Parameter estimation procedure for complex non-linear systems: calibration of ASM No.1 for N-removal in a full-scale oxidation ditch

A. Abusam*, K. J. Keesman*, G. van Straten**, H. Spanjers*** and K. Meinema**

*Systems and Control Group, Wageningen University, Bomenweg 4, 6703 HD Wageningen, The Netherlands
**AEST, Environmental Technology, Wageningen University, P.O. Box 8129, 6700 EV Wageningen, The Netherlands
***DHV Water BV, P.O. Box 484, 3800 AL Amersfoort, Laan 1914 No. 35, The Netherlands

Abstract

When applied to large simulation models, the process of parameter estimation is also called calibration. Calibration of complex non-linear systems, such as activated sludge plants, is often not an easy task. On the one hand, manual calibration of such complex systems is usually time-consuming, and its results are often not reproducible. On the other hand, conventional automatic calibration methods are not always straightforward and often hampered by local minima problems. In this paper a new straightforward and automatic procedure, which is based on the response surface method (RSM) for selecting the best identifiable parameters, is proposed. In RSM, the process response (output) is related to the levels of the input variables in terms of a first- or second-order regression model. Usually, RSM is used to relate measured process output quantities to process conditions. However, in this paper RSM is used for selecting the dominant parameters, by evaluating parameters sensitivity in a predefined region. Good results obtained in calibration of ASM No.1 for N-removal in a full-scale oxidation ditch proved that the proposed procedure is successful and reliable.

Keywords Calibration; carrousel; modelling; oxidation ditches; parameter estimation; ASM No. 1

Introduction

Activated sludge plants are typical examples of complex non-linear systems. For on-line application, there is a need for adjusting the parameters initially obtained from literature or previous experiments, such that the model output fits the available data. The process of parameter estimation, when applied to large simulation models, is also called calibration. Calibration can be done either manually (hand calibration) or through automatic optimisation algorithms. Hand calibration, which is still commonly used in practice, is a trial and error method in which values of the parameters are changed manually and the difference between the measured and predicted values is evaluated visually. Because it consumes a lot of time and the reproduction of its results is always difficult, hand calibration is used only when automatic calibration is not available.

In contrast, the application of automatic calibration is also not always straightforward. For instance, automatic calibration of a model with multiple output variables and more than 5-7 parameters will usually show local minima. Selecting the best identifiable parameters and finding the optimum region(s) to obtain good initial guesses are the difficult parts in the calibration procedure. In order to remedy this a method like the response surface method (RSM) (see for example Box and Draper, 1987) can be useful. Myers and Montgomery (1995) define RSM as “a collection of statistical and mathematical techniques useful for developing, improving and optimising processes.” In RSM, the process response (output) is related to the levels of the input variables in terms of a first- or second-order regression model. In this paper, RSM is used as a “regional sensitivity analysis” tool, where the input
variables are the parameters and the outputs the sum of squares of the residuals. As a result of this, a set of best identifiable parameters can be found, while the optimum region(s) for these parameters can also be allocated from the surface analysis of the relationships found by RSM. Then, within the optimum region(s), any automatic calibration method can be applied for optimising the selected set of parameters.

The objective of this paper is to present a new straightforward automatic procedure, based on RSM, that can be used in parameter estimation for complex non-linear systems. The structure of this paper is as follows. First, the procedure is briefly described. Then, results of applying this procedure in calibration of ASM No.1 for N-removal in a full-scale oxidation ditch is presented. Finally, the paper ends with conclusions about the success and reliability of the proposed method.

Proposed procedure

Generally, RSM is used to relate measured process output quantities to process conditions. However, in this paper RSM will be used to evaluate parameters sensitivity in a predefined region and not, as is common practice, in a local point in the parameter space. On the basis of this, dominant parameters can easily be selected. In addition, optimum region(s) for these parameters can be found by carrying out a surface analysis of the predicted sum of squares of the residuals. The proposed procedure can be summarised in the following six steps: (i) specify plant layout and model, (ii) collect in/output and operational data, (iii) estimate initial state conditions from past data, (iv) use RSM to select dominant parameters (run, for instance, a two-level simulation experiment and select the sensitive parameters from the resulting first- or second-order meta-model), (v) apply formal parameter estimation method and (vi) evaluate estimation results. In the next sections the procedure will be illuminated to a full-scale example. The proposed procedure was applied in the calibration of ASM No.1 for N-removal in a full-scale carrousel type WWTP situated in Rotterdam, the Netherlands.

Step 1: Specify plant layout and model.

