Hybrid neural modelling of an anaerobic digester with respect to biological constraints

A. Karama*, O. Bernard**, J.L. Gouzé**, A. Benhammou* and D. Dochain***

*LAEP (SAC REP), Faculté des Sciences Sémilaia, 40000 Marrakech, Morocco
**COMORE, INRIA, BP93,06 902 Sophia-Antipolis cedex, France
***CESAME-UCL, Av. G. Lematre,4-6, 1348 Louvain-La-Neuve, Belgium

Abstract A hybrid model for an anaerobic digestion process is proposed. The fermentation is assumed to be performed in two steps, acidogenesis and methanogenesis, by two bacterial populations. The model is based on mass balance equations, and the bacterial growth rates are represented by neural networks. In order to guarantee the biological meaning of the hybrid model (positivity of the concentrations, boundedness, saturation or inhibition of the growth rates) outside the training data set, a method that imposes constraints in the neural network is proposed. The method is applied to experimental data from a fixed bed reactor.

Keywords Anaerobic digestion; hybrid model; mass balance modelling; neural network

Introduction
Anaerobic digestion is a biological process in which biodegradable organic materials are decomposed in the absence of oxygen to produce methane. Despite its capacity to degrade difficult substrates, this process is known to become unstable under certain circumstances, and therefore requires a monitoring procedure to predict a destabilisation. This must also be associated to a control action that can avoid the risk of acidification of the fermenter. Nevertheless, to guarantee the success of the supervision of the digester, a reliable model must be available. It turns out that it may be very difficult to design reliable models for biological processes. This is particularly true for anaerobic digestion because of the inherent complex nature of this process involving many different microbial populations and of the difficulties to measure the involved chemical or biological species. One of the major problems in the modelling task remains the difficulty to model the biological kinetics.

Recently, a neural hybrid approach for modelling was successfully introduced for biotechnological engineering (Psichogios and Ungar, 1992; Schubert et al., 1994a; de Azebedo et al., 1997). It consists in combining simultaneously the mass balance description (in order to satisfy the mass conservation constraints) and neural networks (to model the process kinetics).

Neural networks proved to be powerful in representing nonlinear processes. However, they may not be reliable when they are applied beyond their range of experiment. Indeed, there is in general no guarantee that the trajectories of the hybrid model will retain their biological meaning outside the training domain. Concentrations have to remain positive and mass balance have to be respected. Hence, the hybrid model must be constrained in order to ensure these important assumptions.

In our previous study (Karama et al., 2001), a neural hybrid model of an anaerobic digestion plant was proposed without taking these biological constraints into account. Our main interest in this paper is to analyse the constraints that one has to put on the hybrid model to satisfy such biological properties.

The paper is organised as follows: first the used mass balance model is presented. Then the constraints which must be added to the modelling are determined. In the subsequent
section, the choice of the neural networks to satisfy these constraints is explained. Finally, an application of the hybrid model to real experiments obtained for the treatment of vinasses (Bernard et al., 2000) is presented.

Dynamical model

Anaerobic digestion is a biological treatment process for organic wastes with production of methane gas. The underlying model assumes that two main bacterial populations are present (Bernard et al., 1999). The first one, the acidogenic bacteria $X_1$, consumes the organic substrate $S_1$ and produces through an acidogenesis step volatile fatty acids (VFA) $S_2$. The second population (methanogenic bacteria) $X_2$, uses the VFA in a methanogenesis step as substrate for growth and produces methane. The dynamical model of the anaerobic digestion process can be deduced from mass balance considerations and written as follow (Bernard et al., 1999):

\[
\begin{align*}
\frac{dX_1}{dt} &= \mu_1(X_1) - \alpha DX_1 \\
\frac{dS_1}{dt} &= -k_1 \mu_1(X_1) = D(S_{1in} - S_1) \\
\frac{dX_2}{dt} &= \mu_2(X_2) - \alpha DX_2 \\
\frac{dS_2}{dt} &= -k_2 \mu_2(X_2) + k_3 \mu_1(X_1) = D(S_{2in} - S_2) \\
Q_{CH_4} &= k_4 \mu_2(X_2)
\end{align*}
\]

where $D$ is the dilution rate and the terms $S_{1in}$ and $S_{2in}$ are the influent concentrations of $S_1$ and $S_2$ respectively. The $k_i$ represent the yield coefficients associated with bacterial growth, $\mu_1(\cdot)$ and $\mu_2(\cdot)$ are the specific growth rates of $X_1$ and $X_2$ (they depend on some variables that are not explicited here and this point will be discussed later on). $Q_{CH_4}$ denotes the methane flow rate. In the context of the model described in Bernard et al. (1999), $\alpha$ is a parameter reflecting if the biomass is affected or not by the dilution effect: $\alpha=0$ corresponds to an ideal fixed bed reactor, $\alpha=1$ corresponds to an ideal continuous stirred tank reactor (CSTR).

