ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

An Efficient Approach to Automate the Manual Trial and Error Calibration of Activated Sludge Models

Gürkan Sin,[1,2] Dirk J.W. De Pauw,[1] Stefan Weijers,[3] Peter A. Vanrolleghem[1,4]

[1]BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, B-9000, Ghent, Belgium

[2]Department of Chemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby, Denmark; telephone: +45-45252960; fax:+45-45932906; e-mail: gsi@kt.dtu.dk

[3]Waterschap De Dommel, DA Boxtel, The Netherlands

[4]model*EAU*, Département de génie civil, Pavillon Pouliot, Université Laval, Québec, Canada

Received 6 September 2007; revision received 26 November 2007; accepted 28 November 2007

Published online 20 December 2007 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.21769

ABSTRACT: An efficient approach is introduced to help automate the rather tedious manual trial and error way of model calibration currently used in activated sludge modeling practice. To this end, we have evaluated a Monte Carlo based calibration approach consisting of four steps: (i) parameter subset selection, (ii) defining parameter space, (iii) parameter sampling for Monte Carlo simulations and (iv) selecting the best Monte Carlo simulation thereby providing the calibrated parameter values. The approach was evaluated on a formerly calibrated full-scale ASM2d model for a domestic plant (located in The Netherlands), using in total 3 months of dynamic oxygen, ammonia and nitrate sensor data. The Monte Carlo calibrated model was validated successfully using ammonia, oxygen and nitrate data collected at high measurement frequency. Statistical analysis of the residuals using mean absolute error (MAE), root mean square error (RMSE) and Janus coefficient showed that the calibrated model was able to provide statistically accurate and valid predictions for ammonium, oxygen and nitrate. This shows that this pragmatic approach can perform the task of model calibration and therefore be used in practice to save the valuable time of modelers spent on this step of activated sludge modeling. The high computational demand is a downside of this approach but this can be overcome by using distributed computing. Overall we expect that the use of such systems analysis tools in the application of activated sludge models will improve the quality of model predictions and their use in decision making.

Biotechnol. Bioeng. 2008;100: 516–528.

KEYWORDS: activated sludge model (ASM); calibration; validation; Monte Carlo; ASM2d; wastewater treatment

Correspondence to: G. Sin

This article contains Supplementary Material available at http://www.interscience.wiley.com/jpages/0006-3592/suppmat.

Introduction

The typical practice of modeling full-scale activated sludge systems comprises numerous steps starting from (i) the definition of the modeling objective to (ii) data collection and quality check, (iii) formulation of the mathematical models for different plant units/processes, for example for hydraulics, biological, physical and chemical processes, (iv) steady-state and dynamic model calibration, (v) validation and model use (Sin et al., 2005). The relatively short history of dynamic activated sludge modeling (which started at the end of 1970s and was accelerated with the release of ASM1 by Henze et al., 1987) clearly shows that a significant degree of expert and process knowledge was, and still is, being used during different steps in the modeling process, for example which sub-model to choose for describing different plant units/processes, how to fractionate the influent wastewater into the model components, which data are needed for model calibration, which parameters should be calibrated and how, etc. This way or philosophy of modeling can be best described as being predominantly ad hoc and heuristic. It becomes clear in the recently published activated sludge modeling guidelines: particularly in the WERF protocol (Melcer et al., 2003), the STOWA protocol (Hulsbeek et al., 2002) and the HSG guidelines (Langergraber et al., 2003) and to some extent also in the BIOMATH protocol (Petersen et al., 2002; Vanrolleghem et al., 2003). A thorough analysis of the existing modeling guidelines can be found in Sin et al. (2005).

In short, these guidelines aim to help consultants/engineers perform the minimum work that is required to get an adequate level of accuracy from the calibrated model Johnson (2006) in the modelEAU forum (http://www.modeleau.org/forum/viewtopic.php?t=15). This absolute

minimum calibration effort trend has resulted in a model calibration practice where the majority of the model parameters are fixed while only few parameters (and these are mostly the influent fractions) are adjusted on the basis of a relatively small set of experimental data.

Systems analysis offers an alternative approach to this ''heuristic'' modeling practice and promises to ground the modeling exercise on a generally agreed systematic framework (Beck, 1987). However, there are some drawbacks with this approach too. Previous applications of systems analysis to activated sludge models (Abusam et al., 2001; Brun et al., 2002; von Sperling, 1994; Wanner et al., 1992; Weijers and Vanrolleghem, 1997) revealed that these methods are computationally demanding and that the identification of ASMs is currently *practically not possible* (Gujer, 2006). In fact the only study (known to us) that claimed that *all* ASM1 parameters are identifiable was strongly criticized and disapproved since it was shown to be methodologically flawed (De Pauw et al., 2004). The reason for this is twofold: on the one hand the ASM model structure and the selected measured variables (to be used for estimation) do not allow unique identification of all parameters and on the other hand, the quantity and quality of data typically gathered at a treatment plant is far from being sufficient to properly identify all model parameters. Hence the use of classical search algorithms for non-linear estimation of all the model parameters has so far failed due to the complexity of such estimation problems. For example, local optimization algorithms (e.g., the simplex algorithm (Nelder and Mead, 1964)) would get stuck easily in local minima, while global optimization algorithms (e.g., genetic algorithm, simulated annealing, etc.) would require an enormous number of simulations before providing the parameter estimates (Banga et al., 2005). A hybrid approach employing classical local algorithms nested within Monte Carlo method demands even much larger numbers of simulations (von Sperling, 1994).

In the field of hydrological modeling several Monte Carlo based methods have also been successfully developed for model calibration, most notably the GLUE method (Beven and Binley, 1992) and the SCEM-UA algorithm (Vrugt et al., 2003). These methods are powerful for the quantification of parameter and prediction uncertainty (see, e.g., Gallagher and Doherty, 2007; McIntyre et al., 2002). However these methods also come with a heavy computational demand, which becomes infeasible for models whose simulation takes more than even a few seconds (Gallagher and Doherty, 2007). Important to recall is that one simulation of a typical WWTP model has a computational time associated with it in the order of tens of minutes, which makes the application of above-mentioned methods to WWTP models practically infeasible. In short, the existing statistical algorithms known to us for parameter estimation fail to provide a feasible solution for the calibration of the complex activated sludge models.

Having recognized this situation as the prime reason behind the manual calibration of WWTP models in practice (because all the above-mentioned approaches failed for practical reasons), we turn our attention to automate the steps typically involved during the manual calibration by a modeler in this study. To this end, we propose a pragmatic methodology using a Monte Carlo procedure but in a different scope which will become clear below. The emphasis in this pragmatic approach is to make practice move further, by automating the steps involved in the manual trial and error approach of model calibration. Eventually, it is aimed to make this time-consuming step of activated sludge modeling more efficient in view of the modeler's time. This study was performed as part of the MORE project-Model-Based Optimization of Wastewater Treatment Plants, initiated by the Waterboard De Dommel—a water authority responsible for the water management in the South-Eastern part of the province Noord-Brabant (The Netherlands). This project aims at improving the quality and the efficiency of the modeling process and the knowledge about application of models in practice.

