Integration of an artificial neural network and a simulated annealing algorithm for the optimization of the river water pollution index

Iman Ali Abdulkareem, Ammar Salman Dawood, Abdulhussain A. Abbas [*]

University of Basrah, College of Engineering, Civil Engineering Department, Iraq

A B S T R A C T

The Shatt Al-Arab River is the primary source of the water supply in the Al-Basrah province. Therefore, this study aimed to assess the water pollution index (WPI) of the Shatt Al-Arab River at 15 water treatment plants (WTPs) from 2011 to 2020 (except for WTP No. 11, which was sampled from 2012 to 2020). The WPI included 12 physicochemical parameters: turbidity (Tur), pH, electrical conductivity (EC), total dissolved solids (TDS), total hardness (TH), potassium (K^+), sodium (Na^+), magnesium (Mg^{+2}), calcium (Ca^{+2}), alkalinity (Alk), chloride (Cl^-), and sulfate (SO_4^{-2}). Two modeling methods, multiple linear regression (MLR) and an artificial neural network (ANN), were utilized to forecast the minimum value of the WPI. The simulated annealing (SA) technique and an integrated ANN-SA technique were used to estimate the best independent variable values that minimized the value of the WPI. A multilayer feed-forward neural network with a backpropagation algorithm was chosen for this study. A regression technique was employed to generate the WPI predicted equation which was also chosen as an objective function of the SA and combined ANN-SA. For the MLR method, the correlation coefficient (R) and mean squared error (MSE) values for the WPI were 4.746×10^{-7} and 1, respectively. The best ANN structure (10-17-1) predicted a WPI with MSE and R values of 8.851×10^{-11} and 1, respectively, for the training, 1.220×10^{-7} and 1 for the validation, and 1.354×10^{-9} and 1 for the testing. In contrast to the results obtained from the measured data, MLR analysis, and ANN technique, the combined ANN-SA method demonstrated the lowest WPI value at optimal parameters. The minimum WPI value for the integrated ANN-SA was 0.373.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Water is a necessity for all life forms and is as essential to humans as air (Dinka, 2018). Because rivers supply water for domestic, agricultural, and industrial uses, it is crucial to prevent river pollution and have accurate information on river water quality. The physical, chemical, and biological characteristics of water are used to determine its quality (Moyel, 2014). Pollution in rivers has resulted in severe water-borne diseases and health issues that affect the human population (Ahmed and Ismail, 2018).

The Shatt Al-Arab River (SAR) is the main river in Iraq and it is the primary supply of surface water in the province of Basrah. At Qurna, the Shatt Al-Arab River originates from the confluence of the Tigris and Euphrates rivers. Its water is utilized for a variety of purposes, including drinking, irrigation, industrial uses, and navigation. The main reasons for an increase in the salt concentration in the SAR are local and regional causes, the former being water pollution and the latter being a low-level flow from the Tigris and Euphrates rivers (Al-Muhyi, 2015). The largest influxes

of salinity into the SAR are from seawater and salt water that enters the river through sewage channels. The tidal phenomena in the Arabian Gulf have an impact on the hydrological systems of the SAR, including the tidal and seawater flows into the river (Al-Asadi and Alhello, 2019). The high salt content in the SAR poses a serious threat to the environment; it hampers water use and renders it inappropriate for many domestic, industrial, and agricultural purposes (Rahi, 2018).

In recent years, the artificial neural network (ANN) technique has gained popularity in a variety of fields for use in prediction, such as water resources and environmental science (Abyaneh, 2014). An ANN is a mathematical model based on biological neural networks research. According to particular connection patterns, an ANN consists of several interconnected nodes that are analogous to biological neurons (Wong et al., 2013). Simulated Annealing (SA) is a probabilistic search technique predicated on a single solution (Du and Swamy, 2016). In the early 1980s, three IBM researchers (Kirkpatrick et al., 1983) presented the principles of annealing in optimization. These ideas are based on a significant similarity with material annealing; this process comprises heating a solid to an exceptionally high temperature until the structure melts, and then cooling the solid under a

Table 1
Studies of using ANN with their input and target parameters.

References	Input parameters	Target parameters	Methods
Diamantopoulou et al. (2005)	T, pH, EC, DO, HCO_3^-, SO_4^{-2}, Cl^-, Na^+, Mg^{+2}, Ca^{+2}, NO_3^-, NH_4^+, TP, and Q	NO_3^-, EC, DO, Na^+, Mg^{+2}, Ca^{+2}	Multi-layer feed-forward ANN
Zhao et al. (2007)	T, Tur., pH, Alk., NH_4^+, Cl^-, and TH	BOD and COD	BP ANN
Singh et al. (2009)	pH, Alk., TH, TSS, COD, NO_3^-, NH_4^+, Cl^-, PO_4, K^+, and Na^+	DO and BOD	BP ANN
Banejad and Olyaie (2011)	TSS, pH, EC, Tur., Na^+, HCO_3^-, NO_3^-, NH_4^+, and PO_4^{-3}	DO and BOD	BP ANN
Najah et al. (2013)	EC, TDS, TSS, and Tur.	TDS, EC, Tur.	MLP ANN and RBF-NN
Dawood et al. (2016)	TH, EC, Tur., and TDS	EC and Tur.	Multi-layer feed-forward BP ANN
Hamdan and Dawood (2016)	pH, TH, Mg^{+2}, SO_4^{-2}, Cl^-, EC, and Tur.	TDS	Multi-layer feed-forward BP ANN
Gupta et al. (2019)	DO, Ph, Tur., EC, and E Coil	WQI	Multi-layer feed-forward ANN
Khudhur et al. (2020)	pH, Mg^{+2}, EC, TDS, Ca^{+2}, NO_3^-, SO_4^{-2}, and PO_4^{-3}	DO	Multi-layer feed-forward BP ANN
Shahid et al. (2020)	T, EC, DO, and pH	BOD, DO, SAR, TDS, and HCO_3^-	feed-forward BP ANN
Kulisz et al. (2021)	EC, pH, Ca^{+2}, Mg^{+2}, Na^+	WQI	BP ANN

carefully defined temperature reduction plan to create a solid state with the lowest amount of energy (Delahaye et al., 2019). The present study proposes a water pollution index (WPI) that presents a fundamental or adequate understanding of river water quality. The prediction of surface water quality is necessary for appropriate river basin management so that sufficient measures can be implemented to maintain pollution within allowed limits (Najah et al., 2013). Previously, the water quality of the SAR was evaluated by numerous researchers (Dawood, 2017; Dawood et al., 2017, 2018a; Hamdan et al., 2018; Dawood et al., 2018b; Al-Adhab et al., 2019; Dawood et al., 2020). Numerous researchers have employed ANNs in the context of water quality forecasting (Diamantopoulou et al., 2005; Zhao et al., 2007; Singh et al., 2009; Banejad and Olyaie, 2011; Najah et al., 2013; Dawood et al., 2016; Hamdan and Dawood, 2016; Gupta et al., 2019; Khudhur et al., 2020; Shahid et al., 2020; Kulisz et al., 2021). Table 1 presents information on previous studies that used ANNs and their input and target parameters. The integration of ANNs and SA has been the topic of numerous studies in a variety of fields (Zain et al., 2011; Zhang et al., 2011; Chaki and Ghosal, 2011; Bahrami and Ardejani, 2016; Nwobi-Okoye and Ochieze, 2018; Liu et al., 2020; Al-Mahasneh et al., 2021). In the field of water quality prediction, combining ANNs and SA remains a relatively new approach.

The objectives of this research were to assess the WPI in the SAR at numerous sampling locations at water treatment plants (WTPs), identify the optimal topology of the ANN model for the prediction of the WPI value (the results of the ANN model and the MLR model were compared with results of the measured data), and perform an integration of the ANN and SA to estimate the best independent variable values that minimized the value of the WPI.

2. Methodology

2.1. Study area

The SAR originates north of Basrah at the confluence of the Tigris and Euphrates rivers (Maytham et al., 2019). The SAR is located in southern Iraq in the province of Basra between 29° 45′ 00″ to 31° 15′ 00″ N and 47° 10′ 20″ to 48° 45′ 00″ E (Al-Asadi et al., 2020). The SAR is 192 km long, with a width ranging from 300 m at its source to 700 meters in Basrah City, and roughly 800 m at its mouth. The water required to maintain households, agricultural land, industries, ecological systems, transportation, and recreational activities is provided mainly by this river (Allafta and Opp, 2020). The SAR flows at a rate of 25 to 75 m³/s, and the flow of the river varies seasonally depending on the contribution of its tributary rivers, the amount of precipitation, and the effects of the tidal phenomena of the Arabian Gulf (Almuktar et al., 2020).

The freshwater flow rates reaching the river by the tributaries, as well as the flow of saltwater from the Arabian Gulf, determine the hydrological characteristics of the SAR. The Euphrates, Tigris, Karun, and Karkheh rivers are all primary tributaries of the SAR. Since the Karkheh and Karun River flows are diverted within the Iranian borders, and the Euphrates River is cut off before it reaches the Tigris, the hydrological river system is currently in a fundamentally different state. Since 2010, the SAR has been primarily dependent on the freshwater flow of the Tigris River (Al-Asadi and Alhello, 2019).

A decrease in the flow of fresh water from the Euphrates and Tigris rivers over time leads to the saltwater incursion progressing inland, causing the salinity in the SAR to rise (Almuktar et al., 2020). The seawater intrusion from the Arabian Gulf into the SAR could extend as much as 92 km (Al-Asadi et al., 2020). This intrusion corresponds to the following sites in the present study: Mhejran, Hamdan Bridge, Maheilah, and Al-Labanie. The high salinity of the water in the SAR restricted its utilization and rendered it unsuitable for most households and agricultural and industrial uses (Rahi, 2018). Moreover, the growth of aquatic plants has been significantly impacted by the increased salinity in the river. Fisheries have been damaged and biodiversity has been lost due to the increasing salinity in the water. The fish diversity in the Shatt al-Arab ecosystem has decreased from 68 type to 26 form, and fish in the Seebah area fish farms have died due to

Table 2
Statistical analysis of parameters.

Variables	Minimum	Maximum	Mean	Std. deviation
Tur.	0.60	79	15.62	±8.42
pH	7.03	8.47	7.64	±0.23
EC	871	34030	4897.38	±3952.51
TDS	200	22954	3113.79	±2655.54
TH	296	4860	980.59	±516.38
K^+	2.50	123	12.64	±8.11
Na^+	62	6780	713.15	±794
Mg^{+2}	36	590	117.83	±62.73
Ca^{+2}	59	976	199.06	±103.90
Alk.	90	296	157.45	±17.97
Cl^-	104	10300	1118.04	±1209.82
SO_4^{-2}	134	4449	804.76	±495.88

Note: All values are expressed in mg/L except pH (dimensionless), EC (μs/cm) and turbidity (NTU).

the high salinity. Salinity in irrigation water reduces the amount of moisture that is available to cultivated plants, which impacts plant growth. Low river levels and rising salinity have led to the paralysis of economic life in the Basrah province (Yaseen et al., 2016).

In recent years, domestic, industrial, and agricultural activity have increased, resulting in increased pollution, and the SAR has become an accessible dumping ground for all types of waste. The water quality of the SAR has deteriorated due to the above-mentioned types of contamination. Furthermore, effluent from domestic, agricultural, and industrial activities can produce significant alterations in the physical and chemical characteristics of the SAR (Hamdan, 2020).

2.2. Data description

In this study, monthly raw water parameters for the period 2011–2020 were acquired from the Basrah Water Directorate in Basrah, Iraq. A total of 12 parameters were investigated for each water sample: turbidity (Tur), pH, electrical conductivity (EC), total dissolved solids (TDS), total hardness (TH), potassium (K^+), sodium (Na^+), magnesium (Mg^{+2}), calcium (Ca^{+2}), alkalinity (Alk), chloride (Cl^-), and sulfate (SO_4^{-2}). Table 2 shows the descriptive statistics (minimum, maximum, mean, and standard deviation values) of the 12 physical and chemical parameters for the raw water for the period 2011–2020. The dataset used in this study was obtained from our previously published research (Abdulkareem et al., 2022). Table 3 shows the methods and equipment used to measure the parameters, while Figs. 1 and 2 show the measurement frequency of the parameters.

2.3. Sampling sites

The physical and chemical parameters of the river water were measured at 15 WTPs. Traditional water treatment processes were used in all of the WTPs, including pretreatment, chemical coagulation, quick mixing, flocculation, sedimentation, filtration, and disinfection. As a result, the WTPs lacked the necessary equipment to reduce the TDS in the water (Khudair and Eraibi, 2017). The coordinates for these WTPs are shown in Table 4, and Fig. 3 shows their locations. Table 5 shows the distances between the WTPs in this study. According to the TDS values, five categories of WTPs were revealed (Table 6).

2.4. Water pollution index

The water pollution index (WPI) is a new index that can be used to assess the physical, chemical, and biological quality of water sources using known water quality standards for a variety

Table 3
Methods and equipment of the measured parameters in this research.

Parameters	Methods	Equipment
pH	pH meter	pH meter model SD300
Tur.	Nephelometric	Turbidity Meter Lovibond TB 300 IR
EC	Electrometric	(METTLER TOLEDO FIVE GOTM conductivity meter (Mettler Toledo, Columbus, OH, USA)).
TDS	Electrometric	METTLER TOLEDO FIVE GOTM conductivity meter (Mettler Toledo, Columbus, OH, USA)).
Alk.	Titrimetric method	Atomic Absorption Spectrophotometer model SpectroDirect LoviBond (The Tintometer Limited, Amesbury, UK).
TH	EDTA titration	Burette
Mg^{+2}	EDTA titration	Burette
Ca^{+2}	EDTA titration	Burette
Na^+	Flame photometric	Flame photometer model M410
K^+	Flame photometric	Flame photometer model M410
Cl^-	Silver nitrate titration	–
SO_4^{-2}	Spectrophotometrically	Spectrophotometer
	Using barium sulfate turbidity method according to APHA (1998)	

Table 4
Geographic Coordinates of Sampling Sites along the SAR.

No. of water treatment plant	Name of water treatment plant	X	Y
1	Al-Dear	746969.3152	3410480.348
2	Al-Houta	764653.4872	3394077.932
3	Al Basrah Unified	763346.3973	3393806.914
4	Al-Garmma 1	763358.0723	3385353.837
5	Al-Garmma 2	763581.3847	3385795.167
6	Al-Faiha	767627.768	3385217.24
7	Al-Jubailah 1	769800.9144	3383108.889
8	Al-Ribat	771557.1858	3381562.851
9	Al-Bradhiah 1	774079.4702	3377999.902
10	Al-Bradhiah 2	774133.3117	3377949.087
11	Owaisyan	775211.3282	3376911.192
12	Mhejran	776773.9816	3374729.195
13	Hamdan Bridge	778716.0082	3373839.328
14	Maheilah	780719.0404	3373228.353
15	Al-Labanie	786574.7517	3373974.971

Table 5
Distance between water treatment plants in this study.

