Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale,
Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Modelling of agro-zootechnical anaerobic co-digestion for full-scale applications

Davide Carecci[a], Giulia Quarta[a], Arianna Catenacci[b], Gianni Ferretti[a], and Elena Ficara[b]*

[a] Politecnico di Milano, Department of Electronics, Information and Bioengineering (DEIB), Milan, MI, Italy
[b] Politecnico di Milano, Department of Civil and Environmental Engineering (DICA), Milan, MI, Italy
* Corresponding Author: elena.ficara@polimi.it.

ABSTRACT

To match the growing demand for biomethane production, anaerobic digestors need an optimal and time-varying adaptation of the input diet. Dynamic co-digestion constitutes a hard challenge for the limited instrumentation and control equipment typically installed aboard full-scale plants. The development of prediction models is foreseen to support process (optimal) design and control. In this work, a rigorous framework was applied to take full-scale applicability into account while dealing with the design and training of both high-fidelity and control-oriented first-principle/grey-box models, to be used for real-time optimization and process control respectively.

Keywords: Control-oriented modeling, Anaerobic co-digestion, Parameter estimation, Identifiability analysis.

I. INTRODUCTION

To achieve energy independence and environmental sustainability, European bio-methane production is expected to increase over the incoming decades. Anaerobic digestion (AD) processes convert the biodegradable components of organic biomass into bio-methane and biogenic carbon dioxide (biogas), contributing to reduce the greenhouse gas emissions in the agricultural and waste sectors. AD entails a complex sequence of bio-chemical processes (hydrolysis, acidogenesis, acetogenesis, methanogenesis) which requires strict control of the operating conditions to avoid reductions in biogas production, or even complete process failure. Today, the techno-economic performance of bio-methane production via AD is increasingly tied to the plants' capability to embrace co-digestion (AcoD) of multiple co-feedstocks and adapting its input diet to the dynamic supply-chain market (e.g. seasonality). As typical of bio-chemical processes, there are two mains 'time-scales' of optimization tools: (i) high-level scheduling and (ii) low-level process control, to support plant managers and operators respectively. The ability of process models to predict the effect of diet change in advance is of primary importance to pursue those goals. However, the implementation of model-based control is hindered by (i) poor measurement instrumentation (typically gas flowrate and composition being the sole data frequently and automatically (i.e. *online*) measured); (ii) highly non-linear and slow kinetics (time of response from some hours to days); (iii) high parameter uncertainty; (iv) input variability/disturbances; (v) under-actuation. Indeed, literature has primarily focused on developing comprehensive first-principle, grey-box models rather than data-driven ones. The IWA ADM1 model [2] remains the state-of-the-art of AD modeling for process understanding, *offline* optimization and scenario analysis, but it is not suitable to be directly applied to model-based control due to its poor observability/identifiability. The AM2 [3], a simpler two-stage reaction model, was specifically designed with structural local observability and identifiability, to control the AD of simple and soluble feedstocks. After model definition, uncertain parameters shall be estimated: literature is flourish in ADM1 and AM2 tuning for different case studies [9], but their application to dynamic agro-zootechnical AcoD remains scarce [1], as well as the use of general/reliable guidelines for parameter estimation [7].

The aim of this work is to derive a reliable model for control purposes. The approach was to rely as much as possible on a trustworthy 'high-fidelity' model, following the steps: (i) a high-fidelity ADM1-like and a control-oriented (i.e. reduced-order) AM2-like models were extended/adapted for agro-zootechnical AcoD; (ii) with the constraints of industrial practice in mind, data from a

pilot-scale reactor were collected and an identifiable subset of the uncertain high-fidelity model's parameters was selected for estimation, exploiting also batch activity tests for improved practical identifiability; (iii) the reduced-order models' parameters were firstly trained in a 'physically-informed' way with highly informative synthetic data generated by the high-fidelity model, and, later, only a practically identifiable subset of parameters was 'refined' on real data. A model adaptation routine was tested to update the parameters estimated *offline* when model performance is poor. Considerations for the future application of the reduced-order models in control schemes are reported too.

