Water quality modeling for load reduction under uncertainty: A Bayesian approach

Yong Liua,b, Pingjian Yanga, Cheng Huc, Huaicheng Guoa,*

aCollege of Environmental Sciences and Engineering, Peking University, Beijing 100871, China
bSchool of Natural Resources and Environment, University of Michigan, MI 48109, USA
cLiaoning Institute of Environmental Science, Shenyang 110031, China

A Bayesina approach was applied to river water quality modeling (WQM) for load and parameter estimation. A distributed-source model (DSM) was used as the basic model to support load reduction and effective water quality management in the Hun-Taizi River system, northeastern China. Water quality was surveyed at 18 sites weekly from 1995 to 2004; biological oxygen demand (BOD) and ammonia (NH4+) were selected as WQM variables. The first-order decay rate ($k_i$) and load ($L_i$) of the 16 river segments were estimated using the Bayesian approach. The maximum pollutant loading ($L_m$) of NH4+ and BOD for each river segment was determined based on DSM and the estimated parameters of $k_i$. The results showed that for most river segments, the historical loading was beyond the $L_m$ threshold; thus, reduction for organic matter and nitrogen is necessary to meet water quality goals. Then the effects of inflow pollutant concentration ($C_i/C_0$) and water velocity ($v_i$) on water quality standard compliance were used to demonstrate how the proposed model can be applied to water quality management. The results enable decision makers to decide load reductions and allocations among river segments under different $C_i/C_0$ and $v_i$ scenarios.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

In recent decades, deterioration of natural surface water bodies has become one of the major environmental concerns facing national and local governments worldwide. Thus, comprehensive water quality management is required to mitigate the resulting environmental stresses, effectively manage water environments, and better protect aquatic ecosystems. Water quality modeling (WQM) is considered a required element to support water quality management decisions (Vieira and Lijklema, 1989; Zou et al., 2007). For example, simple mechanistic models combined with data and sophisticated statistical methods were highly recommended by the US National Research Council (2001) to support the ongoing total maximum daily load (TMDL) program. Models play a central role in the TMDL program, not only in determining requirements to achieve water quality standards, but also in calculating the effectiveness of actions to limit pollutant sources for the attainment of a designated use (National Research Council, 2001).

Many models, typically classified as watershed or water-body models, have been developed (Chapra, 1997). Most previous studies have involved forward models consisting of two major steps: (1) estimating pollutant loads from a

*Corresponding author. Tel./fax: +86 10 62751921.
Abbreviations: ANOVA, Analysis of variance; CA, Cluster analysis; WQM, Water quality modeling; DSM, Distributed-source model; MCMC, Markov chain Monte Carlo; EPB, Environmental protection bureau.
E-mail addresses: lyong@umich.edu (Y. Liu), hcguo@pku.edu.cn (H. Guo).
0043-1354/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2008.04.007
watershed model and (2) predicting the pollutant concentrations and water quality responses using a waterbody model. Three key issues need to be addressed before WQM can be used for practical management: (1) estimation of pollutant load; (2) estimation of model parameters; and (3) assessment of the uncertainties in model development and application (National Research Council, 2001). In past decades, there have been many attempts at developing the pollutant loading models, from simple export coefficient models and regression models to complex mechanistic models (Ghenu et al., 2008; Shrestha et al., 2008); however, load estimation remains a challenge to molder, especially because of insufficient data. At the same time, the past years have seen rapid development of highly accurate water quality monitoring technologies and improved data sharing. Thus, it would be interesting to determine a robust method for load estimation using the plentiful monitoring data and verified models that meet selection criteria (National Research Council, 2001). Recent developments in inverse methods make such an idea a reality (Wan and Vallino, 2005; Shen et al., 2006; Zou et al., 2007). Inverse methods have been widely applied in practical environmental cases, such as model parameter estimation (Shen and Kuo, 1998; Zou et al., 2007), nonpoint sources estimation (Shen et al., 2006), groundwater (Michalak and Kitanidis, 2004; Vermeulen et al., 2005; Snehalatha et al., 2006), and wastewater systems (Bumgarner and McCray, 2007). Shen et al. (2006) formulated the estimation of nonpoint sources of fecal coliforms as an inverse parameter estimation problem, using observed data from the Wye River (USA) to estimate the allowable load to attain water quality standards.

