



OPTIMAL ESTIMATION FOR KEY PARAMETERS OF THE MARINE QUALITY MODEL USING DATA-DRIVEN NEURAL NETWORK

Ming-Chang Li

Laboratory of Environmental Protection in Water Transport Engineering, Tianjin Research Institute of Water Transport Engineering, Tianjin 300456, China, lmcsq1997@163.com

Shu-Xiu Liang

State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116024, China.

Zhao-Chen Sun

State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116024, China.

Guang-Yu Zhang

Laboratory of Environmental Protection in Water Transport Engineering, Tianjin Research Institute of Water Transport Engineering, Tianjin 300456, China.

Follow this and additional works at: <https://jmstt.ntou.edu.tw/journal>



Part of the [Civil and Environmental Engineering Commons](#)

Recommended Citation

Li, Ming-Chang; Liang, Shu-Xiu; Sun, Zhao-Chen; and Zhang, Guang-Yu (2010) "OPTIMAL ESTIMATION FOR KEY PARAMETERS OF THE MARINE QUALITY MODEL USING DATA-DRIVEN NEURAL NETWORK," *Journal of Marine Science and Technology*. Vol. 18: Iss. 5, Article 19.

DOI: 10.51400/2709-6998.1933

Available at: <https://jmstt.ntou.edu.tw/journal/vol18/iss5/19>

This Research Article is brought to you for free and open access by Journal of Marine Science and Technology. It has been accepted for inclusion in Journal of Marine Science and Technology by an authorized editor of Journal of Marine Science and Technology.

OPTIMAL ESTIMATION FOR KEY PARAMETERS OF THE MARINE QUALITY MODEL USING DATA-DRIVEN NEURAL NETWORK

Acknowledgements

This work was supported by National Nonprofit Institute Research Grant of TIWTE (TKS090204, TKS100217). Authors also deeply appreciate the two anonymous reviewers and editors. Under their suggestions, the paper is improved greatly

OPTIMAL ESTIMATION FOR KEY PARAMETERS OF THE MARINE QUALITY MODEL USING DATA-DRIVEN NEURAL NETWORK

Ming-Chang Li*, Shu-Xiu Liang**, Zhao-Chen Sun**, and Guang-Yu Zhang*

Key words: water quality model, parameter estimation, data-driven model, near-optimal prediction.

ABSTRACT

Marine water quality models are complicated because of their multi-parameter and multi-response characteristics. One major difficulty with water quality models is the accurate estimation of model parameters. In this paper, a new method based on a data-driven model (DDM) is developed to retrieve the value of model parameters. All training data are calculated by numerical water quality models from results of multi-parameter matching design cases so the physical properties are not disturbed. The concept is to find the relationship between model parameters and the pollution concentration values of interior stations. Field data are imported into the relationship for inverting optimal parameters or near-optimal parameters, ultimately an optimal or near-optimal prediction method is applied to validate the long-term stability of inversion results. Case tests were carried out in the Bohai Sea, China. Chemical oxygen demand (COD), dissolved inorganic nitrogen (DIN), chlorophyll (Chl) and their sensitive parameters were considered for validating the present method. The optimal solution determination method is applied for DIN and Chl owing to existence of the same sensitive parameters. Case studies show that the present method can make a more satisfactory estimation for this practical problem.

I. INTRODUCTION

Substantial uncertainties exist in marine ecosystems [21], which mainly have originated from a scarcity of understanding of biochemical processes, the values of pollution source, pollution background field, field data and model parameters [4, 19].

Paper submitted 08/26/09; revised 12/03/09; accepted 01/04/10. Author for correspondence: Ming-Chang Li (e-mail: lmcsq1997@163.com).

*Laboratory of Environmental Protection in Water Transport Engineering, Tianjin Research Institute of Water Transport Engineering, Tianjin 300456, China.

**State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116024, China.

One of the most significant difficulties is the estimation of model parameters for model calibration. Since Shastry *et al.* [17] estimated the parameter of the BOD-DO model, many estimation methods have been developed [1, 2, 11, 13, 24-26]. Trial and error [5] is a widely used technique for model calibration [16]. How close the model comes to real conditions depends on the abundance of field data. Data assimilation methods have been employed for model calibration with the abundance of satellite data, with the adjoint technique being the most widely used one. Lawson *et al.* [9] introduced the adjoint method for data assimilation in a simple predator-prey model, the Lagrange operator method is employed to construct the adjoint equation, model parameter and initial field conditions were estimated successfully. Subsequently Lawson *et al.* [8] used this method into a complex marine ecosystem model that includes five state variables. Performing optimizations with synthetically produced data, they investigated necessary sampling rates to recover the model's parameters values. A large effort by Vallino [20] tested the ability of various data assimilation methods to incorporate mesocosm experiment data into a marine ecosystem model and indicated numerical instability of adjoint approach. A NPZD model with its adjoint equations was applied in the Bohai Sea and the Yellow Sea by Xu *et al.* [22, 23]. Although optimal estimation for parameters, initial and boundary conditions can be obtained by this technique, both water quality model equations and adjoint equations need to be calculated. Adjoint equations are as complicated as water quality model equations, so much time is needed for calibration. When an adjoint technique is chosen for practical engineering, many uncertainties originating from the absence of field data might exist in the calibrated model because of the investment limitation. The application of a data assimilation method is also limited [12].

This paper aims to develop a more practical technique for optimal estimation of model parameters. In the technique, optimal parameters are estimated by a data-driven model [18] based on artificial neural network. The Osaka Daigaku Estuary Model [14] is employed to simulate the marine water quality in interested area.

