Korean J. Chem. Eng., **29**(3), 291-303 (2012)
DOI: 10.1007/s11814-011-0178-2

A systematic model calibration methodology based on multiple errors minimization method for the optimal parameter estimation of ASM1

JungJin Lim*, MinHan Kim*, MinJung Kim*, TaeSuk Oh*, OnYu Kang*, Booki Min*, Ambati Seshagiri Rao, and ChangKyooYoo*,†**

*Department of Environmental Science and Engineering, Center for Environmental Studies, Kyung Hee University, Seocheon-dong 1, Giheung-gu, Yongin-si, Gyeonggi-do 446-701, Korea
**Department of Chemical Engineering, National Institute of Technology (NIT), Tiruchirappalli 620 015, India
(*Received 29 June 2011 • accepted 12 July 2011*)

Abstract−A one-step model calibration methodology of the activated sludge model no. 1 (ASM1) of a full-scale wastewater treatment plant (WWTP) is proposed. First, the key parameters among all parameters of the ASM1 model are selected by sensitivity analysis based on the effluent quality index. Second, multiple response surface methodology (MRSM) is conducted to find the optimal parameter values of the ASM1 model. Lastly, an MRSM analysis is conducted in order to determine the optimal parameter values. This study was conducted in order to develop a new systematic model calibration methodology that can greatly help the modeler to find the optimal solution by selecting the key parameters and optimizing the parameters. In two case studies of simple activated sludge process and a full-scale plant, the experimental results indicated that the calibrated models can improve the prediction quality of the ASM model and the efficiency of the modeling.

Key words: Activated Sludge Model, Multiple Error Minimization, Sensitivity Analysis, Response Surface Methodology, Multiple Response Surface Methodology

INTRODUCTION

Activated sludge processes are most widely used in many biological wastewater treatment plants (WWTPs). Activated sludge models (ASMs) have been developed as bench mark models which are most successful and also describe the industrial standards in WWTPs. These models have proven to be effective for describing carbonaceous, nitrogenous, and phosphorus removal processes in WWTPs [1,2]. To develop control strategies and find optimal parameters in ASMs, determinations of the concentrations of relevant compounds in the wastewater, as well as the stoichiometric and kinetic parameters are needed. However, determination of optimal model parameters is not easy for any specific case [2-4].

Activated sludge models (ASM1, ASM2, ASM2d and ASM3) for nutrient removal in activated sludge process were developed and applied in a variety of WWTPs [1]. Vanrolleghem et al. [5] and Wu et al. [6] conducted respirometric experiments for model calibration, and Vanrolleghem [5] estimated the component concentrations in sludge and influent wastewater, biokinetic and stoichiometric parameters. Gernaey [7] and Sakara [8] studied modeling of activated sludge wastewater treatment plants. Especially, the white-box models and alternative modeling methodologies on the integration of white-box models with artificial intelligence methodologies are indicated in the study [7].

The use of advanced model calibration techniques, including the STOWA, BIOMATH, WERF and HSG protocols, has been recommended in order to solve the problems associated with ASM calibration. These techniques are used to select important parameters and estimate those parameters for the plants. The four calibration methods have their own advantages and disadvantages regarding the applicability, usefulness, accuracy, cost, experimental work and chemical/biological aspects of the model calibration. Recently, Sin et al. [4,9] discussed the comparison of four existing calibration approaches (BIOMATH, HSG, STOWA and WERF) using a SWOT analysis. Kim [10] performed the model-based optimization of the operation in view of the biological performance of N and P removal of a membrane bioreactor (MBR) by means of a two-tier scenario analysis. Kim et al. [11] estimated adequate operating parametric values of WWTP by response surface analysis and optimized the retention time of each reactor to reduce pollution by mixture design. Recently, Sibel [12] investigated NH_4^+-N and PO_4^{3-}-P removal from municipal wastewater by struvite formation, and experimental design was made using Taguchi's technique. However, these studies focused on optimization about a single variable. The process development methods by considering the effect of one variable have some limitations, such as curvature in the system, interaction effect between more than two variables, because, in real plants, there are often cases which need to be considered for several variables, such as total suspended solids (TSS), chemical oxygen demand (COD) and total Kjeldahl nitrogen (TKN) which is the sum of organic nitrogen with ammonia (NH_3) and ammonium (NH_4^+), simultaneously. On the other hand, there are several problems with the model calibration of the ASM families including identifiability, over-parameterization and sparsity. For example, if the parameters are not identifiable then even a low level of noise in the data will result in large variations

†To whom correspondence should be addressed.
E-mail: ckyoo@khu.ac.kr

in the parameters' estimated values, and therefore the parameters cannot be estimated accurately.

The basic assumption on the previous researches of the ASM calibrations is that the parameters are independent of each other. If this assumption is violated, the calibrated parameters may go astray to incorrect position, where the estimated parameters are far from the true parameters. Such an incorrect model calibration may limit the usage of the ASM model.

The multiple response surface methodology (MRSM) can be used for optimizing more than one response variable of the full-scale plants in which many variables should be optimized [13-15]. Several researchers have suggested that by using MRSM, the optimization of the variables can be improved. Vining [13] made use of the dual surfaces of means and variance, and later Lin [14] used an optimization scheme with a cost function to minimize the mean square error by weighting the bias and the variance under the assumption of best target value.

In this study, a new systematic model calibration method for ASM1 parameters is proposed to minimize the modeling error of the plant by using a multiple response surface methodology (MRSM) with a desirability function of the minimum modeling error. Here, the interactions among the modeling errors of COD, TSS and TN concentrations are included in the formulation of the parameter estimation. The proposed systematic calibration method can be applied to a full-scale WWTP.

This paper is organized as follows. Section 2 provides a brief explanation of activated sludge model no. 1 (ASM1), response surface methodology and multiple response surface methodology. In section 3, the proposed calibration scheme of ASM1 is described. Results of the proposed one-step model calibration method are explained in Section 4, followed by conclusions in Section 5.

THEORY

1. Activated Sludge Models No. 1 (ASM1)

Modeling of activated sludge processes has become a common part of the design and operation of WWTP. In 1987, the International Association on Water Quality (IAWQ) task group on mathematical modeling for design and operation of biological wastewater treatment processes presented ASM1. ASM1 is a single-stage activated sludge system that performs simultaneous COD oxidation, nitrification and denitrification processes. After that, the task group developed ASM2 where biological phosphorus removal process model is incorporated, and ASM2d where denitrifying PAOs effects are included, and ASM3 where some defects of ASM1 are enhanced [1-4].

Even though there exist several ASM models (ASM2, ASM2d, ASM3) and the modified ASM model family, still ASM1 is widely used in the biological removal processes since it is well known and easily applicable to modeling of full-scale treatment plants. ASM1 is presented in a matrix format and expressed that the system reaction term, γ_i is obtained by summing the products of the stoichiometric coefficients υ_{ij} and the process rate expression ρ_j for the component i in the mass balance by the following equation.

$$\gamma_i = \sum_j \upsilon_{ij} \rho_j \tag{1}$$

In the ASM1 model, the main 13 components are classified as COD components and nitrogen components. Total COD and total nitrogen balance for the components in ASM1 is defined by the following equations

$$COD_{tot} = S_I + S_S + X_I + X_S + X_{BH} + X_{BA} + X_P \tag{2}$$

Table 1. The typical stoichiometric and kinetic parameters of ASM1

Parameters	Symbol	Unit	Value
Stoichiometric parameters			
Heterotrophic yield	Y_H	g cell COD formed (g COD oxidized)$^{-1}$	0.67
Autotrophic yield	Y_A	g cell COD formed (g N oxidized)$^{-1}$	0.24
Fraction of biomass yielding part.products	f_p	dimensionless	0.08
Mass N/mass COD in biomass	i_{XB}	gN (gCOD)$^{-1}$ in biomass	0.086
Mass N/mass COD in products in biomass	i_{XP}	gN (gCOD)$^{-1}$ in endogenous mass	0.06
Kinetic parameters			
Heterotrophic max. specific growth rate	μ_H	day^{-1}	6.0
Heterotrophic decay rate	b_H	day^{-1}	0.62
Half-saturation coefficient(hsc) for hetero.	K_S	g COD m^{-3}	20
Oxygen hsc for heterotrophs	K_{OH}	g O_2 m^{-3}	0.20
Nitrate hsc for denitrifying heterotrophs	K_{NO}	g NO_3-N m^{-3}	0.50
Autotrophic max. specific growth rate	μ_A	day^{-1}	0.80
Autotrophic decay rate	b_A	day^{-1}	0.20
Oxygen hsc for autotrophs	K_{OA}	g O_2 m^{-3}	0.4
Ammonia hsc for autotrophs	K_{NH}	g NH_3-N m^{-3}	1.0
Correction factor for anoxic growth of hetero.	η_g	Dimensionless	0.8
Ammonification rate	k_a	m^3 (g COD day)$^{-1}$	0.08
Max. specific hydrolysis rate	K_h	g slowly biodeg. COD (g cell COD day)$^{-1}$	3.0
Hsc for hydrolysis of slowly biodeg. Substrate	K_X	g slowly biodeg. COD (g cell COD)$^{-1}$	0.03
Correction factor for anoxic hydrolysis	η_h	dimensionless	0.4