II. PROCESS MODEL

A. High-fidelity: from ADM1 to agri-AcoDM

The high-fidelity model (named agri-AcoDM) is an extension of the ADM1, and resulted in a strongly non-linear differential algebraic equation (DAE) system with 43 state variables. Most of the extensions are described in detail elsewhere [5] and are summarized in the Digital Supplementary Material (DSM). In particular, (i) hydrolysis was extended to consider macromolecules with different times of response (e.g. readily/slowly biodegradable); (ii) salt precipitation and the non-ideality of acid-base equilibrium were considered for a better prediction of pH and inorganic nutrient (nitrogen N, phosphorous P) availability; (iii) the Monod functions used to describe acetate and propionate uptakes were substituted with Haldane ones, since high volatile fatty acids (VFA) concentration was proved to be inhibiting by itself, regardless the pH value [6]. The uncertainty in the model may lay both in the input parameters (characterization of the co-feedstocks in state variables' terms) and in stoichiometric, physico-chemical and kinetic parameters $\boldsymbol{\theta_p}$, but it is common to consider kinetics as the most relevant source of uncertainty. Further details on the system of equation are reported in the DSM.

B. Reduced-order: from AM2 to AM2HN/tan

The reduced-order model extends the AM2HN [8] - itself an extension of the AM2 - with additional state variables and processes for the hydrolysis of the biodegradable particulate fraction of each of the *n* co-feedstocks (i.e. $\boldsymbol{X_h}$ $(g_{VS}\ L^{-1}) \in \mathbb{R}^n$). The resulting ordinary differential equation (ODE) system is reported in the DSM. Since inhibition from free ammonia (S_{nh3}) released by N-rich co-feedstocks (e.g. animal slurries) can be relevant [9], two model versions with ('AM2HNtan') and without ('AM2HN') S_{nh3} inhibition were studied: a state variable (*N*) for the total ammoniacal nitrogen (TAN) was added in the AM2HNtan, along with a multiplicative non-competitive inhibition function in the μ_2 expression, compared to the sole original Haldane function present in the AM2HN.

III. PARAMETER ESTIMATION

Due to the high number of uncertain model's parameters and their collinearity, parameter estimation is all but a trivial task. The availability of data with sufficiently high frequency is of primary importance to stimulate parameter sensitivity, resulting in narrower estimation's confidence intervals. However, this is available in practice only for gas measurements at full-scale and continuous operation data are often poorly informative. Sporadic bio-methane potential tests (BMP) and activity tests are relatively cheap and yet very informative experiments, but the collinearity between $\boldsymbol{\theta_p}$ and unknown state's initial conditions hinders their exploitation. From these realistic considerations it came the intuition of combining data from continuous and batch tests: deriving the inoculum composition from the continuous operation, a mapping from $\boldsymbol{\theta_p}$ to the batch initial conditions is implicitly obtained and the degrees of freedom of unknowns are thus reduced [6]. Batch test's data were used only in the agri-AcoDM $\boldsymbol{\theta_p}$'s estimation, and the knowledge was then 'transferred' to the AM2HN/tan models exploiting synthetic data generated from agri-AcoDM.

A parameter subset selection (PSS) scheme based on both sensitivity and collinearity analysis was run to select a subset $\boldsymbol{\theta}_{PSS} \subseteq \boldsymbol{\theta_p}$ for estimation. For agri-AcoDM, 24 highly uncertain kinetic $\boldsymbol{\theta_p}$ [9] were selected to conduct and compare both linear/local (L.) and global (G.) dynamic sensitivity analysis (see DSM). The relative-relative sensitivity ($\boldsymbol{SI_{rr}}$) matrixes and the practical identifiability of each subset of parameters were computed as in [4]. The average over the data time points of the squared sensitivity values ($\boldsymbol{\delta}^{msqr}$) was selected as metric to rank the parameters, obtaining their ranking indexes $\boldsymbol{\rho}$.

