The impact of objective function formulations on the optimal calibration of QUAL2E

R. K. Goktas[1] and A. Aksoy[2] A.M. ASCE

Department of Environmental Engineering, Middle East Technical University,
06531 Ankara, TURKEY
[1]e-mail: recepkay@metu.edu.tr
[2]e-mail: aaksoy@metu.edu.tr

Abstract

Successful calibration and verification of a model is important to confirm its usage for water quality predictions. Instead of trial-and-error approach, optimization techniques can be used to adjust the biological, chemical, and kinetic parameters in a more efficient way. In this study, a genetic algorithm (GA) is used to calibrate and verify QUAL2E for the reaeration coefficient (K_2) and the sediment oxygen demand rate (K_4) with reference to dissolved oxygen (DO) observation data. Three different objective functions were used to investigate the impact of formulation itself on the performance of the optimization models. The results show that for the river system considered in this study, objective function formulation may have an impact for a successful outcome.

Keywords: calibration, verification, QUAL2E, genetic algorithms, optimization, water quality modeling

Introduction

Simulation models are helpful tools for water quality control and management. Calibration and verification phases of a modeling study are very important in order to confirm the applicability of the model to a specific system. Calibration seeks a match between the observed and simulated water quality variables by tuning the chemical, biological, and kinetic parameters. Verification, on the other hand, is the confirmation of the adjusted parameters for different physical conditions and forcing functions. Following the successful completion of these phases, the model can be

Copyright ASCE 2005

Downloaded 23 Mar 2011 to 194.27.72.122. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

used to make predictions about the faith of water quality in a water body for different scenarios.

Calibration and verification of a model can be tedious, taking the limited amount of data and the complex nature of current water quality simulation models into consideration. Although trial-and-error methods have been widely employed, application of optimization methods has been gaining attention. With optimization, the subjective nature of trial-and-error approach can be avoided.

Several researchers have applied optimization for calibration and verification. Yih and Davidson (1975) adjusted the longitudinal dispersion coefficient in a salinity intrusion model for estuaries using Marquardt's algorithm, conjugate gradient and steepest descent methods. Little and Williams (1992) calibrated the QUAL2E Model using non-linear programming. Cooper et al. (1997) applied shuffled complex evolution, GAs, and simulated annealing for the calibration of a rainfall-runoff model Mulligan and Brown (1998) used GAs to calibrate the Streeter-Phelps Model. Van Griensven and Bauwens (2001) calibrated the SWAT Model with the shuffled complex evolution method. Ng and Perera (2003) studied the importance of GA operator parameters on the calibration of QUAL2E.

QUAL2E is a widely applied stream water quality model that can simulate DO, biological oxygen demand, algae, temperature, nitrogen species, phosphorus species, and others (Brown and Barnwell, 1987). Trial-and-error methods have been a common method for the calibration and verification of the model (Vassilios et al., 1995; Onur et al., 1999). However, benefits of using optimization techniques have also been presented (Little and Williams, 1992; Ng and Perera, 2003). In this study, the impact of objective function formulations on the performance of optimal calibration and verification of QUAL2E is presented.

Methodology

A genetic algorithm (GA) was used as the optimization method. The QUAL2E code was modified to run under the Windows XP system, and then linked to a GA driver (Carroll, 1999). In this structure, the GA generated the potential values of the parameters to be calibrated and verified. QUAL2E returned the response of the river system to these parameters as the water quality constituent concentrations. These simulated concentrations, together with the observed ones, were used to calculate the value of the objective function, therefore the fitness of the potential solution. The search for the best parameter values was conducted using the GA operators. Details about these operators and the search procedure can be found elsewhere (Goldberg, 1989). In this study, although the algorithms got close to the best solution within 100 generations, the search continued for 10,000 generations for additional improvements. Each run was repeated three times with a different initial population of strings. The best solution among them was selected as the optimum.

