Theoretical and Applied Climatology (2022) 147:1725–1742
https://doi.org/10.1007/s00704-021-03913-5

ORIGINAL PAPER

Estimation of solar radiation by joint application of phase space reconstruction and a hybrid neural network model

Mahsa H. Kashani[1] · Samed Inyurt[2] · Mohammad Reza Golabi[3] · Mohammad AmirRahmani[4] · Shahab S. Band[5]

Received: 12 April 2021 / Accepted: 23 December 2021 / Published online: 22 January 2022

Abstract

Estimation of solar radiation can play a key role in environmental management as well as other fields of energy, agriculture, and hydrological and ecological modeling. In some areas, there are not enough solar radiation data due to a lack of pyranometer or its breakdown from time to time. Hence, having an estimation set at hand to estimate solar radiation based on other climatic variables is crucial. In order to develop an estimation tool, two models are applied simultaneously as a new hybrid model for estimation of monthly global solar radiation for three regions in Iran as case studies of this research work: (1): an artificial neural network (ANN) optimized with Harris hawk's optimization (HHO) algorithm (ANNHHO) and (2) phase space reconstruction (PSR) integrated with the ANNHHO hybrid model (PSR-ANNHHO). Monthly meteorological data of minimum temperature (T_{min}), maximum temperature (T_{max}), mean temperature (T_{mean}), sunshine hours (SH), wind speed (U_2), and relative humidity (RH) of 37 years (1985–2018) from three regions in Iran with different climate types were employed for training and testing the developed models. To select appropriate input variables for the models, a relief algorithm was applied. The performance of the new hybrid models is compared with the stand-alone ANN model. The obtained results revealed that although all the intelligent models perform satisfactorily, the hybrid PSR-ANNHHO model outperforms the hybrid ANNHHO and stand-alone ANN models in all regions. The hybrid ANN-HHO model follows the PSR-ANNHHO model as the second most accurate model.

1 Introduction

Measuring solar radiation because of its high importance in meteorology, hydrology, and climatology is an important task for researchers and engineers. Because of its measuring costs, conventional empirical models and then soft computing and intelligent models have been developed as methods for estimation of global solar radiation. The empirical model estimates solar radiation using simple explicit formula. Because of the high uncertainty of the empirical models, artificial intelligence (AI) models are alternative tools that provide solutions to real-world and complex problems (Yacef et al. 2014; Mohanty et al. 2016). In recent years, AI and data-driven models such as multilayer perceptron neural network (MLP), Gaussian process regression (GPR), radial basis function network (RBFN), hybrid networks, genetic

✉ Mahsa H. Kashani
m.hkashani@uma.ac.ir

Samed Inyurt
samed_inyurt@hotmail.com

Mohammad Reza Golabi
hamidgolabi65@gmail.com

Mohammad AmirRahmani
m.amirrahmani@gmail.com

Shahab S. Band
shamshirbands@yuntech.edu.tw

[1] Department of Water Engineering, Water Management Research Center, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran

[2] Department of Geomatics Engineering, Faculty of Engineering and Architecture, Tokat Gaziosmanpasa University, Tokat, Turkey

[3] Faculty of Water Sciences Engineering, University of Shahid Chamran Ahvaz, Ahvaz, Iran

[4] Institute of the Environment, University of Tabriz, Tabriz, Iran

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

programming (GP), support vector machine (SVM), gene expression program (GEP), recurrent neural network (RNN), Elman network, and adaptive neuro-fuzzy inference system (ANFIS) have been employed for solar radiation modeling (Kalogirou 2001; Mellit 2008; Moghaddamnia et al. 2009; Rahimikhoob 2010; Yadav and Chandel 2014; Mohanty et al. 2016; Rohani et al. 2018; Nourani et al. 2019). Also, the AI models have a high ability in coping with missing data rather than empirical techniques which are affected by the missing data and the presence of outliers.

Artificial neural network (ANN) is an effective AI model for function approximation, prediction, simulating complex, and nonlinear problems. Several studies have successfully estimated the global solar radiation using ANNs (Al-Alawi and AL-Hinai 1998; Rehman and Mohandes 2008; Benghanem et al. 2009; Moghaddamnia et al. 2009; Rahimikhoob 2010; Behrang et al. 2010; Moreno et al. 2011; Koca et al. 2011; Yacef et al. 2014; Mohanty (2014); Sharifi et al. 2016; Basaran et al. 2019; Nourani et al. 2019). Rehman and Mohandes (2008) applied ANN to estimate daily solar radiation in Abha city, Saudi Arabia, using relative humidity and air temperature variables as inputs of the model. Moghaddamnia et al. (2009) applied ANN and other intelligent models along with a linear regression model for estimating solar radiation at the Brue catchment, UK, using different meteorological data. Behrang et al. (2010) applied RBF and MLP neural networks to estimate daily global solar radiation in Iran (Dezful city). Moreno et al. (2011) estimated global solar radiation using ANN and compared it with Bristow–Campbell and kernel ridge regression over Spain. Nourani et al. (2019) estimated solar radiation in four regions across Iraq using ANN and compared its performance with empirical and ensemble models. Mohanty (2014) estimated monthly mean global solar radiation applying the ANN model and compared the estimations with those of other AI models and Angstrom–Prescott empirical model in Bhubaneswar, India.

In order to overcome deficiencies of AI methods such as uncertainty, hybrid models have been proposed and developed by different researchers. In this case, ANNs or other standalone intelligent models are combined with different evolutionary algorithms such as invasive weed optimization (IWO), particle swarm optimization (PSO), and genetic algorithms (GA) as optimization tools. Due to the different performances of evolutionary optimization algorithms to the same problem, researchers are continuously investigating the capability of the algorithms and hybrid models on different problems. Mohammadi et al. (2015) applied two different hybrid models for the estimation of monthly mean global solar radiation in different sunny regions of Iran. The first model included the support vector machine and firefly algorithm (FFA-SVM), and the second model was based on the SVM and wavelet transform. Ibrahim and Khatib (2017) developed a new hybrid model for estimation of hourly solar radiation using random forests techniques and firefly algorithm (FFA-RFs) in the Klang Valley region, Malaysia. The integrated FFA-RFs model was compared with RF, ANN, and hybrid FFA-ANN models. Halabi et al. (2018) used the ANFIS model combined with different evolutionary algorithms such as GA, PSO, and differential evolution DE for forecasting monthly global solar radiation in Malaysia. ANFIS-PSO achieved the highest performance followed by ANFIS-GA. Jadidi et al. (2018) proposed a new hybrid model including GA for tuning ANN parameters for one hour ahead for forecasting solar radiation in Carolina, USA. Guermoui et al. (2021) developed a new hybrid artificial bee colony algorithm and least squares-support vector machine (ABC-LS-SVM) for multi-hour ahead forecasting of global solar radiation in Ghardaia, south of Algeria. In all these studies, the results achieved revealed that the hybrid models provide a more accurate performance compared to the stand-alone models.

Chaos theory is a new and powerful tool for predicting complex phenomena. It represents a dynamic system including a set of influential variables located in a phase space diagram. Each point on the phase space diagram describes the system's behavior at any time. Takens (1981) presented a method for reconstructing phase space from any observed time series. Phase space reconstruction (PSR) can be applied for extracting the hidden information and the characteristics of dynamic systems from their time series. The application of PSR in the prediction of hydrological time series has increased recently. In numerous studies such as Lovejoy and Mandelbrot (1985), Farmer and Sidorowich (1987), Casdagli (1989), Rodriguez-Iturbe et al. (1989), Olsson et al. (1993), Porporato and Ridolfi (1996, 1997), Koutsoyiannis and Pachakis (1996), Sivakumar et al. (1998, 1999), Wang and Gan (1998), Pasternack (1999), Sivakumar (2000, 2001), Elshorbagy et al. (2002), Gaume et al. (2006), Ng et al. (2007), Shang et al. (2009), Sun et al. 2010, Dhanya and Kumar (2010, 2011), Baydaroğlu and Koçak, 2014, Golder et al. (2014), and Ghorbani et al. (2018), the PSR has been successfully applied to predict dynamic and nonlinear hydrological time series such as streamflow, evaporation, and rainfall. Fathima et al. (2019) recently analyzed the time series of solar irradiance measurements using chaos theory in Singapore.

Despite a great number of studies that have estimated global solar radiation in different regions all over the world, exploring accurate models for more valid estimations is highly required due to the importance of using accurate solar radiation data. According to the above literature, there are insufficient researches on applications of phase space reconstruction (PSR) in solar radiation estimation. In this study, two newly hybrid methods are proposed to estimate solar radiation: (1) Harris hawk's

optimization (HHO) is employed to optimize a neural network (ANNHHO), and (2) phase space reconstruction (PSR) is employed to optimize the hybrid ANNHHO model (PSR-ANNHHO) for estimation of the solar radiation, and their capability is compared with the stand-alone ANN model. To the best knowledge of the authors, barely study has employed hybrid ANNHHO and PSR-ANNHHO models for the estimation of solar radiation. Furthermore, the current study utilizes a non-linear tool, i.e., relief algorithm to identify the appropriate combination of input variables for simulation of monthly solar radiation. Recently, some studies have been conducted using the Relief algorithm in numerous fields. Almaraashi (2018) estimated the daily solar radiation using neural networks predictor in different regions in Saudi Arabia. To this end, the capability of four different algorithms including the Laplacian score algorithm, Monte Carlo uninformative variable elimination algorithm, random-frog algorithm, and relief algorithm was evaluated in order to select the most appropriate variables. The results proved the importance of applying feature selection methods in order to obtain precise estimations of the solar radiation data. To the best of the authors' knowledge, the relief algorithm applied in this study has barely been applied to this problem before in Iran.

The rest of the article is organized as follows: in Section 2, the study area, the proposed hybrid models, and the performance criteria used in this study are described. In Section 3, the results obtained from all the models in the three regions are presented and discussed. Section 4 explains the final conclusions.

2 Materials and methods

2.1 Study area and data

Iran is located in the western part of Asia with an area of 1,648,195 km^2. It has six neighboring countries, namely, Iraq, Turkey, Armenia, Azerbaijan, Afghanistan, and Pakistan. According to Fig. 1, the latitudes of Iran are between 24 and 40 N, and the longitudes are between 44 and 64 E. The different climate conditions ranging from arid and semi-arid to subtropical (humid) are observed in the northern forests of Iran.

