



# Estimation of daily evapotranspiration in Košice City (Slovakia) using several soft computing techniques

Yunus Ziya Kaya<sup>1</sup> · Martina Zelenakova<sup>2</sup> · Fatih Üneş<sup>3</sup> · Mustafa Demirci<sup>3</sup> · Helena Hlavata<sup>4</sup> · Peter Mesaros<sup>5</sup>

Received: 20 September 2020 / Accepted: 4 January 2021 / Published online: 4 February 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, AT part of Springer Nature 2021

## Abstract

Accurate estimation of evapotranspiration is one of the main aspects of water management. In this study, the capabilities of soft computing techniques for estimating daily evapotranspiration in Košice (Slovakia) were investigated. Daily solar radiation (*SR*), relative humidity (*RH*), air temperature (*T*), and wind speed (*U*) were the meteorological variables used for modeling. Based on the data, different combinations of multilayer perceptron (MLP), support vector regression (SVR), multilinear regression (MLR) models were generated. Model results are compared with each other and with the Hargreaves-Samani, Ritchie, and Turc empirical equations using three statistical criteria, namely mean square error (MSE), mean absolute relative error (MAE), and determination coefficient ( $R^2$ ). Of the empirical formulas applied, the Hargreaves-Samani equation gave the most compatible results with the Penman FAO 56 equation. Error percentage histograms were generated as a reference criterion. Model results show that the MLP model performs better than the other soft computing techniques used.

## 1 Introduction

The increase in world population and climate change ensure that water management issues remain topical and increase their importance day by day. Appropriate regional water management can be assured by considering all parameters of the hydrology cycle, among which evapotranspiration (ET) is probably the main component. Water loss due to ET needs to be calculated or estimated accurately for any efficient water management plan. However, it is not easy to determine ET (Rahimikhoob 2014). The United Nations Food and Agricultural Organization (FAO) set the Penman-Monteith FAO 56 equation as a standard formula for ET calculation;

however, this equation can be applied only in case that all the meteorological variables are recorded in a particular study area, otherwise some variables need to be estimated and the obtained results would be less accurate (Allen et al. 1998).

Potential evapotranspiration represents the upper limit of ET when this process is not limited by water deficit in the soil. Information about spatial and time distribution of potential evapotranspiration is of great importance in treating theoretical and practical problems of agriculture, forest and water management, and protection of the environment. Potential evapotranspiration in Slovakia is calculated according to the empirical or semi-empirical relationships based on measurements of other meteorological elements (Hlaváčiková and

✉ Martina Zelenakova  
martina.zelenakova@tuke.sk

Yunus Ziya Kaya  
yzkoku@outlook.com

Fatih Üneş  
fatih.unes@iste.edu.tr

Mustafa Demirci  
mustafa.demirci@iste.edu.tr

Helena Hlavata  
helena.hlavata@shmu.sk

Peter Mesaros  
peter.mesaros@tuke.sk

<sup>1</sup> Civil Engineering Department, Faculty of Engineering, Osmaniye Korkut Ata University, 80100 Osmaniye, Turkey

<sup>2</sup> Institute of Environmental Engineering, Faculty of Civil Engineering, Technical University of Košice, Vysokoškolská 2, 042 00 Košice, Slovakia

<sup>3</sup> Civil Eng. Department, Faculty of Engineering, Iskenderun Technical University, 31200 Iskenderun, Hatay, Turkey

<sup>4</sup> Slovak Hydrometeorological Institute, branch office, 041 17 Košice, Košice, Slovakia

<sup>5</sup> Department of Construction Technology and Management, Faculty of Civil Engineering, Technical University of Kosice, Košice, Slovakia

Novák 2013). Maps are constructed based on the data calculated using the Budyko-Zubenkova method in 31 climatological stations. The highest annual sums of potential evapotranspiration occur in the Podunajská nížina Lowlands and in the southern part of Slovakia generally, with more than 700 mm per year. Potential evapotranspiration decreases with increasing altitude. In the High Tatras, the northern part of Slovakia, this vertical gradient reaches 18 mm per 100 m per year. The annual course of potential evapotranspiration is very similar to the annual air temperature curve, reaching its maximum during the highest radiative balance (in July) and minimum in the winter (in December and January).

In the past decade, some soft computing techniques have been used in connection with nonlinear hydrological issues, such as evapotranspiration. Chen (2012) used least squares support vector machines to estimate daily reference evapotranspiration, and this type of soft computing technique was compared with the Penman-Monteith equation and artificial neural network models. Kaya et al. estimated ET at St. Johns, FL, USA, using the M5T method and Turc empirical formula (Kaya et al. 2016a). They used 1543 daily solar radiation, air temperature, relative humidity, and wind speed meteorological data readings in that locality. They indicated that the methods and empirical equations developed for ET estimation may have varying outputs with different characteristics of hydrological zones. Kaya et al. (2016b) used an adaptive neuro-fuzzy inference system and the Hargreaves-Samani empirical equation for prediction of daily ET. They calculated the correlation coefficient as 0.874 for the Hargreaves-Samani formula and 0.912 for the adaptive neuro-fuzzy inference system. Kim and Kim (2008) used another soft computing approach, namely neural networks and a genetic algorithm, for nonlinear evaporation and ET modeling. Kisi (2008) investigated the performance of different artificial neural network techniques on ET estimation using multilayer perceptrons, radial basis neural networks, and generalized regression neural network approaches for reference evapotranspiration estimation at the Pomona and Santa Monica weather stations in Los Angeles, USA. Among all the artificial neural network approaches he found that multilayer perceptrons performed better than the others for both stations. Kişi and Çimen (2009) modeled ET using support vector machines (SVM), and they reported that the SVM method can be used for reasonable ET estimation. In the past, some soft computing techniques were used together for pan evaporation estimation (Kisi 2015). Kisi and Zounemat-Kermani (2014) created two different adaptive neuro-fuzzy models for daily reference evapotranspiration estimation. Kumar et al. (2011) also used an artificial neural network method for ET modeling. Landaras et al. (2008) created artificial neural network models and compared their results with empirical and semi-empirical equation outputs for Northern Spain. They found that artificial neural network results are generally better than empirical and

semi-empirical approaches and can be used when it is not possible to make calculations with the Penman-Monteith 56 equation. Pal and Deswal (2009) used the M5T method for modeling daily reference ET. They used meteorological records as inputs and calculated reference evapotranspiration using a relation provided by the California Irrigation Management Information System as output. They compared the models' results with Penman-Monteith FAO 56 and calibrated Hargreaves-Samani equation empirical formulas. Tasar et al. (2018) used an artificial neural network approach to estimate evapotranspiration values. Üneş et al. (2018) used empirical equations and an artificial neural network for daily ET estimation and found that the artificial neural network performed better than the empirical approach. Rahimikhoob (2014) used M5T and neural networks for reference evapotranspiration prediction in an arid area. Üneş et al. (2015) predicted the Millers Ferry reservoir level using neural networks. Traore et al. (2010) used an artificial neural network for modeling reference ET in the Sudano-Sahelian zone. Finally, various other valuable soft computing studies have been employed to identify hydrological problems (Gavili et al. 2018; Gocić et al. 2015; Kumar et al. 2002; Mirás-Avalos et al. 2019; Yihdego and Webb 2018; Zanetti et al. 2007).