The detailed description of the Monte Carlo based approach and its motivation is given below in A Pragmatic Monte Carlo Approach for Calibration of ASM'S Section. The calibration approach was evaluated on a formerly developed ASM2d model of the full-scale (50,000 PE) Haaren WWTP located in The Netherlands. For the re-calibration of the model using the new approach, we have used long-term on-line oxygen, ammonia and nitrate measurements from the (carrousel) aeration tank of the plant as described in Materials and Methods Section. In Results Section, we present the results of the calibration approach and we also perform a validation of the calibrated model. These results are then evaluated and discussed in Discussion Section. Finally, we concluded the manuscript with perspectives and conclusions.

A Pragmatic Monte Carlo Approach for Calibration of ASM'S

Motivation for a Pragmatic Approach

In practice, a step-wise manual calibration procedure typically involves the following operations: (1) the modeler selects a parameter subset [that depends on which guidelines he/she follows during the study, compare STOWA vs. WERF vs. BIOMATH] (2) the modeler runs the model with reference parameter values and visually compares the resulting fits to the available data (3) the modeler changes one parameter value at a time (the way a parameter is changed, e.g., increasing or decreasing depends on the modeler's experience and usually the value will be taken from a parameter range with a lower and upper bound) and runs the model again and compares the resulting fits visually, last (4) the modeler iterates step 3 until a sufficient goodness of fit is obtained—a termination criterion which is subjected to the judgment of the modeler doing the calibration. This procedure is subjective (because it is not a

parameter estimation in a statistical sense) and obviously quite tedious and time consuming since the modeler will have to wait each time for quite a while for the simulation job to be finished before judging the goodness of fit and iterate a few times until terminating the calibration task.

To frame the above-mentioned steps of this manual calibration on an algorithmic basis which can be automated, we have chosen to use the Monte Carlo procedure that also involves a step-wise procedure similar to the above-mentioned steps. In this pragmatic approach, we use the Monte Carlo procedure to produce a number of model simulations which results from exploration of the parameter space selected by the modeler. In this sense, the Monte Carlo procedure is used to efficiently (by an appropriate choice of sampling technique see below) cover the parameter space and thereby provide the corresponding model fits to the available data. These model fits are then presented to the modeler to let him/her evaluate both visually and statistically before making a choice. These steps are summarised in Figure 1 and detailed below.

The difference of this pragmatic approach from the bayesian parameter estimation methods such as SCEM-UA and GLUE is clearly that a parameter search (optimization) algorithm is not implemented nor bayesian inference is employed to update the prior information for the pure reason that their computational burden is too heavy to implement (see Introduction Section).

Further, we have chosen to use the Latin Hypercube Sampling (LHS) technique for sampling the parameter space

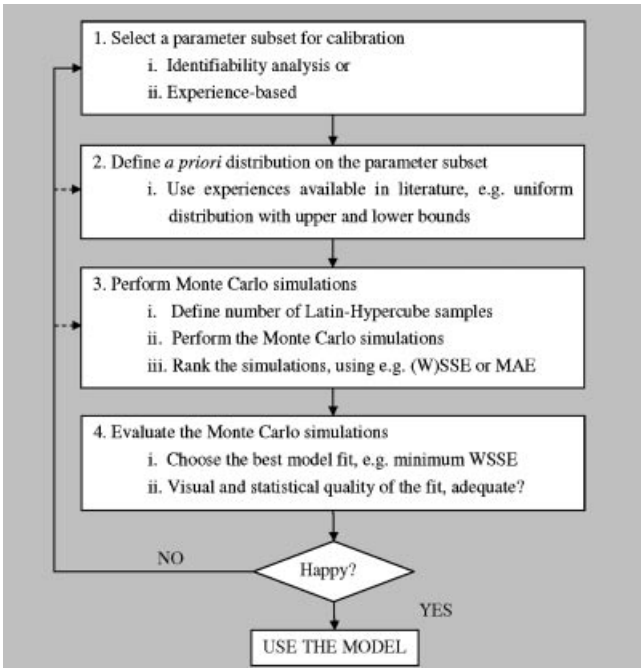

Figure 1. The pragmatic Monte Carlo approach for systematic calibration of ASMs.

(McKay et al., 1979) since it allows an efficient way of covering the parameter space using a relatively small sample size (Helton and Davis, 2003). The latter is important as it determines the computational demand needed to complete the required number of Monte Carlo simulations. Hence, the resulting computational demand of the Monte Carlo simulations can be controlled by the user by decreasing or increasing the number of samples. This is particularly important since performing one long-term WWTP simulation on a Pentium IV type PC may take 1–2 h (at least in the case studied here).

The Monte Carlo Approach for Model Calibration

For notational convenience, let the WWTP model structure be represented by **f** and **g**:

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

where **x**, **u**, **y** and **θ** denote the state vector, the input vector, the output (measured variables) vector, and the parameters vector, respectively. The Monte Carlo procedure typically involves the following steps: (1) specifying the range for each parameter (2) sampling the parameter space and (3) simulating the sampled parameter values through **f** and **g** to obtain the resulting model predictions, **y**.

The LHS of the parameter space is illustrated in Figure 2 for the parameters θ_1 and θ_2. Both parameters are assumed to have a uniform distribution with upper and lower bounds equal to $[0.5 \quad 3.5]$ and $[0.1 \quad 2]$, respectively. To draw a certain number of samples, say N, from the 2×2 parameter space formed by θ_1 and θ_2 the following steps are taken (Helton and Davis, 2003):

(1) First, the range of each parameter is divided into N intervals of equal probability $(1/N)$. For each parameter, one value is randomly selected in each interval. This step results in N randomly selected values for parameters θ_1 and θ_2.
(2) The LHS is then completed by randomly pairing the values of θ_1 and θ_2 leading to N couples of the parameters θ_1 and θ_2. The end result of the sampling is shown in Figure 2 where one can also see that the higher the sampling, the better the coverage of the parameter space will be.

Finally, the steps involved in the pragmatic calibration approach for activated sludge models are summarized below (which was also illustrated in Fig. 1):

(1) *Selection of parameter subset*: This step concerns selection of parameter subset for calibration. This is typically required as only a subset containing few parameters among the many parameters (up to 75) of the ASMs can be identified (Brun et al., 2002; Ruano

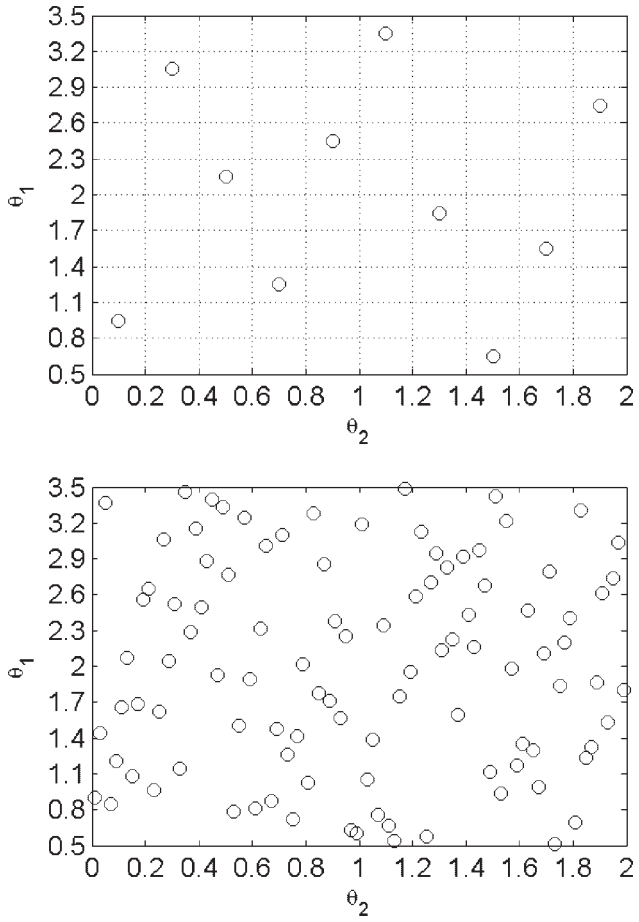

Figure 2. Illustrating the concept of Latin Hypercube Sampling: sampling number of 10 (top) and sampling number of 100 (bottom; see text for explanation).

