Desalination and Water Treatment 2 (2009) 19–23
www.deswater.com

Estimation of margin of safety for Korean TMDL development

Dongil Seo[a*], Eunhyoung Lee[b], Chulhee Lee[c], Ken Reckhow[d]

[a]*Department of Environmental Engineering, Chungnam National University, 220 Jung-dong, Yuseong-gu, Daejeon, Korea*
Tel. +82 42 821 6679; Fax: +82 42 822 6510; email: seodi@cnu.ac.kr
[b]*M-Cubic, Inc, 207 Migun Tech World, Yongsan-dong, Yuseong-gu, Daejeon 305-500, Korea*
[c]*Department of Environmental Engineering, Yeungnam University, 214-2 Daedong, Kyeongsan, Gyeongsangbukdo 712-749, Korea*
[d]*Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708, USA*

Received 30 July 2007; Accepted 14 September 2007

ABSTRACT

The Clean Water Act in the US requires that total daily maximum loads) (TMDLs) shall be established with an assessment of the impact of seasonal variations and of the uncertainty in quantification methods to account for lack of knowledge concerning the relationship between effluent limitations and water quality. Due to uncertainties in model structure and/or model parameters in addition to natural and seasonal variability of forcing functions, estimation of waste loads in a TMDL analysis requires a margin of safety (MOS). The margin of safety is typically incorporated implicitly through conservative assumptions used to develop the TMDL or added explicitly as a separate component of the TMDL. For successful TMDL development, accurate and appropriate predictions of waste loads from a watershed and the variation of water quality in a water body are essential. This paper summarizes the estimation process of the MOS in the Nakdong River, Korea considering variability in observed data and the uncertainty in the model. Approved TMDL reports in eight northeastern states in US were used for guidance.

Keywords: Clean Water Act; TMDL; Waste load allocation; MOS; Margin of safety; Nakdong River

1. Introduction

A total maximum daily load (TMDL) refers to the maximum pollutant load a receiving water body can assimilate and still attain water quality standards. The Clean Water Act in the US requires that TMDLs shall be established in consideration of seasonal variations and the uncertainty in modeling methods to account for lack of knowledge concerning the relationship between effluent limitations and water quality [1]. The standard equation defining a TMDL is as follows:

$$TMDL = WLA + LA + MOS \qquad (1)$$

where *WLA* is the total allowable load allocated to point sources, *LA* the total allowable load allocated to nonpoint sources, and *MOS* is the margin of safety.

The margin of safety can be incorporated either implicitly as conservative assumptions used to develop the TMDL or added explicitly as a separate component of the TMDL. In TMDL development, there can be many different sources of uncertainty; accordingly, Canale and Seo [2] suggested using a probabilistic approach in the analysis of water quality modeling results. Zhang [3] identified four types of uncertainty for water quality and watershed modeling: (1) natural randomness; (2) data; (3) model parameters and (4) model structure. Later, Zhang and Yu [4] observed the significant sources of the uncertainties associated with TMDL estimation, such as

*Corresponding author.

Presented at IWA Efficient 2007, May 20–23, 2007, Jeju, Korea.

linking pollutant load to water quality attainment and estimating loads from storm driven, time varying non-point sources. Dilks and Freedman [5] reported that among 172 TMDLS he surveyed, 120 specified an explicit margin of safety, 40 used an implicit margin of safety, and the remaining 12 did not use a margin of safety at all. Of the 120 TMDLS that used an explicit margin of safety, only one TMDL actually calculated the uncertainty in the analysis for the margin of safety. None of the 40 implicit TMDLs provided any indication of degree of protection provided by the MOS. Dilks and Freedman [5] identified reasons for lack of uncertainty analysis as follows:

1. limited practical experience in defining the uncertainty in the TMDL calculations
2. absence of information regarding the degree of protection provided by the margin of safety; and
3. data-poor/high-uncertainty situations that result in MOS values so large as to make implementation impractical.