Evapotranspiration in Slovakia was studied for example by Hlaváčiková and Novák (2013), who compared the daily reference crop (grass cover) potential evapotranspiration results calculated with two modifications of the Penman-Monteith equation. Their results indicate significant differences in daily and seasonal potential evapotranspiration, stemming from differing net radiation and aerodynamic resistance estimation methods. Fendeková et al. (2018) studied drought occurrence using the Standardized Precipitation and Evapotranspiration Index (SPEI) and the Standardized Precipitation Index (SPI). Their results show that because of continuously increasing air temperature and balance evapotranspiration, there is ongoing decrease in runoff in the territory of Slovakia. Parajka et al. (2004) focused on spatial estimation of long-term mean annual actual (ET) and potential (EP) evapotranspiration in mountainous catchments in Central Slovakia. They compared three methods used for EP and ET estimations in a mapping framework: the modified empirical Turc model, the energy-based SOLEI model, and the continuous water balance simulation using the WASIM model. Hlaváčiková et al. (2004) developed a methodology for estimating monthly potential evapotranspiration (E0) as an input for hydrological balance modeling. Four different methods were used to calculate monthly potential evapotranspiration at six meteorological stations in the catchment studied: the Tomlain method (considered reference method) based on equations of energy balance; the FAO method based on equations of radiation balance and empirical parameters; and two empirical methods, the Thornthwaite and Ivanov

approaches. The genetic algorithm method was used for the calibration. The efficiency of E0 computed using each of the calibrated methods was compared with the Tomlin method, and the results were used for modeling the hydrological balance in the catchment.

This study investigated the abilities of support vector regression, multilayer perceptron, and multilinear regression to provide ET estimation for the city of Košice in eastern Slovakia. While different fields of water-related studies need estimated ET values on different timescales, such as daily, monthly, or annually, the current study aimed to estimate daily values, which could enable planners to estimate monthly or annual sums or averages values. It is known that ET is one of the main parts of the hydrology cycle, and it is crucial to estimate ET to make sustainable irrigation plans, design sustainable water supply systems, or carry out sustainable reservoir management, as the hydrological cycle defines the transformation of water. Thus, the study may be helpful for the management of water resources in the Košice area, as it will allow planners to develop future projections of water resource management/development plans for the region based on ET estimations.

## 2 Methodology

### 2.1 Study area

The area of study corresponds to the environs of the city of Košice (48.66° N, 21.24° E) in Slovakia (Fig. 1). Košice is located in the eastern part of the country, in the Hornad River Valley. It is adjacent to the Slovensky Kras karst area to the south-west, the Slovenske Rudohorie ore mountains in the north-west, and the Slanske vrchy volcanic hills in the east.

Košice has a humid continental climate with relatively severe winters, warm summers, and strong seasonality. The average annual temperature is 8.7 °C. Average monthly temperatures vary by 23 °C. Monthly temperature changes are given in Fig. 2. Total annual precipitation averages 605 mm. On average, there are 2072 h of sunshine per year. See the sunshine and daylight graphs to find monthly details including how high in the sky the sun reaches each month (SHMI 2015).

### 2.2 Datasets

The data for this study was provided by the Slovak Hydrometeorological Institute. For modeling of ET in the Košice area, we used 6497 daily data items recorded between 1995 and 2014. Five thousand two hundred twenty-two daily data items were used for training and the remaining 1276 daily data for testing. Approximately 20% of the total data was chosen as test set. Test set was between 1995 and 2010 and training set was between the years 2011 and 2014. Within the

total records, if a daily data has any of air temperature, relative humidity, wind speed, or solar radiation parameters missing, these daily records are removed from the dataset directly. No assumptions or estimations are used to complete the missing parameters. Extremum records of each parameter were eliminated from the dataset. The reason of elimination of extremum values was minimizing the effect of possible wrong measurements belonging to the used climatic variables. The dataset is a large dataset, and elimination of extremums and missing values did not make a big difference on total data statistically. Total removed values due the missing values and extremums were only approximately 10% of the whole used dataset. Penman-Monteith FAO 56 equation is used for obtaining the daily ET values, and these calculations are used as reference. No normalization was used in the dataset.

Minimum, maximum, mean, and standard deviation statistics are given in Table 1 and Table 2.

Firstly, Košice's daily ET values were calculated using the standard Penman-Monteith FAO 56 equation. As mentioned above, daily air temperature (max-min-mean), solar radiation, relative humidity, and wind speed variables are needed for this calculation. Hargreaves-Samani, Turc, and Ritchie empirical equations were also used for calculation, since they require fewer input variables. Then, different combinations of multilayer perceptron (MLP), support vector regression (SVR), and multilinear regression (MLR) models were created. Each method and the equations used are explained in the following sections.

### 2.3 Multilayer perceptron

Artificial neural networks (ANN) are systems consisting of process elements connected to each other with different weights, inspired by the nerve cell structure in the human brain. Rumelhart et al. advanced a theoretical framework for ANN systems (Rumelhart et al. 1986). Comparing ANN to traditional methods, we find one particular advantage in that they are not limited to the complexity of the structure of the phenomenon. The dataset needs to be divided into training and testing sets for model performance assessment when operating an ANN. MLP are feed-forward networks with a single hidden layer and a back-propagation algorithm (BPA), as shown in Fig. 3.

The MLP structure indicated in Fig. 3 has five inputs and one output.  $W_i$  ( $j,k$ ) stands for related weights, and B for biased. This MLP structure consists of one input layer, one hidden layer, and one output layer. More than one hidden layer is possible, but even possible MLP solutions imply that a single hidden layer is adequate for an MLP to unravel any complex nonlinear phenomenon (Cybenko 1989; Demirci et al. 2015; Hornik et al. 1989; Kisi 2005; Üneş et al. 2015). As it is important to have a sufficient number of hidden-layer node selections to achieve better network efficiency, the most



