Environ Geol (2006) 49: 405–412
DOI 10.1007/s00254-005-0085-0

ORIGINAL ARTICLE

Kyung-Sub Kim
Chung-Hwan Je

# Development of a framework of automated water quality parameter optimization and its application

Received: 20 June 2005
Accepted: 6 September 2005
Published online: 9 November 2005


K.-S. Kim
Department of Environmental Engineering, Hankyong National University, Ansung Kyonggi, Korea

C.-H. Je (✉)
Department of Environmental Health & Safety, University of Nevada, Nevada, Reno 89557, USA
E-mail: cje@unr.edu
Tel.: +1-775-327-2249
Fax: +1-775-784-4553

**Abstract** This paper presents a methodology and framework for the development of an automated least-squares optimization tool for calibrating water quality parameters in QUAL2E. The method has been applied to estimate the optimal water quality parameters in simulation of stream water quality for the Anyang stream in Korea. The Monte Carlo analysis is used to assess the relative importance of model parameters for water quality constituents. It is found that $\mu_{max}$ and $\rho$ are the most influential parameters for Chlorophyll-a modeling and $K_1$ and $K_3$ are critical parameters for variation of DO and BOD in the Anyang stream. A computer program for automated parameter calibration has been developed using a nonlinear GRG optimization algorithm. The application framework provides an intuitive and easy-to-use interface and allows visual evaluation of results. According to the simulation results, the automated approach is computationally efficient for evaluation of model parameters and converges on a best fit more rapidly and reliably than a trial and error method. The methodology proposed herein can be extended to other models to obtain the best possible parameter values.

**Keywords** Water quality · Optimization · Least-squares method · QUAL2E · Model calibration · Korea

## Introduction

Water quality models are used to simulate water quality changes resulting from different water quality management strategies and can also be applied to estimate the target and allowable pollution load in the watershed. The optimization of parameter variables in a given water quality model is an important issue in calibration of a water quality model. In general, the calibration process is an organized procedure to select model coefficients and improve the accuracy of model inputs that model predictions best match measured data and simulated values. According to Little and Williams (1992), model calibration is a critical process and is often ambiguous due to subjective, trial-and–error attempts to match observation and predictions. Traditionally, parameter calibration has been performed manually and repeatedly to determine the optimal parameter values. The process of manual calibration is a very tedious and time-consuming task, depending on the number of correlated parameters. Thus, many mathematical algorithms have been attempted to effectively estimate parameters in quantitative water quality models (Goktas and Aksoy 2004; Melching and Yoon 1996; Mulligan and Brown 1998; Noguchi others 2002; Rinaldi and others 1979; Van-Griensven and Bauwens 2001; Wood and others 1990; Yih and Davidson 1975).

One of the most popular parameter calibration techniques is a least-squares algorithm that has been extensively described in literature (Becker and Yeh 1972; Huang 1996; Je and Kim 2002; Kim others 2004). Least square is a mathematical optimization technique that attempts to find a “best fit” to a set of data by attempting to minimize the sum of the squares of the differences between predicted values and field measurement values. In this paper the same algorithm has been tested for the optimization of QUAL2E water quality model parameters with the Anyang stream data set in Korea. QUAL2E is well known as a comprehensive steady-state one-dimensional stream water quality model. The model, which is applicable to dendritic streams that are well mixed, uses a finite difference solution of the advective and dispersion mass transport and reaction equations using an implicit finite difference technique (Walton and Webb 1994). Reaction rate coefficient and constant values are provided by the user's manual (Brown and Barnwell 1987).

A number of researchers, as mentioned above, have attempted to identify appropriate values of the key water quality parameters in simulation of streams or lakes. However, these approaches are not only limited to use in the field requiring a rapid model calibration, but also do not provide an interactive automated parameter optimization including a graphical user interface. The primary objective of this paper is to optimize key water quality parameter values used in a comprehensive water quality model, QUAL2E, and to provide a framework for easy and rapid estimation of optimum model parameters through a graphical user interface. The framework of the proposed approach consists of three major parts: a QUAL2E model, graphical user-interface (GUI), and nonlinear least-squares optimization. The concentration changes of water quality constituents for a test stream is predicted by QUAL2E, and GUI is used as a tool to display output of the model and results of parameter optimization

## Methodology

### Optimization model development

Mathematical optimization involves determining optimal values for parameter values which will minimize or maximize an objective function subject to a set of constraints. These mathematical optimization techniques are widely applied in many engineering fields. This section describes in detail the model formulation for a least-squares optimization.

Model calibration is the optimization problem to determine model parameters that make small discrepancy between predicted and observed results. Thus, the objective function should be minimized by selecting optimal values of the parameters. In mathematical terms, the objective function is expressed as

$$\min\ z = \sum_{j=1}^{L} \sum_{i=1}^{N} \left(\varepsilon_{i,j}\right)^2, \tag{1}$$

where $\varepsilon$ is normalized residuals of observations and predictions, $i$ is the index for water quality constituents, $j$ is the observation point, $N$ is the number of water quality constituents, and $L$ is the number of observation points.

