



Managing water quality of a river using an integrated geographically weighted regression technique with fuzzy decision-making model

Ajit Pratap Singh  · Kunal Dhadse · Jayant Ahalawat

Received: 17 December 2018 / Accepted: 17 April 2019 / Published online: 19 May 2019
© Springer Nature Switzerland AG 2019

Abstract In the recent times, water quality of most of the rivers in India has been steadily degrading due to increasing numbers of point and non-point sources of pollution. The tremendous increase in population, rapid urbanization, change in irrigation patterns, and unplanned growth of industries without proper enforcement of environmental standards are some of the major causes for poor quality of river water. In addition, unpredictable and scanty rainfall is resulting in uncertain natural stream flow which further leads to uncertainty in assessing and predicting the quality of river water. This paper deals with the assessment of the overall status of water quality of a river by developing a fuzzy-based water quality evaluation system. The quality of water needed for different beneficial uses is based on the value of various parameters. Since the quality attributes of the parameters are fuzzy in nature, they have been described by the linguistic variables. The water quality index of each specific site is then calculated by aggregating the attributes with respect to their degree of importance, which is also expressed in the form of linguistic terms. Finally, a case study of the river Yamuna has been carried out to evaluate the fuzzy comprehensive water quality index (FCWQI). In this study, the FCWQI has been determined only for the use of water for drinking purposes though this model can be applied for other uses as well. The FCWQI developed herein is based on an

integrated approach, which clearly describes the overall state of the water quality by a single rational number. Spatial and parametric sensitivity of the FCWQI model of the river basin is also determined using GIS-based geographically weighted regression technique. The methodology suggests a novel way of introducing parametric sensitivity in defining water quality indices used for surface water quality assessment.

Keywords River water quality · Water quality modeling · Fuzzy water quality index (FWQI) · Fuzzy logic · GIS · Sensitivity analysis

Introduction

Water is essential for sustaining ecological processes that support human survival, aquatic lives, vegetation, wetlands, wildlife, and birdlife. Water is needed for agricultural production, domestic and municipal uses, manufacturing and industrial application, hydroelectric power generation, activities related to recreation and navigation, flow improvisation, natural ecological support to fish and wildlife, drinking and hygiene, etc. The unplanned and reckless utilization of water in these sectors and the absence of effective waste disposal systems have led to emerging water problems, i.e., disproportion in naturally available water quantity and degradation of water quality. While over-exploitation of groundwater has become a serious problem for many Indian states, river water is also being increasingly polluted and utilized inefficiently in various parts of the

A. P. Singh (✉) · K. Dhadse · J. Ahalawat
Civil Engineering Department, Birla Institute of Technology and Science, Pilani, Rajasthan 333031, India
e-mail: aps@pilani.bits-pilani.ac.in

country. River water quality is mostly affected by complex interactions of chemical compositions of compounds available in water and their sensitivity towards the temporally and spatially varying environment. (Awadallah and Yousry 2012; Fulazzaky et al. 2010; Mandal et al. 2009; Singh and Ghosh 2003; Singh et al. 2007; Srinivas et al. 2018).

In the recent years, the increase in the level of water pollution has reached such a stage that the government and the other related agencies have gone for quantitative measures to evaluate water quality. There have been a lot of initiatives by several researchers to address efficient way to represent, analyze, and evaluate water quality by incorporating its characteristics. Introduction of water quality index was among such initiative for the evaluation of water quality status (Nasiri et al. 2007). The indices thus developed by this way are helping to integrate a wide variety of important water quality indicators into a single water quality index that can easily be communicated to the respective policy makers and planners. Practically, a lot of different indices dealing with water quality monitoring and assessment have been introduced.

Since there are many ways to define water quality depending on the various uses of water, for every use, there would be different degree of importance (weights) and tolerance limits (standards) for each indicator parameter. Horton (1965) was among the first researchers to introduce an indexing-based rating system for water quality parameters and their individual importance, and extension of the approach has led to the development of various index methods (Liou et al. 2004). In many traditional methods of evaluation of water quality index, the weighted average of all the normalized parameters has been determined which were then multiplied with their respective weights. For example, Avvannavar and Shrihari (2008) have considered six water quality parameters, namely biochemical oxygen demand (BOD), dissolved oxygen (DO), pH, most probable number (MPN), total dissolved solids (TDS), and turbidity, at selected sampling stations along river Netravathi in South India to develop water quality index (WQI) for drinking purposes. They developed rating curves based on the quality standards for inland waters and impact on health. Each of these parameters was assigned a weight varying from 0 to 1 such that their total sum becomes 1. Harmonic mean and WQI method developed by Bhargava were applied by researchers along the stretch of the river basin. The result obtained is expressed in a

scale of 0 to 100 which was divided into a five-point rating scale: (i) 0–40: poor quality, (ii) 41–50: marginal quality, (iii) 51–70: fair quality, (iv) 71–90: good quality, and (v) 91–100: excellent quality.

Conesa Fernandes-Vitora (1997) has developed a modified version of the previous method of indexing and developed a new index named “Subjective water quality index (WQI_{sub})” by expressing it in the following form:

$$WQI_{sub} = K \frac{\sum_i C_i P_i}{\sum_i P_i} \quad (1)$$

The constant k (0.25–1.00) in the equation represents the degree of pollution from high to low, C_i is the normalized value and P_i represents the relative weight of the i^{th} parameter. Curves suggested by Conesa Fernandes-Vitora (1997) are used to normalize the parameters.

These types of methods can be seen in the works of various researchers. For example, Fulazzaky et al. (2010) presented a model for Selangor river to identify the degradation in water quality along with appropriate suggestions to maintain the quality of river water within the prescribed standards. Water quality evaluation system has been widely used in recent decades and there has always been a number of uncertainties present in the indexing system that are errors in measurement, imprecise model inputs, and interrelationship gaps between dischargers, climatic conditions, and river characteristics. Since the complexity of consideration of uncertainty increases when regional level fuzziness in decision variables varies spatially, the information collected from different sources with respect to these aspects is required to be combined to derive an overall integrated value. It is evident that a traditional crisp weighing system is not very effective to express water quality attributes/parameters in a multiple-usage practice.

The interdisciplinary use of computational technology has aided researchers in developing artificial intelligence tool that is more accurate in defining the water quality index (Chau 2006). Fuzzy logic concepts can be readily used in consideration of uncertainties and it can be used for classification and quantification of the impact of different attributes on quality indices. The results from the fuzzy-based classifications can be later used to the

interpretation of uncertainties in data associated with water quality and later formulate restoration policies.

Water quality management problems are generally associated with uncertainty mainly due to randomness in natural hydrological attributes and imprecision associated with discharge and pollution control. The fuzzy water quality indices have been used to deal with the indicators' uncertainties arising during the quality evaluation process (Ross 2008; Sakawa 1993; Zadeh 1965). The indicators are fuzzified using fuzzy sets theory and several solution techniques are available to perform mathematical analysis on those fuzzy set indicators. These techniques are fuzzy arithmetic (Kaufman and Gupta 1991), fuzzy rule-based analysis (Ross 2008), fuzzy multi-attribute ranking (Singh et al. 2017). The applicability of fuzzy set theory has been advocated by many other investigators (Nasiri et al. 2007; Singh et al. 2007; Srinivas and Singh, 2018a, b). There is enough scope for using fuzzy-based methods in assessing water quality status and suggesting remedial action plans.

Therefore, this study presents a fuzzy comprehensive water quality index (FCWQI) model. Total of eight important water quality parameters, namely, biochemical oxygen demand (BOD), dissolved oxygen (DO), electrical conductivity (EC), fecal coliform (FC), nitrate, pH, temperature, and total coliform (TC), measured at selected sampling stations along the Yamuna river basin. The selection of above-water quality parameters is based on a detailed literature study of various in-practice WQI in various developing countries (Lermontov et al. 2009; Mahapatra et al. 2011). As the study presents a framework that is much needed in developing economies, the water quality parameters in WQI standards from Iran, Malaysia, India, and South Africa have been taken into account along with the expert's opinion to select the above eight water quality parameters. The paper has been organized so as to discuss the conceptual information related to the index development first, followed by the describing FCWQI model. Then, a case study has been taken up to demonstrate its application to quantify year-wise water quality status in a river system and finally the conclusions have been drawn along with the advantages and limitations of the technique. The study is expected to help the water specialists, policymakers, and water resources planners for the overall management of water quality of a river system.

Materials and methods

There are different ways of calculating water quality status at selected sampling stations in a river. However, the traditional methods of calculating the index do not take into account the uncertainties due to data collection errors. In this paper, fuzzy mathematical programming has been employed to account for the uncertainties in the evaluation of water quality index at chosen stations along the river. The details of the techniques have been elaborated by Ross (2008), Sakawa (1993), Srinivas and Singh (2018a, b). However, a brief introduction to fuzzy logic is given in the following paragraphs.