Uncertainty is another issue that requires addressing in model development and application, such as the dynamic interactions existing between pollutant loadings and receiving waters, and the uncertainties in parameters estimation and model outputs (National Research Council, 2001; Qin et al., 2007). A practical water quality management strategy must be based on the effective involvement of (1) the uncertainties in model structure and parameters, (2) the uncertainties in model output, and (3) the risk of water quality violating targeted standards and adaptive management under uncertainty (Liu et al., 2007). Bayesian statistics has been increasingly applied in WQM for its capability of handling uncertainty and incorporating prior information, which allows the use of previous WQM experiences and provides the probabilistic assessment necessary for supporting decision making. Compared with traditional modeling approaches, the estimated model parameters in the Bayesian approach are presented in terms of a posterior joint distribution; thus, it can better predict the frequency of water quality standard violations and provide more information on uncertainty (Qian et al., 2005). The resulting Bayesian credible intervals can then be directly conveyed into the probabilistic risk of decisions in water quality management (Malve and Qian, 2006).

Inverse methods can deal with the bottleneck of traditional methods of load estimation (Shen et al., 2006); and Bayesian models can involve uncertainty and parameter estimation in the modeling process. Thus, the integration of inverse methods and Bayesian modeling should be effective at answering the three WQM questions mentioned above, and has been applied to ground water (Woodbury and Ulrych, 2000), contaminant source identification (Michalak and Kitanidis, 2003) and shellfish aquaculture ecosystem modeling (Dowd and Meyer, 2003). However, this combination has seldom been used in water quality management. Therefore, the objective of this study was to develop a Bayesian approach to river WQM combined with inverse methods to support the practical adaptive water quality management under uncertainty. The heavily polluted Hun-Taizi River system in northeastern China was used as a case study. The model results will help local decision makers reduce and allocate loads in an effective and efficient manner.

2. Materials and methods

2.1. Study area and data source

The Hun-Taizi River system consists of three rivers (Fig. 1) in central Liaoning Province, northeastern China. The Hun River and Taizi River meet at Daliao River and then flow to Bohai Bay, which is under serious pollution and heavy eutrophication from upland nutrient loading, mainly inorganic nitrogen and organic matter (State Environmental Protection Agency [SEPA], 2007). The Hun River is 415 km long, with a basin area of 11,500 km². The Taizi River is 413 km long, with a basin area of 13,900 km². The Daliao River is 94 km long, with a basin area of 1926 km². The Hun-Taizi River system is heavily polluted by point and nonpoint sources. Most of the river segments cannot meet targeted water quality goals set by the local environmental protection bureau (EPB). The Hun-Taizi River is one of the three severely polluted river systems under Chinese Central Government mandate to develop effective water quality management plans and achieve significant water quality improvement by 2020. To protect the regional environment and aquatic ecosystem and to control eutrophication in Bohai Bay, it is essential to control pollutant loads in an effective manner. First, a practical and robust water quality model should be developed. Specifically, in the Hun-Taizi River system, load estimation is important for each river segment, because the segments are under the administration of different counties. Second, model results and the embedded uncertainties and risks should be provided directly to the decision makers.

The EPB of Liaoning Province is responsible for water quality monitoring and management. Data were collected from 18 regular monitoring sites (Fig. 1) for 10 years from 1995 to 2004; monitored parameters included water velocity and 24 required variables mentioned in the Environment Quality Standards for Surface Water Quality (No. GB3838-2002), issued by SEPA of China. Water samples were collected on a weekly basis. Water quality parameters were sampled, preserved, delivered, and analyzed using the standard methods of [the American Public Health Association] APHA (1998) and the standards of GB3838-2002. All data were analyzed; ammonia (NH₃) and biological oxygen demand (BOD) were preferred as prior management targets by the local government, because they are viewed as the primary driving factors for the eutrophication of Bohai Bay (SEPA, 2007). Thus, two separate
water quality models for NH$_4^+$ and BOD were established in this study.

