

# Automatic calibration and selection of optimal performance criterion of a water quality model for a river controlled by total maximum daily load (TMDL)

Jae Heon Cho and Jong Ho Lee

## ABSTRACT

When manually calibrating water quality models, considerable time and attention are required. Hence, developing an automated model that allows for efficient and objective automatic calibration is highly desirable. The QUAL2Kw model calibrates the QUAL2K automatically using a genetic algorithm (GA). This study analyzes auto-calibration results and selects the optimal criterion for each objective function from six performance criteria. Additionally, a multi-objective auto-calibration was conducted using two kinds of performance statistics as the objective function of the GA. The auto-calibration model was applied to the Yeongsan River and the total maximum daily load (TMDL) was established to achieve water quality goals at specific target points of this river. Among the six auto-calibration results based on a single performance criterion, Nash-Sutcliffe model efficiency (NSE) was the best criterion for calculating fitness through auto-calibration. To consider the calibration accuracies of the TMDL target points and the entire river simultaneously, an objective function using multiple performance criteria, specifically the weighted average of the normalized root mean squares error (CV(RMSE)) and the ratio of the RMSE to the standard deviation of the observed data (the RSR), was selected as the final auto-calibration of the model. The model calibration performance was good across the whole region as well as at the target points.

**Key words** | automatic calibration, GA, multiple performance criteria, performance of the model calibration, QUAL2Kw

**Jae Heon Cho** (corresponding author)  
Department of Biosystems and Convergence  
Engineering,  
Catholic Kwandong University,  
24 Beomil-ro, 579 beon-gil, Gangneung-si,  
Gangwon-do 25601,  
Korea  
E-mail: [jhcho@cku.ac.kr](mailto:jhcho@cku.ac.kr)

**Jong Ho Lee**  
Department of Urban Planning and Real Estate,  
Cheongju University,  
298 Daeseongro, Cheongwon-gu, Cheongju,  
Chungbuk 28503,  
Korea

## INTRODUCTION

Manual calibration is commonly used in water quality modeling but it may introduce errors and lacks objectivity. Thus, various optimization techniques such as nonlinear programming, least square method and influence coefficient method are used for automatic correction of a water quality model. [Little & Williams \(1992\)](#) calibrated the QUAL2E model using the least squares. [Kim & Je \(2006\)](#) determined the parameters of biochemical oxygen demand (BOD), dissolved oxygen (DO), and chlorophyll-a content for QUAL2E using a nonlinear programming technique. [Li et al. \(2010\)](#) demonstrated the importance of calibration data length when estimating the optimal parameters and uncertainty performance of conceptual rainfall-runoff models. Particle swarm optimization was used to optimize the parameters of the rainfall-runoff model.

Recently, research on automatic calibration of water quality models using a genetic algorithm (GA) has been actively conducted. [Chlumecký et al. \(2017\)](#) described a novel

methodology and software for the optimization of rainfall-runoff modeling using a GA with the concept of a random number generator, which was at the core of the optimization. Their goal was to optimize the calibration of the model with minimal user interaction. [Cho & Ha \(2010\)](#) developed a parameter optimization method, using an influence coefficient algorithm ([Becker & Yeh 1972](#); [Becker & Yeh 1973](#)) and a GA for the calibration of the QUAL2K model. Many additional mathematical optimization techniques that have been used to estimate the parameters of models in various fields ([Wood et al. 1990](#); [Je & Kim 2004](#); [Goktas & Aksoy 2007](#); [Zou et al. 2009](#)).

[Pelletier & Chapra \(2006\)](#) developed the QUAL2Kw model by integrating the QUAL2K model ([Chapra et al. 2007](#)) and a GA. QUAL2K is one-dimensional river and stream water quality model intended to represent a well-mixed channel (both vertically and laterally) with steady state hydraulics, non-uniform steady flow, and diel heat

budget and water-quality kinetics. Pelletier *et al.* (2006) auto-calibrated the parameters of a small stream using QUAL2Kw, and they studied the improvement in fitness related to various crossover modes, population sizes, random number seeds, and evolutionary strategies. Cho & Lee (2018) applied QUAL2Kw to calibrate water quality model parameters for a river that was greatly influenced by wastewater treatment plant effluent.

Chlumecký *et al.* (2017) used various statistical indicators to validate a rainfall-runoff model. They focused on statistical methods and used the six most fundamental indicators, such as Nash-Sutcliffe model efficiency (NSE), root mean square error (RMSE) and correlation coefficient. Li *et al.* (2010) used the popular NSE and water balance error percentage (WBE) indexes as objective functions in the auto-calibration model. Mannina & Viviani (2010) proposed a simplified river water quality model for small rivers and used generalized likelihood uncertainty estimation (GLUE) to apply this model to a small Italian river. The RMSE, NSE, and  $R^2$  (coefficient of determination) values were reported as statistical performance indicators for four model state variables. The Soil and Water Assessment Tool (SWAT) was applied to a river catchment in England to quantify the long-term impacts of potential changes in agricultural management practices on river water quality (Taylor *et al.* 2016). Model performance was evaluated using three quantitative statistics, i.e., NSE, percent bias (PBIAS) and the ratio of the RMSE to the standard deviation of the observed data (the RSR) (Moriassi *et al.* 2007).

In river water quality modeling, it is common to analyze model performance using one or two of the following criteria: NSE,  $R^2$ , and RMSE. The objectives of this study are to analyze the results of automatic calibration using six different performance criteria, including the three above. First, this is to select the appropriate indicators for the performance analysis of the river water quality model, and second, to select the multi-performance criteria that can increase the precision of calibration at this point considering the importance of the total maximum daily load (TMDL) target point.

Since this study comprehensively analyzes various criteria that can be used for performance analysis of river water quality modeling, it can help select the optimal performance criterion according to the characteristics of the stream being modelled. In addition, it is necessary to improve the precision of calibration for important points of the entire stream area, such as the TMDL target point. In such cases, a methodology for selecting appropriate performance criteria for the stream was proposed through single objective and multiple objective optimization.