$$N_{tot}=S_{NH}+S_{ND}+S_{NO}+X_{ND}+X_{NI}+i_{XB}\cdot(X_{BH}+X_{BA})+i_{XP}\cdot X_P \quad (3)$$

Also, there are different main processes defined in ASM1 such as growth of biomass, decay of biomass, ammonification of organic nitrogen, and hydrolysis of particulate organic matter [1,3,5]. The stoichiometric and kinetic parameters of ASM1 are given in Table 1.

2. Multiple Response Surface Methodology, MRSM

In a full-scale plant, there are many cases where many variables should be optimized. Therefore, the multiple response surface method (MRSM) is used to cope with the problems for more than one variable, and examine the relationship between one or more response variables and a set of quantitative experimental variables or factors [14,15]. These methods are often employed to find the variables' values or settings that optimize the response. MRSM may be employed to find factor settings like operating conditions that produce the best response and satisfy operating or process specifications, and to identify new operating conditions that produce demonstrated improvement in product quality over the quality achieved by current conditions [11,13].

MRSM is the method in which a series of experiments will be conducted first that will yield adequate and reliable measurements of the response of interest. Secondly, a mathematical model will be determined that best fits the data collected from the design chosen in first step, by conducting appropriate tests of hypothesis concerning the model's parameters. Thirdly, optimal settings of the experimental factors that produce the desirability value of the response will be determined [11]. If the obtained best values of the response are beyond the available resources of the experiment, then response surface methods are aimed at obtaining at least a better understanding of the overall process system.

In general, an experimenter's first attempt to approximate the shape is by fitting a 1^{st} order model to the response values in the absence of sufficient knowledge concerning the shape of the true response surface. A linear function of the factors with 1^{st} order model can be written as the following equation.

$$Y=\beta_0+\beta_i x_i+\beta_j x_j+\beta_{ij} x_i x_j+\varepsilon_i \quad (4)$$

where Y is the response variable, x are independent variables, β are regression coefficients, and ε is the error. However, the 1^{st} model suffers from lack of fit arising from the existence of surface curvature, hence a second order model is defined by adding higher-order terms to the first order model and is obtained as following equation.

$$Y=\beta_0+\sum\beta_i x_i+\sum\beta_{ii} x_i^2+\sum\beta_{ij} x_i x_j+\varepsilon_i \quad (5)$$

where Y is the response variable, x_i and x_j are i^{th}, j^{th} coded independent variables, β_0 is the intercept, β_0, β_i, β_{ii}, β_{ij} are the regression coefficients [13,16]. Because many variables have different units and different limits, independent variables are coded in which the original variable X_i is coded as x_i by equation,

$$x_i=\frac{(X_i-X_0)}{\delta X} \quad (6)$$

where x_i is the coded variable, X_i is an independent variable, which is not coded, X_0 is the mean value and δX is a standard deviation [16,17].

The desirability function approach and loss function approach are the most widely used methods for the optimization of MRSM. The desirability function transforms the results of analysis into desirability (d). This approach reduces multiple responses to a single aggregated measure and solves the equation as a single objective optimization. The aggregated measure has often been defined as a desirability function. Desirability assigns a number between 0 and 1 to the possible value of each response Y [16,17].1

To describe the desirability function approach mathematically, let each of the k response variables are related to p independent variables by the following equation.

$$Y_{ij}=f(X_1, X_2, \cdots, X_p)+\varepsilon_{ij}$$
$$i=1, 2, \cdots, k$$
$$j=1, 2, \cdots, p \quad (7)$$

where Y_{ij} is the j^{th} observation on the i^{th} response and f_i denotes the relationship between the i^{th} response, y_i and $\mathbf{x_1}, \cdots, \mathbf{x_p}$. The parameter p is the maximum number of observations for each of the k responses and ε_{ij} is an error term with a mean value of 0 and variance of σ_i^2 [14].

Desirability functions have three different forms: maximized, minimized, assigned or target value [16]. The objective of this study is to find the minimum error differences. To achieve this objective, the desirability function is defined as follows.

$$d=\begin{cases}0 & \hat{Y}_i \le A \\ \left[\dfrac{\hat{Y}_i-A}{B-A}\right]^r & A \le \hat{Y}_i \le B \\ 1 & \hat{Y}_i \le B\end{cases} \quad (8)$$

where A is the minimum allowable value of $\hat{\mathbf{Y}}_i$ and B is the problem-dependent limit value of $\hat{\mathbf{Y}}_i$, that is, the minimum and maximum error between predicted and default values. The weighted value, r, is an assigned number between 0.1 and 10, and is used as an index to express the relative importance of variables. If a response value $\hat{\mathbf{Y}}_i$ is more than B, then d=1 and if $\hat{\mathbf{Y}}_i$ is less than A, then d=0 [14, 15].

The overall desirability (D) is obtained by combining the individual desirability values using geometric mean as equation,

$$D=(d_1\times d_2\times\cdots\times d_k)^{1/k} \quad (9)$$

where k denotes the number of the response. In this study, k is 3, d_1=TSS error, d_2=COD error and d_3=TKN error. This single value of D represents the overall desirability of the combined response levels. The value of D lies between 0 and 1, so, if D is close to 1, it indicates that the balance of the properties becomes more favorable. However, if D=0, it indicates that one of the response variables is unacceptable that is d_i=0, that means the overall product is unacceptable [14,17].

To maximize the total quality of a system, the dual optimization objective function is defined as:

$$\underset{x_i^*}{\text{Max}} D=\left(\prod_{i=1}^{k} d_i\right)^{1/k} \quad (10)$$

s.t.
$d_i=df_i(Y_i)$, desirability function
$Y_i=f_i(X_1, X_2, \ldots, X_p)$

where D is the overall desirability, d_i is individual desirability func-

tion, is the optimal, p is the number of the independent variables, and k is the number of responses [11,13,17].

The outcome of the experiments generally depends on the experimental conditions. This means that the experimental results can be described as a function of the experimental variables by the following equation. In this study, polynomial models of linear, interaction, and quadratic terms are used to describe the relationship between the experimental variables and the responses. For the experimental model, an analysis of variance (ANOVA) method is used to analyze the aspects of the precipitation behavior of alloys for different process combinations. ANOVA can be used to estimate whether the results of the model are significant or not. If the results are not significant, a new model is constructed and analyzed [13,18].

To complete the ANOVA calculations, degrees of freedom (d.f.) should also be considered together with each sum of squares. As ANOVA studies are in fact experimental studies with a certain test error, determination of error variance is an essential step. Similarly, sample variance within the factor levels should be calculated because the sample size establishes the confidence level of the results derived from the analysis. These data are subsequently used to estimate the value F of the Fisher test (F-test). The portion of total variation observed in an experiment attributed to each significant factor and/or interaction is reflected in the percent contribution, (P), which indicates the relative power of a factor and/or interaction to reduce variation, i.e., factors and interactions with substantial percent contributions are most important.

MATERIALS AND METHODS

1. Data Collection and Process Design

General information required for process modeling is collected and compiled from the activated sludge plant database based on designed documents and/or personal communication with the plant operators. The data for the process design are influent and effluent characteristics, physical characteristics such as volumes, compartments, pumping capacities, aerators and pipelines, and various kinetic parameters about microorganisms in the activated sludge process [19]. The selected process is designed using a general purpose simulator of wastewater treatment plant. In this study, the WEST® program is used to model the process.