The plant is designed for dry weather flow of 3583 m$^3$/h and rain flow of 12,800 m$^3$/h (301,500 inhabitants at 54 g BOD$_5$/capita · d). It consists of two parallel lines, each having two primary settlers, one selector, three circular secondary settlers (diameter = 529 m and

---

**Figure 1** Layout of the carousel
depth = 2 m), one carousel of capacity equal to 13,000 m³ (see Figure 1) and four aerators (estimated oxygen input for each is 70.2 kg O₂/h at high speed, and 40.2 kg O₂/h at low speed). The plant was modelled as a loop of 10 equal-volume CSTR’s plus a secondary settler. ASM No. 1 (Henze et al., 1987), which is considered to be the state of the art for dynamic modelling of activated sludge plants with COD and N removal, was used for modelling the biochemical processes, while the double exponential model, developed by Takács et al. (1991) was used for modelling the secondary settler. The number of CSTR’s was chosen on the basis of the number of aerators in the ditch and the ditch layout. This number was limited to 10 CSTR’s, of which six for representing the aerated compartments and four for representing the anoxic compartments. In a previous paper it has been shown that increase of number of CSTR’s beyond 10 does not significantly affect the predicted effluent quality of the oxidation ditch (Abusam and Keesman, 1999).

**Step 2: Collect in/output and operational data**

The calibration data set consist of 10-day measurements performed, at dry weather conditions (DHV Water, 1993), from 27 July to 7 August 1992. Variables measured were daily influent and effluent COD_{tot}, TKN, NH₄-N, NO₃-N and temperature (mean: 22.3°C). COD_{tot} and TKN were calculated from the settleable, colloidal and dissolved fractions. On 29/7, 30/7, 4/8 and 7/8 volume proportional sampling was carried out every 2 hours, from which it was concluded that variations in concentration of COD_{tot} and TKN were very small over the day. Other operational data are as follows. Average daily flow equals to 48681 m³/d; 67% of the carousel influent, on volume basis, is directed to the first aerator and the rest to the fourth aerator (compartment 8, in Figure 1). Further, the rate of waste activated sludge (WAS) is 780–870 m³/d, the rate of recycled activated sludge (RAS) is equal to the influent flow rate, the rate of internal recirculation is about 83 times the influent flow rate, and the sludge age is about 9 days.

The influent composition expressed in ASM No.1 components were obtained using the influent characterisation values provided with the data. Provided data also included the 19 kinetic and stoichiometric parameter values, which were then used as a starting point of the procedure.

**Step 3: Estimate initial state conditions from past data**

Initial substrate and biomass concentrations were unknown. For estimating these, we have performed a 100-day steady state simulation, using the averages of the provided measurements and default parameter values (DHV Water, 1993). The aeration constant used in the steady state simulation was estimated from the given information about oxygen input (based only on COD removal) and it was equal to 150,000 m³/d. Note here that we have estimated the aeration constant (k = K_L a · V_A) because it can be estimated more accurately than K_L or V_A. In fact, neither K_L nor V_A can be identified individually due to the hyperbolic relationship between these parameters (Abusam et al., 1999).

**Step 4: Use RSM to select dominant parameters**

In this study, given the measurements of effluent NH₄-N and effluent NO₃-N, there are two responses; that is, the sum of squared errors for effluent NO₃-N and effluent NH₄-N. However, for total nitrogen removal, the response that needs to be analysed is the weighted sum of squared errors for both effluent NO₃-N and effluent NH₄-N. Note that COD_{tot} has not been taken into account, because it was found to be unreliable.

In order to apply RSM, it is wise to limit the number of computations by restricting the number of parameters to no more than 10–12. A pre-sensitivity analysis could help in performing this pre-selection. Many parameters can be tried for calibrating N-removal in such
cases (Weijers et al., 1996; Keesman et al., 1996). Since our goal here is to illustrate the method, we have restricted the analysis to those parameters that can be of potential interest: $\eta_g$ and $\eta_h$, $K_{NO}$ and $K_{NH}$ (coefficients for growth and hydrolysis under anoxic conditions, nitrate and ammonia half-saturation coefficients, respectively). Since we do not know the actual oxygen input, we will also try to estimate the aeration constant ($k$), which is the product of $K_L a$ and $V_A$ (the actually aerated volume). As mentioned above, $k$ can be estimated more accurately than $K_L a$ or $V_A$. However, instead of estimating $k$ directly, we have estimated a multiplication factor ($m = k/k_0$), which we can use subsequently for the calculation of the actual aeration constant ($k = m \cdot k_0$). Here $k_0$ is the initial guess for the actual aeration constant ($k_0 = 150,000$ m$^3$/d). Thus we have five potential parameters that need to be identified: $m$, $K_{NO}$, $K_{NH}$, $\eta_g$, and $\eta_h$.

