DOI: 10.36253/ijam-3853

Comparison of Bio-Inspired Optimization-Based MLP Models for Monthly Reference Evapotranspiration Estimation in Southeastern Türkiye

Mustafa OZBULDU*

*Department of Biosystems Engineering, Hatay Mustafa Kemal University, Hatay, Türkiye

Corresponding Author: Mustafa OZBULDU, mustafaozbuldu@mku.edu.tr

ORCID : https://orcid.org/0000-0002-5359-8750

ABSTRACT

This study evaluates the performance of hybrid multilayer perceptron (MLP) models with six different biologically inspired meta-heuristic optimization algorithms for the estimation of monthly reference evapotranspiration (ET_0) in Southeast Anatolia, Türkiye. Using NASA POWER data (1984-2024) as main source of climatic data, model input variables for six provinces were determined by Recursive Feature Elimination (RFE) method. Moth-Flame Optimization (MFO), Dolphin Echolocation Optimization (DEO), Firefly Optimization (FOA), Grey Wolf Optimization (GWO), Golden Jackal Optimization (GJO) and Cuckoo Search Optimization (CSO) algorithms were used in the optimization of MLP models. Model performance was evaluated using the coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE), mean absolute error (MAE), RMSE-to-observation standard deviation ratio (RSR), and global performance index (GPI). The results show that DEO-MLP achieved the best overall performance among the investigated hybrid models, with mean values across the six stations of NSE = 0.970, R^2 = 0.972, MAE = 0.39, RSR = 0.17, and GPI = 0.33. The superior performance of DEO-MLP model was demonstrated by low error rates and high accuracy levels compared to the other models. In contrast, the MFO-MLP, CSO-MLP and GJO-MLP models exhibited lower accuracy levels. In general, the results demonstrate that the predictive performance of MLP models varies according to the meta-heuristic optimization algorithm employed, with DEO-MLP showing the best overall performance among the investigated hybrid models. This study is expected to contribute to the development of agricultural water management strategies in Southeast Anatolia and other regions characterized by similar climatic conditions and provides a practical modelling framework for estimating ET_0 when the availability of long-term ground-based meteorological observations is limited.

Keywords: Machine Learning, Evapotranspiration, Global Performance Index, Bio-Inspired Algorithms

1. INTRODUCTION

Evapotranspiration (ET), which refers to the combination of transpiration by plants and evaporation from the soil surface, is one of the key processes of the hydrological cycle. The concept of reference evapotranspiration (ET_0) was first introduced by the Food and Agriculture Organization of the United Nations (FAO) (Allen et al. 1998) for use in agricultural irrigation planning, water budget analyses and climate change studies. The approach has since been accepted as a basic tool for calculating evapotranspiration (Yu et al. 2020; Khairan et al. 2023). In the context of global warming and climate change impacts, accurate estimation of ET_o is an essential element to be considered when developing water management policies (Muhammad et al. 2021). It is imperative to accurately estimate ET_o, as this information is of strategic importance to various sectors, including agriculture and water resources management. Inaccurate ET measurements or ET_o estimates can lead to significant environmental issues, such as inefficient water usage, nutrient loss, and groundwater pollution (Wu et al. 2019; Kisi et al. 2021; Adnan et al. 2021; Mirzania et al. 2023).

Although it is possible to directly measure ET using lysimeters and eddy covariance, the widespread application of these methods is limited due to high costs, technical requirements and operational complexity (Jiang et al. 2020; Liu et al. 2020; Adnan et al. 2021). For this reason, alternative approaches including the FAO-56 Penman-Monteith formula (FAO-56 PM), the Hargreaves equation (Hargreaves and Samani 1985) and the Priestley-Taylor (Priestley and Taylor 1972) equation have been developed to estimate ET_0 indirectly using meteorological data (Landeras et al. 2008; Zhao et al. 2024). Among these approaches, FAO-56 PM is recommended by the FAO as the standard method for estimating ET_0. However, its applicability may be limited in regions where comprehensive and reliable meteorological data are not available, because multiple meteorological variables such as temperature, solar radiation, relative humidity and wind speed are required to evaluate the FAO-56 PM equation (Allen et al. 1998; Valiantzas 2015; Gao et al. 2017).

When complete meteorological datasets are unavailable, empirical and semi-empirical methods requiring fewer meteorological inputs can provide practical alternatives for ET_0 estimation. However, these methods are generally based on predefined functional relationships. Their performance can vary significantly depending on climatic and geographical conditions and may require local validation or calibration in different application areas (Gao et al. 2017; Song et al. 2019; Ndiaye et al. 2020). In contrast to such predefined empirical approaches, data-driven methods are capable of learning the complex and non-linear relationships between meteorological variables and ET_0 from training data, thereby offering a more flexible modelling framework (Tabari and Hosseinzadeh Talaee 2013; Antonopoulos and Antonopoulos 2017).

In recent years, data-driven models such as artificial neural networks (ANN), support vector machines (SVM), adaptive neuro-fuzzy inference systems (ANFIS) and random forests (RF) have been applied to estimate ET_0 under different climatic conditions and combinations of meteorological inputs (Fan et al. 2018; Ferreira et al. 2019; Bakhtiari et al. 2022). Comparative studies have reported that data-driven approaches can achieve performance comparable to, or in some cases better than, empirical ET_0 methods (Landeras et al. 2008; Citakoglu et al. 2014; Ferreira et al. 2019; Reis et al. 2019; Bellido-Jiménez et al. 2022). Nevertheless, their performance may vary depending on the selection of meteorological input variables, the quantity and quality of the training data, model configuration, and the climatic characteristics of the study region (Huo et al. 2012; Antonopoulos and Antonopoulos 2017). Therefore, data-driven approaches should be regarded as flexible modelling alternatives for ET_0 estimation whose suitability depends on data availability, local climatic conditions and application requirements, rather than being assumed to be universally superior to traditional empirical formulations (Huo et al. 2012; Mehdizadeh et al. 2021; Skhiri et al. 2024).

Among neural-network approaches, multilayer perceptron (MLP) models have been widely applied to ET_0 estimation because of their ability to represent complex and nonlinear relationships between meteorological variables and ET_0 and to accommodate different combinations of meteorological inputs (Tabari and Hosseinzadeh Talaee 2013; Antonopoulos and Antonopoulos 2017). However, conventional gradient-based training of neural networks may encounter optimization difficulties such as convergence to local minima and overtraining, potentially affecting model performance (Gao et al. 2021). To address these training limitations, increasing attention has been given to hybrid neural-network models in which metaheuristic optimization algorithms are used to search for suitable network parameters. In ET_0 modelling, neural networks have been coupled with various bio-inspired optimization algorithms, and studies that explicitly compared hybrid models with their standalone counterparts have reported improvements in predictive performance following optimization (Maroufpoor et al. 2020; Tikhamarine et al. 2020; Sayyahi et al. 2021). Because metaheuristic algorithms employ different search mechanisms, their effectiveness in optimizing neural-network parameters may vary across modelling settings. Comparative studies have reported performance differences among alternative ANN–metaheuristic combinations, indicating that the most effective optimization strategy may depend on the characteristics of the modelling problem (Gao et al. 2021; Khairan et al. 2023). Therefore, systematic evaluation of alternative optimization algorithms within a common neural-network framework is important for identifying the most suitable optimization strategy for ET_0 modelling. Accordingly, MLP was selected as a common nonlinear modelling framework that accommodates station-specific predictor subsets identified by RFE and enables the six bio-inspired optimization algorithms to be evaluated under a consistent architecture.

Previous studies in Türkiye have demonstrated the applicability of artificial intelligence and machine-learning approaches for ET_0 estimation (Kisi and Alizamir 2018; Bayram and Çıtakoğlu 2023; Kartal 2024; Hasan and Yüce 2024). Bio-inspired metaheuristic algorithms integrated with MLP have not yet been systematically evaluated for ET_0 estimation in Türkiye, to the authors' knowledge. Therefore, the objective of this study is to develop and comparatively evaluate hybrid prediction models combining MLP with six distinct bio-inspired metaheuristic algorithms for estimating monthly ET_0 at selectedstations in the Southeastern Anatolia Region of Türkiye. Although various optimization algorithms have been examined individually or in different modelling frameworks in previous studies, their systematic comparison within a common MLP architecture remains an important research issue in ET_0 modelling. The hybrid MLP models investigated in this study are included to assess how different optimization strategies affect parameter search and model performance, particularly with respect to the susceptibility of conventional neural-network training to local optima.

A further contribution of the study is the use of multiple statistical performance indicators and a global performance index to provide an integrated assessment of the investigated models. Since it is not feasible to evaluate all available optimization algorithms, six promising bio-inspired algorithms were selected for comparison within the same MLP framework. In addition, the study evaluates a data-driven approach to monthly ET_0 estimation using reanalysis-based meteorological information, particularly in settings where long-term local meteorological observations are incomplete or unavailable. By combining reanalysis data, RFE-based feature selection, and optimized MLP modelling, the study investigates the potential of this framework to support reliable ET_0 estimation and thereby contribute to agricultural water management, irrigation planning, and precision agriculture applications.

2. MATERIALS AND METHODS

2.1. Study Region

The geographic and topographical diversity of Türkiye results in marked regional climatic differences. Southeast Anatolia is characterized by a semi-arid continental climate and is considered particularly vulnerable to climate change (Türkeş and Tatlı 2009). This study covers six provinces (Gaziantep, Adıyaman, Şanlıurfa, Diyarbakır, Siirt and Şırnak), whose geographic and climatic characteristics are summarized in Figure 1 and Table 1. Mean annual temperature ranged from 15.3 to 18.6 °C, while mean annual precipitation ranged from 344.26 to 487.29 mm across the six study sites. Precipitation is concentrated mainly from late autumn to spring, while summers are characterized by high temperatures and very low precipitation (Figure 2). Agriculture is a major land-use and economic activity in Southeastern Anatolia. In this region, approximately 2.99 million ha of land was used for agriculture in 2024, accounting for 12.4% of Türkiye's total agricultural area (GAP Regional Development Administration, 2024). Cereals and pulses are widely cultivated, with wheat, barley and red lentil being among the main crops. Irrigation is especially important under the hot and dry summer conditions of Southeast Anatolia. Cotton and maize are among the important irrigated crops in the region. These characteristics highlight the practical importance of reliable ET_0 estimation for irrigation planning and agricultural water management.

