Designing and performing experiments for model calibration using an automated iterative procedure

D.J.W. De Pauw* and P.A. Vanrolleghem*

*BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, 9000 Gent, Belgium (E-mail: *dirk.depauw@biomath.ugent.be*; *peter.vanrolleghem@biomath.ugent.be*)

Abstract Optimal experimental design for parameter estimation involves complex mathematical and practical steps in order to obtain a model with sufficiently accurate parameters. This paper proposes a methodology where user interaction is only required at the beginning of the experimental design procedure. All subsequent steps are carried out automatically, including: (1) finding the optimal experiment, (2) performing the experiments in practice, and (3) recalibrating the model. A software extension to an existing modelling and simulation package which performs this automatic procedure is also presented. Finally an illustration for the calibration of a one-step nitrification model using respirometric data is given. In this case, the application of the proposed procedure resulted, after three iterations, in considerably better confidence intervals on the parameter estimates, within the desired boundaries.

Keywords Automation; calibration; mathematical modelling; optimal experimental design; parameter estimation

Water Science & Technology **Vol 53 No 1 pp 117–127** © IWA Publishing 2006

Introduction

Experimental data are a prerequisite for the development and correct use of biological models. Indeed, data are required in almost every step of the model building exercise. This involves selecting an appropriate model structure and tuning the model parameters. In this paper we will assume that a correct model structure is available and we will focus more on the aspect of model calibration. Model calibration or parameter estimation is considered here as the exercise of minimizing the weighted sum of squared errors between the model prediction and the available data by changing the model parameters (θ), this to ensure that the model predictions coincide with the experimental data. The quality of the parameter estimates depends largely on the quality and quantity of the available data. Often experiments are performed which deliver data of low quality resulting in a model calibration with very uncertain and correlated parameters. In order to ensure that an experiment will result in high quality data, an experimental design procedure can be used. The term experimental design will be described differently by experimenters and mathematicians. Experimenters often view experimental design as an art involving much intuition while mathematicians tend to view it as a statistical problem. Probably the best experimental design incorporates both visions and combines expert knowledge and intuition with mathematical reasoning.

The nonlinear dynamic systems for which experimental design will be discussed in this paper are described by the general state-output equations:

$$\begin{aligned} \frac{dx}{dt} &= g(x, \theta, u, t), x(t_0) = x_0 \\ y &= f(x, \theta, u, t) \end{aligned} \tag{1}$$

doi: 10.2166/wst.2006.014

where x is a vector of state variables, y a vector of outputs, θ a vector of parameters and u a vector of inputs. For this system parameter estimation can be formulated as the minimization of the following weighted quadratic objective functional:

$$J(\theta) = \sum_{i=1}^{N} (y_i(\theta) - y_i)^T Q_i (y_i(\theta) - y_i) \tag{2}$$

in which y_i and $y_i(\theta)$ are vectors of N measured values and model predictions at times t_i respectively. Q_i is a square matrix of user supplied weights often chosen as the inverse measurement error covariance matrix. The expected value of the objective functional for a parameter set slightly different than the optimal one can be written as (Vanrolleghem *et al.*, 1995):

$$E[J(\theta + \partial\theta)] = E[J(\theta)] + \partial\theta^T \left[\sum_{i=1}^{N} \left(\frac{\partial y}{\partial \theta}(t_i) \right)^T Q_i \left(\frac{\partial y}{\partial \theta}(t_i) \right) \right] \partial\theta \tag{3}$$

A reliable minimum $J(\theta)$, and thus a good model calibration, requires that the difference between the expected value $E[J(\theta + \partial\theta)]$ and $E[J(\theta)]$ is as large as possible. This implies that the term between large square brackets in Equation (3) should be maximized. This term is called the Fisher Information Matrix (FIM) and expresses the information content of the experiment by combining sensitivity functions and measurement errors.

Under certain conditions (uncorrelated white measurement noise), the inverse of the FIM gives the lower bound of the covariance matrix of the estimated parameter vector according to the Cramer–Rao inequality (Ljung, 1999; Walter and Pronzato, 1997). This relationship is illustrated in Figure 1 for a 2-parameter estimation problem. These figures represent the confidence regions of two parameters (θ_1 and θ_2). The size, shape and orientation of the confidence ellipse are determined by the eigenvalues and eigenvectors of the FIM. The largest axis of the confidence ellipse is inversely proportional to the square root of the smallest eigenvalue (λ_{min}), while the smallest axis is inversely proportional to the square root of the largest eigenvalue (λ_{max}). In this way properties of the FIM determine the properties of the confidence region and thus the accuracy of the parameter estimates.

