

Evaporation Estimation from Climatic Factors

Kumar, P.¹, G. Rasul², D. Kumar³

Abstract

This study assessed the ability of two models, Local Linear Regression (LLR) and Artificial Neural Network (ANN) to estimate monthly potential evaporation from Pantnagar, US Nagar (India) which falls under sub-humid and subtropical climatic zone. Observations of relative humidity, solar radiation, temperature, wind speed and evaporation have been used to train and test the developed models. A comparison was made between the estimates provided by the LLR model and ANN model. An extensive review of different methods applied for estimation of evapotranspiration in the sub-continent has also been presented in the study. Results have shown that the models were able to learn well about the events they were trained to recognize. For ANN model, the correlation coefficient for the training period was 0.9311 and for the testing period ended up to 0.9236. The value of Root Mean Square Error (RMSE) for the training period was 1.070 and for the testing period it reached 0.9863. In case of LLR model, the correlation coefficient for the training period was 0.9746 and for the testing period it was 0.9273. The value of RMSE for the training mode was 0.6121 and for the testing period it was 1.5301. The results provided a lot of confidence to carry out further research for different applications in more diverse range of climatic zones to cover the entire region.

Keywords: evaporation, estimation, local linear regression, artificial neural network, correlation coefficient

Introduction

Evaporation refers to water losses from the surface of a water body to the atmosphere. Evaporation occurs when the number of moving molecules that break from the water surface and escape into the air as vapour is larger than the number that re-enters the water surface from the air and become entrapped in the liquid. A sizable quantity of water is lost every year by evaporation from storage reservoirs and evaporation of water from large water bodies influences the hydrological cycle. Among the hydrological cycle, evaporation is perhaps the most difficult to estimate due to complex interactions among the components of land-plant-atmosphere system, (Singh et al., 1997).

The most common and important factors affecting evaporation are solar radiation, temperature, relative humidity, vapour pressure deficit, atmospheric pressure, and wind. Evaporation losses should be considered in the design of various water resources and irrigation systems. In areas with little rainfall, evaporation losses can represent a significant part of the water budget for a lake or reservoir, and may contribute significantly to the lowering of the water surface elevation. Therefore, accurate estimation of evaporation loss from the water body is of primary importance for monitoring and allocation of water resources, at farm scales as well as at regional scales. The rate of evaporation depends on a number of meteorological factors such as solar radiation, air temperature, relative humidity, wind speed, and to some extent atmospheric pressure. Other factors are related to the nature of the evaporating surface and the quality of water. Various studies have been conducted to determine which of these factors have the dominant effect on evaporation. In addition to solar radiation, (Chow et al., 1988) claimed that the mechanism of transporting the vapor from the water surface has also a great effect. Vapor pressure deficit, temperature, barometric pressure, humidity, and wind speed were emphasized by (Singh, 1992) as the controlling factors. (Gupta, 1992) pointed out that relative humidity, wind velocity, and temperature of

¹ Pankaj591@gmail.com, Assistant Professor, Soil & Water Conservation Engineering, College of Technology GBPUA&T Pantnagar-246123, India.

² Chief Meteorologist, Pakistan Meteorological Department, Pitras Bukhari Road, Sector H8/2, Islamabad.

³ Professor & Head, Soil & Water Conservation Engineering and Dean College of Agri Business management, G. B. Pant University of Agriculture & Technology, Pantnagar-263145, US Nagar, Uttarakhand, India.

water and atmosphere are the climatic factors evaporation awfully depends on. In summary, it has been agreed that solar radiation, wind speed, relative humidity, and air temperature have attained special consideration as the most influencing factors by most researchers.

Evaporation and transpiration occur simultaneously therefore there is no easy way of distinguishing between the two processes. Instead of water quantity in the topsoil, the evaporation from a cropped soil is mainly determined by the fraction of the solar radiation reaching the soil surface. When the crop is small, water is predominately lost by evaporation from the soil surface, but once the crop is well developed and completely covers the soil, transpiration becomes the main process (Allen et al., 1996). Estimates of evapotranspiration provide an outlook of soil water balance in association with the amount of precipitation. Such estimates are of immense importance for calculation of water demand of the field crops and irrigation scheduling (Rasul, 1992). It also determines the nature of agro-climate a region has, agro-climatic potential of that region and suitability of crops or varieties, which can be grown successfully with the best economic returns (Rasul et al., 1993).

