Comparison of multi-gene genetic programming and dynamic evolving neural-fuzzy inference system in modeling pan evaporation
Okan Eray, Cihan Mert and Ozgur Kisi

ABSTRACT

Accurately modeling pan evaporation is important in water resources planning and management and also in environmental engineering. This study compares the accuracy of two new data-driven methods, multi-gene genetic programming (MGGP) approach and dynamic evolving neural-fuzzy inference system (DENFIS), in modeling monthly pan evaporation. The climatic data, namely, minimum temperature, maximum temperature, solar radiation, relative humidity, wind speed, and pan evaporation, obtained from Antakya and Antalya stations, Mediterranean Region of Turkey were utilized in the study. The MGGP and DENFIS methods were also compared with genetic programming (GP) and calibrated version of Hargreaves Samani (CHS) empirical method. For Antakya station, GP had slightly better accuracy than the MGGP and DENFIS models and all the data-driven models performed were superior to the CHS while the DENFIS provided better performance than the other models in modeling pan evaporation at Antalya station. The effect of periodicity input to the models’ accuracy was also investigated and it was found that adding periodicity significantly increased the accuracy of MGGP and DENFIS models.

Key words | dynamic evolving neural-fuzzy inference system, genetic programming, modeling, multi-gene genetic programming, pan evaporation, periodicity

INTRODUCTION

Water is the one of the most crucial substances for all living things to maintain their lives. Although it is the most abundant substance on Earth, its usage and conservation have become very important. This has led sustainable water resource management to play a very important role for all existing things. Therefore, several fields of sciences, including hydrology, meteorology, oceanography, limnology, etc., have and are being developed. Among them, hydrology deals with the all forms of water that have great importance for humans and their environment. Therefore, hydrological modeling has been developed that includes solar radiation modeling, stream flow forecasting, rainfall–runoff modeling, and evapotranspiration estimation.

The hydrological cycle is the common point of all these branches of sciences. Evaporation as a basic component of the hydrological cycle is of vital importance in water resources planning and management and also in environmental work. The accurate calculation of evaporation is very important, especially in the arid and semi-arid areas (Malik & Kumar 2015). Direct and indirect approaches are generally used for computing and estimating evaporation. Pan evaporation is one of the direct approaches for evaporation calculation (Goyal & Ojha 2011). Indirect approaches are needed for prediction of pan evaporation because installing evaporation pans and measuring are highly expensive. By utilizing indirect approaches, pan evaporation can be
predicted based on some climatic variables, such as relative humidity (RH), temperature, wind speed, and solar radiation (Kisi et al. 2016).

There are numerous practical issues (e.g., post-mining voids or deep lakes, farm dams or shallow lakes, studies investigating the water balance of a catchment, modeling rainfall–runoff, for areas of small irrigation or for the crops irrigated within a large irrigation district) that require the estimates of daily or monthly actual or potential evaporation. All these issues indicate that the daily or monthly evaporation calculations from climatic data or Class-A pan measurements are necessary in most of the practical issues (McMahon et al. 2017, 2016).

Over the last several decades many empirical and physical models have been utilized (Mustacchi et al. 1979; Yu et al. 2006; Han et al. 2007). Since these processes are nonlinear, complex, temporally and spatially varying, incidental, and unsteady, it is very difficult to model all the physical processes. Later, strong alternative modeling techniques known as artificial intelligence (AI)-based techniques were developed. These AI models showed that they performed better than traditional modeling techniques, especially dealing better with big noisy data.