The structure of this paper is as follows. In Section II, the basic idea and theory of the data-driven model, Back-Propagation Neural Network (BPNN) and water quality model

(ODEM) are introduced. The detailed steps about how to inverse model parameters are described in Section III. In Section IV, the method in Section III is verified with an entitative ocean. In Section V, the conclusions are made.

II. NUMERICAL MODEL

1. Data-Driven Model

So-called data-driven models are different from knowledge-driven models (physically based model) because they are based on limited knowledge of the modeling process and rely purely on the data describing input and output characteristics. They make abstractions and generalizations of the process, and so often play a complementary role to physically based model. Data-driven models can use results from artificial neural networks (ANN), expert systems, fuzzy logic concepts, rule-induction and machine learning systems. The fundamental expression is as follows:

$$(y_1, \dots, y_i, \dots, y_m) = F(x_1, \dots, x_i, \dots, x_n) \quad (1)$$

$(x_1, \dots, x_i, \dots, x_n)$ and $(y_1, \dots, y_i, \dots, y_m)$ are the input and output variables respectively, F is the objective function that needs to be dug by the model. In this paper, the results of ANN are used for the fitting of F .

2. Back-Propagation Neural Network

The BPNN proposed by Rumelhart *et al.* [15] is the most commonly used among the artificial neural networks. The BPNN uses the gradient steepest descent method to determine the weights of connective neurons. The key point is the error back propagation technique. In the learning process of the BPNN, the interconnection weights are adjusted from back layers to front layers to minimize the output error. The merit of the BPNN is that it can approach any nonlinear continuous functions after being trained [7]. Some of the numerical details of the BPNN are described in the following section.

1) Data Normalization

All of the input and output layers data are normalized to a range from 0 to 1 by function (2):

$$\bar{Y}_i = \frac{Y_i - Y_{\min}}{Y_{\max} - Y_{\min}} \quad (2)$$

\bar{Y}_i is the value of data after normalization, Y_i is the value of data before normalization, Y_{\max} is the maximum of all data, and Y_{\min} is the minimum of all data.

To consider the nonlinearity, the sigmoid transfer function is used:

$$f(x) = \frac{1}{1 + e^{-x+\theta}} \quad (3)$$

θ is the threshold value of hidden neurons.

2) Learning Rate η and Appended Momentum

The efficiency and speed of convergence of the BPNN learning algorithm are affected by the learning rate η and appended momentum (L&A). L&A are control parameters of BPNN training algorithms, which control the step size when weights are iteratively adjusted. An L&A that is too low makes the network learn at a very slow pace. An L&A that is too high makes the weights and objective function diverge, therefore, learning is nil. The value of L&A depends on whether the time series change too much. If they do, the value of L&A should be increased. L&A range from 0 to 1. The learning rate is set to be 0.05 and the appended momentum is 0.5 in present research after testing.

3) Error Function

The root mean squared error (RMSE) is used to evaluate the accuracy of prediction and is defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \tilde{Y}_i)^2}{\sum_{i=1}^n Y_i^2}} \quad (4)$$

n is the number of data, Y_i is the value of the field data, and \tilde{Y}_i is the value predicted by the neural network.

4) Structure of the Neural Network

The structure of the neural network, including the number of hidden layers and neurons, is determined by the complexity of the problem to be solved. Although the increase in the number of hidden layers and neurons can reflect the complexity of the problem and decrease the number of iteration steps, it is not helpful in increasing precision and maybe lead to over-fitness. Therefore, throughout the present study, one hidden layer is chosen.

The number of neurons in the hidden layer can be calculated by:

$$\text{NNH} = \frac{\text{NNI} + \text{NNO}}{2} \quad (5)$$

Where NNH is the number of neurons in the hidden layer, NNI is the number of neurons in the input layer, and NNO is the number of neurons in the output layer. Detailed information can be found in Li *et al.* [10].

In this paper, NNI is the number of output data by hours in the water quality model and NNO is the number of sensitive parameters.

3. Water Quality Model

A water quality model is fundamental because it attempts to explain underlying physical processes. There are many water quality models that can retrieve entitative ocean conditions

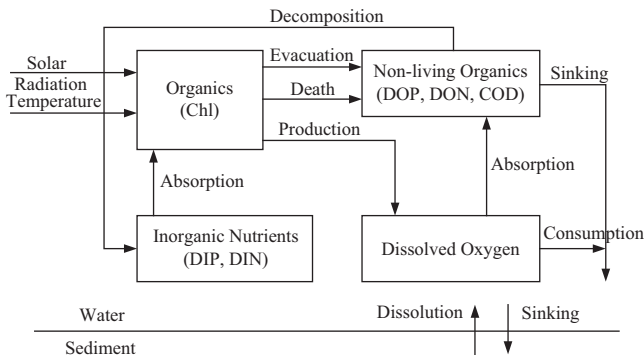


Fig. 1. Schematic diagram of matter circulation in the ODEM.

with great precision. ODEM developed by Nakatuji is one of them. ODEM has been used not only for water quality modeling but for coastal waters studies.

Figure 1 shows the schematic diagram of matter circulation in ODEM. The main processes in the model are that phytoplankton release nonliving organics (organic nitrogen, organic phosphorus, COD) and dissolved oxygen by evacuation, death and production. Parts of nonliving organics sink into sediment with detritus, and others are decomposed into inorganic nutrients combined with dissolved oxygen. The phytoplankton growth is controlled by inorganic nutrients, temperature and solar radiation.