The information content of the experiment can be optimised by considering different measures of the FIM. Table 1 lists these criteria. The D- and A-optimal design criteria aim at minimizing the volume of the confidence ellipse, illustrated in Figure 1a. The modified E design criterion on the other hand aims at reducing parameter correlations by getting the shape of the confidence region as close to a circle as possible (Figure 1b). The best value one can obtain for the modified E criterion is 1, and this has been

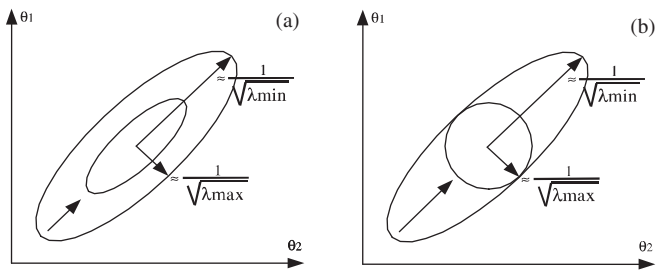

Figure 1 Effect of different FIM design criteria (D-criterion: left, modified E criterion, right) on the size and shape of the parameter confidence region

Table 1 Different optimal design criteria based on FIM properties (Mehra, 1974)

Name	Criterion
A-optimal design	min[tr(FIM-1)]
Modified A-optimal design	max[tr(FIM)]
D-optimal design	max[det(FIM)]
E-optimal design	max[λmin(FIM)]
Modified E-optimal design	min[λmax(FIM)/λmin(FIM)]

achieved for certain experimental design problems. However, one must be aware that such optimum only guarantees that the confidence region is a circle, but it can be a very large circle (Dochain and Vanrolleghem, 2001).

Optimal experimental design based on FIM properties has been applied numerous times using an iterative procedure. The first step of this procedure is the calibration of the model based on initial data. Next, new experiments are proposed, simulated and evaluated until the optimal experiment is found. After the optimal experiment is performed in reality, the model is recalibrated and the procedure repeated if required.

The main reason why the above mentioned procedure is still not used very often is that it involves different, rather complex mathematical and practical steps in order to obtain the calibrated model. This paper will therefore focus on the development of a methodology which automates the entire procedure, making expert interventions almost unnecessary. The paper is organized in three sections. First, the experimental setup and the model used to illustrate the methodology are described. The next section is devoted to the development of the automatic experimental design procedure and software. To conclude, a practical case study is presented which illustrates the proposed procedure.

Methods

In order to illustrate the automatic experimental design procedure, respirometric experiments were performed using an integrated sensor for monitoring aerobic and anoxic activated sludge activities (Sin *et al.*, 2003). Only a short description of the sensor will be given here since the focus of this paper is more on the experimental design procedure. Full details about the experimental setup can be found in the above mentioned paper.

The setup of the integrated sensor shown in Figure 2 consists of an aeration (2.5 l) and a respiration (1.0 l) vessel. A cooling system is used to control the temperature of the reactors. Data acquisition, pH-control and data processing are implemented in LabVIEW (LabVIEW 6.1i, National Instruments). In the aeration vessel dissolved oxygen, nitrate and pH are measured and in the respiration chamber only dissolved oxygen is measured. OUR data were calculated from the two DO-trajectories as described in Petersen *et al.* (2001).

The data acquisition frequency of the sensors is set to 3 seconds. High frequency noise, known to be present in the weak analog signals of the electrodes of the setup, is filtered using a lowpass Savitzky–Golay least square polynomial filter (Press *et al.*, 1992) through a Labview Matlab script node (Matlab R12, The MathWorks Inc.). The pH is controlled within a narrow pH band of ± 0.03. Strong autocorrelations were found in the oxygen uptake rate data up to a lag of ± 40 seconds. This was corrected in LabVIEW by sub-sampling each 13^{th} data point.