The scalar cost function (Eq. 1) was the sum of the percentage mean relative squared (simulation) errors ($MRSE_{\%}$) over the different outputs of the dataset (see Section IV):

$$J = \sum_y w_y \frac{1}{N_y} \sum_{i=1}^{N_y} \left(\frac{y_i - \hat{y}_i}{\bar{y}}\right)^2 \quad (1)$$

where the 'y' nomenclature refers to each data type ('Name' in Table I of the DSM), N_y is the number of time observation points, $\hat{y}_i$ is the 'y' model output at the i^{th} time instant and w_y is a user-configurable weighting parameter to drive the 'efforts' of the minimization problem toward a specific y. The average value $\bar{\boldsymbol{y}}$ normalizes simulation errors for variables of different magnitudes. The cost function was minimized by 30 iterations of the Differential Evolution (DE) algorithm and later 'polished' by the Nelder-Mead algorithm. Parameters' initial guesses (*nominal* $\boldsymbol{\theta}_{PSS}$) were set to literature-suggested values (Table 1-2). The algorithm moved within the feasibility set $\boldsymbol{\Theta} \subset \mathbb{R}_{>0}$ (large yet compact, based on literature values) returned the (*optimal*) minimizer $\boldsymbol{\theta}^*_{PSS}$. Spearman's

coefficient of determination (R^2) and Mean Absolute Relative Error (MARE) were used to quantify model's performances. After the estimation of the $\boldsymbol{\theta}^*_{PSS}$ *expected values*, the Fisher Information Matrix (**FIM**) was evaluated as a lower-bound of the covariance matrix to compute the 95% confidence intervals (CI) of the $\boldsymbol{\theta}^*_{PSS}$ estimates, and T-student hypothesis testing was performed [6]. The uncertainty was then linearly propagated to each *optimal* $\hat{\boldsymbol{y}}^*$ trajectory (see DSM for references).

For the AM2HN/tan models, the estimation entailed two stages: (i) a first 'physically-informed' training on agri-AcoDM-generated synthetic data under highly exciting and well-distributed input profile, and (ii) a 'refinement' on real data of the practically identifiable parameters only. The identifiability analysis was conducted on almost all stoichiometric and kinetic $\boldsymbol{\theta_p}$ (18 in total). Stage '(i)' was designed to *activate*/excite the VFA inhibition to provide information on the Haldane curve's parameters. All the available correspondences between ADM1- and AM2-like variables were exploited in this stage (see DSM Table I) [8]: however, $\boldsymbol{k_h}$ was estimated before stage '(i)', considering the simulation error on the sum of $\boldsymbol{X_h}$ only to limit the effect of other mismatches on its estimate.

IV. THE DATASET

Considering the poor identifiability of agri-AcoDM and the limited budget of companies, the dataset must be minimal yet smartly designed to *activate* the parameters of interest and avoid overfitting. Data from a continuous pilot-scale reactor (12 L working volume, 42°C, initial D = 0.03 d^{-1}, organic loading rate (OLR) = 2.84 g_{COD} $d^{-1}L^{-1}$ with 30, 47 and 23% of maize silage, cow slurry and tomato sauce respectively) resembling the operation of a full-scale plant were considered. Coherently to the industrial practice, the methane and carbon dioxide flowrate (Q_{ch4}, Q_{co2}) and the biogas composition ratio (CO_2/CH_4) only were considered as *online* data. In addition, it is reasonable to assume that some spot *offline* VFA (and TAN) data would be available at full-scale, especially during periods of diet change, so that some measurements of acetate (S_{ac}) and propionate (S_{pro}) concentrations were included. The peculiarity of the entire dataset (15/12/2023-10/04/2024) is that it entails various operative conditions (e.g. impulse and long free-responses (22/12/2023-08/01/2024)), that is an effective way to prevent overfitting. In particular, the *transient* between two different loads and diet compositions was included in the training dataset (05-23/02/2024).

A comprehensive feedstock characterization is usually available in full-scale plants, yet with low frequency, (e.g. seasonal): for each co-feedstock, characterization was performed as in [5] only once, along with BMP tests to derive information about the biodegradability.

On the 10/04/2024 some batch tests were run using the pilot-scale digestate as inoculum: activity on S_{ac} and S_{pro} were tested, with initial substrate concentrations hypothesised to range from limited to inhibited conditions. Hourly Q_{ch4} of all the batch tests was recorded. Additional practical details on the dataset and its exploitation are present in the DSM-Table II.

