Water Resources Management (2019) 33:3127–3141
https://doi.org/10.1007/s11269-019-02289-6

Advanced Evaluation Methodology for Water Quality Assessment Using Artificial Neural Network Approach

Sandeep Bansal[1] · Geetha Ganesan[2]

Received: 1 February 2019 / Accepted: 19 May 2019 /
Published online: 6 June 2019

Abstract
The increasing rate of water pollution and consequent increase of waterborne diseases are compelling evidence of danger to public health and all living organisms. Preservation of flora and fauna by controlling various unexpected pollution activities has become a great challenge. This paper presents an artificial neural network (ANN)-based method for calculating the water quality index (WQI) to estimate water pollution. The WQI is a single indicator representing an overall summary of various water test results. However, selection of the weight values of the water quality parameters for WQI calculation is a tedious task. Therefore, the ANN approach is found to be useful in this study for calculating the weight values and the WQI in an efficient manner. This work is novel because we propose a methodology that uses a mathematical function to calculate the weight values of the parameters regardless of missing values, which were randomly decided in previous work. The results of the proposed model show increased accuracy over traditional methods. The accuracy of the calculated WQI also increased to 98.3%. Additionally, we also designed a web interface and mobile app to supply contamination status alerts to the concerned authorities.

Keywords Water pollution · Water quality · Water quality standards · Artificial neural network · Pollution and health · Water pollution measurement

1 Introduction

Water is the predominant substance on earth and is used in agriculture, industry, commerce, raising of livestock, and generation of hydropower, and well as for drinking and household needs. Variation in water quality levels also affects the climatic patterns.

✉ Sandeep Bansal
er.sandy029@gmail.com

[1] Department of Electronics and Communication Engineering, Lovely Professional University, Jalandhar, Punjab 144401, India

[2] Department of Research and Development, Lovely Professional University, Jalandhar, Punjab 144401, India

Therefore, sufficient availability and quality of the water supply has led to immense development of countries. However, by 2025, it is anticipated that water withdrawals will increase by 50% in the developing countries and by 18% in the developed countries (UNEP 2012). Out of 2.5 lakh gram panchayats in India, only 28,000 have achieved the Nirmal Gram status (Barnard et al. 2013). The World Bank estimates that 21% of contagious diseases in India are related to polluted water and the absence of hygiene practices (WHO 2017). Due to dumping of sewage and garbage into water, industrial waste, oil pollution, acid rain, global warming, eutrophication, etc., the river Ganga and river Yamuna in India are currently the first and the third most polluted rivers in the world (Yeung et al. 2009). Water pollution causes many types of diarrheal diseases as well as cholera, guinea worm diseases, filarial diseases, dysentery, viral gastroenteritis and amebiasis. Therefore, the water quality determines the quality of life for humans, and in short, the survival of humans and other life forms is impossible without water.

The supply and demand of pure water creates a persistent problem. To this end, we need accurate information on the quality of our water, and an urgent need exists to apply environmental engineering aspects to measurement of water pollution and determination of water quality. Many parameters are used to assess water quality, including turbidity, Biological Oxygen Demand BOD, dust, volcanic gases, suspended solids, phytoplankton, algae, temperature, conductivity, solar radiation, etc. As per the WHO guidelines, five of the major parameters necessary for determining water quality are dissolved oxygen (DO), turbidity, temperature, power of hydrogen (pH), and total dissolved solids (TDS) (WHO 2017). Therefore, we considered these five parameters in this study. A large dataset containing 700 samples corresponding to these five parameters was collected from authenticated resources, and this dataset is further used in ANN training. Finally, the trained ANN can be applied to water quality calculation for any sample. The quality of water calculated from the proposed model can further determine its uses, i.e., whether it can be used as drinking water, irrigation purpose or for domestic purposes (ASCE 2000; Wu et al. 2018; Adimalla et al. 2018). The novelty of this work lies in its use of the artificial neural network to calculate weight values for the parameters. Additionally, the designed system is completely automatic and can detect unwanted situations (e.g., filter failure, which lead to deterioration of water quality) and alert the responsible system authorities.

2 Related Work

Water quality monitoring and analysis has been a long-term topic of discussion for researchers. Current drinking water safety evaluation research primarily emphasizes the influence of environmental pollutants (biological, chemical and physical) on water quality and risk evaluation for human health effects. The WHO, EC and USEPA and other organizations have formulated Guidelines for Drinking Water Quality, Instructions on Drinking Water Quality and standards on Drinking Water Quality (BIS 2012; European Union 2014; Ladan 2012; Spry and Branch 2015; UNEP 2012; WHO 2017; Yamamura 2001).

Water quality is monitored with the aid of water quality parameters. Although numerous parameters were selected by standard organizations, various authors have considered different parameters for measuring the water quality (Alam et al. 2007; Naeraa et al. 1966; Nosrati 2017). Once the parameters are decided, the next issue is the source for data collection.

