

Research Article

Practical Identifiability Analysis and Optimal Experimental Design for the Parameter Estimation of the ASM2d-Based EBPR Anaerobic Submodel

Zhenliang Li ^{1,2}, Peili Lu ^{1,3}, Daijun Zhang,^{1,3} and Tian Zhang⁴

¹State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing 400044, China

²Chongqing Research Academy of Environmental Sciences, Chongqing 401147, China

³Department of Environmental Science, Chongqing University, Chongqing 400044, China

⁴Chongqing Industry Polytechnic College, Chongqing 401120, China

Correspondence should be addressed to Peili Lu; lupl@cqu.edu.cn

Received 13 December 2017; Accepted 8 March 2018; Published 11 April 2018

Academic Editor: Salvatore Strano

Copyright © 2018 Zhenliang Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Identifiability analysis is a precondition for reliable parameter estimation. Building on previous work on structural identifiability, this paper focuses on the practical identifiability and optimal experimental design (OED) of the EBPR anaerobic submodel. The nonnegative determinant of the Fisher information matrix (FIM) found in this study clearly demonstrates that the parameters Y_{PO_4} , K_A , q_{PHA} , and X_{PAO} in the submodel are practically identifiable using S_A and S_{PO_4} as the measured variables and fixing K_{PP} as the default value. Furthermore, fixing K_{PP} to study the practical identifiability of the other parameters and to estimate their values is shown to be valid. Subsequently, a modeling-based procedure for the OED for parameter estimation was proposed and applied successfully to anaerobic phosphorus release experiments. According to the FIM D-criterion, the optimal experimental condition was determined to be an initial S_A concentration of 300 mg/L. Under the optimal experimental condition, errors in the values of Y_{PO_4} , K_A , q_{PHA} , and X_{PAO} are all below 20%, and the estimated values were 0.35 ± 0.02 mg P/mg COD, 3.88 ± 0.41 mg COD/L, 3.35 ± 0.27 mg P/(mg COD * d^{-1}), and 1500 ± 72 mg COD/L, respectively. Compared to the results from the nonoptimal experimental condition, the practical identifiability and the estimation precision of the four parameters were improved.

1. Introduction

Over the past 30 years, ASMs have advanced greatly and have become important tools for improving biological wastewater treatment processes. A large number of chemical stoichiometric and kinetic parameters are included in these models. Although default values were suggested for these parameters in ASMs, they might be site-specific in different application. However, the best method for obtaining site-specific values for model parameters remains unresolved, hindering the practical application of ASMs [1]. Parameter estimation (PE), a technique that achieves the best fit between the simulated values and the observed data by choosing a set of parameters to minimize the objective function, is the most frequently adopted method for selecting appropriate parameter values [2, 3]. Reliable PE depends on the model's identifiability, which determines whether it is possible to obtain unique

parameter value according to the type or quality of the model structure (structural identifiability) or the measurement data (practical identifiability), in cases where some state variables can be measured or observed. Therefore, identifiability evaluation is significant for ASMs because of their complicated, nonlinear structure and the occurrence of strong interconnectedness among some parameters.

Structural identifiability is a property of the model structure and can be evaluated without any prior information regarding the values of the parameters and even before collecting any data. A number of methods have been established for the structural identifiability analysis of ASMs, including heterotrophic degradation [4], nitrification and denitrification [5], and hydrolysis [6]. Practical identifiability, however, depends heavily on the quantity and quality of the available experimental data. Thus, parameters that are structurally identifiable may be practically unidentifiable

because of poor or limited data [4]. A method based on the plots of sensitivity functions has been proven to be useful for the analysis of practical identifiability of some Monod-type models [4, 6]. Sensitivity calculations and graphical analysis are extremely difficult to generate for more complex models, however, limiting the application of this technology to some simple models [7]. The inability to quantify the interdependent correlation of the model parameters is also a deficiency of the method. Weijers and Vanrolleghem (1997) developed a procedure based on the FIM to study the identifiability of ASM1 in which D and mode criteria are used to find an identifiable parameter subset among numerous combinations [8]. Currently, this methodology has been successfully applied to ASM2d [9], modified ASM3 [10], and other kinetic models [3, 11, 12].

Considering the fact that practical identifiability and the reliability of PE depend on the quantity and quality of experimental measurements, OED was proposed to generate a high-performance experimental scheme. The FIM forms the core of the OED theory. The practical identifiability of parameters under different experimental conditions is represented by a variety of standard experimental designs that are defined according to different scalar functions of the FIM [13]. Moreover, the FIM inverse matrix is the parameter estimation error covariance matrix. Using the experimental data with the optimized initial substrate concentration [4], sampling frequency [14] and substrate pulse dosage strategy [6] have allowed for improvements on the practical identifiability and the precision of parameter evaluations for a variety of models. Strigul et al. (2009) proposed a practical guide for the optimal design of experiments in the Monod model [15].

