

# Ascertaining Predictability Cognizance for the Prediction of Reference Evapotranspiration

Saravanan Poomalai, Sivapragasam Chandrasekaran



**Abstract:** Reference evapotranspiration ( $ET_0$ ) is a rudimental variable in the estimation of crop water requirement, and preparation of irrigation schedule. Prediction of  $ET_0$  is a necessitous one for estimation of crop water requirement in future time step. In this paper  $ET_0$  is predicted using Artificial Neural Network (ANN) by different inputs Like Temperature, Cloud cover, Vapor pressure, Precipitation and its combinations by various models. Before prediction, the predictability of all the input time series is calculated individually and the effect of predictability on prediction is analyzed in models having single predictor. In spite of inserting additional predictor in input, the reason for increase of Root mean squared error is justified in terms of predictability in the models having multiple predictors. Also it is seen that the performance of models with multiple predictors is better when compared to single predictor models in the estimation of  $ET_0$ .

**Keywords :** Artificial Neural Network, Predictability, Hurst exponent, Reference evapotranspiration

## I. INTRODUCTION

Reference evapotranspiration ( $ET_0$ ) is evapotranspiration from a reference surface having theoretical crop of specified features with enough water [1]. Hence, the factors affecting  $ET_0$  are climatic parameters only. The prediction and estimation of  $ET_0$  must be precise for the exact reckoning of crop water requirement, scheduling irrigation, irrigation water management and water budgets [2]. Generally, computation of  $ET_0$  of future time step with the use of present climatological condition is represented as prediction, while computation of  $ET_0$  of present time step with the use of present climatological condition is estimation.

There are many conventional methods available for the estimation of  $ET_0$ , of which, FAO Penman method is most widely used [3]. Climatological parameters such as air temperature, relative humidity, wind speed, and sunshine are required in estimation of  $ET_0$  using FAO Penman method. Prediction of  $ET_0$  has also received wide attention.

If all the parameters are available for estimation of  $ET_0$  by Penman method, it can be estimated and the estimated past  $ET_0$  can be used to predict future  $ET_0$  by ANN [4, 5]. Chauhan and Shrivastava[4] used past  $ET_0$  as input for ANN prediction of future  $ET_0$  by three different algorithm and found best suitable algorithm for ANN prediction. Trajkovic et al., [5]

also used past  $ET_0$  as input for ANN prediction of future  $ET_0$  and reported ANN is a reliable technique for the prediction of  $ET_0$ . ANN is also reported to be successful in many meteorological and hydrological studies such as runoff modeling, improvement of spatial interpolation of rainfall, bias correction of projected time series, modeling flash flood etc. [6, 7, 8, and 9]. Many authors worked on improving the accuracy of prediction in various Artificial intelligence techniques. For instance, C.Sivapragasam and Shie-Yui Liong [10] improved prediction accuracy by identifying optimal training data set.

Instead of initiating prediction without any strategy, researchers have also proposed ascertaining the cognizance of predictability before prediction [11]. Hurst exponent (H) of the time series represents the predictability of the time series [12]. The exponent is the display of long term memory in the time series [13]. Khalili *et al* [14] used the predictability concept in the prediction of rainfall. Sivapragasam et al., [15] analyzed the truthfulness of the concept of predictability.

Researchers used the knowledge predictability of a predictor in forecast, in cases of predictors and predictand are same (for instance, temperature of present time step is used for the prediction of future temperature) [11, 12, 14]. But in author's knowledge, there is no work has been done as regards the effect of predictability of predictors, if predictors and predictand are unlike (for instance, temperature of present time step is used for the prediction of future precipitation). In this paper, the applicability of predictability concept on the latter case is analyzed systematically.

In addition to the above, there are eight models are developed for estimation of  $ET_0$ . Out of eight models, the reason for better prediction of some models are discussed.