et al., 2007; Weijers and Vanrolleghem, 1997). As mentioned above, it can be based on previous experiences and expert knowledge (e.g., only estimate the influent fractions) or it can (preferably) be based on a sensitivity analysis (identifiability) study or an appropriate combination of both methods (see for a detailed Discussion in Ruano et al. (2007)). It is important to note that all other parameters not selected for calibration are given a default and/or literature value, which for domestic WWTP has been found acceptable (Brun et al., 2002; Ruano et al., 2007).

(2) *Definition of parameter space*: This step involves defining an appropriate range for each of the parameters to be estimated. Similar to the first step, the large body of experiences and knowledge accumulated in literature about ASM parameters can be used to define the upper and lower bounds of realistic values for the parameters to be calibrated (Brun et al., 2002; Cox, 2004). Equally the modeler can make use of process knowledge to define a suitable range on the operational parameters of the plant (e.g., temperature, flow rates, etc.). In this study, we assumed a uniform distribution to be appropriate for each parameter.

(3) *Sampling and running Monte Carlo simulations*: In this step, a specified number of simulations is performed, each with a different set of parameter values sampled from the parameter space using a LHS technique. Each LHS sample of parameter values is then simulated resulting in the so-called set of Monte Carlo simulations. Each of these simulations is compared to the calibration data, and an objective function value can be calculated. Very often, this objective (or cost) function is chosen as the (weighted) sum of squared errors, (W)SSE, or mean absolute error (MAE) and the like.

(4) *Evaluation of the Monte Carlo simulations*: This step involves selecting the parameter set which resulted in the lowest objective value and, thus, for which the model describes the measured data best. A visual and statistical evaluation of the best model fit should be done to decide whether the end result is adequate. In case the evaluation is not positive, then the procedure should be iterated (see below).

In step 4 of the procedure, the good modeling practice requires one to perform a validation to check the prediction power of the model against a data set different than the one used for the calibration. If the quality of the model fit is found insufficient, then the underlying reasons should be investigated and the procedure should be iterated. The lack-of-fit could be due to (i) the use of an unidentifiable subset of parameters [go to step 1] (ii) a too narrow or too large uncertainty assumed on the a priori distributions of one or several parameters (go to step 2), (iii) an inadequate coverage of the parameter space, (increase the number of LHS samples and go back to step 3) or (iv) the model structure is not correct (find out the underlying reasons, modify the structure appropriately and go back to step 1).

As part of this procedure, the initialization of the model's state variable was also performed during each Monte Carlo simulation. This is needed since changing some parameters of the model may influence the initial conditions of the system at the start of the calibration period (e.g., biomass composition). This is illustrated in Figure 3, where one Monte Carlo simulation period consists of two subsequent parts: an initialization period (performed by considering a short history of the plant, i.e., three times SRT) followed by a calibration period in which the model predictions are compared with reality, that is the objective function (e.g., SSEs) is calculated.

Materials and Methods

The Simulation and Calibration Software

The calibration described in this work was performed using a dedicated software tool, MORE, which was specifically developed with automation of the calibration process in

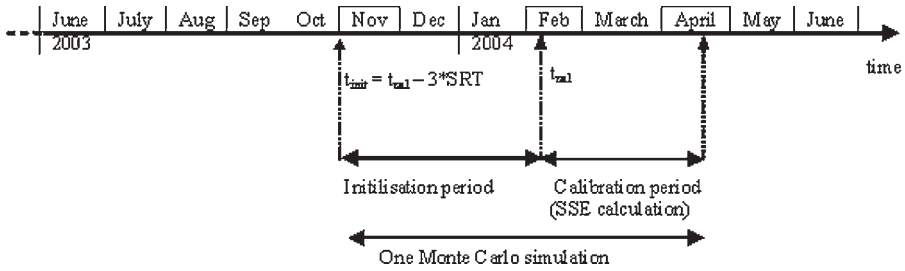

Figure 3. Scheme describing initialization as part of the dynamic calibration method. The time axis indicates the operational period of the plant used in this modeling study.

mind. It is intended to allow for easy data collection and data quality analysis, sensitivity analysis of the model parameters, (steady-state and dynamic) model calibration and scenario analysis. This software is linked through the API module with the WEST® modeling and simulation software (MOSTforWATER NV, Kortrijk, Belgium) in order to perform the model simulations.

Haaren WWTP

The treatment plant under study is located near Haaren, Noord-Brabant, The Netherlands. It is a carrousel type plant receiving the wastewater of 50,000 PE. The recorded average dry weather flow rate of the plant during the study was around 10,000 m^3/day. The plant consists of one anaerobic reactor, two carrousels and four clarifiers (see Fig. 4). The carrousel reactors are operated in parallel and basically the system has two lines after the anaerobic selector. The sludge wastage rate is manually controlled according to off-line daily SVI measurements, except for the weekends where no sludge is wasted from the plant. The volumes for the anaerobic compartment (1), carrousels (2) and the clarifiers (4), are given in total as 1,400, 10,000, and 2,460 m^3, respectively. The operational SRT of the system is around 22 days. The overall hydraulic retention time of the system is 1.4 days.

The Model Structure of the Haaren Plant

The Haaren model used here for evaluation of the new calibration framework is a result of extensive calibration and validation work previously done on the Haaren plant. This included an intensive measurement campaign for influent fractionation, hydraulic profiling, settling characterization, controller model development, solids balancing and calibration of the phosphate and the nitrogen dynamics (Insel et al., 2007; Sin, 2004; Vanrolleghem et al., 2003). The model's general features are outlined below.

The hydraulics was described following the tanks-in-series approach, where a number of four CSTRs was deemed sufficient to describe the selector. This selection was based on the geometry of the tank which had four-compartments and was supported by the phosphate measurements in each of these compartments of the selector. The carrousel was described as a loop of eight equal volume CSTRs of which the two CSTRs (numbered as 1 and 4) represented the location of the two surface aerators. This selection was based on the study of Abusam and Keesman (1999) which found

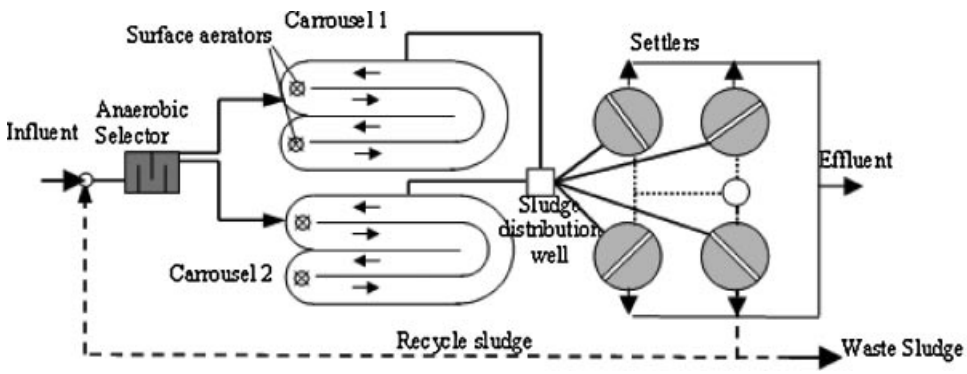

Figure 4. Schematic diagram of the Haaren activated sludge plant.

that a higher number of CSTRs (e.g., 16) does not significantly affect the simulated effluent quality results—a conclusion supported by tracer studies of mixing in carrousel type plants (De Clercq et al., 1999). Since the settling was very good, the clarifiers were described using an ideal point settler model. The biological phosphorus and nitrogen removal processes were described using the ASM2d model. The control strategy of the surface aerators, the details of which are explained elsewhere (Insel et al., 2007; Sin, 2004), was described with a dedicated algorithm implemented in the WEST Model Specification Language (MSL, Vanhooren et al., 2003).