In this study, the QUAL2Kw model was applied to a heavily polluted river in South Korea. The GA of QUAL2Kw was executed with an objective function with several kinds of criteria, the fitness of the GA was calculated using the calibration error of the water quality model. The calibration results were compared and analyzed to identify the best performance criterion.

In our study, auto-calibration was conducted using six performance criteria – CV(RMSE),  $R^2$ , NSE, PBIAS, RSR, and the sum of the squares of the normalized residuals (SSNR); furthermore, multi-objective auto-calibration was also conducted using combinations of two kinds of performance criteria. Since the Korean TMDL is applied to this stream the water quality at the boundary point between the local government districts is important, as well as the BOD and total phosphorus (TP). The target water quality items of the Korean TMDL are more important than other water quality items and so in order to take this into account, weights were placed for each point and quality item in the GA's error calculation. A total of eight calibration results were compared and analyzed to determine the best performance criterion for this stream.

## METHOD

### Application of the water quality model to a polluted river

The QUAL2K model (Chapra *et al.* 2007; Cho 2011) is basically a one-dimensional steady state water quality model with the same characteristics as the QUAL2E model. In QUAL2K, various functions were added and programmed in VBA (Visual Basic for applications), and Excel was used as a GUI (graphic user interface). Unlike its predecessor QUAL2E, QUAL2K can apply different reach lengths and can apply multiple point source loads to a single reach. Carbonaceous BOD (CBOD) is classified into slow CBOD and fast CBOD, and denitrification is modeled as a first-order reaction that becomes pronounced at low oxygen concentrations. Sediment oxygen demand (SOD) and nutrient fluxes are simulated as a function of settling particulate organic matter and the quality of pore water in sediments can be calculated within the model. In QUAL2Kw (Pelletier & Chapra 2006; Pelletier *et al.* 2006), a GA is included to determine the optimum values for the kinetic rate parameters, so the GA of the PIKAIA algorithm (Charbonneau & Knapp 1995) was used for QUAL2Kw. For automatic calibration, we computed the fitness by comparing the water quality observations with the model results

and explored the water quality parameters to maximize the fitness.

The QUAL2Kw model was applied to the Yeongsan River located in southwestern Korea. The river is one of the target areas of the government's TMDL program. The modeling area includes the drainage area upstream of Yeongbon B, and the TMDL target points in the study area are Yeongbon A, Yeongbon B, and Whangyong A. The drainage area upstream of the TMDL site Yeongbon B is approximately 533 km<sup>2</sup>. The population density is high in the middle and lower reaches of the river. The effluent of the Gwangju wastewater treatment plant (WWTP) is discharged directly into the Yeongsan River, and the plant is located immediately above the confluence with the Gwangjucheon River. The effluent of the plant significantly affects the river quality and flow. Yeongbon B point is the boundary between Gwangju City and Jeollanam-do and 6.405 m<sup>3</sup>/sec of sewage effluent flows from Gwangju 1 WWTP, which is a large sewage treatment plant located in Gwangju city, about 12 km upstream. 1.248 m<sup>3</sup>/sec of sewage effluent flows from Gwangju 2 WWTP, about 3 km upstream from Yeongbon B (Table 1). A map of the study area is shown in Figure 1. The mean value of BOD<sub>5</sub> measured in the low flow season from October 2016 to February 2017 at Yeongbon B point, which is one of TMDL's target points, was 4.07 mg /L (6.44 mg /L as ultimate BOD; BOD<sub>u</sub>). The present water pollution level of the river is severe.

Water quality and flow measurements for the TMDL target points were conducted over 8 day intervals by the Ministry of Environment, South Korea. For the water quality survey points excluding the TMDL target points, the Ministry of Environment measures water quality parameters on the main river channel each month. Model calibration was performed using the mean values of the five months of water quality data set from October 2016 to February 2017.

Based on previous research on the GA parameters of the QUAL2Kw model, the population size, the generation number, and the crossover probability were set at 100, 150, and 0.6, respectively, within the GA. Uniform crossover was used as the crossover mode. One-point, fixed-rate mutation was used as the mutation mode, and steady-state-replace-worst was used as the reproduction plan (Cho 2013).

In this study, optimal parameters of the QUAL2Kw model were determined by using the user-defined auto-calibration function of the rates worksheet. Optimal parameters for each of the nine sections were calculated using automatic calibration for 13 items such as slow CBOD hydrolysis rate and slow CBOD oxidation rate, fast CBOD oxidation rate, organic N hydrolysis, organic N settling velocity, ammonium nitrification rate, nitrate denitrification rate, nitrate sediment denitrification transfer coefficient, organic P hydrolysis, organic P settling velocity, inorganic P settling velocity, detritus (POM) dissolution rate and detritus settling velocity. On the other hand, some parameters such as Inorganic suspended solids settling velocity, phytoplankton max growth rate, and phytoplankton respiration rate were calculated by applying a single parameter over the entire range and using automatic correction.

### Evaluation of the performance criteria used as the objective function of a GA in auto-calibration

The calibration performance of the water quality model depends on the fitness of the GA in the auto-calibration model. The optimal objective function of a GA is the fitness calculation formula that minimizes the error between the calculated and observed water quality values. To analyze the model performance, it is necessary to consider the importance of each water quality variable and water quality survey site. The river in this study is currently a part of the Korean TMDL program, and the boundaries between local government districts

**Table 1** | Flow and water quality of the influent and effluent of the WWTPs in the Yeongsan River basin (mean values for 2015)

WWTP		Flow (m <sup>3</sup> /sec)	BOD (mg/L)	COD (mg/L)	SS (mg/L)	TN (mg/L)	TP (mg/L)
Gwangju 1	Influent	–	147.6	79.7	131.3	32.512	3.378
	Effluent	6.405	2.4	7.9	3.1	11.353	0.107
Gwangju 2	Influent	–	120.8	59.9	114.4	37.375	3.772
	Effluent	1.248	1.6	6.9	1.2	9.91	0.083
Damyang	Influent	–	107.7	53.4	75.3	26.784	3.18
	Effluent	0.081	0.7	6.2	1.3	10.156	0.064
Jangseong	Influent	–	95.3	68.7	150.1	17.724	2.755
	Effluent	0.132	0.6	6.3	1.3	6.865	0.223

COD = chemical oxygen demand.