## II. METHODS AND DATA

### A. Hurst Exponent

Harold.E.Hurst, a British Engineer introduced Rescaled range analysis method and estimation of Hurst exponent to understand the patterns in the data. He used Hurst exponent value in the forecast of Nile River flooding.

The value of Hurst exponent value spans between 0 to 1. If the Hurst exponent value is closer to 0.5, it means, there is no long memory effect available in the time series and the predictability of the time series is very low. If the Hurst exponent value is far away from 0.5 (either close to 0 or 1), it means, there is some long memory effect available in the time series and the predictability of the Hurst exponent is higher. This is the concept of Hurst exponent.

Manuscript published on 30 December 2019.

\* Correspondence Author (s)

Saravanan Poomalai,\* Center for Water Technology, Department of Civil Engineering, Kalasalingam Academy of Research and Education, Srivilliputtur, India. Email: p.saravanan@klu.ac.in

Sivapragasam Chandrasekaran, Center for Water Technology, Department of Civil Engineering, Kalasalingam Academy of Research and Education, Srivilliputtur, India. Email: sivapragasam@klu.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

The method for the calculation of Hurst exponent is given below which is based on Peyghami and Khanduzi 2012 [11]:

The Rescaled range of different sizes of time series are calculated after arranging the time series into various sizes (length).

$$\left(\frac{R}{S}\right) = \frac{R_n}{SD_n} \quad \text{for } n = 1, 2, 3, \dots, t \quad (1)$$

Where,

(R/S) = the rescaled range for the given size

SD<sub>n</sub> = the standard deviation for the given range

$$R_n = [\max(y_1, y_2, \dots, y_n) - \min(y_1, y_2, \dots, y_n)] \quad (2)$$

for  $n = 1, 2, \dots, t$

Where:

R<sub>n</sub> = the widest spread in each range

$$y_n = \sum_{i=1}^n (x_i - m) \quad \text{for } n = 1, 2, 3, \dots, t \quad (3)$$

y<sub>n</sub> = the new time series adjusted for deviations from the mean

x<sub>i</sub> = any value in the given range

m = the mean for the range

After calculation of the rescaled range of different sizes, the logarithmic values are to be calculated for the size of each section and for each section's rescaled range. Then, logarithm of the size (x axis) of each series against the logarithm of the rescaled range (y axis) is to be plotted. Slope of this line is called 'Hurst Exponent'

### B. Artificial Neural Network

Artificial Neural Network (ANN) is a kind of Artificial intelligence technique roused from the behaviour of human nerve cells. Like the nerve cells of human, neurons of ANN are connected by synapses or connections. The neurons in the first layer get information from inputs and pass the information to the neurons of the hidden layer where the information is processed and send to the neurons of the last layer. At first, arbitrary weights are doled out in all the connection. The learning of ANN occurred by change of weights in connections. Design of ANN architecture is first in the learning process. Then, number of epochs should be chosen, afterwards, grouping of data sets need to be done. As the grouping of data sets to employ for training, testing and validation is very much influences the proficiency of the model, it should be chosen admirably. Since, the processing of input in the hidden layer is relies upon the activation function, the decision of activation function is very significant. Many strategies of algorithm can be utilized in the learning of ANN. In that, the feed-forward back propagation algorithm is utilized in this investigation. During learning the aggregated errors of all the nodes are determined and by utilizing this determined error, the weights in the connections are acclimated to limit the error. Subsequent to numerous rounds of this procedure, an ideal architecture is formed with negligible error, which can be used to validation set. More information about ANN can be seen in [16, 17, and 18].

### C. Data Set

The historical record of monthly time series for 20 years duration from 1983 to 2002 for temperature, cloud cover, Vapor pressure, Precipitation, and Reference evapotranspiration of Viruthu nagar, Tamil nadu, India are obtained from India Water portal Website.

