Mathematical approach for improving the reliability of parameter calibration in modeling of anaerobic digestion processes

Konrad KOCH1), Tito GEHRING2), Manfred LÜBKEN2), Marc WICHERN2) and Harald HORN3)  
1) Institute of Water Quality Control, Technische Universität München, Am Coulombwall, 85748 Garching, Germany.  
2) Institute of Environmental Engineering, Ruhr-Universität Bochum,Universitätsstraße 150, 44780 Bochum, Germany.  
3) Water Chemistry and Water Technology, Karlsruhe Institute of Technology, Engler Bunte Ring 1, 76131 Karlsruhe, Germany.

wga@bv.tum.de

Keywords  
Mathematical modeling  
ADM1  
Parameter calibration  
Nash-Sutcliffe coefficient  
Equifinality

1. Introduction

Anaerobic digestion is a complex biological process requiring the involvement of a variety of microorganisms. Nowadays, biogas plants are usually operated at relatively low organic loading rates due to fear of overload and failure and the subsequent financial consequences. For example, ten medium-sized agricultural biogas plants in Germany have an average organic loading rate (OLR) of 2.5 kgVS/(m³·d) with a range between 1.5 and 3.8 kgVS/(m³·d) [1]. However, wet fermentation systems could reach OLRs of up to 4 kgVS/(m³·d) [2]. A model based tool which provides plant operators with suggestions concerning ongoing operation may be helpful. Achievement of a stable process at higher organic loading rates would consequently increase the ecological and economical effectiveness of agricultural biogas production.  

Mathematical models can be useful for understanding the numerous processes or microorganisms involved. Thereby, the influences of parameters on the process can be identified and approaches for optimization are revealed. After successful calibration, a mathematical model is able to predict reactor behavior even under varying conditions. Originally, modeling of the anaerobic digestion process was performed with respect to synthetic substrates such as acetate [3], waste material (e.g. animal manure [4]), and waste water sludge [5]. Up to the end of the 20th century, a large
variety of models already existed, however most of them are designed for specific substrates or reactor types. In order to consolidate the mathematical description of the anaerobic process, the IWA task group on the Mathematical Modeling of Anaerobic Digestion Processes was founded in 1997 and resulted in the publication of Anaerobic Digestion Model No. 1 (ADM1) in 2002 [6]. In contrast to most of the parent models, ADM1 is universally applicable and nowadays it is the most complex model for anaerobic digestions modeling and is therefore widely used. But as the model was developed for treatment of activated and primary sludge, model parameters are calibrated for these kinds of substrates. Nevertheless, there is an increasing trend of simulations being used for agricultural substrates [7]. Examples include cattle manure [8, 9], cattle manure and maize [10], cattle manure and co-substrates [11] and grass silage [12, 13]. Due to this development, calibration of kinetic parameters to both different substrates and conditions are required. Especially when using a complex model with several kinetic parameters such as ADM1, sensitivity analyses are valuable in reducing parameter calibration to the sensitive one.

