

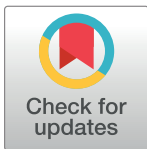
RESEARCH ARTICLE

Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction

Hamidreza Ghazvinian¹, Sayed-Farhad Mousavi¹, Hojat Karami¹, Saeed Farzin¹, Mohammad Ehteram¹, Md Shabbir Hossain², Chow Ming Fai³, Huzafa Bin Hashim⁴, Vijay P. Singh⁵, Faizah Che Ros⁶, Ali Najah Ahmed³, Haitham Abdulmohsin Afan^{4*}, Sai Hin Lai⁴, Ahmed El-Shafie⁴

1 Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan, Iran, **2** School of Energy, Geoscience, Infrastructure and Society, Department of Civil Engineering, Heriot-Watt University, Putrajaya, Malaysia, **3** Institute of Energy Infrastructure (IEI), Department of Civil Engineering, Universiti Tenaga Nasional, Selangor, Malaysia, **4** Department of Civil Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia, **5** Department of Biological and Agricultural Engineering, Zachry Department of Civil Engineering, Texas A&M University, College Station, Texas, United States of America, **6** Department of Environmental and Green Technology, Malaysia-Japan International Institute of Technology (MJIT), Universiti Teknologi Malaysia (UTM), Kuala Lumpur, Malaysia

* haitham.afan@gmail.com



OPEN ACCESS

Citation: Ghazvinian H, Mousavi S-F, Karami H, Farzin S, Ehteram M, Hossain MS, et al. (2019) Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction. PLoS ONE 14(5): e0217634. <https://doi.org/10.1371/journal.pone.0217634>

Editor: Yang Li, Northeast Electric Power University, CHINA

Received: January 29, 2019

Accepted: May 15, 2019

Published: May 31, 2019

Copyright: © 2019 Ghazvinian et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data is provided by the Turkish State Meteorological services of Ministry of Agriculture and Forestry, turkey. (<https://mevbis.mgm.gov.tr/mevbis/ui/index.html#/Workspace>). This data is payable and can be requested from the following contact details and charge of data providing will be applied: General Directorate of Meteorology, 1998-2016, Tel: +90 312 359 75 45, Fax: +90 312 360 25 51, Kütükçü Alibey Caddesi No: 4 06120 Kalaba, Keçiören / ANKARA.

Abstract

Solar energy is a major type of renewable energy, and its estimation is important for decision-makers. This study introduces a new prediction model for solar radiation based on support vector regression (SVR) and the improved particle swarm optimization (IPSO) algorithm. The new version of algorithm attempts to enhance the global search ability for the PSO. In practice, the SVR method has a few parameters that should be determined through a trial-and-error procedure while developing the prediction model. This procedure usually leads to non-optimal choices for these parameters and, hence, poor prediction accuracy. Therefore, there is a need to integrate the SVR model with an optimization algorithm to achieve optimal choices for these parameters. Thus, the IPSO algorithm, as an optimizer is integrated with SVR to obtain optimal values for the SVR parameters. To examine the proposed model, two solar radiation stations, Adana, Antakya and Konya, in Turkey, are considered for this study. In addition, different models have been tested for this prediction, namely, the M5 tree model (M5T), genetic programming (GP), SVR integrated with four different optimization algorithms SVR-PSO, SVR-IPSO, Genetic Algorithm (SVR-GA), FireFly Algorithm (SVR-FFA) and the multivariate adaptive regression (MARS) model. The sensitivity analysis is performed to achieve the highest accuracy level of the prediction by choosing different input parameters. Several performance measuring indices have been considered to examine the efficiency of all the prediction methods. The results show that SVR-IPSO outperformed M5T and MARS.

Funding: The authors would like to appreciate the financial support received from Bold 2025 grant coded RJO 10436494 by Innovation & Research Management Center (iRMC), Universiti Tenaga Nasional and the research grant coded UMRG RPO25A-18SUS to AE-S and BKS008-2016 funded by the University of Malaya.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Solar energy is one of the most important forms of energy. Although fossil fuels can produce a large amount of energy, they cause various kinds of pollution [1,2]. Undoubtedly the knowledge of solar radiation is important as it has direct or indirect impact on the current and future life [3]. This energy affects the agriculture, industry engineering, health and the tourism sector of any nation [2].

Solar radiation (SR), however, does not cause environmental pollution [3,4]. Solar energy can be converted to heat energy or electricity [5,6] and has high potential as an energy supply in various fields. Solar energy can be accessed easily and is not limited to specific regions in the world. Low maintenance cost is one of the most important features of solar energy [7]. Recently, there have been many advances in solar energy generation, such as solar cells [8,9]. The pyranometers or actiongraphs are used for the direct measurement of solar radiation [8]. The measurements can be very accurate with the new sensors, but also the device cost, installation and the maintenance cost to bear is a big drawback for many countries.

Different researchers have used mathematical and statistical methods and artificial intelligence techniques to compute solar radiation energy [3,8]. A combination of artificial intelligence and soft computing with mathematical and statistical methods has led to the development of mathematical models. Increased precision, speed, and ease of computation are considered desirable features for techniques based on artificial intelligence [10]. Statistical regression models have weaknesses that could be overcome with the incorporation of optimization algorithms. For example, SVR could be integrated with PSO to optimally adapt SVR's unknown parameters instead of using a trial-and-error procedure that could lead to non-optimal selection of the unknown parameters and, as a result, poor prediction accuracy.

Background

Artificial neural network (ANN) models were used by [11] to predict SR. The ANN model was compared with SVR, and it was found that multi-layer perceptron neural networks could achieve better RMSE than could SVR and other kinds of ANNs. Latitude, longitude, monthly minimum and maximum temperatures, and relative humidity were used as inputs. Landaras et al. [12] used genetic programming (GP) to compute SR. A radial neural network with four input layers, namely, maximum temperature, hours of sunshine, relative humidity, and minimum temperature, was used. The results based on GP had a higher correlation coefficient than the results of an ANN trained by PSO [13]. The number and structure of the neurons and hidden layers of the ANN model depended on the PSO. The results showed that the improved ANN was more accurate than the classic ANN model with simple architecture. Khatib et al. [14] compared different methods of computing SR and showed that regression methods had drawbacks in terms of identifying the unknown parameters and that artificial intelligence methods generally outperformed them. The unknown parameters of the regression methods were computed by optimization. K-means clustering and ANNs were used by [15] to compute SR. A comparison of the results showed that the clustering method had a smaller RMSE and MAE than the ANN model. SR has also been estimated by a hidden Markov model and a generalized fuzzy model [16].

Different combinations of meteorological parameters have been considered for predicting the SR. Atmospheric pressure, relative humidity, wind speed, and sunshine were used for the SR simulation. The best results were achieved by the combination that used all the inputs except wind speed. The ANN was improved by a genetic algorithm (GA) for computing SR [17]. The GA optimization method computed the number of hidden layers and neurons in the ANN. The results showed that the predictive ability of the ANN was related to the training algorithm and input combinations. A tree regression model and an ANN were used to predict

daily global SR (DGSR) for two locations [18]; the results showed that the ANN estimated DGSR satisfactorily. Mohammadi et al. [19] studied SR based on the combination of SVM and a wavelet method. Air temperature, humidity, and sunshine duration were considered as inputs. The new hybrid SVM achieved more accurate results than simple SVM and ANN models. Olatomive et al. [20] trained a neural network for the computation of solar radiation by computing the number of hidden layers, and the results were compared with those of GP and SVR. The results showed that the trained neural network reduced the RMSE value by 20 and 25%, respectively, compared to the GP and SVR. Premalatha et al. [21] used the Levenberg-Marquardt algorithm, resilient propagation and the scaled conjugate gradient for the development of a neural network and the simulation of solar radiation; the results showed that the mean absolute error of the Levenberg-Marquardt algorithm was 20 and 25% less than those of the resilient propagation and scaled conjugate gradient methods, respectively.

SR has been predicted by a least squares support vector machine (LSSVM) and the firefly optimization algorithm (FFOA) [22] by functionalizing the latter to optimally select the unknown parameters for the LSSVM. The results showed that the new hybrid structure achieved smaller RMSE than did GP and SVR. The accuracies of the three methods (the adaptive neuro-fuzzy interface system (ANFIS), SVR and ANN) were considered for computing SR [23]. The results showed that ANFIS with inputs of the daily maximum and minimum temperatures, hours of sunshine, and rainfall, exhibited better performance than other models. Ibrahim and Khatib [24] used the hybrid structure of Random Forest and Firefly algorithm for the computation of solar radiation; the results showed that the hybrid method had a lower value for the error index than the GP and neural network methods. Meenal et al. [25] used SVR and ANN to simulate solar radiation; the results showed that these models could reduce the RMSE by 20 and 25% over the empirical models. Kumar et al. [26] simulated solar radiation with the different neural networks, and the radial basis neural network had the best results in comparison to other kinds of neural networks; the MAE and RMSE values were negligible for the radial basis neural network. Voyant et al. [27] reviewed several methods for the simulation of solar radiation in the literature; the results showed that the kind of inputs and the accurate estimation of unknown parameters in the regression methods had important effects on the results. Wang et al. [28] trained a system of fuzzy rules using the particle swarm algorithm for the computation of solar radiation and compared the results to those of a neural network and genetic programming. This showed that the improved fuzzy rules increased the correlation coefficient between the observed data and simulated data in comparison to the neural network and GP methods.

Alfadda et al. [29] used the k nearest neighbours and the SVR method to determine solar radiation based on hours of sunshine, maximum temperature, minimum temperature and relative humidity; the results showed that the k nearest neighbours could reduce the RMSE by 20% in comparison to SVR.

Wang et al. [30] used radial basis neural network (RBNN), generalized regression neural network (GRNN) and multilayer perceptron neural network (MPNN) for estimating of solar radiation. The results indicated that MPNN and RBNN can predict more accurately compared with the GRNN and there is significant different among these models.

Rohani et al. [31] used the Gaussian progress regression with K fold cross validation for the estimating of daily solar radiation. The results proved that the new model can be used with small size of data group and it can predict better than empirical model.

However, the general results of the literature review showed that neural networks and fuzzy methods have good performance in estimating solar radiation but require the accurate determination of several parameters, such as weights or the number of hidden layers in a neural network [32,33]. In addition, the structures of these methods are more complex than that of regression methods would be if the regression methods could be applied to solar radiation.

Innovation and objectives

A literature review shows that SVR has been widely used for SR simulation [22,32–36]. Defining the best values of the regression parameters is important for regression models, as this step influences the final prediction accuracy for the whole process. Most of the previous studies determined the best values of the regression parameters using trial and error [22,27,35,37]. Conventionally the integration was basically developed on the basis two different concepts. The first one is to initialize the regularization parameters (three parameters) of the SVR and then identify the best fitness of these parameters by trials and errors for different group of these random initialization. On the other hand, the second concept was developed as modifications for the first concept by initializing the SRV regularization parameters within a predetermined domain to accelerate the training process and by adding k-fold procedure for avoiding the over-fitting for the model performance.

In the current research, the concept of the integration between the SVR and PSO is principally different than the previous ones, as it was developed to compute the objective function for the Root Mean Square Error (RMSE) for the initialized regular parameters of the SVR to search for their optimal values of the initial parameters which are considered as multiple decision variables using PSO algorithm.

In addition, the present study suggests a novel structure for an SVR model that is integrated with Improved PSO (IPSO) as an optimizer. IPSO determines the best values of the regression parameters in SVR, and then the SVR model is used to compute SR. The literature shows that PSO has high potential for use in different optimization applications, such as image processing, dam and reservoir operation, mathematical functions, hydrological prediction, and optimal design [36,38–41]. The new version of algorithm is defined for this paper so that the global ability search for the PSO algorithm increase and the algorithm can escape from local optimums well and thus, a new operator for the algorithm is defined to update the global solutions. Another innovation aspect of the current paper is related to the comprehensive evaluation of different models for the solar estimation in the different climates.