Copyright ASCE 2005

Downloaded 23 Mar 2011 to 194.27.72.122. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

The performance of optimization was tested for three different objective functions. All of them focused on the differences between the observed and simulated concentration values, but in a different form. Since calibration and verification was performed simultaneously, two sets of observation data that represented different field conditions were used. QUAL2E was calibrated for K_2 and K_4, based on the observed DO concentrations at pre-located sampling points. Tested objective function formulations are given below;

Objective Function 1 (OF-1):

$$\text{Min } Z_1 = \left[\sum_{i=1}^{N}\left(C_{Ci} - S_{Ci}\right)^2\right]\left[1 + \sum_{i=1}^{N}\left(C_{Vi} - S_{Vi}\right)^2\right] \tag{1}$$

Objective Function 2 (OF-2):

$$\begin{aligned}\text{Min } Z_2 = &\left[\sum_{i=1}^{N}\left(C_{Ci} - S_{Ci}\right)^2\right]\left[1 + \sum_{i=1}^{N}\left(C_{Vi} - S_{Vi}\right)^2\right] \\ &+ \left[\sum_{i=1}^{N}\left(C_{Vi} - S_{Vi}\right)^2\right]\left[1 + \sum_{i=1}^{N}\left(C_{Ci} - S_{Ci}\right)^2\right]\end{aligned} \tag{2}$$

Objective Function 3 (OF-3):

$$\text{Min } Z_3 = \left[\sum_{i=1}^{N}\left(C_{Ci} - S_{Ci}\right)^2\right]\left[1 + w_c\left(D_{max\,C}\right)\right]\left[1 + w_v\left(D_{max\,V}\right)\right] \tag{3}$$

where, C_{ci} and C_{vi} were the observed DO concentrations at sampling point i for the field conditions of the calibration and verification phases, respectively (M/L^3), S_{ci} and S_{vi} were the simulated (predicted) DO concentrations at sampling point i for the field conditions of the calibration and verification phases, respectively (M/L^3), D_{maxC} and D_{maxV} were the maximum of errors at the sampling point locations for the calibration and verification conditions, respectively (mg/l), w_c and w_v were the weights for the calibration and verification conditions, respectively (-), and N was the number of sampling points (-).

The optimization models were applied to a hypothetical river system where the optimum solution was known a priori. This system was based on 'Sample Data Set 2' (wrkshop2.dat), distributed with the QUAL2E Model. The 45 km-long system was composed of 6 reaches as depicted in Figure 1. Sampling points were placed arbitrarily to have at least one sampling point at a reach. DO concentrations were simulated for the calibration and verification conditions. Different dry and wet bulb temperatures created the variance between these conditions. The concentrations at the sampling points were labeled as the observation data (Goktas and Aksoy, 2004).

Copyright ASCE 2005 EWRI 2005
Downloaded 23 Mar 2011 to 194.27.72.122. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