Because numerous authors have worked with different sources of data collection, a particular source was targeted, and complete analysis was performed on that source (Debels et al. 2005; Ewaid et al. 2018; Gupta et al. 2017; Naidoo and Olaniran 2013). The researchers analyzed the water quality of these sources and proposed different solutions for maintaining the balance of quality.

Water quality is represented in terms of the water quality index or WQI. Numerous authors have worked with different definitions and representations of the WQI (Das Kangabam et al. 2017; Dinius 1987; Said et al. 2004; Srivastava and Kumar 2013). These indices are helpful in analyzing the water quality, but no standard formula exists for choosing the prerequisites of the WQI. Therefore, we selected the neural network approach for automatic generation of the WQI. Numerous authors have worked with ANN techniques for different applications of water quality monitoring (Asce 2000; Maind 2014; Minli and Shanshan 2012), but WQI measurement still requires further attention.

It is observable from previous work that the weight values were randomly decided depending on the parameter selected for study, and the values were not reliable. Therefore, we propose a methodology to calculate the weight values regardless of the number of parameters.

3 Artificial Neural Network

Computers are well suited to addressing algorithmic and mathematical problems, but frequently, problems cannot be well defined with a mathematical process. For instance, language processing and facial recognition are examples of difficulties that cannot simply be addressed using an algorithm, yet these tasks are easy for human beings. The significance of the artificial neural network (ANN) is that it processes the information in a manner similar to that of our biological brain and was designed by drawing inspiration from how our own nervous system functions (Maind 2014). Therefore, ANNs are beneficial implementations for solving problems such as language processing and facial recognition, tasks that our natural brains can perform without any effort.

Benvenuto proposed the backpropagation algorithm (Benvenuto and Piazza 1992), and this algorithm was explained in four major components:

i) Feed-forward computation
ii) Backpropagation to the output layer
iii) Backpropagation to the hidden layer
iv) Weight updates

The general artificial neural network consists of three layers, namely, the input layer, hidden layer and output layer. The inputs are applied at the input layer for further processing. From the input nodes, the information is passed to the hidden layer. In the hidden layer, is placed between the input layer and output layer and thus has no connection with the outside world. The hidden layer performs its process and sends the output to the output layer. In the output layer, the information forwarded from the hidden layer is collected, and the information is transferred to the outside world. In the network, the connection between the layers is enabled by communication lines that carry selected weights. The result of the output layer is compared with the desired output of the circuit, and the error is calculated. In the subsequent step, this error is used in the weight updating process, and the result of the output layer is fed back to the hidden layer. This process continues until the error becomes sufficiently small. In this work, five input nodes are used because five water parameters are considered. Additionally, this process contains 5 hidden nodes and one output node showing the water quality level. The weights are carried through the neural network, and the evaluation is performed within the network.

4 Water Quality Index

The WQI is a single indicator for the entire quality status of the water. The WQI is measured using a scale from 0 to 100. The water quality is defined by several parameters, which are divided into three subcategories, namely, physical, chemical, and biological (Naidoo and Olaniran 2013). Each subcategory is again divided into insignificant categories, e.g., physical parameters including pH, temperature, conductivity, etc.; chemical parameters including hardness, calcium, magnesium, chloride levels, etc.; and biological properties divided into biological oxygen demand (BOD), chemical oxygen demand (COD), phenols, etc.

Therefore, water quality can be estimated by measuring a large number of parameters with a selected specified range defined by reputed organizations such as the World Health Organization (WHO 2017). However, the limits specified for all parameters might vary with the intended application.

It is somewhat easier to measure an individual parameter value and decide whether this parameter is within the defined range or not. Numerous research studies are currently in process in India as well as the United Nations, and these entities finally introduced a term known as the water quality index, which simplifies the problem to measurement of the efficacy of water quality in standard terms (Nosrati 2017). The WQI can be estimated using Eq. 1, and its range is 0–100, where 0–25 defines poor quality, 26–50 denotes below average quality, 51–70 indicates average quality, 71–90 represents good quality and 91–100 defines excellent water quality. (Said et al. 2004).

$$WQI = \sum \mathrm{W_Y Q_Y} / \sum \mathrm{W_Y} \tag{1}$$

where.

y Available parameters

Qy Q-value of available parameters

Wy Weighting factors of available parameters

Many authorities such as the World Health Organization (WHO) and Bureau of Indian Standards (BIS) have set water quality standards for drinking water and other intended uses (BIS 2012; WHO 2017). These organizations set the permissible limits of various parameters for consumer usage to reduce consumer health risks and to ensure clean water throughout the world. To develop our system, after intensive research, we targeted some important physio-chemical parameters such as pH, temperature and turbidity for WQM purposes. Table 1 lists various water quality parameters and their permissible limits for usage as described by various organizations (European Union 2014; Gupta et al. 2017; “NESREA guidelines,” 2007; Spry and Branch 2015; Yamamura 2001), where “-” represents a unitless quantity/parameter.