In a study of the scientific basis of the TMDL program, the National Research Council [6] recommended that uncertainty should be explicitly acknowledged both in the models selected to develop TMDLs and the results generated by those models. The NRC committee further recommended that EPA should end the practice of arbitrary selection of the MOS and instead require uncertainty analysis as the basis for MOS determination. Chapra [7] wrote that a water quality model provides a means to predict water quality as a function of loads and system modifications. He observed that models are not perfect and these imperfections are best represented using probabilistic approaches. He also noted that because of the heavy data requirements for a proper analysis and the time constraints for TMDL development, it is not realistic to require complete uncertainty analysis. As an alternative, Reckhow [8] suggested incorporating more practical but incomplete uncertainty analyses. In this paper, the margin of safety was estimated for the Nakdong River basin, Korea considering data variability and model uncertainty.

2. Study Site — Nakdong River Basin

The Nakdong River is 521.5 km long and is the second largest river in South Korea. Its watershed area is 23,817 km^2, which is equivalent to 1/4 of the country. The Korean government divides this basin area into 41 sub-basins for water quality management purposes.

3. Methods of quantification of margin of safety

The three most common methods for uncertainty analysis are sensitivity analysis, first-order analysis, and Monte Carlo simulation. In general, sensitivity analysis is applied to one variable at a time. If more than two variables are simultaneously evaluated, the relative relationship of those variables can be evaluated (perhaps using generalized sensitivity analysis). However, as the number of variables increases, the level of complexity of the analysis increases.

3.1. First-order error analysis

First-order analysis is based on the approximation of the behavior of the dependent variable with moments of the independent variables or parameters. The first-order error analysis allows the estimation of the overall variance/standard deviation of the model output by using the coefficient of variation of all variables of interest and their sensitivity coefficients. The calculated standard deviation or error bound could be viewed as a surrogate measure of the Margin of Safety in the TMDL equation. By estimating the error bound in the predicted model output, instead of a single value for mean response, the modeler is better able to interpret the model output and provide more useful information for a decision maker. Further, the mean and standard deviation of the model output can be used to parameterize probability distributions that may be used to approximate the cumulative distribution function (CDF) and the probability density function (PDF) of the model output so that the full probability distribution may be used in risk-based design and planning. Reckhow and Chapra [9] characterized the error in a model prediction as a function of individual variable error, a sensitivity factor expressing the "importance" of each variable, and the correlation between variables. First-order error analysis has been successfully applied to the Streeter–Phelps DO model, a seasonal food chain eutrophication model, QUAL2E [10] and a one-dimensional network model [4]. First-order error analysis was also applied to estimate the margin of safety by Zhang [3].

3.2. Monte Carlo simulation

Monte Carlo simulation is an alternative and conceptually simple approach to conduct a prediction error analysis [9]. Under this technique, probability density functions are assigned to each variable or parameter reflecting the uncertainty in that term. Then, with "synthetic sampling," values are randomly chosen from the distribution for each term. These values are inserted into the model, and a prediction is calculated. After this is repeated a large number of times, a distribution of predicted values results, which reflects the combined uncertainties. Parameter distributions are defined from existing data or expert judgment for each uncertain characteristic. QUAL2E-UNCAS [10] includes the Monte Carlo Simulation.

