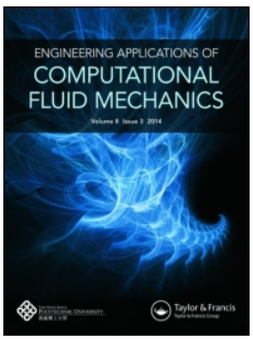

Taylor & Francis
Taylor & Francis Group

Engineering Applications of Computational Fluid Mechanics

ISSN: 1994-2060 (Print) 1997-003X (Online) Journal homepage: https://www.tandfonline.com/loi/tcfm20

Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions

Wang Jing, Zaher Mundher Yaseen, Shamsuddin Shahid, Mandeep Kaur Saggi, Hai Tao, Ozgur Kisi, Sinan Q. Salih, Nadhir Al-Ansari & Kwok-Wing Chau

To cite this article: Wang Jing, Zaher Mundher Yaseen, Shamsuddin Shahid, Mandeep Kaur Saggi, Hai Tao, Ozgur Kisi, Sinan Q. Salih, Nadhir Al-Ansari & Kwok-Wing Chau (2019) Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions, Engineering Applications of Computational Fluid Mechanics, 13:1, 811-823, DOI: 10.1080/19942060.2019.1645045

To link to this article: https://doi.org/10.1080/19942060.2019.1645045

Published online: 08 Aug 2019.

ENGINEERING APPLICATIONS OF COMPUTATIONAL FLUID MECHANICS
2019, VOL. 13, NO. 1, 811–823
https://doi.org/10.1080/19942060.2019.1645045

Taylor & Francis
Taylor & Francis Group

Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions

Wang Jing[a,b], Zaher Mundher Yaseen [c], Shamsuddin Shahid [d], Mandeep Kaur Saggi[e], Hai Tao[a], Ozgur Kisi[f], Sinan Q. Salih[g], Nadhir Al-Ansari[h] and Kwok-Wing Chau[i]

[a]Department of Computer Science, Baoji University of Arts and Sciences, Shaanxi, People's Republic of China; [b]Faculty of Computer Systems & Software Engineering, University Malaysia Pahang, Pahang, Malaysia; [c]Sustainable Developments in Civil Engineering Research Group, Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam; [d]School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia; [e]Department of Computer Science, Thapar Institute of Engineering and Technology, Patiala, India; [f]Faculty of Natural Sciences and Engineering, Ilia State University, Tbilisi, Georgia; [g]Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; [h]Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, Lulea, Sweden; [i]Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China

ABSTRACT

Evapotranspiration is one of the most important components of the hydrological cycle as it accounts for more than two-thirds of the global precipitation losses. Indeed, the accurate prediction of reference evapotranspiration (ET_o) is highly significant for many watershed activities, including agriculture, water management, crop production and several other applications. Therefore, reliable estimation of ET_o is a major concern in hydrology. ET_o can be estimated using different approaches, including field measurement, empirical formulation and mathematical equations. Most recently, advanced machine learning models have been developed for the estimation of ET_o. Among several machine learning models, evolutionary computing (EC) has demonstrated a remarkable progression in the modeling of ET_o. The current research is devoted to providing a new milestone in the implementation of the EC algorithm for the modeling of ET_o. A comprehensive review is conducted to recognize the feasibility of EC models and their potential in simulating ET_o in a wide range of environments. Evaluation and assessment of the models are also presented based on the review. Finally, several possible future research directions are proposed for the investigations of ET_o using EC.

ARTICLE HISTORY
Received 12 May 2019
Accepted 14 July 2019

KEYWORDS
evapotranspiration prediction; state of the art; evolutionary computing models; input variability; future research directions

1. Introduction

Proper monitoring of climate requires comprehensive information about land surface fluxes, particularly the latent and sensible components (Courault, Seguin, & Olioso, 2005). This information is also important for the evaluation of the parameterization schemes in the climate and weather models used for the flux exchange prediction between the surface and the lower atmosphere (Allen, Burt, Solomon, Clemmens, & O'Halloran, 2005; Fisher et al., 2009). In agriculture, it is required for irrigation scheduling and other applications (Farg, Arafat, Abd El-Wahed, & El-Gindy, 2012). It is a major component of the hydrological cycle and, therefore, it has significant implications on water requirements and water resource management (Figure 1). Therefore, it is necessary that water managers and irrigation engineers are provided with an accurate and robust tool for estimating surface fluxes, and especially evapotranspiration (Pereira, Green, & Villa Nova, 2006).

The reference evapotranspiration (ET_o) can be determined either directly through experiments or indirectly through mathematical models. Practically, the major classical methods for ET_o measurement are only obtainable at the field scale (Bowen ratio, eddy correlation system, soil water balance) (Tao, Diop, et al., 2018). However, certain limitations restrict the application of these methods in flux prediction, especially when applied to large spatial scales (Kumar, Jat, & Shankar, 2012). Besides, the field methods are expensive, time consuming and difficult. ET_o is also difficult to measure mathematically because it depends on the interaction of several climatic factors, including temperature, wind speed, humidity and radiation (Tao, Diop, et al., 2018). Hence, evapotranspiration is featured by a nonlinear and complex phenomenon

CONTACT Zaher Mundher Yaseen yaseen@tdtu.edu.vn

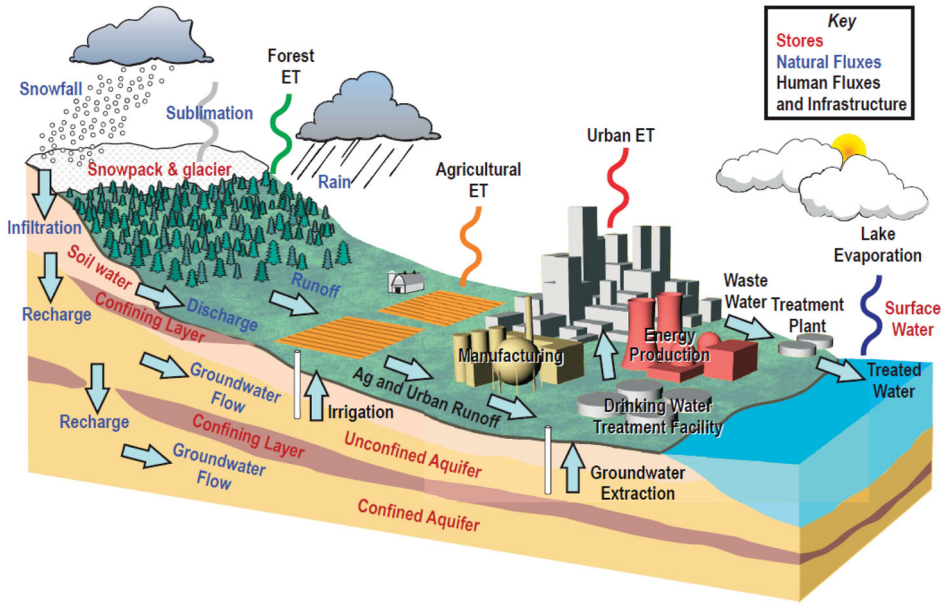

Figure 1. Schematic diagram of hydrological cycle showing the importance of evapotranspiration (ET).

and its determination is based on the availability of several climatic parameters and their mutual interactions with each other.

Experts have developed several ET_o estimation methods over the years; however, the selection of a suitable method depends mainly on the availability of the measured climatic factors. The generalized Penman–Monteith (PM) method is the most adopted method for the estimation of ET_o in agricultural and environmental research, as it tallies well with field observations (Penman, 1948). The PM model has been acknowledged as the standard ET_o estimation method despite the fact that its application requires the availability of a significant amount of climatic data, which may not be available in certain locations, such as in the developing countries. Such cases demand the deployment of alternative methods with less dependence on many weather inputs. The artificial intelligence (AI) methods can model highly nonlinear phenomena with a limited amount of data (Meng et al., 2019). Therefore, the application of AI methods has grown rapidly in recent years for the development of alternative methods of ET_o estimation from limited weather inputs.