(Rasul, 2009) compared following four methods for estimation of potential evaporation and transpiration in different agro-climatic zones of Pakistan to determine the most reliable method to be used in different applications related to hydrology, climatology, agriculture, biodiversity, forestry, aquaculture, water balance and climate system dynamics. Every method has its own specific input but basically the meteorological parameters which are recorded at each standard meteorological observatory.

Radiation Method

It gave better performance for Rawalpindi during all the four seasons in conformity of FAO recommendations. It has already been motioned that Rawalpindi lies in sub humid zone representing vast potohar where mainly cultivation is exercised. The major deviation has been seen in dry summer (pre-monsoon) where it showed little overestimation. In rest of the three climatic zones which vary from Semi-Arid to Arid, the evapotranspiration estimates are highly erratic. During summer underestimation is very high as it reaches 20 % to 25 % less than pan evaporation, however, winter shows comparatively good results.

Modified Penman Method

Penman method was found overestimate evapotranspiration for low evaporative locations like Rawalpindi where summers are moderate and winters are relatively cooler. In Semi-Arid and Arid climates, this method practically failed to give representative results (Hashemi et al 1999). It highly underestimates evapotranspiration when compared with pan evaporation in the same climatic conditions. The deviation in summer season was above 35 % whereas during winter it was around 15 %. The results from this method may mislead the user under Pakistan's climatic conditions as about two-third of Pakistan lies under arid climate.

Blaney-Criddle Method

It shows comparatively closer estimates in variety of climatic conditions of Pakistan. It generally underestimates in summer and overestimates during winter, however the daily average variation range lie between + 1 and – 1 mm. The deviations are relatively large in hot and dry summer as compared to winter. The winter estimates of evapotranspiration practically tally with pan evaporation, however, it underestimates at the rate of 10 % to 15 % during summer. The peak deviation of 20 % was seen during hot and dry summer whereas it was lowered down to 10 % in monsoon season.

Fao Penman-Montith

The results obtained from the Food and Agri organization (FAO) Penman-Monteith in this comparative study may be graded as the best as far as diversified climatic conditions of Pakistan are concerned. It gave more or less similar results as from Blaney- Criddle in case of a sub-humid region represented by Rawalpindi. The computations in Semi-Arid and Arid climates show a little consistency in results. Arid region represented by Quetta, where agricultural practices are carried out

under rain fed conditions, indicate about 10 % underestimation in summer and 3 % overestimation of evapotranspiration in winter. On the other hand similar climatic zone but land irrigated with canal water represented by Tandojam gave altogether different results. In summer, evapotranspiration estimates were 10 % to 15 % lower and 5 % higher than pan evaporation during winter. Faisalabad representing Semi-Arid climate also behaves like Tandojam but amplitude of variations is much less during both the summer and winter seasons.

A large number of experimental formulae exist for evaporation estimation. There are direct and indirect methods available for estimating potential evaporation from free water surfaces. Because evaporation is an incidental, nonlinear, complex, and unsteady process, it is difficult to derive an accurate formula to represent all the physical processes involved. As a result, there are new trend in using data mining techniques such as artificial neural networks techniques to estimate evaporation.

The main objectives of this study were first to investigate the potential of using LLR and ANN models to predict evaporation as affected by climatic factors. Second, is to evaluate the performance of LLR and ANN models in estimating average monthly evaporation in Pantnagar.

Local linear regression (LLR)

LLR technique is a widely studied nonparametric regression method which has been widely used in many low dimensional forecasting and smoothing problems. To make a estimation for a given query point in input space local linear regression (LLR) first finds the k nearest neighbours of the query point from the given data set and then builds a linear model using these k data points. Finally the model is applied to the query point thus producing an estimated output. Consequently local linear regression using the k nearest neighbours (in the training data) of the query point can be accomplished quickly as by (Penrose, 1995) Thus local linear regression is a very fast and capable predictive tool. LLR is most effective in regions of the input space with a high density of data points. For few and far data points in the vicinity of the query point LLR model is not very effective if the underlying function to model is strictly non-linear. LLR model produces accurate predictions in the regions of high data density in input space, but it is predisposed to yield devious results for non-linear functions in regions of low data density. LLR does not generalize fine but is a good interpolative model for large amounts of data. The LLR procedure requires only three data points to obtain an initial prediction and then uses all newly updated data as they becomes available to make further estimation.