The term fuzzy logic was introduced by Zadeh (1965) in his work "fuzzy sets." Fuzzy logic helps in devising tool for modeling scenarios that have inherently natural fluctuations and variability. Fuzzy set theory provides a relatively more flexible and robust expert decision support system in the form of dynamic approximation capabilities to express input and output attributes that can be useful in solving real-life problems. For example, if it is required to assess and classify the objects for a decision-making problem, it is highly unlikely that the attributes for the object would be crisp and exact in nature. In many cases, no well-defined boundary exists among the attributes, for example, the boundary of the class of good water and bad water quality. There is always a possibility of occurrence of ambiguity in this type of cases due to fuzziness in the attributes. Fuzzy set theory can very well define these uncertainties into ranges or "classes" by introducing a degree of belongingness using membership function of qualitative attributes of interest with respect to some well-defined sets. The process of assignment of a fuzzy grade to an attribute using membership function that varies from 0 (not belonging) to 1 (full belongingness) according to the degree of inclination of each attribute. The fuzzy sets allow a partial belongingness by defining appropriate membership functions. In order to express vagueness involved in defining the value of attributes, a number of membership functions have been used. The most commonly used functions are Gaussian, triangular, trapezoidal, and sigmoid functions. In this paper, two types of membership functions (i.e., Gaussian and triangular/trapezoidal membership functions) have been used to define WQI because they are simple to use and are computationally efficient (Phillis and Andriantiatsaholainaina 2001). Triangular membership functions are generally expressed by three variables a , b ,

and c as given in Eq. (2) which can easily be derived from Fig. 1.

$$\mu_A(x, a, b, c) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ \frac{c-x}{c-b} & \text{if } b \leq x \leq c \\ 0 & \text{if } x \geq c \end{cases} \quad (2)$$

Similarly, Gaussian membership function (Fig. 2) are expressed using two parameters, σ and c , which is expressed mathematically in Eq. (3):

$$f(x, \sigma, c) = e^{\left[\frac{-(x-c)^2}{2\sigma^2} \right]} \quad (3)$$

where σ is the width/thickness of the curve and c is the point where the curve has maximum membership value.

As the quality of water can be defined on the basis of its uses, this paper mainly focuses on FCWQI development for drinking water requirement at different sampling stations. However, the technique can also be applied for other beneficial uses of water such as irrigation, industrial cooling, controlled waste disposal, propagation of aquatic life, and outdoor bathing. Further, spatial and parametric sensitivity of the proposed FCWQI model of the river basin is also investigated using GIS-based geographically weighted regression technique. A generalized flow chart representing the entire methodology to describe the fuzzy decision support system for water quality assessment is shown in Fig. 3.

Thus, the water quality indices at different sampling sites along a river have been calculated by incorporating the fuzzy behavior of the decision-making process. This procedure involves the following steps:

- Identification of water quality parameters: A water quality parameter is a factor, which defines water quality in terms of its magnitude and/or

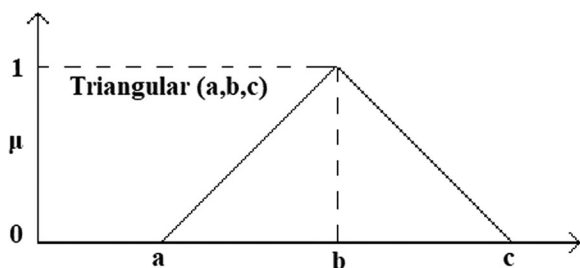


Fig. 1 Triangular membership function

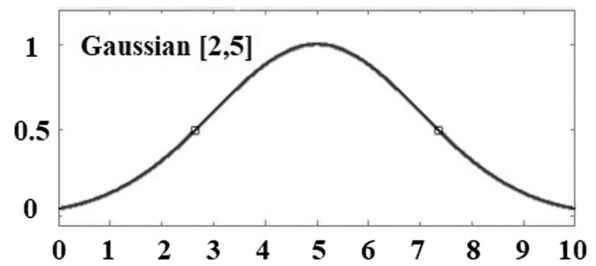
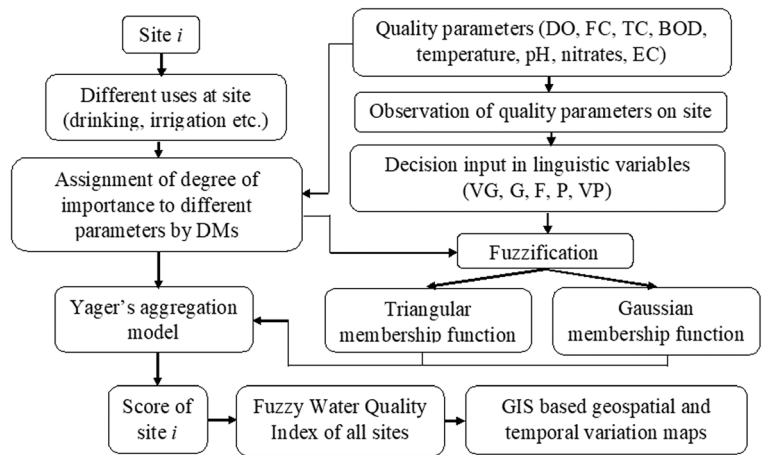


Fig. 2 Gaussian membership function

concentration, which may render the water unfit for use. In this study, eight parameters, namely total coliform (TC), pH, temperature, DO, BOD, fecal coliform (FC), nitrates, and electrical conductivity (EC) have been taken into consideration to derive the water quality index with reference to drinking water as the use. These parameters were monitored consistently at different sampling stations along the stretch of river Yamuna by Central Pollution Control Board (CPCB), New Delhi.

- Parametric estimates and attributes: The next step involves the actual measurement of aforementioned parameters. However, this data may not be that significant for decision makers, especially non-experts unless some attribute is attached to it. These attributes are defined by a group of water quality experts in terms of linguistic variables. For example, a pH of 7.5 has no significance for a decision maker unless it has an attribute attached to it, i.e., whether the given pH is of high magnitude or low magnitude. In this study, water quality parameters have been defined by five attributes, namely very poor (VP), poor (P), fair (F), good (G), and very good (VG). Since these variables are subjective in nature, they are expressed using Gaussian membership functions as given in Fig. 4.
- Degree of importance of parameters: All the parameters listed above do not necessarily affect equally the water quality for a given use. For instance, temperature may not be as important as dissolved oxygen in a specific case. Thus, it is necessary to assign degree of importance to each water quality parameter depending upon its importance. This is also done using linguistic variables, namely unimportant (U), low importance (L), medium importance (M), very important (V), and extremely important (E). These variables are also subjective in

Fig. 3 Water quality assessment in a river using fuzzy decision support system framework



nature since they also depend on the inputs by decision makers. These variables are defined using triangular membership functions as shown in Fig. 5.

The degrees of importance serve as weights for different parameters which are essential to calculate water quality index. However, weights are considered to be crisp. Thus, there is a need to find representative values for aforementioned degrees of importance. Based on centroid method, the representative value (w_i) can be calculated for each fuzzy set W_i using Eq. (4):

$$w_i = \frac{\int_{w=0}^1 w \cdot \mu_{w_i}(w) dw}{\int_{w=0}^1 \mu_{w_i}(w) dw} \tag{4}$$

The representative values for different degrees of importance have been calculated using MATLAB with the integration of Fig. 5 and Eq. (4). In this paper, the centroid method is used to defuzzify the fuzzy set of

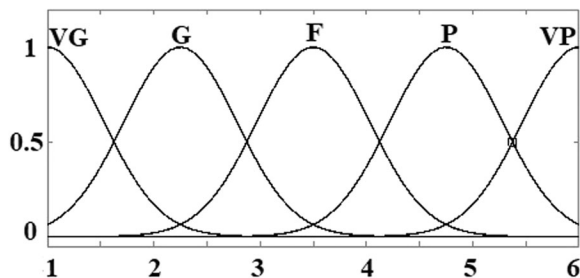


Fig. 4 Gaussian membership function for different attributes

degrees of importance by evaluating the centroid of the area represented by the different curves of U, L, M, V, and E. The centroid (or center of area) method is the most frequently used method for defuzzification which is very consistent in the case of decision-making and assessment studies (Mogharreban and Dilalla 2006; Phillis and Andriantiatsaholiniaina 2001; Tirupathi et al. 2019). However, there exist many other methods to defuzzify the fuzzy set, the details of which can be obtained from Ross (2008). These values of degrees of importance are given in Table 1:

- Water quality index at different stations: Based on the observations of different parameters, the decision makers assign attributes to each parameter at different stations. The final water quality index is calculated by using Yager’s max-min model (Yager