V. RESULTS AND DISCUSSIONS

A. Agri-AcoDM calibration and performance

A reasonably collinear (collinearity index $\gamma \leq 20$) $\boldsymbol{\theta}_{PSS}$ of 14 sensitive ($\boldsymbol{\delta}^{msqr} \geq 0.1$) parameters was returned by the PSS. The ranking indexes $\boldsymbol{\rho}$, *nominal* and *optimal* values (± 95% CI) of each $\boldsymbol{\theta}^*_{PSS}$ are reported in Table 1. The T-student test rejected the null hypothesis for $k_{m,su}$, $k_{m,h2}$, $K_{i,h2,pro}$ and $K_{s,h2}$ (p-value ≤ 0.05), in agreement with previous findings [6]. Figure 1 shows model's performances over the most interesting operative period: although not reported for readability reasons, the propagated uncertainty resulted in the definition of $\hat{\boldsymbol{y}}$ CI that included almost all model mismatches for Q_{ch4} and CO_2/CH_4 (10%-65% and 5%-25% mean-max deviations from the $\hat{\boldsymbol{y}}$ *expected values,* respectively). However, too large and narrow sets for VFAs (120%-230%) and pH ($\leq$ 2%) were returned, respectively. Table 3 reports the metrics of model's performances over training and test periods. Remarkably, the MARE value for Q_{ch4} and S_{ac} decreased by 32% and 92% respectively, compared to the predictions with the *nominal* $\boldsymbol{\theta}_{PSS}$ values over the training period, and R^2 increased substantially. The good performance obtained for CO_2/CH_4 over the test period is very encouraging as well. The high MARE on S_{pro} is due to many data points being near the lower analytical detection limit: $k_{m,pro}$ was reduced to improve the fitting of the high-concentration values and of the activity tests, with a consequent worse fit of low-concentration points. pH values were stably within the 7.1-7.5 range due to the system's high buffering capacity: consequently, the MARE value for pH is low ($\leq$ 4%), but the true challenge of capturing the pH dynamics is reflected by the lousy R^2 (yet a remarkable 0.7 over the training data, but lower than 0.42 over the test data). The validation period was particularly challenging due to the presence of prolonged free-responses (far from actual operation), so that, overall, the performances are satisfactory. In addition, the main falls of model's performance occur over free-responses and transient periods, whereas it is very satisfactory over the two 'steady-states' before and after the diet change (all R^2-MARE higher-lower than 0.6-20% respectively). Indeed, the most interesting experimental finding for control with diet adaptation is the recorded VFA (mainly acetate) increase during the transient period, suggesting unbalanced biological performances which are to be avoided (see Section IV and Figure 1).

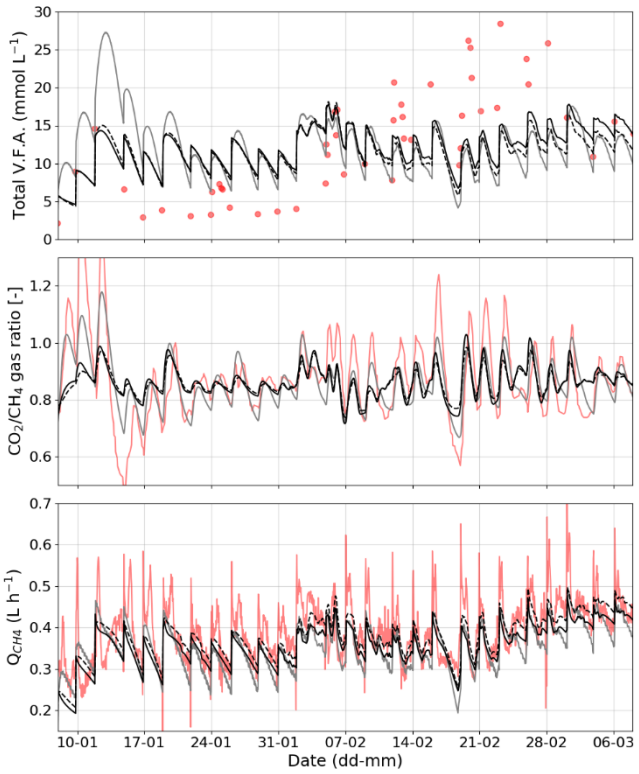

Figure 1. Models' performances: agri-AcoDM (grey line), AM2HN (black line) and AM2HNtan (dashed black line) vs data (red).

The low variability of pH suggests that no pH inhibition of the methanogenesis took place, thus remarking the need to introduce the Haldane functions to justify the VFA increase. Anyway, neither pH nor VFA inhibitions appear to be sufficient to explain such accumulation, since the optimal points of the Haldane functions resulted at values that were not reached during the transient (2.81, 0.26 g_{COD} L^{-1} for S_{ac} and S_{pro} respectively).

