ORIGINAL ARTICLE

Enhancing daily rainfall prediction in urban areas: a comparative study of hybrid artificial intelligence models with optimization algorithms

Yaser Sheikhi[1] · Seyed Mohammad Ashrafi[1] · Mohammad Reza Nikoo[2] · Ali Haghighi[1]

Abstract

Forecasting precipitation is a crucial input to hydrological models and hydrological event management. Accurate forecasts minimize the impact of extreme events on communities and infrastructure by providing timely and reliable information. In this study, six artificial intelligent hybrid models are developed to predict daily rainfall in urban areas by combining the firefly optimization algorithm (FA), invasive weed optimization algorithm (IWO), genetic particle swarm optimization algorithm (GAPSO), neural network (ANN), group method of data handling (GMDH), and wavelet transformation. Optimization algorithms increase forecasting accuracy by controlling all stages. A variety of criteria are used for validating the models, including correlation coefficient (R), root-mean-square error (RMSE), mean absolute error (MAE), critical success index (CSI), probability of detection (POD), and false alarm ratio (FAR). The proposed models are also evaluated in an urban area in Ahvaz, Iran. The GAPSO-Wavelet-ANN model is superior to other models for predicting daily rainfall, with an RMSE of 1.42 mm and an R of 0.9715.

Introduction

Throughout history, understanding weather patterns has been a fundamental human concern. Water is an inevitable part of any country's economic and social life. Since global warming has increased in different parts of the world, forecasting and estimating precipitation have become more critical (Mostaghimzadeh et al. 2021). The accuracy of precipitation forecasting is essential for hydrological models and the management of urban events. Accurate forecasts minimize the impact of extreme weather events on communities and

✉ Seyed Mohammad Ashrafi
ashrafi@scu.ac.ir; semo.ashrafi@gmail.com

Yaser Sheikhi
yasersheikhi70@yahoo.com

Mohammad Reza Nikoo
m.reza@squ.edu.om

Ali Haghighi
a.haghighi@scu.ac.ir

[1] Faculty of Civil Engineering and Architecture, Shahid Chamran University of Ahvaz, Ahvaz, Iran

[2] Department of Civil and Architectural Engineering, Sultan Qaboos University, Muscat, Oman

infrastructure by providing timely and reliable information This is because timely forecasting reduces losses and damages caused by natural disasters, prevents flooding, protects urban and agricultural areas, and ensures proper resource utilization in various sectors, including agriculture and food production (El-shafie et al. 2011; Trinh 2018; Haddad 2011; Samsudin et al. 2010; Adib et al. 2021a, b). There are two main categories for simulating rainfall: numerical simulation and data-driven simulation (Shuman 1989; Olson et al. 1995).

Numerical weather forecasting models are deterministic and approximate complex physical processes for weather forecasting. Several research studies have utilized these models, multiple regressions, and climatological averaging methods. In traditional rainfall models, long-term measurements correlate with other meteorological factors, including air temperature, cloud information, solar radiation, sunshine duration, wind speed, and relative humidity (Mostaghimzadeh et al. 2023). A rainfall model can also be developed using autoregressive moving averages (ARMA) and autoregressive integrated moving averages (ARIMA). These models, however, have limitations in terms of the initial conditions and parameterization of the models, making them less reliable when used for forecasting at a smaller scale (Dabhi

and Chaudhary 2014; Pham et al. 2020; Silvestro and Rebora 2014; DelSole and Shukla 2002; Azadi et al. 2013; Novak et al. 2014; Gouda et al. 2020; Nayagam et al. 2008).

In the past decade, data-driven models have gained significant popularity for their versatility and effectiveness. The models are based on intelligence computing and machine learning techniques and utilize large amounts of data related to a modeled system's physics (e.g., hydraulic and hydrological phenomena). As technology advances, new tools have been developed to improve data collection, analysis, and presentation. In addition, as computer science (hardware and software) continues to evolve and develop, data-driven modeling has become increasingly in recent decades; neural networks have been widely used in time series modeling as part of data-oriented methods. Many researchers have proven the ability of neural networks to model and forecast natural time series. Furthermore, in several studies, artificial neural networks have been used to predict seasonal, monthly, and daily rainfall (Kuligowski and Barros 1998; Abrahart and See 2000; Luk et al. 2000; Luk et al. 2000; Valverde Ramírez et al. 2005; Sedki et al. 2009; Abudu et al. 2010; Wu and Chau 2013; Nastos et al. 2013; Gökbulak et al. 2015; Moazami et al. 2016; Nourani 2017; Qiu et al. 2017; Danladi et al. 2018; Le et al. 2020; Ashrafi et al. 2020; Mostaghimzadeh et al. 2021; Mostaghimzadeh et al. 2022).

An effective rainfall-runoff model relies on accurate and practical input data. These models use hydrological time series that are non-stationary or have seasonal patterns. These time series typically include a wide range of time scales. Preprocessing of data reduces the effects of these factors in modeling. In the context of time–frequency representation, the wavelet transform is a powerful mathematical approach that analyzes a signal in the time domain (Daubechies 1990). In recent years, wavelet transformation has become one of the valuable methods for analysis, such as periodic changes and tendencies (trends) in time series. Researchers have analyzed and reduced the input noise of observations using wavelet transforms for better training and improved prediction results. Nayak et al. (2013) and Partal and Kişi (2007) developed neural fuzzy wavelet hybrid models to forecast daily rainfall in the watersheds of Turkey. They selected the daily rainfall data from three stations in Turkey and used wavelet transformations to decompose it into several subseries, which were then used as inputs to the neuro-fuzzy model to predict the daily rainfall. The neuro-fuzzy wavelet hybrid model performed well, especially for time series with zero rainfall in the summer months and maximum values during the test (handling missing data). As a final step, this model was compared with the neuro-fuzzy model. According to their results, this model makes more accurate predictions than the neuro-fuzzy model. Nourani et al. (2009) used a combined wavelet neural network approach for rainfall-runoff modeling. As a result, rainfall

and runoff time series were divided into subseries by wavelet transformation. This subseries was then entered into the artificial neural network to predict the next day's runoff. The results showed that the model could predict long-term and short-term runoff using multi-scale rainfall-runoff time series as input (Partal and Cigizoglu 2009; Altunkaynak and Nigussie 2015). Farajpanah et al. (2020) stated that wavelet functions play a significant role in improving the performance of the artificial intelligence (AI) models used for estimating daily flow discharge. In the study, five types of mother wavelet functions (MWFs) were employed. Wavelet functions were applied to decompose the input data, particularly the variables of daily flow discharge, temperature, and precipitation, using a technique called discrete wavelet transform (DWT). The DWT breaks down the data into different frequency components, allowing the AI models to analyze and capture the variations in the data at different scales. By incorporating wavelet functions, the researchers aimed to enhance the AI models' ability to capture and represent the complex patterns and variations present in the daily flow discharge data. The combination of AI models with wavelet functions improved the models' performance, as indicated by increased correlation. The wavelet functions played a crucial role in preprocessing the data and providing a more detailed and informative representation of the input variables, enabling the AI models to make more accurate estimations of the daily flow discharge. The study highlights the importance of accurately selecting the appropriate AI-based model, determining the relevant input data, and choosing the most effective MWF to improve the estimation of daily flow discharge. The findings have implications for water resource planning and management. Adib et al. (2021a, b), for estimating snow depth from microwave imager sounder (SSMIS) data, used several wavelet transforms, such as discrete wavelet transforms (DWT), maximum overlap discrete wavelet transforms (MODWT), multiresolution-based MODWT (MODWT-MRA), and wavelet packet transforms (WP), in combination with artificial intelligence models such as multilayer perceptrons, radial basis functions, adaptive neuro-fuzzy inference systems (ANFIS), and gene expression programming. Ashrafi et al. (2020) proposed an efficient operating reservoir rule curve using discrete wavelet transforms and artificial neural networks. A discrete wavelet decomposition model was used to determine the input variables for the artificial neural network training process. Thus, using forecasts, the shortage's severity was generally reduced over a long time 60-year period. Another data-oriented method is the artificial model of the group method of data handling (GMDH), developed by Ivakhnenko (1968), as a multivariate analysis and plan for identifying and modeling complex systems. The GMDH model is a sub-model of the ANN used to model complex systems. With sufficient data, it is possible to model such systems without having prior knowledge of

the processes involved. Various scientific fields have utilized this method to model time series, including science, medical diagnosis, signal processing, engineering, control systems, economics, and water resources management (Nikolaev and Iba 2003; Amanifard et al. 2008; Li et al. 2020). Wang et al. (2021) compared this method with regression neural networks (BPNN) and ARIMA models. Narawi et al. (2022) used the GMDEH model to predict rainfall in a watershed in Malaysia. Lake et al. (2022) in their study discuss various mathematical aspects of GMDH, including data partitioning, partial description synthesis, and using an external criterion for polynomial selection. They also investigate methods for improving modeling accuracy, such as hybridization with most miniature square support vector machines (LSSVM), the application of filters for parameter estimation, and the combination with signal processing techniques like ensemble empirical mode decomposition (EEMD), wavelet transformation (WT), and wavelet packet transformation (WPT). The inclusion of exogenous data and its integration into the GMDH modeling paradigm are also discussed. Alves et al. (2023) presented a methodology for monthly rainfall forecast is presented using the group method of data handling (GMDH) and sea surface temperature (SST). After a variable selection step, the intelligent model gets the mean monthly SST in predefined and temporally lagged areas. For model training, precipitation data from the Climate Prediction Center were used. The methodology was applied in a particular area of the municipality of Marabá, located in the southeastern region of the Pará state. Results show the GMHD's effectiveness for the monthly rainfall prediction, constituting an essential tool for planning and assistance to decision-makers. Mohseni and Muskula (2023) present a study on developing rainfall-runoff models using artificial neural networks (ANNs) in the Yerli sub-catchment. The study covers a period of 36 years, from 1981 to 2016. The ANN models, specifically the feed-forward backpropagation neural network (FFBPNN) and cascade-forward backpropagation neural network (CFBPNN), were used to establish the correlation between input (rainfall) and output (runoff) data sets. The results showed that the FFBPNN model outperformed the other models. The study also compared different training algorithms, including Levenberg–Marquardt (LM), Bayesian regularization (BR), and conjugate gradient scaled (CGS) and found that the LM-trained model with 30 neurons performed the best.

One of the challenges observed in previous studies is the manual selection and optimization of inputs, preprocessing, and simulation parameters, which leads to inaccurate predictions. The present study aims to develop and compare several artificial intelligence hybrid models for predicting daily rainfall in urban basins. The models used in this research include ANN, GMDH simulators, IWO, FA, and GAPSO optimizers using wavelet transformation. It is important to

note that this study is the first to develop a combination of models for predicting daily rainfall in urban basins in order to manage urban floods, which are IWO-Wavelet-ANN (IWA), IWO-Wavelet-GMDH (IWG), FA-Wavelet-ANN (FWA), FA-Wavelet-GMDH (FWG), GAPSO-Wavelet-ANN (GPWA), and GAPSO-Wavelet-GMDH (GPWA). Automatic and optimal control processes of the various modeling stages are included in the proposed models, such as the selection of the decomposition level, mother wavelet, inputs, and simulator parameters that enhance the prediction accuracy.

The rest of this article is organized as follows. First, the algorithms and models used (including GAPSO, FA, IWA, Wavelet, ANN, and GMDH) and other assumptions are discussed in the "Materials and Methods" section. The "Results and Discussion" section describes the study area, input data, and results of developing different hybrid models for daily rainfall in urban areas. Finally, the "Conclusion" section concludes the paper.

