"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.2017.094
DOI: https://doi.org/10.3846/enviro.2017.094

Evapotranspiration Estimation Using Support Vector Machines and Hargreaves-Samani Equation for St. Johns, FL, USA

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

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

Abstract. Information about Evapotranspiration (ET) calculations are not clear enough even it is an important part of hydrological cycle. There are many parameters which effect ET directly or indirectly such as Solar Radiation (SR) and Air Temperature (AT). In this study authors focused on the modelling ET using Support Vector Machines (SVM) method because this method has abilities to solve nonlinear problems. For the training SVM 1158 daily AT, SR, Wind Speed (U) and Relative Humidity (RH) meteorological parameters are used and model is tested using 385 daily parameters. Data set is taken from St. Johns, Florida, USA weather station. To understand the abilities of SVM for ET prediction against Hargreaves-Samani formula, the test set is applied to this empirical equation. Determination coefficient of SVM with observed daily ET values is calculated as 0.913 and determination coefficient of Hargreaves-Samani formula with observed daily ET is found as 0.910. Comparison between both methods is done using Mean Square Error (MSE), Mean Absolute Error (MEA) and determination coefficient statistics. As a result it is seen that SVM method is trustier than Hargreaves-Samani formula for daily ET prediction.

Keywords: evapotranspiration, hargreaves-samani, estimation, modelling.

Conference Topic: Water engineering.

Introduction

Evapotranspiration estimation takes a major role for irrigation management and hydraulic designs. Despite this major role, ET is not understood well enough (Brutsaert 1982). Evapotranspiration is a combination of evaporation and transpiration. When the crop is small in a stated area, evaporation is the main factor of ET but when the crop is well developed then the main factor is going to be transpiration (FAO n. d.). Therefore, to understand the ET mechanism, it is needed to understand evaporation and transpiration. Basically, evaporation takes place at the topsoil when the water is available and transpiration is the removal of vapour to the atmosphere from the crop tissues (Kişi 2007).

Difficulties of ET prediction come from nonlinear direct or indirect effects on ET such as SR, AT, RH meteorological variables. Hence, in the past decade some artificial intelligence and data mining methods are used to estimate daily ET, evaporation and pan evaporation. For instance M5T method is used to estimate Reference ET (Pal, Deswal 2009), artificial neural networks for ET prediction (Kumar *et al.* 2011) and generalized neural networks for ET prediction (Kişi 2006).

In this study, ET is estimated using Support Vector Machines (SVM) as a data mining method. SR, AT, RH, U daily meteorological parameters are used to train SVM model and test set results are obtained. SVM model results are compared with Hargreaves-Samani empirical formula test set results using determination coefficient, Mean Square Error (MSE) and Mean Absolute Error (MAE). Data set is downloaded from USGS website (USGS.gov n. d.) for the St. Johns FL, USA region.

Methodology

Support Vector Machines (SVM)

SVM is a data mining method which is in use for regression and classification. This method is described by Vapnik (Vapnik *et al.* 1996). It is possible to classify variables on a plane by drawing a borderline between them. The borderline which is drawn between variables must be as far as possible to each variable. Here it is SVM defines how to draw this borderline between variables group. SVM working principle is given by Figure 1.

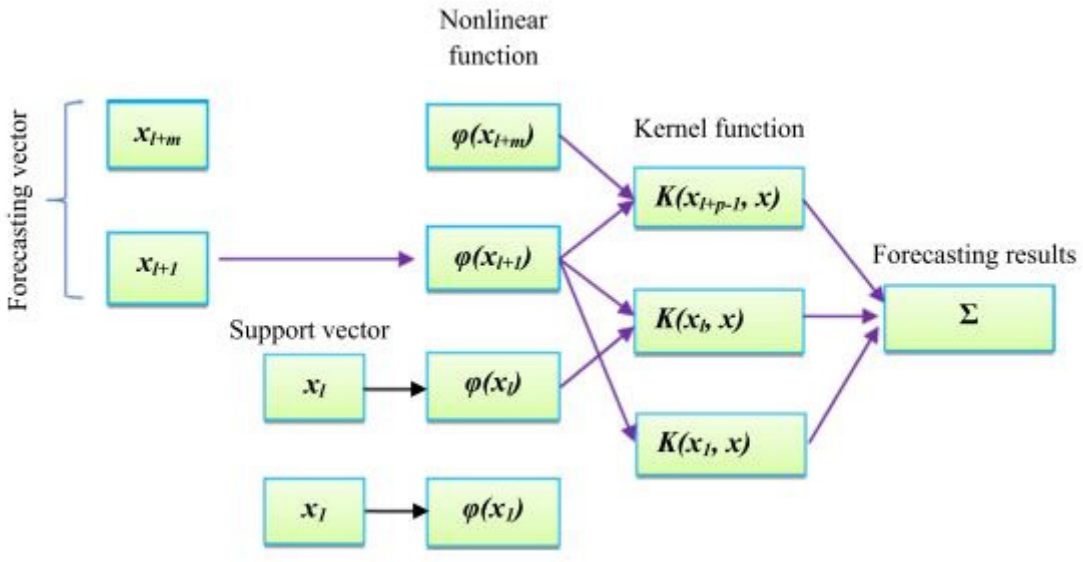

Fig. 1. SVM working principles (Kişi 2015)

Forecasting vector, support vector, nonlinear function, Kernel function and forecasting results are the stages of SVM working principle as given by Figure 1. For further information about Support Vector Machines (SVM) readers are referred to (Vapnik *et al.* 1996).