In the last decades, soft computational approaches have been successfully utilized for prediction of pan evaporation (Sudheer et al. 2002; Kisi 2006, 2015a; Kim & Kim 2008; Shiri et al. 2011; Sanikhani et al. 2012; Lin et al. 2013; Malik & Kumar 2015; Keshtegar et al. 2016; Kisi et al. 2016). Sudheer et al. (2002) predicted daily pan evaporation of Dowleswaram in Andhra Pradesh, India by using artificial neural networks (ANNs) and showed that the ANNs could be used in modeling the evaporation process from the available weather variables. Kisi (2006) investigated the ability of neuro-fuzzy system in estimating pan evaporation of two automated weather stations, Arcata-Eureka and Daggett stations, California, operated by the US Environmental Protection Agency and compared it with ANN. He found the neuro-fuzzy system to be better than the ANN in modeling daily pan evaporation. Kim & Kim (2008) developed a generalized regression neural networks model with genetic algorithm for prediction of pan evaporation and the alfalfa reference evapotranspiration, in the Republic of Korea and obtained promising results. Shiri et al. (2011) used an adaptive neuro-fuzzy inference system (ANFIS) and ANN in prediction of daily pan evaporation of three weather stations, Illinois, USA based on air temperature, wind speed, solar radiation, RH, total rainfall, and surface soil temperature inputs. Sanikhani et al. (2012) applied two ANFIS systems in predicting daily pan evaporation of San Francisco and San Diego, in California, USA and compared them with the ANN method. ANFIS provided better estimates when compared to ANN. Lin et al. (2013) compared the support vector machine (SVM) and ANN in prediction of daily pan evaporation. According to their results, the SVM approach was more appropriate than the ANN with respect to accuracy and efficiency. Malik & Kumar (2015) predicted daily pan evaporation of Pantnagar, located in the foothills of the Himalayas in Uttarakhand state of India using ANN and co-active ANFIS and found ANN to be better than the latter. Kisi (2015a) used least square support vector machine (LSSVM), multivariate adaptive regression splines, and M5 model tree in prediction of pan evaporation and LSSVM was found to perform better than the others using local climatic data. Kisi et al. (2016) predicted pan evaporation of Ankara and Polatlı stations, Turkey, using classification and regression tree, chi-squared automatic interaction detector, and ANNs and comparison of the methods indicated that the ANNs performed better than the others. Keshtegar et al. (2016) used conjugate gradient optimization method for daily pan evaporation prediction and compared it with ANFIS and model tree approaches. The conjugate gradient-based model performed better than the other models. Most of the approaches used in previous studies are based on black box methods in which their formulation is not explicit and cannot be easily used by practical applications. Therefore, in the present study, a multi-gene genetic programming (MGGP)-based model, which had explicit formulation and could be easily used in practice, was developed for prediction of pan evaporation.

There are limited studies in the literature related to the application of the dynamic evolving neural-fuzzy inference system (DENFIS) in water resources and hydrology (Heddam 2014; Heddam & Dechemi 2015; Kwin et al. 2016). DENFIS was applied for modeling dissolved oxygen concentration by Heddam (2014). The study demonstrated the superiority of the DENFIS over the multiple linear regression and ANN models. In a study by Heddam & Dechemi (2015), coagulant dosage in a water treatment
plant in Algeria was modeled using DENFIS and satisfactorily results were obtained. In a study by Kwin et al. (2016), DENFIS was used for rainfall–runoff modeling and compared with the Hydrologic Engineering Center-Hydrologic Modeling System (HEC-HMS) and autoregressive model with exogenous inputs (ARX). It was found that DENFIS estimates were comparable to HEC-HMS and superior to the ARX model. To the best of our knowledge, there is no published study in the literature related to the application of DENFIS in modeling pan evaporation process.

The main aim of this study is: (1) to investigate the accuracy of MGGP and DENFIS approaches in prediction of pan evaporation; (2) to compare their accuracy with genetic programming (GP) and calibrated Hargreaves-Samani (HS) equation; (3) to investigate the effect of periodicity input to the models' accuracy.

**METHODS**

**Genetic programming**

GP is evaluated from computational techniques which are used for applying and solving different kinds of engineering problems (Kaydani et al. 2014). GP is adapted from the natural selection and generated extension of genetic algorithms. First, it was invented by Cramer (1985) and then developed by Koza (1992). Solutions in algorithms are computed by structured trees (Figure 1). Figure 1 illustrates an example of GP binary tree of the function \((x/3 + (-y))\). The tree has two types of nodes which are ‘functional’ and ‘terminals’. Functional nodes are used for arithmetical, logical, and other functions. Terminals are used for variables and constants (Durasevic et al. 2016).

GP algorithms have three genetics operations which are reproduction, mutation, and crossover. Genetic algorithms compute individuals and choose individuals for genetic operations. The initial population of individuals’ outcomes makes a new set of individuals in a new generation. A new generation is produced instead of existing generation at reproduction operation. The branch of the tree is randomly selected from two individuals and swapped. Terminal or functional is selected from the trees and mutated. The goal is to get the best individual from the population with the best fitness value and resulting in GP.