<table>
<thead>
<tr>
<th>Symbols/Oznake</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADM1</td>
<td>Anaerobic Digestion Model No. 1</td>
</tr>
<tr>
<td>COD</td>
<td>Chemical oxygen demand, mg/l</td>
</tr>
<tr>
<td>E_{NSC}</td>
<td>Nash-Sutcliffe model efficiency coefficient</td>
</tr>
<tr>
<td>IWA</td>
<td>International Water Association</td>
</tr>
<tr>
<td>K_{I,H2}</td>
<td>Hydrogen inhibition constant for valerate and butyrate uptake, kgCOD/m³</td>
</tr>
<tr>
<td>K_{I,H2,C4}</td>
<td>Inhibicija konstante vodika za smjesu i unos butirata, kgCOD/m³</td>
</tr>
<tr>
<td>K_{I,H2,Pro}</td>
<td>Hydrogen inhibition constant for propionate uptake, kgCOD/m³</td>
</tr>
<tr>
<td>K_{I,NH3}</td>
<td>Konstanta inhibicije vodika za unos propionata, kgCOD/m³</td>
</tr>
<tr>
<td>k_m_ac</td>
<td>Maximum uptake rate acetate, kgCOD/(kgCOD·d)</td>
</tr>
<tr>
<td>k_m_C4</td>
<td>Maximum uptake rate valerate and butyrate, kgCOD/(kgCOD·d)</td>
</tr>
<tr>
<td>k_m_Pro</td>
<td>Maximum uptake rate propionate, kgCOD/(kgCOD·d)</td>
</tr>
<tr>
<td>K_{S,Ac}</td>
<td>Half-saturation coefficient of acetate, kgCOD/m³</td>
</tr>
<tr>
<td>K_{S,C4}</td>
<td>Half-saturation coefficient of hydrogen, kgCOD/m³</td>
</tr>
<tr>
<td>K_{S,Pro}</td>
<td>Half-saturation coefficient of valerate and butyrate, kgCOD/m³</td>
</tr>
<tr>
<td>K_{I_H2_C4}</td>
<td>Konstanta inhibicije bez amonijaka za unos acetata, kmoL /m³</td>
</tr>
<tr>
<td>OLR</td>
<td>Organic loading rate, kgVS/(m³·d)</td>
</tr>
<tr>
<td>TS</td>
<td>Total solids, %</td>
</tr>
<tr>
<td>VFA</td>
<td>Volatile fatty acids</td>
</tr>
<tr>
<td>VS</td>
<td>Volatile solids, % TS</td>
</tr>
</tbody>
</table>

- K. KOCH et. al., Mathematical approach for improving… Strojarstvo 55 (1) 7-16 (2013)
2. Methods

2.1. Database and model set-up
Application of the modified Nash-Sutcliffe coefficient (E_{NSEC}) is exemplarily demonstrated on data from the mesophilic treatment of grass silage. Two reactors were run as duplicates, indicated in the following as reactor L and reactor R. Details about the reactor set-up and operation can be found in [14] and information concerning modeling in [13]. Simulations and sensitivity analyses were executed with SIMBA 4.2 based on Matlab/Simulink (Version 7.0.4).

2.2. Calibration of kinetic parameters
Most of the kinetic parameters of the ADM1 cannot be estimated independently from others. During anaerobic digestion, hydrogen and in particular volatile fatty acids (VFA) play a central role. An accurate reproduction of both in this respect is important for the whole simulation, as the concentration influences a multitude of other processes. Degradation of acetate, propionate, and butyrate/valerate, respectively, is characterized by the half-saturation coefficient $K_S$, the maximum uptake rate $k_m$ and the inhibition constant $K_I$ for free ammonia (Acetate) and hydrogen (propionate and butyrate/valerate). For example, Equation 1 shows the rate for butyric acid as applied in ADM1. It is assumed that butyric and valeric acid are degraded by the same type of microorganism. Consequently, both parameters are summarized as $C_4$.

$$ \rho_j = k_{m,C_4} \cdot \frac{S_{Bu}}{K_{S,C_4} + S_{Bu}} \cdot \frac{X_{C_4}}{1 + S_{V_4}/S_{Bu}} \cdot \frac{1}{1 + S_{IN}/K_{I,IN} \cdot I_{pH} \cdot I_{IN}} $$

(1)

Where $\rho_j$ is the rate of process $j$ [kg COD/(m³·d)], $k_{m,C_4}$ is the maximum uptake rate for butyrate/valerate [kg COD/(kg COD/m³)], $S_{Bu}$ is the concentration of dissolved butyric acid [kg COD/m³], $K_{S,C_4}$ is the half-saturation coefficient of butyrate/valerate [kg COD/m³], $X_{C_4}$ is the concentration of butyrate/valerate-utilizing bacteria [kg COD/m³], $S_{V_4}$ is the concentration of dissolved valeric acid [kg COD/m³], $S_{IN}$ is the concentration of hydrogen [kg COD/m³], $K_{I,H_2,C_4}$ is the hydrogen inhibition constant for butyrate/valerate [kg COD/m³] and $I_{IN}$ and $I_{pH}$ are inhibition terms for inorganic nitrogen and pH value, respectively. The biomass decay rates were not changed in the simulations presented here ($k_{dec} = 0.2$ d⁻¹ for all biomass groups).