The most usual problem in solving equations (1) is the formulation of reasonable expressions for the corresponding specific growth rates: $\mu_1(\cdot)$ and $\mu_2(\cdot)$. As denoted by Bastin and Dochain (1990), a large number of models have been proposed to describe bioprocess kinetics. So the choice of a growth model for a particular bioreactor is not straightforward at all.

In the hybrid approach modelling, neural networks are used as a nonlinear black box model for the representation of the kinetic expression. The selection of network inputs should rely upon the available prior biological knowledge. In this paper the simplest necessary functional dependency is chosen and $\mu_1=\mu_1(S_1)$ and $\mu_2=\mu_2(S_2)$ is assumed.

Growth rate modelling

It is important to note that all the variables involved in biological models must satisfy some physical properties especially the positivity and the boundness of the concentrations and of the specific growth rates. The main idea of this section is to analyse how to take into account these physical constraints into the hybrid model. In other words, which conditions are needed to be imposed on the neural network to satisfy this biological reality?

Positivity of the variables

Let $\zeta=[X_1, X_2, S_1, S_2]'$ denote the vector of state variables. The hybrid model must be designed in such a way that the variables remain positive for all time if the initial conditions
are positive. Mathematically we have to guarantee that: \((d\xi(t_0))/(dt)\geq 0\) for any time \(t_0\) such that \(\xi(t_0)=0\). This condition specialises as follows for each variable:

- For \(X_i\), if \(X_i=0\), then \(dX_i/dt=0\). The positivity of the variables \(X_1\) and \(X_2\) is verified and is independent of the kinetics.
- For \(S_1\), if \(S_1=0\), then \((dS_1)/(dt)=-k_1 \mu_1(0)X_1+DS_{1,int}\) must be non-negative for any \(D\) and \(S_{1,int}\). This is possible if \((\mu_1(0)=0) (H1)\).
- For \(S_2\), if \(S_2=0\), \((dS_2)/(dt)=-k_2 \mu_2(0) X_2+k_3 \mu_1(0)X_1+DS_{2,int}\) must be non-negative for any \(X_1, \mu_1, D\) and \(S_{2,int}\). This could be achieved if \(\mu_2(0)=0 (H2)\).

In other words, we have two positivity conditions (H1) and (H2). These mean that the specific growth rate is zero whenever the concentration of the substrate is zero. This is in accordance with the biological reality: "when there is no substrate, the microorganisms do not grow".

Qualitative properties of the growth rates
The growth of the acidogenic bacteria is known to be enhanced by the organic matter \(S_1\). The methanogenic bacteria are known to be inhibited when an excess of volatile fatty acid is present in the medium. We want therefore to impose the following properties:

- \(\mu_1\) is a non-negative increasing function of \(S_1\) (H3).
- \(\mu_2\) is a non-negative function of \(S_2\) that tends toward zero for high values of \(S_2\) (H4).

Moreover, the bacterial growth is known to be limited by a maximal growth rate. We want therefore to impose that \(\mu_1\) and \(\mu_2\) are bounded (H5 and H6, respectively).

Choice of the neural networks
In order to guarantee the positivity of the variables and the qualitative properties of the growth rates \(\mu_1\) and \(\mu_2\), conditions (H1) to (H6) must be fulfilled.

Using the advantage that neural networks are not limited in the structure, two different models for \(\mu_1(S_1)\) and \(\mu_2(S_2)\) that satisfy these conditions and model two distinct biological phenomena are proposed.

Acidogenic bacteria
To fulfill the constraints (H1), (H3) and (H5) imposed on the growth rate \(\mu_1\) the following feedforward neural network (Hertz et al., 1991) is selected:

\[
\mu_1(S_1) = NN(S_1) = \sum_{i=1}^{n_h} \omega_i^1 \tanh(v_i^1 S_1) (2)
\]

where \(n_h\) is the number of neurons in the hidden layer. The \(\omega_i^1\)'s and \(v_i^1\)'s are respectively the weights of the output and the hidden layers. The hyperbolic tangent (tanh(.)) is an increasing nonlinear function with saturation.

The weights of the network are chosen to be positive so that conditions (H3) and (H5) are satisfied since \(\mu_1\) is the sum of saturated increasing functions. The condition (H1) is verified because \(\tanh(0)=0\).