Long-Term Data for the Dynamic Calibration

For the re-calibration of the Haaren model using the new framework, a long-term data set with high frequency measurements was used. For the model's initialization, 3 months daily averaged influent load data starting from November 17, 2003 up until February 15, 2004 were used. The high-frequency data set used, started from February 16, 2004 up until May 12, 2004 and consisted of on-line NH_4-N, DO and NO_3-N measurements once every 5 min. The total size of the calibration and validation data set is equal to 24,768 measurements for each variable (86 days $\times$ 1,440 min/1 day $\times$ 1 measurements/5min = 24,768 measurements). Roughly the first 2/3rd of the data set was used for dynamic calibration, while the remaining third of the data were used for the validation of the model.

The influent temperature measurements were also available from the routine plant monitoring and used as input to the model. The dynamic influent load data needed for the long-term dynamic simulation was obtained from the available and also dedicated measurements as follows. First, the daily average influent load was interpolated from the 24 h flow-proportionally sampled bi-weekly measurements of influent COD, TN and TP. Second, the diurnal influent load pattern was obtained by averaging the diurnal influent load profiles obtained from 4-days of intensive measurements (4-h flow-proportional) performed in June 2003 (see Fig. 5). Together with BOD (28 days) measurements, these intensive influent and effluent measurements were used to calculate the average fractions of the influent COD, which were found as 0.05, 0.12, 0.22, 0.28, and 0.33, respectively for the inert soluble COD, volatile fatty acids, degradable soluble COD, inert particulate COD and degradable particulate COD (Insel et al., 2007).

Data Quality Check and the Solids Balance

The quality of the data used during calibration and validation was checked following simple statistical analysis, for example an outlier check and error analysis for the flow balance and correlation analysis between the on-line sensor measurements and control measurements for ammonium and nitrate, respectively (see Sin et al., 2006).

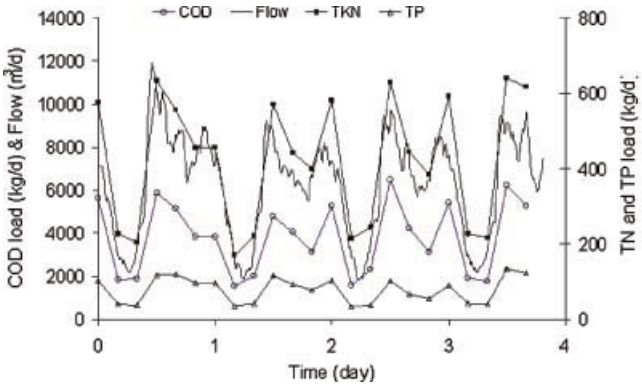

Figure 5. Diurnal influent loading as measured intensively for 4-days during mid-June 2003.

The steady-state solids balance, particularly the sludge wastage rate which determines the SRT of the plant was confirmed to be 22 days. For the steady-state balance, the average influent load observed during 7 months, corresponding to the initialization, calibration and validation periods, was used.

The Modeling Guideline

The BIOMATH protocol (Vanrolleghem et al., 2003) was followed during the different steps of the WWTP modeling, for example the goal definition, data collection and quality check, mathematical formulation of plant units/processes, influent characterization and steady-state modeling (as mentioned above).

Statistical Criteria for Evaluating Model Fits

The weighted SSEs objective function, WSSE is calculated as follows:

$$\mathrm{WSSE}(\boldsymbol{\theta}) = \sum_{k=1}^{m}\sum_{i=1}^{n}\left(\frac{y_{\mathrm{meas},k,i} - y_k(t_i,\boldsymbol{\theta})}{sc_{k,i}}\right)^2 \qquad (2)$$

where $\boldsymbol{\theta}$ is the parameter subset, $y_{\mathrm{meas},k,i}$ is the ith measurement of the kth variable y, and $y_k(t_i, \theta)$ is the corresponding model output at the ith time instance. m is the total number of variables, y, while n is the total number of observations of variable y_k used for model calibration. $sc_{k,i}$ is the scale of the model output y_k at the ith time instance. This term is needed to make the sum term non-dimensional, and is chosen here as the standard errors of the measurements, σ_{yk} and is assumed constant at all time instants. The standard deviation of the on-line sensors of ammonium, σ_{NH_4}, nitrate, σ_{NO_3}, and oxygen, σ_{O_2}, were obtained from the specifications of the manufacturer as follows 0.17, 0.68, and 0.05 mg/L, respectively.

To assess the quality of the model fits, the statistical tests suggested by Power (1993) were used that includes the MAE, root mean squared error (RMSE) and Janus coefficient. These tests are calculated, respectively by Equations (3)–(5). All parameters are as defined above, except for n_cal and n_val, which stand for the total number of measurements used in the calibration and validation periods, respectively

$$\mathrm{MAE} = \frac{1}{n}\sum_{i}^{n} \left| y_{\mathrm{meas},i} - y(t_i,\theta) \right| \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i}^{n} \left(y_{\mathrm{meas},i} - y(t_i,\theta) \right)^2} \tag{4}$$

$$J^2 = \frac{\frac{1}{n\_val}\sum_{i=1}^{n\_val} \left(y_{\mathrm{meas},k,i} - y_k(t_i,\theta) \right)^2}{\frac{1}{n\_cal}\sum_{i=1}^{n\_cal} \left(y_{\mathrm{meas},k,i} - y_k(t_i,\theta) \right)^2} \tag{5}$$

Results

The Modeling Objective

As mentioned above, the previously developed Haaren model was used in this study to evaluate the Monte Carlo approach for calibration. However, to keep the study focused on the calibration approach itself, the modeling objective in this particular study was set to adequately describe *only* the nitrogen (nitrate and ammonium) dynamics in the carrousels of the Haaren plant (the phosphorus removal was described using the previously calibrated PAO parameters of the Haaren model performed in Insel et al. (2007)). In this particular study the model is also required to remain valid for a 1-month time-frame after its calibration, since it is foreseen to be used in day-to-day management of the plant by the operators.

Evaluation of the Monte Carlo Based Calibration Procedure

Step 1: The parameter subset for dynamic calibration was selected following the experience-based approach. The selection was based on process engineering knowledge in biological nutrient removal and the modeling experiences of the authors with the ASM2d model (e.g., WERF; Melcer et al., 2003), STOWA protocol (Hulsbeek et al., 2002), BIOMATH protocol (Insel et al., 2006; Vanrolleghem et al., 2003). The list of selected parameters is given in Table I.