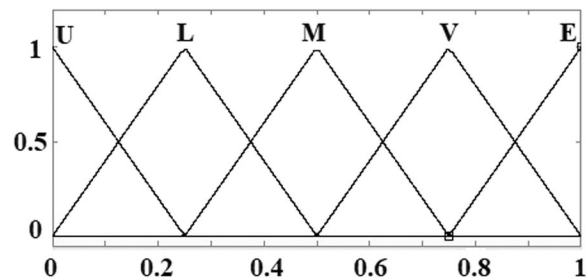


Fig. 5 Triangular membership function for different degrees of importance

Table 1 Importance values of linguistic terms used to define quality parameters

Importance	Importance weight (w_i)
U	0.0833
L	0.2500
M	0.5000
V	0.7500
E	0.9167

1978) shown in Eq. (5). For the n^{th} station, the quality index set $Q\tilde{I}_n$ is defined as:

$$Q\tilde{I}_n = \cap_{i=1,2,\dots,I} S_{i,n}^{w_{i,n}} = \min_{i=1,2,\dots,I} \left\{ \left(\mu_{S_{i,n}}(S) \right)^{w_{i,n}} \right\} \quad (5)$$

where $S_{i,n}$ denotes the score of an attribute assigned by the decision maker to the i^{th} parameter at the n^{th} sampling station. Once the quality index set $Q\tilde{I}_n$ is evaluated, the element associated with the maximum value of membership can be obtained to express reach the quality index (QI_n) at a decision using Yager's algorithm as expressed in eq. (6):

$$QI_n = \left\{ s \mid \mu_{Q_{i,n}}(s) = \text{Max}(QI_n) \right\} \quad (6)$$

Thus, the water quality index at various sampling locations along a river can be calculated. The entire methodology has been explained in the following section.

Case Study

The methodology to quantify the status of water quality has been described using a case study of the Yamuna river. The Yamuna river is a major tributary of the Ganga river. River Yamuna originates from the Yamunotri glacier on the south-western slopes of Banderpooch peak of great Himalayas. The total length of flow of the river is 1376 km, which merges with Ganges at Allahabad (U.P.). It crosses several states which include Himachal Pradesh, Uttarakhand, Haryana, Delhi, and Uttar Pradesh (ADSORBS 2000). It is the major source of drinking water in Delhi while feeding water to several other cities, towns, and villages in the Indian states of Himachal Pradesh, Uttar Pradesh, and Haryana. There are a large number of industrial centers which have

developed either on its banks or in its basin. The Yamuna river basin has fertile agricultural land especially in Haryana and western districts of Uttar Pradesh. This clearly demonstrates that the river Yamuna also plays an important role in the overall economy of India. Though the river Yamuna has been extensively used for irrigation, domestic, and industrial activities, it is also influenced highly by pollution problems caused by rapid industrial development, uncontrolled urban sprawl, and agricultural developments. Keeping in view of these facts, the case study has been taken up to quantify the water quality status of Yamuna river at different locations. The whole catchment area of Yamuna river with its tributaries that is considered for fuzzy water quality index study is shown below (Fig. 6).

A total of 19 sampling stations inside Yamuna river basin and 3 reference stations outside the catchment are selected for the study as shown in Fig. 6. The stations were selected so as to understand the anthropogenic impact on the water quality of Yamuna river. For instance, the sampling stations Shahdara and Agra are near major industrial regions whereas, the stations like Yamunotri is located at relatively low-populated areas without any major industries, and these stations were more pragmatic. All 19 stations inside the basin were selected based on the following criteria: (a) geographic coverage of the whole basin and even special distribution factors has been used to broadly select the sub-basins from which the stations would be further selected; (b) the regions already known for high industrial settlements and relatively high water pollution has been given explicit importance (e.g., downstream of Shahibi and Yamuna at Delhi); (c) The stations present at downstream of river-subsidary junction is chosen over the stations at upstream so as to accurately capture the actual water quality after mixing of stream water; d) the stations are later finalized based on the availability of water quality data in all eight domains and with the minimum outliers. The 3 reference stations outside the basin are used to define a boundary condition in the form of a triangular spatial grid that is essential in the spatial analysis that is performed in the ArcGIS software to interpolate the in-between values. The triangular boundary shape file covers the Yamuna basin and its three vertices represent the 3 reference stations that are used in the study (as shown in Fig. 6). The stations are spatially spread in the basin and covers sampling sites for all the tributaries of Yamuna river. The tributaries of Yamuna basin catchment include river Chambal, Betwa, Ken,

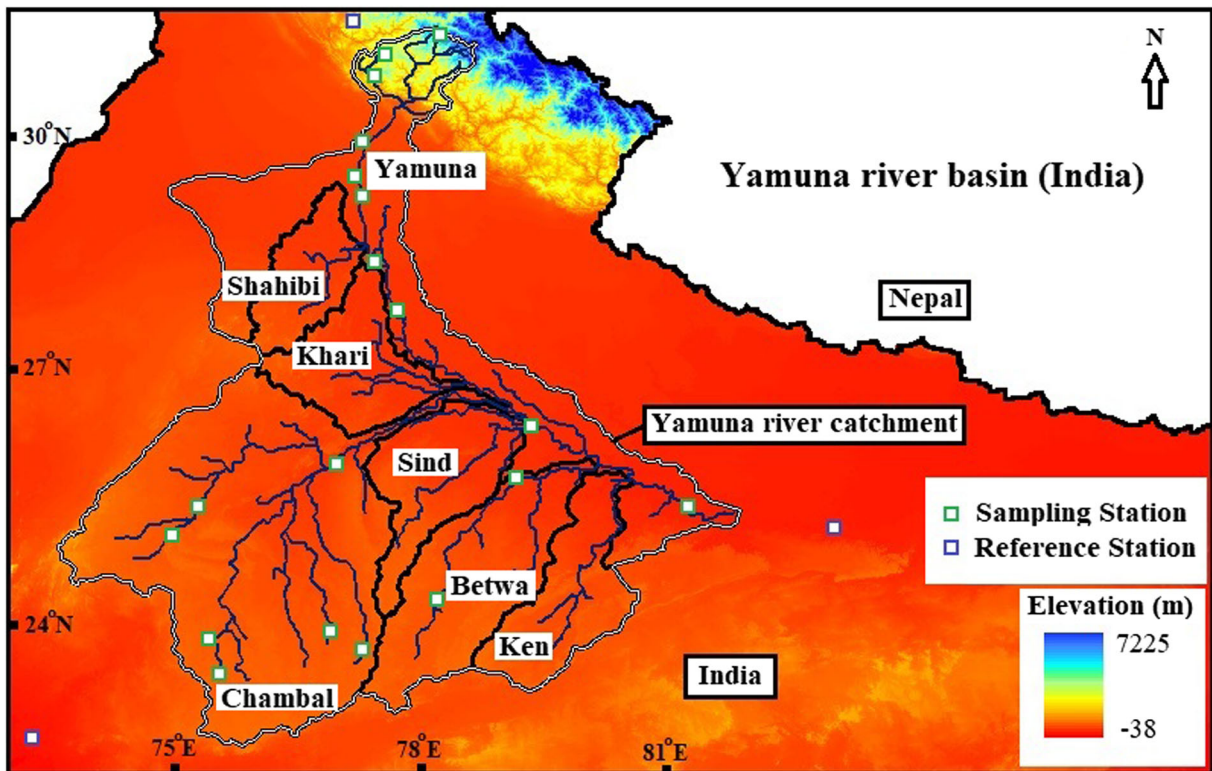


Fig. 6 Yamuna river watershed with catchment boundaries and tributaries

Banas, Khari, Shahibi, Sind, Dhasan, Senger, and Hindon which covers the entire basin. All these rivers form a total of seven catchments in Yamuna basin; namely Yamuna catchment, Ken catchment, Betwa catchment, Chambal catchment, Sind catchment, Khari catchment, and Shahibi catchment.

Results and discussions

The monitoring of water quality of the Yamuna has been carried out by the Central Pollution Control Board, New Delhi, at various locations, on monthly basis since December 1994. The water quality data of 22 stations in Yamuna basin is analyzed using the proposed fuzzy comprehensive water quality index (FCWQI) model. A crisp output ranging from 1 to 5 is obtained from the model with respect to each sampling station and is named as fuzzy water quality index (FWQI) of water at that station, where FWQI = 1 representing good and FWQI = 5 representing bad water quality. A total of eight water quality parameters are considered in obtaining FWQI at each site. These parameters are dissolved oxygen (DO),

fecal coliform (FC), biochemical oxygen demand (BOD), pH, temperature (T), nitrates (N), total coliform (TC), and electrical conductivity (EC).

These water quality parameters have been taken into consideration of a detailed literature study of various in-practice WQI in various developing countries (Lermontov et al, 2009; Mahapatra et al. 2011). As the study presents a framework that is much needed in developing economies, the water quality parameters in WQI standards from Iran, Malaysia, India, and South Africa have been taken into account along with the expert’s opinion to select the above eight water quality parameters. Water quality parameters are classified into five classes by taking into consideration of the opinion of experts and Central Pollution Control Board (CPCB), New Delhi. These classes are very good (VG), good (G), fair (F), poor (P), and very poor (VP) as given in Table 2. Each water quality parameter is assigned weightage by decision makers through triangular fuzzy input (U, L, M, V, E) as described in methodology and is shown in Table 3.

