Stochastic Greybox Modeling for Control of an Alternating Activated Sludge Process

Halvgaard, Rasmus Fogtmann; Vezzaro, Luca; Grum, M.; Munk-Nielsen, Thomas; Tychsen, Peter; Madsen, Henrik

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Stochastic Greybox Modeling for Control of an Alternating Activated Sludge Process

R. Halvgaard\textsuperscript{a,b,*}, L. Vezzaro\textsuperscript{a,b}, M. Grum\textsuperscript{b}, T. Munk-Nielsen\textsuperscript{b}, P. Tychsen\textsuperscript{b}, H. Madsen\textsuperscript{a}

\textsuperscript{a}Technical University of Denmark, Matematiktorvet, DK-2800 Kgs. Lyngby, Denmark
\textsuperscript{b}Krüger A/S, Gladugevej 363, DK-2860 Sborg, Denmark

Abstract

We present a stochastic greybox model of a BioDenitro WWTP that can be used for short time horizon Model Predictive Control. The model is based on a simplified ASM1 model and takes model uncertainty into account. It estimates unmeasured state variables in the system, e.g. the inlet concentration or the sensor measurements in case of temporary sensor faults. This improves control performance without adding additional or redundant sensors. We fitted the parameters of the model to actual plant data and demonstrate the state estimation capabilities with this data set. The model now runs online at a WWTP in Denmark.

Keywords: WWTP, ASM1, Stochastic, Greybox, Alternating, BioDenitro

1. Introduction

Automation and advanced control methods are some of the key tools to improve the performance of existing Waste Water Treatment Plants (WWTPs) \cite{1}. New online control methods improve nutrient recovery, carbon neutrality, or energy neutral plants \cite{2,3,4}.

The control strategies can be evaluated by using the available simulation models based on the family of Activated Sludge Models (ASM) \cite{5,6}, i.e. the Benchmark Simulation Model (BSM) \cite{7}. The ASM models are great tools for simulation and provide a detailed description of the biological processes. However, this level of detail is not well suited for Model Predictive Control (MPC). As mentioned in \cite{1}, current MPCs in WWTP utilize “significantly simple” models. When focusing on alternating plants, that are widely applied in Danish WWTPs, simple linear models of WWTPs were developed by \cite{8,9} including several heuristic model based control strategies \cite{10,11,12,13,14,15}. \cite{16} also developed several models and control strategies for a pilot scale implementation of an alternating process, where \cite{17} specifically focused on the BioDenitro process. \cite{18} investigated a nonlinear observer for an alternating process. \cite{19,20,21,15} provide great summaries of other literature that deals with control and modeling of alternating plants.

Important obstacles on the road to widespread usage of MPC are data quality and measurements encumbered with uncertainty \cite{1}. Although several approaches for data validation are available in literature \cite{22,23,24,25}, their full-scale utilization in combination with MPC is still limited \cite{26}. Focus for allowing the practical implementation of complex MPC approaches in WWTPs are therefore

1. handling of uncertainty
2. fault-tolerant control of sensor failures
3. parameter estimation
4. forecasting of inlet flows and concentrations
5. numerical computational times

These points can be addressed by stochastic greybox models that combine simple and fast model structures with data assimilation routines, e.g. in the form of an extended Kalman filter. \cite{27,28} showed the first applications of greybox models in the 1990s, who applied parameter estimation of reduced ASM1 models on alternating plants with the aim of process prediction and control. \cite{27} concludes that greybox models of the wastewater processes performs significantly better than traditional blackbox models like ARMAX models.

The aim of this paper is to present a stochastic greybox model for Model Predictive Control of a WWTP that also accounts for uncertainty. The model is based on a simplification of the ASM1 model for a control horizon of 24 hours since this short time horizon is the most important for control. The model continuously adapts to the slower time-varying effects of the process described in the ASM1. The proposed model is formulated as a continous-discrete time state space model and is capable of estimating unmeasured variables in the system. This improves control performance without adding additional sensors.

This paper is organized as follows. Section \cite{21} reduces the original ASM1 model to a practical greybox model. Section \cite{3} describes the framework for estimating stochastic greybox model parameters using maximum likelihood estimation and the Extended Kalman Filter for state estimation. Section \cite{4} describes our case study and the treatment process. The case
study applies the greybox model and state estimation to real plant data. Finally, Section 5 provides conclusions.