SS = suspended solids.

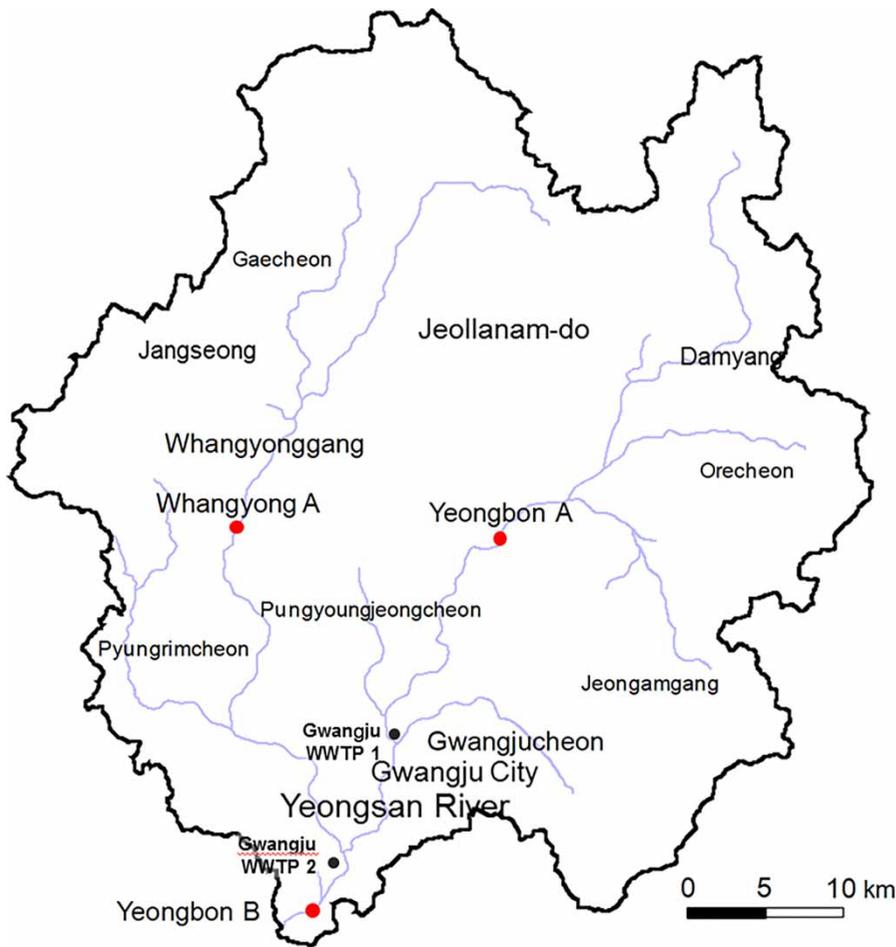


Figure 1 | Study area of the Yeongsan River basin.

have been set as water quality target points. Specifically, BOD and TP are the target water quality variables for water quality management. Hence, among the numerous water quality variables measured at each water quality survey site, BOD and TP were assigned higher weights. Similarly, for the entire river, higher weights were given to target points in terms of achieving the water quality goals. To assess the fitness formula of the GA, CV(RMSE),  $R^2$ , NSE, PBIAS, RSR, and SSNR were used as performance criteria, as they are often used to analyze the performance of water quality and rainfall-runoff models. The calibration results using the six performance criteria were analyzed to identify the best performance criterion. The fitness formula for the GA in QUAL2Kw was modified so that the fitness values for NSE and  $R^2$  were maximized while the errors of the other four performance criteria were minimized.

$$CV(RMSE) = \sum_{i=1}^n w_i \left[ \frac{[\sum_{j=1}^m (P_{ij} - O_{ij})^2 / m]^{1/2}}{(\sum_{i=1}^m O_{ij} / m)} \right] \quad (1)$$

$$R^2 = \frac{[\sum_{i=1}^n (P_i - \bar{P})(O_i - \bar{O})]^2}{\sum_{i=1}^n (P_i - \bar{P})^2 \sum_{i=1}^n (O_i - \bar{O})^2} \quad (2)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (P_i - \bar{O})^2} \quad (3)$$

$$PBIAS = \frac{\sum_{i=1}^n (O_i - P_i) \times 100}{\sum_{i=1}^n O_i} \quad (4)$$

$$RSR = \frac{\left[ \sqrt{\sum_{i=1}^n (O_i - P_i)^2} \right]}{\left[ \sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \right]} \quad (5)$$

Sum of the squares of the normalized residuals

$$= \sum_{i=1}^n \left( \frac{O_i - P_i}{\frac{O_i + P_i}{2}} \right)^2 \quad (6)$$

Additionally, to improve the accuracy of calibration in terms of target points and the entire river, this study further examined the possibility of integrating two performance criteria as the objective function of the GA as follows.

$$\text{Objective function: Minimize } (\alpha(E) + \beta(RSR)) \quad (7)$$

where  $E = CV(RMSE)$  or  $SSNR$  and  $RSR$  is the ratio of the RMSE to the standard deviation of observed data, with the relationship  $\alpha + \beta = 1$ .