Step 2: For each parameter a uniform distribution was chosen since no a priori information is available about the statistical distribution of these parameters in activated sludge models. The upper and lower bounds of the uniform distributions of the biokinetic parameters were defined on the basis of values reported in relevant literature, mainly Brun et al. (2002), Cox (2004) and Henze et al. (2000). The distribution of the parameters related to the operation of the aerators was defined based on manufacturer specifications and dedicated experiments performed with the aerators. The uniform distributions assigned to each of the parameters are given in Table I.

Step 3: For the LHS sampling, 500 samples were taken from the 9×9 dimensional (hyper) space formed by the parameters. As each LHS sample contains randomly selected values for each of the nine parameters, this means 500 simulations were performed using 500 randomly selected values of the nine parameters. The resulting values of the objective obtained from the 500 Monte Carlo simulations provide the squared distance between the model predictions and the measurements. In a way these results also provide a global picture of the objective function indicating how the difference between the model and the data changes as one wanders in parameter space.

Step 4: In Figure 6, the objective function values obtained from the Monte Carlo simulations were sorted from maximum to minimum to help interpret the results. With this ranking done, one can select the parameter sample that provides the minimum objective function value. This LHS sample is then accepted as the calibrated parameter set for the model; hence not an optimal but a pragmatically calibrated model is obtained. The parameter sample is given in Table II and the resulting model fit to the measurements is presented in Figure 7.

In general, Figure 7 suggests that the model is able to follow the measured daily dynamic trends in ammonia and nitrate (see the zoomed day pattern in the same figure). These measured trends are basically the footprints of the aeration controller which periodically controls the oxygen in the carrousel: during full aeration the nitrate rises and the ammonia decreases while in the intermittent aeration period the nitrate is decreased while the ammonia accumulates. Generally, the model match of the measured oxygen was found in good agreement with the measurements.

Table I. The mean, upper and lower ranges of uniform distribution of the parameter subset.

	μ_{AUT} (day^{-1})	$K_{O,AUT}$ (mg/L)	K_O (mg/L)	η_g	k_h (day^{-1})	η_{fe}	K_La_{max} (day^{-1})	K_La_{min1} (day^{-1})	K_La_{min2} (day^{-1})
Upper	2	1	1	1	4	1	350	70	30
Mean	1	0.4	0.3	0.6	3	0.2	300	50	10
Lower	0.8	0.1	0.1	0.2	2	0	200	30	5

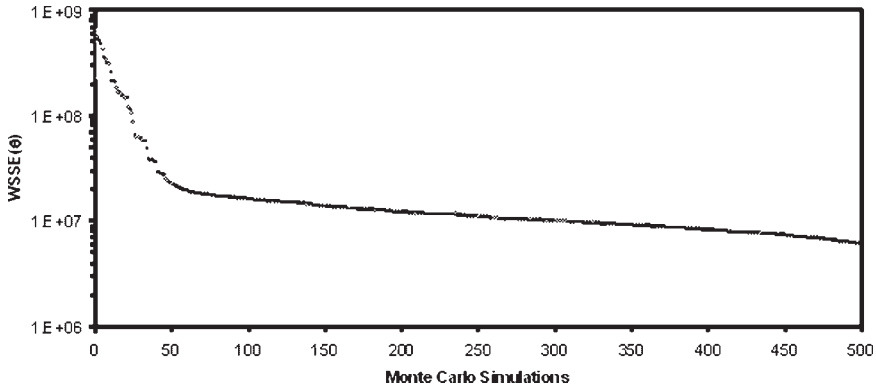

Figure 6. Monte Carlo simulations ranked according to WSSE criteria: the *y*-axis is given on a logarithmic scale to better compare large and small numbers.

One observes a significant mismatch between the model and the ammonium measurements between the 19th and the 23rd of February (see Fig. 7). We believe this mismatch is largely caused by a sudden increase in the influent ammonium load to the plant, which was not measured at the plant nor could it be estimated from the influent load interpolation step (see above).

A detailed statistical analysis of the quality of the model fits is given below in the Discussion Section.

Validation of the Calibrated Model

Evaluation of the calibrated model, which uses the parameter subset corresponding to the least objective function value (see Table II) on a different data set is performed in this section within a model validation framework. For the validation of the model, the next 1-month of data following the calibration period is used (Fig. 8). Visual comparison of the calibrated model predictions with the validation data set indicate that the model could still reproduce the trends in nitrate (especially starting from the April 23 till the May 12). While the general trends were followed by the model predictions (see the zoomed day pattern in Fig. 8), one can observe a systematic difference between the nitrate data and the model, particularly, between the 11th and the 20th of April, which suggests the model has a bias (see below for a detailed quantification). For the oxygen dynamics the general trend, that is full and intermittent aeration periods, also appears to be captured by the calibrated model. A statistical assessment of the model fits is given below.

Discussion

Evaluation of the Calibration Procedure

The WSSE profile obtained from the Monte Carlo procedure is flat which indicates that there exist many parameter values providing a similar fit to the data. The latter points to the identifiability issues typically encountered when employing parameter estimation (see, e.g., Brun et al., 2002; Ruano et al., 2007; Weijers and Vanrolleghem, 1997) whose underlying reasons are explained above (see Introduction Section). Second, the procedure is reproducible when it is run with the same LHS number and provides similar (flat) WSSE profiles (results not shown). In each run, one could obtain equally good model fits to the one shown here but with different parameter values. Hence, it is important to remind that this method is not meant for parameter identification purposes, but for model calibration purposes and as an algorithmic procedure to automate the steps of the manual calibration procedure.

The goodness of model fits was evaluated using several criteria commonly used in modeling including MAE, RMSE and Janus coefficient (Elliott et al., 2000; Power, 1993). The first two criteria indicate the prediction accuracy and should be as small as possible, while the last criterion relates to model validity and should be close to 1. Looking at the MAE criterion in the calibration and validation periods (Table III), we find that there is a systematic difference between the measurements and the predictions. Further, the prediction accuracy of the model for ammonium decreased from 1.4 mg N/L in the calibration period to 1.0 mg N/L in the validation period. The RMSE of the ammonium prediction is also found to decrease in the validation period

Table II. The parameter sample providing the minimum objective function value obtained from the Monte Carlo-based calibration.

