LFT modelling and identification of anaerobic digestion

Alessandro Della Bona [a], Gianni Ferretti [a,*], Elena Ficara [b], Francesca Malpei [b]

[a] Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Piazza Leonardo da Vinci 32, 20133 Milano, Italy
[b] Politecnico di Milano, Dipartimento di Ingegneria Civile e Ambientale, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

Received 5 November 2014
Accepted 25 November 2014

1. Introduction

The need to limit the environmental impact of human activities, particularly in agriculture, and the search for renewable energy sources, have recently favored the spread of biogas power plants, based on the anaerobic digestion process.

The use of heterogeneous substrates, such as agricultural wastes, and the complexity of the biochemical process, subjected to instability, make the control of anaerobic digestion plants a complex task, which needs an accurate process modelling. Moreover, modelling of anaerobic digestion is of fundamental importance not only to monitor and control the plant performance, but also to study the sensitivity of the plant behavior to operational parameters, and to assess to feasibility of the use of new substrates with varying characteristics, biodegradability and operational conditions.

The best known and the most sophisticated model, able to describe the anaerobic degradation of various substrates (even if designed considering activated sludge as substrate), is the Anaerobic Digestion Model No. 1 (ADM1) (Batstone et al., 2002), developed by the IWA Task Group for Mathematical Modelling and later modified by several authors (Blumensaat & Keller, 2005; Galí, Benabdallah, Astals, & Mata-Alvarez, 2009) to improve accuracy and robustness and to fit the model to other specific applications.

While being very detailed in the description of the anaerobic digestion process, it can be hardly used for design and control purposes. In fact, a large number of parameters (about a hundred), depending on the specific substrate, need to be estimated, which is particularly difficult in complex plant operations and also because of the scarce data available in the literature. Moreover, many parameters of the ADM1 model cannot be identified on the basis of field data typically included in conventional protocols applied in digesters operation and monitoring. Therefore, dedicated measurement campaigns are to be planned in order to get the required characterization of the chemical nature of the substrate to be degraded and its specific degradation pathways. Those campaigns are complex and time consuming, although valuable when seeking a deeper process understanding.

These facts have motivated the research of simpler models, focused for example on a few number of processes or specifically designed for particular substrates. Among them, the AMOCO model (Bernard, Hadj-Sadok, Dochain, Genovesi, & Steyer, 2001), reaches a good compro-mise between simplicity and accuracy. The AMOCO model has been developed mainly as a tool to monitor and control the anaerobic digestion process, rather than as a tool for accurate numerical simulation. Nonetheless, this model has been conceived to describe the degradation of soluble organic matter that is readily acidified under anaerobic conditions, and is not adequate to be applied to the degradation of particulate matter that needs an initial disintegration and hydrolysis prior to be acidified. Therefore, in order to widen its field of applicability, a modified AMOCO model has been proposed including a first order hydrolysis step, during which alkalinity is produced because of the release of ammonium from proteins hydrolysis (Allegrini, 2010).

While being of reduced order, the modified AMOCO model is however still nonlinear, therefore the identification of its parameters cannot be performed through classical methods, generally based on a linear time-invariant (LTI) model formulation (Ljung, 1999). In particular, applying a maximum likelihood approach to an output error nonlinear model, as done in this work, results in a

* Corresponding author.
E-mail address: gianni.ferretti@polimi.it (G. Ferretti).

nonlinear least-squares minimization, which can be solved through nonlinear programming.

A main problem with standard nonlinear programming (NLP) methods is the need of gradient and/or Hessian computations (Dennis & Schnabel, 1996), performed in general via computationally inefficient finite-difference approximations. On the other hand, if an equivalent linear fractional transformation (LFT) formulation of the original nonlinear model is considered, the said quantities can be very efficiently computed by simply simulating index-1, semi-explicit, differential-algebraic equations (DAE) systems. Apart from this major advantage, it must be also emphasized that the generality of the LFT approach allows for a treatment of important issues such as identifia-bility, persistence of excitation, convergence, robustness and expe-riment design under a single unified framework (Hsu, Vincent, Wolodkin, Rangan, & Poolla, 2008).

However, manually deriving a LFT model formulation is a non-trivial task in general, so symbolic computing tools are needed. In this work, one of these tools, partially developed in Donida, Romani, Casella, and Lovera (2009) with reference to linear models, has been fully implemented for application to the general non-linear case and, in particular, to the modified AMOCO model.

Two different test cases have been considered. In a first case, the data used for parameter identification have been generated by simulating the ADM1 model, considering waste activated sludge as substrate. A first set of (simulated) data has been considered for the identification of 6 parameters: 5 stoichiometric coefficients and the specific hydrolysis rate, while a second set of data has been used for model validation, obtained by simulating a nominal loading operation.

In a second test case, the experimental data collected on a laboratory scale equipment, performing the anaerobic digestion of ultra-filtered cheese-whey, have been considered for the identification of 4 stoichiometric coefficients.

The paper is organized as follows. Section 2 describes the modified AMOCO model. Section 3 recalls the concepts of LFT modelling and identification. Section 4 discusses the results obtained from the identification of the modified AMOCO model based on the simulation of the ADM1 model. Section 5 shows the results obtained from data collected from the anaerobic digestion of ultra-filtered cheese-whey in a laboratory setup. Finally, Section 6 draws some conclusion.

2. Modified AMOCO model

The original AMOCO model was just developed to support monitoring and control system design, rather than as a tool for numerical simulation of the process behavior, and was mainly focused on the description of the anaerobic digestion of soluble substrates or with a negligible particulate content. Only two bacterial populations were considered, in particular acidogenic and methanogenic, while the hydrolysis and acetogenic phases were no longer considered explicitly. In the first step, the acidogenic bacteria X_1 consume the organic substrate S_1 and produce CO_2 and volatile fatty acids (VFAs) S_2. The population of methanogenic bacteria X_2 uses, in the second step, the VFAs as substrate for growth and produces CO_2 and methane. The bacteria biomass is expressed in terms of volatile solids (VS) which is typically used in environmental engineering to quantify the organic matter content, therefore concentrations of X_1 and X_2 are expressed as gVS L^{-1}. As for the organic substrate S_1, COD (chemical oxygen demand) has been chosen as measuring unit since it reflects the chemical energy content of S_1 that is made available for the growth of X_1 and X_2. Therefore, S_1 concentration is expressed as gCOD L^{-1}. As for VFAs, CO_2 and alkalinity their concentrations are expressed in molar terms (mmol L^{-1}).

In this work, three modifications are introduced with respect to the original model.

At first, the hydrolysis step is taken into account, where the particulate organic matter fed to the digester (X_0) is solubilized into a soluble and degradable organic compound (S_1). The process kinetics is described as first order, having μ_0 as kinetic constant.

Then, a decay term in the growth rates of the biomasses, described by a first order kinetics having k_{d1} and k_{d2} as kinetic constants, has been considered, becoming increasingly important at high hydraulic retention times t_{HR}.

Finally, the contribution of inorganic nitrogen to alkalinity Z (mmol L^{-1}) is taken into account (Ficara et al., 2012). In fact, the AMOCO model considers alkalinity as non-reactive and, consequently, its dynamics is just described by the dilution effect of the reactor. On the contrary, the dynamics of alkalinity is determined by its constituents: bicarbonates, VFAs, hydroxide ions and, above all, free ammonia. The alkalinity state equation has been therefore modified with respect to the original AMOCO model to account for the inorganic nitrogen, introducing as parameters the nitrogen content of the substrate S_1 dependent on its protein content, N_{S1} (gN/gCOD), which would be released into the reactor liquor during the acidogenic process and the nitrogen content in the biomass N_{bac} (gN/gVS) respectively uptaken from or released into the reactor liquor during biomass growth or decay. While nitrogen release was included into the alkalinity mass balance, ammonium has not been added as an additional state variable, since its dynamics was not considered of special interest.

The modified AMOCO model is thus defined by 7 differential equations: 1 for the hydrolysis dynamics, 2 for the mass-balances of the bacterial populations X_1 and X_2, 2 for the organic substrate S_1 and the VFAs S_2 and, finally, 2 for alkalinity Z and inorganic carbon C (mmol L^{-1}):

$$\frac{dX_0}{dt} = \frac{1}{t_{HR}}(X_{0,in} - X_0) - \mu_0 X_0 \tag{1}$$

$$\frac{dS_1}{dt} = \frac{1}{t_{HR}}(S_{1,in} - S_1) + k_0\mu_0 X_0 - k_1\mu_{1,\max}\frac{S_1}{S_1 + K_{S1}}X_1 \tag{2}$$

$$\frac{dS_2}{dt} = \frac{1}{t_{HR}}(S_{2,in} - S_2) + k_2\mu_{1,\max}\frac{S_1}{S_1 + K_{S1}}X_1 - k_3\mu_{2,\max}\frac{S_2}{S_2 + K_{S2} + S_2^2/K_{I2}}X_2 \tag{3}$$

$$\frac{dX_1}{dt} = -\frac{1}{t_{HR}}X_1 + \mu_{1,\max}\left(\frac{S_1}{S_1 + K_{S1}} - k_{d1}\right)X_1 \tag{4}$$

$$\frac{dX_2}{dt} = -\frac{1}{t_{HR}}X_2 + \mu_{2,\max}\left(\frac{S_2}{S_2 + K_{S2} + S_2^2/K_{I2}} - k_{d2}\right)X_2 \tag{5}$$

$$\frac{dZ}{dt} = \frac{1}{t_{HR}}(Z_{in} - Z) + (k_1 N_{S1} - N_{bac})\mu_{1,\max}\frac{S_1}{S_1 + K_{S1}}X_1 - N_{bac}\mu_{2,\max}\frac{S_2}{S_2 + K_{S2} + S_2^2/K_{I2}}X_2 + k_{d1}N_{bac}\mu_{1,\max}X_1 + k_{d2}N_{bac}\mu_{2,\max}X_2 \tag{6}$$

$$\frac{dC}{dt} = \frac{1}{t_{HR}}(C_{in} - C) + k_4\mu_{1,\max}\frac{S_1}{S_1 + K_{S1}}X_1 + k_5\mu_{2,\max}\frac{S_2}{S_2 + K_{S2} + S_2^2/K_{I2}}X_2 - r_C \tag{7}$$

$$\varphi = C + S_2 - Z + K_H P_T + \frac{k_6}{k_{La}}\mu_{2,\max}\frac{S_2}{S_2 + K_{S2} + S_2^2/K_{I2}}X_2 \tag{8}$$

$$r_C = k_{La}\left[(C + S_2 - Z) - \frac{\varphi - \sqrt{\varphi^2 - 4K_H P_T(C + S_2 - Z)}}{2}\right] \tag{9}$$

$$r_{CH_4} = k_6 \mu_{2,\max} \frac{S_2}{S_2 + K_{S2} + S_2^2/K_{I2}} X_2 \tag{10}$$

where μ_0 (d^{-1}) is the specific hydrolysis rate, k_i ($i=0,\ldots,6$) are stoichiometric coefficients, $\mu_{i,\max}$ ($i=1,2$) (d^{-1}) are the maximum specific growth rates, K_{S1} (gCOD L^{-1}) and K_{S2} (mmol L^{-1}) are the half-saturation constants, K_{I2} (mmol L^{-1}) is the inhibition constant, K_H is the Henry's constant for CO_2 (mmol L^{-1} atm^{-1}), $P_T=1$ (atm) is the atmospheric pressure, k_{La} (d^{-1}) is the liquid–gas transfer coefficient, r_C (mmol L^{-1} d^{-1}) quantifies the amount of CO_2 that is transferred by the liquid to the gaseous phase by stripping and r_{CH_4} (mmol L^{-1} d^{-1}) is the methane production rate. The quantities $X_{0,in}$, $S_{1,in}$, $S_{2,in}$, Z_{in} and C_{in} are the concentrations of each component in the influent stream and represent the model inputs.

