Zubaidi, S, Al-Bdairi, NSS, Ortega Martorell, S, Ridha, HM, Al-Bugharbee, H, Hashim, KS and Gharghan, SK

Assessing the Benefits of Nature-Inspired Algorithms for the Parameterisation of ANN in the Prediction of Water Demand

Article

Assessing the Benefits of Nature-Inspired Algorithms for the Parameterization of ANN in the Prediction of Water Demand

Salah L. Zubaidi[1]; Nabeel Saleem Saad Al-Bdairi[2]; Sandra Ortega-Martorell[3]; Hussein Mohammed Ridha[4]; Nadhir Al-Ansari[5]; Hussein Al-Bugharbee[6]; Khalid Hashim[7]; and Sadik Kamel Gharghan[8]

Abstract: Accurate forecasting techniques for a stochastic pattern of water demand are essential for any city that faces high variability in climate factors and a shortage of water resources. This study was the first research to assess the impact of climatic factors on urban water demand in Iraq, which is one of the hottest countries in the world. We developed a novel forecasting methodology that includes data preprocessing and an artificial neural network (ANN) model, which we integrated with a recent nature-inspired metaheuristic algorithm [marine predators algorithm (MPA)]. The MPA-ANN algorithm was compared with four nature-inspired metaheuristic algorithms. Nine climatic factors were examined with different scenarios to simulate the monthly stochastic urban water demand over 11 years for Baghdad City, Iraq. The results revealed that (1) precipitation, solar radiation, and dew point temperature are the most relevant factors; (2) the ANN model becomes more accurate when it is used in combination with the MPA; and (3) this methodology can accurately forecast water demand considering the variability in climatic factors. These findings are of considerable significance to water utilities in planning, reviewing, and comparing the availability of freshwater resources and increasing water requests (i.e., adaptation variability of climatic factors). **DOI: 10.1061/(ASCE)WR.1943-5452.0001602.**

Author keywords: Baghdad City; Climatic factors; Machine learning; Metaheuristic algorithm; Water demand model.

Introduction

Secure clean water availability, quantity, and quality, for all inhabitants under the variability in climate change is fundamental to a resilient environment in modern cities (Tortajada et al. 2019). Freshwater scarcity has appeared as a global challenge because of the impact of climate change and socioeconomic factors. It led to an imbalance between water delivered and water demanded. Worldwide, more than 1 billion individuals lack access to safe potable water (Ahmadi et al. 2020).

Several studies conducted in different areas have shown that the magnitude and pattern of precipitation differ as a result of climate change (Szeląg et al. 2021). The high variability of climate change imposes an increasing challenge for the management of freshwater resources (i.e., due to reduced freshwater availability) (Nunes Carvalho et al. 2021), which highlights the increasing need for protecting the quantity and quality of water resources, particularly in sensitive zones (Lama et al. 2021a, b). Management of municipal water planning considering a nuanced quantitative understanding of water needs is fundamental for solving the problem of water security (Capt et al. 2021). Accordingly, forecasting municipal water consumption in the future with greater precision is essential when designing water distribution networks (Pandey et al. 2021).

In Iraq, located in the fastest-warming area of the world, the temperature reaches 54°C, which is considered one of the highest temperatures ever measured in the Eastern Hemisphere (Salman et al. 2018). Iraq depends on the Tigris and Euphrates Rivers as primary freshwater resources, which originate outside the Iraqi border with Turkey. The discharge rate of these rivers has reduced to less than a third of normal capacity because of the water policies in Turkey, Iran, and Syria. Moreover, investments in industries after 2003 (e.g., the oil industry) has led to increased water consumption (Osman et al. 2017). Studies have been conducted to assess the quality of fresh water in Iraq and have reported an increase in several contaminants (Ewaid et al. 2018). Based on these problems, coupled with others such as continuing wars, embargo, and terrorism, there is an unclear view of whether decision-makers can manage water resources under conditions of decreased availability.

[1]Assistant Professor, Dept. of Civil Engineering, Wasit Univ., Wasit 52001, Iraq; Assistant Professor, College of Engineering, Univ. of Warith Al-Anbiyaa, Karbala 56001, Iraq. Email: salahlafta@uowasit.edu.iq

[2]Assistant Professor, Dept. of Civil Engineering, Wasit Univ., Wasit 52001, Iraq. Email: nsaleem@uowasit.edu.iq

[3]Senior Lecturer, Dept. of Applied Mathematics, Liverpool John Moores Univ., Liverpool L3 3AF, UK. ORCID: https://orcid.org/0000-0001-9927-3209. Email: s.ortegamartorell@ljmu.ac.uk

[4]Ph.D. Student, Dept. of Electrical and Electronics Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia. ORCID: https://orcid.org/0000-0003-4385-7154. Email: hussain_mhammad@yahoo.com

[5]Professor, Dept. of Civil, Environmental and Natural Resources Engineering, Lulea Univ. of Technology, Lulea T3334, Sweden (corresponding author). Email: nadhir.alansari@ltu.se

[6]Assistant Professor, Dept. of Mechanical Engineering, Wasit Univ., Wasit 52001, Iraq. ORCID: https://orcid.org/0000-0003-1746-0137. Email: hrazzaq@uowasit.edu.iq

[7]Assistant Professor, Dept. of Environment Engineering, Babylon Univ., Babylon 51001, Iraq. Email: k.s.hashim@ljmu.ac.uk

[8]Professor, Dept. of Medical Instrumentation Techniques Engineering, Electrical Engineering Technical College, Middle Technical Univ., Baghdad 10022, Iraq. ORCID: https://orcid.org/0000-0002-9071-1775. Email: sadik.gharghan@mtu.edu.iq

Note. This manuscript was submitted on September 29, 2021; approved on June 9, 2022; published online on November 7, 2022. Discussion period open until April 7, 2023; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Water Resources Planning and Management***, © ASCE, ISSN 0733-9496.

Predicting municipal water demand is crucial to enhancing municipal water security, and monthly estimation is vital in the management of dam reservoirs (Ebrahim Banihabib and Mousavi-Mirkalaei 2019). Accurate forecasting of municipal water demand will help utilities recognize the temporal patterns of water needed to satisfy the balance between water delivered and water ordered, which in turn supports the sustainability of the water system (Altunkaynak and Nigussie 2017).

De Souza Groppo et al. (2019) and Ghalehkhondabi et al. (2017) found that forecasting of municipal water consumption has progressed over the last several decades, focusing on machine learning techniques and artificial neural networks (ANNs) as the most popular forecasting techniques. Xenochristou and Kapelan (2020) noted that ANN models have been applied in different fields and have been proven effective in forecasting short-, medium-, and long-term urban water demand (Bata et al. 2020; Tiwari and Adamowski 2015; Zubaidi et al. 2020a). Also, they have been successfully used in ecohydraulic and environmental engineering (Lama et al. 2021a; Pandya et al. 2017; Sadeghifar et al. 2022; Zhu et al. 2022). However, determining the optimum hyperparameters of machine learning models is still considered a substantial challenge. To address this, automated machine learning approaches (such as AutoML) have been proposed to help build hybrid prediction models (He et al. 2021) without extensive knowledge of statistics and machine learning (Zöller and Huber 2021), while reducing human effort and potential bias (Hutter et al. 2019). In addition, recent studies (Archetti and Candelieri 2019; Chatzipavlis et al. 2018; Frazier 2018) have investigated the use of Bayesian Optimization (*BO*) to identify an optimal configuration of the hyperparameters of a machine learning algorithm within a limited number of trials, especially for long-term data.