2. Parameter Selection by Sensitivity Analysis

The optimization of the ASM is problematic due to model complexity because of many components and because of many kinetic and stoichiometric parameters. To improve the prediction efficiency of the model, sensitivity analysis is performed to select the key parameters influencing the removal efficiency, because it is difficult to simultaneously consider all of the parameters when predicting the results at the same time. The phased change method of single parameter is applied between 60-160% scopes of the ASM1 parameters. Then, the most sensitive parameters are determined by sensitivity analysis functions of the measured variables to the model parameters. These selected parameters are used for the optimal parameter estimation using MRSM [19].

3. A Multiple Errors Minimization of ASM1 Model Parameters by Using MRSM

A multiple response surface methodology (MRSM) is conducted to find optimal parameter values of the ASM1 model. Here, the model errors which are obtained by comparing the measured data with the ASM model prediction values are very useful for parameter estimation. These model errors are used as the input of MRSM with a desirability function in order to minimize the modeling errors of chemical oxygen demand (COD), total suspended solid (TSS) and total nitrogen (TN) concentrations. The interactions between the modeling errors of COD, TSS and TN concentrations are included in the formulation of the parameter estimation.

By using MRSM, one can find optimal operating conditions or optimal parameter values through a minimum number of experiments [11]. The experimental design for MRSM uses independent variables which are chosen by sensitivity analysis. The variables were transformed and coded as −1, 0, or 1. The experimental design uses a central composite design (CCD) which results in an efficient quadratic model.

In this study, polynomial models of linear, interaction, and quadratic terms are used to describe the relationship between the experimental variables and the responses. For the experimental model, an analysis of variance (ANOVA) method is used to analyze the aspects of the precipitation behavior of alloys for different process combinations. ANOVA can be used to estimate whether the results of the model are significant or not. If the results are not significant, a new model is constructed and analyzed [13,18].

In this paper, a new objective function of response variables of MRSM for the model parameter estimation is proposed with the following equation.

$$e = y - \hat{y} \tag{11}$$

where y is a vector of the measured data of TSS, COD, and TKN, $\hat{y}$ is a vector of the predicted data and e is a vector of the modeling error of ASM1. The error values are used to confirm the difference between the measured and predicted data. If the errors are closer to zero then it means that the modeling results match with the real process. The optimal values of ASM1 parameters to get minimum value of the error can be determined using MRSM. The interactions be-

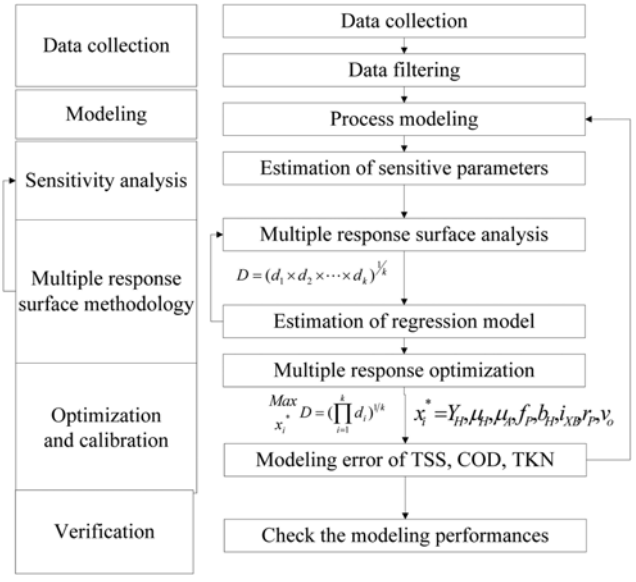

Fig. 1. A proposed systematic model calibration method of the ASM model.

tween the modeling errors of COD, TSS and TN concentrations are included in the formulation of the parameter estimation.

4. Proposed Systematic Model Calibration Methodology for ASM1 Model Parameters

In this study, to find optimal parameters of ASM1, a systematic model calibration method is proposed as shown in Fig. 1. The model calibration method consists of six steps: collecting data, modeling, sensitivity analysis, multiple response surface methodology, optimization and calibration, and verification. First, it gathers key data set from a plant. Second, a model is selected and a treatment plant is modeled by using ASM1. The third step is to select key parameters among all parameters of ASM1 using a sensitivity analysis with respect to the effluent qualities. The fourth and fifth steps are MRSM to find optimal parameters of ASM1, where the parameter searching space is designed using CCD. The following step is based on multiple errors minimization scheme of TSS, COD and TKN. In this step, an individual desirability and a transformation function of TSS, COD and TKN are defined and a multiple optimization framework is performed based on the overall desirability function. The final step is to check whether the modeling performance with the calibrated parameter of ASM1 is satisfactory or not. If the system performance is not satisfactory, the model calibration strategy needs to be iterated.

RESULTS AND DISCUSSION

Two cases are studied in this research. Case study 1 is a simple activated sludge process partitioned into three compartments of an

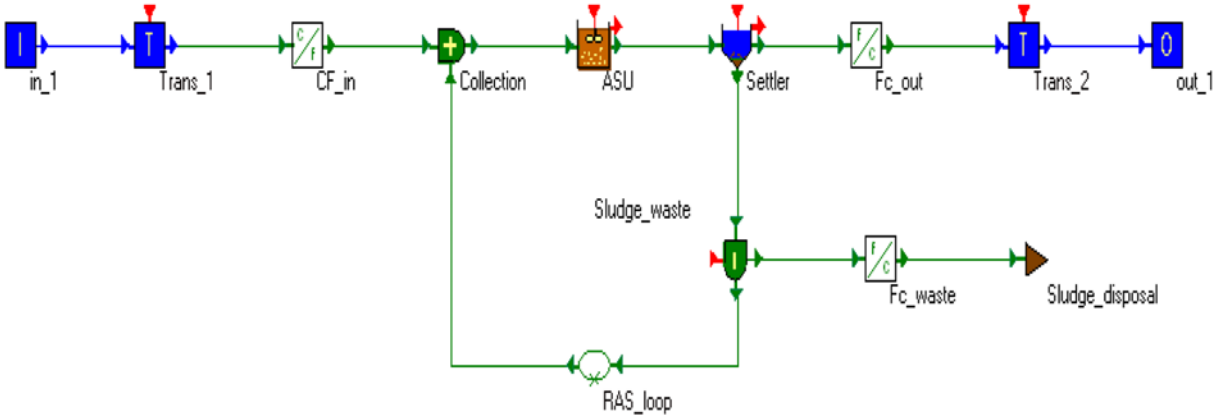

Fig. 2. Layout of activated sludge processes in WEST (case study 1).