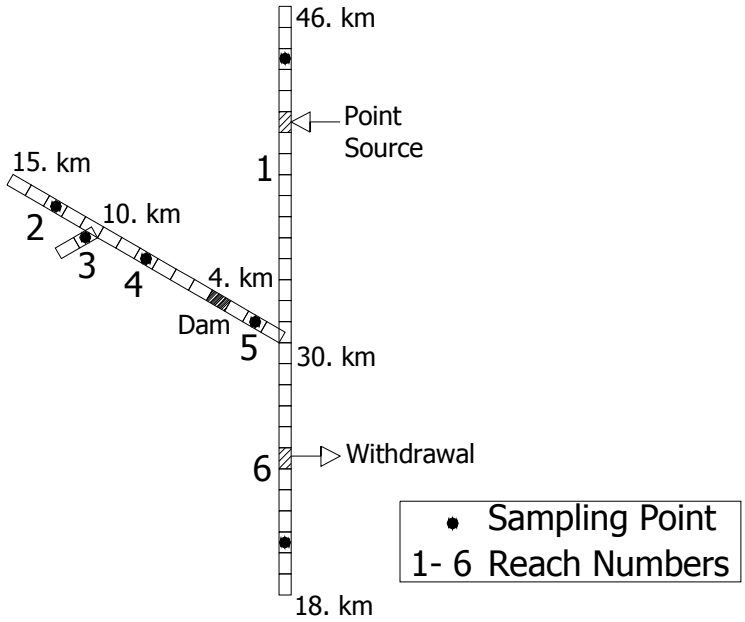


Figure 1. Hypothetical river system and the location of sampling points

The decision variables were encoded in binary GA strings to have a precision of 0.1 day^{-1} and 0.1 g/m^2day for K_2 and K_4, respectively. The search range was 0.0 – 100.0 day^{-1} for K_2, and 0.0 – 1.0 g/m^2-day for K_4. In order to analyze the performances of the optimization models in predicting the reach-variable K_2 and K_4 values (known a priori), the real and optimized parameter values were compared.

Example Results

K_2 values obtained for different objective function formulations are given Figure 2. OF-3 performed better compared to the others in terms of reaching to the goal quantities. The results for OF-2 were slightly superior compared to those for OF-1, the regular sum-of-least-squares formulation. It should be noted that the runs were repeated three times and OF-3 performed better in all of those. As a result, modifications in the regular sum-of-least squares formulation resulted in a better outcome, at least in the convergence rate.

Figure 3 compares K_4 values. As can be seen, the positive impact of OF-3 in GA search was not distinguishable for this parameter. Actually, all optimization models with different objective functions indicated a demand at Reach 2, which was not valid. However, for the river system studied, the impact of K_4 on DO concentrations was not as significant as the impact of K_2.

Copyright ASCE 2005 EWRI 2005
Downloaded 23 Mar 2011 to 194.27.72.122. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

