"Environmental Engineering" 10th International Conference
Vilnius Gediminas Technical University
Lithuania, 27–28 April 2017
http://enviro.vgtu.lt

eISSN 2029-7092 / eISBN 978-609-476-044-0

Article ID: enviro.2017.024
DOI: https://doi.org/10.3846/enviro.2017.024

Evapotranspiration Prediction Using M5T Method and Ritchie Equation for St. Johns, FL, USA

Yunus Ziya Kaya[1], Mustafa Mamak[2], Fatih Üneş[3], Mustafa Demirci[4]

[1,2]Civil Engineering Department, Osmaniye Korkut Ata University, Osmaniye, Turkey
[3,4]Civil Engineering Department, Iskenderun Technical University, Iskenderun, Turkey
E-mails: [1]yunuszkaya@osmaniye.edu.tr; [2]mmamak@osmaniye.edu.tr;
[3]fatihunes66@gmail.com (corresponding author); [4]mustafa.demirci@iste.edu.tr

Abstract. Evapotranspiration (ET) estimation takes an important role in hydraulic designs and agricultural yield. Even it is non-negligible for hydraulic designers and irrigation engineers it is not clear enough to estimate or calculate ET because of direct and indirect parameters effects. In this study Solar Radiation (SR), Air Temperature (AT), Relative Humidity (RH) and Wind Speed (U) meteorological parameters are used to create a M5T model. 1158 daily RH, U, AT and SR records are used to create model and 385 daily values are used to test it. Data set is taken from St. Johns, Florida, USA weather station. The test set is also applied to the Ritchie empirical formula. M5T model and Ritchie formula Results are compared with daily ET records using determination coefficient. Determination coefficient is found 0.966 for M5T model and 0.913 for Ritchie formula. According to the determination coefficient, Mean Square Error (MSE) and Mean Absolute Error (MAE) statistics, it is understood that M5T method can be used for daily ET estimation effectively.

Keywords: evapotranspiration, ritchie, artificial intelligent, estimation, modelling.

Conference Topic: Environmental protection.

Introduction

Water loss due to evaporation and transpiration in an area is called as evapotranspiration (ET). If crop is small in a specific area main part of evapotranspiration is evaporation and if crop is well developed than main part of ET becomes transpiration (FAO 2017). Estimation of ET accurately is an important subject for irrigation engineers. Many methods have been proposed to estimate or calculate ET as defined by Brutsaert (1982) and Jensen *et al.* (1990). Generally aerodynamic equations and rational relationships give most accurate results for ET estimation (Jensen *et al.* 1990).

Artificial intelligence methods such as machine learning and data mining methods have been used for various fields including financial forecasting, controlling, pattern recognition and system modelling. This study focused on to estimate daily ET using M5T data mining method due to this method is in use for the estimation of nonlinear physical cases.

Solar Radiation (SR), Air Temperature (AT), Relative Humidity (RH) and Wind Speed (U) daily meteorological parameters are used for ET estimation. Data set includes 1543 daily data and it is downloaded from USGS website (USGS.gov | Science for a Changing World n. d.). Data set is divided as training and test sets randomly. Close to 25 percent daily data is chosen as test set. M5T model and Ritchie empirical formula results are evaluated according to this chosen test set.

M5T method and empirical formula results are compared using Mean Square Error (MSE), Mean Absolute Error (MAE) and Determination coefficient.

Methodology

M5T Method

This model is a type of binary decision trees which first introduced by Quinlan (1992) that gives the opportunity to the users work with quantitative data, differently than other decision trees. M5 model is in use for clustering and regression analyze. The model has linear regression functions at terminal (leaf) nodes. Building up the tree consists two phases (Solomatine, Xue 2004). First phase of the building up model is splitting data into subsets and creating decision tree which has a root node on top and connections between other decision nodes. The splitting gauge depends on handling the standard deviation of the class values and casting up the expected reduction in this error as a consequent of checking

each attribute at that node (Kişi 2015). The formula to calculate the standard deviation reduction is given by Pal and Deswal (2009) as:

$$SDR = sd(T) - \sum \frac{|Ti|}{|T|} sd(Ti). \quad (1)$$

In equation (1), T is a set of instances that gets at the node, Ti is the subset of instances that have the ith outcome of the potential set and sd is the standard deviation (Rahimikhoob *et al.* 2013); Frank *et al.* 1998).