Materials and methods

$$X_{\text{scaled}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \qquad (1)$$

Figure 1 shows how to predict daily rainfall by the developed models. In this research, precipitation forecasting is carried out using three components: optimizer models, wavelet transforms, and simulators. Optimizers have the role of controlling wavelet transforms and simulators to perform their most optimal function. According to Eq. (1), the available data are scaled from 0 to 1 and 2019 for testing it. The current research's target cost function is the root-mean-square error (RMSE). Next, the optimizer evaluates the cost function's value at a specific decomposition level to select the best mother wavelet, decomposed input subseries, and simulation parameters. In each model execution, the inputs are specified first. Then, a decomposition level between 1 and 10 is selected during the second step. The third step involves choosing the mother wavelet and decomposing the input series into subseries (approximation and detail features).

In the fourth step, the simulation parameters are determined, and the simulator begins working; finally, the amount of the target cost function is checked, and if the stopping criterion has been satisfied, the forecasting process stops, and the outputs are stored for final processing, along with a return to the original scale. Otherwise, the model execution process is repeated from the second stage.

For developing the IWG, FWG, and GPWG models, the IWO, FA, and GAPSO optimizers, discrete wavelet transforms (DWT), and the GMDH simulator are used. This model decomposes input series at ten levels into
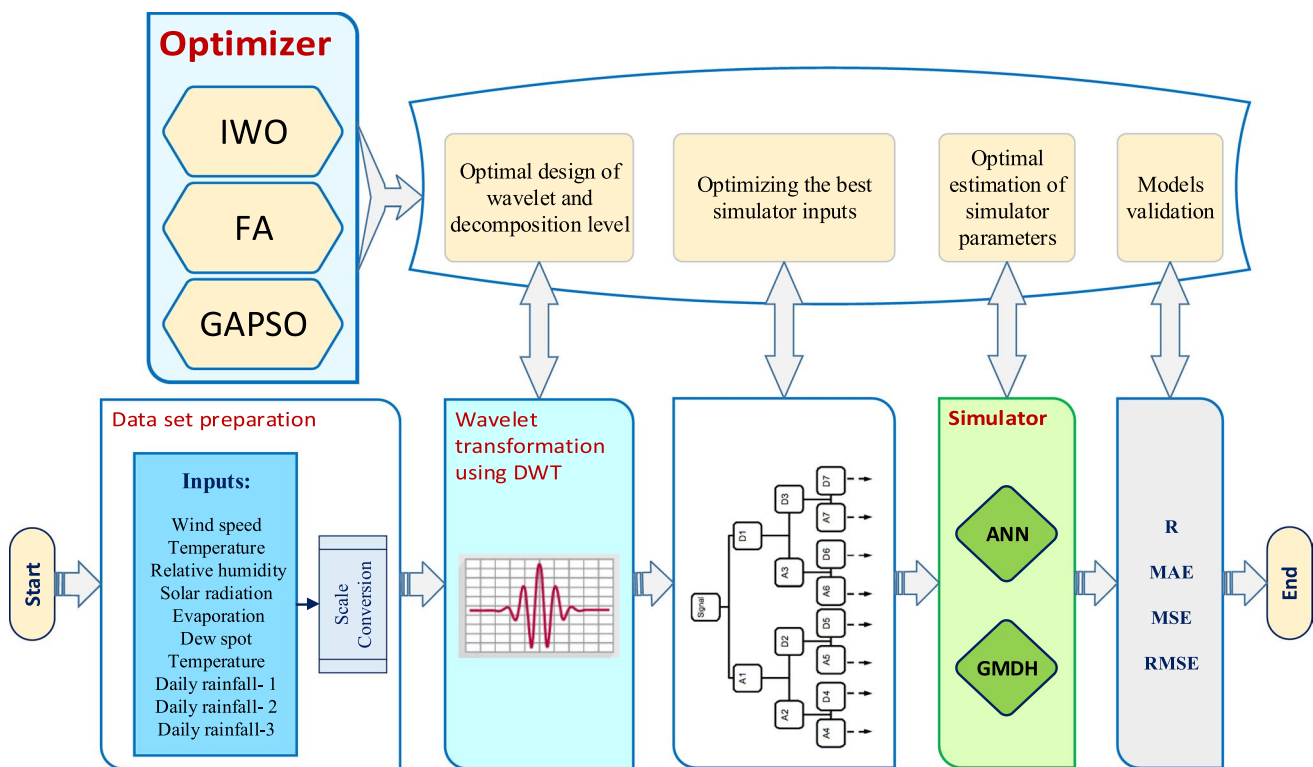
Fig. 1 General schematic of the proposed models

approximate and detailed subseries using discrete wavelet transforms. The model is executed after selecting the initial mother wavelet type, the top ten subseries, and setting the simulator's effective parameters (number of neurons and layers). Then, the optimizer increases the prediction accuracy by optimizing each iteration's model components (mother wavelet type, the top ten subseries, number of neurons, and layers).

The steps of the previous models are repeated in IWA, FWA, and GPWA models, with the difference being that the neural network is used instead of the GMDH simulator. In these models, feed-forward and cascade-feed-forward backpropagation networks are used to construct neural networks with training functions, including quasi-Newton backpropagation (BFGS quasi-Newton backpropagation), scaled conjugate gradient backpropagation, regularization backpropagation, Levenberg–Marquardt backpropagation, and resilient backpropagation. In addition, the training epoch ranges from 1 to 30. Finally, it is essential to note that the model optimization part is responsible for selecting and optimizing all the effective parameters of the model (mother wavelet type, the top ten subseries, network

type, training function, and training epoch) to reduce the RMSE error between predictions and observations.

Group method of data handling (GMDH)

GMDH is a network statistical training technology developed through cybernetics research. GMDH can be applied to various topics, including discovering relationships, forecasting, modeling, optimization, and recognizing nonlinear patterns. GMDH algorithm has several advantages for simulating time series data. It can capture nonlinear relationships, adapt its model structure, handle complex relationships, generate interpretable models, handle small and noisy datasets, work with various types of time series data, avoids overfitting, uses cross-validation to prevent overfitting to training data, fast training, and computationally efficient iterative training procedure. These advantages make GMDH a versatile and powerful tool for simulating time series, especially in scenarios with nonlinear dynamics and limited data availability. GMDH contains a set of layers and neurons that are created by connecting different pairs through a second-degree polynomial. Each layer has two inputs and one output for each processing unit (neurons). As a result of regression

techniques, the coefficients of neural transfer functions are derived as polynomials (Green et al. 1988; Tsai and Yen 2017; Lake, 2022). According to Eq. (2), a polynomial function can be used to express the connection between input and output variables (Zhang et al. 2013):

$$Y = a. + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ijk} x_i x_j x_k \quad (2)$$

In this equation, X_i is the input variable, Y is the estimated target value, and a is the coefficient. The form of the second degree and two variables of these sentences are in the form of Eq. (3):

$$Y = G\left(X_i X_j\right) = a_0 + a_1 X_i + a_2 X_j + a_3 X_i^2 + a_4 X_j^2 + a_5 X_i X_j \quad (3)$$

In this equation, each pair of input variables X_i and X_j are used to calculate the unknown coefficient a_i to minimize the difference between the estimated Y value and the actual target, Y. A group of polynomials is constructed using the above relationship. Then, each unknown coefficient is determined by applying the least squares method (Amiri and Soleimani 2021). Figure 2 represents the structure of a general GMDH neural network.

Artificial neural networks

An artificial neural network is considered a black box method. A network's input and output data allow it to determine relations between them without considering the
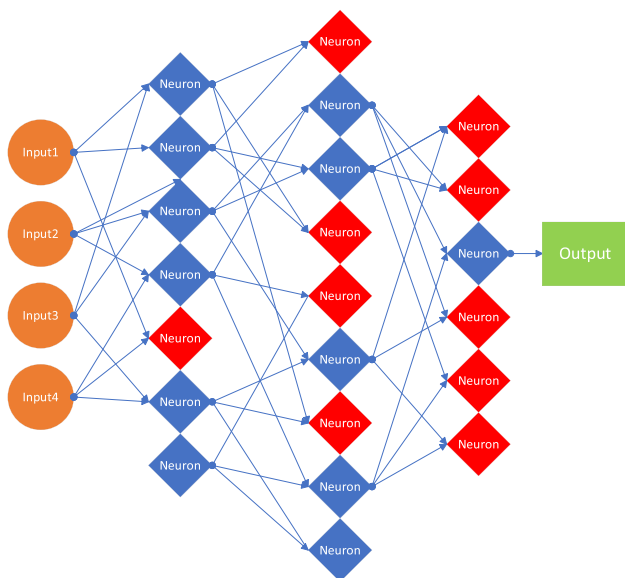

Fig. 2 GMDH network with four inputs and three layers, Blue neurons: verified neurons for the next step, Red neurons: unverified neurons

physical processes that govern the system. Artificial neural networks process experimental data to transfer knowledge or laws hidden behind the data into the network structure. ANNs' models are often characterized by network topology, node properties, and learning rules. The general structure of ANN consists of several layers. The input layer is responsible for spreading the data in the network, and the intermediate (hidden) layers are responsible for their processing. In addition, the output layer also shows the output and performance of the network input vector (Narayanakumar and Raja 2016; Matsumura et al. 2019). Equation (4) shows the relationship between different parts of a neural network:

$$Y(x) = F\left(\sum_{i=1}^{L} W_i(t).X_i(t) + b\right) \quad (4)$$

where X is the inputs variables vector, Y represents the predicted/simulated target value, F stands for the activation function, W is the vector of connection weights, L represents the hidden cells number, and b is the bias value (Ashrafi et al. 2020). Based on the selection of the optimizer, backpropagation neural networks, either feed-forward or cascade-feed-forward, are used as simulation models in this study. The feed-forward neural network is the most widely used type of neural network for rainfall-runoff prediction, and its advantages have been explored in previous rainfall forecasting studies (Hung et al. 2009; Valverde Ramírez et al. 2005; Khalili et al. 2016). The feed-forward neural network offers advantages for simulating time series data. These advantages include arbitrary connections allowing modeling complex time series, tapped delay lines capture temporal relationships, modeling non-stationary time series, enabling continuous adaptation on streaming data, robust activation functions handle noisy data, lagged connections providing insights into temporal dependencies, weight factorization speeds up training, and explicitly designed for one-step ahead forecasting. Here are the mathematical formulations of a feed-forward neural network:

The forward propagation process involves computing the activations of neurons in each layer of the network, starting from the input layer and progressing through the hidden layers to the output layer.

For a given layer l and neuron j, the activation is computed as follows:

$$y_j^l = \sum_{ij}^{L} (W_{ij}^l * a_{ij}^{l-1} + b_j^l) \quad (5)$$

$$a_j^l = \sigma\left(z_j^l\right) \quad (6)$$

Here, W_{ij}^l represents the weight connecting neuron i in layer $(l-1)$ to neuron j in layer l, a_{ij}^{l-1} is the activation of neuron i

in layer $(l$-1$)$, b_j^l is the bias term for neuron j in layer l, σ is the activation function, and $\sum$ denotes summation over all input neurons in layer $(l$-1$)$.

Backpropagation is used to compute the gradients of the weights and biases with respect to the loss function, which allows for their update during the training process.