## III. DEVELOPMENT OF ANN MODELS

### A. Prediction models

There are eight models developed for prediction of monthly reference evapotranspiration. Since, the aim is the prediction of reference evapotranspiration, the output is reference evapotranspiration of the month ahead (ET<sub>0 t+1</sub>) in all the eight models. The inputs are current month temperature (T<sub>t</sub>), current month cloud cover (C<sub>t</sub>), current month vapor pressure (V<sub>t</sub>), current month precipitation (P<sub>t</sub>), current month reference evapotranspiration (ET<sub>0</sub>) in the models M1P, M2P, M3P, M4P, and M8P respectively. T<sub>t</sub> and C<sub>t</sub> are the inputs for model M5P; T<sub>t</sub>, C<sub>t</sub>, and V<sub>t</sub> are the inputs for the M6P model; T<sub>t</sub>, C<sub>t</sub>, V<sub>t</sub>, and P<sub>t</sub> are the inputs for M7P model.

Out of available 240 data, 42% is used for training and another 42 % is used for testing and the remaining 16 % of data is used for validation. For all the eight models, standard three layered feed forwarded back propagation algorithm is used. 11 neurons in the hidden layer and logistic activation function are set to obtain optimal network by trial and error method. The inputs and output of all the models are given in Table 2.

### B. Estimation models

Estimation models: There are seven models developed for the estimation of monthly reference evapotranspiration. Since the aim is estimation of reference evapotranspiration, the output is reference evapotranspiration of the current month (ET<sub>0t</sub>) in all the seven models. The inputs are current month temperature (T<sub>t</sub>), current month cloud cover (C<sub>t</sub>), current month vapor pressure (V<sub>t</sub>), current month precipitation (P<sub>t</sub>) in the models M1E, M2E, M3E and M4E respectively. T<sub>t</sub> and C<sub>t</sub> are the inputs for model M5E; T<sub>t</sub>, C<sub>t</sub>, and V<sub>t</sub> are the inputs for the M6E model; T<sub>t</sub>, C<sub>t</sub>, V<sub>t</sub>, and P<sub>t</sub> are the inputs for M7E model. Out of available 239 data, 42 % is used for training and another 42 % is used for testing and the remaining 16 % of data is used for validation. For all the eight models standard three layered feed forwarded back propagation algorithm is used. 11 neurons in the hidden layer and logistic activation function are set to obtain optimal network by trial and error method. The inputs and output of all the models are given in Table 3.

## IV. RESULTS AND DISCUSSIONS

Though the aim is to predict only the reference evapotranspiration, the predictability is estimated for all the 5-time series, because the predictability of other variables may also affect the prediction of ET<sub>0</sub>, when it is used as input in ANN either separately or combined with other variables [19]. From Table 1, it is seen that, the average temperature has very little predictability. Meanwhile, the predictability of vapour pressure and precipitation is moderate.

The predictability of cloud cover and reference evapotranspiration time series are very high compared to all other time series. Hence, it is expected that Reference evapotranspiration can be predicted without much endeavor.

**Table- I: Hurst exponent of different variables**

Variables	Hurst exponent
Average Temperature	0.5
Cloud Cover	0.31
Vapour Pressure	0.43
Precipitation	0.46
Reference Evapotranspiration	0.32

Out of eight prediction models, the models M1P, M2P, M3P, M4P, M8P have single predictor variable, which are,  $T_t$ ,  $C_t$ ,  $V_t$ ,  $P_t$  and  $ET_{ot}$  respectively and remaining models have combination of more than one predictor variable. In M1P, M4P, and M8P models, the Hurst exponent of the predictor variable are 0.5, 0.46, 0.32 respectively and the corresponding RMSE value are 0.103, 0.095, and 0.037 respectively. This is clearly evident that if the predictor time series has low predictability the corresponding RMSE is high and vice versa. But, the models M2P and M3P have Hurst exponent value of 0.31 and 0.43, though the corresponding RMSE are 0.056 and 0.053. This reveals that the Hurst exponent concept is working well in most of the cases but not always. In all the models having single predictor variable, a

relationship between the Hurst exponent of predictor time series and the corresponding RMSE can be seen. Hence, one can conclude that, the predictability of predictor variable also affects the prediction.