**Multi-gene genetic programming**

In this method, the initial population achievement is measured on the training data based on fitness function minimized to get a better solution. Typically, the fitness function used is special root mean squared error (SRMSE), given by

\[
SRMSE = \sqrt{\frac{\sum_{i=1}^{N} |G_i - A_i|^2}{N}} 
\]

where \(G_i\) is predicted value of \(i^{th}\) data sample by MGGP, \(A_i\) is an actual value of \(i^{th}\) sample of data, and \(N\) is number of training samples (Garg et al. 2014).

MGGP is derived from GP which is used for advancing empirical mathematical modeling. This model is a weighted number of GP trees combination of set genes and designed by regression genes which use the least squares method and estimation of coefficients \(S_0, S_1, S_2,\) and \(S_3\). The model predicts output value depending on two input variables which are \(x_1\) and \(x_2\). A typical MGGP formulation mechanism is shown in Figure 2. In this figure, Gene 1, Gene 2, and Gene 3 express the functions \(\cos(x_1)/x_2 + (-5), x_2 + x_1*2/(-5),\) and \((x_1-5)*(x_2 + 3)\), respectively. Values of \(G_{max}\) and \(d_{max}\) are specified by the user to produce control over the MGGP model complexity. The \(G_{max}\) and \(d_{max}\) parameters influence the size and the number of...
models to be searched in the global space. Thus, there are ideal values of $G_{\text{max}}$ and $d_{\text{max}}$ which generate a comparatively compact model (Searson et al. 2010).

The initial population in MGGP is built by creation of individuals which exist in GP trees with randomly selected different genes. During MGGP execution, genes are added or deleted by crossover operator which is described as a two-point high-level crossover which allows gene exchange between individuals.

A gene is randomly selected from each parent. Then, standard subtree crossover operator is applied and the comprised tree is swapped with parent trees. Probability of crossover operator is defined by the user (Gandomi & Alavi 2012). MGGP algorithm follows these steps:

Step 1: Define the problem.
Step 2: Set initial parameters (population size, maximum number of genes, generations, etc.)
Step 3: Build a model using least square method.
Step 4: Evaluate performance of models (based on SRMSE).
Step 5: Apply genetic operators and construct new population.
Step 6: Evaluate performance of individuals in new population.
Step 7: Finish the steps if termination criteria are supplied. Otherwise, go to step 5.

Detailed information about MGGP can be obtained from previous studies (Searson 2009; Searson et al. 2010).

Dynamic evolving neural-fuzzy inference system

DENFIS is a new type of fuzzy inference system proposed by Song & Kasabov (2000) for adaptive offline and online learning. DENFIS uses and improves dynamic features of evolving fuzzy neural network (EFuNN) so that DENFIS becomes suitable for online adaptive systems.

DENFIS model uses evolving clustering method (ECM) and evolving clustering method with constrained minimization (ECMc) for partitioning the input space. In DENFIS, the rules are created and updated in this partitioning. ECM is an online evolving fast clustering method that is based on maximum distance. By using the maximum
The monthly maximum temperature (Tmax), minimum temperature (Tmin), solar radiation (Rs), wind speed (Ws), RH, and pan evaporation data from Antakya (latitude 36.33°N, longitude 36.30°E, altitude 100 m) and Antalya (latitude 36.89°N, longitude 30.68°E, altitude 47 m) stations located in the Mediterranean Region of Turkey were used in the present study. Related data were obtained from the Turkish State Meteorological Service (TSMO). More information about data measurements is available from the TSMO website (https://www.mgm.gov.tr/eng/forecast-5days.aspx?g=0).

The climate of this region has cool, rainy winters and moderately dry and hot summers. Yearly rainfall ranges from 580 to 1,300 mm. For the Antakya and Antalya stations, 203 (from 1983 to 2010) and 362 (from 1967 to 2006) monthly data were available. It should be noted that an evaluation of data uncertainty and its effect on the results (Beven et al. 2008; Wang et al. 2015) were not considered in this study. For each station, the first 80% of the whole data was used for training and the remaining 20% was used for testing the obtained models. The brief statistical properties of the used data sets are reported in Table 1. In this table, the Xmean, Xmin, Xmax, Sx and Csx indicate the mean, minimum, maximum, standard deviation, and skewness, respectively. It is clear from the table that data of both stations have different distributions to each other. Pan evaporation ranges are 0.9–9.8 and 1.3–12.4 for the Antakya and Antalya stations, respectively, and data of the latter station show more skewed distribution (Csx = 0.54).