On the basis of the subjective judgments of the experts and classes defined in Table 2, attributes were

Table 2 Linguistic classification of water quality parameters

Parameter	Very good	Good	Fair	Poor	Very poor
DO (mg/l)	8 <	4–7.9	3–3.9	2–2.9	1–1.9
FC (MPN/100 ml)	0–499	500–999	1000–1999	2000–2500	2500 <
BOD (mg/l)	0–0.9	1–1.9	2–2.9	3–4.9	5 <
pH	7–8	6.5–6.9 & 8.1–8.5	6–6.4 & 8.6–8.9	5.5–5.9 & 9–9.5	0–5.4 & 9.6–14
T (°C)	23–28	20–22.9 & 28.1–30	15–19.9 & 30.1–32	10–14.9 & 32.1–34	0–9.9 & 34.1 <
N (mg N/l)	0–0.1	0.11–0.50	0.51–1.50	1.51–5	5.1 <
TC (MPN/100 ml)	0–999	1000–1999	2000–3999	4000–4999	5000 <
EC (µmhos/cm)	150–499	100–149 & 500–699	0–99 & 700–999	1000–1999	2000 <

assigned fuzzy values with respect to all the water quality parameters as shown in Table 4. Finally, Yager’s max-min algorithm has been applied to find the FCWQI for all the selected stations of river Yamuna. For example, by Yager’s algorithm, the value of FWQI for sampling station Yamunotri is given by Eq. (7):

$$FWQI = \alpha(VG)^V (VG)^E (G)^L (VG)^E (VG)^M (VG)^V (VP)^M (VG)^M \tag{7}$$

The spatial variation of each water quality parameter is produced using geospatial analysis performed in ArcGIS 10.3. To assess the critical areas in Yamuna basin affected by one or more water quality issue(s), each sampling station is assigned its location in the basin based on their corresponding coordinates using geographical coordinate system of World Geodetic System (WGS)—1984. To quantify the impact on study area, the Yamuna basin is divided into distinct 1157 cells using “GIS tool: Spatial Fish Net” and is shown in Fig. 7. The fishnet grid for the whole study area is developed and the geo cells overlapping with the basin boundary is cropped using Data Management Tool of ArcGIS.

Actual numerical values of each water quality attribute are assigned to each of 19 sampling stations and 3

Table 3 Importance values of linguistic terms used to define quality parameters

Parameter	Importance	Parameter	Importance
DO	E (0.9167)	T	M (0.5000)
FC	E (0.9167)	N	M (0.5000)
BOD	V (0.7500)	TC	M (0.5000)
pH	V (0.7500)	EC	L (0.2500)

reference stations and the data is interpolated spatially using GIS inverse reverse distance (IDW) interpolation. Using spatial zonal statistics with geospatial fishnet as input shape class, interpolated data is distributed and classified into five quality standard classes (very poor, poor, fair, good, and very good). Figure 8 showing dissolved oxygen, fecal coliform, and biochemical oxygen demand and Fig. 9 showing pH, temperature, and nitrate and Fig. 10 showing total coliform and electrical conductivity zonal statistics are presented below.

DO is essential to support aquatic life and is highly sensitive with changes in the environment. Low-dissolve oxygen levels result in anaerobic conditions which cause bad odors. The spatial distribution shows that there is a DO deficit and critical BOD levels in downstream of Delhi. These critical levels of DO and BOD are an indication of high effluent discharge from Delhi. Data from the literature shows that at current generation rate of sewage and maximum Delhi’s treatment capacity, even if the effluent from Delhi (joining river Yamuna) is treated with a target BOD value of 10 mg/l, it may produce about 179 t/d of BOD from Delhi alone. This may result in BOD levels of around 46 mg/l in the final effluent that will pollute the Yamuna river beyond allowable quality standards. Another data shows that if hypothetically the entire sewage of Delhi is treated to low levels of 5 mg/l of BOD, still the BOD effluent load in the final discharge would be 19.4 t/d, which is still well beyond allowable standards (Upadhyay and Rai 2013).

High-temperature imbalance can be seen in Yamuna river at D/S Delhi and in U/S Betwa catchment. River temperature is one of the most important criteria needed to be addressed while assessing water quality as it indirectly affects other parameters in

Table 4 Analytical data of water samples in Yamuna basin (2016)

Station name	X	Y	BOD	DO	EC	FC	N	pH	T	TC	FWQI
Yamunotri	78.45	31.02	VG	VG	G	VG	VG	VG	VP	VG	2.65
Dak patthar	77.80	30.50	G	VG	G	VG	VG	VG	G	VG	1.88
Paonta sahib	77.60	30.45	G	G	F	VG	F	G	G	VG	2.25
Sonepat	77.02	28.99	VP	VG	P	VP	G	VG	VG	VP	2.88
Shahdara drain	77.32	28.62	VP	VP	P	VP	VG	VG	F	VP	3.08
Vrindavan	77.66	27.57	VP	VG	P	VP	VG	G	P	VP	2.91
Agra	78.01	27.18	VP	VG	P	VP	P	G	G	VP	2.91
Allahabad	81.85	25.44	F	VG	G	VP	F	G	G	VP	2.81
Etawah	79.00	26.81	VP	VG	P	VP	P	VG	F	VP	3.13
Saharanpur	77.55	29.97	VP	VP	F	VP	VG	VG	G	VP	2.88
Hamirpur	76.52	31.69	P	G	VG	P	VG	G	P	VP	3
Beena	78.20	24.18	F	VG	VG	VG	VP	F	P	VG	2.99
Raisen	77.78	23.33	G	G	VP	VG	G	VG	G	G	2.26
Bhojpur	84.52	25.47	F	G	G	VG	P	F	G	VG	2.67
Vidisha	77.81	23.53	P	VG	G	VG	G	F	G	VG	2.86
Ujjain	75.78	23.18	VP	G	P	G	P	F	VG	G	2.96
Nagda	75.42	23.45	VP	G	VP	F	VP	G	VG	G	2.85
Kota	75.86	25.21	G	G	VG	VG	P	VP	F	VG	2.85
Rampura	75.44	24.47	F	G	G	VG	P	G	F	VG	2.67
Dholpur	77.89	26.70	F	VG	F	VG	P	F	G	VG	2.76
Jaipur	75.79	26.91	VP	G	P	VG	VP	P	F	VG	2.94
Vasad	73.06	22.46	P	VG	VG	VG	F	F	G	VG	2.86