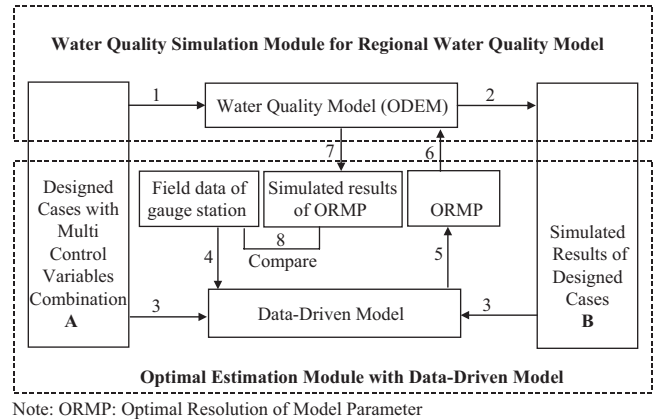
III. OPTIMAL ESTIMATION FOR MODEL PARAMETERS

1. Optimal Estimation Method

Tens or even hundreds integral computations have been repeated in the process of trial and error for the estimation of model parameters. The results, however, are only approximate. In this paper, a new technique was developed that combines a data-driven model with a water quality model automatically. In the technique, the water quality model repeats a series of designed computations. Then, a data set that contains the corresponding relationship between model parameters' values $[(x_1, \dots, x_i, \dots, x_n)]$ in (1) and the values of interior stations for pollution (state variables) $[(y_1, \dots, y_i, \dots, y_m)]$ in (1) are stored. The task of the data-driven model is to find the relationship $[F$ in (1)] between $(x_1, \dots, x_i, \dots, x_n)$ and $(y_1, \dots, y_i, \dots, y_m)$. After the field data are transported into the model, optimal model parameter values will be inverted. The detailed technique is as follows:

Step 1: Choice of control variables

There are many parameters in the water quality model and they act with each other. If all of them are included, the computation cost is excessive and uncertainty increases [3], therefore sensitive model parameters have been analyzed that aim to select control variables [6].



Note: ORMP: Optimal Resolution of Model Parameter

Fig. 2. Diagram of optimal estimation of model parameter using data-driven model combined with water quality model.

Step 2: Cases computation by the water quality model

In water quality numerical models, the governing equations have to be discretized into computation domain. In this paper, initial guess values for all the control variables are assumed and their corresponding ranges are set. If the number of control variables is m , and n values are taken for one control variable, there are as many as $\prod_{i=1}^m [C_n^1]_i = n^m$ designed cases.

All the designed cases are computed by ODEM one by one. The results of pollution concentration are output and stored for the data-driven model.

Step 3: Model parameters' estimation by the data-driven model

The results of pollution concentrations at interior stations and their corresponding parameters' values are input into data-driven model. After training, the relationship of interior stations and model parameters is generalized.

Input the field data of interior stations into the above relationship and obtain the optimal solution.

Step 4: Verification of optimal solution

Input the optimal solution into the water quality model and repeat the computation. The RMSE between measurement and results of numerical computation is calculated.

In Fig. 2, the process of model parameter estimation is described. The sequence is one through eight. In the entire process, there are two modules—water quality and optimal estimation. The computation of the designed cases and the final verification are finished by the water quality module. The optimal estimation module is responsible for the analysis of water quality model results and generalization of the relationship between model parameters and interior stations. A and B comprise the database of the two modules.

2. The Optimal Solution Determination Method

The marine water quality model has the characteristics of being multi-parameter and multi-response. Multi-state variables may have common sensitive parameters, so the inverse

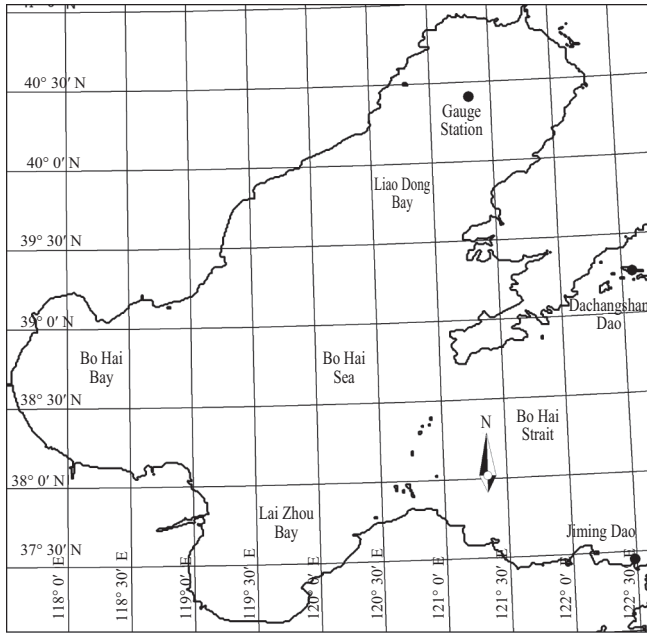


Fig. 3. The location of computation domain.

problem has a characteristic of being multi-solutions, which are near-optimal solutions for the water quality model. Equation (6) is used to obtain a model optimal solution based on multi-near-optimal solutions, in which state variables practical engineering attention extent and sensitivity extent for parameters are applied for calculating weight factor q .

$$[v_1^o, \dots, v_j^o, \dots, v_n^o] = \sum_{i=1}^m q_i \times [v_{1i}, \dots, v_{ji}, \dots, v_{ni}] \quad (6)$$

In (6), n is the number of sensitive parameters, m is the number of state variables, v_j^o is the optimal solution of sensitive parameter j and v_{ji} is the i near-optimal solution of the sensitive parameter j .

The optimal solution solving through optimization theory cannot be the real solution of the inversion problem, so a near-optimal and optimal prediction method is applied to verify the long-term stability of solutions.