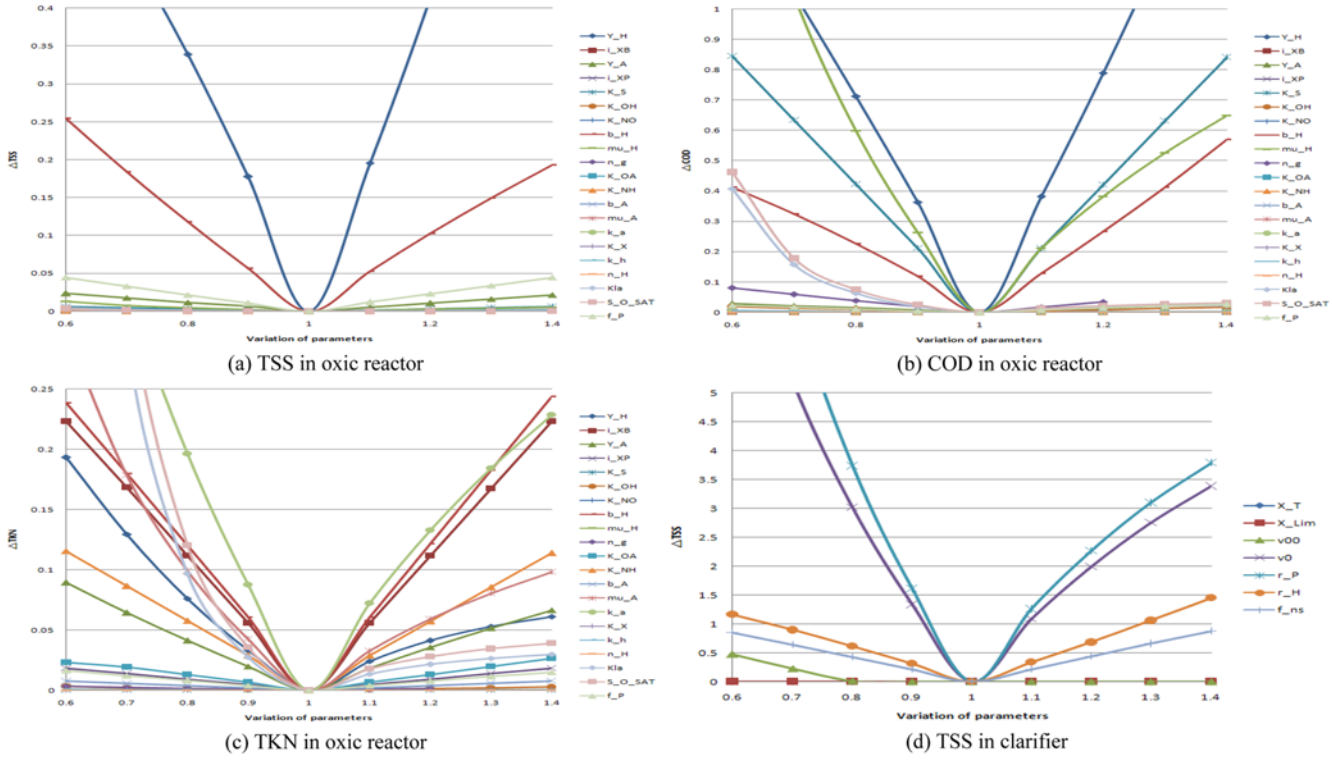

Fig. 3. Sensitivity analysis for each parameter in case study 1, (a) TSS, (b) COD, (c) TKN in oxic reactor, and (d) TSS in clarifier.

oxic reactor, a clarifier and an external recycle. Case study 2 is a Daewoo Nutrient Removal (DNR) process which is an advanced biological treatment process with four basins of anaerobic, two anoxic, oxic processes and a secondary clarifier. To model these two case studies, the WEST® program has been used as a general purpose simulator.

- Case study 1: Simple activated sludge process.
- Case study 2: DNR process in a full-scale treatment plant.

1. Case Study 1

A simple activated sludge process consists of oxic reactor, clarifier and recycled sludge in clarifier. The influent conditions in this study are as follows: flow 2,000 m^3/d; TSS 149 g/m^3; COD 398 g/m^3; and TKN 18 g/m^3 and operational conditions are as follows: oxic reactor volume 1,000 m^3; clarifier volume 300 m^3; recycled sludge 2,000 m^3/d; and wasted sludge 40 m^3/d. Fig. 2 shows the layout of the first case study used in WEST program simulation.

Fig. 3 shows the relative sensitivity about TSS, COD and TKN. Heterotrophic yield coefficient Y_H, which affects the sludge production and nitrogen content of the biomass I_{XB}, was sensitively represented as change of TSS in Fig. 3(a). Heterotrophic yield coefficient Y_H, which affects the oxygen demand and maximum specific heterotrophic growth rate μ_H and Half-saturation coefficient (hsc) for heterotrophic growth rate K_S, was sensitively represented as change of COD in Fig. 3(b). Ammonification rate constant K_A and heterotrophic decay coefficient b_H were sensitively represented as change of TKN in Fig. 3(c). Finally, Y_H, μ_H, μ_A, b_H and i_{XB} selected in association with organic matter, nutrient removal, growth and decay of biomass in aeration tank are sensitive to TSS in Fig. 3(d). Also settling parameter in low concentration r_P and maximum theoretical settling velocity v_0 are sensitively represented as change of TSS in Fig. 3(d).

Response surface methodology (RSM) is conducted to use central composite design (CCD) and to find optimal parameters. Table 2 shows the coded and actual levels of seven key variables, and the model has run through 152 experimental trials. Experimental results are subjected to multiple linear regression analysis using MINITAB 14 statistical software to approach the optimal response region for minimum errors of TSS, COD and TKN. ANOVA analysis is conducted for estimating error between real data and model data for TSS, COD and TKN. All the *p*-values are less than 0.01, R^2 values obtained for TSS, COD and TKN are TSS=0.973, COD=0.972, TKN=0.979. Then, nonlinear regression coefficients are estimated and are given in Table 3.

To know whether the model is over fitted, verification has been carried out by using training data and test data. The training data is defined as the data which is used for obtaining the model equation in this study. On the other hand, the test data is comprised of non-used data for obtaining the model equation in the process.

In Fig. 4 and Fig. 5, comparison of experimental data and predicted data from 1st and 2nd order model equation for training and test data are presented. The results of 2nd order model equations are closer to real data than 1st order model equations. Also, a root mean square error (RMSE) value is used to find out the model accuracy.

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(X_i-Y_i)^2}{n-1}} \tag{12}$$

where, X_i is experimental value, Y_i is prediction value, n is the number of experimental values. Table 4 shows RMSE values for each

Table 2. Experimental design for MRSM in case study 1

Independent variables	Range and level		
	−1	0	1
Y_H	0.402	0.67	0.938
μ_H	2.4	4	5.6
b_H	0.24	0.4	0.56
μ_A	0.33	0.55	0.77
i_{XB}	0.0516	0.086	0.1204
r_P	0.001716	0.00286	0.004004
v_0	284.4	474	663.6

Table 3. Nonlinear regression coefficients in case study 1

	TSS	COD	TKN
Term	Coefficient	Coefficient	Coefficient
Constant	−102.75	−138.35	−13.11
Y_H	−0.63	−3.12	1.75
μ_H	0.09	2.22	0.09
b_H	7.21	−7.01	−0.85
μ_A	−1.45	5.26	4.77
i_{XB}	21.03	−20.39	0.08
r_P	31099.10	41890.50	3231.19
v_0	0.15	0.21	0.02
Y_H*Y_H	−4.11	−4.92	−1.14
μ_H*μ_H	−0.05	−0.28	−0.01
b_H*b_H	−6.75	−10.81	−0.72
μ_A*μ_A	−2.93	−4.08	−2.39
i_{XB}*i_{XB}	−120.12	−167.42	−13.29
r_P*r_P	−2891418.00	−3872908.00	−287974.00
v_0*v_0	0.00	0.00	0.00
Y_H*μ_H	0.44	0.63	0.03
Y_H*b_H	−7.18	−9.80	1.11
Y_H*μ_A	2.75	−1.21	−1.71
Y_H*i_{XB}	−16.88	9.01	2.74
Y_H*r_P	1339.25	1547.59	118.87
Y_H*v_0	0.00	0.00	0.00
μ_H*b_H	0.77	2.76	−0.03
μ_H*μ_A	−0.46	−0.89	−0.03
μ_H*i_{XB}	2.93	5.79	0.14
μ_H*r_P	−106.35	−203.53	−14.26
μ_H*v_0	0.00	0.00	0.00
b_H*μ_A	4.54	5.84	1.46
b_H*i_{XB}	−29.64	−38.10	−18.44
b_H*r_P	820.96	2217.82	32.93
b_H*v_0	0.00	0.00	0.00
μ_A*i_{XB}	26.72	39.94	1.47
μ_A*r_P	−621.63	−1480.60	−135.82
μ_A*v_0	0.01	0.00	0.00
i_{XB}*r_P	3964.21	9343.56	693.61
i_{XB}*v_0	−0.04	−0.02	0.00
r_P*v_0	−15.95	−22.02	−1.71