For the uptake of acetate, the inhibition term for hydrogen is substituted by one for free ammonia. Furthermore, the term which affects the competitive inhibition by the concurrence of butyrate and valerate disappears. Already from the equation, the parameters that will probably be sensitive when enough substrate is present ($X_{C_4} >> 0$) can be deduced. The change of $k_m$ will always have an influence on simulation results; however $K_S$ will only affect the process when the concentration of the corresponding acids is quite low. In high concentrations, the influence of $K_S$ is not as great because the Monod term is converging to 1. The $K_I$ Monod term is reacting the opposite way. Its sensitivity increases when the concentration of hydrogen (or free ammonia in case of acetate) increases.

2.3. Modified Nash-Sutcliffe coefficient
Besides a mathematical description of important processes, the calibration of model parameters is fundamental for a successful simulation. The quality of the model results depends strongly on the calibration of kinetic parameters, whose estimation is usually done manually by visual comparison to experimental results.
but also to quantitatively describe the accuracy of the model outputs. In its modified version, the term of the squared differences is replaced by their absolute values to avoid sensitivity to the outliers. This modification guarantees that a difference is as dominant as two differences, each with half the value:

$$E_{NSC} = 1 - \frac{\sum_{i=1}^{n} |X_{i,\text{measured}} - X_{i,\text{simulated}}|}{\sum_{i=1}^{n} |X_{i,\text{measured}} - \overline{X}_{\text{measured}}|}$$

(3)

Where $X_{i,\text{measured}}$ and $X_{i,\text{simulated}}$ are the measured and simulated values respectively, and $\overline{X}_{\text{measured}}$ is the average of all measurements.

An efficiency of 1 corresponds to a perfect match of the modeled parameters to the observed data, where an efficiency of 0 indicates that the model prediction is as accurate as the mean of the observed data. The values below 0 imply that the simulation was worse than the simple average of the measurements.

3. RESULTS AND DISCUSSION

3.1. Application of $E_{NSC}$

The applied methodology to adapt ADM1 to grass silage followed a two-step procedure. First, all kinetic parameters were set to the suggested ADM1 values for mesophilic sludge digestion. Second, simulations were run by varying the kinetic parameters (maximum uptake rate $k_m$, half-saturation coefficient $K_S$ and inhibition constant $K_I$) of the corresponding component. The agreement of the simulation results compared to measurements has been evaluated on the basis of the modified Nash-Sutcliffe coefficient. Figure 1 depicts the measured data for the two reactors and the simulation results with two different hydrogen inhibition constants $K_{I,H_2,C4}$.

Figure 1. Simulation results and measured data for the concentration of butyrate for two different hydrogen inhibition constants $K_{I,H_2,C4}$.

Slika 1. Rezultati simulacije i mjerene vrijednosti koncentracije butirata za dvije različite konstante inhibicije vodika $K_{I,H_2,C4}$.

It is obvious that the simulation on the right expresses the dynamics of the measurements much better than the left one. Applying the modified $E_{NSC}$ results in a value of 0.24 (0.15) for the left picture and a value of 0.67 (0.57) for the right with reference to reactor L (reactor R). The visual impression could hence be quantified by the coefficient.

3.2. Identification of kinetic parameters

If the determination of the modified Nash-Sutcliffe coefficient is implemented in the model software, a value is obtained for every combination of parameters. In principle, all influencing parameters can be varied and a coefficient is obtained for every combination. For a better visualization, a variation of only two parameters is advisable.

The following simulation studies have been done based on the measurements for reactor L treating grass silage under mesophilic conditions. Graphs are structured in such a way that on the abscissa and ordinate, the variation of one kinetic parameter around a default value (multiplication by 1.0) is displayed and the agreement (as $E_{NSC}$) is depicted in grey. As it is impossible to visualize more than two parameters, the third one remained at its default value (may already be adapted). To emphasize the highest values, the negative values have been zeroed (equal to white) and the maximum value was set to black. Therefore, grey scale differs between each figure. Furthermore, only highly dynamic areas are presented.