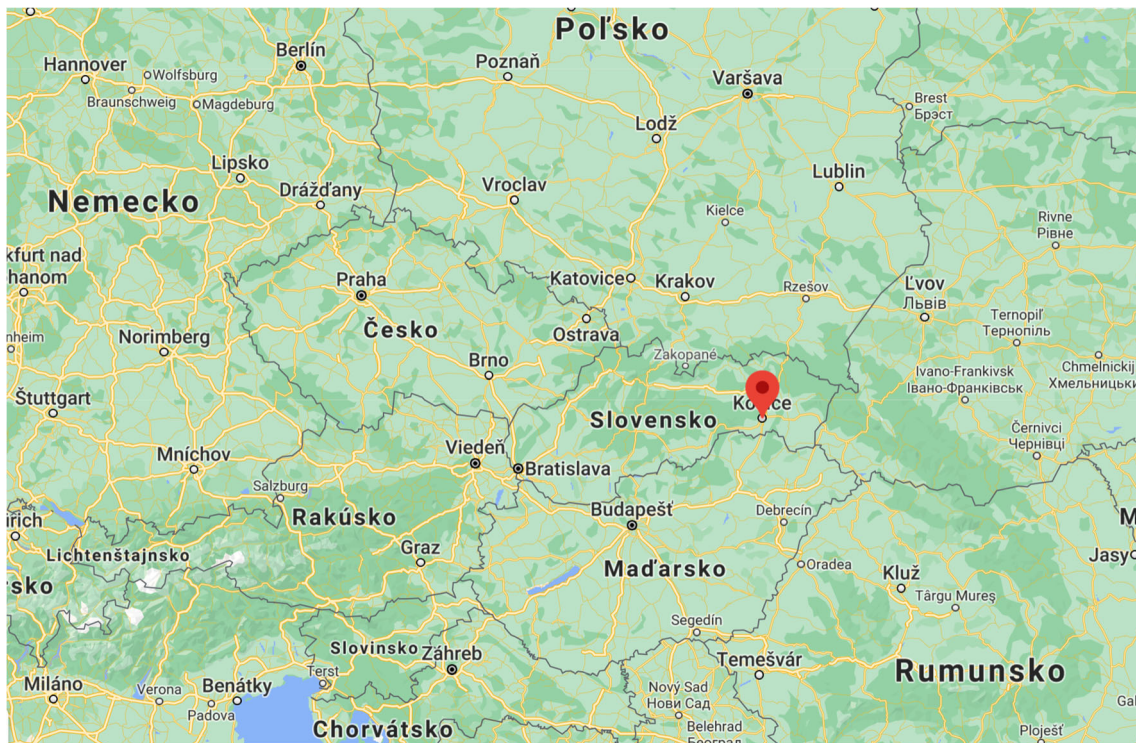
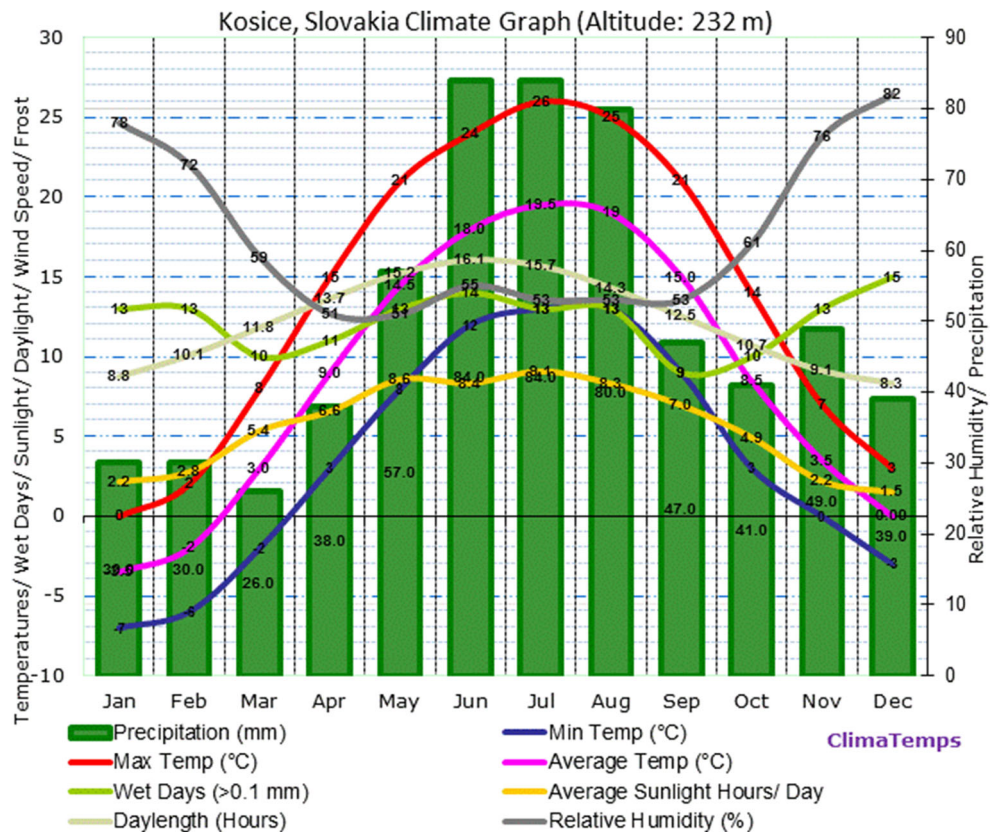


Fig. 1 Location of the study area (<https://www.google.com/maps>)

Fig. 2 Climate conditions of the Kosice City area (<http://www.kosice.climatemps.com/>)



**Table 1** Statistical information of the daily training set used for modeling soft computing techniques.

Parameter	Unit	Minimum	Maximum	Mean	Standard deviation
<i>T</i>	°C	- 13.9	28.7	10.5	8.8
<i>RH</i>	%	29.0	98.7	71.9	12.3
<i>U</i>	m/s	0.0	14.3	3.1	1.9
<i>SR</i>	MJ/m <sup>2</sup> /day	0.3	89.8	15.2	12.4
<i>ET</i>	mm	0.0	16.0	2.8	2.5

*T* air temperature (°C), *RH* relative humidity (%), *U* wind speed (m/s), *SR* solar radiation (MJ/m<sup>2</sup>/day), *ET* evapotranspiration (mm)

suitable hidden layer must be selected after several attempts, which is known as training. The first step in the process is the forward feed stage where the input values are connected to the hidden layer and the weights. The next step is the reverse propagation process which adjusts the weights. The next stage is the process of back propagation, which changes the weights in line with the differences between prediction and observation. MLP uses Bayesian regularization for training. The Bayesian regularization technique changes weights and bias values according to optimization of the Levenberg-Marquardt algorithm (Kışı 2004; Toprak and Cigizoglu 2008). Further details about ANN and MLP can be found in Bishop (1995) or Haykin (1999).

### 2.4 Support vector regression

Generalization capability and performance of support vector machines (SVMs) make them a popular and well-developed subject of machine learning. Since they were first introduced by Cortes and Vapnik in 1995, they have started being used in many different fields to analyze nonlinear problems (Cortes and Vapnik 1995). The performance of SVMs is generally better than that of neural networks on small datasets, because they proceed on the basis of minimization of structural risk instead of empirical risk minimization. Distinguishing two datasets as accurately as possible is the main purpose of the SVM method. Hyperplane or decision limits need to be determined to achieve this purpose. However, SVMs are not able to draw a linear hyperplane in a nonlinear dataset. To handle this disadvantage, Kernel numbers are used. The Kernel method SVM estimator can be written as Eq. (1):

$$y = (K_{xi} \cdot W_{jk}) + b \tag{1}$$

where *b* is the bias term of the SVM network, and *W<sub>jk</sub>* is called the weight vector. *K<sub>xi</sub>* is a nonlinear function which maps input vectors to a high-dimensional property field.

Figure 4 shows the structure of a support vector regression model. The framework consists of three layers: inputs, kernel functions, outputs. There are several common types of kernels, namely linear, polykernel, and functions with a radial basis. In this analysis, the polykernel function was selected as the most suitable kernel function.

### 2.5 Multilinear regression

This is one of the most prominent and rapidly growing types of regression. It is used to try to explain the relationship between one dependent variable and two or more independent variables. If it is supposed that “*u*” dependent variable is affected by *x*<sub>1</sub>, *x*<sub>2</sub>, ..., *x*<sub>*m*</sub> independent variables, the equation which defines the relationship between the mentioned variables can be simply written as follows:

$$u = a + b_1x_1 + b_2x_2 + \dots + b_mx_m \tag{2}$$

Similar to simple regression *a*, *b*<sub>1</sub>, *b*<sub>2</sub>, ..., *b*<sub>*m*</sub> are coefficients of regression, which are obtained by minimizing the amount of “*e<sub>yi</sub>*” distance between the plane regression equations and the observation points (Bayazit 1998).