Although several automated machine learning approaches have been applied in the last decades in forecasting water demand, there is still room for improvement (De Souza Groppo et al. 2019). For example, Candelieri and Archetti (2018) reported a substantial improvement in forecast precision over that reported in other research studies (Candelieri 2017; Shabani et al. 2018). Furthermore, Candelieri and Archetti (2018) intend to utilise the other automated machine learning techniques in other application fields. These research studies highlight the importance of continuing the investigation of new methodologies that may offer useful scientific insights to policymakers. Based on recent literature (Archetti and Candelieri 2019; Chatzipavlis et al. 2018; Frazier 2018), Bayesian optimization (*BO*) can identify an optimal configuration of the hyperparameters of a machine learning algorithm within a limited number of trials, especially for long-term data.

In our study, five nature-inspired optimization algorithms were integrated in the ANN model to simulate monthly stochastic water demand data. These algorithms include (1) the slime mold algorithm (SMA), which was proposed by Li et al. (2020) and successfully applied in feature selection (Abdel-Basset et al. 2021), wind power prediction (Yan and Wu 2020), and image segmentation (Abdel-Basset et al. 2020a); (2) the marine predators algorithm (MPA), which was proposed by Faramarzi et al. (2020) and effectively applied in COVID-19 detection (Abdel-Basset et al. 2020b), engineering applications (Ghafil and Jármai 2020), and tensile behavior prediction (Abd Elaziz et al. 2020); (3) multiverse optimizer (MVO), which was efficiently utilized in engineering optimization (Sulaiman et al. 2020), streamflow prediction modeling (Mohammadi et al. 2020), and design optimization of a cam-follower mechanism (Abderazek et al. 2020); (4) backtracking search algorithm (BSA), which was successfully used in identification of soil parameters (Jin and Yin 2020), parameter estimation of power signals (Mehmood et al. 2020), and optimization of photovoltaic models (Zhang et al. 2020); and (5) crow search algorithm (CSA), which was effectively applied in feature selection (Ouadfel and Abd Elaziz 2020), reinforced concrete applications (Sultana et al. 2020), and solving optimal control issues (Turgut et al. 2020).

Currently, urban water demand forecasting is extremely challenging for water companies that are struggling to adapt water systems, specifically in terms of increasing concerns about the impact of climate change and water security. Additionally, there are very few studies about forecasting the stochastic signal of water needed, based on climatic factors. Consequently, considerable uncertainty still exists concerning the unexpected growth of stochastic patterns in water demand resulting from the stochastic impact of climatic factors (Zubaidi et al. 2018, 2020c). Based on the literature review, the innovation of this research is to (1) assess, for the first time in Iraq, the extent to which climatic factors have driven urban water demand; (2) integrate the ANN model and the recently developed MPA algorithm to create MPA-ANN, the first application in the field of urban water demand forecasting; (3) compare MPA-ANN with four nature-inspired optimization algorithms (SMA, MVO, BSA, and CSA) to increase the forecasting range and decrease uncertainty; (4) apply a novel methodology (data preprocessing and hybrid model) to forecast the monthly stochastic pattern of water demand; and (5) offer a scientific view to decision-makers of the impact of climatic factors on water demand to satisfy sustainability in a country that faces a unique environment of climate change and water scarcity.

Study Area

Iraq is one of the Arab countries located in an arid to semiarid area in the Middle East. Its capital is Baghdad City, which is situated in the center of Iraq, covering an area of around 204.2 km^2 (Fig. S1). Baghdad City suffered from sectarian violence from 2004 to 2017 that impacted the pattern and rate of its population growth. However, Iraq had a rapid population growth rate of 2.5% in 2018, with more than 8.5 million inhabitants living in Baghdad. The Mayoralty of Baghdad City has ten water treatment projects to treat and deliver potable water from the Tigris River to residential, institutional, industrial, and commercial customers. The predominant climate in Iraq is dry and hot to extremely hot in summer, and cold and wet in winter. The country faces considerable climate change that causes extreme heat waves and decreases the amount and changes the pattern of precipitation. Hence, the capacity of freshwater resources has decreased and the municipal water system is under stress (Chabuk et al. 2020; Ewaid et al. 2018; Zubaidi et al. 2019).

Methodology

The urban water demand methodology suggested here allows medium-term time series demand forecasting to be calculated based on climatic factors. Fig. 1 shows the steps needed to build it.

Forecast Model Data

Historical data can assist in estimating and extrapolating possible impacts in the future, and the forecast will contribute to building a looked-for future (Partidário 2007). Development of the urban water demand-forecast model necessitates the availability of historical water consumption and climatic factors time series data. Accordingly, in this study nine climatic factors were used to simulate monthly municipal water demand (million cubic meters; MCUM)

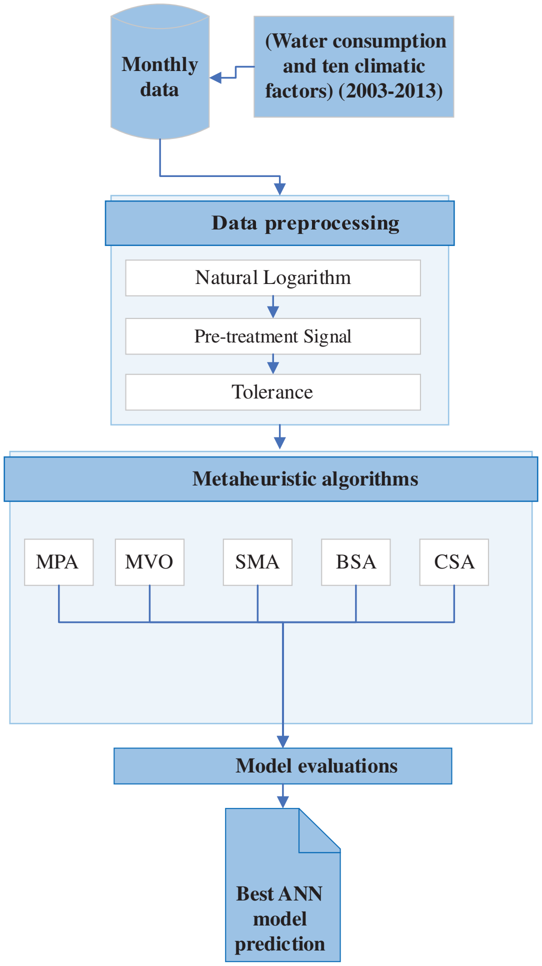

Fig. 1. (Color) Flowchart for forecasting future municipal water demand.

over 11 years (2003–2013) in Baghdad City. These climatic factors have been used effectively to forecast water demand in different scenarios in studies focusing on different regions. Here they included maximum temperature, $T_{\max}$ (°C), minimum temperature, $T_{\min}$ (°C), mean temperature, T_{mean} (°C), precipitation, P (mm), wind, W (m/s), solar radiation, S_{rad} (MJ/m^2), relative humidity, RH (%), dew point temperature, T_{dp} (°C), and surface pressure, S_p (kPa).

Socioeconomic factors (e.g., population) are deterministic components (Rasifaghihi et al. 2020; Zubaidi et al. 2020c) and therefore out of the scope of this study, which focused on the impact of climatic factors, which have stochastic behavior, on water consumption.

Data Preprocessing

Data preprocessing is a substantial phase that brings the data to a state that enables to easy and accurate forecasting by the developed model. It can be divided into normalization, cleaning, and choice of model input (Tabachnick and Fidell 2013). Haque et al. (2018) claimed that time series should be scaled down (normalized) to make the output space smoother and reduce the impact of outliers; Cleophas and Zwinderman (2016) suggested applying natural logarithm to normalize the time series.

Data cleaning means decomposing time series trends, seasonal (nonstationary) components, the stochastic (stationary) component, and noise, and then selecting the stochastic component only for dependent and independent factors because of the stochastic relationship between climatic factors and water consumption (Zubaidi et al. 2020c). Thus, the pretreatment signal approach was used here to implement this step.