In this study, a total of 407 monthly meteorological data (1985–2018) from three stations (Ardabil, Rasht, and Zabol) with different climate conditions were collected from the Climatology Organization including minimum temperature

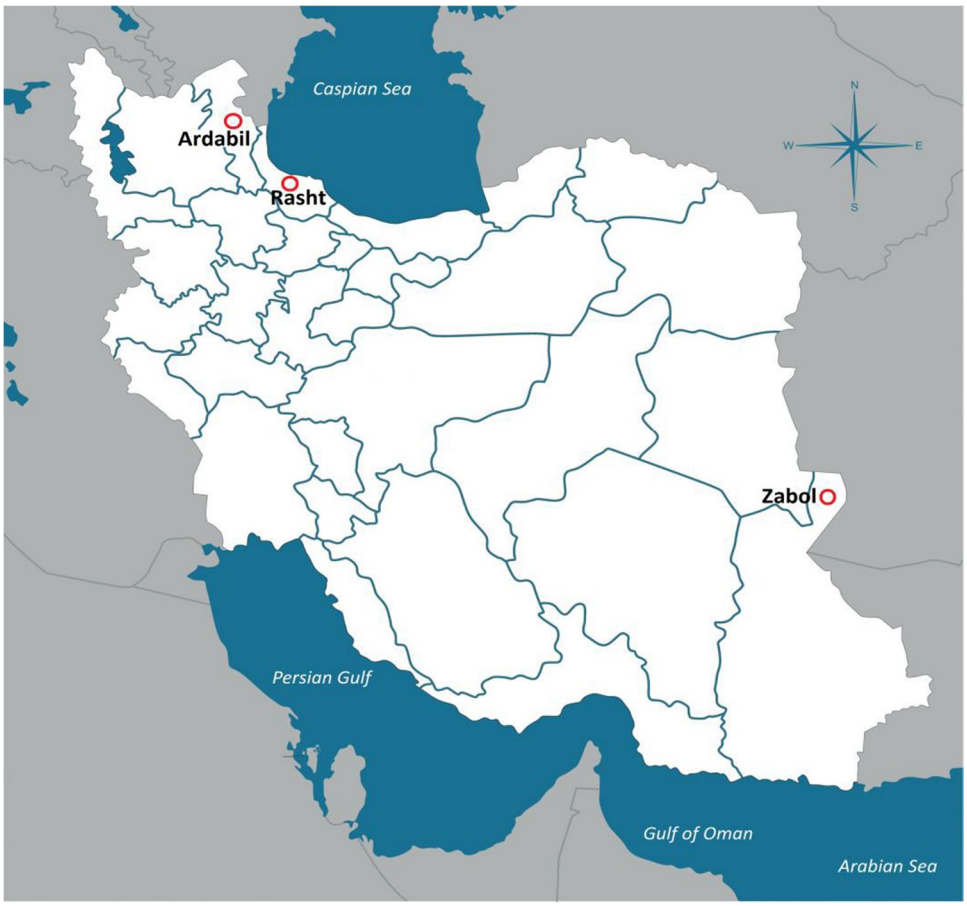

Fig. 1 Location of study regions in Iran

(T_{min} in°C), maximum temperature (T_{max} in°C), mean temperature (T_{mean} in°C), solar radiation (R_s in MJ/m^2/month), wind speed (U_2 at 2 m above the earth surface, in m/s), sun hours per day (S.H in hours), and relative humidity (RH) (in %). Based on the Domarton index, Ardabil's climate is semi-arid, Rasht's climate is humid, and Zabol's climate is arid. The data were divided into 70% (285 samples) for training and 30% (121 samples) for testing the models. Table 1 presents the geographical features of the study stations and the statistics of the climatology data used in this study. As shown in Table 1, all the regions have different parameter values. So, it seems that for different regions, different variables will be important for modeling solar radiation and so the best model for each area may be a bit different from the other locations.

2.2 Relief algorithm

Relief algorithm is a feature selection algorithm applied to reduce the size of the problem and was first proposed and developed by Kira and Rendell (1992). The positive characteristics of this algorithm are its simplicity, the need for a small amount of training data, usability for continuous data, and solvability with low-order polynomial functions. In a dataset with the number of N samples (observational data) and the number of P attributes belonging to two different classes, each attribute must be in the range (0 and 1). The algorithm is repeated n times, each time using a different weight vector starting from zero. In each iteration, the algorithm selects the X attribute vector belonging to a random sample and the attribute vectors of the closest sample to the X sample in the desired class by the Euclidean distance function. Finally, after n repeated times, each element of the weight vector is divided by n. Then, a related vector is obtained and one attribute is selected if the value of the related vector of the attribute exceeds the defined threshold (Kira and Rendell 1992). As mentioned, one of the most important features and highlights of the relief algorithm is its suitability for datasets with a small number of training samples. Therefore, in this study, due to the relatively short statistical length of the data, this method is applied to determine the appropriate parameters for solar radiation estimation. For more information about the relief algorithm, see Urbanowicz et al. (2018).

2.3 Artificial neural network (ANN)

The multi-layer perceptron (MLP) ANN is one of the most versatile algorithms and has proven to be able to simulate highly nonlinear and complex relationships between a set

Table 1 Geographical and statistical characteristics of the study regions and data

Region	Location	Coordinates	Parameter	Unit	Training				Testing			
					Min	Max	Mean	Standard deviation	Min	Max	Mean	Standard deviation
Ardabil	Latitude	38 22' N	Tmin	°C	−19.9	15.4	2.67	7.14	-9.9	15.6	4.34	6.30
			Tmax	°C	−4.32	30.64	15.46	8.31	0.74	29.31	16.71	7.90
	Longitude	48 32' E	Tmean	°C	−10.95	21.34	9.06	7.61	-4.15	22.45	10.53	7.05
			RH	%	55	90	74.29	6.95	48.95	87.61	71.20	7.52
			U2	m/s	1.3	9.5	3.99	1.20	2.05	7.44	3.74	0.92
	Altitude	1335.2 m	S.H	Hour	2	11.58	6.78	2.11	3.19	11.95	6.99	2.12
			Rs	MJ/m^2/month	5.9	27	15.51	6.05	6.8	27	15.62	6.17
Rasht	Latitude	37 32' N	Tmin	°C	-0.76	22.86	12.01	6.92	-0.13	24.81	12.66	6.81
			Tmax	°C	6.95	33.62	20.85	7.30	8.85	33.61	21.19	7.30
			Tmean	°C	3.99	27.20	15.92	7.04	3.94	28.65	16.29	7.11
	Longitude	49 62' E	RH	%	72.34	93.19	83.45	4.73	71.27	92.12	82.91	5.27
			U2	m/s	0.13	2.47	1.10	0.40	0.57	2.56	1.46	0.32
	Altitude	−8.6 m	S.H	Hour	0.19	9.99	4.67	1.99	1.03	10.46	4.71	2.28
			Rs	MJ/m^2/month	4.9	24.3	13.15	5.40	5.7	24.6	13.13	5.83
Zabol	Latitude	31 09' N	Tmin	°C	−0.75	30.77	15.06	9.19	0.31	30.91	15.42	9.75
			Tmax	°C	8.08	44.53	29.92	9.46	15.19	44.57	30.94	9.46
			Tmean	°C	6.13	37.72	22.27	9.37	8.38	37.84	22.541	9.53
	Longitude	61 54' N	RH	%	13	78.38	37.26	13.60	10.31	64.83	29.74	13.47
			U2	m/s	0.91	13.84	5.89	3.04	2.03	13.81	5.57	2.45
	Altitude	489.2 m	S.H	Hour	3.29	12.02	8.67	1.70	5.80	12.03	8.97	1.58
			Rs	MJ/m^2/month	8.7	28	19.11	5.26	10.4	28	19.36	5.35

of input variables (predictors) and the output data (predictand) (McClelland and Rumelhart 1989). The network is trained using some training data and adjusting its parameters (weights and biases) with the assistance of the Levenberg–Marquardt backpropagation algorithm. The efficiency of the MLP network is determined by the network architecture, which involves a set of processing units (neurons), a specific topology of weighted links (synapses) that connects the neurons, and the learning function that updates the connection weights. The inputs of the MLP are weighted by the weight matrix, summated with a bias term, and provided to the transfer or activation function. The transfer function makes a non-linear decision boundary via linear combinations of the weighted inputs and then uses a threshold to transform the net inputs from the entire neuronal unit into an output (Haykin 2009). The backpropagation algorithm incrementally adjusts the network parameters (weight and bias) to minimize the mean square error (MSE) criterion of the network. At every epoch, the learning rate parameter determines the quantum of progressions in adjusting the synaptic weights and biases. Smaller values of the learning rate parameter result in longer training time but create stability steering to minimize errors. In this study, the Levenberg–Marquardt algorithm is used to train the three-layered MLP with a varying number of neurons for the simulation of monthly global solar radiation. The sigmoid and linear functions are used as the transfer functions of the hidden and output layers, respectively. The optimal number of hidden nodes is determined using a trial-and-error procedure.

2.4 Harris hawks' optimizer (HHO)

HHO algorithm, which was developed by Heidari et al. (2019), is an optimization algorithm resembling the Harris' hawks' behavior for solving optimization problems. In HHO, hawks hunt in three main stages including exploration, transferring, and exploiting. Waiting, seeking, and discovering possible prey is included in the first phase which is named exploration. Equation (1) determines the position of hawks:

$$Y(iter+1)=\begin{cases} Y_{rand}(iter)-r_1|Y_{rand}(iter)-2r_2Y(iter) & ifq\geq 0.5 \\ (Y_{prey(iter)-Y_m(iter)})-r_3\left(LB+r_4(UB-LB)\right) & ifq<0.5 \end{cases} \quad (1)$$

in which Y_{rand} represents a hawk randomly selected, Y_{prey} is the position of prey, and $r_i(i=1,2,3,4,q)$ is an accidental number ranging in (0,1). Y_m is the average position and is calculated as below:

$$Y_m(iter)=\frac{1}{N}\sum_1^N Y_i(iter) \quad (2)$$

The next stage is termed transition and focuses on modeling the energy of the prey as below:

$$E=2E_0\left(1-\frac{iter}{T}\right) \quad (3)$$

where E is the energy of the prey, E_0 is the initial energy of the prey and ranges between − 1 and 1 (−1, 1), and T is the number of total iterations. This equation indicates the more the prey escapes, the more its energy falls. Based on the value of computed E, the hawk makes a decision on how to continue its aim. It can decide to search for new locations or to use the neighborhood of the solutions:

$$\begin{cases} \textit{Start exploration phase if } |E|\geq 1 \\ \textit{Exploiting the neighborhood if } |E|<1 \end{cases}$$

In the exploiting phase, the value of $|E|$ plays an important role in the decision of the hawk to have a soft or hard besiege. Whenever $|E|$ is more than 0.5, the prey can escape by enough energy which it still has; hence, some jumps with the purpose of misleading it could be effective in making it fail, thus, a soft besiege works. Conversely, $|E|<0.5$ is an indication of fatigue less of the prey and the low chance to escape; therefore, Harris' hawks' hard besiege and performing the surprise pounce can work (Heidari et al. 2019).

2.5 Hybrid model: artificial neural network with Harris hawks' optimizer (ANNHHO)

The main purpose of the recommended hybrid model is to enhance the performance of the ANN by applying the HHO algorithm by discovering the optimal parameters (weights and biases) of the ANN, so, it is named ANNHHO. The hybrid ANNHHO model was used to estimate solar radiation (Fig. 2) (Sammen et al. 2020).

2.6 Phase space reconstruction (PSR)

A chaotic system displays a relatively complex behavior by the dynamics of a nonlinear system. The orbits of the chaotic system attract a strange attractor, which is a complex higher-dimensional subset and is important due to its extensive occurrences in the real world (Huang et al. 2010). Every dynamic system may be chaotic, stochastic, or deterministic, which can be identified by applying the PSR concept. PSR is the foundation for forecasting chaotic time series, in which a univariate time series is reconstructed using all the variable information in the dynamic system of the time series (e.g., Koçak et al. 2004). Each point in phase space is a description of the state of the system and each trajectory describes the time evolution of the system, which corresponds to diverse initial conditions.