The gradient of the loss function with respect to a weight W_{ij}^l is given by:

$$\partial E / \partial w_{ij}^l = \delta_j^l + a_i^{l-1} \tag{7}$$

Here, E represents the error or loss function, δ_j^l is the error term of neuron j in layer l, and a_i^{l-1} is the activation of neuron i in layer $(l$-1$)$.

The weights and biases of the neural network are updated using an optimization algorithm such as gradient descent or one of its variants. The update rule typically follows the form:

$$W_{ij}^l(t+1) = \left(W_{ij}^l(t) - \eta * \frac{\partial E^l}{\partial w_{ij}} \right) \tag{8}$$

Here, $W_{ij}^l(t+1)$ is the updated weight, $W_{ij}^l(t)$ is the current weight, η is the learning rate, and $\partial E / \partial w_{ij}^l$ represents the gradient of the loss function with respect to the weight.

Algorithm of weeds

In the agricultural sector, weeds are plants that grow in unwanted places and pose a severe threat to the growth of agricultural plants. Mehrabian and Lucas (2006) introduced and used the invasive weeds algorithm with inspiration from nature. This algorithm has several advantages, including its speed and effectiveness in finding the optimal points. It maintains diversity in the population, adapts to different problem types, and does not require gradient information. It is robust to parameter choice, suitable for parallel computing, and uses a dynamic population size to reduce memory usage. The system has a very high success rate because the system is based on weeds' primary and natural characteristics, such as growth, seed production, and survival conflict within colonies. These advantages make IWO effective in solving optimization problems with non-differentiable, discontinuous, and noisy objective functions while being versatile and efficient in various problem domains. The steps for implementing the weed algorithm are as follows:

Generating the primary population: A certain number of answers are selected with uniform distribution in the search space of the question.
Calculating the degree of fit: The compatibility of weeds is measured with the fit function.

Determining the number of new seeds: According to each parent weed's competence level, the number of new seeds around each weed is determined.
The relationship of seed number production is in the form of Eq. (9):

$$N = \frac{f - f_{\min}}{f_{\max} - f_{\min}} \left(S_{\max} - S_{\min} \right) + S_{\min} \tag{9}$$

where N is the number of seeds produced, f is the compatibility of the current weed, and $f_{\max}$ and $f_{\min}$ are, respectively, the most and least compatible of the current population. $S_{\max}$ and $S_{\min}$ are the maximum and minimum possible seed production amounts, respectively.
Determination of new seeds: In this stage, the produced seeds are randomly dispersed throughout the multidimensional problem space during this stage. Usually, the random distribution function has an average value of zero and a standard deviation that fluctuates from stage to stage. Therefore, it guarantees that the randomly distributed seeds are very close to the parent weed.

The value of the standard deviation of this distribution is determined by Eq. (10):

$$\sigma_{itr} = \frac{\left(itr_{\max} - itr \right)^n}{\left(itr_{\max} \right)^n} \left(\sigma_{\text{initial}} - \sigma_{\text{final}} \right) + \sigma_{\text{final}} \tag{10}$$

In this regard, $itr_{\max}$ is the maximum number of repetitions, σ_{itr} is the standard deviation value of each stage, σ_{initial} is the standard deviation of the seed propagation around the parent weeds in the first propagation stage, and σ_{final} is the standard deviation of the seed propagation in the last propagation stage in this research (Mehrabian and Lucas 2006).

Firefly algorithm

An optimization algorithm based on collective intelligence, the firefly algorithm (FA), was introduced by Yang in 2008. FA can converge relatively quickly to an optimal or near-optimal solution. It achieves this by leveraging the communication and attraction between fireflies. The attractiveness between fireflies influences their movement, promoting convergence toward better solutions over time. FA applies to a wide range of problem sizes and complexities. It can handle both small-scale and large-scale optimization problems, making it versatile in various domains. In summary, the FA algorithm works this way; it begins by randomly dispersing several artificial oscillating fireflies within the search area. There is a relationship between the intensity of the light emitted by each firefly and the degree of optimality of its location. In order to compare the light intensities of the different fireflies, each one is constantly compared

with the light intensities of other fireflies. Therefore, the less bright firefly is attracted to the more brilliant fireflies. At the same time, the best firefly also moves randomly in the search domain to increase the chance of finding the best solution. Finally, fireflies exchange information with each other by emitting light. The combination of these group operations causes the general tendency of the fireflies to be more optimal points (Yang 2009). The calculation of the light intensity at the distance r is based on Eq. (11):

$$I = I_0 e^{-\gamma r} \tag{11}$$

Accordingly, I_0 shows the primary light intensity, and γ is a medium (such as air) that has a constant coefficient of light absorption. However, sometimes a function is needed that decreases uniformly and slower. In such a case, Eq. (12) is used:

$$I(r) = \frac{I_0}{1 + \gamma r^2} \tag{12}$$

A firefly's attractiveness is proportional to the intensity of light observed by its neighbors, so its attractiveness parameter β can be determined by Eq. (13):

$$\beta(r) = \frac{\beta_0}{1 + \gamma r^2} \tag{13}$$

where β_0 stands for the attractiveness of the firefly at a distance is $r = 0$. The motion of a firefly attracted by a brighter firefly is defined by Eq. (14):

$$X_i = X_i + \beta_0 e^{-\gamma r_{ij}^2} (X_i - j) + \alpha \left(\text{rand} - \frac{1}{2}\right) \tag{14}$$

In this regard, a significant role is played by α in the randomization process. Additionally, the rand expression generates random numbers between 0 and 1. Experiments in solving relatively complex continuous optimization problems show that even when faced with such issues, the FA algorithm can quickly and most likely find the optimal solution (Yang 2009).

GAPSO hybrid algorithm

The GAPSO algorithm is a combination of two genetic algorithms (GA) and the particle swarm optimization (PSO) algorithm, which was first introduced in 2004 by Juang (Juang 2004). Consequently, the advanced GAPSO algorithm merges the advantages of GA and PSO algorithms. By combining the sensitivity and accuracy of the genetic algorithm (GA) for population selection with the rapidity of the particle swarm optimization (PSO) in determining solutions, the newly proposed algorithm demonstrates superiority over both methods. In this hybrid algorithm, GA operators are first applied, and then PSO

operators are applied to the generated population members. In this algorithm, two types of search techniques are used: The first is a local search technique, in which the solutions obtained by each group improve their position compared to the best solution, and the second is a method of exchanging information between groups. After each local search in each group, the obtained information is compared between the groups. The advantage of this algorithm is its rapid convergence. GAPSO hybrid programming algorithm has already been successfully used in various scientific fields due to its high efficiency (Kao and Zahara 2008; Wu et al. 2009; Jeong et al. 2009). This algorithm is schematically illustrated in Fig. 3.

The mathematical formulations of the GAPSO algorithm involve the concept of particle positions and velocities being represented as chromosomes and the update equations for particle positions and velocities being updated according to t.

The update equation for the position of particle i in the GA component of GAPSO is given by:

$$X_{i,t+1} = X_i, t + M_{i,t} \tag{15}$$

Here, $M_{i,t}$ represents a random mutation that is generated based on a specified probability distribution, such as a normal distribution.

The best mutation, X_{g,t}, is selected based on the highest objective function value:

$$(X_{g,t} = \arg \max(f(X_i, t)), \ i = 1, \quad , N \tag{16}$$

In this formulation, $f(X_i, t)$ represents the objective function, and $X_{g,t}$ is the mutation with the highest objective

Fig. 3 GAPSO flowchart

function value. These formulations are iteratively applied in each iteration of the algorithm until a near-optimal solution is obtained.

The velocity update equation for particle i in the PSO algorithm is given by:

$$V_{i,t+1} = w * V_{i,t} + c1 * \text{rand1} * (P_{i,t} - X_{i,t}) \\ + c2 * \text{rand2} * (P_{g,t} - X_{i,t}) \quad (17)$$

$$X_{i,t+1} = X_i, t + V_{i,t+1} \quad (18)$$

In these equations, $V_{i,t+1}$ represents the velocity of particle i at time t, X_i, t represents the position of particle i at time t, $P_{i,t}$ represents the best position achieved by particle i up to that time, $P_{g,t}$ represents the best position achieved by any particle up to that time, and w, $c1$, and $c2$ are control parameters of the algorithm. rand1 and rand2 are random numbers between 0 and 1 used to introduce diversity in the optimization process.

Wavelet transformation

Many researchers have used wavelet transforms in recent years (since their inception in the early 1980s), and their popularity has increased (Nourani et al. 2014). Wavelet transformation is a powerful mathematical tool that offers numerous advantages in data analysis. It enables multiresolution analysis, localization in time and frequency domains, adaptability to different signal characteristics, efficient data compression, denoising capabilities, and feature extraction. In the context of rainfall analysis, wavelet analysis is particularly useful. It allows for the identification of variations at different temporal scales, detection of specific
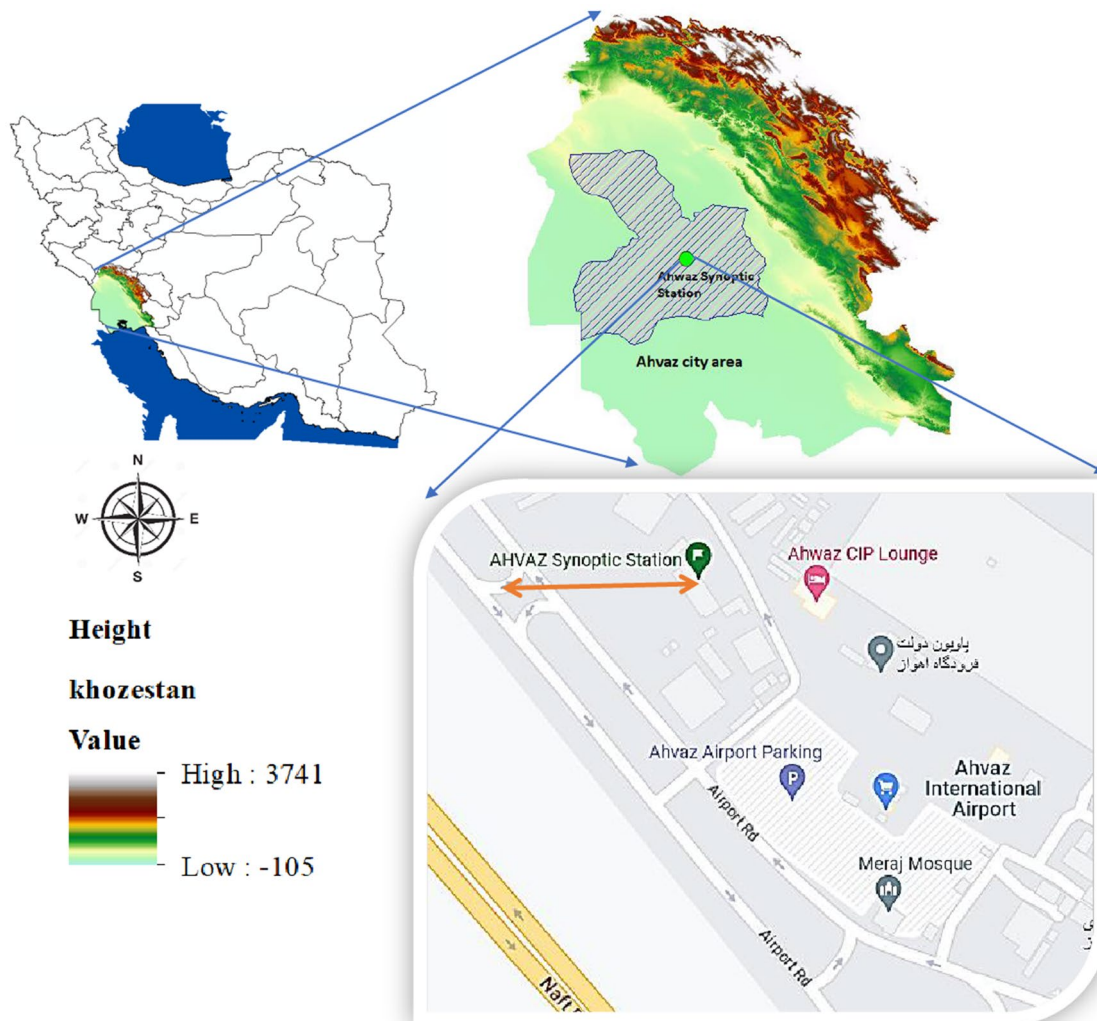