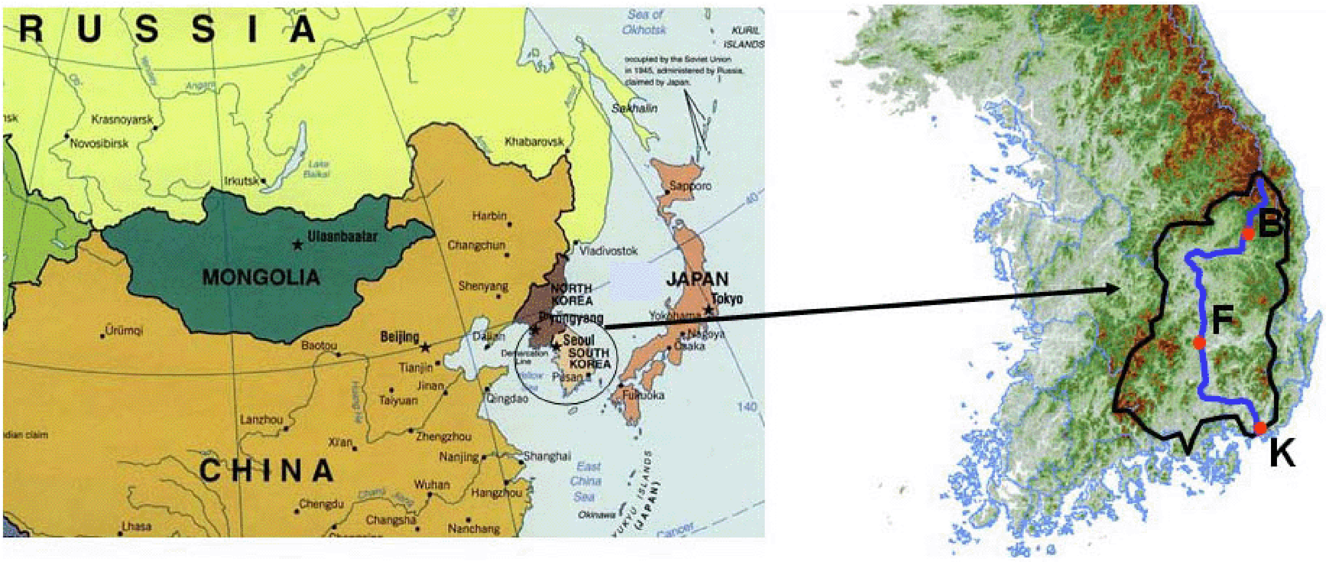

Fig. 1. Nakdong River Basin and locations of study sites (B, F and K stand for the end-points of the upper, middle and lower basin areas, respectively).

3.3. Walker's method

Walker's method can be used to consider the variability of measured data and model uncertainties by assuming the errors of the data follow a lognormal distribution. If a variable y can be expressed as $y = \exp(x)$ and if x has a normal distribution, $N(\mu,\sigma^2)$, then y is said to have a lognormal distribution $LN(\mu,\sigma^2)$. In general, errors in water quality concentration data and waste loads tend to follow a lognormal distribution of $LN(0,\sigma^2)$. Based on this, Walker [11] developed a method of estimation of the margin of safety for lakes in New England using the expressions given below. He also developed a spreadsheet program to support the calculations. Walker [12] suggested dividing the MOS into a variability component and an uncertainty component as follows:

$$MOS = MOV + MOU \tag{2}$$

where MOV is the margin of variability (kg/year) and MOU the margin of uncertainty (kg/year). He suggested using the following approach by considering compliance rate (β) with a confidence interval (α):

$$LA = TMDL\ Fv\ Fu \tag{3}$$

$$Fv = \exp(-Z_\beta\, Sv) \tag{4}$$

$$Fu = \exp(-Z_\alpha\, Su) \tag{5}$$

where LA is the allocated long-term-average load ($TMDL - MOS$), Fv the factor accounting for year-to-year variability in a dependent variable (or a target concentration), Sv the year-to-year coefficient of variation (CV) of a dependent variable (or a target concentration), Fu the factor accounting for uncertainty in the predicted variable (or a target concentration), Su the model error coefficient of variation (CV) for the predicted dependent variable (or a target concentration), Z_β the standard normal variate with upper tail probabilityβ, Z_α the standard normal variate with upper tail probability α; β is the assumed compliance rate and α is the assumed confidence level. By combining the above equations, the MOS, MOU and MOV can be explicitly estimated as follows:

$$LA = TMDL - MOS = TMDL\ Fv\ Fu \tag{6}$$

$$MOS = TMDL\ (1 - Fv\ Fu) \tag{7}$$

$$MOU = MOS\ (1 - Fu)/(2 - Fu - Fv) \tag{8}$$

$$MOV = MOS - MOU \tag{9}$$

4. Results and discussion

Table 1 summarizes the MOS calculation results for various water quality constituents in 295 approved TMDL reports during 1999–2005 in eight states in the eastern part of the US including MA, CT, NY, PA, MD, VA, NC, and OH. For explicit margins of safety, the MOS percentage of the TMDL varied from 3, 5, 10, 13, 15 and 17%. In these explicit MOS cases, 79 used 5% for the MOS and 86 used 10% (out of 295 reports). The MOS was less than 10% in 179 out of the 188 explicit MOS cases for various water quality parameters (in the 259 TMDL reports). Some exceptions were: (1) La Trappe Creek, MA, allocated an MOS of 25% of the difference between weekly discharge