Points in a phase space forming trajectories as attractors yield a quantitative estimate of the complexity of a system using the attractor's properties and determine the

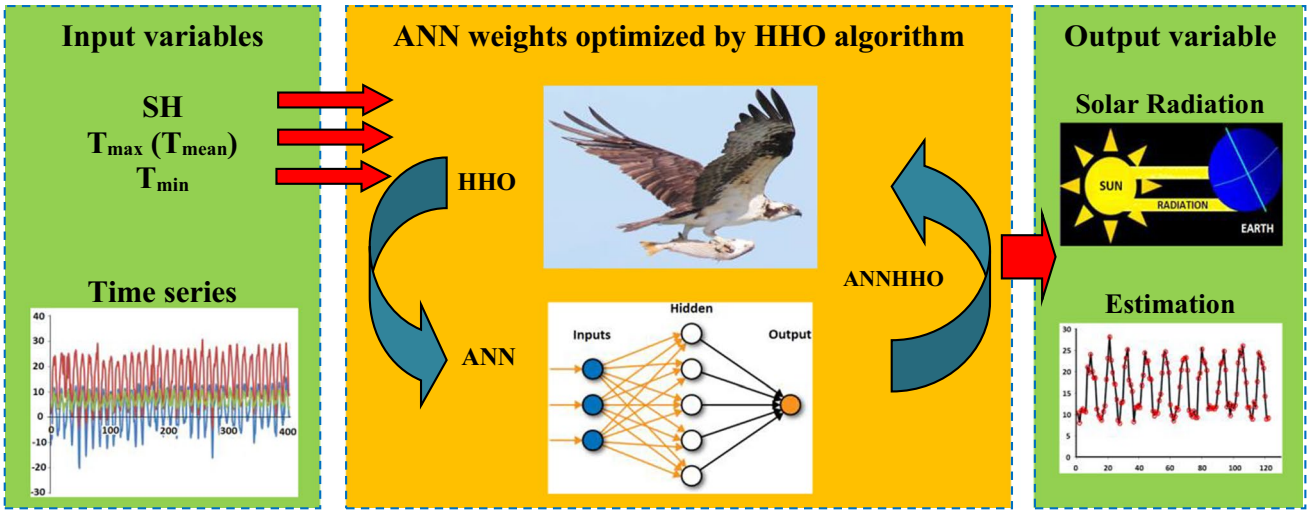

Fig. 2 Schematic presentation of the applied hybrid ANNHHO model in this study

nature of the dynamic behaviors. One way of reconstructing a phase space for a one-dimensional chaotic time series $S_t = (X_1, X_2, X_3, \ldots, X_n)$ in terms of the phase space vectors $\boldsymbol{X}_t$ is through the Takens time-delay embedding theorem. It is as follows (Takens 1981):

$$X_t = \left(X_t, X_{t-\tau}, X_{t-2\tau}, \ldots, X_{t-(m-1)\tau}\right) \quad t = 1,2,\ldots,M; M = n-(m-1)\tau \tag{4}$$

where τ is the delay time, m is the dimension of the phase-space reconstruction or embedding dimension, and M is the number of points in the reconstructed phase-space. The trajectories in the phase space diagram give information about the dynamics of a chaotic system (Sivakumar and Jayawardena 2002). The underlying structures in the chaotic time series can be observed when unfolded into the reconstructed phase space. The components of the current phase-space vectors, i.e., $X_t, X_{t-\tau}, X_{t-2\tau}, \ldots, X_{t-(m-1)\tau}$, are considered as the inputs of the network while keeping the future value X_{t+T} as the desired (output) response. Given the time series S_t, an m-dimensional PSR can be presented as follows (Takens 1981):

$$PSR = \begin{bmatrix} x_1 & x_{1+\tau} & x_{1+2\tau} & x_{1+3\tau} & \cdots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & x_{2+2\tau} & x_{2+3\tau} & \cdots & x_{2+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{n_m} & x_{n_m+\tau} & x_{n_m+2\tau} & x_{n_m+3\tau} & \cdots & x_{n_m+(m-1)\tau} \end{bmatrix} \tag{5}$$

Selecting an optimum value of embedding dimension m and delay time τ is an important step in reconstructing an appropriate phase space. The property of this optimum value of delay time is that the values of $X(i)$ and $X(i+\tau)$ is sufficiently independent of each other to be useful as coordinates in a time delay vector but not so independent that having no connection with each other at all.

2.6.1 Mutual information function (MIF)

MIF or average mutual information is often used to identify the optimum delay time for attractor reconstruction because it is better than the other alternatives such as autocorrelation function (e.g., Holzfuss and Mayer-Kress 1986), which is suitable to linear properties, and the correlation integral (e.g., Liebert and Schuster 1989), which requires more information and data. In this study, the first MIF minimum $I(\tau)$ is used to determine τ value as follows:

$$I(\tau) = \sum_{t=1}^{N-\tau} P\left(x_t, x_{t+\tau}\right).log\left(\frac{P\left(x_t, x_{t+\tau}\right)}{P\left(x_t\right).P\left(x_{t+\tau}\right)}\right) \tag{6}$$

where $I(\tau)$ is a statistical criterion of dependence of the reconstruction variables on each other, $P\left(x_t\right)$ is the probability density of x_t, and $P\left(x_t, x_{t+\tau}\right)$ is the joint probability density of x_t and $x_{t+\tau}$. If the reconstruction variables are statistically independent variables, then $I(\tau) = 0$. Complete dependence between the variables results in $I(\tau) = \infty$. An optimum value of delay time is obtained when the MIF is minimum (Fraser and Swinney 1986).

2.6.2 False nearest neighbour (FNN)

The FNN method, developed by Kennel et al. (1992), is used to identify the minimum embedding dimension m, as the number of effective variables for representing the system dynamics (Elshorbagy et al. 2002; Ng et al. 2007; Huang et al. 2010; Khatibi et al. 2012). In the phase space, points of the trajectories on the attractor have neighbors, which their behavior yields valuable information to find out the

evolution of neighborhoods for producing prediction equations (Abarbanel 1996). In the FNN algorithm, five stages are performed as follows: (i) searching for nearest neighbor $\boldsymbol{Z}_j$ of each point $\boldsymbol{Z}_i$ in the time series and an m-dimensional space; (ii) computing the distance $\|\boldsymbol{Z}_i - \boldsymbol{Z}_j\|$; and (iii) calculating R_i as follows:

$$R_i = \frac{\|\boldsymbol{Z}_{i+1} - \boldsymbol{Z}_{j+1}\|}{\|\boldsymbol{Z}_i - \boldsymbol{Z}_j\|} \tag{7}$$

(iv) if the value of R_i exceeds a given heuristic threshold (R_t), then the point $\boldsymbol{Z}_i$ is marked as having a false nearest neighbor; and (v) since the embedding dimension should be high enough, the fraction of points for which $R_i > R_t$ should be zero, or at least enough small (Huang et al. 2010).

2.7 Hybrid model: phase space reconstruction with ANNHHO model (PSR-ANNHHO)

The paper develops a hybrid model of Phase Space Reconstruction (PSR) with ANNHHO (Fig. 3). According to Fig. 3, the hybrid model involves two main steps: (i) identifying the optimum values of delay time and embedding dimension of the models' input variables to reconstruct their phase spaces; (ii) the dominant phase space signals and the original input variables are entered into the ANNHHO model to find out the relation between input and target (output) variables. These steps are similar to Uyumaz et al. (2014) and Ghorbani et al. (2018).

2.8 Performance criteria

To analyze the efficiency and performance of the applied models, six performance criteria, namely, coefficient of determination (CD), coefficient of efficiency (COE or Nash–Sutcliffe efficiency criterion), root mean square error (RMSE), mean absolute error (MAE), mean absolute percent error (MAPE), and Kling-Gupta efficiency (KGE), are used:

Coefficient of determination (CD) (Malik et al. 2020):

$$CD = \frac{\left[\sum_{i=1}^{N}\left(R_i - \overline{R_i}\right)\left(\widehat{R}_i - \overline{\widehat{R}_i}\right)\right]^2}{\sum_{i=1}^{N}\left(R_i - \overline{R_i}\right)\sum_{i=1}^{N}\left(\widehat{R}_i - \overline{\widehat{R}_i}\right)} \quad (0 < \mathrm{CD} < 1) \tag{8}$$

Coefficient of efficiency (COE) (Malik et al. 2020):

$$COE = 1 - \frac{\sum_{i=1}^{N}\left(R_i - \widehat{R}_i\right)^2}{\sum_{i=1}^{N}\left(R_i - \overline{R}\right)^2} \quad (-\infty < \mathrm{COE} < 1) \tag{9}$$

Root mean squared error (RMSE) (Malik et al. 2020):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(R_i - \widehat{R}_i\right)^2}{N}} \quad (0 < \mathrm{RMSE} < \infty) \tag{10}$$

Mean absolute error (MAE) (Kashani et al. 2020):

$$MAE = \frac{\sum_{i=1}^{N}\left|R_i - \widehat{R}_i\right|}{N} \tag{11}$$

Mean absolute percent error (MAPE) (Malik et al. 2020):

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{N}\left|\frac{R_i - \widehat{R}_i}{R_i}\right|}{N} \times 100 \quad (0 \le \mathrm{MAPE} < \infty) \tag{12}$$

Kling-Gupta efficiency (KGE):

$$KGE = 1 - \sqrt{\left(R^2 - 1\right)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \quad \beta = \frac{\overline{R_i}}{\overline{\widehat{R}_i}} \quad \gamma = \frac{CV_{R_i}}{CV_{\widehat{R}_i}} = \frac{\frac{\sigma_{R_i}}{\overline{R_i}}}{\frac{\sigma_{\widehat{R}_i}}{\overline{\widehat{R}_i}}} \tag{13}$$

where R_i and $\widehat{R}_i$ are the observed and estimated values of monthly solar radiation for an i^{th} dataset; N indicates the number of observations; and $\overline{R_i}$ and $\widehat{R}_i$ are the mean of observed and estimated values of monthly solar radiation, and σ_{R_i} and $\sigma_{\widehat{R}_i}$ are the standard deviation of observed and estimated values of monthly solar radiation, respectively. The models which are having a higher value of CD, COE, and KGE, and lower values of MAE, RMSE, and MAPE are considered relatively better models for monthly solar radiation estimation.

3 Results and discussion

One of the most important duties in any AI-based modeling is the selection of the best and most important input variables and discarding the less effective variables in order to obtain optimum results. In view of this, the relief algorithm was used in order to identify the most dominant input variables for the monthly global solar radiation modeling. Six variables were involved in the analysis, including $T_{\max}$, $T_{\min}$, U_2, RH, SH, and T_{mean}. Table 2 shows the results of the relief algorithm. According to Table 2, it is obvious that the SH, $T_{\max}$, and $T_{\min}$ with ranking orders of 1, 2, and 3 have high and positive values of the weights for Ardabil, so these variables are chosen as the input variables of the models. Similarly, for the Rasht region, the SH, T_{mean}, and $T_{\min}$ are the most effective variables for solar radiation estimation. Also, for the Zabol region, the SH, T_{mean}, and $T_{\min}$ are the most effective variables for solar radiation estimation. In general, the effective variables are almost the same for the three regions. It seems that the climate conditions do not significantly influence the determination of the dominant variables for solar radiation estimation.