Root mean square error (RMSE), mean absolute error (MAE), Nash–Sutcliffe efficiency (NSE) coefficient, and determination coefficient (R2) were used for evaluation of the applied models. The RMSE, MAE, and NSE can be expressed as:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (PE_{M,i} - PE_{O,i})^2}
\]

(3)

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |PE_{M,i} - PE_{O,i}|
\]

(4)

\[
NSE = 1 - \frac{\sum_{i=1}^{n} (PE_{M,i} - PE_{O,i})^2}{\sum_{i=1}^{n} (PE_{O,i} - \bar{PE}_O)^2}
\]

(5)

<table>
<thead>
<tr>
<th>Station</th>
<th>Variable</th>
<th>Xmean</th>
<th>Xmin</th>
<th>Xmax</th>
<th>Sx</th>
<th>Csx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antakya</td>
<td>Rs (Langley)</td>
<td>107</td>
<td>30</td>
<td>179</td>
<td>39.8</td>
<td>-0.25</td>
</tr>
<tr>
<td></td>
<td>Tmax (°C)</td>
<td>24.1</td>
<td>10</td>
<td>32.2</td>
<td>5.67</td>
<td>-0.80</td>
</tr>
<tr>
<td></td>
<td>Tmin (°C)</td>
<td>16.2</td>
<td>0.6</td>
<td>28.3</td>
<td>7.55</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>RH (%)</td>
<td>69.1</td>
<td>30.8</td>
<td>80.2</td>
<td>5.03</td>
<td>-0.49</td>
</tr>
<tr>
<td></td>
<td>Ws (m/s)</td>
<td>3.45</td>
<td>1.3</td>
<td>7.1</td>
<td>1.26</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Pan evaporation (mm)</td>
<td>4.59</td>
<td>0.9</td>
<td>9.8</td>
<td>2.15</td>
<td>-0.03</td>
</tr>
<tr>
<td>Antalya</td>
<td>Rs (Langley)</td>
<td>119</td>
<td>33.3</td>
<td>215</td>
<td>47.3</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>Tmax (°C)</td>
<td>23.0</td>
<td>11.1</td>
<td>37.2</td>
<td>6.90</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Tmin (°C)</td>
<td>14.1</td>
<td>0.6</td>
<td>27.4</td>
<td>7.31</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>RH (%)</td>
<td>63.3</td>
<td>42.2</td>
<td>80.4</td>
<td>6.98</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>Ws (m/s)</td>
<td>2.92</td>
<td>1.7</td>
<td>5.8</td>
<td>0.65</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Pan evaporation (mm)</td>
<td>5.31</td>
<td>1.3</td>
<td>12.4</td>
<td>2.61</td>
<td>0.54</td>
</tr>
</tbody>
</table>
where $PE_{O,i}$ and $PE_{M,i}$ are the observed and estimated pan evaporation rates, $n$ is the number of time steps, $PE_O$ is the mean of the observed pan evaporation.

The MGGP and DENFIS were applied for estimating pan evaporation and compared with GP and calibrated version of the HS empirical equation. The Hargreaves & Samani (1985) equation is given as:

$$ET_0 = 0.0023R_a\left(\frac{T_{\text{max}} + T_{\text{min}}}{2}\right) + 17.8\left(T_{\text{max}} - T_{\text{min}}\right)^{0.5} \quad (6)$$

where $ET_0$ = reference evapotranspiration (mm day$^{-1}$); $T_{\text{max}}$ and $T_{\text{min}}$ = maximum and minimum temperature (°C) and $R_a$ = extraterrestrial radiation (mm day$^{-1}$). In this study, $R_a$ was calculated as described by Allen et al. (1998). Calibration of the HS equation was made for estimating pan evaporation by using the data used for the training of AI models. Rahimikhoob (2009) also previously used this equation for pan evaporation modeling. The optimal $a$, $b$, and $c$ parameters given in Equation (7) were calculated by using genetic algorithm. Mutation rate, population size, and error were set to 0.075, 100, and 0.0001, respectively.

$$ET_0 = aR_a\left(\frac{T_{\text{max}} + T_{\text{min}}}{2} + b\right)\left(T_{\text{max}} - T_{\text{min}}\right)^{c} \quad (7)$$

Thus, the calibrated HG (CHS) estimates were compared with the AI methods.