2. Model of the activated sludge process

Since one of the most promising fields for MPC in WWTP is energy optimization, we focus our model on nitrogen removal. This removal process requires aeration of wastewater, which is the most energy demanding process in treatment plants. Among the available models for modelling processes in WWTPs [3], we selected the ASM1 model [29], since the additional processes included in the later ASM models (e.g. P removal) are not relevant for our purpose. The ASM1 and its expansions [6] serve as great tools for detailed simulation studies of activated sludge processes, and are widely used for benchmarking control strategies [7]. However, the full model (see section 6) is not well-suited for practical online process control purposes in WWTPs. Partly due to the large amount of parameters entering its nonlinear model equations that change over different time scales, and partly because online sensors for measuring all variables in the model are simply not technologically available, too costly, or too difficult to maintain in practice. So all parameters can not be identified by statistical methods and the currently available measurements [23]. Fortunately, the most important parameters in the processes can be identified in reduced versions of the ASM. Ammonium NH$_4$, nitrate NO$_3$ and phosphate PO$_4$ (which is not included in ASM1) can be measured online and combined with stochastic greybox models to provide a sufficient model for control and short term prediction.

2.1. Reduced greybox model of BioDenitro process

Based on the assumptions made in Section 6.2, the equations for simulating removal of nitrogen in WWTP can be reduced to the following:

\[ \dot{S}_{NH} = -k_{NH}^O \rho - k_{NH}^{NO} \rho - k_{NH}^{PNH} \]  
\[ \dot{S}_{NO} = -k_{NO}^{O} \rho + k_{NO}^{NH} \rho \]  
\[ \dot{S}_O = -k_{O}^{O} \rho - k_{O}^{NO} \rho \]  

Where the reactions rates are

\[ \rho_O = \frac{S_O}{K_{O,OH} + S_O} \]  
\[ \rho_{NO} = \frac{S_O}{K_{O,OH} + S_O + K_{NO,OH} + S_{NO}} \]  
\[ \rho_{NH} = \frac{S_O}{K_{O,ANO} + S_O + K_{NH,ANO} + S_{NH}} \]

This simplified model utilizes the half-saturation constants $K$, which can be compared to physical values from literature, while a range of newly introduced constants $k$ cannot be translated to physically interpretable parameters but can still be calculated by using the equations listed in [4]. The simplified model has only three state variables (which can be measured online): oxygen $S_O$, nitrate and nitrite $S_{NO}$, and ammonia $S_{NH}$ concentrations.

The dynamics of the dissolved oxygen $S_O$ are very fast and often a local control loop regulates the oxygen concentration to a setpoint. [30, 31, 32, 33] model this control loop by including $K_{La}$ in the $S_O$ equation. We assume a PID controller follows a setpoint and controls $S_O$ directly. This reduces the model to only two states.

If we insert typical ASM1 values for all these parameters we see that $k_{NH}^{PNH}$ and $k_{NO}^{NH}$ are much bigger than the other terms and dominate the equations. Consequently, we end up with

\[ \dot{S}_{NH} = -k_{NH}^{PNH} \rho \]  
\[ \dot{S}_{NO} = k_{NH}^{PNH} \rho \]  
\[ \rho_{NH} = \frac{S_O}{K_{O,ANO} + S_O + K_{NH,ANO} + S_{NH}} \]

We use this reduced model for estimation in the remaining part of the paper.

2.2. Adding transport flows

The nitrogen removal process takes place in a tank reactor with constant volume $V$, which is assumed to be a Continuously Stirred Tank Reactor (CSTR). Under these common assumptions the mass balance for $S_j$ with process reaction rate $\rho_j$ is

\[ \dot{S}_j = \frac{Q}{V} (S_j^m - S_j) + \rho_j \]  

With no reaction, $\rho_j = 0$, and constant flow $Q$ this is simply a first order low pass filter smoothing the inlet concentration. The time constant or hydraulic retention time of the tank is $\tau = \frac{V}{Q}$. Obviously, a higher flow yields a faster response towards a steady state where $S_j = S_j^m$. For a given treatment process the transport flows can be added to each state variable in the reduced ASM1 model [3].

3. Stochastic greybox model likelihood parameter estimation

The developed dynamical model of the activated sludge process is time-varying and nonlinear. We wish to estimate the model parameters online from data and evaluate the model uncertainty, so we can use it for prediction and control. The model fits the general model structure for continuous-discrete stochastic state space models, i.e. a model of the state variables in continuous time and discrete time samples measurements of some of the states.

\[ dx_t = f(x_t, u_t, t, \theta) dt + \sigma(u_t, t, \theta) d\omega_t \]  
\[ y_k = h(x_t, u_t, t, \theta) + e_k \]
where \( t \in \mathbb{R} \) is time and \( k \in \mathbb{Z} \) is the discrete time index of the sampled time instants \( t_k \) spaced with sampling period \( T_s \), \( x_t \in \mathbb{R}^n \) contains the \( n \) state variables, \( u_t \in \mathbb{R}^m \) is the \( m \) input variables, \( y_t \in \mathbb{R}^l \) is the \( l \) output variables while \( \theta \in \mathbb{R}^p \) contains the \( p \) model parameters to be estimated. \( f(\cdot) \in \mathbb{R}^n \), \( \sigma(\cdot) \in \mathbb{R}^{n \times m} \), and \( h(\cdot) \in \mathbb{R}^m \) are nonlinear functions. When these functions are linear \( \mathbb{E} \) reduce to linear state space models. \( \omega_t \) is an \( n \)-dimensional standard Wiener process and \( e_t \) is an \( l \)-dimensional white noise process with \( e_t \in N(0, S(u_t, l_t, \theta)) \) characterizing the measurement noise.