The oxygen uptake rate (OUR) measured in the respiration chamber was used to calibrate a one-step nitrification model based on ASM1 (Henze *et al.*, 2000) where ammonium (NH_4^+-N) is converted to nitrate (NO_3^--N) assuming that the NH_4^+-N to NO_2^--N (nitrite)

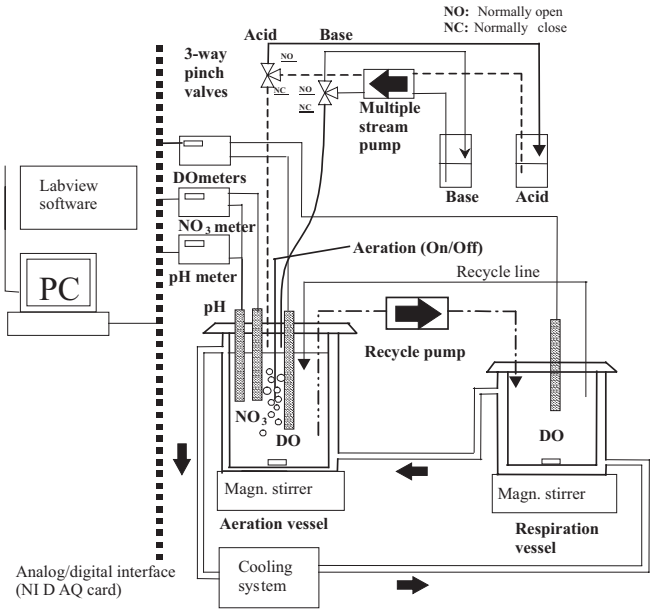

Figure 2 Illustration of the integrated sensor set-up (from Sin *et al.*, 2003)

conversion is the rate limiting step. The OUR can be considered as the sum of the exogenous OUR and the endogenous OUR given by Equation 4.

$$\underbrace{OUR = \frac{4.57 - Y_{NH_4}}{Y_{NH_4}} \times \left(1 - e^{-\frac{t - t_{pulse}}{\tau}}\right) \times \mu_{\max,NH_4} \times \frac{S_{NH_4}}{K_{NH_4} + S_{NH_4}} \times X_{NH_4}}_{OUR_{exogenous}} + \underbrace{(1 - f_P) \times b_H \times X_H}_{OUR_{endogenous}} \quad (4)$$

In this equation Y_{NH_4} (mgCOD.mg^{-1}N) is the yield of the autotrophic biomass X_{NH_4} (mgCOD.l^{-1}); $\mu_{\max,NH_4}$ (min^{-1}) the maximum growth rate of the autotrophic biomass; K_{NH_4} (mgN.l^{-1}) the Monod half-saturation constant for ammonium; τ (min) an OUR transient term *time constant* (Vanrolleghem *et al.*, 2004); S_{NH4} (mgN.l^{-1}) the ammonium concentration; f_P (-) the inert particulate fraction of the biomass and b_H (min^{-1}) the decay coefficient of the heterotrophic biomass X_H (mgCOD.l^{-1}).

Experiments were performed with pulse additions of NH_4^+-N added to the aeration vessel using a pump. For these experiments biomass from the Maria Middelares WWTP (Gent, Belgium) was used.

Results and discussion

Development of the automatic optimal experimental design procedure and software

Figure 3 illustrates the proposed automated optimal experimental design procedure. In order to apply this procedure in practice an extension was programmed to the modelling and simulation software package West (Vanhooren *et al.*, 2003). This extension makes use of existing West software modules like simulation, optimization and sensitivity analysis.