From–To	Distance (km)
Al-Dear–Al-Houta	24.16
Al-Houta–Al Basrah Unified	1.33
Al Basrah Unified–Al-Garmma 1	8.01
Al-Garmma 1–Al-Garmma 2	0.52
Al-Garmma 2–Al-Faiha	4.09
Al-Faiha–Al-Jubailah 1	3.01
Al-Jubailah 1–Al-Ribat	2.36
Al-Ribat–Al-Bradhiah 1	4.36
Al-Bradhiah 1–Al-Bradhiah 2	0.07
Al-Bradhiah 2–Owaisyan	1.50
Owaisyan–Mhejran	2.68
Mhejran–Hamdan Bridge	2.13
Hamdan Bridge–Maheilah	2.10
Maheilah–Al-Labanie	5.90
Al-Labanie–Coast	103.62
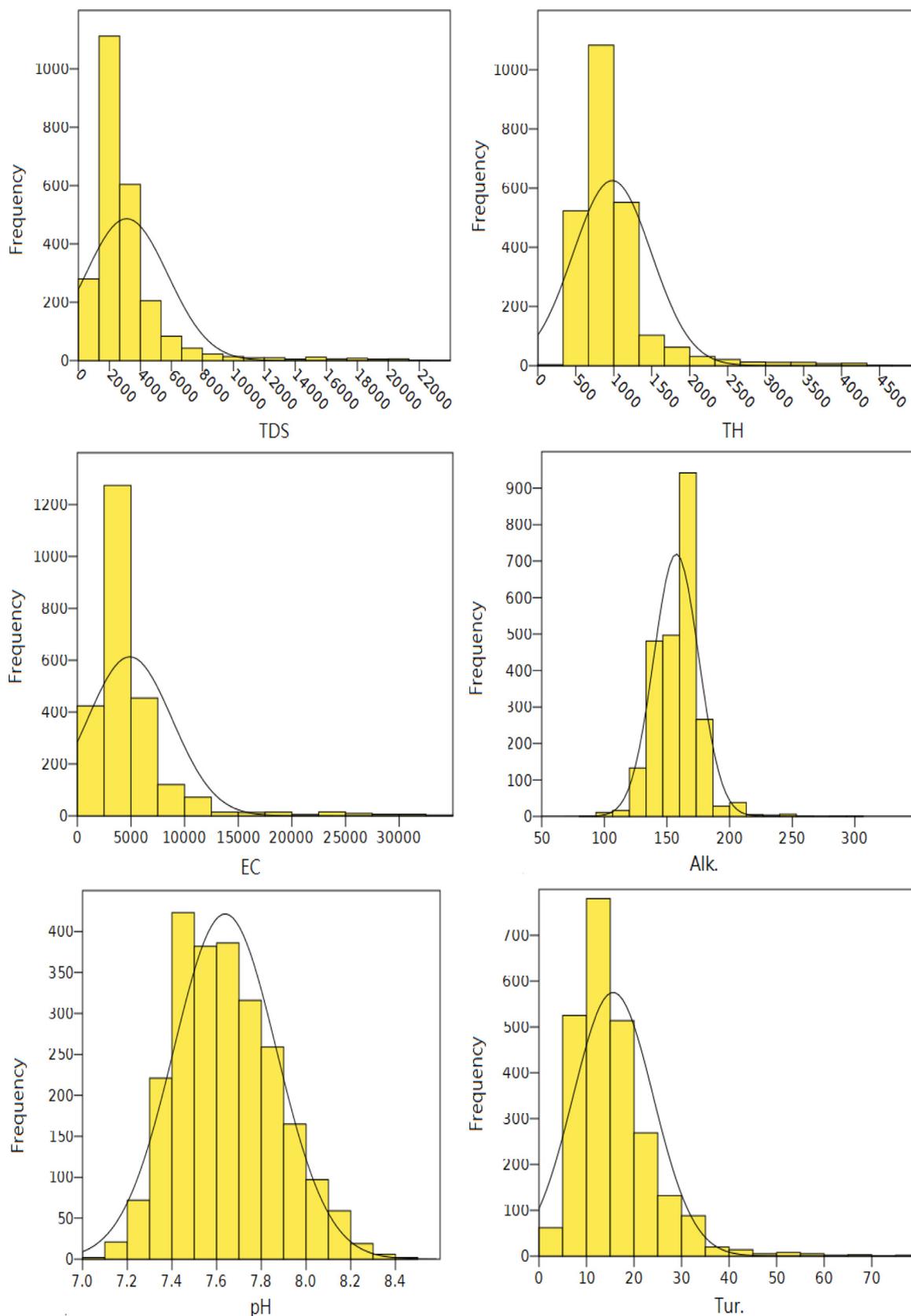
Fig. 1. Histograms of the TDS, TH, EC, Alk, pH and Tur parameters.

of purposes. In this context of any designated use, the WPI can present a basic or general understanding of the water quality state (Hossain and Patra, 2020).

A total of 12 water quality parameters (TDS, Tur, Mg^{+2}, pH, K^+, Cl^-, Na^+, EC, SO_4^{-2}, TH, Ca^{+2}, and Alk) were chosen to estimate the water pollution in the SAR using the WPI, according
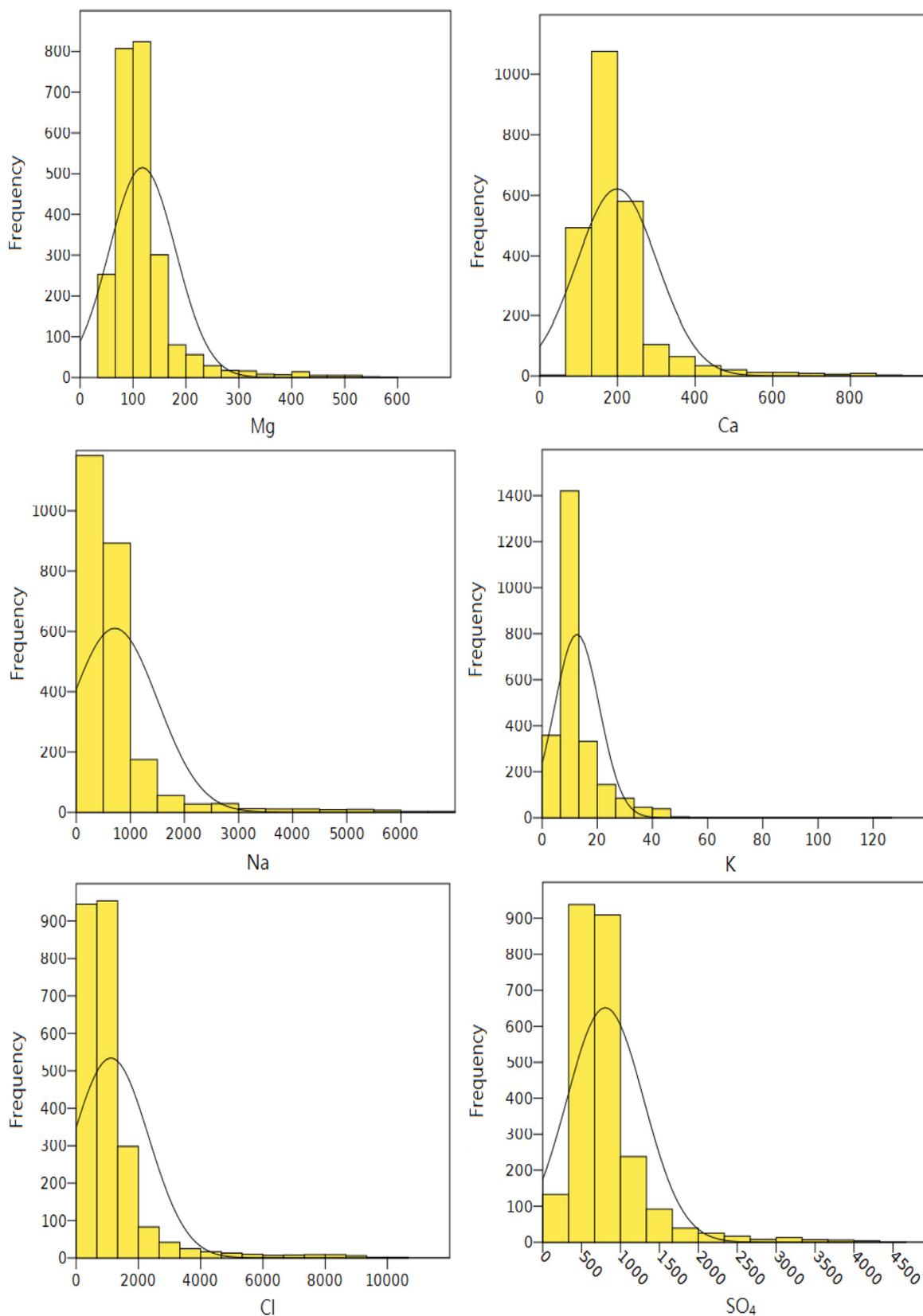
Fig. 2. Histograms of the Mg, Ca, Na, K, Cl, and SO₄ parameters.

to their standard permitted limits as stated by the World Health Organization (WHO, 2011) (Table 7). These standard permitted limits were used in several previous studies (Ewaid and Abed, 2017; Abbas et al., 2017; Ghalib, 2017; Mahmood et al., 2019).

The steps used to calculate the WPI were as follows (Hossain and Patra, 2020):

The pollution load (PL_i) for each parameter was computed in the first step using the formula presented below (Hossain and
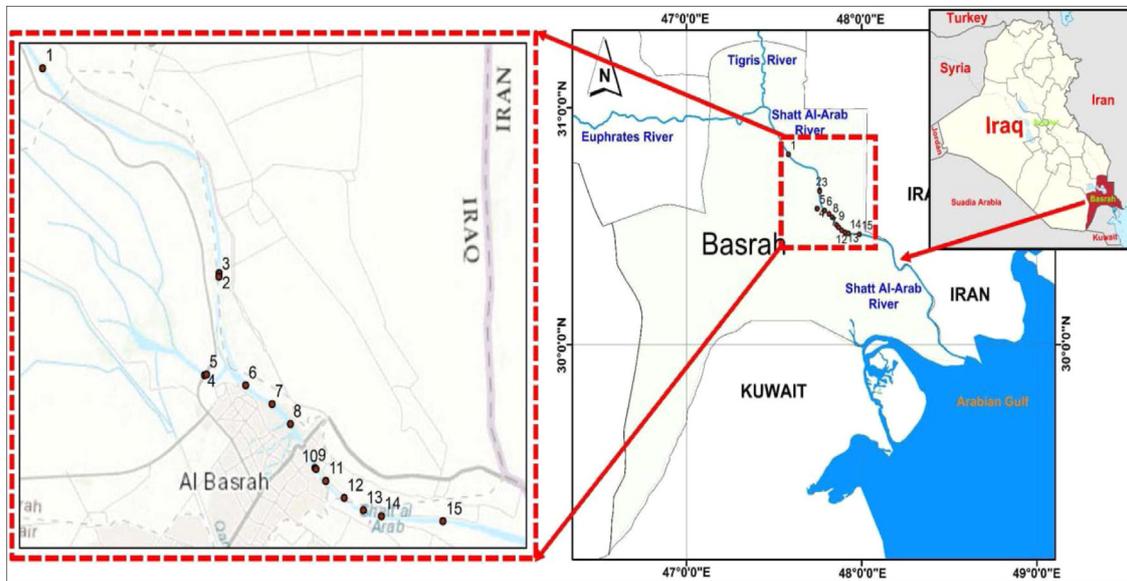
Fig. 3. Map of Basrah, Iraq, with an inset map showing the locations of the WTPs surveyed.

Table 6

Water treatment plants classification based on TDS values.

TDS (mg/l) Values	200–5000	5000–12000	12000–19000	19000–21000	>22000
WTPs	Al-Dear	Al-Houta	Al-Jubailah 1	Al-Bradhiah 1	Mhejran
		Al Basrah	Al-Faiha	Al-Bradhiah 2	Hamdan Bridge
		Unified	Al-Ribat	Owaisyan	
			Al-Garmma 1	Maheilah	Al-Labanie
			Al-Garmma 2		

Patra, 2020):

$$PL_i = 1 + \frac{Ci - Si}{Si} \tag{1}$$

Where Ci is the measured concentration for the specified parameter and Si denotes the standard permitted limit for the specified parameter.

In the case of pH, the following equations are advised for various pH ranges; if the pH is less than 7, Eq. (2) is suggested (Hossain and Patra, 2020). Where Si_c is the minimum pH value that may be accepted (6.5).

$$PL_i = \frac{Ci - 7}{Si_c - 7} \tag{2}$$

If the pH is greater than 7, Si_d is the highest pH value that can be accepted (8.5), and the proposed equation is given below (Hossain and Patra, 2020):

$$PL_i = \frac{Ci - 7}{Si_d - 7} \tag{3}$$

Finally, the overall WPI was computed by combining all the values of the pollution load and then dividing that by the number of parameters, as specified in the formula below (Hossain and Patra, 2020):

$$WPI = \frac{\sum_{i=1}^{n} PL_i}{n} \tag{4}$$

The values of the WPI can be categorized depending on n number of parameters in four categories (Table 8) (Hossain and Patra, 2020).

Table 7

Maximum allowable values of water quality variables given by WHO 2011.

Variables	WHO 2011
pH	6.5-8.5
Tur.	5
EC	1500
TDS	1000
TH	500
K^+	12
Na^+	200
Mg^{+2}	100
Ca^{+2}	75
Alk.	200
Cl^-	250
SO_4^{-2}	250

Table 8

WPI classification.