Given a model structure and a set of \( N + 1 \) measurements \( \mathcal{Y}_N = \{y_N, y_{N-1}, \ldots, y_1, y_0\} \) we wish to find the parameters \( \theta \) that maximizes the likelihood function

\[
L(\theta, \mathcal{Y}_N) = p(\mathcal{Y}_N|\theta) = p(y_0|\theta) \prod_{k=1}^{N} p(y_k|\mathcal{Y}_{k-1}, \theta) \tag{6}
\]

\( p(y_k|\mathcal{Y}_{k-1}, \theta) \) is a conditional density denoting the probability of observing \( y_k \) given the previous observation set \( \mathcal{Y}_{k-1} \) and parameters \( \theta \). For sufficiently fast sampled measurements we approximate this conditional density with a Gaussian distribution such that

\[
p (y_k|\mathcal{Y}_{k-1}, \theta) = \frac{\exp\left( -\frac{1}{2} \frac{R_{k-1}^T e_k}{R_{k-1}} \right)}{\sqrt{2\pi} R_{k-1}} \tag{7}
\]

where

\[
e_k = y_k - \hat{y}_{k-1} \tag{8a}
\]

\[
\hat{y}_{k-1} = E[y_k|\mathcal{Y}_{k-1}, \theta] \tag{8b}
\]

\[
R_{k-1} = V[y_k|\mathcal{Y}_{k-1}, \theta] \tag{8c}
\]

defines the residual output \( \epsilon \) and estimated mean output \( \hat{y} \) with covariance \( R \). A Kalman filter provides exactly these estimates for Gaussian noise densities, even for nonlinear models in the extended linearizing version the Extended Kalman Filter (EKF) \([13]\).

If we assume the initial probability density \( p(y_0|\theta) \) to be known and put (7) into the likelihood function (6), then we can find a set of parameters \( \theta \) that maximizes the likelihood function by solving the following nonlinear optimization problem

\[
\max_{\theta \in \Theta} \log[L(\theta, \mathcal{Y}_N|x_0)] \tag{9}
\]

Including the logarithm decomposes the likelihood function into sums instead of products that is easier to handle when solving the optimization problem. We assume each parameter to take values within predefined bounds with equal probability, so the parameters are constrained to the set \( \Theta \). The problem can also be extended to maximum a posteriori (MAP) estimation, where instead of giving simple parameter bounds, a prior probability density function limits the parameter value. This is useful if some prior knowledge about the probability density is available.

### 3.1. Software

We use R and the package CTSM-R \([35, 36, 37]\) to model and estimate the parameters using Maximum Likelihood Estimation. CTSM-R solves the optimization problem (9) and provides estimates of the parameters and their uncertainty. It also includes outlier detection, handles irregular sampling and can combine multiple independent data sets into the same estimation procedure. Finally, the implemented Extended Kalman Filter estimates any missing observations caused by calibrating sensors or communication faults.

### 4. Case study: Kolding WWTP

Kolding WWTP in Denmark runs a BioDenitro™ process with a treatment capacity of 125,000 PE and a load of 75,000 PE. We use data from this plant as a case study for validating our greybox model and running the state estimation. We wish to estimate the NH\(_4\) concentration in both tanks. Estimating unmeasured states accurately leads to better control performance without adding costly additional sensors.

#### 4.1. BioDenitro operation

The BioDenitro process involves two hydraulically connected tank reactors that are operated with equal loading in counter phase. The continuous influent wastewater is divided between the reactors while the operating conditions alternates between aerobic or anoxic conditions. The influent primarily flows to the anoxic reactor to utilize its organic carbon content for denitrification. The reactor role is toggled as soon as the NH\(_4\) and NO\(_3\) concentrations in the anoxic reactor reach a set point or if a process time limit is exceeded. Each reactor is equipped with actuators for aeration and propellers for mixing. The BioDenitro process approaches a batch process, because the phases are short relative to the hydraulic retention time and the inlet wastewater volume is small compared to the total tank volume. Figure 2 shows the operation cycle of a typical BioDenitro plant with indication of flow patterns and phases. A RBC strategy that changes the cycle lengths depending on online measurements \([38]\) is the commercially available control system STAR for the BioDenitro process developed by Krüger A/S. After its development in the early 1990s \([39, 40, 41]\) approximately 400 plants primarily in Denmark run the BioDenitro process (or BioDeniPho if Phosphorus is treated as well). The process is very flexible and can adapt to different waste water compositions. The periodically higher concentrations yields faster reaction times and the continuous alternating excitation of the dynamics makes it easier to estimate the process rates through mathematical system identification. However, the phase control design is quite complex where nitrification and denitrification are performed sequentially in a semi-batch manner by cyclically switching flow and aeration pattern.