Note that the inorganic carbon is assumed to be mainly composed of dissolved carbon dioxide (mmol L^{-1}) and bicarbonate (mmol L^{-1}), neglecting the amount of carbonate in the normal operating conditions,[1] while VFAs are assumed as fully dissociated.

3. LFT modelling and parameter identification

In order to deal with parameter identification, in this work a linear fractional transformation (LFT) model formulation, a widely used formalism both in modern control (Hecker, Varga, & Magni, 2004; Zhou, Doyle, & Glover, 1996) and identification (Casella & Lovera, 2008; Hsu, Poolla, & Vincent, 2008; Hsu, Vincent et al., 2008; Lee & Poolla, 1999), is investigated.

The general form of a LFT model can be written as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}_1\mathbf{w} + \mathbf{B}_2\boldsymbol{\zeta} + \mathbf{B}_3\mathbf{u} \tag{11}$$

$$\mathbf{z} = \mathbf{C}_1\mathbf{x} + \mathbf{D}_{11}\mathbf{w} + \mathbf{D}_{12}\boldsymbol{\zeta} + \mathbf{D}_{13}\mathbf{u} \tag{12}$$

$$\boldsymbol{\omega} = \mathbf{C}_2\mathbf{x} + \mathbf{D}_{21}\mathbf{w} + \mathbf{D}_{22}\boldsymbol{\zeta} + \mathbf{D}_{23}\mathbf{u} \tag{13}$$

$$\mathbf{y} = \mathbf{C}_3\mathbf{x} + \mathbf{D}_{31}\mathbf{w} + \mathbf{D}_{32}\boldsymbol{\zeta} + \mathbf{D}_{33}\mathbf{u} \tag{14}$$

$$\mathbf{w} = \boldsymbol{\Delta}\mathbf{z} = \mathrm{diag}\{\delta_1\mathbf{I}_{r_1}, \ldots, \delta_q\mathbf{I}_{r_q}\}\mathbf{z} \tag{15}$$

$$\boldsymbol{\zeta} = \boldsymbol{\Theta}(\boldsymbol{\omega}) \tag{16}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^p$, $\mathbf{u} \in \mathbb{R}^m$ are the state, output and input vectors respectively, $\mathbf{z} \in \mathbb{R}^{n_z}$, $\boldsymbol{\omega} \in \mathbb{R}^{n_\omega}$, $\mathbf{w} \in \mathbb{R}^{n_w}$, $\boldsymbol{\zeta} \in \mathbb{R}^{n_\zeta}$ are vectors of auxiliary variables, δ_i ($i=1,\ldots,q$) are the unknown parameters, $\mathbf{A}$, $\mathbf{B}_i$, $\mathbf{C}_i$, $\mathbf{D}_{ij}$ are known constant matrices, r_i are the sizes of the corresponding identity matrices $\mathbf{I}_{r_i}$ in the $\boldsymbol{\Delta}$ block and $\boldsymbol{\Theta}(\boldsymbol{\omega}) : \mathbb{R}^{n_\omega} \to \mathbb{R}^{n_\zeta}$ is a known nonlinear vector function.

It must be pointed out that deriving such formulation is non-trivial if carried out manually. To this aim, a symbolic computing approach, partially developed in Donida et al. (2009) with reference to linear models, has been first fully implemented for application to the general non-linear case in this work.

It is also worth mentioning that, in order to deal with parameter estimation, it is essential to rewrite model (11)–(16) by introducing normalized unknown parameters $\bar{\delta}_i$, varying between ± 1 as parameter δ_i varies between a maximum $\delta_{i,\max}$ and a minimum $\delta_{i,\min}$.[2]

The problem of model identification formulated over LFT model structures has been a subject of active research for more than 10 years, see, e.g., Lee and Poolla (1999), Demourant and Ferreres (2002), and Hsu, Vincent et al. (2008). In particular, the parameter estimation method proposed in Hsu, Vincent et al. (2008) is extended here to account for the nonlinear functions $\Theta(\omega)$.

First, with a slight abuse of notation, denote with $\mathbf{y}(k)$ and $\mathbf{u}(k)$ the values of $\mathbf{y}(t)$ and $\mathbf{u}(t)$ at the sampling instant t_k then, the identification problem can be formulated as follows: given model (11)–(16) and given input and output measurements $\{\mathbf{u}(k), \mathbf{y}(k)\}$, $k=1,\ldots,N$, generated by the plant, find the values of the unknown parameters $\tilde{\boldsymbol{\delta}}$ minimizing the cost function:

$$J(\boldsymbol{\delta}) = \frac{1}{2}\langle \mathbf{e}(\boldsymbol{\delta}), \mathbf{e}(\boldsymbol{\delta})\rangle_N = \frac{1}{2N}\sum_{k=1}^{N} \mathbf{e}^T(k,\boldsymbol{\delta})\mathbf{e}(k,\boldsymbol{\delta}) \tag{17}$$

where $\mathbf{e}(k,\boldsymbol{\delta}) = \mathbf{y}(k) - \hat{\mathbf{y}}(k,\boldsymbol{\delta})$ is the prediction error between the measured output $\mathbf{y}(k)$ and the output $\hat{\mathbf{y}}(k,\boldsymbol{\delta})$, predicted by the model.

As it is well known, $\tilde{\boldsymbol{\delta}}$ is a *maximum-likelihood* estimate of the model parameters $\boldsymbol{\delta}$ for output-error plants (Ljung, 1999), and can be obtained through well known iterative optimization procedures such as, for example, the Gauss–Newton algorithm:

$$\hat{\boldsymbol{\delta}}(\nu+1) = \hat{\boldsymbol{\delta}}(\nu) - \alpha(\nu)\hat{\mathbf{H}}^{-1}(\hat{\boldsymbol{\delta}}(\nu))\mathbf{g}(\hat{\boldsymbol{\delta}}(\nu)) \tag{18}$$

where ν is the iteration number, $\alpha(\nu)$ is the step size, $\mathbf{g}(\delta) : \mathbb{R}^q \to \mathbb{R}^q$ and $\hat{\mathbf{H}}(\delta) : \mathbb{R}^q \to \mathbb{R}^{q\times q}$ are respectively the gradient vector and a positive semi-definite approximation of the Hessian of the cost function with respect to the unknown parameters:

$$\mathbf{g}(\boldsymbol{\delta}) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{E}^T(k,\boldsymbol{\delta})\mathbf{e}(k,\boldsymbol{\delta}), \tag{19}$$

$$\hat{\mathbf{H}}(\boldsymbol{\delta}) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{E}^T(k,\boldsymbol{\delta})\mathbf{E}(k,\boldsymbol{\delta}) \tag{20}$$

where $\mathbf{E}(k,\boldsymbol{\delta}) \in \mathbb{R}^{p\times q}$ is the Jacobian of $\mathbf{e}(k,\boldsymbol{\delta})$ and is given by

$$\mathbf{E}(k,\boldsymbol{\delta}) = \left[\frac{\partial \mathbf{e}(k,\boldsymbol{\delta})}{\partial \delta_1} \quad \frac{\partial \mathbf{e}(k,\boldsymbol{\delta})}{\partial \delta_2} \quad \cdots \quad \frac{\partial \mathbf{e}(k,\boldsymbol{\delta})}{\partial \delta_q}\right] \tag{21}$$

In this work, the optimization has been performed through the MATLAB function `fminunc`, receiving as input parameters the cost function $J(\boldsymbol{\delta})$, the gradient $\mathbf{g}(\boldsymbol{\delta})$ and the approximated Hessian $\hat{\mathbf{H}}(\boldsymbol{\delta})$. In turn, the said parameters have been computed without resorting to finite differences, by exploiting the LFT formulation of the model, as described in the following.

The computation of the predicted output $\hat{\mathbf{y}}(k,\boldsymbol{\delta})$ (first stage of the scheme in Fig. 1) can be dealt with by rewriting model (11)–(16) as follows:

$$\mathbf{M}\dot{\tilde{\mathbf{x}}} = \mathbf{f}(\tilde{\mathbf{x}}, \mathbf{u}) \tag{22}$$

$$\mathbf{y} = \mathbf{g}(\tilde{\mathbf{x}}, \mathbf{u}) \tag{23}$$

where $\tilde{\mathbf{x}} = \begin{bmatrix}\mathbf{x}^T & \mathbf{z}^T & \boldsymbol{\omega}^T\end{bmatrix}^T$ and

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times n_z} & \mathbf{0}_{n\times n_\omega} \\ \mathbf{0}_{n_z\times n} & \mathbf{0}_{n_z\times n_z} & \mathbf{0}_{n_z\times n_\omega} \\ \mathbf{0}_{n_\omega\times n} & \mathbf{0}_{n_\omega\times n_z} & \mathbf{0}_{n_\omega\times n_\omega} \end{bmatrix} \tag{24}$$

$$\mathbf{f}(\tilde{\mathbf{x}}, \mathbf{u}) = \begin{bmatrix} \mathbf{A}\mathbf{x} + \mathbf{B}_1\boldsymbol{\Delta}\mathbf{z} + \mathbf{B}_2\boldsymbol{\Theta}(\boldsymbol{\omega}) + \mathbf{B}_3\mathbf{u} \\ \mathbf{C}_1\mathbf{x} + (\mathbf{D}_{11}\boldsymbol{\Delta} - \mathbf{I}_{n_z})\mathbf{z} + \mathbf{D}_{12}\boldsymbol{\Theta}(\boldsymbol{\omega}) + \mathbf{D}_{13}\mathbf{u} \\ \mathbf{C}_2\mathbf{x} + \mathbf{D}_{21}\boldsymbol{\Delta}\mathbf{z} + \mathbf{D}_{22}\boldsymbol{\Theta}(\boldsymbol{\omega}) - \boldsymbol{\omega} + \mathbf{D}_{23}\mathbf{u} \end{bmatrix} \tag{25}$$

$$\mathbf{g}(\tilde{\mathbf{x}}, \mathbf{u}) = \mathbf{C}_3\mathbf{x} + \mathbf{D}_{31}\boldsymbol{\Delta}\mathbf{z} + \mathbf{D}_{32}\boldsymbol{\Theta}(\boldsymbol{\omega}) + \mathbf{D}_{33}\mathbf{u} \tag{26}$$

thus by sampling the output of a dynamic system defined by the algebraic transformation output (23) and by an index-1, semi-explicit DAE system defined by Eq. (22), fed by the sampled input $\mathbf{u}(k)$.