The concentrations of hydrogen and the organic acids depend first of all on the production of the corresponding component itself and in addition to those factors discussed in the following chapter, on the decay rate of the utilizing bacteria. All biomass decay rates remained at their suggested ADM1 values.
### 3.2.1. Hydrogen

The concentration of hydrogen is only influenced by the maximum uptake rate $k_{m,H2}$ and the half-saturation coefficient $K_{S,H2}$. For creation of Figure 2, $k_{m,H2}$ has been varied around its suggested value (35 d$^{-1}$), while $K_{S,H2}$ has already been adapted to $5.6 \cdot 10^{-5}$ kgCOD/m$^3$ (suggested ADM1 value: $7 \cdot 10^{-6}$ kgCOD/m$^3$). The range is quite similar to previous studies [12], where it has already been stated that the hydrogen inhibition constant applied for propionate and valerate/butyrate uptake for grass silage is up to two orders of a magnitude lower than that for activated sludge digestion. The colour gradient reveals that both parameters are similarly sensitive. Because as few as possible parameters should be modified and $K_{S,H2}$ has already been changed, $k_{m,H2}$ should remain at the suggested value. Hence, the optimum can be found by multiplication with 1.0 (white line). The highest value and hence, the highest degree of agreement is achieved for about $1.0 \cdot K_{S,H2}$, which justifies the previously realized modification by a factor of 8. Nevertheless, several other combinations resulted in comparably high coefficients. Although a couple of maxima (not only in the depicted area) exist, a tuple of parameters needs to be chosen for running a simulation. In order to limit the amount of combinations, as few as possible parameters are calibrated. Finally, the suggested ADM1 parameter values are based on widespread research reflecting the complex biological processes, and any unnecessary change should be avoided.

![Figure 2. Sensitivity analysis for the concentration of hydrogen by varying $k_{m,H2}$ and $K_{S,H2}$.](image)

### 3.2.2. Acetate

For this simulation study all three parameters remained at their suggested ADM1 values. As can be seen in the top graphs of Figure 3, $K_{S,Ac}$ is – in contrast to $k_{m,Ac}$ and $K_{I,NH3}$ – quite a non-sensitive parameter. This corresponds to the previous considerations based on Equation 1 that the half-saturation coefficient will only be sensitive at low concentrations of the associated acid. However, acetate reached concentrations of 5,000 mg/L during this experiment. If $K_{S,Ac}$ remains unchanged and the two sensitive parameters are varied, the down graph is obtained. According to the convention that as the least possible parameters should be adapted, intersections with standard value (multiplication by 1.0) should be preferred. Furthermore, the change of a parameter should be as small as possible, too. As the Nash-Sutcliffe coefficient for $0.55 \cdot k_{m,Ac}$ is higher as for $0.4 \cdot K_{I,NH3}$ as well, a calibration of the maximum uptake rate $k_{m,Ac}$ on 55 % of its suggested value is advised. However, other combinations (1.8 $k_{m,Ac}$ and 0.2 $K_{I,NH3}$, for example) also lead to similar coefficients. A definitive calibration of parameters is once again not possible.
Figure 3. Sensitivity analysis for the concentration of acetate by varying $k_{m, Ac}$, $K_{S, Ac}$ and $K_{I, NH3}$.