Similar parameters were set for the GP and MGGP methods and given in Table 2. Figure 3 illustrates the variation of the mean and best fitness with the generation number and the statistical characteristics of the MGGP model evolved (training data) for Antakya station. As clearly seen from Figure 3(a), the weight values of the bias term and Gene 2 are higher than those of the other genes. Figure 3(b) gives the significance degree of each gene in respect to p values. Low p values in this figure indicate that the contribution of the genes to explain variations in pan evaporation is very high. Also, Figure 4 demonstrates the populations of the models evolved in terms of their node numbers indicating complexity together with their fitness values. From this figure, the best model according to complexity (less complex) and population (less population) can be determined. A big black circle in the figure indicates the best model with respect to population. A light grey circle shows the models which are not strongly dominated by other models in the populations in terms of model complexity and fitness. The obtained MGGP equation is given in Table 3. For the DENFIS models, selection of distance threshold value (Dthr) is very important (Heddam & Dechemi 2013). In this study, various Dthr values were tried to decide the optimal one. The optimal Dthr value obtained for the DENFIS model is 0.02 for Antakya station.

Training and test results of the optimal MGGP, GP, DENFIS, and CHS models are compared in Table 4. It is apparent from the results that the GP model performs slightly better than the MGGP and DENFIS models with respect to RMSE, MAE, NSE, and $R^2$. Training results show that the MGGP model approximates pan evaporation better than the GP model. In the test stage, also, closer estimates are obtained for the MGGP, GP, and DENFIS. All the AI models have better accuracy than the CHS empirical equation. Figure 5 illustrates the scatterplot comparison of applied models in estimating pan evaporation. It is apparent that the MGGP provides less scattered estimates with a slope constant and bias value, respectively, closer to 1 and 0 than those of the GP and DENFIS.

The variation of the fitness vs generations and the statistical properties of the MGGP model in training are illustrated in Figure 6(a) for Antalya Station. From the figure, it is apparent that Gene 3 has the highest weight followed by the bias and Gene 7. The other genes have

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function set</td>
<td>+, −, x, /, √, exp, ln</td>
</tr>
<tr>
<td>Population size</td>
<td>100</td>
</tr>
<tr>
<td>Number of generations</td>
<td>100</td>
</tr>
<tr>
<td>Maximum number of genes allowed in an individual</td>
<td>8</td>
</tr>
<tr>
<td>Maximum tree depth</td>
<td>4</td>
</tr>
<tr>
<td>Tournament size</td>
<td>12</td>
</tr>
<tr>
<td>Elitism</td>
<td>0.01% of population</td>
</tr>
<tr>
<td>Probability of GP tree mutation</td>
<td>0.1</td>
</tr>
<tr>
<td>Probability of GP tree crossover</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Table 2 | Parameters of the GP and MGGP models
considerably lower weights. The significance degree of each gene is provided in Figure 6(b). Similar too is the Antakya station, here also the genes’ contribution to explain variations in pan evaporation are significantly high, because of their corresponding very low p values. The populations of the optimal MGGP models in terms of their complexity and fitness – Antakya station.

Table 3 | The equations of the optimal MGGP models

| Antakya |  
| Pan evaporation | $= \left( \frac{2.504}{x_2 - x_5} - 0.0004908(x_2)^2 - 0.0705 \right) x_1 + 0.07009 x_3$
- $0.1727 x_2 + 0.09533 x_5 + 0.0003387 \log(x_2)$
+ $0.00658(x_1)^2 - 0.00002331(x_2)^2 + 0.00002331(x_1)^2(x_4)^2$
+ $0.0003387 x_3 x_5 + 0.0003387 x_5^2$
- $0.0001844 x_5 x_5 (x_5 - \log(x_2)) - 4.592$

| Antalya |  
| Pan evaporation | $= 1.691 \sqrt{x_1 \log(x_5)} + 26.55 \tanh \left( \frac{x_1}{x_5} \right)$
- $0.00008704(x_3)^2 + 0.0005501 x_1 x_3 + \frac{2.38 x_5}{x_5 + x_3 + 79.4 x_1}$
- $- 0.924 x_1 \tanh \left( \frac{x_1}{x_5} \right) + 0.00309$ if $\tanh(x_1) > 0.3719$
\( (x_1)^2, (x_2)^2, - x_3 - 4.767 \) - 21.27

\( P_{x_1}, x_2, x_3, x_4, x_5 \) indicate the maximum temperature, minimum temperature, solar radiation, wind speed, and relative humidity, respectively. \( \text{if} \ A < B \text{ then } C \text{ else } D \) on an element by element basis.
the models evolved are shown in Figure 7. As can be seen, the best model with respect to complexity and population has 46 nodes shown by a big black circle. Table 3 provides the equations of the optimal MGGP model for both stations. Table 5 compares the MGGP, GP, DENFIS, and CHS models in estimating pan evaporation of Antalya station. From the results, it is clear that the DENFIS has the best accuracy in both training and test periods. CHS is also better than the GP in the test period with respect to RMSE and NSE statistics. Figure 8 demonstrates scatter-plots for the test results of the optimal models. It is apparent from the graphs that the DENFIS performs better than the other models with respect to data dispersion.

