ORIGINAL PAPER

A novel hybrid AIG-SVR model for estimating daily reference evapotranspiration

Ehsan Mirzania[1] · Dinesh Kumar Vishwakarma[2] · Quynh-Anh Thi Bui[3] · Shahab S Band[4] · Reza Dehghani[5]

Abstract

Reference evapotranspiration (ET_o) simulation is of great importance for various procedures of the hydrological cycle, irrigation, agronomic management, and planning and management of water resources. Agricultural water management and irrigation scheduling require accurate estimation of crop water requirements. The reference evapotranspiration (ET_o), which is used to calculate the evapotranspiration of each crop, is estimated first. The purpose of this study was to develop a new coupled model for estimating daily ET_o. In this research, differing artificial intelligence (AI) approaches including support vector regression (SVR), algorithm of innovative gunner (AIG), and hybrid algorithm of innovative gunner-support vector regression (AIG-SVR) models were used in two different climatic conditions to estimate reference evapotranspiration, one station in the dry region (Marree Aero) and one station in the humid region (St Helen Aerodrome) in Australia. The results of the AIG-SVR model were compared with those of the conventional support vector regression (SVR) model using several performance evaluation methods comprising the statistical criteria including correlation coefficient (R), root mean square error (RMSE), Nash–Sutcliffe coefficient (NS), and RMSE-observations standard deviation ratio (RSR)). Evaluation results showed that the developed coupled models yielded better results than the classic SVR, with the AIG-SVR outperforming the SVR. The results showed that ETo values predicted by all AIG-SVR models agreed well with the corresponding observed values, with R, RMSE (mm day^{-1}), NS, and RSR = 0.945, 1.124, 0.894, 0.325 respectively in Marree Aero station and 0.951, 0.476, 0.905, and 0.307 respectively in St Helen Aerodrome station in testing data sets. In the local scenario, RMSE and RSR reduced by up to 15.805% (1.335 to 1.124) and 15.803% (0.386 to 0.325), respectively, and NSE and R increased by up to 5.176 (0.850 to 0.894) and 2.383% (0.923 to 0.945), respectively, in testing data sets at Marree Aero station whereas RMSE and RSR reduced by up to 19.594% (0.592 to 0.476) and 19.633% (0.382 to 0.307), respectively, and NSE and R increased by up to 6.096 (0.853 to 0.905) and 2.922% (0.924 to 0.951), respectively, in testing data sets at St Helen Aerodrome station. The findings reveal that the AIG-SVR model performs better than the SVR model. As a result of this study, AIG-SVR and SVR models can both provide important insights into how ET_o simulating can be improved. Based on the results obtained from the hybrid model in this research. ET_o in Australia is possibly considered an application system in Australia.

Keywords Algorithm of the innovative gunner · Hybrid model · Reference evapotranspiration · Support vector regression

Introduction

Evapotranspiration (ET_o) is the procedure by which water derived from the soil and vegetation layer is released into the atmosphere. It represents the release of water from a vegetative surface as a result of the integrated the transpiration cycle of plants and the evaporation of soil and atmospheric moisture (Elbeltagi et al. 2022c; Raza et al. 2022). It represents the amount of water used by plants. It is proportionate to consumption. ET_o analyses are crucial in studies pertaining to food security, the management and maintenance of land, pollution detection, irrigated agriculture scheduling,

✉ Ehsan Mirzania
    ehsan_mirzania@yahoo.com

✉ Shahab S Band
    shahab@yuntech.edu.tw

Extended author information available on the last page of the article

and water balance investigation (Ayaz et al. 2021; Vishwakarma et al. 2022b; Dorji et al. 2016; Zhang et al. 2021; Fu et al. 2022; Kadam et al. 2021). Furthermore, ET_o provides potential benefits for irrigation management (Kushwaha et al. 2021; Vishwakarma et al. 2022b; Elbeltagi et al. 2022b). Therefore, an accurate ET_o calculation is vital to improving irrigation efficiency, water reuse, and controlling infiltration (Singh and Marcy 2017; Shiri 2019; Kushwaha et al. 2021). Estimating reference evapotranspiration (ET_o) accurately is one of the major challenges of estimating water demand in agriculture. The evapotranspiration process is highly complex, its dependence on meteorological variables is growing, and a general model is difficult to generalize to different climates all factors that make an authentic estimation of this variable challenging (Kumar et al. 2011a, b). ET_o can be estimated using direct and indirect methods as well as mathematical models (Jing et al. 2019; Vishwakarma et al. 2022b).

The lysimeter, covariance of the eddy, and balance of energy according to the Bowen ratio method are among the direct measurement methods that can be used to determine ET_o values (Zhang et al. 2013; Kool et al. 2014; Sagar et al. 2022). They are expensive and are not widely accessible in several countries (Allen et al. 1998). In many parts of the world, where direct measurement of ET_o is either not feasible or is unavailable for financial reasons, the indirect approach is widely employed to estimate ET_o using an assemblage of readily accessible regional climate factors. One of the most common and important techniques used to evaluate ET_o is the equation FAO-56 Penman–Monteith (FAO-56 PM) (Allen et al. 1998). In spite of the fact that physically based and conceptualizations are trustworthy equipment for examining the phenomenon's actual physics, they do have constraints in practice. We may find black-box models useful when accuracy is more important than understanding the underlying physical principles.

Developed countries and renowned universities across the world develop new algorithms and models each year with the aim of enhancing accuracy, reducing error and cost, etc. Therefore, many researchers make use of such algorithms for developing models in various fields. They strive to examine the efficiency of their models for use on industrial and global scales. The primarily developed algorithms were not well addressed and known and there was no comprehensive knowledge about them. Hence, definite and accurate discussions of their pros and cons in the early years are very rare.

Since the last decades, artificial intelligence techniques have gained substantial attention in all areas of engineering and agriculture (Shukla et al. 2021; Elbeltagi et al. 2022a, c; Kumar et al. 2022; Vishwakarma et al. 2022a; Singh et al. 2022a, b). Several studies had been done in the field of ET by machine learning (ML) (Farias et al. 2020; Chen et al. 2020; Niaghi et al. 2021; Dias

et al. 2021). Over the past few years, the use of artificial intelligence (AI), for example, support vector regression (SVR), has been widely used to predict ET_o, and numerous research articles have been published through it (Elbeltagi et al. 2022b). Mohammadi and Mehdizadeh (2020) conducted an analysis of the daily ET_o in the semi-arid climate of Urmia, the arid climate of Isfahan, and the hyper-arid climate of Yazd in Iran using artificial intelligence. Roy et al. (2022) applied SVR, long short-term memory network (LSTM), sequence-to-sequence regression LSTM network (SSR-LSTM), bi-directional, and long short-term memory network (Bi-LSTM) and M5 model tree (M5Tree) models for daily forecast and an intricate process forward forecasting of ET_o. Wu et al. (2021) used SVR, random forest (RF), multilayer perceptron neural network (MLP), and K-nearest neighbor regression (KNN) for estimating daily ET_o. Yin et al. (2017) forecast the future variation of ET_o in a watershed of the mountainous inland in North-West China, using the SVR, extreme-learning machine (ELM), global climate model (GCM), and coupled model inter-comparison project phase 5 (CMIP5) models. As you already know, the presented algorithm was developed over the years 2018–2020, and there have even been a few studies dealing with it in the fields of electricity and computers, some of which are introduced in the following. It is worth mentioning that no research has been performed in the field of ET_0 and the presented study is the first one in the field of water resources management. With algorithm of innovative gunner (AIG) model, no study has been done on reference evapotranspiration, but similar studies that have been done with the AIG model can be referred to several studies. Dehghani and Torabi Poudeh (2021) investigated the effectiveness of hybrid artificial neural network (ANN) models with the black widow optimization (BWO), algorithm of innovative gunner (AIG), and chicken swarm optimization (CSO) algorithms to estimate monthly groundwater levels in Selseleh Plain, Iran. Dehghani and Poudeh (2021) investigated the river flow simulation of the Karkheh basin area located in Iran using artificial neural network (ANN) models accompanied by genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization (ACO), harmonic balance algorithm (HB), and algorithm of innovative gunner (AIG). In this research, the AIG algorithm has been used to increase the accuracy and effectiveness of SVR and its avoidance of being trapped in local optimal solutions. On the other hand, hybrid models can perform better than single models, but there is usually an exchange between training time and model performance. Therefore, a model that is both highly efficient and can be trained quickly is recommended for model development. Using an artificial algorithm helped the researchers

resolve some weaknesses of SVR, including being trapped in local optimal solutions, which reduces accuracy and increases error. The proposed new model reduces smart model development costs. According to the research done, there is no conclusive and sufficient knowledge of this algorithm yet, and in this research, its pros and cons were investigated in application to the estimation of daily reference evapotranspiration. However, the application of the AIG-SVR hybrid model in predicting ET_o has not yet been investigated and requires further study to assess the model's effectiveness with respect to water resource problems.