Kolding WWTP has four aeration tanks, i.e. two BioDenitro tank sets, equipped with two \( O_2 \) sensors each. Figure 1 shows the plant layout from above. Figure 2 shows the flow directions and sensors in one BioDenitro tank set at Kolding. All possible flow combinations and aeration on/off patterns are described in \([39]\). At Kolding only one tank, i.e. the master tank A, has a NH\(_4\) and a NO\(_3\) sensor, and in this case study we estimate the NH\(_4\) in tank B. Estimating the concentration in
tank B will increase the control performance in terms of better effluent quality and lower power consumption without an expensive physical sensor. As the two tanks sets are disconnected and only one tank set have NH4 and NO3 sensors we can not estimate the concentrations in the second tank set. So the other tank set simply copy the control action from tank A and B in an open loop manner.

4.2. Data

We have measurements of NH4, NO3, O2 and flows between the tanks from Kolding WWTP. In total we use a 48 hour data set for the greybox model estimation. The sensors sample every 5 min while the process control system samples every 2 min. So the individual sensor measurements are collected multiple times and repeated in the data either two or three times. Consequently, to pick out the actual observations, we excluded all points as missing observations when their finite difference was very low, ideally zero. CTSM−R easily handles these missing observations and non-equidistant samples.

The sensors also calibrate automatically 3−4 times per day leaving the control with no measurements up to 1 hour.

For the given data set we time shifted the observations 8 samples to account for sensor and actuation delays. The delay might vary over time and can be included in the model or estimated independently.

4.3. Reduced greybox model of BioDenitro process

We build a model of the NH4 concentrations in tank A and B using \( S_{NH}^{in} \). As we assume that the incoming waste water contains no NO3 or O2 such that \( S_{O}^{in} = 0 \) and \( S_{O}^{in} = 0 \).

\[
\begin{align}
S_{NH}^{A} &= \theta_1 Q \left[ S_{NH}^{in} w_i + S_{NH}^{B} \bar{w}_i - S_{NH}^{A} \right] - \\
\theta_2 &K_{NH}^{A} + S_{NH}^{A} K_{O}^{A} + S_{NH}^{A} \\
S_{NH}^{B} &= \theta_1 Q \left[ S_{NH}^{in} \bar{w}_i + S_{NH}^{A} w_i - S_{NH}^{B} \right] - \\
\theta_4 &K_{NH}^{B} + S_{NH}^{B} K_{O}^{B} + S_{NH}^{B}
\end{align}
\]  

(10a)

\[
\begin{align}
\dot{x} &= \begin{bmatrix} S_{NH}^{A} \quad S_{NH}^{B} \quad \theta_1 Q \quad \theta_2 \quad \theta_3 \quad \theta_4 \quad K_{NH}^{A} \quad K_{NH}^{B} \quad \sigma_3 \end{bmatrix} \end{align}
\]

The parameters \( \theta \) contain the tank volume and we have the unknown parameters: \( \theta = (\theta_1, \theta_2, \theta_3, \theta_4, K_{NH}^{A}, K_{NH}^{B}, K_{NH}^{A}, K_{NH}^{B}, \sigma_3) \).

4.4. Estimating unmeasured NH4 concentrations

Unlike the inlet flow input, the other important part of the waste water load, namely the NH4 inlet concentration, is not measured. We know that this concentration follows the inlet flow with a diurnal profile. We estimate it using the EKF and the greybox model by writing the model equations (10) as a set of stochastic differential equations on the form (11).

\[
dS_{NH}^{in} = 0 dt + \sigma_3 d\omega_3, \quad \sigma_3 > 0
\]

(11)

This is a random walk model for \( S_{NH}^{in} \) that allows \( S_{NH}^{in} \) to vary in a flexible manner. Any model discrepancies will be transferred to this state and might disrupt the physical interpretation of the concentration. Since the NH4 concentration in tank B, \( S_{NH}^{B} \), is also not measured, we also get an estimate of this concentration.

Figure 3 shows the model fit on the estimation data set. The one-step ahead predictions fit the data very well with low residuals. All missing observations and unmeasured states are estimated by the EKF. As expected the NH4 concentration in tank B is similar to tank A but time-shifted due to the alternating pattern. The diurnal pattern of the NH4 inlet concentration is also clearly evident but not completely smooth. This is mainly due to the sensor placement and practical non-ideal mixing conditions. Also the NH4 concentrations are controlled to low values due to the treatment process, so consequently we see only short periods of mixing where the load goes towards the higher NH4 inlet concentration.