The objectives of this paper are i) to evaluate the ability of a new version of SVR to predict SR, ii) to compare the new method to the M5T, GP and multivariate adaptive regression models, and iii) to investigate the effect of different input variables on the models' predictive ability. Two case studies in Turkey are used to validate the proposed prediction model.

Methods

Support vector regression

For SVR, a linear function is defined that the related independent and dependent variables. A linear equation is used as the main equation in SVR and is expressed as

$$f(x) = W^{Tr} \cdot x + b \quad (1)$$

where x is the input variable, W is the weighting vector, b is the bias, Tr is the transpose, and $f(x)$ is the output variable. Vapnik et al. [42] suggested the following error function to prevent an overfitting deficit. The function is defined based on the following equation and is known as the epsilon intensive function:

$$|y - f(x)| = \begin{cases} 0 & \leftarrow \text{if } |y - f(x)| \leq \kappa \\ |y - f(x)| - \kappa & = \xi \end{cases} \quad (2)$$

is subject to

$$\begin{aligned} (w_i \cdot x_i + b) - y_i &< \kappa + \xi_i^+ \\ y_i - (w_i \cdot x_i + b) &\leq \kappa + \xi_i^- \end{aligned} \tag{3}$$

$$\xi_i^+, \xi_i^-$$

where ξ_i^+ and ξ_i^- are violations of the *i*th training data that are below and above $(0, \kappa + \kappa)$, κ is the permissible error threshold, y_i is the output variable, x_i is the input variable, w_i is the weight vector, ξ is the computed penalty for the estimated error, and b and w are two decision variables. The values of b and w are computed when the SVR completes the training level. The values of b and w are inserted into Eq (1), and $f(x)$ is computed. There are several kernel functions that can convert linear Eq (1) into nonlinear forms. The radio kernel function has been widely used in previous articles [22,27,35,37].

$$f(x) = w^{Tr} \cdot K(x, x_i) + b \tag{4}$$

$$K(x, x_i) = \exp\left(-\frac{|x - x_i|^2}{2\gamma^2}\right) \tag{5}$$

where $K(x, x_i)$ is the kernel function and γ is a parameter. The most important duty of SVR is to compute the values of the parameters κ and γ . Huang and Wang [43] found that these parameters have an important effect on the results.

Particle Swarm Optimization

The basis of PSO is the group behaviour of particles in a search space. In addition, members of the community can profit from the experiences of other members. An important feature of the PSO algorithm is social behaviour, in that it directs members towards the best place in the search space. Each particle in PSO is known as one solution candidate in the domain of possible solutions. The *i*th member of the population is represented by a D-dimensional vector $X_i = (x_{i1}, x_{i2}, \dots, x_{id})^T$. In addition, $V_i = (v_{i1}, v_{i2}, \dots, v_{id})^T$ is the velocity of the particle. The best previously computed position of the *i*th particle is $P_i = (p_{i1}, p_{i2}, \dots, p_{id})^T$, and the index g in the equations indicates a global guide for the particle in the population. The positions and velocities of the particles are updated based on the following equations:

$$v_{id}^{n+1} = \chi \left[iw v_{id}^n + c_1 r_1 \frac{(p_{id}^n - x_{id}^n)}{\Delta t} + c_2 r_2 \frac{(p_{gd}^n - x_{gd}^n)}{\Delta t} \right] \tag{6}$$

$$x_{id}^{n+1} = x_{id}^n + \Delta t (v_{id}^{n+1}) \tag{7}$$

Here, $d = 1, 2, \dots, D$, χ is a constriction coefficient, iw is the inertia weight, c_1 and c_2 are acceleration coefficients, r_1 and r_2 are random parameters, and Δt is the time interval.

The global best particle is considered as solution candidate, it guides the other particle toward to the other neighbours, and thus, this issue can cause that the particles trap in the local optimums. As a result, many particles will not have chance of a comprehensive search of large problem space. Thus, an effective strategy can cause the particles to get rid of local optimums and a global strategy, based on following equation, is used for improving the model's efficiency.

$$p_{gd}^n = p_{gd}^n \times (1 + \lambda \times N(0, 1)) \tag{8}$$

Where, λ : disturbance factor, $N(0, 1)$: the normal distribution. λ was tested with the different values (i.e. 0.01, 0.02, 0.05, 0.10, 0.15 and 0.2) and there is not significant in the results and thus, it was considered 0.1. the p_{gd}^n is replaced with p'_{gd} when the p'_{gd} has the better value compared to the p_{gd}^n .

SVR and IPSO

The hybrid structure of SVR and IPSO was considered through the following steps:

1. Determine the input variables for data collection and processing.
2. Consider the initial values of the SVR parameters.
3. Consider the training level of the SVR and compute the objective function (RMSE) for the input variables.
4. If the stopping criterion is satisfied, the optimal values of the coefficients are extracted for the test level. Otherwise, the algorithm goes to the next level.
5. The parameters are considered decision variables and are inserted into the IPSO. The velocities and positions of these variables are updated, and the algorithm returns to the third step. [Fig 1](#) shows the performance flowchart for SVR-IPSO.

Genetic programming

Genetic programming is a successful and widely used method in hydrological simulation. The method searches for a good relationship between the input and output variables. [Fig 2b and 2c](#) shows that the method acts based on tree structures. There are different nodes and several branches that connect them to one another. The terminal and function sets are used in the nodes. The terminal sets consist of numerical and non-numerical variables, and the function sets consist of automatic operators ($\pm \times \div$), mathematical functions (e.g., sin, cos), Boolean operators and logical expressions. The search process proceeds by generating random trees. Each tree has an objective function with a corresponding error function. A ranking method is used for the selection of trees with better objective functions. Crossover and mutation operators prepare the trees for the next iteration, as shown in [Fig 2d and 2e](#). Two trees are designed, and some branches are considered for them. In addition, the swapping of parent subtrees can be used to generate two new trees. [Fig 2d and 2f](#) show the condition of two new trees after crossover. A mutation operator is another genetic operator that exchanges nodes using a random variable or operator. [Fig 2g and 2h](#) shows the condition of trees before and after mutation. Two arithmetic operators (\pm and \times) and three mathematical functions (sin, cos and power (x^y)) are considered for the GP; a Gaussian membership function has been shown to give good results for solar radiation [[19,44](#)].

Case study

The present study deals with computing solar radiation in a Mediterranean region of Turkey. The Adana and Antakya stations are considered ([Fig 3](#)). The major climatic features of this region are inclination towards rainy winters and hot summers. The Adana station is located at latitude 37.22°N , longitude 35.40°E , and an altitude of 20 m; Antakya is located at latitude 36.22°N , longitude 35.40°E , and an altitude of 20 m. The climatic conditions of the region are affected by a winter season with high rainfall as well as hot summers. The SR distribution

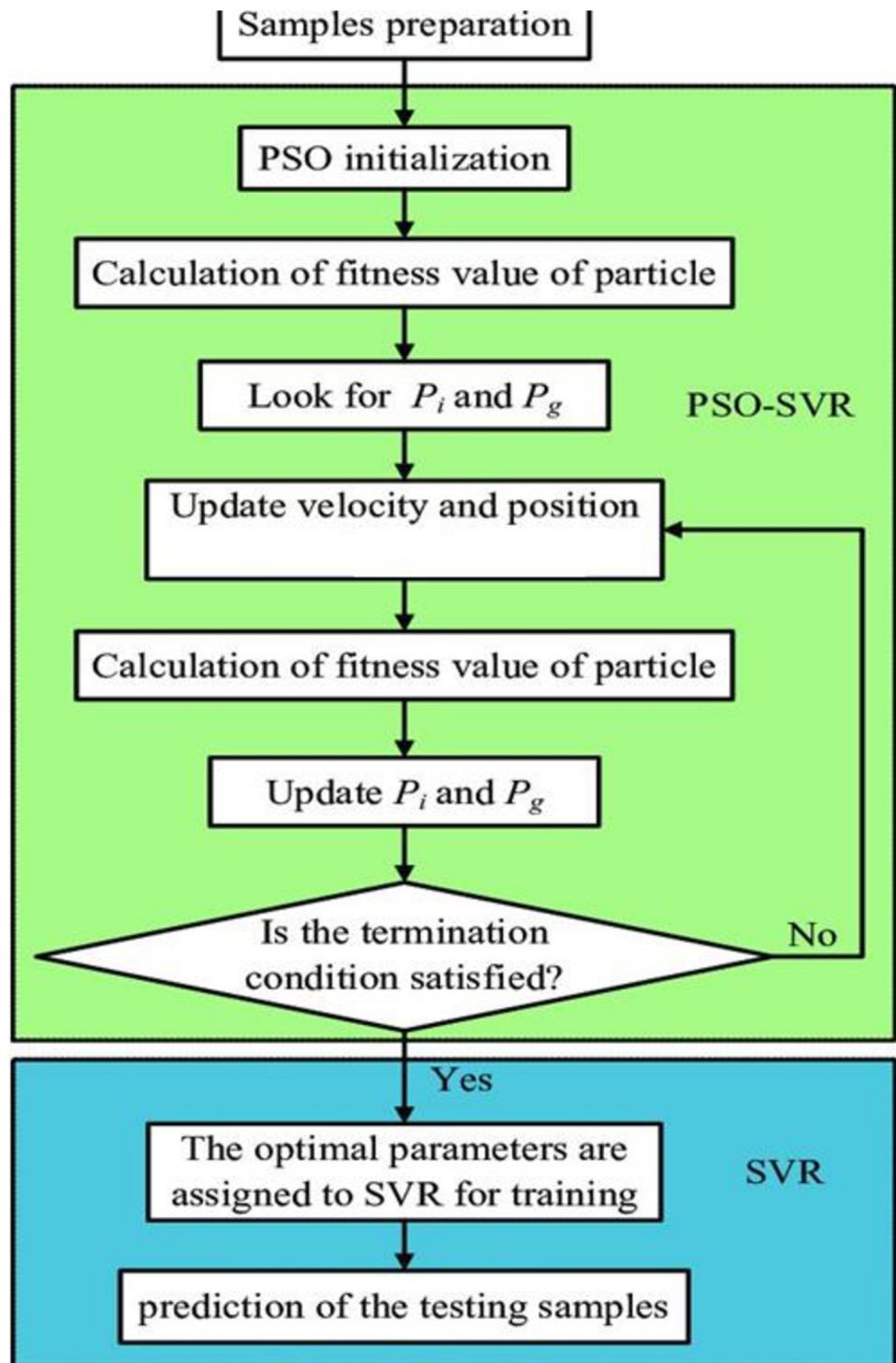


Fig 1. Structure of the hybrid SVR and IPSO.

<https://doi.org/10.1371/journal.pone.0217634.g001>

shows that the region has high solar energy potential and Turkey has high potential for solar energy because it is in the northern hemisphere. The most value for solar radiation for two stations are observed in July. The annual solar radiation for Antakya is 10.89 MJ/m²/day and it is 12.23 MJ/m²/day for Adana station.

