Sequential calibration of a water quality model using reach-specific parameter estimates

Shushobhit Chaudhary, C. T. Dhanya and Arun Kumar

ABSTRACT

Calibration is the most critical phase in any water quality modelling process. This study proposes a sequential calibration methodology for any water quality model using reach-specific estimates of model parameters, which would aid in the improved prediction of river water quality characteristics. The proposed methodology accounts for the heterogeneity of river reaches, i.e., diverse characteristics of different reaches on the river stretch. The water quality model, QUAL2K, is coupled with MATLAB, a computing platform, to facilitate sequential estimation of reach-wise model parameters using a grid-based weighted average optimization. The Delhi segment of the Yamuna River is selected as study river stretch. Observations of water quality variables, dissolved oxygen and biochemical oxygen demand are used to calibrate and validate QUAL2K. Desirable performance measures are obtained during the calibration and the validation period. The methodology proves superior to the existing calibration methodologies applied over the study region. The proposed technique also captures the system behaviour effectively, through a systematic, efficient and user-friendly way. The proposed approach is expected to aid decision-makers in formulating better reach-wise management decisions and treatment policies by providing a simpler and efficient way to simulate water quality parameters.

Key words | parameter estimation, QUAL2K, water quality simulation, Yamuna River

Shushobhit Chaudhary
C. T. Dhanya (corresponding author)
Arun Kumar
Department of Civil Engineering,
Indian Institute of Technology Delhi,
New Delhi 110016,
India
E-mail: *dhanya@civil.iitd.ac.in*

INTRODUCTION

Recent decades have witnessed a vast increase in the literature on water quality modelling softwares and applications thereof, all over the world. This can be attributed to the improved awareness about the dwindling water resources and over-exploitation of river reaches, which in turn is due to the steep rise in demand and augmented pollution levels arising from the rapid increase in population. These issues have led to the development, modification and review of various river water quality models (Brown & Barnwell 1987; Whitehead *et al.* 1997; Chapra *et al.* 2008; Lindström *et al.* 2010; Lázár *et al.* 2012). Subsequently, the water quality modelling segment has become an integral part of water resources and environmental management studies (Lázár *et al.* 2012; Liu *et al.* 2015; Chiu *et al.* 2016; Lindström 2016; Wang *et al.* 2016). The majority of these models have many inherent coefficients and rate constants (together known as model parameters) that customize the model for different systems with diverse characteristics. However, most of the time, because of data unavailability and system constraints, these model parameters may not be directly relatable to any physical system components or may not be even measurable physically. Since the availability of physical data of any river system is constrained in water quality modelling studies, the calibration procedure has been given importance (Wood *et al.* 1990; Janssen & Heuberger 1995; Ng & Perera 2003; Liu *et al.* 2007).

Traditionally, calibration has been performed by employing a 'trial-and-error' approach in most water quality modelling studies, where various combinations of model parameters are considered in a trial-and-error manner,

doi: 10.2166/nh.2017.246

until the best fit is obtained. While this approach is quick and effective for small parameter sets, the method becomes cumbersome, sluggish and computationally demanding for relatively large parameter sets (Janssen & Heuberger 1995). Moreover, this calibration approach has a high risk of missing the global optima (Liu *et al.* 2007). In order to overcome the drawbacks of the trial-and-error calibration approach, various automated calibration techniques have been proposed, for example, sequential extended Kalman filters (Bowles & Grenney 1978), knowledge-based expert systems (Wood *et al.* 1990), inverse modelling technique (Shen & Kuo 1996), genetic algorithm (GA) (Mulligan & Brown 1998), shuffled complex evolution algorithm (Van Griensven & Bauwens 2003), hybrid genetic-k-nearest neighbour algorithm (Ostfeld & Salomonst 2005), modified simplex algorithm (Marsili-Libelli & Giusti 2008) and particle swarm optimization approach (Afshar *et al.* 2011). Among these methods, GA, a quick and robust evolutionary algorithm approach, has gained much attention (Goldberg 1989; Ng & Perera 2003; Pelletier *et al.* 2006; Liu *et al.* 2007). GA is powerful optimization technique based on the concepts of natural selection and genetics, and has outperformed traditional optimization approaches in many applications (Goldberg 1989; Mulligan & Brown 1998; Ng & Perera 2003). Mulligan & Brown (1998) revealed that GA is able to accurately determine model parameter sets in different scenarios, although it possesses greater computational demands than the traditional approaches. GA provides additional information about the search space, unlike traditional approaches. Efficient application of GA requires proper selection of GA operators, which are essentially the components that make up the overall GA process (Ng & Perera 2003). However, modern GA approaches can be used as black-boxes and are very easy for application (Pelletier *et al.* 2006).

Although advanced auto-calibration techniques based on different high-level optimization algorithms are available, there is always room for simpler methods, if they are shown to be fit for the purpose. Therefore, in this study we have focused on a simple and systematic method of sequential calibration of a water quality model using distinct model parameter estimates. Sequential calibration means calibration of the segments of river one-by-one. Segmentation of river reaches is done by identifying the length of river stretches with similar hydraulic characteristics. A segment of river with constant hydraulic characteristics is termed as a reach. Sequential calibration is applied in our study by first calibrating the first reach, and using the model parameters of the first reach to calibrate the second reach, and the process is repeated until the entire river stretch is calibrated. The advantage of this systematic method of calibration is that it conserves the heterogeneity of river stretches by giving them an opportunity to adopt different model parameter values for different reaches. Consideration of heterogeneity of river reaches and model parameters is of utmost importance while modelling complex river systems with distinct pollution sources entering at different points along the stretch (Sharma & Singh 2009; Parmar & Keshari 2012; Zhang *et al.* 2012).

The present study aims to develop a methodologic framework for systematic and sequential calibration of a water quality model, considering the heterogeneity of river reaches. The proposed methodology couples a frequently used water quality simulation tool, QUAL2K (Kalburgi *et al.* 2010; Rehana & Mujumdar 2011; Vasudevan *et al.* 2011; Zhang *et al.* 2012; Walling *et al.* 2017), with a computing platform, MATLAB (MATrixLABoratory). The efficacy of the proposed methodology is demonstrated on the Delhi segment of the Yamuna River. A critical comparison of the proposed methodology with the existing and popular approaches in estimating model parameters is presented and the performance metrics are discussed.

STUDY AREA

Yamuna River originates from Yamunotri glacier of the lower Himalayas in the state of Uttaranchal, northern India. The stretch of Yamuna River passing through the Delhi region is considered for this study (Figure 1). This stretch is 21.9 km in length, flowing between Wazirabad barrage and Okhla barrage, and is polluted by 16 drains, which primarily discharge domestic wastewater into the river, as shown in Figure 1. Of these, Najafgarh drain (labelled as Drain 1 in Figure 1) and Hindon cut (labelled as Drain 16 in Figure 1) are the two dominant pollution sources. Other drains are minor drains and contribute only small quantities of wastewater into the Yamuna River (CPCB 2005).