Figure 2 highlights the relevant improvements in model's predictions over the activity tests: MARE decreased on average by 55%. For the high substrate's initial concentration tests, R^2 were 0.92 and 0.64 for S_{ac} and S_{pro} respectively. For the BMP test on cow slurry, mean-max deviations were 28%-35%, and 40%-70% and 100%-200% for low and high initial concentration's activity tests respectively.

Table 1: Parameter estimates for the high-fidelity model.

Name (Unit)	L. $\rho(\theta_{PSS})$; G. $\rho(\theta_{PSS})$	***Nominal****	***Optimal* ±95%CI**
$k_{m,ac}$ $(g_{COD}g_{COD}^{-1}d^{-1})$	1; 5	8	4.9 ± 2.3
$k_{m,pro}$ $(g_{COD}g_{COD}^{-1}d^{-1})$	2; 24	13	5.3 ± 1.9
$K_{I,nh3}$ $(mol_N L^{-1})$	3; 15	1.8e^{-3}	2.7e^{-3} ± 1.6e^{-3}
$K_{I,hpro,pro}$ $(g_{COD}L^{-1})$	4; 18	2.2**	3.3 ± 0.9
$K_{I,hac,ac}$ $(g_{COD}L^{-1})$	5; 12	2.2**	7.9 ± 1.3
$pH_{UL,ac}$ $(-)$	7; 9	7	7.96 ± 0.33
$K_{s,ac}$ $(g_{COD}L^{-1})$	9; 21	0.15	0.44 ± 0.08
$k_{La,batch}(d^{-1})$	10; 13	10**	1.05 ± 0.05
$k_{m,su}$ $(g_{COD}g_{COD}^{-1}d^{-1})$	12; 7	30	29.5 ± 118
$k_{m,h2}$ $(g_{COD}g_{COD}^{-1}d^{-1})$	13; 11	35	21.4 ± 24.9
$K_{I,h2,pro}$ $(g_{COD}L^{-1})$	14; 16	3.5e^{-6}	2.2e^{-6} ± 2.2e^{-6}
$K_{s,h2}$ $(g_{COD}L^{-1})$	17; 1	7e^{-6}	1.3e^{-6} ± 2.3e^{-5}
$K_{s,pro}$ $(g_{COD}L^{-1})$	18; 4	0.1	0.08 ± 0.03
$k_{hyd,xchr}$ (d^{-1})	19; 19	1**	0.65 ± 0.15

**ADM1 default values. **Literature-suggested values.*

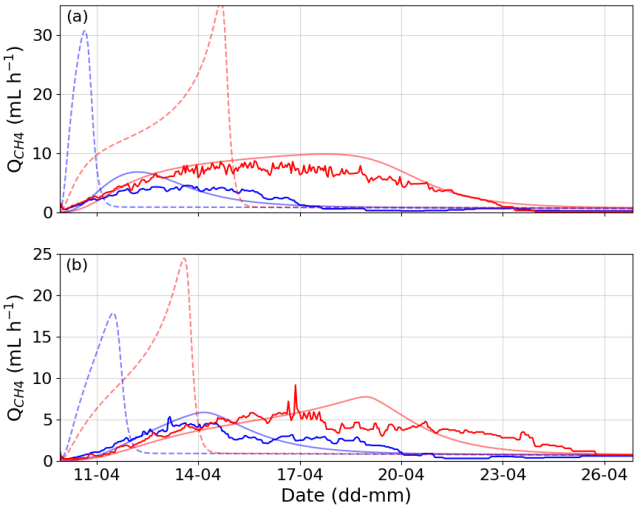

Figure 2. Agri-AcoDM model's performance for the batch tests on S_{ac} (blue) and S_{pro} (red): *nominal* (dashed line) and *optimal* (thinner continuous line) $\boldsymbol{\theta}_{PSS}$ vs data (thicker continuous line). Tests with high (a; ≥ 6 g L^{-1}) and low substrate's initial concentrations (b; 3 g L^{-1}).