2. State of the art: evolutionary computing (EC) models for ET_o simulation

The evolution of computer technology has increased the applicability of AI in different scientific fields (Baghban, Jalali, Shafiee, Ahmadi, & Chau, 2019; Chau, 2017; Chuntian & Chau, 2002; Haie, Pereira, Machado, & Shahidian, 2019; Wu & Chau, 2011). AI is a connection of several processors, distributed in parallel and composed of simple processing units with the natural capability to store experimental information and produce it whenever needed (Danandeh Mehr et al., 2018; Yaseen, Sulaiman, Deo, & Chau, 2019). Since AI models require few inputs and can map input–output relationships without having prior knowledge of the physical processes involved, they are considered as effective tools for nonlinear process modeling. Consequently, several alternative intelligent computational models for the estimation of ET_o have been developed over the last few decades (Chen, Chen, & Chen, 2018; Kumar, Raghuwanshi, & Singh, 2011; Sun et al., 2019).

EC approaches, such as gene expression programming (GEP) and genetic programming (GP), are one of several types of AI model which have been considered appropriate for the modeling of ET_o (Danandeh Mehr et al., 2018). These models were originally developed by Koza (1992), when the concept of GP was first introduced. They were conceptually developed based on the genetic algorithm (GA), an algorithm used for the implementation of symbolic regression when trying to establish a mathematical function that fits a data set (Maulik & Bandyopadhyay, 2000; Najafzadeh & Kargar, 2019). Several attempts have been made regarding the feasibility of using EC to solve problems such as circuit design, multi-agent strategies and time-series prediction (Dineva et al., 2019; Mirjalili, Mirjalili, Saremi, & Mirjalili, 2020). Similarly, several researchers have examined the applicability of EC to solve hydrological, climatological, environmental and ecological problems over the past two decades (Yaseen, El-shafie, Jaafar, Afan, & Sayl, 2015).

Within the scope of modeling the ET_o process, a number of studies have been conducted using EC for different climatological scenarios. To investigate the effect of different climatic and topographic conditions (Parasuraman, Elshorbagy, & Carey, 2007), ET_o was estimated using GP from ground temperature, wind speed, eddy-covariance-measured latent heat as a function of net radiation, relative humidity and air temperature. Based on the predictability comparison with the classical artificial neural network (ANN) and PM empirical formulation, it was found to be a potential model (Parasuraman et al., 2007).

Guven, Aytek, Yuce, and Aksoy (2008) investigated the effect of ET_o using the daily atmospheric variables collected from the California Irrigation Management Information System (CIMIS) database for Davis, Hasting, Suisun, Dixon and Oakville stations. The authors validated the performance of the GA model against several empirical models such as the Food and Agriculture Organization of the United Nations Penman–Monteith (FAO56-PM), the PM, the Jensen–Haise, the Hargreaves–Samani (HS) equation, the Jones–Ritchie, the Turc method and solar radiation-based models. It was found that the GA model performed better than the considered empirical models, with relatively low error and high correlation metrics.

Kim and Kim (2008) proposed three type of GA model based on the generalized regression neural networks model (GRNNM) namely, COMBINE-GRNNM-GA, EXTREME-GRNNM-GA and Average-GRNNM-GA, for the estimation of pan evaporation and alfalfa reference evapotranspiration (ET_{ref}) for the period 1985–1992 at 14 meteorological stations located in the Republic of Korea. Based on statistical results, they found COMBINE-GRNNM-GA to be the best performing of these models. Kisi and Guven (2010) investigated the accuracy of linear genetic programming (LGP) with GEP for modeling ET_o at three stations, and indicated the superior performance of LGP over GEP.

El-Baroudy, Elshorbagy, Carey, Giustolisi, and Savic (2010) reported two case studies of the South West Sand Storage and the South Bison Hill, Mildred Lake mine, Canada, for the estimation of ET_o using three intelligent models, namely evolutionary polynomial regression (EPR), ANN and GEP. The authors found better performance of the EPR model over the other models in estimation of ET_o.

Izadifar and Elshorbagy (2010) estimated ET_o using a physical model, i.e. HYDRUS-1D, at Alberta, Canada, and compared its performance with different AI (ANN and GP) and statistical models. The results indicated that the multiple regression and GP models performed better than the ANN model in estimating ET_o.

Kisi (2010) used GA for the calibration of the fuzzy membership function for the modeling of daily ET_o at three stations located in central California, USA, for the period 1998–2007. They employed the fuzzy-genetic (FG) model to estimate the ET_o obtained using the FAO56-PM method and compared its performance with the Penman, Hargreaves, Ritchie, Turc and ANN methods. It was observed that the FG model, which uses only two parameters, performed better in estimating ET_o compared to other empirical models.

Shiri et al. (2012) applied the GEP model to determine the magnitude of ET_o in the Basque Country (northern Spain). The models were developed using weather parameters, namely relative humidity, solar radiation, air temperature and wind speed. The capability of the GEP model was validated against the other AI and empirical models. Overall, the results showed that the GEP model achieved better accuracy than the Priestley–Taylor (PT), HS and adaptive neuro-fuzzy inference system (ANFIS) models.

The generalizability of the GEP model was scrutinized by Traore and Guven (2012) for modeling the ET_o in Sahelian country, Burkina Faso. They reported that the incorporation of several weather variables provided the best accuracy using the GEP model. Jean et al. (2012) proved that the combination of GEP and FAO56-PM models can better estimate evaporation and seepage for Datong Basin. In another major study, Eslamian, Gohari, Zareian, and Firoozfar (2012) developed hybrid ANN and GA (ANN-GA) and ANN models for the estimation of ET_o. They found that the ANN-GA could calculate the ET_o more accurately than the other models.

Shiri et al. (2013) investigated the performance of GEP and ANFIS models for ET_o estimation at five stations in Iran. The authors found that the GEP model outperformed the empirical formulations (e.g. HS, Makkink and Turc) and ANFIS model.

The GEP model was developed for the estimation of ET_o over West Africa (Traore & Guven, 2013). The authors reported a high performance of the algebraic formulation obtained using GEP for ET_o simulation over the Sub-Saharan African regions.

A comparative analysis using several intelligent models, namely ANN, ANFIS, support vector machine (SVM) and GEP to estimate the ET_o was investigated by Shiri, Nazemi, et al. (2014). In addition, the capacity of the established AI models was validated against several empirical formulations, namely PT, HS, Turc and Makkink. The results showed the superiority of the GEP model over the other AI and empirical methodologies.

Shiri, Sadraddini, et al. (2014) applied GEP to estimate ET_o with different combinations of meteorological variables as input attributes to construct a predictive model.

The study reported an excellent performance of GEP using all available meteorological variables.

The performance of ANN, an integrative ANFIS with subtractive clustering and grid partition, and GEP models was inspected to simulate the long-term monthly ET_o at 50 stations in Iran (Kisi, Sanikhani, Zounemat-Kermani, & Niazi, 2015). The authors reported the feasibility of the GEP model over the other developed AI models. They also reported a remarkably consistent predictability in ET_o simulation using GEP.

Martí, González-Altozano, López-Urrea, Mancha, and Shiri (2015) conducted a study to compare the performance of GEP in estimation of ET_o with lysimeter data at two locations in Spain: Las Tiesas (Albacete) and La Orden (Badajoz). The study reported that the best results in the estimation of ET_o were obtained using GEP.