μ_{AUT} (day^{-1})	$K_{O,AUT}$ (mg/L)	K_O (mg/L)	η_g	k_h (d^{-1})	η_{fe}	K_La_{max} (day^{-1})	K_La_{min1} (day^{-1})	K_La_{min2} (day^{-1})
1.15	0.52	0.20	0.65	2.97	0.20	326.0	37.8	7.82

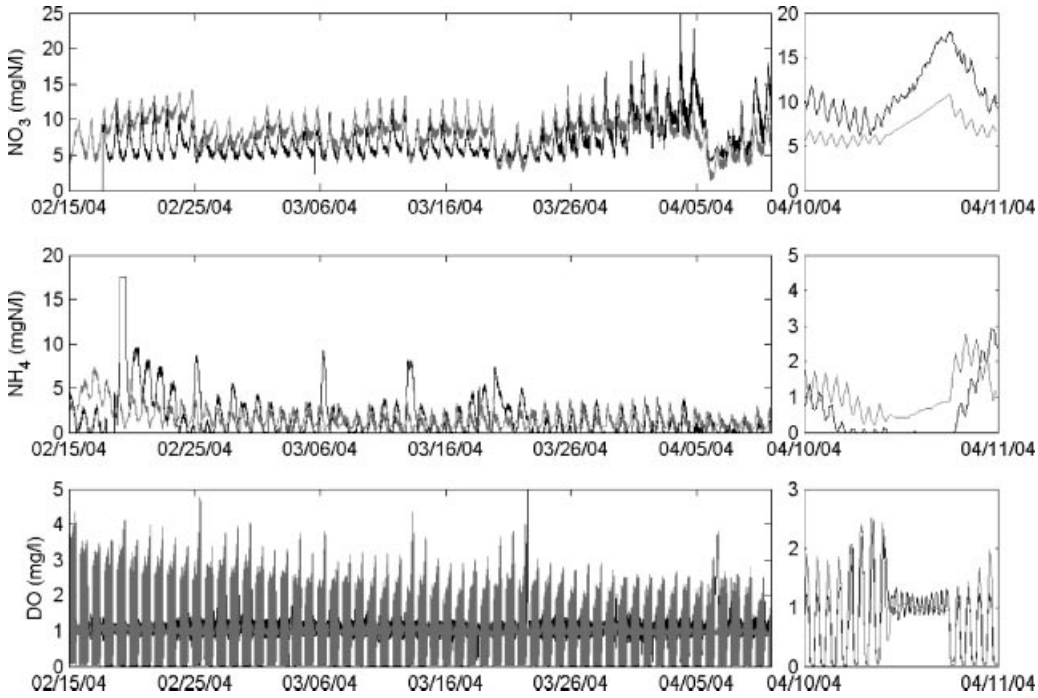

Figure 7. The best model fit obtained from 500 LHS samples out of the parameter space: measurements (black line) simulation (gray line). The last day was zoomed in to provide a better visualization of the data and the simulation (right).

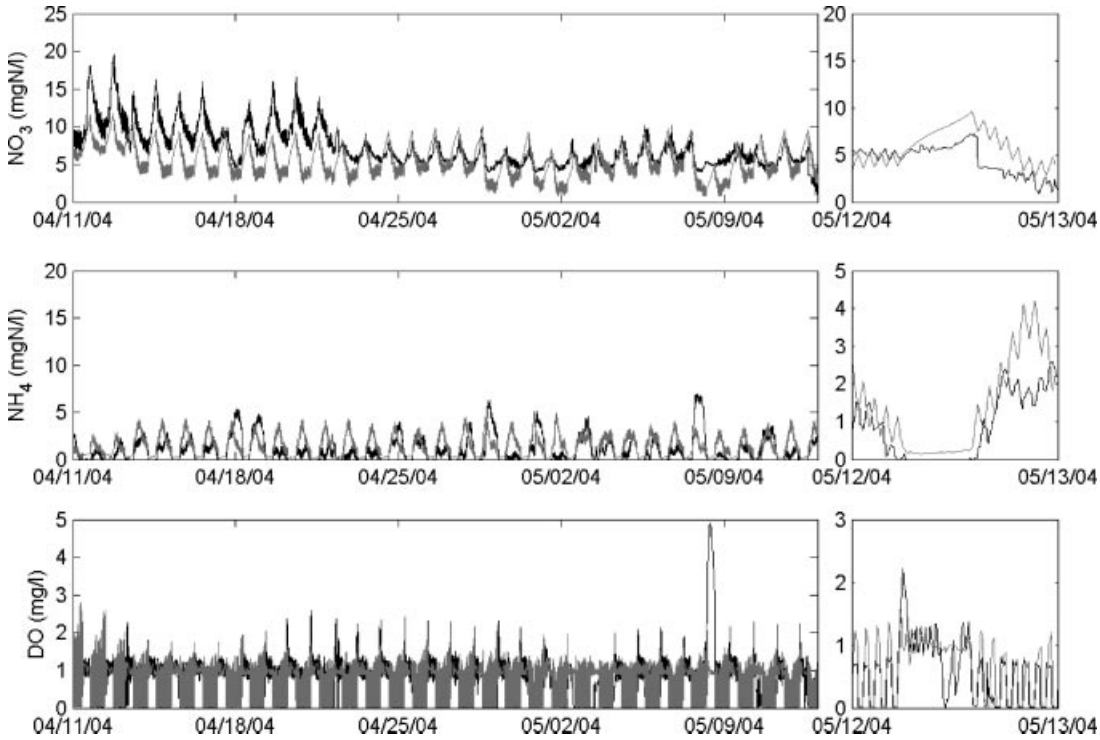

Figure 8. Model validation results using 1 month of independent data: measurements (black line) simulation (gray line). The last day was zoomed in to provide a better visualization of the data and the simulation (right).

Table III. Comparison of the model fits during calibration and validation.

	Ammonium		Nitrate		Oxygen	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
Total no. of data	16,117	9,217	16,117	9,217	16,117	9,217
MAE	1.39	0.98	2.56	2.31	0.48	0.29
RMSE	2.55	1.40	3.03	2.87	0.80	0.51
Janus coefficient		0.6		0.9		0.7

(see Table III) supporting the MAE criterion. These results are meaningful since it means that as the prediction accuracy increases, the mean absolute difference between the measurements and the simulation also decrease. The Janus coefficient for the ammonium is found to be 0.6, which means that the mean prediction error decreased in the validation period compared to the calibration period. It also suggests that the model structure remained more or less the same as J is still in the neighborhood of 1. In short, based on these statistical criteria, one may conclude that the structure and the statistical accuracy of the model remained valid for ammonium in both calibration and validation periods.

The underlying reason why the model performed unexpectedly better in the validation period could be partly because the calibration data set was more difficult to describe (compare data set in Figs. 7 and 8). This challenge was due to the rainfall event that occurred during the first week of the calibration period (see the ammonium data in Fig. 7) but which were not measured with sufficient frequency (only routine bi-weekly influent measurements were available). Hence, the model could not reproduce this week resulting in a higher discrepancy as reflected by the statistical tests.

For the nitrate predictions, the MAE and RMSE were found around 2.5 and 3 mg N/L, respectively. Both criteria were observed to slightly decrease in the validation period (see Table III). These results suggest that also for this output both the mean absolute difference and the prediction accuracy remained more or less similar in both calibration and validation periods. This is confirmed by the Janus coefficient which is found quite close to 1, that is 0.9 (see Table III). Overall the validation tests are successfully passed for this modeling study. This means that the model structure remained valid for nitrate during the validation periods, albeit with a systematic (absolute) error of around 2.5 mg N/L.

The quality of the model fits to oxygen measurements in both the calibration and validation periods are equally good. The MAE of the oxygen predictions is between 0.5 and 0.3 mg O_2/L and is accompanied with a good prediction accuracy (RMSE), between 0.8 and 0.5 (much smaller compared to the ammonium and nitrate predictions). This is rather acceptable, particularly when one considers the fact that the oxygen is a controlled variable which creates an additional difficulty in terms of correct approximation since the adequate description of the highly dynamic inputs of the controller in reality is a challenging task. The Janus coefficient was found to be around 0.7, which confirms again that the model to a large extent remained valid for the oxygen predictions.