The main aim of a factor choice procedure is to find the right independent factors, which have a significant effect on the dependent factor and yield a robust forecast model (Seo et al. 2018). In this research, the tolerance technique was used to select the model input factors by avoiding multicollinearity. Each independent factor in the best scenario had to have a tolerance coefficient of more than 0.2 to ensure that there was no collinearity (Cleophas and Zwinderman 2016).

Artificial Neural Network

ANN is currently the most common machine learning technique applied in the hydrological area, particularly learning using a feedforward back-propagation (FFBP) structure The FFBP was used in precisely simulating municipal water needed across various spatiotemporal scales because of its ability to map the nonlinear (i.e., trend and seasonal) behavior of water data (Shirkoohi et al. 2021; Zounemat-Kermani et al. 2020).

The Levenberg-Marquardt (*LM*) algorithm was used to train the ANN approach because it is known to minimize prediction error and efficiently simulate any predictor/response map (Bayatvarkeshi et al. 2018; Zare Abyaneh et al. 2016). As in Zubaidi et al. (2020c), the topology of the ANN was classified into four layers of neurons, including the input layer, which contained the predictor factors (i.e., climatic factors), two hidden layers, and the output layer, which contained the response factors (i.e., water demand) (Fig. S2). The tansigmoidal activation function was chosen in the first and second hidden layers, while the linear activation function was employed in the output layer. The process of ANN training was repeated many times over an epoch (i.e., 1,000 iterations) until the error between the actual and simulated urban water time series data reached its minimum. In this study, for each variable, 70% of the data set (92 out of 132 data points) was used for training, 15% was used as the test set (20 out of 132 data points), and 15% was used for validation (20 out of 132 data points). Choosing these percentages followed earlier studies, (e.g., Chyad et al. 2022; Zubaidi et al. 2020b, c).

Indeed, ANN performance relies on the optimization of its hyperparameters, which define topology options and ANN learning. Recently, ANN models were successfully integrated by various metaheuristic algorithms to select the best hyperparameters for the short and long term. However, these combined techniques were applied in limited way in the urban water demand field and additional research effort is required to develop more effective and precise combined models in the future (Shirkoohi et al. 2021; Zounemat-Kermani et al. 2020).

In this study, the ANN model was combined with the MPA algorithm to select the learning rate (*Lr*) and the number of hidden neurons (*N*1 and *N*2) for the first and second hidden layers instead of the trial-and-error approach. The MPA-ANN algorithm was compared with SMA-ANN, CSA-ANN, BSA-ANN, and MVO-ANN to increase the forecasting range and decrease the uncertainty.

Marine Predators Algorithm

The marine predator algorithm (MPA) is a novel metaheuristic optimization algorithm in which the behavior of ocean creatures in their search for food is simulated. These creatures include sharks, monitor lizards, sunfish, equine fishes, and swordfish. In the ocean, both predators and prey strive to get food to survive. Their behavior inspired researchers to follow this approach to derive a sound algorithm in terms of fitness. Formulation of MPA begins with assigning an initial random set of solutions based on the search space, as illustrated in Eq. (1)

$$Z = \boldsymbol{X}_{\mathbf{lower}} + \text{rand} \times (\boldsymbol{X}_{\mathbf{upper}} - \boldsymbol{X}_{\mathbf{lower}}) \tag{1}$$

where $\boldsymbol{X}_{\mathbf{lower}}$ and $\boldsymbol{X}_{\mathbf{upper}}$ = lower and upper bound of the search space, respectively; and rand = random number with a range of [0, 1].

Two matrices must be defined in MPA because of the nature of the algorithm, in which both predator and prey are looking for food. Therefore, both are considered search agents. These two matrices are referred to as elite (for predator) and prey, respectively, as shown in Eqs. (2) and (3). According to the concept of "survival of the fittest," the top predators should be the ones with higher hunting kills and merits in the search space. As such, the elite matrix should include only the fittest agents in the search space (predators). Then, depending on the prey positions, the matrix is updated. The dimensions of the prey matrix must be the same as for the elite matrix. In Elite matrix, the predator updates its position based on Prey matrix

$$\mathbf{Elite} = \begin{bmatrix} \boldsymbol{X}_{11}^1 & \boldsymbol{X}_{12}^1 & . & \boldsymbol{X}_{1d}^1 \\ \boldsymbol{X}_{21}^1 & \boldsymbol{X}_{22}^1 & \cdot & \boldsymbol{X}_{2d}^1 \\ \cdot & \cdot & \cdot & \cdot \\ . & . & . & . \\ \boldsymbol{X}_{n1}^1 & \boldsymbol{X}_{n2}^1 & \cdot & \boldsymbol{X}_{nd}^1 \end{bmatrix} \tag{2}$$

$$\mathbf{Prey} = \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} & \cdot & \boldsymbol{X}_{1d} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} & \cdot & \boldsymbol{X}_{2d} \\ \cdot & \cdot & \cdot & \cdot \\ \boldsymbol{X}_{n1} & \boldsymbol{X}_{n2} & \cdot & \boldsymbol{X}_{nd} \end{bmatrix} \tag{3}$$

where $\boldsymbol{X}_1^1$ = optimal predator vector; n = number of search agents; and d = number of dimensions.

In the two matrices, the positions of the predators and preys are updated according to three phases. These phases are merely dependent on the velocity difference between the two. To simulate the whole life of both predator and prey in nature, a designated number of iterations should be assigned in each phase. The details of each phase are discussed in the following subsections.

Phase 1: High-Velocity Ratio

In Phase 1, the predator is faster than the prey. This phase occurs in one-third of the total number of iterations (i.e., $1/3t_{\max}$). The step size of prey movement is updated as in Eq. (4)

$$S_i = \boldsymbol{R_B} \otimes (\mathbf{Elite}_i - \boldsymbol{R_B} \otimes \mathbf{Prey}_i), \quad i = 1, 2, \ldots, n \tag{4}$$

$$\mathbf{Prey}_i = \mathbf{Prey}_i + P.\boldsymbol{R} \otimes S_i \tag{5}$$

where $\boldsymbol{R}$ = random vector with a range of [0, 1]; $P = 0.5 = a$ constant number; $\boldsymbol{R_B}$ = random vector referring to Brownian motion; and $\otimes$ = element-wise multiplication.

Phase 2: Unit Velocity Ratio

In Phase 2, the predator and prey are moving at the same pace. The prey is represented by Levy flight while the predator is represented by Brownian motion. This phase occurs in the second third of the total iterations (i.e., $1/3t_{\max} < t < 2/3t_{\max}$). The following equations are applied to the first half of the population

$$S_i = R_L \otimes (\mathbf{Elite}_i - R_L \otimes \mathbf{Prey}_i), \quad i = 1, 2, \ldots, n \tag{6}$$

$$\mathbf{Prey}_i = \mathbf{Prey}_i + P.\boldsymbol{R} \otimes S_i \tag{7}$$

where R_L= numbers following Levy distribution. The second half of the population is subjected to the following equations:

$$S_i = \boldsymbol{R_B} \otimes (\boldsymbol{R_B}\mathbf{Elite}_i - \mathbf{Prey}_i), \quad i = 1, 2, \ldots, n \tag{8}$$

$$\mathbf{Prey}_i = \mathbf{Prey}_i + P.CF \otimes S_i, \qquad CF = \left(1 - \frac{t}{t_{\max}}\right)^{2\left(\frac{t}{t_{\max}}\right)} \tag{9}$$

where CF= parameter controlling the movement step size of the predator.

Phase 3: Low-Velocity Ratio

Phase 3, the final phase of the optimization, simulates predator movements when the predator is faster than the prey. It occurs in the final third of the total iterations (i.e., $2/3t_{\max}$)

$$S_i = R_L \otimes (R_L \otimes \mathbf{Elite}_i - \mathbf{Prey}_i), \quad i = 1, 2, \ldots, n \tag{10}$$

$$\mathbf{Prey}_i = \mathbf{Prey}_i + P.CF \otimes S_i, \qquad CF = \left(1 - \frac{t}{t_{\max}}\right)^{2\left(\frac{t}{t_{\max}}\right)} \tag{11}$$

Details on both Levy and Brownian motions are further discussed in the following paragraphs.