The performance of the PM model with four soft computing approaches, namely SVM–firefly algorithm (SVM-FFA), ANN, SVM-wavelet and GP, in the estimation of ET_o in Serbia was assessed by Gocić et al. (2015). They reported that SVM-wavelet and SVM-FFA performed better than GP and ANN.

Yassin, Alazba, and Mattar (2016b) developed mathematical equations using GEP for the estimation of ET_o from meteorological data at 13 stations in the Kingdom of Saudi Arabia, and reported accurate estimation of ET_o using the GEP model. However, they found a slightly better performance using ANN compared to GEP.

Alazba, Yassin, and Mattar (2016) developed ET_o models for eight combinations of inputs using GEP for 19 meteorological stations located in a hyper-arid region. The performance of the GEP model was compared with the PM model, and better formation of the GEP model with eight input variables was found.

Yassin et al. (2016b) investigated the performance of GEP and ANN in modeling daily ET_{ref} at 19 meteorological stations in Saudi Arabia for the period 1980–2010. The results showed that the ANN performed better than the GEP in estimating ET_{ref}.

Kumar, Adamowski, Suresh, and Ozga-zielinski (2016) investigated the performance of extreme learning machine (ELM), ANN, GP and SVM in estimating daily ET_o in North Bihar, India, for the period 2001–2005. They reported that the ELM model performed better than the other machine learning and soft computing approaches.

Karimi, Kisi, and Kim (2017) developed heuristic models combining SVM and GEP for the estimation of ET_o in the Republic of Korea. The result showed that the GEP model outperformed the SVM model in the local and cross-station scenarios.

Kiafar et al. (2017) derived equations for the estimation of ET_o at two meteorological stations located in a hyper-arid region of Iran and at two stations in a humid region of Spain. The authors compared the performance of GEP with three empirical models to calculate ET_o, namely the mass transfer, temperature-based and radiation-based models. They reported better performance of the GEP model with four inputs compared to the empirical models.

Mehdizadeh, Behmanesh, and Khalili (2017) investigated the performance of SVM [Poly and radial basis function (RBF)], GEP, multivariate adaptive regression spline (MARS) and the empirical models to determine the ET_o at 44 meteorological stations in Iran. The study reported that the MARS and SVM-RBF models performed better than the SVM-Poly and GEP models.

The daily scale of ET_o at Gaoyou station in China was predicted using the GEP algorithm (Traore, Luo, & Fipps, 2017). The study suggested that the GEP model can be used for estimation of ET_o and also it can be employed as a tool for short-term irrigation scheduling for decision making.

Different empirical and semi-empirical approaches namely, Kimberly–Penman, temperature-based and radiation-based models, were applied along with GEP by Shiri (2017) for the estimation of daily ET_o at five meteorological stations in Iran. The author divided the data set into three segments for model calibration, testing and validation. The comparison of the results showed that the temperature-based empirical model developed using GEP was the most accurate in estimating ET_o.

The random forest model (RFM) and generalized regression neural network (GRNN) models were employed by Feng, Cui, Gong, Zhang, and Zhao (2017) to estimate the ET_o for the period 2009–2014 at two stations in China. The results revealed that the temperature-based RFM and GRNN models could be applied successfully for daily ET_o estimation. However, the RFM could perform slightly better than the GRNN in estimating ET_o.

The performance of AI-based models such as ANN and GEP and ancillary/external approaches were investigated by Landeras et al. (2018) for the estimation of ET_o at four stations in Ghana. They used 7 years' data (2006–2009) for the training and 4 years' data (2009–2012) for the testing of the models. The study reported that the GEP and ancillary model could be applied for better estimation of ET_o in West Africa.

Mattar and Alazba (2018) examined the performance of the GEP model with eight input combinations for the simulation of monthly ET_o at 27 meteorological stations in Egypt using the CLIMWAT meteorological data. The authors found that the air temperature, wind speed and relative humidity are required for accurate modeling of ET_o. The results also indicated that multiple linear regression (MLR) and GEP with mean humidity and wind

speed at 2 m height as inputs provide the best estimation of ET_o.

Mattar (2018) implemented eight combinations of GEP models to calculate the ET_o at 32 stations in Egypt using the CLIMWAT data. The comparison of the performance of GEP model with other empirical models showed that the GEP model provides promising outcomes in the modeling of ET_o.

Mehdizadeh (2018) performed extensive research using MARS and GEP to determine the daily ET_o at six stations in Iran, namely Isfahan and Shiraz (arid), Urmia and Tabriz (semi-arid) and Yazd and Zahedan (hyper-arid), for the period 2000–2014. They used daily meteorological data and their lags for the modeling of ET_o, and concluded that MARS–autoregressive conditional heteroscedasticity (MARS-ARCH) and GEP-ARCH showed superior results compared to the standalone MARS and GEP models.

The potential of the GEP model for the estimation of ET_o was analyzed by Jovic, Nedeljkovic, Golubovic, and Kostic (2018). The models were built based on diverse meteorological variables including minimum and maximum temperature, vapor pressure, wind speed, sunshine hours and relative humidity. The authors evidenced the improved capability of the EC model through the incorporation of more climatic information.

Mohammad, Pour, Piri, and Kisi (2018) investigated the performance of three data-driven models, SVM, ANFIS and GEP, to estimate the ET_o. They used five different combinations of input for the simulation of ET_o for the period 1970–2010 in south-eastern Iran. The results showed that the SVM model had superior performance with the input combination of average air temperature, relative humidity, wind speed and sunshine hours.

Six AI models, namely multilayer perceptron, GRNN, radial basis neural network (RBNN), ANFIS with grid partition (ANFIS-GP), ANFIS with subtractive clustering (ANFIS-SC) and GEP, were investigated by Sanikhani, Kisi, Maroufpoor, and Yaseen (2019) for the determination of ET_o at two meteorological stations located in the Mediterranean region, at Antalya and Isparta in Turkey. The prediction results revealed that the best performance belonged to the GRNN and GEP models at Antalya station, and ANFIS-SC and RBNN models at Isparta station.

Daily ET_o was calculated using PM-FAO56, radiation-based, mass transfer-based and temperature-based models at five stations in Iran (Shiri, 2019). The performance of the empirical models was validated against the predominant EC model (i.e. GEP). The author reported the better performance of the mass transfer-based model and relatively lower performance of GEP, temperature-based and radiation-based models.

Shiri (2019) presented the results of ET_o models at five stations in Iran. The author calculated the daily ET_o with FAO56-PM and compared the results with the temperature-, radiation- and mass transfer-based ET_o equations and GEP-derived models. Overall, the results indicated that the temperature-based and the radiation-based empirical and GEP models could not provide satisfactory results to simulate ET_o, while the models derived from mass transfer equations could provide more accurate results.

Shiri, Marti, Karimi, and Landeras (2019) introduced a new approach for the estimation of ET_o by splitting the data set and incorporating external ancillary inputs. They applied GEP with the temperature-based HS model and the radiation-based PT model for the estimation of ET_o from the local meteorological inputs. Comparison of the performance accuracy of the models revealed that the GEP-based model produced the most accurate results among all the applied approaches.

The present review also investigated the successful application of evolutionary algorithms in other fields. For instance, Danandeh Mehr, Kahya, and Ozger (2014) introduced a new explicit gene-wavelet model for drought forecasting, Mehr (2018) developed a new hybrid GA combined with GEP for stream flow forecasting in intermittent streams, and other studies investigated the capability of GP in advances in rainfall runoff modeling (e.g. Jayawardena, Muttil, & Lee, 2006; Nourani, Komasi, & Alami, 2013).