The next step is to design a simulation experiment such that a second-order regression-type meta-model can be fitted. As recommended by Box and Draper (1987), a two-level factorial design with cubic, star and centre portions has been chosen. The normalised second-order composite design around the nominal parameter vector has been specified as follows: cube portion: ($\pm 1$, $\pm 1$, $\pm 1$, $\pm 1$, $\pm 1$), star portion: ($\pm \alpha$, 0, 0, 0, 0), (0, $\pm \alpha$, 0, 0, 0), (0, 0, $\pm \alpha$, 0, 0), (0, 0, 0, $\pm \alpha$, 0), (0, 0, 0, 0, $\pm \alpha$) and centre portion: (0, 0, 0, 0, 0). Here $\alpha = \sqrt{5}$, so that all points, except the centre point, are situated on a ball in the parameter space. Hence the total number of simulation runs equals to 43 ($= 2^5 + 5.2 + 1$).

Table 1 presents the coded levels for the five chosen parameters. As it can be seen from this table, values of the parameters $K_{NO}$ and $K_{NH}$ were limited to the ranges reported in literature, while relatively wide ranges are explored for $m$ (indirectly the aeration constant) $\eta_g$ and $\eta_h$.

From the 43 simulation runs the following second-order regression model resulted for the weighted sum of squares

\[
V = 187.8 -611.4x_1 -120.4x_2 +184.0x_3 +21.3x_4 +34.2x_5 \\
= 187.8 -611.4(\pm 29.8) -120.4(\pm 29.8) +184.0(\pm 29.8) +21.3(\pm 29.8) +34.2(\pm 29.8) \\
+369.7x_1^2 +63.1x_2^2 +18.7x_3^2 +3.9x_4^2 +3.7x_5^2 +165.2x_1x_2 \\
(\pm 46.0) (\pm 46.0) (\pm 46.0) (\pm 46.0) (\pm 46.0) (\pm 46.0) \\
-130.2x_1x_3 +42.5x_1x_4 +30.7x_1x_4 -6.3x_2x_3 +37.7x_2x_4 +35.7x_3x_5 \\
(\pm 34.1) (\pm 34.1) (\pm 34.1) (\pm 34.1) (\pm 34.1) (\pm 34.1) \\
+37.9x_3x_4 +36.7x_3x_5 +36.3x_4x_5 \\
(\pm 34.1) (\pm 34.1) (\pm 34.1) \\
\text{(1)}
\]
where the values between the parenthesis are the standard deviations, and $V_{\text{NO}_3}$ and $V_{\text{NH}_4}$ are the sum of squared errors for effluent NO$_3$–N and effluent NH$_4$–N, respectively.

Comparison of the standard deviations (shown between parenthesis) and corresponding coefficients (which are the parameter sensitivity coefficients) in the model (Eq. (1)) indicates which parameters are sensitive and which are not (coefficient < 2 std. dev.). Equation (1) shows that the sensitive parameters are only $x_1$, $x_2$, $x_3$, $x_1^2$, $x_2^2$, and the interactions $x_1x_2$ and $x_1x_3$. Thus, it can be concluded that the sensitive parameters here are: $x_1$, $x_2$ and $x_3$. Consequently $x_4$ and $x_5$ are insensitive and can be set to default values.

For further analysis and evaluation of this result, the general format for this second-order model:

$$V = C + Lx + x^T H x,$$

will be used, where $C$ is a scalar term, $L$ is a column vector consisting of the coefficients of the linear terms and $H$ is a symmetrical matrix formed from the coefficients of the quadratic terms.

Figure 2 presents the contour plots (of Eq. (1)) of the weighted sum of squared errors for both NO$_3$–N and NH$_4$–N, where the weighting factor is equal to 10. The factor of 10 was chosen because individual responses indicated that the sum of squared errors for NO$_3$–N is about 10 times higher than that for NH$_4$–N. In fact, this is an engineering guess.

Careful examination of Figure 2 reveals that the optimum value of $m$ is between 1.25 to 2.125 ($x_1: -0.5$ to $1.25$), optimum value of $\eta_g$ is between 0.3 and 1.0 ($x_2: -1.5$ to $2.24$), and optimum value of $\eta_h$ is between 0.2 and 0.9 ($x_3: -2.24$ to $1.5$). In addition to this, directions of the principal axes of the ellipses can be obtained from the eigenvalue decomposition of the matrix $H$ for this so called meta model (see Appendix B), and the centre of the model response surface ($x_S = -\text{inv}(H) L/2$) (Box and Draper, 1987) can be used as an initial guess. This is especially useful when the number of parameters is large. Since it exponentially increases with the increase in number of parameters, computational time will be the real problem when the number of parameters become very high. But this problem can be solved, as suggested by Box and Draper (1987), by using the so-called fractional factorial designs instead of the full factorial design. In the next step, this information about the optimum regions will be used within a constrained optimisation routine for estimating the parameters $m$, $\eta_g$ and $\eta_h$. 