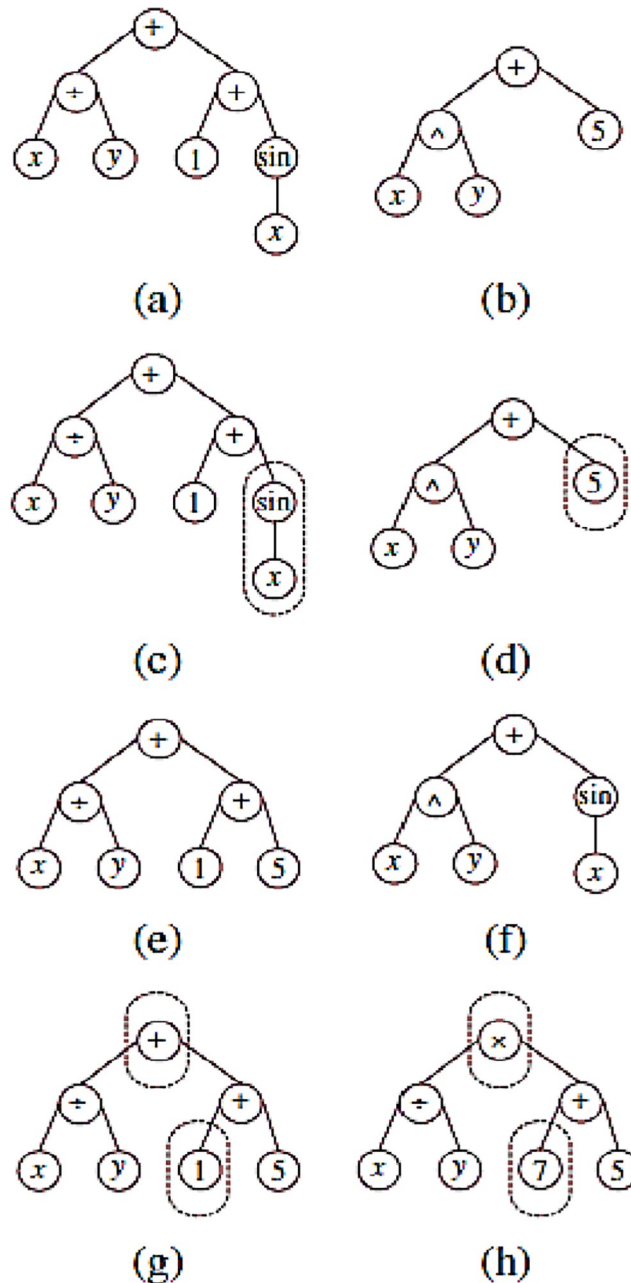


Fig 2. Genetic programming structure and search process.

<https://doi.org/10.1371/journal.pone.0217634.g002>

The data is provided by the Turkish State Meteorological services of Ministry of Agriculture and Forestry, turkey [45]. Data were collected from 1981 to 2016. Based on previous studies, 75% of the data were used for training and 25% for testing [22,27,35,37]. Yearly rainfall varies from 580 to 1300 mm. There are many stations in turkey measure solar radiation based on Siap, Muller and Fuess actiongraphs and there are 11 other stations that have pyranometers for measuring the solar radiation.

Table 1 shows the statistical data for the stations. The highest skew distribution is related to the wind speed, followed by the relative humidity and maximum temperature for both

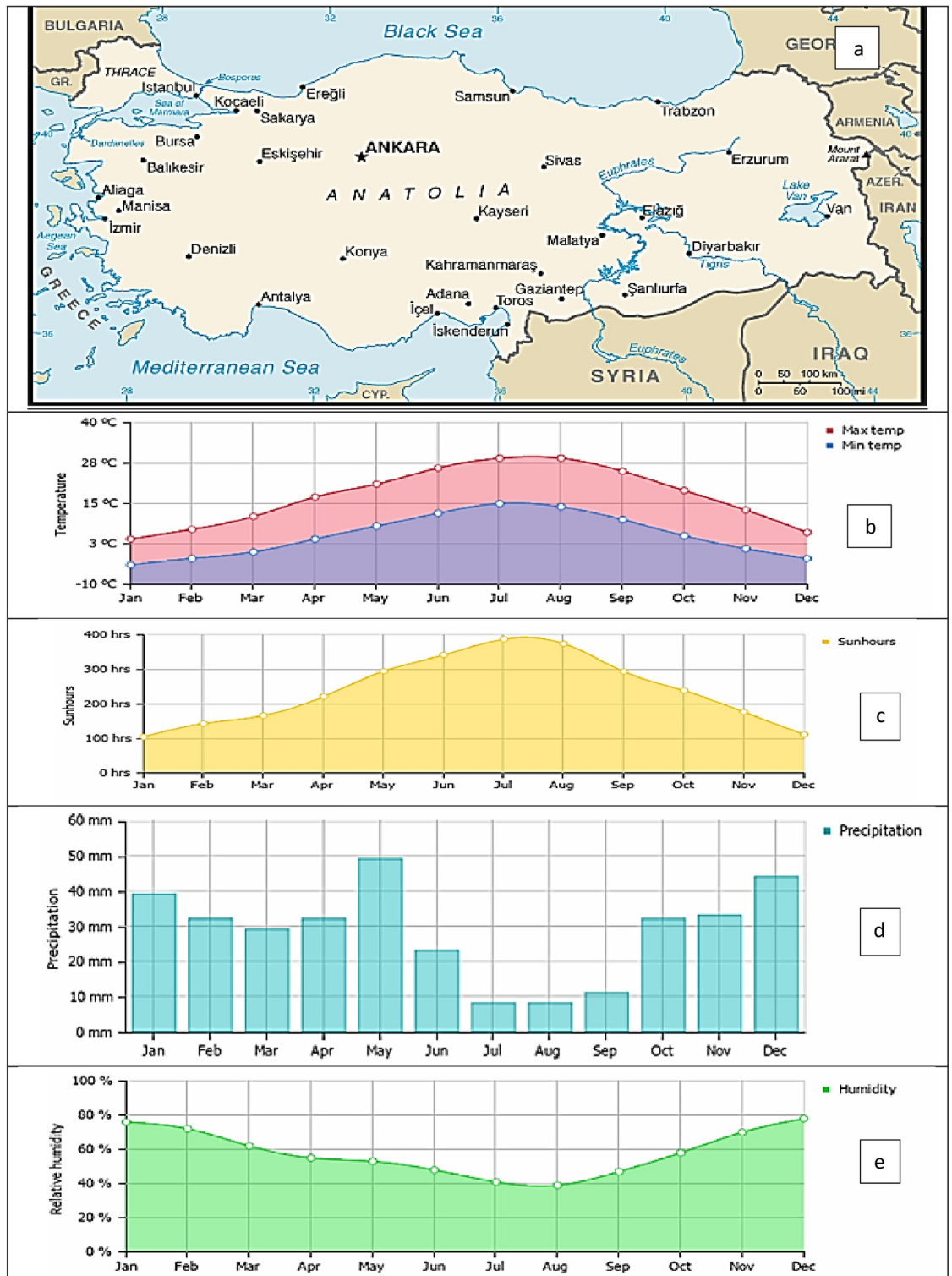


Fig 3. a) Location of the studied stations b: Temperature Average c: Sun hours Average, d: Rainfall average and e: Humidity average.

<https://doi.org/10.1371/journal.pone.0217634.g003>

Table 1. Monthly statistical parameters of each climatological data set.

Station	Data set	Unit	x_{mean}	x_{min}	x_{max}	Correlation with SR
Adana	T_{max}	°C	31.32	17	43.8	0.802
	T_{min}	°C	9.32	-6.4	23.4	0.794
	H_s	H	223	0	365	0.899
	W_s	m/s	1.34	0.10	2.3	0.277
	R_h	%	66.4	46.4	80.8	0.152
	S_r	Langley	110	36.3	195	1.000
Antakya	T_{max}	°C	30.10	14.4	42.6	0.798
	T_{min}	°C	9.38	-4.6	24.8	0.818
	H_s	H	229	29.6	384	0.923
	W_s	m/s	3.15	1.1	7.1	0.865
	R_h	%	69.8	49.4	85.1	-0.104
	S_r	Langley	98.6	27.5	179	1.000

T_{max} = maximum temperature, T_{min} = minimum temperature, H_s = sunshine duration, W_s = wind speed, R_h = relative humidity, S_r = solar radiation, x_{min} = minimum value, x_{max} = maximum value and x_{mean} = average value

<https://doi.org/10.1371/journal.pone.0217634.t001>

stations. The highest correlation coefficient is for hours of sunshine; thus, there is a high correlation between the hours of sunshine and solar radiation. The necessary data were collected at a geophysics institute for the Antakya and Adana stations from 1981 to 2016.

Model structure and performance indicators

Evolutionary algorithms, such as IPSO, have parameters whose best values can be reported based on a literature review or experimentation. We set an interval for the random parameters and evaluate the variations in the objective function for various values of the parameters. Then, the best values of the parameters (c_1 and $c_2 = 2$, $w = 0.6$ and population size for particle swarm = 40) are selected when the objective function converges to its minimum value [38,39,46]. A sensitivity analysis is considered for determining the most suitable parameters and the variation of objective function values is observed accordingly. In this regard, the least objective function value was preferred.

Several scenarios based on different inputs have been proposed based on the correlation coefficients between different input variables and solar radiation. Therefore, four scenarios based on four different input combinations were considered in this study. Keshtgar et al. [3] reported the same inputs for the two stations over the same period (1981–2016).

1. Maximum and minimum temperature
2. Maximum temperature, minimum temperature and sunshine duration
3. Maximum temperature, minimum temperature, sunshine duration and wind speed
4. Maximum temperature, minimum temperature, sunshine duration, wind speed and relative humidity

Keshtgar et al. [3] estimated solar radiation by the M5Tree model (M5T) and multivariate adaptive regression splines using the same inputs. The results were then compared with the results of previous studies. The M5T divided the search space into subspaces and developed a linear regression model for each one [3]. The M5T model divided the data into several sub-collections and then generated decision trees. Each tree had a node on the top and branches connected to other nodes.

The division rule relies on decreasing the standard deviation of the category values that reach a node as an error index for that node [3]. Piecewise linear splines were used in a multi-variate adaptive regression model (MARS) known as nonlinear and non-parametric regression. The method could model nonlinear relationships between dependent and independent variables.

To evaluate and examine the performance of the proposed prediction model, several performance indicators have been calculated. The following indices were used to evaluate the developed models:

Root mean square error (RMSE) as the objective function:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (SRm_i - SRO_i)^2}{N}} \tag{9}$$

Mean absolute error (MAE):

$$MAE = \frac{\sum_{i=1}^N |SRm_i - SRO_i|}{N} \tag{10}$$

Mean bias error:

$$MBE = \frac{1}{N} \sum_{i=1}^N \frac{SRm_i - SRO_i}{SRO_i} \tag{11}$$

Nash-Sutcliff efficiency:

$$NSE = 1 - \frac{\sum_{i=1}^N (SRm_i - SRO_i)^2}{\sum_{i=1}^N (SRO_i - SR_{mean})^2} \tag{12}$$

Here, Srm_i is the estimated solar radiation, SRO_i is the observed solar radiation, SR_{mean} is the average observed solar radiation, and N is number of data points.

Results and discussion

Antakya station

Table 2 shows the performance of different methods in the test stage for the Antakya station. The MAE index shows that the fourth input combination resulted in the best performance for SVR-PSO among all input combinations. The MAE index for SVR-PSO (4) was 43, 37, and 16% less than those for SVR-PSO (1), SVR-PSO (2) and SVR-PSO (3), respectively. The other indices supported this finding. For example, the RMSE index for SVR-PSO (4) was 52, 28 and 8.79% less than those for SVR-PSO (1), SVR-PSO (2) and SVR-PSO (3), respectively. Thus, increasing the number of input data points for SVR-PSO led to improved results for Antakya station.

SVR-IPSO was compared to M5T, MARS and SVR-PSO models for the Antakya station. The results showed that the improved PSO and SVR predicted better compared to the other models. For example, RMSE for SVR-IPSO (4) is 20%, 57%, 55% and 25% less than SVR-PSO (4), M5T (4), MARS (4) and GP (4). Also, the best input combination for the SVR-IPSO is the

Table 2. Comparison of statistical indices for different methods for estimation of solar radiation in Antakya station.