WPI value	Category
<0.5	Excellent water
0.5–0.75	Good water
0.75–1	Mildly polluted water
>1	Highly polluted water

2.5. Artificial neural networks

Artificial neural networks are among the finest reliable and commonly used prediction models that have shown successful application for the prediction of social, economic, and engineering aspects, etc. (Khashei and Bijari, 2010). In comparison to traditional computers, ANNs manage problem-solving in a unique
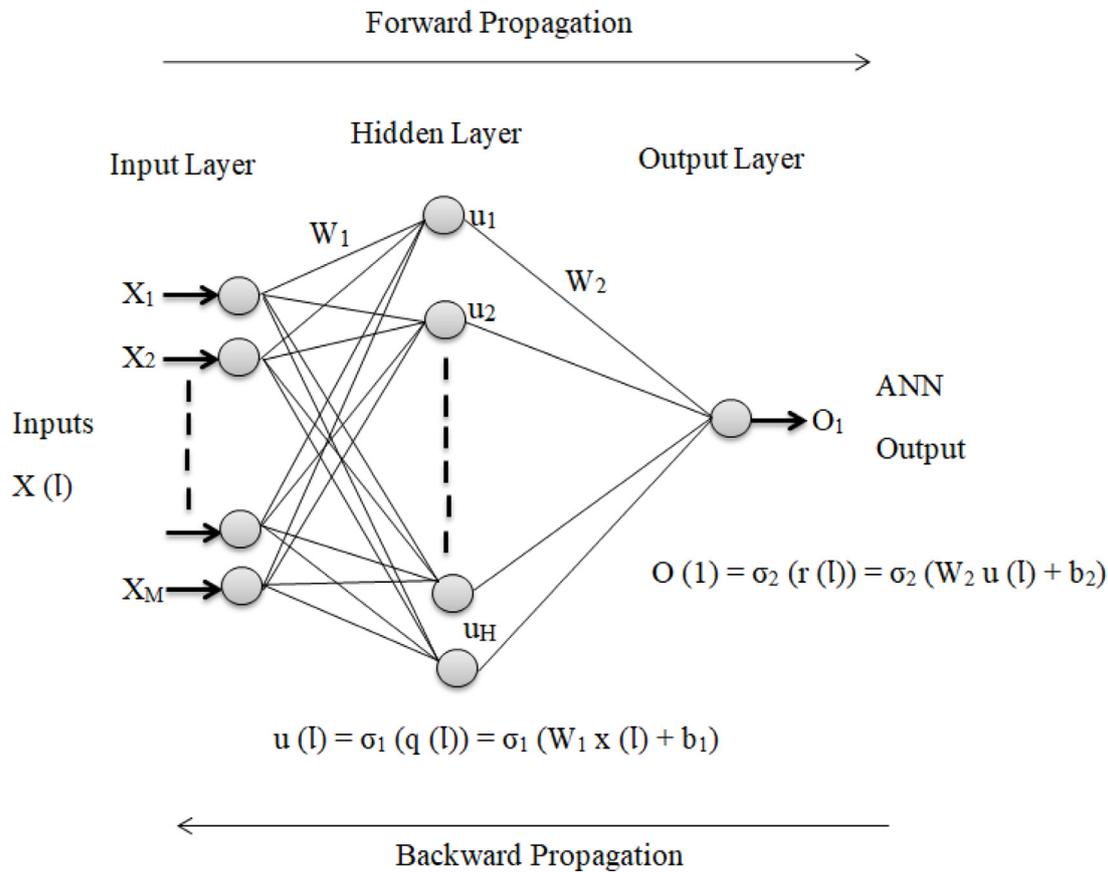
Fig. 4. Structure of the multilayer artificial neural network with the backpropagation algorithm.

way. Traditional computers employ an algorithmic technique to solve problems, and a set of steps are implemented. If the precise procedures that the computer must use are unknown, the computer will be unable to solve the problem. As a result, the problem-solving capabilities of traditional computers are limited to problems that humans already understand and can solve. Conversely, computers would be much more efficient if they could perform tasks that humans cannot. ANNs are analogous to the human brain when it comes to processing information (Maind and Wankar, 2014). Depending on the communication pattern, the network architecture can be separated into two types: a feed-forward neural network (FFNN) and a feed-backward neural network (FBNN) (Gupta, 2013). Feed-forward neural networks permit signals to transit in one direction only, from input to output (Eluyode and Akomolafe, 2013). The first and most simple type of ANN model was the FFNN. Without cycles or loops, information travels from the input layer to any hidden layers and then to the output layer in the FFNN (Abiodun et al., 2019). An input layer, one or more hidden layers, and an output layer of neurons constitute a multilayer feed-forward network. The most popular ANN in use presently is the multilayer ANN (Zhang, 2018).

2.5.1. Backpropagation algorithm

The backpropagation (BP) algorithm is a supervised learning algorithm that has been utilized in prediction and classification apps and is implemented with known data of input and output samples (Hameed et al., 2016). In Fig. 4 below, $\sigma 1$ is the activation function for the hidden layer and $\sigma 2$ is the activation function for the output layer, W1 represents the matrices involving the weights of the connections among the input nodes and the hidden nodes, and W2 represents the matrices involving the weights of the connections among the hidden nodes and the output nodes,

whereas b1 is the bias vector for the nodes in the hidden layer and b2 is the bias vector for the nodes in the output layer (Khan et al., 2020). The forward phase is a two-step procedure. The first involves obtaining the values of the nodes in the hidden layer, and the second step involves using those values to determine the output layer value. The input data are sent to the hidden neurons, where they are multiplied by the weights of the linked neurons. The computed values are then added to one another, and each hidden neuron computes its output using an activation function. The network spreads forward once the values of the hidden layer are computed, transmitting values from the hidden neurons to the output neuron. To calculate the current network error, the network result is compared to the corresponding output pattern (Cilimkovic, 2015). In the backpropagation phase, the weights correction is calculated using the specified minimum locating rule and the correction is implemented into the weights of the layer (Gallo, 2015).

2.5.2. Performance measurement

The output of the ANN was evaluated in terms of the mean squared error (MSE) and the correlation coefficient (R) values. The MSE performed well and the calculation was straightforward. The difference between both items of the output vector and goal vector was used to calculate the MSE, as shown in Eq. (5) (Kubat, 2015). The R was computed using Eq. (6) (Kulisz et al., 2021).

$$MSE = \frac{\sum_{i=1}^{n}(t_i - y_i)^2}{n} \tag{5}$$

$$R = \frac{\sum_{i=1}^{n}(t_i - \bar{t})(y_i - \bar{y})}{\left(\sqrt{\sum_{i=1}^{n}(t_i - \bar{t})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}\right)} \tag{6}$$
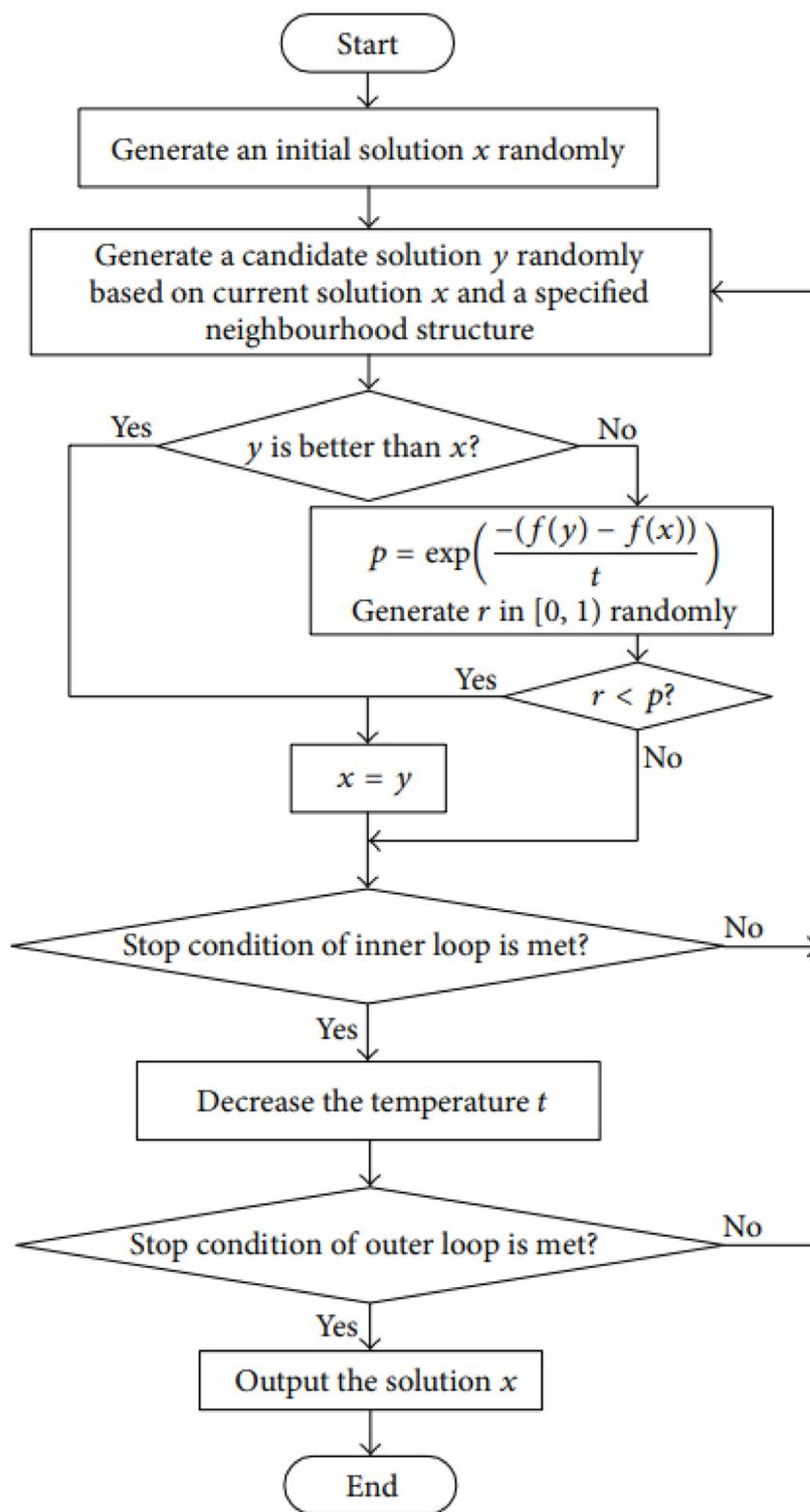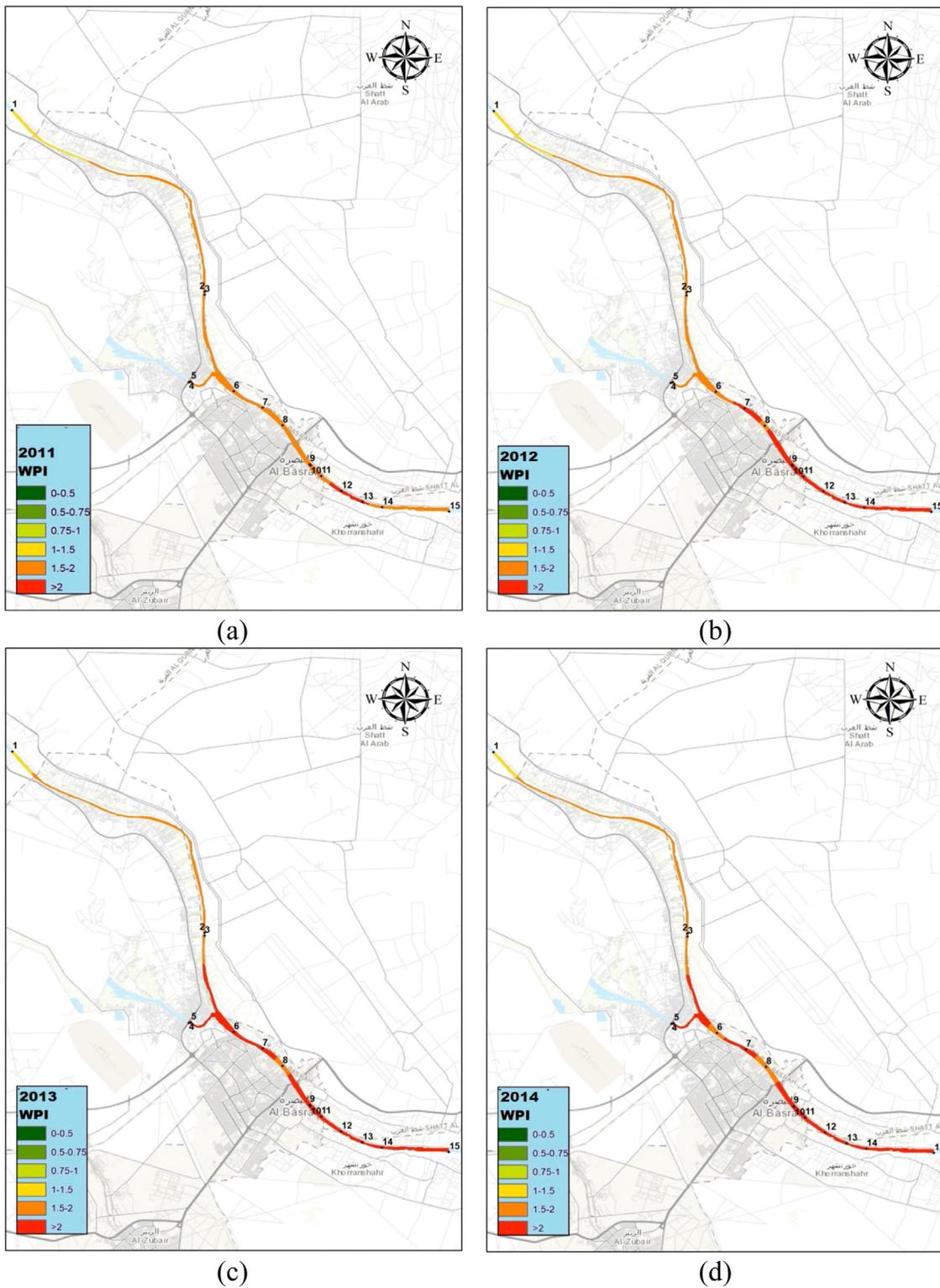
Fig. 5. Flowchart of simulated annealing algorithm (Zhan et al., 2016).