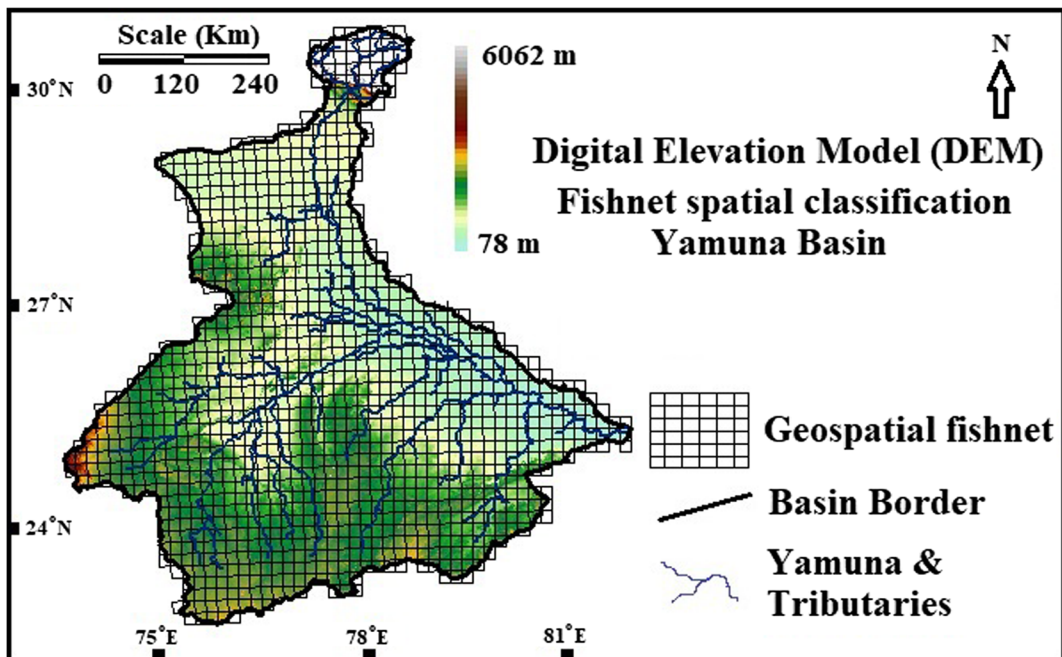


Fig. 7 Geospatial classification of Yamuna basin using multiple polygon features

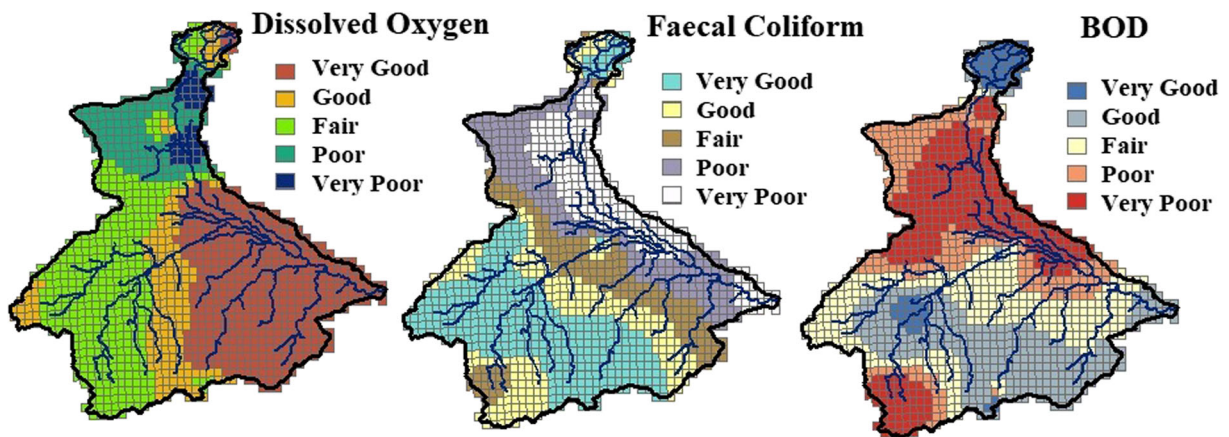


Fig. 8 Spatial variation of water quality parameters (DO, FC, and BOD)

terms of severity and impact on river ecosystem. River temperature influences physio-chemical characteristics of water like solubility of oxygen (directly affecting DO) and other gases, rate of chemical reactions and toxicity of chemical compounds, and metabolisms of aquatic microbial organisms (Dallas and Day 2004). Vulnerability and tolerance of aquatic microorganisms toward toxic compounds and metals (cyanide, zinc, phenol, xylene) get critically affected with the rise in river temperature (Bhadja and Vaghela 2013 and Wurts 2012). High river water temperature accelerates the process of oxygen depletion in case of high organic loading (Washington State Department of Ecology 1991). Nitrogen-based compounds are well known for their toxicity in high pH solvents, but high temperature also influences chronic criteria concentrations of these compounds (EPA 2013).

Alarmingly, high nitrate concentrations can be found at Yamuna–Tributaries junction in rivers originating from Rajasthan and Madhya Pradesh (Chambal, Sind, and Betwa catchment). The presence of nitrates in large amount may result in mechanisms of accelerated eutrophication, which indirectly leads to anoxic conditions (algae boom) in the water streams (Elmanama et al. 2006). Agricultural runoff and urban land use are major sources of nitrate pollution in surface water bodies, through sediment runoff mechanism nitrates can get be excessively loaded in rivers receiving the catchment water output. The excess of this nitrogen loading results in an excessive algae growth and mineral imbalance in river aquatic habitats (Elliott et al. 2005). Nitrates indirectly influence BOD levels through biological influences that reinforces micro-biological metabolisms and thus increases biochemical oxygen demand (Jannasch 1968).

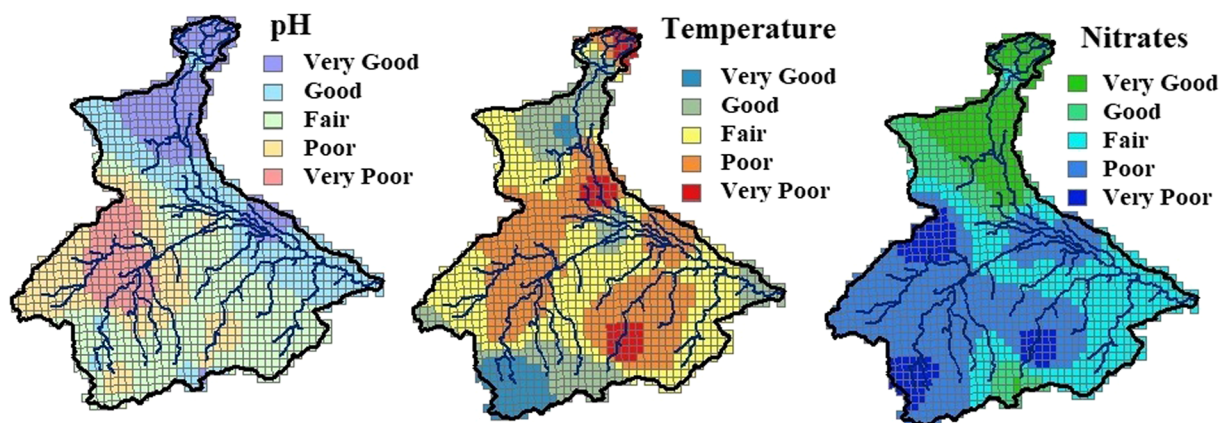


Fig. 9 Spatial variation of water quality parameters (pH, T, and N)

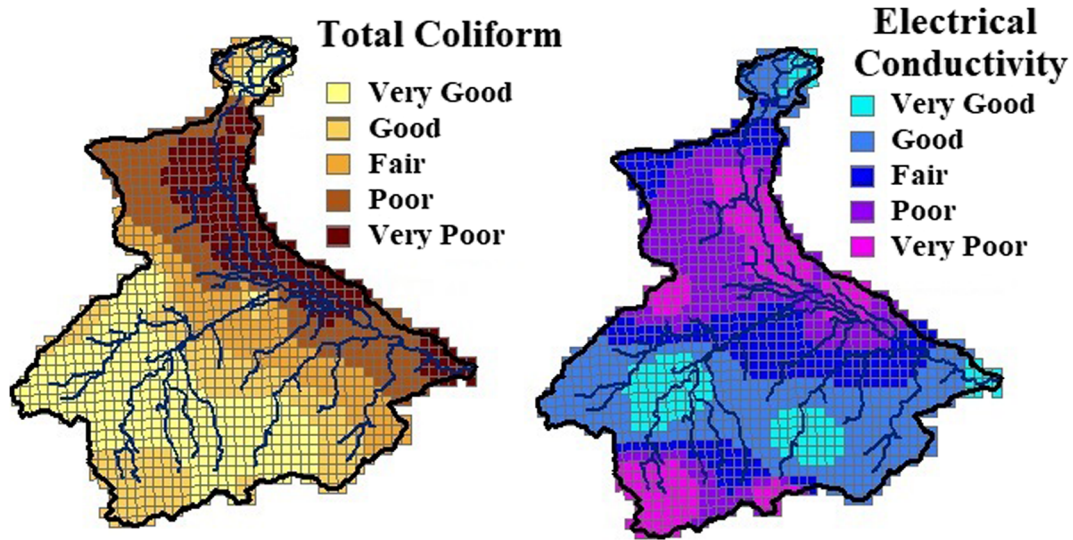


Fig. 10 Spatial variation of water quality parameters (TC and EC)

High count of FC and TC in middle Yamuna stretch and D/S of Delhi and a corresponding high EC value can also be inferred from the spatial distribution shown in Fig. 10. Currently, 8500 MLD of sewage and industrial wastewater is generated in the Yamuna basin, out of which more than 4500 MLD of sewage is disposed of directly in Yamuna river and about 1200 MLD is dumped in its subsidiaries (Chambal, Betwa, Ken). Remaining 2800 MLD of effluent is either disposed of on land or used for irrigation purpose (Upadhyay and Rai 2013). Due to rapid population growth, this wastewater generation is expected to increase significantly with estimated prediction of 17,200 MLD by year 2030. Delhi having a very high population density generates about 3800 MLD of wastewater effluent, out of which 1700 MLD account for sewage discharge (CPCB 2017).

Similar to the water quality spatial distribution, zonal statistics for fuzzy water quality index for each fishnet polygon feature class is obtained from interpolated raster data. The spatial variation is later represented by mapping normalized FWQI for each polygon feature class of the Yamuna basin as shown in Fig. 11.