Method	Input variables	MAE (Langley)	RMSE (Langley)	MBE	NSE	Time (s)
SVR-IPSO (1)	T_{max}, T_{min}	9.01	19.10	0.141	0.72	10
SVR-PSO (1) $\kappa = 0.05, C = 61, \gamma = 0.167$	T_{max}, T_{min}	9.12	19.11	0.142	0.701	10
M5T (1) [3]	T_{max}, T_{min}	17.83	23.80	0.176	0.639	12
MARS (1) [3]	T_{max}, T_{min}	16.75	21.74	0.166	0.699	14
GP(1)	T_{max}, T_{min}	16.55	20.72	0.154	0.654	11
SVR-IPSO (2)	T_{max}, T_{min}, H_s	8.12	12.52	0.101	0.843	12
SVR-PSO (2) $\kappa = 0.05, C = 60, \gamma = 0.166$	T_{max}, T_{min}, H_s	8.25	12.54	0.103	0.842	12
M5T (2) [3]	T_{max}, T_{min}, H_s	12.09	15.58	0.106	0.873	15
MARS (2) [3]	T_{max}, T_{min}, H_s	11.17	14.10	0.112	0.845	17
GP (2)	T_{max}, T_{min}, H_s	8.29	12.59	0.107	0.844	14
SVR-IPSO (3)	$T_{max}, T_{min}, H_s, W_s$	6.01	10.11	-0.022	0.942	16
SVR-PSO (3)	$T_{max}, T_{min}, H_s, W_s$	6.12	10.12	-0.023	0.917	16
M5T (3) [3] $\kappa = 0.05, C = 55, \gamma = 0.164$	$T_{max}, T_{min}, H_s, W_s$	9.45	12.58	-0.037	0.899	18
MARS (3) [3]	$T_{max}, T_{min}, H_s, W_s$	9.88	11.56	-0.097	0.915	20
GP (3)	$T_{max}, T_{min}, H_s, W_s$	6.72	10.45	-0.022	0.824	17
SVR-IPSO (4)	$T_{max}, T_{min}, H_s, W_s, R_h$	4.10	9.01	-0.020	0.954	18
SVR-PSO (4) $\kappa = 0.04, C = 54, \gamma = 0.167$	$T_{max}, T_{min}, H_s, W_s, R_h$	5.12	9.23	-0.021	0.924	19
M5T (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h$	9.35	12.35	-0.024	0.903	22
MARS (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h$	9.26	11.12	-0.086	0.921	25
GP (4)	$T_{max}, T_{min}, H_s, W_s, R_h$	5.55	9.39	-0.025	0.922	20

<https://doi.org/10.1371/journal.pone.0217634.t002>

fourth combination. Also, the NSE coefficient for the SVR-IPSO performed better than the other methods.

SVR-PSO was compared to M5T and MARS for the Antakya station. The results showed that SVR-PSO outperformed MARS and M5T models. For example, the MBE for SVR-PSO (4), the best SVR-PSO model, was 12 and 75% less than those for M5T (4) and MARS (4), which were the best MARS and M5T models. The NSE for SVR-PSO (4) was 0.924, which was 0.3 and 2.2% greater than those for MARS (4) and M5T (4), the best MARS and M5T models. Table 2 also shows that the results of SVR-PSO are better than those of genetic programming (GP). For example, the MAE and RMSE for SVR-PSO (4) are 7.7 and 1.7%, which are less than those for GP. The best combination of inputs for GP occurs in GP (4); the results show that this GP performs better than MARS and M5T for all inputs. For example, the MAE and RMSE for GP (4) are 66 and 15%, which are less than the MARS indices. An increase in inputs to SVR-IPSO results in good performance.

Based on the comparison of the performances, as shown in the Table 2, it is noticeable that M5T (1) model has the lowest accuracy, with the first inputs. Thus, RMSE, MAE, NSE and MBE had the highest values in comparison to the SVR-PSO and MARS models. Increasing the number of inputs improved the performance of the M5T model for all indices. For example, the RMSE for M5T (4) was 48, 20 and 1.8% less than those for M5T (3), M5T (5) and M5T (1), respectively. Fig 4 shows the values of RMSE (the objective function) for different models.

The RMSE indices for the SVR-IPSO method, based on all inputs, were less than those for the MARS and M5T models (Fig 4). In addition, the lowest value of the RMSE index was exhibited by SVR-IPSO (4) and the highest value was exhibited by M5T (1). The general results for the Antakya station showed that SVR-PSO had the best performance in comparison to the other models. Fig 5 shows the R² coefficients for different SVR-IPSO models. Based on its

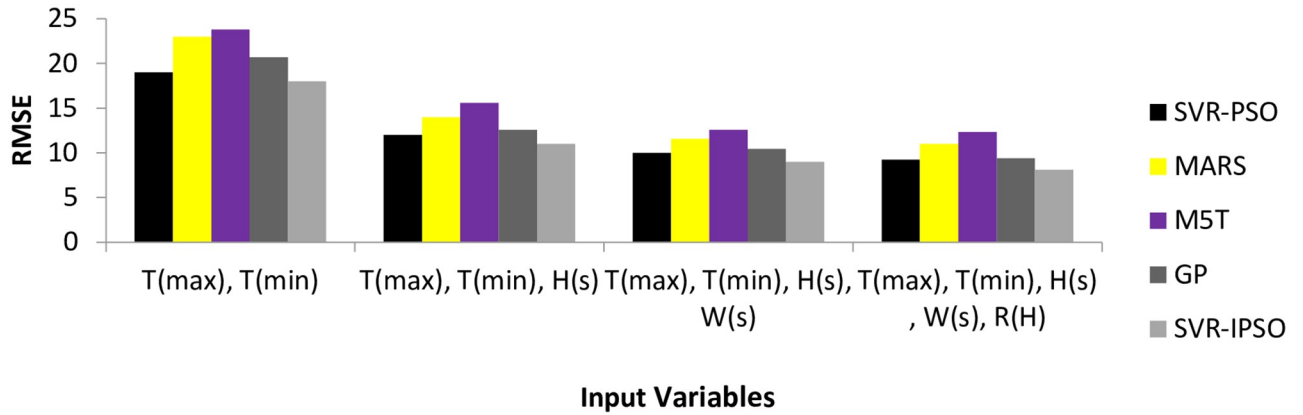


Fig 4. Computed RMSE for Antakya station.

<https://doi.org/10.1371/journal.pone.0217634.g004>

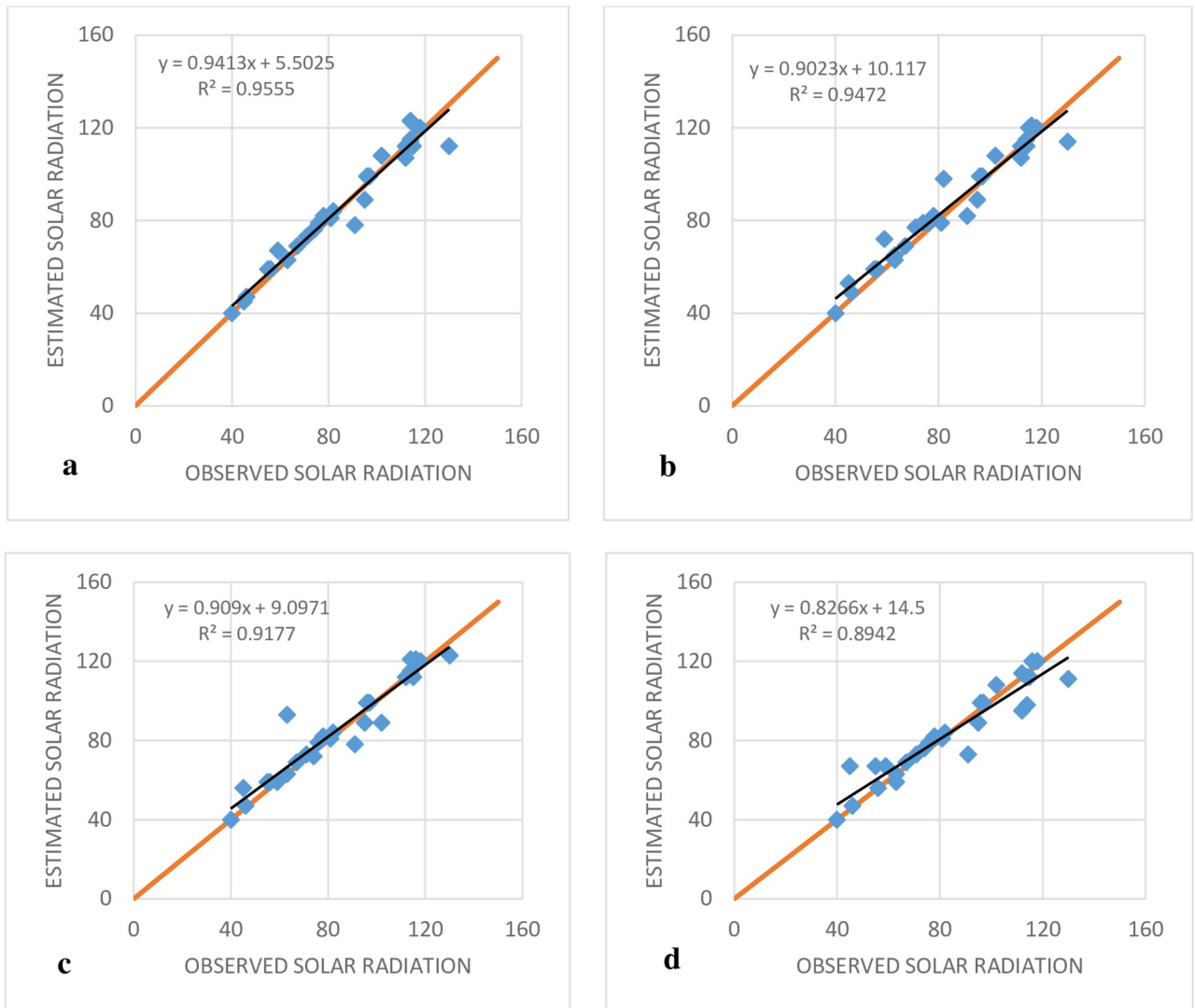


Fig 5. Observed and estimated SR of Antakya station at the test stage. (A) SVR-IPSO (4), (B) SVR-IPSO (3), (C) SVR-IPSO (2) and (D) SVR-IPSO (1).

<https://doi.org/10.1371/journal.pone.0217634.g005>

Table 3. Comparison of statistical indices for different models for estimation of solar radiation in the test stage for Adana station.

Method	Input variables	MAE (Langley)	RMSE (Langley)	MBE	NSE	Time (s)
SVR-IPSO (1)	T_{max}, T_{min}	17.01	21.12	0.042	0.701	9
SVR-PSO (1) $\kappa = 0.05, C = 69, \gamma = 0.117$	T_{max}, T_{min}	17.54	22.12	0.045	0.699	10
M5T (1) [3]	T_{max}, T_{min}	21.39	26.54	0.053	0.611	12
MARS (1) [3]	T_{max}, T_{min}	19.67	23.46	0.049	0.696	14
GP (1)	T_{max}, T_{min}	17.61	22.02	0.042	0.612	11
SVR-IPSO (2)	T_{max}, T_{min}, H_s	11.10	12.21	0.032	0.734	12
SVR-PSO (2) $\kappa = 0.05, C = 62, \gamma = 0.167$	T_{max}, T_{min}, H_s	11.23	14.23	0.035	0.717	14
M5T (2) [3]	T_{max}, T_{min}, H_s	15.65	19.29	0.061	0.795	16
MARS (2) [3]	T_{max}, T_{min}, H_s	12.88	15.49	0.039	0.868	18
GP (2)	T_{max}, T_{min}, H_s	11.36	14.35	0.037	0.719	15
SVR-IPSO (3)	$T_{max}, T_{min}, H_s, W_s$	10.12	12.01	0.032	0.761	15
SVR-PSO (3) $\kappa = 0.05, C = 60, \gamma = 0.166$	$T_{max}, T_{min}, H_s, W_s$	10.25	12.11	0.034	0.711	16
M5T (3) [3]	$T_{max}, T_{min}, H_s, W_s$	15.36	19.13	0.079	0.798	17
MARS (3) [3]	$T_{max}, T_{min}, H_s, W_s$	11.91	14.69	0.036	0.881	19
GP (3)	$T_{max}, T_{min}, H_s, W_s$	10.27	12.22	0.39	0.722	17
SVR-IPSO (4)	$T_{max}, T_{min}, H_s, W_s, H_s$	9.01	10.02	0.034	0.871	17
SVR-PSO (4) $\kappa = 0.09, C = 62, \gamma = 0.165$	$T_{max}, T_{min}, H_s, W_s, H_s$	9.87	10.32	0.044	0.709	18
M5T (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h$	15.97	20.89	0.072	0.759	20
MARS (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h$	12.79	15.40	0.021	0.869	22
GP (4)	$T_{max}, T_{min}, H_s, W_s$	9.99	10.35	0.046	0.712	24

<https://doi.org/10.1371/journal.pone.0217634.t003>

higher value for the R^2 coefficient (Fig 5), SVR-IPSO (4) performed better than the other SVR-IPSO models. In general, the kind of inputs and the number of inputs affect the accuracy of the results. The consideration of parameters with high correlations, such as maximum and minimum temperature, can be good choices for inputs. In addition, one of the most important factors for the evaluation of the different methods is the computational time. Table 2 shows that SVR-IPSO could obtain the desired outputs with less computational time. For example, the computational time for SVR-IPSO (4) is, 5.2% 13, 24 and 5% less than those for SVR-PSO (4), M5T (4), MARS (4) and GP (4), respectively. The SVR-IPSO can improve the computational time because the convergence velocity is increased by this method.