IV. CASE STUDY

Case tests were carried out in the Bohai Sea, China, which is a semi-closed sea with a mean depth of 18.7 m and an area of more than 80000 km². The sea is divided into four parts: Laizhou Bay, Bohai Bay, Liaodong Bay and the central part. Its bottom is very flat with average slope is 28°. Figure 3 shows the location of computation domain.

In the numerical simulation, the area was discretized as 4 km × 4 km horizontally and 17 levels in depth. To save computational time and improve accuracy, the level thickness varied non-uniformly in water depth. The level thickness was

Table 1. State variables and sensitive parameters.

State variable	Sensitive parameters	Unit
COD	Non-organism settlement velocity (NOSV)	1/d
	Decomposed rate of COD (DRCD)	1/d
DIN	Fastest growth rate of phytoplankton (VMAX)	1/d
	Most suitable growth temperature of phytoplankton (TEMPS)	°C
	Decomposed rate of DIN (DRDN)	1/d
Chl	Fastest growth rate of phytoplankton (VMAX)	1/d
	Most suitable growth temperature of phytoplankton (TEMPS)	°C

Table 2. Values of sensitive parameters.

State variable	Sensitive parameter	Values		
COD	NOSV	0.18	0.3	0.42
	DRCD	0.012	0.02	0.028
DIN Chl	VMAX	1.44	2.5	3.36
	TEMPS	15	25	35
	DRDN	0.012	0.022	0.028

4 m × 1, 2 m × 4, 3 m × 5, 4 m × 3, 5 m × 2, 6 m × 2 from top to bottom respectively. The model time step is 30s. The surface height along the open boundary is given by interpolating results between Dachen Dao (39°16'N, 122°35'E) and Jimo Dao (37°27'N, 122°35'E). Five main tidal constituents M₂, S₂, K₁, O₁, and N₂ are input in two open boundary control stations. The gauge station (40°24'31"N, 121°19'12"E) locates in the Liaodong Bay.

1. Choices of Control Variable

In this paper, COD, DIN and Chl as state variables are considered to verify the present optimal estimation method. Coefficient of variation is computed by the Monte Carlo method [6] to validate the sensitivity of model parameters. State variables and its sensitive parameters are listed in Table 1.

In Table 1, DIN and Chl have the same two sensitive parameters, so optimal solution determination method is used for VMAX and TEMPS. Control variable corresponding ranges are set among initial guess values in Table 2.

In Table 2, the values of control variables are listed. For each control variable, 3 values are taken in its range, 9 designed cases are obtained for COD and 27 designed cases for DIN and Chl.

2. Optimal Estimation

Nine or 27 cases of pollution concentration data are acquired by 9 or 27 times model computation of 80 hours. After inputting the 9 or 27 COD, Chl and DIN data of 80 hours and

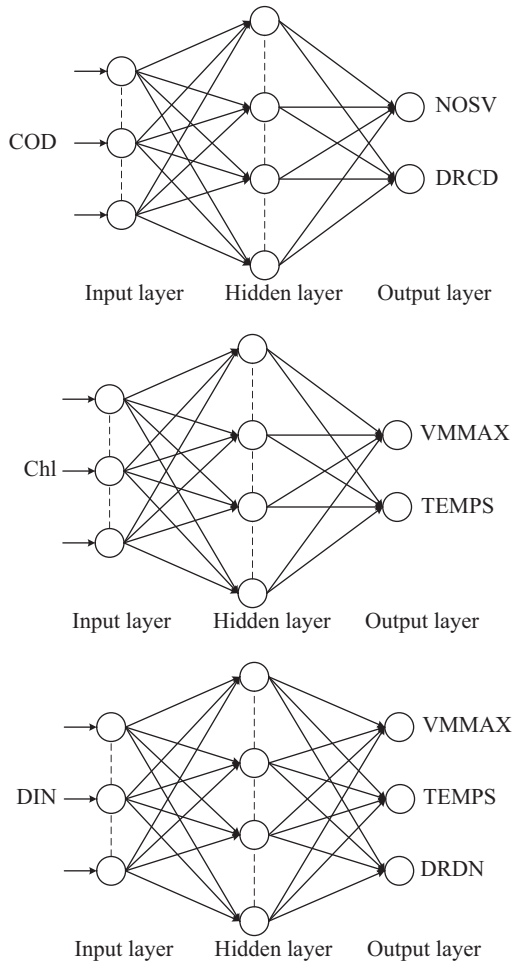


Fig. 4. Network structure for data-driven model.

relative sensitive parameters data in Table 2 into the data-driven model, the relationship is generalized. Figure 4 shows the network structure for the data-driven model, and the number of neurons in the three layers is shown in Table 3.

In this paper, the so-called “twin experiment” method [3] is used to verify the efficiency of optimal estimation method.

The parameters in Table 4 are as real values and were input into ODEM for the pollution concentration data as pseudo-field data. The optimal model parameters in Table 5 are inverted by inputting the pseudo-field data into the above relationship.

The optimal and near-optimal solutions are listed in Table 5. Because VMAX and TEMPS are common sensitive parameters of DIN and Chl, there are two near-optimal solutions. Their optimal solution (OS) can be obtained by Eq. (6).

3. Verification of Optimal Solution

Pseudo-field data in the validation part are transported into the relationship in step 3 for model parameters and used to verify the optimal solution as step 4. Pseudo-field data in the forecasted part are used to validate the long-term stability of the optimal solution for water quality model.

Table 3. The number of neurons.

Network	Neurons' number		
	Input layer	Hidden layer	Output layer
COD	80	41	2
Chl	80	41	2
DIN	80	42	3

Table 4. Design of twin experiment.

State variable	Sensitive parameter		
	NOSV	DRCD	
COD	0.2	0.015	
DIN	VMAX	TEMPS	DRDN
Chl	2.4	26	0.02

Table 5. Optimal and near-optimal solution of control variables.