The numerical integration of the DAE system (22) has been dealt with in MATLAB through the `ode15s.m` function, which implements a variable order BDF method and allows to define separately the mass matrix $\mathbf{M}$ and the vector function $\mathbf{f}(\tilde{\mathbf{x}}, \mathbf{u})$.

[1] Values of *pH* range between 6 and 8, and temperature between 35 and 38 °C.

[2] $\delta_i = (\delta_{i,\max} + \delta_{i,\min})/2 + \bar{\delta}_i(\delta_{i,\max} - \delta_{i,\min})/2$.

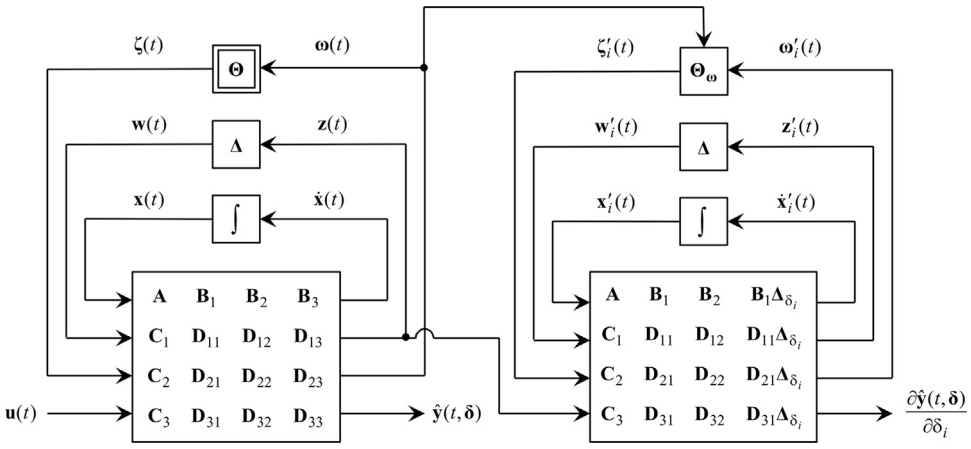

Fig. 1. Simulation scheme for the computation of the sensitivity functions.

Moreover, in order to improve reliability and efficiency, the Jacobian matrix $\partial\mathbf{f}/\partial\tilde{\mathbf{x}}$ has been analytically computed.

It must be also pointed out that since matrix $\mathbf{M}$ is singular, thus properly defining (22) as an index-1 DAE system, a consistent value for $\dot{\tilde{\mathbf{x}}}(0)$ must be supplied,[3] i.e.

$$\mathbf{M}\dot{\tilde{\mathbf{x}}}(0) = \mathbf{f}(\tilde{\mathbf{x}}(0), \mathbf{u}(0)) \tag{27}$$

In turn, the sensitivity $\partial\hat{\mathbf{y}}(k, \boldsymbol{\delta})/\partial\delta_i = -\partial\mathbf{e}(k, \boldsymbol{\delta})/\partial\delta_i$ of the predicted output $\hat{\mathbf{y}}(k, \boldsymbol{\delta})$ with respect to parameter δ_i can be computed (Appendix A) by sampling the output $\mathbf{y}'_i(t)$ of the following LFT system (second stage of the scheme in Fig. 1):

$$\dot{\mathbf{x}}'_i = \mathbf{A}\mathbf{x}'_i + \mathbf{B}_1\mathbf{w}'_i + \mathbf{B}_2\boldsymbol{\zeta}'_i + \mathbf{B}_1\boldsymbol{\Delta}_{\delta_i}\mathbf{z} \tag{28}$$

$$\mathbf{z}'_i = \mathbf{C}_1\mathbf{x}'_i + \mathbf{D}_{11}\mathbf{w}'_i + \mathbf{D}_{12}\boldsymbol{\zeta}'_i + \mathbf{D}_{11}\boldsymbol{\Delta}_{\delta_i}\mathbf{z} \tag{29}$$

$$\boldsymbol{\omega}'_i = \mathbf{C}_2\mathbf{x}'_i + \mathbf{D}_{21}\mathbf{w}'_i + \mathbf{D}_{22}\boldsymbol{\zeta}'_i + \mathbf{D}_{21}\boldsymbol{\Delta}_{\delta_i}\mathbf{z} \tag{30}$$

$$\mathbf{y}'_i = \mathbf{C}_3\mathbf{x}'_i + \mathbf{D}_{31}\mathbf{w}'_i + \mathbf{D}_{32}\boldsymbol{\zeta}'_i + \mathbf{D}_{31}\boldsymbol{\Delta}_{\delta_i}\mathbf{z} \tag{31}$$

where

$$\boldsymbol{\Delta}_{\delta_i} = \frac{\partial\boldsymbol{\Delta}}{\partial\delta_i} = \mathrm{diag}\{\mathbf{0}_{r_1\times r_1}, \ldots, \mathbf{I}_{r_i}, \ldots, \mathbf{0}_{r_q\times r_q}\} \tag{32}$$

$$\mathbf{w}'_i = \boldsymbol{\Delta}\mathbf{z}'_i \tag{33}$$

$$\boldsymbol{\zeta}'_i = \frac{\partial\boldsymbol{\Theta}(\boldsymbol{\omega})}{\partial\boldsymbol{\omega}}\boldsymbol{\omega}'_i = \boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\boldsymbol{\omega}'_i \tag{34}$$

By substituting (33) and (34) in (28)–(31) and solving (29) and (30) with respect to $\mathbf{z}'_i$ and $\boldsymbol{\omega}'_i$ the following time-variant, linear system is obtained:

$$\dot{\mathbf{x}}'_i = \tilde{\mathbf{A}}(\boldsymbol{\omega})\mathbf{x}'_i + \tilde{\mathbf{B}}(\boldsymbol{\omega})\boldsymbol{\Delta}_{\delta_i}\mathbf{z} \tag{35}$$

$$\mathbf{y}'_i = \tilde{\mathbf{C}}(\boldsymbol{\omega})\mathbf{x}'_i + \tilde{\mathbf{D}}(\boldsymbol{\omega})\boldsymbol{\Delta}_{\delta_i}\mathbf{z} \tag{36}$$

where

$$\tilde{\mathbf{A}}(\boldsymbol{\omega}) = \mathbf{A} + \begin{bmatrix}\mathbf{B}_1\boldsymbol{\Delta} & \mathbf{B}_2\boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\end{bmatrix}\mathbf{W}(\boldsymbol{\omega})\begin{bmatrix}\mathbf{C}_1\\ \mathbf{C}_2\end{bmatrix} \tag{37}$$

$$\tilde{\mathbf{B}}(\boldsymbol{\omega}) = \mathbf{B}_1 + \begin{bmatrix}\mathbf{B}_1\boldsymbol{\Delta} & \mathbf{B}_2\boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\end{bmatrix}\mathbf{W}(\boldsymbol{\omega})\begin{bmatrix}\mathbf{D}_{11}\\ \mathbf{D}_{21}\end{bmatrix} \tag{38}$$

[3] Anyway, if $\dot{\tilde{\mathbf{x}}}(0)$ and $\tilde{\mathbf{x}}(0)$ are not consistent, the solver treats them as guesses, tries to compute consistent values close to the guesses, and then goes on with the integration.

$$\tilde{\mathbf{C}}(\boldsymbol{\omega}) = \mathbf{C}_3 + \begin{bmatrix}\mathbf{D}_{31}\boldsymbol{\Delta} & \mathbf{D}_{32}\boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\end{bmatrix}\mathbf{W}(\boldsymbol{\omega})\begin{bmatrix}\mathbf{C}_1\\ \mathbf{C}_2\end{bmatrix} \tag{39}$$

$$\tilde{\mathbf{D}}(\boldsymbol{\omega}) = \mathbf{D}_{31} + \begin{bmatrix}\mathbf{D}_{31}\boldsymbol{\Delta} & \mathbf{D}_{32}\boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\end{bmatrix}\mathbf{W}(\boldsymbol{\omega})\begin{bmatrix}\mathbf{D}_{11}\\ \mathbf{D}_{21}\end{bmatrix} \tag{40}$$

$$\mathbf{W}(\boldsymbol{\omega}) = \begin{bmatrix}\mathbf{I}_{n_z} - \mathbf{D}_{11}\boldsymbol{\Delta} & -\mathbf{D}_{12}\boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\\ -\mathbf{D}_{21}\boldsymbol{\Delta} & \mathbf{I}_{n_\omega} - \mathbf{D}_{22}\boldsymbol{\Theta}_{\omega}(\boldsymbol{\omega})\end{bmatrix}^{-1} \tag{41}$$

In order to account for the fact that the time instants t_e selected by the solver when integrating (22) are different in general from the time instants t_{s_i} selected by the solver when integrating (35), while also improving numerical efficiency, system (35) is rewritten as

$$\dot{\mathbf{x}}'_i = \boldsymbol{\Gamma}(\boldsymbol{\omega})\begin{bmatrix}\mathbf{x}'_i\\ \boldsymbol{\Delta}_{\delta_i}\mathbf{z}\end{bmatrix} \tag{42}$$

with

$$\boldsymbol{\Gamma}(\boldsymbol{\omega}) = \begin{bmatrix}\tilde{\mathbf{A}}(\boldsymbol{\omega}) & \tilde{\mathbf{B}}(\boldsymbol{\omega})\end{bmatrix} \tag{43}$$

and matrix $\boldsymbol{\Gamma}(\boldsymbol{\omega})$ is stored as a function of the time instants t_e during the computation of the prediction error. Then, when integrating (42), the matrix $\boldsymbol{\Gamma}(\boldsymbol{\omega}(t_{s_i}))$ is computed from the stored values $\boldsymbol{\Gamma}(\boldsymbol{\omega}(t_e))$ by interpolation, while also exploiting its sparsity.