Adana station

Table 3 shows the performance of different models for the Adana station. All indices indicated that SVR-IPSO has a remarkable advantage over the MARS, SVR-PSO and M5T models. For example, the MAE for SVR-IPSO (4) was 47%, 19 and 11% less than those for SVR-IPSO (1), SVR-IPSO (2) and SVR-IPSO (3), respectively. Increasing the number of inputs for the Adana station improved the results, as for the Antakya station. For example, the RMSE for SVRI-PSO (4) was 10.02, approximately 52, 19 and 17% lower than those for SVR-IPSO (1), SVR-IPSO (2) and SVR-PISO (3), respectively. Other indices also indicated the superiority of SVR-IPSO (4) over other SVR-IPSO models.

The SVR-IPSO, SVR-PSO, MARS and M5T models were compared for the Adana station. Multiple indices indicated the superiority of the SVR-IPSO model over the SVR-PSO MARS and M5T models. For example, the M5T (3) model had the lowest MAE (15.36) in comparison to the other M5T models, and the MARS (3) model exhibited the best performance of the MAE index (11.91) in comparison to the other MARS models. However, the MAE values for

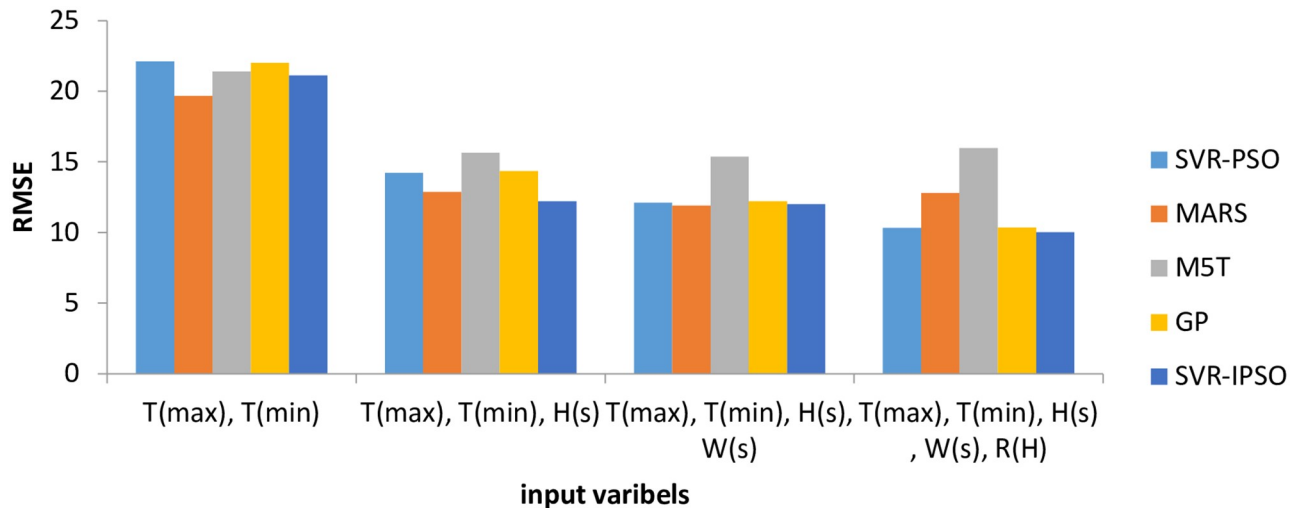


Fig 6. Computed RMSE for Adana station.

<https://doi.org/10.1371/journal.pone.0217634.g006>

these two models were 36 and 18% more than SVR-PSO (4). Other indices also indicated the superiority of SVR-IPSO (4) over the MARS and M5T models. However, a comparison of the results for GP and SVR-IPSO shows better performance for GP. For example, the MAE and RMSE values for SVR-IPSO (4) are 1.20 and 3.1% less than those for GP (4). In addition, the GP model has the best performance in comparison to the MARS and M5T models. Increasing the inputs for the different models shows that models with 5 inputs have the best performance.

The various indices did not unanimously indicate the superiority of a specific model as the best among the M5T models. For example, based on RMSE and MAE, M5T (3) had the best performance in comparison to the other M5T models. However, the MBE showed better performance for M5T (1) in comparison to the other M5T models. Fig 6 shows the performance of different models based on the RMSE (objective function) for the Adana station. The SVR-IPSO models for all input combinations had the best performance in comparison to the SVR-PSO, MARS and M5T models. The lowest value for RMSE was exhibited by SVR-IPSO (4) and the worst performance was exhibited by M5T (4) (Fig 6). Fig 7 shows the R^2 coefficients for the SVR-IPSO models. The results showed that SVR-IPSO (4) performed better than the other SVR-IPSO models based on its higher value for the R^2 coefficient (Fig 7). In addition, the computational times in Table 3 show that SVR-IPSO performed better than the other methods.

Periodic model for computing solar radiation at the Antakya station

A periodic forecast entails the addition of the months as an input. Many studies have shown that periodic prediction can improve the results of precise forecasts of hydrological and climatic data. All simulation models in Table 4 are based on the addition of the month as an input. The results in Table 4 tangibly show the superiority of the SVR-IPSO model over the other models. For example, the MAE index indicated that SVR-IPSO (4), SVR-IPSO (3), M5T (4) and MARS (4) were the best in comparison to the other SVR-IPSO, SVR-PSO, M5T and MARS models. The MAE for SVR-IPSO (4) was 4.01; this was 26, 6.7 and 1.3% better than those of the M5T (4), SVR-PSO and MARS (4) models and showed the superiority of SVR-IPSO (4) over the other models. The NSE index also indicated the superiority of SVR-IPSO over the other models. For example, the highest value for the NSE (0.911) was

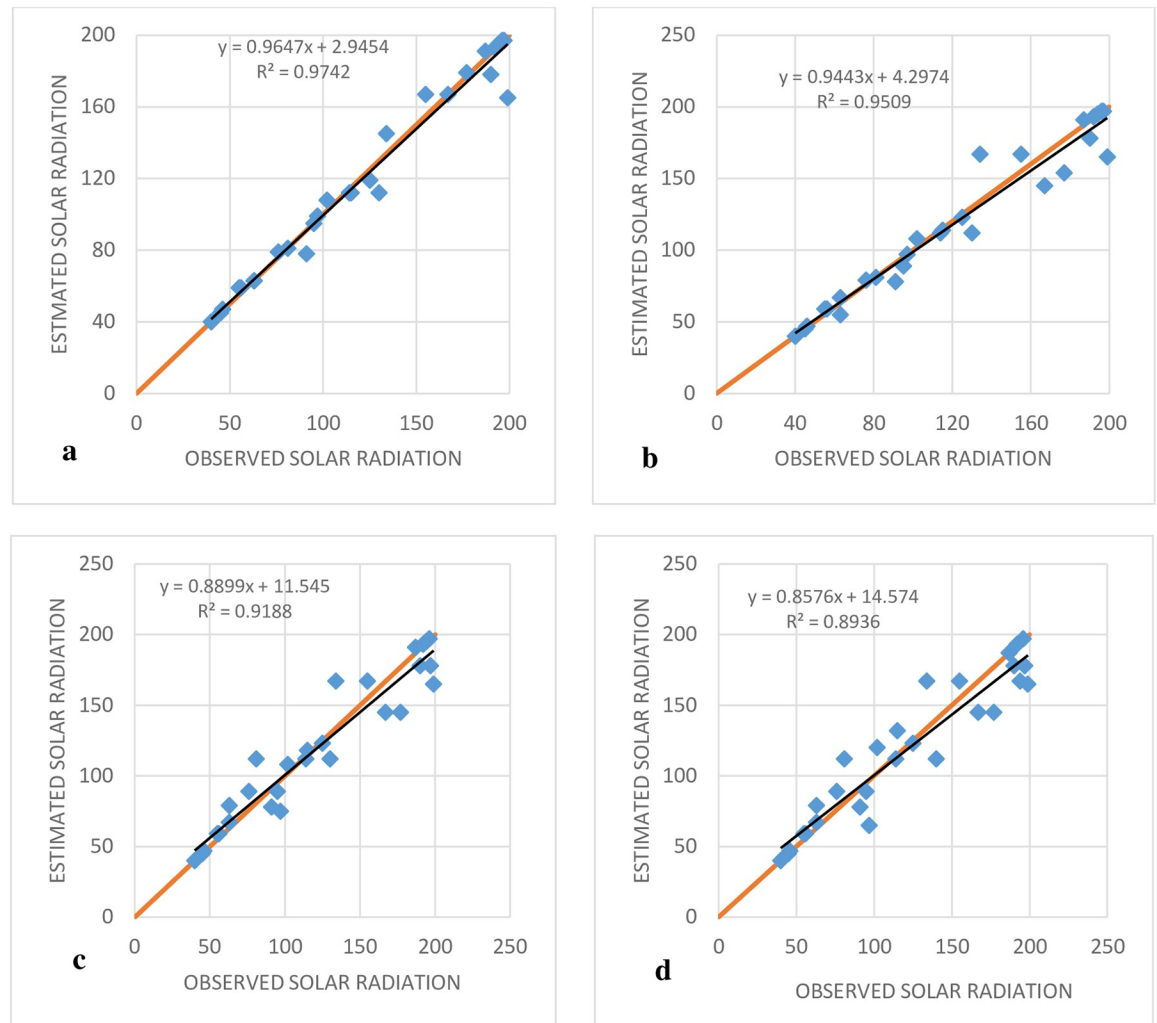


Fig 7. Observed and estimated SR of Adana station at the test stage. (A) SVR-IPSO (4), (B) SVR-IPSO (3), (C) SVR-IPSO (2) and (D) SVR-IPSO (1).

<https://doi.org/10.1371/journal.pone.0217634.g007>

exhibited by SVR-PSO (4), while the best values of this index for the M5T and MARS models were 0.963 and 0.981, respectively. In addition, a comparison of these results with the GP model shows that SVR-IPSO has the best performance; the error indices of SVR-IPSO are less than those of GP. For example, the MAE values for SVR-IPSO (4), SVR-IPSO (3), SVR-IPSO (2) and SVR-IPSO (1) are 7.1, 0.92, 2.8% and 1.3% less than those for GP (1), GP (2), GP (3), and GP (4), respectively. Increasing the number of inputs has a good effect on the simulation results; the fourth combination of each methods exhibits the best results. This shows that all the parameters, namely the maximum temperature, minimum temperature, wind speed, relative humidity and month, affect the results.

The trend of the results showed that adding inputs improved the performance of SVR-IPSO for all indices; this was not true for the M5T and MARS models. Another important point was a comparison of the forecasts with/without adding the month. For example, the RMSE for SVR-IPOS (4) based on the periodic model and Table 4 was 5.02, while it was 9.01 based on Table 2 without adding the month as the input. Other indices for the other models indicated that the value of the error indices could be significantly reduced by adding the month as an

Table 4. Comparison of statistical indices for different methods for estimation of solar radiation at the test stage based on periodic models for Antakya station.