Table 1 Standards for drinking water

Parameters	Units	WHO	EU	USEPA	Canada	NESREA	ICMR
PH	–	6.5–8.5	6.5–9.5	6.5–8.5	6.5–7.5	6.5–8.5	6.0–8.5
Temperature	ºC				<15	Ambient	
Total solids	mg/l	500	500	500	500	500	500
Turbidity	NTU	<5	4	0.5–1.0	0.1–1.0	<5	2.5–9
DO	mg/l						5

It can be observed from Eq. 1 that two major factors are required for calculating WQI:

- Quality factor (Q/q-Factor).
- Weighting factor

4.1 Q-Factor

The Q-Factor indicates the quality of water for an individual parameter, where 100 indicates excellent water quality and 0 indicates poor water quality. The value of each parameter is given in standard and different terms, e.g., pH is unitless, temperature is measured in degrees Celsius, TDS is measured in ppm, turbidity units are measured in NTU, etc. Although these are standard units, we cannot combine data with different units into a single formula for WQI. Therefore, we need a common term for all parameters to include them in a single formula. To this end, each parameter value is converted into a q-factor. We designed equations to calculate the q-factor from the standard terms for different parameters. The Q-factor for different water quality parameters can be calculated using the following equations:

$$\mathrm{pH_Q} = \left(\left(-0.776289 \times \mathrm{pH}^3\right) + \left(11.6597 \times \mathrm{pH}^2\right) - (37.9749 \times \mathrm{pH}) + 32.1271\right) \tag{2}$$

$$\mathrm{Temp_Q} = \left(\left(0.00723232 \times t^3\right) - \left(0.261429 \times t^2\right) - (0.844661 \times \mathrm{t}) + 80.6558\right) \tag{3}$$

$$\mathrm{TDS_Q} = \left(\left(1.37374 \times 0.000001 \times \mathrm{tds}^3\right) - \left(0.00108052 \times \mathrm{tds}^2\right) + (0.128167 \times \mathrm{tds}) + 80.0909\right) \tag{4}$$

For turbidity, the conversion from NTU to Q-factor is performed by mapping using Table 2.

For dissolved oxygen, the conversion from PPM to Q-factor is performed by first converting it to % saturation and subsequently to the Q-factor by mapping the values given in Table 3.

4.2 Weight Factor

The weight factor defines the importance of each parameter in deciding the water quality. Although a number of studies have been performed in this field to decide the weights of the parameters for calculation of WQI, no standard formula exists for assigning the weights to the parameters. Different researchers have chosen different values of the weight factors for various parameters (Das Kangabam et al. 2017; Debels et al. 2005; Dinius 1987; Ewaid et al. 2018). A summary of various studies related to water quality weights is shown in Table 4.

Table 4 shows the weights assigned to each parameter by different researchers. Although the number of parameters differed in each study, the mean deviation of the weight from value 1

Table 2 Conversion from NTU to Q-factor for turbidity

Transparency (cm)	Turbidity (NTU)	
Reading from Tube	Use in database	Q-Value
150	0	97
120	5	84
90	10	76
>60 (turb tube)	<15 (turb tube)	70
60	15	68
30	20	62
27.5	25	57
25	30	53
22.5	35	48
20	40	45
15	50	39
12.5	60	34
10	70	28
7.5	80	25
5	90	22
2.5	100	17
<2.5	>100	5

Table 3 Conversion from ppm to Q factor for DO

Temp °C	Solubility (mg/L)
0	14.6
1	14.2
2	13.8
3	13.5
4	13.1
5	12.8
6	12.5
7	12.2
8	11.9
9	11.6
10	11.3
11	11.1
12	10.9
13	10.6
14	10.4
15	10.2
16	10
17	9.8
18	9.6
19	9.4
20	9.2
21	9
22	8.9
23	8.7
24	8.6
25	8.4
26	8.2
27	8.1
28	7.9
29	7.8
30	7.7