Hargreaves-Samani Formula

Necessary parameters for calculation of daily ET with Hargreaves-Samani equation are daily maximum temperature (T_{max}), daily minimum temperature (T_{min}) and extraterrestrial solar radiation (R_s) (Hargreaves, Samani 1985). The equation which is used for calculation is given below;

$$ET = 0.0135 \times 0.408\ \mathrm{Rs}\left(\mathrm{T} + 17.8\right), \tag{1}$$

where: *T* represents daily mean temperature and "Rs" extraterrestrial solar radiation in Hargreaves-Samani equation.

Evaluation Criteria

Mean Square Error (MSE), Mean Absolute Error (MAE) and determination coefficient statistics are calculated using equations (2, 3 & 4).

$$MSE = \frac{1}{N} * \sum_{i=1}^{n} \left(fi - yi \right)^2 ; \tag{2}$$

$$MAE = \frac{1}{N} * \sum_{i=1}^{n} \left| fi - yi \right| ; \tag{3}$$

$$R = \frac{\sum_{i=1}^{n} (x_i -).\left(y_i -\right)}{\sqrt{\sum_{i=1}^{n} \left(x_i -\right)^2} .\sqrt{\sum_{i=1}^{n} \left(y_i -\right)^2}}, \tag{4}$$

where: "f_i" represents predicted values and "y_i" represents daily observed values for equation (2&3) "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 (4).

Results

This study is focused on SVM method abilities on ET estimation. Hence, a SVM model is created by using SR, AT, RH, U daily meteorological parameters as input. Model test results are compared with Hargreaves-Samani empirical equation. Comparison is done with Mean Square Error (MSE), Mean Absolute Error (MAE) and determination coefficient statistics.

Distribution graph and scatter chart of SVM model and Hargreaves-Samani formula results are given separately.

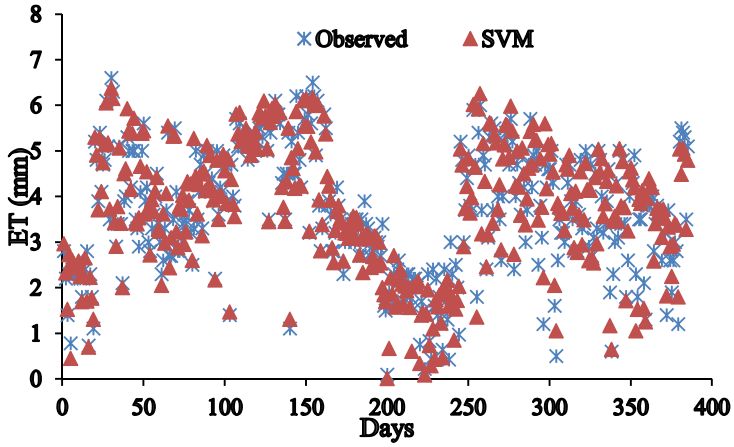

Fig. 2. SVM test results distribution graph

It is shown in Figure 2 that distribution of SVM model results is in same direction with observed daily values for test set.

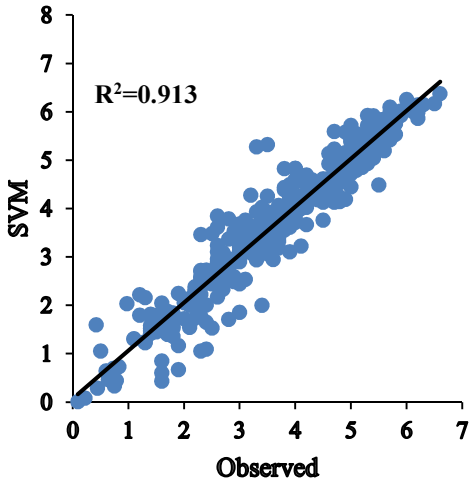

Fig. 3. SVM test results scatter chart

Determination coefficient of SVM method is calculated as 0.913 which is also given in Figure 3.