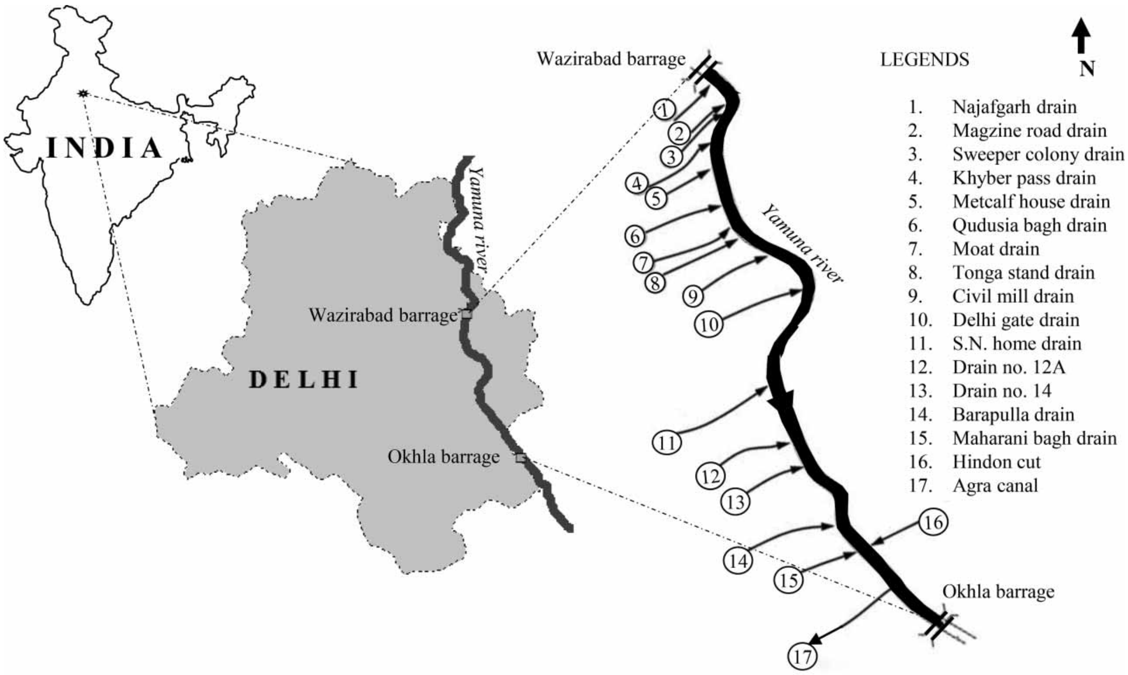

Figure 1 | Schematic showing different drains contributing to the Delhi segment of the Yamuna River.

In addition, water is abstracted through the Agra Canal for irrigation. During the dry season, low flow conditions prevail in the river stretch owing to low rainfall in the catchment area and huge diversions through the upstream barrages (CPCB 1999–2000).

Different water quality modelling techniques have been applied over the study region in the past. Paliwal *et al.* (2007) and Singh *et al.* (2007) assumed water quality model parameters based on the findings of studies such as CPCB (1999–2000), Kazmi (2000) and O'Connor & Dobbins (1958). Parmar & Keshari (2012) used a semi-empirical approach to calibrate the QUAL2E model, in which separate equations were used to determine a few parameters; however, the rest of the parameters were assumed to be constant. Further, Sharma & Singh (2009) and Walling (2014) applied a trial-and-error approach to calibrate STREAM-II and QUAL2K models, respectively.

DATA USED

Details of streamflow, temperature and water quality characteristics (dissolved oxygen (DO) and biochemical oxygen demand (BOD)) at the beginning and the end of the river stretch and its individual segments were obtained. Information on hydraulic characteristics of the river stretch such as length, elevation and location of individual reaches were collected from various sources (Delhi Jal Board 2005; Parmar & Keshari 2012). The average cross-sectional width of the river varies from 60 m to 272 m and the average cross-sectional depth of the river varies from 0.4 m to 6 m along this stretch. Meteorological information of the study area was obtained from the Indian Meteorological Department. Details about the pollution sources, its location, flow rates and water quality parameters were collected from the Central Pollution Control Board (CPCB), India (CPCB 2005; Parmar & Keshari 2012). Average monthly water quality characteristics during the dry season periods of 15 March to 15 June 2002 and February 2003 (CPCB 2005; Parmar & Keshari 2012) were considered in this study (shown in Supporting information; Table S1, available with the online version of this paper). The seasonal period of 15 March to 15 June 2002 was selected as the calibration period and February 2003 as the validation period. The calibration and validation periods were kept the same as those of previous studies such as Parmar & Keshari (2012) and Walling (2014) to ensure a one-to-one comparison of the calibration procedures adopted by previous studies with the proposed methodology.

METHODOLOGY

The water quality modelling tool QUAL2K was selected for application in this study. QUAL2K is a one-dimensional river water quality model in which river stretch is divided into a series of smaller reaches which are further subdivided into smaller computational elements for internal analysis (Chapra *et al.* 2008). QUAL2K has been widely calibrated using a trial-and-error approach (Kalburgi *et al.* 2010; Rehana & Mujumdar 2011; Vasudevan *et al.* 2011; Walling *et al.* 2017). However, in the present study, we adopted a sequential, reach-specific and automatic calibration approach for QUAL2K by coupling it with the computational platform MATLAB. Figure 2 presents a schematic of the proposed approach, which is further explained below.

River segmentation

The 21.9 km Delhi stretch of the Yamuna River was divided into 16 reaches considering the diverse hydraulic characteristics and locations of pollution sources (Figure 3). These reaches are of varying lengths and varying hydraulic characteristics. However, within each reach, the hydraulic and biochemical characteristics were assumed to be constant.

Initialization of QUAL2K

The QUAL2K model was initialized with river hydraulic characteristics, meteorological characteristics, flow and water quality characteristics of March–June 2002. Headwater boundary conditions were incorporated in terms of flow and pollution load specifications of river flow entering at the upstream of Wazirabad barrage. Also, the flow and pollution load specifications of drains, which describe the energy and mass transfer within the river reach, were input to the model.

Selection of model parameters

The choice of model parameters in a water quality modelling approach lies in the underlying assumption of the hydrological and biochemical processes considered for simulation. For example, for simulation of DO and BOD

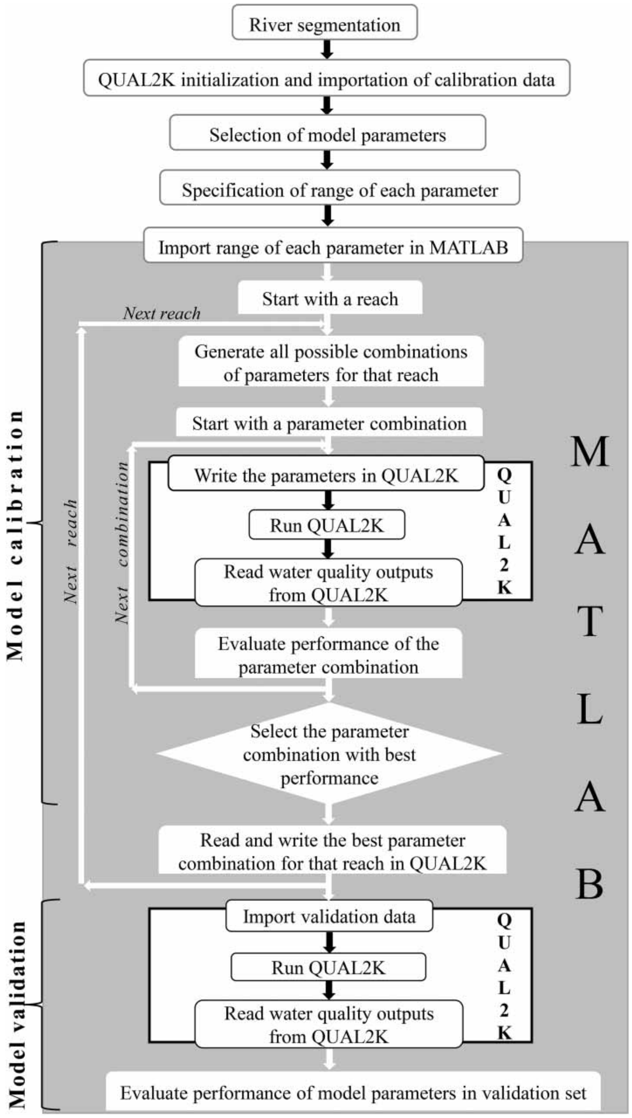

Figure 2 | Schematic of the proposed sequential calibration technique.