The obtained “*e<sub>yi</sub>*” is given in Eq. (3).

$$\sum_{i=1}^N e_{yi}^2 = \left[ \sum_{i=1}^N (y_i - a - b_1x_{1i} - b_2x_{2i} - b_mx_m)^2 \right] \tag{3}$$

**Table 2** Statistical information of the daily test set used for modeling soft computing techniques and calculation of the empirical equations

Parameter	Unit	Minimum	Maximum	Mean	Standard deviation
<i>T</i>	°C	- 12.6	28.3	11.7	8.7
<i>RH</i>	%	36.0	97.3	69.4	13.4
<i>U</i>	m/s	0.3	12.3	3.2	2.0
<i>SR</i>	MJ/m <sup>2</sup> /day	0.3	30.4	13.5	8.0
<i>ET</i>	mm	0.0	8.6	2.7	2.0

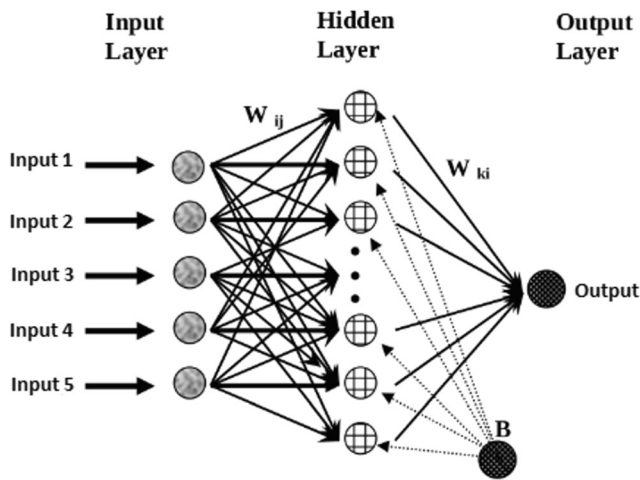


Fig. 3 Systematic diagram of a general multilayer perceptron structure

The calculation of “ $e_{yi}$ ” is based on minimizing error values in the regression analysis. Minimizing the distance between observed values and regression estimations leads to more accurate results.

**2.6 Penman-Monteith FAO 56 equation**

The PM FAO 56 equation is the standard equation suggested by FAO for ET calculation. This equation was derived based on the original Penman equation [41], and Allen et al. (1998) described it thus:

$$ET = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} U_2 (e_w - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (4)$$

where  $\Delta$  [kPa °C<sup>-1</sup>] is the slope of the vapor pressure curve,  $\gamma$  [kPa °C<sup>-1</sup>] is the psychrometric constant,  $R_n$  [MJ m<sup>-2</sup> day<sup>-1</sup>] is the net radiation,  $u_2$  [m s<sup>-1</sup>] is the wind speed at 2 m height,

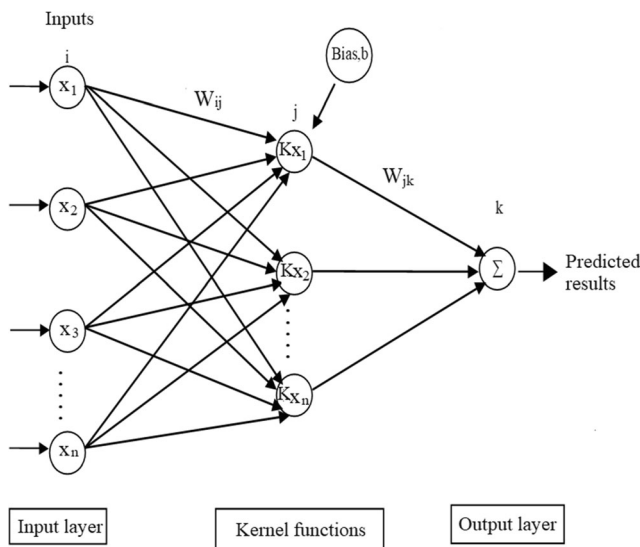


Fig. 4 Schematic diagram of support vector regression structure

$G$  [MJ m<sup>-2</sup> day<sup>-1</sup>] is the soil heat flux density,  $T$  [°C] is the mean air temperature,  $e_w$  [kPa] is the saturation vapor pressure, and  $e_a$  [kPa] is the actual vapor pressure in Eq. (4).

**2.7 Hargreaves-Samani formula**

The Hargreaves-Samani formula was developed by Hargreaves and Samani (1985) and it is presented as Eq. (5):

$$ET = 0.0023R_s \left( \frac{T_{max} + T_{min}}{2} + 17.8 \right) \sqrt{T_{max} - T_{min}} \quad (5)$$

where  $T$  [°C] represents daily mean temperature and  $R_s$  [MJ m<sup>-2</sup> day<sup>-1</sup>] extraterrestrial solar radiation.

**2.8 Ritchie formula**

Ritchie equation was developed by Jones and Ritchie in 1990 (Jones and Ritchie 1990) and is given below.

$$ET = \alpha_1 [3.87 \times 10^{-3} \times SR(0.6T_{max} + 0.4T_{min} + 29)] \quad (6)$$

In Eq. (6),  $SR$  [MJ m<sup>-2</sup> day<sup>-1</sup>] symbolizes solar radiation,  $T_{max}$  [°C] and  $T_{min}$  [°C] represent maximum and minimum temperatures, and  $\alpha_1$  is a coefficient which is calculated as follows:

$$\text{if } 5^\circ\text{C} < T_{max} < 35^\circ\text{C} \text{ then } \alpha_1 = 1.1 \quad (7)$$

$$\text{if } T_{max} > 35^\circ\text{C} \text{ then } \alpha_1 = 1.1 + 0.05 [T_{max} - 35] \quad (8)$$

$$\text{if } T_{max} < 5^\circ\text{C} \text{ then } \alpha_1 = 0.01 \times \exp[0.18 (T_{max} + 20)] \quad (9)$$

$\alpha_1$  coefficient calculations depend on maximum air temperature.

**2.9 Turc equation**

Daily mean air temperature, solar radiation, and daily percentage relative humidity parameters are essential for Turc equation (Turc 1961) calculation. Two different equations are offered based on the relative humidity percentage value, i.e., Eqs. (10 and 11).

$$RH > 50\% \rightarrow ET = 0.0133 \frac{T}{T + 15} (SR + 50) \quad (10)$$

$$RH < 50\% \rightarrow ET = 0.0133 \frac{T}{T + 15} (SR + 50) \left( 1 + \frac{50 - RH}{70} \right) \quad (11)$$

In order to determine the goodness of fit of the models used to estimate ET values, root mean square error (RMSE), relative absolute error (RAE), mean absolute error (MAE), and determination coefficient  $R^2$  given in Eq. (12), (13), and (14) were used. Here  $n$  represents the



number of data items.

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (S_a - S_e)^2} \tag{12}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |S_a - S_e| \tag{13}$$

$$RAE = \frac{[\sum_{i=1}^n (S_e - S_a)^2]^{\frac{1}{2}}}{[\sum_{i=1}^n S_a^2]^{\frac{1}{2}}} \tag{14}$$

In Eqs. (12, 13, and 14), the notation is as follows:  
 $S_a$ : actual value,  $S_e$ : estimated value.