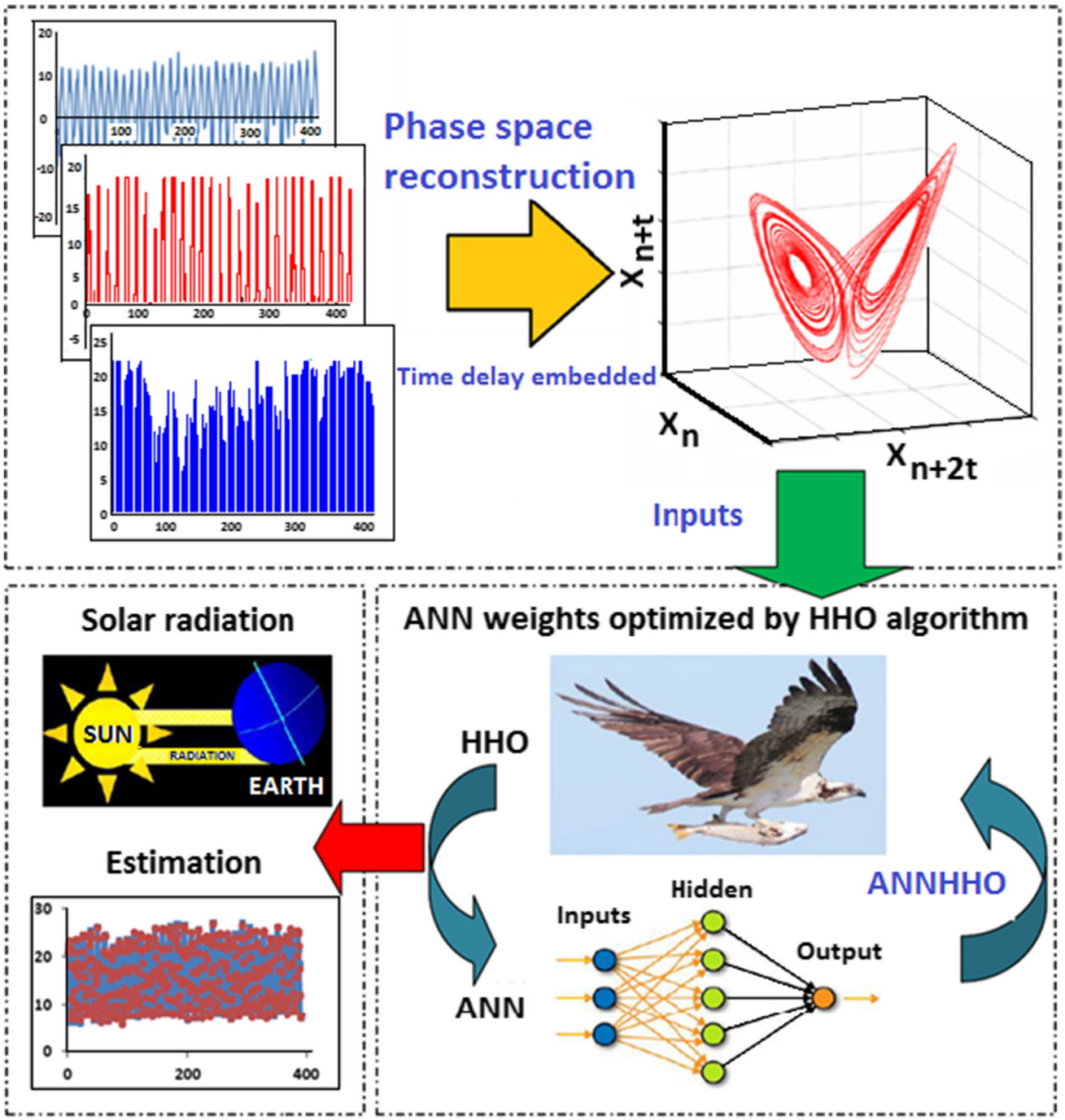

Fig. 3 Schematic presentation of the proposed PSR-ANNHHO model in this study

3.1 Simulation of solar radiation using heuristic models in Ardabil region

The optimum values of the delay time derived by MIF and Eq. (6), and embedding dimension obtained by FNN and Eq. (7), for Ardabil station, were obtained as 6 and 3, respectively, for the PSR-ANNHHO model. The estimations of monthly solar radiation using ANN (MLP), ANNHHO, and PSR-ANNHHO models were evaluated based on CD, COE, RMSE, MAE, and MAPE and during training and testing periods in the Ardabil region. The values of CD, COE, RMSE, MAE, and MAPE during the testing phase for ANN (MLP), ANNHHO, and PSR-ANNHHO models are presented in Table 3. As evaluated for Ardabil region based on Table 3, the ANN (MLP), ANNHHO, and PSR-ANNHHO models provide CD = 0.882, 0.888, and 0.990,

Table 2 Relief weights and ranking orders for all the variables

Ardabil			Rasht			Zabol		
Ranking order	Variable	Relief weight	Ranking order	Variable	Relief weight	Ranking order	Variable	Relief weight
1	*SH*	0.0127	1	*SH*	0.0083	1	*SH*	0.0211
2	T_{max}	0.0017	2	T_{mean}	0.0035	2	T_{mean}	0.0060
3	T_{min}	0.0014	3	T_{min}	0.0033	3	T_{min}	0.0035
4	T_{mean}	−0.0001	4	T_{max}	0.0011	4	T_{max}	0.0007
5	*RH*	−0.0025	5	*RH*	0.0002	5	U_2	0.0006
6	U_2	−0.0028	6	U_2	−0.0050	6	*RH*	−0.0023

COE = 0.842, 0.882, and 0.989, and KGE = 0.847, 0.859, and 0.978, RMSE = 2.244, 2.117, and 0.632, MAE = 1.832, 1.750, and 0.490, and MAPE = 15.126, 14.297, and 3.858 during the testing period. Table 3 indicates that all heuristic and hybrid models perform successfully. However, the PSR-ANNHHO model performance showed the best simulation based on the statistical criteria during the testing phase. Therefore, the PSR-ANNHHO model was selected as the best model among other applied models. The ANNHHO model followed the PSR-ANNHHO as the second rank model.

The temporal variations between the observed and simulated monthly solar radiation values along with scatter plots (modeled solar radiation against corresponding measured solar radiation) for the ANN, ANNHHO, and PSR-ANNHHO models during the testing phase are plotted in Fig. 4. This figure indicates that the estimated values of solar radiation are close to the observed values for all the models. However, it seems that the ANNHHO outperforms the ANN model in the estimation of solar radiation. It is obvious that the performance of the PSR-ANNHHO model is much better than the other models and estimates the high and low values of solar radiation more precisely. According to the scatter diagrams, a large number of data points are close to the best-fit line (1:1) and so display a good correlation for all the models, but there is some relatively strong discordant estimation at relatively higher ranges for the ANN and ANNHHO models. The scatter diagram of the PSR-ANNHHO model indicates the excellent performance of this model in estimating different ranges of solar radiation.

The performance of the models at their testing phase is also shown in Fig. 5, in which it shows that estimation error (modeled solar radiation minus corresponding measured solar radiation) ranges are between −7 and +4, −6 and +5, and −2 and +1 for the ANN, ANNHHO, and PSR-ANNHHO models, respectively. It is obvious that the estimation errors of the PSR-ANNHHO model are very lower than those of the other models. The estimation error ranges of the ANN and ANNHHO models are almost similar to each other.

Figure 6a and b show box and whisker plots and PDF plots of the actual and estimated values of the solar radiation for all the models. A box and whisker plot is a graphical approach for displaying variation in a set of data. Figure 6a indicates that the minimum and maximum values of the estimated solar radiation of the PSR-ANNHHO model are almost equal to the box whisker of the actual solar radiation data. The minimum and maximum values of the box whisker of the ANN and ANNHHO models are somewhat different from the actual ones. Moreover, a variation of the estimated data in the first 25%, second 25%, third 25%, and fourth 25%

Table 3 Comparison of performance of the applied models based on statistical criteria at the testing period

Region	Models	Statistical criteria					
		CD	COE	KGE	RMSE	MAE	MAPE
Ardabil	ANN (MLP)	0.882	0.842	0.847	2.244	1.832	15.126
	ANNHHO	0.888	0.882	0.859	2.117	1.750	14.297
	PSR-ANNHHO	0.990	0.989	0.978	0.632	0.490	3.858
Rasht	ANN (MLP)	0.903	0.807	0.772	2.445	2.023	21.864
	ANNHHO	0.907	0.907	0.936	1.778	1.327	13.337
	PSR-ANNHHO	0.994	0.994	0.989	0.442	0.328	2.990
	ANN (MLP)	0.947	0.940	0.933	1.308	1.113	6.226
Zabol	ANNHHO	0.947	0.944	0.921	1.261	1.053	6.005
	PSR-ANNHHO	0.998	0.998	0.938	0.255	0.203	1.201

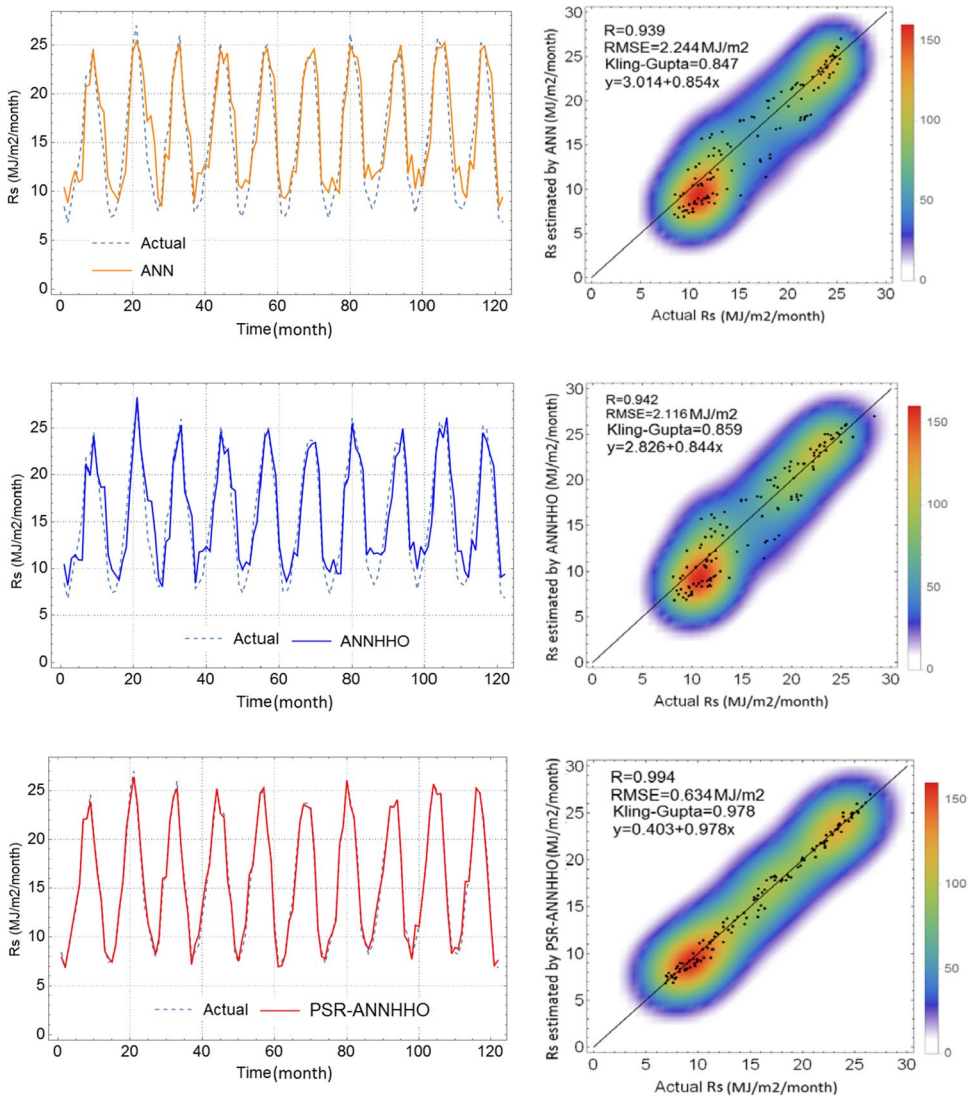

Fig. 4 Time series and scatter plots of the actual and estimated solar radiations of the models at the testing period in Ardabil region