An important issue is the notion of identifiability itself, particularly in the case of nonlinearly parameterized models (Dötsch & Van den Hof, 1996; Van Doren, Van den Hof, Hansen, & Bosgra, 2008). Local identifiability in $\hat{\boldsymbol{\delta}}$, minimizing (17), can be verified by testing if the Hessian $\hat{\mathbf{H}}$ is positive definite in $\hat{\boldsymbol{\delta}}$, i.e. if its rank is equal to q. If the Hessian $\hat{\mathbf{H}}$ in $\hat{\boldsymbol{\delta}}$ is not full rank then the estimated parameters are not unique and the system is not identifiable. In this case, to select the identifiable parameter space, we can consider the singular value decomposition (SVD) of the Hessian:

$$\hat{\mathbf{H}} = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_1^T\\ \mathbf{V}_2^T\end{bmatrix} \tag{44}$$

where the separation between $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ is chosen in such way that the singular values in $\boldsymbol{\Sigma}_2$ are considerably smaller than those in $\boldsymbol{\Sigma}_1$. Accordingly, the column space of $\mathbf{U}_1$ represents the subspace of the original parameters δ_i that will be identifiable from the measurements.

4. Parameter identification based on ADM1 model simulation data

In this work, the data used for parameter identification have been first generated by simulating the ADM1 model, which has been therefore considered as the "plant". Specifically, the ADM1 model version proposed in Blumensaat and Keller (2005) and Rosen, Vrecko, Gernaey, Pons, and Jeppsson (2006) was considered, applied to the anaerobic digestion of waste sludge in a single stage CSTR (completely stirred tank reactor) with constant liquid volume and temperature and with no biomass retention. The model consists of a DAE system of 35 differential and 1 algebraic equation: 29 state variables are given by the concentrations in the liquid outflow and in the gas outflow, the other 6 variables are given by the concentrations of ionized volatile fatty acids, bicarbonate and free ammonia. The differential equations are given by the mass balances of the dynamic (state) variables, and involve 19 biochemical processes, 3 gas–liquid transfer processes and 6 additional acid–base dissociation processes. According to the authors, the input values may not be completely realistic for all variables but they have been chosen so that every input is active (i.e. non-zero) and able to excite all internal modes of the ADM1 model.

Since the modified AMOCO model (1)–(10) has a cascade structure, where the first block is made up by the first 5 equations and its outputs, namely X_0, S_1, S_2, X_1, X_2, appear as inputs for the remaining equations, the identification can be focused on the parameters governing the dynamics of substrates and biomasses only.

Then, the following reduced model can be considered, obtained by simply rewriting Eqs. (1)–(5) and (10) accordingly to the parameters chosen for the identification and to the measurements assumed available ($d = 1/t_{HR}$):

$$\dot{x}_1 = -dx_1 - \delta_6 x_1 + du_1 \tag{45}$$

$$\dot{x}_2 = -dx_2 + \delta_1\delta_6 x_1 - \delta_2 \mu_{1,\max} \frac{x_2}{x_2 + K_{S1}} x_4 + du_2 \tag{46}$$

$$\dot{x}_3 = -dx_3 + du_3 + \delta_3 \mu_{1,\max} \frac{x_2}{x_2 + K_{S1}} x_4 - \delta_4 \mu_{2,\max} \frac{x_3}{x_3 + K_{S2} + x_3^2/K_{I2}} x_5 \tag{47}$$

$$\dot{x}_4 = -dx_4 + \mu_{1,\max} \left(\frac{x_2}{x_2 + K_{S1}} - k_{d1} \right) x_4 \tag{48}$$

$$\dot{x}_5 = -dx_5 + \mu_{2,\max} \left(\frac{x_3}{x_3 + K_{S2} + x_3^2/K_{I2}} - k_{d2} \right) x_5 \tag{49}$$

$$y_1 = x_3 \tag{50}$$

$$y_2 = x_1 + x_2 + c(x_4 + x_5) \tag{51}$$

$$y_3 = \delta_5 \mu_{2,\max} \frac{x_3}{x_3 + K_{S2} + x_3^2/K_{I2}} x_5 \tag{52}$$

where

$$x = [x_1\ x_2\ x_3\ x_4\ x_5]^T = [X_0\ S_1\ S_2\ X_1\ X_2]^T \tag{53}$$

$$u = [u_1\ u_2\ u_3]^T = [X_{0,in}\ S_{1,in}\ S_{2,in}]^T \tag{54}$$

$$\delta = [\delta_1\ \delta_2\ \delta_3\ \delta_4\ \delta_5\ \delta_6]^T = [k_0\ k_1\ k_2\ k_3\ k_6\ \mu_0]^T \tag{55}$$

Of course, it is assumed that all inputs and outputs are measurable, while the parameters reported in Table 1 are assumed fixed and known. In particular, the measurable outputs are respectively given by the volatile fatty acids concentration $y_1 = x_3 = S_2$ (mmol L^{-1}), the sum of the others substances concentrations $y_2 = x_1 + x_2 + c(x_4 + x_5) \triangleq S_3$ (gCOD L^{-1}), where c is a scaling factor from (gVS L^{-1}) to (gCOD L^{-1}), and the methane flow rate $y_3 = r_{CH_4}$ (mmol L^{-1} d^{-1}). Accordingly, the relevant LFT model is defined in Appendix B.

Table 1
Fixed parameters for identification based on ADM1 model simulations.

d (d^{-1})	$\mu_{1,\max}$ (d^{-1})	$\mu_{2,\max}$ (d^{-1})
0.05	0.206	0.2
K_{S1} (kgCOD m^{-3})	K_{S2} (mmol L^{-1})	K_{I2} (mmol L^{-1})
1.096	6.86	433,968
k_{d1}	k_{d2}	c (gCOD $(gVS)^{-1}$)
0.1	0.1	1.55

A first main issue in obtaining synthetic data from the ADM1 simulation is the need of lumping several variables of the ADM1 model into single variables of the modified AMOCO model (see also Della Bona, Ferretti, Ficara, & Malpei, 2015).

First of all, having extended the model to account for the hydrolysis step, the particulate substrate, described by composite, carbohydrates, proteins and lipids concentrations, has been aggregated in the variable X_0.

The soluble organic matter (sugar, amino acids and fatty acids concentrations) corresponding to the ADM1 variables, S_{su}, S_{aa}, S_{fa}, has been aggregated in the variable S_1.

The variable S_2 has been considered accounting for the total concentration of VFAs, given by the soluble compounds valeric, butyric, propionic and acetic acids.

The ADM1 bacterial populations in charge of the degradation of sugars, amino acids, fatty acids and volatile acids (with exception of acetic acid) can be grouped into the variable X_1.

Finally, the concentration of biomasses converting acetic acid and hydrogen into methane can be grouped into the variable X_2.

A sequence of steps of type "3-2-1-1"[4] in the input variable u_1 has been considered as test signal, shown in Fig. 2. An overall simulation interval of 40 days has been considered while, for the sake of realism, a white noise disturbance, with a signal-to-noise ratio of 30, has been added to the simulated outputs, sampled with a period of 1 day.

Figs. 3, 4 and 5 compare the simulated measurements y_1, y_2 and y_3 (denoted by circles), obtained by a simulation of the ADM1 model, with the simulated outputs of the identified modified AMOCO model $\hat{y}_1 = \hat{S}_2$, $\hat{y}_2 = \hat{S}_3$ and $\hat{y}_3 = \hat{r}_{CH_4}$ (solid line) respectively. As it is apparent, a good correspondence has been obtained.

In terms of estimated parameters, the results of the identification are reported in Table 2, together with the initial parameter values. As for parameters k_1–k_6 they represent the stoichiometry of the processes of acidogenesis and methanogenesis. For this reason, a first guess for their values can be obtained by referring to average stoichiometric values suggested in the ADM1 model for the acidogenic processes and for the methanogenic process, and by taking into account the appropriate change in measuring units. As reported in Table 2, final estimates do not differ significantly from initial values, suggesting that, although the modified AMOCO model is lumping several substrates and bacteria populations in few variables, the general meaning and order of magnitude of the stoichiometric coefficients is, in this case, preserved. This cannot however be considered as a general conclusion, since stoichiometry is strictly

[4] The numbers denoting the test signal indicate the duration of the "high" and "low" signal intervals. Such a shape of signal is frequently used in the aeronautical field (Klein & Morelli, 2006), where the period "2" correspond to the semiperiod of the expected mode to be identified, while in our case is equal to half t_{HR}, thus 10 days.

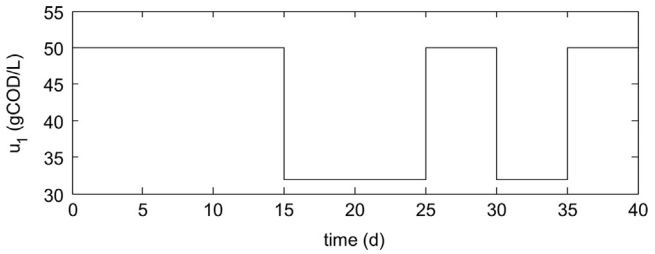

Fig. 2. Input u_1 for identification.

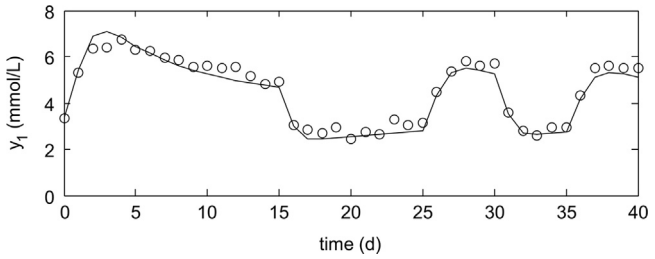

Fig. 3. ADM1 output y_1 (circles) and reduced model output $\hat{y}_1 = \hat{S}_2$ (solid line).

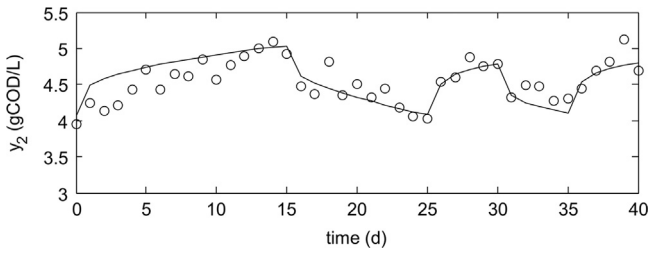

Fig. 4. ADM1 output y_2 (circles) and reduced model output $\hat{y}_2 = \hat{S}_3$ (solid line).

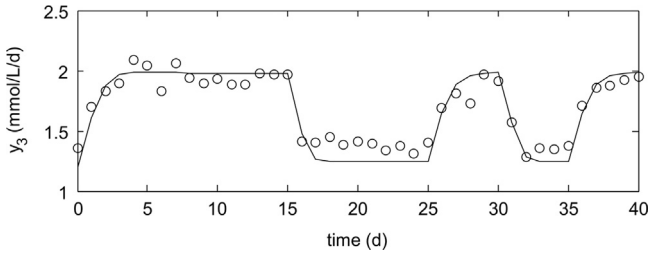

Fig. 5. ADM1 output y_3 (circles) and reduced model output $\hat{y}_3 = \hat{r}_{CH_4}$ (solid line).

