Optimization of an Ecosystem Model Simulating Water Quality in a Eutrophic Pond using Simple Genetic Algorithm

Do Thuy Nguyen*, Harada Masayoshi**, Hiramatsu Kazuaki**

1. Introduction

Ecosystem model approach coupling population models with biochemical models is an effective tool to provide insight into the relationship between water quality and its biological components in standing waters. Such kind of model commonly has a large number of parameters, of which values are inevitable to be well optimized within the reference range in order to yield a good description and prediction of the actual water quality. Therefore, application of effective optimization techniques on handling the model calibration and searching the optima of parameter value may be considered a promising solution to water quality modeling research. This study presents an attempt to calibrate an ecosystem model simulating water quality in a eutrophic pond using a Simple Genetic Algorithm (SGA) method.

2. Materials and methods

The study was conducted in Okubo pond, an agricultural pond located in Itoshima peninsula. An ecosystem model has been constructed and calibrated with four-month water quality data from August to December 2008 (Do *et al.* 2008). Results of sensitivity analysis tested on simulated data at that time suggested a few parameters imposing the most effect on model performance. By using this sensitivity analysis results and the same model construction, the present research tried the SGA method optionally to explore optimal solutions for the parameters to which the model was most sensitive. The final purpose aimed to achieve desirable calibration for the year 2009 water quality data lasting from June to December.

The model is a one-box ecosystem model obtaining 2 control variables including solar radiation, water temperature and 14 state variables consisting of four algal groups (green algae, blue green algae, diatom dinoflagellates and cryptophytes), zooplankton, total organic carbon (TOC), dissolved organic carbon (DOC), total nitrogen (TN), total phosphorus (TP), ammonia (NH_4^+-N), nitrite (NO_2^--N), nitrate (NO_3^--N), phosphate (PO_4^{3-}) and dissolved oxygen (DO). Simulated results were calibrated by 181-day measured field data. Relative mean squared error ($RMSE_{m,g}$) was used to estimate the simulated results of the model. Totally, 105 parameters were involved in the model but only 48 most sensitivity parameters were adopted to be optimized by the SGA, the remaining were set to the values of parameters in previous model. Each selected model parameter was then encoded in the form of a binary sub-string consisting of “0” and “1”. The length of a sub-string was assigned to 4 bit-length. A population size of 50 strings (individuals) was randomly initialized and the fitness value ($Fitness_{m,g}$) for each string was then calculated as below:

$$RMSE_{m,g} = \sum_{k=1}^{K}\left(\frac{1}{N}\sum_{n=1}^{N}\frac{\left(M_{n,k,m,g} - S_{n,k,m,g}\right)^2}{M_{n,k,m,g}}\right) \quad ; \quad Fitness_{m,g} = \frac{1}{RMSE_{m,g}} \; ; Fitmax_g = \max\left(Fitness_{m,g}\right)$$

In which, M= measured data, S= simulated data, n= calculated time n^{th} of the model, k= state variable k^{th}, m= individual m^{th}, and g= generation g^{th} of the model.

The overall quality of the string was indicated by the magnitude of this fitness value. The optimization process was designed to generate a better quality string reproduced by exchanging and randomly mutating some bits between strings via two operators: crossover and mutation with the probability of 0.5 and 0.05, respectively. Otherwise, good strings were preserved by elitism mechanism. Each generation, the greatest values ($Fitmax_g$) among $Fitness_{m,g}$ or the “best fit” parameters were investigated. This process was iterated until the stop criterion, which was set as number of iteration reached 100^{th}.

3. Results and discussions

There were several indicators of eutrophication threatening the water quality of Okubo pond last year (Do *et al.* 2008), however the observation within six months in 2009 indicated water quality just under eutrophic status. The concentration of total nitrogen, total phosphorus and nutrient ion in the water was recorded mostly low value compared to those at the same time last year. Three algal groups: blue green algae, diatom and dinoflagellates and cryptophtes hardly could bloom throughout the period meanwhile green algae appeared to be likely the most dominant group strongly increasing its flourish ability during June to August before falling

* Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University **Faculty of Agriculture, Kyushu University
Keywords: Eutrophication, Agricultural reservoir, Dynamic of water qualities, Algal groups, Numerical simulation