Also, given that in recent years, many researches on different hybrid models have been done in this field, in the present study, the performance of the offered model was evaluated in comparison with highly accurate models. Positive and negative points of the models are mentioned in the following; however, presenting a discussion of different views requires further research in this area. In meta-exploration algorithms, the values may be unknowingly transferred outside the realms of their definitions since velocities with random values are added to the problem variables. In other algorithms, the solutions obtained in all repetitions are placed in the scope of the problem by using discrete values for the problem variables and, while obtaining an optimal solution will take a long time, the problem is located in the optimal solutions domain. Therefore, the AIG algorithm, which is a combination of continuous and discrete optimization, reduces the time in which one can achieve an optimal solution in a widespread search area as it avoids the local optimal solutions. This makes the algorithm acceptable for solving nonlinear problems with large dimensions at an appropriate speed in convergence towards an optimal solution. Therefore, in this study, a combination of AIG algorithm and SVR was developed and offered.

However, the above studies have concluded that the integrated model AIG-SVR is more accurate than the classic SVR model in terms of predictive power.

The aim of this research is to compare the abilities of the AIG ensemble ET_o model with that of the SVR model to predict the ET_o in arid and humid regions of Australia. For this purpose, data on meteorology were collected from two stations (namely, Marree Aero and St Helen Aerodrome station) situated in two states in Australia. The highest accuracy and least error criterion were used for identifying the best mode to predict the ET_o using the conventional method over SVR and hybrid AIG-SVR models. The remainder of the paper is organized as continuing to follow: methodology is described in the "Materials and methods" section, outlining the research field and the data utilized. The "Model performance assessment" section displays and discusses the acquired results. Finally,

the conclusions are provided to show the efficacy of the models created for ET_o prediction.

Materials and methods

Data used and case study

The data required for this research, featuring daily minimum and maximum air temperatures (T_{min}, T_{max}) and reference evapotranspiration (ET_o), were gathered from 1 January 2009 to 31 December 2020 records from two meteorological stations in 2 states from Australia. The data were obtained from the Australian Meteorological Agency (BOM) (http://www.bom.gov.au). ET_o data obtained from automated meteorological station records and satellite measurements. There are two stations in Australia, St. Helens Aerodrome and Marree Aero in two different states. Station St Helen Aerodrome is located in Tasmania State (latitude: 41.3379°S, longitude: 148.2854°E, and elevation: 48 m). Tasmania, formerly Van Diemen's Land, island state of Australia, has an area of 26,410 square miles (68,401 km²). The state comprises a main island called Tasmania. Tasmania, located in the midlatitude westerly wind belt and dominated by southern maritime air masses, generally enjoys a moist, equable climate, with mild to warm summers, mild winters in most settled areas, and rain during all seasons. Tasmania, which is located in the western belts and includes the air masses of the southern sea, generally has a humid climate and mild summers, and it rains in all seasons of the year. Average annual precipitation exceeds 100 inches (2500 mm) on the western ranges and declines eastward to less than 20 inches (510 mm) in some places; along the north coast, it exceeds 30 in (760 mm) in all locations (Scott and Roe 2022). Marree Aero Station is located in the state of South Australia (longitude: 138.07°E, latitude: 29.66°S, and elevation: 50 m). South Australia state is located in south-central Australia and occupies one of the continent's driest and barren regions. The area is 379,725 square miles (983,482 km²) (McCaskill and Richards 2023). The state of South Australia has wide plains and low elevations, and the surface of the lakes is covered by salt or clay-encrusted and rarely contain water. Only about one-fifth of the area receives annual precipitation of more than 10 in (250 mm), and less than half of that has more than 16 in (400 mm) (McCaskill and Richards 2023). According to Table 1, daily data at the two stations have the following statistical characteristics. Eighty percent and 20% of daily data were used for training and testing the models, respectively. The current study used the MATLAB R2018b software to implement the SVR and AIG-SVR models. Figure 1 depicts the regions and stations that were considered for this study. Figure 2 shows a time series diagram of the observational data for the selected stations, i.e.,

Table 1 Statistics for daily meteorological data drawn from two stations in Australia

Data set	Statistics	Marree Aero			St Helens Aerodrome		
		T_{min}	T_{max}	ET	T_{min}	T_{max}	ET
All data	Minimum	−3.361	11.9	0	−5.593	8.687	0
	Mean	13.71	29.151	6.898	9.09	18.269	2.769
	Maximum	34.9	48.3	19.4	21.2	41.8	11
	Deviation	7.372	7.98	3.412	4.323	4.303	1.574
	Skewness	0.125	0.173	0.366	0.012	0.72	0.792
Training	Minimum	−2.4	11.9	0	−0.9	8.687	0
	Mean	13.745	28.989	6.785	9.121	18.278	2.762
	Maximum	34.2	48.3	19.2	21.2	41.8	11
	Deviation	7.3	7.975	3.391	4.339	4.301	1.581
	Skewness	0.117	0.167	0.364	0.016	0.654	0.75
Testing	Minimum	−3.361	11.9	0	−5.593	9.6	0
	Mean	13.63	29.798	7.347	8.966	18.23	2.796
	Maximum	34.9	48.1	19.4	20.2	36.8	10.1
	Deviation	7.597	7.968	3.458	4.257	4.309	1.548
	Skewness	0.139	0.2	0.372	−0.006	0.98	0.971

**Fig. 1** Locations of the stations under study in Australia
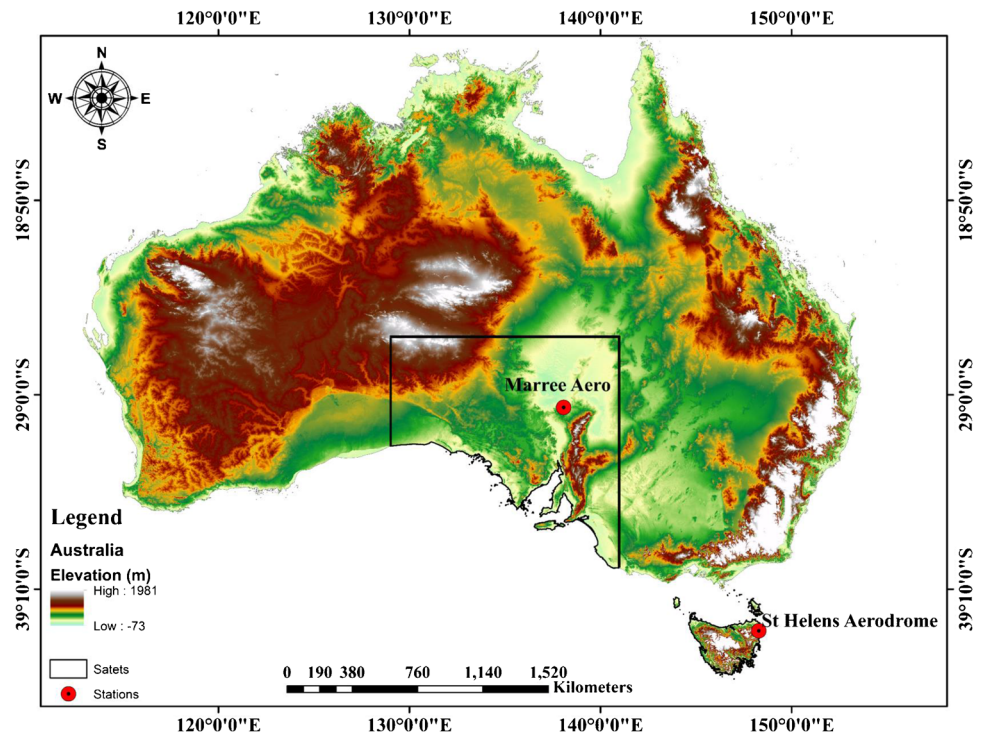
Marree Aero and St Helens Aerodrome. Figures 3 and 4 show the time series plots of T_{max} and T_{min} parameters for both stations.