Artificial neural networks

ANN was first introduced as a mathematical aid by (McCulloch et al., 1943). They were inspired by the neural structure of the brain. Figure 1 is a general architecture of a Feed Forward ANN, with one hidden layer. Most ANNs have three layers or more: an input layer, which is used to present data to the network; an output layer, which is used to produce an appropriate response to the given input; and one or more intermediate layers, which are used to act as a collection of feature detectors. The ability of a neural network to process information is obtained through a learning process, which is the adaptation of link weights so that the network can produce an approximate output. In general, the learning process of an ANN will reward a correct response of the system to an input by increasing the strength of the current matrix of nodal weights.

There are several features in ANN that distinguish it from the empirical models. First, neural networks have flexible nonlinear function mapping capability which can approximate any continuous measurable function with arbitrarily desired accuracy, whereas most of the commonly used empirical models do not have this property. Second, being non-parametric and data-driven, neural networks impose few prior assumptions on the underlying process from which data are generated. Because of these properties, neural networks are less susceptible to model misspecification than most parametric nonlinear methods.

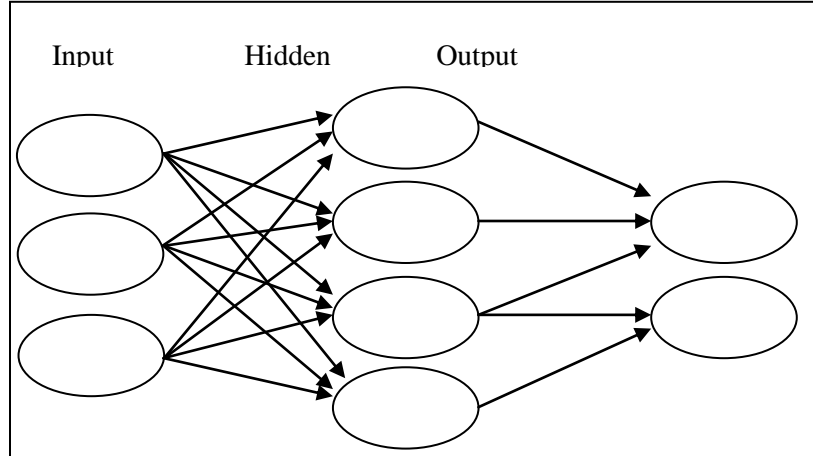


Figure 1: Architecture of multilayer feed forward neural network

There are a wide variety of algorithms available for training a network and adjusting its weights. In this study, an adaptive technique momentum Levenberg–Marquardt based on the generalized delta rule was adopted.

Let x_i ($i = 1, 2, \dots, n$) are inputs and w_i ($i = 1, 2, \dots, n$) are respective weights. The net input to the node can be expressed as

$$net = \sum_{i=1}^n x_i w_i \quad (1)$$

The net input is then passed through an activation function $f(\cdot)$ and the output y of the node is computed as

$$y = f(net)$$

Sigmoid function is the most commonly used nonlinear activation function which is given by

$$y = f(net) = \frac{1}{1 + e^{-net}} \quad (2)$$

Throughout all ANN simulations, the adaptive learning rates were used for increasing the convergence velocity. For each epoch, if the performance decreases toward the goal, then the learning rate is increased by the factor of learning increment. If the performance increases, the learning rate is adjusted by the factor of learning decrement.

Study Area

The weekly evaporation data for the year 1990 to 2009 (236 months) approximately 19 years and 8 months were collected from Meteorological Observatory, G.B. Pant University of Agriculture and Technology, Pantnagar, District Udham Singh Nagar, India. Pantnagar falls in sub-humid and subtropical climatic zone and situated in Tarai belt of Shivalik range, of foot hills of Himalayas. Geographically it is located at 29°N latitude and 79.29°E longitude and an altitude of 243.84 m above mean sea level. Generally, monsoon starts in the last of June and continues up to September. The mean annual rainfall is 1364 mm of which 80-90 percent occurs during June to September. May to June is the hottest months and December and January the coldest. The mean relative humidity remains almost 80-90 percent from mid-June to February end.