3. Survey evaluation and assessment

The summary of the research which implemented the EC models for the estimation of ET_o over the period 2007–2018 is reported in Table 1. It was observed that a large number of studies has been conducted on the implementation of GP and GEP models for estimation of ET_o. The main reason for preferring evolutionary models in estimating ET_o their capability to provide explicit formulation and ease of application. The formulation provided by GEP can be simply used in practical applications.

Most studies have demonstrated a limitation in the predictability of EC models using fewer meteorological variables as input. This was observed more prominently in the regions characterized by arid and semi-arid climates. This can best be explained by the ET_o processes being influenced by multiple climate variables and thus varying from one case to another. Hence, it can be concluded, based on the observations, that ET_o models in such regions need several pieces of climate information to attain a reliable predictive capability.

Table 1. Short review of the implementation of evolutionary computing models for the estimation of reference evapotranspiration (2007–2019).

Authors (year)	Predictive models	Input/output variables	Timescale	Region	Best predictive model	Performance metrics
Parasuraman et al. (2007)	GP, ANN	NR, ST, T, WS, RH/EC-LE	Hourly	Canada	GP	RMSE, MARE, R
Guven et al. (2008)	GEP	SR, T, ST, RH, VP, WS, WD, RF/ET_o	Daily	California	GA	MSE, R^2
Kim and Kim (2008)	COMBINE-GRNNM-GA, EXTREME-GRNNM-GA and Average-GRNNM-GA based on GRNN and GA	T_{max}, T_{mean}, T_{min}, RH_{min}, RH_{mean}, SD, WS_{mean}, WS_{max}, SD, DP_{mean}	Daily	Republic of Korea	COMBINE-GRNNM-GA	RMSE, MAE, E, CC
Kisi and Guven (2010)	LGP, SVR, ANN	T_{max}, RH, SR, WS/ET_o	Daily	Windsor, Oakville, Santa Rosa	LGP	RMSE, R^2, MAE
El-Baroudy et al. (2010)	GEP, EPR, ANN	NR, T, ST, WS, RH/ET_o	Hourly	Canada	EPR	RMSE, MARE, R
Izadifar and Elshorbagy (2010)	GP, ANN, statistical models, Penman	NR, ST, T, RH, WS/EC-AET	Hourly	Canada	GP	RMSE, MARE, R
Kisi (2010)	FG1, FG2, ANN, Hargreaves, Ritchie, Turc	T_{max}, RH, SR, WS/ET_o	Daily	California	FG2	RMSE, MAE, R
Shiri et al. (2012)	GEP, ANFIS, empirical formulations	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Daily	Spain	GEP	RMSE, R^2, SI
Traore and Guven (2012)	GEP	NR, T_{mean}, RH, WS/ET_o	Daily	Burkina Faso	GEP	RMSE, R^2
Jean et al. (2012)	GEP, PM	T_{mean}, SR, soil HF, NR, WS, evaporation seepage loss	Monthly	Datong Basin	GEP	–
Eslamian et al. (2012)	ANN-GA, ANN	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Monthly	Iran	ANN-GA	MSE, NMSE, MAE, min/max absolute error, R^2
Shiri et al. (2013)	GEP, ANFIS, empirical formations	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Daily	Iran	GEP	MAE, SI, NS, R^2
Traore and Guven (2013)	GEP	T_{max}, T_{min}, RF, RH, Sh, WS/ET_o	Daily	Burkina Faso	GEP	RMSE, R^2
Shiri, Nazemi, et al. (2014)	ANN, ANFIS, SVR, GEP, empirical formulations	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Daily	Iran	GEP	RMSE, R^2, SI, MAE
Shiri, Sadraddini, et al. (2014)	GEP_{1-3}	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Daily	Iran	GEP	RMSE, R^2, AARE, MAE
Kisi et al. (2015)	ANN, ANFIS-GP, ANFIS-SC, GEP	Data periodic, station latitude, longitude and altitude/ET_o	Monthly	Iran	ANFIS-GP, ANFIS-SC	RMSE, R^2
Martí et al. (2015)	GEP, empirical formulations	ET_o/ET_o	Daily	Spain	GEP	RMSE, R^2, MAE
Gocić et al. (2015)	GP, SVR-FFA, ANN, WA-SVR, Penman	T_{max}, T_{min}, VP, Sh, WS/ET_o	Monthly	Serbia	SVM-wavelet	RMSE, R^2, r, MAE, MAPE
Yassin, Alazba, and Mattar (2016a)	ANN, GEP, Penman	T_{max}, T_{mean}, T_{min}, RH_{max}, RH_{mean}, RH_{min}, SR, WS/ET_o	Daily	Saudi Arabia	ANN	RMSE, R^2, OI, MAE
Alazba et al. (2016)	GEP	T_{max}, T_{mean}, T_{min}, RH_{max}, RH_{mean}, RH_{min}, SR, WS/ET_o	Daily	Saudi Arabia	GEP	RMSE, R^2, MAE
Yassin et al. (2016b)	GEP	T_{max}, T_{mean}, T_{min}, RH_{max}, RH_{mean}, RH_{min}, SR, WS/ET_o	Daily	Saudi Arabia	GEP_{MOD8}	RMSE, R^2, OI, MAE
Kumar et al. (2016)	ELM, ANN, SVM, GP	T_{max}, T_{min}, RH_{max}, RH_{min}, SR, IS, rainfall, WS/ET_o	Daily	India	ELM	RMSE, R^2, Time
Karimi et al. (2017)	SVM, GEP	T_{max}, T_{min}, RH, SR, WS/ET_o	Daily	Republic of Korea	GEP	MAE, CRM

(*continued*).

Table 1. Continued.

Authors (year)	Predictive models	Input/output variables	Timescale	Region	Best predictive model	Performance metrics
Kiafar et al. (2017)	GEP, mass transfer, temperature- and radiation-based models	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Daily	Iran, Spain	GEP	RMSE, R^2, CRM
Mehdizadeh et al. (2017)	MARS, SVR, GEP	T_{max}, T_{mean}, T_{min}, RH, SR, WS/ET_o	Monthly	Iran	MARS	RMSE, R^2, MAE
Traore et al. (2017)	GEP, Penman	T_{max}, T_{mean}, T_{min}, RH, Sh, WS/ET_o	Daily	China	GEPMLP	RMSE, *R*, MAE, RRSE
Shiri (2017)	KP temperature-based radiation-based and GEP_{1-8}	T_{max}, T_{min}, T_{mean}, RH_{mean}, SR, WS/ET_o	Daily	Iran	GEP	SI, MAE, NS
Feng et al. (2017)	RF, GRNN	T_{max}, T_{min}, RH, SR, SD, WS/ET_o	Daily	China	RF	RRMSE, MAE, NS
Landeras et al. (2018)	ANN, GEP	T_{max}, T_{min}, T_{mean}, RH_{max}, RH_{min}, RH_{mean}, SD, WS/ET_o	Daily	Ghana	GEP	RMSE, MAE, SI
Mattar and Alazba (2018)	GEP, MLR, Penman	T_{max}, T_{min}, RH_{mean}, SR, WS/ET_o	Monthly	Egypt	GEP	RMSE, *R*, MAE
Mattar (2018)	GEP, empirical formulations	T_{max}, T_{min}, RH_{mean}, SR, WS/ET_o	Monthly	Egypt	GEP	RMSE, OI, IAs
Mehdizadeh (2018)	MARS, GEP	T, RH, SR, WS/ET_o	Daily	Iran	GEP1-ARCH	RMSE, R^2, MAE, MAPE
Jovic et al. (2018)	GP	T_{max}, T_{mean}, T_{min}, RH, Sh, WS/ET_o	Monthly	–	GP_6	RMSE, R^2
Mohammad et al. (2018)	SVM, ANFIS, GEP	T_{max}, T_{mean}, T_{min}, RH, Sh, WS/ET_o	Daily	Zabol, Zahedan, Iranshahr, Chabahar	SVM	RMSE, R^2, MAE
Shiri (2019)	GEP	T_{max}, T_{mean}, T_{min}, RH, Sh, WS/ET_o	Monthly	Island	E1-GEP4, E2-GEP4	IOA, RMSE, NS
Sanikhani et al. (2019)	MLP, GRNN, RBFNN, ANFIS-GP, ANFIS-SC, GEP	T_{max}, T_{mean}, T_{min}, RH, Sh, WS/ET_o	Monthly	Antalya, Isparta	RBNN, GEP	RMSE, R^2, MAE, CRM, NS
Shiri et al. (2019)	GEP, temperature- and radiation-based models	T_{max}, T_{mean}, T_{min}, Ra, SR/ET_o	Daily	North-western Iran	GEP	SI, NS