State variable	Sensitive parameter		
	NOSV	DRCD	
COD	0.1881	0.018	
	VMAX	TEMPS	DRDN
DIN	2.533114	26.56910	0.02002
Chl	2.537788	26.93573	
OS	2.535451	26.752415	0.02002

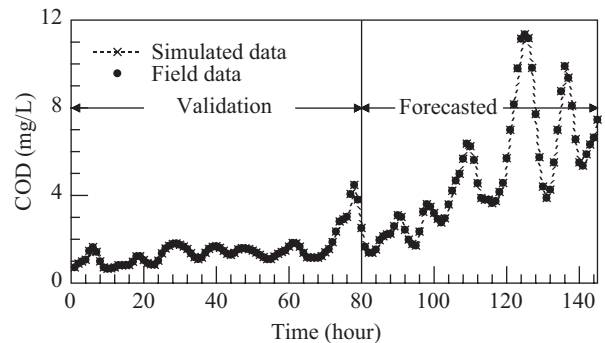


Fig. 5. The validation and forecasted results of COD optimal solution.

COD and its two sensitive parameters are verified in Fig. 5, the validation and forecasted data are closed to field data and the RMSE is 0.000625. Figure 6 is the corresponding scatter plot of Fig. 5. The correlation coefficient computed by (7)-(9) is 0.9999 in Fig. 6.

$$CC = \frac{\sum(\eta_o - \bar{\eta}_o)(\eta_p - \bar{\eta}_p)}{\sqrt{\sum(\eta_o - \bar{\eta}_o)^2 \sum(\eta_p - \bar{\eta}_p)^2}} \tag{7}$$

$$\bar{\eta}_o = \frac{\sum \eta_o}{n} \tag{8}$$

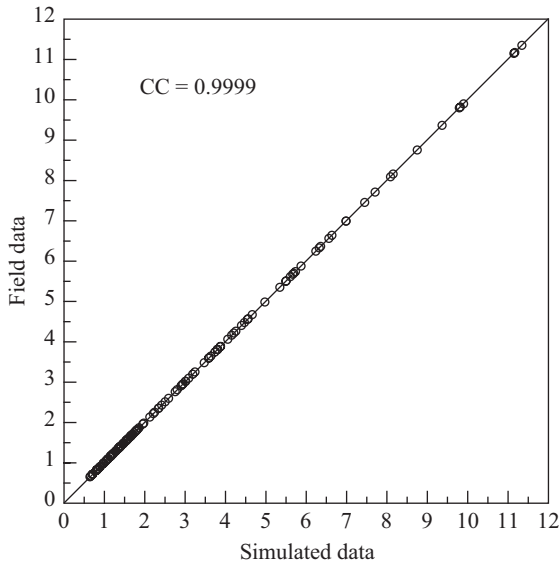


Fig. 6. The validation and forecasted correlation coefficient results of COD optimal solution.

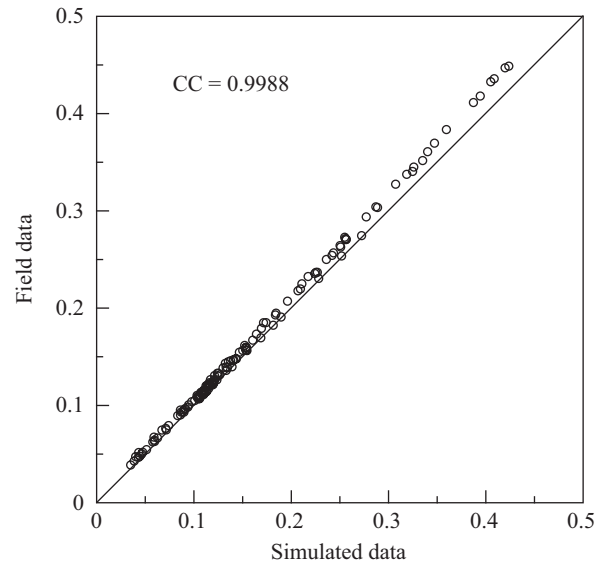


Fig. 9. The validation and forecasted correlation coefficient results of DIN optimal parameter (simulated data 2).

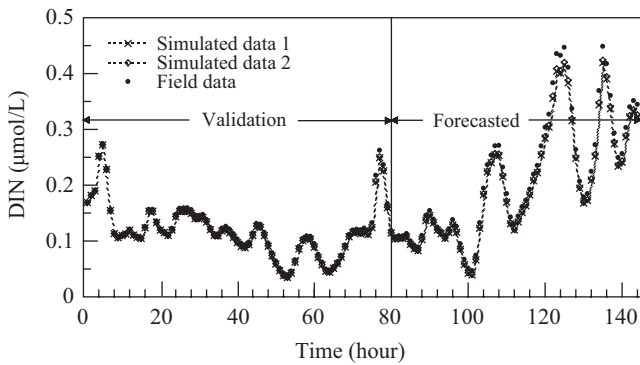


Fig. 7. The validation and forecasted results of DIN optimal parameter.

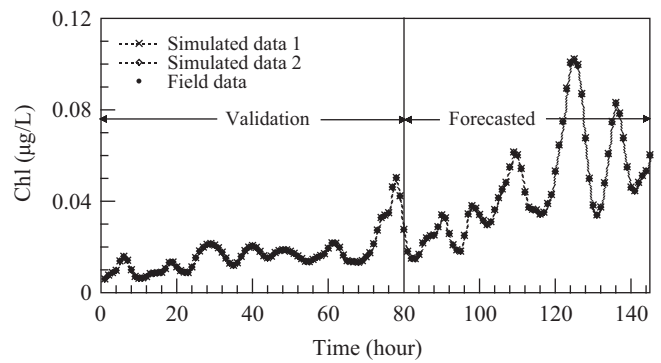


Fig. 10. The validation and forecasted results of Chl optimal parameter.