Fig. 4 Location of the Ahvaz meteorological synoptic station

rainfall events, efficient data compression, and the detection of trends and changes in rainfall patterns over time. Wavelet analysis is a versatile tool with broad applications in data analysis and processing, including rainfall analysis for hydrology, climate studies, and water resource management. A wavelet is a small wave with three characteristics: the limited number of oscillations, the fast return to zero in both positive and negative directions in its range, and the mean of zero. The wavelet function, Ψ, is called the mother wavelet. The term "mother" is used for the reason that different functions (wavelets) created based on scale and

transfer parameters are all derived from the primary function (mother wavelet) and are called daughter wavelets. In other words, the mother wavelet is the main wavelet for generating other window functions. In addition, the mother wavelet has shock characteristics and can quickly decrease to zero. The mother wavelet function is defined in the mathematical form of Eq. (19):

Table 1 Statistical characteristics of the data used

Parameter	Mean Wind speed (m/s)	Mean temperature (°C)	Relative humidity (fraction)	Solar radiation (MJ/m^2)	Evaporation	Dew spot (°C)	Soil temperature (°C)	Daily rainfall (mm)
Minimum	0.00	3.60	6.63	0.00	0.00	−9.18	−5.00	0.00
Maximum	8.13	44.00	96.38	13.00	29.00	25.98	34.00	73.70
Average	2.22	26.81	42.13	8.65	8.15	9.95	16.03	0.53
Standard deviation	1.14	9.74	19.79	3.44	5.35	5.05	8.63	3.33

Table 2 Results of the IWGM model at each level of decomposition

| Decomposition level	The selected input subseries										Wave*	Neurons in a layer	Layers
1	1	9	6	4	18	15	14	17	16	13	FK	7	3
2	11	29	23	5	27	5	20	14	1	22	FK	5	2
3	6	34	35	1	30	12	25	9	34	25	FK	3	4
4	35	42	5	27	5	43	20	15	37	43	FK	6	3
5	44	40	4	50	55	42	37	46	57	19	S	2	5
6	59	24	30	58	2	51	40	55	30	20	FK	20	2
7	42	62	66	60	41	28	50	58	67	18	FK	2	5
8	62	12	33	47	29	61	72	75	74	62	FK	17	3
9	83	72	82	78	57	6	96	26	30	55	FK	15	3
10	49	79	51	90	11	5	54	27	32	40	FK	5	2

*FK: Fejer-Korovkin, S: Symlet
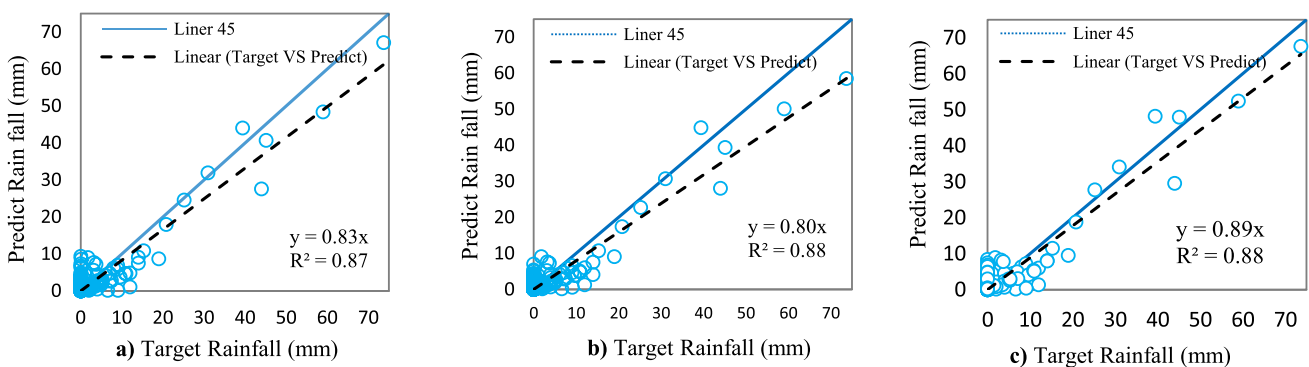
Fig. 5 Predicted versus target rainfall for the testing dataset using: **a** IWG, **b** FAWG, and **c** GPWG models

$$\int_{-\infty}^{+\infty} \psi(t)\mathrm{d}t = 0 \tag{19}$$

There are two types of wavelet transformation: 1. continuous wavelet transform (CWT) and 2. the discrete wavelet transform (DWT). Due to the absence of additional components in the transformed data, DWT is more suitable for the processing and decomposition of time series than CWT. Therefore, each frequency and time data group can be changed inversely (Toufani et al. 2011). The discrete wavelet transform of a time series, f, is defined by Eq. (20):

$$f(a.b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-b}{a}\right)\mathrm{d}t \tag{20}$$

$f(a, b)$ can simultaneously reflect the original time series features in the time b and frequency domain. In the decomposition stage, the discrete wavelet transform decomposes the time series into a set of high frequencies (detail signal) and low frequencies (approximate indication). In multi-stage decomposition, the wave decomposition continues after the first decomposition stage, with the re-decomposition of the approximate part (Merry and Steinbuch 2005; Venkata Ramana et al. 2013; Goyal 2014; Mostaghimzadeh et al. 2023). As a part of this study, CWT is employed to decompose the input time series into subseries (detail and approximation). In a subsequent stage, optimizers (IWO, FA, and GAPSO) select ten subseries and enter them into the simulators (ANN and GMDH). Developed models use six wavelets (Haar, Meyer, Symlet, Coiflet, Daubechies, and Fejer-Korovkin) at different decomposition levels (1 to 10). In the

Table 3 Results of the IWG, GPWG, and FAWG models at ten levels of decomposition

Model	IWG		FWG		GPWG	
Index	Average	best	Average	best	Average	Best
RMSE (mm)	2.0677	1.8400	2.2006	1.8371	1.7610	1.7610
MAE (mm)	0.8556	0.8059	0.9436	0.7834	0.8250	0.8250
R	0.9136	0.9319	0.8988	0.9356	0.9368	0.9368

Table 4 Results of IWG, GPWG, and FAWG models

Model	The Selected input subseries	Decomposition level										Wave[*]	Neurons	Layers
IWG	59	24	30	58	2	51	40	55	30	20	6	FK	20	2
FWG	35	18	31	34	27	16	34	7	12	30	3	FK	12	5
GPWG	92	56	37	98	90	91	107	8	41	79	4	FK	6	3

[*]FK: Fejer-Korovkin

Table 5 Best results of the GPWA model at each level of decomposition

Decomposition level	The Selected input subseries										Wave name[*]	Train Function[**]	Net[***]	Epoch
1	18	11	1	10	14	4	7	14	16	4	FK	LM	F	14
2	23	26	21	29	19	22	27	24	3	22	S	LM	F	19
3	17	32	14	35	14	22	39	36	30	35	M	LM	C	14
4	37	37	48	10	20	42	47	43	41	21	FK	BR	C	20
5	47	46	12	41	22	50	29	44	35	51	M	LM	C	22
6	6	59	58	42	24	44	18	11	4	53	M	BR	C	24
7	3	16	39	40	14	3	58	17	68	40	M	LM	C	14
8	23	24	39	87	19	65	88	84	67	67	M	LM	C	19
9	93	72	61	17	10	94	86	92	8	93	M	BR	F	10
10	93	72	61	17	10	94	86	92	8	101	C	BR	C	10

[*]FK: Fejer-Korovkin, C: Coiflet, M: Meyer, S: Symlet

[**]BR: Bayesian regularization backpropagation, LM: Levenberg–Marquardt backpropagation

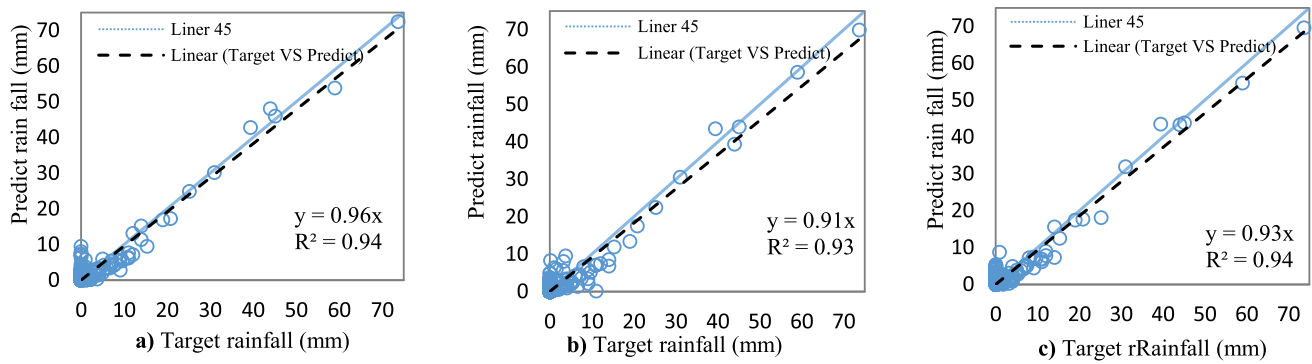[***]C: Cascade-forward backpropagation, F: Feed-forward backpropagation

Fig. 6 Predicted versus target rainfall for the testing dataset using: **a** IWA, **b** FAWA, and **c** GPWA models

Table 6 Results of the IWA, GPWA, and FAWA models at ten levels of decomposition

Model	IWA		FWA		GPWA	
Index	Average	Best	Average	Best	Average	Best
RMSE	1.4394	1.2384	1.5896	1.3790	1.3633	1.1953
MAE	0.6121	0.5742	0.6933	0.6298	0.6364	0.5731
R	0.9583	0.9694	0.9500	0.9618	0.9635	0.9715

Table 7 Results of IWG, GPWG, and FAWG models

Wave[*]	Train function[**]	Network[***]										Epoch	Wave[*]	Train function[**]	Network[***]
IWA	23	27	26	30	13	20	1	24	9	22	2	M	BR	F	12
FWA	37	37	48	10	14	42	47	43	41	36	4	FK	BR	C	20
GPWA	26	10	24	29	22	27	23	20	1	5	2	M	BR	C	10

*M: Meyer, FK: Fejer-Korovkin

**BR: Bayesian regularization backpropagation

***C: Cascade-forward backpropagation, F: Feed-forward backpropagation

proposed model, the optimizer is responsible for selecting the best subseries, mother wavelet, and decomposition level.