Table 1. Geographic coordinates and long-term climatic characteristics of the six stations considered for this study

Station	Lat (°N)	Lon (°E)	Mean annual temperature (°C)	Mean annual precipitation (mm)
Adıyaman	37.7558	38.2726	17.4	401.80
Diyarbakır	37.8861	40.1997	16.0	466.53
Gaziantep	37.0684	37.3858	15.3	434.26
Siirt	37.9228	41.9318	16.3	487.29
Şanlıurfa	37.1554	38.7862	18.6	344.26
Şırnak	37.5228	42.4536	15.4	418.36

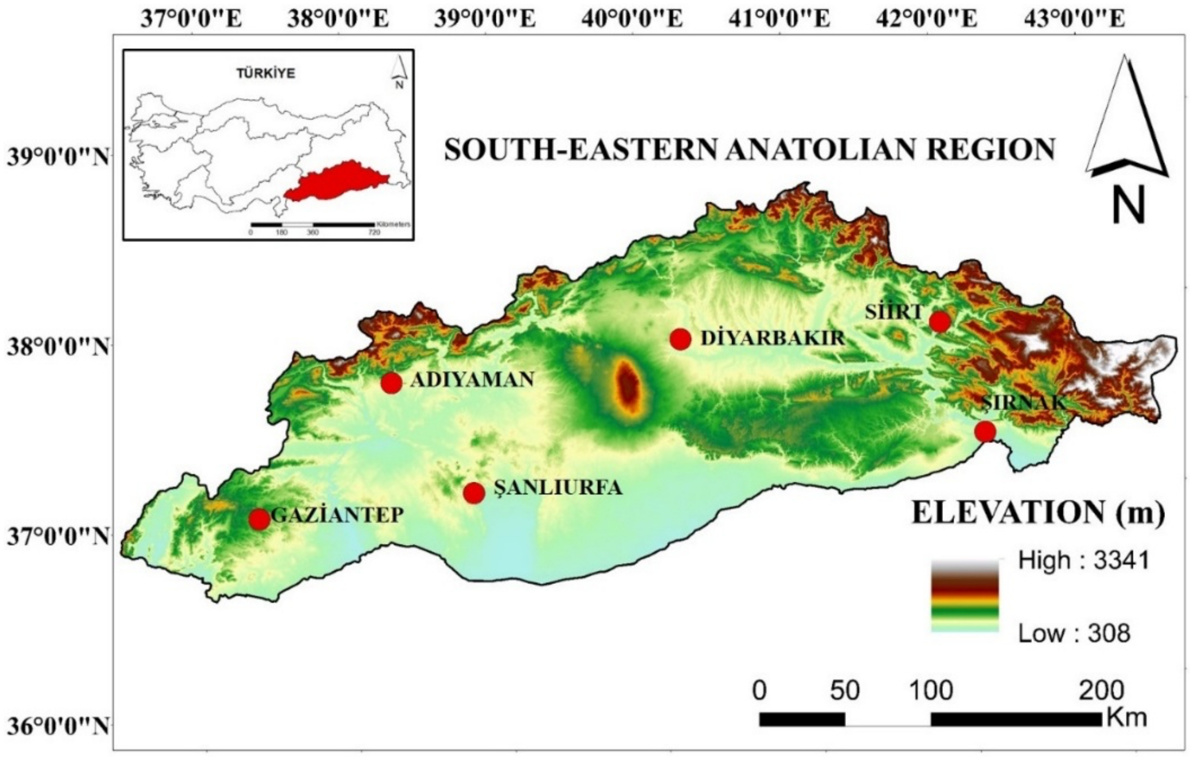

Figure 1. Study area and location of the stations

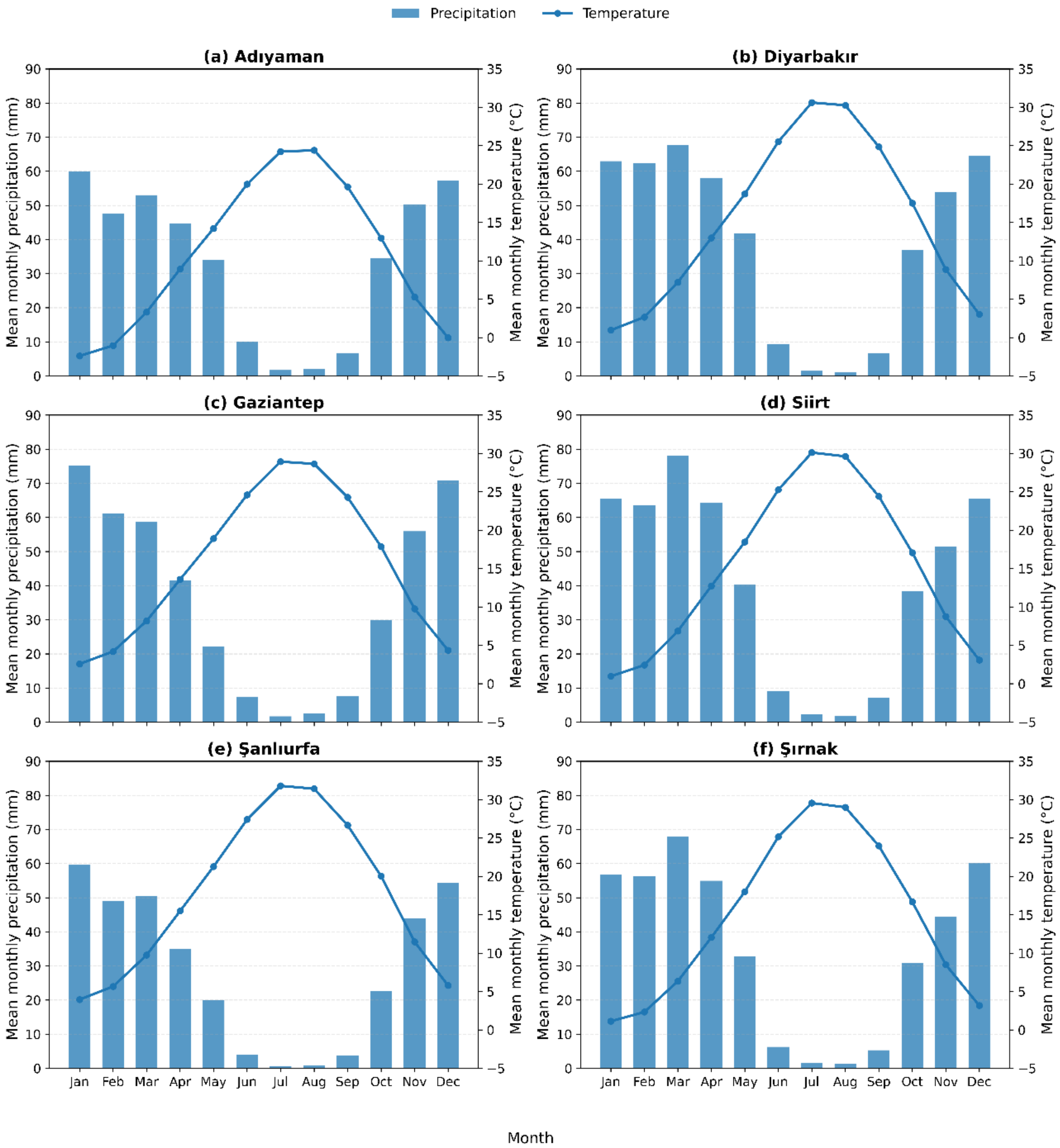

Figure 2. Long-term monthly temperature and precipitation climatology at the six study sites based on NASA POWER data for the period 1984-2024

2.2. Weather Data

Limited availability of long-term and continuous meteorological observations can constrain applications such as irrigation planning, water resources management and climate analyses. The NASA Prediction of Worldwide Energy Resources (POWER) dataset provides a globally accessible source of long-term meteorological information that can help address limitations in ground-based observations (Rodrigues and Braga 2021). NASA POWER provides reanalysis-based meteorological data at a spatial resolution of 0.5° and includes variables derived from satellite observations, ground-based measurements and atmospheric modelling (Tayyeh and Mohammed 2023). The availability of these data since 1981 enables their use in long-term climatic analyses. In agricultural applications, NASA POWER has been extensively utilized for the estimation of reference evapotranspiration (ETo), crop yield analysis and irrigation scheduling (Delgado-Ramírez et al. 2023; Khairan et al. 2023).

Previous studies have reported satisfactory performance of NASA POWER meteorological data across tropical, Mediterranean and continental climatic conditions (Monteiro et al. 2018; Rodrigues and Braga 2021). Rodrigues and Braga (2021) found that NASA POWER data were largely in agreement with ground-based observations under Mediterranean climatic conditions. Furthermore, a study conducted in the Euphrates River Basin reported good agreement between NASA POWER and ground-based temperature and precipitation data (Tayyeh and Mohammed 2023).

In this study, daily NASA POWER meteorological data for the period 1984–2024 were downloaded for the six study sites from the NASA POWER Data Access Viewer (https://power.larc.nasa.gov/data-access-viewer). Information on NASA POWER variables used in the study is presented in Table 2. Monthly climatologies were derived for the 1984-2024 period from these daily data. The same daily dataset was also used to characterize the long-term climate of the six study sites. Long-term mean annual temperature was calculated from annual T2M values, whereas long-term mean annual precipitation was calculated from annual totals evaluated from PRECTOTCORR (see Table 2 for abbreviations).

Table 2. Available NASA POWER variables and abbreviations used in the study

Variables	Unit	Abbreviation
Temperature at 2 Meters Maximum	°C	T2M_MAX
Temperature at 2 Meters Minimum	°C	T2M_MIN
Temperature at 2 Meters	°C	T2M
Dew/Frost Point at 2 Meters	°C	T2MDEW
Wet Bulb Temperature at 2 Meters	°C	T2MWET
Earth Skin Temperature	°C	TS
All Sky Surface Longwave Downward Irradiance	MJ m^{-2}	ALLSKY_SFC_LW_DWN
All Sky Surface Shortwave Downward Irradiance	MJ m^{-2}	ALLSKY_SFC_SW_DWN
Specific Humidity at 2 Meters	g kg^{-1}	QV2M
Relative Humidity at 2 Meters	%	RH2M
Surface Pressure	kPa	PS
Wind Speed at 2 Meters	m s^{-1}	WS2M
Wind Speed at 10 Meters	m s^{-1}	WS10M
Precipitation	mm day^{-1}	PRECTOTCORR
Surface Soil Wetness	-	GWETTOP
Root Zone Soil Wetness	-	GWETROOT
Profile Soil Moisture	-	GWETPROF

2.3. FAO-56 Penman-Monteith equation

The FAO-56 Penman-Monteith equation is a physically based model that was proposed as a standard by the Food and Agriculture Organization of the United Nations (FAO) and is widely utilized for the calculation of ET_o (Allen et al. 1998). The method is founded on the principles of energy balance and aerodynamic transfer, with the objective of estimating water loss from an idealized, well-watered grass surface. The FAO-56 Penman–Monteith equation used to calculate reference evapotranspiration is given in Equation 1.