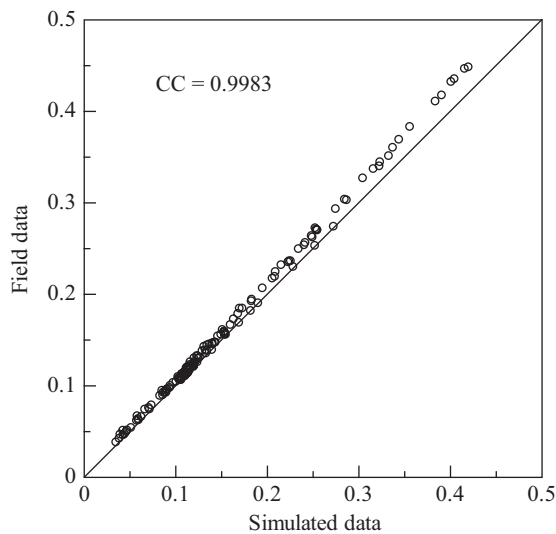


Fig. 8. The validation and forecasted correlation coefficient results of DIN optimal parameter (simulated data 1).

$$\bar{\eta}_p = \frac{\sum \eta_p}{n} \tag{9}$$

Where CC is the correlation coefficient, η_o is field data, η_p is simulated data, and n is the number of data. The over bar indicates the mean value.

Two near-optimal solutions of VMMA and TEMPS are imported into ODEM for verification of DIN and Chl, near-optimal solution inverted by Chl are computed for simulated data 1 and DIN for simulated data 2. Therefore for DIN, the precision of simulated data 2 is superior to simulated data 1 in Figs. 7-9, but the precision of simulated data 2 is same as simulated data 1 for Chl in Figs. 10-12. This is due to parameter sensitivity of VMMA and TEMPS for DIN being higher than for Chl.

The OS of VMMA and TEMPS obtained by (6) is validated in Figs. 13-16, the simulated accuracy indicates the present optimal estimation method can inverse realistic model

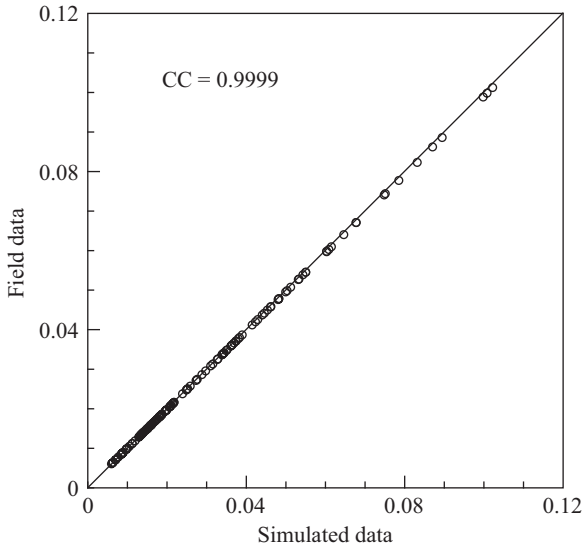


Fig. 11. The validation and forecasted correlation coefficient results of Chl optimal parameter (simulated data 1).

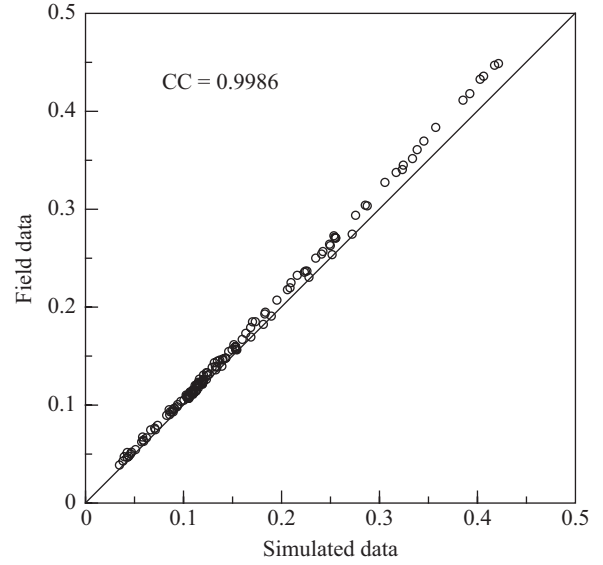


Fig. 14. The validation and forecasted correlation coefficient results of DIN optimal solution.

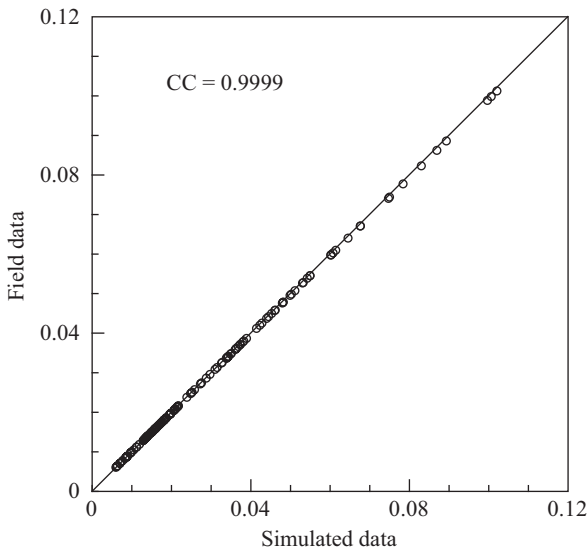


Fig. 12. The validation and forecasted correlation coefficient results of Chl optimal parameter (simulated data 2).