### 3 Results and discussion

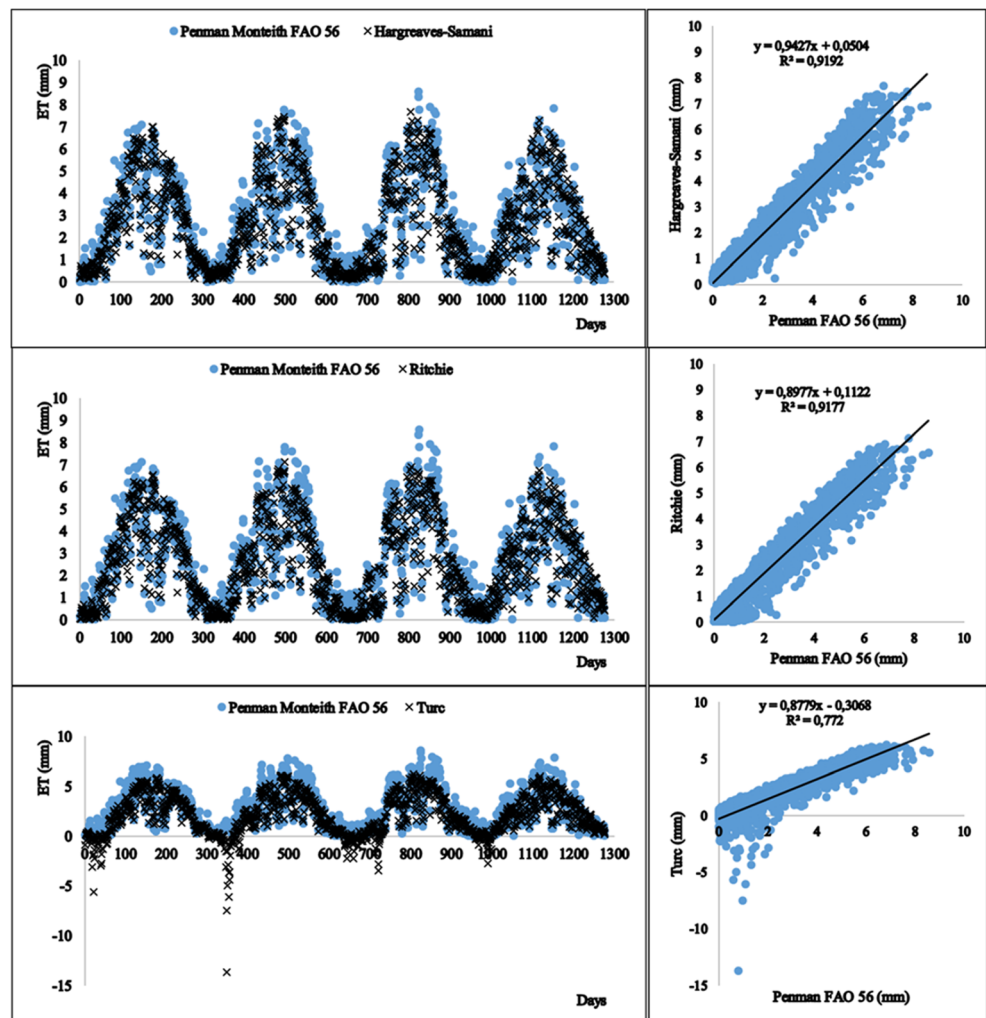
In this study, daily ET was estimated using MLP, SVR, and MLR soft computing techniques. The Penman-Monteith FAO 56 equation was accepted as the reference equation, and other empirical equation results (Hargreaves-Samani, Ritchie, Turc)

were compared with PM FAO 56. Figure 5 shows distributions and scatter charts of the empirical equation results.

Table 3 shows the results of empirical equations based on the error and determination coefficient calculations. Considering its high determination coefficient and low errors, the Hargreaves-Samani formula produced better results in the field of this study within experimental equations, while the worst performance belongs to the Turc equation.

The Ritchie and Hargreaves-Samani equation results are close to each other, based on their high determination coefficient and low error values. Both equation results are also parallel to the Penman-Monteith FAO 56 standard equation results: the distribution of results is shown in Fig. 5. However, it is not possible to claim the same performance for the Turc equation, as it has a low determination coefficient and high error values. Furthermore, comparison of the Turc values with the Penman-Monteith FAO 56 equation results showed that the Turc empirical equation results were not usable for the study area. It is known that the empirical Turc formula was originally developed for southern France and northern Africa, and since there are so many negative values of mean air

Fig. 5 Distribution graph and scatter chart of empirical equation results



**Table 3** Goodness of fit indicators for the different empirical equations used in this study for the estimation of reference evapotranspiration

Formula	Necessary parameters for calculation	MAE	RMSE	RAE (%)	R <sup>2</sup>
Hargreaves-Samani	$T_{max}, T_{min}, T, SR$	0.441	0.587	25.587	0.919
Ritchie	$T_{max}, T_{min}, SR$	0.449	0.605	26.015	0.918
Turc	$T, SR, RH$	0.771	1.184	44.662	0.772

MAE mean absolute error, RMSE root mean square error, RAE relative absolute error, R<sup>2</sup> determination coefficient

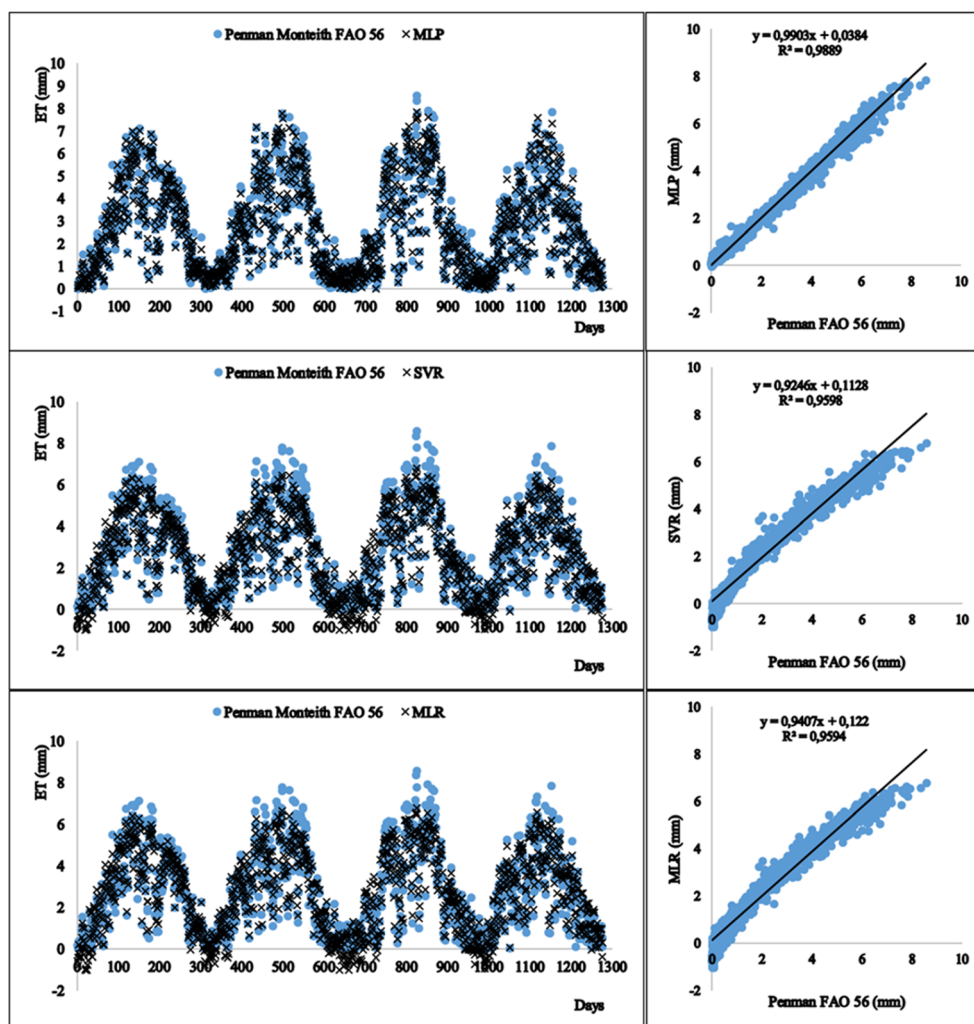
temperature in Košice, the weaker performance of the Turc formula is understandable. When the Turc formula is examined it will be seen that negative values of mean air temperature give negative ET values.

Distributions and scatter charts of the support vector regression, multilayer perceptron, and multilinear regression are presented in Fig. 6. Different combinations of individual input parameters were generated in the modeling process.