Materials and Methods

As far as the significance of individual meteorological parameters is concerned, the study revealed that the highest value of correlation coefficient and least value of root mean square error were obtained for evaporation with air temperature, followed by using wind speed and relative humidity (Table1). While the lowest correlation coefficient was obtained with sunshine hours, which mean bright sunshine hours alone does not appear to influence the evaporation significantly. The effect of air temperature, wind speed and sunshine hours was found to be positive; whereas a negative correlation exists between evaporation and relative humidity (that is evaporation decreases with increase in relative humidity).It is a natural fact that the climatic/meteorological factors in general act in concert. Therefore, it is pertinent to take into account the combined influence of all the meteorological parameter on evaporation. By various trials it was suggested that a combination of temperature, wind speed, sunshine hour and humidity provides a maximum value of correlation coefficient with minimum values of root mean square error in comparison to other inputs combinations.

Table 1: Statistical analysis of the total monthly weather data

S. No	Data	Maximum	Minimum	Correlation coefficient with evaporation
1	Air temperature (°C)	32.35	10.45	0.7625
2	Relative humidity (%)	89	38.5	- 0.640
3	Wind velocity (m/s)	14.2	0.7	0.6612
4	Sunshine hours (hour)	10.5	3	0.4931
5	Evaporation (mm)	13.1	1.1	1.00

The input combinations used in this application to estimate evaporation for Pantnagar station were Air temperature (°C), Relative humidity (%), Wind velocity (m/s) and Sunshine hours (hour) of a month t and Evaporation (mm) of that month t was considered as output of the models.

Different LLR architectures were tried using these inputs and the appropriate numbers of nearest neighbors. Various LLR models were tested for highest correlation coefficient and lowest root mean square error statistics. 157 data sets were used for training and 79 months data were used for testing for ANN models.

Different Feed Forward ANN architectures were tried using these inputs and the appropriate model structures were determined for each input combination. Then, the ANN models were tested and the results were compared by means of correlation and coefficient root mean square error statistics.

For best fit ANN model in present study multilayer perception with one hidden layer and with a sigmoid activation function was used as it works well for this data set. Other user-defined parameters used were momentum learning rate and step size = 0.1, momentum = 0.700, hidden layer nodes = 4 and iterations = 1000. These values were obtained after a number of trials by using different combination of these parameters carried out on data set.

Results and Discussion

The correlation coefficient and root mean square error values of developed model in the training period as well as in testing period are given in Table 2. It can be seen from the table that for ANN model the correlation coefficient for training period is 0.9311 and for testing period is 0.9236. The value of root mean square error for training period is 1.070 and for testing period it is 0.9863. For LLR model the correlation coefficient for training period is 0.9746 and for testing period is 0.9273 and value of root mean square error for training period is 0.6121 and for testing period it is 1.5301 respectively. It is clear from Table 2 that the higher values of correlation coefficients and lower values of root mean square error suggests the applicability of LLR model for evaporation estimation over the ANN model.

The observed and predicted monthly evaporation by ANN model (Figure 2) and LLR model (Figure 3) incorporates all the data points. There is a very strong correlation depicted in the output of both the models under discussion. Most of the points lie near the regression line with a few exceptions of greater deviations in the testing mode. However, the correlation above 90 % and standard error mainly less than 1 encourages this technique to be used in practical applications.

Table 2: Performance evolution of model on training and testing period

Models		Correlation coefficient	Root mean square error
Training	ANN	0.9311	1.070
	LLR	0.9746	0.6121
Testing	ANN	0.9236	0.9863
	LLR	0.9273	1.5303

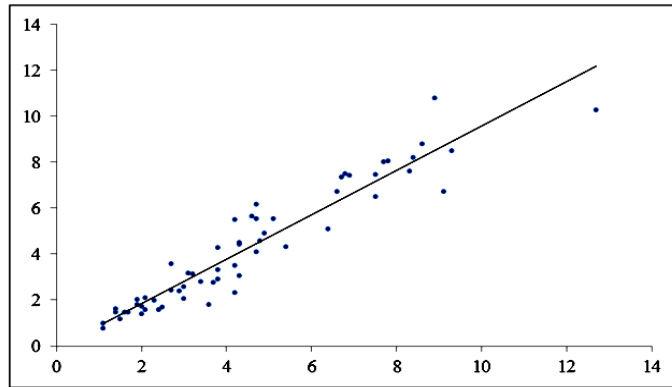


Figure 2: Observed and predicted monthly evaporation for ANN model in testing period