Method	Input variables	MAE	RMSE	MBE	NSE	Time (s)
SVR-IPSO (1)	$T_{max}, T_{min}, month$	6.40	7.20	0.065	0.961	19
SVR-PSO (1) $\kappa = 0.05, C = 60, \gamma = 0.165$	$T_{max}, T_{min}, month$	6.45	7.23	0.067	0.959	20
M5T (1) [3]	$T_{max}, T_{min}, month$	10.06	12.1	0.091	0.954	22
MARS (1) [3]	$T_{max}, T_{min}, month$	6.83	8.48	0.077	0.906	24
GP (1)	$T_{max}, T_{min}, month$	6.49	7.29	0.069	0.961	21
SVR-IPSO (2)	$T_{max}, T_{min}, H_s, month$	6.05	7.10	0.060	0.955	20
SVR-PSO (2) $\kappa = 0.04, C = 60, \gamma = 0.166$	$T_{max}, T_{min}, H_s, month$	6.20	7.12	0.063	0.962	22
M5T (2) [3]	$T_{max}, T_{min}, H_s, month$	11.8	11.8	0.090	0.912	24
MARS (2) [3]	$T_{max}, T_{min}, H_s, month$	8.18	8.18	0.085	0.957	26
GP(2)	$max, T_{min}, H_s, month$	6.23	7.91	0.089	0.967	23
SVR-IPSO (3)	$T_{max}, T_{min}, H_s, W_s, month$	5.28	5.40	0.001	0.986	24
SVR-PSO (3) $\kappa = 0.05, C = 54, \gamma = 0.166$	$T_{max}, T_{min}, H_s, W_s, month$	5.34	5.45	0.002	0.981	25
M5T (3) [3]	$T_{max}, T_{min}, H_s, W_s, month$	7.63	7.63	0.003	0.963	27
MARS (3) [3]	$T_{max}, T_{min}, H_s, W_s, month$	5.83	5.83	0.029	0.978	29
GP (3)	$T_{max}, T_{min}, H_s, W_s, month$	5.39	5.55	0.002	0.979	26
SVR-IPSO (4)	$T_{max}, T_{min}, H_s, W_s, R_h, month$	4.01	5.02	0.002	0.991	26
SVR-PSO (4) $\kappa = 0.05, C = 62, \gamma = 0.167$	$T_{max}, T_{min}, H_s, W_s, R_h, month$	4.30	5.12	0.002	0.985	28
M5T (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h, month$	5.87	7.68	0.005	0.962	30
MARS (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h, month$	4.36	5.42	-0.019	0.981	32
GP (4)	$T_{max}, T_{min}, H_s, W_s, R_h, month$	4.32	5.18	0.004	0.989	29

<https://doi.org/10.1371/journal.pone.0217634.t004>

input. For example, the NSE index showed that, assuming periodic simulation, the M5T (3) model had the best performance in comparison to the other M5T models; the value of the index for this model was 0.963, while the value of the index for the best M5T model without adding the month was 0.903 based on Table 2. Thus, prediction can improve the results for all models. Fig 8 shows the R^2 coefficients for the SVR-IPSO models. Fig 8 indicates that the periodic SVR-IPSO models perform better than the non-periodic models shown in Fig 5. In addition, the trend in computational time for the different methods shows that increasing the number of inputs increases the computational time and that SVR-IPSO has the best performance in comparison to the other methods.

Periodic model for computing solar radiation at the Adana station

Table 5 shows the performance of different models with periodic prediction for the Adana station. The MAE for SVR-IPSO (4) was 4.12, which was 30, 37 and 57% less than those for SVR-IPSO (3), SVR-IPSO (2) and SVR-IPSO (1), respectively. Based on MAE and Table 6, M5T (3), SVR-PSO (4) and MARS (2) performed better than the other MARS and M5T models. The MAE for SVR-IPSO (4) was 37, 47 and 42% less than those for the, SVR-PSO (4), M5T (3) and MARS (2) models, indicating the superiority of the SVR-IPSO model. A reduction in the error indices can be achieved by increasing the number of inputs for the SVR-PSO model, as in the previous sections. The NSE for SVR-IPSO (4) was 0.992, while the best values of this coefficient for MARS and M5T were 0.942 and 0.942; this indicates the superiority of the SVR-IPSO model over the other models. In addition, the results show that the GP model performs better than MARS and M5T, but SVR-IPSO performs better than GP. For example, the RMSE for SVR-IPSO (4), SVR-IPSO (3), SVR-IPSO (2) and SVR-IPSO (1) were 12, 14, 3.2 and 2.6% less than those for GP (4), GP (3), GP (2) and GP (1), respectively.

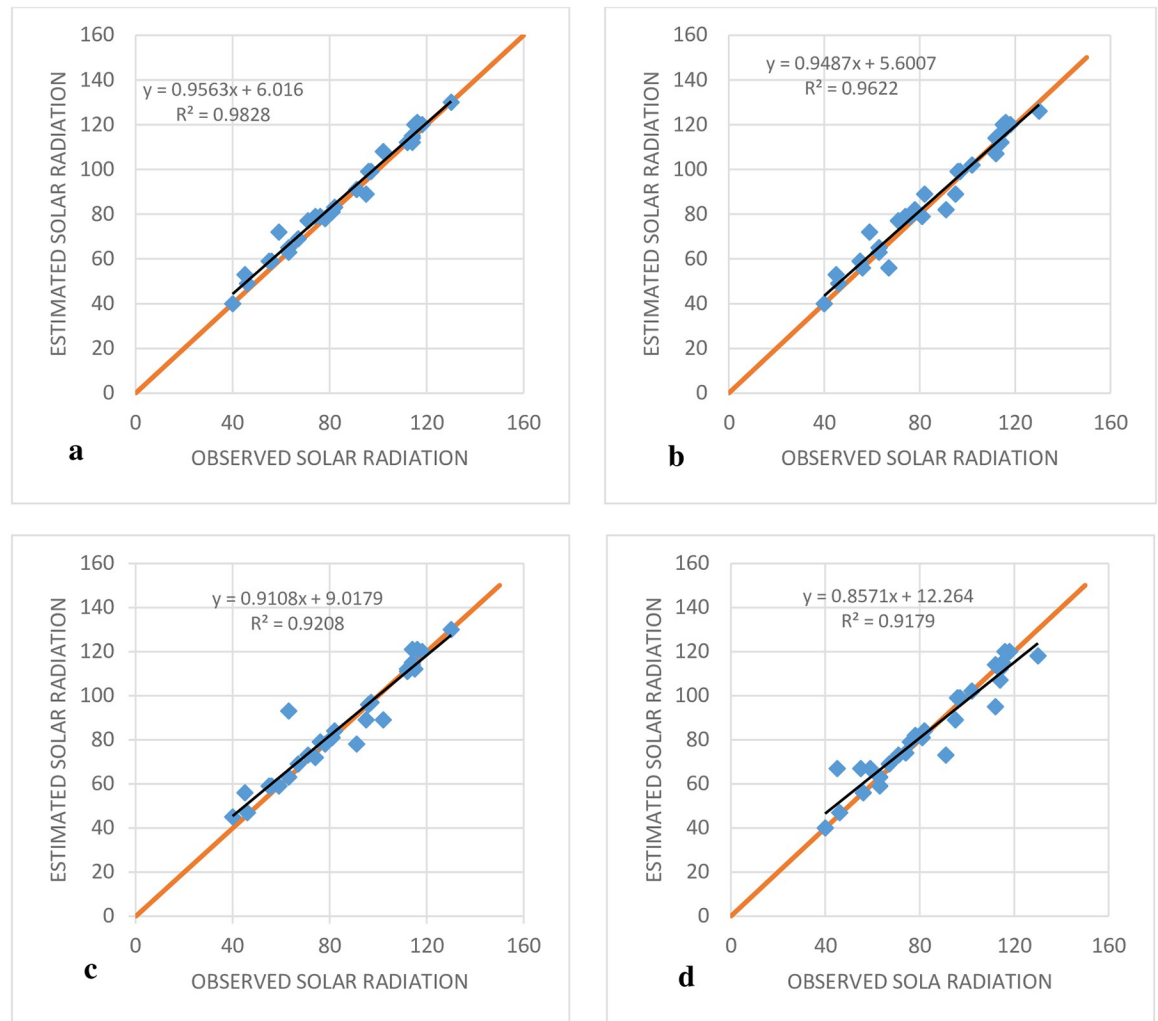


Fig 8. Observed and estimated SR of Antakya station at test stage based on the periodic model. (A) SVR-IPSO (4), (B) SVR-IPSO (3), (C) SVR-IPSO (2) and (D) SVR-IPSO (1).

<https://doi.org/10.1371/journal.pone.0217634.g008>

In addition, periodicity improved the results. For example, the RMSE for SVR-PSO (4) without periodicity, according to Table 3, was 10.02, while it was 7.64 for the periodic SVR-IPSO (4) model. Other indices for other models indicated similar effects. For example, M5T (2) exhibited the lowest value for RMSE in comparison to the other M5T models (based on Table 5), which was 9.66. On the other hand, the value of RMSE for the best model without periodicity (M5T), based on Table 4, was 19.13. Thus, it is obvious that periodicity improved the results. Furthermore, the results show that the addition of the month as an input resulted in the best outcome for GP; the index error was reduced in comparison to the simulation results for which the month was not an input. In addition, SVR-IPSO exhibited decreased computational time.

Fig 9 shows the R^2 coefficients for different SVR-IPSO models. SVR-IPSO (4) has the highest value; furthermore, the values of R^2 for the models in Fig 7 was better than those for the models in Fig 9 (without periodic prediction).

The results indicated that the SVR-IPSO act better than other models. There is more challenge for solar radiation estimation such as daily solar radiation or solar radiation for different

Table 5. Comparison of statistical indices for different models for estimation of solar radiation at the test stage based on periodic models for Adana station.

Method	Input variables	MAE	RMSE	MBE	NSE	Time (s)
SVR-IPSO (1)	$T_{max}, T_{min}, month$	9.52	10.82	-0.012	0.878	24
SVR-PSO (1) $\kappa = 0.05, C = 60, \gamma = 0.166$	$T_{max}, T_{min}, month$	9.54	10.9	-0.014	0.933	25
M5T (1) [3]	$T_{max}, T_{min}, month$	10.2	13.2	-0.016	0.903	27
MARS (1) [3]	$T_{max}, T_{min}, month$	9.17	11.2	-0.031	0.930	29
GP (1)	$T_{max}, T_{min}, month$	9.59	11.12	-0.015	0.932	26
SVR-IPSO (2)	$T_{max}, T_{min}, H_s, month$	6.56	8.84	-0.090	0.962	25
SVR-PSO (2) $\kappa = 0.05, C = 62, \gamma = 0.165$	$T_{max}, T_{min}, H_s, month$	7.01	9.12	-0.010	0.950	27
M5T (2) [3]	$T_{max}, T_{min}, H_s, month$	9.26	9.66	-0.011	0.924	28
MARS (2) [3]	$T_{max}, T_{min}, H_s, month$	7.10	11.72	-0.026	0.949	30
GP (2)	$T_{max}, T_{min}, H_s, month$	7.05	9.14	0.011-	0949	29
SVR-IPSO (3)	$T_{max}, T_{min}, H_s, W_s, month$	5.89	8.02	-0.009	0.962	27
SVR-PSO (3) $\kappa = 0.05, C = 60, \gamma = 0.166$	$T_{max}, T_{min}, H_s, W_s, month$	6.94	9.01	-0.010	0.951	28
M5T (3) [3]	$T_{max}, T_{min}, H_s, W_s, month$	7.79	10.3	-0.015	0.942	31
MARS (3) [3]	$T_{max}, T_{min}, H_s, W_s, month$	7.68	10.3	-0.037	0.942	33
GP (3)	$T_{max}, T_{min}, H_s, W_s, month$	6.99	9.23	-0.010	0.949	29
SVR-IPSO (4)	$T_{max}, T_{min}, H_s, W_s, R_h, month$	4.12	7.64	-0.008	0.992	29
SVR-PSO (4) $\kappa = 0.05, C = 60, \gamma = 0.166$	$T_{max}, T_{min}, H_s, W_s, R_h, month$	6.45	8.67	-0.009	0.970	30
M5T (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h, month$	8.46	10.40	-0.032	0.940	32
MARS (4) [3]	$T_{max}, T_{min}, H_s, W_s, R_h, month$	7.84	10.1	-0.018	0.944	33
GP (4)	$T_{max}, T_{min}, H_s, W_s, R_h, month$	6.49	8.69	-0.009	0.971	31

<https://doi.org/10.1371/journal.pone.0217634.t005>

climates. Thus, the combination of Tables 2 and 3 were used to estimate daily solar radiation for better evaluation of new model. Also, previous studies suggested that the SVR with the other optimization algorithms. Genetic algorithm and firefly algorithms were used to determine the SVR parameters. The firefly algorithm (FFA) acts based on social behavior in the fireflies. The attractiveness of each firefly is based on its light intensity. The light intensity for each firefly is considered as objective function and the social behavior for the fireflies is used for the

Table 6. Comparison of statistical indices for different methods for estimation of daily solar radiation in Konay station.