Sanikhani & Kisi (2012) and Kisi (2015b) have examined the effect of periodicity in forecasting monthly streamflows by including periodicity component (month number) as input to the AI-based models and they have reported that

<table>
<thead>
<tr>
<th>Model</th>
<th>Training RMSE</th>
<th>Training MAE</th>
<th>Training NSE</th>
<th>Training $R^2$</th>
<th>Test RMSE</th>
<th>Test MAE</th>
<th>Test NSE</th>
<th>Test $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGGP</td>
<td>0.712</td>
<td>0.514</td>
<td>0.898</td>
<td>0.910</td>
<td>0.674</td>
<td>0.533</td>
<td>0.837</td>
<td>0.948</td>
</tr>
<tr>
<td>GP</td>
<td>0.790</td>
<td>0.594</td>
<td>0.875</td>
<td>0.875</td>
<td>0.644</td>
<td>0.495</td>
<td>0.851</td>
<td>0.903</td>
</tr>
<tr>
<td>DENFIS</td>
<td>0.874</td>
<td>0.560</td>
<td>0.863</td>
<td>0.864</td>
<td>0.663</td>
<td>0.510</td>
<td>0.842</td>
<td>0.919</td>
</tr>
<tr>
<td>CHS</td>
<td>3.107</td>
<td>2.490</td>
<td>−0.63</td>
<td>0.002</td>
<td>2.994</td>
<td>2.469</td>
<td>−0.65</td>
<td>0.493</td>
</tr>
</tbody>
</table>
it significantly increased the models’ accuracy. Therefore, here, also the effect of periodicity on models’ accuracy was investigated by adding month number of the year as input to the models for each data set. The same parameters were set for the MGGP and GP methods, as provided in Table 2. The comparison results of the periodic models are provided in Table 5 for Antakya station. It is apparent from the table that the periodic DENFIS model has better performance than the periodic MGGP and GP models. MGGP also performs better than the GP in estimation of pan evaporation. Comparison with Table 4 clearly indicates that adding periodicity input to the models obviously increases their performances in estimating pan evaporation. Periodic DENFIS model increases the RMSE, MAE, and NSE accuracy of the DENFIS by 24, 22, and 8%, respectively. The results are graphically compared in Figure 9. It can be seen from the fit line equation that the DENFIS estimates are closer to the exact line \((y = x)\). Comparison with Figure 5 obviously indicates that the periodicity input increases the accuracy of DENFIS and GP methods while the MGGP accuracy decreases in estimating monthly pan evaporation. Table 6 compares the training and test results of the periodic models for Antalya station. Unlike the Antakya, here, periodic MGGP provides better accuracy than the GP and DENFIS models with respect to various statistics. Comparison with Table 5 obviously shows that including periodicity component as input considerably increases the accuracy of DENFIS and MGGP. Periodic MGGP model increases the RMSE, MAE, and NSE accuracy of the MGGP by 51, 53, and 144%, respectively. Periodic models’ estimates are shown in Figure 10 for Antalya station.

In a study by Bruton et al. (2000), neural network models were used for estimating pan evaporation of Pome, Plains, and Watkinsville, Georgia using climatic data (e.g., temperature, rainfall, solar radiation, RH, and wind speed). \(R^2\) was calculated as 0.717 for the most accurate model. In a study by Dogan et al. (2007), feed forward neural networks (FFNNs) and radial basis neural networks (RBNNs) were applied for estimating pan evaporation of Lake Sapanca, Turkey using climatic data of min and max temperature, RH, wind speed, real solar period, and maximum solar period. \(R^2\) was computed as 0.651 and 0.716 for the best FFNNs and RBNNs models, respectively. From the tables (Tables 4–7), it is apparent that the obtained MGGP, GP, and DENFIS models provide accurate results in estimating pan evaporation process from the \(R^2\) viewpoint.