2.2. Distributed-source model (DSM)

WQM has been used frequently in past years to support river basin environmental management. The DSM was used here to support effective water quality management of the Hun-Taizi River system, considering a main loading contribution from nonpoint sources. The basic assumption for DSM is distributed sources, which means the source enters a river in a diffuse and uniform manner (Chapra, 1997). Sixteen river segments were used for the Hun-Taizi River system, each segment between two adjacent sampling sites (Fig. 1).

The DSM can be written as (Chapra, 1997)

$$\frac{dC}{dx} = -kC + L$$

(1)
where \( v \) is the average water velocity in the river segment (m d\(^{-1}\)), \( C \) the concentration (mg l\(^{-1}\)), \( x \) the distance (m), \( k \) the first-order decay rate (d\(^{-1}\)), and \( L \) the diffuse source loading (g m\(^{-1}\)d\(^{-1}\)). Then the concentration \( C \) at any distance \( x \) can be expressed as

\[
C = f(k, v, x, C_{\text{inflow}}, L)
\]

where \( C_{\text{inflow}} \) is the inflowing pollutant concentration. Fig. 2 shows a conceptual description of river segmentation that is related to Eq. (2). Specifically, for the \( i \)th river segment \( S_i \), the out-flowing concentration \( C_i \) should be

\[
C_i = f(k, v, x, C_{i-1}, L_i) = C_{i-1}e^{-k(v)/x} + \frac{L_i}{k} (1 - e^{-k(v)/x})
\]

The model in Eq. (3) can be viewed as a forward model, from \( L_i \) to \( C_i \). The reverse model for Eq. (3) is

\[
k, L_i = f^{-1}(v, x, C_{i-1})
\]

Thus, the problem of load and decay rate estimation can be transformed into an inverse model with the aim of finding a set of \( k \) and \( L_i \) jointly so that a defined objective function can be met. The Bayesian approach was then applied for \( k \) and \( L_i \) estimation.

### 2.3 Bayesian approach

Bayesian statistics has been increasingly used since the 1990s (Qian et al., 2003). In the case of Bayesian statistics, all unknown parameters are treated as random variables and their distributions are derived from known information (Borsuk et al., 2001). Thus, Bayesian statistics provides a rigorous method for uncertainty analysis and presents key information for management decision making (Reckhow, 1994). Bayesian inference is based on Bayes theorem as (Gill, 2002)

\[
p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} = \frac{\int \rho(\theta)p(y|\theta)d\theta}{p(y)}
\]

where \( p(y|\theta) \) is the posterior probability of \( \theta \), which is the conditional distribution of the parameters after observed data, such as the observed concentration \( C_i \) and \( C_{i-1} \) in Eq. (3), \( \theta \) the parameter needing estimated, including \( L \) and \( k \) in Eq. (3), \( p(\theta) \) the prior probability of \( \theta \), and \( p(y|\theta) \) the likelihood function, which incorporates the statistical relationships as well as the mechanistic relationships among the predictor and variables. Usually, it is impossible to obtain the analytically summarizing posterior distributions, which limits the practical implementation of the Bayesian approach. In recent decades, however, the Monte Carlo method, in particular the Markov chain Monte Carlo (MCMC) algorithm, has been applied to obtain the numerical summarization of parameters (Qian et al., 2003; Qian and Shen, 2007).

There are three major steps in the Bayesian model using the MCMC sampling method (Malve and Qian, 2006): (a) Formulation of prior probability distributions for targeted parameters; (b) specification of the likelihood function; and (c) MCMC sampling for the posterior probability distributions, in which Gibbs sampling and Metropolis–Hastings algorithms are popularly used. For the MCMC algorithm, the modeler needs to provide starting values for the model burn-in period.