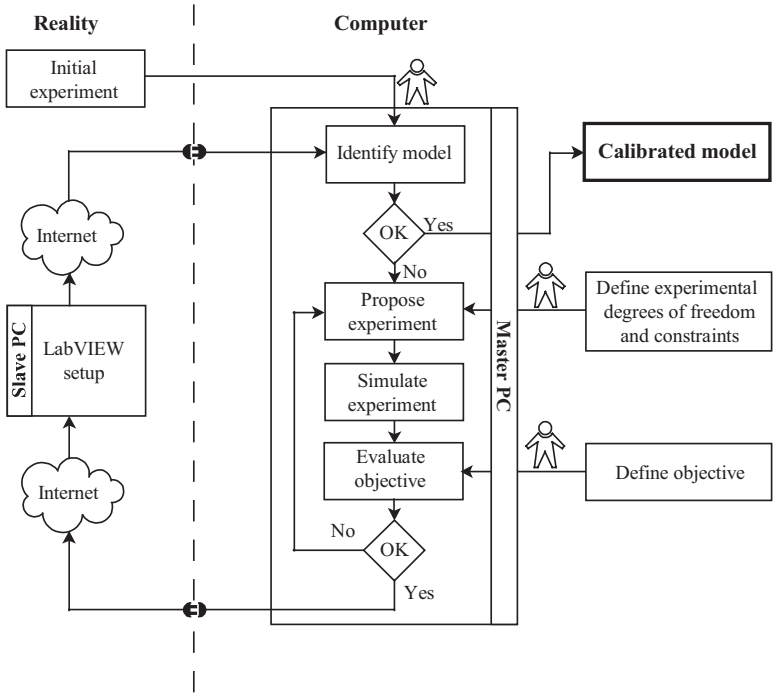

Figure 3 General automatic optimal experimental design procedure (adapted from Dochain and Vanrolleghem, 2001)

Two main parts can be distinguished in the procedure. The left side of Figure 3 represents reality where experiments are conducted and data collected. The right side of the figure represents the computer system where models are calibrated and experiments are simulated and their potential information content evaluated. User interaction is only required once at three points in the proposed procedure, these are indicated by a human symbol in Figure 3.

The starting point of the experimental design procedure is the identification of the model using data of an initial experiment. At this point user interaction is still very important as automation seems as yet unfeasible. First, the user has to select a model structure which is appropriate. Next, the identifiability of the proposed model should be studied in order to find the parameter combinations which can be uniquely estimated based on the available data. This can be done by analysing the output sensitivities of the measured variables to the parameters (Brun *et al.*, 2002; Weijers and Vanrolleghem, 1997). Parameters for which the output sensitivities are very small or highly correlated should not be estimated but fixed at a default value. Once the parameter subset is determined, the model can be fitted to the data of the initial experiment.

User interaction is also required to specify the experimental degrees of freedom and constraints. These can be classified into two types: measurements and manipulations. The user needs to specify which variables can be measured and which measurement frequencies and measurement locations in time or space are allowed possibly within certain constraints. Also, the measurement error characteristics of each measurement need to be specified. Experimental manipulations concern all actions that can be taken to manipulate the experimental setup in some way. These include the initial conditions and experimental inputs like imposed flow, temperature or concentration profiles.

The choice of the experimental degrees of freedom has important implications on the choice of optimizer that has to be used to solve the design problem. For continuous optimization problems classical optimization algorithms can be used. However, if experimental degrees of freedom have been defined which can only take certain discrete values, other algorithms which can cope with this kind of restrictions have to be used. Ideal candidates for this are genetic algorithms (Goldberg, 1989) because of their ability to optimize combined continuous and discrete problems.

The final decision the user has to make is to define the objective that will be used to evaluate the proposed experiments. These objectives can be any of the FIM properties that are listed in Table 1. In addition to that, multiple objectives (including, for instance, experimental cost) can be evaluated using a multi-objective genetic algorithm optimizer (Fonseca and Fleming, 1998). In contrast to single objective optimization, multi-objective optimization does not give a single solution but a set of solutions that can be Pareto optimal. This means that each solution of this set is optimal in the sense that no improvement can be achieved in one objective that does not lead to a degradation in at least one of the remaining objectives. However, since this type of optimization produces multiple solutions, it cannot be used in an automatic experimental design procedure.

Once the experimental degrees of freedom/constraints and the objective are specified the experimental design procedure can be started. Based on the initially calibrated model new experiments are proposed using the specified experimental degrees of freedom. These experiments are simulated and the objective evaluated. The optimization of the experimental degrees of freedom is continued until an optimal experiment is found, optimal in the sense that it minimizes or maximizes the objective, depending on the choice of the objective.

If the optimal experiment has been found, it can be performed in reality. The developed software is able to transfer the chosen values for the experimental degrees of freedom of the optimal experiment through TCP-sockets over the internet. In this way it can communicate with every software package which is equipped with this form of communication interface. In this work LabVIEW (National Instruments) is used to perform the experiment and collect the data. The experimental design module transfers the experimental degrees of freedom and signals LabVIEW to perform the experiment automatically. After the experiment is finalized LabVIEW transfers the acquired data through the same network interface to the West experimental design module. Once the data is received, the model is recalibrated using the parameter values of the already available model as initial guesses. It is then decided if the procedure needs to be repeated. Depending on the chosen objective criterion, this decision can be made on the basis of the accuracy of the parameter estimates or the correlations among them. In this work, the proposed procedure is stopped when the largest correlation is lower than 0.5 if the modified-E criterion is used, or when the largest 95% confidence interval is smaller than $\pm 20\%$ of the parameter value if another FIM criterion is used. The procedure is also stopped when the correlation or confidence intervals for the latest experiment show no further improvements compared to the previous one. These proposed stopping criteria are useful in practice but further research is certainly desirable.