Parameter values are constrained to obtain the optimal solution in a reasonable range. The constraint of Eq. 1 is expressed by the following

$$p_{m,l} \le p_m \le p_{m,u} \quad m = 1, 2, \ldots, M, \tag{2}$$

where $p$ is parameter value, $m$ is a specific parameter, $l$ and $u$ are lower and upper limits, $M$ is the number of parameters, respectively.

Then, approximating to multi-parameter first-order Taylor's series expansion to minimize residuals obtained from new parameters, the equation can be obtained as

$$\varepsilon_{i,j}^{k} = \varepsilon_{i,j}^{k-1} + \left(p_1^k - p_1^{k-1}\right)\frac{\partial \varepsilon_{i,j}^{k-1}}{\partial p_1^{k-1}} + \cdots + \left(p_M^k - p_M^{k-1}\right)\frac{\partial \varepsilon_{i,j}^{k-1}}{\partial p_M^{k-1}}, \tag{3}$$

where superscript represents the iteration step. The partial derivative can be expressed as

$$\frac{\partial \varepsilon_{i,j}^{k-1}}{\partial p_m^{k-1}} \approx \frac{\Delta \varepsilon_{i,j}^{k-1}}{\Delta p_m^{k-1}}. \tag{4}$$

Incorporating Eqs. 3 and 4 into Eq. 1 and rearranging the equation gives

$$\min\ z = \sum_{j=1}^{L} \sum_{i=1}^{N} \left[\varepsilon_{i,j}^{k-1} + \left(p_1^k - p_1^{k-1}\right)\frac{\Delta \varepsilon_{i,j}^{k-1}}{\Delta p_1^{k-1}} + \cdots + \left(p_M^k - p_M^{k-1}\right)\frac{\Delta \varepsilon_{i,j}^{k-1}}{\Delta p_M^{k-1}}\right]^2. \tag{5}$$

The new parameters, $p_{\mathrm{m}}^{k}$, can be estimated from Eq. 5 with constraints. The objective function of Eq. 5 is nonlinear in the parameters to be estimated.

### Software description

The application software, QUAL2E_OPT, is designed to determine the best possible water quality parameter values using a nonlinear least-squares parameter optimization technique. The software has a user friendly interface written in the Visual Basic programming environment with a GRG (Generalized Reduced Gradient) optimization algorithm which is widely used to solve small-to-medium-sized nonlinear problems (Je and Kim 2004; Smith and Lasdon 1992). In addition, the system is flexibly designed to update or add new water quality constituents for parameter optimization because the system structure is modulated. To handle required water quality parameter values, QUAL2E_OPT provides reasonable default ranges provided in the QUAL2E user's manual and allows users to change the range of water quality parameter values suggested. The initial values are set to be the median value in the range of water quality parameter values. The observed water quality data used to calculate the least-squares error can be imported in Excel format.

QUAL2E_OPT can directly run a comprehensive water quality model, QUAL2E, written in FORTRAN codes as a subroutine in succession until the sum of squares error (SSE) between the predicted and observed data is achieved to a minimum under given constraints. While the optimization module is being executed, the number of iteration and sum of squares error value are dynamically displayed in the window form. Once the automated optimization process is complete, the user can view a graphical output for predicted concentrations and observed water quality data for use in analyzing goodness-of-fit.

As previously stated, the automated parameter estimation is to determine the best fit by finding the minimum SSE value under constraints. In this study, parameter estimation procedure consists of four main categories: (1) observation data to compare model prediction; (2) QUAL2E model to simulate stream water quality; (3) optimization module to determine optimal parameter values; (4) graphical output to display the goodness-of-fit between observed data and simulation results. The following is a proposed procedure to perform the automated parameter optimization for QUAL2E model. The detailed procedure diagram is shown in Figure 1.

1. Prepare a QUAL2E input data file before executing application.
2. Select a water quality constituent from a dropdown box.
3. Set a range of water quality parameters based on available information such as QUAL2E manual or literature values.
4. Import an observation data file in Excel format.
5. Run QUAL2E.
6. Calculate the SSE value.
7. Run non-linear optimization module to obtain the optimal parameters.
8. Update the model input parameters with new values obtained from optimization module.
9. Run QUAL2E and non-linear optimization module repeatedly until the minimum SSE value is achieved.