Fig. 6. Annual mean values of the water pollution index (WPI) for the water treatment plants (WTPs) investigated in this study during (a) 2011, (b) 2012, (c) 2013, and (d) 2014.
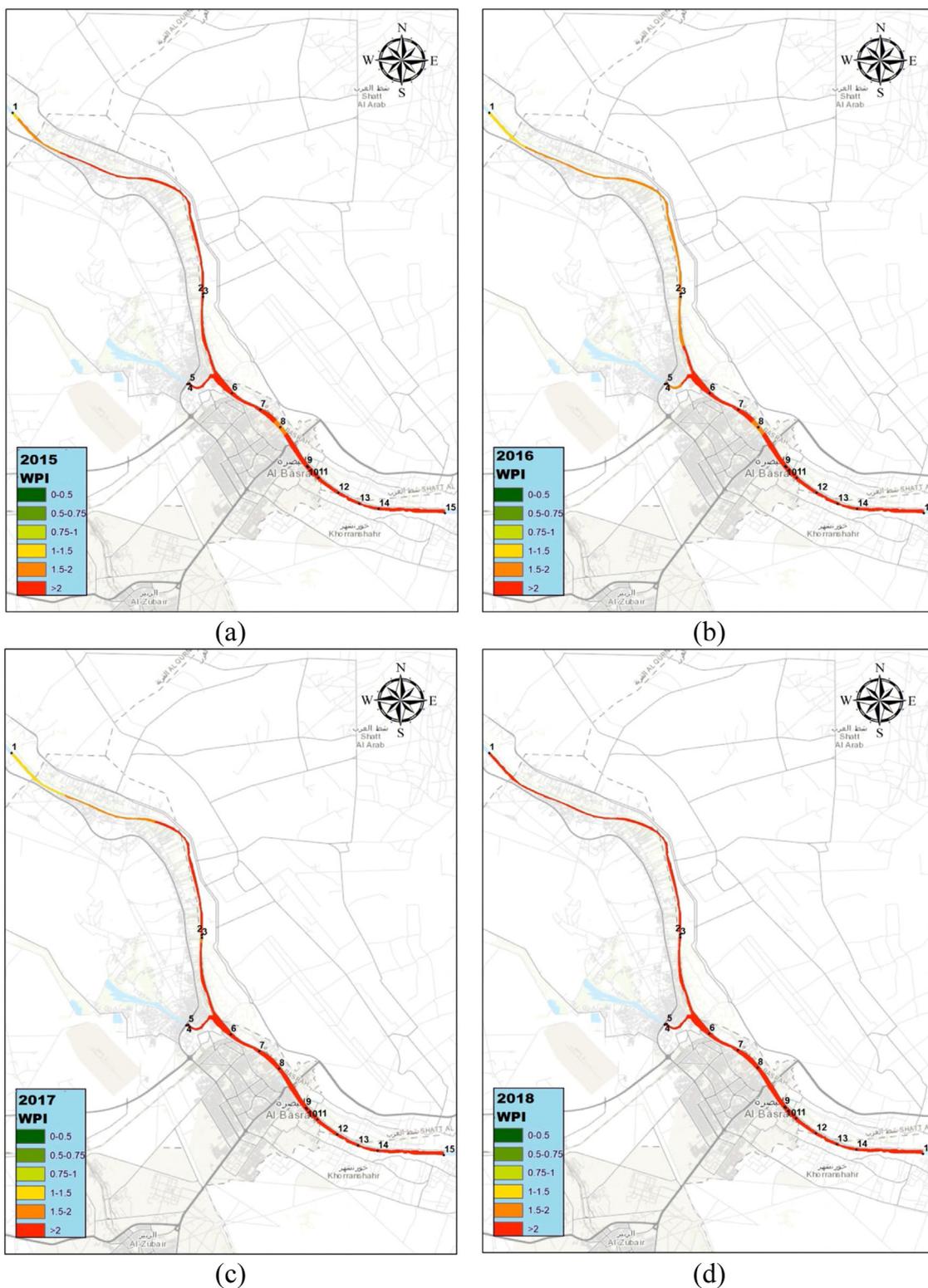
Fig. 7. Annual mean values of the WPI for the WTPs investigated in this study during (a) 2015, (b) 2016, (c) 2017, and (d) 2018.
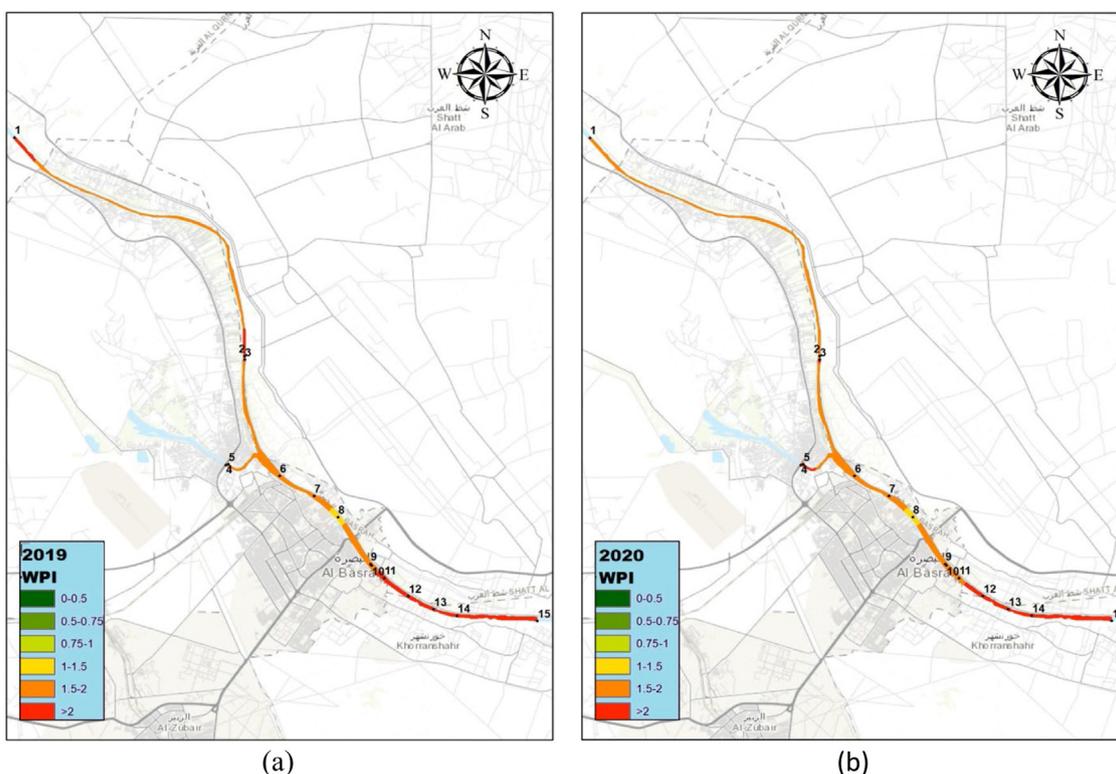
Fig. 8. Annual mean values of the WPI for the WTPs investigated in this study during (a) 2019 and (b) 2020.

WTPs	WPI for the year									
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1	1.4131	1.2845	1.4489	1.4401	1.4903	1.3606	1.1044	2.2657	2.0107	1.6146
2	1.5343	1.5973	1.8006	1.8267	2.8651	1.8421	2.0637	3.748	2.1908	1.7977
3	1.5359	1.677	1.9188	1.7748	3.0545	1.9583	1.9719	3.4401	1.8152	2.0453
4	1.5389	1.7947	2.5979	2.3049	3.4918	2.1802	2.2214	4.9532	1.6755	1.8812
5	1.5242	1.7304	2.0666	2.3096	3.4666	1.8119	2.2043	3.7436	1.7419	2.1368
6	1.5169	1.6453	2.3879	1.9419	3.4013	2.1517	2.2472	4.9107	1.9409	1.8178
7	1.6732	2.1374	2.2415	2.3393	2.8708	2.141	2.2349	3.8809	1.5051	1.5285
8	1.6214	1.9793	1.8382	1.4986	1.6952	1.9472	2.2045	3.0729	1.4287	1.4226
9	1.7189	2.2198	2.2747	2.1817	3.211	2.2853	2.3913	6.0261	1.9512	1.903
10	1.7129	2.2197	2.3607	2.2859	2.9667	2.0393	2.2934	5.7628	1.9049	1.8775
11	1.9444	2.7175	2.8225	2.5647	3.5805	2.3243	2.4649	5.8843	2.0662	1.8561
12	2.1759	2.7026	3.2634	3.0563	3.6339	2.9475	2.991	6.9553	2.9476	3.0249
13	2.0194	3.1313	3.1735	2.7887	4.3888	2.6474	2.8694	6.0943	2.5202	2.6797
14	1.9257	2.2282	2.3112	2.8107	3.4469	2.4354	2.3294	5.9148	2.0529	2.4778
15	1.9733	2.3322	3.0218	2.9281	3.6734	2.2765	2.16	6.3578	2.0516	2.2917

Color						
WPI value	< 0.5	0.5 – 0.75	0.75 – 1	1-1.5	1.5-2	> 2
Category of water	Excellent	Good	Moderately	Highly polluted		

Fig. 9. WPI annual values of WTPs for all study years (2011–2020).

Where n is the number of data, t_i is the target data, y_i is the network predicted, $\bar{t}$ is the average value of the target data, and $\bar{y}$ is the average value of the output network.

2.6. Simulated annealing

Simulated annealing is a simulation of the physical gradual cooling operation, also known as annealing, which generates better crystals and improves the strength properties of a metal. The transition method between the states and the cooling timetable are the two primary steps in SA, and the goal is to determine the state with the least amount of energy. Using precise temperature control during the annealing process, a perfect crystal can be formed. Kirkpatrick et al. (1983) and Cerny (1985) independently proposed this optimization technique; they discovered that alternate physical states of matter are similar to the solution space for an optimization method and that the fitness function for an
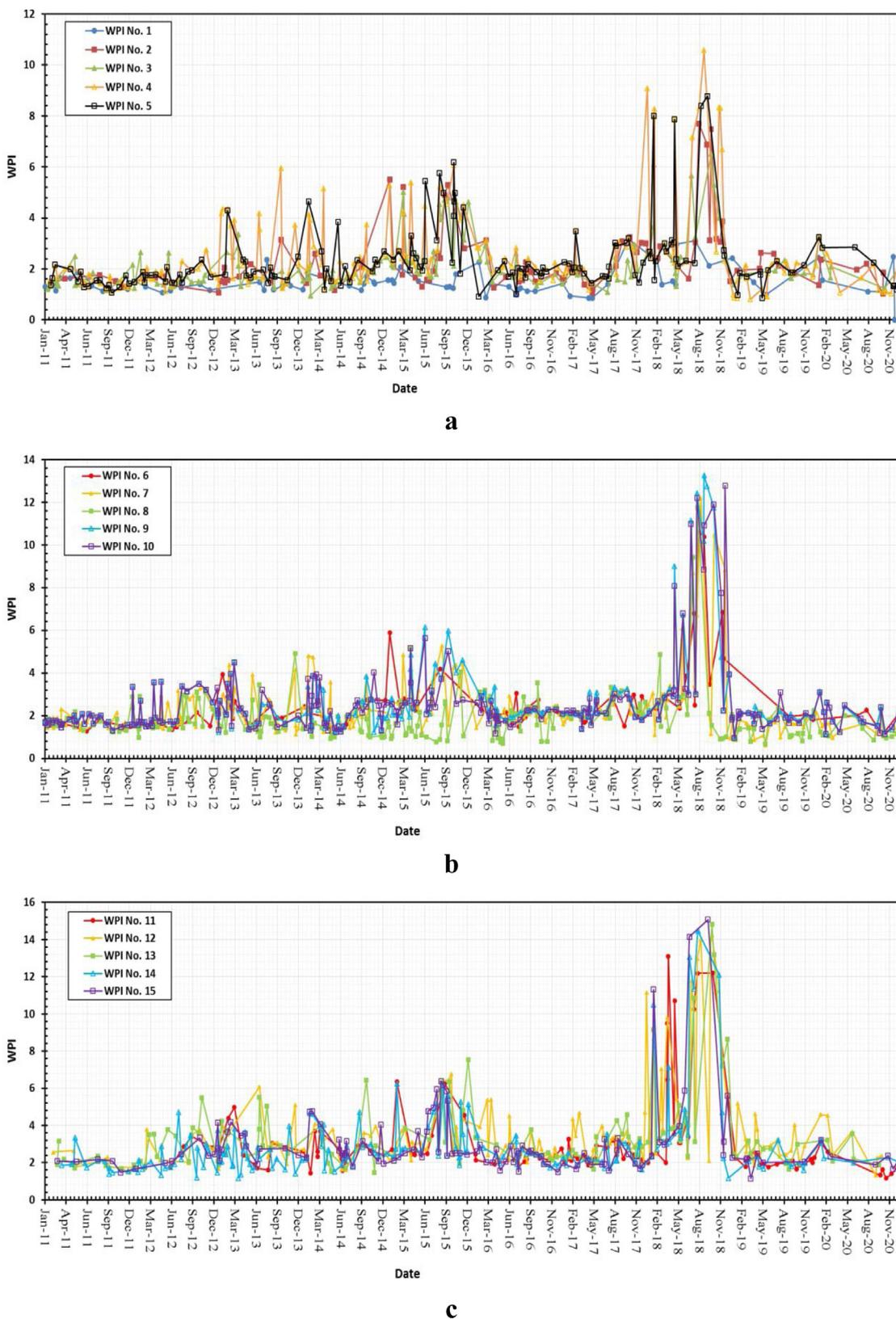
Fig. 10. Monthly values of the WPI from 2011 to 2020 for (a) WTP No. 1 to WTP No. 5, (b) WTP No. 6 to WTP No. 10, and (c) WTP No. 11 to WTP No. 15.

optimization method matches the free energy of the material (El-Naggar et al., 2012). The initial temperature is selected to include the entire solution space. Depending on the type of issue, the initial temperature is determined through trial and error (Suman and Kumar, 2006). The most widely utilized procedure for the SA algorithm is as follows (Sibalija, 2018):

Fig. 11. Correlation coefficients for the water parameters and the WPI.
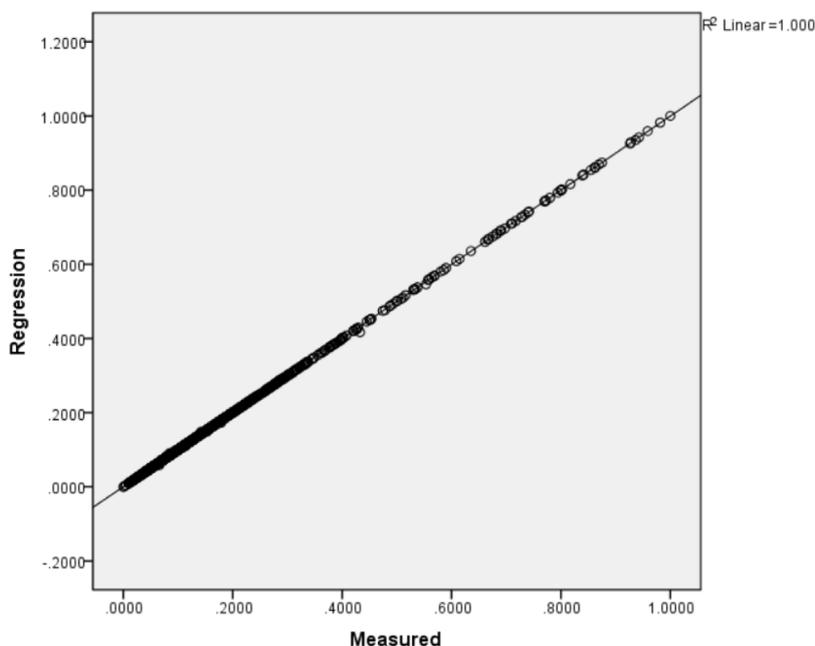

Fig. 12. The results of the regression model compared to the results of the measured data.