Despite the recognition that the estimated parameter values are unreliable for an unidentifiable model, many ASMs practitioners have ignored the investigation of practical identifiability to focus only on how well their simulation results fit the experimental results. More importantly, relatively less work has been devoted to the identifiability and OED/PE of EBPR submodel than the other ASMs submodels. Our previous work revealed that all the parameters of the EBPR anaerobic submodel are structurally identifiable if prior information of one of the three parameters, K_{pp} , X_{PAO} , and q_{PHA} , is available [16]. Based on this finding, the current paper evaluates the practical identifiability of the EBPR anaerobic submodel in ASM2d by the FIM using data from anaerobic phosphorus release batch experiments and assuming the default value for K_{pp} . A modeling-based OED procedure is proposed, and the optimal experimental conditions are obtained according to the FIM D-criterion. Improvements in both the practical identifiability and PE precision through optimized experiments are demonstrated. Additionally, the bias problem associated with fixing the parameter K_{pp} is also assessed.

2. Materials and Methods

2.1. Model Description. The submodel of the EBPR anaerobic phosphorus release process is summarized in (1) according to the simplifications considered in [16]:

(1) Limitations or inhibitions of electron acceptor (DO or nitrate), nutrients, and alkalinity on process rates are not considered. These simplifications are frequently applied in a batch experiment.

(2) It is assumed that the storage process of X_{pp} would not be affected by the phosphorus content in PAOs.

(3) The model only considers substrate degradation processes. This means that the effect of biomass decay processes on the available data (measurements) could be excluded.

(4) Growth process of PAOs is neglected.

$$\begin{aligned} \frac{dS_{PO4}}{dt} &= Y_{PO4} q_{PHA} \frac{S_A}{K_A + S_A} \frac{X_{pp}/X_{PAO}}{K_{pp} + X_{pp}/X_{PAO}} X_{PAO} \\ \frac{dS_A}{dt} &= -q_{PHA} \frac{S_A}{K_A + S_A} \frac{X_{pp}/X_{PAO}}{K_{pp} + X_{pp}/X_{PAO}} X_{PAO}. \\ \frac{dX_{pp}}{dt} &= -Y_{PO4} q_{PHA} \frac{S_A}{K_A + S_A} \frac{X_{pp}/X_{PAO}}{K_{pp} + X_{pp}/X_{PAO}} X_{PAO} \end{aligned} \quad (1)$$

There are five parameters, Y_{PO4} , K_A , q_{PHA} , X_{PAO} , and K_{pp} , which need to be identified in this submodel. Zhang et al. (2010) demonstrated that parameters Y_{PO4} and K_A are globally structurally identifiable and that all the parameters are structurally identifiable, provided that one of q_{PHA} , K_{pp} , or X_{pp} is known [16]. In this paper, K_{pp} was selected as the “known” parameter to evaluate the practical identifiability of the other four parameters using S_{PO4} and S_A as experimental data available.

2.2. FIM-Based Method of Practical Identifiability Evaluation.

The FIM measures the variation in the output variables caused by a variation in the model parameters and thus summarizes the influence of each model parameter on the outputs. The details on FIM can be found elsewhere. Algebraically, the FIM is represented as follows:

$$FIM = \sum_{i=1}^N Y_p(t_i, p)^T Q_i^{-1} Y_p(t_i, p). \quad (2)$$

A FIM calculated for r output variables and p parameters is a $p \times p$ matrix, where i represents each sampling data point and N is the total number of samples.

As each column of the matrix represents a parameter, the determinant and the condition number of the FIM provide a reasonable measurement of the correlation of a set of parameters. Hence, less correlated parameters will easily provide a diagonal-dominant matrix. The FIM determinant can be used as a criterion (the D-criterion) for the practical identifiability evaluation of the parameters [9]: when $\text{Det}(FIM) = 0$, the unique parameter values cannot be obtained from the experimental data because some sensitivity functions are linear combination of each other; otherwise ($\text{Det}(FIM) \neq 0$), the parameters are practically identifiable. For identifiable parameters, a greater determinant value indicates smaller parameter estimation error.

Although numerical values for sensitivity are generally much less informative than an algebraic relation, algebraic sensitivity analysis is not feasible if the equations of the

model are complicated [9, 17]. Therefore, the Y_p values were determined numerically using the finite differences method according to (3). The central difference approach with 10^{-4} (0.01%) was selected as the perturbation factor for the sensitivity calculations according to Machado et al. (2009) [9].

$$Y_p(t_i, p) \approx \frac{y(t_i, p_j + \Delta p_j) - y(t_i, p_j - \Delta p_j)}{2\Delta p_j}. \quad (3)$$

2.2.1. Determination of Q. The measurement error covariance matrix Q of S_A and S_{PO4} is defined by (4), and the measurement error of S_A and S_{PO4} can be obtained by data fitting according to (5).

$$Q = \begin{bmatrix} s_A^2 & s_A s_{PO4} \\ s_A s_{PO4} & s_{PO4}^2 \end{bmatrix} \quad (4)$$

$$s^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N - 1}, \quad (5)$$

where y_i represents the experimental data with repeated analytical determination, \bar{y} represents the average of the measured values, and N represents the number of measurements.