Table 4 Mean deviation of weights assigned for different parameters

Parameter	(Das Kangabam et al. 2017)		(Ewaid et al. 2018)		(Debels et al. 2005)		(Dinius 1987)	
	Relative Weight	% Mean Deviation	Relative Weight	% Mean Deviation	Relative Weight	% Mean Deviation	Relative Weight	% Mean Deviation
pH	0.091174	0.026536364	0.091	0.063636364	0.1	1.7	0.077	0.633333333
TDS	0.100291	0.938236364	0.1	0.963636364	–	–	–	–
Temperature	–	–	–	–	0.1	1.7	0.077	0.633333333
Turbidity	0.087527	0.338163636	0.087	0.336363636	–	–	–	–
Hardness	0.051057	3.985163636	0.051	3.936363636	–	–	0.065	1.833333333
Nitrate	0.109409	1.850036364	0.109	1.863636364	0.07	4.7	0.09	0.666666667
Nitrite	0.093727	0.281836364	0.093	0.263636364	0.07	4.7	–	–
Orthophosphate	–	–	–	–	0.12	0.3	–	–
E. coli	–	–	–	–	–	–	0.116	3.266666667
Sodium	0.058351	3.255763636	0.058	3.236363636	–	–	–	–
EC	0.116703	2.579436364	0.116	2.563636364	0.06	5.7	0.079	0.433333333
Alkalinity	–	–	–	–			0.063	2.033333333
Ammonia	–	–	–	–	0.13	1.3	–	–
BOD	0.072939	1.796963636	0.072	1.836363636	0.17	5.3	0.097	1.366666667
Chloride	–	–	–	–	–	–	0.074	0.933333333
COD	0.072939	1.796963636	0.072	1.836363636	0.17	5.3	–	–
Coli	–	–	–	–	–	–	0.09	0.666666667
Color	–	–	–	–	–	–	0.063	2.033333333
DO	0.145878	5.496936364	0.145	5.463636364	0.18	6.3	0.109	2.566666667

(the sum of the weights is 1) should be same for a parameter in every study. This table shows the % mean deviation of the assigned relative weight. As observed from Table 4, the mean deviations of pH in different studies are 0.02, 0.06 and 0.6, which is not even closer to each other. The same analysis can be conducted on other parameters as well. This situation ultimately leads to a probable/estimated error rate in calculating the WQI. Hence, we need a standard mechanism for estimating the weights and the WQI. According to the situation and application, the back propagation method is best suited for this purpose.

5 Design Methodology

Algorithm 1 gives the complete procedure for carrying out this process. The complete process is divided into the following phases:

5.1 Water Quality Parameter Estimation and Mapping

In the initial step, we use various sensors to estimate the value of the five parameters of water, i.e., pH, turbidity, temperature, TDS and DO. We used heavy-duty sensors for different parameters, e.g., SEN0161 for pH, DS18B20 for temperature, SEN0189 for turbidity, SEN0244 for TDS and SEN0237 for DO. These sensors can be used in critical conditions as well, and their reliability is much greater than that of any other sensor. In the following step, each parameter value is converted into a q-value using Eqs. 2–4, Table 2 and Table 3.

5.2 ANN Validation

In this step, approximately 696 samples were collected from various authenticated sources. The.

Algorithm 1

Input: *Water quality parameter values with WQI for training, Q-factor mapping values, water sample to be tested.*

Output: Þ: *Modified accurate weights, water quality index (WQI) for the sample to be tested.*

1. *Calculate the q-factor values for the dataset.*
2. *Create a report for the q-factor values of the water parameter values with their WQI.*
3. *Feed the database into the multi-backpropagation tool to train the model.*
4. *Collect the parameter values using the designed hardware consisting of various sensors according to the water quality parameters to be tested.*
5. *Apply the sample values collected in step 4 to the trained model generated in step 3 to find the WQI.*
6. *Finally, calculate the reliable parameter weights using the soft max function.*

WQI value for these parameters was calculated in a lab. Finally, a dataset was prepared consisting of the individual parameter value and lab-calculated value of WQI for each sample. This dataset is fed to the neural network for training. We used the backpropagation neural network for this purpose. The main goal of the backpropagation algorithm is to optimize the weights to make the neural network more efficient in learning how to correctly map the inputs to the outputs. The accuracy of this function is measured in terms of the cost function or error/loss function (not to be confused with the Gaussian error function). The loss function is a function that maps values of one or more variables onto a real number that intuitively represents a "cost" associated with those values (Minli and Shanshan 2012). This error is calculated by comparing the expected output with the output of the outer layer, and the difference is represented in terms of a loss function. The backpropagation method feeds back the error and output to the hidden layer, and this process is repeated until the error is reduced to the minimum possible range. Finally, the weight values from the input layer to the hidden layer and from the hidden layer to the output layer are calculated. The soft max function (De Frahan et al. 2019) is applied to calculate the normalized value of the weights for individual parameters, as given in Eq. 5.

$$W_i = e^x / \sum e^x \tag{5}$$

5.3 Alert Generation and Notification

This step is the final step of the proposed methodology. In this phase, the data (for which the WQI must be calculated) are fed to the trained ANN network, and the result in terms of WQI is displayed on the LED screen. The alert generation feature is also introduced in the designed system to set alarms for any unusual tracing of WQI, e.g., any filter that is not working, such that action can be taken accordingly. Additionally, the data are also stored on the cloud, and a mobile app was prepared for monitoring of all activities and dataset collection. A Wi-Fi module is used to store the data into an online database. This database can be further used in analysis of historical data.