1. Adjust the parameters specific to the algorithm, such as the start point, the initial temperature, the annealing schedule (which includes both the annealing and temperature update functions), the reannealing period, and the termination condition (for example, the ultimate temperature, the number of iterations, the alteration in the objective function).

2. Compute the objective function of the initial point, f(x).

3. Compute a new neighbor point, defined by a probability distribution commensurate to the existing temperature, and calculate the objective function of the new point, f(x).

4. Compute the difference in the objective functions among the new and existing points (ΔE).

5. Accept the new point if ΔE is less than zero. If $\Delta E \geq 0$, generate a stochastic number r into the range (0, 1) and verify if $r \leq \exp(-\frac{\Delta E}{T})$. If yes, continue to step 6. If not, start with a new point and move on to step 3.

6. Reduce the temperature regularly using the provided temperature update function. If the present temperature is greater than the final temperature, proceed to step 3. If not, proceed to step 7.

7. Terminate the operation if the termination criteria are satisfied (for example, if the present temperature is less than
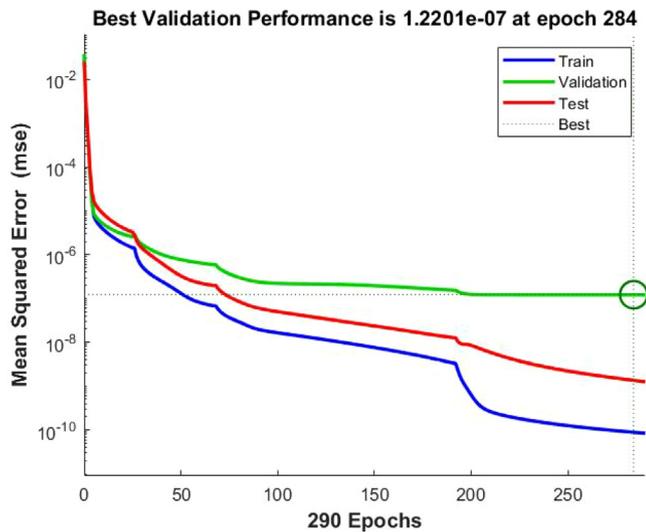
Fig. 13. MSE vs. epoch numbers for the proposed ANN model.

Table 9
Values of coefficient from MLR analysis for twelve independent parameters and WPI as dependent variable.

Model of MLR	Unstandardized coefficients		t	Significant
	B	Standard Error		
Constant	−0.019	0.001	−31.485	0
K^+	0.062	0.002	33.776	0
Na^+	0.219	0.005	41.767	0
TDS	0.233	0.023	10.327	0
SO_4^{-2}	0.109	0.003	33.056	0
Cl^-	0.140	0.011	12.397	0
Mg^{+2}	0	0.011	0.032	0.975
TH	0	–	0.039	0.969
Ca^{+2}	0.129	0.011	11.385	0
Alk.	0.005	0.001	4.053	0
EC	0.213	0.022	9.781	0
pH	0.007	0.001	9.863	0
Tur.	0.100	0.001	101.498	0

or equal to the set final temperature, the number of iterations is completed, and the alteration in the objective function is lower than the stated value).

2.7. Normalization of the data

Normalization is the process of converting data into a certain range, such as between 0 and 1 or −1 and 1 (Ali and Faraj, 2014). The normalization of all the input and output data is the initial step in the calculation. Eq. (7) is the normalization equation for data of the input values within the range of 0 to 1, whereas Eq. (8) is the normalization equation for data of the output values within the range of 0 to 1 (Zain et al., 2009).

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{7}$$

$$t_n = \frac{t - t_{min}}{t_{max} - t_{min}} \tag{8}$$

Where X is the input variable, X_n denotes the normalized value for the variable X, X_{min} is the lowest value of the variable X, and X_{max} is the highest value of the variable X. While t is the output variable, t_n denotes the normalized value of the variable t, t_{min} is the lowest value of the variable t, and t_{max} is the highest value of the variable t.

3. Results and discussion

3.1. Water pollution index

The water of the SAR is highly polluted according to the classification of the WPI for all the WTPs sampled in this study between 2011 and 2020, as shown in Figs. 6 to 9. These figures illustrate the annual mean values of the WPI for all the WTPs investigated in this study. It was observed that 2018 was the most polluted year for all the sampling stations. The salinity in the SAR showed significant changes in 2018 due to a sharp reduction in the flow and a salinity intrusion that extended to the upper reaches of the river. According to the data obtained from the Basrah Water Directorate, the maximum values of the TDS and EC in 2018 reached 22954 mg/l and 34030 μs/cm, respectively.

The variations in the measured parameters and the WPI are shown in Figs. A.1 to A.3 in Appendix.

Fig. 10 depicts the monthly values of the WPI from 2011 to 2020 for each WTP that was investigated in this study. WTP No. 7 (Al-Jubailah 1) showed the lowest WPI value (0.634543 in the year 2019), while WTP No. 15 (Al-Labanie) showed the highest WPI value (15.06927 in the year 2018). The most polluted WTPs were located in the Abu Al Kaseeb area (WTP No. 11, 12, 13, 14, and 15). Extremely high pollution levels were observed in 2018. The pollution began to increase from the sixth month until the end of the year, with the greatest increase in pollution occurring in the ninth month. Fig. 11 illustrates the correlations between the water parameters and the WPI. The results showed strong correlations between the EC, Na, TDS, SO_4, Cl, and Ca and the WPI.

3.2. Prediction of the WPI by MLR

The aim of MLR analysis is to forecast the value of the dependent variable given a set of predictor variables (independent variables).

In the present study, the WPI was the dependent variable, while the TDS, Tur, Mg^{+2}, pH, K^+, Cl^-, Na^+, EC, SO_4^{-2}, TH, Ca^{+2}, and Alk were the independent variables. The SPSS software program was used to analyze the data, the multiple correlation coefficient (R) of our model was 0.999 and the coefficient of determination (R^2) was 0.998. This model had a success rate of 99.8% and an error of 0.2%.

The data presented in Table 9 shows that most of the independent variables (TDS, Tur, pH, K^+, Cl^-, Na^+, EC, SO_4^{-2}, Ca^{+2}, and Alk) showed a significant effect on the WPI (p < 0.001; the significance level was 0.05). However, the Mg^{+2} showed no effect on the WPI ($p = 0.975$). Additionally, the TH showed no effect on the WPI ($p = 0.969$).

Therefore, the variables that were used to predict the WPI were TDS, Tur, pH, K^+, Cl^-, Na^+, EC, SO_4^{-2}, Ca^{+2}, and Alk. The SPSS program was used to analyze the data. The R of our model was 1, and the R^2 was 1, indicating that it was optimal for forecasting. The data presented in Table 10 indicate that all the independent variables showed a significant effect on the WPI (p < 0.001). In this study, the MLR analysis estimated the regression function to be:

$$\begin{aligned} WPI = {} & -0.020 + 0.067 \, K^+ + 0.237 \, Na^+ + 0.248 \, TDS \\ & + 0.118 \, SO_4^{-2} + 0.156 \, Cl^- + 0.44 \, Ca^{+2} + 0.006 \, Alk \\ & + 0.225 \, EC + 0.007 \, pH + 0.019 \, Tur \end{aligned} \tag{9}$$

Two of the variables showed a positive correlation, as presented in Table 11. In Fig. 12, a comparison of the predicted results of the MLR model and the measured results is depicted.
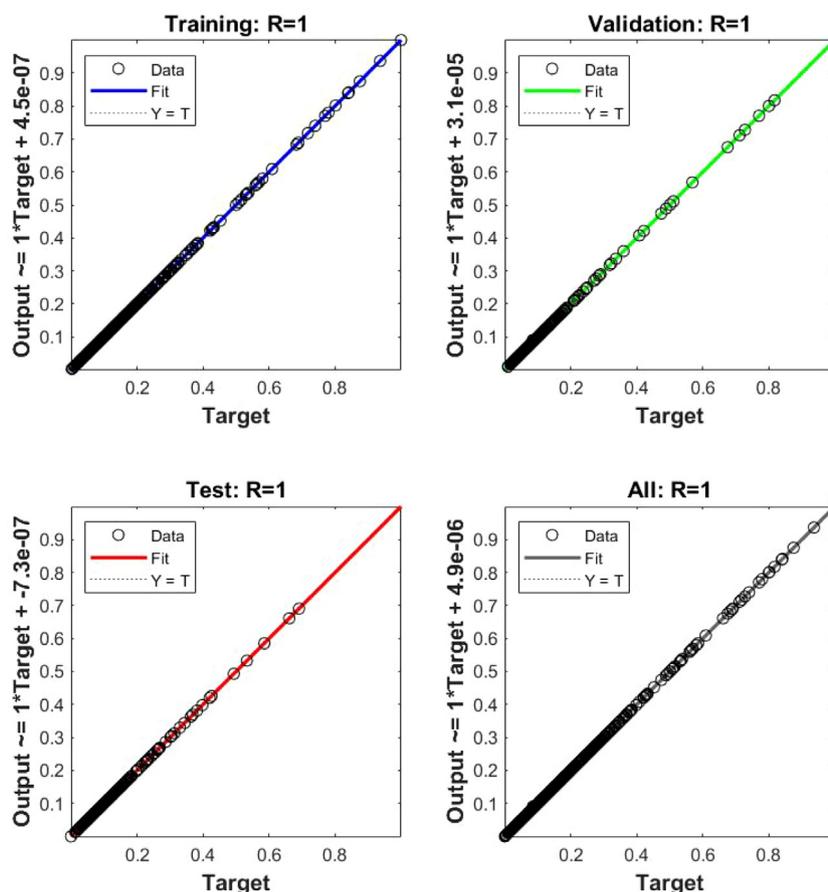
Fig. 14. Regression analysis for the proposed ANN model.

Table 10
Values of coefficient from MLR analysis for ten independent parameters and WPI as dependent variable.

Model of MLR	Unstandardized coefficients		t	Significant.
	B	Standard Error		
Constant	−0.020	0	−248.831	0
K^+	0.067	0	273.184	0
Na^+	0.237	0.001	337.755	0
TDS	0.248	0.003	81.966	0
SO_4^{-2}	0.118	0	266.611	0
Cl^-	0.156	0.002	104.265	0
Ca^{+2}	0.044	0.001	51.752	0
Alk.	0.006	0	32.315	0
EC	0.225	0.003	77.085	0
pH	0.007	0	76.702	0
Tur.	0.109	0	823.701	0

3.3. Prediction of the WPI by the ANN

An ANN with a backpropagation algorithm was chosen for this study. The following 10 variables were used as input variables for the ANN model: Tur, Alk, pH, EC, TDS, Na^+, K^+, Ca^{+2}, Cl^-, and SO_4^{-2}. One variable was selected as the output: the WPI. The total number of data points used in this research was 2430. These data were divided into three categories: 75% for the training set, 10% for the validation set, and 15% for the testing set since this ratio generated the optimum results regarding the lowest MSE and highest R values. Therefore, a dataset containing 1823 values was selected as the training set, a dataset containing 243 values was selected as the validation set and a dataset containing 364 values was selected as the testing set (Salami et al., 2016).

One of the important aspects in the design of an ANN is the identification of the appropriate number of hidden layers and hidden neurons; a trial-and-error method was used in the present study because no universal rule exists for determining the characteristics of the hidden layer and neurons. The network size that resulted in the minimum MSE and maximum regression in the training, validation, and testing sets was chosen as the optimal network. Tables 12 and 13 present the best results of an ANN model that employed a single layer and two hidden layers, respectively, with different numbers of nodes and epochs in which the training functions provided the minimum MSE and maximum R values for the training set, validation set, and testing set.

Tables 12 and 13, present the ideal neural network architecture employed in the present study. The Levenberg–Marquardt algorithm (trainlm) was employed to train this network because this algorithm provided the best results regarding the minimum MSE and maximum R values compared with the other algorithms. The network consisted of an input layer, one hidden layer, and the output layer; the hidden layer had 17 neurons. The logsig activation function was employed between the input layer and the hidden layer, whereas the purelin was utilized between the hidden layer and the output layer. This optimal network structure provided the best results, with an MSE of 8.851×10^{-11}, 1.220×10^{-7}, and 1.354×10^{-9} for the training, validation, and testing sets, respectively (Fig. 13), and an R value of 1, 1, and 1 for the training, validation, and testing sets, respectively (Fig. 14). The ideal neural network architecture employed in this study is depicted in Fig. 15.
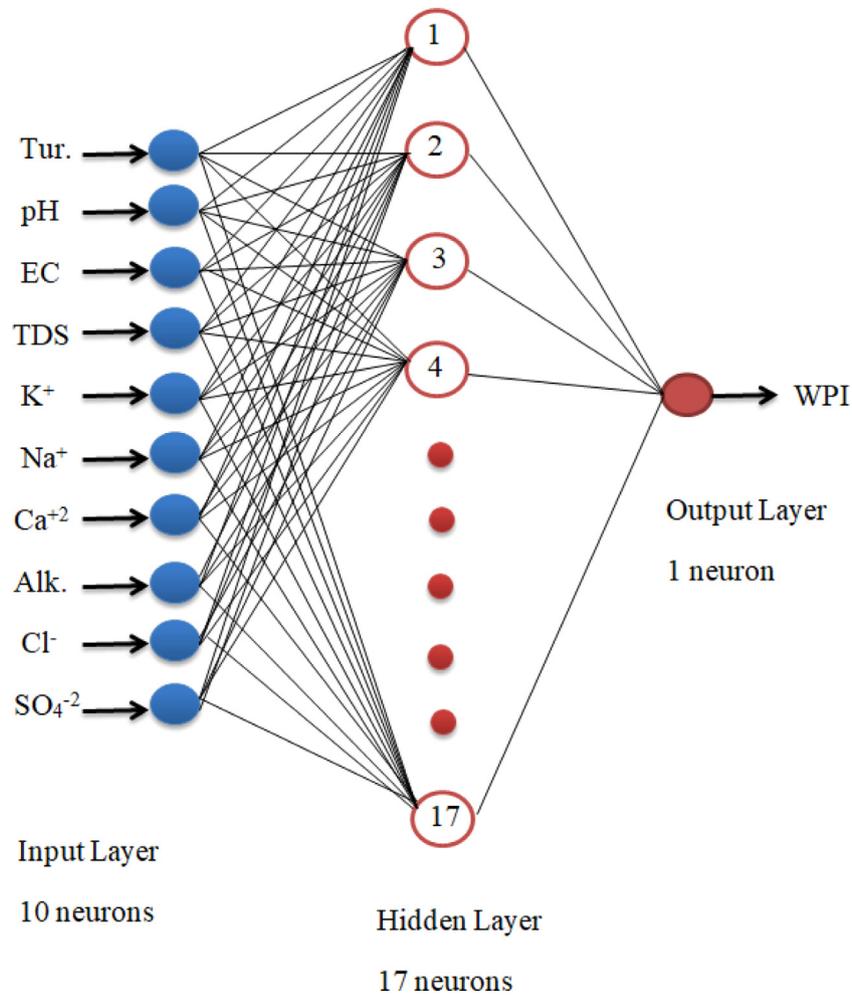
Fig. 15. Structure of the suggested ANN model.