of the whole data points of the PSR-ANNHHO model is almost the same as those of the actual box. This is not very true for the other two models. All the box plots are right-skewed; however, this is more sensible for the ANN and ANNHHO models. Figure 6b shows that the PDF plot of the PSR-ANNHHO model is very similar and close to the PDF plot of the actual solar radiation values. Here again, the performance and PDF plots of the ANN and ANNHHO models are similar to each other and somewhat different from the actual PDF plot. None of the PDF plots shows the normal distribution of the actual and modeled solar radiations of the models.

3.2 Simulation of solar radiation using heuristic models in Rasht region

The optimum values of the delay time derived by MIF and Eq. (6), and embedding dimension determined by FNN and Eq. (7), for the Rasht station, were obtained as 4 and 3, respectively, for the PSR-ANNHHO model. The estimations of monthly solar radiation using ANN (MLP), ANNHHO, and PSR-ANNHHO models were evaluated based on CD, COE, RMSE, MAE, and MAPE and during training and testing periods in Rasht region. The values of CD, COE, RMSE, MAE, and MAPE during the testing phase for ANN (MLP), ANNHHO, and PSR-ANNHHO models are presented in Table 3. As evaluated for Rasht region based on Table 3, the ANN (MLP), ANNHHO, and PSR-ANNHHO models provided CD = 0.903, 0.907, and 0.994, COE = 0.807, 0.907, and 0.994, KGE = 0.772, 0.936, and 0.989, RMSE = 2.445, 1.778, and 0.442, MAE = 2.023, 1.327, and 0.328, and MAPE = 21.864, 13.337, and 2.990 during the testing period. Table 3 indicates that all heuristic and hybrid models perform successfully. However, the PSR-ANNHHO model outperformed the other models based on the statistical criteria during the

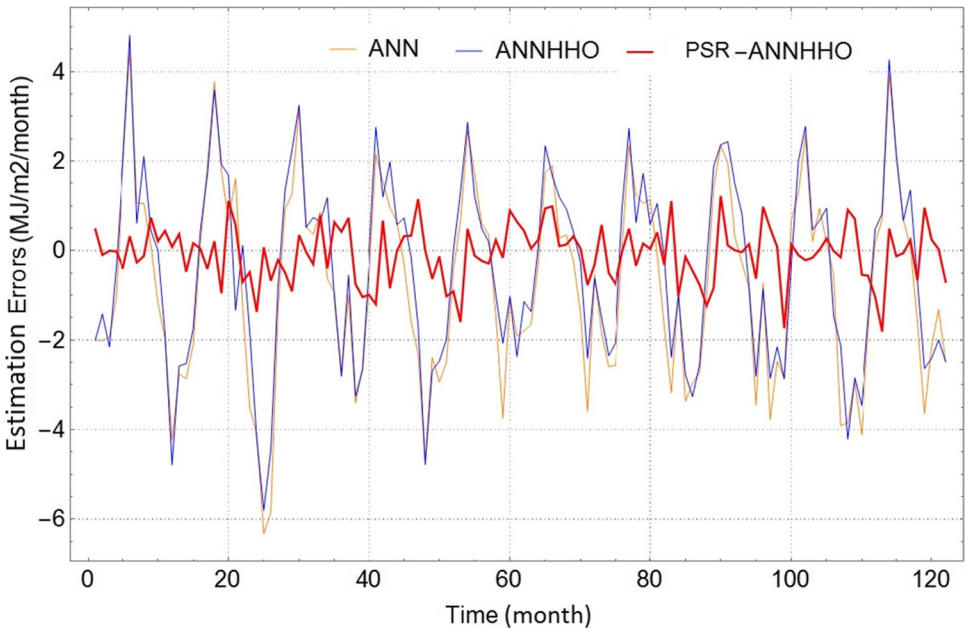

Fig. 5 Estimation error plot of the models at the testing period in Ardabil region

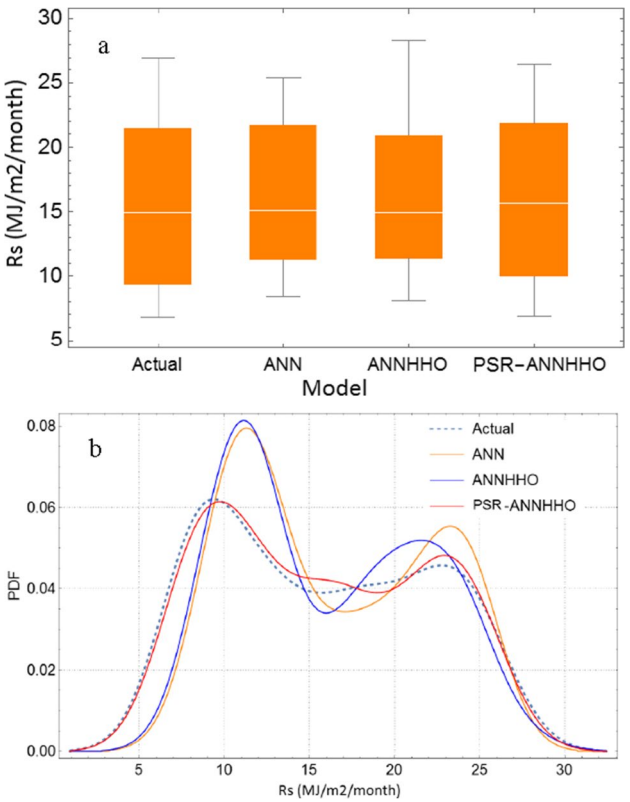

Fig. 6 Box and whisker plot (a) and PDF plot (b) of the models at the testing period in Ardabil region

testing period. Therefore, the PSR-ANNHHO model was selected as the best model among other intelligent models. The ANNHHO model followed the PSR-ANNHHO as the second rank model.

The temporal variations between the observed and simulated monthly solar radiation values for the ANN, ANNHHO, and PSR-ANNHHO models during the testing phase are plotted in Fig. 7. This figure indicates that the estimated values of solar radiation are close to the observed values for all the models. However, the ANNHHO outperforms the ANN model in estimating low and high values of solar radiation. It is obvious that the performance of the PSR-ANNHHO model is much better than the other hybrid and stand-alone models and estimates the high and low values of solar radiation more accurately. According to the scatter diagrams, a large number of data points are close to the best-fit line (1:1) and so display a good correlation for all the models, but there are some relatively strong discordant estimations at medium ranges for the ANN and ANNHHO models. The scatter diagram of the PSR-ANNHHO model indicates the excellent performance of this model in estimating different ranges of solar radiation.

The performance of the models at their testing phase is also shown in Fig. 8, which shows that estimation error ranges are between −6 and +4, −6 and +6, and −2 and +1 for the ANN, ANNHHO, and PSR-ANNHHO models, respectively. It is obvious that the estimation errors of the PSR-ANNHHO model are lower than those of the other models. The estimation error ranges of the ANN and ANNHHO models are almost similar to each other.

Figure 9a and b show box and whisker plots and PDF plots of the actual and estimated values of the solar radiation for all the models. Figure 9a indicates that the minimum and maximum values of the estimated solar radiation of the PSR-ANNHHO and ANNHHO models are

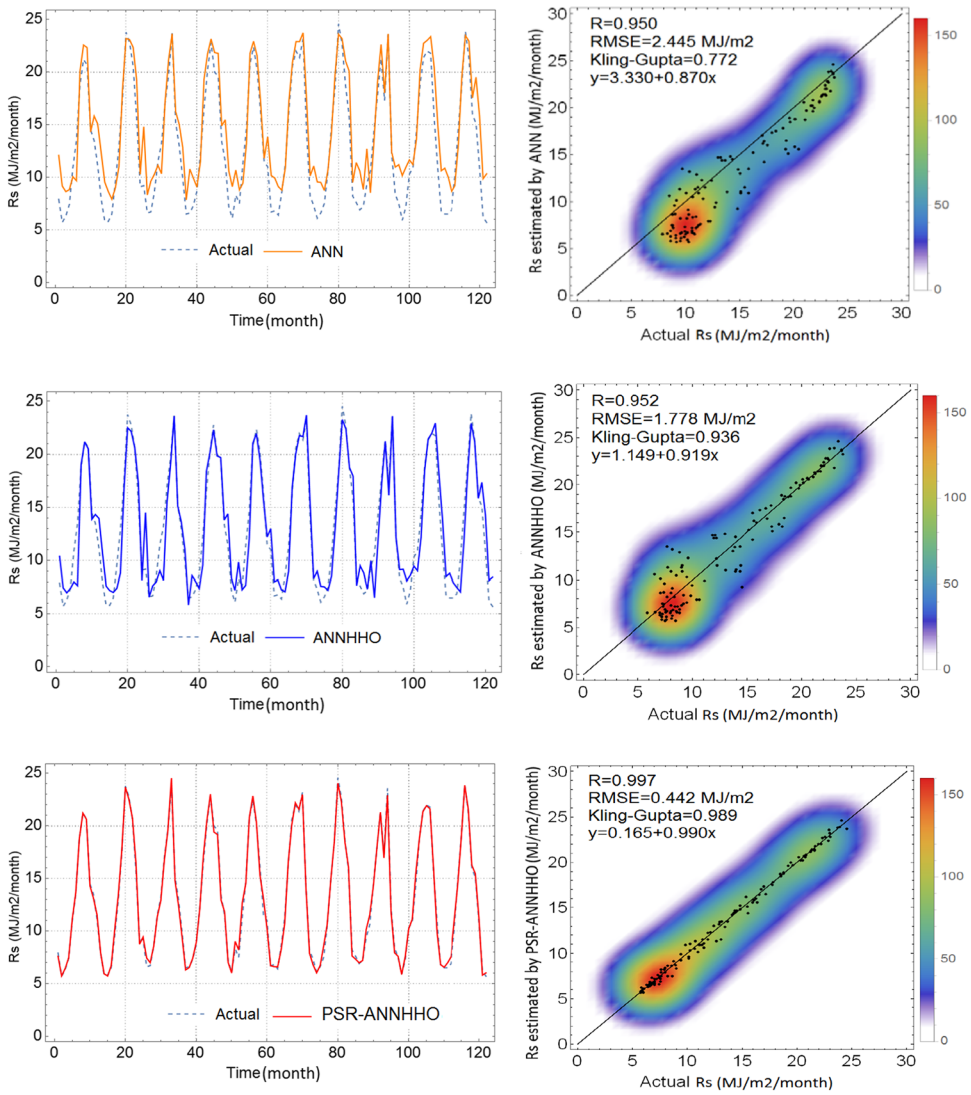

Fig. 7 Time series and scatter plots of the actual and estimated solar radiations of the models at the testing period in Rasht region