Brownian motion is inspired by normal distribution with a mean of 0 ($\mu = 0$) and a variance of 1 ($\sigma^2 = 1$). To determine the probability density function (PDF) corresponding to this motion at Point x, the following formula should be used:

$$P_B(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{x^2}{2}\right) \tag{12}$$

Levy flight is a stochastic and random step size in which Levy distribution is followed. The probability function of Levy distribution is formulated as

$$L(x_j) \approx |x_j|^{1-\alpha} \tag{13}$$

where x_j = flight length; and α = exponent of the power law that has the range ($1 < \alpha \leq 2$). The PDF of the Levy distribution is formulated as

$$P_L(x, \mu, \sigma) = \frac{1}{\pi} \int_0^\infty \exp(-\gamma q^\alpha) \cos(qx) dq \tag{14}$$

where γ = scale unit. The integral form can be used if α falls within its normal range (1, 2). As such, a Gaussian distribution is obtained if α equals 2 while Cauchy distribution can be obtained if α is 1. Higher values of x require series expansion method

$$P_L(x, \mu, \sigma) = \frac{\gamma\Gamma(1+\alpha)\sin\left(\frac{\pi\alpha}{2}\right)}{\pi x^{(1+\alpha)}}, \quad x = \infty \tag{15}$$

where Γ = gamma function in which $\Gamma(1+\alpha) = \alpha!$. Here, α ranges from 0.3 to 1.99. The present study followed Levy distribution to generate a random number

$$\text{Levy}(\alpha) = 0.05 \times \frac{x}{|y|^{\left(\frac{1}{\alpha}\right)}} \tag{16}$$

where y and x = variables with normal distributions of $x = \text{Normal}(0, \sigma_x^2)$ and $y = \text{Normal}(0, \sigma_y^2)$, where σ^2 can be determined as

$$\sigma_x = \left[\frac{\Gamma(1+\alpha)n\left(\frac{\pi\alpha}{2}\right)}{\Gamma\left(\frac{(1+\alpha)}{2}\right)\alpha 2^{\left(\frac{\alpha-1}{2}\right)}}\right]^{\frac{1}{\alpha}} \tag{17}$$

where $\sigma_y = 1$; and $\alpha = 1.5$.

Eddy Formation and Effect of FAD

It should be noted that in formulating the MPA, the surrounding environment can play a vital role in terms of its impacts on the behavior of prey, specifically eddy formulation and fish-aggregating devices (FADs). This effect can be presented as

$$\mathbf{Prey}_i = \begin{cases} \mathbf{Prey}_i + \boldsymbol{CF}[\boldsymbol{X}_{\mathbf{min}} + \boldsymbol{R} \otimes (\boldsymbol{X}_{\mathbf{max}} - \boldsymbol{X}_{\mathbf{min}})] \otimes \bar{\boldsymbol{U}} & \textit{if } \boldsymbol{r} < \text{FAD} \\ \mathbf{Prey}_i + [\mathbf{FADs}(1-\boldsymbol{r}) + \boldsymbol{r}](\mathbf{Prey}_{r1} - \mathbf{Prey}_{r2}) & \textit{if } \boldsymbol{r} > \text{FAD} \end{cases} \tag{18}$$

where $\boldsymbol{r}$ = random value ranging 0–1; $\boldsymbol{r}1$ and $\boldsymbol{r}2$ = random indices of prey matrix; FADs = 0.2 = probability of FADs effect; $\bar{\boldsymbol{U}}$ = binary vector; and $\boldsymbol{X}_{\mathbf{max}}, \boldsymbol{X}_{\mathbf{min}}$ = vectors of lower and upper bounds of dimensions.

In the MPA technique, memory saving should be done so that the old position of the prey can be saved to compare the fitness values of the old position with other successive solutions in which prey update their positions during the simulation. The flowchart of the MPA-ANN algorithm is presented in Fig. S3.

Performance Evaluation Criteria

The parameters of statistical criteria indicate the accuracy of prediction, so forecast error plays a significant role in the choice of an appropriate model that diminishes deviations in future forecasts (Donkor et al. 2014). Because there are no global performance criteria, it is essential to select criteria that are proper for a particular application (Seo et al. 2018). In this study, different performance criteria were considered for assessing the performance of the model. These criteria included coefficient of determination (R^2), coefficient of efficiency (CE), Nash-Sutcliffe index (NSI), root mean-square error (RMSE), mean absolute error (MAE), and mean bias error (MBE). The forecast technique has good accuracy and high performance in simulating the water advance time when satisfying one of these values of R^2, when CE is more than 0.9, when RMSE, MAE, and MBE approach zero (Dawson et al. 2007; Li et al. 2013), or if the value of NSI approaches 1 (Jain and Sudheer 2008).

The augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test were conducted to examine and determine the stochastic signal of all-time series of the dependent and independent variables.

Results

Data Preprocessing

The first step in data preprocessing is to normalize the time series of dependent and independent variables and detect and treat outliers (as discussed in "Methodology"). Figs. 2 and 3 highlight the differences between the raw data and the normalized and cleaned data.

Our main emphasis in this study was on the stochastic component only (as earlier discussed). To develop a water demand model based on climatic factors, water consumption, and climatic factor time series should first be decomposed using the pretreatment signal approach. Fig. 4 shows the normalized and cleaned water time series coupled with the first four signals (trend, seasonal, stochastic, and noise). The stationarity of the stochastic signal for the all-time series is assessed and confirmed by the ADF and KPSS tests.

Table 1 shows the difference in correlation coefficient between dependent and independent variables in raw and stochastic stages. It is obvious that the values corresponding to climatic factors in the stochastic stage are much higher than the values in the raw data. For example, the R between water consumption and precipitation increased from −0.535 to −0.931.

In the final part of data preprocessing, it is necessary to determine the highly correlated predictors (climatic factors) and at the same time avoid multicollinearity. According to the tolerance technique described in the section "Data Preprocessing," the scenario of selecting the best model input is repeated several times to choose predictors with a tolerance coefficient not less than 0.2. Accordingly, Table 2 reveals P, S_{rad}, and T_{dp} as the optimum scenario with coefficients of more than 0.2, meaning that there was no multicollinearity.

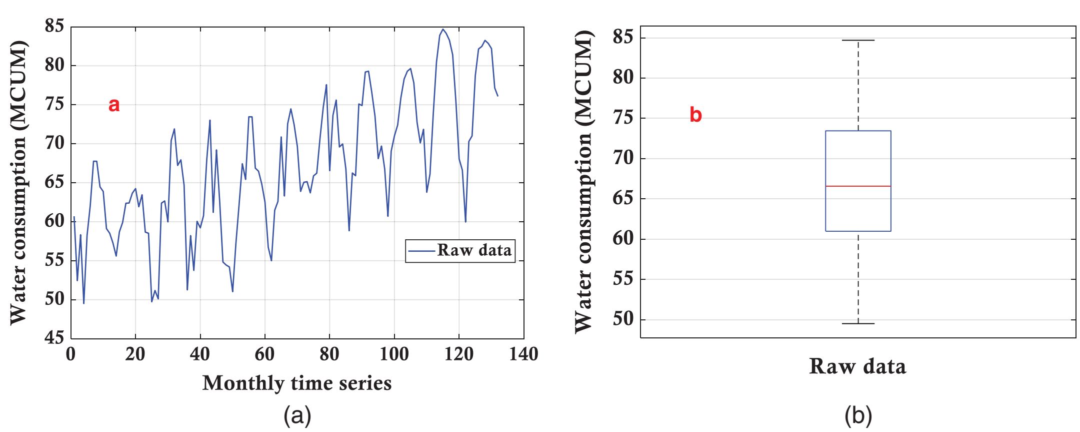

Fig. 2. (Color) (a) Monthly raw time series; and (b) urban water consumption for Baghdad City.