levels in rivers, kinetic interactions among various sources and internal sinks of DO and organic matter in the river play a pivotal role (Thomann & Mueller 1987; Chapra *et al.* 2008). Therefore, DO and BOD levels would depend on re-aeration rate, BOD decay rate, benthic oxygen demand, nitrification rate, phytoplankton respiration and growth rate, benthic algal growth and respiration rate and zooplankton respiration rate, BOD settling and decay rate, zooplankton death and excretion rates, phytoplankton death rate and benthic algal death rate. Depending on regional conditions, some of the above-mentioned

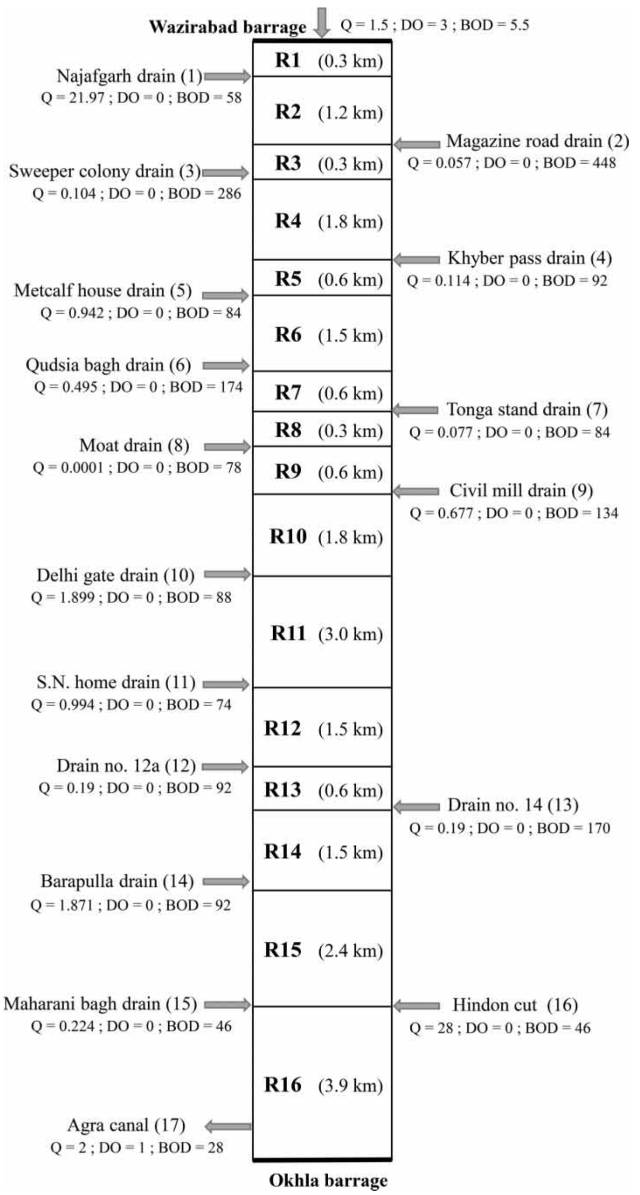

Figure 3 | Segmentation of the Delhi stretch of the Yamuna River. Details of discharge (Q in $m^3 s^{-1}$) and water quality variables – DO (mg L^{-1}) and BOD (mg L^{-1}) during average conditions of March–June 2002 are given.

parameters might have negligible roles. Therefore, information on regional conditions is required for the selection of significant parameters.

For the Yamuna stretch, preliminary information about model parameters was obtained from previous studies on this region. A review of these studies (Table 1) suggests that some parameters play negligible roles in DO and BOD simulations of the river stretch considered. Oxygen re-aeration rate, BOD hydrolysis rate and BOD oxidation rate are found to be significant parameters over the study region and, therefore, considered for calibration in this study.

Table 1 | Model parameters with negligible roles in DO and BOD simulations for the Yamuna River (Delhi stretch)

S. No.	Model parameters	Reference
1	BOD settling rate	CPCB (1999–2000), Kazmi (2000) and Vasudevan et al. (2011)
2	Phytoplankton growth rate, death rate and respiration rate	Kazmi (2000)
3	Zooplankton respiration rate, death rate and excretion rate	Kumar (2001)
4	Sediment oxygen demand	Kazmi (2000) and Paliwal et al. (2007)
5	Algal growth rate, death rate and respiration rate	Parmar & Keshari (2012)

Specification of range of model parameters

Once the significant model parameters were selected, the next step was to specify the range of values for each parameter. Direct empirical equations were available for some parameters; however, they are highly biased towards the ambient conditions in which these equations are derived (Melching & Flores 1999). Parmar & Keshari (2012) empirically derived a re-aeration equation for the Delhi segment of the Yamuna River, but the performance of this equation was found to be poor for downstream reaches. Therefore, instead of using a single site-specific equation, a review of the literature was carried out to quantify the approximate range of model parameters.

We considered the user manuals of QUAL2E (Brown & Barnwell 1987), QUAL2K (Chapra et al. 2008) and QUAL2Kw (Pelletier et al. 2006), Environment Protection Agency Guiding document (USEPA 1985) and Zhang et al. (2012) for obtaining ranges of model parameters. Oxygen re-aeration rate (day^{-1}) of any river was found to vary between 0 and 100. However, in order to shorten the range specific to the study area, various equations from the relevant literature (shown in Table 2) were also

Table 2 | Equations for estimating oxygen re-aeration rate (K_a)

Equation for oxygen re-aeration rate, K_a	Reference	Application in DO modelling of rivers
$K_a = 3.93\frac{U^{0.5}}{H^{1.5}}$	O'Connor & Dobbins (1958)	Singh et al. (2007)[a] and Paliwal et al. (2007)[a]
$K_a = 5.026\frac{U}{H^{1.67}}$	Churchill et al. (1962)	Sharma & Singh (2009)[a]
$K_a = 5.14\frac{U}{H^{1.33}}$	Langbein & Durum (1967)	Wallace & Daque (1973)
$K_a = 5.32\frac{U^{0.67}}{H^{1.85}}$	Owens et al. (1964)	Ghosh & McBean (1998)
$K_a = 5.792\frac{U^{0.5}}{H^{0.25}}$	Jha et al. (2005)	Jha et al. (2005)
$K_a = 4.27\frac{U^{0.47}}{H^{2.09}}$	Parmar & Keshari (2012)	Parmar & Keshari (2012)[a]

U is average flow velocity in ms^{-1}; H is average depth of flow over the river in m.
[a]Indicates studies conducted on the Yamuna River.

considered and, finally, a smaller range of values of re-aeration rate was estimated. Thus, oxygen re-aeration rate (day^{-1}) was considered to lie between 0.02 and 4.1. Values of BOD hydrolysis rate (day^{-1}) and BOD oxidation rate (day^{-1}) were found to lie between 0.04 and 4.2.