Figure 6 is created using a different combination of SR, RH, U, and T input parameters for each method. Combination error results and determination coefficients

were computed, as presented in Table 4. Different combinations of meteorological parameters were used in order to understand the impact of each parameter in the modeling process. The combination which gives the best results for each model is that of SR, T, RH, and U, as was expected. The highest determination coefficient for a single variable was calculated for the solar radiation parameter in each soft computing method separately, indicating that SR has a major effect on ET. The highest determination coefficient was calculated as 0.891 for SR in all three methods. However, error calculations

**Fig. 6** Distribution graphs and scatter charts of soft computing techniques (SR, RH, U, T combinations)





**Table 4** Goodness of fit indicators for the different soft computing techniques used in this study for the estimation of reference evapotranspiration

Combination of parameters	MLP				SVR				MLR			
	MAE	RMSE	RAE (%)	R <sup>2</sup>	MAE	RMSE	RAE (%)	R <sup>2</sup>	MAE	RMSE	RAE (%)	R <sup>2</sup>
<i>T</i>	1.470	1.754	85.196	0.665	1.028	1.270	59.571	0.609	1.036	1.289	60.026	0.609
<i>SR</i>	<b>0.674</b>	<b>0.927</b>	<b>39.075</b>	<b>0.891</b>	<b>0.575</b>	<b>0.814</b>	<b>33.332</b>	<b>0.888</b>	<b>0.615</b>	<b>0.847</b>	<b>35.631</b>	<b>0.888</b>
<i>RH</i>	1.291	1.537	74.829	0.545	1.100	1.393	63.777	0.536	1.129	1.435	65.409	0.536
<i>U</i>	1.737	2.018	100.692	0.008	1.727	2.090	100.110	0.013	1.732	2.009	100.378	0.013
<i>T-SR</i>	0.593	<b>0.712</b>	34.348	0.903	0.519	0.702	30.079	0.905	0.518	<b>0.676</b>	30.009	0.899
<i>T-RH</i>	<b>0.585</b>	0.753	<b>33.933</b>	0.911	0.613	0.753	35.501	0.865	0.702	0.864	40.716	0.866
<i>T-U</i>	1.283	1.553	74.361	0.716	0.959	1.181	55.602	0.664	0.967	1.200	56.021	0.664
<i>SR-RH</i>	0.980	1.111	56.776	0.895	<b>0.477</b>	<b>0.664</b>	<b>27.668</b>	0.905	<b>0.497</b>	0.682	<b>28.782</b>	0.899
<i>SR-U</i>	0.804	0.988	46.583	<b>0.913</b>	0.506	0.726	29.349	<b>0.910</b>	0.572	0.800	33.146	<b>0.910</b>
<i>RH-U</i>	1.262	1.513	73.162	0.551	1.096	1.387	63.538	0.540	1.125	1.430	65.187	0.540
<i>T-SR-RH</i>	<b>0.305</b>	<b>0.378</b>	<b>17.686</b>	<b>0.971</b>	<b>0.367</b>	<b>0.491</b>	<b>21.245</b>	<b>0.945</b>	<b>0.370</b>	<b>0.476</b>	<b>21.461</b>	<b>0.945</b>
<i>T-SR-U</i>	0.402	0.489	23.276	0.957	0.417	0.566	24.191	0.939	0.433	0.566	25.071	0.932
<i>T-RH-U</i>	0.550	0.707	31.878	0.925	0.600	0.734	34.778	0.872	0.694	0.848	40.196	0.873
<i>SR-RH-U</i>	1.082	1.197	62.702	0.904	0.448	0.635	25.942	0.916	0.483	0.668	28.021	0.905
<i>T-SR-RH-U</i>	0.150	0.213	8.696	0.989	0.307	0.422	17.810	0.960	0.315	0.411	18.263	0.959

MAE mean absolute error, RMSE root mean square error, RAE relative absolute error, R<sup>2</sup> determination coefficient

indicated that SVR was slightly better than other methods with regard to single parameters. The second most effective meteorological parameter was found to be air temperature in MLP, SVR, and MLR models. *SR* and *RH* combination gave the best performance in ET estimation for SVR and MLR techniques within binary combinations, but *T-RH* binary combination had a determination coefficient in the MLP method as high as 0.911, and lower error calculations. Investigation of triple combinations of inputs revealed that *T-SR-RH* was more powerful than the others in all the three soft computing techniques used. The results of all these given combinations also confirmed that the wind speed value (*U*) alone was less influential compared to the other parameters. However, applying the *U* meteorological variable together with other factors increased the models' accuracy emphatically. Detailed performance analysis was done using error percentage frequencies and histograms, and is plotted in Fig. 7.

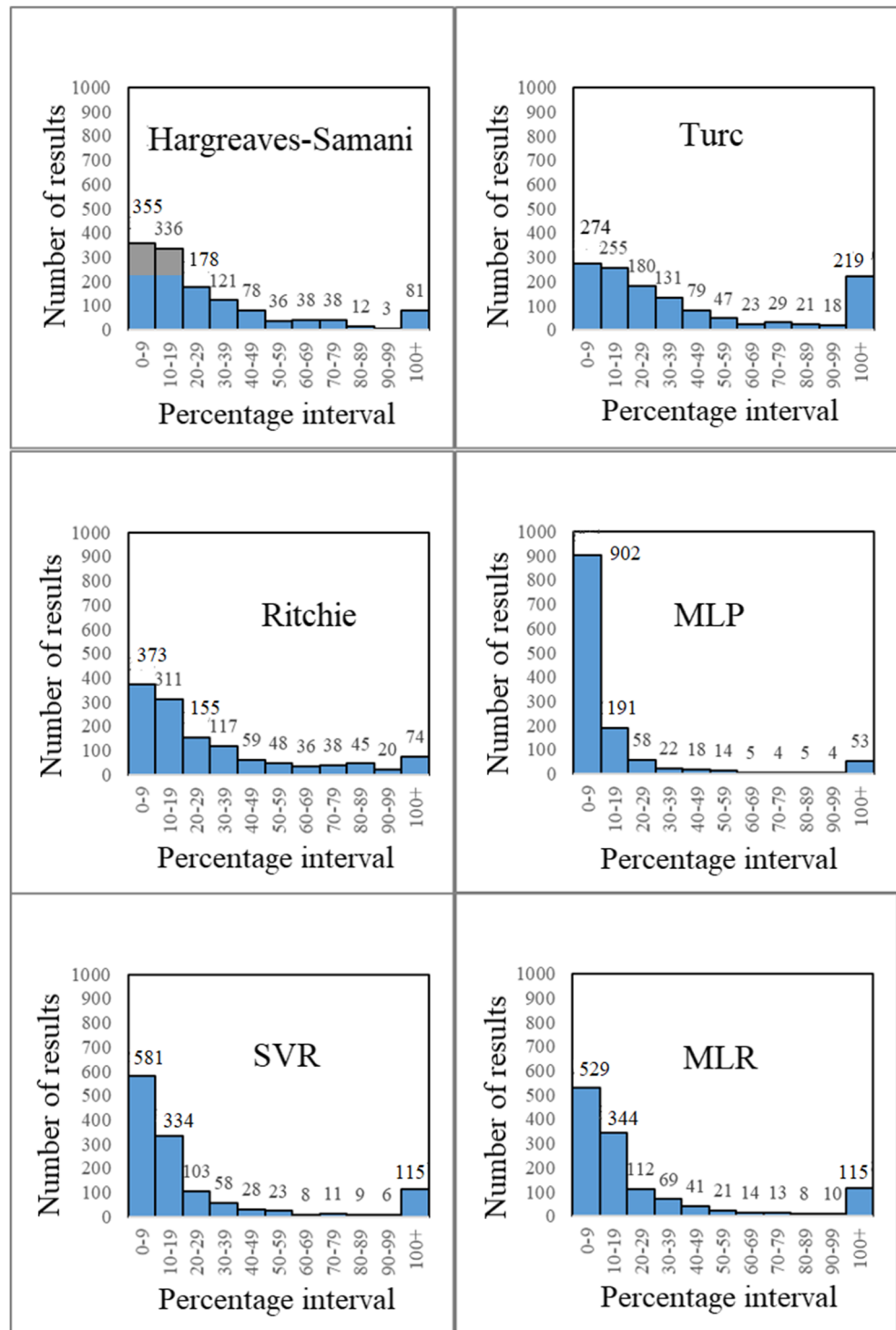
The best statistics of each meteorological parameter combinations and methods are marked as bold in Table 4. Based on the percentage error histograms, it can be underlined that 90.1% of test set results have percentage error lower than 30% for the *SR-T-RH-U* combination in MLP. When the same calculation was done for the other methods and equations, the following results were obtained: 79.7% for SVR *SR-T-RH-U* combination, 77.1% for MLR *SR-T-RH-U* combination, 68.1% for Hargreaves-Samani formula, 65.7 for Ritchie