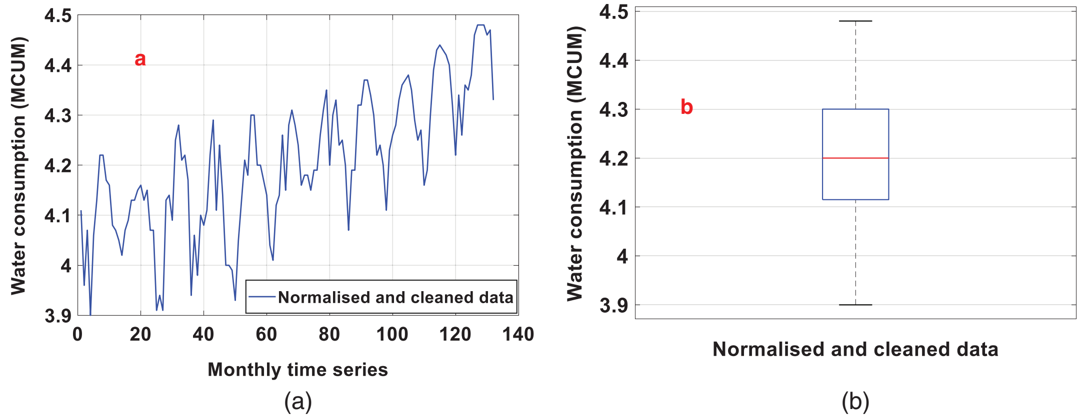

Fig. 3. (Color) (a) Monthly normalized and clean time series; and (b) box-plot of urban water consumption.

Model Configuration

The systematic configuration of the ANN model rather than by trial and error is necessary to build an accurate water demand prediction model. Accordingly, five hybrid metaheuristic algorithms (MVO-ANN, SMA-ANN, BSA-ANN, CSA-ANN, and MPA-ANN) were used to locate the ANN model's optimal hyperparameters (Lr, $N1$, and $N2$). Five swarm sizes (10, 20, 30, 40, and 50) were attempted by combining different algorithms with ANN, and each swarm for each algorithm was implemented five times to obtain the optimal solution (e.g., the MPA-ANN algorithm in Fig. S4, the best swarms are 10-3, 20-2, 30-3, 40-5, and 50-4). After that, the optimal swarm for each algorithm was selected to compare it with other swarms for the same algorithm (Fig. S5). From the figure, one can see that the best swarms are 30 for CSA-ANN, 40 for MPA-ANN, and 50 for SMA-ANN, MVO-ANN, and BSA-ANN.

Among all five hybrid metaheuristic algorithms, MPA-ANN was superior (Fig. 5). It yielded the lowest fitness function (RMSE) of 0.003993 after 42 iterations (fewer than the other algorithms). As such, adopting MPA-ANN is feasible and warranted. The 40 swarms of the MPA-ANN algorithm yielded Lr, $N1$, and $N2$ values of 0.213, 7, and 1, respectively.

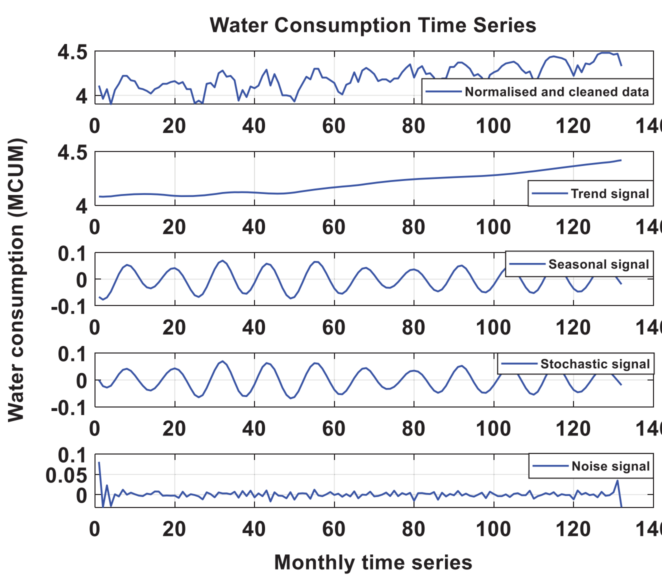

Fig. 4. (Color) Normalized and cleaned data and the first four signals obtained by pretreatment signal.

Evaluation of Model Performance

After integrating the ANN approach by determining the optimum hyperparameters, the model was run several times to locate a better network (weights and biases) that can precisely forecast the monthly stochastic signal of water demand. Various kinds of statistical tests were carried out to evaluate the ANN's ability to generalize stochastic water demand data depending on climatic factors in the validation phase.

Five statistical indicators were used to gauge the performance of the model as presented in Table 3. CE and NSE assessed the linear dependency between observed and predicted water demand, while MAE, RMSE, and MBE evaluated the nonlinear dependency between observed and predicted water demand. According to the limitation in the section "Performance Evaluation Criteria," the ANN model offered good accuracy.

The estimated model was further validated to double-check the model's power to accurately predict water consumption in the city of Baghdad. The target data of water consumption (x-axis) was plotted against simulated data (y-axis), with a 95% confidence interval (CI) (Fig. 6). It is noticeable that the target and simulated data

Table 2. Collinearity statistics for chosen predictors

Predictor	Tolerance coefficient
P	0.35
S_{rad}	0.23
T_{dp}	0.33

Table 1. Correlation of dependent and independent factors in raw and stochastic stages

Data	$T_{\max}$	$T_{\min}$	T_{mean}	P	W	S_{rad}	T_{dp}	RH	S_p
Raw	0.558	0.585	0.571	−0.535	0.396	0.453	0.376	−0.541	−0.523
Stochastic	0.92	0.93	0.926	−0.931	0.835	0.728	0.794	−0.917	−0.869

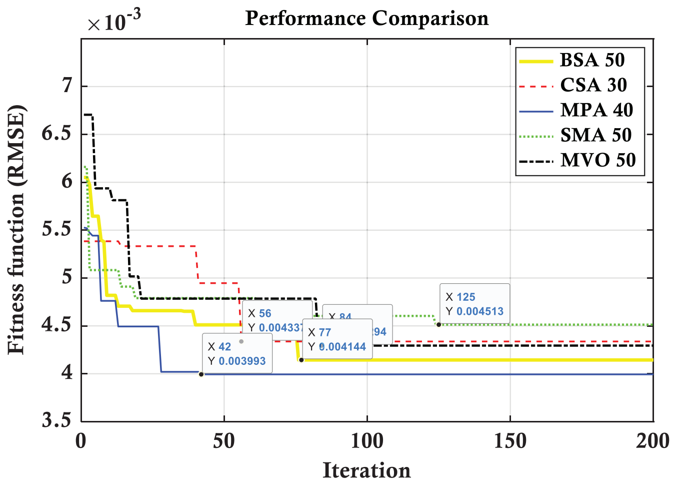

Fig. 5. (Color) Performance comparison of the five hybrid algorithms.