## RESULTS AND DISCUSSION

### Performance statistics calculated by auto-calibration using a single performance criterion

The rates worksheet in Figure 2 selects whether or not to automatically calibrate parameters that are commonly applied across the target area of the river. Phytoplankton parameters related to phytoplankton growth and death, such as max growth rate, respiration rate, and death rate, and the equation for the reaeration model are selected. In the right-hand part, parameters of GA to be applied to

automatic calibration such as random number seed, generation number, population size, and crossover capability are entered. In the upper part, the fitness of the GA calculated according to the performance criterion used for the automatic calibration is displayed, and the values of each performance criterion are displayed. Figure 3 shows the 13 parameters of the water quality model determined using the user-defined auto-calibration function of the rates worksheet.

The auto-calibration performance statistics based on various performance criteria are shown in Table 2, which presents the calculation results of six different performance statistics, i.e.,  $CV(RMSE)$ , PBIAS, RSR, SSNR, NSE and  $R^2$  values, for the auto-calibration results of water quality models using eight performance criteria as the objective function.

The fitness of the model for  $CV(RMSE)$ , PBIAS, RSR, SSNR was calculated as the inverse of each performance statistic using measured and calculated water quality values. And the fitness of the model for NSE and  $R^2$  was calculated by their own statistics. High weights were given to CBOD and TP based on the water quality goals established by the TMDL program. For rivers with established TMDLs in Korea, a higher calibration accuracy is required for TMDL

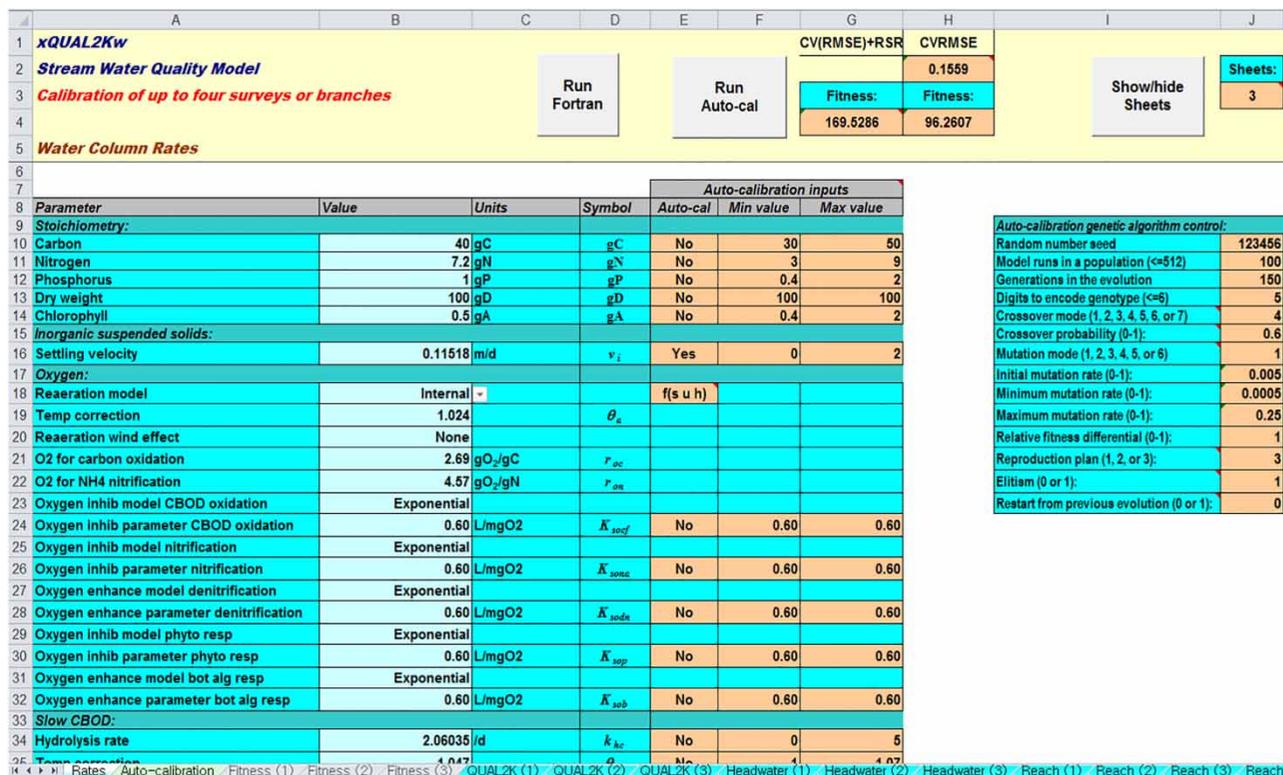


Figure 2 | Rates worksheet and performance in the QUAL2Kw.