formula, and 55.5% for Turc formula test set results. In Fig. 7, MLP, SVR, and MLR results are taken from *T*, *SR*, *RH*, and *U* combination.

Kişı found MLP to be a well-performing approach during his study comparing the performances of different artificial neural network models (Kisi 2008). He also used CIMIS Penman, Hargreaves, and Ritchie empirical formulas for comparison with results obtained from the Penman-Monteith FAO 56 equation. According to his comparison, he found that MLP was the best approach for estimating daily ET in the study area out of the artificial neural network approaches applied. With regard to the Pomona station, he calculated the determination coefficient as 0.991 for MLP, 0.981 for non-modified Hargreaves, and 0.985 for non-modified Ritchie. Then, with regard to the Santa Monica station, he computed the determination coefficient as 0.997 for MLP, 0.818 for non-modified Hargreaves, and 0.780 for non-modified Ritchie. It is clear from his results that empirical equations' results can differ even for two nearby stations. However, the MLP method has quite good results for both stations. He calculated mean absolute error as 0.140 and 0.048 for the Pomona and Santa Monica stations, respectively.

In the present study, MLP, SVR, and MLR methods were used as soft computing approaches and MLP gave results most compatible with the Penman-Monteith FAO 56 evapotranspiration results. The determination coefficient was calculated as 0.989, MAE as 0.150, RMSE as 0.213, and RAE as

**Fig. 7** Percentage error frequencies for the methods and equations used



8.696 for the MLP method, using  $T$ ,  $SR$ ,  $RH$ , and  $U$  parameters as input. Comparison with earlier studies shows that the calculated statistics for MLP and the other approaches are relevant for ET estimation of Košice.

Performances of empirical equations set against related combinations of soft computing techniques, based on the same parameter combination inputs, are given in Table 5.

According to the parameter combination input comparison, it is seen that the Hargreaves-Samani and Ritchie empirical equations produced results more in line with the Penman-Monteith FAO 56 equation results in the study area when the soft computing techniques used only the  $T$  and  $SR$  combination as input. But if the soft computing models were operated with the  $T$ ,  $SR$ , and  $RH$  combination, then all soft

**Table 5** Parameter effects on evaluation of each method

Parameters used for calculation/input combinations	Method	MAE	RMSE	RAE	$R^2$
$T_{\max}$ , $T_{\min}$ , $T$ , $SR$	Hargreaves-Samani	0.441	0.587	25.587	0.919
$T_{\max}$ , $T_{\min}$ , $SR$	Ritchie	0.449	0.605	26.015	0.918
$T$ , $SR$	MLP	0.593	0.712	34.348	0.903
	SVR	0.519	0.702	30.079	0.905
	MLR	0.518	0.676	30.009	0.899
$T$ , $SR$ , $RH$	MLP	0.305	0.378	17.686	0.971
	SVR	0.367	0.491	21.245	0.945
	MLR	0.370	0.476	21.461	0.945
	Turc	0.771	1.184	44.662	0.772

computing techniques showed better performance than the empirical equations. In any case, the Turc empirical formula did not produce satisfactory results in the study area.

## 4 Conclusions

In this study, the authors used SVR, MLP, and MLR models and most commonly used empirical equations for daily ET estimation in the Košice City area in eastern Slovakia.

Solar radiation was identified as the parameter most affecting ET in all three soft computing models, while wind speed was found to be the one with least impact. All soft computing techniques used performed well when compared to the Penman FAO 56, as suggested by the goodness-of-fit indicators used. Ultimately, the multilayer perceptron method ability to estimate ET in the study area was found to be the best. MLP results were more compatible with Penman-Monteith FAO 56 than those of the other models and the empirical formulas. All soft computing techniques gave better performance than the empirical formulas used. The Hargreaves-Samani equation results were much closer to the Penman FAO 56 results than the other empirical equations. The non-modified Turc empirical formula is not recommended for use in making ET predictions for the study area. Negative daily mean air temperature values caused its ET calculations to be significantly lower than in the other approaches. Another important outcome of the study is that using double combinations of  $T$ - $RH$  or  $T$ - $SR$  is quite useful for the prediction of ET when applying the MLP method.  $SR$ - $RH$  combination results are acceptable when using the SVR and MLR methods. The results of these combinations of meteorological parameters are compatible with the Penman-Monteith FAO 56 equation results. Thus, it is possible to state that implementing soft computing techniques gives researchers opportunities to make quite accurate ET estimations while using fewer recorded meteorological parameters.

**Acknowledgments** The authors would like to thank the General Directorate for State Hydraulic Works (DSI) for sharing their measurements.

**Author contributions** Conceptualization, F.Ü. and B.T.; compiling data, Y.Z.K.; formal analysis, M.D.; funding acquisition, M.D.; investigation, H.V.; methodology, H.V.; project administration, M.Z.; resources, M.D.; software, Y.Z.K.; supervision, F.Ü., B.T., M.Z., and F.Ü.; validation, F.Ü.; visualization, Y.Z.K.; writing—original draft preparation, Y.Z.K. and H.V.; writing—review and editing, M.Z.

**Funding** This work was supported by the project of the Ministry of Education of the Slovak Republic VEGA 1/0308/20: Mitigation of hydrological hazards—floods and droughts—by exploring extreme hydroclimatic phenomena in river basins and the project of the Slovak Research and Development Agency APVV-17-0549: Research of knowledge and virtual technologies supporting intelligent design and implementation of buildings with emphasis on their economic efficiency and sustainability.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration—guidelines for computing crop water requirements. FAO - Food and Agriculture Organization of the United Nations, Rome
- Bayazit MOB (1998) Probability and statistics for engineers. Birsen Publishing House, Istanbul, Turkey
- Bishop C (1995) Neural networks for pattern recognition. Oxford: University Press
- Chen D (2012) Daily reference evapotranspiration estimation based on least squares support vector machines. In: IFIP Advances in Information and Communication Technology. [https://doi.org/10.1007/978-3-642-27278-3\\_7](https://doi.org/10.1007/978-3-642-27278-3_7)
- Cortes, C., Vapnik, V., 1995. Support-Vector Cortes, C., & Vapnik, V.. Support-vector networks. Machine Learning, 20(3), 273–297. doi: <https://doi.org/10.1023/A:1022627411411>
- Cybenko G (1989) Approximation by superpositions of a sigmoidal function. Math Control Signals, Syst 2. <https://doi.org/10.1007/BF02551274>