### Table 1 – The one-way ANOVA results for monitoring data of water velocity, BOD, and \( \text{NH}_4^+ \)

<table>
<thead>
<tr>
<th></th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Water velocity at different segments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>307083.90</td>
<td>18063.76</td>
<td>8.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>346137.87</td>
<td>2136.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>653221.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Water velocity at different year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>43519.64</td>
<td>4835.52</td>
<td>1.35</td>
<td>0.216</td>
</tr>
<tr>
<td>Within groups</td>
<td>609702.13</td>
<td>3586.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>653221.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) BOD at different year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>3307.08</td>
<td>367.45</td>
<td>1.09</td>
<td>0.372</td>
</tr>
<tr>
<td>Within groups</td>
<td>57230.03</td>
<td>336.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>60537.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) BOD at different segments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>41259.97</td>
<td>2427.06</td>
<td>20.40</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>19277.14</td>
<td>119.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>60537.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) ( \text{NH}_4^+ ) at different segments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>4042.78</td>
<td>237.81</td>
<td>24.20</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Within groups</td>
<td>1591.75</td>
<td>9.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5634.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) ( \text{NH}_4^+ ) at different year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>229.12</td>
<td>25.457</td>
<td>0.801</td>
<td>0.616</td>
</tr>
<tr>
<td>Within Groups</td>
<td>5405.41</td>
<td>31.797</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5634.52</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Convergence can be reached after a sufficient burn-in period. A random sample from the posterior distribution can then be provided and used for further estimation. The practical implementation of MCMC, using Gibbs sampling in this study, is based on a specialized and free software package, WinBUGS (Lunn et al., 2000); the WinBUGS code used in this study can be found in the Supporting Materials.

To incorporate DSM into Eq. (5), a normally distributed error term, \( \epsilon \), was added to the model, with zero mean and variance of \( \sigma^2 \) (Eq. (6)). The value of \( \sigma \) can be estimated in the Bayesian model based on the prior distribution:

\[
C_i = C_{i-1}e^{-k_i} + \frac{L_i}{k_i}(1 - e^{-k_i}) + \epsilon
\]

(6)

The likelihood function for all 10 years \( (j = 1, \ldots, 10) \); where 1995 \( j = 1 \) and 16 river segments of the Hun-Taizi River system in Fig. 1 \( (i = 1, \ldots, 16) \) is

\[
\prod_{i=1}^{16} \prod_{j=1}^{20} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( \frac{(C_{ij} - C_{i-1}e^{-k_i}) - \frac{L_i}{k_i}(1 - e^{-k_i}) - \epsilon}{2\sigma^2} \right)
\]

(7)

where \( C_{ij} \) and \( C_{i-1} \) is the observed value.

Fig. 3 – The model fitting results for (b) BOD and (a) NH₄⁻, between observed data and the modeled mean, median and 2.5% and 97.5% confidence level values.

Estimation of parameter \( k \), the first-order decay rate in Eq. (1), is a key step for WQM. There are two alternatives on \( k \) for the Hun-Taizi River system: (1) \( k_i \) assuming \( k \) is spatially variable for each segment, but similar for each year from a common prior distribution and (2) \( k_j \), assuming \( k \) is temporally variable for each period, but similar for each segment from a common prior distribution. Thus, there are two alternative models for estimating \( k \) in Eqs. (8) and (9) in which subscripts \( i \) and \( j \) refer to river segment and time, respectively:

\[
C_{ij} = C_{i-1}e^{-k_j} + \frac{L_j}{k_j}(1 - e^{-k_j}) \quad \forall i = 1, 2, \ldots, 16 \quad (8)
\]

or

\[
C_{ij} = C_{i-1}e^{-k_i} + \frac{L_i}{k_i}(1 - e^{-k_i}) \quad \forall j = 1, 2, \ldots, 10 \quad (9)
\]