Table 3. MPA-ANN model statistical indicators in validation phase

Data	MAE	RMSE	*CE*	NSE	MBE
Validation	0.0057	0.0071	0.998	0.975	−0.0007

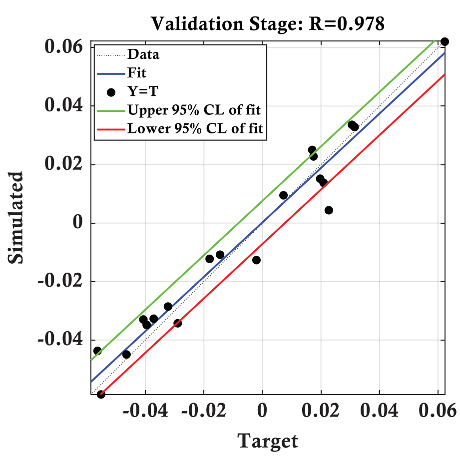

Fig. 6. (Color) Target water consumption data versus simulated data in the validation stage.

reveal a high level of consistency with $R = 0.978$, which supports the accuracy of the prediction model based on the limitation described in the section "Performance Evaluation Criteria."

According to the statistical tests used, the model demonstrated good performance in forecasting water consumption data in the validation stage.

Discussion

The selection of the stochastic component improved the correlation coefficients to climatic factors much higher than the counterpart values in the raw data. For example, between water consumption and precipitation, R increased from −0.535 to −0.931. The tolerance technique was shown to be very helpful in selecting the best model input among the nine independent variables. The climatic factors P, S_{rad}, and T_{dp} were selected to be the optimum scenario with tolerance coefficients of more than 0.2, which means no multicollinearity existed.

When the five metaheuristic algorithms were combined with the ANN for obtaining the hyperparameters at various numbers of swarms, the optimum swarm size was different for each algorithm based on the RMSE value. The performance of the metaheuristic algorithms was then compared at these optimum swarm sizes as there was no direct guide for selecting a swarm size for all of them. Comparing the performance of hybridized ANN, it was observed that MPA-ANN provided the highest prediction accuracy with the lowest RMSE value in fewer iterations than needed by the other hybrid algorithms. Consequently, the ANN optimum hyperparameters values were determined. Model validation process showed the model's very good performance in forecasting future values of water consumption with a coefficient of determination value of $R^2 = 0.978$.

Wolpert and Macready (1997) noted that, according to the "no free lunch" (NFL) concept, there is no specific theorem that can deliver the best solution compared with other theorems for all optimization problems. Using NFL, Faramarzi et al. (2020) developed the combined MPA theorem for guaranteeing a global solution depending on several strategies and techniques during optimization. Different foraging strategies have inspired MPA in the biological interaction between the prey and predators. Consequently, the Brownian and *LF* distributions were designed not only to have a systematic, efficient explorer-exploiter tendency but also to significantly enhance the search capability in each implementation. These permitted the MPA algorithm to precisely locate the global optima to solve the optimization issues considered in this research.

As a final note, since the size of the data set used in this study was relatively small, Bayesian optimization could have been used in conjunction with the MPA algorithm to increase execution speed and accuracy. Also, further methodological advances in ANN may substantially increase model performance after a limited number of iterations (i.e., faster convergence time). Since computation time was not a critical consideration in our study given that the measured data was obtained offline, we did not require Bayesian optimization. Bayesian optimization methods become relevant when using online data, which involves prolonged training and is computationally expensive. The main objective of our study was to reduce the error between the measured and simulated data.

Conclusion

Precise water demand prediction has received significant attention from water companies in the last few decades due to water scarcity and the rapid growth of water consumption. We used a novel methodology to estimate monthly stochastic municipal water demand based on some climatic factors by employing data over 11 years in Baghdad City. Ours is the first study to focus on in Iraq, which is one of the hottest countries in the world. The methodology contains data preprocessing and five metaheuristic algorithms (MPA, SMA, CSA, BSA, and MVO) combined with an ANN model. Considering the outcomes, data preprocessing was found to be a powerful technique for analyzing and selecting the stochastic component of any time series by applying pretreatment signals, and to determine the best model input scenario using tolerance. Accordingly, we have provided a guide for choosing suitable parameters that drive water demand. MPA was found to be a robust optimization algorithm that selects the best hyperparameters of ANN. The developed methodology can accurately forecast the monthly stochastic signal of urban water demand based on various statistical tests. These findings are of considerable significance to water utilities in planning, reviewing, and comparing the availability of freshwater resources against increasing water demand. Finally, our methodology can be applied to other cities in surrounding countries at various scales.

Data Availability Statement

Some or all data, models, or code used during the study were provided by the Mayoralty of Baghdad. Direct request for these materials may be made to the provider as indicated in the Acknowledgments.

Acknowledgments

The authors would like to thank the Mayoralty of Baghdad for providing data.

Supplemental Materials

Figs. S1–S5 are available online in the ASCE Library (www.ascelibrary.org).

References

Abd Elaziz, M., T. A. Shehabeldeen, A. H. Elsheikh, J. Zhou, A. A. Ewees, and M. A. A. Al-qaness. 2020. "Utilization of random vector functional link integrated with marine predators algorithm for tensile behavior prediction of dissimilar friction stir welded aluminum alloy joints." *J. Mater. Res. Technol.* 9 (5): 11370–11381. https://doi.org/10.1016/j.jmrt.2020.08.022.

Abdel-Basset, M., V. Chang, and R. Mohamed. 2020a. "HSMA_WOA: A hybrid novel Slime mould algorithm with whale optimization algorithm for tackling the image segmentation problem of chest X-ray images." *Appl. Soft Comput.* 95 (10): 106642. https://doi.org/10.1016/j.asoc.2020.106642.

Abdel-Basset, M., R. Mohamed, R. K. Chakrabortty, M. J. Ryan, and S. J. C. Mirjalili. 2021. "An efficient binary slime mould algorithm integrated with a novel attacking-feeding strategy for feature selection." *Comput. Ind. Eng.* 2021 (Jan): 153. https://doi.org/10.1016/j.cie.2020.107078.

Abdel-Basset, M., R. Mohamed, M. Elhoseny, R. K. Chakrabortty, and M. Ryan. 2020b. "A hybrid COVID-19 detection model using an improved marine predators algorithm and a ranking-based diversity reduction strategy." *IEEE Access* 8 (20): 79521–79540. https://doi.org/10.1109/ACCESS.2020.2990893.

Abderazek, H., A. R. Yildiz, and S. Mirjalili. 2020. "Comparison of recent optimization algorithms for design optimization of a cam-follower mechanism." *Knowl.-Based Syst.* 191 (Mar): 105237. https://doi.org/10.1016/j.knosys.2019.105237.

Ahmadi, M. S., J. Sušnik, W. Veerbeek, and C. Zevenbergen. 2020. "Towards a global day zero? Assessment of current and future water supply and demand in 12 rapidly developing megacities." *Sustainable Cities Soc.* 61 (Oct): 102295. https://doi.org/10.1016/j.scs.2020.102295.

Altunkaynak, A., and T. A. Nigussie. 2017. "Monthly water consumption prediction using season algorithm and wavelet transform–based models." *J. Water Resour. Plann. Manage.* 143 (6): 1–10. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000761.

Archetti, F., and A. Candelieri. 2019. *Bayesian optimization and data science*. Berlin: Springer.

Bata, M. T. H., R. Carriveau, and D. S. K. Ting. 2020. "Short-term water demand forecasting using nonlinear autoregressive artificial neural networks." *J. Water Resour. Plann. Manage.* 146 (3): 04020008. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001165.

Bayatvarkeshi, M., K. Mohammadi, O. Kisi, and R. Fasihi. 2018. "A new wavelet conjunction approach for estimation of relative humidity: Wavelet principal component analysis combined with ANN." *Neural Comput. Appl.* 21 (9): 4989–5000. https://doi.org/10.1007/s00521-018-3916-0.

Candelieri, A. 2017. "Clustering and support vector regression for water demand forecasting and anomaly detection." *Water* 9 (3): 224. https://doi.org/10.3390/w9030224.

Candelieri, A., and F. Archetti. 2018. "Global optimization in machine learning: The design of a predictive analytics application." *Soft Comput.* 23 (9): 2969–2977. https://doi.org/10.1007/s00500-018-3597-8.