## Example application

### Description of the study area

This study has been conducted in the Anyang stream downstream of the Han River in Korea. The Anyang

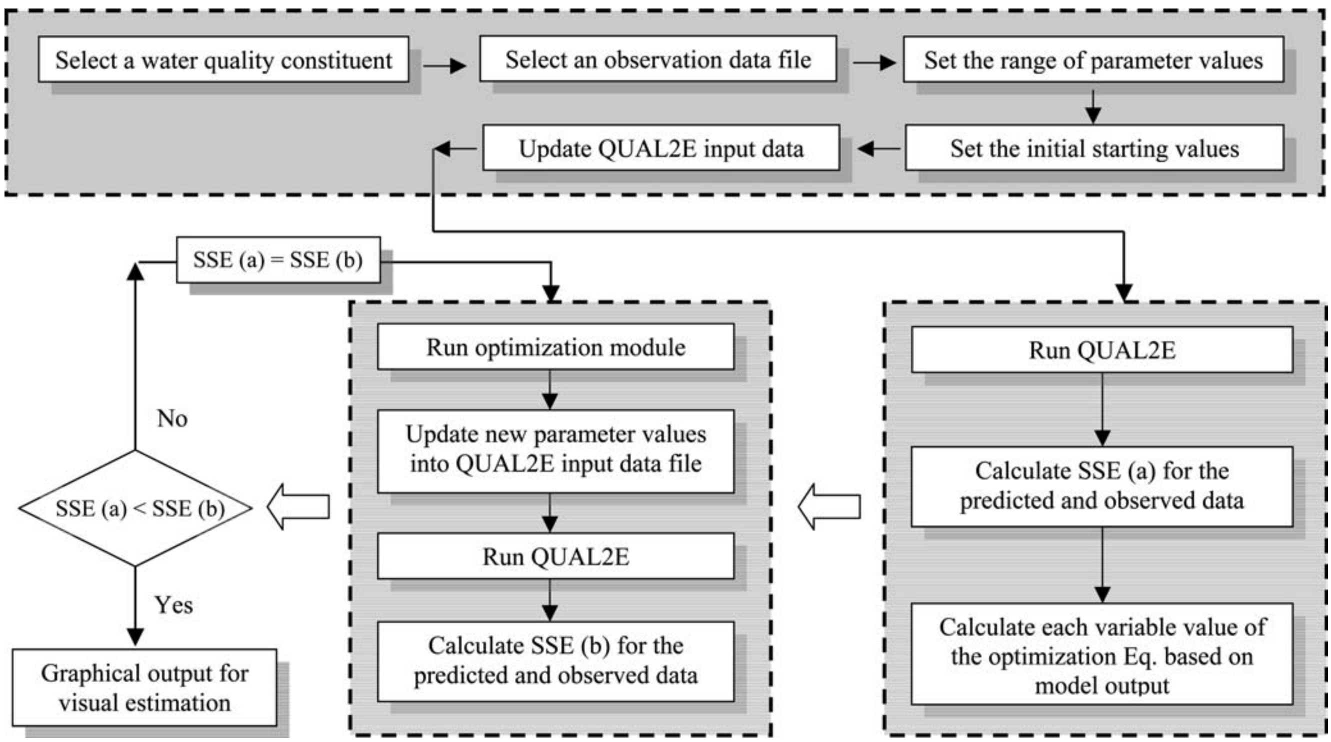


**Fig. 1** Schematic of parameter optimization procedure

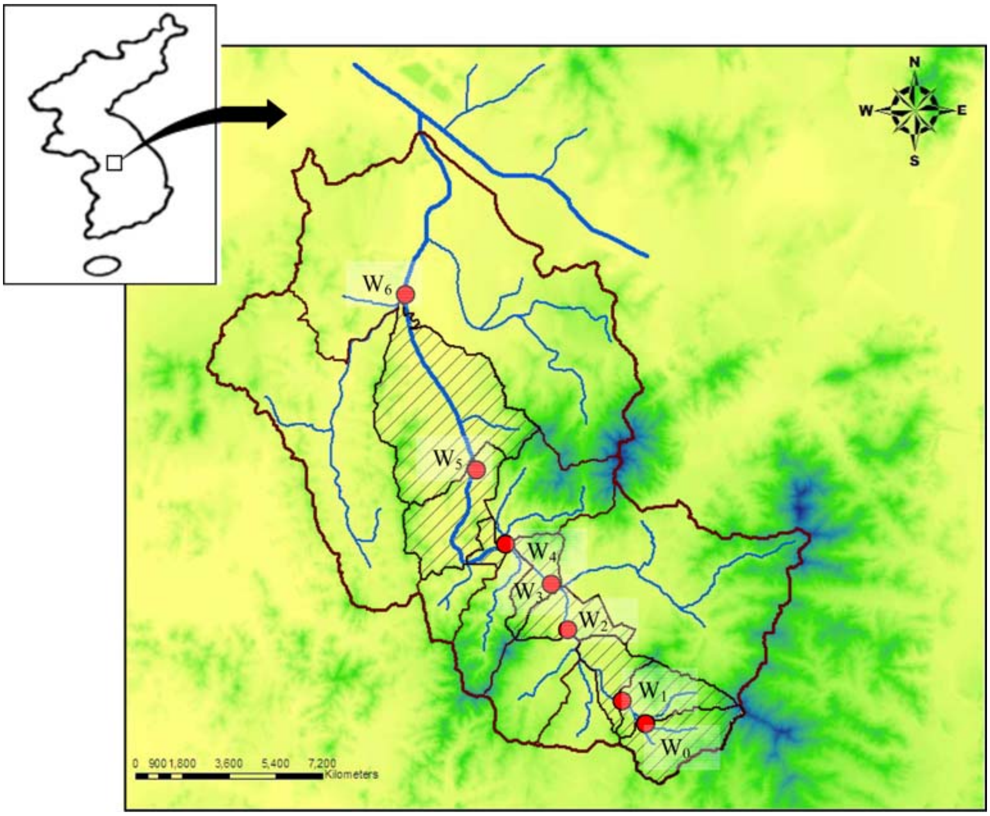


**Fig. 2** Map of the Anyang stream showing the monitoring locations

**Table 1** Water quality monitoring data at the Anyang stream in Korea

| Sampling Location | BOD | DO | Org-N | $NH_3$-N | $NO_3$-N | Org-P | Dis-P | Chl-a |
|---|---|---|---|---|---|---|---|---|
| $W_0$ | 2.97 | 8.50 | 8.10 | 0.41 | 0.63 | 0.08 | 0.43 | 7.61 |
| $W_1$ | 6.52 | 8.04 | 10.6 | 0.70 | 2.43 | 0.14 | 0.19 | 4.73 |
| $W_2$ | 32.3 | 3.10 | 25.7 | 12.79 | 1.10 | 0.44 | 0.62 | 6.58 |
| $W_3$ | 10.8 | 6.40 | 40.6 | 4.08 | 1.64 | 0.32 | 0.34 | 1.49 |
| $W_4$ | 18.5 | 5.20 | 10.5 | 9.45 | 1.17 | 0.56 | 0.46 | 3.37 |
| $W_5$ | 7.88 | 7.70 | 7.20 | 10.59 | 9.91 | 1.28 | 1.06 | 2.37 |
| $W_6$ | 8.23 | 6.10 | 10.7 | 11.24 | 5.89 | 1.05 | 0.63 | 2.88 |

**Table 2** Coefficient of variation of model parameters for water quality constituents

| Parameter | DO | BOD | Org-N | $NH_3$-N | $NO_3$-N | Org-P | Dis-P | Chl-a |
|---|---|---|---|---|---|---|---|---|
| $K_1$ | 0.010 | 0.060 | | | | | | |
| $K_3$ | | 0.012 | | | | | | |
| $\sigma_4$ | | | 0.002 | | | | | |