Practical illustration of the developed procedure

To illustrate this procedure it is applied to the model and experimental setup described at the beginning of this paper. As a first step in the procedure, the model was calibrated with OUR data of an initial experiment (Figure 4). For this experiment, lasting 35 minutes, a pulse of 0.5 mg NH_4^+-N was added to the aeration vessel at 12 minutes after the start of the experiment.

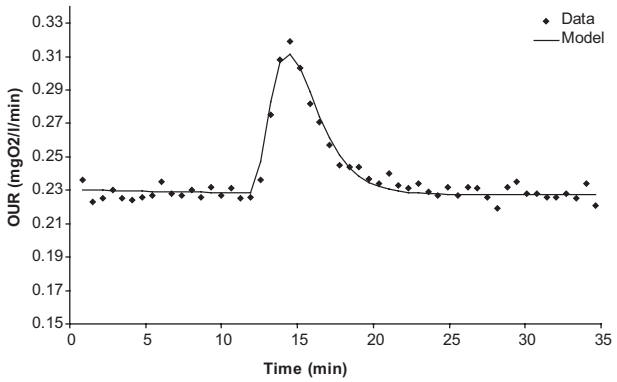

Figure 4 OUR data and model fit for the initial experiment

Five parameters were selected for parameter estimation: $\mu_{max,NH4}$, K_{NH4}, Y_{NH4}, τ, X_H. This selection was made based on the fact that the measured variable (OUR) was sufficiently sensitive to these parameters. In order to fit the endogenous respiration rate, the heterotrophic biomass concentration X_H was selected for parameter estimation while b_H and f_P were kept at their default value of 1.39E-04 min^{-1} and 0.2 respectively. The autotrophic biomass concentration X_{NH4} is also assumed to be known and was fixed here at a value of 40 $mg.l^{-1}$.

Figure 4 shows the model fit to the data of the initial experiment. The calibrated parameters and their confidence intervals are listed in Table 2. From the table it is clear that the confidence intervals of the parameters are extremely large. In order to improve this, the automatic optimal experimental design procedure was run.

The next step in the proposed procedure is to fix the experimental degrees of freedom and constraints. It was decided to perform an experiment with two pulse additions of NH_4^+-N where the total amount of NH_4^+-N was limited to 5 mg in order to limit biomass growth. The duration of the experiment was fixed to 110 minutes and the experimental degrees of freedom to be optimized were chosen to be: (1) the amount of the second addition; (2) the time instant of the second pulse addition. The first pulse of NH_4^+-N was always added at the start of the experiment.

Next the objective function was selected. The D-optimal design criterion was chosen in order to decrease the overall confidence region of the parameter estimates. Since the optimization problem only consists of continuous experimental degrees of freedom, the Simplex method (Nelder and Mead, 1965) could be used as optimization algorithm. This algorithm converges faster for this type of problem.

The experiment that was found to be optimal for the model with parameters obtained from the first experiment was the addition of all 5 mg NH_4^+-N in the first pulse at the start of the experiment while a negligible amount had to be added using a second pulse. For this experiment the D criterion was 2.07E+15.

Once the optimal experiment was found, the values of the experimental degrees of freedom for the optimal experiment were transferred automatically to the LabVIEW setup (slave PC) and the experiment was started. When the experiment ended 110 minutes later, the data were transferred to the master PC. As soon as the data were received, the model was recalibrated using the parameter values of the already available model as initial estimates. Again, the calibration was performed using the Simplex optimization method. Figure 5 shows the data of this experiment together with the model fit. Table 2 shows the parameter values found after calibration together with their 95% confidence

Table 2 Values and confidence intervals for the estimated parameters based on the data of the initial and the designed experiments