Validation criteria

Models are validated using various criteria, including mean absolute error (MAE), correlation coefficient (R), root-mean-square error (RMSE), probability of detection (POD), critical success index (CSI), and false alarm ratio (FAR). When using the R function, a zero value indicates no correlation between observable and predicted results, while a value of one indicates a complete correlation. Therefore, the model's performance will improve if its RMSE and MAE values are close to zero. The formulas for calculating possible POD, CSI, and POD indices are as relations 21–23:

$$POD = \frac{Hits}{Hits + Misses} \tag{21}$$

$$CSI = \frac{Hits}{Hits + Misses + False\ Alarm} \tag{22}$$

$$FAR = \frac{False\ Alarm}{Hits + False\ Alarm} \tag{23}$$

In terms of *Hits*, this refers to the number of days in which the output of the models predicts precipitation, and the observations indicate that rainfall has also occurred. The number of *Misses* is the number of days when rain is observed, but the models have predicted no rain. Finally, *False Alarm* is a number of days observed without rain but predicted rain by the models. (Chokngamwong and Chiu 2008; Golian et al. 2011).
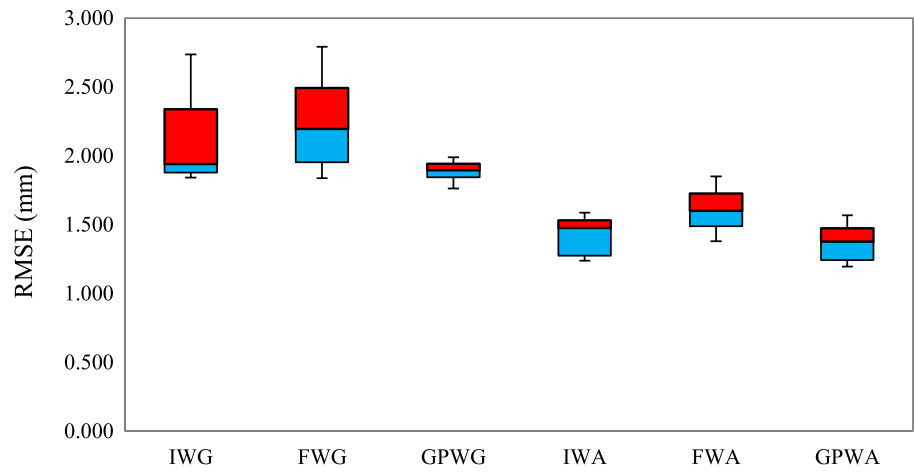
Fig. 7 Boxplot of models results

Fig. 8 Cumulative graph of error with different levels of decomposition
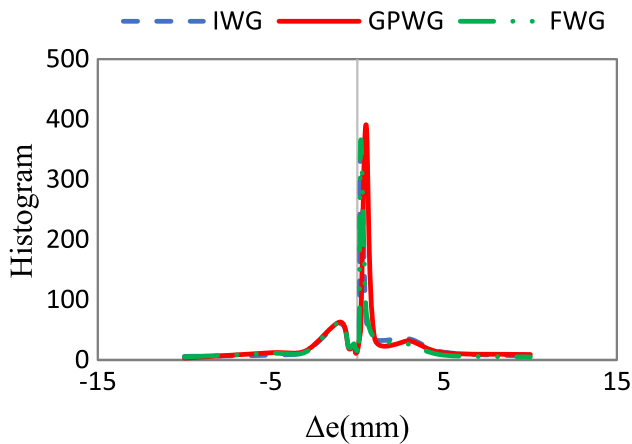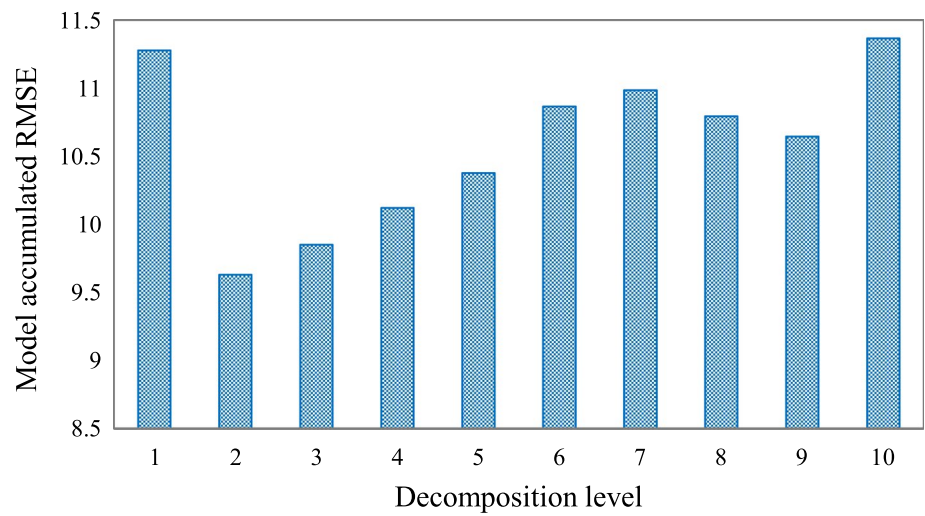

Fig. 9 Prediction error histogram in IWG, GPWG and FWG models
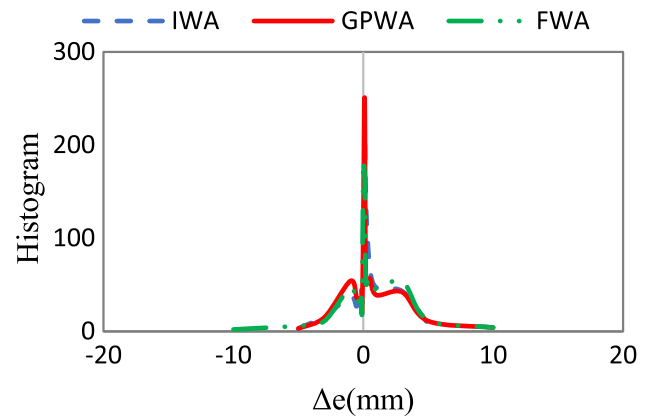

Fig. 10 Prediction error histogram in IWA, GPWA and FWA models

Study area

For this study, the meteorological data of the Ahvaz synoptic weather station, which is located at a latitude of 31.345 m, a longitude of 48.744 m, and an altitude of 22.5 m above sea level, were used to predict urban rainfall. The city of Ahvaz,

Table 8 Error distribution of intelligent models

Index/model	IWG	FWG	GPWG	IWA	FWA	GPWA
Q25	0.061	0.017	0.064	−0.049	−0.073	−0.070
Q50	0.185	0.188	0.213	0.085	0.008	0.029
Q75	0.186	0.189	0.214	0.187	0.164	0.140
Mean	−0.067	−0.156	−0.036	0.079	−0.012	−0.039
STD	1.839	1.830	1.761	1.236	1.379	1.195

the capital of Khuzestan province and the seventh most populous in Iran, is situated in a plain region at 31 degrees 30 min north latitude and 45 degrees 65 min east longitude, with an altitude of 22 m. In Fig. 4, the location of the Ahvaz synoptic station is shown. Input time series include average daily temperatures, average daily wind speeds, average daily relative humidity, solar radiation, evaporation, dew point, soil temperatures, and precipitation with a 1- to 3-day delay, and output data include daily rainfall. Between January 1, 2010, and December 31, 2019, 3652 data samples were collected. Table 1 presents the statistical characteristics of the data used in this study.

Results and discussion

IWG, FWG, and GPWG models

In the hybrid models of IWG, FWG, and GPWG, GMDH simulator models are used. The number of neurons, layers and inputs is optimized using optimization algorithms IWO, FA, and GAPSO. For example, Table 2 shows the best results of the IWGA hybrid model at different decomposition levels. As shown in Table 2, the most optimal set of ten subseries of the model input, the most optimal
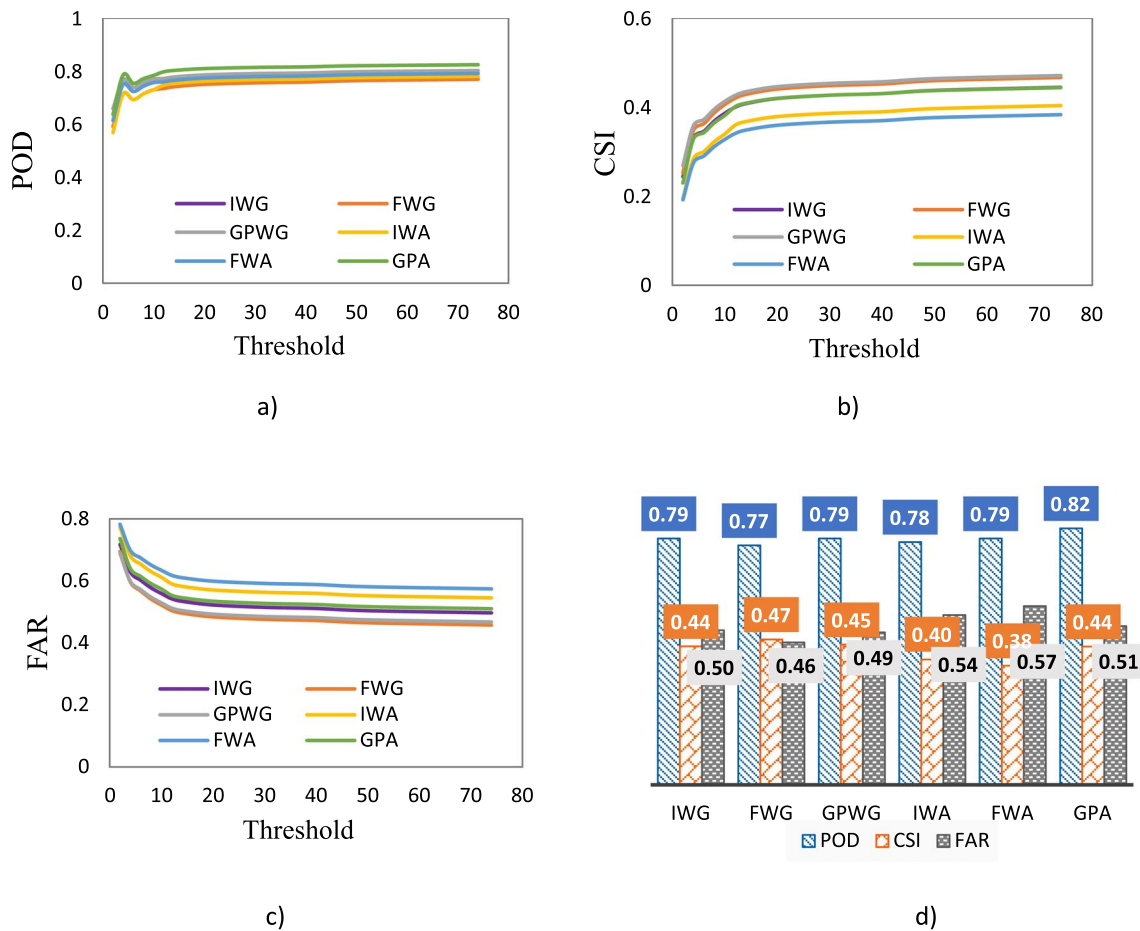
Fig. 11 Evaluation of indicators: **a** POD, **b** CSI, **c** FAR, **d** General diagram of POD, CSI, and FAR
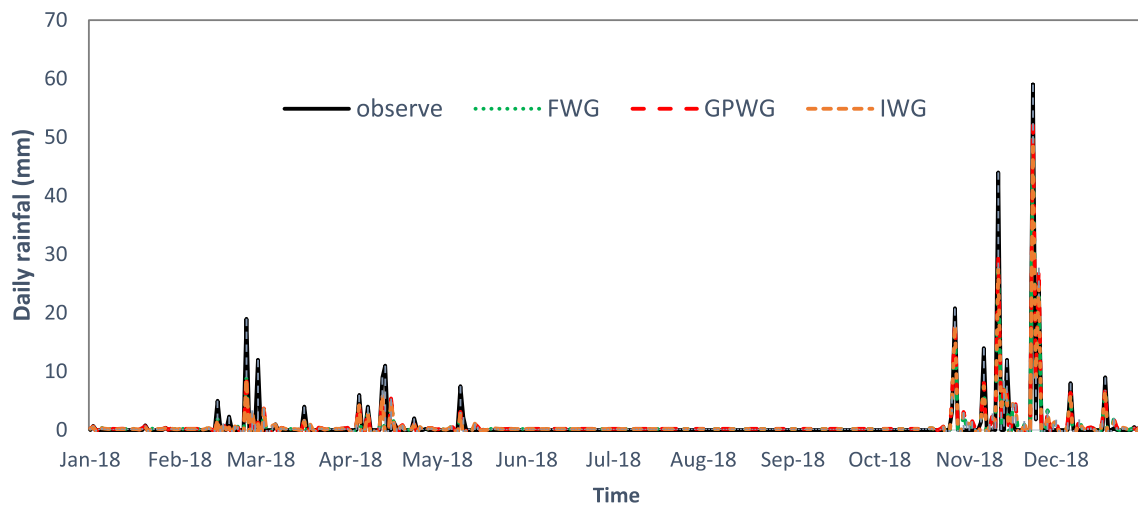
(a)
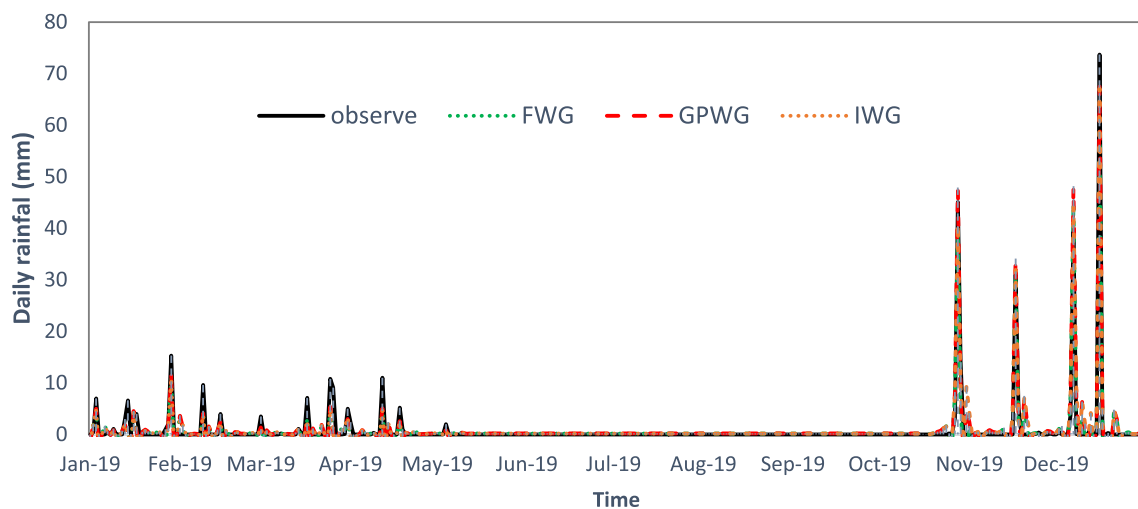
(b)