almost equal to the box whisker of the actual solar radiation data. The minimum and maximum values of the box whisker of the ANN model are somewhat different from the actual one. Moreover, a variation of the estimated data in the first 25%, second 25%, third 25%, and fourth 25% of the whole data points of the PSR-ANNHHO and ANNHHO models is almost the same as those of the actual box. This is not very true for the ANN model. All the box plots are right-skewed; however, this is more sensible for the ANN model. Figure 9b shows that the PDF plot of the PSR-ANNHHO model is very similar and close to the PDF plot of the actual solar radiation values. Here again, the performance and PDF plots of the ANN and ANNHHO models are almost similar to each other and somewhat different from the actual PDF plot especially the ANN model. None of the PDF plots shows the normal distribution of the actual and modeled solar radiations of the models.

3.3 Simulation of solar radiation using heuristic models in Zabol region

The optimum values of the delay time derived by MIF and Eq. (6), and embedding dimension determined by FNN and Eq. (7), for Zabol station, were obtained as 11 and 3, respectively, for the PSR-ANNHHO model. The estimations of monthly solar radiation using ANN (MLP), ANNHHO, and PSR-ANNHHO models were evaluated based on CD, COE, RMSE, MAE, and MAPE and during training and testing periods in Zabol region. The values of CD, COE, RMSE, MAE, and MAPE during the testing phase for ANN (MLP), ANNHHO, and PSR-ANNHHO models are also presented in Table 3. As evaluated for Zabol region based on Table 3, the ANN (MLP), ANNHHO, and PSR-ANNHHO models provide CD = 0.947, 0.947, and 0.998, COE = 0.940, 0.944, and 0.998, KGE = 0.933, 0.921, and 0.938, RMSE = 1.308, 1.261, and 0.255, MAE = 1.113, 1.053, and 0.203, and

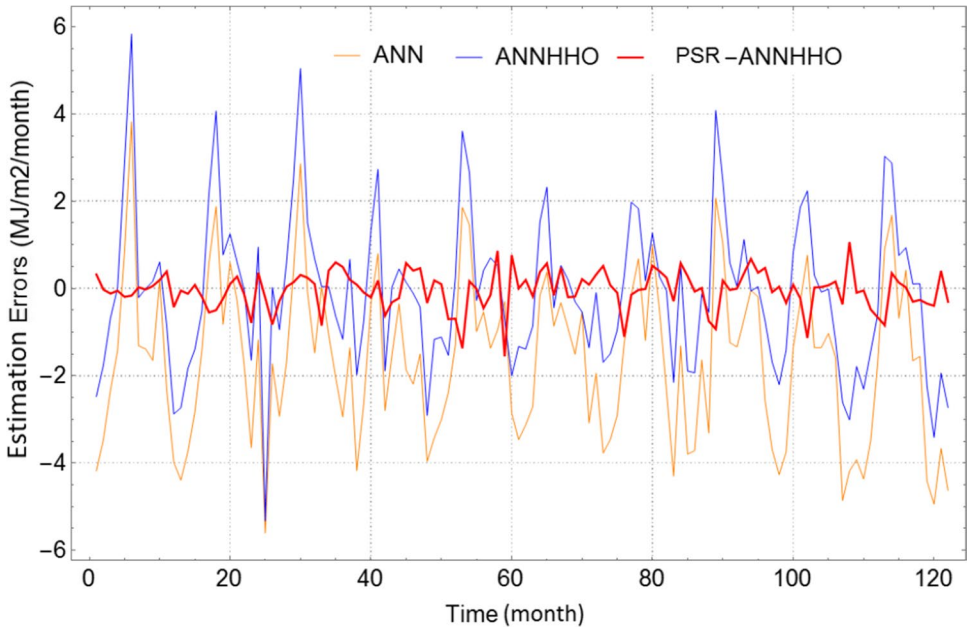

Fig. 8 Estimation error plot of the models at the testing period in Rasht region

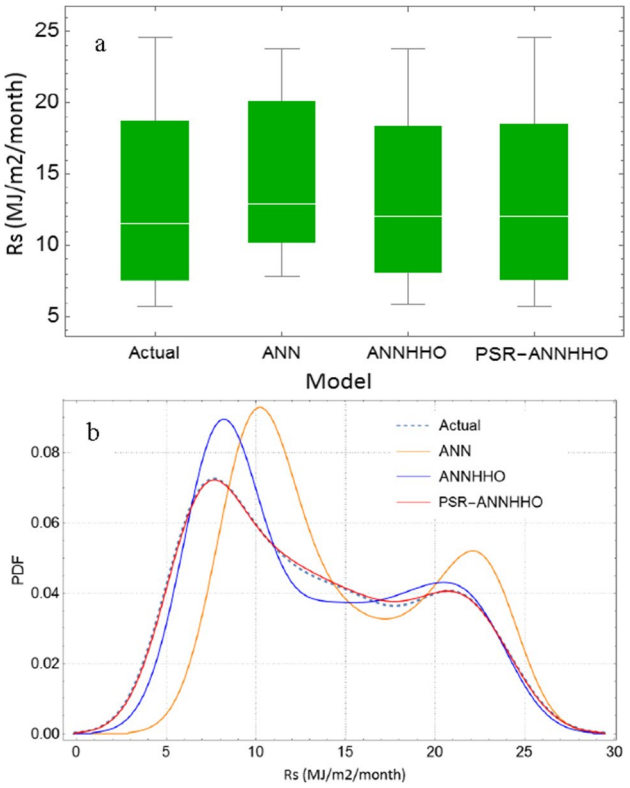

Fig. 9 Box and whisker plot (**a**) and PDF plot (**b**) of the models at the testing period in Rasht region

MAPE = 6.226, 6.005, and 1.201 during the testing period. Table 3 indicates that all heuristic and hybrid models perform successfully. However, the PSR-ANNHHO model outperformed the other models based on the statistical criteria during the testing period. Therefore, the PSR-ANNHHO model was selected as the best model among other intelligent models. The ANNHHO model followed the PSR-ANNHHO as the second rank model.

The temporal variations between the observed and simulated monthly solar radiation values for the ANN, ANNHHO, and PSR-ANNHHO models during the testing phase are plotted in Fig. 10. This figure indicates that the estimated values of solar radiation are close to the observed values for all the models, especially for the PSR-ANNHHO. In other words, the performance of the PSR-ANNHHO model is much better than the other hybrid and stand-alone models and estimates the high and low values of solar radiation more precisely. According to the scatter diagrams, a large number of data points are close to the best-fit line (1:1) and so display a good correlation for all the models, but there are some relatively strong discordant estimations at medium ranges for the ANN and ANNHHO models. The scatter diagram of the PSR-ANNHHO model indicates the excellent performance of this model in estimating different ranges of solar radiation.

The performance of the models at their testing phase is also shown in Fig. 11, which shows that estimation error ranges are approximately between − 3 and + 3, − 3 and + 3, and − 1 and + 1 for the ANN, ANNHHO, and PSR-ANNHHO models, respectively. It is obvious that the estimation errors of the PSR-ANNHHO model are lower than those of the other models. The estimation error ranges of the ANN and ANNHHO models are almost identical.

Figure 12a and b show box and whisker plots and PDF plots of the actual and estimated values of the solar radiation for all the models. Figure 9a indicates that the minimum and maximum values of the estimated solar radiation of the

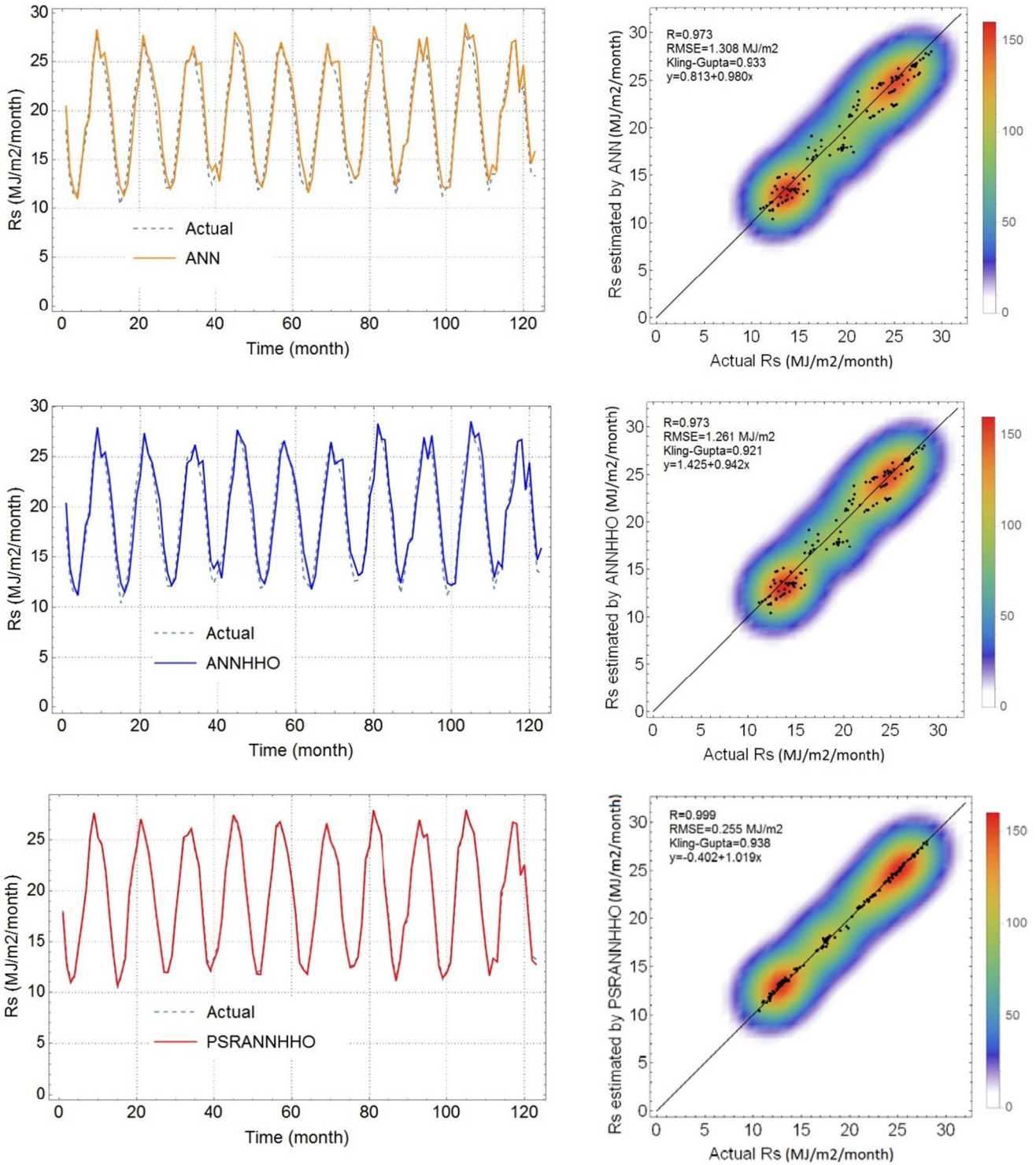

Fig. 10 Time series and scatter plots of the actual and estimated solar radiations of the models at the testing period in Zabol region