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\left(\frac{900}{T+273}\right)u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad \textbf{(1)}$$

where ET_0 is the reference evapotranspiration ($mm\ day^{-1}$); R_n is the net radiation at the crop surface (MJ m^{-2} day^{-1}); G is the soil heat flux density (MJ m^{-2} day^{-1}); T is the mean air temperature at 2 m height (°C); u_2 is the wind speed at 2 m height ($m\ s^{-1}$); e_s is the saturation vapor pressure (kPa); e_a is the actual vapor pressure (kPa); $\Delta$is the slope of the saturation vapor pressure curve (kPa $°C^{-1}$); and γ is the psychrometric constant (kPa $°C^{-1}$).

2.4. Recursive Feature Elimination

Recursive Feature Elimination (RFE) is an iterative feature-selection method used to identify the most informative predictors and eliminate variables that make a limited contribution to model performance (Darst et al. 2018). The procedure recursively removes less informative variables and evaluates the remaining predictor subsets until the combination providing the best predictive performance is identified (Li et al. 2017; Huang et al. 2024). In the present study, the dataset for each station was first divided into training (70%) and testing (30%) subsets. RFE was then applied only to the training data using a Random Forest regressor with five-fold cross-validation. The initial predictor pool was intentionally constructed to include a broad range of meteorological and land-surface variables available from NASA POWER (Table 2). Therefore, related atmospheric variables and auxiliary variables such as soil moisture variables were not excluded a priori. RFE was instead used to identify the most informative station-specific predictor subset. The testing data were not involved in the feature-selection procedure. Once the predictor subset had been identified for a given station, it was kept unchanged across all six hybrid MLP models. Thus, the optimization algorithms were compared using identical input information within each station.

2.5. Multi-Layer Perceptron (MLP)

Artificial neural networks (ANNs) are computational models inspired by the learning mechanisms of the human brain and are widely used for pattern recognition, forecasting, and modelling complex nonlinear relationships, particularly in meteorological and hydrological applications (Kumar et al. 2011; Bellido-Jiménez et al. 2022; Yin et al. 2023; Elbeltagi et al. 2024; Ozbuldu and Irvem 2024). Among ANN architectures, the Multi-Layer Perceptron (MLP) is widely employed for prediction and modelling because of its ability to capture nonlinear relationships (Kumar et al. 2011; Gao et al. 2021). A typical MLP consists of input, hidden, and output layers, in which neurons process weighted inputs through nonlinear activation functions and propagate the resulting information through the network (Shiri et al. 2012; Bellido-Jiménez et al. 2022). Model learning is achieved through backpropagation, whereby network weights are iteratively adjusted to minimize the error between predicted and observed values (Kumar et al. 2011). In this study, an MLP architecture was employed for ET_0 modelling because of its ability to represent nonlinear relationships (Zhu et al. 2020). The number of hidden-layer neurons was set to one more than the number of input variables (Kumar et al. 2011), and ReLU (Rectified Linear Unit) was used as the activation function. The MLP models were implemented in Python using the Scikit-learn library (Pedregosa et al. 2011)

2.6. Optimization Algorithms

Six bio-inspired metaheuristic optimization algorithms were employed to optimize the MLP model parameters: Moth-Flame Optimization (MFO; Mirjalili 2015), Dolphin Echolocation Optimization (DEO; Kaveh and Farhoudi 2013), Firefly Optimization (FOA; Yang et al. 2010), Grey Wolf Optimization (GWO; Mirjalili et al. 2014), Golden Jackal Optimization (GJO; Chopra and Ansari 2022), and Cuckoo Search Optimization (CSO; Yang and Deb 2009). These algorithms employ different exploration and exploitation mechanisms to search the solution space and identify suitable model parameters. In the present study, they were evaluated within the same MLP framework under consistent optimization conditions to enable a comparative assessment of their performance. Detailed descriptions, mathematical formulations, and additional relevant references for the six algorithms are provided in Supplementary Text S1.

2.7. Hybrid Optimization Model

The station-specific predictor subsets identified by RFE were used as inputs to the hybrid MLP models. The input data were scaled to the range [0,1] using Min-Max normalization. Six bio-inspired optimization algorithms were then used to optimize the weights and biases of the MLP models. To ensure a consistent comparison among the algorithms, the population size and maximum number of iterations were fixed for all optimization methods, and Mean Squared Error (MSE) was used as the objective function. The optimized models were subsequently evaluated on the test data using the selected statistical performance criteria. Model development and analysis were performed in Python.

2.8. Evaluation of Model Performances

In this study, the performance of the proposed MLP models was evaluated using different statistical metrics including the coefficient of determination (R^2), the mean absolute error (MAE), the root-mean square error (RMSE), the RMSE-to-observation standard deviation ratio (RSR), the Nash-Sutcliffe efficiency (NSE), the mean bias error (MBE) and the scatter index (SI). The mathematical formulations of the performance evaluation metrics are given in Equations 2-8.

$$R^2 = \frac{\left[\sum_{i=1}^{n}(O_i-\bar{O})(M_i-\bar{M})\right]^2}{\left[\sum_{i=1}^{n}(O_i-\bar{O})^2\right]\left[\sum_{i=1}^{n}(M_i-\bar{M})^2\right]} \quad \textbf{(2)}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | O_i - M_i | \quad \textbf{(3)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - M_i)^2} \quad \textbf{(4)}$$

$$RSR = \frac{RMSE}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i-\bar{O})^2}} \quad \textbf{(5)}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - M_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \quad \textbf{(6)}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(M_i - O_i) \quad \textbf{(7)}$$

$$SI = \frac{RMSE}{\bar{O}} \quad \textbf{(8)}$$

where O_i represents the i-th observation data, $\bar{O}$ represents the average of the observation data, M_i represents the i-th data of the model, $\bar{M}$ represents the average of the model estimations and n represents the number of data in the dataset. In general, low RMSE and MAE values, MBE values closer to zero, and high R^2 values indicate better predictive performance. The performance evaluation classification for NSE, RSR and SI criteria is given in the Table 3 (Moriasi et al. 2007, Li et al. 2013, Mirzania et al. 2023).

Table 3. Criteria for the evaluation of model performance for the three metrics NSE, RSR and SI

Performance	NSE	RSR	SI
Very Good	0.75 - 1.00	0.00 - 0.50	0.00 - 0.10
Good	0.65 - 0.75	0.50 - 0.60	0.10 - 0.20
Satisfactory	0.50 - 0.65	0.60 - 0.70	0.20 - 0.30
Unsatisfactory	< 0.50	> 0.70	> 0.30

Since individual statistical metrics describe different aspects of predictive performance, the global performance index (GPI) was used to integrate the selected metrics into a single criterion and provide an overall ranking of the models (Despotovic et al. 2015; Zhu et al. 2020). In this study, the GPI was calculated from six performance indicators (MAE, RSR, MBE, SI, R^2 and NSE) using Equation 9.

$$GPI_i = \sum_{j=1}^{m} \alpha_j \left(s_j - r_{ij}\right) \qquad \textbf{(9)}$$

where GPI_i denotes the global performance index of model i; αj is a constant with a value of 1 for MAE, RSR, MBE and SI and -1 for R^2 and NSE; s_j is the median of the normalized values of indicator j and r_{ij} is the normalized value of indicator j for model i. Because the GPI is a relative ranking measure, it has no universal performance thresholds; a higher GPI indicates better overall performance within the set of models compared.

In addition to the statistical performance metrics, Taylor diagrams (Taylor 2001) were used to provide a graphical comparison of model performance. Taylor diagrams simultaneously evaluate model performance in terms of correlation, standard deviation, and centered RMSE. They are especially useful in the field of climatic and hydrological prediction studies (Taylor 2001; Zhou et al. 2021). In this study, Taylor diagrams were constructed for the six stations to compare the performance of the investigated hybrid MLP models.

3. RESULTS

The selected predictor subsets contained between seven and eleven variables across the six stations (Table 4). Eight predictors were retained for Adıyaman and Diyarbakır, eleven for Gaziantep and Şanlıurfa, and seven for Siirt and Şırnak. Five predictors, TS, T2M_MAX, RH2M, ALLSKY_SFC_SW_DWN and PS, were retained at all six stations. In addition to the five predictors retained at all stations, WS2M and WS10M were each selected at four stations, whereas T2M, PRECTOTCORR, GWETTOP and T2MWET were each retained at three stations. GWETROOT and T2MDEW showed more station-specific selection patterns, being retained only for Gaziantep and Şanlıurfa, respectively. In contrast, T2M_MIN, ALLSKY_SFC_LW_DWN, QV2M, and GWETPROF were not retained at any of the six stations. Overall, the results indicate the presence of a common core predictor set together with station-specific differences in the final variable combinations.

Table 4. Predictor variables retained by RFE for each study station

Station	Number of retained predictors	Predictors retained by RFE
Adıyaman	8	TS, T2M_MAX, GWETTOP, RH2M, ALLSKY_SFC_SW_DWN, PS, WS2M, WS10M
Diyarbakır	8	T2M, PRECTOTCORR, ALLSKY_SFC_SW_DWN, PS, RH2M, TS, T2M_MAX, T2MWET
Gaziantep	11	T2M, GWETTOP, PRECTOTCORR, ALLSKY_SFC_SW_DWN, WS10M, WS2M, GWETROOT, TS, RH2M, T2M_MAX, PS
Siirt	7	TS, T2M_MAX, ALLSKY_SFC_SW_DWN, T2MWET, RH2M, WS10M, PS
Şanlıurfa	11	T2M, GWETTOP, PRECTOTCORR, ALLSKY_SFC_SW_DWN, WS10M, WS2M, PS, TS, RH2M, T2MDEW, T2M_MAX
Şırnak	7	TS, T2M_MAX, T2MWET, RH2M, ALLSKY_SFC_SW_DWN, PS, WS2M

The statistical evaluation showed that the six hybrid MLP models generally achieved strong predictive performance. Detailed station-specific performance metrics for all six hybrid models are provided in Table S1. In particular, DEO-MLP showed the strongest overall performance across the six study stations. Based on the arithmetic mean of the station-specific performance metrics reported in Table S1, DEO-MLP achieved mean values of NSE = 0.970, R^2 = 0.972, MAE = 0.39, RSR = 0.17, and SI = 0.109. The mean MBE was close to zero (0.004), although the station-specific MBE values ranged from −0.098 to 0.061, indicating small but directionally varying biases among stations. These results, together with the highest mean GPI value (0.33), indicate that DEO-MLP provided the best overall performance among the investigated hybrid models. In contrast, MFO-MLP, CSO-MLP, and GJO-MLP showed comparatively weaker overall performance than the other models. CSO-MLP yielded NSE values of 0.89–0.95 and R^2 values of 0.92–0.96 across the six stations. Although the NSE values fell within the "Very Good" category, the error-based metrics and GPI indicated larger prediction errors than those of the best-performing models at several stations. GJO-MLP also showed comparatively larger errors at some stations, with MAE values ranging from 0.55 to 0.84 and MBE values ranging from −0.23 to 0.29.