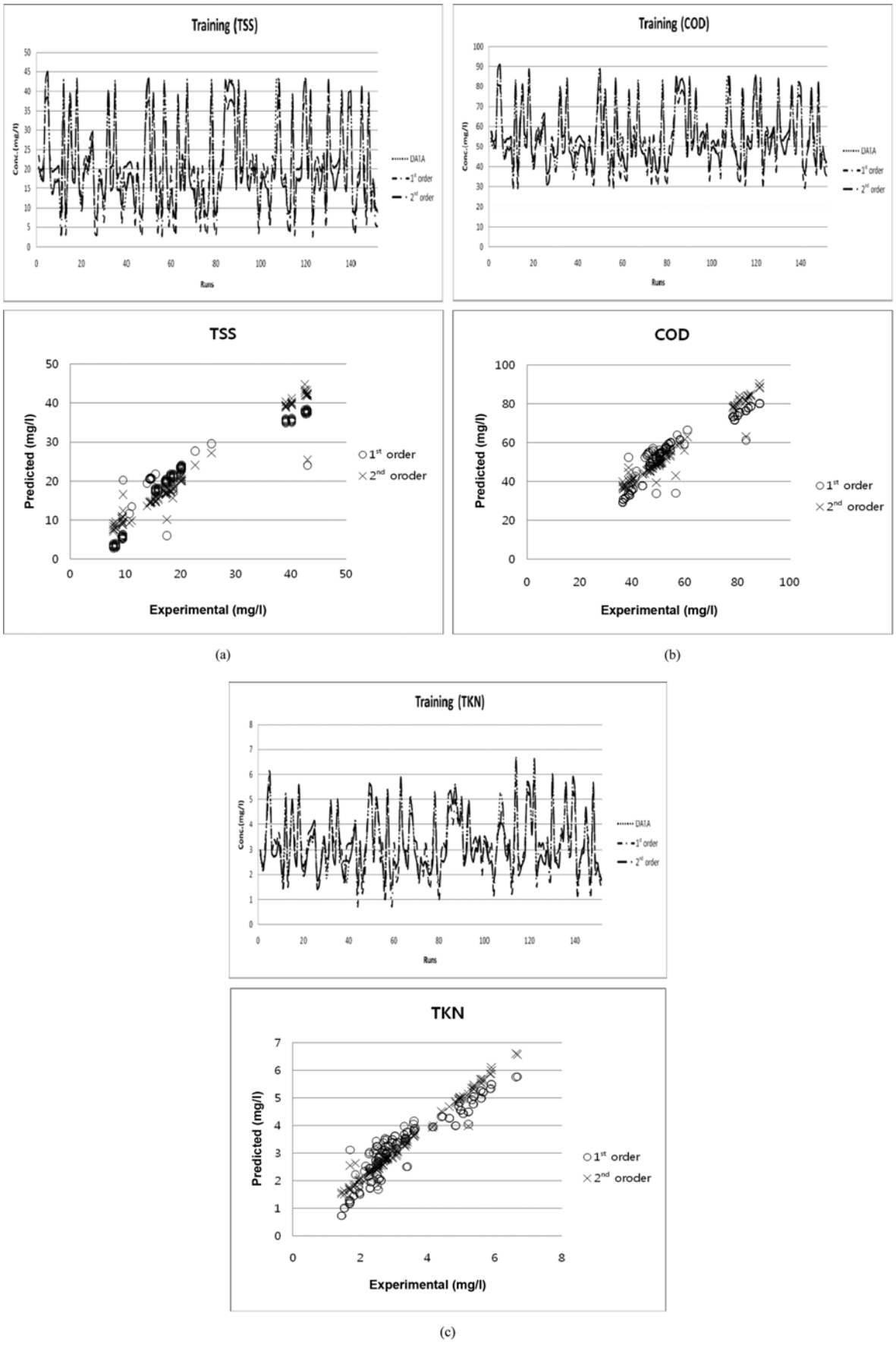

Fig. 4. The comparison of experimental and predicted data for training data (case study 1), (a) TSS, (b) COD, and (c) TKN.

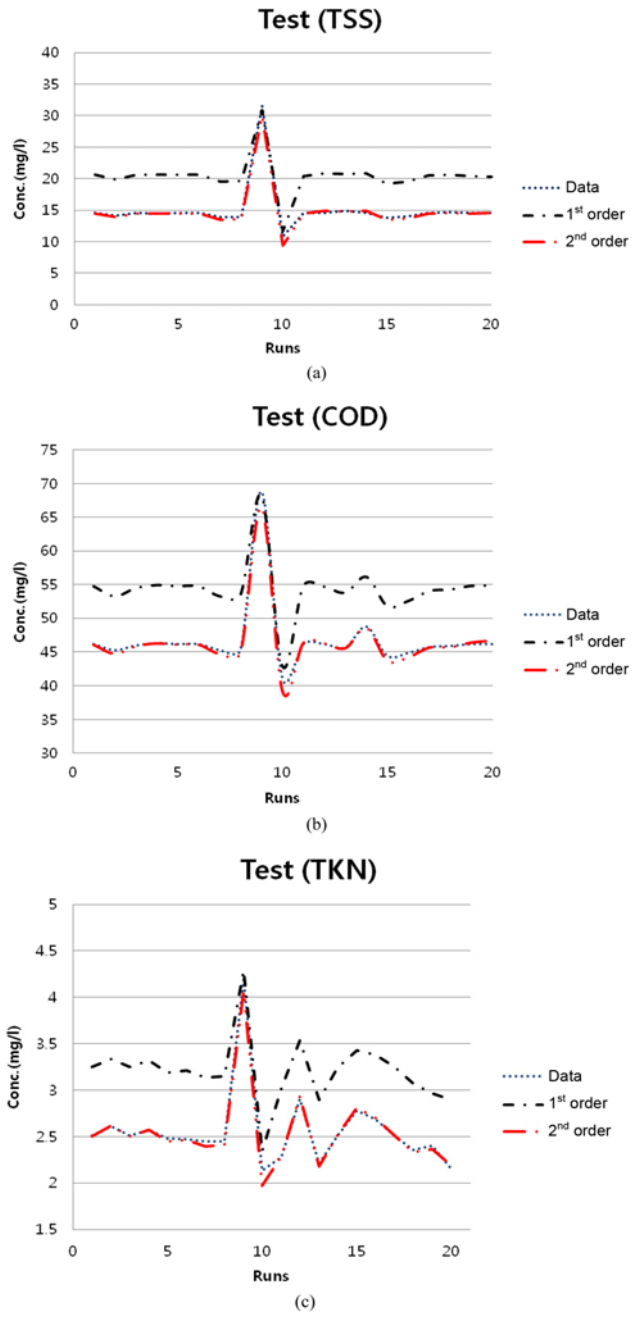

Fig. 5. The comparison of experimental and predicted data for test data (case study 1), (a) TSS, (b) COD, and (c) TKN.

Table 4. RMSE values between training and test data in case study 1

		TSS (mg/L)	COD (mg/L)	TKN (mg/L)
Training data	1^{st} Order	6.33	4.52	0.51
	2^{nd} Order	2.63	1.89	0.18
Test data	1^{st} Order	12.27	8.40	1.73
	2^{nd} Order	4.30	6.11	1.09

component, and the accuracy of the model was favorable when the RMSE is close to zero. RMSE values of 2^{nd} order model equations are less than RMSE values of 1^{st} order model equations.

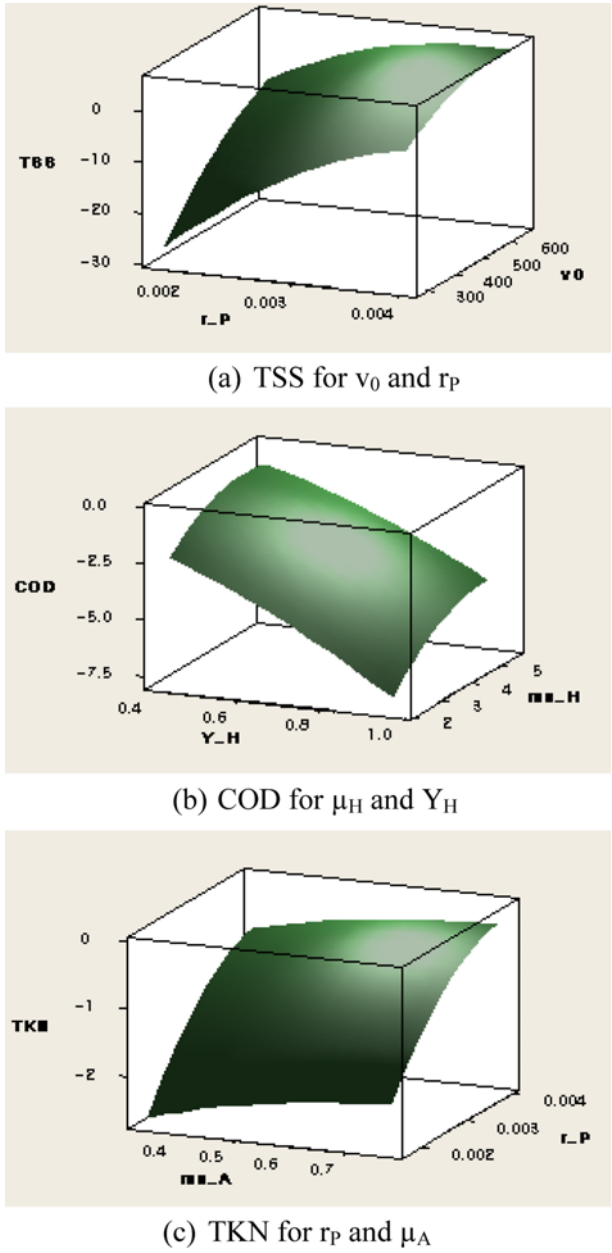

Fig. 6. Response surface plots of (a) TSS, (b) COD, and (c) TKN with six kinetic parameters of ASM1.