Model calibration

A sequential calibration approach was adopted in which 16 reaches were calibrated one at a time, until the entire stretch was calibrated (as shown in Figure 4). Steps involved in the sequential calibration approach were as follows.

Step 1: Calibration of the first reach

(a) The selected model parameters were discretized at a smaller interval of 0.1 within their specified range. This was done in order to explore the possibility of each of the discretized values in representing the best performing model parameters. A river reach will not necessarily have the same value for all the model parameters. This approach generates numerous possible combinations of discretized model parameters values for a particular reach. For instance, if p number of parameters are discretized to take k different values, a total of p^k combinations would be generated for the test. In the present study, for calibration of three model parameters, i.e., oxygen re-aeration rate, BOD hydrolysis rate and BOD oxidation rate, at each reach, a total of 77,658 possible combinations ($42 \times 43 \times 43$) of model parameters was generated. MATLAB was used to generate all possible combinations of model parameters.

(b) The performance of each of these combinations in simulating the historic water quality conditions of DO and BOD was analysed using different performance measures. MATLAB provided a platform to develop different user-defined performance measurement functions for calibration.

(c) Three model performance metrics, namely, index of agreement (IOA), correlation coefficient (R) and coefficient of efficiency (E) were considered (details in Table 3) in this study. IOA is the standardized measure of the degree of error in the model simulations. The denominator in IOA represents the largest 'potential error', i.e., the squared sum of absolute deviations of observed and modelled values from the observational mean. IOA values closer to 1 indicate better model performance. R is the measure of the degree of linear relationship between the observed and simulated data with an ideal value of 1. E represents the degree of improvement in model simulations from the observations. E ranges from $-\infty$ to 1; a negative value indicates that observed mean is a better predictor than the model, 0 indicates observed mean is as good a predictor as the model and positive value indicates the model is a better predictor than the observed mean.

(d) Finally, the best parameters' combination for the first reach (r_1) were selected by maximizing the optimization function (OF_{r_1}) defined as:

$$OF_{r_1} = maximize\,[\{IOA + R + E\}_{DO_{r_1}} + \{IOA + R + E\}_{BOD_{r_1}}]$$

This corresponds to optimization based on a grid search since OF_{r_1} is maximized on different values of model parameters generated at specific grids. This is a simple approach and is only really feasible for fast models with few parameters (and not too fine a discretization).

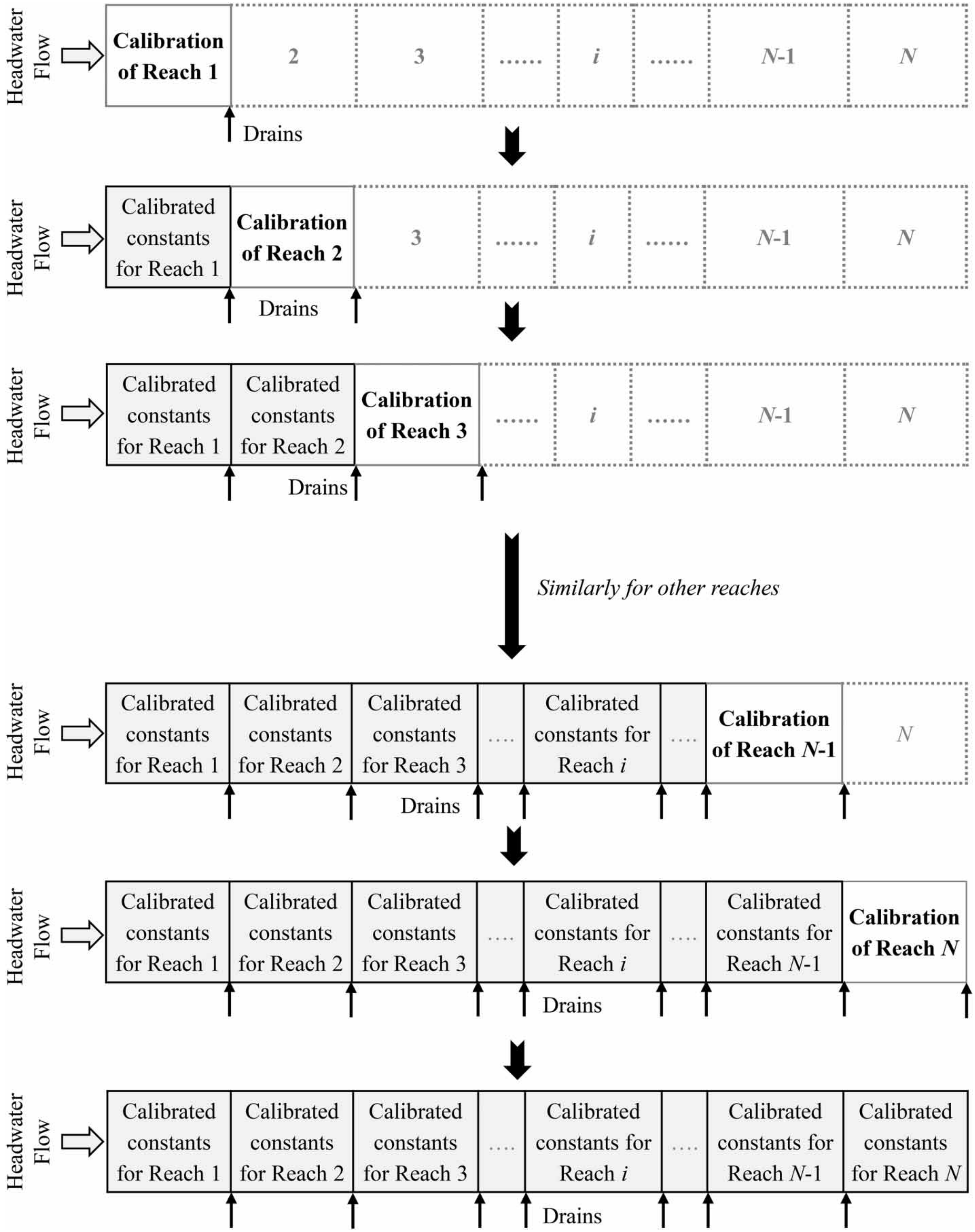

Figure 4 | Steps for conducting sequential calibration of all reaches (N = total number of reaches in the river stretch).

Table 3 | Performance measures used for optimization

Model performance measures	Equation	Range	Ideal value
IOA	$IOA = 1 - \frac{\sum_{i=1}^{n}(p_i - o_i)^2}{\sum_{i=1}^{n}(\lvert p_i - \bar{o}\rvert + \lvert o_i - \bar{o}\rvert)^2}$	[0,1]	1
Correlation coefficient (*R*)	$R = \frac{n\sum p_i \times o_i - (\sum p_i) \times (\sum o_i)}{\left\{\sqrt{n \times (\sum p_i^2) - (\sum p_i)^2}\right\}\left\{\sqrt{n \times (\sum o_i^2) - (\sum o_i)^2}\right\}}$	[−1,1]	1
Coefficient of efficiency (*E*)	$E = 1 - \frac{\sum_{i=1}^{n}(p_i - o_i)^2}{\sum_{i=1}^{n}(p_i - \bar{o}_i)^2}$	(−∞,1]	1

p_i and o_i are the predicted and the observed value respectively; $\bar{o}$ is mean of the observed variable; n is the total number of values.