The parameter estimation for 48 hours of data using CTSMR took around 10 seconds on a laptop but is running online on a server every 2 minutes as part of the plant control software.
Figure 3: Estimated one-step predictions of NH₄ concentrations in both tanks and the inlet. The dotted line in the upper plot is the intermittent aeration pattern.

Figure 4: Parameter estimates.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{NH₄}^{o}$</td>
<td>2.29e+00</td>
<td>6.99e-02</td>
</tr>
<tr>
<td>$S_{NH₄}^{ purchaser}$</td>
<td>5.35e-01</td>
<td>1.25e+00</td>
</tr>
<tr>
<td>$S_{NH₄}^{in}$</td>
<td>9.66e+00</td>
<td>2.96e+00</td>
</tr>
<tr>
<td>$K_{NH}$</td>
<td>3.49e+01</td>
<td>9.06e+02</td>
</tr>
<tr>
<td>$K_{NH}$</td>
<td>1.44e+00</td>
<td>9.26e-01</td>
</tr>
<tr>
<td>$K_{A}$</td>
<td>1.76e+01</td>
<td>8.39e-02</td>
</tr>
<tr>
<td>$K_{OA}$</td>
<td>1.74e+00</td>
<td>1.17e+00</td>
</tr>
<tr>
<td>$\sigma_{NH₄}$</td>
<td>3.78e-04</td>
<td>3.21e-04</td>
</tr>
<tr>
<td>$\sigma_{A}$</td>
<td>3.38e-02</td>
<td>2.37e-03</td>
</tr>
<tr>
<td>$\sigma_{B}$</td>
<td>7.09e-07</td>
<td>2.25e-05</td>
</tr>
<tr>
<td>$\sigma_{NH₄min}$</td>
<td>2.07e-01</td>
<td>7.82e-02</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>5.96e-02</td>
<td>2.05e-02</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>7.32e-02</td>
<td>7.18e-03</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>6.39e-02</td>
<td>1.52e-02</td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>9.34e-01</td>
<td>8.09e-01</td>
</tr>
</tbody>
</table>

5. Conclusion

The method provides a greybox model for the process which describes the uncertainty, and state estimation which are critical for Model Predictive Control purposes. We built a stochastic greybox model of a BioDenitro process aiming for a time horizon below 24 hours. An R-package, CTSMR, estimates the model parameters using maximum likelihood optimization and the Extended Kalman Filter. If run repeatedly, e.g. in moving 48 hours windows, the algorithm adaptively fits the model parameters to the current operating conditions and fills in missing sensor measurements. This is particularly important in the frequent and long periods where the sensors are calibrating and not providing measurements to the control system. Since we estimate missing observations, the inlet load, and the concentrations in the sensorless tank, we could obtain a more accurate control performance. The model is an important piece in predicting the process and applying Model Predictive Control.

References


URL http://linkinghub.elsevier.com/retrieve/pii/S0261093003002044


URL http://linkinghub.elsevier.com/retrieve/pii/0043135494900043


B. Chachuat, N. Rouche, M. A. Latifi, Reduction of the ASM1 model for optimal control of small-size activated sludge treatment plants.


The state variables are presented in columns, while processes are in the rows. The mass balance for the model can be described by 9 variables and 8 processes (see Table 1).

\[
\begin{align*}
\dot{x}_i &= \sum_{j=1}^{8} r_{ij} - y_i, \\
\text{Input} - \text{Output} &= \sum_{i=1}^{9} (r_{ii} - y_i)
\end{align*}
\]

The reaction rate \( r_i \) [\( \text{m}^3 \text{g}^{-1} \text{d}^{-1} \)] for each \( i \)-th state variable is then calculated as

\[
r_i = \sum_{j=1}^{8} r_{ij}
\]

The rates for the \( i \)-th process. By using this framework, the ASM1 model can be described by 9 variables and 8 processes (see Table 1).

### 6.1. Activated Sludge Model no. 1 (ASM1)

In the following sections we firstly introduce the ASM1 model and then we simplify it based on several assumptions that are not when looking at predictive control and probabilistic forecasting up to 24 hours ahead.

### 6.1.1. ASM1 differential equations

6.1.1. ASM differential equations

Based on the tables provided in the previous section, we can write up the set of first order ordinary differential equations as:

\[
\begin{align*}
\dot{X}_1 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_2 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_3 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_4 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_5 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_6 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_7 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_8 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0, \\
\dot{X}_9 &= \mu_{B,0,1} X_0 S_0 - \mu_{B,0,1} X_0 Y_0 - b_{N, Bio} X_0
\end{align*}
\]

\[
\begin{align*}
\dot{Y}_0 &= b_{N, Bio} X_0, \\
\dot{Y}_1 &= b_{N, Bio} X_0, \\
\dot{Y}_2 &= b_{N, Bio} X_0, \\
\dot{Y}_3 &= b_{N, Bio} X_0, \\
\dot{Y}_4 &= b_{N, Bio} X_0, \\
\dot{Y}_5 &= b_{N, Bio} X_0, \\
\dot{Y}_6 &= b_{N, Bio} X_0, \\
\dot{Y}_7 &= b_{N, Bio} X_0, \\
\dot{Y}_8 &= b_{N, Bio} X_0
\end{align*}
\]

For practical purposes, the ASM models are commonly presented by using the Gujer matrix representation (see Table 1).