Table 1
Analysis of MOS in TMDL reports in eight states in the US

	Implicit MOS only	Explicit MOS only	Explicit margin of safety magnitude				
			3%	5%	10%	13, 15, 17%	Others
Alkalinity	1						
Bacteria	3	6		6			
C_{BOD}	1	7	1	4			2
Chlordane	3						
Chlorine	1						
Cu, Mn, Pb, Zn	8	2		1	1		
Dissolved oxygen	4	4		3	1		
E. Coli (1)[a]	33	8			8		
Fecal coliform (2)[a]	9	53		28	18	7	
Mercury	8	2			1		1
N_{BOD}		2		1			1
Nitrogen	11	20	9	15	1		3
Organic solids		1		1			
PCB		2			2		
pH (1)[a]	1	1					1
Phosphorus (2)[a]	12	39	4	17	16		2
Sediment	5	31		3	28		
Total dissolved solids	1	1			1		
Turbidity		9			9		
Total: 295[a]	101	188	14	79	86	7	10

[a]Number of No MOS case.

limits and monthly discharge limits of C_{BOD}, N_{BOD}, and P; (2) the Newport Bay System, MD and Swan Creek, MD, used 5% of the nonpoint source load as their MOS; and (3) the Upper Naugatuck River, CT, determined the MOS as the difference between the point source load and the nonpoint source load. Kane [13] suggested a method to calculate the MOS for phosphorus load as function of data variability; however, this method is basically the same as the explicit method that selects an MOS between 10-20 % of the total TMDL based on monitored data.

In this study, Walker's method and QUAL2E-UNCAS [10] were used to estimate the margin of safety for the Nakdong River. During the year of 2005, 40 sets of weekly flow rate and water quality data were collected, including BOD and TP. A spreadsheet was developed using Walker's method and the confidence level and the compliance level were assumed as 90% and 75%, respectively. The coefficients of variation for a Monte Carlo simulation were chosen from QUAL2E-UNCAS [10].

Table 2 shows results of the MOV calculations using Walker's method and the MOU calculations using QUAL2E-UNCAS for major points in the Nakdong River. As shown in the table, values of the MOV from Walker's method for both variables are 26% and 23% for BOD and TP, respectively. According to Walker's method, this will result in an MOS of greater than 36% and 34% for BOD and TP, respectively. These values are significantly higher than those in the approved TMDL reports as shown in Table 1.

In general, Walker's method tends to result in high MOS values since data variability and model uncertainty are considered together. Variability in water quality data is a function of many factors including watershed waste load delivery characteristics, accidental spills, and mixing characteristics in the water. Therefore, prediction of data variability can be very complicated and accurate estimation can be very difficult.

The Monte Carlo simulation method has been widely used for estimation of model uncertainty [2]. As shown in Table 2 when MOU was used alone, the margin of safety values for BOD and TP for the study site are 10% and 11%, respectively. These values are less than the maximum value of the MOS in Table 1. The current MOS value in Korean TMDL reports is 10%. Though Monte Carlo simulation results require further validation, it is interesting to note the similarity of the MOS values. However, for large complex dynamic models with disperse parameter distributions, Monte Carlo simulation may be computationally infeasible. Further, estimation of probability distributions is difficult to do reliably and may need a number of assumptions. To overcome the difficulties of Monte Carlo simulation, application of some advanced alternative

Table 2
Estimation of margin of safety using Walker's method and QUALE2E-UNCAS

	Walker's method		QUAL2E-UNCAS		MOS	
MOS	MOV		MOU		MOV+MOU	
Sites	BOD	TP	BOD	TP	BOD	TP
ND_A	0.23	0.29	0.12	0.13	0.35	0.42
ND_B	0.29	0.32	0.11	0.13	0.4	0.45
ND_D	0.22	0.25	0.08	0.09	0.3	0.34
ND_E	0.29	0.15	0.08	0.10	0.37	0.25
ND_F	0.28	0.22	0.08	0.09	0.36	0.31
ND_G	0.26	0.17	0.10	0.11	0.36	0.28
Average	0.26	0.23	0.10	0.11	0.36	0.34

methods have been suggested including "Generalized Sensitivity Analysis" [14,15], "Interval Spaced Sensitivity Analysis" [16], and Generalized Likelihood Uncertainty Estimation (GLUE; [17]). Stow et al. [18] discussed the advantages of the Markov Chain Monte Carlo method in detail, and they proposed use of the method in MOS estimation for TMDL development.