The generalization capacity of the applied models with respect to number of parameters or weights was also investigated. For this, Akaike information criterion (AIC) defined

<table>
<thead>
<tr>
<th>Model</th>
<th>Training RMSE</th>
<th>MAE</th>
<th>NSE</th>
<th>(R^2)</th>
<th>Test RMSE</th>
<th>MAE</th>
<th>NSE</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGGP</td>
<td>1.311</td>
<td>1.016</td>
<td>0.736</td>
<td>0.931</td>
<td>2.243</td>
<td>1.869</td>
<td>0.347</td>
<td>0.929</td>
</tr>
<tr>
<td>GP</td>
<td>0.727</td>
<td>0.564</td>
<td>0.919</td>
<td>0.919</td>
<td>1.374</td>
<td>1.116</td>
<td>0.755</td>
<td>0.945</td>
</tr>
<tr>
<td>DENFIS</td>
<td>0.707</td>
<td>0.545</td>
<td>0.923</td>
<td>0.925</td>
<td>1.163</td>
<td>0.944</td>
<td>0.824</td>
<td>0.943</td>
</tr>
<tr>
<td>CHS</td>
<td>1.000</td>
<td>5.014</td>
<td>0.846</td>
<td>0.864</td>
<td>1.320</td>
<td>4.767</td>
<td>0.775</td>
<td>0.925</td>
</tr>
</tbody>
</table>
Akaike (1974) was used. This criterion is shown in the following equation:

$$AIC = n \ln(MSE) + 2k$$

where \( n \) is the number of samples in the testing set, \( MSE \) is mean square error, and \( k \) is the number of model parameters. This criterion can be successfully used for the evaluation of soft computing models together with their system size (Kisi & Guven, 2010). In addition to RMSE, MAE, NSE, and \( R^2 \), another new criterion was used for evaluation of the models. This criterion combines the RMSE, MAE, and \( R^2 \) statistics and provides general evaluation of the applied models similar to the ideal point error (IPE) (Domínguez et al., 2011). This criterion can be seen in the following equation:

$$CA = 0.33(RMSE + MAE + (1 - R^2))$$

where \( CA \) is the combined accuracy. The AIC and CA of the applied models are reported in Table 8 for both stations. It is apparent from the table that the MGGP and GP have a lesser number of parameters compared to DENFIS. Therefore, their AIC are less than the latter method. It can be said that the DENFIS method has a more complex structure and more uncertainty exists for this method compared to the others. The MGGP, GP, and DENFIS models have less CA than the CHS in the test period. Adding periodicity component generally increases models' accuracy while an accuracy decrement

---

**Table 6** | Comparison of statistical errors for periodic MGGP, GP, DENFIS, and HS in estimation of pan evaporation – Antakya station

<table>
<thead>
<tr>
<th>Model</th>
<th>Training</th>
<th></th>
<th></th>
<th></th>
<th>Test</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAE</td>
<td>NSE</td>
<td>( R^2 )</td>
<td>RMSE</td>
<td>MAE</td>
<td>NSE</td>
<td>( R^2 )</td>
</tr>
<tr>
<td>Periodic MGGP</td>
<td>0.905</td>
<td>0.654</td>
<td>0.835</td>
<td>0.889</td>
<td>0.574</td>
<td>0.444</td>
<td>0.882</td>
<td>0.899</td>
</tr>
<tr>
<td>Periodic GP</td>
<td>0.804</td>
<td>0.595</td>
<td>0.870</td>
<td>0.870</td>
<td>0.611</td>
<td>0.523</td>
<td>0.866</td>
<td>0.952</td>
</tr>
<tr>
<td>Periodic DENFIS</td>
<td>0.894</td>
<td>0.599</td>
<td>0.839</td>
<td>0.840</td>
<td>0.503</td>
<td>0.396</td>
<td>0.909</td>
<td>0.794</td>
</tr>
</tbody>
</table>

---

**Figure 8** | Scatterplots of observed and predicted pan evaporation by GP, MGGP, DENFIS, and CHS models – Antalya station.

**Figure 9** | Scatterplots of observed and predicted pan evaporation by periodic GP, MGGP, and DENFIS models – Antakya station.
is seen for the DENFIS method. The reason for this may be the high number of parameters which implies that the DENFIS has a highly complex structure.