Notes: AARE = average absolute relative error; ANFIS = adaptive neuro-fuzzy inference system; ANFIS-GP = adaptive neuro-fuzzy inference system integrated with grid partition; ANFIS-SC = adaptive neuro-fuzzy inference system integrated with subtractive clustering; ANN = artificial neural network; ARCH = autoregressive conditional heteroscedasticity; CC = correlation coefficient; CRM = coefficient of residual mass; DP = dew point; E = efficiency; EC-AET = eddy covariance-measured actual evapotranspiration; EC-LE = eddy-covariance-measured latent heat; ELM = extreme learning machine; EPR = evolutionary polynomial regression; ET_o = reference evapotranspiration; FFA = firefly algorithm; FG = fuzzy genetic; GA = genetic algorithm; GEP = gene expression programming; GP = genetic programming; GRNN = generalized regression neural network; GRNNM = generalized regression neural networks model; HF = heat flux; IAs = index of agreements; IOA = index of agreement; SI = scatter index; KP = Kimberly–Penman; LGP = linear genetic programming; MAE = mean absolute error; MAPE = mean absolute percentage error; MARE = mean absolute relative error; MARS = multivariate adaptive regression spline; MLP = multilayer perceptron; MLR = multiple linear regression; MSE = mean square error; NMSE = normalized mean square error; NR = net radiation; NS = Nash–Sutcliffe efficiency coefficient; OI = overall index; PM = Penman–Monteith; Ra = extraterrestrial radiation; R^2 = coefficient of determination; RBFNN = radial basis function neural network; RBNN = radial basis neural network; RF = rainfall; RH = relative humidity; RMSE = root mean square error; RRMSE = relative root mean square error; RRSE = root relative squared error; SD = sunshine duration; Sh = sunshine; SI = scatter index; SR = solar radiation; ST = soil temperature; SVM = support vector machine; SVR = support vector regression; T = temperature; T_{max}, T_{mean} and T_{min} = maximum, mean and minimum air temperature; VP = vapor pressure; WA = wavelet; WD = wind direction; WS = wind speed.

In Table 1, the GP or GEP models have been compared with several other methods in estimating ET_o for different timescales (hourly, daily and monthly) and different climatic regions, ranging from hyper-arid to humid. In most of the studies, such models were found to be superior compared to other empirical and AI models. Only three studies reported a lower accuracy of GP/GEP compared to other AI methods in the estimation of ET_o. El-Baroudy et al. (2010) compared GEP with EPR and ANN for modeling hourly ET_o, and reported marginal differences among the models. However, a limited number of data was used in the study and, therefore, generalization of the obtained results is difficult. The data division rule (65% for training and 35% for testing) was also different from the related literature. Mehdizadeh (2018) compared GEP, SVM and MARS for the estimation of long-term monthly ET_o (mean values for the period 1951–2010 and, thus, 60 values for each month) and found that MARS showed superior performance to the other methods. Kisi et al. (2015) compared the performance of ANN and ANFIS with GEP for the estimation of long-term monthly ET_o (mean values for the period 1956–2010 and, thus, 55 values for each month) without climatic data as input, and found GEP to provide less accuracy in ET_o estimation compared to ANN and ANFIS. Therefore, it may be remarked that GEP may not have good capability in the modeling of long-term monthly ET_o in some climatic environments. The review also revealed that the training span of data has a substantial influence on the model's predictability.

Most of the studies reviewed in the present study (Table 1) used full climatic data as inputs to calculate ET_o by the FAO56-PM method. The main reason for searching for alternatives to the standard method is that it requires a high number of climatic data as inputs and such data are not available at many stations, especially those located in developing countries. Therefore, the development of AI models using all available data, similar to the standard method (FAO56 PM), is not worthwhile. Because ET_o can be calculated using the FAO56 PM method if the data of all meteorological variables are available, the alternative methods are not necessary for the replacement of this standard method. As an example, Sanikhani et al. (2019) used several AI methods including the EC method for the development of the ET_o model using only temperature data, and their performances were compared with the corresponding HS empirical method, which uses the same inputs. Promising results, such as determination coefficient higher than 0.900, were obtained using the models, which means that the models could explain 90% of variations in the test data. However, the development of new models to attain similar accuracy to the established model using the same input variables is not worthwhile. Rather, emphasis should be placed on the development of models with better accuracy than the established model or the development of models with fewer inputs but similar accuracy.

The GEP models developed for the estimation of ET_o in all the studies are very complex. Many mathematical and logical operators and constants were incorporated in equations when attempting to develop models using few meteorological variables. For example, Mattar and Alazba (2018) utilized eight operators [+, −, ×, %, $\sqrt{}$, x^2, power and ln(x)]. In some other studies, trigonometric (sin, cos and tan) and logical (AND and OR) operators were also included to generate the ET_o equation. This has made the equations very complex and discouraged hydrologists and water professionals from implementing them in the field. Emphasis should be placed on developing equations with fewer operators to make them more usable in the field. In addition, uncertainty in the model output can be estimated and provided to end users to make them aware of the range of error in the estimated ET_o values using GEP-derived equations.

Another important issue is the development of separate ET_o models using EC for each station. To date, a limited number of studies has considered the development of a generalized model which can be used for the estimation of ET_o accurately at all stations in a region. This is very important, especially for developing countries where the meteorological variables required for the estimation of ET_o are missing or not available for long periods for technical reasons. In such cases, generalized models including explicit equations obtained by pooled data or calibrated with data from nearby stations are very useful. Therefore, alternative methods which use information (e.g. limited climatic input such as temperature data) from both local and nearby stations should be further investigated for the modeling of ET_o (Sanikhani et al., 2019).

As briefly mentioned earlier, various data-division scenarios have been applied in previous studies. For example, Shiri, Sadraddini, et al. (2014) and Yassin et al. (2016b) divided data into two sets, training and test or calibration and validation, while other researchers used three data divisions, training, validation and test, in their applications. Others applied leave-one-out process (LOOP) or cross-validation (CV) methods for the development and evaluation of the model. A two-data-division procedure was preferred by most researchers in modeling ET_o. The main drawback of this type of modeling is the evaluation of methods without independent data sets. By applying three data divisions, the methods are calibrated using the first part (training data) of the whole data set, the optimal ones are decided using the second part (validation data) and the optimal models

are tested using the last part (test data) of the data set. The main disadvantages of the LOOP or CV approach are the use of fewer data for calibration and the requirement for a longer time for calibration. Therefore, it is recommended that the three-data-division rule or different training–testing scenarios should be used to obtain more robust models.

4. Possible future research directions

A comprehensive review is conducted in this study to assess the potential of EC techniques in modeling ET_o. The study revealed that EC has been used for the development of ET_o models for a wide climatic range. The evaluation of the performance of the models as presented in this study revealed the high potential of EC in modeling ET_o. Although significant progress has been achieved in the development of ET_o models using EC, the synthesis of current understanding based on the literature reviewed in the present study indicates the need for future research in the following directions.