Table 11
Correlation and statistic.

Variables	N	Mean	Std. deviation	Std. error mean	MSE	Correlation
Measured	2430	0.13390	0.11979	0.00243	4.746×10^{-7}	1
Regression	2430	0.13406	0.11973	0.00242		

Table 12
Minimum (MSE) and maximum (R) for the training functions, with one hidden layer.

Training functions	Number of nodes	Training set's MSE value	Validation set's MSE value	Testing set's MSE value	Testing set's R value	Number of Epochs
trainbfg	10	1.933×10^{-5}	5.893×10^{-5}	2.923×10^{-5}	0.9983	91
traincgb	20	2.443×10^{-5}	1.334×10^{-5}	3.409×10^{-5}	0.9987	86
traincgf	19	1.380×10^{-5}	2.359×10^{-5}	1.621×10^{-5}	0.9992	230
traincgp	16	1.432×10^{-4}	1.476×10^{-4}	1.263×10^{-4}	0.992	69
traingdm	20	0.0018	0.0012	0.0026	0.8884	5000
traingda	13	0.0010	8.710×10^{-4}	8.995×10^{-4}	0.9356	183
traingdx	17	5.654×10^{-4}	5.841×10^{-4}	4.989×10^{-4}	0.9747	154
trainlm	17	8.851×10^{-11}	1.220×10^{-7}	1.354×10^{-9}	1	290
trainoss	21	4.904×10^{-5}	5.668×10^{-5}	1.063×10^{-4}	0.9967	101
trainrp	13	8.115×10^{-5}	8.889×10^{-5}	9.315×10^{-5}	0.9933	278
trainscg	19	1.101×10^{-5}	2.023×10^{-5}	1.319×10^{-5}	0.9993	197

3.4. Simulated annealing optimization solution

The regression model presented in Eq. (9) was selected as the fitness function of the SA solution and in the following way:

Minimize : WPI (K^+, Na^+, TDS, SO_4^{-2}, Cl^-, Ca^{+2}, Alk, EC, pH, Tur) $= \min(-0.020 + 0.067\ K^+ + 0.237\ Na^+ + 0.248\ TDS + 0.118\ SO_4^{-2} + 0.156\ Cl^- + 0.044\ Ca^{+2} + 0.006\ Alk + 0.225\ EC + 0.007\ pH + 0.019\ Tur)$ (10)

The minimization of the fitness function value was restricted by the constraints of the independent variables. The range of

Table 13
Minimum (MSE) and maximum (R) for the training functions, with two hidden layer.

Training functions	Number of nodes	Training set's MSE value	Validation set's MSE value	Testing set's MSE value	Testing set's R value	Number of Epochs
trainbfg	(11 28)	3.474×10^{-6}	2.757×10^{-6}	4.713×10^{-6}	0.9998	136
traincgb	(12 12)	9.531×10^{-6}	1.101×10^{-5}	1.953×10^{-5}	0.9992	105
traincgf	(17 13)	1.006×10^{-5}	1.205×10^{-5}	1.966×10^{-5}	0.9989	237
traincgp	(11 16)	1.143×10^{-4}	1.203×10^{-5}	7.916×10^{-5}	0.9952	63
traingdm	(10 10)	7.032×10^{-4}	4.229×10^{-4}	5.199×10^{-4}	0.9704	5000
traingda	(14 21)	9.406×10^{-4}	0.0029	6.509×10^{-4}	0.9562	438
traingdx	(12 10)	2.427×10^{-4}	2.473×10^{-5}	2.251×10^{-4}	0.9912	263
trainlm	(10 10)	2.510×10^{-8}	2.783×10^{-9}	1.057×10^{-8}	1	177
trainoss	(20 10)	8.979×10^{-5}	1.334×10^{-4}	9.903×10^{-5}	0.9952	79
trainrp	(15 14)	2.909×10^{-5}	2.987×10^{-4}	3.443×10^{-5}	0.9975	198
trainscg	(16 20)	6.592×10^{-5}	8.196×10^{-5}	7.133×10^{-5}	0.9958	108
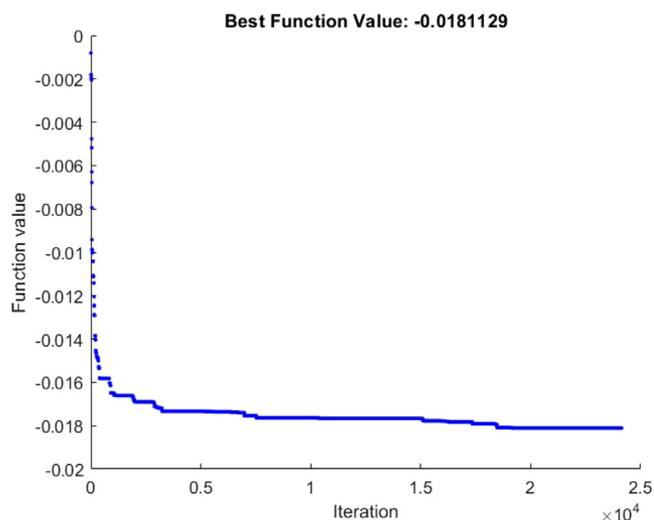
Fig. 16. Best fitness value and mean fitness value for the SA.
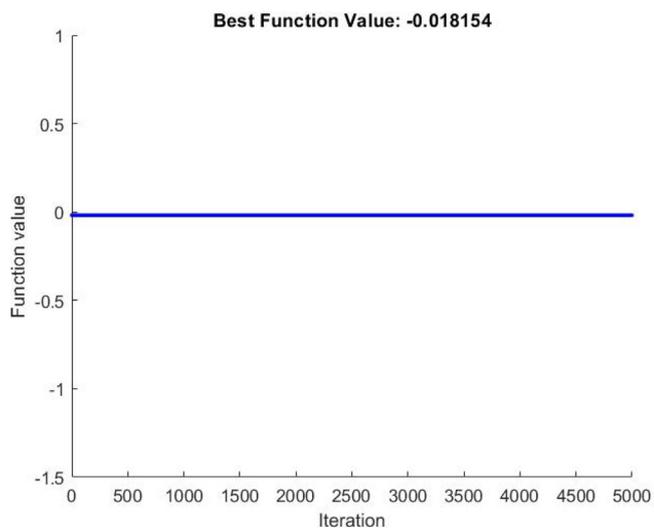
Fig. 17. Best fitness value and mean fitness value for the ANN-SA.

Table 14
Limitations and start points of independent variables (SA).

Parameters	Range (normalized values)	Start point
K^+	$0 \leq K^+ \leq 1$	0.008
Na^+	$0 \leq Na^+ \leq 1$	0.003
TDS	$0 \leq TDS \leq 1$	0.017
SO_4^{-2}	$0 \leq SO_4^{-2} \leq 1$	0.005
Cl^-	$0 \leq Cl^- \leq 1$	0.005
Ca^{+2}	$0 \leq Ca^{+2} \leq 1$	0.007
Alk.	$0 \leq Alk. \leq 1$	0.165
EC	$0 \leq EC \leq 1$	0.002
pH	$0 \leq pH \leq 1$	0.333
Tur.	$0 \leq Tur. \leq 1$	0.080

Table 15
Conditions to define limitations of integrated ANN-SA (Zain et al., 2011).

Condition	Decision	
	Lower limit	Upper limit
(Opt-ANN) < (Opt-SA)	Opt-ANN	Opt-SA
(Opt-ANN) > (Opt-SA)	Opt-SA	Opt-ANN

Table 16
Limitations and start points of independent variable (ANN-SA).

Parameter	Range (normalized values)	Start point
K^+	$0 \leq K^+ \leq 0.008$	0
Na^+	$0 \leq Na^+ \leq 0.003$	0
TDS	$0 \leq TDS \leq 0.017$	0
SO_4^{-2}	$0 \leq SO_4^{-2} \leq 0.005$	0
Cl^-	$0 \leq Cl^- \leq 0.005$	0
Ca^{+2}	$0.001 \leq Ca^{+2} \leq 0.007$	0.001
Alk.	$0.116 \leq Alk. \leq 1.165$	0.116
EC	$0 \leq EC \leq 0.002$	0
pH	$0.158 \leq pH \leq 0.333$	0.158
Tur.	$0 \leq Tur. \leq 0.080$	0

values for the measured independent variables (normalized values) was taken to represent the limitations of the optimization solution, as presented in Table 14. Additionally, the parameters that resulted in the lowest WPI of the regression model were selected as the start points for the SA solution, as presented in Table 14 below.

By employing the fitness function in Eq. (10), the limitations of the independent variables in Table 14, the start points in Table 14, and the MATLAB optimization toolbox were used to determine the lowest value of the WPI at the best points. Several trials were conducted to obtain the minimum WPI values using the MATLAB optimization toolbox.

Fig. 16 presents the optimal results of the MATLAB optimization toolbox. As shown in Fig. 16, the minimum observed WPI value was −0.01811. The set values (normalized values) of the independent parameters that resulted in the lowest WPI value were 0 for K^+, 0 for Na^+, 0 for TDS, 0 for SO_4^{-2}, 0 for Cl^-, 0.001 for Ca^{+2}, 0.116 for Alk, 0 for EC, 0.158 for pH, and 0 for Tur. The optimal solution was indicated at iteration number 24141 of the
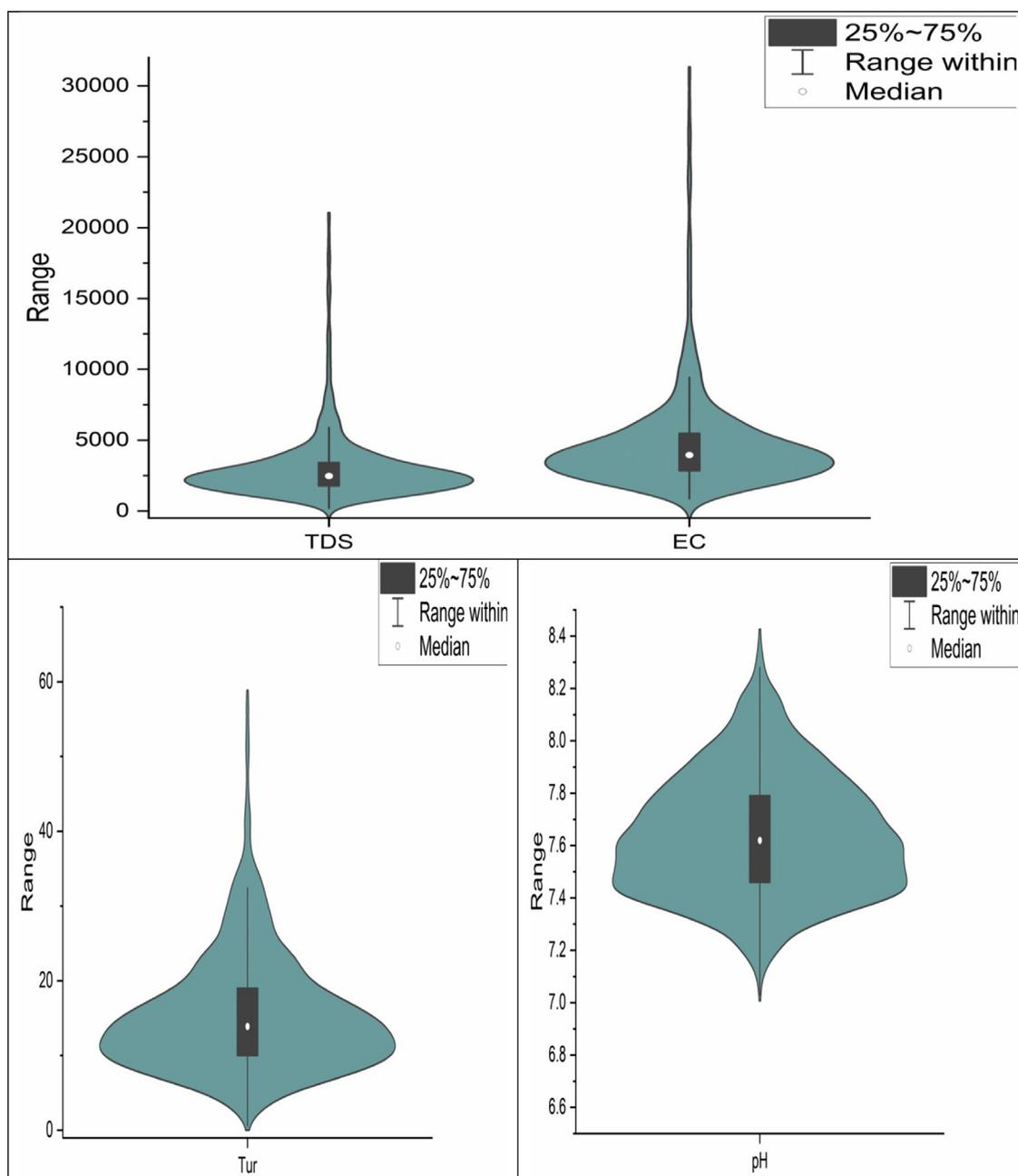
Fig. A.1. Variations of the TDS, EC, Tur., and pH parameters.

simulated annealing, the annealing function used was Boltzmann annealing, and the initial temperature was 0.04 °C.