| $\beta_1$ | | | | 0.016 | 0.007 | | | |
| $\beta_2$ | | | | | 0.008 | | | |
| $\beta_3$ | | | 0.009 | 0.022 | | | | |
| $\sigma_5$ | | | | | | 0.002 | | |
| $\beta_4$ | | | | | | 0.013 | 0.011 | |
| $\rho$ | | | | | | | | 0.006 |
| $\mu_{max}$ | | | | | | | | 0.005 |

stream had been severely impaired in 70's and 80's by rapid growth of population and economic activity. Even though public concern of environment has improved the water quality of the Anyang stream in recent past decades, the unappropriated control of domestic and industrial wastewaters has hampered the water use.

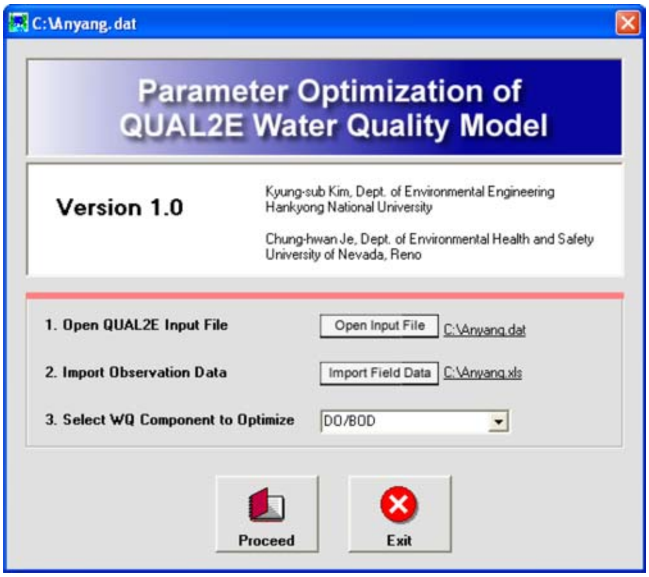


**Fig. 3** Initial menu of parameter optimization

Water quality control and management of this stream still take a public attention. QUAL2E input file was created to have 11 reaches and 117 computational elements each 0.2 km long to simulate approximately 23.4 km of the main stem Anyang stream. Water quality was monitored at seven monitoring stations. Figure 2 shows the map of the Anyang stream and the monitoring stations. The water quality monitoring results were shown in Table 1. As an example, the model prediction of eight water quality constituents (chlorophyll-a, dissolved oxygen (DO), biochemical oxygen demand (BOD), Org-N, $NH_3$-N, $NO_3$-N, Dis-P, and Org-P), which are important in the Anyang stream, was considered in this study.