(1) In most of the previous studies, ET_o models were developed with a large number of input variables. Many of these variables are not easily available in most of the meteorological stations, particularly in developing countries. The major challenge in ET_o modeling is the reliable estimation of ET_o with easily available meteorological variables. Future research should be directed towards the development of ET_o models with easily available meteorological variables, such as maximum and minimum temperature.

(2) Although some temperature-based empirical models are available for the estimation of ET_o, these are not useful for the projection of ET_o owing to climate change. All the temperature-based models show a decrease in ET_o under climate change scenarios, even though the temperature would rise. This is due to the higher influence of diurnal temperature range on ET_o, which is projected to decline as a result of climate change. Therefore, it is important to develop simple temperature-based ET_o models which can be used easily for the reliable projection of ET_o under climate change scenarios.

(3) Most of the ET_o models developed using EC are station specific. To increase the applicability of EC-based ET_o models, it is important to develop a generalized model which can be used for the reliable estimation of ET_o for a whole country, or at least in a homogeneous climatic region. This needs calibration and validation of the EC-based model with all the station data available in a region.

(4) Although reliable accuracy has been attained in the estimation of ET_o using EC, there is still room to improve the accuracy of the models through optimization of EC parameters using sophisticated optimization algorithms. A possible direction is the use of other recent EC algorithms which have not yet been employed for modeling the ET_o process, such as particle swarm optimization (Ali Ghorbani, Kazempour, Chau, Shamshirband, & Taherei Ghazvinei, 2018; Chau, 2007) and ELM (Taormina & Chau, 2015; Yaseen, Sulaiman, et al., 2019). Many recently developed and popular EC algorithms have yet to be applied to this area, and their performance on this particular topic is yet to be determined and compared with the performance of their counterparts.

(5) While some hybrid EC models have been developed individually for ET_o modeling, the effectiveness of many other feasible combinations that have been applied to other fields (Ghorbani, Deo, Karimi, Yaseen, & Terzi, 2018; Ghorbani, Deo, Yaseen, & Kashani, 2018; Moazenzadeh, Mohammadi, Shamshirband, & Chau, 2018; Tao, Sulaiman, et al., 2018) is yet to be explored. It is worth investigating comprehensively how to employ hybridization with other algorithms to enhance their overall search performance by effectively coupling the strengths of each intelligent algorithm. Hybridizing nature-inspired optimization algorithm can positively enhance the internal hyperparameters mechanism of the EC models, particularly for such a complex hydrological process (Al Sudani, Salih, & Yaseen, 2019; Yaseen, Mohtar, et al., 2019).

(6) Symbolic regression models developed so far using EC are very complex, as they use a large number of mathematical operators. Research could be conducted into developing simple formulae with fewer computational steps for easy estimation of ET_o in the field.

(7) With recent advances in computational power, in terms of parallel processing, software and hardware, a new machine learning paradigm, named the deep learning technique, is now feasible to address millions or even billions of weights among neurons for better learning of behaviors. It is recognized to have started a revolutionary era since it is able to address the problems that have resisted AI for a long time. It was proven to outperform other conventional machine learning algorithms in language translation (Sutskever et al., 2014), natural language understanding (Am, 2013), speech recognition (Hinton et al., 2012), financial markets (Fischer & Krauss, 2018), etc. However, the application of deep learning

to hydrological prediction is still limited. The deep learning technique could be a promising field to explore for the modeling of ET_o.

(8) In recent decades, geographical information system (GIS) and satellite data have been widely used in water resources management and hydrological modeling issues. High-resolution meteorological data from satellites may be used for ET_o modeling. This type of modeling could be especially beneficial for places where climatic stations are rare or required data are missing. Satellite data can be calibrated by comparing the corresponding data acquired from the stations. In this way, better spatial modeling could be possible.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Zaher Mundher Yaseen http://orcid.org/0000-0003-3647-7137
Shamsuddin Shahid http://orcid.org/0000-0001-9621-6452

References

Alazba, A. A., Yassin, M. A., & Mattar, M. A. (2016). Modeling daily evapotranspiration in hyper-arid environment using gene expression programming. *Arabian Journal of Geosciences*, *9*(3), 202. doi:10.1007/s12517-015-2273-x

Ali Ghorbani, M., Kazempour, R., Chau, K.-W., Shamshirband, S., & Taherei Ghazvinei, P. (2018). Forecasting pan evaporation with an integrated artificial neural network quantum-behaved particle swarm optimization model: A case study in Talesh, Northern Iran. *Engineering Applications of Computational Fluid Mechanics*, *12*(1), 724–737. doi:10.1080/19942060.2018.1517052

Allen, R. G., Burt, C. M., Solomon, K., Clemmens, A. J., & O'Halloran, T. (2005). Prediction accuracy for project wide evapotranspiration using crop coefficients and reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, *131*, 24–36. doi:10.1061/(asce)0733-9437(2005)131:1(24)

Al Sudani, Z. A., Salih, S. Q., & Yaseen, Z. M. (2019). Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. *Journal of Hydrology*, *573*(June), 1–12. doi:10.1016/j.jhydrol.2019.03.004

Am, S. (2013). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, *12*, 2493–2537. doi:10.1.1.231.4614

Baghban, A., Jalali, A., Shafiee, M., Ahmadi, M. H., & Chau, K. (2019). Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids. *Engineering Applications of Computational Fluid Mechanics*, *13*(1), 26–39. doi:10.1080/19942060.2018.1542345

Chau, K. (2017). Use of meta-heuristic techniques in rainfall-runoff modelling. *Water*, *9*(3), 1–6.

Chau, K. W. (2007). A split-step particle swarm optimization algorithm in river stage forecasting. *Journal of Hydrology*, *346*(3-4), 131–135. doi:10.1016/j.jhydrol.2007.09.004

Chen, L.-H., Chen, J., & Chen, C. (2018). Effect of environmental measurement uncertainty on prediction of evapotranspiration. *Atmosphere*, *9*(10), 400. doi:10.3390/atmos9100400

Chuntian, C., & Chau, K. W. (2002). Three-person multiobjective conflict decision in reservoir flood control. *European Journal of Operational Research*, *142*(3), 625–631. doi:10.1016/S0377-2217(01)00319-8

Courault, D., Seguin, B., & Olioso, A. (2005). Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrigation and Drainage Systems*, *19*(3), 223–249. doi:10.1007/s10795-005-5186-0

Danandeh Mehr, A., Kahya, E., & Ozger, M. (2014). A gene-wavelet model for long lead time drought forecasting. *Journal of Hydrology*, *517*, 691–699. doi:10.1016/j.jhydrol.2014.06.012

Danandeh Mehr, A., Nourani, V., Kahya, E., Hrnjica, B., Sattar, A. M. A., & Yaseen, Z. M. (2018). Genetic programming in water resources engineering: A state-of-the-art review. *Journal of Hydrology*, *566*(November), 643–667. doi:10.1016/j.jhydrol.2018.09.043

Dineva, A., Mosavi, A., Ardabili, S. F., Vajda, I., Shamshirband, S., Rabczuk, T., & Chau, K. (2019). Review of soft computing models in design and control of rotating electrical machines. *Energies*, *12*(6), 1049.

El-Baroudy, I., Elshorbagy, A., Carey, S. K., Giustolisi, O., & Savic, D. (2010). Comparison of three data-driven techniques in modelling the evapotranspiration process. *Journal of Hydroinformatics*, *12*(4), 365–379. doi:10.2166/hydro.2010.029

Eslamian, S. S., Gohari, S. A., Zareian, M. J., & Firoozfar, A. (2012). Estimating Penman–Monteith reference evapotranspiration using artificial neural networks and genetic algorithm: A case study. *Arabian Journal for Science and Engineering*, *37*(4), 935–944.