Support vector regression (SVR)

An SVR is a method for solving nonlinear regression problems based on the SVM concept. Existing applications based on SVM, such as the SVR model, take operating risks into consideration so that they reduce the error among the observed and calculated values, not the error between observed and calculated values, as is the case with a plethora of different black-box forecasting techniques. SVR first applies linear regression to the data, The output is then passed through a nonlinear kernel to capture the data patterns with nonlinearities. Assuming the training data set $(X_1, Z_1), …, (X_n, Z_n)$ where x is consisted of the input vector, Z represents the output value that corresponds to the input

Fig. 2 Time series of daily ETo data for the training and testing phases in both stations
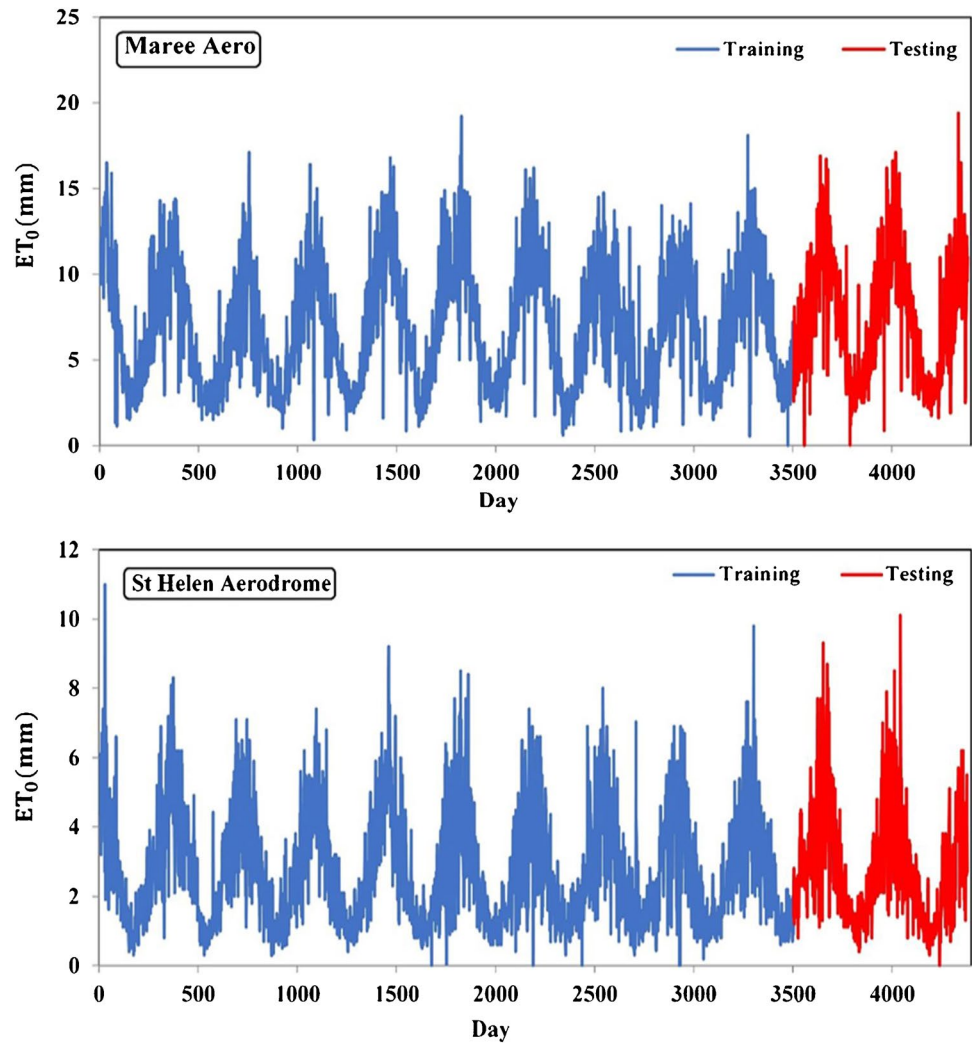
value and *N* denotes the total number of datasets. The overall performance of SVR is followed:

$$f(x) = w.\varnothing(x) + b \qquad (1)$$

where *w* is the vector of weight, *b* indicates the bias, and $\varnothing$ represents a nonlinear transmission function. The purpose of $\varnothing$ aims to map the vectors as inputs onto a feature space with a high degree of dimension by nonlinear mapping (Vapnik 1998). Figure 5 illustrates an overview of the SVR employed in this research. Table 2 shows the parameters used in the SVR model development.

Algorithm of the innovative gunner (AIG)

AIG invented by Pijarski and Kacejko (2019) solves different optimization problems using a new, efficient, and quick meta-heuristic optimization algorithm. The algorithm has some advantages, including the following: high convergence, finding an optimal solution with minimum cost, and high

accuracy in a short period of time. The algorithm discussed uses a plethora of heuristics in the response vectors that are used to search space using a swarm approach to find and converge to the optimal solution. The AIG is highly capable of avoiding local optimal solutions. This algorithm is predicted to provide more accurate results than other swarm intelligence algorithms (for example, particle swarm, genetic algorithm, shrimp, grasshopper, etc.). Figure 6 illustrates a flowchart diagram form of the AIG-SVR used throughout the research. Table 3 shows the parameters used in the AIG-SVR model development. This algorithm is run in the following stages (Dehghani and Torabi Poudeh 2021):

(i)   Randomly determine the first bullet's primary value
(ii)   Adjust the shooting distance (shooting distance to the target)
(iii) Create the bullet that was formed (in the third stage, the first bullet is used to create the second bullet.)
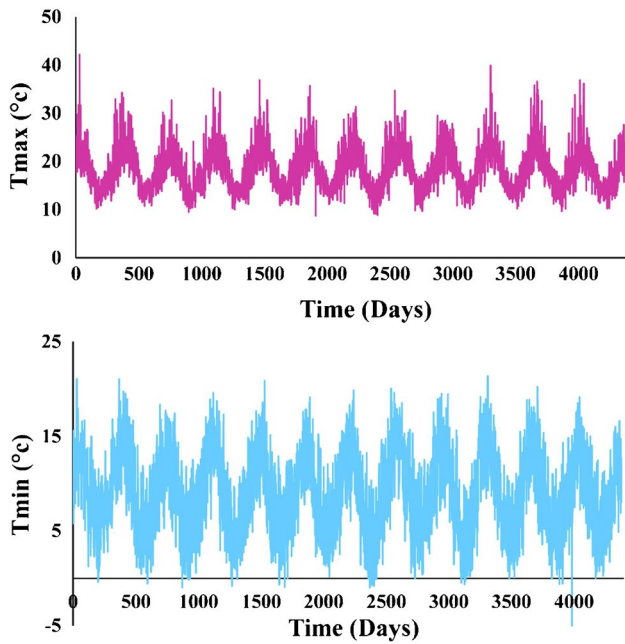(iv)   Investigate the likelihood that the bullet will hit the target (was it precisely targeted?)

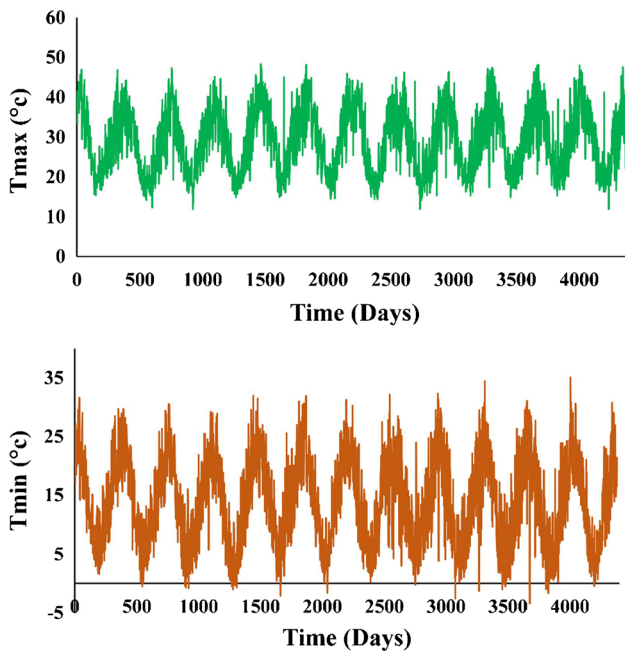**Fig. 3** Time series plot of Tmax and Tmin for St Helen Aerdrome station

Fig. 4 Time series plot of Tmax and Tmin for Marree Aero station

(v) *N* random bullets will be selected as the original bullets (in the event of an appropriate clash with the target)