- Demirci M, Üneş F, Aköz MS (2015) Prediction of cross-shore sandbar volumes using neural network approach. *J. Mar. Sci. Technol.* 20: 171–179. <https://doi.org/10.1007/s00773-014-0279-9>
- Fendeková M, Gauster T, Labudová L, Vrablíková D, Danáčová Z, Fendek M, Pekárová P (2018) Analysing 21st century meteorological and hydrological drought events in Slovakia. *J. Hydrol. Hydromechanics* 66:393–403. <https://doi.org/10.2478/johh-2018-0026>
- Gavili S, Sanikhani H, Kisi O, Mahmoudi MH (2018) Evaluation of several soft computing methods in monthly evapotranspiration modelling. *Meteorol. Appl.* 25:128–138. <https://doi.org/10.1002/met.1676>
- Gocić M, Motamedi S, Shamsirband S, Petković D, Ch S, Hashim R, Arif M (2015) Soft computing approaches for forecasting reference evapotranspiration. *Comput. Electron. Agric.* 113:164–173. <https://doi.org/10.1016/j.compag.2015.02.010>
- Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* 1(2):96–99
- Haykin S (1999) *Neural networks: a comprehensive*. Pearson Education. 13:409–412. <https://doi.org/10.1017/S0269888998214044>
- Hlaváčiková H, Novák V (2013) Comparison of daily potential evapotranspiration calculated by two procedures based on Penman-Monteith type equation. *J. Hydrol. Hydromechanics* 61:61–176. <https://doi.org/10.2478/johh-2013-0022>
- Hlavčová K, Kalaš M, Kohnová S, Szolgay J, Danihlík R (2004) Modelling of monthly potential evapotranspiration and runoff in the Hron Basin. *J. Hydrol. Hydromech* 52:255–266
- Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural Networks* 2:359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Jones JW, Ritchie JT (1990) Crop growth models. In: Hoffman GJ, Howel TA, Solomon KH (eds) *Management of Farm Irrigation System*, pp 63–89
- Kaya YZ, Mamak M, Unes F (2016a) Evapotranspiration prediction using M5T data mining method. *Int J Adv Eng Res Sci.* 3. <https://doi.org/10.22161/ijaers/3.12.40>
- Kaya YZ, Üneş F, Mamak M (2016b) Estimating evapotranspiration using adaptive neuro-fuzzy inference system and Hargreaves-Samani method. In: *Book of abstracts of the International Conference on Engineering and Natural Sciences (ICENS) 2016*. Sarajevo, pp 1983–1989
- Kim S, Kim HS (2008) Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *J. Hydrol.* 351:299–317. <https://doi.org/10.1016/j.jhydrol.2007.12.014>
- Kişi Ö (2004) River flow modeling using artificial neural networks. *J. Hydrol. Eng.* 9:60–63. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2004\)9:1\(60\)](https://doi.org/10.1061/(ASCE)1084-0699(2004)9:1(60))
- Kisi O (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrol Sci J* 50. <https://doi.org/10.1623/hysj.2005.50.4.683>
- Kisi O (2008) The potential of different ANN techniques in evapotranspiration modelling. *Hydrol. Process.* 22:2449–2460. <https://doi.org/10.1002/hyp.6837>
- Kisi O (2015) Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *J. Hydrol.* 528:312–320. <https://doi.org/10.1016/j.jhydrol.2015.06.052>
- Kişi O, Çimen M (2009) Evapotranspiration modelling using support vector machines / Modélisation de l'évapotranspiration à l'aide de 'support vector machines'. *Hydrol. Sci. J.* 54:918–928. <https://doi.org/10.1623/hysj.54.5.918>
- Kisi O, Zounemat-Kermani M (2014) Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. *Water Resour. Manag.* 28:2655–2675. <https://doi.org/10.1007/s11269-014-0632-0>
- Kumar M, Raghuvanshi NS, Singh R, Wallender WW, Pruitt WO (2002) Estimating evapotranspiration using artificial neural network. *J. Irrig. Drain. Eng.* 128:224–233. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2002\)128:4\(224\)](https://doi.org/10.1061/(ASCE)0733-9437(2002)128:4(224))
- Kumar M, Raghuvanshi NS, Singh R (2011) Artificial neural networks approach in evapotranspiration modeling: a review. *Irrig. Sci.* 29: 11–25. <https://doi.org/10.1007/s00271-010-0230-8>
- Landeras G, Ortiz-Barredo A, López JJ (2008) Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agric. Water Manag.* 95:553–565. <https://doi.org/10.1016/j.agwat.2007.12.011>
- Mirás-Avalos JM, Rubio-Asensio JS, Ramírez-Cuesta JM, Maestre-Valero JF, Intrigliolo DS (2019) Irrigation-advisor-a decision support system for irrigation of vegetable crops. *Water (Switzerland)* 11. <https://doi.org/10.3390/w11112245>
- Pal M, Deswal S (2009) M5 model tree based modelling of reference evapotranspiration. *Hydrol. Process.* 23:1437–1443. <https://doi.org/10.1002/hyp.7266>
- Parajka J, Szolgay J, Meszaros I, Kostka Z (2004) Grid-based mapping of the long-term mean annual potential and actual evapotranspiration in upper Hron River basin. In *J Hydrol Hydromech, ÚH SAV*, no. 4, 2004, 239–254
- Rahimikhoob A (2014) Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water Resour. Manag.* 28:657–669. <https://doi.org/10.1007/s11269-013-0506-x>
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL (eds) *Parallel distributed processing: explorations in the microstructure of cognition*. MIT Press, Cambridge, pp 318–362
- SHMI. 2015. *Climate atlas of Slovakia*.
- Tasar B, Üneş F, Demirci M, Kaya YZ (2018) Yapay sinir ağları yöntemi kullanılarak buharlaşma miktarı tahmini. *DÜMF Mühendislik Derg.* 91(1):543–551
- Toprak ZF, Cigizoglu HK (2008) Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods. *Hydrol. Process.* 22:4106–4129. <https://doi.org/10.1002/hyp.7012>
- Traore S, Wang YM, Kerh T (2010) Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Agric. Water Manag.* 97:707–714. <https://doi.org/10.1016/j.agwat.2010.01.002>
- Turc L (1961) Evaluation des besoins en eau d'irrigation, évapotranspiration potentielle, formulation simplifiée et mise à jour. *Ann. Agron.* 12:13–49
- Üneş F, Demirci M, Kişi Ö (2015) Prediction of millers ferry dam reservoir level in USA using artificial neural network. *Period. Polytech. Civ. Eng.* 59:309–318. <https://doi.org/10.3311/PPci.7379>
- Üneş F, Doğan S, Tasar B, Kaya YZ, Demirci M (2018) The evaluation and comparison of daily reference evapotranspiration with ANN and empirical methods. *Nat Eng Sci* 3:54–64
- Yihdego Y, Webb JA (2018) Comparison of evaporation rate on open water bodies: energy balance estimate versus measured pan. *J Water Clim Chang* 9:9–111. <https://doi.org/10.2166/wcc.2017.139>
- Zanetti SS, Sousa EF, Oliveira VPS, Almeida FT, Bernardo S (2007) Estimating evapotranspiration using artificial neural network and minimum climatological data. *J Irrig Drain Eng.* 133:83–89. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2007\)133:2\(83\)](https://doi.org/10.1061/(ASCE)0733-9437(2007)133:2(83))

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.