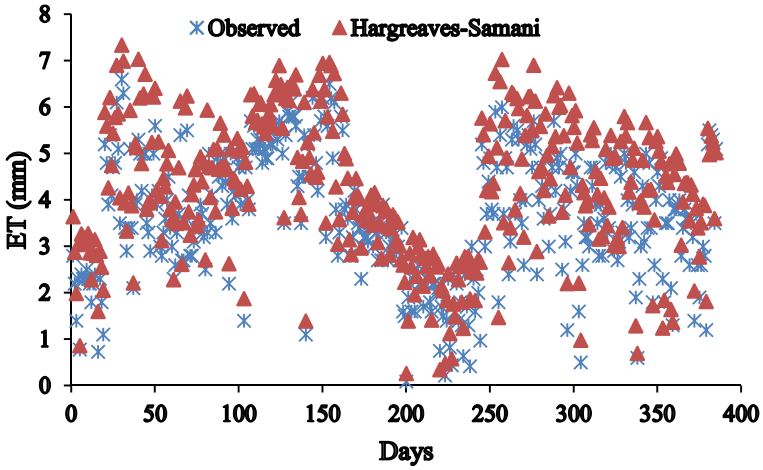

Fig. 4. Distribution graph of Hargreaves-Samani equation test results

Daily distribution graph of Hargreaves-Samani which is given by Figure 4 shows that there is high correlation between empirical formula and observed daily values.

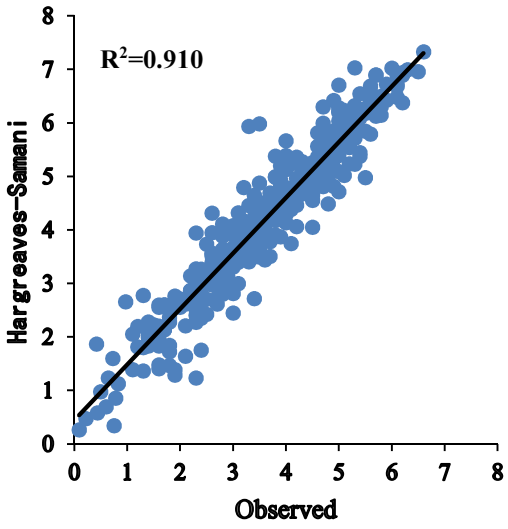

Fig. 5. Scatter chart of Hargreaves-Samani equation test set results

Scatter chart of Hargreaves-Samani empirical formula is given by Figure 5. Determination coefficient is calculated as 0.910 which that means determination coefficient is almost same as SVM model determination coefficient.

The determination coefficient, MSE and MAE values of SVM model and Hargreaves-Samani equation are given in Table 1. Although the determination coefficients of both methods seem to be equal, MSE and MAE values of SVM model are less than Hargreaves-Samani method.

Table 1. Comparison statistics

Method	Parameters Used	Determination Coefficient	MSE	MAE
SVM	SR, AT, RH, U	0.913	0.178	0.298
Hargreaves-Samani	AT, SR	0.910	0.550	0.635

Conclusion

Authors study on ET estimation using SVM model for St. Johns, FL, USA region. It is understood that SVM method gives quite close results to Hargreaves-Samani empirical equation results. MSE, MAE error calculations present that both method could be employed ET prediction successfully, but according to the error calculations it is clear that SVM method results are better than empirical equation results. This study is carried out for a particular region St. Johns, FL, USA. That is why authors proposed that SVM method should be applied to different regions to verify SVM method ability on ET estimation.

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. n. d. *Chapter 1 – Introduction to evapotranspiration* [online], [cited 01 December 2017]. Available from Internet: http://www.fao.org/docrep/x0490e/x0490e04.htm

Hargreaves, G. H.; Samani, Z. A. 1985. Reference crop evapotranspiration from temperature, *Applied Engineering in Agriculture* 1: 96–99. https://doi.org/10.13031/2013.26773

Kişi, O. 2015. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, *Journal of Hydrology* 528: 312–320. https://doi.org/10.1016/j.jhydrol.2015.06.052

Kişi, O. 2007. Evapotranspiration modelling from climatic data using a neural computing technique, *Hydrological Processes* 21: 1925–1934. https://doi.org/10.1002/hyp.6403

Kişi, Ö. 2006. Generalized regression neural networks for evapotranspiration modelling, *Hydrological Sciences Journal* 51: 1092–1105. https://doi.org/10.1623/hysj.51.6.1092

Kumar, M.; Raghuwanshi, N. S.; Singh, R. 2011. Artificial neural networks approach in evapotranspiration modeling: a review, *Irrigation Science* 29: 11–25. https://doi.org/10.1007/s00271-010-0230-8

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

USGS.gov. n.d. *Science for a changing world* [online], [cited 12 January 2017]. Available from Internet: https://www.usgs.gov/

Vapnik, V.; Vapnik, V.; Golowich, S. E.; Smola, A. 1996. Support vector method for function approximation, regression estimation, and signal processing, *Advances in Neural Information Processing Systems* 9: 281–287.