Step 2

Once the first reach was calibrated, the second reach (r_2) along with the first reach (with fixed parameter values) was considered. Steps (a), (b), (c) and (d) were repeated considering the two segments together. The optimization function (OF_{r_2}) was modified as:

$$OF_{r_2} = maximize\ [OF_{r_1} + \{IOA + R + E\}_{DO_{r_2}} + \{IOA + R + E\}_{BOD_{r_2}}]$$

Step 3

The steps were repeated by adding the other segments one-by-one, until all the 16 segments were calibrated sequentially. The optimization function (OF_{r_i}) for the i^{th} reach, in general, is:

$$OF_{r_i} = maximize\ [OF_{r_{i-1}} + \{IOA + R + E\}_{DO_{r_i}} + \{IOA + R + E\}_{BOD_{r_i}}]$$

The calibration of the whole stretch was thus accomplished, ensuring the heterogeneity of each stretch. For calibration, the average conditions of March–June 2002 were used.

Model validation

The QUAL2K model was validated for average conditions of February 2003. The effectiveness of calibrated model parameters in simulating the historic water quality conditions for the validation period was evaluated using the same performance metrics.

RESULTS

The proposed sequential calibration technique was implemented for modelling the water quality of a stretch of the Yamuna River.

Calibration of QUAL2K model

Model parameters obtained after the calibration of all 16 reaches of the study stretch are shown in Table 4. Performance measures estimated during the calibration period using the best model parameter combination for the entire river stretch are shown in Table 5. Oxygen re-aeration rate (day^{-1}) varies from 0.2 to 4.1, with higher values for latter stretches. Walling (2014) also observed an increase in the re-aeration rate for latter stretches. The increase in oxygen re-aeration rate for latter stretches may be attributed to an observed steep temperature increase in the downstream of river reaches (shown in Supporting information, available with the online version of this paper). Previous studies on the Yamuna River have also reported that re-aeration rate (day^{-1}) varied from 0.02 to 4.0 (Paliwal *et al.* 2007; Singh *et al.* 2007; Parmar & Keshari 2012). Reach-to-reach variation in re-aeration rate may be attributed to a change in variables such as flow velocity, water temperature, etc. (shown in Supporting information; Figures S1 and S2).

BOD hydrolysis rate (day^{-1}) was found to be 0.6 for most of the stretches. However, a maximum value of 3.6 and a minimum value of 0.2 were observed for some

Table 4 | Calibrated model parameters for the 16 reaches

Reach number	O_2 re-aeration rate at 20 °C (day^{-1})	BOD hydrolysis rate at 20 °C (day^{-1})	BOD oxidation rate at 20 °C (day^{-1})
1	0.40	0.60	4.20
2	0.40	0.60	4.20
3	0.40	0.60	4.20
4	0.40	0.60	0.40
5	0.20	0.20	0.40
6	1.80	3.00	1.20
7	0.80	3.60	0.80
8	1.20	3.60	4.20
9	1.80	0.60	0.60
10	0.60	0.60	0.60
11	1.20	0.60	0.60
12	1.80	0.60	0.60
13	4.10	0.60	4.20
14	3.60	0.60	0.60
15	3.60	0.60	0.60
16	3.60	1.20	1.20

segments of the stretch. Similarly, BOD oxidation rate (day^{-1}) varied from 0.4 to 4.2, with most of the reaches taking a value of 0.6. However, a few reaches show a steep increase in BOD oxidation rate to a value of 4.2 (day^{-1}), which then quickly declined to 0.6 (day^{-1}). These variations in BOD hydrolysis rate and BOD oxidation rate along different stretches of the river can be attributed to (i) the difference in biochemical characteristics of pollution loads contributed by different drains and (ii) variations in water temperature (Bowie *et al.* 1985).

The three model parameters selected in this study were also subjected to local sensitivity analysis to find any dormant parameter, if any, which may have affected the calibration process. Each of the parameters was varied one at a time within their domain and the corresponding effect on DO and BOD simulation evaluated. However, all three parameters were found to be sensitive for the simulation of either DO or BOD of the study river stretch (shown in Supporting information; Figures S3–S5).

Table 5 | Model performance measures for calibration and validation periods

	Calibration period (March–June 2002)		Validation period (February 2003)	
Model performance measures	DO	BOD	DO	BOD
IOA	0.995	0.980	0.972	0.955
Correlation coefficient (R)	0.994	0.961	0.954	0.919
Coefficient of efficiency (E)	0.982	0.918	0.904	0.839

The simulated DO and BOD profiles for the calibration period (March–June 2002), with the best combination of parameters, are shown in Figure 5, which indicates an acceptable match between simulated and observed data. The model performance measures (Table 5) also indicate a good fit in DO and BOD simulations. It is observed that due to huge load discharge from the Najafgarh drain into the Yamuna River, DO concentration drops drastically from 5.5 mg L^{-1} to 0 mg L^{-1} in the upstream stretches. Reduction in BOD load for the latter part of the river segment results in lesser consumption of oxygen for pollutants' decomposition. Also, higher re-aeration rates in

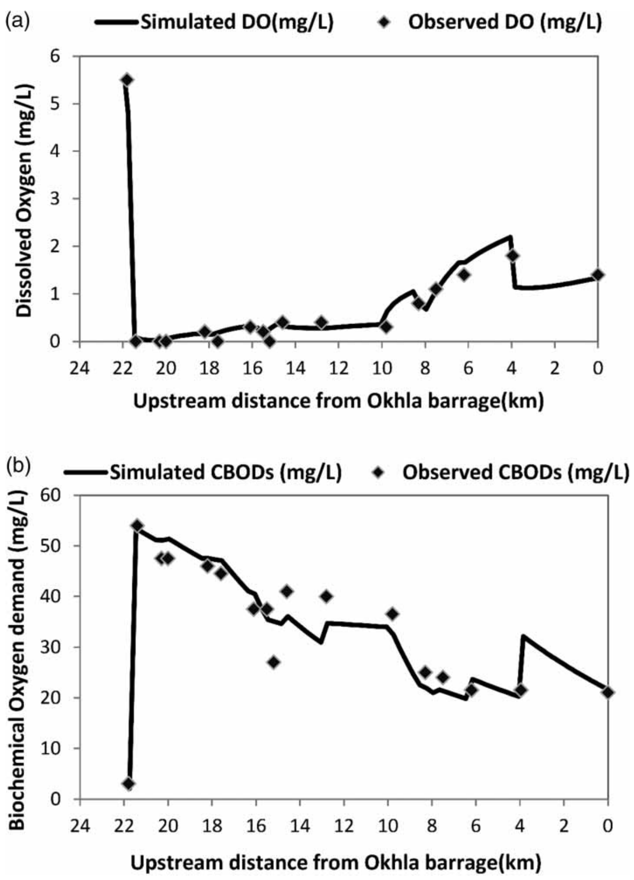

Figure 5 | Simulated (a) DO and (b) BOD profiles for the calibration period (March–June 2002).

latter stretches increase DO values, which results in a moderate gain in DO at latter stretches. Except for the upstream segment of the Yamuna River stretch, the DO concentration falls below 2 mg L^{-1} reaching a negligible DO concentration in most of the reaches. The BOD profile showed a sharp increase, 64 mg L^{-1} at entry which further decreased to 20 mg L^{-1} for the subsequent reaches. The permissible DO and BOD concentrations for the aquatic ecosystem, i.e., DO values greater than 4 mg L^{-1} and BOD lesser than 5 mg L^{-1} (river water class type 'D'; CPCB 1999–2000) was found to be violated throughout the river stretch.