2.3. Parameter Estimation and Optimal Experimental Design

(1) Cost Function and Algorithm for PE. The parameter estimation procedure was based on the minimization of a cost function (see (6)), calculated as the sum of squared error (SSE) between the experimental data and the model predictions.

$$J(p) = \sum_{i=1}^N (y(t_i) - y_m(t_i))^2. \quad (6)$$

A genetic algorithm using roulette choice, uniform cross, and simple mutation was applied to estimate the parameters. Crossover probability was 0.7, and variation probability was 0.1. The ranges of parameters Y_{PO4} , K_A , and q_{PHA} were determined as [0.20~0.60] mg P/mg COD, [2.00~6.00] mg COD/L, and [1.50~4.50] mg P/[mg COD · d⁻¹], respectively. These ranges were selected to represent a 50% deviation from the typical values according to the parameters uncertainty [18]. X_{PAO} was assumed to be in the range of [0~3000] mg COD/L. The initial population was set to 30. The initial population was defined with Latin hypercube sampling to make it evenly distributed in the parametric space. The reciprocal of the objective function was taken as the fitness function; thus, when the parameters were closer to the optimal solution, the objective function was smaller and the fitness value was greater. The algorithm was terminated after 300 iterations, and the optimal solution was the output along with the maximum fitness value.

(2) Confidence Interval of PE. The inverse matrix of the FIM (FIM^{-1}) is the parameter estimation error covariance matrix $COV(p_i)$, giving the Cramer-Rao lower bound of the parameter estimation error [19], as shown by (7). Therefore, the

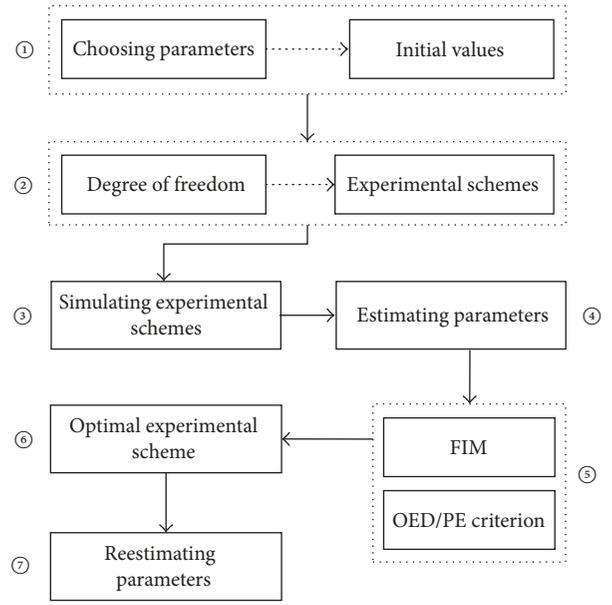


FIGURE 1: Optimal experimental design for parameter estimation procedure.

standard deviation of parameter estimation is the diagonal elements' square root of FIM^{-1} . The confidence interval of p_i is shown in (8).

$$COV(p_i) \geq FIM^{-1} \quad (7)$$

$$\left[p_i - t_{\alpha; (N-p)} \sqrt{FIM^{-1}_{ii}}, p_i + t_{\alpha; (N-p)} \sqrt{FIM^{-1}_{ii}} \right], \quad (8)$$

where t is the statistical Student's t -test with $\alpha = 95\%$ of confidence and $N - p$ degrees of freedom (number of experimental data points minus p parameters).

(3) Methodology for Optimal Experimental Design. The proposed methodology combines modeling technology with the theory of OED. In this technique, all experimental schemes will be first simulated on a computer before real experiments are conducted (Figure 1). Briefly, the proposed procedure for OED includes (1) selecting parameters to be estimated and determining their initial values, (2) choosing the experimental degree of freedom and the experimental schemes to be optimized, (3) simulating the experimental schemes, (4) estimating the parameters using the simulated data and the generated virtual data, (5) calculating the FIM and the values of the adopted OED/PE criterion, (6) determining the optimal experimental scheme according to the OED/PE criterion and performing the experiment, and (7) reestimating the parameters using the data from the optimal experiment. Details of the methodology will be further presented through its application to the design of an experiment on anaerobic phosphorus release. All calculations and simulations were implemented in MATLAB.

2.4. Experimental Procedure for Anaerobic Phosphorus Release. Batch experiments of anaerobic phosphorus release

were carried out in a cylindrical, plexiglass reactor with a working volume of 2.5 L. Activated sludge rich in polyphosphate was taken from a lab-scale A²/O activated sludge system treating synthetic wastewater and performing well in microbial phosphorus removal. Synthetic wastewater containing designed COD concentration was mixed with the activated sludge by magnetic stirring. Nitrogen gas was flushed into the reactor to maintain anaerobic conditions. Acetate was used as the carbon source. During phosphorus release, samples were taken at selected intervals, and the concentrations of COD (S_A) and P-PO₄³⁻ (S_{PO_4}) were determined offline.