Table 2
AMOCO initial and estimated parameters from identification based on ADM1 model simulations.

Parameter	Initial value	Optimal value
k_0	1	0.7387
k_1 (gCOD gVS^{-1})	15	14.8410
k_2 (mmol gVS^{-1})	200	200.1184
k_3 (mmol gVS^{-1})	440	321.0709
k_6 (mmol gVS^{-1})	27	26.1949
μ_0 (d^{-1})	0.5	5.8874

dependent with the substrate composition and complexity and is therefore case specific.

Note that, in order to test the performance of the identification algorithm, a much lower initial value for the parameter μ_0, with respect to the actual behavior of the ADM1 model, has been considered. This test can be considered as representative of the case where a more easily hydrolyzable load is introduced in the digester, marked just by a rapid decrease of the parameter μ_0.

A different variation of the input variable u_1 has been considered for the sake of validation of the identified model, obtained by increasing the said variable at a rate of 50%/day along 6 days (Fig. 6). The obtained transients of the aggregate ADM1 model variables y_1, y_2 and y_3 (circles), and of the simulated outputs of the identified modified AMOCO model $\hat{y}_1$, $\hat{y}_2$ and $\hat{y}_3$ (solid lines) are shown in Figs. 7, 8 and 9 respectively, where a quite good matching between the ADM1 model and the identified modified AMOCO model can still be appreciated.

It must be finally pointed out that the choice of the number of identified parameters has been actually validated through a local identifiability analysis, based on the analysis of the Hessian condition number. Fig. 10 shows the trend of the said condition number with respect to the identification steps, considering the assumed uncertain parameters (solid line), and the same trend obtained by identifying also the parameters $\mu_{1,max}$ and $\mu_{2,max}$ (dashed line). As it is apparent, with a greater number of parameters the higher value

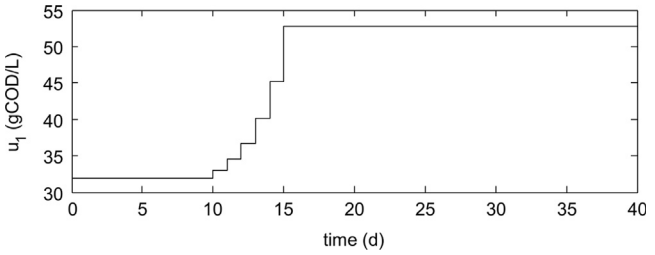

Fig. 6. Input u_1 for validation.

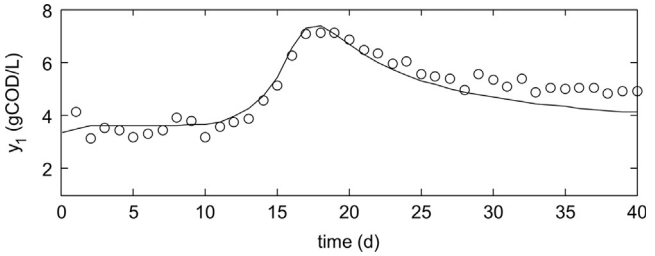

Fig. 7. ADM1 output y_1 (circles) and reduced model output $\hat{y}_1 = \hat{S}_2$ (solid line).

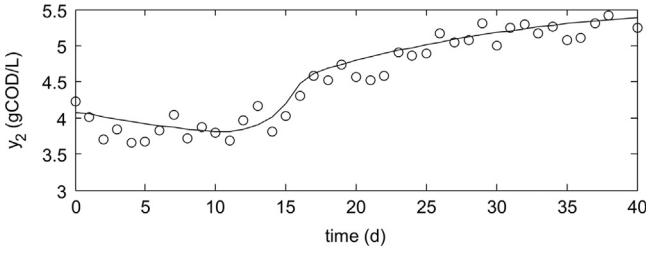

Fig. 8. ADM1 output y_2 (circles) and reduced model output $\hat{y}_2 = \hat{S}_3$ (solid line).

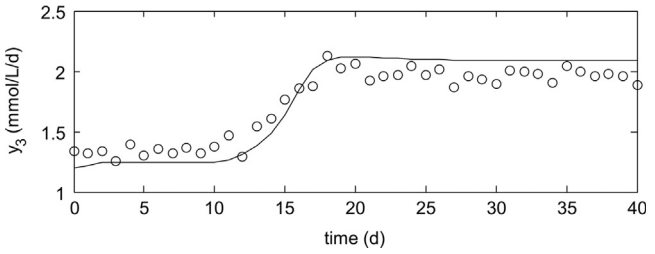

Fig. 9. ADM1 output y_3 (circles) and reduced model output $\hat{y}_3 = \hat{r}_{CH_4}$ (solid line).

of the condition number values indicates a worst identifiability condition.

5. Parameter identification based on experimental data

A semi-batch experiment was performed on the anaerobic digestion of ultra-filtered cheese-whey (UF-CW). Cheese whey is the main residue of cheese production, representing up to the 90% of the volume of processed milk. Cheese whey is mainly made of lactose, milk proteins and salts. Valuable milk proteins can be recovered by a well-established membrane separation process (ultra-filtration), producing a residual lactose stream that can be fed to anaerobic digestion.

The experiment was performed by using a laboratory scale equipment made of two glass bottles of 1140 mL each one, connected as sketched in Fig. 11. Each bottle is endowed with three openings: one in the central bottle-neck and two lateral ones. Bottle A is filled with the anaerobic suspension and it is mixed by means of a magnetic mixer. The central opening is provided with a sampling port for sampling/feeding, and one of the lateral openings is connected to a gas line that connects the head space of bottle A to that of bottle B, through one of the lateral opening of bottle B. The second lateral opening of bottle A is sealed by a gas-tight rubber septum. Bottle B is filled with an alkaline solution for CO_2 absorption and a pressure transducer with a data logger (OxiTop Control System, WTW) is accommodated in the central opening. The second lateral opening allows for the manual gas discharge, once the maximum overpressure is achieved. The gas is released in a water bath to prevent oxygen penetration. The whole equipment is kept in a thermostated chamber at a temperature of 35 ± 0.5 °C. Manually operated valves are available on the sampling port and on the gas lines.

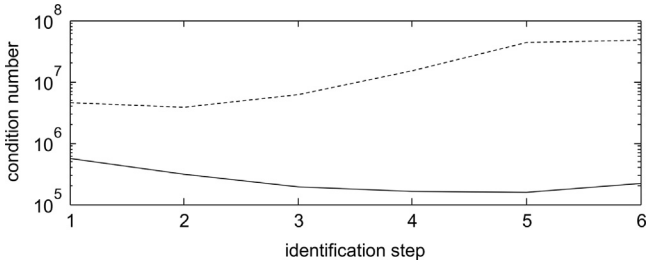

Fig. 10. Condition number with 6 (solid line) and 8 (dashed line) parameters.

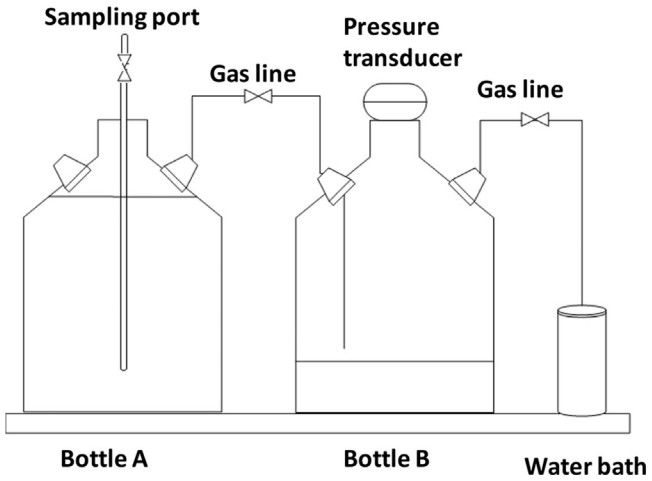

Fig. 11. Experimental setup scheme.

Two identical setups (later referred to as setup 1 and setup 2) were operated in parallel in a semi-continuous mode, according to a spiking strategy described in Fig. 12. Both were started by filling bottle A with a digested sludge from the anaerobic digester of a municipal wastewater treatment plant. The digested sludge was diluted 1:1 with tap water to a total solid content of 10 gVS/L. Each bottle B was filled with 200 mL of a 3 M NaOH solution.

The feeding solution was freshly prepared before each spike and was made of lactose (44.6 g/L), nitrogen (0.3 gN-NH_4Cl/L), and NaOH (0.065 M). Nitrogen and alkalinity were added in order to avoid nutrients or pH limitation. Before each spike, a volume of 1/20 of the anaerobic suspension was withdrawn from the reactor through the sampling port of bottle A and substituted by an equivalent volume of the feeding solution.

Then, setup 1 was left unmodified until the next spike, and the measure of the overpressure due to methane production was collected by the pressure transducer. By knowing the overall headspace of bottles A and B and the operational temperature, the cumulated volume of methane produced was assessed from pressure data, by applying the ideal gas law equation. Finally, the methane production rate was assessed by a linear regression on the data of cumulated methane production versus time.

Setup 2 was used to assess the time trend of the relevant suspension parameters. To this purpose, samples of 25–45 mL each one were taken at defined intervals from the central opening of bottle A. Overpressure data were not used for the computation of the methane production because of the interferences on pressure data generated by each sampling event. The sampling procedure in setup B caused a decrease of the anaerobic suspension volume from the initial 1000 mL to 375 mL at the end of the experiment. Therefore, the volume of the feeding solution was reduced during the course of the experiment in order to maintain the same volumetric loading rate applied to setup A.

The following parameters were measured on samples collected from setup 2:

- total suspended solids (TSS) and volatile suspended solids (VSS) assessed according to APHA, AWWA, WEF (2005);
- chemical oxygen demand (COD) on 0.45 μm filtered samples (APHA, AWWA, WEF, 2005);
- volatile fatty acids by using the H-lange spectrophotometric kit (LCK 365).