The integrated GPI rankings of the six hybrid MLP models are presented in Table 5. According to the GPI-based ranking results, DEO-MLP achieved the highest overall GPI value (0.33) and was therefore ranked as the best-performing model, followed by GWO-MLP (0.18) and FOA-MLP (0.07). In contrast, CSO-MLP (−0.22), GJO-MLP (−0.17) and MFO-MLP (−0.29) obtained lower overall GPI values, indicating comparatively weaker performance within the investigated model set. Since GPI is a relative ranking measure, negative values should be interpreted in relation to the performance of the other models rather than as an absolute indication of unacceptable model performance.

Table 5. GPI results for the six bio-inspired meta-heuristic optimization algorithms

Stations/Model	GWO-MLP	CSO-MLP	GJO-MLP	FOA-MLP	DEO-MLP	MFO-MLP
Adıyaman	0.17	0.09	-0.34	-0.02	0.48	-0.36
Diyarbakır	0.13	-0.18	-0.01	0.06	0.25	-0.31
Gaziantep	0.08	-0.54	0.20	0.16	0.45	-0.27
Siirt	0.32	-0.49	-0.26	0.41	0.50	-0.29
Şanlıurfa	0.18	0.15	-0.74	-0.29	0.11	-0.08
Şırnak	0.19	-0.37	0.12	0.12	0.20	-0.42
Overall	0.18	-0.22	-0.17	0.07	0.33	-0.29

Figure 3 presents scatter plots comparing observed and predicted ET_0 values for all models across the selected stations. The results obtained from the graphs reveal that the prediction accuracy of the models varies spatially, with each model exhibiting a unique error profile for different stations. While the GWO-MLP and DEO-MLP models show the highest prediction accuracy overall, the FOA-MLP and CSO-MLP models exhibit systematic prediction errors in specific regions. Specifically, at the Şanlıurfa and Adıyaman stations, FOA-MLP and CSO-MLP models frequently over-predict at low ET values, while GJO-MLP and MFO-MLP models under-predict at high ET values. These station and magnitude-dependent deviations indicate that the error patterns of the hybrid models were not uniform across the range of ET_0 values, with some models showing greater deviations particularly at the lower or upper ends of the observed ET_0 range.

Figure 4 presents the Taylor diagrams for the six study stations, revealing differences in the agreement between the hybrid MLP predictions and the reference ET_0 values. Overall, DEO-MLP was generally positioned close to the reference point, indicating a combination of high correlation, a standard deviation comparable to that of the reference data, and relatively low centered RMSE. GWO-MLP also showed close agreement with the reference data at several stations. In contrast, some of the remaining models were positioned farther from the reference point and exhibited greater station-to-station variability in their correlation, standard deviation, and centered RMSE. These results indicate that the relative performance of the optimization algorithms was not uniform across the study locations. Nevertheless, the overall pattern observed in the Taylor diagrams was consistent with the statistical performance indicators, which identified DEO-MLP as the strongest overall model among the investigated hybrid MLP configurations.

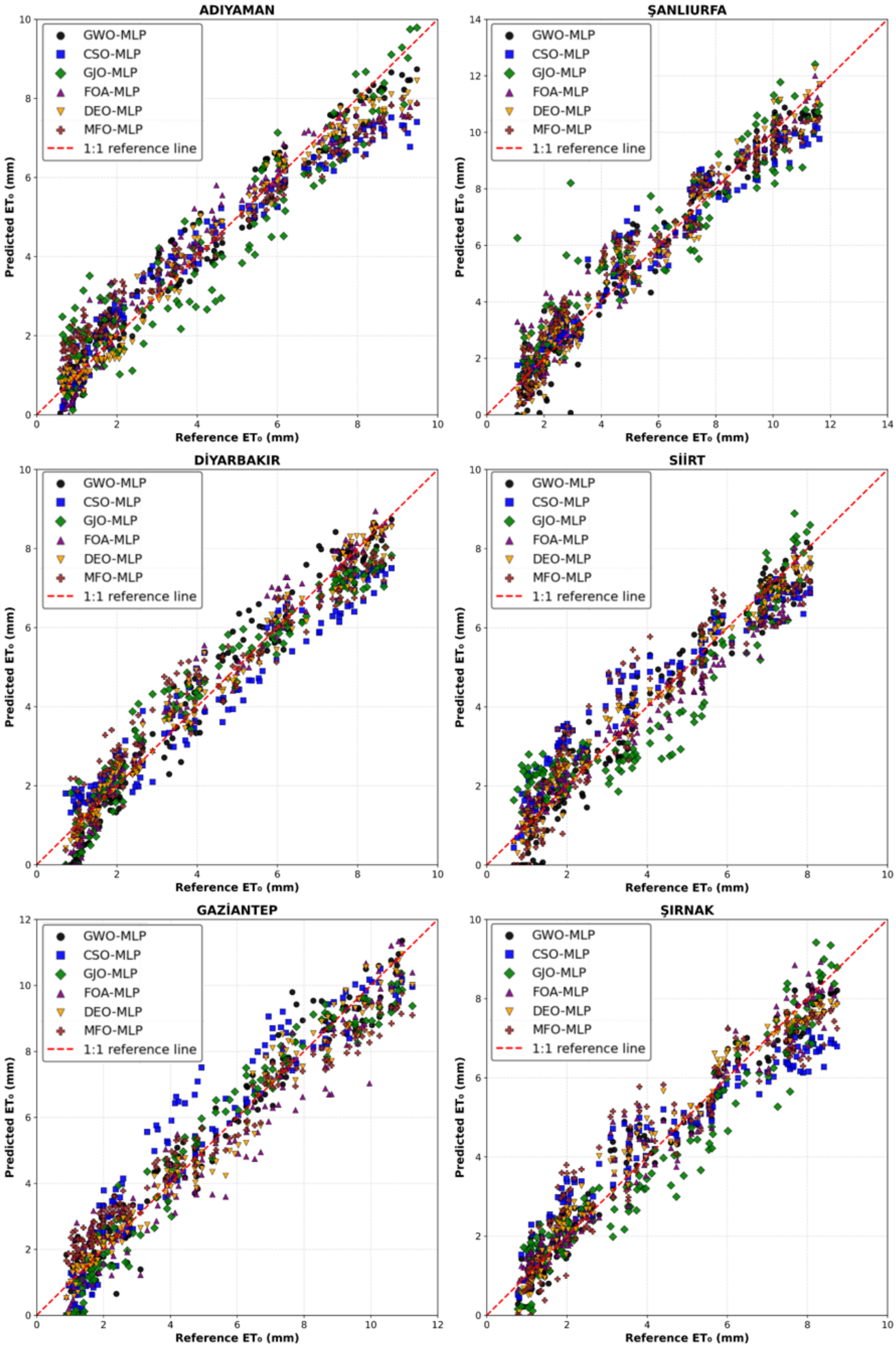

Figure 3. Scatter plots of predicted (y-axis) versus reference ETo (x-axis) calculated with the FAO56 PM equation for the six study sites

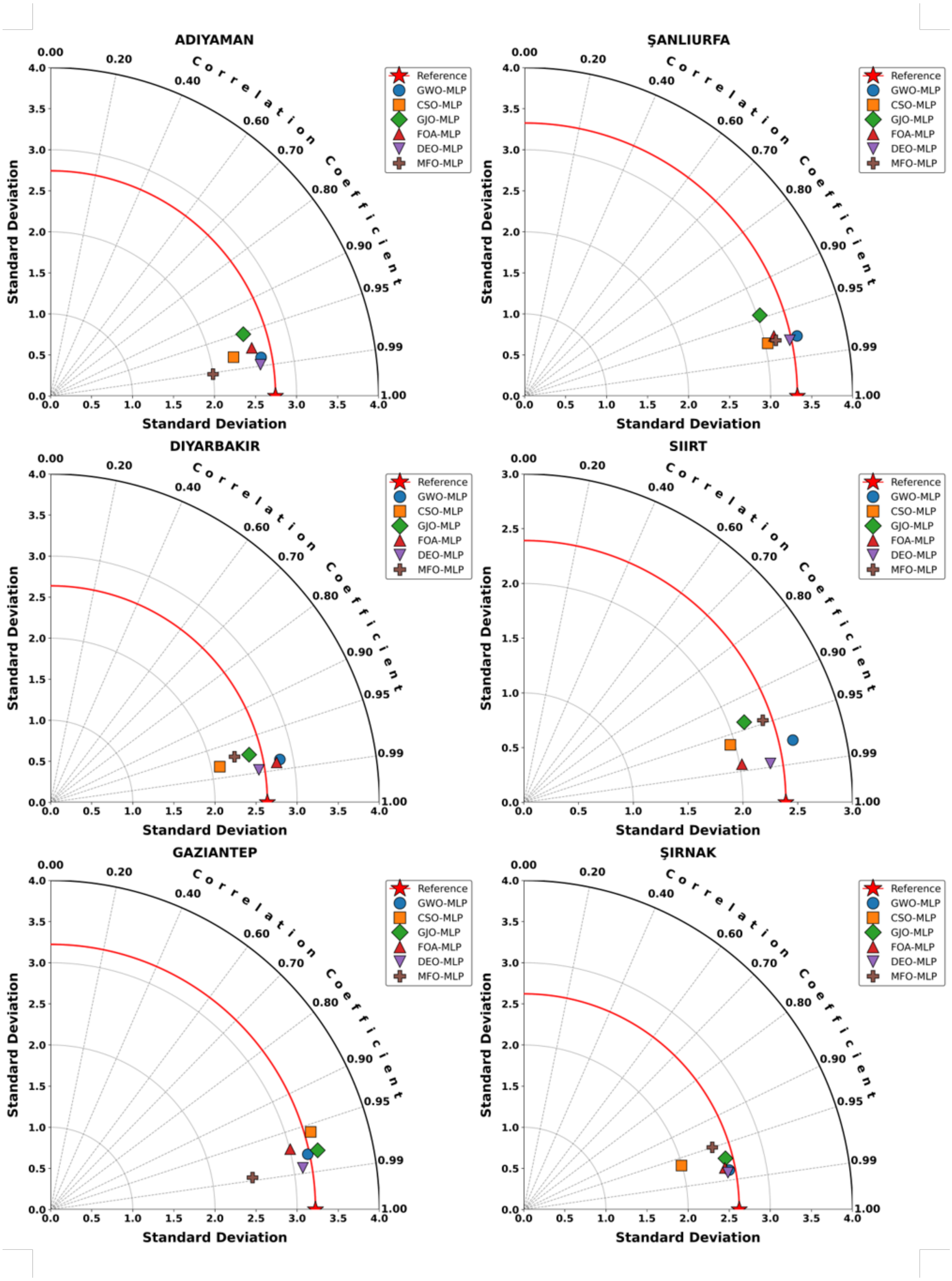

Figure 4. Taylor diagrams illustrating the performance of the examined optimization algorithms at the six study stations