One ideally wishes the bias to be zero for the model predictions. In reality, however, model building has to be done against a background of many assumptions and uncertainties (in influent measurements and estimation of its fractions (e.g., only bi-weekly influent measurements were available), mixing, default parameter values, model structure, etc). Practical experiences with the use of activated sludge models have reported a rather wide range of model deviations deemed acceptable, with relative errors ranging from 10% up to 40% (Melcer et al., 2003). Therefore, we believe that the model bias found in this study for the ammonium, the nitrate and the oxygen concentrations, around 1, 2.5 mg N/L and 0.4 mg O_2/L, respectively, is acceptable for several model applications, especially when the emphasis is mostly put on comparing different alternatives/scenarios rather than predicting the absolute values.

Efficiency of the Calibration Procedure

The Monte Carlo based calibration approach developed here makes use of relatively large numbers of model simulations to obtain a calibrated model. Since these simulations can be performed without user intervention by a computer, a considerable time of the modeler/engineer is thus saved, which can be used to analyze and interpret the predictions, which is the ultimate aim of using a model.

At the downside of our method lies essentially the requirement of computational power to perform the lengthy simulations in a short-time. In this particular case, one simulation run (the simulated time period is ca. 5 months long) required 45 min of PC-time (using a Pentium IV 3 GHz PC). Therefore, 500 Monte Carlo runs would require 2 weeks of computation time. Fortunately, this simulation time can be considerably decreased (e.g., by an order of magnitude) by using distributed computing, as shown in Claeys et al. (2006). Indeed, Monte Carlo simulations allow the distribution of the simulations on the computers available in a computer cluster, grid or even using the computers of an organization in their down-time. The investment into or hiring a distributed computation service may increase the cost of using our method but given the low cost of computers this cost will be considerably lower than

the cost of the time a modeler otherwise has to invest in performing the calibration manually.

Future Perspectives: Where Do We Go From Here?

We believe the pragmatic Monte Carlo approach to the calibration of ASM is useful (1) to successfully automate the manual trial and error model calibration now in use in activated sludge modeling practice but also (2) as it introduces the WWTP modeling community to the framework of Monte Carlo based model calibration. The latter can be extended further with parameter search (optimization) and bayesian inference algorithms for parameter and prediction uncertainty estimation as the computational demand eases on these methods in future. The development of an efficient global parameter estimation method remains desirable as it will allow obtaining unbiased estimates rather than the approximate estimates the traditional trial and error methods provide. Within the proposed method, the selection of an appropriate parameter subset to be calibrated, for example by using global sensitivity analysis such as Morris screening (Morris, 1991) or standardized regression coefficients (SRCs), a priori knowledge on the distribution of the model parameters including the correlation matrix and the required number of Monte Carlo samples are among the important issues that need further research.

Within the framework of full-scale model application, on the other hand, an important issue that certainly needs ample consideration is the fact that only a small number of parameters are used for calibration, while all others remain at their default or literature value. Fixing parameters at unrealistic or inappropriate values may have serious consequences on the extrapolation power of the model outside the calibration conditions as the error/bias of the predictions will increase considerably (see, e.g., Omlin and Reichert, 1999; Vanrolleghem, 2007). Therefore, it would be useful to perform an uncertainty analysis to assess the magnitude of the prediction errors when using the model in practice with possibly wrong settings of the fixed parameters. In fact, since the calibration procedure employs a Monte Carlo framework, it can easily be extended for uncertainty analysis (Saltelli et al., 2005). However, such extension of the procedure requires further research.

Conclusions

In this contribution, we have proposed and evaluated a pragmatic Monte Carlo based approach to replace manual trial and error calibration of activated sludge models. The approach essentially structures the manual trial and error approach into a systematic framework using four steps: (1) parameter subset selection, (2) setting ranges for each parameter in the subset, (3) LHS out of the defined parameter space and running the Monte Carlo simulations, and (4) evaluation of the Monte Carlo simulations and choosing the best parameter sample.

The approach was successfully used to calibrate an ASM2d model for a domestic wastewater treatment plant using long-term oxygen, ammonium and nitrate data collected at minute-scale frequency. Confronting the calibrated model with one month of independent validation data showed that the model was able to provide statistically accurate and valid predictions for oxygen, ammonia and nitrate with an acceptable bias, as confirmed by their Janus coefficients close to 1.

Overall, the Monte Carlo-based approach allows replacing the manual trial and error calibration approach in use today with an automated approach for the calibration of ASM parameters. In this way, it relieves the modeler from a large part of the time consuming and tedious task of calibration and hands it over to a (distributed) computer (cluster or network).

Nomenclature

API	application programming interface
ASM	activated sludge model
ASM1	activated sludge model number 1
ASM2d	activated sludge model number 2d
COD	chemical oxygen demand (mg COD/L)
CSTR	completely stirred tank reactors
DO	dissolved oxygen (mg O_2/L)
J	Janus coefficient
k_h	hydrolysis rate constant of slowly biodegradable COD (day^{-1})
K_O	oxygen affinity constant of heterotrophic biomass (mg COD/L)
$K_{O,AUT}$	oxygen affinity constant of autotrophic biomass (mg COD/L)
K_La_{max}	maximum oxygen mass transfer coefficient of the aerators (day^{-1})
K_La_{min1}	minimum oxygen mass transfer coefficient of the aerators during the anoxic phase of the intermittent aeration before the introduction of propellers (day^{-1})
K_La_{min2}	minimum oxygen mass transfer coefficient of the aerators during the anoxic phase of the intermittent aeration after the introduction of propellers (day^{-1})
LHS	Latin hypercube sampling
MAE	mean absolute error
NH_4-N	ammonium nitrogen (mg N/L)
NO_3-N	nitrate nitrogen (mg N/L)
RMSE	root mean squared error
SRT	solids retention time (day)
SSE	sum of squared errors
TN	total nitrogen (mg N/L)
TP	total phosphate (mg P/L)
WSSE	weighted sum of squared errors
WWTP	wastewater treatment plant
y	model output

Greek Letters

f_{XI}	fraction of inert particulate COD
f_{ns}	fraction of non-settleable solids
μ_{AUT}	maximum growth rate of biomass (day^{-1})

η_g	reduction factor heterotrophic growth under anoxic conditions
η_{fe}	reduction factor for hydrolysis under anaerobic conditions
θ	model parameter
σ	standard error of measurements

The authors would like to thank Waterboard De Dommel (Boxtel, The Netherlands) for the financial support through the MORE (Model-based Optimization of WWTPs) project (Sin et al., 2007). We kindly thank as well the operation team of the Haaren WWTP namely John Simons, Ad Berkel and Mark Janssens for their valuable cooperation during the project. Peter Vanrolleghem holds the Canada Research Chair in Water Quality Modeling and Gürkan Sin is a H.C. Ørsted postdoctoral fellow at the Technical University of Denmark.

References

Abusam A, Keesman KJ. 1999. Effect of number of CSTR's on the modelling of oxidation ditches: Steady state and dynamic analysis. Med Fac Landbouw Univ Gent 64/5a:91–94.

Abusam A, Keesman KJ, van Straten G, Spanjers H, Meinema K. 2001. Parameter estimation procedure for complex non-linear systems: Calibration of ASM No.1 for N-removal in a full-scale oxidation ditch. Wat Sci Technol 43(7):357–365.

Banga JR, Balsa-Canto E, Moles CG, Alonso AA. 2005. Dynamic optimization of bioprocesses: Efficient and robust numerical strategies. J Biotechnol 117:407–419.

Beck MB. 1987. Water quality modelling: A review of the analysis of uncertainty. Water Resour Res 23:1393–1442.