Two derived parameters were applied in this study to test the above alternatives in Eqs. (8) and (9): (1) the deviance information criterion (DIC), a Bayesian measure of how well
the model fits the data, where a larger DIC indicates a worse fit (Malve and Qian, 2006) and (2) the coefficient of determination, $R^2$, between the original data and predicted values. For a Bayesian model with data $y$, unknown parameters $\theta$ and the likelihood function $p(y|\theta)$, the deviance is defined as (Gelman et al., 2004)

$$D(\theta) = -2 \log(p(y|\theta)) + c$$

where $c$ is a constant. The expectation of $D(\theta)$ is

$$\mathbb{E}[D(\theta)]$$

The effective number of parameters in the model is

$$pD = \mathbb{E}[D(\theta)]$$

where $\theta$ is the expectation of $y$. DIC is calculated as

$$DIC = pD + \mathbb{E}[D(\theta)]$$

$R^2$ is defined as

$$R^2 = 1 - \frac{SS_E}{SS_T} = 1 - \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

where $SS_E$ and $SS_T$ are the sum of squared errors and total sum of squares, respectively; $y_i$ and $\bar{y}$ are the original data values and modeled values, respectively; $\bar{y}$ is the mean of the observations $y$.

Posterior simulations, based on the Bayesian model results, were performed to better understand the load reduction for each river segment required to attain the targeted water quality goal. The maximum pollutant loading ($L_m$) was calculated as follows:

$$L_{m,i} = \frac{k_i(C_i - C_{g,i-1}e^{-k_i/v_{a,i}t_i})}{1 - e^{-k_i/v_{a,i}t_i}}$$

where $L_{m,i}$ is the maximum pollutant loading for river segment $i$ (g m$^{-3}$ d$^{-1}$); $C_{g,i}$ and $C_{g,i-1}$ are the targeted water quality goals for sites $i$ and $i-1$ (mg l$^{-1}$); respectively; $v_{a,i}$ is the historical annual average water velocity (m d$^{-1}$); and $k_i$ is the mean value from the Bayesian model results.

From Eq. (3), the water quality of segment $S_i (i=1, 2, ..., 16)$ is affected by two important factors, the inflow water velocity $v_i$ and inflow pollutant concentration $C_{g,i}$. Thus for effective water quality management and to attain water quality goals, the effects of $v_i$ and $C_{g,i}$ should be addressed. The effects of various values of $v_i$ and $C_{g,i}$, usually expressed as intervals at several confidence levels, on the loads can be plotted. Decision makers are thereby informed how to reduce loads under different $v_i$ and $C_{g,i}$ scenarios.

3. Results and discussion

3.1. Statistical analysis

Basic statistical analysis was conducted to the raw data before the application of the Bayesian approach (Table S1, Fig. S1 and Fig. S2 in Supporting Materials), including statistical description, hierarchical cluster analysis (CA), and correlation analysis. It was aimed for primary verification of model structure in Eq. (1) and the two alternatives for $k$ in Eqs. (8) and (9). One-way analysis of variance (ANOVA) was then conducted to examine differences in water velocity, BOD concentration, and NH$_4^+$ concentration among the various years or sampling sites (Bordalo et al., 2001). The results can be helpful in determining the two alternative models for $k$.

Given a confidence level of $p = 0.001$, the variance analysis in Table 1 indicated significant differences in water velocity, BOD, and NH$_4^+$ concentration among the various years or sampling sites (Bordalo et al., 2001). The results can be helpful in determining the two alternative models for $k$. Given a confidence level of $p = 0.001$, the variance analysis in Table 1 indicated significant differences in water velocity, BOD, and NH$_4^+$ at various sites, but no significant difference among the various years, which fits the CA results of Fig. S1 in Supporting Materials.
3.2. Estimation of $k_i$ and $L$ in DSM

The MCMC simulations for the BOD and NH$_4^+$ models were carried out separately in WinBUGS. All river segments were modeled in the same WinBUGS program. The determination of the prior distribution of load $L_i$ was based on historical local statistics; and $k$ was determined using interval values from the literature at [0.01, 0.70] d$^{-1}$ for NH$_4^+$ and [0.05, 0.60] d$^{-1}$ for BOD (Chapra, 1997; Alexander et al., 2000). The MCMC was carried out in WinBUGS with four chains, each with 10,000 iterations (first 5000 discarded after model convergence); and 1000 samples for each unknown quantity were taken from the next 5000 iterations to reduce autocorrelations of the data after Malve and Qian (2006). The two estimation alternatives for $k$ were then compared based on the Bayesian approach, using DIC and $R^2$ (Table S2 in Supporting Materials). The results indicated that $k_i$ is a more precise fit for NH$_4^+$ and BOD than $k_j$, which means that the first-order decay rate ($k$) differs significantly among various river segments. This is a direct result of the different physical and chemical properties of each river segment, including oxygen concentrations, organic content of benthic sediments, river depth, and water velocity (Alexander et al., 2000). This interpretation also corresponds well with the ANOVA results in Table 1. Thus, the model results below were based on $k_i$ in Eq. (8).