Capt, T., A. Mirchi, S. Kumar, and W. S. Walker. 2021. "Urban water demand: Statistical optimization approach to modeling daily demand." *J. Water Resour. Plann. Manage.* 147 (2): 04020105. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001315.

Chabuk, A., Q. Al-Madhlom, A. Al-Maliki, N. Al-Ansari, H. M. Hussain, and J. J. A. J. Laue. 2020. "Water quality assessment along Tigris River (Iraq) using water quality index (WQI) and GIS software." 13 (14): 1–23. https://doi.org/10.1007/s12517-020-05575-5.

Chatzipavlis, A., G. E. Tsekouras, V. Trygonis, A. F. Velegrakis, J. Tsimikas, A. Rigos, T. Hasiotis, and C. Salmas. 2018. "Modeling beach realignment using a neuro-fuzzy network optimized by a novel backtracking search algorithm." *Neural Comput. Appl.* 31 (6): 1747–1763. https://doi.org/10.1007/s00521-018-3809-2.

Chyad, M. H., S. K. Gharghan, H. Q. Hamood, A. S. H. Altayyar, S. L. Zubaidi, and H. M. Ridha. 2022. "Hybridization of soft-computing algorithms with neural network for prediction obstructive sleep apnea using biomedical sensor measurements." *Neural Comput. Appl.* 34 (11): 8933–8957. https://doi.org/10.1007/s00521-022-06919-w.

Cleophas, T. J., and A. H. Zwinderman. 2016. *SPSS for starters and 2nd levelers*. Berlin: Springer.

Dawson, C. W., R. J. Abrahart, and L. M. See. 2007. "HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts." *Environ. Modell. Software* 22 (7): 1034–1052. https://doi.org/10.1016/j.envsoft.2006.06.008.

De Souza Groppo, G., M. A. Costa, and M. Libânio. 2019. "Predicting water demand: A review of the methods employed and future possibilities." *Water Supply* 19 (8): 2179–2198. https://doi.org/10.2166/ws.2019.122.

Donkor, E. A., T. H. Mazzuchi, R. Soyer, and J. A. Roberson. 2014. "Urban water demand forecasting: Review of methods and models." *J. Water Resour. Plann. Manage.* 140 (2): 146–159. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000314.

Ebrahim Banihabib, M., and P. Mousavi-Mirkalaei. 2019. "Extended linear and non-linear auto-regressive models for forecasting the urban water consumption of a fast-growing city in an arid region." *Sustainable Cities Soc.* 48 (Jul): 101585. https://doi.org/10.1016/j.scs.2019.101585.

Ewaid, S. H., S. A. Abed, and S. A. Kadhum. 2018. "Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis." *Environ. Technol. Innov.* 11 (Aug): 390–398. https://doi.org/10.1016/j.eti.2018.06.013.

Faramarzi, A., M. Heidarinejad, S. Mirjalili, and A. H. Gandomi. 2020. "Marine predators algorithm: A nature-inspired metaheuristic." *Expert Syst. Appl.* 152 (Aug): 113377. https://doi.org/10.1016/j.eswa.2020.113377.

Frazier, P. I. 2018. "Bayesian optimization." In *Recent advances in optimization and modeling of contemporary problems*, 255–278. Washington, DC: Informs.

Ghafil, H. N., and K. Jármai. 2020. "Dynamic differential annealed optimization: New metaheuristic optimization algorithm for engineering applications." *Appl. Soft Comput.* 93 (Aug): 106392. https://doi.org/10.1016/j.asoc.2020.106392.

Ghalehkhondabi, I., E. Ardjmand, W. A. Young, and G. R. Weckman. 2017. "Water demand forecasting: Review of soft computing methods." *Environ. Monit. Assess.* 189 (7): 1–13. https://doi.org/10.1007/s10661-017-6030-3.

Haque, M. M., A. Rahman, D. Hagare, and R. K. Chowdhury. 2018. "A comparative assessment of variable selection methods in urban water demand forecasting." *Water* 10 (419): 1–15. https://doi.org/10.3390/w10040419.

He, X., K. Zhao, and X. Chu. 2021. "AutoML: A survey of the state-of-the-art." *Knowl.-Based Syst.* 212 (Apr): 106622. https://doi.org/10.1016/j.knosys.2020.106622.

Hutter, F., L. Kotthoff, and J. Vanschoren. 2019. *Automated machine learning: Methods, systems, challenges*. Berlin: Springer.

Jain, S. K., and K. P. Sudheer. 2008. "Fitting of hydrologic models: A close look at the Nash-Sutcliffe index." *J. Hydrol. Eng.* 13 (10): 981–986. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:10(981).

Jin, Y. F., and Z. Y. Yin. 2020. "Enhancement of backtracking search algorithm for identifying soil parameters." *Int. J. Numer. Anal. Methods Geomech.* 44 (9): 1239–1261. https://doi.org/10.1002/nag.3059.

Lama, G. F. C., M. Crimaldi, V. Pasquino, R. Padulano, and G. B. Chirico. 2021a. "Bulk drag predictions of riparian *Arundo donax* stands through UAV-acquired multispectral images." *Water* 13 (10): 1333. https://doi.org/10.3390/w13101333.

Lama, G. F. C., M. Rillo Migliorini Giovannini, A. Errico, S. Mirzaei, R. Padulano, G. B. Chirico, and F. Preti. 2021b. "Hydraulic efficiency of green-blue flood control scenarios for vegetated rivers: 1D and 2D unsteady simulations." *Water* 13 (19): 2620. https://doi.org/10.3390/w13192620.

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

Li, S., H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili. 2020. "Slime mould algorithm: A new method for stochastic optimization." *Future Gener. Comput. Syst.* 111 (Mar): 300–323. https://doi.org/10.1016/j.future.2020.03.055.

Mehmood, A., P. Shi, M. A. Z. Raja, A. Zameer, and N. I. Chaudhary. 2020. "Design of backtracking search heuristics for parameter estimation of power signals." *Neural Comput. Appl.* 33 (5): 1479–1496. https://doi.org/10.1007/s00521-020-05029-9.

Mohammadi, B., F. Ahmadi, S. Mehdizadeh, Y. Guan, Q. B. Pham, N. T. T. Linh, and D. Q. Tri. 2020. "Developing novel robust models to improve the accuracy of daily streamflow modeling." *Water Resour. Manage.* 34 (10): 3387–3409. https://doi.org/10.1007/s11269-020-02619-z.

Nunes Carvalho, T. M., F. D. A. de Souza Filho, and V. C. Porto. 2021. "Urban water demand modeling using machine learning techniques: Case study of Fortaleza, Brazil." *J. Water Resour. Plann. Manage.* 147 (1): 05020026. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001310.

Osman, Y. Z., M. Abdellatif, N. Al-Ansari, S. Knutsson, and S. Jawad. 2017. "Climate change and future precipitation in an arid environment of the Middle East: Case study of Iraq." *J. Environ. Hydrol.* 25 (8): 1–18.

Ouadfel, S., and M. Abd Elaziz. 2020. "Enhanced crow search algorithm for feature selection." *Expert Syst. Appl.* 159 (Jun): 113572. https://doi.org/10.1016/j.eswa.2020.113572.

Pandey, P., N. D. Bokde, S. Dongre, and R. Gupta. 2021. "Hybrid models for water demand forecasting." *J. Water Resour. Plann. Manage.* 147 (2): 04020106. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001331.

Pandya, D. A., B. H. Dennis, and R. D. Russell. 2017. "A computational fluid dynamics based artificial neural network model to predict solid particle erosion." *Wear* 378–379 (Feb): 198–210. https://doi.org/10.1016/j.wear.2017.02.028.

Partidário, M. R. 2007. "Scales and associated data—What is enough for SEA needs?" *Environ. Impact Assess. Rev.* 27 (5): 460–478. https://doi.org/10.1016/j.eiar.2007.02.004.