(vi) Assessment and update of the bullet's location and coordinates when it collides with the target (In the case of a bullet at the target's center, a final set of
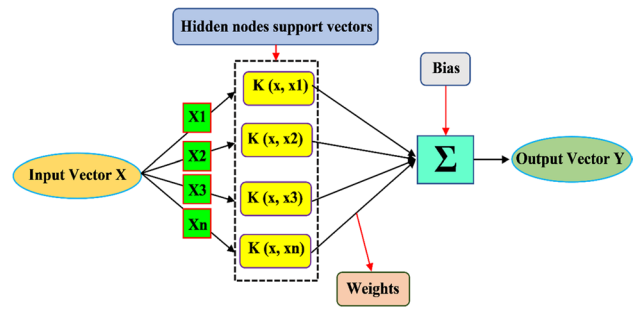

Fig. 5 A schematic structure of the SVR model

Table 2 Parameters used SVR model

Model parameters	SVR
Function estimation	f
The value of the gamma variable	10
Kernel type	RBF_Kernel
Kernel parameter	0.2

requirements is realized, and after that, everything is completed, but if the bullet does not strike the target, the early value must be re-evaluated)

(vii) Define the finest circumstance that has been recorded

(viii) End

Model performance assessment

Four statistical evaluation metrics, i.e., correlation coefficient (*R*); root mean square error (RMSE); Nash–Sutcliffe efficiency coefficient (NS), and RMSE-observation standard deviation ratio (RSR) were conducted to assess the performance proposed models in Eqs. 2-5 (Nash and Sutcliffe 1970; Vishwakarma et al. 2018, 2022b; Kushwaha et al. 2021; Elbeltagi et al. 2022b). Table 4 describes how the NSE and RSR results were analyzed.

Correlation coefficient (*R*)

$$\text{R} = \frac{\sum_{i=1}^{N}\left(\text{EO}-\overline{\text{EO}}\right)\left(\text{EP}-\overline{\text{EP}}\right)}{\sqrt{\sum_{i=1}^{N}\left(\text{EO}-\overline{\text{EO}}\right)^2 \cdot \sum_{i=1}^{N}\left(\text{EP}-\overline{\text{EP}}\right)^2}} \quad -1 \le R \le 1 \quad (2)$$

Root mean square error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(EP - \text{EO})^2} \quad 0 \le RMSE \le \infty \quad (3)$$

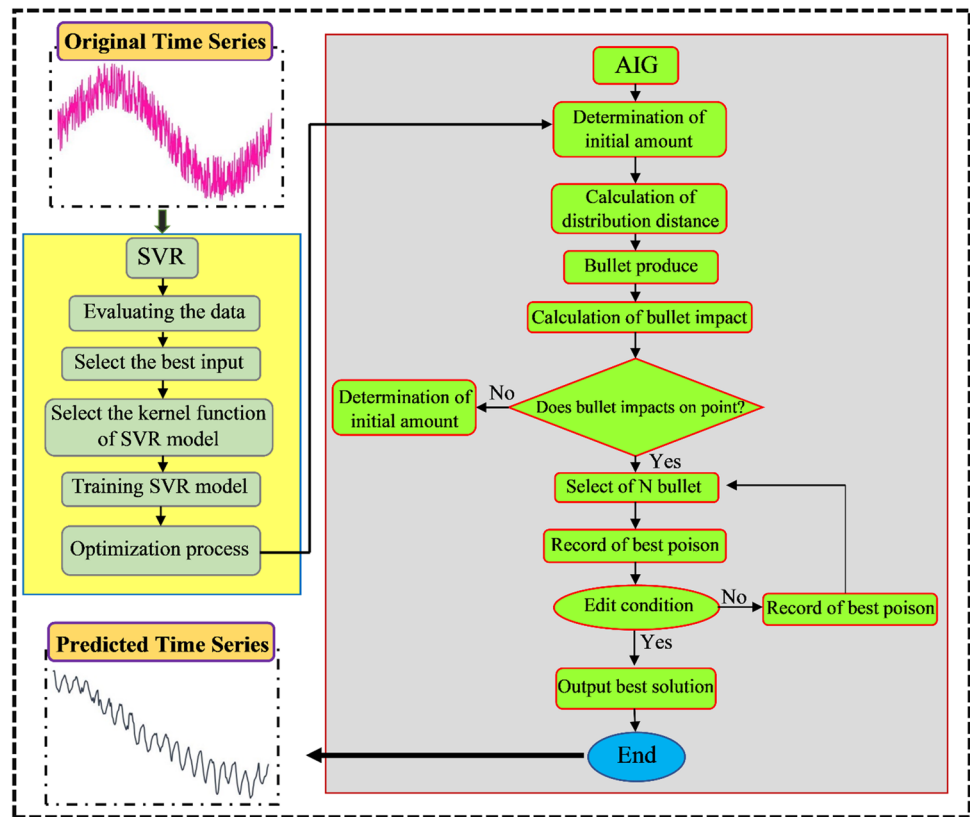Nash–Sutcliffe efficiency coefficient (NS)

Fig. 6 Flowchart of AIG-SVR

Table 3 Parameters used for the AIG-SVR model

Model parameters	AIG-SVR
Angles α	$0.995 < \alpha \leq 1$
Correction angles α_{max}	90
Correction angles β_{max}	90

Table 4 The characteristics of acceptable results for RSR and NS

Performance rating	RSR	NS
Very good	$0.00 \leq RSR \leq 0.50$	$0.75 < NS \leq 1.00$
Good	$0.50 < RSR \leq 0.60$	$0.65 < NS \leq 0.75$
Satisfactory	$0.60 < RSR \leq 0.70$	$0.50 < NS \leq 0.65$
Unsatisfactory	$RSR > 0.70$	$NS \leq 0.50$

$$NS = 1 - \left[\frac{\sum_{i=1}^{N}(EO-EP)^2}{\sum_{i=1}^{N}\left(EO-\overline{EO}\right)^2} \right] \quad -\infty < NS \leq 1 \tag{4}$$

RMSE-observation standard deviation ratio (RSR)

$$RSR = \frac{\sqrt{\sum_{i=1}^{N}(EO-EP)^2}}{\sqrt{\sum_{i=1}^{N}\left(EO-\overline{EO}\right)^2}} \quad 0 \leq RSR \leq \infty \tag{5}$$

where in Eqs. 2-5, EO and EP are the observational and predicted ET_o data, N denotes the number of observations, $\overline{EO}$ and $\overline{EP}$ are the means of observational and predicted ET_o data. Figure 7 illustrates the study's overall methodology.

Application and results

Studies of evapotranspiration estimation in different regions of the world have received attention in the past years (Zhou et al. 2020; Yan and Mohammadian 2020; Yurtseven and Serengil 2021) that shows the importance of this subject. Table 5 summarizes the statistical results of the support vector regression (SVR) and innovative gunner (SVR-AIG) models in order to forecast daily ET_o at two stations in the training and testing phases. All five developed SVR and AIG-SVR models developed precise daily ET_o estimations during training and testing with a range of R, RMSE (mm day^{-1}), NS, and RSR between 0.917–0.951, 0.476–1.34 (mm day^{-1}), 0.841–0.905, and 0.307–0.398 respectively. The values of NS for both models were higher than 0.75, showing high precision in predicting daily ET_o. The MAE value was lower than 1.34 and 0.629 for Marree Aero and St Helen Aerodrome sites respectively. The amount of RSR for both models was lower than 0.50, stating excellent accuracy in predicting daily ET_o.

Fig. 7 Schematic diagram of the proposed methodology
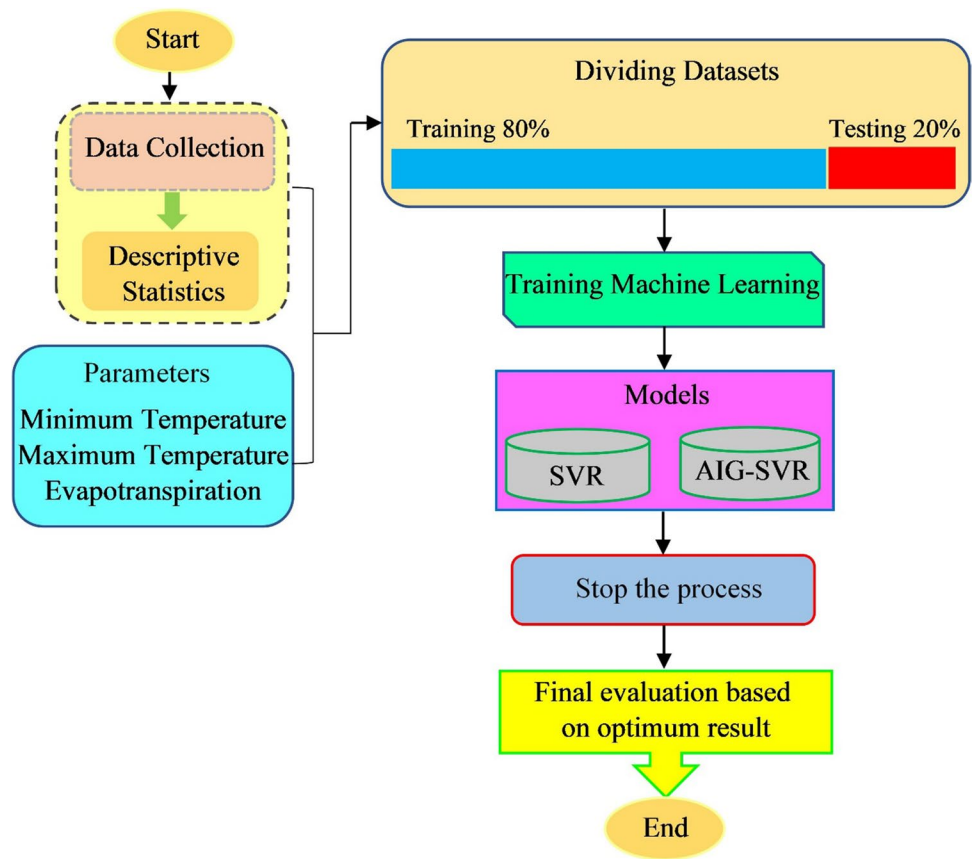
Table5 Statistical performance of the SVR and AIG-SVR hybrid models

