A Global Sensitivity Analysis Methodology for Anaerobic Digestion Models through Functional Principal Components Projection

Dhan Lord Fortela[1], Alyssa DeLattre[2], Spencer Kowalski[2], Wayne Sharp[2], Emmanuel Revellame[2], Rafael Hernandez[2], Daniel Gang[2], and Mark Zappi[2]

[1]University of Louisiana at Lafayette College of Engineering
[2]University of Louisiana at Lafayette

May 5, 2020

Abstract

Sensitivity analysis (SA) for the influence of model parametric constants has been integral in the use of mathematical kinetic models for design and operation of various anaerobic digestion applications. Using Anaerobic Digestion Model No. 1 (ADM1) as case study, this work aimed to broaden the approach for SA on the time-dependent model outputs of anaerobic digestion models by demonstrating the use of functional principal component analysis (fPCA) scores as input analysis variables into global SA (GSA) for the influence of stoichiometric parameters in ADM1. The methodology involved the following: Morris' screening design as the GSA technique; ADM1 biomass yield and product yield coefficients as GSA parameters; and ADM1 outputs transformation via fPCA to generate principal component (PC) scores for GSA. Results indicate that 95-99% of the variations in the time-dependent outputs can be captured by the PCs after fPCA transformation, and that the first PC is sufficient to represent the model outputs. Ranked Morris sensitivity indices calculated from the first PC scores revealed the stoichiometric parameters that dominantly affect kinetic responses and those that are least sensitive. The ranking of stoichiometric sensitivities can be used for various purposes including driving mechanisms identification, and mathematical model modification.

1. INTRODUCTION

Anaerobic digestion (aka. digestion) is a treatment bioprocess that uses a consortium of anaerobic microorganisms to treat the organic fraction of wastewaters via a stepwise degradation pathway through volatile organic acids into the primary end-products of methane and carbon dioxide (biogas), which can fuel electric generators, cookers, boilers, and/or heaters [1-3]. Digester performance can vary dramatically based on the composition of the wastewater influent and/or digestion operating conditions employed [2, 4]. Hence, any refinement of models that are capable of predicting both technical and economic performance are of great value to designers and operators considering options for both waste treatment and power generation using digestion.

Mathematical models have been instrumental in research and development of many bioprocesses. A consequence of the accumulation of findings on modeled processes and of elucidation of the fundamental structures of underlying mechanisms is the transformation of simple models into complex mathematical systems [5]. Application dictates how these models are manipulated: (1) fast responding models, which usually are the simple models, are preferred in process control [6]; (2) complex models are preferred during process and system design [7]; and (3) varying model complexities are typically encountered during parameter calibration with empirical data [8]. Thus, model reduction techniques are essential in using complex bioprocess models into specific types of applications [8, 9]. Closely related and often times inherent to model formulation and reduction problems is sensitivity analysis (SA) [10, 11]. SA measures how sensitive model outputs are to

perturbation of inputs (initial conditions, external forcing factors, model parameters, etc.) [11], which lends itself as a computational technique for uncertainty analysis, model calibration, model diagnostic evaluation, and surrogate model formulation [10]. The following SA techniques have been implemented in biological wastewater treatment models (anaerobic digestion and activated sludge): derivative-based local sensitivity analysis [6, 12-14], factorial sensitivity analysis [15], Standard Regression Coefficient (SRC) [16, 17], Morris screening method [18], and Extended Fourier Amplitude Sensitivity Testing (Extended-FAST) method [18]. At the center of these methods are Monte-Carlo type simulations of a model subjected to input perturbations, hence, simulating various response levels of model outputs [16]. These previous works dealt directly with the time-dependent model outputs – an approach that may be unsatisfactory due to the following reasons: (1) the SA input variables (time-dependent model outputs) are highly correlated with one another, (2) the pointwise (timewise) results can be difficult to interpret for the underlying physical or modelling problem, and (3) the different simulation runs may not generate outputs at the same points [19]. Others have tried to overcome the time-varying character of sensitivity indices by computing the overall average of the time-series sensitivity indices [20]. An alternative approach mathematically demonstrated by Campbell, McKay and Williams [19] is for the time-dependent outputs to be transformed into an appropriate functional coordinate system to generate a new set of functional representation (basis functions) that includes time-independent representation, e.g., coefficients of the set of basis functions . This concept and associated mathematical technique [19] was adopted by Sumner, Shephard and Bogle [21] to address the need for SA of time-dependent outputs in bioprocess modelling. The approach combines global sensitivity analysis (GSA) with functional principal component analysis (fPCA), which results into performing SA on the coefficients (scores) of the basis functions, also called functional principal components (PCs) derived from the model outputs. An overview of the methodology is schematically depicted in Figure 1. An advantage of this fPCA transformation over the typical averaging of sensitivity indices is the quantification of the time-dependent outputs variations captured by the PCs prior to the calculation of the sensitivity indices using the time-independent PC scores [19, 21].