Validation of QUAL2K model

The calibrated model wass then validated using the data for February 2003. Figure 6 presents the profiles of simulated and observed DO and BOD. A good agreement is observed between the simulated and observed data of DO and BOD values, with high model performance measures (as shown in Table 5). The validation results indicate that the calibrated model is able to effectively simulate the system characteristics, along with its heterogeneity.

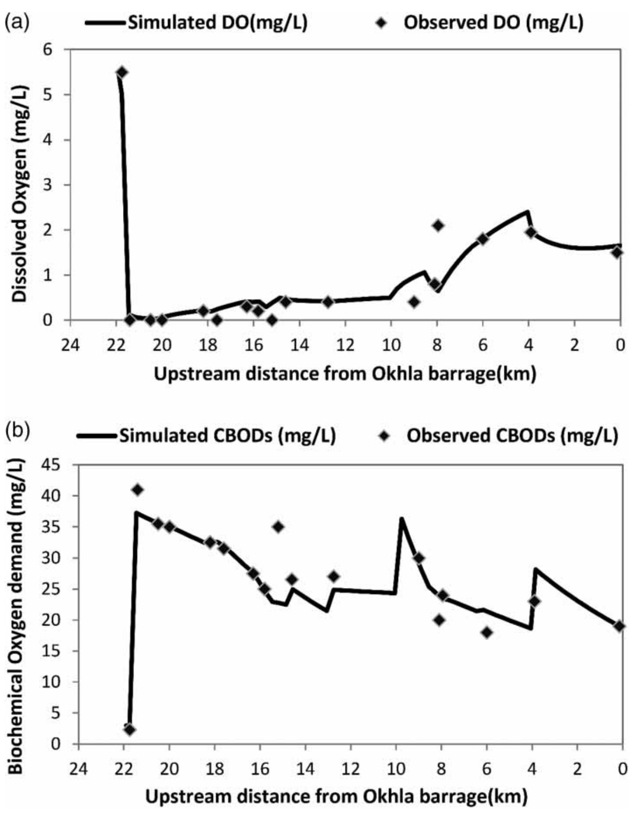

Figure 6 | Simulated (a) DO and (b) BOD profiles for the validation period (February 2003).

Comparison of simulated DO and BOD values with previous studies

The simulated DO and BOD profile from the present approach were compared with those obtained from the existing approaches over the study region for two different periods of March–June 2002 and February 2003. The simulations of the present approach gave the closest match to the observed DO and BOD values for both the periods (Figure 7). Table 6 shows various model performance measures from present and past studies for the selected periods.

In the present study, DO and BOD have been jointly considered for determination of model parameters, which was the case of other studies too. Therefore, in order to compare different calibration approaches, the collective model performance in simulating DO and BOD should be taken into consideration. The calibration approach of the present study gave the highest *IOA* values of 0.988 and 0.964 for the periods of March–June 2002 and February 2003 respectively. Similarly, for both the periods, highest R^2 values (0.964 and 0.878) were observed in the present study, followed by Parmar & Keshari (2012), Walling (2014) and O'Connor & Dobbins (1958). Therefore, on an overall scale, the model parameters of the present study simulated DO and BOD better than existing studies for both the analysis periods. Improvements in the performance measures of the present approach, when compared with those from previous studies, may be attributed to the consideration of heterogeneity of river reaches while determining model parameters for the entire river stretch.

DISCUSSION

Application of sequential calibration techniques over the Delhi segment of the Yamuna River yields different model parameters for different reaches of the river. Variation was most prominent in oxygen re-aeration rate followed by BOD oxidation rate and BOD hydrolysis rate; however,

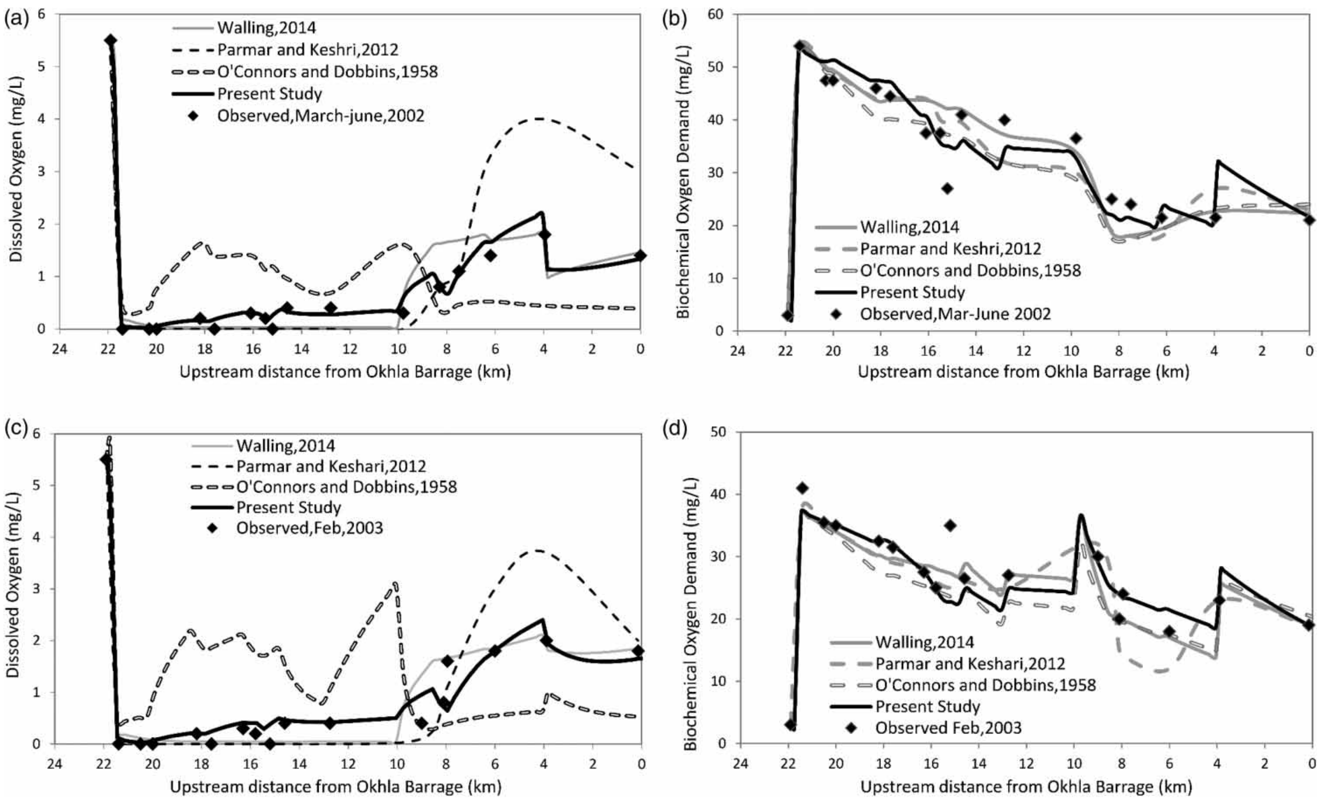

Figure 7 | Comparison of simulated (a) DO and (b) BOD profiles using the present and previous approaches for the period March–June 2002. Comparison of simulated (c) DO and (d) BOD profiles using the present and previous approaches for the period February 2003.