Farg, E., Arafat, S. M., Abd El-Wahed, M. S., & El-Gindy, A. M. (2012). Estimation of evapotranspiration ETcand crop coefficient Kcof wheat, in south Nile Delta of Egypt using integrated FAO-56 approach and remote sensing data. *Egyptian Journal of Remote Sensing and Space Science*, *15*(1), 83–89. doi:10.1016/j.ejrs.2012.02.001

Feng, Y., Cui, N., Gong, D., Zhang, Q., & Zhao, L. (2017). Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agricultural Water Management*, *193*, 163–173. doi:10.1016/j.agwat.2017.08.003

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. doi:10.1016/j.ejor.2017.11.054

Fisher, J. B., Malhi, Y., Bonal, D., Da Rocha, H. R., De Araújo, A. C., Gamo, M., … Von Randow, C. (2009). The land-atmosphere water flux in the tropics. *Global Change Biology*, *15*(11), 2694–2714. doi:10.1111/j.1365-2486.2008.01813.x

Ghorbani, M. A., Deo, R. C., Karimi, V., Yaseen, Z. M., & Terzi, O. (2018). Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. *Stochastic Environmental Research and Risk Assessment*, *32*(6), 1683–1697. doi:10.1007/s00477-017-1474-0

Ghorbani, M. A., Deo, R. C., Yaseen, Z. M., & Kashani, M. H. (2018). Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: Case study in North Iran. *Theoretical and Applied Climatology*, *133*(3–4), 1119–1131. doi:10.1007/s00704-017-2244-0

Gocić, M., Motamedi, S., Shamshirband, S., Petković, D., Ch, S., Hashim, R., & Arif, M. (2015). Soft computing approaches for forecasting reference evapotranspiration. *Computers and Electronics in Agriculture*, *113*, 164–173. doi:10.1016/j.compag.2015.02.010

Guven, A., Aytek, A., Yuce, M. I., & Aksoy, H. (2008). Genetic programming-based empirical model for daily reference evapotranspiration estimation. *Clean - Soil, Air, Water*, *36*(10–11), 905–912. doi:10.1002/clen.200800009

Haie, N., Pereira, R. M. S., Machado, G. J., & Shahidian, S. (2019). An introduction to the hyperspace of Penman-Monteith reference evapotranspiration. *International Journal of Hydrology Science and Technology*, *9*(1), 48–64.

Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A. R., Jaitly, N., … Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, *29*(6), 82–97. doi:10.1109/MSP.2012.2205597

Izadifar, Z., & Elshorbagy, A. (2010). Prediction of hourly actual evapotranspiration using neural networks, genetic programming, and statistical models. *Hydrological Processes*, *24*(23), 3413–3425. doi:10.1002/hyp.7771

Jayawardena, A. W., Muttil, N., & Lee, J. H. W. (2006). Comparative analysis of data-driven and GIS-based conceptual rainfall-runoff model. *Journal of Hydrologic Engineering*, *11*(February), 1–11.

Jean, M., Paix, D., Lanhai, L. I., Jiwen, G. E., Jean, N., & Verenyam, A. (2012). Water losses in arid and semi-arid zone: Evaporation, evapotranspiration and seepage. *Journal of Mountain Science*, *9*, 256–261. doi:10.1007/s11629-012-2186-z

Jovic, S., Nedeljkovic, B., Golubovic, Z., & Kostic, N. (2018). Evolutionary algorithm for reference evapotranspiration analysis. *Computers and Electronics in Agriculture*, *150*(April), 1–4. doi:10.1016/j.compag.2018.04.003

Karimi, S., Kisi, O., & Kim, S. (2017). Modelling daily reference evapotranspiration in humid locations of South Korea using local and cross-station data management scenarios. *International Journal of Climatology*, *37*(November 2016), 3238–3246. doi:10.1002/joc.4911

Kiafar, H., Babazadeh, H., Marti, P., Kisi, O., Landeras, G., Karimi, S., & Shiri, J. (2017). Evaluating the generalizability of GEP models for estimating reference evapotranspiration in distant humid and arid locations. *Theoretical and Applied Climatology*, *130*(1–2), 377–389. doi:10.1007/s00704-016-1888-5

Kim, S., & Kim, H. S. (2008). Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *Journal of Hydrology*, *351*(3–4), 299–317. doi:10.1016/j.jhydrol.2007.12.014

Kisi, O. (2010). Fuzzy genetic approach for modeling reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, *136*(March), 175–183.

Kisi, O., & Guven, A. (2010). Evapotranspiration modeling using linear genetic programming technique. *Journal of Irrigation and Drainage Engineering*, *136*(October), 715–723.

Kisi, O., Sanikhani, H., Zounemat-Kermani, M., & Niazi, F. (2015). Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. *Computers and Electronics in Agriculture*, *115*, 66–77. doi:10.1016/j.compag.2015.04.015

Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection*. Cambridge, MA: MIT Press.

Kumar, D., Adamowski, J., Suresh, R., & Ozga-zielinski, B. (2016). Estimating evapotranspiration using an extreme learning machine model: Case study in North Bihar, India. *Journal of Irrigation and Drainage Engineering*, *142*(9), 1–9. doi:10.1061/(ASCE)IR

Kumar, R., Jat, M. K., & Shankar, V. (2012). Methods to estimate irrigated reference crop evapotranspiration - A review. *Water Science and Technology*, *66*(3), 525–535. doi:10.2166/wst.2012.191

Kumar, M., Raghuwanshi, N. S., & Singh, R. (2011). Artificial neural networks approach in evapotranspiration modeling: A review. *Irrigation Science*, *19*(1), 11–25. doi:10.1007/s00271-010-0230-8

Landeras, G., Bekoe, E., Ampofo, J., Logah, F., Diop, M., Cisse, M., & Shiri, J. (2018). New alternatives for reference evapotranspiration estimation in West Africa using limited weather data and ancillary data supply strategies. *Theoretical and Applied Climatology*, *132*, 701–716. doi:10.1007/s00704-017-2120-y

Martí, P., González-Altozano, P., López-Urrea, R., Mancha, L. A., & Shiri, J. (2015). Modeling reference evapotranspiration with calculated targets. Assessment and implications. *Agricultural Water Management*, *149*, 81–90. doi:10.1016/j.agwat.2014.10.028

Mattar, M. A. (2018). Using gene expression programming in monthly reference evapotranspiration modeling: A case study in Egypt. *Agricultural Water Management*, *198*, 28–38. doi:10.1016/j.agwat.2017.12.017

Mattar, M. A., & Alazba, A. A. (2018). GEP and MLR approaches for the prediction of reference evapotranspiration. *Neural Computing and Applications*, 1–13. doi:10.1007/s00521-018-3410-8

Maulik, U., & Bandyopadhyay, S. (2000). Genetic algorithm-based clustering technique. *Pattern Recognition*, *33*(9), 1455–1465. doi:10.1016/S0031-3203(99)00137-5

Mehdizadeh, S. (2018). Estimation of daily reference evapotranspiration (ETo) using artificial intelligence methods: Offering a new approach for lagged ETo data-based modeling. *Journal of Hydrology*, *559*, 794–812. doi:10.1016/j.jhydrol.2018.02.060

Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Computers and Electronics in Agriculture*, *139*, 103–114. doi:10.1016/j.compag.2017.05.002

Mehr, A. D. (2018). An improved gene expression programming model for streamflow forecasting in intermittent streams. *Journal of Hydrology*, *563*(June), 669–678. doi:10.1016/j.jhydrol.2018.06.049

Meng, E., Huang, S., Huang, Q., Fang, W., Wu, L., & Wang, L. (2019). A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. *Journal of Hydrology*, *568*(November 2018), 462–478. doi:10.1016/j.jhydrol.2018.11.015

Mirjalili, S. M., Mirjalili, S. Z., Saremi, S., & Mirjalili, S. (2020). Sine cosine algorithm: Theory, literature review, and application in designing bend photonic crystal waveguides. In *Nature-inspired optimizers* (pp. 201–217). Cham: Springer.