Estimated parameters	Initial experiment		First designed experiment		Second designed experiment	
	Value	95% confidence interval	Value	95% confidence interval	Value	95% confidence interval
μ_{max,NH_4} (min^{-1})	3.28E-03	3.05E-02 (927%)	9.41E-04	3.23E-05 (3.44%)	7.31E-04	4.73E-05 (6.47%)
K_{NH_4} ($mg.l^{-1}$)	4.28	2.18E+02 (5118%)	8.65E-03	1.82E-03 (21.1%)	6.66E-03	1.55E-03 (23.3%)
Y_{NH_4} ($mg.mg^{-1}$)	4.16E-02	1.95 (4682%)	1.01	3.45E-02 (3.40%)	8.24E-01	5.36E-02 (6.50%)
τ (min)	1.93	2.04 (106%)	1.39	1.52E-01 (10.9%)	1.04	1.21E-01 (11.6%)
X_H ($mg.l^{-1}$)	1854	18.89 (1.02%)	1472	11.75 (0.82%)	1561	16.22 (1.04%)

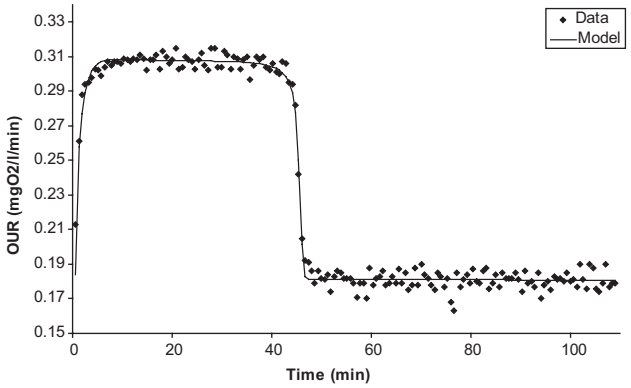

Figure 5 Model calibration based on the OUR data for the first designed experiment

interval. From the table it is clear that the confidence in the parameters has increased considerably.

Next, a second iteration of the procedure was started because the largest estimation error was still above 20%. Another optimal experiment was found and this time it consisted of one pulse of 4.17 mg NH_4^+-N at the start of the experiment and another pulse of 0.83 mg NH_4^+-N after 48.32 minutes. This corresponds to an addition just after the first pulse of NH_4^+-N is completely taken up by the biomass. The FIM D criterion for this experiment increased to 1.26E+23. Vanrolleghem *et al.* (1995) also obtained an optimal experiment with extra pulse addition after the substrate was almost exhausted for a respirometric experiment similar to the one described here. Again, the experimental degrees of freedom were transferred to the slave PC and the experiment was performed. After the experiment, the data were sent to the master PC and the model was recalibrated. Figure 6 and Table 2 show the data and the results of the model calibration. From the table it is clear that the parameters have not changed too much but, probably surprisingly, the errors on the parameter estimates have increased. Although the information content of the second optimal experiment was higher (see above), the model fit was less good than the fit on the first optimal experiment as a result of a model structure problem. Indeed, the model was unable to optimally fit the acquired data and the increased residual error propagated in the higher parameter estimation errors. Since no further improvement was made in parameter confidence, the optimal experimental design procedure was stopped at this point.

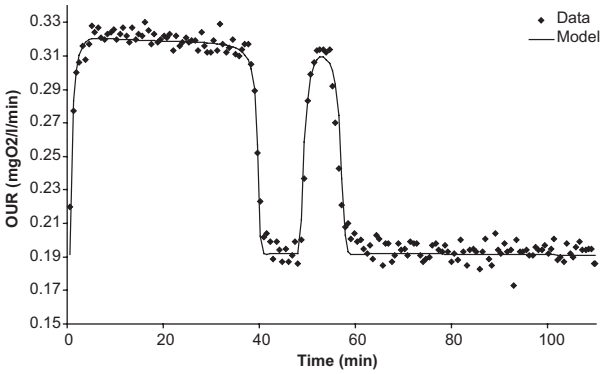

Figure 6 Model calibration based on the OUR data for the second designed experiment

Looking at the parameter values in Table 2, it is clear that the yield parameter has a quite different value from the often reported 0.24 $mgCOD.mg^{-1}N$ for autotrophic biomass. The area below the OUR profile corresponds to $(4.57\text{-}Y_{NH}) \times S_{NH}(0)$ where $S_{NH}(0)$ is the added NH_4^+-N concentration which is assumed to be known (Petersen *et al.*, 2001). For the first designed experiment, the integrated surface corresponds to 5.55 $mgO_2.l^{-1}$ which is identical to the calculated value (5.56 $mgO_2.l^{-1}$) using a yield of 1.01 mgCOD.mgN-1 and assuming a known concentration of 1.65 $mgN.l^{-1}$ (5 $mgNH_4^+$-N in 3.2 l). If we assume a default yield of 0.24 $mgCOD.mg^{-1}N$ and calculate the NH_4^+-N concentration that would correspond to an oxygen utilization of 5.55 $mgO_2.l^{-1}$, it becomes 1.28 $mgN.l^{-1}$, significantly less than the assumed value. The gap in NH_4^+-N could be explained by several things: (1) adsorption of ammonium on biomass, (2) assimilation of ammonium by heterotrophic biomass, (3) inaccuracies of the ammonium addition pump. Indeed, the effect of these processes can be significant because only a small amount of NH_4^+-N is added to the system.