Since in this case no hydrolysis takes place, the model is characterized by 4 state equations, moreover, the flow rate q_{in} ($m^3\,d^{-1}$) is now considered as an input ($q_{in} = u_1$) therefore, the dilution rate d in Eqs. (45)–(49) has been replaced by $d = q_{in}/V$, where V=0.001 m^3 is the reactor volume, while the organic substrate $y_1 = x_2 = S_1$ (gCOD L^{-1}), the volatile fatty acids concentration $y_2 = x_3 = S_2$ (mmol L^{-1}), and the methane flow rate $y_3 = r_{CH_4}$ (mmol $L^{-1}\,d^{-1}$) have been measured as outputs. Accordingly, Eqs. (2)–(5) and (10) have been rewritten as follows:

$$\dot{x}_1 = \frac{u_1}{V}(u_2 - x_1) - k_1 \frac{x_1}{x_1 + \delta_1}\delta_3 x_3 \tag{56}$$

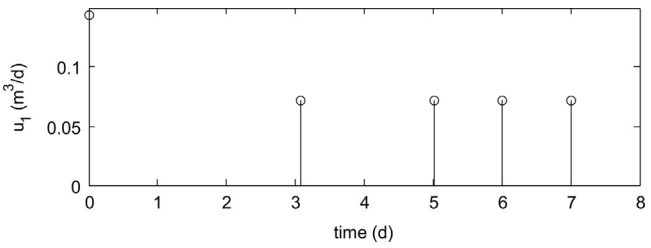

Fig. 12. Input flow rate variation.

$$\dot{x}_2 = \frac{u_1}{V}(u_3 - x_2) + k_2 \frac{x_1}{x_1 + \delta_1}\delta_3 x_3 - k_3 \frac{x_2}{x_2 + \delta_2 + x_2^2/K_{I2}}\delta_4 x_4 \quad (57)$$

$$\dot{x}_3 = -\frac{u_1}{V}x_3 + \delta_3\left(\frac{x_1}{x_1 + \delta_1} - k_{d1}\right)x_3 \quad (58)$$

$$\dot{x}_4 = -\frac{u_1}{V}x_4 + \delta_4\left(\frac{x_2}{x_2 + \delta_2 + x_2^2/K_{I2}} - k_{d2}\right)x_4 \quad (59)$$

$$y_1 = x_1 \quad (60)$$

$$y_2 = x_2 \quad (61)$$

$$y_3 = \delta_5\delta_4 \frac{x_2}{x_2 + \delta_2 + x_2^2/K_{I2}}x_4 \quad (62)$$

where:

$$x = [x_1\ x_2\ x_3\ x_4]^T = [S_1\ S_2\ X_1\ X_2]^T \quad (63)$$

$$u = [u_1\ u_2\ u_3]^T = [q_{in}\ S_{1,in}\ S_{2,in}]^T \quad (64)$$

$$\delta = [\delta_1\ \delta_2\ \delta_3\ \delta_4\ \delta_5]^T = [K_{S1}\ K_{S2}\ \mu_{1,\max}\ \mu_{2,\max}\ k_6]^T \quad (65)$$

the fixed parameters are reported in Table 3, the relevant LFT model is defined in Appendix C, while the initial and estimated values of the parameters δ_i are reported in Table 4.

It is worth noting that all kinetic parameters but the K_{I2} were selected for identification. This latter parameter is included in the Haldane model, that is in turn used to describe the dependence of the X_2 growth from its substrate S_2. Specifically, K_{I2} quantifies the negative effect that an excess of substrate S_2 may have on biomass growth, so that any increase in substrate concentration above the threshold value of $\sqrt{K_{I2}K_{S2}}$ would cause a decrease in the biomass growth rate, typically leading to process instability and, on the long-term, to biomass washout . Therefore, the effect of K_{I2} becomes relevant at high S_2 concentrations (of the order of magnitude of $\sqrt{K_{I2}K_{S2}}$), making it identifiable only under process conditions at which inhibition becomes relevant. Since this was not the case for the available experimental data, the K_{I2} was not included in the list of the parameters to be identified. As a matter of fact, the experimentation was planned so that usual operational conditions were represented, which means that the anaerobic reactor was maintained around a stable equilibrium point.

Fig. 12 shows the variation of the input u_1, while Figs. 13, 14 and 15 compare the measurements of y_1, y_2 and y_3 (denoted by circles) with the simulated outputs of the identified modified AMOCO model $\hat{y}_1 = \hat{S}_1$, $\hat{y}_2 = \hat{S}_2$ and $\hat{y}_3 = \hat{r}_{CH_4}$ (solid line) and with the simulated outputs obtained with the initial values of the parameters (dashed line) respectively.

Table 3
Fixed parameters for identification based on experimental data.

k_1 (gCOD gVS^{-1})	k_2 (mmol gVS^{-1})	k_3 (mmol gVS^{-1})
15	200	400
K_{I2} (mmol L^{-1})	k_{d1}	k_{d2}
120	0.1	0.1

Table 4
Initial and estimated parameters from identification based on experimental data.

Parameter	Initial value	Optimal value
K_{S1} (kgCOD m^{-3})	1.5	1.6533
K_{S2} (mmol L^{-1})	7.5	15.6259
$\mu_{1,\max}$ (d^{-1})	0.2	0.57853
$\mu_{2,\max}$ (d^{-1})	0.2	0.19518
k_6 (mmol gVS^{-1})	27	437.6786

As for the organic substrate (Fig. 13), a quite satisfactory correspondence can be observed, while the experimental trends of VFAs (Fig. 14) and methane production rate (Fig. 15) deserve more detailed comments.

The data suggest that the biological response changed slightly from spike to spike, although the same experimental conditions were applied. Specifically, a faster and more substantial response was observed in the last 3 spikes. This biochemical effect is likely due to an *acclimation* phenomenon, that is typical of anaerobic sludge samples in response to a relevant change in the nature of the substrate to be degraded. Actually, this is what happened in this anaerobic digestion test, where an anaerobic sludge inoculum, taken from a full scale anaerobic digester fed on waste activated sludge, was used, i.e. a complex mixture of organic substrates chemically very different from deprotenized cheese whey. Acclimation is a transient phenomenon that is neither described in the ADM1 model nor in the modified AMOCO model. For this reason, the model is not fully capable to interpret all data-points but it allows anyway to give a satisfactory description of the general trend.

It is also worth noting that a residual VFA concentration of about 1 mmol/L is found at the end of each spike, which could be considered as an offset not accounted for by the model. Nonetheless, the dynamic responses of VFAs and methane production rate are fairly described, despite the strong simplifying assumptions underlying the use of a simplified model such as the modified AMOCO.

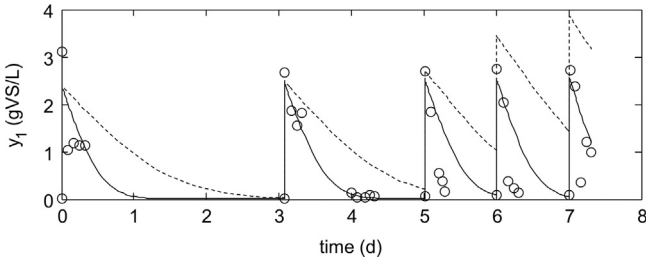

Fig. 13. Measurements y_1 (circles) and simulated model output $\hat{y}_1 = \hat{S}_1$ (solid line: estimated parameters values, dashed line: initial parameters values).

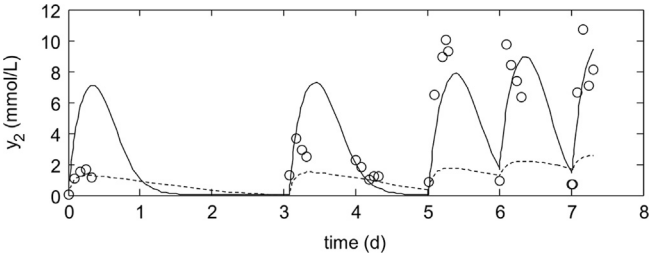

Fig. 14. Measurements y_2 (circles) and simulated model output $\hat{y}_2 = \hat{S}_2$ (solid line: estimated parameters values, dashed line: initial parameters values).

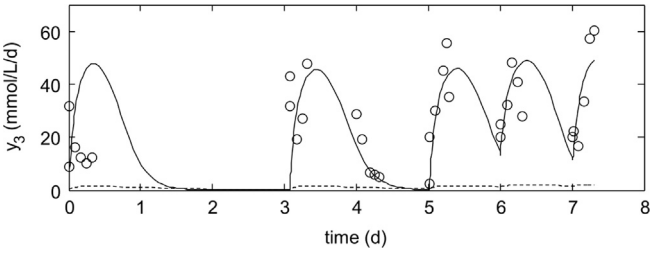

Fig. 15. Measurements y_3 (circles) and simulated model output $\hat{y}_3 = \hat{r}_{CH_4}$ (solid line: estimated parameters values, dashed line: initial parameters values).

As stated before, the available experimental data set was inadequate for the identification of the parameter K_{I2}. Therefore, one further case was conceived in order to test the feasibility of identifying a full set of kinetic parameters, including K_{I2}, once an adequate data set is available. To this purpose, a synthetic data set was obtained by simulating the degradation behavior of cheese whey, as predicted by the previously described AMOCO model.

In this case however, the ADM1 model was considered as unsuitable to generate artificial data. In fact, in this model, the inhibition of methanogenic bacteria by high VFA levels is described as due to the decrease in the suspension pH, which follows VFA accumulation, according to an empirical S-shaped equation. In contrast, the AMOCO model describes this inhibition phenomenon as directly related to the high concentration of acid, this mismatch being due to the fact that no consensus has yet been achieved in the literature on the most correct way to describe the inhibition of methane production under VFA accumulation conditions. Because of the mismatch, the same AMOCO model was used to generate the artificial data set.

In this experiment, as for the anaerobic reactor, a CSTR was again considered with an average t_{HR} of 10 days, fed on S_1 at a concentration of 50 gVS/L, until a steady state was achieved. The simulation of a large increase in the organic load fed to the digester has been then simulated in order to cause VFA accumulation, by modifying the influent concentration according to the pattern shown in Fig. 16. A sampling time of 1 h was assumed and, in order to better match a real experiment, a white noise disturbance, with a signal-to-noise ratio of 20, 20, and 25, has been added to the simulated outputs y_1, y_2 and y_3 respectively. As a matter of fact, when the S_2 concentration approaches the limit concentration of $\sqrt{K_{S2}K_{I2}}$ (i.e. the concentration above which any further S_2 increase would result in a slower growth of X_2), the inhibiting effect caused by VFA accumulation starts affecting the digester behaviors and, therefore, the outputs y_i. These data sets are therefore expected to be adequate when both K_{S2} and K_{I2} are to be identified.