3.5. Integration of the ANN with the SA (ANN-SA)

The optimum independent variable values of the SA were merged with the nonoptimal independent variable values of the ANN method to set the upper and lower limits for the optimization solution was the procedure used to apply the merged ANN-SA technique. The nonoptimal values of the parameters that resulted in the lowest predicted WPI value in the ANN model were $K^+ = 0.008$, $Na^+ = 0.003$, TDS $= 0.017$, $SO_4^{-2} = 0.005$, $Cl^- = 0.004$, $Ca^{+2} = 0.007$, Alk $= 0.165$, EC $= 0.002$, pH $= 0.333$, Tur $= 0.080$. The ideal values of the parameters from the SA were $K^+ = 0$, $Na^+ = 0$, TDS $= 0$, $SO_4^{-2} = 0$, $Cl^- = 0$, $Ca^{+2} = 0.001$, Alk $= 0.116$, EC $= 0$, pH $= 0.158$, Tur $= 0$.

Classification of the nonoptimal parameter values for the ANN model (Opt-ANN) and the optimum parameter values for the SA (Opt-SA) are presented in Table 15, under two conditions (Zain et al., 2011). The second condition was met.

The limitations of K^+, Na^+, TDS, SO_4, Cl^-, Ca^{+2}, Alk, E.C, pH and Tur., parameters, respectively, are presented in Table 16. To identify the start points for the merged ANN-SA, the ideal parameter values that resulted in the lowest WPI for the SA were chosen, as presented in Table 16.

By employing the fitness function in Eq. (10), the limitations of the parameters in Table 16, the start points in Table 16, and the MATLAB optimization toolbox were used to obtain the minimum values of the WPI at the optimal points.

The minimum observed WPI value was -0.01815, as shown in Fig. 17. The specified values of the independent parameters that
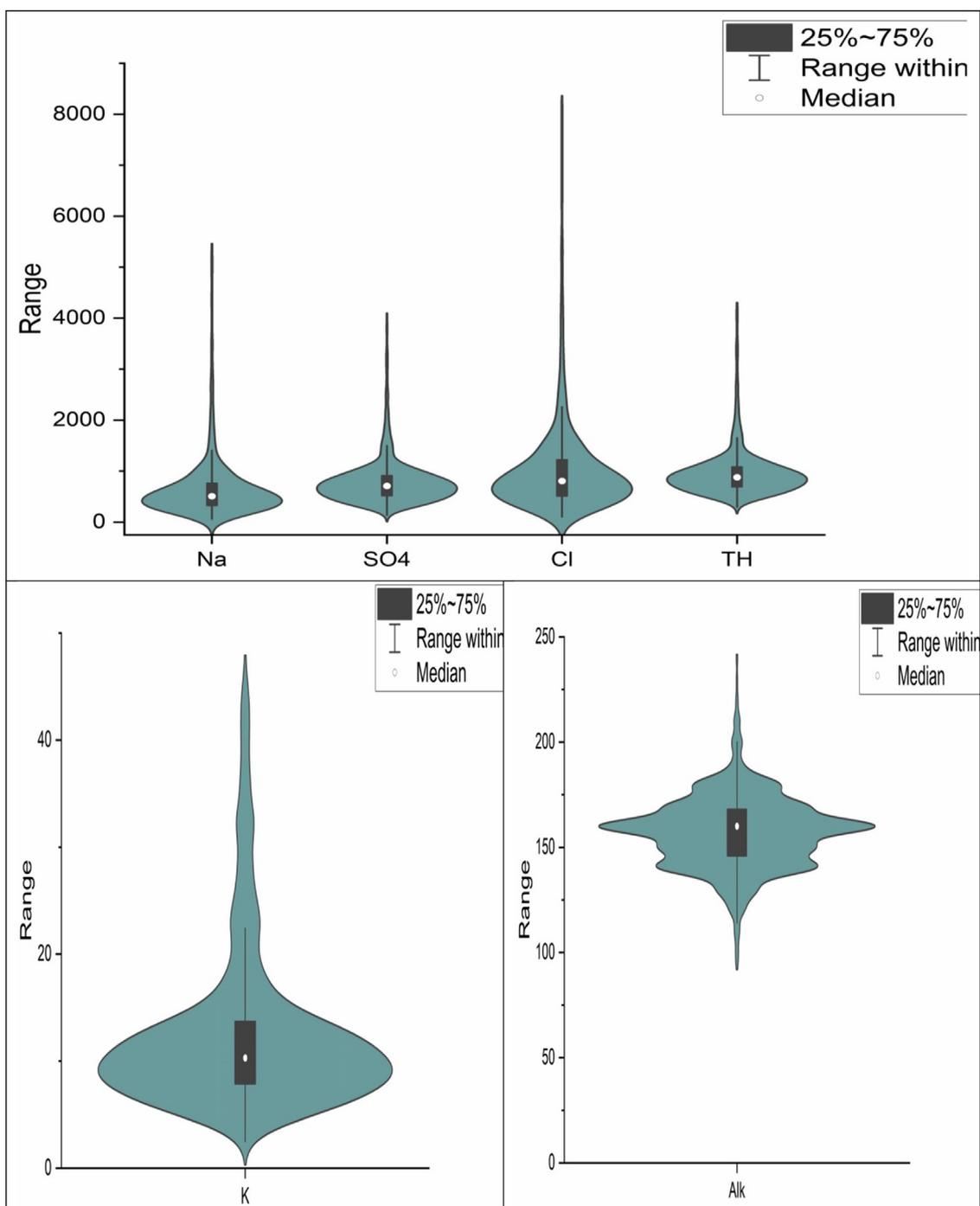
Fig. A.2. Variations of the Na, SO$_4$, Cl, TH, K, and Alk. Parameters.

resulted in the minimum WPI value were 0 for K$^+$, 0 for Na$^+$, 0 for TDS, 0 for SO$_4^{-2}$, 0 for Cl$^-$, 0.001 for Ca^{+2}, 0.116 for Alk, 0 for EC, 0.158 for pH, and 0 for Tur. The optimal solution was indicated at iteration number 5000 of the ANN-SA.

3.6. Results

In contrast to the results obtained from the measured data, MLR analysis, ANN technique, and SA technique, the combined ANN-SA provided the lowest WPI value at the optimal parameters. The minimum WPI value for the integrated ANN-SA was 0.373, as presented in Table 17.

4. Conclusion

The WPI is beneficial to the public in terms of assessing the state of the surface water quality and may be utilized as a powerful tool in the development of pollution control plans regarding treatment requirements at various levels. The water of the SAR is highly polluted according to the classification of the WPI for all the WTPs investigated in this study between 2011 and 2020. The effectiveness of the MLR and ANN models at predicting the WPI in the SAR were investigated. The results demonstrated that these models were quite effective at predicting the WPI (the value of the R for both models was 1). The 10-17-1 network structure was
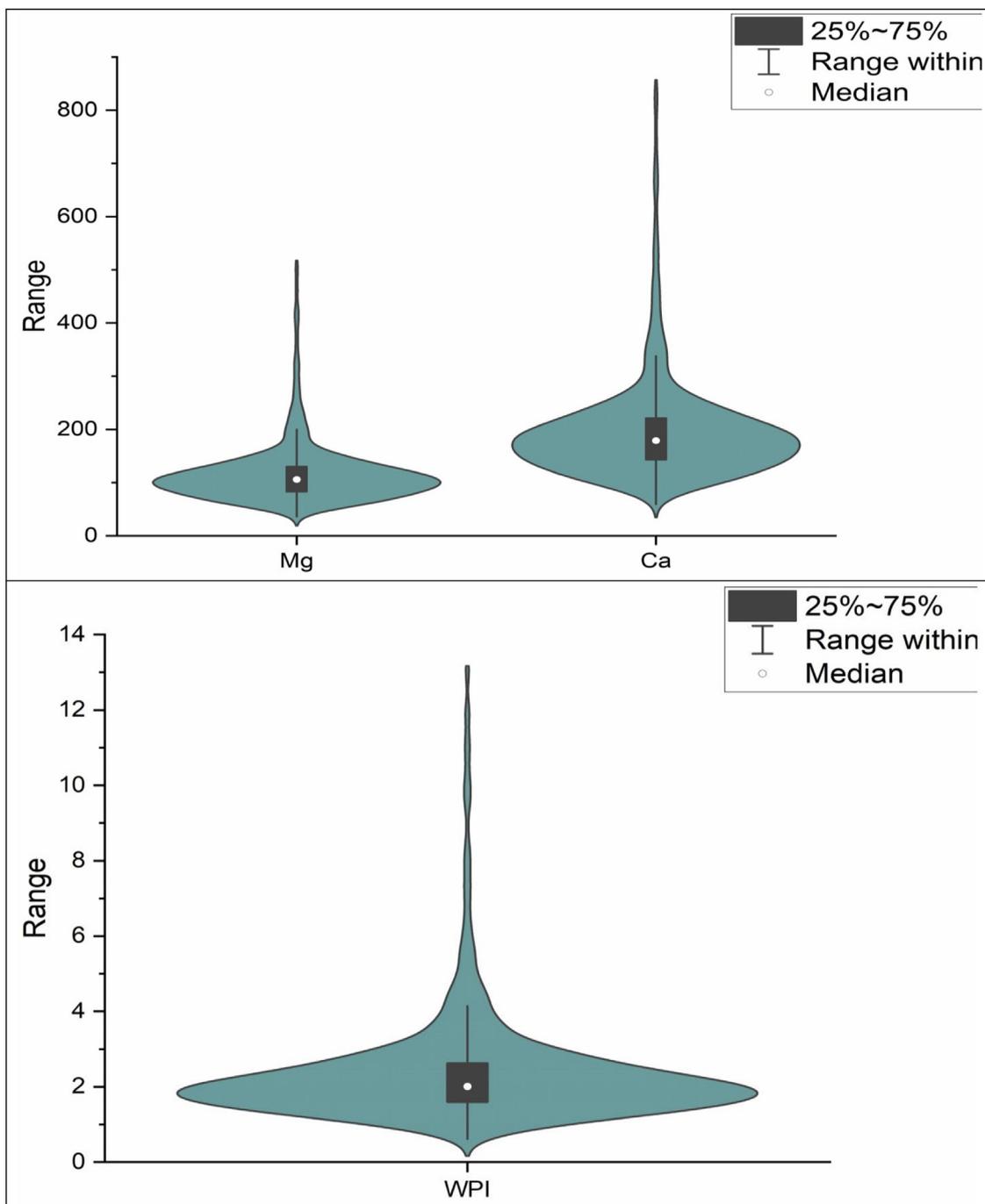
Fig. A.3. Variations of the Mg, Ca, and WPI.

Table 17
Presentation of research's results.

Methods	Measured data	MLR model	ANN model	SA	ANN-SA
WPI minimum normalized value	0	−0.00032	0.00002	−0.01811	−0.01815
WPI minimum real value	0.63543	0.63081	0.63572	0.37403	0.37346
The percentage reduction of WPI compared to the measured data	–	0.727%	−0.046%	41.137%	41.227%
Tur. (NTU)	6.9	6.9	6.9	0.6	0.6
pH	7.51	7.51	7.51	7.26	7.26

(*continued on next page*)

Table 17 (continued).

Methods	Measured data	MLR model	ANN model	SA	ANN-SA
EC (μs/cm)	942	942	942	871	871
TDS (Mg l^{-1})	584	584	584	200	200
K (Mg l^{-1})	3.5	3.5	3.5	2.5	2.5
Na (Mg l^{-1})	84	84	84	62	62
Ca (Mg l^{-1})	65	65	65	59.92	59.92
Alk. (Mg l^{-1})	124	124	124	113.9	113.9
Cl (Mg l^{-1})	150	150	150	104	104
SO$_4$ (Mg l^{-1})	156	156	156	134	134

the ideal neural network architecture employed in this study. This optimal network structure with the algorithm (trainlm) provided the best result with MSE values of 8.851×10^{-11}, 1.220×10^{-7} and 1.354×10^{-9} for the training, validation and testing sets, respectively. Furthermore, R values of 1, 1, and 1 were obtained for the training, validation, and testing sets, respectively. The results of this study indicated that the merged ANN-SA was the most successful approach for determining the minimum value of the WPI when compared to the measured, MLR, ANN, and SA results. The minimum WPI value for the integrated ANN-SA was 0.373459. The MLR, ANN, SA, and ANN-SA reduced the lowest WPI value of the measured data by 0.738%, −0.006%, 41.137%, and 41.228%, respectively. No significant difference was observed (0.15%) between the minimum WPI value presented by the SA technique and the minimum WPI value presented by the combined ANN-SA technique. However, the number of iterations provided by the combined ANN-SA was less than that provided by the SA.

CRediT authorship contribution statement

Iman Ali Abdulkareem: Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization, Funding acquisition. **Ammar Salman Dawood:** Conceptualization, Methodology, Study design, Software, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Abdulhussain A. Abbas:** Conceptualization, Methodology, Study design, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

Appendix. Variations of the water parameters and wpi

See Figs. A.1–A.3.

References

Abbas, A.A., Dawood, A.S., Al-Hasan, Z.M., 2017. Evaluation of groundwater quality for drinking purpose in basrah governorate by using application of water quality index. Kufa J. Eng. 8 (1), 65–78.

Abdulkareem, I.A., Abbas, A.A., Dawood, A.S., 2022. Modeling pollution index using artificial neural network and multiple linear regression coupled with genetic algorithm. J. Ecol. Eng. 23 (3), 236–250. http://dx.doi.org/10.12911/22998993/146177.

Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Umar, A.M., Linus, O.U., Kiru, M.U., 2019. Comprehensive review of artificial neural network applications to pattern recognition. IEEE Access 7, 158820-158846. http://dx.doi.org/10.1109/ACCESS.2019.2945545.

Abyaneh, H.Z., 2014. Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. J. Environ. Health Sci. Eng. 12 (1), 1–8. http://dx.doi.org/10.1186/2052-336X-12-40.

Ahmed, S., Ismail, S., 2018. Water pollution and its sources, effects and management: A case study of Delhi. Int. J. Current Adv. Res. 7 (2), 10346–10442.

Al-Adhab, H., Salman, A., Sagban, A., 2019. Using multivariate statistical methods to Evaluate water quality in some of Basrah province locations. Proc. ICCEET 65–72.

Al-Asadi, S.A., Al-Qurnawi, W.S., Al Hawash, A.B., Ghalib, H.B., Alkhlifa, N.H.A., 2020. Water quality and impacting factors on heavy metals levels in Shatt Al-Arab River, Basra, Iraq. Appl. Water Sci. 10 (5), 1–15. http://dx.doi.org/10.1007/s13201-020-01196-1.