## Estimation of optimal model parameters

QUAL2E model was calibrated for water quality constituent data collected during a field study of the Anyang stream in August 2002. The Monte Carlo uncertainty analysis was used as a screening method. The coefficient of variation (CV) was investigated to assess the relative importance of model parameters for water quality constituents in the Anyang stream. The range of variation in each parameter was obtained from the QUAL2E manual and then the Monte Carlo simulation of 2000 times was conducted to obtain a more stable estimation in the model output. The results after discarding the meaningless small values are summarized in Table 2. The high value of the coefficient of variation indicates a large fraction of the output variability. Even though the settling velocity of Chlorophyll-a ($\sigma$), the reaeration coefficient ($K_2$), and the sediment oxygen demand ($K_4$) have an impact on the water quality constituents, these parameters were not considered because of a very small value in the coefficient of variation (CV). Especially, $K_4$ is only applied in reach No.3 over 2 km long to consider the tributary input of highly settling organic matter.

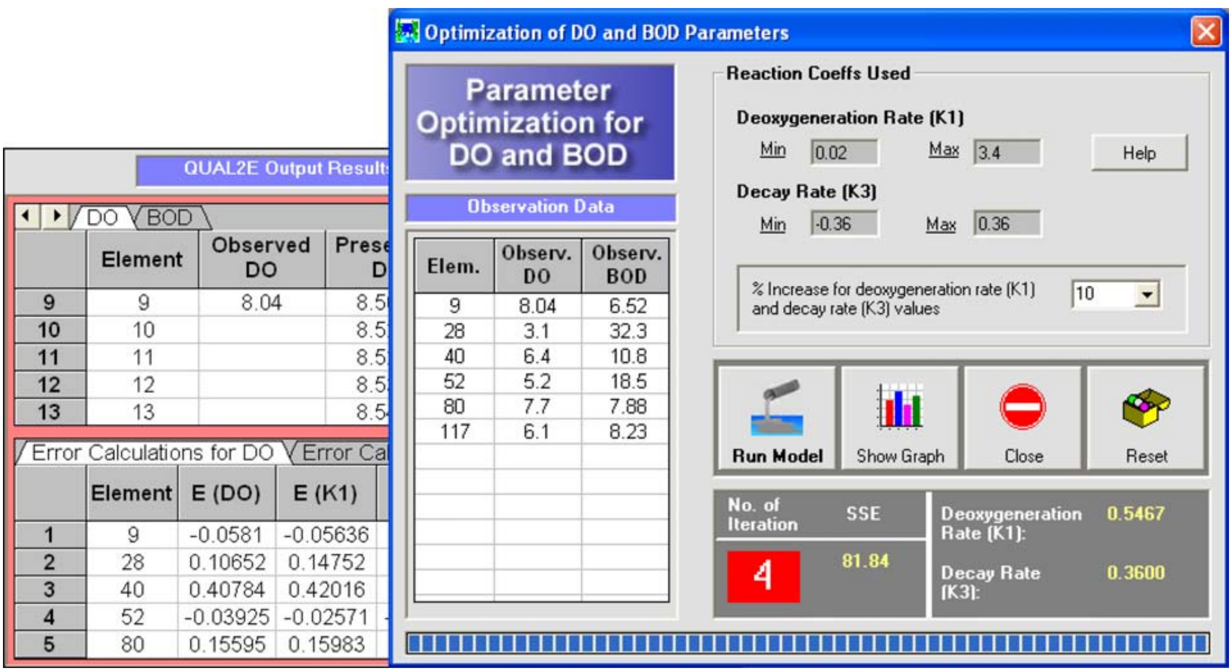


**Fig. 4** User interface of DO and BOD parameter optimization

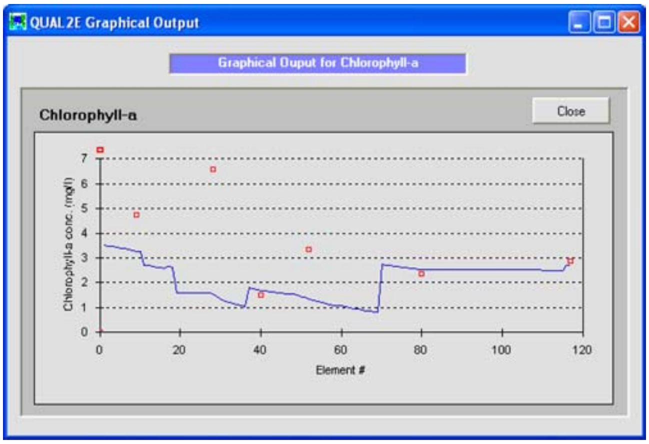


**Fig. 5** Graphical comparison of parameter optimization results for Chlorophyll-a

Overall, ten parameters were taken as main input model parameters in an optimization analysis. The key water quality parameters used for parameter optimization were selected from Table 2.