Optimal operational conditions in Case 1 are determined by using MRSM. The target values of response variables for the desirability function are close to zero because of minimum error between experimental and predicted data. The response surface plots to estimate the COD error, Y_H and μ_H (keeping other variables as constants) are shown in Fig. 6. The optimal parameter values obtained after doing multiple response optimizations are: Y_H 0.938; μ_H 5.567; b_H 0.240; μ_A 0.506; i_{XB} 0.120; r_P 0.003 and v_0 663.600. Also the errors of TSS, COD and TKN obtained are 0.0412, −0.1913 and −0.4369, respectively and given in Fig. 7.

Then the models are calibrated by optimal parameters which are found from MRSM in case 1. Table 5 shows the comparison of real data, default results and calibration results. The errors between real data and calibration data are less than using default values of ASM1

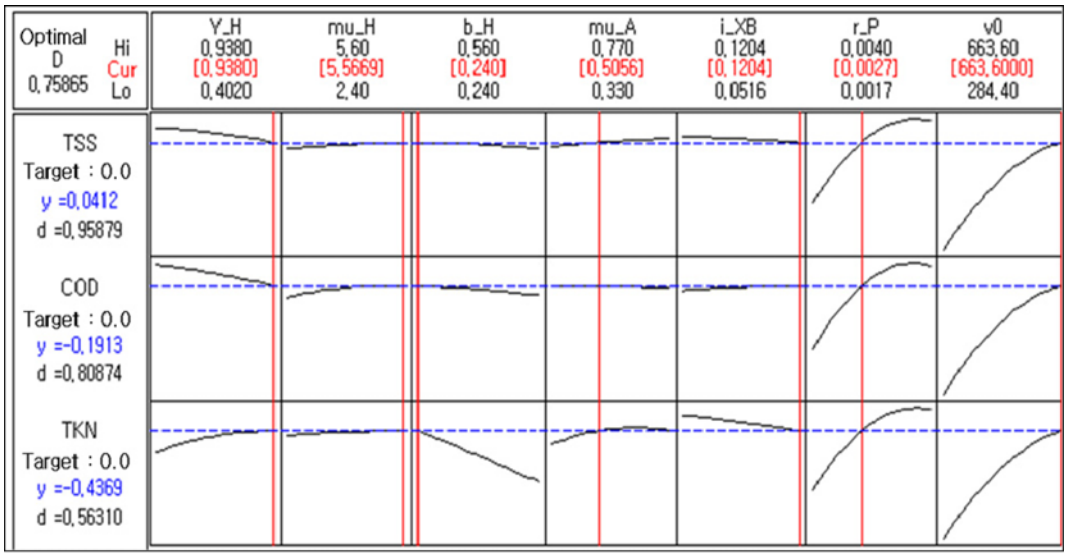

Fig. 7. The composite desirability and optimal parameter values of seven kinetic parameters of ASM1 (case study 1).

Table 5. The calibration results in case study 1

	Real data	Default	Calibration
TSS (mg/L)	12.82	14.56	12.79
COD (mg/L)	43.09	46.14	42.83
TKN (mg/L)	1.557	2.509	2.079

parameters.

2. Case Study 2

The proposed method is applied to a full-scale plant of a DNR process. The DNR process is an advanced biological nutrient removal process and consists of two anoxic, anaerobic and oxic reactor and clarifier. The influent conditions used are as follows: flow 8,767 m^3/d; TSS 114 g/m^3; BOD_5 108 g/m^3; COD 206 g/m^3; and TKN 26 g/m^3. Also, the WEST® program is used for modeling and

Table 6. The results of sensitivity analysis in case study 2

	Sensitive variables
Denitrification reactor	Y_H, b_H, f_P
Anaerobic reactor	Y_H, b_H, f_P
Anoxic reactor	Y_H, b_H, f_P
Oxic reactor	K_S, μ_H, Y_H
Secondary clarifier	v_0, r_P

Fig. 8 shows the layout of DNR process.

To find optimal parameters in ASM1, the parameters and procedures are the same as carried out in case 1. Sensitivity analysis of TSS, COD and TKN is performed at each reactor. The results of sensitivity analysis at each reactor are indicated in Table 6. Based on the results of sensitivity analysis, five parameters of b_H, K_S, μ_H,

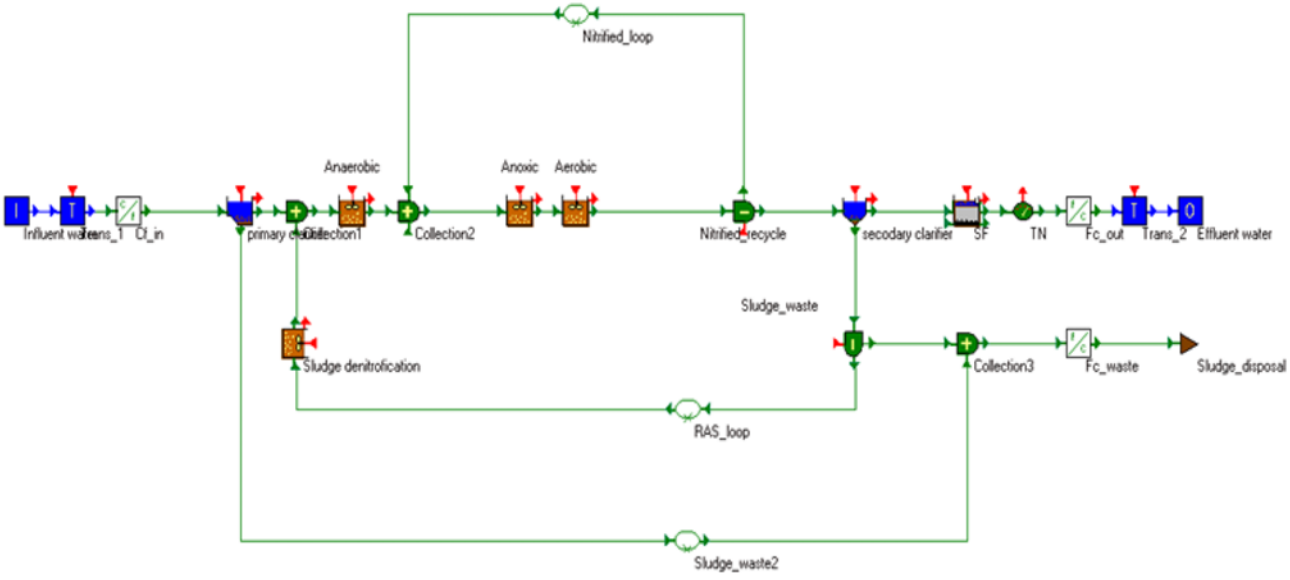

Fig. 8. Layouts of activated sludge processes in WEST (case study 2).

Table 7. Experimental design of five key parameters in case study 2

Independent variables	Range and level		
	−1	0	1
b_H	0.24	0.4	0.56
K_S	12	20	28
μ_H	2.4	4	5.6
v_0	284	474	664
r_P	0.00172	0.00286	0.00400

v_0 and r_p, are selected as the key parameters for DNR modeling. And then MRSM is performed to find optimal parameters.