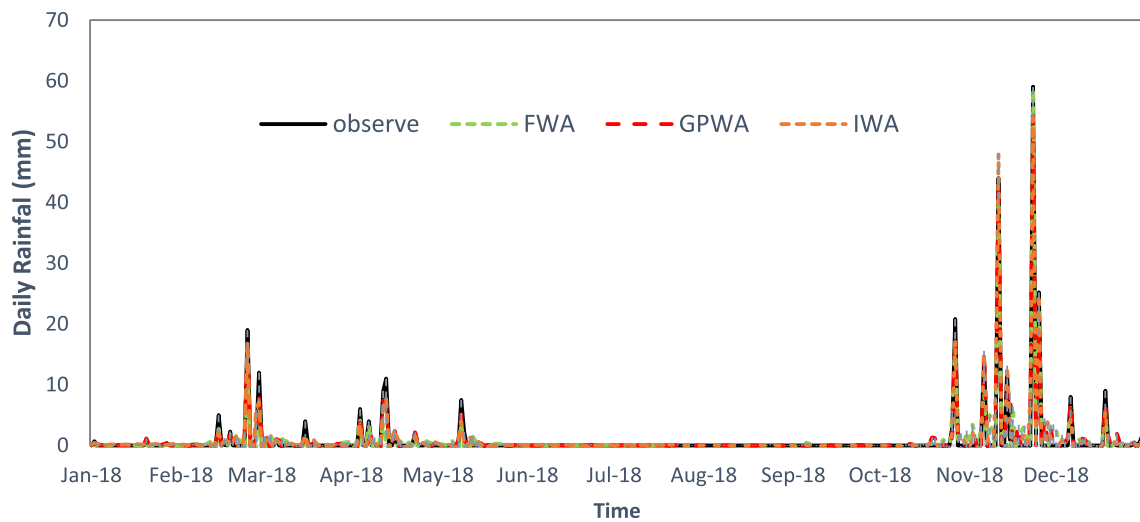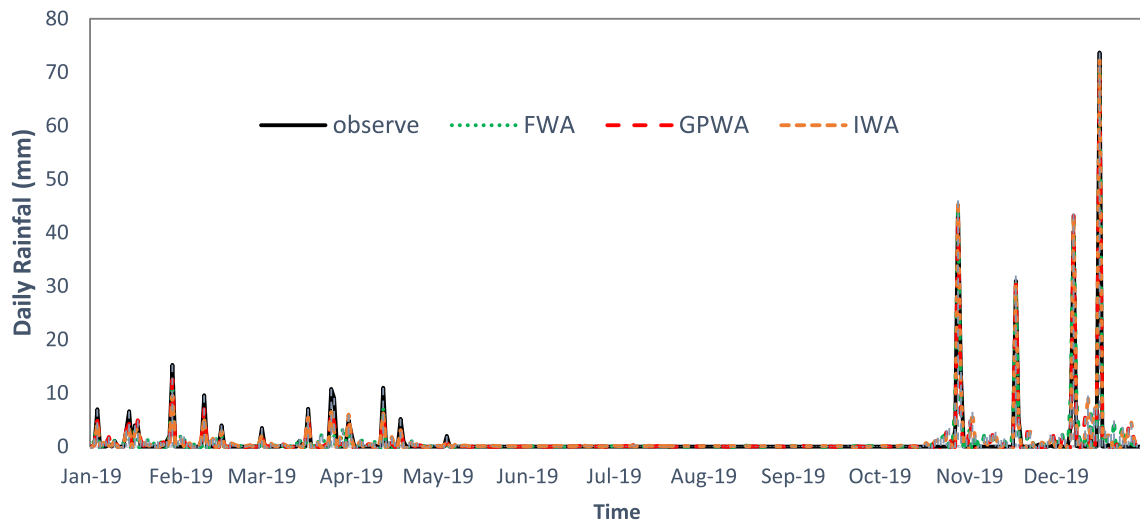**Fig. 12** **a** Prediction results of **I**WG, FWG, and GPWG models versus observational data (the year 2018) (b) Prediction results of **I**WG, FWG, and GPWG models versus observational data (the year 2019)

mother wavelet, the most optimal number of neurons, and the most optimal number of GMDH simulator layers are displayed at each level of decomposition. The scatter diagram in Fig. 5 and the final results of each artificial intelligence prediction model are presented in Tables 3 and 4. In comparing the average results of each model, GPWG has a better overall performance than IWG and FWG models with RMSE values of 2 mm, MAE equal to 0.8442 mm, and R equivalent to 0.92 (Table 3). Moreover, the best results are related to the same model with decomposition level 3, Fejer-Korovkin mother wavelet, and the number of 12 neurons in 5 layers, which has the lowest RMSE,

equivalent to 1.7610 mm. For each model, according to Fig. 5, the linear fit equation indicates that the coefficient of determination exceeds $R^2 = 86$, and the angle between the linear fit and the 45-degree line is 17 degrees for the IWG, 20 degrees for the FWG, and 11 degrees for the GPWG models. In other words, all three models performed appropriately within 80% confidence intervals, whereas the GPWG model performed within an 89% confidence bond. All three models have been developed using the same simulator, and the GPWG model is superior to the other two due to the use of the GP optimizer, which combines GA and PSO features to achieve better results.

(a)

(b)

Fig. 13 **a** Prediction results of IWA, FWA, and GPWA models versus observational data (the year 2018) **b** Prediction results of **I**WA, FWA, and GPWA models versus observational data (the year 2019)

IWA, FWA, and GPWA models

For developing the hybrid models of IWA, FWA, and GPWA, neural networks have been used in both feed-forward and cascade-feed-forward backpropagation modes, with different training functions and training epochs from 1 to 30 as determined by the optimizer. As shown in Table 5, the GPWA model results include the best set of ten decomposed subseries, the best mother wavelet, the best network function, and the epoch displayed at each decomposition level. The scatter diagram in Fig. 6 and the final results of each three artificial intelligence prediction models are presented in Tables 6 and 7.

Comparing the average results of each model, GPWA with RMSE values of 1.19 mm, MAE equal to 0.57 mm, and R equal to 0.97 perform better than IWG and FWG (Table 6). For all three models, the linear fit equation indicates that R^2 is greater than 92%, and the angle between the linear fit and the 45-degree line is 4 degrees for IWA, 9 degrees for FWA, and 7 degrees for GPWA, respectively. These results are within a 90% confidence band, and all three models demonstrate appropriate performance. It is evident from Table 6 that the most optimal model is GPWA, which has RMSE values of 1.19 mm, MAE values of 0.57 mm, and R values of 0.9785. This model is optimized in level 2 decomposition

Fig. 14 Comparison of the proposed model of the present research with previous research
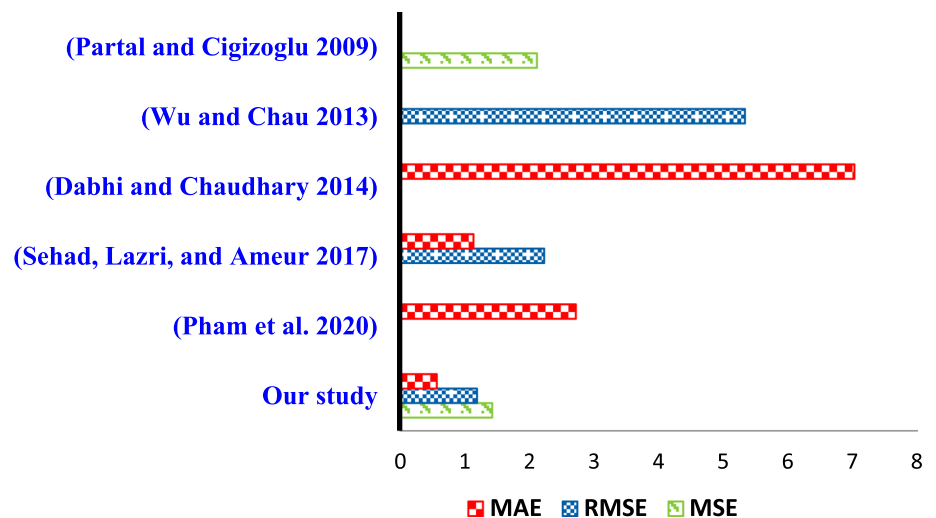
using a Fejer-Korovkin wavelet and a cascade-feet-forward backpropagation network with ten epochs.