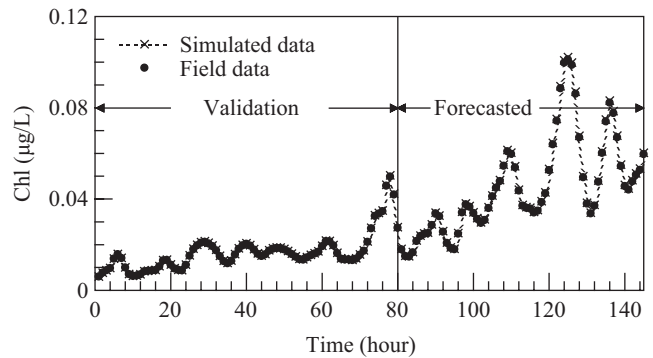


Fig. 15. The validation and forecasted results of Chl optimal solution.

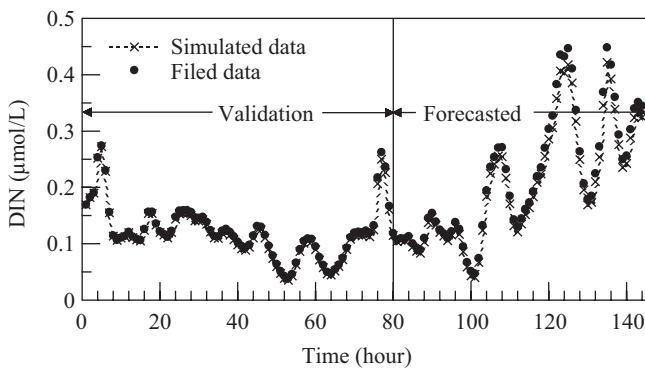


Fig. 13. The validation and forecasted results of DIN optimal solution.

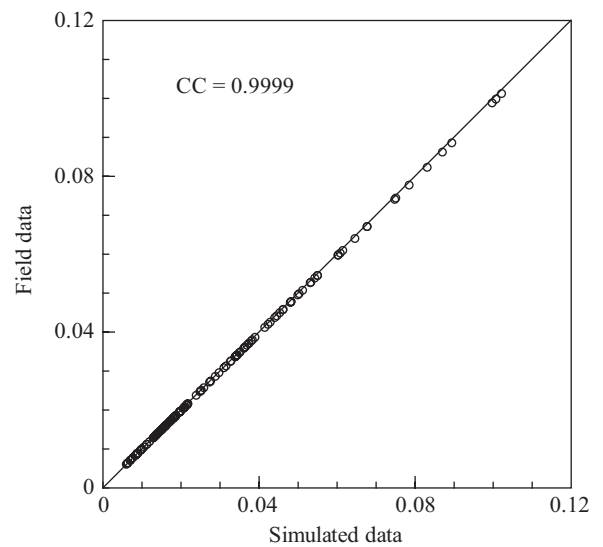


Fig. 16. The validation and forecasted correlation coefficient results of Chl optimal solution.

Table 6. Error measures.

Error measure	Fig. 5	Fig. 7		Fig. 10		Fig. 13	Fig. 15
		S1	S2	S1	S2		
RMSE	0.000625	0.059836	0.050869	0.011202	0.009514	0.05534	0.010358
CC	0.9999	0.9983	0.9988	0.9999	0.9999	0.9986	0.9999

parameters. The correlation coefficient for DIN is 0.9986 in Fig. 14 lies between 0.9983 in Fig. 8 and 0.9988 in Fig. 9, because of comprehensive consideration for every near-optimal solution in (6).

The RMSE and CC of Figs. 5, 7, 10, 13, and 15 are shown in the Table 6, S1 and S2 stand for the errors of simulation data 1 and simulation data 2 respectively.

Compared with the adjoint method, the present method has two superiorities. One is simplicity. There is no need to deduce and solve complicated adjoint equations. The data-driven model based on BPNN is easy to develop. The other is its flexibility. In the adjoint method, the different adjoint equations need to be deduced according to different numerical models. If the basic equations change, the adjoint equations need to alter accordingly. In present method, the data-driven model can be kept unchanged when different water quality models are used. The adjoint technique has to repeat the computation for both water quality model equations and adjoint equations. Even with good initial guesses, more time is consumed compared with the present method.

V. CONCLUSION

In this paper, a new method is developed to estimate model parameters' values. In this method, the data-driven model and water quality model are coupled automatically. Water quality model repeats a number of computations for designed cases, the results of pollution concentration data are output and stored for data-driven model. The data-driven model generalizes the relationship between model parameters and interior stations. After the field data are imported, optimal solutions are obtained.

In realistic case study, pseudo-field data of concentration and optimal solution determination method are both used to estimate the model parameters. A near-optimal and optimal prediction method is applied to verify long-term stability of multi-solutions. The results show the present estimation method is suitable for the inverse problem.

Compared with the adjoint method, the present method is simple, flexible and less time-consuming.

ACKNOWLEDGMENTS

This work was supported by National Nonprofit Institute Research Grant of TIWTE (TKS090204, TKS100217). Authors also deeply appreciate the two anonymous reviewers and editors. Under their suggestions, the paper is improved greatly.