4. DISCUSSION

The results demonstrated differences in predictive performance among the investigated hybrid MLP configurations. DEO-MLP showed the strongest overall performance across the study stations, with NSE and R^2 values of approximately 0.97-0.98 together with comparatively low MAE and MBE values. The strong predictive performance observed in the present study is consistent with previous studies reporting high predictive accuracy for metaheuristic–artificial intelligence hybrid models (Mehdizadeh et al. 2020; Safari et al. 2020; Roy et al. 2021; Adnan et al. 2021; Mirzania et al. 2023). The comparatively strong performance of DEO-MLP and GWO-MLP may partly be associated with their mechanisms for balancing global exploration and local exploitation. DEO progressively shifts from global to local search (Kaveh and Farhoudi 2013), whereas GWO uses a hierarchical search mechanism that supports a balance between exploration and exploitation (Kiani et al. 2022). This search behaviour may also help reduce premature convergence by progressively refining promising solutions (Zhang and Zhou 2017). The comparatively weaker performance of some GJO-MLP, MFO-MLP, and CSO-MLP configurations may be related to differences in their exploration and exploitation behavior under the present modelling conditions. Previous studies have discussed algorithm-specific search characteristics, including escape-energy dynamics in GJO, convergence behavior in MFO, and Lévy-flight-based exploration in CSO (Li et al. 2016; He et al. 2018; Jia et al. 2019; Mohapatra and Mohapatra 2023; Ma and Yu 2022).

Although the hybrid MLP models achieved high predictive performance, alternative methods for estimating ET_o that are less input demanding can also provide practical alternatives to the standard approach under data-limited conditions. However, their performance may vary with climatic and regional conditions and may require local validation or calibration (Gao et al. 2017; Song et al. 2019; Ndiaye et al. 2020). Accordingly, MLP was considered here as a flexible nonlinear framework incorporating station-specific predictors selected by RFE rather than as a universally superior alternative. This common MLP architecture also enabled a consistent comparison of the six bio-inspired optimization algorithms. Since conventional empirical methods were not included as benchmarks, the results should be interpreted as a comparison among the investigated hybrid MLP configurations rather than as evidence that MLP outperforms traditional ET_o estimation methods.

The RFE results revealed both common and station-specific predictor patterns across the study region. TS, T2M_MAX, RH2M, ALLSKY_SFC_SW_DWN and PS were retained at all six stations. This is consistent with previous studies showing that temperature, humidity and radiation variables are important predictors of ET_o (Ravindran et al. 2021; Ahmadi et al. 2021; Bashir et al. 2024). Differences in the final predictor subsets among stations are also consistent with previous findings that predictor importance can vary with climatic and geographical conditions (Fan et al. 2018; Ahmadi et al. 2021). Because the initial predictor pool intentionally included a broad range of meteorological and auxiliary land-surface variables, RFE was used to identify informative predictor subsets for each station. Accordingly, the selected variables should be interpreted in terms of predictive relevance rather than as independent physical controls on ET_o.

Direct numerical comparison of the present results with previous ET_o modelling studies should be made cautiously because of differences in geographical settings, input datasets, temporal scales, and modelling frameworks (Zhu et al. 2020; Gao et al. 2021; Elbeltagi et al. 2024). Nevertheless, the performance metrics obtained in the present study were generally within the ranges reported in these recent ET_o modelling studies. Previous studies have reported high predictive performance for hybrid models based on artificial bee colony, particle swarm optimization, whale optimization, crow search, bat algorithms, water wave optimization, and other metaheuristic methods (Zhu et al. 2020; Gao et al. 2021; Sayyahi et al. 2021; Khairan et al. 2023; Elbeltagi et al. 2024). The strong predictive performance of the hybrid models investigated here is therefore consistent with the broader literature on metaheuristic–ANN modelling of ET_0. However, because a standalone MLP was not evaluated in the present study, the performance gain attributable specifically to metaheuristic optimization cannot be quantified from the present results.

A fixed population size and number of iterations were used for all optimization algorithms to ensure a consistent comparison. However, different meta-heuristic algorithms may respond differently to parameter settings, and no algorithm-specific sensitivity analysis was performed in the present study. Therefore, future studies could investigate adaptive parameter tuning and algorithm-specific optimization settings to determine whether further performance improvements are possible. Accurate monthly ET_0 estimation may contribute to more informed irrigation planning and agricultural water management, particularly in regions where continuous long-term ground-based meteorological observations are limited. However, the broader applicability of the proposed framework should be examined across different climatic and agro-environmental conditions and further evaluated through direct comparisons with alternative ET_0 estimation approaches.

5. CONCLUSIONS

This study developed and comparatively evaluated hybrid models combining MLP with six bio-inspired metaheuristic algorithms for monthly ET_0 estimation in Southeastern Anatolia. DEO-MLP provided the strongest overall predictive performance among the investigated models, with mean NSE and R^2 values of 0.970 and 0.972, respectively, together with low MAE and MBE close to zero. GWO-MLP and FOA-MLP also showed relatively stable performance across the study stations, whereas MFO-MLP, CSO-MLP, and GJO-MLP exhibited comparatively higher errors at some stations. RFE identified a common core of predictors across the study stations while also revealing station-specific differences in the selected input subsets. These findings show that predictive performance varied with the optimization algorithm used within the common MLP framework.

Since conventional empirical ET_0 methods were not included as benchmark models, the present findings should be interpreted as a comparative assessment of the investigated hybrid MLP configurations. The proposed framework demonstrated strong performance under the input conditions evaluated in this study, while its robustness under incomplete-data scenarios should be further investigated. Overall, the results demonstrate the potential of combining NASA POWER meteorological data, RFE-based feature selection, and optimized MLP models for monthly ET_0 estimation. Future studies should extend the evaluation to different climatic regions, directly compare the hybrid models with reduced-input empirical methods, and investigate their performance under systematically controlled incomplete-data scenarios.

Data Availability

Data will be made available on request.

Declaration of Competing Interest

The author declares having no known competing financial interests or personal relationships that could influence the work reported in this article.

Funding Declaration

This research received no funding.

References

Adnan R.M., Mostafa R.R., Islam A.R.M.T., Kisi O., Kuriqi A., Heddam S., 2021. Estimating reference evapotranspiration using hybrid adaptive fuzzy inferencing coupled with heuristic algorithms. Comput Electron Agric 191: 106541. https://doi.org/10.1016/j.compag.2021.106541.

Ahmadi F., Mehdizadeh S., Mohammadi B., Pham Q.B., Doan T.N.C., Vo N.D., 2021. Application of an artificial intelligence technique enhanced with intelligent water drops for monthly reference evapotranspiration estimation. Agric Water Manag 244: 106622. https://doi.org/10.1016/j.agwat.2020.106622.

Allen R.G., Pereira L.S., Raes D., Smith M., 1998. Crop evapotranspiration—Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56. Rome: Food and Agriculture Organization of the United Nations.

Antonopoulos V.Z., Antonopoulos A.V., 2017. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Comput Electron Agric 132:86–96.https://doi.org/10.1016/j.compag.2016.11.011.

Bakhtiari B., Mohebbi-Dehaghani A., Qaderi K., 2022. Comparative analysis of data-driven methods for daily reference evapotranspiration estimation of Southern Caspian Sea. Meteorol Appl 29:e2091. https://doi.org/10.1002/met.2091.

Bashir R.N., Mzoughi O., Shahid M.A., Alturki N., Saidani O., 2024. Principal Component Analysis (PCA) and feature importance-based dimension reduction for reference evapotranspiration (ET_o) predictions of Taif, Saudi Arabia. Computers and Electronics in Agriculture 222:109036. https://doi.org/10.1016/j.compag.2024.109036.

Bayram S., Çıtakoğlu H., 2023. Modeling monthly reference evapotranspiration process in Turkey: application of machine learning methods. Environ Monit Assess 195(67). https://doi.org/10.1007/s10661-022-10662-z.

Bellido-Jiménez J.A., Estévez J., García-Marín A.P., 2022. A regional machine learning method to outperform temperature-based reference evapotranspiration estimations in Southern Spain. Agric Water Manag 274:107955.https://doi.org/10.1016/j.agwat.2022.107955.

Chopra N., Ansari M.M., 2022. Golden jackal optimization: A novel nature-inspired optimizer for engineering applications. Expert Syst Appl 198:116924.https://doi.org/10.1016/j.eswa.2022.116924.

Citakoglu H., Cobaner M., Haktanir T., Kisi O., 2014. Estimation of monthly mean reference evapotranspiration in Turkey. Water Resour Manag 28(1):99–111.https://doi.org/10.1007/s11269-013-0474-1.

Darst B.F., Malecki K.C., Engelman C.D., 2018. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC Genet 19:1–6. https://doi.org/10.1186/s12863-018-0633-8.

Delgado-Ramírez G., Bolaños-González M.A., Quevedo-Nolasco A., López-Pérez A., Estrada-Ávalos J., 2023. Estimation of reference evapotranspiration in a semi-arid region of Mexico. Sensors 23(15):7007. https://doi.org/10.3390/s23157007.

Despotovic M., Nedic V., Despotovic D., Cvetanovic S., 2015. Review and statistical analysis of different global solar radiation sunshine models. Renew Sustain Energy Rev 52:1869–1880. https://doi.org/10.1016/j.rser.2015.08.035.

Elbeltagi A., Katipoğlu O.M., Kartal V., Danandeh Mehr A., Berhail S., Elsadek E.A., 2024. Advanced reference crop evapotranspiration prediction: A novel framework combining neural nets, bee optimization algorithm, and mode decomposition. Appl Water Sci 14:256. https://doi.org/10.1007/s13201-024-02308-x.