| Models	Parameter		Sites	
			Marree Aero	St Helen Aerodrome
SVR	**R**	**Train**	0.918	0.917
		Test	0.923	0.924
	RMSE (mm)	**Train**	1.340	0.629
		Test	1.335	0.592
	NS	**Train**	0.843****	0.841****
		Test	0.850****	0.853****
	RSR	**Train**	0.395****	0.398****
		Test	0.386****	0.382****
AIG-SVR	**R**	**Train**	**0.931**	**0.937**
		Test	**0.945**	**0.951**
	RMSE (mm)	**Train**	**1.232**	**0.549**
		Test	**1.124**	**0.476**
	NS	**Train**	**0.867******	**0.879******
		Test	**0.894******	**0.905******
	RSR	**Train**	**0.363******	**0.347******
		Test	**0.325******	**0.307******

NS and RSR: ****very good, ***good, **satisfactory, *unsatisfactory

In addition, Table 5 illustrates the yield of the statistical parameter of daily ET_o estimation in the training phase at the Marree Aero stations which indicates $R = 0.918$ and 0.931; RMSE (mm day^{-1}) = 1.340 and 1.232; NS = 0.843 and 0.867; and RSR = 0.395 and 0.363 by SVR and SVR-AIG models respectively. In the testing phase (Table 5), the $R = 0.923$ and 0.945; RMSE (mm day^{-1}) = 1.335 and 1.124; NS = 0.850 and 0.894; and RSR = 0.386 and 0.325 by SVR and SVR-AIG models respectively. As Table 5 clearly shows that the AIG-SVR model outperformed in both the training and testing phase followed by the SVR model.

The statistical parameter of daily ET_o estimation in the training phase at the St Helens Aerodrome stations which indicates $R = 0.917$ and 0.937; RMSE (mm day^{-1}) = 0.629 and 0.549; NS = 0.841 and 0.879; and RSR = 0.398 and 0.347 by SVR and SVR-AIG models respectively. In the testing phase (Table 5), the $R = 0.924$ and 0.951; RMSE (mm day^{-1}) = 0.592 and 0.476; NS = 0.853 and 0.905; and RSR = 0.382 and 0.307 by SVR and SVR-AIG models respectively. As Table 5 clearly shows that the AIG-SVR model outperformed in both the training and testing phase followed by the SVR model at both stations. The AIG-SVR model adhered to the best statistical indicator for (i.e., minimum value of RMSE and SRS and maximum value of R and NS) for the training and testing phase and was selected as the

**Fig. 8** Results of SVR and AIG-SVR models for Marree Aero station
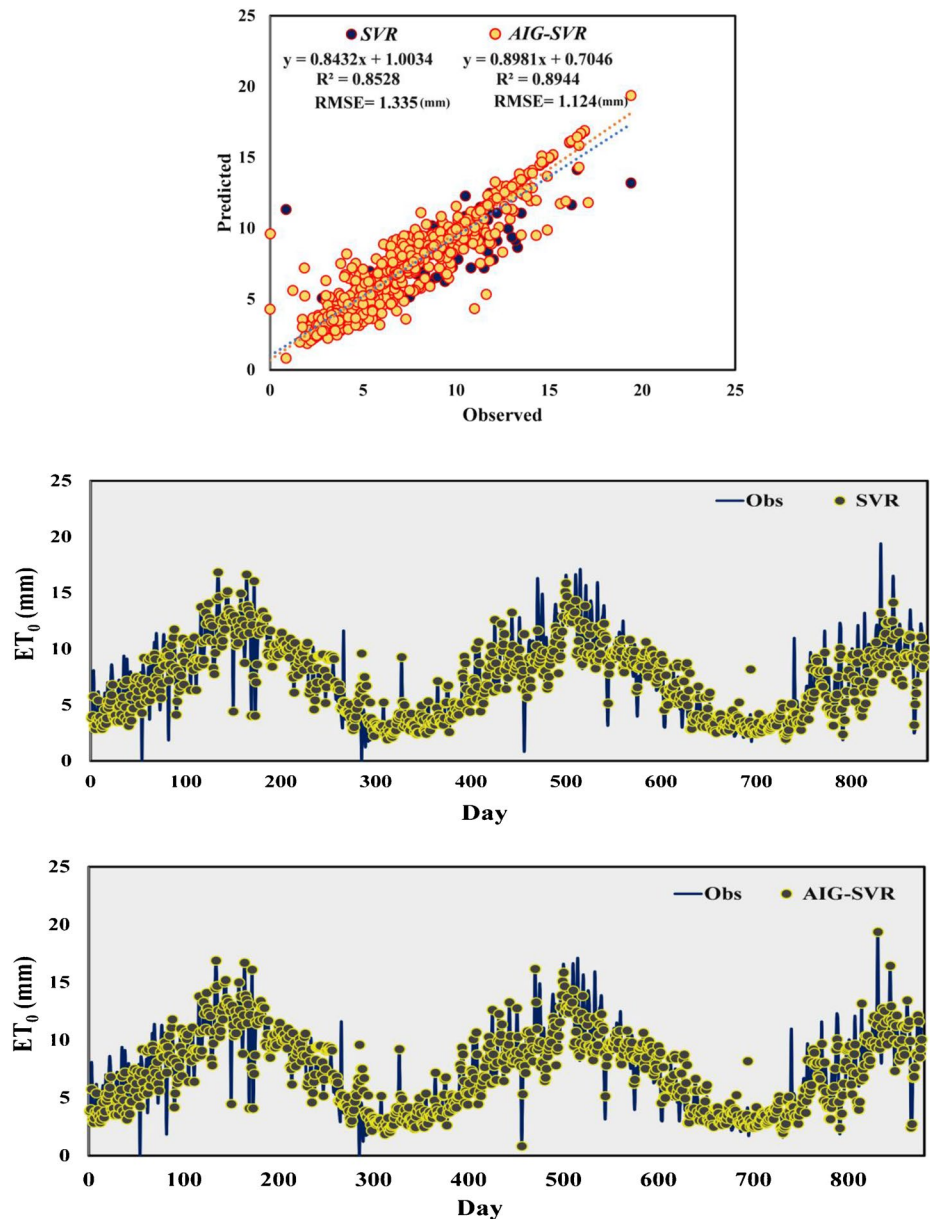
best both models. Figure 8 shows the scatter and time series diagram of observed ET_o values and estimated daily ET_o by SVR and AIG-SVR models throughout the years 2009–2020 for Marree Aero. Based on the scatter plot, the regression line (RL) offers the coefficient of determination (R^2) for both models. In the scatter plot, the value of $R^2 = 0.8528$ and RMSE = 1.335 mm day^{-1} were found in the SVR model and the value of $R^2 = 0.8944$ and RMSE = 1.124 mm day^{-1} were found in AIG-SVR model during the study period data sets. Figure 9 shows the scatter and time series diagram of observed ET_o values and estimated daily ET_o by SVR and AIG-SVR models throughout the years 2009–2020 for St Helens Aerodrome. According to scatter plots, the line of regression (RL) displays the coefficient of