Beven K, Binley A. 1992. The future of distributed models—Model calibration and uncertainty prediction. Hydrol Process 6:279– 298.

Brun R, Kuhni M, Siegrist H, Gujer W, Reichert P. 2002. Practical identifiability of ASM2d parameters—Systematic selection and tuning of parameter subsets. Wat Res 36:4113–4127.

Claeys F, Chtepen M, Benedetti L, Dhoedt B, Vanrolleghem PA. 2006. Distributed virtual experiments in water quality management. Wat Sci Technol 53(1):297–305.

Cox CD. 2004. Statistical distributions of uncertainty and variability in activated sludge model parameters. Water Environ Res 76:2672–2685.

De Clercq B, Coen F, Vanderhaegen B, Vanrolleghem PA. 1999. Calibrating simple models for mixing and flow propagation in waste water treatment plants. Wat Sci Technol. 39(4):61–69.

De Pauw DJW, Sin G, Insel G, Van Hulle S, Vandenberghe V, Vanrolleghem PA. 2004. Discussion of 'assessing parameter identifiability of activated sludge Model No. 1' by Pedro Afonso and Maria da Conceicao Cunha. J Environ Eng 130:110–112.

Elliott JA, Irish AE, Reynolds CS, Tett P. 2000. Modelling freshwater phytoplankton communities: An exercise in validation. Ecol Modell 128:19–26.

Gallagher M, Doherty J. 2007. Parameter estimation and uncertainty analysis for a watershed model. Environ Modell Softw 22:1000–1020.

Gujer W. 2006. Activated sludge modelling: Past, present and future. Wat Sci Technol 53(3):111–119.

Helton JC, Davis FJ. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Safe 81:23–69.

Henze M, Grady CPL, Jr., Gujer W, Marais GvR, Matsuo T. 1987. Activated sludge model No.1. IAWPRC Scientific and Technical Report No. 1, IAWPRC, London, UK.

Henze M, Gujer W, Mino T, van Loosdrecht MCM. 2000. Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. IWA Scientific and Technical Report No. 9. IWA Publishing. London, UK.

Hulsbeek JJW, Kruit J, Roeleveld PJ, van Loosdrecht MCM. 2002. A practical protocol for dynamic modelling of activated sludge systems. Wat Sci Technol 45(6):127–136.

Insel G, Sin G, Lee DS, Nopens I, Vanrolleghem PA. 2006. A calibration methodology and model-based system analysis for nutrient removing SBRs under reduced aeration. J Chem Tech Biotech 81(4):679–687.

Insel G, Sin G, Lee DS, De Pauw DJW, Weijers S, Vanrolleghem PA. 2007. Evaluation of prediction performance of ASM2d three years after its calibration for an intermittently aerated carousel type EBPR plant. In: Proceedings of the 10th IWA Specialised Conference on Design, Operation and Economics of Large Wastewater Treatment Plants. Vienna, Austria, September 9–13 2007 (on CD-ROM).

Johnson BR. 2006. Engineering practise in North America. In discussions: Academia versus practice or how to merge different worlds? modelEAU forum (http://www.modeleau.org/Forum/viewtopic.php?t=15).

Langergraber G, Rieger L, Winkler S, Alex J, Wiese J, Owerdieck C, Ahnert M, Simon J, Maurer M. 2003. A guideline for simulation studies of wastewater treatment plants. Wat Sci Technol 50(7):131–138.

McIntyre N, Wheater H, Lees M. 2002. Estimation and propagation of parametric uncertainty in environmental models. J Hydroinform. 4:177–198.

McKay MD, Beckman RJ, Conover WJ. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2):239–245.

Melcer H, Dold PL, Jones RM, Bye CM, Takacs I, Stensel HD, Wilson AW, Sun P, Bury S. 2003. Methods for wastewater characterisation in activated sludge modelling. Alexandria, VA, USA: Water Environment Research Foundation (WERF).

Morris MD. 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33:161–174.

Nelder JA, Mead R. 1964. A simplex method for function minimization. Computer J 7:308–313.

Omlin M, Reichert P. 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecol Modell 115:45–59.

Petersen B, Gernaey K, Henze M, Vanrolleghem PA. 2002. Evaluation of an ASM1 model calibration procedure on a municipal–industrial wastewater treatment plant. J Hydroinform 4:15–38.

Power M. 1993. The predictive validation of ecological and environmental models. Ecol Modell 68:33–50.

Ruano MR, Ribes J, De Pauw DJW, Sin G. 2007. Parameter subset selection for the dynamic calibration of Activated Sludge Models (ASMs): Experience versus systems analysis. Wat Sci Technol 56(8):107–115.

Saltelli A, Ratto M, Tarantola S, Campolongo F. 2005. Sensitivity analysis for chemical models. Chem Rev 105:2811–2827.

Sin G. 2004. Systematic calibration of activated sludge models. PhD. Thesis. Faculty of Agricultural and Applied Biological Sciences. Ghent University, Gent, Belgium. pp 372.

Sin G, Van Hulle SWH, De Pauw DJW, van Griensven A, Vanrolleghem PA. 2005. A critical comparison of systematic calibration protocols for activated sludge models: A SWOT analysis. Wat Res 39:2459–2474.

Sin G, De Pauw DJW, Weijers SR, Vanrolleghem PA. 2006. Global and long-term calibration methodology of ASMs for full-scale WWTPs. In: Proceedings of the World Wide West (WWWest) users meeting. Kortrijk, Belgium, May 8–9, 2006.

Sin G, De Pauw DJW, Weijers S, Vanrolleghem PA. 2007. Developing a framework for continuous use of models in daily management and operation of WWTPs: A life cycle approach. In: Proceedings the 10th IWA Specialised Conference on Design, Operation and Economics of Large Wastewater Treatment Plants. September 9–13, 2007 Vienna, Austria. (on CD-ROM).

Vanhooren H, Meirlaen J, Amerlinck Y, Claeys F, Vangheluwe H, Vanrolleghem PA. 2003. WEST: Modelling biological wastewater treatment. J Hydroinform 5:27–50.

Vanrolleghem PA. 2007. On the problem of identifying overparameterised wastewater treatment models with prior information regarding a subset of the parameters. In: Proceedings of the International Workshop on Advances in Hydroinformatics 2007 (HIW07). Niagara Falls, Canada, June 4–7, 2007 (on CD-ROM).

Vanrolleghem PA, Insel G, Petersen B, Sin G, De Pauw D, Nopens I, Weijers S, Gernaey K. 2003. A comprehensive model calibration

procedure for activated sludge models. In Proceedings of the 76th Annual WEF Technical Exhibition and Conference (WEFTEC2003). Los Angeles, CA, USA, October 11–15, 2003 (on CD-ROM).

von Sperling M. 1994. Calibration of poorly identifiable systems: application to activated sludge model. J Environ Eng-ASCE 120(3):625–644.

Vrugt JA, Gupta HV, Bouten W, Sorooshian S. 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 39: 1201–1215.

Wanner O, Kappeler J, Gujer W. 1992. Calibration of an activated sludge model based on human expertise and on a mathematical optimization technique—A comparison. Wat Sci Technol 25(6):141–148.

Weijers SR, Vanrolleghem PA. 1997. A procedure for selecting best identifiable parameters in calibrating activated sludge model no. 1 to full scale plant data. Wat Sci Technol 36(5):69–79.