Row	Parameter Name	Value	Unit
133	<b>User-defined auto-calibration parameters (optional)</b>		
134	Slow CBOD Hydrolysis Rate	0.65205	5
135	Slow CBOD Hydrolysis Rate	0.08	5
136	Slow CBOD Hydrolysis Rate	2.9656	5
137	Slow CBOD Hydrolysis Rate	3.01735	5
138	Slow CBOD Hydrolysis Rate	3.5843	5
139	Slow CBOD Hydrolysis Rate	3.88815	5
140	Slow CBOD Hydrolysis Rate	3.7857	5
141	Slow CBOD Hydrolysis Rate	0.2677	5
142	Slow CBOD Hydrolysis Rate	1.81505	5
143	Slow CBOD Oxidation Rate	0.50665	5
144	Slow CBOD Oxidation Rate	4.8384	5
145	Slow CBOD Oxidation Rate	4.9633	5
146	Slow CBOD Oxidation Rate	1.84065	5
147	Slow CBOD Oxidation Rate	3.80955	5
148	Slow CBOD Oxidation Rate	4.8734	5
149	Slow CBOD Oxidation Rate	0.6902	5
150	Slow CBOD Oxidation Rate	1.46985	5
151	Slow CBOD Oxidation Rate	0.9297	5
152	Fast CBOD Oxidation rate	0.8575	5
153	Fast CBOD Oxidation rate	4.90465	5
154	Fast CBOD Oxidation rate	1.45115	5
155	Fast CBOD Oxidation rate	4.508	5
156	Fast CBOD Oxidation rate	4.49905	5
157	Fast CBOD Oxidation rate	4.995	5
158	Fast CBOD Oxidation rate	0.77385	5
159	Fast CBOD Oxidation rate	0.49845	5
160	Fast CBOD Oxidation rate	0.0209	5
161	Organic N Hydrolysis	2.1113	5
162	Organic N Hydrolysis	4.87205	5
163	Organic N Hydrolysis	1.93935	5
164	Organic N Hydrolysis	4.94205	5
165	Organic N Hydrolysis	4.1063	5
166	Organic N Hydrolysis	4.0818	5
167	Organic N Hydrolysis	1.73485	5
168	Organic N Hydrolysis	2.4631	5
169	Organic N Hydrolysis	2.89605	5
170	Organic N Settling velocity	1.51632	2
171	Organic N Settling velocity	1.75182	2
172	Organic N Settling velocity	1.22002	2
173	Organic N Settling velocity	1.8346	2
174	Organic N Settling velocity	1.19844	2
175	Organic N Settling velocity	0.7333	2
176	Organic N Settling velocity	0.16342	2
177	Organic N Settling velocity	1.78974	2
178	Organic N Settling velocity	0.5376	2
179	Ammonium Nitrification	9.0238	10

Figure 3 | User-defined auto-calibration parameters in the QUAL2Kw.

Table 2 | Performance statistics determined from the auto-calibration using various performance criteria

Performance criteria for auto-calibration	NSE	R <sup>2</sup>	CV(RMSE)	PBIAS	RSR	SSNR (site weighted)
CV(RMSE) (site weighted)	0.60	0.72	0.16	8.73	0.22	0.06
R <sup>2</sup>	-2.35	0.73	0.80	49.46	0.56	1.00
NSE	0.68	0.79	0.18	4.74	0.20	0.06
PBIAS	-0.73	0.41	0.54	4.19	0.42	0.35
RSR	0.67	0.76	0.19	4.40	0.51	0.06
SSNR (site weighted)	0.59	0.72	0.16	9.02	0.01	0.05
SSNR (site weighted) + RSR	0.57	0.72	0.17	10.18	0.22	0.05
CV(RMSE) (site weighted) + RSR	0.65	0.77	0.16	7.19	0.21	0.05

target points. In the Yeongsan River, Yeongbon A (17 km from the headwaters) and Yeongbon B (43 km from the headwaters), shown in Figure 1, are the main target points. The BOD and TP, target water quality items of the Korean TMDL, were each given 100 weights. Phytoplankton was given 10, and DO and TN were given 1 weight each. The weights for the water quality monitoring site were given 100 to the TMDL Target points Yeongbon A and Yeongbon B, respectively, and the remaining points were given the weights of 10 and 5, respectively, depending on their importance.

NSE is a normalized statistic that describes the degree of the goodness of fit between model predictions and observations and can vary between  $-\infty$  and 1, where a value of 1 represents a perfect fit. An NSE value between 0 and 1 is generally recognized to indicate acceptable model performance, and the model performance associated with an NSE value greater than 0.5 is satisfactory in simulating discharge, nitrate and TP loads at a monthly time step (Kollat *et al.* 2012; Taylor *et al.* 2016). PBIAS is described as the average tendency of simulated data to overestimate or

underestimate a variable relative to observations. The optimum value of PBIAS is zero, indicating perfect agreement between model simulations and observations. A negative PBIAS value indicates overestimation and a positive value indicates underestimation (Taylor *et al.* 2016). RSR can vary from an optimum value of zero, indicating that there is no error between measured and simulated data, up to large positive values. A small RSR indicates good model performance (Moriassi *et al.* 2007; Taylor *et al.* 2016). For RMSE and  $R^2$ , the absolute error of the measured and calculated values greatly affects the error analysis. In this study, SSNR is used to ensure that unusually large measurements do not significantly affect the overall error. Normalizing the RMSE facilitates the comparison between datasets or models with different scales. When normalizing by the mean value of the measurements, the term coefficient of variation of the RMSE, CV(RMSE) may be used.

The calibration results of the water quality model for the Yeongsan River and the Whangyonggang River are shown in Figures 4 and 5. The square points on the figures of the

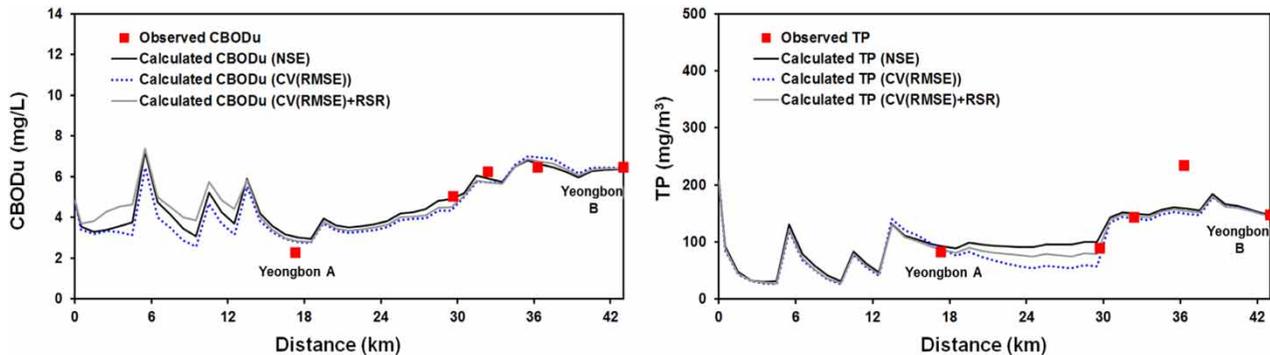


Figure 4 | Calibration results for the main stem of the Yeongsan River.