6 Results and Discussions

In this work, the hardware and backpropagation algorithm were integrated, and significant improvements are noted in the results compared with those of traditional methods. The root mean square (RMS) error curve is shown in Fig. 1 for the training data water quality information. This error is calculated during training of the ANN, and this curve is referred to as the error curve showing the root mean square error. The error decreases with the increasing rate of the training set. In the graph, the X-axis displays the number of times that the neural network is trained, and the Y-axis shows the root mean square error. Figure 1 illustrates that the error is reduced to a negligible amount.

When the training set is applied to the software, it undergoes a loop or a cycle and applies methods that reduce the error rate. This training process applies 298 loops and shows a rapidly decreasing error rate. Comparing the first and last iteration, the error rate gradually decreases and becomes notably small at the end of the loop.

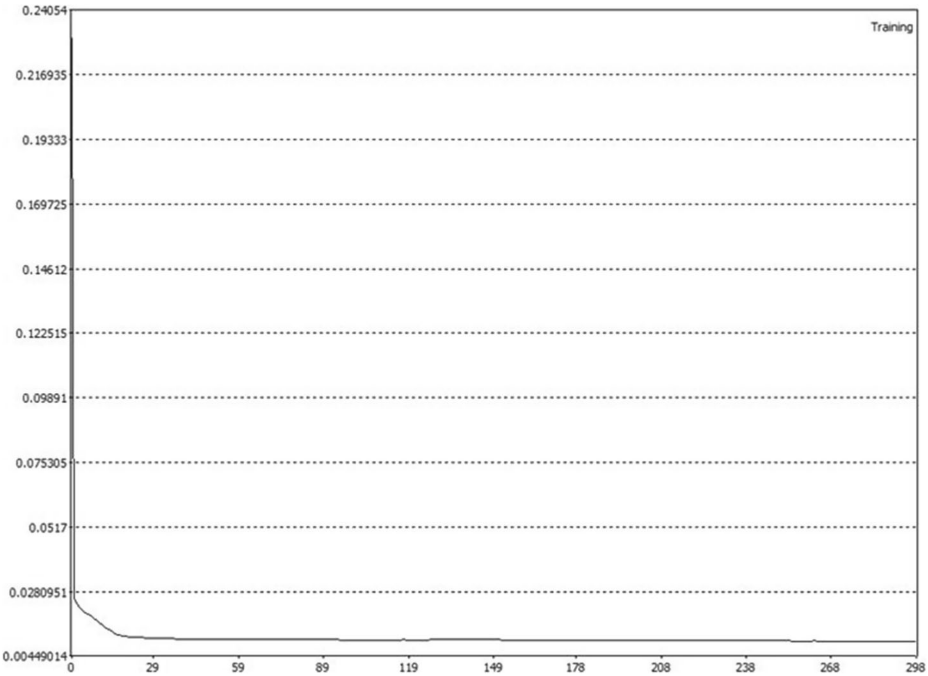


Fig. 1 Training of back propagation neural network measured by RMS error

Figure 2 shows the graph generated between the network output and desired output. It can be observed from this analysis that the error rate is minimal with the neural network approach for measurement of water quality information. After the analysis and measurement via the neural network, all of the information is also transmitted to the mobile app and to online data storage for further processing.

We visited nearby places to measure the water quality index of different sources of water. Although these sources are located near each other, notably large differences in their WQI values occur, as shown in Table 5. We attempted to find the possible sources of pollution in these water resources. The study reveals that untreated industrial waste from Ludhiana industries is dumped at Budha Nallah, and hence, its water quality is

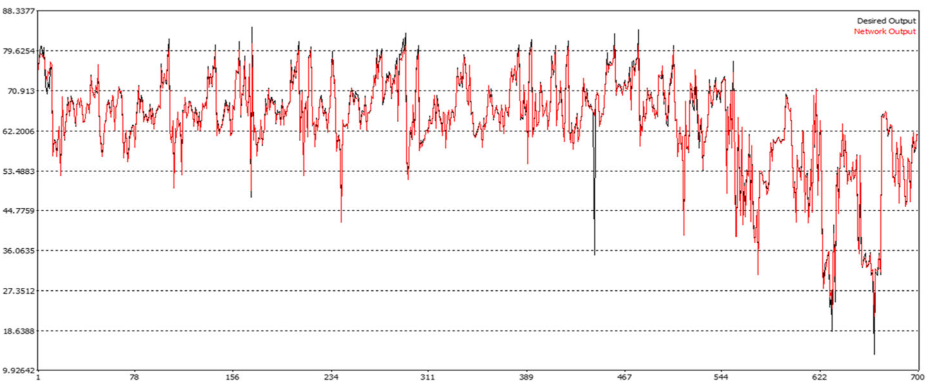


Fig. 2 Output vs. desired (Training data)