Table 6 | Comparison of different calibration approaches

		Performance measures for different periods			
		March–June 2002		February 2003	
References	Objective water quality parameters	IOA	R^2	IOA	R^2
Walling (2014)	DO	0.999	0.951	0.996	0.932
	BOD	0.842	0.779	0.876	0.733
	DO and BOD	0.921	0.865	0.936	0.833
Parmar & Keshari (2012)	DO	0.976	0.898	0.954	0.897
	BOD	0.843	0.838	0.712	0.849
	DO and BOD	0.910	0.868	0.833	0.873
O'Connor & Dobbins (1958)	DO	0.861	0.487	0.784	0.198
	BOD	0.962	0.850	0.938	0.776
	DO and BOD	0.912	0.669	0.861	0.487
Present study	DO	0.995	0.988	0.972	0.910
	BOD	0.980	0.924	0.955	0.845
	DO and BOD	0.988	0.956	0.964	0.878

studies in the past have ignored these variations. BOD decay (oxidation and hydrolysis) rate depends on the nature of the organic matter added to the river, which in turn depends on the source of organic matter (Bowie *et al.* 1985). In Delhi, drains are the major contributor of organic matter and carry pollutants from different industrial, domestic or agricultural sources. The decay rate of organic matter from these industrial, domestic or agricultural sources would necessarily vary. Therefore, the BOD decay rate should not be conveniently assumed to be constant throughout the Delhi stretch. This is equally applicable in the case of other parameters. The proposed technique offers a local calibration approach in which the river reaches are individually calibrated in sequence. In the past, automated calibration techniques proposed by various studies offered a global method of calibration, i.e., the whole stretch was calibrated together. The proposed technique offers a simpler approach to determine the model parameters for simulating river water quality conditions.

It is important to note here that the number of model parameters selected for calibration of the water quality model does not constrain the utility and applicability of the proposed framework. However, the inclusion of greater numbers of model parameters would not only augment the complexity and execution time of the calibration process, but might also introduce identifiability problems. Therefore, it is suggested to identify the parameters playing a negligible role and exclude them (or give less priority) from the calibration process. The final selection of the best-fitting model parameters largely depends on the user-specified ranges of model parameters. Therefore, it is essential to accurately specify the ranges of model parameters selected for calibration. Inclusion of very broad ranges of model parameters would merely add to the complexity and time-consumption factors, without benefitting the quality of the results.

CONCLUSIONS

A sequential calibration technique was proposed to consider the heterogeneity of river reaches while determining model parameters in water quality modelling studies. This approach takes into consideration the diverse characteristics of polluting sources and reach-specific behaviour by considering different parameter values for different river reaches. The proposed methodology was applied to simulate the water quality characteristics of the Yamuna River passing through Delhi. Three model parameters, namely, oxygen re-aeration rate, BOD hydrolysis rate and BOD oxidation rate were considered in the calibration process. The proposed approach proved superior to the previous calibration approaches over the study region. This may be attributed to the (1) consideration of reach-specific model parameters and (2) the adoption of sequential calibration methodology. These two innovations significantly contributed to the existing calibration framework by providing a simple, effective and systematic alternative of calibration. The effects of other parameters such as nitrification rate, phytoplankton respiration and growth rate, benthic algal growth and respiration rate, zooplankton respiration rate, zooplankton death and excretion rate, phytoplankton death rate and benthic algal death rate were not considered for DO and BOD simulation, because they were found to have a negligible role in the present study region. However, over other regions, these parameters may be incorporated in modelling DO and BOD for obtaining improved water quality estimates. A sensitivity analysis of model parameters can be adopted to choose the major parameters affecting the DO and BOD simulations. The consideration of other model parameters while modelling river water quality can be considered in future works. In addition, there is also the need to address the issue of identifiability or uncertainty of model parameters in the calibration framework.

The present calibration approach is data-intensive and needs water quality data downstream of each river reach, in order to calibrate it specifically. The approach also assumes the data to be error-free for every reach and selected parameters only play a significant role in DO and BOD simulation process over the study stretch. The range of selected model parameters derived in the present study are specific to study region and their upper or lower bound may vary for other regions. In the present study, weighted average optimization function was adopted, in which some of its component objectives fit better than others. Further investigation of trade-off between component objectives can be carried out in future works.

Nevertheless, the proposed approach is generic and can be implemented to calibrate water quality models of any river stretch for any number of model parameters. The approach presented here can also be used for simulating other water quality characteristics as well. The step-wise framework of the proposed approach helps in setting up the structure and considering the heterogeneity of reaches. The proposed approach substantially improves the efficiency of the water quality model by closely replicating the physical system, which in turn aids in the efficient management of water resources and in deriving reliable treatment policies for any river stretch.

ACKNOWLEDGEMENTS

The authors would like to thank the Department of Science and Technology (India) for partly supporting this study through grant no. SB/S3/CEE/045/2014 to Dr Dhanya C. T. The authors sincerely thank the Editor and the anonymous reviewers for reviewing the manuscript and providing insightful comments. The authors also express thanks to the Indian Institute of Technology (IIT), Delhi for supporting this work.

REFERENCES

Afshar, A., Kazemi, H. & Saadatpour, M. 2011 Particle swarm optimization for automatic calibration of large scale water quality model (CE-QUAL-W2): application to Karkheh reservoir, Iran. *Water Resources Management* **25** (1), 2613–2632.

Bowie, G. L., Mills, W. B., Porcella, D. B., Campbell, C. L., Pagenkopf, J. R., Rupp, G. L., Johnson, K. M., Chan, P. W. H., Gherini, S. A. & Chamberlin, C. E. 1985 *Rates, Constants, and Kinetics Formulations in Surface Water Quality Modeling.* US Environmental Protection Agency, EPA/600/3-85/040.

Bowles, D. S. & Grenney, W. J. 1978 Steady state river quality modelling by sequential extended Kalman filters. *Water Resources Research* **14** (1), 84–96.

Brown, L. C. & Barnwell Jr, T. O. 1987 *The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual.* USEPA, Environmental Research Laboratory, Athens, GA, EPA/600/3-87/007.

Chapra, S. C., Pelletier, G. J. & Tao, H. 2008 *QUAL2K: A Modeling Framework for Simulating River and Stream Water Quality, Version 2.11: Documentation and Users Manual.* Civil and Environmental Engineering Department, Tufts University, Medford, MA, USA.

Chiu, Y. C., Chiang, C. W. & Lee, T. Y. 2016 Time series modeling of biochemical oxygen demand at the upstream catchment of Feitsui Reservoir, Taiwan. *Hydrology Research* **47** (5), 1069–1085.

Churchill, M. A., Elmore, H. L. & Buckingham, R. A. 1962 The prediction of stream reaeration rates. *Journal of the Sanitary Engineering Division, ASCE* **88** (4), 1–46.

CPCB 1999–2000 *Water Quality Status of Yamuna River.* ADSORBS/32, Central Pollution Control Board, Delhi, India.