Fan J., Yue W., Wu L., Zhang F., Cai H., Wang X., Lu X., Xiang Y., 2018. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. Agricultural and Forest Meteorology 263:225–241. https://doi.org/10.1016/j.agrformet.2018.08.019.

Ferreira L.B., Cunha F.F.D., Oliveira R.A.D., Filho E.I.F., 2019. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM—A new approach. J Hydrol 572:556–570. https://doi.org/10.1016/j.jhydrol.2019.03.028.

Gao F., Feng G., Ouyang Y., Wang H., Fisher D., Adeli A., Jenkins J., 2017. Evaluation of reference evapotranspiration methods in arid, semiarid, and humid regions. J Am Water Resour Assoc 53(4):791–808. https://doi.org/10.1111/1752-1688.12530.

Gao L., Gong D., Cui N., Lv M., Feng Y., 2021. Evaluation of bio-inspired optimization algorithms hybrid with artificial neural network for reference crop evapotranspiration estimation. Comput Electron Agric 190:106466.https://doi.org/10.1016/j.compag.2021.106466.

GAP Regional Development Administration, 2024. 2024 GAP son durum. Şanlıurfa, Türkiye: Southeastern Anatolia Project Regional Development Administration.

Hargreaves, G.H. and Samani, Z.A., 1985: Reference Crop Evapotranspiration from Temperature. Applied Engineering in Agriculture, 1, 96-99. https://doi.org/10.13031/2013.26773

Hasan I.A., Yuce M.I., 2024. Prediction of potential evapotranspiration via machine learning and deep learning for sustainable water management in the Murat River Basin. Sustainability 16(24):11077. https://doi.org/10.3390/su162411077.

He Z., Xia K., Niu W., Aslam N., Hou J., 2018. Semisupervised SVM based on cuckoo search algorithm and its application. Math Probl Eng 243764: 1–13. https://doi.org/10.1155/2018/8243764.

Huang J., Liu J., Ye Y., Jiang Y., Lai Y., Qin X., Zhang L., Jiang Y., 2024. Mapping soil properties in the Haihun River sub-watershed, Yangtze River Basin, China, by integrating machine learning and variable selection. Sensors 24(12):3784. https://doi.org/10.3390/s24123784.

Huo Z., Feng S., Kang S., Dai X., 2012. Artificial neural network models for reference evapotranspiration in an arid area of northwest China. J Arid Environ 82:81–90. doi:10.1016/j.jaridenv.2012.01.016.

Jia H., Ma J., Song W., 2019. Multilevel thresholding segmentation for color image using modified moth-flame optimization. IEEE Access 7:44097–44134. https://doi.org/10.1109/access.2019.2908718.

Jiang S., Liang C., Cui N., Zhao L., Liu C., Feng Y., Hu X., Gong D., Zou Q., 2020. Water use efficiency and its drivers in four typical agroecosystems based on flux tower measurements. Agric For Meteorol 295:108200. https://doi.org/10.1016/j.agrformet.2020.108200.

Kartal V., 2024. Prediction of monthly evapotranspiration by artificial neural network model development with Levenberg–Marquardt method in Elazig, Turkey. Environ Sci Pollut Res 31:20953–20969. https://doi.org/10.1007/s11356-024-32464-1.

Kaveh A,. Farhoudi N., 2013. A new optimization method: Dolphin echolocation. Adv Eng Softw 59:53–70. https://doi.org/10.1016/j.advengsoft.2013.03.004.

Khairan H.E., Zubaidi S.L., Raza S.F., Hameed M., Al-Ansari N., Ridha H.M., 2023. Examination of single- and hybrid-based metaheuristic algorithms in ANN reference evapotranspiration estimating. Sustainability 15(19):14222. https://doi.org/10.3390/su151914222.

Kiani F., Seyyedabbasi A., Nematzadeh S., Candan F., Çevik T., Anka F.A., Randazzo G., Lanza S., Muzirafuti A., 2022. Adaptive metaheuristic-based methods for autonomous robot path planning: Sustainable agricultural applications. Appl Sci 12(3):943. https://doi.org/10.3390/app12030943.

Kisi O., Alizamir M., 2018. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs wavelet neural networks. Agric For Meteorol 263:41–48. https://doi.org/10.1016/j.agrformet.2018.08.007.

Kisi O., Keshtegar B., Zounemat-Kermani M., Heddam S., Trung N.T., 2021. Modeling reference evapotranspiration using a novel regression-based method: Radial basis M5 model tree. Theor Appl Climatol 145:639–659. https://doi.org/10.1007/s00704-021-03645-6.

Kumar M., Raghuwanshi N.S., Singh R., 2011. Artificial neural networks approach in evapotranspiration modeling: A review. Irrig Sci 29(1):11–25. https://doi.org/10.1007/s00271-010-0230-8.

Landeras G., Ortiz-Barredo A., López J.J., 2008. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). Agric Water Manag 95(5):553–565. https://doi.org/10.1016/j.agwat.2007.12.011.

Li M.F., Tang X.P., Wu W., Liu H.B. 2013. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. Energy Convers Manag 70:139–148.https://doi.org/10.1016/j.enconman.2013.03.004.

Li Y., Li T., Liu H., 2017. Recent advances in feature selection and its applications. Knowl Inf Syst 53(3):551–577.https://doi.org/10.1007/s10115-017-1059-8.

Li Z., Zhou Y., Zhang S., Song J., 2016. Lévy-flight moth-flame algorithm for function optimization and engineering design problems. Math Probl Eng 423930, 1–22. https://doi.org/10.1155/2016/1423930.

Liu C., Cui N., Gong D., Hu X., Feng Y., 2020. Evaluation of seasonal evapotranspiration of winter wheat in humid region of East China using large-weighted lysimeter and three models. J Hydrol 590:125388. https://doi.org/10.1016/j.jhydrol.2020.125388.

Ma L., Yu J., 2022. Hierarchical clustering cuckoo search optimization implemented in optimal setting of directional overcurrent relays. Math Probl Eng 966149, 1–18. https://doi.org/10.1155/2022/8966149.

Maroufpoor S., Bozorg-Haddad O., Maroufpoor E., 2020. Reference evapotranspiration estimating based on optimal input combination and hybrid artificial intelligent model: Hybridization of artificial neural network with grey wolf optimizer algorithm. J Hydrol 588:125060. https://doi.org/10.1016/j.jhydrol.2020.125060.

Mehdizadeh S., Mohammadi B., Pham Q.B., Duan Z., 2021. Development of boosted machine learning models for estimating daily reference evapotranspiration and comparison with empirical approaches. Water 13(24):3489. https://doi.org/10.3390/w13243489.

Mehdizadeh S., Mohammadi B., Pham Q.B., Khoi D.N., Linh N.T.T., 2020. Implementing novel hybrid models to improve indirect measurement of the daily soil temperature: Elman neural network coupled with gravitational search algorithm and ant colony optimization. Measurement 165:108127. https://doi.org/10.1016/j.measurement.2020.108127.

Mirjalili S., 2015. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowl-Based Syst 89: 228–249. https://doi.org/10.1016/j.knosys.2015.07.006.

Mirjalili S., Mirjalili S.M., Lewis A., 2014. Grey wolf optimizer. Adv Eng Softw 69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

Mirzania E., Kashani M.H., Golmohammadi G., Ibrahim O.R., Saroughi M., 2023. Hybrid COOT-ANN: A novel optimization algorithm for prediction of daily crop reference evapotranspiration in Australia. Theor Appl Climatol 154:201–218.https://doi.org/10.1007/s00704-023-04552-8.

Mohapatra S., Mohapatra P., 2023. An improved golden jackal optimization algorithm using opposition-based learning for global optimization and engineering problems. Int J Comput Intell Syst 16:147. https://doi.org/10.1007/s44196-023-00320-8.

Monteiro L.A., Sentelhas P.C., Pedra G.U., 2018. Assessment of NASA/POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation. Int J Climatol 38:1571–1581. https://doi.org/10.1002/joc.5282.

Moriasi D.N., Arnold J.G., Van Liew M.W., Bingner R.L., Harmel R.D., Veith T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50(3):885–900. https://doi.org/10.13031/2013.23153.

Muhammad M.K.I., Shahid S., Ismail T., Harun S., Kisi O., Yaseen Z.M., 2021. The development of evolutionary computing model for simulating reference evapotranspiration over Peninsular Malaysia. Theor Appl Climatol 144:1419–1434. https://doi.org/10.1007/s00704-021-03606-z.

Ndiaye P.M., Bodian A., Diop L., Deme A., Dezetter A., Djaman K., 2020. Evaluation and calibration of alternative methods for estimating reference evapotranspiration in the Senegal River Basin. Hydrology 7(2):24. https://doi.org/10.3390/hydrology7020024.

Ozbuldu M., Irvem A., 2024. Assessment of seven different global climate models for historical temperature and precipitation in Hatay, Türkiye. Int J Environ Sci Technol. https://doi.org/10.1007/s13762-024-06033-5.

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E., 2011. Scikit-learn: Machine learning in Python. J Mach Learn Res 12:2825–2830.

Priestley, C.H.B. and Taylor, R.J. (1972) On the Assessment of Surface Heat Flux and Evaporation Using Large Scale Parameters. Monthly Weather Review, 100, 81-92. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Ravindran S.M., Bhaskaran S.K.M., Ambat S.K.N., 2021. A deep neural network architecture to model reference evapotranspiration using a single input meteorological parameter. Environmental Processes 8:1567–1599. https://doi.org/10.1007/s40710-021-00543-x.

Reis M.M., da Silva A.J., Zullo Junior J., Tuffi Santos L.D., Azevedo A.M., Lopes E.M.G., 2019. Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. Comput Electron Agric 165:104937. http://dx.doi.org/10.1016/j.compag.2019.104937.

Rodrigues G.C., Braga R.P., 2021. Estimation of daily reference evapotranspiration from NASA POWER reanalysis products in a hot summer Mediterranean climate. Agronomy 11(10):2077. https://doi.org/10.3390/agronomy11102077

Roy D.K., Lal A., Sarker K.K., Saha K.K., Datta B., 2021. Optimization algorithms as training approaches for prediction of reference evapotranspiration using adaptive neuro fuzzy inference system. Agric Water Manag 255:107003. https://doi.org/10.1016/j.agwat.2021.107003.