Due to splitting stages, child nodes always have less standard deviation than their parent nodes and it means child nodes are simpler. This splitting process builds a huge tree and it may cause poor generalization. That is why tree must be pruned to have finally linear regression functions at terminal leafs. For further information about M5T method readers are referred to Quinlan (1992).

Ritchie Formula

Necessary parameters for Ritchie formula calculations are solar radiation and air temperature, and the equation is given below (Jones, Ritchie 1990);

$$ET = \alpha_1[3.87 \times 10^{-3} \times R_s(0.6T_{max} + 0.4T_{min} + 29)]. \quad (2)$$

At equation (2); "R_s" symbolized solar radiation, "T_{max}" and "T_{min}" represent maximum and minimum temperatures. "α_1" is a coefficient which is calculated as below;

$$if\ 5C^0 < T_{max} < 35C^0\ then\ \alpha_1 = 1.1 \quad (3)$$

$$if\ T_{max} \geq 35C^0\ then\ \alpha_1 = 1.1 + 0.05[T_{max} - 35] \quad (4)$$

$$if\ T_{max} < 5C^0\ then\ \alpha_1 = 0.01 \times \exp[0.18(T_{max} + 20)] \quad (5)$$

Evaluation Criteria

Results of M5T model and Ritchie equation are evaluated using MSE, MAE, and determination coefficient statistics. These values are calculated as follows:

$$MSE = \frac{1}{N} * \sum_{i=1}^{n} (fi - yi)^2 \quad (6)$$

$$MAE = \frac{1}{N} * \sum_{i=1}^{n} |fi - yi| \quad (7)$$

$$R = \frac{\sum_{i=1}^{n} [(x]_i - \text{EMBED Equation.3})(y_i - \text{EMBED Equation.3})}{\sqrt{\sum_{i=1}^{n} (x_i - \text{EMBED Equation.3})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \text{EMBED Equation.3})^2}} \quad (8)$$

where, "f_i" represents predicted values and "y_i" represents actual values for equations (6&7). "x_i" shows *i*th actual value, "y_i" shows *i*th predicted value, "$\bar{x}$"represents x_{mean} and "$\bar{y}$"represents y_{mean} at equation (8).

Results

Distribution graphs and scatter charts are drawn for M5T model results and Ritchie empirical formula.

Distribution of daily M5T model test set results against observed daily test set ET values are given by Fig. 1. According to the distribution, it is possible to see high correlation between both values.

Scatter chart of M5T model test set results against daily observed ET values are given by Fig. 2.

Determination coefficient for M5T model and observed values are calculated as 0.966.

Ritchie formula test set results distribution graph is given by Fig. 3.