![Figure 2](https://iwaponline.com/wst/article-pdf/43/7/357/429891/357.pdf)
Step 5: Apply formal parameter estimation method

All the 19 parameters in ASM No.1 were set to default values obtained from DHV Water (1993), except the three best identifiable parameters: $m$, $\eta_g$, and $\eta_h$. Values of these parameters were constrained as specified in the previous step.

Figure 3 presents the results of the second least-squares optimisation trial with the following results: $m = 1.84$ ($k = 1.84 \times 150,000 = 276,000$ m$^3$/d), $\eta_g = 1.0$ and $\eta_h = 0.32$. Two trials of optimisation were conducted because it appeared that the parameter values of the first trial ($m = 1.89$, $\eta_h = 0.92$ and $\eta_h = 0.31$) led to significantly different initial state conditions when compared to those obtained from prior knowledge. Using these recalculations with conditions led to a relatively poor fit. In the second trial, using the recalculated initial conditions from the first trial, however, there was no change in the fit even when the results of least-square optimisation were used again for recalculating the initial conditions. From these two trials, we noted that the estimated initial concentrations of heterotrophic and autotrophic biomass are almost the same for both trials. However, the initial concentration of the slowly biodegradable substrate ($X_s$) obtained from steady state simulation using the default parameters was very high (418 mg COD/l) compared to that obtained using parameters from the first trial (64 mg COD/l) or the second trial (68 mg COD/l). Hence, there was clearly a need for recalculating the initial conditions.

In order to see the effect of the initial concentration of NO$_3$–N and NH$_4$–N, we carried out a number of simulations with different initial values for NO$_3$–N and NH$_4$–N. Results of these simulations indicated that over a reasonable range initial nitrogen concentrations have no effect on the curve fit. From steady state simulations, however, we have noted that the predicted initial nitrogen concentration (about 9.5 mg NO$_3$–N/l and 0.5 mg NH$_4$–N/l) is

<table>
<thead>
<tr>
<th>Number of CSTR's</th>
<th>$m$</th>
<th>$\eta_g$</th>
<th>$\eta_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.89</td>
<td>0.99</td>
<td>0.35</td>
</tr>
<tr>
<td>10</td>
<td>1.84</td>
<td>1.0</td>
<td>0.32</td>
</tr>
<tr>
<td>16</td>
<td>1.85</td>
<td>0.99</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Figure 3 Predicted and measured concentration in effluent (a) nitrate and (b) ammonia
always different from the measured value (5.21 mg NO$_3$-N/l and 2.57 mg NH$_4$-N/l). If we again have a look at Figure 3, we can see that nitrogen concentrations in the first day seem to be inconsistent with the rest of the data. Therefore, we indicate that nitrogen values reported for the first day may be incorrect. The reason for that is probably an experimental error.

Furthermore, the effect of the number of CSTR’s on the parameter estimation was evaluated. For that we have constructed two extra models that consist of 8 and 16 CSTR’s. Subsequently, the same formal optimisation procedure was applied. As shown in Table 2 the number of CSTR’s has no significant effect on the estimated values.

**Step 6: Evaluate estimation results**

First of all, evaluation of the residuals (Figure 3) shows that the amplitude of the residual is acceptable. However, systematic errors appear indicating incorrectness of model structure and/or experimental data. The data is too short to deduce more profound conclusions. In addition, uncertainty analysis of the estimates was carried out (see also Lukasse et al., 1996, 1997). Neglecting the residual correlation structure, local parameter estimates

$$\text{Cov}(\hat{\theta}_N) = \hat{\sigma}_\epsilon^2 (J^T J)^{-1},$$

(3)

uncertainties can be approximately found from the resulting covariance matrix:

$$\frac{d e(t_j | \theta)}{d \theta_j} \quad \text{with } k = 1, ..., N; \ j = 1, ..., p, \ \text{and } \hat{\sigma}_\epsilon^2 \quad \text{is the variance of the residuals},$$

where $J$ is the Jacobi matrix

$$M^T \text{Cov} \hat{\theta}_N M = \Lambda.$$

(4)

In order to evaluate the estimation uncertainty in more detail, the dominant parameter directions were investigated via eigenvalue decomposition of the covariance matrix; that is where $M$ is the orthogonal matrix of eigenvectors and $\Lambda$ is a diagonal matrix with eigenvalues.