5. Conclusions

It is essential to consider model uncertainty and seasonal variability of data when estimation of allowable waste load for a water body. Margin of safety (MOS) is considered for such a purpose but its basis of quantification has not been discussed in depth. In this paper, MOS estimation methods were evaluated using literature survey of TMDL reports in US, Walker's method and Monte Carlo Method in QUAL2E-UNCAS model. It seems there is no certain rule in the selection of MOS in literatures. Walker's method seems to over-estimate MOS, therefore it is difficult to apply this method directly. Monte Carlo Method seems to provide a reasonable estimate. However as input variables may become more complex, the application of Monte Carlo Method can be prohibitive due to calculation requirement. It is recommended to investigate the possibility of employing more advanced statistical techniques, such as GLUE (Generalized Likelihood Uncertainty Estimate) or MCMC (Markov Chain Monte Carlo) Technique.

References

[1] USEPA, Guidance for water-quality-based decisions: The TMDL process., USEPA Washington, DC, 1991.

[2] R. Canale and D. Seo, Performance, reliability and uncertainty of total phosphorus models for lakes--II. Stochastic analyses. Water Res., 30(1) (1996) 95–102.

[3] H. Zhang, The critical flow-storm approach and uncertainty analysis for the TMDL development process, PhD Dissertation, Department of Civil Engineering, University of Virginia, Charlottesville, VA, 2000.

[4] H. Zhang and S. Yu, Applying the first-order error analysis in determining the margin of safety for total maximum daily load computations. J. Environ. Eng., 130(6) (2004) 664–673.

[5] D. Dilks and P. Freedman, Improved consideration of the margin of safety in total maximum daily load development. J. Environ. Eng., 130(6) (2004) 690–694.

[6] NRC, The National Academy of Sciences' National Research Council report on assessing the scientific basis of the total maximum daily load approach to water quality management : hearing before the Subcommittee on Water Resources and Environment of the Committee on Transportation and Infrastructure, House of Representatives, United States Congress. House. Committee on Transportation and Infrastructure. Subcommittee on Water Resources and Environment, Washington, DC, 2001.

[7] S. Chapra, Engineering water quality models and TMDLs. J. Water Res. Planning Manage., 129(4) (2003) 247–256.

[8] K. Reckhow, On the need for uncertainty assessment in TMDL modeling and implementation. J. Water Res. Planning Manage., 129(4) (2003) 245–246.

[9] K. Reckhow and S. Chapra, Engineering Approaches For Lake Management Volume 1: Data Analysis and Empirical Modeling, Butterworth, 1983.

[10] L. Brown and T. Barnwell, The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual, EPA/600/3-87/007, 1987.

[11] W. Walker, Quantifying uncertainty in phosphorus TMDL's for Lakes, W.W.J. Walker, ed., NE Interstate Water Pollution Control Commission and USEPA Region I, 2001, p. 22.

[12] W. Walker, Consideration of variability and uncertainty in phosphorus total maximum daily loads for lakes. J. Water Res. Planning Manage., 129(4) (2003) 337–344.

[13] K. Kane, Methodology for calculating Phase II total maximum daily loads (TMDLs) of phosphorus for New York City drinking water reservoirs, New York City Department of Environmental Protection, 1999.

[14] R. Spear and G. Hornberger, Eutrohpication in Peel Inlet-II. Identification of critical uncertainties via generalized sensitivity analysis. Water Res., 14 (1980) 43–49.

[15] B. Van Harn Adams, Parameter Distributions for Uncertainty Propagation in Water Quality Modeling, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC, 1998.

[16] J. Benaman and C.A. Shoemaker, Methodology for analyzing ranges of cncertain model parameters and their impact on total maximum daily load process. J. Environ. Eng., 130(6) (2004) 648–656.

[17] K. Beven and A. Binley, The future of distributed models—model calibration and uncertainty prediction. Hydrol. Pro., 6(3) (1992) 279–298.

[18] C. Stow, K. Reckhow, S. Qian, C. Lamon, G. Arhonditsis, M. Borsuk and D. Seo, Evaluating parameter uncertainty in water quality models: An imperative for adaptive TMDL implementation. J. Am. Water Res. Assoc., 43(6) (2007) 1499–1507.