Moazenzadeh, R., Mohammadi, B., Shamshirband, S., & Chau, K. (2018). Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran. *Engineering Applications of Computational Fluid Mechanics*, *12*(1), 584–597. doi:10.1080/19942060.2018.1482476

Mohammad, O., Pour, R., Piri, J., & Kisi, O. (2018). Comparison of SVM, ANFIS and GEP in modeling monthly potential evapotranspiration in arid region (case study: Sistan and baluchestan province –Iran). *Water Science & Technology*, *19*(2), 392–403. doi:10.2166/ws.2018.084

Najafzadeh, M., & Kargar, A. R. (2019). Gene-expression programming, evolutionary polynomial regression, and model tree to evaluate local scour depth at culvert outlets. *Journal of Pipeline Systems Engineering and Practice*, *10*(3). doi:10.1061/(ASCE)PS.1949-1204.0000376

Nourani, V., Komasi, M., & Alami, M. T. (2013). Geomorphology-based genetic programming approach for rainfall–runoff modeling. *Journal of Hydroinformatics*, *15*, 427–445. doi:10.2166/hydro.2012.113

Parasuraman, K., Elshorbagy, A., & Carey, S. K. (2007). Modelling the dynamics of the evapotranspiration process using genetic programming. *Hydrological Sciences Journal*, *52*(3), 563–578. doi:10.1623/hysj.52.3.563

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, *193*(1032), 120–145. doi:10.1098/rspa.1948.0037

Pereira, A. R., Green, S., & Villa Nova, N. A. (2006). Penman-Monteith reference evapotranspiration adapted to estimate irrigated tree transpiration. *Agricultural Water Management*, *83*(1-2), 153–161. doi:10.1016/j.agwat.2005.11.004

Sanikhani, H., Kisi, O., Maroufpoor, E., & Yaseen, Z. M. (2019). Temperature-based modeling of reference evapotranspiration using several artificial intelligence models: Application of different modeling scenarios. *Theoretical and Applied Climatology*, *135*(1-2), 449–462. doi:10.1007/s00704-018-2390-z

Shiri, J. (2017). Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. *Agricultural Water Management*, *188*, 101–114. doi:10.1016/j.agwat.2017.04.009

Shiri, J. (2019). Modeling reference evapotranspiration in island environments: Assessing the practical implications. *Journal of Hydrology*, *570*(January), 265–280. doi:10.1016/j.jhydrol.2018.12.068

Shiri, J., Kişi, Ö, Landeras, G., López, J. J., Nazemi, A. H., & Stuyt, L. C. P. M. (2012). Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). *Journal of Hydrology*, *414–415*, 302–316. doi:10.1016/j.jhydrol.2011.11.004

Shiri, J., Marti, P., Karimi, S., & Landeras, G. (2019). Data splitting strategies for improving data driven models for reference evapotranspiration estimation among similar stations. *Computers and Electronics in Agriculture*, *162*(March), 70–81. doi:10.1016/j.compag.2019.03.030

Shiri, J., Marti, P., Kisi, O., Landeras, G., Fard, A. F., Nazemi, A. H., & Sadraddini, A. A. (2013). Evaluation of different data management scenarios for estimating daily reference evapotranspiration. *Hydrology Research*, *44*(6), 1058–1070. doi:10.2166/nh.2013.154

Shiri, J., Nazemi, A. H., Sadraddini, A. A., Landeras, G., Kisi, O., Fakheri Fard, A., & Marti, P. (2014). Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Computers and Electronics in Agriculture*, *108*, 230–241. doi:10.1016/j.compag.2014.08.007

Shiri, J., Sadraddini, A. A., Nazemi, A. H., Kisi, O., Landeras, G., Fakheri Fard, A., & Marti, P. (2014). Generalizability of gene expression programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. *Journal of Hydrology*, *508*, 1–11. doi:10.1016/j.jhydrol.2013.10.034

Sun, H., Yang, Y., Wu, R., Gui, D., Xue, J., Liu, Y., & Yan, D. (2019). Improving estimation of cropland evapotranspiration by the bayesian model averaging method with surface energy balance models. *Atmosphere*, *10*(4), 188.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in neural information processing systems* (pp. 3104–3112). Snowbird: NeurIPS.

Tao, H., Diop, L., Bodian, A., Djaman, K., Ndiaye, P. M., & Yaseen, Z. M. (2018). Reference evapotranspiration prediction using hybridized fuzzy model with firefly algorithm: Regional case study in Burkina Faso. *Agricultural Water Management*, *208*, 140–151.

Tao, H., Sulaiman, S. O., Yaseen, Z. M., Asadi, H., Meshram, S. G., & Ghorbani, M. A. (2018). with SVM-FFA data-intelligence model? Application of rainfall forecasting over regional scale. *Water Resources Management*, *32*(12), 3935–3959. doi:10.1007/s11269-018-2028-z

Taormina, R., & Chau, K. W. (2015). Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and extreme learning machines. *Journal of Hydrology*, *529*, 1617–1632. doi:10.1016/j.jhydrol.2015.08.022

Traore, S., & Guven, A. (2012). Regional-specific numerical models of evapotranspiration using gene-expression programming interface in sahel. *Water Resources Management*, *26*(15), 4367–4380. doi:10.1007/s11269-012-0149-3

Traore, S., & Guven, A. (2013). New algebraic formulations of evapotranspiration extracted from gene-expression programming in the tropical seasonally dry regions of West Africa. *Irrigation Science*, *31*(1), 1–10. doi:10.1007/s00271-011-0288-y

Traore, S., Luo, Y., & Fipps, G. (2017). Gene-expression programming for short-term forecasting of daily reference evapotranspiration using public weather forecast information. *Water Resources Management*, *31*(15), 4891–4908. doi:10.1007/s11269-017-1784-5

Wu, C. L., & Chau, K. W. (2011). Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology*, *399*(3–4), 394–409. doi:10.1016/j.jhydrol.2011.01.017

Yaseen, Z. M., El-shafie, A., Jaafar, O., Afan, H. A., & Sayl, K. N. (2015). Artificial intelligence based models for stream-flow forecasting: 2000–2015. *Journal of Hydrology*, *530*, 829–844. doi:10.1016/j.jhydrol.2015.10.038

Yaseen, Z., Mohtar, W. H. M. W., Ameen, A. M. S., Ebtehaj, I., Razali, S. F. M., Bonakdari, H., ... Shahid, S. (2019). Implementation of univariate paradigm for streamflow simulation using hybrid data-driven model: Case study in tropical region. *IEEE Access*, *7*, 74471–74481. doi:10.1109/ACCESS.2019.2920916

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., & Chau, K.-W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, *569*, 387–408. doi:10.1016/j.jhydrol.2018.11.069

Yassin, M. A., Alazba, A. A., & Mattar, M. A. (2016a). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agricultural Water Management*, *163*, 110–124. doi:10.1016/j.agwat.2015.09.009

Yassin, M. A., Alazba, A. A., & Mattar, M. A. (2016b). Comparison between gene expression programming and traditional models for estimating evapotranspiration under hyper arid conditions. *Water Resources*, *43*(2), 412–427. doi:10.1134/S0097807816020172