From the covariance matrix and eigenvalue decomposition results, which are presented in Appendix C, we can conclude the following. Firstly, a small value of the elements of $\Lambda$ (eigenvalues) indicates that all the three parameters can be estimated with a high reliability. This is also indicated by the small standard deviations (square root of the diagonal of the covariance matrix). Secondly, the third element of $\Lambda$ has the smallest value (0.0063), which indicates that $\eta_h$ slightly dominates the model behaviour. Finally, there is no relatively big difference between the values of the elements in $\Lambda$. This simply means that all the three parameters are sensitive, a result that has also been found from the previous RSM analysis.

The aeration constant ($k$, related to $m$) can also be estimated roughly in another way from the average $C$ and $N$ removals (about 215 g COD/m$^3$ and 27 g N/m$^3$, respectively). This provides a cross-check about the adequacy of the method. Using Eq. (5), the oxygen input by one aerator at full capacity can be estimated to be 1774 kg O$_2$/d.

$$O_2 = Q(COD_{tot} \cdot (1 - Y_H) + 4.57TN) / f,$$

(5)

where $Q=(48681/2)$ m$^3$/d, $Y_H = 0.67$ and $f = 2.76$. Here $f$ is the sum of the relative working capacity per day for all the aerators (see Appendix A). From simulations it has been found that the average oxygen deficit, around the first and second aerators, is about 7.2 g/m$^3$. Thus the $k$ (oxygen input/oxygen deficit) is about 246,389 m$^3$/d. Comparison of this rough estimate (246,389 m$^3$/d) with the value obtained in step 5 (276,000 m$^3$/d) indicates that the aeration constant was estimated very accurately.
Conclusions

A new procedure was proposed for estimating parameters of a non-linear system. Success in calibrating ASM No. 1 (a typical non-linear system) for N-removal in a full-scale carousel WWTP proved that the proposed procedure is successful and reliable. The loop-of-CSTR’s model proved to be an acceptable model for describing the behaviour of oxidation ditches under process conditions, with respect to influent and effluent quality. As under clean water conditions (Abusam et al., 1999), the aeration constant \(k = K_L a V_A\) can be estimated with a very high accuracy \((k = 276,000 \pm 63.16 \text{ m}^3/\text{d})\) under process conditions also.

Acknowledgements

Discussions by Dr. John B. Copp, Co-ordinator of IAWQ Task Group on Respirometry, are appreciated.

Appendix A

Given \(H\) (see Eq. (2)), eigenvalue decomposition of \(H\) gives:

\[
\begin{pmatrix}
0.96 & 0.13 & -0.22 & -0.15 & -0.02 \\
0.24 & -0.67 & 0.70 & -0.08 & -0.03 \\
-0.16 & -0.45 & -0.47 & -0.72 & -0.16 \\
0.06 & -0.41 & -0.36 & 0.63 & -0.58 \\
0.04 & -0.41 & -0.34 & 0.25 & 0.80
\end{pmatrix}
\]

where \(M\) is the orthogonal matrix of eigenvectors, \(\Lambda\) is a diagonal matrix with eigenvalues for \(V\) and “\(\text{diag}\)” defines diagonal matrix operations.

The rows of \(V\) correspond with the parameter sequence of Table 1. Eigenvectors give directions of the principal axes of the ellipses, where lengths of these principal axes are inversely proportional to the square root of the absolute magnitude of corresponding eigenvalue. Hence, it can be concluded that the eigenvector \([0.96 \ 0.24 \ -0.16 \ 0.06 \ 0.04]^T\), which is roughly dominated by the first two elements, defines a dominant direction. Notice that
this is also verified by Figure 2 (plot $x_1 - x_2$), using the rough simulation data.

$$\text{Cov} \hat{\Theta}_N = 10^{-6} \begin{bmatrix} 0.1773 & -0.0354 & 0.0189 \\ -0.0354 & 0.0175 & -0.0068 \\ 0.0189 & -0.0068 & 0.0105 \end{bmatrix} M = \begin{bmatrix} -0.9718 & 0.2332 & 0.0345 \\ 0.2075 & 0.7766 & 0.5949 \\ -0.1119 & -0.5853 & 0.8031 \end{bmatrix}$$

$$\Lambda = \text{diag}(10^{-6}[0.1870 \ 0.0120 \ 0.0063])$$

Appendix C: Evaluation of parameter uncertainties
where “diag” defines diagonal matrix operations. The rows of $M$ correspond to weights on $m$, $\eta_g$ and $\eta_h$, respectively.

References