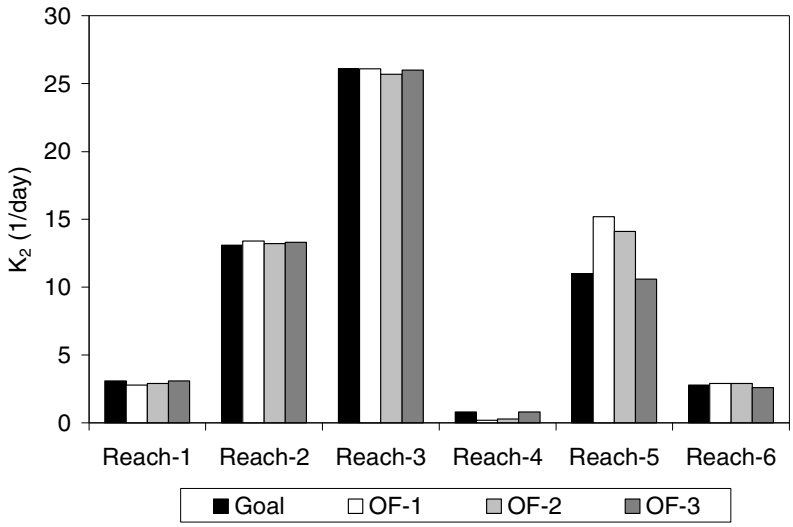


Figure 2. K_2 values obtained for different objective functions

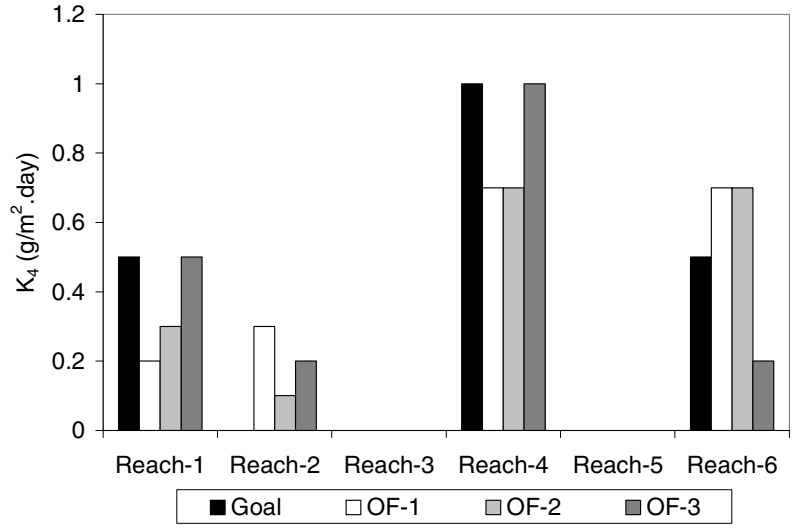


Figure 3. K_4 values obtained for different objective functions

Following the determination of K_2 and K_4 values for different objective functions, DO profiles were simulated. The better performance of OF-3 was also reflected in these profiles. The maximum errors in DO concentrations within the whole river system were 0.9 mg/L, 0.7 mg/L, and 0.02 mg/L for OF-1, OF-2, and OF-3, respectively.

Runs were also performed for biased observation data that contained up to 20% error. The impact of the objective function formulations on the performance of optimization diminished with the error in the observation data. The quality of the

Copyright ASCE 2005 EWRI 2005
Downloaded 23 Mar 2011 to 194.27.72.122. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org

data, rather than the formulation itself became a factor in reaching to the goal parameter values.

Conclusion

In this study, the impact of different objective function formulations on the calibration and verification of QUAL2E Model with GA optimization is presented. It was seen that the formulation itself can have a significant impact at least on the convergence rate. However, the extent of this effect is also related with the observation data quality that leads the search.

References

Brown, L.C., Barnwell, T.O. (1987). *The enhanced stream water quality models Qual2e and Qual2e-Uncas: Documentation and user manual*. EPA Office of Research and Development.

Carroll, D.L. (1999). "FORTRAN genetic algorithm driver in Fortran", version 1.7.1. http://cuaerospace.com/carroll/ga.html.

Cooper, V.A., Nguyen V.T.V. and Nicell, J.A. (1997). "Evaluation of global optimization methods for conceptual rainfall-runoff model calibration." *Water Science and Technology*, 36 (5), 53-60.

Goktas, R.K., and Aksoy, A. (2004). "Application of genetic algorithms for calibration and verification of QUAL2E Model" *Proceedings of World Water and Environmental Resources Congress 2004*, Sehlke, G., Hayes, D.F., and Stevens, D.K., eds. June 27 – July 1, 2004, Salt Lake City: ASCE.

Goldberg, D.E. (1989). *Genetic algorithms in search, optimization and machine learning*. Massachusetts: Addison Wesley Longman.

Little, K.W., and Williams, R.E. (1992). "Least-squares calibration of QUAL2E." *Water Environment Research*, 64 (2), 179-185.

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

Ng, A.W.M., and Perera, B.J.C. (2003). "Selection of genetic algorithm operators for river water quality model calibration." *Engineering Applications of Artificial Intelligence*, 16(5-6), 529-541.

Onur, A.K., Ekemen, E., Soyupak, S., Yurteri, C. (1999) "Management strategies for the Lower Seyhan catchment." *Water Science and Technology*, 40 (10), 177-184.

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

Vassilios, A., Fuentes, H.R., and Rodriquez, L. (1995). "Calibration and verification of QUAL2E water quality model in sub-tropical canals." *International Water Resources Engineering Conference - Proceedings*, v1, pp 214-218.

Yih, S.M., and Davidson, B. (1975). "Identification in nonlinear, distributed parameter water quality models." *Water Resources Research*, 11 (5), pp 693-704.

Copyright ASCE 2005 EWRI 2005

Downloaded 23 Mar 2011 to 194.27.72.122. Redistribution subject to ASCE license or copyright. Visithttp://www.ascelibrary.org