3. Results

3.1. Practical Identifiability of the EBPR Anaerobic Submodel. S_A and S_{PO_4} were both measured 30 times in the same sample to determine the measurement error, as shown in Figure 2. According to these measurements, the measurement error covariance matrix Q of S_A and S_{PO_4} , which is a diagonal matrix due to S_A and S_{PO_4} 's independence, was determined using the following equation:

$$Q = \begin{bmatrix} s_A^2 & 0 \\ 0 & s_{PO_4}^2 \end{bmatrix} = \begin{bmatrix} 1.014 \times 10^2 & 0 \\ 0 & 9.796 \times 10^{-3} \end{bmatrix}. \quad (9)$$

In addition to Q , the initial values of all the parameters are needed for calculation of the FIM. Instead of the default values used in ASM2d, estimated values were used in this study. Therefore, the profiles of S_A and S_{PO_4} obtained from the anaerobic phosphorus release experiment were simulated

(Figure 3), using the model in (1), through the method described above in *Materials and Methods*. According to the results of structural identifiability [16], K_{PP} was fixed at the ASM2d default value of 0.01 mg X_{PP} /mg X_{PAO} , and the other parameters were estimated as follows: Y_{PO_4} : 0.36 mg P/mg COD, K_A : 3.27 mg COD/L, q_{PHA} : 3.41 mg P/[mg COD * d⁻¹], and X_{PAO} : 1560 mg COD/L.

The FIM of the EBPR anaerobic submodel was then calculated as in (10) in the sequence of Y_{PO_4} , K_A , q_{PHA} , and X_{PAO} . The determinant of this matrix was 8.048×10^5 .

$$\begin{bmatrix} 3.743 \times 10^6 & -8.492 \times 10^3 & 3.867 \times 10^5 & 5.923 \times 10^2 \\ -8.492 \times 10^3 & 2.150 \times 10 & -8.800 \times 10^2 & -1.337 \\ 3.867 \times 10^5 & -8.800 \times 10^2 & 4.015 \times 10^4 & 6.162 \times 10 \\ 5.923 \times 10^2 & -1.337 & 6.162 \times 10 & 9.521 \times 10^{-2} \end{bmatrix}. \quad (10)$$

Based on the knowledge that a full-rank FIM (Det(FIM) $\neq 0$) indicates practical identifiability and that a larger determinant value represents better practical identifiability, the determinant value of 8.048×10^5 clearly demonstrates that the parameters Y_{PO_4} , K_A , q_{PHA} , and X_{PAO} in the EBPR anaerobic submodel are practically identifiable using S_A and S_{PO_4} as the measured variables.

3.2. Reliability of Parameter Estimation and the Bias Problem Associated with Fixed K_{PP} . In addition to the practical identifiability, the reliability of parameter estimation can also be evaluated based on the FIM. The parameter estimation error covariance matrix is first calculated as the inverse matrix of the FIM.

$$E = \begin{bmatrix} 6.133 \times 10^{-5} & -2.807 \times 10^{-4} & -7.807 \times 10^{-4} & 1.198 \times 10^{-1} \\ -2.807 \times 10^{-4} & 5.211 \times 10^{-1} & 3.191 \times 10^{-2} & -1.159 \times 10 \\ -7.807 \times 10^{-4} & 3.191 \times 10^{-2} & 1.519 \times 10^{-2} & -4.528 \\ 1.198 \times 10^{-1} & -1.159 \times 10 & -4.528 & 2.033 \times 10^3 \end{bmatrix}. \quad (11)$$

Next, the confidence interval for parameter estimation with a 95% confidence level was calculated according to (8), and the results are shown in Table 1 in the *Nonoptimal experiment* column.

Generally, estimation errors within 20% of the parameter value are considered to be acceptable in the field of activated sludge process modeling. Obviously, the parameters Y_{PO_4} , q_{PHA} , and X_{PAO} were estimated with acceptable error levels in this study, as all three have error lower than 10%. However, the estimation of K_A needs to be improved due to its error of 39.34%.

To investigate the probable bias problem caused by fixing K_{PP} , the variations in the estimated parameter values and errors were calculated when the value of K_{PP} was increased by 10%, and the results are presented in Table 2. The very slight resulting variations of the estimated values of the four parameters illustrate that the fixed value of K_{PP} has almost

no effect on the parameter estimates. Although comparatively higher variations of the confidence intervals of Y_{PO_4} and X_{PAO} were observed, the influence of the fixed value of K_{PP} on the reliability of the parameter estimation is still insignificant. The bias problem would, therefore, not be evident when the value of K_{PP} is altered, and it is acceptable to fix K_{PP} to study the practical identifiability of the other parameters and estimate their values.

3.3. Optimal Design of the Anaerobic Phosphorus Release Experiment. The proposed OED/PE procedure was applied to the experimental design of the anaerobic phosphorus release process to improve the practical identifiability of the EBPR anaerobic submodel. The estimated values of the four identifiable parameters, listed in Table 1 in the *Nonoptimal experiment* column, were set as initial values. Previous studies have reported that the practical identification of a model