Assuming the following vector of unknown parameters:

$$\delta = \begin{bmatrix} \delta_1 & \delta_2 & \delta_3 & \delta_4 & \delta_5 & \delta_6 \end{bmatrix}^T = \begin{bmatrix} K_{S1} & K_{S2} & \mu_{1,\max} & \mu_{2,\max} & k_6 & 1/K_{I2} \end{bmatrix}^T \quad (66)$$

while maintaining the same values for the other fixed parameters (Table 5), the rewriting of Eqs. (2)–(5) and (10) gives

$$\dot{x}_1 = \frac{u_1}{V}(u_2 - x_1) - k_1 \frac{x_1}{x_1 + \delta_1}\delta_3 x_3 \quad (67)$$

$$\dot{x}_2 = \frac{u_1}{V}(u_3 - x_2) + k_2 \frac{x_1}{x_1 + \delta_1}\delta_3 x_3 - k_3 \frac{x_2}{x_2 + \delta_2 + x_2^2\delta_6}\delta_4 x_4 \quad (68)$$

$$\dot{x}_3 = -\frac{u_1}{V}x_3 + \delta_3\left(\frac{x_1}{x_1 + \delta_1} - k_{d1}\right)x_3 \quad (69)$$

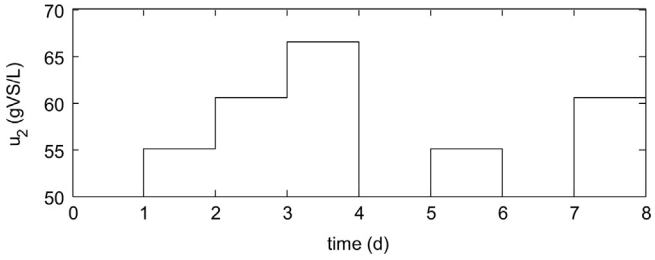

Fig. 16. Input S_1 concentration.

$$\dot{x}_4 = -\frac{u_1}{V}x_4 + \delta_4\left(\frac{x_2}{x_2 + \delta_2 + x_2^2\delta_6} - k_{d2}\right)x_4 \quad (70)$$

$$y_1 = x_1 \quad (71)$$

$$y_2 = x_2 \quad (72)$$

$$y_3 = \delta_5\delta_4 \frac{x_2}{x_2 + \delta_2 + x_2^2\delta_6}x_4 \quad (73)$$

where state and input vectors x and u are still defined as in (63) and (64), and the relevant LFT model is reported in Appendix D.

Identification results are reported in Table 5, while the comparison between simulated measurements and model outputs is shown in Figs. 17–19. As one can see from Table 5, the optimal value is very close to the true one for all parameters but the K_{Si} estimates. This again is related to the practical identifiability issue already commented earlier. As a matter of fact, this experiment

Table 5
Initial, estimated and true parameters including K_{I2}.

Parameter	Initial value	Optimal value	True value
K_{S1} (kgCOD m^{-3})	1	1.2454	1.5
K_{S2} (mmol L^{-1})	10	10.7323	7.5
$\mu_{1,\max}$ (d^{-1})	0.3	0.18626	0.2
$\mu_{2,\max}$ (d^{-1})	0.3	0.216	0.2
k_6 (mmol gVS^{-1})	30	27.0219	27
K_{I2} (mmol L^{-1})	150	108.3032	120

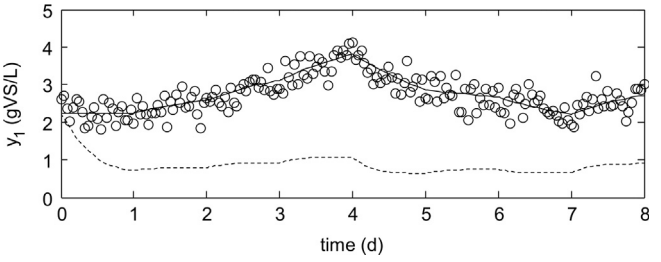

Fig. 17. Simulated measurements y_1 (circles) and simulated model output $\hat{y}_1 = \hat{S}_1$ (solid line: estimated parameters values, dashed line: initial parameters values).

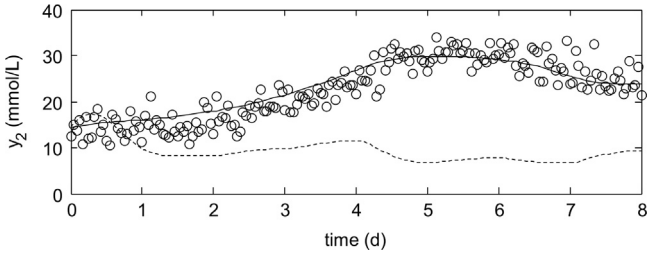

Fig. 18. Simulated measurements y_2 (circles) and simulated model output $\hat{y}_2 = \hat{S}_2$ (solid line: estimated parameters values, dashed line: initial parameters values).

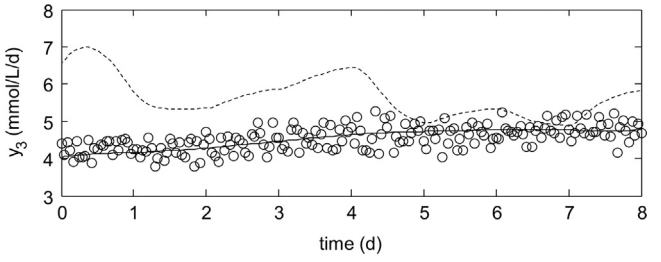

Fig. 19. Simulated measurements y_3 (circles) and simulated model output $\hat{y}_3 = \hat{r}_{CH_4}$ (solid line: estimated parameters values, dashed line: initial parameters values).

was meant to gather experimental data allowing for the estimation of K_{I2}, therefore forcing the digester to operate under higher S_1 and S_2 concentrations. Under those operational conditions, the effect of K_{Si} on the outputs is much less relevant, making them more difficult to identify.

On the other hand, it is well known that a reliable identification of Monod-like or Haldane-like parameters is not an easy task, and that the practical identifiability of these parameters is strongly dependent on data quality and reliability and on the experimental design, as already stressed by other authors (Nihtilä & Virkkunen, 1977; Vanrolleghem, Daele, & Dochain, 1995; Versyck, Claes, & Van Impe, 1998). In this respect, the practical identifiability issue cannot be overcome by applying the LFT technique, since it is related to the quality of the data set and on the experimental design, whose optimization is out of the scope of this work.

6. Conclusion

A reduced order model of anaerobic digestion has been first proposed in this paper, with the main goal to develop an efficient tool for process monitoring and control. To this aim, the AMOCO model, originally conceived to describe the degradation of soluble organic matter, has been modified to include a first order hydrolysis step.

In order to perform parameter estimation, the model has been then rewritten in a LFT formulation, using a symbolic tool originally developed for linear models and modified for the processing of nonlinear models. Based on this reformulation, the output error and the sensitivities with respect to the unknown parameters, needed in a maximum-likelihood estimation, can be computed respectively by simulating an index-1, semi-explicit DAE system and some linear but time-variant filters.

Two different test cases have been considered for the estimation of the uncertain parameters of the modified AMOCO model. In a first case, the data used for parameter identification have been generated by simulating the well known and more complex ADM1 model, considering waste activated sludge as substrate. In a second case, experimental data were collected on a laboratory scale equipment, operated in a semi-batch experiment, performing the anaerobic digestion of ultra-filtered cheese-whey.

The LFT formulation of the modified AMOCO model is currently under consideration for the implementation of MPC strategies, aimed at preventing instabilities due to sudden changes of the influent substrate.

Acknowledgment

This work has been performed within the project ID 2009-3170 La Fabbrica della Bioenergia, cofunded by Fondazione Cariplo, and with the support of Regione Lombardia, through the project ID 40602141 CITY WISE-NET.

Appendix A. Derivation of the sensitivity functions

Differentiating Eqs. (11)–(16) with respect to parameter δ_i yields

$$\frac{\partial \dot{\mathbf{x}}}{\partial \delta_i} = \mathbf{A}\frac{\partial \mathbf{x}}{\partial \delta_i} + \mathbf{B}_1\frac{\partial \mathbf{w}}{\partial \delta_i} + \mathbf{B}_2\frac{\partial \boldsymbol{\zeta}}{\partial \delta_i}$$

$$\frac{\partial \mathbf{z}}{\partial \delta_i} = \mathbf{C}_1\frac{\partial \mathbf{x}}{\partial \delta_i} + \mathbf{D}_{11}\frac{\partial \mathbf{w}}{\partial \delta_i} + \mathbf{D}_{12}\frac{\partial \boldsymbol{\zeta}}{\partial \delta_i}$$

$$\frac{\partial \boldsymbol{\omega}}{\partial \delta_i} = \mathbf{C}_2\frac{\partial \mathbf{x}}{\partial \delta_i} + \mathbf{D}_{21}\frac{\partial \mathbf{w}}{\partial \delta_i} + \mathbf{D}_{22}\frac{\partial \boldsymbol{\zeta}}{\partial \delta_i}$$

$$\frac{\partial \hat{\mathbf{y}}}{\partial \delta_i} = \mathbf{C}_3\frac{\partial \mathbf{x}}{\partial \delta_i} + \mathbf{D}_{31}\frac{\partial \mathbf{w}}{\partial \delta_i} + \mathbf{D}_{32}\frac{\partial \boldsymbol{\zeta}}{\partial \delta_i}$$

$$\frac{\partial \mathbf{w}}{\partial \delta_i} = \frac{\partial \boldsymbol{\Delta}}{\partial \delta_i}\mathbf{z} + \boldsymbol{\Delta}\frac{\partial \mathbf{z}}{\partial \delta_i}$$

$$\frac{\partial \boldsymbol{\zeta}}{\partial \delta_i} = \frac{\partial \boldsymbol{\Theta}(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}\frac{\partial \boldsymbol{\omega}}{\partial \delta_i}$$

which gives Eqs. (28)–(34) by renaming the partial derivatives:

$$\mathbf{x}_i' = \frac{\partial \mathbf{x}}{\partial \delta_i}, \quad \mathbf{z}_i' = \frac{\partial \mathbf{z}}{\partial \delta_i}, \quad \boldsymbol{\omega}_i' = \frac{\partial \boldsymbol{\omega}}{\partial \delta_i}, \quad \mathbf{y}_i' = \frac{\partial \hat{\mathbf{y}}}{\partial \delta_i}, \quad \mathbf{w}_i' = \frac{\partial \mathbf{w}}{\partial \delta_i}, \quad \boldsymbol{\zeta}_i' = \frac{\partial \boldsymbol{\zeta}}{\partial \delta_i}$$

Appendix B. LFT model for parameter identification based on ADM1 model simulation data