The fit between the observed data and modeled values is demonstrated in Fig. 3, with mean and median values and two confidence levels at 2.5% and 97.5%. The model results generally fit the observed data, with the exception of some extremely high BOD values. The results are acceptable, considering the complexities of water quality changes in rivers and the weakness of DSM; thus, the model can be applied to practical water quality management. However, despite the good model fit for NH$_4^+$, we have likely over-simplified the sources of NH$_4^+$ in this study. Usually, NH$_4^+$ can originate from point and nonpoint sources, as well as ammonification and river sediment release (Alexander et al., 2000). Thus, the estimated load $L_i$ may not solely represent nonpoint sources and may therefore overestimate the non-point source load.

The posterior distributions of $k_i$ in the NH$_4^+$ and BOD models are described in Fig. 4, including the mean value and 50% and 95% posterior credible intervals. The Monte Carlo error (MC error) and sample standard deviation (S.D.), along with the confidence level at 2.5% and 97.5%, for posterior distributions of $L_i$ for all segments in different years are summarized in Table S3 in the Supporting Materials. The above model results show that the parameters $k_i$ and $L_i$ have small MC errors, indicating that the model converged well (Huang and McBean, 2007).

3.3. Water quality management under uncertainty

The main focus of water quality management is to control pollutant loading into river segments to ensure that water quality meets the targeted goals. The water quality goals for each segment and sampling site are from the EPB of Liaoning Province, China (Table S4 in the Supporting Materials). The maximum pollutant loading ($L_{max}$) for NH$_4^+$ and BOD of each segment, based on Eq. (15), are shown in Fig. 5 and Table 2.

<table>
<thead>
<tr>
<th>Segment</th>
<th>$L_{max}^{NH_4^+}$</th>
<th>$L_{max}^{BOD}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1.13</td>
<td>1.85</td>
</tr>
<tr>
<td>S2</td>
<td>1.08</td>
<td>1.88</td>
</tr>
<tr>
<td>S3</td>
<td>1.06</td>
<td>1.86</td>
</tr>
<tr>
<td>S4</td>
<td>1.04</td>
<td>1.89</td>
</tr>
<tr>
<td>S5</td>
<td>1.00</td>
<td>1.93</td>
</tr>
<tr>
<td>S6</td>
<td>0.98</td>
<td>1.93</td>
</tr>
<tr>
<td>S7</td>
<td>0.96</td>
<td>1.93</td>
</tr>
<tr>
<td>S8</td>
<td>0.94</td>
<td>1.93</td>
</tr>
<tr>
<td>S9</td>
<td>0.92</td>
<td>1.93</td>
</tr>
<tr>
<td>S10</td>
<td>0.90</td>
<td>1.93</td>
</tr>
<tr>
<td>S11</td>
<td>0.88</td>
<td>1.93</td>
</tr>
<tr>
<td>S12</td>
<td>0.86</td>
<td>1.93</td>
</tr>
<tr>
<td>S13</td>
<td>0.84</td>
<td>1.93</td>
</tr>
<tr>
<td>S14</td>
<td>0.82</td>
<td>1.93</td>
</tr>
<tr>
<td>S15</td>
<td>0.80</td>
<td>1.93</td>
</tr>
<tr>
<td>S16</td>
<td>0.78</td>
<td>1.93</td>
</tr>
</tbody>
</table>