Safari M.J.S., Mohammadi B., Kargar K., 2020. Invasive weed optimization-based adaptive neuro-fuzzy inference system hybrid model for sediment transport with a bed deposit. J Clean Prod 276:124267. https://doi.org/10.1016/j.jclepro.2020.124267.

Sayyahi F., Farzin S., Karami H., 2021. Forecasting daily and monthly reference evapotranspiration in the Aidoghmoush Basin using multilayer perceptron coupled with water wave optimization. Complexity 2021:6683759. https://doi.org/10.1155/2021/6683759.

Shiri J., Kisi Ö., Anderas 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). J Hydrol 414–415:302–316. https://doi.org/10.1016/j.jhydrol.2011.11.004.

Skhiri A., Ferhi A., Bousselmi A., Khlifi S., Mattar M.A., 2024. Artificial neural network for forecasting reference evapotranspiration in semi-arid bioclimatic regions. Water 16(4):602.https://doi.org/10.3390/w16040602.

Song X., Lu F., Xiao W., Zhu K., Zhou Y., Xie Z., 2019. Performance of 12 reference evapotranspiration estimation methods compared with the Penman–Monteith method and the potential influences in northeast China. Meteorol Appl 26:83–96. https://doi.org/10.1002/met.1739.

Tabari H., Hosseinzadeh Talaee P., 2013. Multilayer perceptron for reference evapotranspiration estimation in a semiarid region. Neural Comput Appl 23:341–348. https://doi.org/10.1007/s00521-012-0904-7.

Taylor K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J Geophys Res Atmos 106(D7):7183–7192. https://doi.org/10.1029/2000JD900719.

Tayyeh H.K., Mohammed R., 2023. Analysis of NASA POWER reanalysis products to predict temperature and precipitation in Euphrates River basin. J Hydrol 619:129327. https://doi.org/10.1016/j.jhydrol.2023.129327.

Tikhamarine Y., Malik A., Souag-Gamane D., Kisi O., 2020. Artificial intelligence models versus empirical equations for modeling monthly reference evapotranspiration. Environ Sci Pollut Res Int 27:30001–30019. https://doi.org/10.1007/s11356-020-08792-3.

Türkeş M., Tatlı H., 2009. Use of the standardized precipitation index (SPI) and a modified SPI for shaping the drought probabilities over Turkey. Int J Climatol 29:2270–2282. https://doi.org/10.1002/joc.1862.

Valiantzas J.D., 2015. Simplified limited data Penman's ET_0 formulas adapted for humid locations. J Hydrol 524:701–707. https://doi.org/10.1016/j.jhydrol.2015.03.021.

Wu L., Zhou H., Ma X., Fan J., Zhang F., 2019. Daily reference evapotranspiration prediction based on hybridized extreme learning machine model with bio-inspired optimization algorithms: Application in contrasting climates of China. J Hydrol 577:123960. https://doi.org/10.1016/j.jhydrol.2019.123960.

Yang X.S., 2010. Firefly algorithm, stochastic test functions and design optimisation. Int J Bio-Inspired Comput 2(2):78–84. https://dx.doi.org/10.1504/IJBIC.2010.032124.

Yang X.S., Deb S., 2009. Cuckoo search via Lévy flights. Proc World Congr Nat Biol Inspired Comput (NaBIC 2009):210–214. IEEE Publ. https://doi.org/10.1109/NABIC.2009.5393690.

Yin L., Wang L., Li T., Lu S., Yin Z., Liu X., Zheng W., 2023. U-Net-STN: A novel end-to-end lake boundary prediction model. Land 12(8):1602. https://doi.org/10.3390/land12081602.

Yu J., Zheng W., Xu L., Zhangzhong L., Zhang G., Shan F., 2020. A PSO-XGBoost model for estimating daily reference evapotranspiration in the solar greenhouse. Intell Autom Soft Comput 26:989–1003. https://doi.org/10.32604/iasc.2020.010130.

Zhang S., Zhou Y., 2017. Template matching using grey wolf optimizer with lateral inhibition. Optik 130:1229–1243. https://doi.org/10.1016/j.ijleo.2016.11.173.

Zhao L., Qing S., Li H., Qiu Z., Niu X., Shi Y., Chen S., Xing X., 2024. Estimating maize evapotranspiration based on hybrid back-propagation neural network models and meteorological, soil, and crop data. Int J Biometeorol 68:511–525. https://doi.org/10.1007/s00484-023-02608-y.

Zhou Q., Chen D., Hu Z., Chen X., 2021. Decompositions of Taylor diagram and DISO performance criteria. Int J Climatol 41:5726–5732. https://doi.org/10.1002/joc.7149.

Zhu B., Feng Y., Gong D., Jiang S., Zhao L., Cui N., 2020. Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. Comput Electron Agric 173:105430. https://doi.org/10.1016/j.compag.2020.105430.

Supplementary Material

Supplementary Text S1. Detailed Description of the Bio-Inspired Optimization Algorithms

S1.1. Moth-Flame Optimization (MFO)

The Moth-Flame Optimization (MFO) algorithm is a nature-inspired heuristic optimization method that mimics moth behavior (Mirjalili 2015). The algorithm provides a balance of exploration and exploitation by modelling moths as candidate solutions and flames as optimal solutions (Darwish 2018; Adnan et al. 2021). The movement of moths to flames is expressed by Equation S1.

$$S(M_i, F_j) = D_{ij} \times e^{b \times t} \times \cos(2\pi t) + F_j \qquad \textbf{(S1)}$$

where M_i denotes the position of moth i, F_j represents the position of flame j, $D_{ij}=|F_j-M_i|$ is the Euclidean distance between moth i and flame j, b is a constant coefficient determining the tightness of the spiral curve, and t is a random number in the interval [-1, 1] (Darwish 2018; Adnan et al. 2021). This movement pattern enables moths to initially scan a large search space and move closer to the optimal solutions as the iteration progresses. The MFO algorithm enhances the precision of the optimization process by dynamically adjusting the number of flames during the iteration procedure. The adaptive reduction of the number of flames is expressed by Equation S2.

$$Flame_N = N - n \times \frac{N-1}{T} \qquad \textbf{(S2)}$$

where N is the initial total number of flames, n is the current number of iterations and T is the maximum number of iterations (Lin et al. 2020). This mechanism enables the algorithm to undertake large-scale exploration in the initial stages and to focus on the optimal solutions with fewer flames in the later stages. Details of MFO can be found in Mirjalili (2015) and Adnan et al. (2021).

S1.2. Dolphin Echolocation Optimization (DEO)

The Dolphin Echolocation Optimization (DEO) algorithm is a meta-heuristic optimization method that models the echolocation mechanism of dolphins (Kaveh and Farhoudi 2013). The algorithm is based on the principle of position determination by analyzing the reflection of sound waves and provides an effective search strategy for complex optimization problems (Neda 2020; Alkasassbeh et al. 2023).

The algorithm utilizes high-frequency sound waves emitted by dolphins to systematically explore the solution space. The algorithm is composed of three fundamental components: Echolocation Process, Solution Update and Exploration & Exploitation. Initially, dolphins generate sound waves to evaluate possible solutions in the target region. By analyzing the time and amplitude of the reflected signals, the distance to the optimal solution is determined, and the dolphins update their position and move towards better solutions.

While DEO initially uses high-frequency signals to cover a large search area, it performs a more precise search by changing the frequency and amplitude parameters as it approaches the optimal solution. This adaptive mechanism supports convergence to the global solution by optimizing the algorithm's balance of exploration and exploitation (Darvishpoor et al. 2023). Detailed information on DEO can be found in Kaveh and Farhoudi (2013) and Darvishpoor et al. (2023).

S1.3. Firefly Optimization (FOA)

The Firefly Optimization (FOA) Algorithm is a meta-heuristic optimization method bio-inspired by the mechanism of light emission and attraction exhibited by fireflies in their natural habitat. This algorithm presents a population-based approach with global search capability for optimizing complex functions. (Yang et al. 2010). The FOA is theorized to be contingent on the intensity of the light emitted by fireflies and the gravitational force they exert on one another. The algorithm is predicated on three fundamental principles: Firstly, the attraction of fireflies to each other is gender-neutral. Secondly, there is a bias towards brighter (better solution) fireflies. Thirdly, the light intensity varies depending on the structure of the objective function (Chao and Horng 2015).

The optimization process of FOA functions on the principle of determining the attractiveness levels of the individuals representing the optimal solutions in the solution space and moving towards these individuals. The attractiveness of a firefly is directly proportional to light intensity and inversely proportional to distance. Consequently, superior solutions attract other fireflies, thereby leading the population to converge on the optimal solution (Roy et al. 2021). However, in the absence of a brighter individual, a firefly exhibits random behavior and attempts to identify new potential solutions. The algorithm's objective is to identify the optimum solution by establishing a balanced mechanism between exploration and exploitation processes (Yang et al. 2010). More details of FOA are referred to Yang et al. (2010) and Chao and Horng (2015).

S1.4. Grey Wolf Optimization (GWO)

Grey Wolf Optimization (GWO) is a meta-heuristic algorithm that models the social hunting strategies of grey wolves and provides effective solutions to multidimensional and nonlinear optimization problems (Mirjalili et al. 2014). The algorithm is based on four different hierarchical groups: alpha (α), beta (β), delta (δ) and omega (ω). Alpha wolves are designated as leaders and are tasked with hunting and orienting activities, while beta and delta wolves assume supportive roles. The fourth category, omega wolves, is responsible for maintaining balance within the group (Darwish 2018; Maroufpoor et al. 2020).

The GWO is comprised of three primary phases: reconnaissance, encirclement, and attack. Initially, the solution space is explored by randomly placing wolves. Subsequently, the three optimal solutions (α, β, δ) are identified, and the remaining wolves adjust their positions according to these solutions. As the location of the prey (target solution) becomes apparent, the algorithm is iteratively directed to the best solution (Darwish 2018; Faris et al. 2018).

Mathematically, each wolf is modelled as a candidate solution, with their initial positions determined randomly. The determination of the optimal solutions (α, β, δ) is achieved by updating the wolves' positions in accordance with the objective function. The population-based structure of the algorithm prevents it from becoming trapped in local minima by ensuring the maintenance of global search capability (Mirjalili et al. 2014). The detailed information can be found in Mirjalili et al. (2014).

S1.5. Golden Jackal Optimization (GJO)

The Golden Jackal Optimization (GJO) algorithm is an optimization method inspired by the hunting strategies of golden jackals. The algorithm's objective is to deliver effective solutions to global optimization problems through the mathematical modelling of the cooperative pursuit, encirclement and attack phases of jackals. In their natural habitat, jackals typically hunt in pairs, with both male and female individuals participating in the pursuit and attack of their prey, ensuring that the prey expends its energy before being captured (Chopra and Ansari 2022).