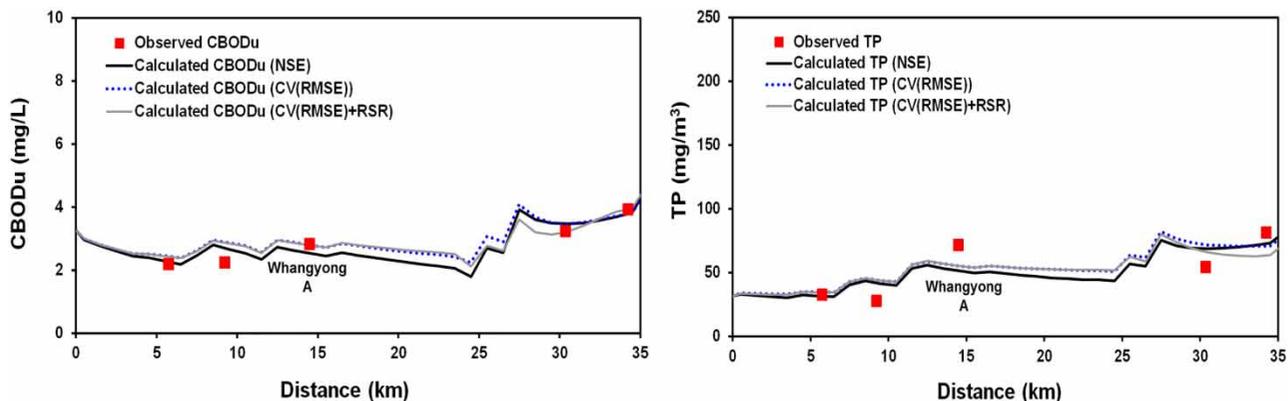


Figure 5 | Calibration results for the Whangyonggang River.

calibration results indicate the mean observed values of the low flow season at the TMDL target points and other water quality survey points from the Ministry of Environment.

In the course of the automatic calibration of this study, the target points of TMDL and target water quality items of TMDL were given relatively high weights. The target points of TMDL are 17 km and 43 km from headwaters of Yeongsan River and 17 km from headwaters of Whangyong-gang River. As shown in Figures 4 and 5, the mean observed values and calculated values of CBOD<sub>u</sub> and TP at the TMDL target points show good correspondence. Due to the fact that phytoplankton is influenced by many parameters compared to other water quality items, it is common that there is a relatively large error compared to BOD and TP. This study also showed that the phytoplankton error was relatively large because of the higher weight of BOD and TP, the target water quality of Korean TMDL.

As shown in Table 2, among the six auto-calibration results based on a single performance criterion, the model's performance was highest when NSE was used as the performance criterion to calculate fitness through auto-calibration, resulting in a generally low calibration error. The performances of NSE and  $R^2$  were good, with values of 0.68 and 0.79, respectively. The values of PBIAS, RSR and SSNR were 4.74, 0.2 and 0.06, respectively, and the performances associated with these criteria were relatively good in comparison with the calibration results of the other performance criteria. When CV(RMSE) was used as the performance criterion, the model performance was greatest at the TMDL target points (Yeongbon A, 17 km from the headwaters; Yeongbon B, 43 km from the headwaters), as shown in Figure 4. Moreover, the calibration error was fairly small at the target points when SSNR was used as the performance criterion.

### Model performance calculated by auto-calibration using multiple performance criteria

Due to the fact that the TMDL target point, with a designated water quality goal, is situated at the border of two local government districts means that the interests of two local governments are highly intertwined at this point. Since the Korean TMDL is being applied at the target point, the calibration accuracy at this point is the most important across all sections of the river. For the water quality target point, the fitness values and calibration accuracy tended to be high when CV(RMSE) and SSNR were used individually as the performance criteria. Therefore, this study examines a calibration method for the water quality

model using the sum of these two performance criteria as the GA objective function to improve the calibration accuracy both at the target points and along the entire river. The value for both  $\alpha$  and  $\beta$  in Equation (7) was set to 0.5.

If NSE is used, the overall calibration accuracy of the river tends to be higher. However, since the maximum value of NSE is 1, it is difficult to combine NSE with other variables, such as CV(RMSE) or SSNR, for use as the objective function to improve the accuracy at the target points. When calculating the fitness, CV(RMSE) (or SSNR) were given weights according to the importance of the water quality survey site. Therefore, the calibration accuracy of the water quality calculation is higher at target points of higher importance. For RSR, the overall calibration error is low, although no weights were given to different survey points. Hence, the performance of the water quality model after calibration using an objective function based on two performance criteria (Equation (7)), namely CV(RMSE) (or SSNR) and RSR, was calculated and compared with the results with the aforementioned results based on the six single performance criteria.

The calibration accuracy of the target points was significantly higher when the objective function involved a combination of CV(RMSE) and RSR than when the objective function involved a single performance criterion. In this case, the CV(RMSE) and SSNR values, calculated with weighted values for each water quality survey point, were 0.16 and 0.05. The model calibration performance was good at not just the target points but also along the entire river as the model yielded an NSE of 0.65 and an  $R^2$  of 0.77. The results of SSNR + RSR were also fairly good, although the calibration results of CV(RMSE) + RSR were far greater. When the auto-calibration results of CV(RMSE) + RSR was compared with the auto-calibration results of NSE based on a single performance criterion, the performance statistics of  $R^2$ , NSE, and RSR were fairly similar. With values weighted per the importance of each TMDL target point, the auto-calibration of CV(RMSE) + RSR yielded slightly better results than the auto-calibration of SSNR + RSR. When calibrated with SSNR + RSR as performance criteria, the values of NSE,  $R^2$ , CV(RMSE), PBIAS, RSR, and SSNR were calculated as 0.57, 0.72, 0.17, 10.18, 0.22 and 0.05, respectively. Therefore, this study chose the multiple performance criteria combination of CV(RMSE) + RSR as the final calibration.