determination (R^2) for both models. In the scatter plot, the value of $R^2 = 0.8539$ and RMSE = 0.9520 mm day^{-1} were found in the SVR model, and the values of $R^2 = 0.9057$ and RMSE = 0.4760 mm day^{-1} were found in AIG-SVR model during the study period data sets. The figures demonstrated the reasonable forecasting accuracy of the suggested models. Especially, the AIG-SVR model had similar excellence in performance in estimating ET_o to the SVM in both study sites. in Fig. 10, a Taylor diagram (TD) is used to evaluate the quantitative performance of the SVM and AIG-SVR models across the spatial distribution of observed and estimated daily ET_o values measured at Marree Aero and St Helens Aerodrome station during the testing period respectively. A Taylor diagram is a type of polar diagram that is
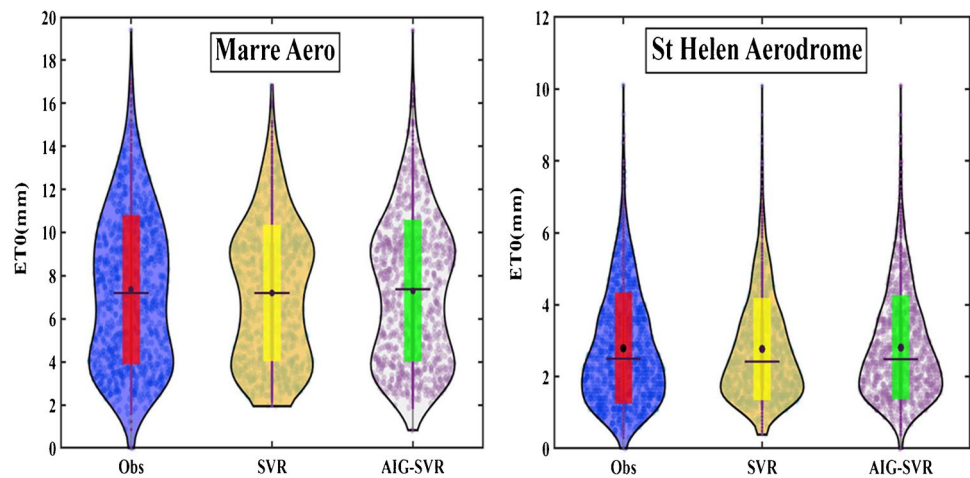
Fig. 9 Results of SVR and AIG-SVR models for St Helen Aerodrome station
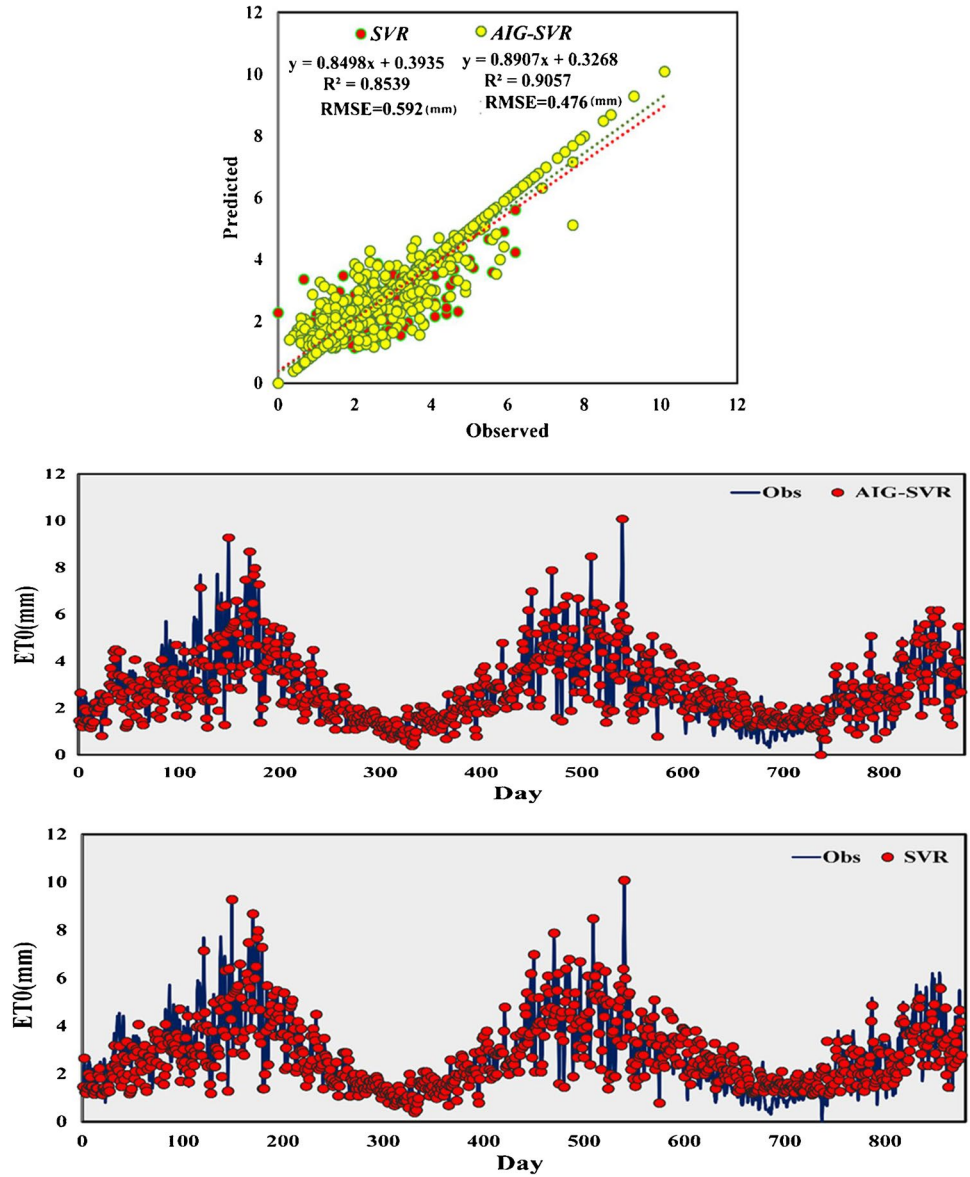
Fig. 10 Taylor diagram for statistical comparisons of the most accurate models for estimating ETo during the testing phase
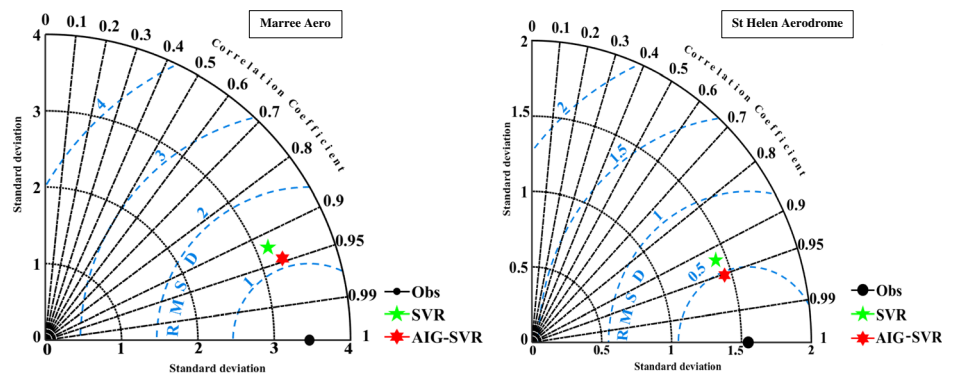
Fig. 11 Violin plots of the SVR and AIG-SVR models for two stations

used to examine the coefficient of correlation (R), standard deviation (SD), and root mean square deviations (RMSD) of the model. As shown in Fig. 10, the AIG-SVR model provides the estimates very close to the daily ET_o values observed in both sites. Additionally, the violin diagrams in Fig. 11 show the observed values for all models and the results of the box diagrams. The figure below illustrates that the distribution of the proposed AIG-SVR is closer to the distribution of the observed values in both sites.

Discussion

As a result of the only T_{max} and T_{min} input combinations, the results of hybrid models for simulating ET_o in Table 5 indicate that the algorithm of innovative gunner-support vector regression (AIG-SVR) performed higher than other support vector regression (SVR) models. As a case study for SVR-AIG, we obtained $R = 0.945$, RMSE $= 1.124$ (mm day^{-1}), NS $= 0.894$, and RSR $= 0.325$ in the course of the testing phase, and $R^2 = 0.8944$ and RMSE $= 1.1240$ (mm day^{-1}) in the overall data set at Marree Aero station. While the SVR–AIG model obtained $R = 0.951$, RMSE $= 0.476$ (mm day^{-1}), NS $= 0.905$, and RSR $= 0.307$ in the course of the testing phase and $R^2 = 0.9057$ and RMSE $= 0.4760$ (mm day^{-1}) in the overall data set at St Helen Aerodrome station. The results obtained from this study are very promising and promising in the sense that they are very encouraging. In general, RMSE and RSR reduced by up to 15.805% (1.335 to 1.124) and 15.803% (0.386 to 0.325), respectively, and NSE and R increased by up to 5.176 (0.850 to 0.894) and 2.383% (0.923 to 0.945), respectively, in testing data sets at Marree Aero station whereas RMSE and RSR reduced by up to 19.594% (0.592 to 0.476) and 19.633% (0.382 to 0.307), respectively, and NSE and R increased by up to 6.096 (0.853 to 0.905) and 2.922% (0.924 to 0.951), respectively, in testing data sets at St Helen Aerodrome station which are superior to the values reported by Dehghani et al. (2021) and Roshni et al. (2022).