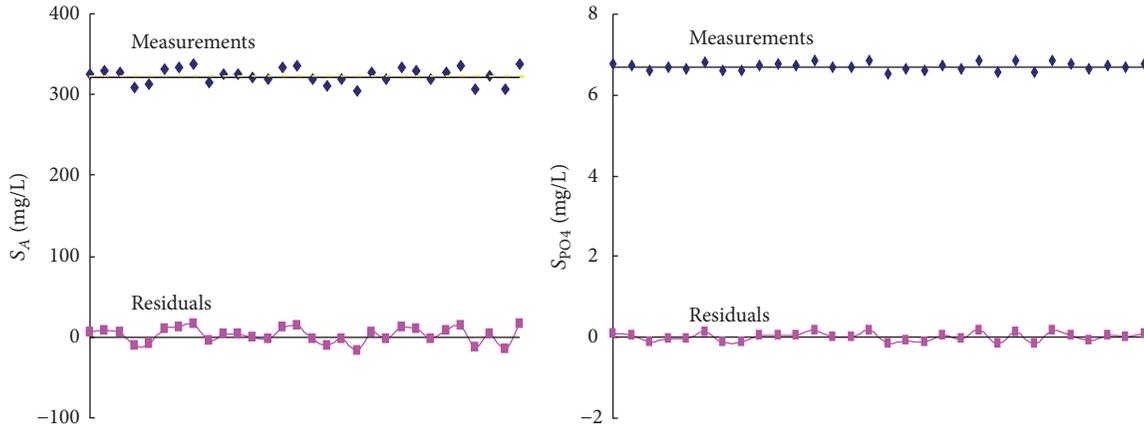


FIGURE 2: Data and residuals of measurement of S_A and S_{PO_4} .

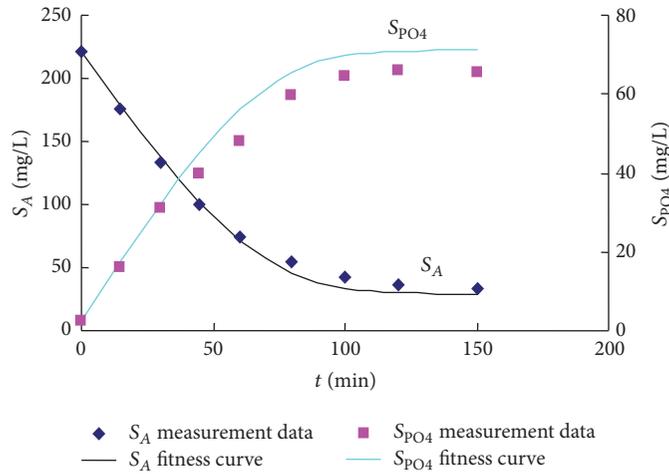


FIGURE 3: Experimental data and model fit.

TABLE 1: Comparison of the parameter estimates results.

Parameters (unit)	Nonoptimal experiment		Optimal experiment	
	Parameter value \pm confidence intervals	Error (percentage of parameter values)	Parameter value \pm confidence intervals	Error (percentage of parameter values)
Y_{PO_4} (mg (P)/mg (COD))	0.36 ± 0.01	3.88%	0.35 ± 0.02	4.86%
K_A (mg (COD)/L)	3.27 ± 1.29	39.34%	3.88 ± 0.41	10.67%
q_{PHA} (mg (P)/(mg (COD) * d ⁻¹))	3.41 ± 0.22	6.44%	3.35 ± 0.27	8.15%
X_{PAO} (mg (COD)/L)	1560 ± 80	5.15%	1500 ± 72	4.79%

TABLE 2: Variations of parameter estimates when K_{pp} was changed.

Parameter	Parameter value \pm confidence intervals	Variation of estimated value	Variation of error
Y_{PO_4}	0.36 ± 0.02	0%	14.43%
K_A	3.39 ± 1.31	+3.67%	-2.16%
q_{PHA}	3.43 ± 0.22	+0.59%	-2.17%
X_{PAO}	1520 ± 91	-2.56%	16.31%

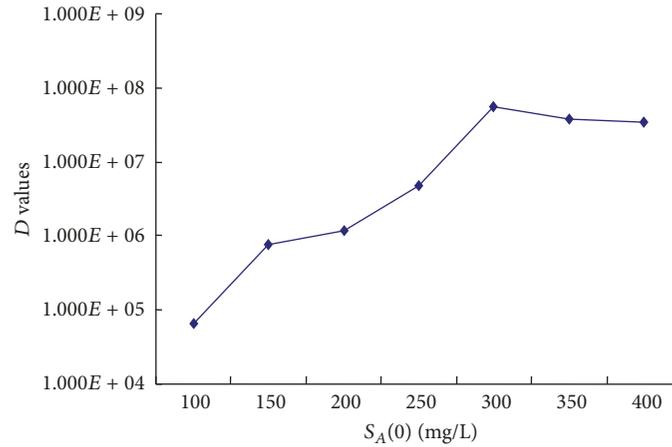


FIGURE 4: Values of D-criterion for the 7 experiments with different $S_A(0)$.