PSR-ANNHHO model are almost equal to the box whisker of the actual solar radiation data. The minimum values of the box whisker of the ANN and ANNHHO models are somewhat different from the actual ones. Moreover, a variation of the estimated data in the first 25%, second 25%, third 25%, and fourth 25% of the whole data points of the PSR-ANNHHO model is almost the same as those of the actual box. All the box plots are right-skewed; however, this is more sensible for the ANN model. Figure 9b shows that the PDF plot of the PSR-ANNHHO model is very similar and

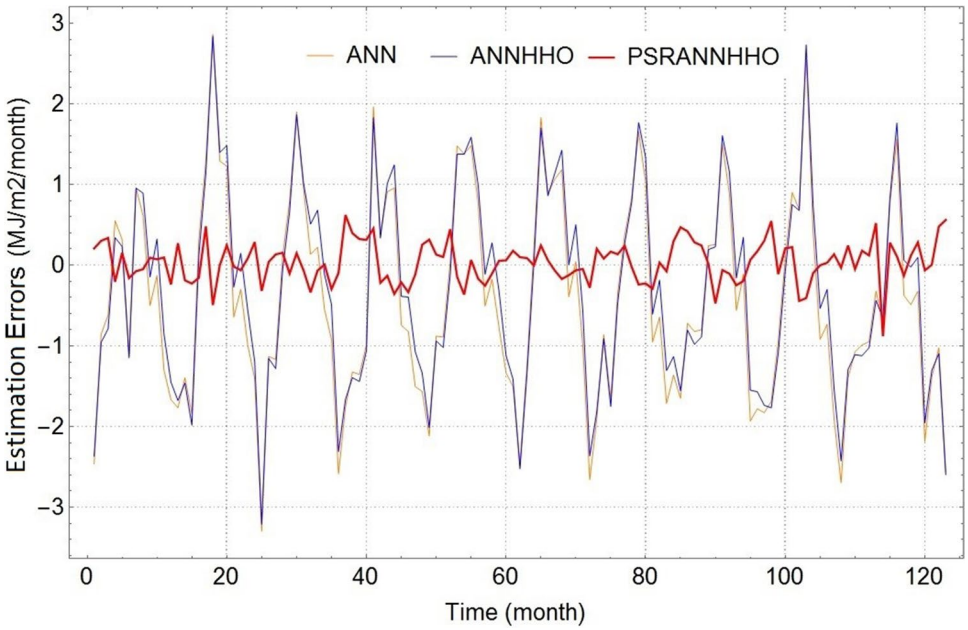

Fig. 11 Estimation error plot of the models at the testing period in Zabol region

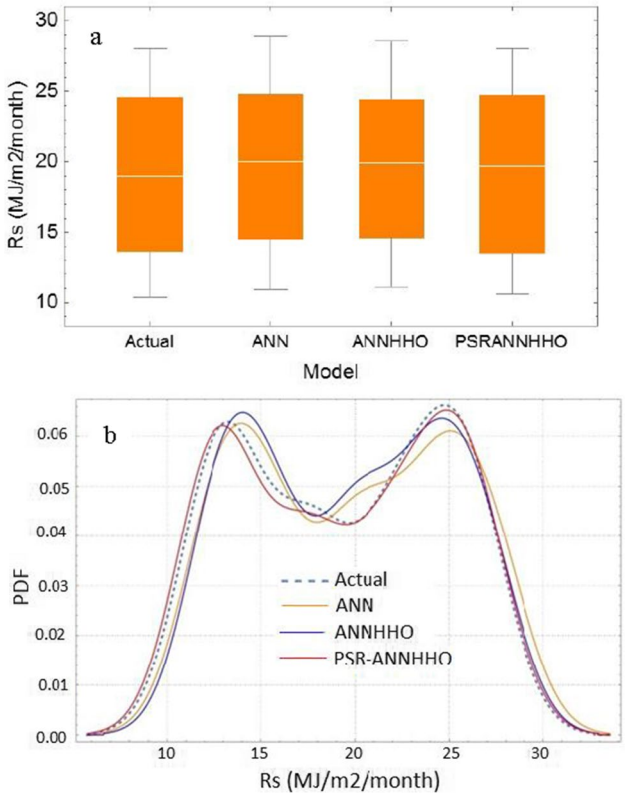

Fig. 12 Box and whisker plot (a) and PDF plot (b) of the models at the testing period in Zabol region

close to the PDF plot of the actual solar radiation values. Here again, the PDF plots of the ANN and ANNHHO models are almost similar to each other and somewhat different from the actual PDF plot. None of the PDF plots shows the normal distribution of the actual and modeled solar radiations of the models.

In general, all the heuristic stand-alone and hybrid models estimated solar radiation successfully in three regions with different climate conditions. However, the ANNHHO models showed high capability than the stand-alone ANN (MLP) model for solar radiation estimation in the regions, especially in Rasht. This proves the high capability of the HHO algorithm in optimizing the ANN's parameters. The hybrid PSR-ANNHHO model outperformed both the ANN and ANNHHO models in all regions and estimated monthly solar radiation more precisely especially in Zabol region. This indicates that applying chaos theory in the simulation of complex phenomena such as solar radiation improves estimation accuracy significantly. The stand-alone ANN model is the third-best model in solar radiation modeling in all three regions. It should be noted that all models' errors in Zabol region with arid climate conditions are less than Ardabil and Rasht regions with semi-arid and humid climate conditions, respectively. This means that the performances of the models are more accurate in arid regions. This is in accordance with the findings of Nourani et al. (2019). They concluded that the models show fewer errors in daily solar radiation estimation in arid regions of Iraq. Benghanem et al. (2009) applied ANN models for estimating daily solar radiation in Madinah with an arid climate and obtained satisfactory results and R^2 values of about 0.94. Acceptable performances of all the models in solar radiation simulation show that the input variables of the models have been selected rightly in all regions. In other words, the relief algorithm has high efficiency in determining the dominant input variables

in all regions with different climate conditions. This is in accordance with Almaraashi (2018).

4 Conclusion

In this research, the capability of two new hybrid models were studied including (1) artificial neural network with Harris hawks' optimizer (ANNHHO) and (2) artificial neural network with Harris hawks' optimizer and phase space reconstruction (PSR-ANNHHO) in monthly solar radiation estimation in three different regions (Ardabil, Rasht, and Zabol) of Iran with different climate conditions (semi-arid, humid, and arid). To this end, the performance of the new hybrid models was compared with the stand-alone ANN (MLP) model. Different climatic variables such as T_{max}, T_{min}, T_{mean}, RH, SH, U_2, and R_s were collected from the regions for the period of 1985–2018. According to the results of the relief algorithm, T_{max}, T_{min}, T_{mean}, and SH are the most effective variables in monthly global solar radiation simulation in all regions. Based on the statistical criteria, CD, COE, KGE, RMSE, MAE, and MAPE, this study showed all the heuristic models perform satisfactorily; however, PSR-ANNHHO and ANNHHO models are superior to the stand-alone ANN model in all the regions. The estimation accuracy of the hybrid PSR-ANNHHO model is higher than the ANNHHO model, so it is strongly recommended to be applied for solar radiation estimation in these regions. The HHO algorithm showed more capability in improving the stand-alone MLP model in all the regions, and the phase space reconstruction was identified as a powerful tool in improving the estimation accuracy of the hybrid ANNHHO model. The stand-alone ANN model showed relatively poor performance in the estimation of solar radiations rather than the other hybrid models in all regions. In general, climate conditions had not significantly influenced the models' input variables and models' performances. However, the models' performance was more precise in Zabol region with arid climate conditions. The uncertainty in the results of the models can be due to possible errors in measuring meteorological variables.

This study can be rather a valid reference in which the PSR-ANNHHO model is recommended to be used for solar radiation estimation in different climate conditions all over the world. For future studies, it is recommended to use other input variables such as the month of the year, latitude, and longitude and examine their influence on monthly solar radiation estimation. Moreover, it is recommended to use these powerful tools (ANNHHO and PSR-ANNHHO models) to predict electricity, which is generated by photovoltaics.

Acknowledgements The authors would like to thank the editors and the unknown reviewers for improving the quality of the paper.

Author contribution Conceptualization, Writing-review and editing: Mahsa H. Kashani.
Formal analysis and investigation: Samed Inyurt.
Data collection and data analysis, Methodology: Mohammad Reza Golabi.
Methodology: Mohammad AmirRahmani.
Methodology and Revising manuscript: Shahab S. Band.
All authors read and approved the final manuscript.

Data availability The datasets used in this study, were compiled and supplied by the Meteorology Organization in Iran. They are available from the corresponding author on reasonable request.

Code availability The code generated in this study is available from the corresponding author on reasonable request.

Declarations

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest/Competing interests (include appropriate disclosures) The authors declare no competing interests.