The parameter optimization process begins with the selection of a water quality constituent from the drop-down list in the initial menu as illustrated in Figure 3. The parameter optimization of the selected water quality constituent can be repeatedly run in the system until the SSE value between model output and field measurement is reached to a minimum in the least-squares error. Figure 4 shows a user-interface to run optimization process for DO and BOD as a typical example. QUAL2E_OPT provides visual inspections of results obtained from the optimization process. Graphical comparisons of parameter optimization results for eight water quality constituents are also shown in Figures 5 though 8. The solid line on each graph represents the line of a best fit by minimizing the sum of squares error and the rectangular symbol indicates the actual data observed in the Anyang stream. In addition, Help provides the reaction coefficient ranges of each water quality constituent parameter for running a model and is shown in Figure 9. The following gives a brief discussion of parameter optimization results obtained by QUAL2E_OPT.

$\mu_{max}$ and $\rho$ are selected as key parameters of Chlorophyll-a modeling based on the Monte Carlo uncertainty analysis of Anyang stream mentioned previously. It is shown that the minimum SSE between predicted and observed data is obtained when $\rho$ value approaches

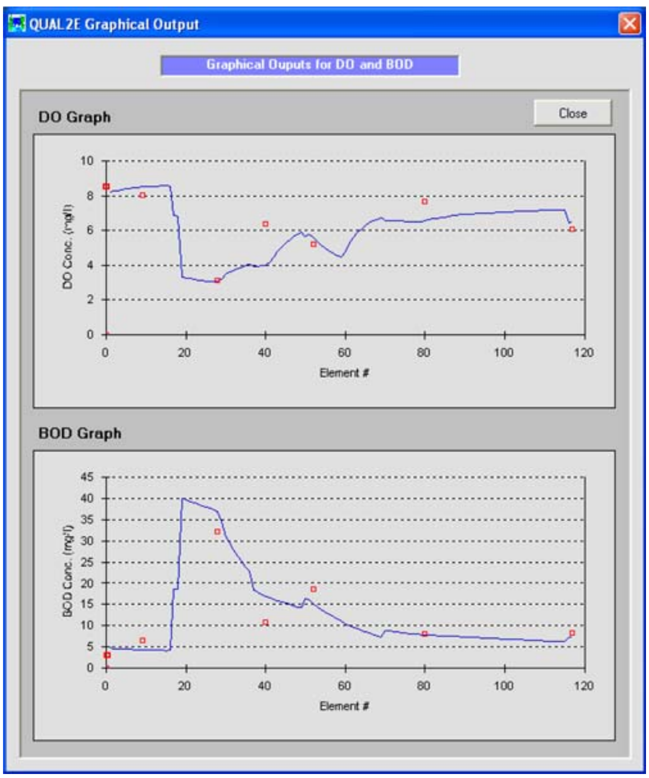


**Fig. 6** Graphical comparisons of parameter optimization results for DO and BOD

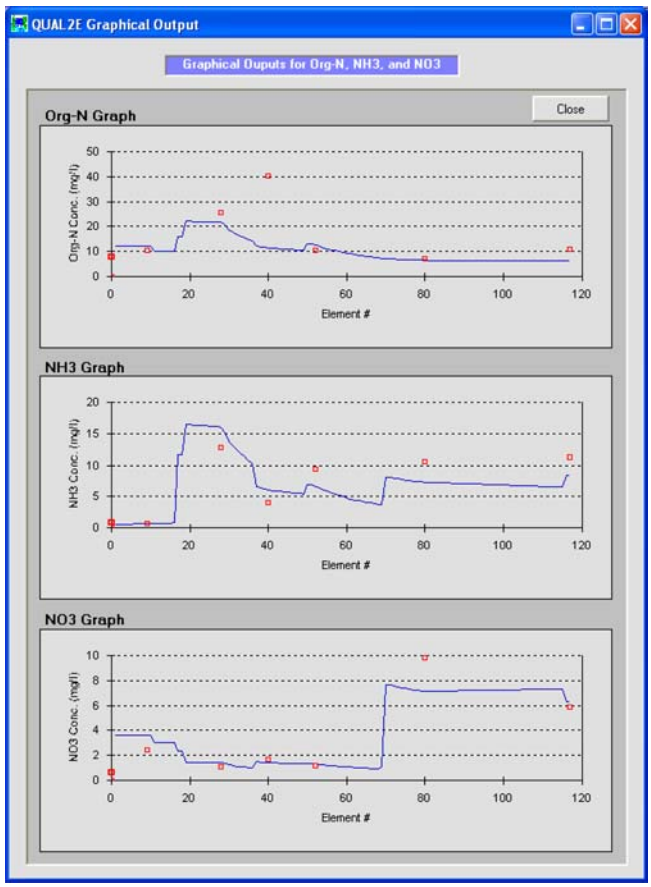


**Fig. 7** Graphical comparisons of parameter optimization results for Org-N, $NH_3$ and $NO_3$

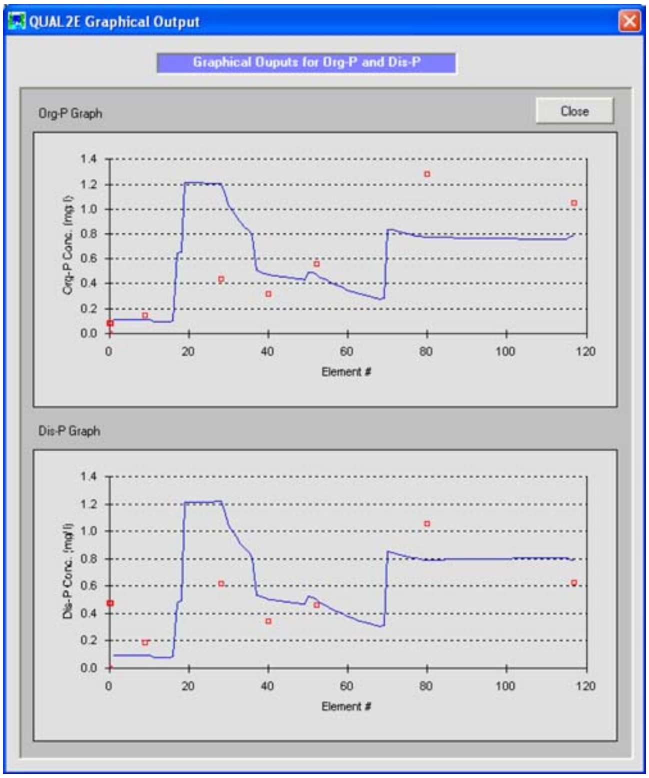