#### 6.2. Appendix

- **Table 1:** Gujer matrix for ASM1 (from [5], adapted with the new notation from [7])

<table>
<thead>
<tr>
<th>Component →</th>
<th>Process ↓</th>
<th>( S_0 )</th>
<th>( X_0 )</th>
<th>( X_{OHO} )</th>
<th>( X_{ANO} )</th>
<th>( S_0 )</th>
<th>( S_{NO} )</th>
<th>( S_{B, org} )</th>
<th>( S_{X, org} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aerobic Growth of heterotrophs</td>
<td>1</td>
<td>1</td>
<td>( \frac{1}{Y_{OHO}} )</td>
<td>( \frac{1}{Y_{OHO}} )</td>
<td>( -b_{N, Bio} )</td>
<td>( \frac{S_0}{K_{S, OHO} + S_0} )</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Anoxic Growth of heterotrophs</td>
<td>1</td>
<td>1</td>
<td>( \frac{1}{Y_{OHO}} )</td>
<td>( \frac{1}{Y_{OHO}} )</td>
<td>( -b_{N, Bio} )</td>
<td>( \frac{S_0}{K_{S, OHO} + S_0} )</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Aerobic Growth of autotrophs</td>
<td>1</td>
<td>1</td>
<td>( \frac{4.37 - Y_{ANO}}{Y_{ANO}} )</td>
<td>( \frac{1}{Y_{ANO}} )</td>
<td>( -b_{N, Bio} )</td>
<td>( \frac{S_0}{K_{S, OHO} + S_0} )</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Decay of heterotrophs</td>
<td>1</td>
<td>1</td>
<td>( -f_{B, X} )</td>
<td>( f_{B, X} )</td>
<td>( b_{B, X} )</td>
<td>1</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Decay of autotrophs</td>
<td>1</td>
<td>1</td>
<td>( -f_{B, X} )</td>
<td>( f_{B, X} )</td>
<td>( b_{B, X} )</td>
<td>1</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Ammonification of soluble organic nitrogen</td>
<td>1</td>
<td>1</td>
<td>( -1 )</td>
<td>( -1 )</td>
<td>( q_{am} )</td>
<td>1</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Hydrolysis of entrapped organics</td>
<td>1</td>
<td>1</td>
<td>( -1 )</td>
<td>( -1 )</td>
<td>1</td>
<td>1</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Hydrolysis of entrapped organic nitrogen</td>
<td>1</td>
<td>1</td>
<td>( -1 )</td>
<td>( -1 )</td>
<td>1</td>
<td>1</td>
<td>( X_{OHO} )</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: List of state variables for ASM1 (from [6]).

<table>
<thead>
<tr>
<th>State Variable</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_B$</td>
<td>gCODm$^{-1}$</td>
<td>Readily biodegradable substrate</td>
</tr>
<tr>
<td>$X_U$</td>
<td>gCODm$^{-1}$</td>
<td>Slowly biodegradable substrate</td>
</tr>
<tr>
<td>$X_{OHO}$</td>
<td>gCODm$^{-1}$</td>
<td>Ordinary heterotrophic organisms</td>
</tr>
<tr>
<td>$X_{ANO}$</td>
<td>gCODm$^{-1}$</td>
<td>Autotrophic nitifying organisms</td>
</tr>
<tr>
<td>$S_O$</td>
<td>gNm$^{-3}$</td>
<td>Oxygen (negative COD)</td>
</tr>
<tr>
<td>$S_{NO}$</td>
<td>gNm$^{-3}$</td>
<td>Nitrate and nitrite nitrogen</td>
</tr>
<tr>
<td>$S_{NH}$</td>
<td>gNm$^{-3}$</td>
<td>Ammonium plus ammonia nitrogen</td>
</tr>
<tr>
<td>$S_{NBorg}$</td>
<td>gNm$^{-3}$</td>
<td>Soluble biodegradable organic nitrogen</td>
</tr>
<tr>
<td>$X_{NBorg}$</td>
<td>gNm$^{-3}$</td>
<td>Particulate biodegradable organic nitrogen</td>
</tr>
</tbody>
</table>