B. AM2HN/tan calibration and performance

Results from the first-stage estimation for the reduced-order models on synthetic data were very satisfactory: R^2 was higher than 0.6 for all states and outputs; MARE below 10% for X_1 and X_2, and below 30% for S_1 and S_2 (total VFA). A reasonably collinear $\boldsymbol{\theta}_{PSS}$ ($\gamma \leq 10$) of 4 sensitive parameters was returned by the PSS for the second-stage estimation. The *optimal* parameter values are reported in Table 2, along with the SI_{rr} ranking indexes $\boldsymbol{\rho}$ of the AM2HNtan: the $\boldsymbol{\theta}^*_{PSS}$ values are reported within brackets ± 95% CI. Two mid-low r parameters (k_{La}, $K_{s,2}$) were selected instead of a more sensitive one (e.g. $\mu_{max,2}$ was not selected due to its collinearity with k_6) disclosing the main drawback of this PSS method i.e. its strong dependence on the user choice of a γ's upper threshold. Compared to agri-AcoDM, due to the lower dimensionality of $\boldsymbol{\theta}_{PSS}$, the propagated uncertainty resulted in narrower $\hat{\boldsymbol{y}}$ CI (9%-28%, 3%-18%, 14%-55% of mean-max deviation from the $\hat{\boldsymbol{y}}$ *expected values* of Q_{ch4}, CO_2/CH_4 and S_2 respectively). Table 3 and Figure 1 show model performances after both stages of estimation: compared to agri-AcoDM, similar performances for Q_{ch4} and total VFA were obtained, and yet less ability to reproduce the CO_2/CH_4 dynamics. Interestingly, for both

reduced models, the Haldane curves have similar total VFA's (S_2) optimal value compared to the one on S_{ac} in agri-AcoDM (difference ≤ 30%), whereas the maximum methanogenesis Haldane's rate are different (up to 70% difference), probably due to the higher complexity of agri-AcoDM (see DSM). The introduction of N release improved the fitting of Q_{co2} by 5-10% only, whereas modeling the hydrolysis was of great relevance. AM2HNtan reported a better performance over pH (0.7% vs 1.5% MARE): overall, the higher complexity of AM2HNtan may be justified when (i) the system is poorly buffered and thus a better pH prediction may be more relevant, and/or (ii) the load of N-rich co-feedstocks is used as control action.

Table 3: Models' performances over train/test periods.

	R²; MARE (%)		
	Q_{ch4}	CO_2/CH_4	$S_2(S_{ac},S_{pro})$
agri-AcoDM train	0.80; 23	0.73; 4	(0.29,0.81); (45,143)
agri-AcoDM test	0.80; 37	0.67; 6	(0.11,0.32); (65,109)
AM2HN train	0.73; 18	0.28; 11	0.29; 51
AM2HN test	0.71; 33	0.13; 15	0.27; 57
AM2HNtan train	0.72; 17	0.21; 12	0.25; 52
AM2HNtan test	0.71; 32	0.10; 18	0.21; 55

C. Recursive AM2HN parameter update

To reduce the relevant model mismatch over the diet transient, a modeler could either choose to add physical knowledge to the agri-AcoDM (e.g. lactate metabolism) or to embrace the philosophy for which *'no model is perfect':* in this work, a model adaptation scheme was tested for the AM2HN, via *'online'* recursive estimation of time-varying parameters over 10-days-long time windows (sliding, not overlapping) of the pilot-scale dataset. An *arrival* cost was added to the objective function to penalize 'useless' explorations far from the previously learned parameter values, thus preventing the risk to overfit process drifts and of catastrophic forgetting of the 'reference' parameters learned *offline* (Table 2). A weighting approach ($\boldsymbol{W_\theta}$) based on the $\boldsymbol{\rho}$ of the local $\boldsymbol{SI_{rr,k}}$ computed on the k^{th} window was applied to the *arrival* cost (see DSM): the resulting optimization problem on each k^{th} window returned the minimizer $\boldsymbol{\theta^*_{p,k}}$ expressed by Eq.2, where J_k is Eq.1 evaluated on the k^{th} window and $\boldsymbol{\theta^*_{p,k-1}}$ is the vector of *optimal* parameter values learned on the previous k^{th-1} window.