This study has developed six artificial intelligence models named IWG, FWG, GPWG, IWA, FWA, and GPWA to predict daily rainfall. The error boxplots of the models are shown in Fig. 6 at ten different decomposition levels. GPWG and GPWA models provide similar and more stable results at different decomposition levels, as shown in Fig. 6. Figure 7 illustrates the cumulative error value of the models based on the decomposition level. According to Fig. 7, cumulative error increases in the beginning levels and decreases in levels eight and nine, but as more decomposition is performed, the error rates of the models increase again. Levels 2 and 3 are the best decomposition levels with the lowest cumulative error rate (Fig. 8).

Figures 9 and 10 show the histograms of the models' prediction difference error (Δe) and the relevant statistical data in Table 8. The formula Δe can be calculated by using Eq. (24):

$$\Delta e = \text{Predicted} - \text{Target} \tag{24}$$

According to the information in Figs. 9, 10, and Table 8, IWA, GPWA, and FWA models generally have lower error values than IWG, GPWG, and FWG models in all indicators. GPWA and FWA models have the lowest error, with an average of -0.039 and -0.012, respectively. In the first quadrant, the FWG model, in the second or middle quadrant, the FWA model, and third quadrant, the GPWA model have the lowest error rate, close to zero. Also, the GPWA model has the lowest error standard deviation. Considering all the indicators in Table 8, the GPWA model performs best with the lowest error, close to zero. The main reason for the superiority of the GPWA model, in addition to the use of the artificial neural network (which has better results compared

to GMDH), is the GAPSO hybrid optimizer. As optimizers play a significant role in this research (selection and control of all forecasting stages), GAPSO performs better than other optimizers by combining the advantages of GA and PSO optimizers.

Artificial intelligent models were also evaluated using probability indicators (POD, CSI, and FAR). The daily rainfall threshold ranges from 2 to 74 mm. Figure 11 shows the evaluation results of CSI, POD, and FAR. It can be seen from Fig. 11 that as the precipitation threshold increases, the CSI and POD values increase, while the FAR value decreases. Therefore, based on the POD index, the GPWA model performs better, with a value of 0.82, while based on the CSI and FAR indices, the GPWG and FWG models perform better, respectively, with values of 0.38 and 0.46. Therefore, as a result of the evaluation based on these indicators, it is concluded that all artificial intelligence models have accurate predictions.

As mentioned, the results have been verified using 20% of the available data (2018–2019). In Figs. 12, 13, 14, the prediction results of the intelligent models developed in this research are plotted against the observational data.

One of the capabilities of rainfall forecasting models is the ability to transfer from non-rainy days to rainy days. From May to November, there are often no daily rainfalls in the region under study, which makes the ability to transfer the model in this region vital. The results shown in Figs. 12 and 13 indicate that the developed models demonstrate excellent performance in transitioning from non-rainy to rainy days despite the sudden changes in rainfall. A critical factor in this can be found in the optimal selection of the inputs, optimal decomposition level, and mother wavelet by the model's optimization part, which significantly impacts predicting the critical points of the

rainfall time series. According to Figs. 12 and 13, it is clear that IWA, FWA, and GPWA models have a better performance in predicting rainfall than IWG, FWG, and GPWG models, especially at peak points. Accurate prediction of peak values is a crucial and difficult stage of intelligent water resources management (Mostaghimzadeh et al. 2022).

According to the current research, GPWA was the best model for forecasting daily rainfall based on its RMSE value of 1.19 mm and R-value of 97.15, as well as its superiority over other models on the indicators evaluated. The proposed GPWA model results were compared with previous research for further evaluation. The results presented in Fig. 14 indicate that the proposed model performs better than other research in predicting daily rainfall. The significant advantage of the presented models lies in their automatic and optimal selection of the inputs, decomposition level, mother wavelet, and simulators' effective parameters.

There are some limitations to the models that are being used in current research; it is because they are based on historical data and assumptions. In this way, the ability of these instruments to record events related to severe or unprecedented weather is limited. Moreover, regional or temporal changes in rainfall patterns can affect models' performance, underscoring the need for continuous validation and adaptation to local conditions. Future climate change may also affect model performance, indicating the need to combine current models with climate change models in future research.

Conclusion

In this study, six artificial intelligence models were developed (IWG, FWG, GPWG, IWA, FWA, and GPWA) by combining FA, IWO, and GPSO optimization algorithms with ANNs and GMDH simulators and wavelet transformation to predict daily rainfall. Also, they are used to predict the daily rainfall of Ahvaz city, Iran, to measure their efficiency. The significant advantage of the presented models is their automatic and optimal selection of the inputs, decomposition level, mother wavelet, and simulators' effective parameters to reduce RMSE, which are handled by optimizers that, like an intelligent brain, enhance prediction accuracy by controlling all the model components. The results of the models were validated using RMSE, MAE, R, POD, CSI, and FAR indicators. Our finding shows the proposed models are highly appropriate for predicting rainfall. The GPWA model has the highest accuracy and stability of the prediction results. Furthermore, based on the daily rainfall forecasts for Ahvaz city provided by the GPWA model for 2018 and 2019, the model can effectively transfer between non-rainy and rainy days and predict peak points more accurately. Thus, the GPWA model

is the best current research model recommended for predicting daily rainfall. To facilitate the development of models, it is recommended to forecast rainfall data over an extended duration. Additionally, it is feasible to integrate existing research models with simulations of urban events, specifically focusing on the prediction of urban floods, which will be our future study. Finally, it is important to acknowledge the limitations of these models. One limitation is the reliance on historical data and assumptions, which may restrict their performance in capturing unprecedented or extreme weather events. Additionally, the models' performance could be affected by regional or temporal variations in rainfall patterns, emphasizing the need for continuous validation and adaptation to local conditions. Further research and development are recommended to address these limitations and enhance the models' applicability for urban flood prediction.

Declarations

References

Abrahart RJ, See L (2000) Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrol Process 14(11–12):2157–2172

Abudu S, Cui CL, King JP, Abudukadeer K (2010) Comparison of performance of statistical models in forecasting monthly streamflow of Kizil River. China Water Sci. Eng. 3(3):269–281. https://doi.org/10.3882/j.issn.1674-2370.2010.03.003

Adib A, Soori Damirchi Sofla M, Mirzaei SY, Mahmoudian Shoushtari M, Liaghat A (2021a) Calibration of seasonal transfer equation (Z–R) by data of Doppler weather radar, rainfall gauging station and genetic algorithm method in the Abolabbas watershed (in southwest of Iran). Water Supply 21(2):567–580. https://doi.org/10.2166/ws.2020.328

Adib A, Zaerpour A, Kisi O, Lotfirad M (2021b) A rigorous wavelet-Packet transform to retrieve snow depth from SSMIS data and evaluation of its reliability by uncertainty parameters. Water Resour Manage 35:2723–2740

Altunkaynak A, Nigussie TA (2015) Prediction of daily rainfall by a hybrid wavelet-season-neuro technique. J Hydrol 529:287–301. https://doi.org/10.1016/j.jhydrol.2015.07.046

Alves E, Lopes M, Sales F, Balieiro A, Leal A (2023) A GMDH approach for forecast monthly rainfall in southeast of pará. IEEE Lat Am Trans 21(6):707–714. https://doi.org/10.1109/TLA.2023.10172135

Amanifard N, Nariman-Zadeh N, Farahani MH, Khalkhali A (2008) Modelling of multiple short-length-scale stall cells in an axial compressor using evolved GMDH neural networks. Energy Convers Manage 49(10):2588–2594. https://doi.org/10.1016/j.enconman.2008.05.025

Amiri M, Soleimani S (2021) ML-based group method of data handling: an improvement on the conventional GMDH. Complex Intell Syst 7(6):2949–2960. https://doi.org/10.1007/s40747-021-00480-0

Ashrafi SM, Mostaghimzadeh E, Adib A (2020) Applying wavelet transformation and artificial neural networks to develop forecasting-based reservoir operating rule curves. Hydrol Sci J 65(12):2007–2021. https://doi.org/10.1080/02626667.2020.1784902

Azadi M, Taghizadeh E, Memarian MH, Dmitrieva-Arrago LR (2013) Comparing the results of precipitation forecast based on mesoscale models on the territory of Iran during the cold season. Russ Meteorol Hydrol 38(9):605–613. https://doi.org/10.3103/S1068373913090033

Chokngamwong R, Chiu LS (2008) Thailand daily rainfall and comparison with TRMM products. J Hydrometeorol 9(2):256–266

Dabhi VK, Chaudhary S (2014) Hybrid wavelet-postfix-GP model for rainfall prediction of anand region of India. Adv Artif Intell 2014:3–3

Danladi A, Stephen M, Aliyu BM, Gaya GK, Silikwa NW, Machael Y (2018) Assessing the influence of weather parameters on rainfall to forecast river discharge based on short-term. Alex Eng J 57(2):1157–1162. https://doi.org/10.1016/j.aej.2017.03.004

Daubechies I (1990) The wavelet transform, time-frequency localization and signal analysis. IEEE Trans Inf Theory 36(5):961–1005. https://doi.org/10.1109/18.57199

DelSole T, Shukla J (2002) Linear prediction of Indian monsoon rainfall. J Clim 15(24):3645–3658. https://doi.org/10.1175/1520-0442(2002)015%3c3645:LPOIMR%3e2.0.CO;2

El-Shafie AH, El-Shafie A, El-Mazoghi HG, Shehata A, Taha MR (2011) Artificial neural network technique for rainfall forecasting applied to Alexandria. Egypt Int J Phys Sci 6(6):1306–1316. https://doi.org/10.5897/IJPS11.143

Farajpanah H, Lotfirad M, Adib A, Esmaeili-Gisavandani H, Kisi Ö, Riyahi MM, Salehpoor J (2020) Ranking of hybrid wavelet-AI models by TOPSIS method for estimation of daily flow discharge. Water Supply 20(8):3156–3171. https://doi.org/10.2166/ws.2020.211

Gökbulak F, Şengönül K, Serengil Y, Yurtseven İ, Özhan S, Cizigoglu HK, Uygur B (2015) Comparison of rainfall-runoff relationship modeling using different methods in a forested watershed. Water Resour Manage 29(12):4229–4239. https://doi.org/10.1007/s11269-015-1056-1

Golian S, Saghafian B, Elmi M, Maknoon R (2011) Probabilistic rainfall thresholds for flood forecasting: evaluating different methodologies for modelling rainfall spatial correlation (or dependence). Hydrol Process 25(13):2046–2055. https://doi.org/10.1002/hyp.7956

Gouda SG, Hussein Z, Luo S, Yuan Q (2020) Review of empirical solar radiation models for estimating global solar radiation of various climate zones of China. Progress Phys Geograp Earth Environ 44(2):168–188. https://doi.org/10.1177/0309133319867213

Goyal MK (2014) Monthly rainfall prediction using wavelet regression and neural network: an analysis of 1901–2002 data, Assam, Indian. Theor Appl Climatol 118(1):25–34. https://doi.org/10.1007/s00704-013-1029-3