Method	Input variables	MAE%	RMSE%	MBE%	NSE	Time (s)
SVR-IPSO (1)	T_{max}, T_{min}	12	14	8	0.84	15
SVR-PSO (1)	T_{max}, T_{min}	14	16	9	0.80	18
SVR-FFA (1)	T_{max}, T_{min}	17	20	10	0.79	22
SVR-GA (1)	T_{max}, T_{min}	22	23	12	0.77	25
SVR-IPSO (2)	T_{max}, T_{min}, H_s	10	12	7	0.89	17
SVR-PSO (2)	T_{max}, T_{min}, H_s	12	14	8	0.83	20
SVR-FFA (2)	T_{max}, T_{min}, H_s	15	18	10	0.82	25
SVR-GA (2)	T_{max}, T_{min}, H_s	16	21	11	0.81	27
SVR-IPSO (3)	$T_{max}, T_{min}, H_s, W_s$	9	10	5	0.91	19
SVR-PSO (3)	$T_{max}, T_{min}, H_s, W_s$	11	12	7	0.85	21
SVR-FFA (3)	$T_{max}, T_{min}, H_s, W_s$	12	16	8	0.83	26
SVR-GA (3)	$T_{max}, T_{min}, H_s, W_s$	14	18	9	0.82	28
SVR-IPSO (4)	$T_{max}, T_{min}, H_s, W_s, R_h$	7	9	3	0.95	20
SVR-PSO (4)	$T_{max}, T_{min}, H_s, W_s, R_h$	10	10	5	0.90	22
SVR-FFA (4)	$T_{max}, T_{min}, H_s, W_s, R_h$	11	14	6	0.88	27
SVR-GA (4)	$T_{max}, T_{min}, H_s, W_s, R_h$	12	16	7	0.86	30

<https://doi.org/10.1371/journal.pone.0217634.t006>

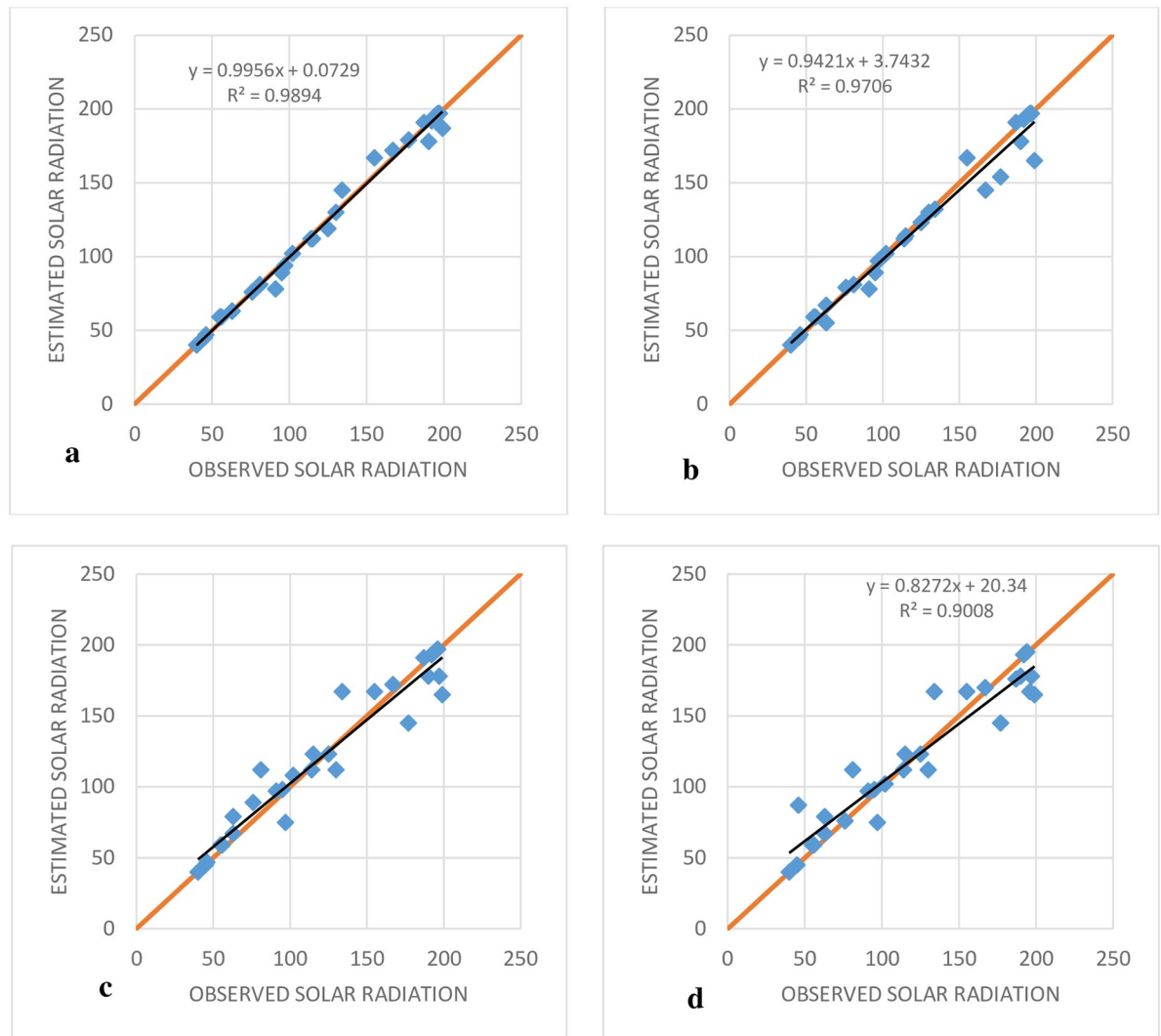


Fig 9. Observed and estimated SR of Adana station based on periodic prediction at the test stage. (A) SVR-IPSO (4), (B) SVR-IPSO (3), (C) SVR-IPSO (2) and (D) SVR-IPSO (1).

<https://doi.org/10.1371/journal.pone.0217634.g009>

optimization algorithm. More details can be seen in the [33]. Also, the genetic algorithm acts based on natural selection and chromosomes. The mutation and crossover and selection operators were used to improve the solutions and more details can be seen in the [47]. Thus, the SVR-IPSO, SVR-PSO, SVR-FFA and SVR-GA are used for the estimation of daily solar radiation for different station with different climatic condition. Konya station is selected as station for evaluating of validation of SVR-IPSO. This station Konya is located at the 37°58'N, 32°32'E in a semi dry climate. Also, the average temperature for this station is 23.3°C. There are 2860 sunshine duration per year averagely. The estimation of daily solar estimation for period 1981 to 2016 is considered as more challenging for the SVR-IPSO model.

The RMSE, MAE and MBE index are computed on percentage. Table 6 shows the daily solar radiation for Konya Station 1981 to 2016. The results indicated that the SVR-IPSO has the less RMSE and MAE value and thus, the fourth combination for SVR-IPSO is better compared to the SVR-GA, SVR-PSO and SVR-FFA. Also, the worst performance is also observed and recorded for the SVR-GA. The first input for the SVR-GA has the worst performance

Table 7. The enhancement rate for best model (SVR-IPSO (4)) comparing with the previous study [3].

Method	Enhancement Rate	
	Adana	Antakya
M5T (1) [3]	42.12%	58.51%
MARS (1) [3]	31.78%	40.80%
M5T (2) [3]	20.91%	57.45%
MARS (2) [3]	34.81%	38.63%
M5T (3) [3]	25.82%	34.20%
MARS (3) [3]	25.82%	13.89%
M5T (4) [3]	26.53%	34.63%
MARS (4) [3]	24.35%	7.38%

<https://doi.org/10.1371/journal.pone.0217634.t007>

compared to SVR-GA (2), SVR-GA (3) and SVR-GA (4). The NSE value for SVR-IPSO (4) has the most value compared to the other model and least value for the NSE is related to the SVR-GA. However, the results indicate that the SVR-IPSO can show a good performance for different climates for daily or monthly solar estimation.

It is true that the accuracy of the any forecasting model will be affected positively or negatively, when using future input data (in this study metrological data). Therefore, there is a need to keep updating the forecasting model by feeding it with the new patterns of the future data. However, it should be noted that as long as the future data is similar with the used data pattern, the model's performance will not change much. Based on the RMSE for the best model (SVR-IPSO (4)), the enhancement rate has been investigated by comparing with different studies in the literature. Table 7 summarize the enhancement that has been accomplished by the proposed model for two stations (Adana and Antakya). The proposed model shows a significant enhancement in the performance for both stations. It can be determined that the significant of performance is (42.12%–20.91%) for Adana and (58.51%–7.38%) for Antakya.

Conclusion

The present study introduced a new prediction model for SR. The model is essentially based on an improved SVR integrated with PSO; PSO determines the optimum values of the unknown SVR parameters. The proposed model was applied to two stations from Turkey for evaluation against the previously developed SVR-PSO, MARS, GP and M5T models, which have been applied to the same stations. Based on the proposed performance indicators, increasing the number of inputs improved the results of the SVR-IPSO model. In addition, the application of SVR-IPSO to the Antakya station showed the superiority of SVR-IPSO over the other models. The proposed SVR-IPSO models for the two stations achieved better performance than the MARS, GP, SVR-PSO and M5T models for different input scenarios. Furthermore, an additional input variable representing the month of the year resulted in improvements over previous input scenarios. In conclusion, the proposed SVR integrated with IPSO (SVR-IPSO) can be considered an effective tool for solar radiation prediction that could help decision-makers create efficient plans for renewable energy production. A few important variables were lacking in the selected stations and hence could not be examined in this study. Also, the SVR-IPSO, was validated for Konya station and the results were compared with the SVR-GA, SVR-PSO and SVR-FFA. The results showed that the SVR-IPSO model has best performance comparing with all the presented models.

The proposed model could be improved by adding other input variables that directly influence solar radiation. Future studies should consider additional input variables that might

improve the accuracy of predicting solar radiation. In addition, an integration of the SVR and advanced meta-heuristic optimization algorithms should be investigated as it might improve the forecasting accuracy for SR. The sun radiation has a direct effect on the climate condition therefore, it is essential to consider further evaluation for the proposed model in different climatic zone.

Acknowledgments

The authors would like to thank so much the data supplier. We also thank all reviewers and the editor in chief for their insightful comments that have improved the quality of the final manuscript.

Author Contributions

Formal analysis: Md Shabbir Hossain, Haitham Abdulmohsin Afan, Ahmed El-Shafie.

Methodology: Hamidreza Ghazvinian, Sayed-Farhad Mousavi, Mohammad Ehteram, Ali Najah Ahmed.

Resources: Hojat Karami.

Supervision: Ahmed El-Shafie.

Validation: Chow Ming Fai, Faizah Che Ros.

Writing – original draft: Saeed Farzin.

Writing – review & editing: Huzaifa Bin Hashim, Vijay P. Singh, Haitham Abdulmohsin Afan, Sai Hin Lai.