### CONCLUSION

The study compared the ability of MGGP and DENFIS in modeling pan evaporation and compared them with GP and CHS equation. The monthly maximum temperature, minimum temperature, solar radiation, wind speed, RH, and pan evaporation data from Antakya and Antalya stations, Mediterranean Region of Turkey were used in the applications. The influence of periodicity component on the models’ prediction accuracy was also examined in the study. Involving periodicity in the inputs considerably improved the accuracy of DENFIS and MGGP models in Antakya and Antalya stations, respectively. The DENFIS model with periodic input performed superior to the periodic MGGP and GP models in Antakya station, while the periodic MGGP model provided better accuracy than the periodic DENFIS and GP models at Antalya station. Periodic DENFIS model decreased the RMSE of the periodic GP model from 0.611 mm to 0.503 mm at Antakya station. For Antalya, the RMSE of the periodic GP model was decreased from 1.386 mm to 1.089 mm using periodic MGGP. The models' complexity was also investigated by using AIC and it was seen that the

### Table 7 | Comparison of statistical errors for periodic MGGP, GP, DENFIS and HS in estimation of pan evaporation – Antalya station

<table>
<thead>
<tr>
<th>Model</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAE</td>
</tr>
<tr>
<td>Periodic MGGP</td>
<td>0.571</td>
<td>0.437</td>
</tr>
<tr>
<td>Periodic GP</td>
<td>0.695</td>
<td>0.543</td>
</tr>
<tr>
<td>Periodic DENFIS</td>
<td>0.715</td>
<td>0.565</td>
</tr>
</tbody>
</table>

### Table 8 | Test results of the MGGP, GP, DENFIS, and CHS with respect to CA and AIC in estimation of pan evaporation

<table>
<thead>
<tr>
<th>Model</th>
<th>k</th>
<th>Antakya</th>
<th></th>
<th>Antalya</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>AIC</td>
<td>CA</td>
<td>MSE</td>
</tr>
<tr>
<td>MGGP</td>
<td>14</td>
<td>0.454</td>
<td>-4.40</td>
<td>0.415</td>
<td>9</td>
</tr>
<tr>
<td>GP</td>
<td>2</td>
<td>0.415</td>
<td>-32.1</td>
<td>0.408</td>
<td>3</td>
</tr>
<tr>
<td>DENFIS</td>
<td>138</td>
<td>0.440</td>
<td>242</td>
<td>0.414</td>
<td>148</td>
</tr>
<tr>
<td>CHS</td>
<td>3</td>
<td>8.964</td>
<td>45.1</td>
<td>2.180</td>
<td>3</td>
</tr>
<tr>
<td>Periodic MGGP</td>
<td>11</td>
<td>0.329</td>
<td>-23.5</td>
<td>0.369</td>
<td>12</td>
</tr>
<tr>
<td>Periodic GP</td>
<td>2</td>
<td>0.373</td>
<td>-36.4</td>
<td>0.390</td>
<td>3</td>
</tr>
<tr>
<td>Periodic DENFIS</td>
<td>161</td>
<td>0.253</td>
<td>266</td>
<td>0.711</td>
<td>139</td>
</tr>
</tbody>
</table>

---

Image: Figure 10 | Scatterplots of observed and predicted pan evaporation by periodic GP, MGGP, and DENFIS models – Antalya station.
DENFIS model has a highly complex structure and high number of parameters. The main advantage of the MGGP in addition to its high accuracy is that it has a very simple structure, and therefore, it could be easily utilized in practical applications. The CHS model provided the worst accuracy in both stations with respect to a new criterion which combines the RMSE, MAE, and R$^2$ statistics. The obtained results in the two stations are different to each other, especially in terms of the RMSE, MAE, and R$^2$ statistics. The obtained results in the both stations with respect to a new criterion which combines structure, and therefore, it could be easily utilized in practical applications in addition to its high accuracy is that it has a very simple number of parameters. The main advantage of the MGGP requires more comparison using different data from different regions for justifying their generalization. The MGGP and DENFIS models may be incorporated as modules in general hydrological analysis models.

REFERENCES


Beven, K., Smith, P. & Freer, J. 2008 So just why would a modeller choose to be incoherent? J. Hydrol. 354, 15–32.


Kisi, O. 2015a Pan evaporation modelling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. J. Hydrol. 528, 512–320.


Searson, D. P. 2009 GPTIPS: Genetic Programming & Symbolic Regression for MATLAB.


First received 13 April 2017; accepted in revised form 5 October 2017. Available online 14 November 2017.