REFERENCES

- Arhonditsis, G. B., Paerl, H. W., Valdes-Weaver, L. M., Stow, C. A., Steinberg, L. J., and Reckhow, K. H., "Application of Bayesian structural equation modeling for examining phytoplankton dynamics in the Neuse River Estuary (North Carolina, USA)," *Estuarine Coastal and Shelf Science*, Vol. 72, pp. 63-80 (2007).
- Fasham, M. J. R. and Evans, G. T., "The use of optimization techniques to model marine ecosystem dynamics at the JGOFS station at 47° N 20° W," *Philosophical Transaction of the Royal Society of London B*, Vol. 348, pp. 203-209 (1995).
- Friedrichs, M. A. M., "A data assimilation marine ecosystem model for the central equatorial Pacific: Numerical twin experiments," *Journal of Marine Research*, Vol. 59, pp. 859-894 (2001).
- Ganoulis, J. G., *Engineering Risk Analysis of Water Pollution*, Wiley-VCH, New York (1994).
- Gerritsen, H., De Vries, H., and Philippart, M., "The dutch continental shelf model. Quantitative skill assessment for coastal ocean models," *Coastal Estuarine Studies*, Vol. 47, pp. 425-468 (1995).
- Hakanson, Lars., "The role of characteristic coefficients of variation in uncertainty and sensitivity analyses, with examples related to the structuring of lake eutrophication models," *Ecological Modelling*, Vol. 131, pp. 1-20 (2000).
- Hornik, K., Stinchcombe, M., and White, H., "Multilayer feedforward networks are universal approximators," *Neural Networks*, Vol. 2, No. 5, pp. 359-475 (1988).
- Lawson, L. M., Hofmann, E. E., and Spitz, Y. H., "Time series sampling and data assimilation in a simple marine ecosystem model," *Deep Sea Research II*, Vol. 43, No. 2-3, pp. 625-651 (1996).
- Lawson, L. M., Spitz, Y. H., Hofmann, E. E., and Long, R. B., "A data assimilation technique applied to a predator-prey model," *Bulletin of Mathematical Biology*, Vol. 57, pp. 593-617 (1995).
- Li, M. C., Liang, S. X., and Sun, Z. C., "Application of artificial neural networks to tide forecasting," *Journal of Dalian University of Technology*, Vol. 47, No. 1, pp. 101-105 (2007). (in Chinese)
- Liu, J. G., Liu, J. H., and Guo, H. G., "Study on parametric recognition & comparison of improved S-P model," *Journal of Changchun Normal University*, Vol. 25, No. 6, pp. 16-21 (2006). (in Chinese)
- Ma, Z. P. and Jing, A. Q., "Data assimilation method applied in marine science—its significance, system configuration and development situation," *Coastal Engineering*, Vol. 24, No. 4, pp. 83-99 (2005). (in Chinese)
- Malve, O., Laine, M., Haario, H., Kirkkala, T., and Sarvala, J., "Bayesian modeling of algal mass occurrences using adaptive MCMC methods with a lake water quality model," *Environmental Modeling & Software*, Vol. 22, pp. 966-977 (2007).
- Nakastuji, K., Yamanaka, R., and Nishida, S., "Numerical simulation of seasonal baroclinic circulation and dispersion process of COD in the Bohai Sea," *The First Asian and Pacific Coastal Engineering Conference*, Dalian, China, pp. 369-378 (2001).
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J., "Learning representations by back propagating errors," *Nature*, Vol. 323, pp. 533-536 (1986).
- Seo, D.-II. and Canale, R. P., "Performance, reliability and uncertainty of total phosphorus models for lakes—I: deterministic analyses," *Water Research*, Vol. 30, pp. 83-94 (1996).
- Shastry, J. S., Fan, L. T., and Erickson, L. E., "Nonlinear parameter estimation in water quality modeling," *Journal of the Environmental Engineering Division, ASCE*, Vol. 99, No. 3, pp. 315-331 (1973).

18. Solomatine, D. P., "Data-driven modelling: Paradigm, methods, experiences," *Proceedings of the 5th International Conference on Hydroinformatics*, Cardiff, London, UK, pp. 757-763 (2002).
19. Takyi, A. K. and Lence, B. J., "Surface water quality management using a multiple realization chance constrained method," *Water Resources Research*, Vol. 35, No. 5, pp. 1657-1669 (1999).
20. Vallino, J. J., "Improving marine ecosystem models: use of data assimilation and mesocosm experiments," *Journal of Marine Research*, Vol. 58, pp. 117-164 (2000).
21. Wan, X. T., "Research and application on water quality monitor in rivers," *Jiangsu Environment Science and Technology*, Vol. 2, pp. 1-4 (1999). (in Chinese)
22. Xu, Q., Liu Y. G., Cheng, Y. C., and Lv, X. Q., "Adjoint method in marine ecosystem models," *Marine Science Bulletin*, Vol. 24, No. 6, pp. 58-64 (2005). (in Chinese)
23. Xu, Q., Liu, Y. G., Cheng, Y. C., and Lv, X. Q., "Adjoint assimilation technique in a marine ecosystem model in the Bohai Sea and the Yellow Sea: control variables and twin experiment," *High Technology Letters*, Vol. 16, No. 1, pp. 78-83 (2006). (in Chinese)
24. Yang, M. D., Sykes, R. M., and Merry, C. J., "Estimation of algal biological parameters using water quality modeling and SPOT satellite data," *Ecological Modeling*, Vol. 125, pp. 1-33 (2000)
25. Zhang, F. W., Chen, D. X., and Wang, E., "On new method and its significance of BOD-DO model parametric estimation," *Journal of Liaocheng University*, Vol. 8, No. 1, pp. 27-29 (2005). (in Chinese)
26. Zhang, M. Y., "Estimation method of oxygen consumption coefficient K1 in the water quality model," *Water Resources & Hydropower of Northeast*, Vol. 22, No. 250, pp. 42-44 (2005). (in Chinese)