$$\boldsymbol{\theta^*_{p,k}} = \underset{\boldsymbol{\theta_p}\in\Theta}{argmin}\ J_k + ||\boldsymbol{\theta_{p,k}} - \boldsymbol{\theta^*_{p,k-1}}||_{2,\boldsymbol{W_\theta}}\ \forall_k \qquad (2)$$

The capability of the AM2HN to improve the fitting of S_2 data over the transient period is shown in Table 3 and Figure 3: considering the 'posterior' model trajectories (i.e. evaluated with the $\boldsymbol{\theta^*_{p,k}}$ on each window) the S_2 MARE reduced on average by 50% compared to the model's performance with the fixed parameters learned *offline*. The average relative standard deviation of the $\boldsymbol{\theta^*_p}$ over the windows was lower than 6%, with the only exception of $\mu_{max,2}$ that varied on average by 26%: it is likely that the latter triggered the most the adaptation to the S_2 transient. Although the most sensitive parameters were always $\mu_{max,2}$, k_5 and k_6, the $K_{I,2}$ jumped forward in the $\boldsymbol{SI_{rr,k}}$ ranking when S_2 reached high values thus activating the inhibition function (e.g. 5^{th} window). It follows that, in the prospective of a nonlinear model predictive controller (NMPC) design, the update could be beneficial to prevent overloading the system (e.g. on the 4^{th} and 5^{th} windows, where S_2 accumulation is foreseen by the 'prior' model trajectories i.e. evaluated with the parameters estimated on the 3^{th} and 4^{th} windows ($\boldsymbol{\theta^*_{p,3}}$ and $\boldsymbol{\theta^*_{p,4}}$)). Eventually, this AM2HN structure seems to be suited for *online* control purposes owing to its fair prediction ability and restricted complexity (i.e. optimization's duration etc.).

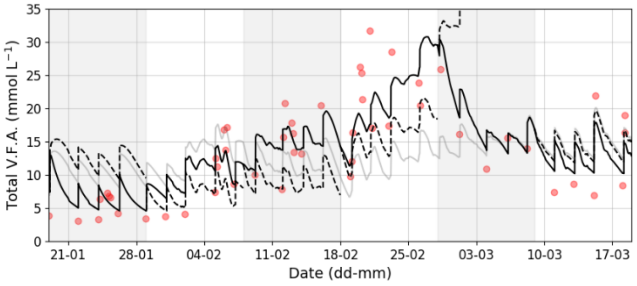

Figure 3. AM2HN model's performance with the *optimal offline* (grey line), the 'prior' (dashed black line) and the 'posterior' (black line) time-varying $\boldsymbol{\theta^*_p}$ vs data (red). Time windows are identified by the background color.

VI. CONCLUSIONS AND FUTURE WORK

Both high-fidelity and reduced-order models were tuned on an informative and challenging agro-zootechnical anaerobic co-digestion dataset. The results encourage the use of the AM2HN/AM2HNtan in model-based control applications. However, the models' current structure appeared to fail in describing the transient dynamics of a diet change: a model adaptation scheme was tested to reduce the model's mismatch, 'following' the transient. It is difficult to say whether the cause of such mismatch was due to a misleading agri-AcoDM's reconstruction of the unmeasured states, then inherited by AM2HN/tan, or to time-varying parameters. To put the record straight and to move toward NMPC design, further efforts will focus on the design of state observers. For such scope, an asynchronous separation between unmeasured states and slowly time-varying $\boldsymbol{\theta_p}$'s estimations can be beneficial, due to the strong collinearity between the two on the outputs (e.g. $\mu_{max,2}$ and X_2 on Q_{ch4}). The ultimate goal is the design and validation of a robust NMPC scheme over similar diet changes, to serve as a digital twin for a concrete and transparent improved management of full-scale reactors.

Table 2: Parameter estimates for both the reduced-order models (AM2HN and AM2HNtan).