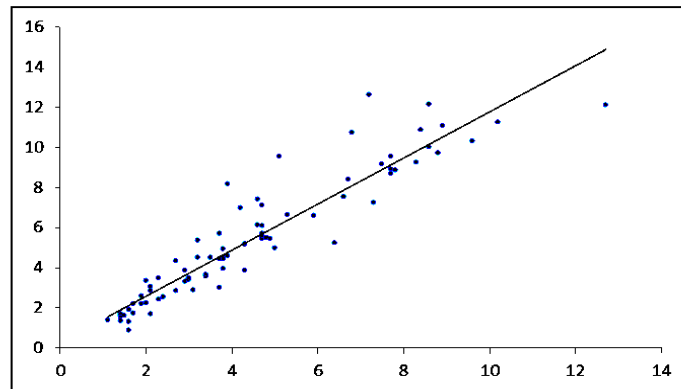


Figure 3: Observed and predicted monthly evaporation for LLR model in testing period

Conclusion

The present study discusses the application and usefulness of LLR and ANN based modeling approaches in estimation of evaporation losses over a region. The results are quite encouraging and suggest the usefulness of both modeling technique in estimation of the evaporation. In case of LLR model the values of correlation coefficient are better than ANN model for both training and testing periods. Root mean square error value for training period is better for LLR model and for testing period root mean square values of ANN model is found slightly greater than LLR model. Due to having less complexity and less computational efforts to execute LLR model is found appropriate overfeed Forward ANN. This study also

concludes that a combination of mean air temperature, wind speed, sunshine hour and mean relative humidity provides better performance in predicting the evaporation losses.

References

- Chow, V., D. Maidment and, L. Mays, 1988:** Applied Hydrology. N.Y.: McGraw-Hill Pub.
- Firat, M., M. Turan and M. A. Yurdusev, 2009:** Comparative analysis of fuzzy inference systems for water consumption time series prediction. Journal of Hydrology Vol. 374, p.235–241(6 pp).
- Gupta, B., 1992:** Engineering Hydrology. Jain, India: N.C.,
- Jones, A. J., 2004:** New tools in non-linear modeling and prediction. Computational Management Science, Vol1, p. 109–149(58 pp).
- Kisi, O., 2006:** Daily pan evaporation modelling using a neuro-fuzzy computing technique. Journal of Hydrology Vol 329, p. 636–646 (10 pp).
- Kissi, O., 2004:** River flow modeling using artificial neural networks. Journal of Hydrologic Engg. ASCE Vol9, p. 60–63(3 pp).
- McCulloch, W. S. and W. Pitts, 1943:** A logic calculus of the ideas immanent in nervous activity. Bulletin of Mathematical and Biophysics Vol5, p. 115-133(18 pp).
- Penrose, R., 1955:** A generalized inverse for matrices.Proceedings of the Cambridge Philosophical Society.Vol.51, p. 406-413 (19 pp).
- Rasul, G., 1992:** Water Requirement of Wheat Crop in Pakistan. J. of Engg. & App. Sci.Vol. 3 No. 1. pp. 65-79.
- Rasul, G. and A. B. Farooqi, 1993:** Water Requirement of Cotton Crop in Pakistan. J. of Engg. & App. Sci. Vol. 4 No. 2. pp. 154-165.
- Rasul, G. and A. Mahmood, 2009:** Performance Evaluation of Different Methods for Estimation of Evapotranspiration in Pakistan's Climate. Pak. J. of Met. Vol. 5, Issue 10. Pp 22-34.
- Singh, V., 1992:** Elementary Hydrology. NJ, U.S.A.: Prentice Hall Inc.
- Singh, V. P. and C. Y. Xu, 1997:** Evaluation and generalization of 13 mass transfer equations for determining free water evaporation. Hydrological Processes Vol 11, p.311-323(12 pp).
- Sudheer, K. P., A. K. Gosain, R. D. Mohan and S. M. Saheb, 2002:** Modelling evaporation using an artificial neural network algorithm, Hydrological Processes Vol.16, p. 3189-3202(13 pp).
- Wang, W. C., K. W. Chau, C. T. Cheng and Q. Lin, 2009:** A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of hydrology Vol. 374, p. 294-306 (12 pp).