The comparison with historical $L$ ranges with the estimated maximum pollutant loading of BOD under four river velocity scenarios, including mean value, median value, and values at 2.5% and 97.5% confidence levels are shown in Table 2.
respectively, assuming the upper segment meets its set goal. The aim of the results is to inform the decision makers how to control the pollutant loading for targeted water quality goals under different hydrological conditions. The historical loading estimated from the Bayesian model was compared with the $L_m$ values under three river velocity scenarios (mean, 2.5%, and 97.5% confidence level). For most river segments, the historical loading is above the $L_m$ threshold. Thus, pollutant reduction for organic matter and nitrogen is necessary to meet the water quality goals. The reduction amount required for each segment can be seen in Fig. 5 and Table 2.

The effects of inflowing water velocity $v_i$ and inflowing pollutant concentration $C_i$ on $C_i$ were measured. Here, segment 15 (S15) and sampling site DL2 were demonstrated as an example for detailed water quality management under uncertainty, based on the model results. The graphs in Fig. 6 show, for S15, how to control $L_i (i = 15)$ to reach targeted water quality goals under different scenarios of inflowing pollutant concentration ($C_{i-1}$) and water velocity ($v$). Fig. 6(a) and (c) demonstrates how $L$ changes with $C_{i-1}$ under four $v_i$ scenarios. The $v_i$ changes were determined from historical data and include the mean, median, and 2.5% and 97.5% confidence levels. Thus, decision makers can be informed from the graph that at a certain value of $v_i$, when the inflowing pollutant concentration $C_i/C_0$ equals some value of $C_i$, the maximum load can be discharged to the river without violating the water quality goal. The maximum load, the threshold of $L_i$, can be read in Fig. 6a and c as the values A–D for the four $v_i$ scenarios. The graph can also be used to decide the maximum inflow pollutant concentration $C_i/C_0$ that is the minimum water quality requirement for the upper streams when $L_i = 0$ (E–H in Fig. 6a and c). The graph is also useful for load allocation when two river segments are not in the same administrative zone. Two similar graphs, Fig. 6(b) and (d) shows the effects of water velocity ($v$) on $L$ under different inflowing pollutant concentrations ($C_{i-1}$). Here, three $C_{i-1}$ scenarios were examined: (1) $C_{i-1}$ equals the targeted water quality goal of $C_i$; (2) $C_{i-1}$ exceeds the targeted water quality goal of $C_i$; and (3) $C_{i-1}$ is less than the targeted water quality goal of $C_i$.
Two models were used to estimate first-order decay rate management. To the stakeholders' preferences, to improve water quality. Additional graphs can be used in a similar fashion, according to the stakeholders' preferences, to improve water quality management.

4. Conclusions

(1) A Bayesian approach was applied to improve the water quality management of the Hun-Taizi River system, China. The model was fitted to water quality monitoring data from 1995 to 2004. The proposed model can integrate the three key issues in WQM, including pollutant load estimation, parameter estimation, and uncertainties in model development and application. It should be a useful tool for decision making in water quality management under uncertainty.

(2) The hierarchical CA and ANOVA results indicate significant differences in water velocity, BOD, and NH$_4^-$ among various sites; there was no difference among years.

(3) Two models were used to estimate first-order decay rate ($k$). The results indicated that $k$ and load rate ($L$) have very small MC errors and the model converged well. The comparison showed that for most of the studied river segments, the historical loading exceeded the maximum pollutant loading. Thus, reduction of organic matter and nitrogen input is necessary to meet water quality goals.

(4) The basic water quality model applied in this study is relatively simple with an assumption of distributed sources. The model result is based on annual average data of a 10-year dataset. Thus, despite the model's capacity to handle uncertainties in WQM, the model can be improved by incorporating a longer-period dataset for better interpretation of water quality changes and seasonal variations. Pollutant loading could then be estimated on a seasonal scale or an even shorter period.

Acknowledgments

The author would like to appreciate Dr. Song S. Qian, the editors and the anonymous reviewers for their helpful comments on the paper. This paper was supported by the "National Basic Research (973) Program" Project (No. 2005CB724205) of the Ministry of Science and Technology of China.

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.watres.2008.04.007.

References