CPCB 2005 *Status of Sewage and Sewage Treatment Plants in Delhi.* CUPS/57, Central Pollution Control Board, Delhi, India.

Delhi Jal Board (DJB) 2005 *Delhi Water Supply and Sewerage Project. Final report – Project preparation study*, Part C: Sewerage.

Ghosh, N. C. & Mcbean, E. A. 1998 Water quality modelling of the Kali River, India. *Water, Air, and Soil Pollution* **102**, 91–103.

Goldberg, D. E. 1989 *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, Boston, MA.

Janssen, P. H. M. & Heuberger, P. S. C. 1995 Calibration of process-oriented models. *Ecological Modelling* **83** (1), 55–66.

Jha, R., Ojha, C. S. P. & Bhatia, K. K. S. 2005 Water quality and flow simulation in River Kali, India. *Journal of the Institution of Engineers (India): Environmental Engineering Division* **85**, 54–61.

Kalburgi, P. B., Shivayogimath, C. B. & Purandara, B. K. 2010 Application of QUAL2K for water quality modeling of River Ghataprabha (India). *Journal of Environmental Sciences and Engineering* **4** (12), 6–11.

Kazmi, A. A. 2000 Water quality modelling of river Yamuna. *Journal of Institute of Engineers, India* **81**, 17–22.

Kumar, P. 2001 Cost of wetland conversion: A case study of floodplain wetland ecosystems along the Yamuna river corridors of Delhi region. In: *Open Meeting of the Global Environmental Change Research Community*, 6–8 October 2001, Rio de Janeiro, Brazil.

Langbein, W. B. & Durum, W. H. 1967 *The Aeration Capacity of Streams.* US Geological Survey, Washington, DC.

Lázár, A. N., Wade, A. J., Whitehead, P. G., Neal, C. & Loewenthal, M. 2012 Reconciling observed and modelled phytoplankton dynamics in a major lowland UK river, the Thames. *Hydrology Research* **43** (5), 576–588.

Lindström, G. 2016 Lake water levels for calibration of the S-HYPE model. *Hydrology Research* **47** (4), 672–682.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J. & Arheimer, B. 2010 Development and testing of the HYPE (Hydrological predictions for the environment) water quality model for different spatial scales. *Hydrology Research* **41** (3–4), 295–319.

Liu, S., Butler, D., Brazier, R., Heathwaite, L. & Khu, S. T. 2007 Using genetic algorithms to calibrate a water quality model. *Science of the Total Environment* **374** (2), 260–272.

Liu, J., Song, X., Wang, Z., Yang, L., Sun, Z. & Wang, W. 2015 Variations of carbon transport in the Yellow River, China. *Hydrology Research* **46** (5), 746–762.

Marsili-Libelli, S. & Giusti, E. 2008 Water quality modelling for small river basins. *Environmental Modelling & Software* **23** (4), 451–463.

Melching, C. S. & Flores, H. E. 1999 Reaeration equations derived from U.S. Geological Survey database. *Journal of Environmental Engineering* **125** (5), 407–414.

Mulligan, A. E. & Brown, L. C. 1998 Genetic algorithms for calibrating water quality models. *Journal of Environmental Engineering* **124** (3), 202–211.

Ng, A. W. M. & Perera, B. J. C. 2003 Selection of genetic algorithm operators for river water quality model calibration. *Engineering Applications of Artificial Intelligence* **165** (6), 529–541.

O'Connor, D. J. & Dobbins, W. E. 1958 Mechanism of reaeration in natural streams. *Transactions of the American Society of Civil Engineers* **123**, 641–684.

Ostfeld, A. & Salomonst, S. 2005 A hybrid genetic-instance based learning algorithm for CE-QUALW2 calibration. *Journal of Hydrology* **310**, 122–142.

Owens, M., Edwards, R. W. & Gibbs, J. W. 1964 Some reaeration studies in streams. *Air and Water Pollution* **8**, 469–486.

Paliwal, R., Sharma, P. & Kansal, A. 2007 Water quality modelling of the River Yamuna (India) using QUAL2E-UNCAS. *Journal of Environmental Management* **83** (2), 131–144.

Parmar, D. L. & Keshari, A. K. 2012 Sensitivity analysis of water quality of the River Yamuna, India. *Environmental Monitoring Assessment* **184**, 1487–1508.

Pelletier, G. J., Chapra, S. C. & Tao, H. 2006 QUAL2Kw: a framework for modelling water quality in streams and rivers using a genetic algorithm for calibration. *Environmental Modelling and Software* **21** (3), 419–425.

Rehana, S. & Mujumdar, P. P. 2011 River water quality response under hypothetical climate change scenarios in Tunga-Bhadra River, India. *Hydrological Processes* **25** (22), 3373–3386.

Sharma, D. & Singh, R. K. 2009 DO-BOD modelling of River Yamuna for national capital territory, India using STREAM II, a 2D water quality model. *Environmental Monitoring Assessment* **159** (1–4), 231–240.

Shen, J. & Kuo, A. Y. 1996 Inverse estimation of parameters for an estuarine eutrophication model. *Journal of Environmental Engineering* **122** (11), 1031–1040.

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** (2), 515–553.

Thomann, R. V. & Mueller, J. A. 1987 *Principles of Surface Water Quality Modeling and Control*. Harper-Collins, New York.

USEPA 1985 *Rates, Constants and Kinetics Formulations in Surface Water Quality*, 2nd edn. EPA 600/3-85-040. US Environmental Protection Agency, Athens, GA, USA.

Van Griensven, A. & Bauwens, W. 2003 Multiobjective autocalibration for semidistributed water quality models. *Water Resources Research* **39** (12), 9.

Vasudevan, M., Nambi, I. M. & Kumar, S. G. 2011 Application of QUAL2K for assessing waste loading scenario in river Yamuna. *International Journal of Advanced Engineering Technology* **2**, 336–344.

Wallace, D. A. & Daque, R. R. 1973 Modelling of land runoff effects on dissolved oxygen. *Journal (Water Pollution Control Federation)* **45**, 1795–1809.

Walling, B. 2014 *Impact of climate change on river water quality: A case study on Yamuna River*. M. Tech Thesis, Department of Civil Engineering, Indian Institute of Technology, Delhi.

Walling, B., Chaudhary, S., Dhanya, C. T. & Kumar, A. 2017 Estimation of environmental flow incorporating water quality and hypothetical climate change scenarios. *Environment Monitoring Assessment* **189**, 225. doi: 10.1007/s10661-017-5942-2.

Wang, A., Tang, L. & Yang, D. 2016 Spatial and temporal variability of nitrogen load from catchment and retention along a river network: a case study in the upper Xin'anjiang catchment of China. *Hydrology Research* **47** (4), 869–887.

Whitehead, P. G., Williams, R. J. & Lewis, D. R. 1997 Quality simulation along river systems (QUASAR): model theory and development. *Science of the Total Environment* **194**, 447–456.

Wood, D. M., Houck, M. H. & Bell, J. M. 1990 Automated calibration and use of a stream quality simulation model. *Journal of Environmental Engineering* **116** (2), 236–249.

Zhang, R., Qian, X., Yuan, X., Ye, R., Xia, B. & Wang, Y. 2012 Simulation of water environmental capacity and pollution load reduction using QUAL2K for water environmental management. *International Journal of Environmental Research and Public Health* **9**, 4504–4521.

First received 26 September 2016; accepted in revised form 6 March 2017. Available online 12 May 2017