**Fig. 8** Graphical comparisons of parameter optimization results for Dis-P and Org-P

the upper bound of the range and $\mu_{max}$ value goes to the lower bound of the range. As shown in Table 3, the SSE value is not decreased within the limited range of $\mu_{max}$ and $\rho$ parameter values.

$K_1$ and $K_3$ are significant variables which affect DO and BOD concentrations in the Anyang stream. Especially, $K_3$ is a critical parameter that causes variation of DO and BOD, simultaneously. However, $K_1$ appeared to be the influential parameter only for DO. The result shows that the minimum SSE is obtained when $K_3$ value approaches the upper limit of the range and $K_1$ value goes to the lower bound of the range (Table 3). The SSE value is not decreased any more within the limited range of $K_3$ parameter value. Indeed, the minimum SSE is achieved by iteration 5 but the iteration can be carried out one more time to ensure that no further decrease in the SSE value is obtained. The DO and BOD plots illustrated in Figure 6 show adequate goodness-of-fits between predicted and observed data. In addition, Org-N, $NH_3$-N, and $NO_3$-N are calibrated using four water quality parameters selected, $\beta_1$, $\beta_2$, $\beta_3$, and $\sigma_4$, simultaneously. Minimum SSE value is obtained at the first iteration. No more decrease is found in the SSE value. The obtained calibration results show good agreement with observed data as shown in Figure 7. Finally, Org-P and Dis-P are also calibrated using $\beta_4$ and $\sigma_5$ in QUAL2E_OPT and the running results are shown in Figure 8.

## Conclusions

A methodology for calibrating the key water quality parameters in QUAL2E model has been presented along with a procedure for parameter optimization. The calibration procedure is based on a least-squares algorithm that determines optimal water quality parameters by minimizing the SSE between observed and predicted data. The Monte Carlo simulation is used to assess the