Table 7 shows the experimental design of five key parameters which are coded with three levels of −1/0/1 and has run through 32 experimental trials. Design of experiments of five key parameters of ASM1 model is conducted through the combinations of 5 key parameters which can search for the optimal parameters with minimum modeling errors of TSS, COD and TKN in a treatment plant. The 2nd order model regression coefficients in case 2 are given in Table 8. ANOVA results of nonlinear regression represents that all of the p-values are less than 0.01, and their R^2 values are TSS= 0.993, COD=0.993, TKN=0.981.

Fig. 9 shows the composite desirability and optimal parameter values of five kinetic parameters of ASM1. By using multiple response optimization, the optimal parameters of ASM1 by using the proposed systematic model calibration method are found as b_H 0.56; K_S 28.0; μ_H 3.30; v_0 664.0; and r_P 0.004. RMSE values of ASM1 for training data in Table 9 are 0.47 for TSS, 0.65 for COD and 0.12 for TKN, respectively. The data used for test period is not used for model building step. RMSE values of ASM1 for test data are 1.45, 1.96, and 0.34 for TSS, COD and TKN, respectively. Fig. 10 and Fig. 11 show the prediction results of ASM1 model with the defaults and calibrated parameters for training and test data to compare with the measured data of TSS, COD, and TKN. Table 9, Fig. 10 and

Table 8. The 2nd order model regression coefficients in case study 2

	TSS	COD	TKN
Term	Coefficient	Coefficient	Coefficient
Constant	171.494	18.216	119.432
b_H	−6.263	1.574	−9.954
K_S	0.041	−0.040	−0.070
μ_H	−1.142	−0.183	−0.279
v_0	−0.197	−0.023	−0.149
r_P	−42030.600	−4685.600	−31766.000
b_H*b_H	12.507	2.346	6.900
K_S*K_S	0.003	0.001	0.002
$\mu_H*\mu_H$	0.275	0.018	0.048
v_0*v_0	0.000	0.000	0.000
r_P*r_P	3897027.000	408837.000	2908634.000
b_H*K_S	0.392	0.000	−0.002
$b_H*\mu_H$	−2.422	0.002	0.020
b_H*v_0	0.003	−0.001	0.002
b_H*r_P	512.183	−174.142	409.747
$K_S*\mu_H$	−0.046	0.002	−0.004
K_S*v_0	0.000	0.000	0.000
K_S*r_P	−0.819	−0.205	0.273
μ_H*v_0	0.000	0.000	0.000
μ_H*r_P	10.244	1.024	−2.732
v_0*r_P	19.708	2.832	15.556

Table 9. Comparison of the measured data of TSS, COD, and TKN and the predicted value of ASM1 with the defaults and calibrated parameters

	Real data	Default	Calibration
TSS (mg/*l*)	1.83	15.23	8.79
COD (mg/*l*)	11.91	35.87	30.66
TKN (mg/*l*)	6.25	3.37	4.01

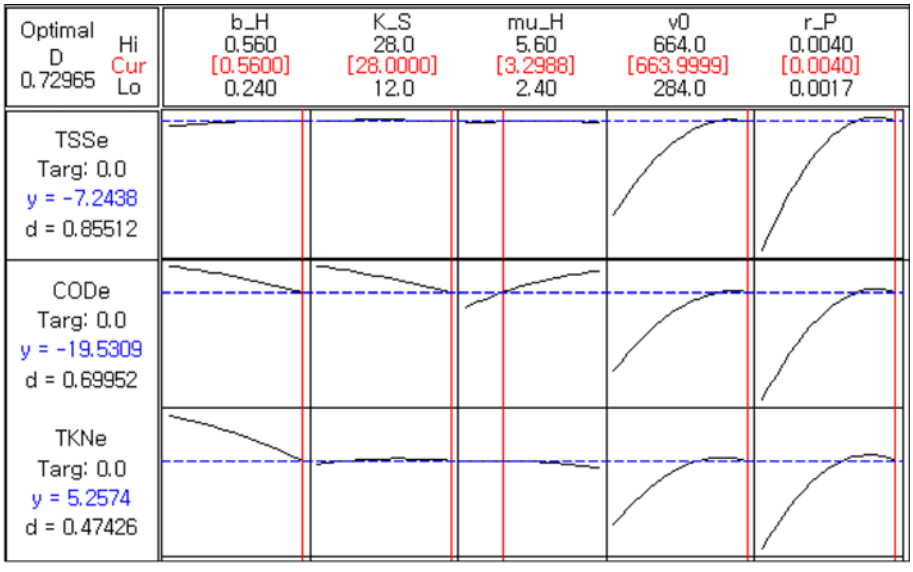

Fig. 9. The composite desirability and optimal parameter values of five kinetic parameters of ASM1 (case study 2).

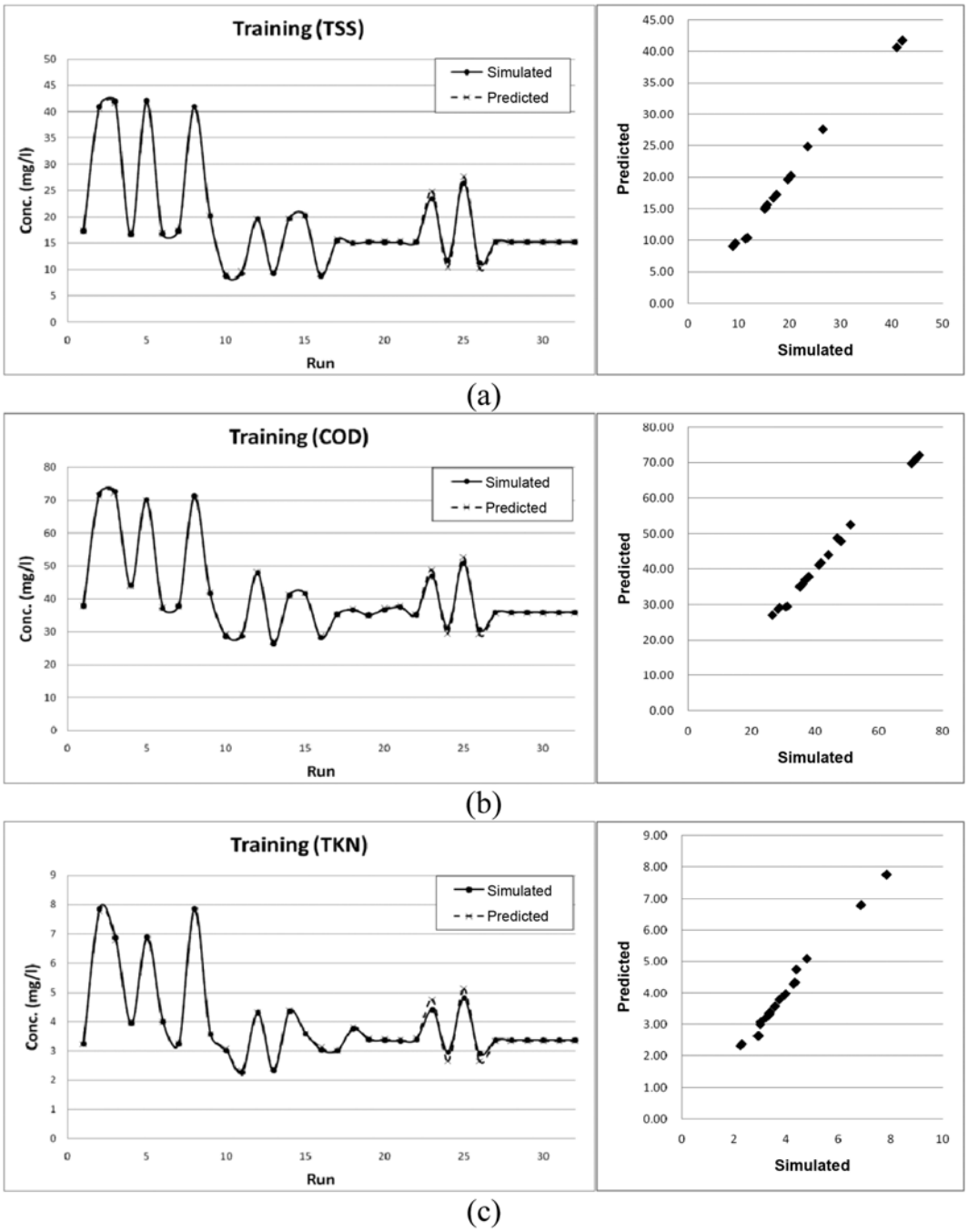

Fig. 10. The comparison of measured and predicted data for training data (case study 2), (a) TSS, (b) COD, and (c) TKN.