In general, both the models perform best as if they followed the statistical criteria, but overall, it can be said that the AIG-SVR model had the performance superior to the SVR model. In a recently published paper, Roshni et al. (2022) reported that groundwater level can be forecasted with sufficient accuracy by hybrid machine learning combined with signal decomposition, i.e., SVR-AIG, and it was found that the high values of the R, NSE, RMSE, RSR, and Legates and McCabe's index (E_{LM}) were 0.995, 0.99, 0.151 (m), 0.096, and 0.916 in the training stage, and 0.941, 0.879, 0.146 (m), 0.346, and 0.660, respectively, in the testing stage which were significantly less than the obtained values in our present study. In another study, Dehghani et al. (2021) compared support vector regression (SVR) hybridized with metaheuristic algorithms of chicken swarm optimization (CSO), social ski-driver (SSD) optimization, black widow optimization (BWO), and the algorithm of the innovative gunner (AIG) for predicting dissolved oxygen concentration, and they reported that none of the reported SVR–AIG models was able to reduce the RMSE below the level of 0.644 mg/l. A similar result was found by Dehghani and Torabi Poudeh (2021) for groundwater level forecasting using ANN and ANN-AIG, ANN-BWO, and ANN-CSO and they reported ANN-AIG model was the only one possessing all statistical indices (e.g., $R^2 = 0.964$–0.995, RMSE $= 0.273$–0.71 m, MAE $= 0.219$–0.059 m, NSE $= 0.886$–0.978, and PBIAS $= 0.002$–0.001) in the validation stage, therefore highlighting the extent and the importance of the modeling framework reported in the present study. Dehghani et al. (2022a, b) and Adnan et al. (2022) also reported AIG-SVR has the greatest efficacy with the least error and increases the model efficiency in comparison with another model's hybridizers.

In accordance with the achieved estimation accuracy, the minimum and maximum air temperatures were used for 17 years of weather the data span. On average, the 17-year daily-scale climatic information examined was sufficient to construct predictive models. AIG-SVR has also been noted

for its remarkable ability to replicate the actual daily ET_o process.

In a decade of research, artificial intelligence models have been used to estimation of daily ET_o utilizing statistical metrics to validate against literature studies, and the AIG-SVR model presented here demonstrated its predictability capability (Dehghani and Torabi Poudeh 2021).

Conclusions

Many climatic variables are required to compute ET_o when using the P–M model, such as wind speed, solar radiation, relative humidity, and air temperature. P-M models are therefore difficult and laborious to apply to numerous areas where comprehensive input data is unavailable. An attempt was made in the present study to improve the modeling accuracy of the AIG-SVR in daily ET_o estimation. This study assessed, for the first time, based on only two meteorological data sets (T_{max} and T_{min}) as only two inputs in different climates of Australia (Marree Aerodrome and St Helen Aerodrome) to estimate ET_o. This paper investigated the utilization of two machine learning algorithms, such as the AIG-SVR and SVR models, to accurately estimate daily ET_o. The classic SVR was coupled with optimization algorithms such as the algorithm of innovative gunner (AIG). So, novel hybrid AIG-SVR models were proposed and implemented. Furthermore, using the evaluation criteria, we concluded that both models were able to estimate the relatively high ET_o level accurately. While the SVR–AIG model exhibited a higher level of accuracy and fewer errors than the SVR models, it also showed higher variance. As a result of the study, the algorithm AIG outperformed all others (based on correlation and RMSE criteria). In the study, results indicated the possibility of hybridizing the SVR model to improve daily ET_o prediction in climates with contrasting climates and insufficient meteorological data. However, when minimum and maximum temperature variables were available, the AIG-SVR model could predict ET_o with a higher accuracy than SVM and empirical equations beyond the study areas. Future research works could implement a variety of hybrid models for ET_o and other hydrological modeling through coupling the other AI models with the other types of bio-inspired optimizers.

Author contribution Ehsan Mirzania: conceived the problem, conducted the data collection, and designed the analysis.

Quynh-Anh Thi Bui and Shahab S. Band: prepared draft and graphical design

Dinesh Kumar Vishwakarma and Reza Dehghani: contributed data and analysis tools and writing review and editing

Dinesh Kumar Vishwakarma: supervision

Declarations

Ethics approval Not applicable.

Consent to participate The authors express their consent to participate in research and review.

Consent for publication The authors express their consent for publication of research work.

Conflict of interest The authors declare no competing interests.

References

Adnan RM, Dai H-L, Mostafa RR et al (2022) Modeling multistep ahead dissolved oxygen concentration using improved support vector machines by a hybrid metaheuristic algorithm. Sustainability 14:3470. https://doi.org/10.3390/su14063470

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration-guidelines for computing crop water requirements-FAO irrigation and drainage paper 56, 300th edn. Fao, Rome, p D05109

Ayaz A, Rajesh M, Singh SK, Rehana S (2021) Estimation of reference evapotranspiration using machine learning models with limited data. AIMS Geosci 7:268–290. https://doi.org/10.3934/geosci.2021016

Chen Z, Zhu Z, Jiang H, Sun S (2020) Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. J Hydrol 591:. https://doi.org/10.1016/j.jhydrol.2020.125286

Dehghani R, Poudeh HT (2021) Applying hybrid artificial algorithms to the estimation of river flow: a case study of Karkheh catchment area. Arab J Geosci 14:768. https://doi.org/10.1007/s12517-021-07079-2

Dehghani R, Babaali H, Zeydalinejad N (2022a) Evaluation of statistical models and modern hybrid artificial intelligence in the simulation of precipitation runoff process. Sustain Water Resour Manag 8:154. https://doi.org/10.1007/s40899-022-00743-9

Dehghani R, Poudeh HT, Izadi Z (2022) The effect of climate change on groundwater level and its prediction using modern meta-heuristic model. Groundw Sustain Dev 16:100702. https://doi.org/10.1016/j.gsd.2021.100702

Dehghani R, Torabi Poudeh H (2021) Application of novel hybrid artificial intelligence algorithms to groundwater simulation. Int J Environ Sci Technol.https://doi.org/10.1007/s13762-021-03596-5

Dehghani R, Torabi Poudeh H, Izadi Z (2021) Dissolved oxygen concentration predictions for running waters with using hybrid machine learning techniques. Model Earth Syst Environ 1–15. https://doi.org/10.1007/s40808-021-01253-x

Dias SHB, Filgueiras R, Filho EIF, Arcanjo GS, Silva GHd, Mantovani EC, Cunha FFd (2021) Reference evapotranspiration of Brazil modeled with machine learning techniques and remote sensing. PLoS ONE 16(2):e0245834. https://doi.org/10.1371/journal.pone.0245834

Dorji, U., Olesen, E.J., and Seidenkrantz, S.M. (2016). Water balance in the complex mountainous terrain of Bhutan and linkages to

land use. J Hydrol: Reg Stud 7:. https://doi.org/10.1016/j.ejrh.2016.05.001

Elbeltagi A, Raza A, Hu Y et al (2022c) Data intelligence and hybrid metaheuristic algorithms-based estimation of reference evapotranspiration. Appl Water Sci 12:152. https://doi.org/10.1007/s13201-022-01667-7

Elbeltagi A, Kumar M, Kushwaha NL, et al (2022a) Drought indicator analysis and forecasting using data driven models: case study in Jaisalmer, India. Stoch Environ Res Risk Assess. https://doi.org/10.1007/s00477-022-02277-0

Elbeltagi A, Kushwaha NL, Rajput J, et al (2022b) Modelling daily reference evapotranspiration based on stacking hybridization of ANN with meta-heuristic algorithms under diverse agro-climatic conditions. Stoch Environ Res Risk Assess. https://doi.org/10.1007/s00477-022-02196-0