Rasifaghihi, N., S. S. Li, and F. Haghighat. 2020. "Forecast of urban water consumption under the impact of climate change." *Sustainable Cities Soc.* 52 (Apr): 101848. https://doi.org/10.1016/j.scs.2019.101848.

Sadeghifar, T., G. F. C. Lama, P. Sihag, A. Bayram, and O. Kisi. 2022. "Wave height predictions in complex sea flows through soft-computing models: Case study of Persian Gulf." *Ocean Eng.* 245 (Jan): 110467. https://doi.org/10.1016/j.oceaneng.2021.110467.

Salman, S. A., S. Shahid, T. Ismail, K. Ahmed, and X.-J. Wang. 2018. "Selection of climate models for projection of spatiotemporal changes in temperature of Iraq with uncertainties." *Atmos. Res.* 213 (Jun): 509–522. https://doi.org/10.1016/j.atmosres.2018.07.008.

Seo, Y., S. Kwon, and Y. Choi. 2018. "Short-term water demand forecasting model combining variational mode decomposition and extreme learning machine." *Hydrology* 5 (4): 54. https://doi.org/10.3390/hydrology5040054.

Shabani, S., A. Candelieri, F. Archetti, and G. Naser. 2018. "Gene expression programming coupled with unsupervised learning: A two-stage learning process in multi-scale, short-term water demand forecasts." *Water* 10 (2): 142. https://doi.org/10.3390/w10020142.

Shirkoohi, M. G., M. Doghri, and S. Duchesne. 2021. "Short-term water demand predictions coupling an artificial neural network model and a genetic algorithm." *Water Supply* 21 (5): 2374–2386. https://doi.org/10.2166/ws.2021.049.

Sulaiman, M. H., Z. Mustaffa, M. M. Saari, and H. Daniyal. 2020. "Barnacles mating optimizer: A new bio-inspired algorithm for solving engineering optimization problems." *Eng. Appl. Artif. Intell.* 87 (Jun): 3330. https://doi.org/10.1016/j.engappai.2019.103330.

Sultana, N., S. M. Z. Hossain, M. S. Alam, M. M. A. Hashish, and M. S. Islam. 2020. "An experimental investigation and modeling approach of response surface methodology coupled with crow search algorithm for optimizing the properties of jute fiber reinforced concrete." *Constr. Build. Mater.* 243 (May): 118216. https://doi.org/10.1016/j.conbuildmat.2020.118216.

Szeląg, B., R. Suligowski, J. Drewnowski, F. De Paola, F. J. Fernandez-Morales, and Ł. Bąk. 2021. "Simulation of the number of storm overflows considering changes in precipitation dynamics and the urbanisation of the catchment area: A probabilistic approach." *J. Hydrol.* 598 (Jul): 126275. https://doi.org/10.1016/j.jhydrol.2021.126275.

Tabachnick, B. G., and L. S. Fidell. 2013. *Using multivariate statistics*. London: Pearson Education.

Tiwari, M. K., and J. F. Adamowski. 2015. "Medium-term urban water demand forecasting with limited data using an ensemble wavelet-bootstrap machine-learning approach." *J. Water Resour. Plann. Manage.* 141 (2): 04014053. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000454.

Tortajada, C., F. González-Gómez, A. K. Biswas, and J. Buurman. 2019. "Water demand management strategies for water-scarce cities: The case of Spain." *Sustainable Cities Soc.* 45 (Nov): 649–656. https://doi.org/10.1016/j.scs.2018.11.044.

Turgut, M. S., O. E. Turgut, and D. T. Eliiyi. 2020. "Island-based crow search algorithm for solving optimal control problems." *Appl. Soft Comput.* 90 (May): 106170. https://doi.org/10.1016/j.asoc.2020.106170.

Wolpert, D. H., and W. G. Macready. 1997. "No free lunch theorems for optimization." *IEEE Trans. Evol. Comput.* 1 (1): 67–82. https://doi.org/10.1109/4235.585893.

Xenochristou, M., and Z. Kapelan. 2020. "An ensemble stacked model with bias correction for improved water demand forecasting." *Urban Water J.* 17 (3): 212–223. https://doi.org/10.1080/1573062X.2020.1758164.

Yan, H., and Z. Wu. 2020. "A hybrid short-term wind power prediction model combining data processing, multiple parameters optimization and multi-intelligent models apportion strategy." *IEEE Access* 8 (12): 227126–227140. https://doi.org/10.1109/ACCESS.2020.3046001.

Zare Abyaneh, H., M. Bayat Varkeshi, G. Golmohammadi, and K. Mohammadi. 2016. "Soil temperature estimation using an artificial neural network and co-active neuro-fuzzy inference system in two different climates." *Arabian J. Geosci.* 9 (5): 2388. https://doi.org/10.1007/s12517-016-2388-8.

Zhang, Y., C. Huang, and Z. Jin. 2020. "Backtracking search algorithm with reusing differential vectors for parameter identification of photovoltaic models." *Energy Convers. Manage.* 223 (Nov): 113266. https://doi.org/10.1016/j.enconman.2020.113266.

Zhu, L., W. Zhang, and G. Tu. 2022. "Generalization enhancement of artificial neural network for turbulence closure by feature selection." *Adv. Aerodyn.* 4 (1): 1–24. https://doi.org/10.1186/s42774-021-00088-5.

Zöller, M.-A., and M. F. Huber. 2021. "Benchmark and survey of automated machine learning frameworks." *J. Artif. Intell. Res.* 70 (Jan): 409–472. https://doi.org/10.1613/jair.1.11854.

Zounemat-Kermani, M., E. Matta, A. Cominola, X. Xia, Q. Zhang, Q. Liang, and R. Hinkelmann. 2020. "Neurocomputing in surface water hydrology and hydraulics: A review of two decades retrospective, current status and future prospects." *J. Hydrol.* 588 (2): 125085. https://doi.org/10.1016/j.jhydrol.2020.125085.

Zubaidi, S. L., H. Al-Bugharbee, Y. R. Muhsen, K. Hashim, R. M. Alkhaddar, and W. H. Hmeesh. 2019. "The prediction of municipal water demand in Iraq: A case study of Baghdad Governorate." In *Proc., 12th Int. Conf. on Developments in eSystems Engineering (DeSE)*, 274–277. New York: IEEE.

Zubaidi, S. L., J. Dooley, R. M. Alkhaddar, M. Abdellatif, H. Al-Bugharbee, and S. Ortega-Martorell. 2018. "A novel approach for predicting monthly water demand by combining singular spectrum analysis with neural networks." *J. Hydrol.* 561 (Mar): 136–145. https://doi.org/10.1016/j.jhydrol.2018.03.047.

Zubaidi, S. L., K. Hashim, S. Ethaib, N. S. S. Al-Bdairi, H. Al-Bugharbee, and S. K. Gharghan. 2020a. "A novel methodology to predict monthly municipal water demand based on weather variables scenario." *J. King Saud Univ. Eng. Sci.* 2020 (9): 1–7. DOI:10.1016/j.jksues.2020.09.011.

Zubaidi, S. L., S. Ortega-Martorell, H. Al-Bugharbee, I. Olier, K. S. Hashim, S. K. Gharghan, P. Kot, and R. M. Alkhaddar. 2020b. "Urban water demand prediction for a city that suffers from climate change and population growth: Gauteng Province case study." *Water* 12 (7): 1885. https://doi.org/10.3390/w12071885.

Zubaidi, S. L., S. Ortega-Martorell, P. Kot, R. M. Alkhaddar, M. Abdellatif, S. K. Gharghan, M. S. Ahmed, and K. Hashim. 2020c. "A method for predicting long-term municipal water demands under climate change." *Water Resour. Manage.* 34 (3): 1265–1279. https://doi.org/10.1007/s11269-020-02500-z.