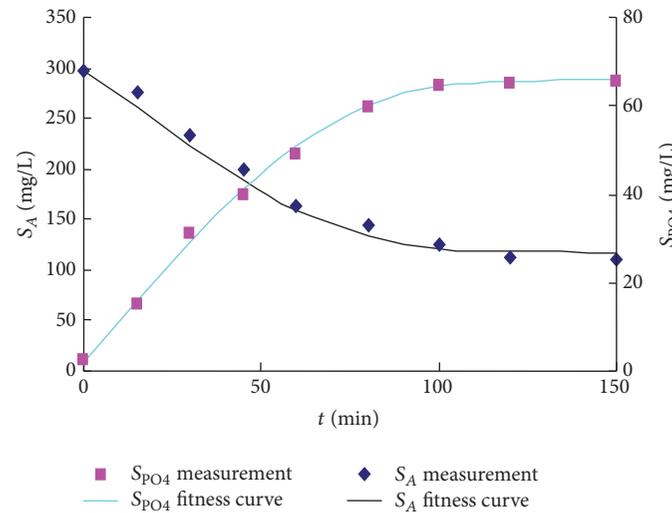


FIGURE 5: Experimental data and model fit for the optimal experiment.

parameter based on batch experiments is related to the initial substrate concentration [4]; therefore, the initial COD concentration $S_A(0)$ in the anaerobic phosphorus release experiment was set as the degree of freedom of the optimal experiment. Seven different $S_A(0)$ values, 100, 150, 200, 250, 300, 350, and 400 mg/L, were adopted. Virtual anaerobic phosphorus release experiments with the selected initial S_A concentrations and parameter values were run on a computer. “Noise” was added to the simulated data using a normal distribution to generate a virtual experimental dataset, which was then used as measured data to reestimate the parameters. Figure 4 shows the values of the D-criterion for different initial S_A concentrations. When $S_A(0)$ varied between 100 mg/L and 400 mg/L, the value of the D-criterion changed from 6.558×10^4 to 5.635×10^7 . According to the D-criterion, the optimal experimental condition was determined to be the initial S_A concentration of 300 mg/L, corresponding to the maximum D-criterion value of 5.635×10^7 .

A real batch experiment of anaerobic phosphorus release was carried out under the optimal experimental condition. By fitting the measured data (Figure 5), the parameters were

estimated, as shown in Table 1 in the *Optimal experiment* column.

Compared to the results from the nonoptimal experimental condition, the error in the estimated value of K_A was reduced significantly from 39.34% to 10.67%. Despite the slight increase of the error of Y_{PO4} and q_{PHA} , the overall practical identifiability and estimation precision of the parameter vector were improved through the optimal experimental design. Errors in the four parameter values are all approximately 10% or lower and reliable parameter estimation is achieved. The estimated values of the parameters differ only slightly from the ASM2d default values.

4. Discussion

Phosphorus removal is becoming an increasingly important goal for wastewater treatment systems. ASMs have been proven to be beneficial tools for the design and operational control of wastewater treatment systems. Careful calibration of these models might be essential for their correct application, and their calibration must be based on the evaluation

of model identifiability. However, insufficient research has focused on the identifiability of the EBPR model, although Zhang et al. (2010) have contributed important work on its structural identifiability [16]. Building on this work, the practical identifiability of the EBPR anaerobic submodel was investigated in this paper using a method based on the FIM. FIM analysis provides a quantitative evaluation of the practical identifiability of the model's parameters as well as a criterion for OED. Due to both its simple implementation and the low impact of subjective choices on the results [9, 20], this methodology has been widely applied to a variety of models. This method was used to determine the practically identifiable parameter subset of a modified ASM3 with a two-step nitrification-denitrification process [10]. In addition to ASMs, methods based on FIM analysis have been frequently adopted in integrated urban drainage water-quality modeling [20, 21]. This paper demonstrates unambiguously that the parameters Y_{PO_4} , K_A , q_{PHA} , and X_{PAO} in the EBPR anaerobic submodel are practically identifiable using S_A and S_{PO_4} as the measured variables, provided that X_{PP} is known. The precision of the parameter estimation was also acceptable. This finding indicates that, for reliable calibration of anaerobic P release model, it is necessary to conduct anaerobic phosphorus release experiments and simultaneous measurements of S_A and S_{PO_4} . Although X_{PP} can also be selected as a measurement variable for the parameter estimation, as the index of poly-P content in PAOs, its determination is more difficult than S_A and S_{PO_4} . Although some variables defined in ASMs can not be measured directly, these state variables might be determined through softy-sensor based on specific models and corresponding parameters estimates. Therefore, improvement in the precision and frequency of S_A and S_{PO_4} analysis is beneficial not only to the calibration of EBPR model but also to the indirect measurements of some state variables such as X_{PAO} and X_{PP} . Since one of the three parameters K_{PP} , X_{PAO} , and q_{PHA} needs to be "known" for the identification of the other parameters, fixing K_{PP} might be more reasonable because this parameter represents the maximum capacity of poly-P in PAOs, which is more constant than the other two process variables. The same default value of K_{PP} (0.01) in ASM2d under different temperature (10°C and 20°C) might be taken as a clue. In this paper, the calculations also show that it is valid to fix K_{PP} as default value to identify and estimate the other four parameters.

In addition to the FIM-based method, Brun et al. (2001) reported a systematic approach for evaluating the practical identifiability of the parameters of large environmental models [22], and this method was applied to a large lake model [11]. However, the application of this approach to ASM2d resulted in different conclusions: one study reported that a maximum of nine model parameters should be (conditionally) identifiable from the data available [18], while another reported that 13 model parameters are identifiable [23].