Generally, it is expected that, if the inputs of ANN model is increased, the prediction would be better. The above said hypothesis is ascertained by experiments. The model M1P, M5P, M6P, M7P have single, double, triple and quadruple predictor variables and the corresponding RMSE are 0.103, 0.038, 0.045 and 0.039, respectively. In M6P the total input variables are  $T_t$ ,  $C_t$ ,  $V_t$  whereas in M5P it is  $T_t$ ,  $C_t$  only. By the addition of Vapor pressure as extra input, the Model M6P is expected to produce better RMSE when compared to M5P. But instead better RMSE, it become worse. This may due to the poor predictability of Vapor Pressure time series. In contrast to the M6P case, addition of precipitation as extra input in M7P, the RMSE is reduced considerably, when compared to M6P. Though, the predictability of Precipitation is less, a meaningful relation exist between precipitation and reference evapotranspiration, which improved the prediction in the M7P. Hence, the decision of input variables cannot be decided only by the knowledge of predictability.

From Table 2 and Table 3, it is observed that wherever there are models of single input variables prediction is better than estimation and the cases of more than one input variable, estimation is better than prediction. This resembles better performance of FAO-Penman method (4 input variables) over Blaney-Criddle method (1 input variable) obtained by R.Kumar et al., [3].

**Table- II: Models developed for the prediction of  $ET_{ot}$**

Model	Inputs	Output	RMSE
M1P	$T_t$	$ET_{ot+1}$	0.103
M2P	$C_t$	$ET_{ot+1}$	0.056
M3P	$V_t$	$ET_{ot+1}$	0.053
M4P	$P_t$	$ET_{ot+1}$	0.095
M5P	$T_t, C_t$	$ET_{ot+1}$	0.038
M6P	$T_t, C_t, V_t$	$ET_{ot+1}$	0.045
M7P	$T_t, C_t, V_t, P_t$	$ET_{ot+1}$	0.039
M8P	$ET_{ot}$	$ET_{ot+1}$	0.037

**Table- III: Models developed for the estimation of  $ET_{ot}$**

Model	Inputs	Output	RMSE
M1E	$T_t$	$ET_{ot}$	0.032
M2E	$C_t$	$ET_{ot}$	0.122
M3E	$V_t$	$ET_{ot}$	0.072
M4E	$P_t$	$ET_{ot}$	0.135
M5E	$T_t, C_t$	$ET_{ot}$	0.015
M6E	$T_t, C_t, V_t$	$ET_{ot}$	0.011
M7E	$T_t, C_t, V_t, P_t$	$ET_{ot}$	0.013

## V. CONCLUSION

- a) Before prediction the knowledge of predictability plays major role in selection of predictable time series.
- b) A predictor with more predictability gives better prediction in most of the time in case of models having single predictor as input
- c) Knowledge of predictability may not be able to guide researcher in the selection of input variables in case of models having multiple predictors as input. In such cases, the knowledge of physical relation between predictor and predictand needs to be used to select inputs.
- d) In the estimation of reference evapotranspiration, the models having multiple inputs performs better than models having single input, though the predictability of single input is more.