**Slika 3.** Analiza osjetljivosti koncentracije acetata promjenom $k_{m, Ac}$, $K_{S, Ac}$ i $K_{I, NH3}$.

### 3.2.3. Propionate

In contrast to the simulation studies for acetate, the hydrogen inhibition constants for propionate and butyrate/valerate have already been adapted due to reasons of visualization. $K_{I, H2, Pro}$ has been reduced by a factor of 0.013 compared to its suggested value [13], but is set as the standard value (multiplication by 1.0) in the following graphs. Similar to Figure 3 and due to a high concentration of propionate (up to 8,000 mg/L) once more, the sensitivity of $K_{S, Pro}$ is marginal as depicted in Figure 4 (top graphs). As an adaption of $K_{I, H2, Pro}$ was already necessary, $k_{m, Pro}$ remains at its suggested value. Although the lower graph shows that the optimum should be rather in the field of $1.3\cdot K_{I, H2, Pro}$ (white line), this discrepancy is due to the fact that both reactors were considered for parameter calibration, but this simulation study refers only to reactor L.
3.2.4. Butyrate
Similar to the studies of propionate, the hydrogen inhibition constant for butyrate and valerate uptake $K_{I\text{H}_2\text{C}_4}$ has been reduced by a factor of 0.005 compared to its suggested value, but is set as the standard value. Again this change overlaps with the ranges proposed by previous research for the adaption of ADM1 to grass silage [12]. The concentration of butyrate in reactor L acted as a reference for the calculation of the ENSC. Once more, the maximum uptake rate $k_{m\text{C}_4}$ and the hydrogen inhibition constant $K_{I\text{H}_2\text{C}_4}$ are the most sensitive parameters. Nevertheless, due to the relatively low concentration of butyric acid (less than 2,000 mg/L), the influence of $K_{S\text{C}_4}$ increased as can be seen by the shift of $E_{\text{NSC}}$ relative to the $K_{S\text{C}_4}$-axis. In contrast to Figure 3, where a variation of $K_{S\text{Ac}}$ had nearly no influence on the goodness of fit, changing $K_{S\text{C}_4}$ by constant $1.2 \cdot K_{I\text{H}_2\text{C}_4}$ caused a change from nearly zero for low factors up to the maximum Nash-Sutcliffe coefficients for higher factors.

**Figure 4.** Sensitivity analysis for the concentration of propionate by varying $k_{m\text{Pro}}, K_{S\text{Pro}}$ and $K_{I\text{H}_2\text{Pro}}$.

**Slika 4.** Analiza osjetljivosti koncentracije propionata promjenom $k_{m\text{Pro}}, K_{S\text{Pro}}$ and $K_{I\text{H}_2\text{Pro}}$. 
3.2.5. Summary

Figures 2 to 5 clarify that the kinetic parameters have different impacts on the model results. They further indicate that the model results are not bijective. On the one hand, a pair of parameters led to a definitive $E_{NSC}$ value. But on the other hand, an arbitrary $E_{NSC}$ was reached by a couple of combinations of the parameters. The phenomenon that the same result can be achieved in different ways is called equifinality. This means that there is not simply one optimal parameter set that represents a system, but rather that several combinations of parameter values for a chosen model structure may fit the data equally well. On this account, several authors have already started to specify parameters not by single values, but by confidence regions [19, 20]. Thus, not the parameter value itself is important, but rather the combination. Knowledge on equifinality is insofar basically for modelling as calibration of parameters is usually not definite. It is hence possible that the model’s output is equal although parameter sets were different. Maybe some parameters compensate for each other due to the overdetermined system. Proportional to its complexity, parameter calibration in ADM1 requires a large database. In further research, adapted parameters must be proven by independent datasets.

4. CONCLUSIONS

The modified Nash-Sutcliffe model efficiency coefficient was applied as an alternative methodology for the calibration of kinetic parameters of the ADM1. Sensitivity analyses highlighted different influences on the model’s output by what the number of parameters for calibration could be reduced to the most sensitive ones. As has already been shown [13], only changes of hydrogen inhibition constants and maximum uptake of acetate compared to the ADM1 suggested values were necessary in order to fit the measurements.
The approach presented is universally applicable and hence transferable to any other model where a calibration of parameters is necessary and an independent determination is not possible. A quality assessment procedure based on the ENSC implemented in the model allows the software to automatically search for the parameter combination with the best fit. In this context, the phenomena of equifinality should be considered.

Applying the Nash-Sutcliffe coefficient enables parameter sensitivity analysis and parameter estimation as well. Thereby,
- extensive manual analysis of model output is avoided,
- percentile quantification of the degree of agreement is enabled,
- comparison of the different goodness of fit values (independent of their units) is possible,
- a new approach for parameter sensitivity analysis is presented, and
- finally, the reliability of model calibration is significantly improved.

Acknowledgements
Funding for this study was provided by German Agency for Renewable Resources (FNR) - FKZ: 22011905. Furthermore, the German Academic Exchange Service (DAAD) is acknowledged for funding the attendance at the 6th Dubrovnik Conference on Sustainable Development of Energy, Water and Environmental Systems (FKZ: 521/dmü).

REFERENCES