Methanogenic bacteria
A neural model for \(\mu_2\) that accommodates for hypotheses (H2), (H4) and (H6) is decomposed as follows:

\[
\mu_2(S_2) = \nu(S_2) S_2 (3)
\]

where

\[
\nu(S_2) = NN(S_2) = \sum_{i=1}^{n_h} \omega_i^2 \sigma(v_i^2 S_2) (4)
\]
with $\sigma(x)=1/(1+\exp(-x))$ (sigmoidal function) (Hertz et al., 1991), the $\omega^2_i$'s are positive and the $\nu^2_i$'s are negative.

Equation (3) guarantees that the condition (H2) is verified. The positivity of $\sigma(.)$ and of the weights $\omega^2_i$'s implies the positivity of $\mu_2(S_2)$.

The function $\phi(x)=x\sigma(-x)=x/(1+\exp(x))$ is non-negative, upper bounded and tends toward zero when $x$ tends to infinity. Therefore, if we choose the $\nu^2_i$'s negative, (3) is a sum of parametrised functions $\phi(x)$. Therefore conditions (H4) and (H6) are fulfilled.

Application to experimental data

Presentation

Experiments performed by the LBE-INRA (Narbonne, France) in a 948 litres fixed bed reactor with vinasses are considered to illustrate the applicability of the proposed approach. A detailed presentation of this process and the experimental conditions can be found in Bernard et al. (2000). The yield parameters are summarized in Table 1 they can be identified without the knowledge of $\mu_i$ (Bernard et al., 1999).

Now the problem is to find an expression of the kinetics for $\mu_1(.)$ and $\mu_2(.)$ from the available data in order to use model (1). The difficulties are connected with the rather restricted information concerning the state variables. Since only $S_1(t)$, $S_2(t)$ and $Q_{CH_4}(t)$ are measured, we propose a method allowing to decouple the identification of the two specific growth rates $\mu_1$ and $\mu_2$.

The proposed method is organized as follows:

• In a first step $\mu_1$ is estimated by considering only the subsystem corresponding to the bacterial degradation of $S_1$. Using the structure of the neural net chosen in (2), an hybrid model is built. The training of the neural net is performed by comparing the estimates of $S_1$ given by the hybrid model with the measurements.

• The preceding algorithm can be used to fit the neural net (4), but it leads to an estimate of $\mu_2$ depending on the estimation of $\mu_1$. To decouple the two neural net identification procedures, we first design an estimator which will provide us with a set of couple ($S_2(t)$, $\mu_2(t)$). These estimations are then used to fit the neural net (4).

Training of $\mu_1$

We shall consider the following subsystem of (1) to train $\mu_1$:

$$\frac{dX_1}{dt} = \mu_1(S_1)X_1 - \alpha DX_1$$

$$\frac{dS_1}{dt} = -k_1\mu_1(S_1)X_1 = D(S_{in} - S_1)$$

(5)

As $\mu_1$ is not measured, the training of the neural network is based on a criterion

<table>
<thead>
<tr>
<th>Unit</th>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>/</td>
<td>Yield for COD degradation</td>
</tr>
<tr>
<td>$k_2$ (mmol/g)</td>
<td>Yield for VFA consumption</td>
<td>7.8</td>
</tr>
<tr>
<td>$k_3$ (mmol/g)</td>
<td>Yield for VFA production</td>
<td>611.2</td>
</tr>
<tr>
<td>$k_4$ (mmol/g)</td>
<td>Yield for $CH_4$ production</td>
<td>1139.2</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>/</td>
<td>Proportion of attached bacteria</td>
</tr>
</tbody>
</table>
performance of the overall hybrid model. So the network must be trained using the output variables of the process. As usual for wastewater processes, the biomass concentration $X_1$ is not measured, and only the substrate concentration $S_1$ is measured. The network weights are determined in such a way that we minimise the sum of squared errors:

$$J_1(p^1) = \frac{1}{2} \sum_{k=1}^{N} \left( S_{1k,\text{exp}} - S_{1k}(p^1) \right)^2$$

where $p^1 = [\omega^1, v^1]'$.

The sensitivity procedure (Schubert et al., 1994b; Karama et al., 2001) can be readily applied to minimise $J_1(p^1)$.

Training of $\mu_2$

We propose here to build an observer that will provide us with an estimate of $\mu_2(t)$ along the experiment (in fact it is an observer for $\nu(t)$ as defined in Eq. (4)). For this purpose, first an asymptotic observer of the methanogenic biomass $X_2$ is developed from the measurements of the methane flow rate $Q_{CH_4,\text{exp}}$:

$$\frac{d\tilde{X}_2}{dt} = \frac{Q_{CH_4,\text{exp}}}{k_4} - \alpha D \tilde{X}_2$$

The dynamics of the estimation error ($\epsilon = X_2 - \tilde{X}_2$) are governed by the following equation:

$$\dot{\epsilon} = -\alpha D \epsilon$$

proving that the error converges exponentially toward zero ($\alpha$ and $D$ are positive).