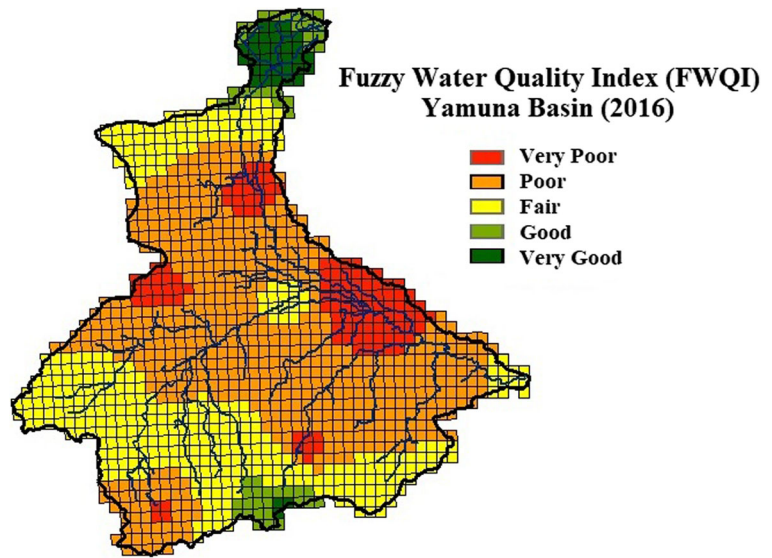
River Shahibi (D/S), river Hindon, river Betwa (U/S), and river Chambal (U/S) along with river Yamuna (Middle river junction) lie in critical FWQI class/zone. High effluent discharge from Delhi and pollutant loading from tributaries at Yamuna junction justify the criticality of regions across Yamuna river. A more detailed approach toward the significance of each water

quality parameter corresponding to calculated FWQI number is required to pinpoint the most critical parameter for further analysis. Next section aims to delineate the sensitivity of the FCWQI model toward the spatial location of interest and the parameter dominant in that spatial region (in terms of pollution and criticality).

Spatial and parametric sensitivity of FWQI model

In the above section, an attempt has been made to assess critical quantization of water quality in various regions of Yamuna basin in terms of a unique fuzzy water quality index. However, it is essential to determine dominating parameters which are responsible to cause criticality in that spatial region. Therefore, for pinpointing the parameter(s) for degraded FWQI at any given spatial location, the study uses approach of geographically weighted regression. Considering unique FWQI for each fishnet polygon as the dependent variable and all other parameters (taken one at a time) as predictor variables, an R^2 spatial distribution has been obtained. The distribution of regression coefficient (R^2) and standard residual for geographically weighted regression with DO as predictor variable is shown in Fig. 12. It can be seen that the maximum R^2 value class obtained is in the order of 0.23–0.33, although DO was given highest weightage while formulating FCWQI model. The results indicate that the obtained FWQI values are less sensitive to DO as a water quality

Fig. 11 Spatial variation of FWQI across Yamuna basin



parameter. Although DO is given high weightage while modeling, but due to very less spatial variation in Yamuna basin, its significance toward deciding FWQI is normalized. The maximum sensitivity ($R^2 = 33\%$) of FCWQI model can be seen in the south-east region of the basin in river Ken catchment.

Similarly, spatial sensitivity of FCWQI model over Yamuna basin is determined in terms of spatial

distribution of regression R^2 value. Maximum spatial correlation from geographically weighted regression model has found out to be 0.91, 0.89, 0.83, 0.79, and 0.79 for parameters BOD, EC, pH, FC, and TC, respectively. For these parameters, the most critically sensitive spatial region found in the basin is the Yamuna catchment outlet and Yamuna river junction. The results indicate that

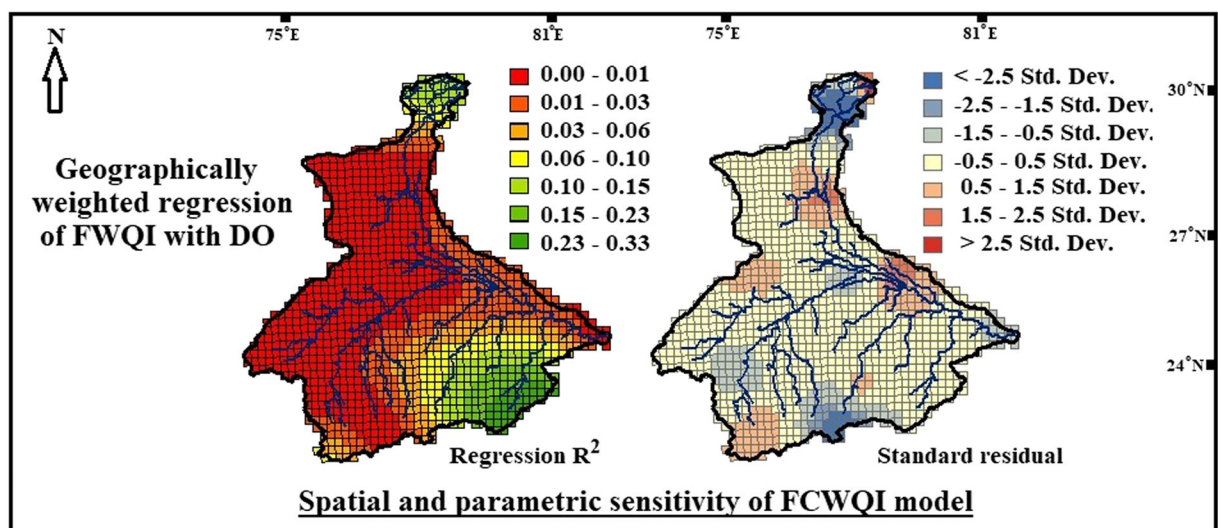


Fig. 12 Spatial and parametric sensitivity of FCWQI model using geographically weighted regression

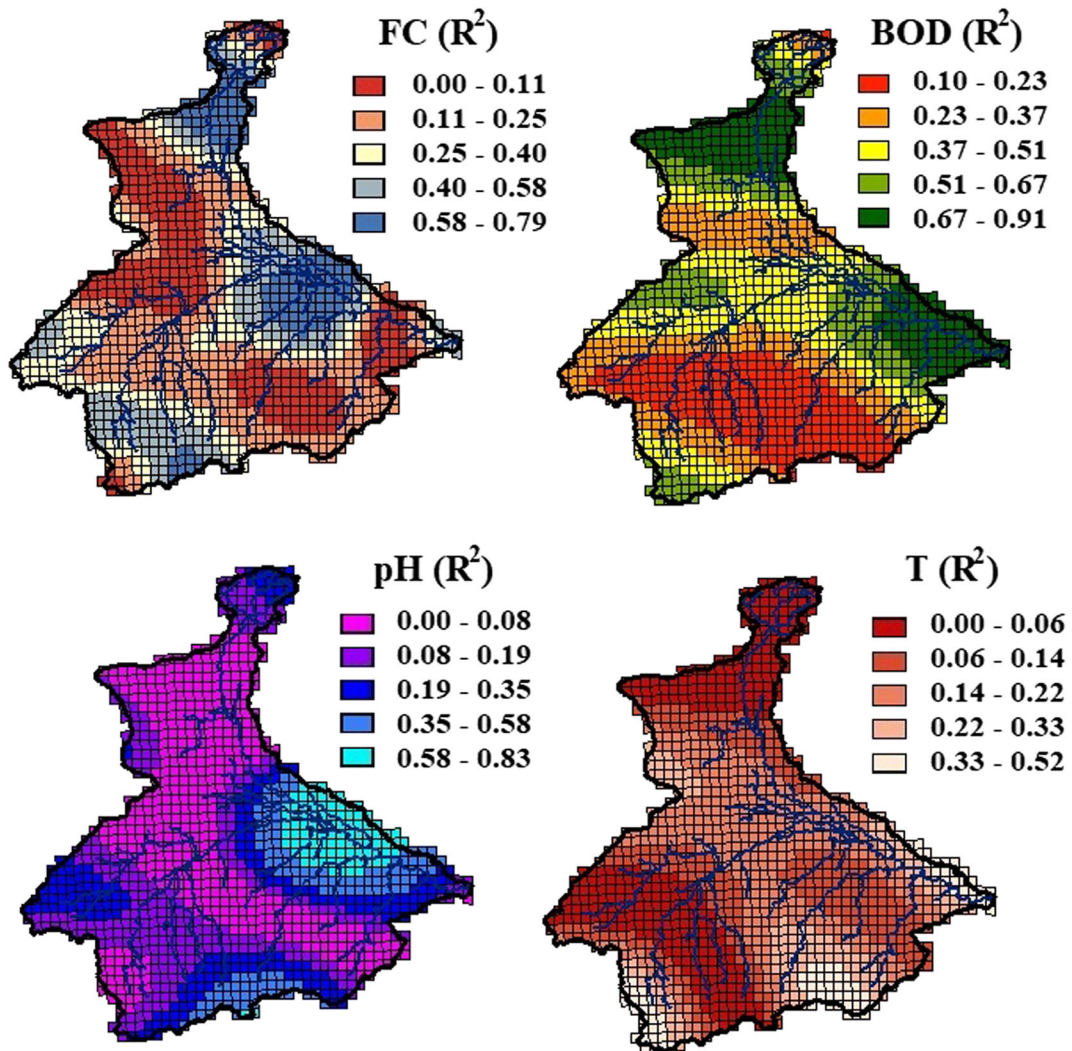


Fig. 13 Spatial sensitivity of FCWQI model with respect to parameters FC, BOD, pH, and T

overall status of a river basin water quality can be best described by measuring parameters like BOD, EC, pH, FC, and TC at the catchment outlet and at river junction of the parent river. These parameters are less sensitive towards their surrounding environment and relatively stable when compared with parameters like DO and T. Variation of spatial sensitivity for these parameters in Yamuna basin is shown in Figs. 13 and 14.