**Fig. 9** Typical reaction rate coefficients used in QUAL2E model

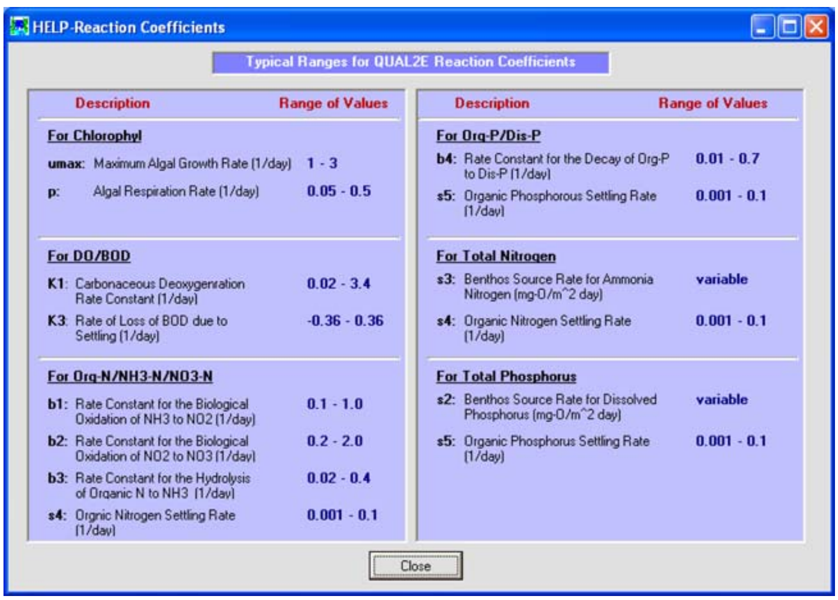

**Table 3** Estimated optimal water quality parameters

| Iteration | $\mu_{\max}$ | $\rho$ | SSE |
|---|---|---|---|
| For Chlorophyll-a | | | |
| 1 | 2 | 0.1581 | 31.82 |
| 2 | 3 | 0.0500 | 30.24 |
| Iteration | $K_1$ | $K_3$ | SSE |
| For DO and BOD | | | |
| 1 | 0.2608 | 0.0000 | 129.27 |
| 2 | 0.6365 | 0.0653 | 92.34 |
| 3 | 0.5346 | 0.3600 | 82.26 |
| 4 | 0.5453 | 0.3600 | 81.91 |
| 5 | 0.5467 | 0.3600 | 81.84 |

relative importance of model parameters for water quality constituents. It is found that $\mu_{\max}$ and $\rho$ are the most influential parameters for Chlorophyll-a modeling and $K_1$ and $K_3$ are critical parameters for variation of DO and BOD in the Anyang stream. In addition, it is shown that Org-N, $NH_3$-N, and $NO_3$-N concentration changes are dominated by $\beta_1$, $\beta_2$, $\beta_3$, and $\sigma_4$. It should be noted that the range of variation in each parameter should be determined carefully based on reliable data available in the literature and then the sensitivity test should be done to know the relative importance of parameters.

A computer program for automated water quality parameter calibration, QUAL2E_OPT, has been developed using a nonlinear GRG optimization algorithm. The optimization equation is derived from a multi-parameter first-order Taylor's series approximation. The application software has simple and intuitive user interface written in the VB programming environment to assist in the rapid implementation and ease of use. It has been applied to water quality parameter analyses for the Anyang stream in Korea. The results show that the automated approach is computationally efficient for estimation of model parameters and converges on a best fit more rapidly and reliably than a trial and error method. Thus, the automated water quality parameter optimization system can be used as a tool in rapidly and effectively calibrating the key parameters of a QUAL2E model. Furthermore, the methodology proposed herein could also be extended to other models to obtain the best possible parameter values.

## References


Becker L, Yeh WG (1972) Identification of parameters in unsteady open channel flow. WRR 8(4):956–965

Brown LC, Barnwell TO (1987) Documentation and user manual for the enhanced stream water quality models QUAL2E and QUAL2E-UNCAS. USEPA Report EPA/600/3-87/007, Environmental Research Laboratory, Athens, GA

Goktas RK, Aksoy A (2004) Applications of genetic algorithm for calibration and verification of QUAL2E model. In: Proceedings of the 2004 World Water and Environmental Resources Congress, Salt Lake City, Utah. ASCE, NY

Huang GH (1996) IPWM: an interval parameter water quality management model. Engineering Optimization 26:79–103

Je CH, Kim KS (2002) Evaluation of mathematical models for analyzing flocculent settling data. Environmental Progress 21(4):123–456

Je CH, Kim KS (2004) Web-based application for estimating water quality impacts due to environmental dredging. Environmental Geology 46(2):123–234

Kim KS, Yoon DG, Lee GY (2004) Calibration of parameters in QUAL2E using the least-squares method. J of Korea Water Res Assoc 37(9):719–727

Little KW, Williams RE (1992) Least-squares calibration of QUAL2E. Water Environment Research 64(2):179–185

Melching C, Yoon CG (1996) Key sources of uncertainty in QUAL2E model of Passive River. Journal of Water Resources Planning and Management 122:105–113

Mulligan A, Brown LC (1998) Genetic algorithms for calibrating water quality models. Journal of Environmental Engineering 124(3):202–211

Noguchi M, Solomatine DP, Nishida W (2002) Calibration of water quality model by global optimization techniques. In: Proceedings of the 5th International Conference on Hydroinformatics, Cardiff, UK. ASCE, NY

Rinaldi S, Romano P, Soncini-Sessa R (1979) Parameter estimation of Street-Phelps models. Journal of Environmental Engineering 105(1):75–88

Smith S, Lasdon L (1992) Solving sparse nonlinear programs using GRG. ORSA Journal on Computing 4(1):1–15

Van-Griensven A, Bauwens W (2001) Integral water quality modeling of catchments. Water Science and Technology 43(7):321–328

Walton R, Webb M (1994) QUAL2E simulations of pulse loads. Journal of Environmental Engineering 120(5):1017–1031

Wood D, Houck MH, Bell JM (1998) Automated calibration and use of stream quality simulation model. Journal of Environmental Engineering 116(2):236–248

Yih SM, Davidson B (1975) Identification in nonlinear distributed parameter water quality models. Water Resources Research 11(5):6693–704