Table 1 Model parameters optimized by the SGA method

Group	Parameter	G	B	D	C	Unit
1	Optimum temperature for algal growth	28.7	27.3	11.3	22	°C
	Effect of a temperature above optimum temperature for algal growth	0.001	0.014	0.008	0.034	none
	Effect of a temperature below optimum temperature for algal growth	0.034	0.014	0.014	0.041	none
2	Maximal growth rate of algae at 20°C	2.5	1.9	1.9	2.2	day^{-1}
	Respiratory rate of algae at 20°C	0.1	0.06	0.02	0.02	day^{-1}
	Phosphorus half saturation for algae	1.7	2.2	3.0	2.5	mg/m^3
3	Settling velocity for algae	0.012	0.074	0.1624	0.049	m/day
	Chlorophyll/Carbon ratio of algae	0.020	0.029	0.021	0.029	none
	Phosphorus/Carbon ratio of algae	0.017	0.005	0.010	0.006	none
	Nitrogen/Carbon ratio of algae	0.3	0.19	0.15	0.14	none
4	Nitrogen/Carbon ratio of zooplankton	0.15				none
	Phosphorus/Carbon ratio of zooplankton	0.015				none
	Nitrogen/Carbon ratio of POC	0.14				none
	Phosphorus/Carbon ratio of POC	0.035				none
	Nitrogen/Carbon ratio of DOC	0.00848				none
	Phosphorus/Carbon ratio of DOC	0.0043				none
	Sediment exchange flux of PO_4^{3-} at 0°C	0.19				day^{-1}
	Sediment exchange flux of NH_4^+ at 0°C	6.5				day^{-1}

G= Green algae, B=Blue green algae, D=Diatom and dinoflagellates, C=Cryptophytes

down until the end of the year.

Table 1 shows all parameters optimized by SGA. First trial of the SGA which set all 48 parameters to 1 string didn't lead to expected result. In the second trial, we divided 48 parameters into 4 groups, each one consisted of 12 parameters and was encoded by 1 string, of which length was set to 48 bits. By ranking in descending order of sensitive magnitude of each parameter, four groups were optimized in turn by the SGA (from group 1 to group 4 in **Table 1**). In this case, value of $Fitmax_g$ was improved significantly from 0.054 at the first generation to 0.18 at the final generation and the results of validation suggested the model performed well with the calibrated parameters. As can be seen in **Fig. 1**, results of the calibration presented a good fit between observed and predicted values. Relative mean squared error was generally low and the simulated trend also followed the similar way of actual water quality. However, at the end of the period, there was slightly raising of chlorophyll-a concentration, which the model found it difficult to simulate. Moreover, the concentration of dissolved oxygen and total phosphorus in simulation results showed few differences with the observation.

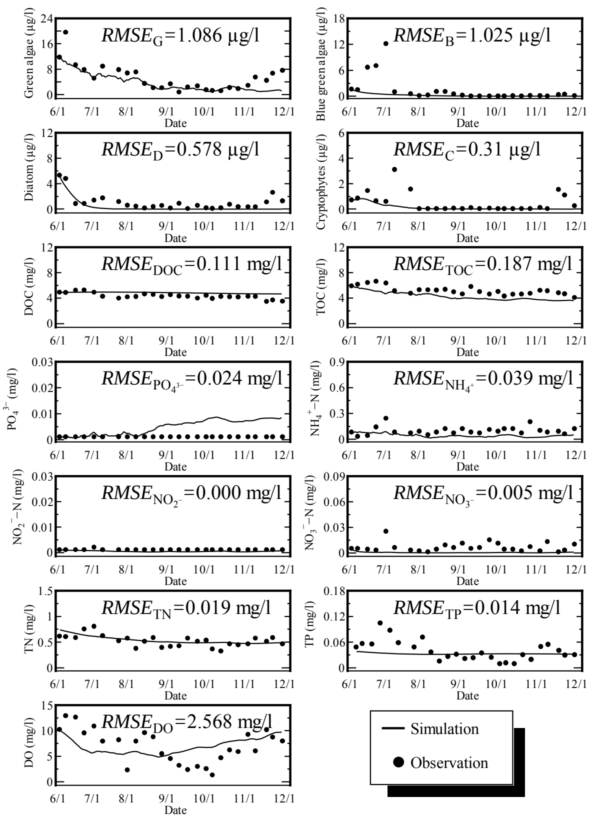

Fig. 1 Observed and simulated results in the Okubo pond

4. Conclusion

As validated by the 2009 water quality data in Okubo pond, the SGA method showed a significant increase of fitness value and it proved itself as a great solution for searching the good optima for parameters in case of numerous parameters ecosystem model.

Acknowledgement

This research is partly supported by the JSPS Grant-in-Aid for Scientific Research (B) (Project number: 19380138) and Kyushu University New Campus Planning Office.

Reference:

Do, T. N., Harada, M. and Hiramatsu, K.: Numerical Simulation of Water Quality in an Agricultural Pond using a One-box Ecosystem Model, *Proceedings of the 6th Japan-Korea International Joint Symposium*, CD-Paper No.WS2A-1KU 2009