## REFERENCES

1. R.G.Allen, L.S. Pereira, D.Raes, and M. Smith, Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome, 300(9), 1998,p.D05109.
2. J.Shiri, A.A. Sadraddini, A.H. Nazemi, O. Kisi, P. Marti, A.F. Fard, and G.Landeras. Evaluation of different data management scenarios for estimating daily reference evapotranspiration. Hydrology Research, 44(6),2013 pp.1058-1070.
3. R.Kumar, M.K. Jat, and V. Shankar, Methods to estimate irrigated reference crop evapotranspiration-a review. Water Science and Technology, 66(3),2012. pp.525-535.
4. S.Chauhan, R.K. Shrivastava. Reference evapotranspiration forecasting using different artificial neural networks algorithms. Canadian Journal of Civil Engineering, 36(9), 2009.pp.1491-1505.
5. S.Trajkovic, B. Todorovic, and M.Stankovic. Forecasting of reference evapotranspiration by artificial neural networks. Journal of Irrigation and Drainage Engineering, 129(6),2003 pp.454-457.
6. A.S.Tokar, and P.A.Johnson. Rainfall-runoff modeling using artificial neural networks. Journal of Hydrologic Engineering, 4(3), 1999.pp.232-239.
7. C.Sivapragasam, V.M. Arun, and D.Giridhar,. A simple approach for improving spatial interpolation of rainfalling using ANN. Meteorology and Atmospheric Physics, 109(1-2),2010. pp.1-7.  
P.Saravanan, C. Sivapragasam, M. Nitin, S. Balamurali, R.K. Ragul, S.S.Prakash, G.S. Ganesan, and V.V.Murugan, December. ANN-Based Bias Correction Algorithm for Precipitation in the Yarra River Basin, Australia. In International Conference on Theoretical Computer Science and Discrete Mathematics 2016.(pp. 362-370). Springer, Cham.
8. C. Dinu, R. Drobot, C. Pricop, and T.V.Blidaru. Flash-flood modelling with artificial neural networks using radar rainfall estimates. Mathematical Modelling in Civil Engineering, 13(3), 2017.pp.10-20.
9. C.Sivapragasam, and S.Y.LIONG. Identifying Optimal Training Data Set-A New Approach. In Hydroinformatics: (In 2 Volumes, with CD-ROM)2004. (pp. 1621-1628).
10. M.R.Peyghami, and R.Khanduzi. Predictability and forecasting automotive price based on a hybrid train algorithm of MLP neural network. Neural Computing and Applications, 21(1), 2012.pp.125-132.
11. B.Qian, and K.Rasheed, K. November. Hurst exponent and financial market predictability. In IASTED conference on Financial Engineering and Applications 2004. (pp. 203-209).
12. H.Tatli. "Detecting persistence of meteorological drought via the Hurst exponent." Meteorological Applications 22, no. 4,2015. 763-769.
13. N.Khalili, S.R. Khodashenas, K.Davary, M.M. Baygi, and F.Karimaldini. Prediction of rainfall using artificial neural networks for synoptic station of Mashhad: a case study. Arabian Journal of Geosciences, 9(13), 2016.p.624.
14. C.Sivapragasam, P. Saravanan, S. Balamurali, S.Sumila, S. Keerthi and A.F.Farveen. An Investigation on Relationship between Hurst exponent and Predictability of a Rainfall Time Series. Meteorological Applications.26,2019. (pp511– 519).
15. ASCE Task Committee. a. Artificial neural networks in hydrology. I: Preliminary concepts. J. Hydrol. Eng. 5:2000.115-123.
16. ASCE Task Committee. b. Artificial neural networks in hydrology. II: Hydrologic applications. J. Hydrol. Eng. 5:2000.124-137
17. Haykin S. Neural network: a comprehensive foundation, 2nd edn. Prentice Hall, New Jersey.1999
18. Rangarajan G, Sant DA. A climate predictability index and its applications. Geophys. res. lett. 24:1239-1242.1997.

## AUTHORS PROFILE



**Dr.C.Sivapragasam**, has completed his graduation and post-graduation from IIT-Roorkee and IIT-Delhi respectively. He completed his PhD from NUS, Singapore. He has 2 years of industry experience and more than 17 years of teaching experience. He has completed 3 sponsored research projects from various funding agencies and published more than 70 papers in peer reviewed journals and conferences.



**Mr.P.Saravanan** did his Masters in Hydrology and Water Resources Engineering from College of Engineering, Guindy, Chennai. and doing PhD at Kalasalingam Academy of Research and Education. He has published 10 International Journal papers.