Table 3: List of stoichiometric parameters for ASM1 (from [6]).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{B,a}$</td>
<td>–</td>
<td>Fraction of biomass yielding particulate products</td>
</tr>
<tr>
<td>$1s_{B,a}$</td>
<td>gN(gCOD)$^{-1}$</td>
<td>Mass N / Mass COD in biomass</td>
</tr>
<tr>
<td>$1y_{B,a}$</td>
<td>gN(gCOD)$^{-1}$</td>
<td>Mass N / Mass COD in products from biomass</td>
</tr>
<tr>
<td>$Y_{ANO}$</td>
<td>g(XO/ANO)(gN)$^{-1}$</td>
<td>Autotrophic yield</td>
</tr>
<tr>
<td>$Y_{OHO}$</td>
<td>g(XO/OHO)(gN)$^{-1}$</td>
<td>Heterotrophic yield</td>
</tr>
</tbody>
</table>

Table 4: List of kinetic parameters for ASM1 (from [6]).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_{ANO}$</td>
<td>–</td>
<td>Autotrophic yield</td>
</tr>
<tr>
<td>$Y_{OHO}$</td>
<td>–</td>
<td>Heterotrophic yield</td>
</tr>
<tr>
<td>$K_{NH}$</td>
<td>gNO/m$^{-3}$</td>
<td>Ammonium and amino nitrogen half-saturation coefficient for autotrophs</td>
</tr>
<tr>
<td>$K_{NO,ANO}$</td>
<td>gNO/m$^{-3}$</td>
<td>Nitrate and nitrite half-saturation coefficient for autotrophs</td>
</tr>
<tr>
<td>$K_{NO,OHO}$</td>
<td>gNO/m$^{-3}$</td>
<td>Oxygen half-saturation coefficient for autotrophs</td>
</tr>
<tr>
<td>$K_{O,H}$</td>
<td>gCOD/m$^{-1}$</td>
<td>Oxygen half-saturation coefficient for heterotrophs</td>
</tr>
<tr>
<td>$K_{NO,H}$</td>
<td>gCOD/m$^{-1}$</td>
<td>Oxygen half-saturation coefficient for heterotrophs</td>
</tr>
<tr>
<td>$q_{ANO}$</td>
<td>–</td>
<td>Autotrophic nitifier growth rate</td>
</tr>
<tr>
<td>$q_{OHO}$</td>
<td>–</td>
<td>Heterotrophic growth rate</td>
</tr>
<tr>
<td>$q_{Am}$</td>
<td>m$^3$COD$^{-1}$</td>
<td>Ammonification rate</td>
</tr>
<tr>
<td>$q_{Hyd}$</td>
<td>g slowly biodegrad. COD (g cell COD d)$^{-1}$</td>
<td>Hydrolysis rate</td>
</tr>
</tbody>
</table>

scribing the ASM1 dynamics

$$S_{NH} = -i_{XB}(p_1 + p_2) - \left( i_{XB} + \frac{1}{Y_A} \right) p_3 + k_6 S_{ND} X_{BH}$$ (14a)

$$S_{NO} = -\frac{1}{Y_H} Y_H p_2 + \frac{1}{Y_A} p_3$$ (14b)

$$S_{O} = -\frac{1}{Y_H} Y_H p_1 - 4.57 - Y_A p_3$$ (14c)

$$\dot{S} = \rho_f - \frac{1}{Y_H} (p_1 + p_2)$$ (14d)

$$\dot{X} = (1 - f_p) (b_H X_{BH} + b_A X_{BA}) - \rho_f$$ (14e)

$$X_{BH} = p_1 + p_2 - b_H X_{BH}$$ (14f)

$$X_{BA} = p_3 - b_A X_{BA}$$ (14g)

$$\dot{S}_{ND} = p_8 - k_6 S_{ND} X_{BH}$$ (14h)

$$\dot{X}_{ND} = (i_{XB} - f_p i_{XP})(b_H X_{BH} + b_A X_{BA}) - p_8$$ (14i)

The four remaining original ASM1 variables $S_L$, $X_I$, $X_P$, and $S_{ALK}$ do not affect any other state variables in (14) and were left out. [29] provides details about all the variables and parameters and these will not be repeated here. The state variables are influenced by the nonlinear process rates

$$\rho_1 = \mu_H S_{S} \frac{S_{O}}{K_{O,H} + S_{O}} X_{BH}$$ (15a)

$$\rho_2 = \mu_H S_{S} \frac{S_{O}}{K_{O,H} + S_{O}} \frac{S_{NO}}{S_{NO} + S_{NO}} X_{BH}$$ (15b)

$$\rho_3 = \mu_A S_{NH} \frac{S_{NO}}{S_{NO}} X_{BA}$$ (15c)

$$\rho_f = k_b 1 \frac{S_{NH} X_{BH}}{X_{BH}} \left( \frac{S_{NO}}{K_{O,H} + S_{O}} \frac{S_{NO}}{S_{NO} + S_{NO}} X_{BH} \right)$$ (15d)

$$\rho_f = \rho_f (X_{ND}/X_{S})$$ (15e)

The Monod-expressions serve as continuous concentration-dependent switching functions for the different process rates. Mainly to distinguish between aerobic ($S_O > 0$) and anoxic ($S_O = 0$, $S_{NO} > 0$) conditions. If the variable in the Monod-expression is rarely in the nonlinear switching region, i.e. near the half-saturation constant $K$, the Monod-expression can be approximated by a binary signal.