The study follows the rational arguments of defining parametric sensitivity with geographically weighted R^2 value and considering the region with highest correlation class as the most sensitive region with respect to a given parameter. Sensitivity of FWQI toward all considered water quality parameters has been calculated

and critical spatial regions for that parameter have been found, which are tabulated in Table 5.

From fuzzy comprehensive water quality index model, it can be inferred that the model is the most sensitive towards BOD, pH, and EC. FCWQI model is moderately sensitive toward FC and TC and is relatively less sensitive toward T, N, and DO due to their low variability across the basin. Results obtained in Table 5 indicate the dominating pollution parameters at critical locations as specified earlier in Fig. 11. Some of the justifications of criticality at various locations are as follows:

- BOD, pH, and EC for Delhi (D/S) and Yamuna junction: In Yamuna action plan—I (1993), 16 low cost and low energy upflow anaerobic sludge

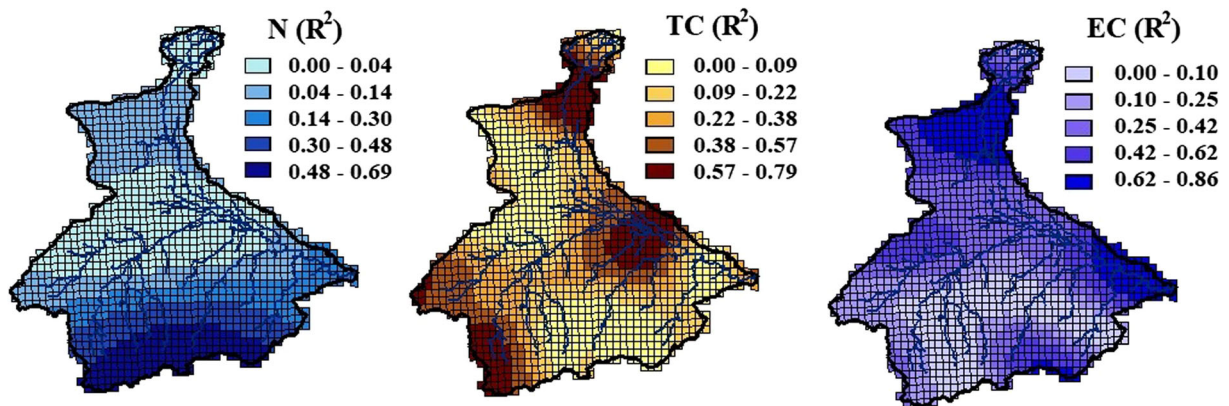


Fig. 14 Spatial sensitivity of FCWQI model with respect to parameters N, TC, and EC

blanket reactors (UASBR) were constructed along Yamuna river to restore its water quality. Unfortunately, UASBR plants yields in an extremely poor quality of effluent discharge that are high in BOD, EC, and FC levels (Mittal 2008). Decade-old non-sustainable treatment infrastructures along Yamuna river need to be modified with high-end sludge treatment technologies under strict effluent discharge regulations. It can help in improving the water quality in this region.

- BOD, T, and EC for Yamuna catchment outlet: Restrictions on high-temperature discharge from industrial outlets and regulations of barrages (that disturbs natural flow of Yamuna) can help in improving water quality of this region in an optimal way.
- T for Chambal river origin: Strict regulations on industrial discharge with specialized care in discharging high-temperature effluents in Yamuna would restore water quality within the prescribed standards.

- N for Betwa river origin: Proper management of nutrient loading from agricultural runoff by adopting organic farming, farmer education on fertilizers use, etc. would assess water quality criticality in this region.

It is to be noted that western critical area of Yamuna basin could not be located in sensitivity based criticality analysis, concluding that standard water quality distribution from Figs. 8, 9, and 10 need to be considered for assessing the water quality issue (pH, N, and EC are found out to be dominant cause by this argument).

The results obtained using the proposed methodology are static in nature; dynamic factors like climate change and geospatial climatic viabilities are not included in representing the water quality status of the study area. However, time-series data inputs and forecasted water quality model inputs into the model can produce a dynamic, realistic, and more representative results using the same methodology. The water quality forecasting models like the one

Table 5 Degree of sensitivity of parameters with critically sensitive region in Yamuna basin

Parameter	R-squared value	Model sensitivity	Most sensitive region
DO	0.23–0.33	Extremely low	Ken and Yamuna river origin
FC	0.58–0.79	Moderate	Delhi (U/S) and Yamuna junction
BOD	0.79–0.91	Extremely high	Yamuna catchment outlet and Delhi (D/S)
pH	0.64–0.83	High	Yamuna junction
T	0.35–0.52	Low	Chambal and Ken river origin and Yamuna catchment outlet
N	0.54–0.69	Low	Kali Sind, Ken and Betwa river origin
TC	0.64–0.79	Moderate	Yamuna junction and Delhi (U/S) and Chambal river origin
EC	0.65–0.86	High	Delhi (U/S) and Yamuna catchment outlet

developed by Ahn et al. (2016) and Rehana and Mujumdar (2012) for river water streams can be integrated with the model presented in the study for a generating priority maps for policymakers. The presented methodology is not limited to surface water quality studies, it can readily be used in groundwater quality assessment. Recent work by Bhakar and Singh (2019), Gorai et al. (2016), and Vadiati et al. (2016) can be improvised by implementing geographic weighted regression to define index sensitivity with respect to individual parameters.

Conclusions

The FCWQI has been formulated for different locations and years on the river Yamuna and the water quality index have been calculated. FCWQI determines the water quality taking into consideration all parameters, and thus, if one parameter is in excess, it might get balanced by another parameter which is scarce. The problem of finding the FCWQI of the stations involved many parameters. A fuzzy approach was used in calculating the FCWQI of the stations. The inputs given by decision makers were in the form of linguistic variables. These were fuzzified with the help of fuzzy membership functions. The importance of fuzzy set theory lies in the fact that the inputs given by the decision makers are subjective in nature. Thus, fuzzy membership functions help in capturing the vagueness of these inputs. A case study of the river Yamuna was performed to demonstrate the applicability of this method. It was found that the normalized FCWQI at Nizamuddin was the least thus indicating poorest quality of water. The normalized FCWQI was the highest at Hathnikund thus indicating best water quality. The index can be very useful in tracking water quality changes at a given sampling site over time and space by comparing them directly among different sites. A threshold value of FCWQI can be fixed. If the FCWQI at any point goes beyond this value, appropriate measures can be taken to restore the water quality. This process can be used for any beneficial use of water as specified by the decision makers. Thus, this methodology has numerous applications and can be used for effective water quality management.

Since the results from the analysis are obtained in linguistic terms and with a regional specific classification of water quality, the results can be readily adopted

by regional as well central governing agencies for developing sustainable policies regarding water quality improvement. Methodology for water quality assessment is not strictly based on a number of preselected water-quality parameters. Hence, the methodology provides a flexible and modifiable water quality assessment framework that can be easily redesigned and implemented on any spatial region and with various other water quality criteria which are not included in the study. The study is a part of a developing country, policies in developing countries face a challenge of economic viability. Thus, the regression maps allow planners to devise realistic and achievable objectives by optimizing their priorities with respect to their implementation capabilities. The limitation of the study can be summarized by stating that the GIS data requirements and computational capacity requirements for application of the methodology drastically increase as we try to adapt the method for more detailed and regional specific analysis.

Compliance with ethical standards

Conflict of interest and acknowledgements The authors declare that they have no conflict of interest. Authors are thankful to Central Pollution Control Board (CPCB), New Delhi, India for sharing information on river Yamuna. All references cited in the text have provided the detailed insight of this important research and therefore are greatly acknowledged. The first author is extremely grateful to Mr. Rajendra Prasad Singh, Retd. Sr. Faculty, J.I.C., Arkha, Raebareli who has taught him basic fundamentals of geographical features of river basin planning in a very conceptualised manner. Authors also express our sincere thanks to the anonymous reviewers and editors for their valuable comments and time.