References

1. Sengupta M, Xie Y, Lopez A, Habte A, Maclaurin G, Shelby J. The National Solar Radiation Data Base (NSRDB). *Renew Sustain Energy Rev*. Pergamon; 2018; 89: 51–60. <https://doi.org/10.1016/J.RSER.2018.03.003>
2. Khosravi A, Koury RNN, Machado L, Pabon JJG. Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms. *J Clean Prod*. Elsevier; 2018; 176: 63–75. <https://doi.org/10.1016/J.JCLEPRO.2017.12.065>
3. Keshtegar B, Mert C, Kisi O. Comparison of four heuristic regression techniques in solar radiation modeling: Kriging method vs RSM, MARS and M5 model tree. *Renew Sustain Energy Rev*. Pergamon; 2018; 81: 330–341. <https://doi.org/10.1016/J.RSER.2017.07.054>
4. Qazi A, Fayaz H, Wadi A, Raj RG, Rahim NA, Khan WA. The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. *J Clean Prod*. Elsevier; 2015; 104: 1–12. <https://doi.org/10.1016/J.JCLEPRO.2015.04.041>
5. Qin W, Wang L, Lin A, Zhang M, Xia X, Hu B, et al. Comparison of deterministic and data-driven models for solar radiation estimation in China. *Renew Sustain Energy Rev*. Pergamon; 2018; 81: 579–594. <https://doi.org/10.1016/J.RSER.2017.08.037>
6. Vakili M, Sabbagh-Yazdi SR, Khosrojerdi S, Kalhor K. Evaluating the effect of particulate matter pollution on estimation of daily global solar radiation using artificial neural network modeling based on meteorological data. *J Clean Prod*. Elsevier; 2017; 141: 1275–1285. <https://doi.org/10.1016/J.JCLEPRO.2016.09.145>
7. Loghmari I, Timoumi Y, Messadi A. Performance comparison of two global solar radiation models for spatial interpolation purposes. *Renew Sustain Energy Rev*. Pergamon; 2018; 82: 837–844. <https://doi.org/10.1016/J.RSER.2017.09.092>
8. Bou-Rabee M, Sulaiman SA, Saleh MS, Marafi S. Using artificial neural networks to estimate solar radiation in Kuwait. *Renew Sustain Energy Rev*. Pergamon; 2017; 72: 434–438. <https://doi.org/10.1016/J.RSER.2017.01.013>

9. Renno C, Petite F, Gatto A. ANN model for predicting the direct normal irradiance and the global radiation for a solar application to a residential building. *J Clean Prod.* Elsevier; 2016; 135: 1298–1316. <https://doi.org/10.1016/J.JCLEPRO.2016.07.049>
10. Zhang S, Kang L, Conservation BZ-, Water C, Soil A&, 2017 undefined. Parameter estimation of nonlinear Muskingum model with variable exponent using adaptive genetic algorithm. *books.google.com*.
11. Ozgoren M, Bilgili M, Sahin B. Estimation of global solar radiation using ANN over Turkey. *Expert Syst Appl.* Pergamon; 2012; 39: 5043–5051. <https://doi.org/10.1016/J.ESWA.2011.11.036>
12. Landeras G, López JJ, Kisi O, Shiri J. Comparison of Gene Expression Programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain). *Energy Convers Manag.* Pergamon; 2012; 62: 1–13. <https://doi.org/10.1016/J.ENCONMAN.2012.03.025>
13. Mohandes MA. Modeling global solar radiation using Particle Swarm Optimization (PSO). *Sol Energy.* Pergamon; 2012; 86: 3137–3145. <https://doi.org/10.1016/J.SOLENER.2012.08.005>
14. Khatib T, Mohamed A, Sopian K. A review of solar energy modeling techniques. *Renew Sustain Energy Rev.* Pergamon; 2012; 16: 2864–2869. <https://doi.org/10.1016/J.RSER.2012.01.064>
15. Benmouiza K, Cheknane A. Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. *Energy Convers Manag.* Pergamon; 2013; 75: 561–569. <https://doi.org/10.1016/J.ENCONMAN.2013.07.003>
16. Bhardwaj S, Sharma V, Srivastava S, Sastry OS, Bandyopadhyay B, Chandel SS, et al. Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model. *Sol Energy.* Pergamon; 2013; 93: 43–54. <https://doi.org/10.1016/J.SOLENER.2013.03.020>
17. Yadav AK, Chandel SS. Solar radiation prediction using Artificial Neural Network techniques: A review. *Renew Sustain Energy Rev.* Pergamon; 2014; 33: 772–781. <https://doi.org/10.1016/J.RSER.2013.08.055>
18. Amrouche B, Le Pivert X. Artificial neural network based daily local forecasting for global solar radiation. *Appl Energy.* 2014; 130: 333–341. <https://doi.org/10.1016/j.apenergy.2014.05.055>
19. Mohammadi K, Shamshirband S, Tong CW, Arif M, Petković D, Ch S. A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation. *Energy Convers Manag.* Pergamon; 2015; 92: 162–171. <https://doi.org/10.1016/J.ENCONMAN.2014.12.050>
20. Olatomiwa L, Mekhilef S, Shamshirband S, Mohammadi K, Petković D, Sudheer C. A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Sol Energy.* Pergamon; 2015; 115: 632–644. <https://doi.org/10.1016/J.SOLENER.2015.03.015>
21. Premalatha N, Valan Arasu A. Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. *J Appl Res Technol.* Elsevier; 2016; 14: 206–214. <https://doi.org/10.1016/J.JART.2016.05.001>
22. Shamshirband S, Mohammadi K, Khorasanizadeh H, Yee PL, Lee M, Petković D, et al. Estimating the diffuse solar radiation using a coupled support vector machine–wavelet transform model. *Renew Sustain Energy Rev.* Pergamon; 2016; 56: 428–435. <https://doi.org/10.1016/J.RSER.2015.11.055>
23. Quej VH, Almorox J, Arnaldo JA, Saito L. ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. *J Atmos Solar-Terrestrial Phys.* Pergamon; 2017; 155: 62–70. <https://doi.org/10.1016/J.JASTP.2017.02.002>
24. Ibrahim IA, Khatib T. A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. *Energy Convers Manag.* Pergamon; 2017; 138: 413–425. <https://doi.org/10.1016/J.ENCONMAN.2017.02.006>
25. Meenal R, Selvakumar AI. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. *Renew Energy.* Pergamon; 2018; 121: 324–343. <https://doi.org/10.1016/J.RENENE.2017.12.005>
26. Kumar N, Sinha UK, Sharma SP, Nayak YK. International journal of renewable energy research IJRER. [Internet]. International Journal of Renewable Energy Research (IJRER). Gazi Univ., Fac. of Technology, Dep. of Electrical et Electronics Eng; 2017. <https://www.ijrer.org/ijrer/index.php/ijrer/article/view/5988>
27. Elsevier Science (Firm) C, Notton G, Kalogirou S, Nivet M-L, Paoli C, Motte F, et al. Renewable energy. [Internet]. Renewable Energy. Pergamon Press; 1991. https://econpapers.repec.org/article/eeerene/v_3a105_3ay_3a2017_3ai_3ac_3ap_3a569-582.htm
28. Wang L, Kisi O, Zounemat-Kermani M, Zhu Z, Gong W, Niu Z, et al. Prediction of solar radiation in China using different adaptive neuro-fuzzy methods and M5 model tree. *Int J Climatol.* John Wiley & Sons, Ltd; 2017; 37: 1141–1155. <https://doi.org/10.1002/joc.4762>
29. Alfadda A, Rahman S, Pipattanasomporn M. Solar irradiance forecast using aerosols measurements: A data driven approach. *Sol Energy.* Pergamon; 2018; 170: 924–939. <https://doi.org/10.1016/J.SOLENER.2018.05.089>

30. Wang L, Kisi O, Zounemat-Kermani M, Salazar GA, Zhu Z, Gong W. Solar radiation prediction using different techniques: model evaluation and comparison. *Renew Sustain Energy Rev*. Elsevier BV; 2016; 61: 384–397. <https://doi.org/10.1016/j.rser.2016.04.024>
31. Rohani A, Taki M, Abdollahpour M. A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). *Renew Energy*. Elsevier BV; 2018; 115: 411–422. <https://doi.org/10.1016/j.renene.2017.08.061>
32. Shang C, Wei P. Enhanced support vector regression based forecast engine to predict solar power output. *Renew Energy*. Pergamon; 2018; 127: 269–283. <https://doi.org/10.1016/J.RENENE.2018.04.067>
33. Moazenzadeh R, Mohammadi B, Shamshirband S, Chau K. Coupling a firefly algorithm with support vector regression to predict evaporation in northern Iran. *Eng Appl Comput Fluid Mech*. Taylor & Francis; 2018; 12: 584–597. <https://doi.org/10.1080/19942060.2018.1482476>
34. Voyant C, Notton G, Kalogirou S, Nivet M-L, Paoli C, Motte F, et al. Machine learning methods for solar radiation forecasting: A review. *Renew Energy*. Pergamon; 2017; 105: 569–582. <https://doi.org/10.1016/J.RENENE.2016.12.095>
35. Belaid S, Mellit A. Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate. *Energy Convers Manag*. 2016; 118: 105–118. <https://doi.org/10.1016/j.enconman.2016.03.082>
36. Zhou P, Huang J, Pontius RG, Hong H. New insight into the correlations between land use and water quality in a coastal watershed of China. *Science of the Total Environment*. 2016. pp. 591–600. <https://doi.org/10.1016/j.scitotenv.2015.11.063> PMID: 26615482
37. Deo RC, Wen X, Qi F. A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Appl Energy*. Elsevier; 2016; 168: 568–593. <https://doi.org/10.1016/J.APENERGY.2016.01.130>
38. Awange JL, Paláncz B, Lewis RH, Völgyesi L. Particle Swarm Optimization. *Mathematical Geosciences*. Cham: Springer International Publishing; 2018. pp. 167–184.
39. Nouri M, Bekrar A, Jemai A, Niar S, Ammari AC. An effective and distributed particle swarm optimization algorithm for flexible job-shop scheduling problem. *J Intell Manuf*. Springer US; 2018; 29: 603–615. <https://doi.org/10.1007/s10845-015-1039-3>
40. Liu H-H, Chang L-C, Li C-W, Yang C-H. Particle Swarm Optimization-Based Support Vector Regression for Tourist Arrivals Forecasting. *Comput Intell Neurosci*. Hindawi; 2018; 2018: 1–13. <https://doi.org/10.1155/2018/6076475> PMID: 30327666
41. Lou I, Xie Z, Ung WK, Mok KM. Integrating Support Vector Regression with Particle Swarm Optimization for numerical modeling for algal blooms of freshwater. *Appl Math Model*. Elsevier; 2015; 39: 5907–5916. <https://doi.org/10.1016/J.APM.2015.04.001>
42. Vapnik VN, Vapnik V. *Statistical learning theory*. Wiley New York; 1998.
43. Huang C-L, Wang C-J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst Appl*. Pergamon; 2006; 31: 231–240. <https://doi.org/10.1016/J.ESWA.2005.09.024>
44. Ramezani F, Nikoo M, Nikoo M. Artificial neural network weights optimization based on social-based algorithm to realize sediment over the river. *Soft Comput*. 2014; 19: 375–387. <https://doi.org/10.1007/s00500-014-1258-0>
45. Turkish State Meteorological Service. Meteoroloji Genel Müdürlüğü [Internet]. 2018. <https://mevbis.mgm.gov.tr/mevbis/ui/index.html#Workspace>
46. Zhou Y, Guo S, Chang FJ, Liu P, Chen AB. Methodology that improves water utilization and hydropower generation without increasing flood risk in mega cascade reservoirs. *Energy*. 2018; 143: 785–796. <https://doi.org/10.1016/j.energy.2017.11.035>
47. Tang X, Wang L, Cheng J, Chen J, Sheng VS. Forecasting Model Based on Information-Granulated GA-SVR and ARIMA for Producer Price Index. *Comput Mater Contin*. 2019; 58: 463–491.