Farias DBSd, Althoff D, Rodrigues LN, Filgueiras R (2020) Performance evaluation of numerical and machine learning methods in estimating reference evapotranspiration in a Brazilian agricultural frontier. Theor Appl Climatol 142:1481–1492. https://doi.org/10.1007/s00704-020-03380-4

Fu J, Wang W, Shao Q, Xing W, Cao M, Wei J, Chen Z and Nie W (2022). Improved global evapotranspiration estimates using proportionality hypothesis-based water balance constraints, Remote Sens Environ 279:. https://doi.org/10.1016/j.rse.2022.113140

Jing W, Yaseen ZM, Shahid S et al (2019) Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions. Eng Appl Comput Fluid Mech 13:811–823. https://doi.org/10.1080/19942060.2019.1645045

Kadam SA, Gorantiwar SD, Mandre NP, Tale DP (2021) Crop coefficient for potato crop evapotranspiration estimation by field water balance method in semi-arid region, Maharashtra, India. Potato Res 64:421–433. https://doi.org/10.1007/s11540-020-09484-8

Kool D, Agam N, Lazarovitch N et al (2014) A review of approaches for evapotranspiration partitioning. Agric for Meteorol 184:56–70. https://doi.org/10.1016/j.agrformet.2013.09.003

Kumar A, Singh VK, Saran B et al (2022) Development of novel hybrid models for prediction of drought- and stress-tolerance indices in teosinte introgressed maize lines using artificial intelligence techniques. Sustainability 14:2287. https://doi.org/10.3390/su14042287

Kumar R, Shankar V, Kumar M (2011a) Modelling of crop reference evapotranspiration: a review. Univers J Environ Res Technol 1:239–246

Kumar R, Shankar V, Kumar M (2011b) Evaluation of evapotranspiration models for pea (Pisum Sativum) in Mid Hill Zone-India. Univers J Environ Res Technol 1:

Kushwaha NL, Rajput J, Elbeltagi A et al (2021) Data intelligence model and meta-heuristic algorithms-based pan evaporation modelling in two different agro-climatic zones: a case study from northern India. Atmosphere (basel) 12:1654. https://doi.org/10.3390/atmos12121654

McCaskill M, Richards ES (2023) South Australia. Encyclopedia Britannica. https://www.britannica.com/place/South-Australia. Accessed 7 Apr 2023

Mohammadi B, Mehdizadeh S (2020) Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. Agric Water Manag 237:106145. https://doi.org/10.1016/j.agwat.2020.106145

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—a discussion of principles. J Hydrol 10:282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Niaghi AR, Hassanijalilian O, Shiri J (2021) Estimation of reference evapotranspiration using spatial and temporal machine learning approaches. Hydrology 8:25. https://doi.org/10.3390/hydrology8010025

Pijarski P, Kacejko P (2019) A new metaheuristic optimization method: the algorithm of the innovative gunner (AIG). Eng Optim 51:2049–2068. https://doi.org/10.1080/0305215X.2019.1565282

Raza A, Al-Ansari N, Hu Y et al (2022) Misconceptions of reference and potential evapotranspiration: a PRISMA-guided comprehensive review. Hydrology 9:153. https://doi.org/10.3390/hydrology9090153

Roshni T, Mirzania E, HasanpourKashani M et al (2022) Hybrid support vector regression models with algorithm of innovative gunner for the simulation of groundwater level. Acta Geophys 70:1885–1898. https://doi.org/10.1007/s11600-022-00826-3

Roy DK, Sarkar TK, Kamar SS et al (2022) Daily prediction and multistep forward forecasting of reference evapotranspiration using LSTM and Bi-LSTM models. Agronomy 12:594

Sagar A, Hasan M, Singh DK et al (2022) Development of smart weighing lysimeter for measuring evapotranspiration and developing crop coefficient for greenhouse chrysanthemum. Sensors 22:6239. https://doi.org/10.3390/s22166239

Scott P, Roe M (2022) Tasmania, Encyclopedia Britannica, https://www.britannica.com/place/Tasmania. Accessed 7 Apr 2023

Shiri J (2019) Evaluation of a neuro-fuzzy technique in estimating pan evaporation values in low-altitude locations. Meteorol Appl 26:204–212. https://doi.org/10.1002/met.1753

Shukla R, Kumar P, Vishwakarma DK, et al (2021) Modeling of stage-discharge using back propagation ANN-, ANFIS-, and WANN-based computing techniques. Theor Appl Climatol. https://doi.org/10.1007/s00704-021-03863-y

Singh SK, Marcy N (2017) Comparison of simple and complex hydrological models for predicting catchment discharge under climate change. AIMS Geosci 3:467–497. https://doi.org/10.3934/geosci.2017.3.467

Singh AK, Kumar P, Ali R et al (2022a) An integrated statistical-machine learning approach for runoff prediction. Sustainability 14:8209. https://doi.org/10.3390/su14138209

Singh VK, Panda KC, Sagar A et al (2022b) Novel Genetic algorithm (GA) based hybrid machine learning-pedotransfer function (ML-PTF) for prediction of spatial pattern of saturated hydraulic conductivity. Eng Appl Comput Fluid Mech 16:1082–1099. https://doi.org/10.1080/19942060.2022.2071994

Vapnik V (1998) Statistical learning theory. John Wiley & Sons, Inc., Oxford

Vishwakarma DK, Kumar R, Pandey K et al (2018) Modeling of rainfall and ground water fluctuation of Gonda District Uttar Pradesh, India. Int J Curr Microbiol Appl Sci 7:2613–2618. https://doi.org/10.20546/ijcmas.2018.705.302

Vishwakarma DK, Pandey K, Kaur A et al (2022) Methods to estimate evapotranspiration in humid and subtropical climate conditions. Agric Water Manag 261:107378. https://doi.org/10.1016/j.agwat.2021.107378

Vishwakarma DK, Ali R, Bhat SA, et al (2022a) Pre- and post-dam river water temperature alteration prediction using advanced machine learning models. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-022-21596-x

Wu T, Zhang W, Jiao X et al (2021) Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration. Comput Electron Agric 184:106039. https://doi.org/10.1016/j.compag.2021.106039

Yan X, Mohammadian A (2020) Estimating future daily pan evaporation for Qatar using the Hargreaves model and statistically downscaled global climate model projections under RCP climate change scenarios. Arab J Geosci 13:938. https://doi.org/10.1007/s12517-020-05944-0

Yin Z, Feng Q, Yang L et al (2017) Future projection with an extreme-learning machine and support vector regression of reference evapotranspiration in a mountainous inland watershed in North-West China. Water 9:880

Yurtseven I, Serengil Y (2021) Comparison of different empirical methods and data-driven models for estimating reference evapotranspiration in semi-arid Central Anatolian Region of Turkey. Arab J Geosci 14:2033. https://doi.org/10.1007/s12517-021-08150-8

Zhang B, Liu Y, Xu D et al (2013) The dual crop coefficient approach to estimate and partitioning evapotranspiration of the winter wheat–summer maize crop sequence in North China Plain. Irrig Sci 31:1303–1316. https://doi.org/10.1007/s00271-013-0405-1

Zhang L, Schlaepfer RD, Chen Z, Zhao L, Li Q, Gu S and Lauenroth KW (2021). Precipitation and evapotranspiration partitioning on the three-river source region: a comparison between water balance and energy balance models, Journal of Hydrology: Reg Stud 38:. https://doi.org/10.1016/j.ejrh.2021.100936

Zhou Z, Zhao L, Lin A, Qin W, Lu Y, Li J, Zhong Y, He L (2020) Exploring the potential of deep factorization machine and various gradient boosting models in modeling daily reference evapotranspiration in China. Arab J Geosci 13:1287. https://doi.org/10.1007/s12517-020-06293-8

Authors and Affiliations

Ehsan Mirzania[1] · **Dinesh Kumar Vishwakarma[2]** · **Quynh-Anh Thi Bui[3]** · **Shahab S Band[4]** · **Reza Dehghani[5]**

Dinesh Kumar Vishwakarma
dinesh.vishwakarma4820@gmail.com

Quynh-Anh Thi Bui
quynhanhbt@utt.edu.vn

Reza Dehghani
reza.dehghani67@yahoo.com

[1] Department of Water Engineering, University of Tabriz, Tabriz, Iran

[2] Department of Irrigation and Drainage Engineering, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, Uttarakhand, India

[3] University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Hanoi 100000, Vietnam

[4] Future Technology Research Center, College of Future, National Yunlin University of Science and Technology, 123 University Road, Section 3Douliou, Yunlin 64002, Taiwan, Republic of China

[5] Faculty of Agriculture, Lorestan University, Khorramabad, Iran