Figure 6 shows the change in GA fitness when three objective functions, i.e., NSE, CV(RMSE), and CV(RMSE) + RSR, were used in the calculation for up to 150 generations. When NSE was used as the objective function, the

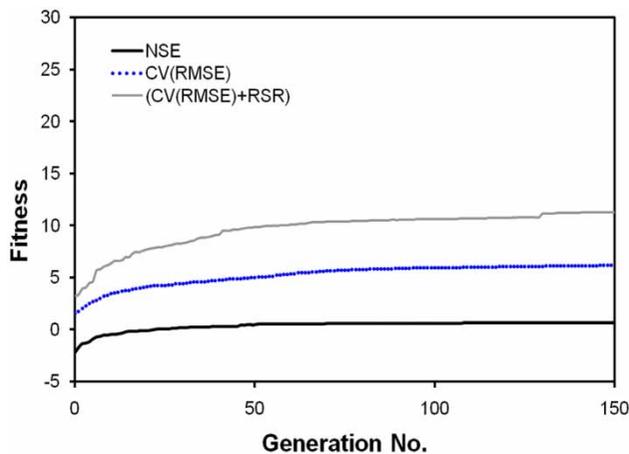


Figure 6 | Changes in GA fitness according to the number of generations.

final value was reached in a fairly early generation, whereas the calculation using CV(RMSE) + RSR and CV(RMSE) required at least 150 generations to reach the final value.

The water quality model with the auto-calibration using multiple performance criteria yielded slow CBOD hydrolysis rates of 0.08–3.88/d, slow CBOD oxidation rates of 0.51–4.96/d, fast CBOD oxidation rates of 0.02–4.90/d, organic P hydrolysis rates of 0.001–2.60/d, organic P settling velocities of 0.0001–1.97 m/d, inorganic suspended solids

settling velocity of 0.12 m/d, and phytoplankton max growth rate of 2.92/d (Table 3).

The verification of the model was carried out using the mean values of water quality measured by the Ministry of Environment in March 2017. The values NSE,  $R^2$ , CV (RMSE), and RSR were calculated as 0.44, 0.74, 0.27, 0.27, respectively. The performance of the model verification was relatively good.

## CONCLUSION

Auto-calibration of the parameters in a water quality model of a polluted Korean river in which TMDL limits have been implemented was conducted for water quality management purposes. The GA technique, along with various performance criteria, was used in the calibration process, and the calibration results were analyzed to identify the best performance criterion for use as the objective function of the GA. QUAL2Kw was executed with several kinds of criteria as the objective function. Auto-calibration was conducted using the six performance statistics, and a multi-objective auto-calibration was also conducted through the integration of two kinds of performance criteria. Overall,

Table 3 | Estimated parameter values using QUAL2Kw and their minimum and maximum ranges

Model parameter	Range of calibrated values					Min.	Max.
	CV(RMSE)	NSE	RSR	SSNR	CV(RMSE) + RSR		
Slow CBOD hydrolysis rate (/d)	0.52–4.97	0.02–4.23	0.039–4.76	0.008–4.45	0.08–3.88	0	5
Slow CBOD oxidation rate (/d)	0.003–4.98	0.06–4.78	0.0067–4.99	0.27–4.97	0.51–4.96	0	5
Fast CBOD oxidation rate (/d)	0.07–4.99	0.12–4.99	0.04–4.99	0.28–4.99	0.02–4.90	0	5
Organic N hydrolysis rate (/d)	0.35–4.75	0.06–4.22	0.074–3.84	0.014–4.94	1.73–4.94	0	5
Organic N settling velocity (m/d)	0.27–1.75	0.05–1.58	0.067–1.98	0.055–1.34	0.16–1.83	0	2
Ammonium nitrification rate (/d)	0.88–7.89	0.02–7.61	0.53–8.35	0.16–7.70	1.41–9.02	0	10
Nitrate denitrification rate (/d)	0.75–1.99	0.45–1.52	0.11–1.72	0.036–1.99	0.43–1.38	0	2
Nitrate sed. denitrification transfer coeff. (m/d)	0.03–0.83	0.02–0.96	0.0024–0.97	0.018–0.99	0.16–0.95	0	1
Organic P hydrolysis rate (/d)	0.001–4.12	0.02–4.37	0.016–0.68	0.006–4.35	0.001–2.60	0	5
Organic P settling velocity (m/d)	0–1.97	0.001–1.67	0.0012–1.94	0.0001–1.67	0.0001–1.97	0	2
Inorganic P settling velocity (m/d)	0.001–1.99	0.001–1.99	0.0002–1.98	0.004–1.96	0.0004–1.97	0	2
Detritus dissolution rate (/d)	0.08–4.69	0.60–4.99	0.15–3.85	0.021–4.73	0.15–4.33	0	5
Detritus settling velocity (m/d)	0.64–4.99	0.04–4.93	0.015–4.82	0.003–4.92	0.0001–4.98	0	5
Inorganic suspended solids settling velocity (m/d)	0.222	0.16	0.96	0.24	0.12	0	2
Phytoplankton max growth rate (/d)	2.764	2.69	2.99	2.90	2.92	1.5	3
Phytoplankton respiration rate (/d)	0.089	0.029	0.0016	0.065	0.077	0	1
Phytoplankton death rate (/d)	0.199	0.10	0.21	0.28	0.23	0	1

the calibration accuracy of the BOD and TP was satisfactory.

Among the auto-calibrations based on a single performance criterion, using the NSE criterion as the objective function yielded the highest model performance among all six auto-calibrated results for the entire river. When automatically calibrated with NSE as the performance criterion, the values NSE,  $R^2$ , CV (RMSE), PBIAS, RSR, and SSNR were calculated as 0.68, 0.79, 0.18, 4.74, 0.20 and 0.06, respectively.