Table 5 Comparison of WQI from different water resources

S. No.	Sampling location	Coordinates	Average WQI using ANN
1.	Beas river, Kapurthala	Latitude: 31.368489 Longitude: 75.346263	61.03
2.	Budha Nallah, Ludhiana	Latitude: 30.931188 Longitude: 75.714377	5.66
3.	River Sutlej (20 km downstream from Phillaur)	Latitude: 30.973369 Longitude: 75.623442	17.27
4.	Sirhind Canal, Sirhind	Latitude: 30.586669 Longitude: 76.40295	64.84
5.	Ground water	Latitude: 31.372463 Longitude: 75.552464	68.04

quite poor. This polluted water is dumped into the Sutlej 15 km downstream from Phillaur, and we collected samples 5 km downstream of this place. Therefore, the quality of this source is also poor. The water quality of the Sirhind canal and Beas river is comparatively better than that of the previous resources and could be used in irrigation, but it is not suitable for drinking purposes. The groundwater is the has the best quality among all resources tested thus far.

Table 6 shows the comparison results for different optimization techniques. We applied the PID3 (He et al. 2012) and decision tree (Chandanapalli et al. 2018) approaches on our dataset (696 samples) and calculated the accuracy for all three techniques. We show 5 samples from a large population of 696 samples in Table 6.

Table 7 shows the weight values obtained from the backpropagation method. The final output layer weight was taken for the individual parameter, and the soft max function is applied on these values to obtain the final weight values shown in Table 8. Figure 3 shows a subset of the analysis graphs and visuals from the mobile app and web storage data. Graphical

Table 6 Comparison of optimization techniques over the same dataset

Optimization method	Water sample	WQI/ water quality classification using standard formulae and lab tests	Estimated WQI/ water quality classification using different optimization techniques	Accuracy (%)
PID3	Sample 1	40.4/Below Average	Below Average	96.23
	Sample 2	26.4/Below Average	Poor	
	Sample 3	68.4/Average	Average	
	Sample 4	55.43/Average	Average	
	Sample 5	62.3/Average	Average	
Decision Tree	Sample 1	40.4/Below Average	Below Average	97.4
	Sample 2	26.4/Below Average	Poor	
	Sample 3	68.4/Average	Average	
	Sample 4	55.43/Average	Average	
	Sample 5	62.3/Average	Average	
Proposed Approach	Sample 1	40.4	40.14	98.3
	Sample 2	26.4	26.56	
	Sample 3	68.4	68.23	
	Sample 4	55.43	54.78	
	Sample 5	62.3	61.97	

Table 7 Weight values from input layer to hidden layer and from hidden layer to output layer

	From Input Layer				
To Hidden Layer	1st Neuron	2nd Neuron	3rd Neuron	4th Neuron	5th Neuron
1st Neuron	−0.418938	−0.87416	−0.421608	−0.0697913	−0.0965121
2nd Neuron	1.16881	0.147598	0.592252	0.723902	0.949286
3rd Neuron	0.398993	0.634462	−0.606276	0.676783	0.559296
4th Neuron	−0.0597412	−0.328126	−0.219032	1.54461	0.613572
5th Neuron	−0.198823	0.420973	0.760568	0.868954	0.933312
	From Hidden Layer				
To Output Layer	1st Neuron	2nd Neuron	3rd Neuron	4th Neuron	5th Neuron
Output Neuron	−2.06164	2.77701	0.789522	1.46769	1.14347

analysis is also possible from web data, and alarming situations can be noted and information supplied tothe concerned authorities.

7 Application of Proposed Approach

We have developed a prototype to showcase the operation of the proposed approach. Once this prototype is implemented on a large scale, it can be installed on the outlets of water resources where water treatment plants have been established. The designed hardware measures the water quality at regular intervals. This information can be sent to administrators and management authorities using the mobile app, as shown in Fig. 3. When the water quality drops below the permissible range, management can take action accordingly, a process that was delayed in previous approaches due to the carelessness of intermediaries. Hence, the proposed approach is the best possible solution for reducing human error and addressing the lack of knowledge of manual water quality estimation.

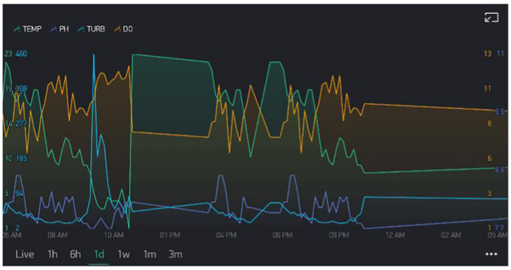
8 Conclusions and Future Work

Different physical and chemical parameters are considered in this work for monitoring real-time water quality information. Five main parameters, pH, temperature, TDS, DO and turbidity, are considered to assess water quality. This study reveals that the neural network approach shows a significant improvement regarding the tested values compared with traditional methods. This study focuses on the backpropagation algorithm for training and weight value calculation. Selection formulations were presented for automatically calculating the q-value from the parameter value. The water quality index