References

- ADSORBS. (2000). Water quality status of Yamuna river. Assessment and development of river basin series. In *ADSORBS/32/1999–2000*. New Delhi: Central Pollution Control Board.
- Ahn, J. M., Lee, S. J., & Kang, T. (2016). Development of water quality forecasting system with ensemble stream prediction method in the Geum River Basin, Korea. *Desalination and Water Treatment*, 57(2), 670–683.
- Avvannavar, S. M., & Shrihari, S. (2008). Evaluation of water quality index for drinking purposes for river Netravathi, Mangalore, South India. *Environmental Monitoring Assessment*, 143, 279–290.

- Awadallah, A. G., & Yousry, M. (2012). Identifying homogeneous water quality regions in the Nile river using multivariate statistical analysis. *Water Resources Management*, 26, 2039–2055.
- Bhadja, P., & Vaghela, A. (2013). Effect of temperature on the toxicity of some metals to Labeo bata. *International Journal of Advanced Life Sciences*, 6(3), 252–254.
- Bhakar, P., & Singh, A. P. (2019). Groundwater quality assessment in a hyper-arid region of Rajasthan, India. *Natural Resources Research*, 28(2), 505–522. <https://doi.org/10.1007/s11053-018-9405-4>.
- Chau, K. (2006). A review on integration of artificial intelligence into water quality modeling. *Marine Pollution Bulletin*, 52, 726–733.
- Conesa Fernandes-Vitora, V. (1997). *Methodological guide for environmental impact evaluation* (3rd ed.p. 412). Spain: Mundi-Prensa, Madrid.
- CPCB. (2017). Assessment of pollution of drains carrying sewage/ industrial effluent joining River Ganga and its tributaries. In *A joint report by Uttar Pradesh Jal Nigam (UPJN)*. Uttar Pradesh Pollution Control Board (UPPCB), Lucknow and Central Pollution Control Board (CPCB) New Delhi: India.
- Dallas, H. & Day, J. (2004). *The effect of water quality variables on aquatic ecosystems: a review*. WRC Technical Report No. 224/04, Water Research Commission, Pretoria, South Africa.
- Elliott, A. H., Alexander, R. B., Schwartz, G. E., Shanker, U., Sukias, J. P. S., & McBride, G. B. (2005). Estimation of nutrient sources and transport for New Zealand using the hybrid mechanistic-statistical model SPARROW. *Journal of Hydrology (NZ)*, 44(1), 1–27.
- Elmanama, A. A., Afifi, S., & Bahr, S. (2006). Seasonal and spatial variation in the monitoring parameters of Gaza Beach during 2002–2003. *Environmental Research*, 101(1), 25–33.
- EPA. (2013). *Aquatic life ambient water quality criteria for Ammonia – freshwater*. Office of Water: In U.S. Environmental Protection Agency Retrieved from <http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/ammonia/upload/AQUATIC-LIFE-AMBIENT-WATER-QUALITY-CRITERIA-FOR-AMMONIA-FRESHWATER-2013.pdf>. Accessed 14 Sept 2018.
- Fulazzaky, M. A., Seong, T. W., & Masirin, M. I. M. (2010). Assessment of water quality status for the Selangor river in Malaysia. *Water, Air and Soil Pollution*, 205(1–4), 63–77.
- Gorai, A. K., Hasni, S. A., & Iqbal, J. (2016). Prediction of ground water quality index to assess suitability for drinking purposes using fuzzy rule-based approach. *Applied Water Science*, 6(4), 393–405.
- Horton, R. K. (1965). An index number system for rating water quality. *Water Pollution Control Fed.*, 37(3), 300–305.
- Jannasch, H. (1968). Competitive elimination of Enterobacteriaceae from sea water. *Applied Microbiology*, 16, 1616–1618.
- Kaufman, A., & Gupta, M. M. (1991). *Introduction to fuzzy arithmetic, theory and application*. New York: Van Nostrand Reinhold.
- Lermontov, A., Yokoyama, L., Lermontov, M., & Machado, M. A. S. (2009). River quality analysis using fuzzy water quality index: Ribeira do Iguape river watershed, Brazil. *Ecological Indicators*, 9, 1188–1197.
- Liou, S. M., Lo, S. L., & Wang, S. H. (2004). A generalized water quality index for Taiwan. *Environmental Monitoring and Assessment*, 96, 35–52.
- Mahapatra, S. S., Kumar, S., & Panigrahy, B. K. (2011). A cascaded fuzzy inference system for Indian river water quality prediction. *Advances in Engineering Software*, 42(10), 787–796.
- Mandal, P., Upadhyay, R., & Hasan, A. (2009). Seasonal and spatial variation of Yamuna river water quality in Delhi, India. *Environment Monitoring and Assessment*, 170(1–4), 661–670.
- Mittal, A. K. (2008). *Urban run-off: source of major microbial load on Indian Rivers*. (December 15–16) paper presented at the national Conference on Hydraulics, Water Resources and Environment, HYDRO-2008. Jaipur City.
- Mogharreban, N., & Dilalla, L. F. (2006). *Comparison of defuzzification techniques for analysis of non-interval data*. In Fuzzy information processing society, 2006. NAFIPS 2006. Annual meeting of the north American (pp. 257–260). IEEE.
- Nasiri, F., Maqsood, I. H. G., & Fuller, N. (2007). Water quality index: a fuzzy river-pollution decision support expert system. *Water Resources Planning and Management*, ASCE, 133(2), 95–105.
- Phillis, Y. A., & Andriantiatsaholiniaina, L. A. (2001). Sustainability: an ill-defined concept and its assessment using fuzzy logic. *Ecological Economics*, 37(3), 435–456.
- Rehana, S., & Mujumdar, P. P. (2012). Climate change induced risk in water quality control problems. *Journal of Hydrology*, 444–445, 63–77. <https://doi.org/10.1016/j.jhydrol.2012.03.042>.
- Ross, T. J. (2008). *Fuzzy logic with engineering applications* (2nd ed.). New Delhi: John Wiley and Sons.
- Sakawa, M. (1993). *Fuzzy sets and interactive multi objective optimization*. New York: Plenum Press.
- Singh, A. P., & Ghosh, S. K. (2003). Uncertainty analysis in river basin water quality management in integrated water resources planning and management. In K. Srinivasa Raju, A. K. Sarkar, & M. L. Dash (Eds.), *Integrated water resources planning and management* (pp. 260–268). New Delhi: Jain Brothers.
- Singh, A. P., Ghosh, S. K., & Sharma, P. (2007). Water quality management of a stretch of river Yamuna: an interactive fuzzy multi-objective approach. *Water Resources Management*, 21, 515–532.
- Singh, A. P., Chakrabarti, S., Kumar, S., & Singh, A. (2017). Assessment of air quality in Haora river basin using fuzzy multiple-attribute decision making techniques. *Environmental Monitoring and Assessment*, 189(373), 373. <https://doi.org/10.1007/s10661-017-6075-3>.
- Srinivas, R., & Singh, A. P. (2018a). Impact assessment of industrial wastewater discharge in a river basin using interval-valued fuzzy group decision-making and spatial approach. *Environment, Development and Sustainability*, 20(5), 2373–2397. <https://doi.org/10.1007/s10668-017-9994-9>.
- Srinivas, R., & Singh, A. P. (2018b). An integrated fuzzy-based advanced eutrophication simulation model to develop the best management scenarios for a river basin. *Environmental Science and Pollution Research*, 25(9), 9012–9039. <https://doi.org/10.1007/s11356-018-1206-0>.

- Srinivas, R., Singh, A. P., Dhadse, K., Garg, C., & Deshmukh, A. (2018). Sustainable management of a river basin by integrating an improved fuzzy based hybridized SWOT model and geo-statistical weighted thematic overlay analysis. *Journal of Hydrology*, 563, 92–105. <https://doi.org/10.1016/j.jhydrol.2018.05.059>.
- Tirupathi, C., Shashidhar, T., Pandey, V. P., & Shrestha, S. (2019). Fuzzy-based approach for evaluating groundwater sustainability of Asian cities. *Sustainable Cities and Society*, 44, 321–331. <https://doi.org/10.1016/j.scs.2018.09.027>.
- Upadhyay, A., & Rai, R. K. (2013). Water management and public participation. *Springer Briefs in Earth Sciences*. https://doi.org/10.1007/978-94-007-5709-7_2.
- Vadiati, M., Asghari-moghaddam, A., Nakhaei, M., & Adamowski, J. (2016). A fuzzy-logic based decision-making approach for identification of groundwater quality based on groundwater quality indices. *Journal of Environmental Management*, 184(2), 255–270.
- Washington State Department of Ecology. (1991). Chapter 2 – lakes: temperature in lakes. In *A Citizen's guide to understanding and monitoring lakes and streams* Retrieved from <http://www.ecy.wa.gov/programs/wq/plants/management/foysmanual/temperature.html>.
- Wurts, W. (2012). Pond pH and Ammonia Toxicity. *World Aquaculture*, 34(2), 20–21.
- Yager, R. R. (1978). Fuzzy decision making including unequal objectives. *Fuzzy Sets and Systems*, 1, 87–95.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.