Optimal experimental design can undoubtedly be useful in reducing the number of necessary experimental measurements and improving the precision of model parameter value determinations. In addition to linear regression models and several simple nonlinear regression models [24], the application of local optimal experimental design

for different modifications of the Monod model has been recommended [25]. Furthermore, a complete theoretical examination of optimal designs for the Monod model has also been presented [24]. In this paper, the practical identifiability of the parameters in the EBPR anaerobic submodel was significantly improved through the optimal design of the anaerobic phosphorus release experiment, which provided a two-order-of-magnitude increase in the D-criterion value. According to D-criteria, the higher initial S_A was used in the optimal experiment and the fitness of S_{PO_4} was improved obviously. Consequently, the estimation accuracy of K_A was improved, which might counteract the possible adverse effect of the higher final S_A in the optimal experiment. The overall precision of the estimated parameter set was also improved by using data from the optimized experiment. The estimated values of the parameters differ only slightly from the ASM2d default values. However, OED procedures are not yet widely used in practice, most likely because practitioners often lack the necessary mathematical background or experience in numerical methods. Therefore, the development of a practical guide and application of commercial ASMs simulators would allow for more widespread application of the OED technique.

The selection of an appropriate criterion is an important issue in the field of OED. A different choice of criterion produces different optimal experiments. In addition to the D-criterion used in this paper, other FIM-based criteria include the A criterion, the modA criterion, the E criterion, and the modE criterion. In some cases, normalization would be necessary to weaken the dependence of the criterion value on the magnitude of the involved parameters, allowing comparisons among subsets with the same size but with different parameters. A new RDE criterion calculated as the ratio of the normalized D-criterion to the modE criterion has been proposed and applied successfully to the study of the IWA-ASM2d model [9]. In conclusion, the adoption of a suitable criterion is essential to the successful application of OED. Therefore, further investigation on the application of other optimal design criteria to the OED of anaerobic biological P release process is necessary.

5. Conclusions

The FIM found in this study was a full-rank matrix with a determinant value of 8.048×10^5 , clearly demonstrating that the parameters Y_{PO_4} , K_A , q_{PHA} , and X_{PAO} in the EBPR anaerobic submodel are practically identifiable using S_A and S_{PO_4} as the measured variables and assuming that K_{PP} was determined as a default value. It was found that fixing K_{PP} to study the practical identifiability of the other parameters and to estimate their values was valid. The proposed OED procedure for parameter estimation was applied successfully to anaerobic phosphorus release experiments by using a computer model to optimize the experiment. The practical identifiability and the estimation precision of the four parameters in the EBPR anaerobic submodel were significantly improved through the optimal experimental design procedure. The estimated values of the parameters differed only slightly from the ASM2d default values. Errors in the estimated values of

the four parameters were all approximately 10% or below, and reliable parameter estimation was achieved.

Nomenclature

ASMI:	Activated sludge model number 1
ASM2d:	Activated sludge model number 2d
ASM3:	Activated sludge model number 3
ASMs:	Activated sludge models
COV(P_0):	Parameter estimation error covariance matrix
EBPR:	Enhanced biological phosphorus removal
FIM:	Fisher information matrix
FIM ⁻¹ :	Inverse matrix of FIM
K_A :	Half-saturation coefficient for S_A (g COD m ⁻³)
K_{pp} :	Half-saturation coefficient for polyphosphate (g X_{pp} g X_{PAO}^{-1})
OED:	Optimal experimental design
p :	Vector of p parameters
PAOs:	Phosphorus accumulating organisms
PE:	Parameter estimation
PHA:	Polyhydroxyalkanoates
Q_i^{-1} :	Covariance matrix of the measurement noise
q_{PHA} :	Rate constant for the storage of X_{PHA} (gPHA g X_{PAO}^{-1} d ⁻¹)
S_A :	Concentration of fermentation products (g COD m ⁻³)
S_{PO_4} :	Concentration of inorganic soluble phosphorus (g P m ⁻³)
t :	Time
$t_{\alpha;(N-p)}$:	Statistical Student's t -test, with $\alpha = 95\%$ of confidence and $N - p$ degrees of freedom
X_{PAO} :	Concentration of phosphorus accumulating organisms (g COD m ⁻³)
y :	Output variable
Y_p :	Output sensitivity function matrix
X_{pp} :	Concentration of polyphosphate (g P m ⁻³)
Y_{PO_4} :	PP requirement per PHA stored (g P g COD ⁻¹).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was partially supported by the National Natural Science Foundation of China (51609028), Natural Science Foundation of Chongqing (CSTC2016jcyjA0506 and 2015CSTC-JBKY-01605), and Chongqing Postdoctoral Science Foundation (no. Xm2017141).

