *Journal of Irrigation Drainage Engineering,* 131(1),73–84.
[14] Trajkovic S. (2007) *Journal of Irrigation Drainage Engineering,* 133(1), 38–42.
[15] Xu C.Y. and Singh V.P. (1998) *Hydrological Processes,* 12, 429–442.