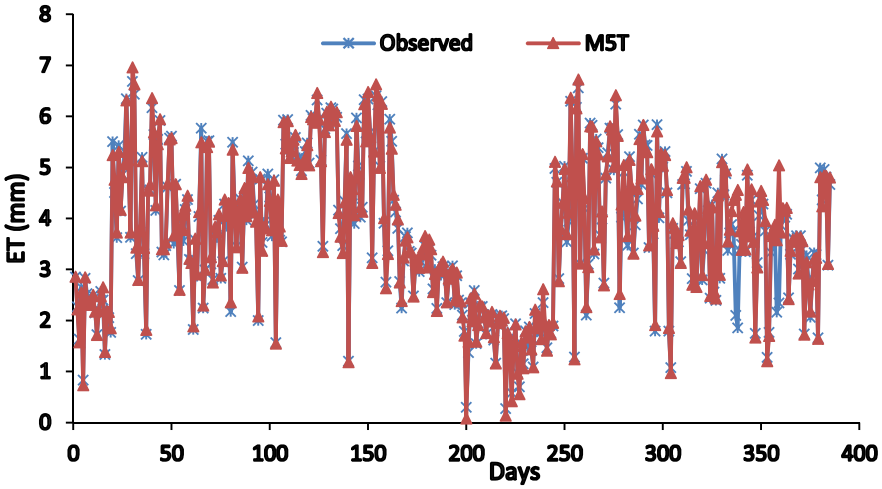

Fig. 1. Distribution graph of M5T method (Kaya *et al.* 2016)

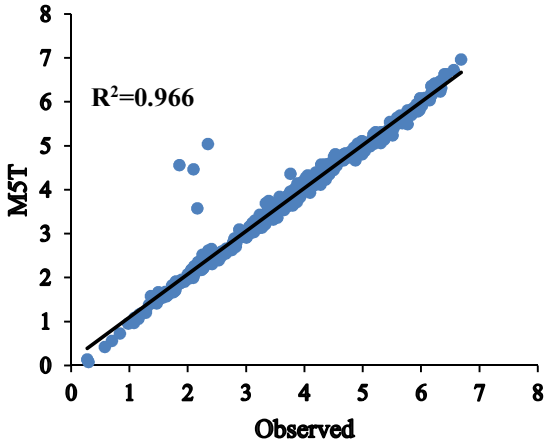

Fig. 2. Scatter chart of M5T method (Kaya *et al.* 2016)

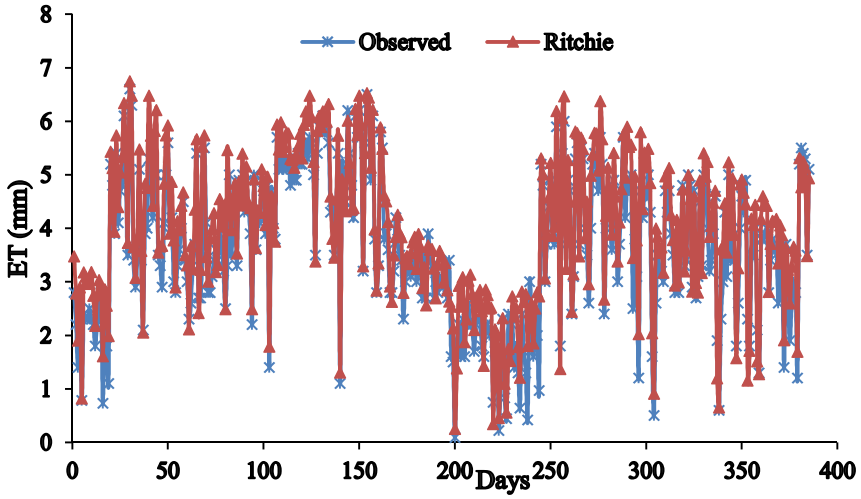

Fig. 3. Distribution graph of Ritchie formula

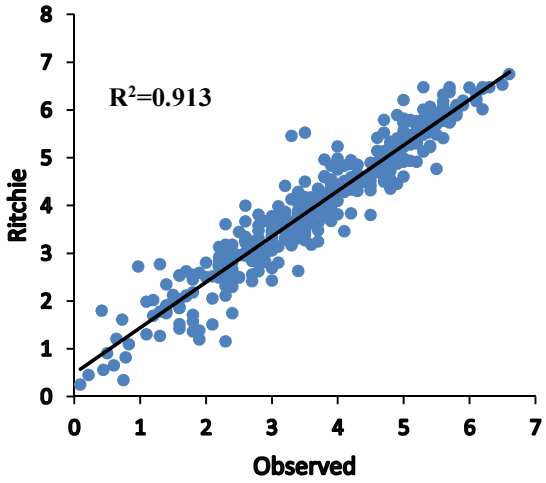

Fig. 4. Scatter chart of Ritchie formula

Scatter chart of Ritchie formula test set results against observed values is given by Fig. 4. Determination coefficient is calculated as 0.913.