When CV(RMSE) was used as the performance criterion, the model performance was highest at the TMDL target points. Furthermore, when SSNR was used as the performance criterion, the calibration error was small at the target points.

When CV(RMSE) and RSR were used together as an objective function, the auto-calibration performance was good in terms of the CV(RMSE) and SSNR values calculated with weighted values for the target points, and the calibration accuracy was greater than that of the calibrations based on other performance criteria. When automatically calibrated with CV(RMSE) + RSR as the performance criterion, the values of NSE,  $R^2$ , CV (RMSE), PBIAS, RSR, and SSNR were calculated as 0.65, 0.77, 0.16, 7.19, 0.21 and 0.05, respectively. The NSE and  $R^2$  values were satisfactory, and the performances of the model calibrations were good for all sections of the river. The performance of SSNR + RSR was fairly satisfactory, and it was close to the results of CV(RMSE) + RSR. Considering the importance of the TMDL target points, this study selected the auto-calibration results that utilized the multiple performance criteria combination of CV(RMSE) + RSR as the final calibration. When TMDL target points are not a concern, using the NSE performance criterion as the objective function for auto-calibration is more appropriate.

## ACKNOWLEDGEMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: NRF-2017R1D1A1B03032816).

## REFERENCES

- Becker, L. & Yeh, W. W. 1972 Identification of parameters in unsteady open channel flows. *Water Resources Research* 8 (4), 956–965.
- Becker, L. & Yeh, W. W. 1973 Identification of multiple reach channel parameters. *Water Resources Research* 9 (2), 326–335.
- Chapra, S. C., Pelletier, G. J. & Tao, H. 2007 *QUAL2K: A Modeling Framework for Simulating River and Stream Water Quality*. Version 2.07: Documentation and Users Manual, Civil and Environmental Engineering Dept., Tufts University, Medford, MA.
- Charbonneau, P. & Knapp, B. 1995 *A User's Guide to PIKAIA 1.0*, NCAR Technical Note 418 + 1A. National Center for Atmospheric Research, Boulder. <http://www.hao.ucar.edu/modeling/pikaia/pikaia.php#sec1/> (accessed 12 February 2019)
- Chlumecký, M., Buchtele, J. & Richta, K. 2017 Application of random number generators in genetic algorithms to improve rainfall-runoff modelling. *Journal of Hydrology* 553, 350–355.
- Cho, J. H. 2011 Application of the QUAL2Kw model to a polluted river for automatic calibration and sensitivity analysis of genetic algorithm parameters. *Journal of Environmental Impact Assessment* 20 (3), 357–365.
- Cho, J. H. 2013 Waste load allocation method for total maximum daily load program of a polluted river. *Journal of Environmental Impact Assessment* 22 (2), 157–170.
- Cho, J. H. & Ha, S. R. 2010 Parameter optimization of the QUAL2K model for a multiple-reach river using an influence coefficient algorithm. *Science of The Total Environment* 408 (8), 1985–1991.
- Cho, J. H. & Lee, J. H. 2018 Automatic calibration and performance evaluation of a water quality model for a river greatly influenced by wastewater treatment plant effluent. In: *Proceedings of HIC 2018, 13th International Conference on Hydroinformatics, Palermo*, pp. 447–455.
- Goktas, R. K. & Aksoy, A. 2007 Calibration and verification of QUAL2E using genetic algorithm optimization. *Journal of Water Resources Planning and Management-ASCE* 133 (2), 126–136.
- Je, C. & Kim, K. 2004 Web-based application for estimating water quality impacts due to environmental dredging. *Environmental Geology* 46 (2), 217–225.
- Kim, K.-S. & Je, C.-H. 2006 Development of a framework of automated water quality parameter optimization and its application. *Environmental Geology* 49 (3), 405–412.
- Kollat, J. B., Reed, P. M. & Wagener, T. 2012 When are multiobjective calibration trade-offs in hydrologic models meaningful? *Water Resources Research* 48 (3), 0043–1397.
- Li, C. Z., Wang, H., Liu, J., Yan, D. H., Yu, F. L. & Zhang, L. 2010 Effect of calibration data series length on performance and optimal parameters of hydrological model. *Water Science and Engineering* 3 (4), 378–393.
- Little, K. W. & Williams, R. E. 1992 Least-squares calibration of QUAL2E. *Water Environment Research* 64, 179–185.
- Mannina, G. & Viviani, G. 2010 Water quality modelling for ephemeral rivers: model development and parameter assessment. *Journal of Hydrology* 393 (3–4), 186–196.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in

- watershed simulations. *Transactions of the ASABE* **50** (3), 885–900.
- Pelletier, G. J. & Chapra, S. C. 2006 QUAL2Kw theory and documentation (version 5.1): A modeling framework for simulating river and stream water quality. <http://www.ecy.wa.gov/programs/eap/models/>.
- Pelletier, G. J., Chapra, S. C. & Tao, H. 2006 QUAL2Kw – A framework for modeling water quality in streams and rivers using a genetic algorithm for calibration. *Environmental Modelling and Software* **21** (3), 419–425.
- Taylor, S. D., He, Y. & Hiscock, K. M. 2016 Modelling the impacts of agricultural management practices on river water quality in Eastern England. *Journal of Environmental Management* **180**, 147–163.
- Wood, D. M., Houck, M. H. & Bell, J. M. 1990 Automated calibration and use of stream quality simulation model. *Journal of Environmental Engineering* **116** (2), 236–249.
- Zou, R., Lung, W. S. & Wu, J. 2009 Multiple-pattern parameter identification and uncertainty analysis approach for water quality modeling. *Ecological Modelling* **220** (5), 621–629.

First received 14 February 2019; accepted in revised form 19 June 2019. Available online 26 June 2019