References

- [1] E. Liwarska-Bizukolc and S. Ledakowicz, "Determination of kinetic and stoichiometric Parameters of activated sludge models," *Environment Protection Engineering*, vol. 37, no. 3, pp. 73–83, 2011.
- [2] R. Urban and R. Szetela, "Calibration of the activated sludge model with genetic algorithms. Part I. Calibration results," *Environment Protection Engineering*, vol. 33, no. 1, pp. 31–49, 2007.
- [3] R. Urban and R. Szetela, "Calibration of the activated sludge model with genetic algorithms. Part II. Analysis of results," *Environment Protection Engineering*, vol. 33, no. 1, pp. 51–64, 2007.
- [4] A. Holmberg, "On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities," *Mathematical Biosciences*, vol. 62, no. 1, pp. 23–43, 1982.
- [5] S. Julien, J. P. Babary, and P. Lessard, "Theoretical and practical identifiability of a reduced order model in an activated sludge process doing nitrification and denitrification," *Water Science and Technology*, vol. 37, no. 12, pp. 309–316, 1998.
- [6] B. Petersen, K. Gernaey, and P. A. Vanrolleghem, "Practical identifiability of model parameters by combined respirometric-titrimetric measurements," *Water Science and Technology*, vol. 43, no. 7, pp. 347–355, 2001.
- [7] N. Checchi, E. Giusti, and S. Marsili-Libelli, "PEAS: A toolbox to assess the accuracy of estimated parameters in environmental models," *Environmental Modeling and Software*, vol. 22, no. 6, pp. 899–913, 2007.
- [8] S. R. Weijers and P. A. Vanrolleghem, "A procedure for selecting best identifiable parameters in calibrating activated sludge model no 1 to full-scale plant data," *Water Science and Technology*, vol. 36, no. 5, pp. 69–79, 1997.
- [9] V. C. Machado, G. Tapia, D. Gabriel, J. Lafuente, and J. A. Baeza, "Systematic identifiability study based on the Fisher Information Matrix for reducing the number of parameters calibration of an activated sludge model," *Environmental Modeling and Software*, vol. 24, no. 11, pp. 1274–1284, 2009.
- [10] E. Giusti, S. Marsili-Libelli, and A. Spagni, "Modelling microbial population dynamics in nitrification processes," *Environmental Modeling and Software*, vol. 26, no. 7, pp. 938–949, 2011.
- [11] P. Reichert and P. Vanrolleghem, "Identifiability and uncertainty analysis of the River Water Quality Model No. 1 (RWQM1)," *Water Science and Technology*, vol. 43, no. 7, pp. 329–338, 2001.
- [12] S. Marsili-Libelli and E. Giusti, "Water quality modelling for small river basins," *Environmental Modeling and Software*, vol. 23, no. 4, pp. 451–463, 2008.
- [13] A. Munack, "Optimal feeding strategy for identification of Monod-type models by fed-batch experiments," in *Computer Applications in Fermentation Technology Modelling and Control of Biotechnological Processes*, pp. 195–204, 1989.
- [14] C. Vialas, A. Cheruy, and S. Gentil, "An experimental approach to improve the Monod model identification," in *Modelling and Control of Biotechnological Processes: 1st IFAC Symposium*, pp. 155–160, 1985.
- [15] N. Strigul, H. Dette, and V. B. Melas, "A practical guide for optimal designs of experiments in the Monod model," *Environmental Modeling and Software*, vol. 24, no. 9, pp. 1019–1026, 2009.
- [16] T. Zhang, D. Zhang, Z. Li, and Q. Cai, "Evaluating the structural identifiability of the parameters of the EBPR sub-model in ASM2d by the differential algebra method," *Water Research*, vol. 44, no. 9, pp. 2815–2822, 2010.
- [17] J. P. Norton, "Algebraic sensitivity analysis of environmental models," *Environmental Modeling and Software*, vol. 23, no. 8, pp. 963–972, 2008.

- [18] R. Brun, M. Kühni, H. Siegrist, W. Gujer, and P. Reichert, "Practical identifiability of ASM2d parameters - Systematic selection and tuning of parameter subsets," *Water Research*, vol. 36, no. 16, pp. 4113–4127, 2002.
- [19] E. Walter and L. Pronzato, *Identification of parametric models from experimental data*, Springer Verlag, Heidelberg, Germany, 1999.
- [20] G. Freni, G. Mannina, and G. Viviani, "Identifiability analysis for receiving water body quality modelling," *Environmental Modeling and Software*, vol. 24, no. 1, pp. 54–62, 2009.
- [21] G. Freni, G. Mannina, and G. Viviani, "Assessment of the integrated urban water quality model complexity through identifiability analysis," *Water Research*, vol. 45, no. 1, pp. 37–50, 2011.
- [22] R. Brun, P. Reichert, and H. R. Künsch, "Practical identifiability analysis of large environmental simulation models," *Water Resources Research*, vol. 37, no. 4, pp. 1015–1030, 2001.
- [23] M. V. Ruano, J. Ribes, D. J. W. De Pauw, and G. Sin, "Parameter subset selection for the dynamic calibration of activated sludge models (ASMs): Experience versus systems analysis," *Water Science and Technology*, vol. 56, no. 8, pp. 107–115, 2007.
- [24] H. Dette, V. B. Melas, A. Pepelyshev, and N. Strigul, "Robust and efficient design of experiments for the Monod model," *Journal of Theoretical Biology*, vol. 234, no. 4, pp. 537–550, 2005.
- [25] I. Y. M. Smets, K. J. E. Versyck, and J. F. M. Van Impe, "Optimal control theory: A generic tool for identification and control of (bio-)chemical reactors," *Annual Reviews in Control*, vol. 26, pp. 57–73, 2002.



Hindawi

Submit your manuscripts at
www.hindawi.com