Determination coefficient, MSE and MAE statistics are computed for comparison and all statistics are given in Table 1.

Table 1. Comparison statistics

Method	Parameters Used	Determination Coefficient	MSE	MAE
M5T (Kaya, Mamak, and Unes, 2016)	SR, AT, RH, U	0.966	0.071	0.115
Ritchie	AT, SR, U	0.913	0.274	0.415

Conclusion

Main target of this study is evaluating M5T model performance against Ritchie empirical formula for prediction daily evapotranspiration. Daily SR, AT, U, RH meteorological parameters are used for creating M5T model and determination coefficient is calculated as 0.966. Determination coefficient for Ritchie equation is computed as 0.913. According to the statistical calculations it is seen that M5T model has higher accuracy for daily ET estimation. Also it is understood that if there is just AT and SR meteorological parameters recorded in an area then it is possible to use Ritchie empirical formula for daily ET estimation. Authors suggest that M5T model should be applied to different regions having different climatic conditions to verify M5T ability on daily ET prediction.

References

Brutsaert, W. 1982. *Evaporation into the atmosphere: theory, history and applications*. Springer Netherlands. https://doi.org/10.1007/978-94-017-1497-6

FAO. 2017. "Chapter 1 – Introduction to Evapotranspiration" [online], [cited 12 January 2017]. Available from Internet: http://www.fao.org/docrep/x0490e/x0490e04.htm.

Frank, E.; Wang, Y.; Inglis, S.; Holmes, G. and Witten, I. H. 1998. Using model trees for classification, *Machine Learning* 32(1): 63–76. https://doi.org/10.1023/A:1007421302149

Jensen, M. E.; Burman, R. D. Allen, R. G. 1990. Evapotranspiration and irrigation water requirements, *ASCE Manuals and Reports on Engineering Practice No. 70*. Vol. 1.

Jones, J. W.; Ritchie, J. T. (Ed.). 1990. Crop growth models: management of farm irrigation system, in G. J. Hoffman, T. A. Howel, K. H. Solomon (Eds.). *ASAE monograph* No. 9, ASAE. St. Joseph, Mich., 63–89.

Kaya, Y. Z.; Mamak, M.; Unes, F. 2016. Evapotranspiration prediction using M5T data mining method, *International Journal of Advanced Engineering Research and Science* 3(12): 225–29. https://doi.org/10.22161/ijaers/3.12.40

Kişi, O. 2015. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, *Journal of Hydrology* 528. Elsevier B.V., 312–20.

Pal, M.; Deswal, S. 2009. M5 model tree based modelling of reference evapotranspiration, *Hydrological Processes* 23(10): 1437–1443. https://doi.org/10.1002/hyp.7266

Quinlan, J. R. 1992. Learning with continuous classes, *Machine Learning* 92: 343–48.

Rahimikhoob, A.; Asadi, M.; Mashal, M. 2013. A comparison between conventional and M5 model tree methods for converting pan evaporation to reference evapotranspiration for semi-arid region, *Water Resources Management* 27(14). Springer Netherlands: 4815–26.

Solomatine, D. P.; Xue, Y. 2004. M5 model trees and neural networks: application to flood forecasting in the upper reach of the Huai river in China, *Journal of Hydrologic Engineering* 9(6): 491–501. https://doi.org/10.1061/(ASCE)1084-0699(2004)9:6(491)

USGS.gov | Science for a Changing World [online]. n. d. [cited 12 January 2017]. Available from Internet: https://www.usgs.gov/.