The GJO algorithm is comprised of three fundamental stages. Firstly, the prey location and orienting of the jackals towards it is undertaken. Following this is the encircling of the prey and the consumption of the escape energy. The third stage is the attack process, the aim of which is to reach the optimum solution. The algorithm dynamically manages the exploration and exploitation processes by modelling the movements of jackals with the Levy flight mechanism and adaptive escape energy concepts.

Mathematically, the new positions of the jackals are updated according to the position of the prey identified as the optimal solution. The update method is delineated in Equation S3.

$$Y(t+1) = \frac{1}{2}\left[Y_M(t) - E\left|RL \times Y_M(t) - Prey(t)\right| + Y_{FM}(t) - E\left|RL \times Y_{FM}(t) - Prey(t)\right|\right] \quad \textbf{(S3)}$$

where $Y(t+1)$ denotes the updated position at iteration $t+1$; $Y_M(t)$ and $Y_{FM}(t)$ represent the current positions of the male and female golden jackals, respectively; $Prey(t)$ represents the prey position; E is the prey evasion energy; RL is a random vector generated based on Lévy motion and t denotes the current iteration. Details of GJO can be found in Chopra and Ansari (2022).

S1.6. Cuckoo Search Optimization (CSO)

The Cuckoo Search Optimization (CSO) algorithm is a meta-heuristic optimization method inspired by biological systems and is based on the egg-laying and reproduction strategies of cuckoos. This algorithm aims to identify the optimal solution through the utilization of candidate solutions, which are randomly positioned within the solution space (Yang and Deb 2009). As is the case in natural processes, more effective solutions are preserved while weaker solutions are eliminated and substituted by new, potentially superior solutions (Gao et al. 2021).

One of the most significant components of CSO is its utilization of a distinctive random walk mechanism known as Levy flights. This approach offers a dynamic equilibrium between large-scale search and local solution refinement processes by integrating long-distance and short-distance steps. Through the process of Levy flights, the algorithm enhances its global optimization capacity by mitigating the risk of becoming trapped in local minima. Mathematically, the position update equation for cuckoo i at time t is expressed by Equation S4 (Yang and Deb 2009).

$$X_i^t = X_i^{t-1} + \alpha \times Levy(\lambda) \quad \textbf{(S4)}$$

where X_i^t denotes the new position of the cuckoo, X_i^{t-1} represents its position in the previous iteration, α is a coefficient that determines the step size, and Levy(λ) denotes a random step based on the Levy distribution. This step enables solutions to be searched for extensively across a large area and subsequently refined by concentrating them in specific regions. The Levy distribution is modelled as shown in Equation S5.

$$Levy(\lambda) = t^{-\lambda}, \qquad 1 < \lambda \leq 3 \quad \textbf{(S5)}$$

where λ determines the nature of the random walks and helps to balance large distance explorations with shorter distance precision searches. The details of CSO are referred to Yang and Deb (2009).

Supplementary Text S2. Detailed evaluation results of performance of the six optimization algorithms, individually for each of the six selected sites.

Table S1. Performance metrics (for abbreviations and formulas see main text)

Stations	Models	RSR	SI	MBE	MAE	NSE	R^2	RMSE	GPI
Adıyaman	GWO-MLP	0.18	0.13	0.05	0.43	0.97	0.97	0.50	0.17
	CSO-MLP	0.26	0.18	-0.14	0.53	0.93	0.96	0.71	0.09
	GJO-MLP	0.31	0.21	-0.05	0.71	0.90	0.91	0.85	-0.34
	FOA-MLP	0.24	0.16	0.03	0.50	0.94	0.95	0.65	-0.02
	DEO-MLP	0.16	0.11	-0.10	0.33	0.97	0.98	0.43	0.48
	MFO-MLP	0.30	0.20	0.09	0.69	0.91	0.98	0.81	-0.36
Diyarbakır	GWO-MLP	0.21	0.13	-0.10	0.45	0.96	0.97	0.49	0.13
	CSO-MLP	0.28	0.17	-0.13	0.64	0.92	0.96	0.88	-0.18
	GJO-MLP	0.24	0.15	-0.11	0.52	0.94	0.95	0.68	-0.01
	FOA-MLP	0.19	0.12	0.07	0.39	0.96	0.97	0.53	0.06
	DEO-MLP	0.16	0.10	0.03	0.32	0.98	0.98	0.47	0.25
	MFO-MLP	0.26	0.16	0.11	0.56	0.93	0.94	0.82	-0.31
Gaziantep	GWO-MLP	0.21	0.14	0.08	0.54	0.96	0.96	0.68	0.08
	CSO-MLP	0.30	0.20	0.26	0.74	0.91	0.92	0.98	-0.54
	GJO-MLP	0.23	0.15	-0.14	0.61	0.95	0.95	0.73	0.20
	FOA-MLP	0.26	0.17	-0.22	0.64	0.93	0.94	0.82	0.16
	DEO-MLP	0.17	0.11	-0.06	0.42	0.97	0.97	0.53	0.45
	MFO-MLP	0.27	0.18	0.12	0.75	0.93	0.98	0.86	-0.27
Siirt	GWO-MLP	0.25	0.15	-0.15	0.48	0.94	0.95	0.59	0.32
	CSO-MLP	0.33	0.20	0.30	0.64	0.89	0.93	0.79	-0.49
	GJO-MLP	0.35	0.21	-0.05	0.68	0.88	0.88	0.82	-0.26
	FOA-MLP	0.23	0.14	-0.15	0.45	0.95	0.97	0.55	0.41
	DEO-MLP	0.16	0.10	0.04	0.32	0.97	0.98	0.38	0.50
	MFO-MLP	0.33	0.20	0.08	0.64	0.89	0.89	0.78	-0.29
Şanlıurfa	GWO-MLP	0.22	0.14	-0.04	0.56	0.95	0.95	0.72	0.18
	CSO-MLP	0.22	0.14	-0.01	0.56	0.95	0.96	0.73	0.15
	GJO-MLP	0.34	0.21	0.29	0.84	0.89	0.89	1.12	-0.74
	FOA-MLP	0.25	0.16	0.27	0.66	0.94	0.95	0.82	-0.29
	DEO-MLP	0.21	0.13	0.06	0.57	0.96	0.96	0.68	0.11
	MFO-MLP	0.22	0.14	0.17	0.61	0.95	0.95	0.74	-0.08
Şırnak	GWO-MLP	0.19	0.12	0.01	0.40	0.96	0.96	0.49	0.19
	CSO-MLP	0.34	0.21	-0.07	0.71	0.89	0.93	0.88	-0.37
	GJO-MLP	0.26	0.16	-0.23	0.55	0.93	0.94	0.68	0.12
	FOA-MLP	0.21	0.13	0.05	0.40	0.96	0.96	0.53	0.12
	DEO-MLP	0.18	0.11	0.05	0.38	0.97	0.97	0.47	0.20
	MFO-MLP	0.31	0.20	0.05	0.66	0.90	0.90	0.82	-0.42

Additional references

Adnan R.M., Mostafa R.R., Islam A.R.M.T., Kisi O., Kuriqi A., Heddam S., 2021. Estimating reference evapotranspiration using hybrid adaptive fuzzy inferencing coupled with heuristic algorithms. Comput Electron Agric 191: 106541. https://doi.org/10.1016/j.compag.2021.106541.

Alkasassbeh A., Yasin B., Almasaeid H.H., 2023. Tuned mass damper parameters design by means of meta-heuristic optimization algorithms. Civil Eng Archit 11(3): 1523–1537. https://doi.org/10.13189/cea.2023.110333.

Chao C., Horng M., 2015. The construction of support vector machine classifier using the firefly algorithm. Comput Intell Neurosci 2015:212719:1–8. https://doi.org/10.1155/2015/212719.

Chopra N., Ansari M.M., 2022. Golden jackal optimization: A novel nature-inspired optimizer for engineering applications. Expert Syst Appl 198:116924.https://doi.org/10.1016/j.eswa.2022.116924.

Darvishpoor S., Darvishpour A., Escarcega M., Hassanalian M., 2023. Nature-inspired algorithms from oceans to space: a comprehensive review of heuristic and meta-heuristic optimization algorithms and their potential applications in drones. Drones 7(7):427. https://doi.org/10.3390/drones7070427.

Darwish A., 2018. Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications. Future Comput Inform J 3(2):231–246.https://doi.org/10.1016/j.fcij.2018.06.001.

Kaveh A,. Farhoudi N., 2013. A new optimization method: Dolphin echolocation. Adv Eng Softw 59:53–70. https://doi.org/10.1016/j.advengsoft.2013.03.004.

Lin G.Q., Li L.L., Tseng M.L., Liu H.M., Yuan D.D., Tan R.R., 2020. An improved moth-flame optimization algorithm for support vector machine prediction of photovoltaic power generation. J Clean Prod 253:119966. https://doi.org/10.1016/j.jclepro.2020.119966.

Maroufpoor S., Bozorg-Haddad O., Maroufpoor E., 2020. Reference evapotranspiration estimating based on optimal input combination and hybrid artificial intelligent model: Hybridization of artificial neural network with grey wolf optimizer algorithm. J Hydrol 588:125060. https://doi.org/10.1016/j.jhydrol.2020.125060.

Mirjalili S., 2015. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowl-Based Syst 89: 228–249. https://doi.org/10.1016/j.knosys.2015.07.006.

Mirjalili S., Mirjalili S.M., Lewis A., 2014. Grey wolf optimizer. Adv Eng Softw 69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

Neda O.M., 2020. Optimal coordinated design of PSS and UPFC-POD using DEO algorithm to enhance damping performance. Int J Electr Comput Eng 10(6):6111–6121. https://doi.org/10.11591/ijece.v10i6.pp6111-6121.

Roy D.K., Lal A., Sarker K.K., Saha K.K., Datta B., 2021. Optimization algorithms as training approaches for prediction of reference evapotranspiration using adaptive neuro fuzzy inference system. Agric Water Manag 255:107003. https://doi.org/10.1016/j.agwat.2021.107003.

Yang X.S., Deb S., 2009. Cuckoo search via Lévy flights. Proc World Congr Nat Biol Inspired Comput (NaBIC 2009):210–214. IEEE Publ. https://doi.org/10.1109/NABIC.2009.5393690.

Yang X.S., 2010. Firefly algorithm, stochastic test functions and design optimisation. Int J Bio-Inspired Comput 2(2):78–84. https://dx.doi.org/10.1504/IJBIC.2010.032124.