Al-Asadi, S.A., Alhello, A.A., 2019. General assessment of Shatt Al-Arab River, Iraq. Int. J. Water 13 (4), 360–375. http://dx.doi.org/10.1504/IJW.2019.106049.

Al-Mahasneh, M., Al-U'datt, M., Rababah, T., Al-Widyan, M., Abu Kaeed, A., Al Mahasneh, A.J., Abu-Khalaf, N., 2021. Classification and prediction of bee honey indirect adulteration using physiochemical properties coupled with K-means clustering and simulated annealing-artificial neural networks (SA-ANNs). J. Food Qual. 1–9. http://dx.doi.org/10.1155/2021/6634598-w.

Al-Muhyi, A.H.A., 2015. The challenges facing Shatt Al Arab River in present and future. In: 7th National Conference of the Environment and Natural Resources, pp. 1–22.

Ali, P.J., Faraj, R.H., 2014. Data normalization and standardization: A technical report. Mach. Learn. Tech. Rep. 1 (1), 1–6.

Allafta, H., Opp, C., 2020. Spatio-temporal variability and pollution sources identification of the surface sediments of Shatt Al-Arab River, Southern Iraq. Sci. Rep. 10 (1), 1–16. http://dx.doi.org/10.1038/s41598-020-63893-w.

Almuktar, S., Hamdan, A.N.A., Scholz, M., 2020. Assessment of the effluents of Basra City main water treatment plants for drinking and irrigation purposes. Water 12 (12), 1–26.

Bahrami, S., Ardejani, F.D., 2016. Prediction of pyrite oxidation in a coal washing waste pile using a hybrid method, coupling artificial neural networks and simulated annealing (ANN/SA). J. Clean. Prod. 137, 1129–1137. http://dx.doi.org/10.1016/j.jclepro.2016.08.005.

Banejad, H., Olyaie, E., 2011. Application of an artificial neural network model to rivers water quality indexes prediction–a case study. J. Am. Sci. 7 (1), 60–65.

Cerny, V., 1985. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. J. Optim. Theory Appl. 45 (1), 41–51.

Chaki, S., Ghosal, S., 2011. Application of an optimized SA-ANN hybrid model for parametric modeling and optimization of LASOX cutting of mild steel. Prod. Eng. 5 (3), 251–262. http://dx.doi.org/10.1007/s11740-011-0298-x.

Cilimkovic, M., 2015. Neural networks and back propagation algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North Dublin, Ireland, pp. 1–12.

Dawood, A.S., 2017. Using multivariate statistical methods for the assessment of the surface water quality for a river: A case study. Int. J. Civ. Eng. Technol. 8 (10), 588–597.

Dawood, A.S., Akesh, A.A., Khudier, A.S., 2018a. Study of surface water quality and trends assessment at Shatt Al-Arab river in Basrah province. J. Kerbala Univ. 14 (1), 215–231.

Dawood, A.S., Faroon, M.A., Yousif, Y.T., 2020. The use of multivariate statistical techniques in the assessment of river water quality. Anbar J. Eng. Sci. 8 (3), 93–203.

Dawood, A.S., Hamdan, A.N., Khudier, A.S., 2017. Assessment of physico-chemical parameters and water quality index for Shatt Al Arab River, Iraq. Int. J. Adv. Sci. Eng. Technol. 5 (4), 62–67.

Dawood, A.S., Hamdan, A.N., Khudier, A.S., 2018b. Assessment of water quality index with analysis of physiochemical parameters. Case study: The Shatt Al-Arab River, Iraq. Int. Energy Environ. Found. 9, 3–106.

Dawood, A.S., Hussain, H.K., Hassan, A., 2016. Modeling of river water quality parameters using artificial neural network-a case study, in: Proceedings of 40th ISERD International Conference, Cairo, Egypt, pp. 37–41.

Delahaye, D., Chaimatanan, S., Mongeau, M., 2019. Simulated annealing: From basics to applications. In: Gendreau, M., Potvin, J.Y. (Eds.), Handbook of Metaheuristics. International Series in Operations Research & Management Science. vol. 272, Springer, Cham, http://dx.doi.org/10.1007/978-3-319-91086-4_1.

Diamantopoulou, M.J., Papamichail, D.M., Antonopoulos, V.Z., 2005. The use of a neural network technique for the prediction of water quality parameters. Oper. Res. 5 (1), 115–125. http://dx.doi.org/10.1007/BF02944165.

Dinka, M.O., 2018. Safe drinking water: Concepts, benefits, principles and standards. In: Glavan, M. (Ed.), Water Challenges of an Urbanizing World. IntechOpen, London, pp. 163–181. http://dx.doi.org/10.5772/intechopen.71352.

Du, K.L., Swamy, M.N.S., 2016. Search and Optimization By Metaheuristics. Techniques and Algorithms Inspired by Nature, first ed. Birkhäuser, Cham, http://dx.doi.org/10.1007/978-3-319-41192-7.

El-Naggar, K.M., AlRashidi, M.R., AlHajri, M.F., Al-Othman, A.K., 2012. Simulated annealing algorithm for photovoltaic parameters identification. Sol. Energy 86 (1), 266–274. http://dx.doi.org/10.1016/j.solener.2011.09.032.

Eluyode, O.S., Akomolafe, D.T., 2013. Comparative study of biological and artificial neural networks. Eur. J. Appl. Eng. Sci. Res. 2 (1), 36–46.

Ewaid, S.H., Abed, S.A., 2017. Water quality index for Al-Gharraf river, southern Iraq. Egypt. J. Aquat. Res. 43 (2), 117–122. http://dx.doi.org/10.1016/j.ejar.2017.03.001.

Gallo, C., 2015. Artificial neural networks tutorial. In: Encyclopedia of Information Science and Technology, third ed. IGI Global, pp. 179–189. http://dx.doi.org/10.4018/978-1-4666-5888-2.ch626.

Ghalib, H.B., 2017. Groundwater chemistry evaluation for drinking and irrigation utilities in east Wasit province, Central Iraq. Appl. Water Sci. 7 (7), http://dx.doi.org/10.1007/s13201-017-0575-8.

Gupta, N., 2013. Artificial neural network. Netw. Complex Syst. 3 (1), 24–28.

Gupta, R., Singh, A.N., Singhal, A., 2019. Application of ANN for water quality index. Int. J. Mach. Learn. Comput. 9 (5), 688–693. http://dx.doi.org/10.18178/ijmlc.2019.9.5.859.

Hamdan, A.N.A., 2020. Assessment of heavy metals pollution in the Shatt Al-Arab River, Basra-Iraq. In: AIP Conference Proceedings. vol. 2213, (1), AIP Publishing LLC, pp. 1–10. http://dx.doi.org/10.1063/5.0000300.

Hamdan, A.N.A., Dawood, A.S., 2016. Neural network modelling of Tds concentrations in Shatt Al-Arab River water. Eng. Technol. J. 34 (2), 334–345.

Hamdan, A., Dawood, A., Naeem, D., 2018. Assessment study of water quality index (WQI) for Shatt Al-arab River and its branches, Iraq. In: MATEC Web of Conferences. vol. 162, pp. 1–7. http://dx.doi.org/10.1051/matecconf/201816205005.

Hameed, A.A., Karlik, B., Salman, M.S., 2016. Back-propagation algorithm with variable adaptive momentum. Knowl.-Based Syst. 114, 79–87. http://dx.doi.org/10.1016/j.knosys.2016.10.001.

Hossain, M., Patra, P.K., 2020. Water pollution index–A new integrated approach to rank water quality. Ecol. Indic. 117, http://dx.doi.org/10.1016/j.ecolind.2020.106668.

Khan, F.N., Fan, Q., Lu, C., Lau, A.P.T., 2020. Machine learning methods for optical communication systems and networks. In: Willner, A.E. (Ed.), Optical Fiber Telecommunications VII. Academic Press, pp. 921–978. http://dx.doi.org/10.1016/B978-0-12-816502-7.00029-4.

Khashei, M., Bijari, M., 2010. An artificial neural network (p, d, q) model for timeseries forecasting. Expert Syst. Appl. 37 (1), 479–489. http://dx.doi.org/10.1016/j.eswa.2009.05.044.

Khudair, K.M., Eraibi, N.A., 2017. Environmental impact of RO units installation in main water treatment plants of Basrah city/south of Iraq. Desalination 404, 270–279. http://dx.doi.org/10.1016/j.desal.2016.11.020.

Khudhur, Z.A., Dawood, A.S., Arab, S.A., 2020. Forecasting of dissolved oxygen in Shatt Al-Arab River based on parameters of water quality using artificial neural networks. In: IOP Conference Series: Materials Science and Engineering. pp. 1–12. http://dx.doi.org/10.1088/1757-899X/888/1/012049.

Kirkpatrick, S., Gelatt Jr., C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.

Kubat, M., 2015. An Introduction To Machine Learning, first ed. Springer International Publishing, http://dx.doi.org/10.1007/978-3-319-63913-0.

Kulisz, M., Kujawska, J., Przysucha, B., Cel, W., 2021. Forecasting water quality index in groundwater using artificial neural network. Energies 14 (18), 1–16.

Liu, B., Wang, R., Zhao, G., Guo, X., Wang, Y., Li, J., Wang, S., 2020. Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm. Tunn. Undergr. Space Technol. 95, 1–12. http://dx.doi.org/10.1016/j.tust.2019.103103.

Mahmood, W., Ismail, A.H., Shareef, M.A., 2019. Assessment of potable water quality in Balad city, Iraq. In: IOP Conference Series: Materials Science and Engineering. vol. 518, (2), pp. 1–10. http://dx.doi.org/10.1088/1757-899X/518/2/022002.

Maind, S.B., Wankar, P., 2014. Research paper on basic of artificial neural network. Int. J. Recent Innov. Trends Comput. Commun. 2 (1), 96–100.

Maytham, A.A., Hammadi, N.S., Abed, J.M., 2019. Environmental study of zooplankton in the middle part of the Shatt Al-Arab River, Basrah, Iraq. Basrah J. Agric. Sci. 32 (2), 85–96. http://dx.doi.org/10.37077/25200860.2019.259.

Moyel, M.S., 2014. Assessment of water quality of the Shatt Al-Arab River, using multivariate statistical technique. Mesopotamia Environ. J. 1 (1), 39–46.

Najah, A., El-Shafie, A., Karim, O.A., El-Shafie, A.H., 2013. Application of artificial neural networks for water quality prediction. Neural Comput. Appl. 22 (1), 187–201. http://dx.doi.org/10.1007/s00521-012-0940-3.

Nwobi-Okoye, C.C., Ochieze, B.Q., 2018. Age hardening process modeling and optimization of aluminum alloy A356/Cow horn particulate composite for brake drum application using RSM, ANN and simulated annealing. Def. Technol. 14 (4), 336–345. http://dx.doi.org/10.1016/j.dt.2018.04.001.

Rahi, K.A., 2018. Salinity management in the Shatt Al-Arab River. Int. J. Eng. Technol. 7 (4.20), 128–133. http://dx.doi.org/10.14419/ijet.v7i4.20.25913.

Salami, E.S., Salari, M., Ehteshami, M., Beadokhti, N.T., 2016. Application of artificial neural networks and mathematical modeling for the prediction of water quality variables (case study: southwest of Iran). J. Desalination Water Treat. 57 (56), http://dx.doi.org/10.1080/19443994.2016.1167624.

Shahid, E.S., Salari, M., Ehteshami, M., Sheibani, S.N., 2020. Artificial neural network (ANN) modeling of cavitation mechanism by ultrasonic irradiation for cyanobacteria growth inhibition. J. Environ. Treat. Tech. 8, 625–633.

Sibalija, T., 2018. Application of simulated annealing in process optimization: a review. In: Scollen, A., Hargraves, T. (Eds.), Simulated Annealing: Introduction, Applications and Theory. Nova science publishers, pp. 1–14.

Singh, K.P., Basant, A., Malik, A., Jain, G., 2009. Artificial neural network modeling of the river water quality—a case study. Ecol. Model. 220 (6), 888–895. http://dx.doi.org/10.1016/j.ecolmodel.2009.01.004.

Suman, B., Kumar, P., 2006. A survey of simulated annealing as a tool for single and multiobjective optimization. J. Oper. Res. Soc. 57 (10), 1143–1160. http://dx.doi.org/10.1057/palgrave.jors.2602068.

Wong, W.K, Guo, Z.X., Leung, S.Y.S., 2013. Optimizing Decision Making in the Apparel Supply Chain using Artificial Intelligence (AI): From Production To Retail, first ed. Woodhead Publishing.

Yaseen, B.R., Al Asaady, K.A., Kazem, A.A., Chaichan, M.T., 2016. Environmental impacts of salt tide in Shatt al-Arab-Basra/Iraq. IOSR J. Environ. Sci. Toxicol. Food Technol. 10 (1–2), 35–43. http://dx.doi.org/10.9790/2402-10123543.

Zain, A.M., Haron, H., Sharif, S., 2009. Application of regression and ANN techniques for modeling of the surface roughness in end milling machining process. In: Third Asia International Conference on Modelling and Simulation. IEEE, pp. 188–193. http://dx.doi.org/10.1109/AMS.2009.76.

Zain, A.M., Haron, H., Sharif, S., 2011. Estimation of the minimum machining performance in the abrasive waterjet machining using integrated ANN-SA. Expert Syst. Appl. 38 (7), 8316–8326. http://dx.doi.org/10.1016/j.eswa.2011.01.019.

Zhan, S.H., Lin, J., Zhang, Z.J., Zhong, Y.W., 2016. List-Based Simulated Annealing Algorithm for Traveling Salesman Problem. Hindawi Publishing Corporation, Computational intelligence and neuroscience.

Zhang, Z., 2018. Multivariate Time Series Analysis in Climate and Environmental Research, first ed. Springer International Publishing, http://dx.doi.org/10.1007/978-3-319-67340-0.

Zhang, Y.J., Niu, B., Zhuang, X.L., Liao, H.C., 2011. Water content ratio measurement with neural network based on simulated annealing. In: Seventh International Conference on Natural Computation. vol. 2, IEEE, pp. 878–881. http://dx.doi.org/10.1109/ICNC.2011.6022215.

Zhao, Y., Nan, J., Cui, F.Y., Guo, L., 2007. Water quality forecast through application of BP neural network at yuqiao reservoir. J. Zhejiang Univ.-Sci. A 8 (9), 1482–1487. http://dx.doi.org/10.1631/jzus.2007.A1482.