6.2. Short time horizon ASM1 assumptions

The Monod kinetics operate as switching functions, and if the variable in the Monod-expression is rarely in the nonlinear switching region, i.e. near the half-saturation constant $K$, then the Monod-expression can be approximated by a binary signal.

A great number of variables in the ASM1 change very little on a daily basis. [42] [43] manually estimate and tune the parameters while [43] [44] use a linear model. In this paper, we aim to build a model that is accurate on an hourly timescale in dry weather operation and we assume that

$$X_{OHO} \approx 0 \quad X_{ANO} \approx 0 \quad S_B \approx 0 \quad X_U \approx 0$$
i.e., we assume that the biomass (\(X_{\text{NH}}\) and \(X_{\text{ANO}}\)) could be replaced by simple functions of measured SS concentration. This slowly varying parameter could be added to model longer term prediction or rain weather operation. The substrate (\(S_B\) and \(X_U\)) can be lumped into one state variable, assuming that measurements of chemical oxygen demand COD cannot distinguish between soluble and particulate components, i.e. COD is measured as total concentration and then subdivided into the two components by assuming a fraction \(f\). The two variables accounting for substrate can therefore be expressed as

\[
S_B = fS_{\text{COD}}\,\text{COD}_{\text{tot}} \quad X_U = fX_{\text{COD}}\,\text{COD}_{\text{tot}}
\]

We further assume that there is no ammonium production from organic bound nitrogen. Consequently, \(S_{NB,\text{org}} = 0\).

This eliminates \(X_{NB,\text{org}}\) as well, leaving only the three most important and often measured state variables \(S_{NH}, S_{NO},\) and \(S_O\). Most control-oriented models contain only these three states \([45][46][47][48][49]\), some also \(S_B\) \([50][51]\).

The model complexity reduces significantly when the slowly varying states are assumed locally constant and is routinely done for real-time control-oriented ASP models \([52][53]\). The reduced model is then only valid for shorter prediction horizons. However, if the model parameters are re-calibrated frequently the model will adapt to the slowly changing processes. We re-calibrate the model by estimating all parameters and states using the latest measurement window.

Based on these assumptions, the ASM1 model (Table 1) becomes a model with only three variables and three removal processes (Table 5), where we can define new constants related to the original ASM1 parameters:

\[
\begin{align*}
  k^0_{NH} &= i_{OHO}k_{NH}, \\
  k_{NH} &= \hat{\mu}_{OHO,\text{max}} \left( \frac{S_B}{K_S + S_B} \right) X_{OHO} \\
  k^0_{NO} &= i_{OHO}k_{NO}, \\
  k_{NO} &= \eta_{OHO}k_{NH} \\
  k^0_{NH} &= \left( i_{OHO} + \frac{1}{Y_{ANO}} \right) k_O \\
  k_{O} &= \hat{\mu}_{ANO,\text{max}} X_{ANO} \\
  k^0_{NO} &= 1 - \frac{Y_{OHO}}{2.68Y_{OHO}}k_{NO} \\
  k_{O} &= 1 - \frac{Y_{OHO}}{Y_{OHO}}k_{NH} \\
  k^0_{NH} &= \frac{1}{Y_{ANO}} k_O \\
  k_{NO} &= \frac{4.57 - Y_{ANO}}{Y_{ANO}}k_O
\end{align*}
\]

<table>
<thead>
<tr>
<th>Component →</th>
<th>Process ↓</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Acrobic Growth of heterotrophs</td>
<td>(1 - Y_{OHO} - \frac{Y_{ANO}}{4.57} - \frac{Y_{ANO}}{2.68Y_{OHO}})</td>
<td>(-\frac{Y_{ANO}}{4.57} - \frac{Y_{ANO}}{2.68Y_{OHO}})</td>
<td>(-\frac{Y_{ANO}}{4.57} - \frac{Y_{ANO}}{2.68Y_{OHO}})</td>
</tr>
<tr>
<td>2</td>
<td>Anaerobic Growth of heterotrophs</td>
<td>(1 - Y_{OHO})</td>
<td>(-\frac{Y_{ANO}}{4.57})</td>
<td>(-\frac{Y_{ANO}}{4.57})</td>
</tr>
<tr>
<td>3</td>
<td>Anaerobic Growth of autotrophs</td>
<td>(-\frac{Y_{ANO}}{4.57})</td>
<td>(-\frac{Y_{ANO}}{4.57})</td>
<td>(-\frac{Y_{ANO}}{4.57})</td>
</tr>
</tbody>
</table>

Table 5: Guyer matrix for the reduced model of the BioDenitro process.