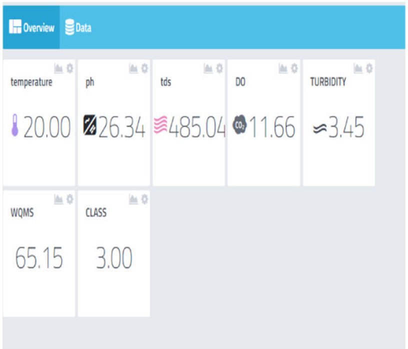
Table 8 Final weight values for individual parameters

S. No.	Parameter	Value x	e^x	Weight value
1	Temperature	−2.06164	0.127245117	0.00491724
2	pH	2.77701	16.07089706	0.62104172
3	TDS	0.789522	2.202343454	0.08510708
4	DO	1.46769	4.339200003	0.167683498
5	Turbidity	1.14347	3.137637098	0.121250453

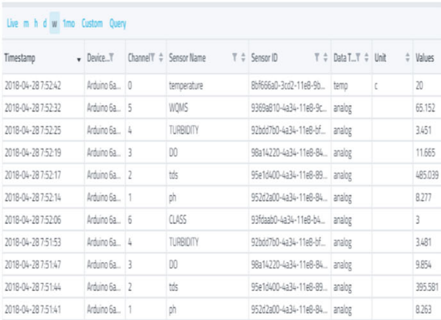
(a) Snapshot of working mobile app interface



(b) Snapshot of generated graph analysis for water quality parameters



(c) Snapshot of water quality information through web server.


Live m h d w 1mo Custom Query

Timestamp	Device...	Channel	Sensor Name	Sensor ID	Data T...	Unit	Values
2018-04-28 7:52:42	Arduino 6a...	0	temperature	8bf666a0-3cd2-11e8-9b...	temp	c	20
2018-04-28 7:52:32	Arduino 6a...	5	WQMS	9369a810-4a34-11e8-9c...	analog		65.152
2018-04-28 7:52:25	Arduino 6a...	4	TURBIDITY	92bdd7b0-4a34-11e8-bf...	analog		3.451
2018-04-28 7:52:19	Arduino 6a...	3	DO	98a14220-4a34-11e8-84...	analog		11.665
2018-04-28 7:52:17	Arduino 6a...	2	tds	95e1d400-4a34-11e8-89...	analog		485.039
2018-04-28 7:52:14	Arduino 6a...	1	ph	952d2a00-4a34-11e8-84...	analog		8.277
2018-04-28 7:52:06	Arduino 6a...	6	CLASS	93fdaab0-4a34-11e8-b4...	analog		3
2018-04-28 7:51:53	Arduino 6a...	4	TURBIDITY	92bdd7b0-4a34-11e8-bf...	analog		3.481
2018-04-28 7:51:47	Arduino 6a...	3	DO	98a14220-4a34-11e8-84...	analog		9.854
2018-04-28 7:51:44	Arduino 6a...	2	tds	95e1d400-4a34-11e8-89...	analog		395.581
2018-04-28 7:51:41	Arduino 6a...	1	ph	952d2a00-4a34-11e8-84...	analog		8.263

(d) Snapshot of data storage in cloud.

Fig. 3 Alert generation and cloud storage

(WQI) was also calculated during this study. The proposed scheme showed 98% accuracy over the traditional methods. We also automatically calculated the weight values for each parameter. This is the first time that such a methodology has been proposed to calculate reliable weight values instead of assigning random weights to the water quality parameters. Additionally, we developed a web interface to analyze historical data for future decision-making and a mobile app for alert generation if the water quality falls below the allowable limits. In future work, we plan to use deep learning techniques or genetic algorithms to perform the same analysis. We can also integrate a data analysis module into this methodology to completely automate the process.

Compliance with Ethical Standards

Conflict of Interest None.

References

Adimalla N, Li P, Venkatayogi S (2018) Hydrogeochemical evaluation of groundwater quality for drinking and irrigation purposes and integrated interpretation with water quality index studies. Environ Process 5(2):363–383

Alam MJB, Islam MR, Muyen Z, Mamun M, Islam S (2007) Water quality parameters along rivers. J Environ Sci Technol 4(1):159–167. https://doi.org/10.1007/BF03325974

Asce. (2000). Artificial neural networks in hydrology by the ASCE task committee on application of artificial neural networks in hydrology 1. J Hydrol Eng, 5(]), 124–137. https://doi.org/10.5121/ijsc.2012.3203

Barnard S, Routray P, Majorin F, Peletz R, Boisson S, Sinha A, Clasen T (2013) Impact of Indian Total sanitation campaign on latrine coverage and use: a cross-sectional study in Orissa three years following Programme implementation. PLoS One 8(8):e71438. https://doi.org/10.1371/journal.pone.0071438

Benvenuto N, Piazza F (1992) The backpropagation algorithm. IEEE Trans Signal Process 40(4):967–969

B.I.S. (Bureau of Indian Standards) (2012) Drinking Water Specification, 2nd revision, IS:10500.