Name (Unit)	**L.** $\rho(\theta_{PSS})$	**G.** $\rho(\theta_{PSS})$	***Nominal****	***Optimal* AM2HN**	***Optimal* AM2HNtan**
k_1 $(g_{COD}\ g_{VS}^{-1})$	n.d.	n.d.	42.1	14.5	16.2
k_2 $(mmol\ g_{VS}^{-1})$	n.d.	n.d.	117	154	206
$k_3(mmol\ g_{VS}^{-1})$	n.d.	n.d.	268	335	326
$k_4(mmol_C\ g_{VS}^{-1})$	5	7	50.6	34.5	39.4
$k_5(mmol_C\ g_{VS}^{-1})$	2	2	344	330 (391 ± 2)	313 (341 ± 1)
$k_6(mmol_C\ g_{VS}^{-1})$	1	1	453	408 (485 ± 2)	334 (422 ± 1)
$k_{La}(d^{-1})$	10	9	19.8	7.3 (19.3 ± 3.6)	2.8 (6.8 ± 0.6)
$K_{I,2}(mmol\ L^{-1})$	12	14	256	55.3	78.5
$K_{s,1}(mmol\ L^{-1})$	15	12	7.1	6.4	13.5
$K_{s,2}\ (mmol\ L^{-1})$	6	6	9.28	17.1 (27.7 ± 0.6)	6.5 (12.8 ± 0.3)
$\mu_{max,1}(d^{-1})$	14	13	1.2	8.1	13
$\mu_{max,2}(d^{-1})$	3	4	0.74	0.14	0.17
$k_{d,1}(d^{-1})$	13	15	0.057**	0.032	0.021
$k_{d,2}(d^{-1})$	7	8	0.019**	0.011	0.017
$k_h(d^{-1})$	[4, 9, 11]	[3, 5, 10]	[0.30, 0.18, 0.32]***	[0.22 ± 0.01, 0.14 ± 0.02, 0.48 ± 0.13]	
$K_{I,nh3}(mmol_N\ L^{-1})$	8	11	1.8	-	8.7

**From [3], with α = 0.5; **From [8]; ***From BMP tests' fit with first-order models.*
The vectorization within square brackets refers to the list of co-feedstocks ordered as: '[maize silage, cow slurry, tomato sauce]'.

DIGITAL SUPPLEMENTARY MATERIAL

Link to the Digital Supplementary Material: https://psecommunity.org/LAPSE:2025.0002

ACKNOWLEDGEMENTS

This work was supported by A2A S.p.A., the European Union (Next-Generation EU, Agritech National Research Centre for Agricultural Technologies) and the Italian Ministry of University and Research (PNRR).

REFERENCES

1. J. A. Arzate et al., Anaerobic digestion model (AM2) for the description of biogas processes at dynamic feedstock loading rates. Chemie Ingenieur Technik, 89(5):686–695, 2017. https://doi.org/10.1002/cite.201600176
2. D. Batstone et al., The IWA anaerobic digestion model No 1 (ADM1). Water Science and technology, 45(10):65–73, 2002. https://doi.org/10.2166/wst.2002.0292
3. O. Bernard et al., Advanced monitoring and control of anaerobic wastewater treatment plants: software sensors and controllers for an anaerobic digester. Water Science and Technology, 43(7):175–182, 2001. https://doi.org/10.2166/wst.2001.0418
4. R. Brun, P. Reichert, and H. R. Künsch. Practical identifiability analysis of large environmental simulation models. Water Resources Research, 37(4):1015–1030, 2001.
5. D. Carecci et al., A plant-wide modelling framework to describe microalgae growth on liquid digestate in agro zootechnical biomethane plants. Chemical Engineering Journal, 485:149981, 2024. https://doi.org/10.1016/j.cej.2024.149981
6. A. Catenacci, D. Carecci, A. Leva, A. Guerreschi, G. Ferretti, and E. Ficara. Towards maximization of parameters identifiability: Development of the CalOpt tool and its application to the anaerobic digestion model. Chemical Engineering Journal, 499:155743, 2024.
7. A. Donoso-Bravo, J. Mailier, C. Martin, J. Rodríguez, C. A. Aceves-Lara, and A. V. Wouwer. Model selection, identification and validation in anaerobic digestion: A review. Water Research, 45(17):5347–5364, 2011.
8. A. Della Bona, G. Ferretti, E. Ficara, and F. Malpei, LFT modelling and identification of anaerobic digestion. Control Engineering Practice, 36, 2015.
9. R. Mo, W. Guo, D. Batstone, J. Makinia, and Y. Li. Modifications to the anaerobic digestion model No. 1 (ADM1) for enhanced understanding and application of the anaerobic treatment processes – a comprehensive review. Water Research, 244:120504, 2023.