The estimates $\tilde{\nu}$ is computed then with the available measurements of the substrate $S_{2,\text{exp}}$ by the following expression:

$$\tilde{\nu} = \frac{Q_{CH_4,\text{exp}}}{k_4 \tilde{X}_2 S_{2,\text{exp}}}$$

Now this estimate of $\nu$ can be used to adjust the network weights by minimising the following criteria (with $p^2 = [w^2, v^2]'$):

$$J_2(p^2) = \frac{1}{2} \sum_{k=1}^{N} \left( \nu_k(p^2) - \tilde{\nu}_k \right)^2$$

For this purpose, the standard backpropagation algorithm can be used (Rumelhart and McClelland, 1986; Leonard and Kramer, 1990).

Derivation of the training algorithm

Both the backpropagation rule and the sensitivity approach are considered as a steepest descent solution to a nonlinear least square minimisation problem. The training rule can be obtained by defining the objective function to be minimised as a function according to the weights $J(p)$. It can be written as follows:

$$p_j(n + 1) = p_j(n) - \eta \frac{dJ}{dp_j}(p(n))$$

According to the previous analysis, the training rule is needed to be modified so that the network weights satisfy the constraints imposed on the sign of $p'$. To achieve this, some conditions must be imposed in order to guarantee that for $\text{sign}(p) = \epsilon$ ($\epsilon$ is a vector of +1 and -1).

Let us look for $\theta_j$ such that $p_i = \epsilon_i (\theta_j)^2$; thus an equivalent form of the constrained minimisation problem is to minimise the unconstrained objective function $\Psi(\theta) = J(\theta)$. By applying the chain rule, the gradient is given by:
The term \( \frac{d\Psi}{dp_i} \frac{dp_i}{d\theta_i} \) can be readily computed by using the standard backpropagation method or the sensitivity procedure developed in Rumelhart and McClelland (1986), Leonard and Kramer (1990) and Karama et al. (2001).

**Results**

The experimental data of 80 days of experiments are used to establish the hybrid model as the basis of a kinetic and state estimation software. The last 55 days’ measurements were used for the network training and the first 25 days for validating the hybrid model. The obtained networks for \( \mu_1(S_1) \) and \( \mu_2(S_2) \) consist of one input, a single hidden layer with four neurons and one output.

The derived training method modifies the backpropagation or the sensitivity procedure by projecting the adjusted weight into the feasible region satisfying the imposed conditions. This means that positivity of the process variables and qualitative criteria (increasing, saturation, inhibition) for the growth rates are guaranteed by the hybrid model.

The dynamics of substrates concentrations as well as methane flow rate are well estimated by the hybrid model compared to those respectively measured (Figure 1(a,b,c)). The adequacy between the model and measurements is good. The results obtained with the hybrid approach are better than the classical model in Bernard et al. (2000) during the last period of destabilisation (particularly for \( S_1 \)). During the other time periods the performances of both models are quite similar. The results are globally satisfactory and demonstrate not only the ability of the hybrid model to reconstruct the variables trajectories from initial conditions but also the possibility to guarantee a biological meaning even when the system is drifted into an area from where no data had been used during training.

**Figure 1** Comparison between the neural hybrid model (solid), the classical model in Bernard et al. (2000) (dashed) and the measured variables (dotted) for \( S_1, S_2 \) and \( Q_{CH4} \).
Figure 2 shows the estimated functions $\mu_1(S_1)$ and $\mu_2(S_2)$ respectively. These functions are very close to the classical Monod and Haldane models used respectively for acidogenic and methanogenic bacteria. It can be verified that the inhibition of methanogenesis by high VFA values is fully identified by the neural net. It also validates the simple structure retained for the neural network.

Conclusion

We demonstrated how to take into account some physical properties (positivity of concentrations and saturation of growth rates...) in the design of an hybrid model. This method guarantees to use the main advantages of the neural net approach (flexibility, set up rapidity, high number of toolboxes...) without the main drawback: the loss of the network control in regions far from the training data set. Therefore, the model will keep a biochemical meaning even if it has to run far from its calibration data set. These ideas have been applied to a neural hybrid modelling for an anaerobic digestion process. The hybrid model has been successfully validated with data from a real plant and it estimates in a reliable way the value of the process variables. It turns out to be easier to calibrate than the standard model based on Monod or Haldane kinetics. Since its behaviour will always remain acceptable, the resulting neural structure can provide basis for further development of control strategies.

Acknowledgements

This paper presents research results of the Belgium Programme on Inter University Poles of Attraction initiated by the Belgium State, Prime Minister’s Office for Sciences, Technology and Culture. The scientific responsibility rests with its authors.

References