Conclusions

An automatic optimal experimental design procedure was proposed. In this procedure, user interaction and expert knowledge is only required at the beginning and all subsequent steps can be performed automatically. These include: (1) finding the optimal experiment; (2) performing the experiments in practice; and (3) recalibrating the model. The proposed procedure was implemented in an existing modelling and simulation package (West) that interacts with a slave LabVIEW-PC. This slave PC controls the experiment and can be located anywhere in the world.

The proposed procedure was applied successfully to the calibration of a one-step nitrification model using respirometric data. After an initial model calibration, an optimal experiment for the parameter values obtained, was found and performed automatically by transferring the experimental degrees of freedom through the internet to the LabVIEW experimental setup. The acquired data was sent back to the experimental design module where the model was automatically recalibrated and a second and final experiment designed, performed and used for a final recalibration. This resulted in significantly smaller confidence intervals for the estimated parameters.

Acknowledgements

This research was undertaken with the support of the Special Research Fund of Ghent University and the IST Program of the European Union (TELEMAC, project no. IST-2000-28156). We would also like to thank Gürkan Sin for his valuable contribution.

References

Brun, R., Kuhni, M., Siegrist, H., Gujer, W. and Reichert, P. (2002). Practical identifiability of ASM2d parameters – systematic selection and tuning of parameter subsets. *Water Research*, **36**(16), 4113–4127.

Dochain, D. and Vanrolleghem, P.A. (2001). *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, London, UK.

Fonseca, C.M. and Fleming, P.J. (1998). Multiobjective optimization and multiple constraint handling with evolutionary algorithms – Part I: A unified formulation. *IEEE Transactions on Systems, Man, and Cybernetics*, **28**, 26–37.

Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA, USA.

Henze, M., Gujer, W., Mino, T. and Loosdrecht, M.C.M. (2000). *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*, IWA Scientific and Technical Report no. 9. IWA, London, UK.

Ljung, L. (1999). *System Identification: Theory for the User*, Prentice Hall, Englewood Cliffs, NJ, USA.

Mehra, R.K. (1974). Optimal input signal for parameter estimation in dynamic systems – survey and new results. *IEEE Transactions on Automatic Control.*, **19**(6), 753–768.

Nelder, J.A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, **7**, 308–313.

Petersen, B., Gernaey, K. and Vanrolleghem, P.A. (2001). Practical identifiability of model parameters by combined respirometric-titrimetric measurements. *Water Science and Technology*, **43**(7), 347–355.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition. Cambridge University Press, Cambridge, UK.

Sin, G., Malisse, K. and Vanrolleghem, P.A. (2003). An integrated sensor for the monitoring of aerobic and anoxic activated sludge activities in biological nitrogen removal plants. *Water Science and Technology*, **47**(2), 141–148.

Vanhooren, H., Meirlaen, J., Amerlinck, Y., Claeys, F., Vangheluwe, H. and Vanrolleghem, P.A. (2003). WEST: Modelling biological wastewater treatment. *Journal of Hydroinformatics*, **5**(1), 27–50.

Vanrolleghem, P.A., Van Daele, M. and Dochain, D. (1995). Practical identifiability of a biokinetic model of activated sludge respiration. *Water Research*, **29**(11), 2561–2570.

Vanrolleghem, P.A., Sin, G. and Gernaey, K.V. (2004). Transient response of aerobic and anoxic activated sludge activities to sudden substrate concentration changes. *Biotechnology and Bioengineering*, **86**(3), 277–290.

Weijers, S.R. and Vanrolleghem, P.A. (1997). A procedure for selecting best identifiable parameters in calibrating activated sludge model no.1 to full-scale plant data. *Water Science and Technology*, **36**(5), 69–79.

Walter, E. and Pronzato, L. (1997). *Identification of Parametric Models: From Experimental Data*. Springer Verlag, Heidelberg, Germany.