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\\ \dot{x}_5\end{bmatrix} = \begin{bmatrix}-dx_1 + w_6 + du_1\\ -dx_2 - w_1 - \mu_{1,\max}w_2 + du_2\\ -dx_3 + \mu_{1,\max}w_3 + \mu_{2,\max}w_4 + du_3\\ -(d + k_{d1}\mu_{1,\max})x_4 + \mu_{1,\max}\zeta_1\\ -(d + k_{d2}\mu_{2,\max})x_5 + \mu_{2,\max}\zeta_2\end{bmatrix}, \quad \begin{bmatrix}z_1\\ z_2\\ z_3\\ z_4\\ z_5\\ z_6\end{bmatrix} = \begin{bmatrix}w_6\\ \zeta_1\\ \zeta_1\\ -\zeta_2\\ -\zeta_2\\ -x_1\end{bmatrix},$$

$$\begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \omega_4\end{bmatrix} = \begin{bmatrix}x_5\\ x_3\\ x_4\\ x_2\end{bmatrix}, \quad \begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix} = \begin{bmatrix}x_3\\ x_1 + x_2 + c(x_4 + x_5)\\ -\mu_{2,\max}w_5\end{bmatrix},$$

$$\begin{bmatrix}w_1\\ w_2\\ w_3\\ w_4\\ w_5\\ w_6\end{bmatrix} = \begin{bmatrix}z_1\delta_1\\ z_2\delta_2\\ z_3\delta_3\\ z_4\delta_4\\ z_5\delta_5\\ z_6\delta_6\end{bmatrix}, \quad \begin{bmatrix}\zeta_1\\ \zeta_2\end{bmatrix} = \begin{bmatrix}\frac{\omega_3\omega_4}{K_{S1}+\omega_4}\\ \frac{\omega_1\omega_2}{K_{S2}+\omega_2+\omega_2^2/K_{I2}}\end{bmatrix}$$

Appendix C. LFT model for parameter identification based on experimental data

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix} = \begin{bmatrix}\zeta_3\\ \zeta_4\\ \zeta_5\\ \zeta_6\end{bmatrix}, \quad \begin{bmatrix}z_1\\ z_2\\ z_3\\ z_4\\ z_5\end{bmatrix} = \begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\\ \zeta_2\end{bmatrix},$$

$$\begin{bmatrix}\zeta_1\\ \zeta_2\\ \zeta_3\\ \zeta_4\\ \zeta_5\\ \zeta_6\end{bmatrix} = \begin{bmatrix}\frac{\omega_1^2}{\omega_1^2+\omega_2}\omega_3\\ \frac{\omega_4^2}{\omega_4^3/K_{I2}+\omega_4^2+\omega_5}\omega_6\\ \frac{\omega_7}{V}(\omega_8-\omega_1) - k_1\omega_9\\ \frac{\omega_7}{V}(\omega_{10}-\omega_4) + k_2\omega_9 - k_3\omega_{11}\\ -\frac{\omega_7}{V}\omega_{12} + \omega_9 - k_{d1}\omega_3\\ -\frac{\omega_7}{V}\omega_{13} + \omega_{11} - k_{d2}\omega_6\end{bmatrix}, \quad \begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \omega_4\\ \omega_5\\ \omega_6\\ \omega_7\\ \omega_8\\ \omega_9\\ \omega_{10}\\ \omega_{11}\\ \omega_{12}\\ \omega_{13}\end{bmatrix} = \begin{bmatrix}x_1\\ w_1\\ w_3\\ x_2\\ w_2\\ w_4\\ u_1\\ u_2\\ \zeta_1\\ u_3\\ \zeta_2\\ x_3\\ x_4\end{bmatrix}$$

$$\begin{bmatrix}w_1\\ w_2\\ w_3\\ w_4\\ w_5\end{bmatrix} = \begin{bmatrix}z_1\delta_1\\ z_2\delta_2\\ z_3\delta_3\\ z_4\delta_4\\ z_5\delta_5\end{bmatrix}, \quad \begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix} = \begin{bmatrix}x_1\\ x_2\\ w_5\end{bmatrix},$$

Appendix D. LFT model for identification of parameter K_{I2}

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix}=\begin{bmatrix}\zeta_4\\ \zeta_5\\ \zeta_6\\ \zeta_7\end{bmatrix},\quad \begin{bmatrix}z_1\\ z_2\\ z_3\\ z_4\\ z_5\\ z_6\end{bmatrix}=\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\\ \zeta_3\\ \zeta_2\end{bmatrix},$$

$$\begin{bmatrix}\zeta_1\\ \zeta_2\\ \zeta_3\\ \zeta_4\\ \zeta_5\\ \zeta_6\\ \zeta_7\end{bmatrix}=\begin{bmatrix}\frac{\omega_1^2}{\omega_1^2+\omega_2}\omega_3\\ \omega_4^3\\ \frac{\omega_6^2}{\omega_5+\omega_6^2+\omega_7}\omega_8\\ \frac{\omega_9}{V}(\omega_{10}-\omega_1)-k_1\omega_{11}\\ \frac{\omega_9}{V}(\omega_{12}-\omega_6)+k_2\omega_{11}-k_3\omega_{13}\\ -\frac{\omega_9}{V}\omega_{14}+\omega_{11}-k_{d1}\omega_3\\ -\frac{\omega_9}{V}\omega_{15}+\omega_{13}-k_{d2}\omega_8\end{bmatrix},\quad \begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \omega_4\\ \omega_5\\ \omega_6\\ \omega_7\\ \omega_8\\ \omega_9\\ \omega_{10}\\ \omega_{11}\\ \omega_{12}\\ \omega_{13}\\ \omega_{14}\\ \omega_{15}\end{bmatrix}=\begin{bmatrix}x_1\\ w_1\\ w_3\\ x_2\\ w_6\\ x_2\\ w_2\\ w_4\\ u_1\\ u_2\\ \zeta_1\\ u_3\\ \zeta_3\\ x_3\\ x_4\end{bmatrix}$$

$$\begin{bmatrix}w_1\\ w_2\\ w_3\\ w_4\\ w_5\\ w_6\end{bmatrix}=\begin{bmatrix}z_1\delta_1\\ z_2\delta_2\\ z_3\delta_3\\ z_4\delta_4\\ z_5\delta_5\\ z_5\delta_6\end{bmatrix},\quad \begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}=\begin{bmatrix}x_1\\ x_2\\ w_5\end{bmatrix},$$

References

Allegrini, A. (2010). *Anaerobic digestion modelling: A comparison between ADM1 and AMOCO* (Master's thesis). Politecnico di Milano.

APHA, AWWA, WEF. (2005). *Standard methods for the examination of water and wastewater* (21st ed.). Washington, DC: American Public Health Association, American Water Works Association, Water Environment Federation. Batstone, D. J., Keller, J., Angelidaki, I., Kalyuzhnyi, S. V., Pavlostathis, S. G., Rozzi, A., et al. (2002). The IWA anaerobic digestion model no. 1 (ADM1). *Water Science and Technology*, *45*(10), 65–73.

Bernard, O., Hadj-Sadok, Z., Dochain, D., Genovesi, A., & Steyer, J. (2001). Dynamical model development and parameter identification for an anaerobic wastewater treatment process. *Biotechnology and Bioengineering*, *75*(4), 424–438. Blumensaat, F., & Keller, J. (2005). Modelling of two-stage anaerobic digestion using the IWA anaerobic digestion model no. 1 (ADM1). *Water Research*, *39*(1), 171–183.

Casella, F., & Lovera, M. (2008). LPV/LFT modelling and identification: Overview, synergies and a case study. In *IEEE international conference on computer-aided control systems, 2008, CACSD2008* (pp. 852–857).

Alessandro, Della Bona, Gianni, Ferretti, Elena, Ficara, & Francesca, Malpei (2015). Parameter identification of a reduced order LFT model of anaerobic digestion. In: Marco Lovera (Ed.), *Control-oriented modelling and identification: Theory and Practice. Stevenage: Institution of Engineering and Technology* (pp. 303–329) (2014.03.10).

Demourant, F., & Ferreres, G. (2002). Closed loop identification of a LFT model. *Journal Européen des Systémes Automatisés*, *36*(3), 449–464.

Dennis, J. E., Jr., & Schnabel, R. B. (1996). Numerical methods for unconstrained optimization and nonlinear equations. In *Classics in applied mathematics, Vol. 16*. Society for Industrial & Applied Mathematics, Philadelphia, PA.

Donida, F., Romani, C., Casella, F., & Lovera, M. (2009). Integrated modelling and parameter estimation: An LFT-modelica approach. In *2009 IEEE conference on decision and control*, December (pp. 8357–8362).

Dötsch, H. G. M., & Van den Hof, P. M. J. (1996). Test for local structural identifiability of high-order non-linearly parametrized state space models. *Automatica*, *32*(July (6)), 875–883.

Ficara, E., Hassam, S., Allegrini, A., Leva, A., Malpei, F., & Ferretti, G. (2012). Anaerobic digestion models: A comparative study. In *7th Vienna conference on mathematical modelling, MATHMOD 2012*, February 15–17.

Galí, A., Benabdallah, T., Astals, S., & Mata-Alvarez, J. (2009). Modified version of ADM1 model for agro-waste application. *Bioresource Technology*, *100*(11), 2783–2790.

Hecker, S., Varga, A., & Magni, J. F. (2004). Enhanced LFR-toolbox for MATLAB. In *Proceedings of 2004 IEEE international symposium on computer aided control systems design*. Taipei, Taiwan, September (pp. 25–29).

Hsu, K., Poolla, K., & Vincent, T. (2008). Identification of structured nonlinear systems. *IEEE Transactions on Automatic Control*, *53*(11), 2497–2513.

Hsu, K., Vincent, T., Wolodkin, G., Rangan, S., & Poolla, K. (2008). An LFT approach to parameter estimation. *Automatica*, *44*(12), 3087–3092.

Klein, V., & Morelli, E. A. (2006). Aircraft system identification: Theory and practice. In *AIAA education series*.

Lee, L. H., & Poolla, K. (1999). Identification of linear parameter-varying systems using nonlinear programming. *Journal of Dynamic Systems, Measurement, and Control*, *121*(March), 71–78.

Ljung, L. (1999). *System identification: Theory for the user*. Upper Saddle River, NJ, USA: Prentice Hall.

Nihtilä, M., & Virkkunen, J. (1977). Practical identifiability of growth and substrate consumption models. *Biotechnology and Bioengineering*, *19*(12), 1831–1850. Rosen, C., Vrecko, D., Gernaey, K., Pons, M., & Jeppsson, U. (2006). Implementing ADM1 for plant-wide benchmark simulations in Matlab/Simulink. *Water Science and Technology*, *54*(4), 11–19.

Van Doren, J.F.M., Van den Hof, P. M.J., Jansen, J.D., & Bosgra, O.H. (2008). Determining identifiable parameterisations for large-scale physical models in reservoir engineering. In: *2008 IFAC world congress*. Seoul, South Korea.

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

Versyck, K. J., Claes, J. E., & Van Impe, J. F. (1998). Optimal experimental design for practical identification of unstructured growth models. *Mathematics and Computers in Simulation*, *46*(5), 621–629.

Zhou, K., Doyle, J. C., & Glover, K. (1996). *Robust and optimal control*. Upper Saddle River, NJ, USA: Prentice-Hall.