Fig. 11 clearly show that the calibration results of ASM1 are closer to the measured data than using default parameters.

CONCLUSION

A new systematic calibration method for the ASM model based on multiple error minimization values of TSS, COD, and TKN is proposed to find optimal parameter values of ASM1. The proposed method is validated in two case studies. Case studies represent that each wastewater plant has different sensitive parameters of ASM1 model. In case study 1, Y_H, μ_H, b_H, μ_A, i_{XB}, r_P and v_0 are selected as sensitive parameters, while in case study 2, b_H, K_S, μ_H, v_0 and r_P are determined as sensitive parameters. The nonlinear regression coefficients for each component are estimated, and model equations are used to predict effluent components in training and test data. The results of the calibrated ASM1 parameters show better modeling performance with low RMSE values for both training and test data set than using default parameter values. We expect that the proposed method can be applied to calibrate the model parameters of other processes such as anaerobic digestion model and river water quality model. Moreover, our on-going research is focused to further upgrade and optimize a full-scale WWTP.

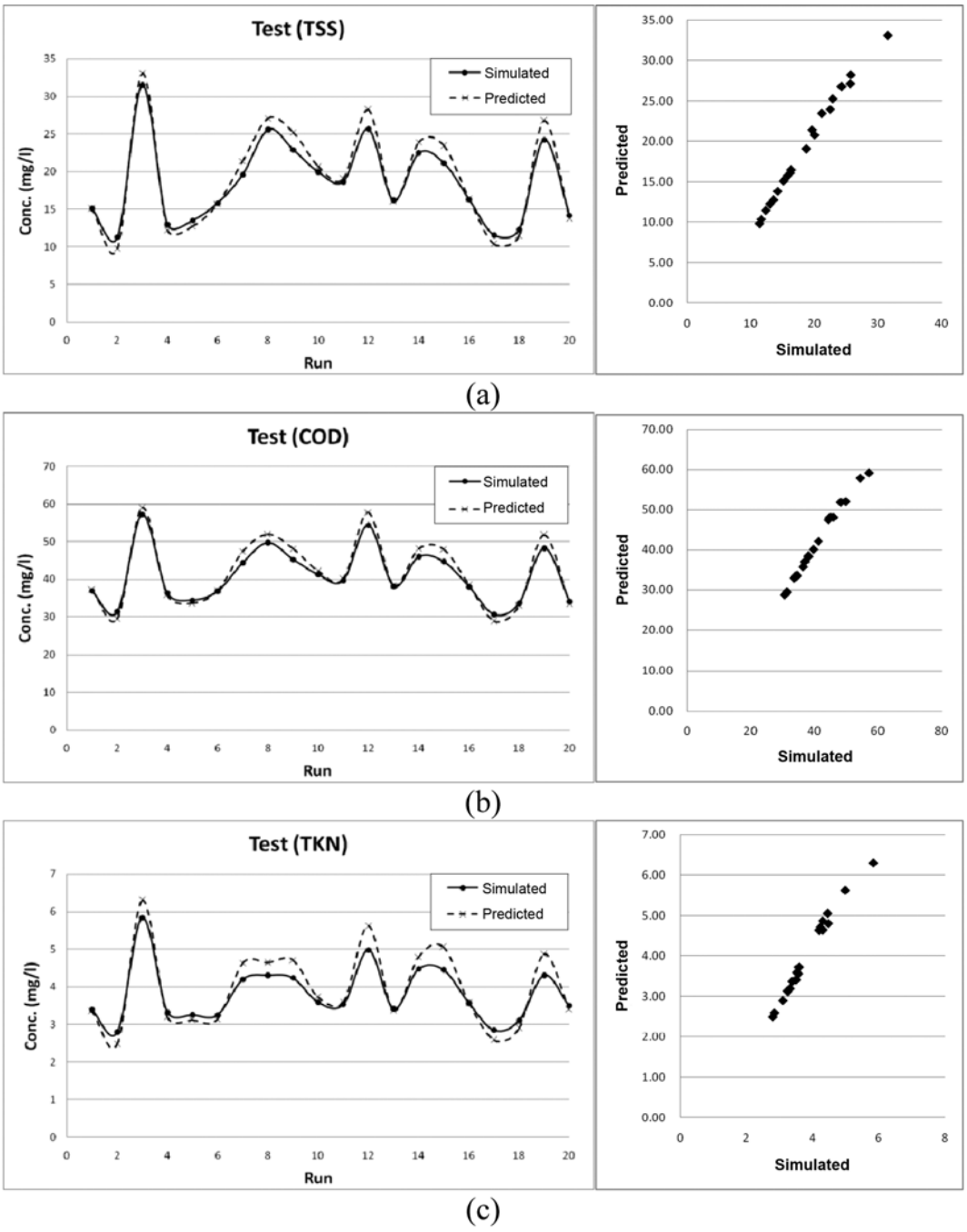

Fig. 11. The comparison of measured and predicted data for test data (case study 2), (a) TSS, (b) COD, and (c) TKN.

ACKNOWLEDGEMENTS

This work was supported by the Korea Science and Engineering Foundation (**KOSEF**) grant funded by the Korea government (**MEST**) (KRF-2009-0076129) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0001031).

REFERENCES

1. M. Henze, W. Gujer, T. Mino and M. Loosdrecht, *Activated sludge models ASM1, ASM2, ASM2d and ASM3, IWA scientific and technical report No. 9*, IWA, U.K. (1999).
2. G. Olsson and B. Newell, *Wastewater treatment system - modeling, diagnosis and control*, IWA, U.K. (1999).
3. P. Britta, Ph. D. Thesis, Ghent University, Belgium (2000).
4. G. Sin, S. W. H. Van Hulle, D. J. W. DePauw, A. V. Griensven and P. A. Vanrolleghem, *Water Res.*, **39**, 2459 (2009).
5. P. A. Vanrolleghem, S. Henri, P. Britta, G. Philippe and T. Imre, *Water Sci. Technol.*, **39**, 195 (1999).
6. J. Wu, X. Jiang and A. Wheatley, *Desalination*, **249**, 969 (2009).
7. K. V. Gernaey, M. C. M. Loosdrecht, M. Henze, M. Lind and S. B.

Jorgensen, *Environ. Model. Soft.*, **19**, 763 (2004).
8. U. Sarkar, D. Dasgupta, T. Bhattacharya, S. Pal and T. Chakroborty, *Desalination*, **252**, 120 (2009).
9. G. Sin, G. Insel, D. S. Lee and P. A. Vanrolleghem, *Water Sci. Technol.*, **50**, 97 (2004).
10. I. Nopens, G. Sin, T. Jian and P. A. Vanrolleghem, *Water Sci. Technol.*, **56**, 135 (2007).
11. S. L. Kim, Ph. D. Thesis, Pukyong National University, Korea (2007).
12. U. D. A. Sibel, *Environ. Eng. Sci.*, **25**, 1 (2008).
13. G. G. Vining and R. H. Myers, *J. Quality Technol.*, **22**, 38 (1990).
14. D. K. J. Lin and W. Tu, *J. Quality Technol.*, **27**, 34 (1995).
15. S. H. R. Pasandideh and S. T. A. Niaki, *Appl. Math. Comput.*, **175**, 366 (2006).
16. J. B. Kwon, J. S. Lee, S. H. Lee, C. H. Jun and K. J. Kim, *J. KORMS*, **30**, 95 (2005).
17. M. H. Kim, A. S. Rao and C. K. Yoo, *Ind. Eng. Chem. Res.*, **48**, 6363 (2009).
18. J. R. Kim, Ph. D. Thesis, Pusan University, Korea (2006).
19. P. A. Vanrolleghem, *Systematic calibration protocol for actuated sludge models, BIOMATH Report*, Ghent University, Belgium (2006).