Green DG, Reichelt RE, Bradbury RH (1988) Statistical behaviour of the GMDH algorithm. Biometrics. https://doi.org/10.2307/2531895

Haddad MS (2011) Capacity choice and water management in hydro-electricity systems. Energy Econom 33(2):168–177. https://doi.org/10.1016/j.eneco.2010.05.005

Hung NQ, Babel MS, Weesakul S, Tripathi NK (2009) An artificial neural network model for rainfall forecasting in Bangkok. Thailand Hydrol Earth Syst Sci 13(8):1413–1425. https://doi.org/10.5194/hess-13-1413-2009

Ivakhnenko AG (1968) The group method of data handling, a rival of the method of stochastic approximation. Soviet Autom Control 13(3):43–55

Jeong S, Hasegawa S, Shimoyama K, Obayashi S (2009) Development and investigation of efficient GA/PSO-hybrid algorithm applicable to real-world design optimization. IEEE Comput Intell Mag 4(3):36–44. https://doi.org/10.1109/MCI.2009.933099

Juang CF (2004) A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. IEEE Trans Syst Man Cyber Part B 34(2):997–1006. https://doi.org/10.1109/TSMCB.2003.818557

Kao YT, Zahara E (2008) A hybrid genetic algorithm and particle swarm optimization for multimodal functions. Appl Soft Comput 8(2):849–857. https://doi.org/10.1016/j.asoc.2007.07.002

Khalili N, Khodashenas SR, Davary K, Baygi MM, Karimaldini F (2016) Prediction of rainfall using artificial neural networks for synoptic station of Mashhad: a case study. Arab J Geosci 9(13):1–9. https://doi.org/10.1007/s12517-016-2633-1

Kuligowski RJ, Barros AP (1998) Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. Weather Forecast 13(4):1194–1204. https://doi.org/10.1175/1520-0434(1998)013%3c1194:LPFFAN%3e2.0.CO;2

Lake RW, Shaeri S, Senevirathna STMLD (2022) Review of the limitations and potential empirical improvements of the parametric group method of data handling for rainfall modelling. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-022-23194-3

Le TT, Pham BT, Ly HB, Shirzadi A, Le LM (2020) Development of 48-hour precipitation forecasting model using nonlinear autoregressive neural network. In: CIGOS 2019, innovation for sustainable infrastructure. Springer, Singapore. pp 1191–1196. https://doi.org/10.1007/978-981-15-0802-8_191

Li Y, Shi H, Liu H (2020) A hybrid model for river water level forecasting: cases of Xiangjiang River and Yuanjiang River, China. J Hydrol 587:124934. https://doi.org/10.1016/j.jhydrol.2020.124934

Luk KC, Ball JE, Sharma A (2000) A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. J Hydrol 227(1–4):56–65. https://doi.org/10.1016/S0022-1694(99)00165-1

Matsumura R, Harada K, Domae Y, Wan W (2019) Learning based industrial bin-picking trained with approximate physics simulator. In: Intelligent autonomous systems 15: proceedings of the 15th international conference IAS-15. Springer International Publishing. pp 786–798, https://doi.org/10.1007/978-3-030-01370-7_61

Mehrabian AR, Lucas C (2006) A novel numerical optimization algorithm inspired from weed colonization. Eco Inform 1(4):355–366. https://doi.org/10.1016/j.ecoinf.2006.07.003

Merry RJE, Steinbuch M (2005) Wavelet theory and applications. Literature study, Eindhoven university of technology, Department of mechanical engineering, Control systems technology group.

Moazami S, Abdollahipour A, Zakeri Niri M, Ashrafi SM (2016) Hydrological assessment of daily satellite precipitation products over a basin in Iran. J Hydraul Struct 2(2):35–45. https://doi.org/10.22055/JHS.2016.12850

Mohseni U, Muskula SB (2023) Rainfall-runoff modeling using artificial neural network—a case study of purna sub-catchment of Upper Tapi Basin, India. Environ Sci Proc 25(1):1. https://doi.org/10.3390/ECWS-7-14232

Mostaghimzadeh E, Ashrafi SM, Adib A, Geem ZW (2021) Investigation of forecast accuracy and its impact on the efficiency of data-driven forecast-based reservoir operating rules. Water 13(19):2737. https://doi.org/10.3390/w13192737

Mostaghimzadeh E, Adib A, Ashrafi SM, Kisi O (2022) Investigation of a composite two-phase hedging rule policy for a multi reservoir system using streamflow forecast. Agric Water Manag 265:107542. https://doi.org/10.1016/j.agwat.2022.107542

Mostaghimzadeh E, Ashrafi SM, Adib A, Geem ZW (2023) A long lead time forecast model applying an ensemble approach for managing the great Karun multi-reservoir system. Appl Water Sci 13(6):124. https://doi.org/10.1007/s13201-023-01924-3

Narawi A, Abang Jawawi DN, Samsudin R (2022) Rainfall forecasting using the group method of data handling model: a case study of sarawak, Malaysia. In: International conference of reliable information and communication technology. Springer, Cham. pp 129–140, https://doi.org/10.1007/978-3-030-98741-1_12

Narayanakumar S, Raja K (2016) A BP artificial neural network model for earthquake magnitude prediction in Himalayas, India. Circ. Syst 7(11):3456–3468. https://doi.org/10.4236/cs.2016.711294

Nastos PT, Moustris KP, Larissi IK, Paliatsos AG (2013) Rain intensity forecast using artificial neural networks in Athens, Greece. Atmos Res 119:153–160. https://doi.org/10.1016/j.atmosres.2011.07.020

Nayagam LR, Janardanan R, Mohan HR (2008) An empirical model for the seasonal prediction of southwest monsoon rainfall over Kerala, a meteorological subdivision of India. Int J Climatol A J Royal Meteorol Soc 28(6):823–831. https://doi.org/10.1002/joc.1577

Nayak P, Venkatesh B, Krishna B, Jain SK (2013) Rainfall-runoff modeling using conceptual, data driven, and wavelet based computing approach. J Hydrol 493:57–67. https://doi.org/10.1016/j.jhydrol.2013.04.016

Nikolaev NY, Iba H (2003) Polynomial harmonic GMDH learning networks for time series modeling. Neural Netw 16(10):1527–1540. https://doi.org/10.1016/S0893-6080(03)00188-6

Nourani V (2017) An emotional ANN (EANN) approach to modeling rainfall-runoff process. J Hydrol 544:267–277. https://doi.org/10.1016/j.jhydrol.2016.11.033

Nourani V, Komasi M, Mano A (2009) A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resour Manage 23(14):2877–2894. https://doi.org/10.1007/s11269-009-9414-5

Nourani V, Baghanam AH, Adamowski J, Kisi O (2014) Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. J Hydrol 514:358–377. https://doi.org/10.1016/j.jhydrol.2014.03.057

Novak DR, Bailey C, Brill KF, Burke P, Hogsett WA, Rausch R, Schichtel M (2014) Precipitation and temperature forecast performance at the weather prediction center. Weather Forecast 29(3):489–504. https://doi.org/10.1175/WAF-D-13-00066.1

Olson DA, Junker NW, Korty B (1995) Evaluation of 33 years of quantitative precipitation forecasting at the NMC. Weather Forecast 10(3):498–511. https://doi.org/10.1175/1520-0434(1995)010%3c0498:EOYOQP%3e2.0.CO;2

Partal T, Cigizoglu HK (2009) Prediction of daily precipitation using wavelet—neural networks. Hydrol Sci J 54(2):234–246. https://doi.org/10.1623/hysj.54.2.234

Partal T, Kişi Ö (2007) Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. J Hydrol 342(1–2):199–212. https://doi.org/10.1016/j.jhydrol.2007.05.026

Pham BT, Le LM, Le TT, Bui KTT, Le VM, Ly HB, Prakash I (2020) Development of advanced artificial intelligence models for daily rainfall prediction. Atmos Res 237:104845. https://doi.org/10.1016/j.atmosres.2020.104845

Qiu M, Zhao P, Zhang K, Huang J, Shi X, Wang X, Chu W (2017) A short-term rainfall prediction model using multi-task convolutional neural networks. In: 2017 IEEE international conference on data mining (ICDM). IEEE. pp 395–404, https://doi.org/10.1109/ICDM.2017.49

Ramirez MCV, de Campos Velho HF, Ferreira NJ (2005) Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. J Hydrol 301(1–4):146–162. https://doi.org/10.1016/j.jhydrol.2004.06.028

Samsudin R, Saad P, Shabri A (2010) A hybrid least squares support vector machines and GMDH approach for river flow forecasting. Hydrol Earth Syst Sci Discuss 7(3):3691–3731

Sedki A, Ouazar D, El Mazoudi E (2009) Evolving neural network using real coded genetic algorithm for daily rainfall–runoff forecasting. Expert Syst Appl 36(3):4523–4527. https://doi.org/10.1016/j.eswa.2008.05.024

Shuman FG (1989) History of numerical weather prediction at the national meteorological center. Weather Forecast 4(3):286–296. https://doi.org/10.1175/1520-0434(1989)004%3c0286:HONWPA%3e2.0.CO;2

Silvestro F, Rebora N (2014) Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain. J Hydrol 519:1052–1067. https://doi.org/10.1016/j.jhydrol.2014.07.042

Toufani P, Mosaedi A, Fakheri FA (2011) Prediction of precipitation applying wavelet network model (case study: Zarringol station, Golestan province, Iran). https://doi.org/10.22067/jsw.v0i-.11260

Trinh TA (2018) The impact of climate change on agriculture: findings from households in Vietnam. Environ Resource Econ 71(4):897–921. https://doi.org/10.1007/s10640-017-0189-5

Tsai TM, Yen PH (2017) GMDH algorithms applied to turbidity forecasting. Appl Water Sci 7(3):1151–1160. https://doi.org/10.1007/s13201-016-0458-4

Venkata Ramana R, Krishna B, Kumar SR, Pandey NG (2013) Monthly rainfall prediction using wavelet neural network analysis. Water Resour Manage 27:3697–3711. https://doi.org/10.1007/s11269-013-0374-4

Wang W, Du Y, Chau K, Chen H, Liu C, Ma Q (2021) A Comparison of BPNN, GMDH, and ARIMA for Monthly Rainfall Forecasting Based on Wavelet Packet Decomposition. Water 13(20):2871. https://doi.org/10.3390/w13202871

Wu CL, Chau KW (2013) Prediction of rainfall time series using modular soft computingmethods. Eng Appl Artif Intell 26(3):997–1007. https://doi.org/10.1016/j.engappai.2012.05.023

Wu X, Wang Y, Zhang T (2009) An improved GAPSO hybrid programming algorithm. In: 2009 International conference on information engineering and computer science. IEEE. pp 1–4, https://doi.org/10.1109/ICIECS.2009.5365983

Yang XS (2009) Firefly algorithms for multimodal optimization. In: Stochastic algorithms: foundations and applications: 5th international symposium, SAGA 2009, Sapporo, Japan, October 26–28, 2009. Proceedings 5. Springer Berlin Heidelberg. pp 169–178, https://doi.org/10.1007/978-3-642-04944-6_14

Zhang H, Liu X, Cai E, Huang G, Ding C (2013) Integration of dynamic rainfall data with environmental factors to forecast debris flow using an improved GMDH model. Comput Geosci 56:23–31. https://doi.org/10.1016/j.cageo.2013.02.003