Chandanapalli SB, Reddy ES, Lakshmi DR (2018) DFTDT: distributed functional tangent decision tree for aqua status prediction in wireless sensor networks. Int J Mach Learn Cybern 9(9):1419–1434

Das Kangabam R, Bhoominathan SD, Kanagaraj S, Govindaraju M (2017) Development of a water quality index (WQI) for the Loktak Lake in India. Appl Water Sci 7(6):2907–2918. https://doi.org/10.1007/s13201-017-0579-4

De Frahan MTH, Yellapantula S, King R, Day MS, Grout RW (2019) Deep learning for presumed probability density function models. arXiv preprint arXiv:1901.05557

Debels P, Figueroa R, Urrutia R, Barra R, Niell X (2005) Evaluation of water quality in the Chillán River (Central Chile) using physicochemical parameters and a modified water quality index. Environ Monit Assess 110(1–3):301–322. https://doi.org/10.1007/s10661-005-8064-1

Dinius SH (1987) Design of an Index of water quality. JAWRA J Am Water Resour Assoc 23(5):833–843. https://doi.org/10.1111/j.1752-1688.1987.tb02959.x

European Union (European Union Drinking water regulations) (2014) Drinking water regulations, S.I. No. 122, 1–34. http://www.irishstatutebook.ie/eli/2014/si/122/made/en/pdf . Accessed 9 Mar 2018

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

Gupta N, Pandey P, Hussain J (2017) Effect of physicochemical and biological parameters on the quality of river water of Narmada, Madhya Pradesh, India. Water Sci 31(1):11–23. https://doi.org/10.1016/j.wsj.2017.03.002

He Q, Dong Z, Zhuang F, Shang T, Shi Z (2012) Parallel decision tree with application to water quality data analysis. In International symposium on neural networks (pp. 628–637). Berlin: Springer

Ladan, M. T (2012) Review of NESREA act 2007 and regulations 2009-2011: a new Dawn in environmental compliance and enforcement in Nigeria. Law Env't & Dev. J., 8, 116.

Maind MSB (2014) Research paper on basic of artificial neural network. Int J Recent Innov Trends Comput Commun 2(1):96–100

Minli Z, Shanshan Q (2012) Research on the application of artificial neural networks in tender offer for construction projects. Phys Procedia 24:1781–1788. https://doi.org/10.1016/j.phpro.2012.02.262

Naeraa N, Petersen ES, Boye E, Severinghaus JW, O'Driscoll R, O'Driscoll B et al (1966) pH and molecular CO2 components of the Bohr effect in human blood. Scand J Clin Lab Invest 18(1):96–102. https://doi.org/10.1183/20734735.001415

Naidoo S, Olaniran AO (2013) Treated wastewater effluent as a source of microbial pollution of surface water resources. Int J Environ Res Public Health 11(1):249–270. https://doi.org/10.3390/ijerph110100249

Nosrati K (2017) Identification of a water quality indicator for urban roof runoff. Sustainability Water Qual Ecol 9–10:78–87. https://doi.org/10.1016/j.swaqe.2017.07.001

Said A, Stevens DK, Sehlke G (2004) An innovative index for evaluating water quality in streams. Environ Manag 34(3):406–414. https://doi.org/10.1007/s00267-004-0210-y

Spry D, Branch T (2015) An Overview of Canadian Water Quality Guidelines. USEPA Expert Meeting, Washington DC

Srivastava G, Kumar P (2013) Water quality index with missing parameters. Int J Res Eng Technol 2(4):609–614. http://www.ijret.org . Accessed 10 Jan 2018

UNEP. (2012). The Emissions Gap Report 2012. Retrieved from http://www.unep.org/pdf/2012gapreport.pdf

WHO (2017) Guidelines for drinking-water quality: first addendum to the fourth edition. https://doi.org/10.1016/S1462-0758(00)00006-6

Wu Z, Wang X, Chen Y, Cai Y, Deng J (2018) Assessing river water quality using water quality index in Lake Taihu Basin, China. Sci Total Environ 612:914–922

Yamamura S (2001) Drinking water guidelines and standards, In United Nations Synthesis Report on Arsenic in Drinking Water, 1-18. https://www.who.int/water_sanitation_health/dwq/arsenicun5.pdf. Accessed 12 Feb 2018

Yeung LWY, Yamashita N, Taniyasu S, Lam PKS, Sinha RK, Borole DV, Kannan K (2009) A survey of perfluorinated compounds in surface water and biota including dolphins from the Ganges River and in other waterbodies in India. Chemosphere 76(1):55–62. https://doi.org/10.1016/j.chemosphere.2009.02.055