References

Abarbanel HDI (1996) Choosing the dimension of reconstructed phase space. In: Analysis of observed chaotic data. Institute for nonlinear science. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-0763-4_4

Al-Alawi SM, Al-Hinai HA (1998) An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. Renew Energy 14(1–4):199–204

Almaraashi M (2018) Investigating the impact of feature selection on the prediction of solar radiation in different locations in Saudi Arabia. Appl Soft Comput 66:250–263

Basaran K, Özçift A, Kılınç D (2019) A new approach for prediction of solar radiation with using ensemble learning algorithm. Arab J Sci Eng. https://doi.org/10.1007/s13369-019-03841-7

Baydaroğlu Ö, Koçak K (2014) SVR-based prediction of evaporation combined with chaotic approach. J Hydrol 508:356–363. https://doi.org/10.1016/j.jhydrol.2013.11.008

Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR (2010) The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Sol Energy 84(8):1468–1480

Benghanem M, Mellit A, Alamri SN (2009) ANN-based modelling and estimation of daily global solar radiation data: A case study. Energy Convers Manage 50(7):1644–1655

Casdagli M (1989) Nonlinear prediction of chaotic time series. Phys D Nonlin Phenom 35:335–356

Dhanya CT, Kumar DN (2010) Nonlinear ensemble prediction of chaotic daily rainfall. Adv Water Resour 33:327–347

Dhanya CT, Kumar DN (2011) Multivariate nonlinear ensemble prediction of daily chaotic rainfall with climate inputs. J Hydrol 403:292–306

Elshorbagy A, Simonovic SP, Panu US (2002) Estimation of missing streamflow data using principles of chaos theory. J Hydrol 255:123–133. https://doi.org/10.1016/S0022-1694(01)00513-3

Farmer JD, Sidorowich JJ (1987) Predicting chaotic time series. Phys Rev Lett 59:845

Fraser AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. Phys Rev A 33:1134–1140. https://doi.org/10.1103/PhysRevA.33.1134

Fathima TA, Nedumpozhimana V, Lee YH, Winkler S, Dev S (2019) A chaotic approach on solar irradiance forecasting. Photonics & Electromagnetics Research Symposium - fall (PIERS - Fall), Xiamen, China 2724–2728, https://doi.org/10.1109/PIERS-Fall48861.2019.9021305

Gaume E, Sivakumar B, Kolasinski M, Hazoumé L (2006) Identification of chaos in rainfall temporal disaggregation: application of the correlation dimension method to 5-minute point rainfall series measured with a tipping bucket and an optical raingage. J Hydrol 328:56–64

Ghorbani MA, Khatibi R, Mehr AD, Asadi H (2018) Chaos-based multigene genetic programming: a new hybrid strategy for river flow forecasting. J Hydrol 562:455–467

Golder J, Joelson M, Neel MC, Di Pietro L (2014) A time fractional model to represent rainfall process. Water Sci Eng 7:32–40

Guermoui M, Gairaa K, Boland J, Arrif T (2021) A novel hybrid model for solar radiation forecasting using support vector machine and bee colony optimization algorithm: review and case study. J Solar Energy Eng 143(2):020801

Halabi LM, Mekhilef S, Hossain M (2018) Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. Appl Energy 213:247–261. https://doi.org/10.1016/j.apenergy.2018.01.035

Haykin S (2009) Neural Networks and Learning Machines, 3rd edn. McMaster University, Canada

Heidari AA, Mirjalili S, Farisetal H (2019) Harris hawks' optimization: algorithm and applications. Future Gener Comput Syst. https://doi.org/10.1016/j.future.2019.02

Holzfuss J, Mayer-Kress G (1986) An approach to error-estimation in the application of dimension algorithms, in: Dimensions and Entropies in Chaotic Systems. Springer, 114–122

Huang SC, Chuang PJ, Wu CF, Lai HJ (2010) Chaos-based support vector regressions for exchange rate forecasting. Expert Syst Appl 37:8590–8598

Ibrahim IA, Khatib T (2017) A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. Energy Convers Manage 138:413–425. https://doi.org/10.1016/j.enconman.2017.02.006

Jadidi A, Menezes R, de Souza N, de Castro LA (2018) A hybrid GA–MLPNN model for one-hour-ahead forecasting of the global horizontal irradiance in Elizabeth city, North Carolina. Energies 11(10):2641. https://doi.org/10.3390/en11102641

Kalogirou SA (2001) Artificial neural networks in renewable energy systems applications: A review. Renew Sust Energ Rev 5(4):373–401

Kashani MH, Ghorbani MA, Shahabi M, Naganna SR, Diop L (2020) Multiple AI model integration strategy - application to saturated hydraulic conductivity prediction from easily available soil properties. Soil Tillage Res 196:104449

Kennel MB, Brown R, Abarbanel HDI (1992) Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A 45:3403–3411. https://doi.org/10.1103/PhysRevA.45.3403

Khatibi R, Sivakumar B, Ghorbani MA, Kisi O, Koçak K, Farsadi Zadeh D (2012) Investigating chaos in river stage and discharge time series. J Hydrol 414–415:108–117. https://doi.org/10.1016/j.jhydrol.2011.10.026

Kira K, Rendell LA (1992) The feature selection problem: traditional methods and a new algorithm. AAAI-92 Proceedings of the tenth national conference on Artificial intelligence, 129–134

Koca A, Oztop HF, Varol Y, Koca GO (2011) Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey. Expert Syst Appl 38(7):8756–8762

Koçak K, Şaylan L, Eitzinger J (2004) Nonlinear prediction of near-surface temperature via univariate and multivariate time series embedding. Ecol Modell 173:1–7. https://doi.org/10.1016/S0304-3800(03)00249-7

Koutsoyiannis D, Pachakis D (1996) Deterministic chaos versus stochasticity in analysis and modeling of point rainfall series. J Geophys Res Atmos 101:26441–26451

Liebert W, Schuster HG (1989) Proper choice of the time delay for the analysis of chaotic time series. Phys Lett A 142:107–111

Lovejoy S, Mandelbrot BB (1985) Fractal properties of rain, and a fractal model. Tellus A 37:209–232

Malik A, Kumar A, Kim S, Kashani MH, Karimi V, Ghorbani MA, Al-Ansari N, Salih SQ, Yaseen ZM, Chau KW (2020) Modeling monthly pan evaporation process over the Indian central Himalayas: application of multiple learning artificial intelligence models. Eng Appl Comput Fluid Mech 14(1):323–338

McClelland JL, Rumelhart DE (1989) Explorations in parallel distributed processing: A handbook of models, programs, and exercises. MIT Press, Cambridge MA

Mellit A (2008) Artificial Intelligence technique for modeling and forecasting of solar radiation data: A review. J Artif Intell Soft Comput Res 1(1):52–76

Moghaddamnia A, Remesan R, Kashani MH, Mohammadi M, Han D, Piri J (2009) Comparison of LLR, MLP, Elman, NNARX and ANFIS Models—with a case study in solar radiation estimation. J Atmos Solar-Terres Physics 71(8–9):975–982

Mohammadi K, Shamshirband S, Danesh AS, Zamani M, Sudheer C (2015) Horizontal global solar radiation estimation using hybrid SVM-firefly and SVM-wavelet algorithms: a case study. Natural Hazardshttps://doi.org/10.1007/s11069-015-2047-5

Mohanty S (2014) ANFIS based prediction of monthly average global solar radiation over Bhubaneswar (State of Odisha). Int J Ethics Eng Manage Edu 1(5):2348–4748

Mohanty S, Patra PK, Sahoo SS (2016) Prediction and application of solar radiation with soft computing over traditional and conventional approach—a comprehensive review. Renewable Sustain Energy Rev 56:778–796

Moreno A, Gilabert MA, Martı´nez B (2011) Mapping daily global solar irradiation over Spain: A comparative studyof selected approaches. Sol Energy 85(9):2072–2084

Ng WW, Panu US, Lennox WC (2007) Chaos based analytical techniques for daily extreme hydrological observations. J Hydrol 342:17–41. https://doi.org/10.1016/j.jhydrol.2007.04.023

Nourani V, Elkiran G, Abdullahi J, Tahsin A (2019) Multi-region modeling of daily global solar radiation with artificial intelligence ensemble. Natural Resour Res. https://doi.org/10.1007/s11053-018-09450-9

Olsson J, Niemczynowicz J, Berndtsson R (1993) Fractal analysis of high-resolution rainfall time series. J Geophys Res Atmos 98:23265–23274

Pasternack GB (1999) Does the river run wild? Assessing chaos in hydrological systems. Adv Water Resour 23:253–260

Porporato A, Ridolfi L (1996) Clues to the existence of deterministic chaos in river flow. Int J Mod Phys B 10:1821–1862

Porporato A, Ridolfi L (1997) Nonlinear analysis of river flow time sequences. Water Resour Res 33:1353–1367. https://doi.org/10.1029/96WR03535

Rahimikhoob A (2010) Estimating global solar radiation using artificial neural network and air temperature data in a semiarid environment. Renewable Energy 35(9):2131–2135

Rehman S, Mohandes M (2008) Artificial neural network estimation of global solar radiation using air temperature and relative humidity. Energy Policy 36(2):571–576

Rodriguez-Iturbe I, Febres De Power B, Sharifi MB, Georgakakos KP (1989) Chaos in rainfall. Water Resour Res 25:1667–1675

Rohani A, Taki M, Abdollahpour M (2018) A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). Renew Energy 115:411–422

Sammen SS, Ghorbani MA, Malik A, Tikhamarine Y, AmirRahmani M, Al-Ansari N, Chau KW (2020) Enhanced artificial neural network with harris hawks optimization for predicting scour depth downstream of ski-jump spillway. Appl Sci 10(15):5160

Shang P, Na X, Kamae S (2009) Chaotic analysis of time series in the sediment transport phenomenon. Chaos Soli Fractals 41:368–379

Sharifi SS, Rezaverdinejad V, Nourani V (2016) Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: a comparative study of selected temperature-based approaches. J Atmos Solar-Terres. Physics. https://doi.org/10.1016/j.jastp.2016.10.008

Sivakumar B, Liong SY, Liaw CY (1998) Evidence of chaotic behavior in Singapore rainfall. J Am Water Resour Assoc 34:301–310

Sivakumar B, Liong SY, Liaw CY, Phoon KK (1999) Singapore rainfall behavior: chaotic? J Hydrol Eng 4:38–48

Sivakumar B (2000) Chaos theory in hydrology: important issues and interpretations. J Hydrol 227:1–20

Sivakumar B (2001) Rainfall dynamics at different temporal scales: a chaotic perspective. Hydrol Earth Syst Sci Discuss 5:645–652

Sivakumar B, Jayawardena AW (2002) An investigation of the presence of low-dimensional chaotic behaviour inthe sediment transport phenomenon. Hydrol Sci J 47:405–416. https://doi.org/10.1080/02626660209492943

Sun Y, Babovic V, Chan ES (2010) Multi-step-ahead model error prediction using time-delay neural networks combined with chaos theory. J Hydrol 395:109–116. https://doi.org/10.1016/j.jhydrol.2010.10.020

Takens F (1981) Detecting strange attractors in turbulence. In: Rand D, Young L-S (eds) Dynamical systems and turbulence, Warwick, 1980: Proceedings of a Symposium Held at the University of Warwick 1979/80. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 366–381

Urbanowicz RJ, Meeker M, Cava WL, Olson RS, Moore JH (2018) Relief-based feature selection: Introduction and review. J Biomed Inform 85:189–203. https://doi.org/10.1016/j.jbi.2018.07.014

Uyumaz A, Danandeh Mehr A, Kahya E, Erdem H (2014) Rectangular side weirs discharge coefficient estimation in circular channels using linear genetic programming approach. J Hydroinform 16:1318–1330. https://doi.org/10.2166/hydro.2014.112

Wang Q, Gan TY (1998) Biases of correlation dimension estimates of streamflow data in the Canadian prairies. Water Resour Res 34:2329–2339

Yacef R, Mellit A, Belaid S, Şen Z (2014) New combined models for estimating daily global solar radiation from measured air temperature in semi-arid climates: application in Ghardaïa, Algeria. Energy Convers Manag 79:606–615. https://doi.org/10.1016/j.enconman.2013.12.057

Yadav AK, Chandel SS (2014) Solar radiation prediction using artificial neural network techniques: A review. Renew Sust Energ Rev 33:772–781