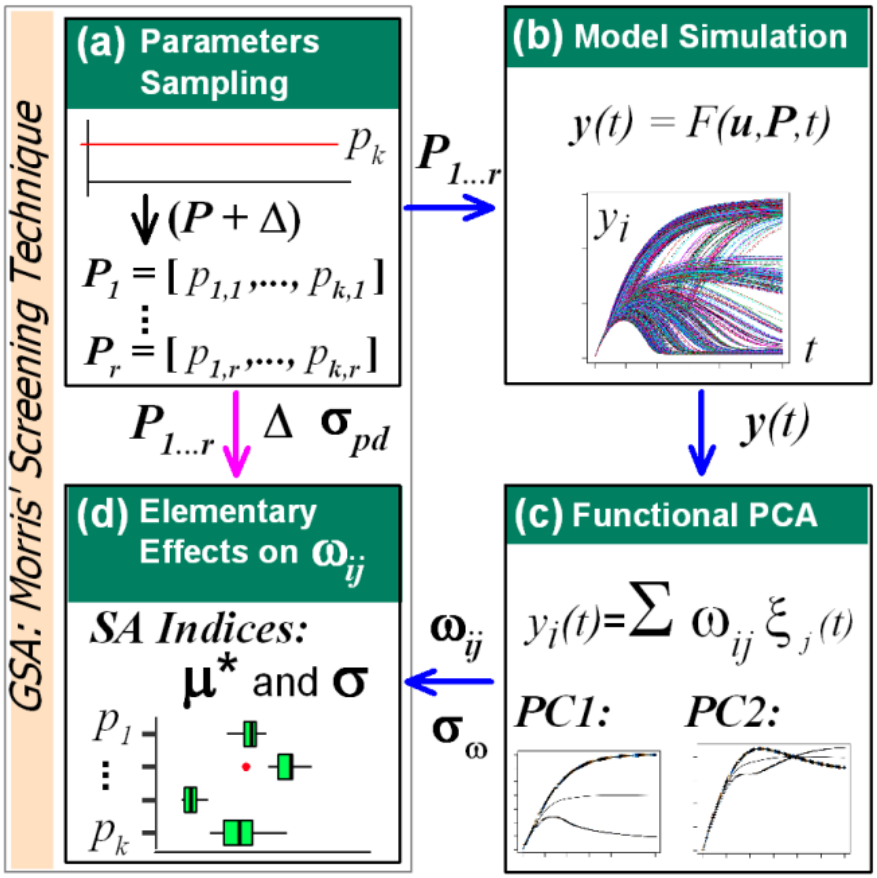
This work demonstrates a GSA-fPCA approach in the area of anaerobic digestion kinetics modelling and simulation. The approach combined fPCA with conventional global SA (Morris' technique) with the intent of establishing a comprehensive GSA for this particular bioprocess application. Of the many possible kinetic models for anaerobic digestion available in the literature, the Anaerobic Digestion Model No. 1 (ADM1) [22] was chosen due to its standardized consolidation of kinetic data from multitudes of research works, and for its complexity in terms of the integrated differential and algebraic equations, and of the nonlinear relations of model parameters. ADM1 describes the dynamics of 24 biochemical species (time-dependent) and 19 bioconversion processes together with key physico-chemical processes such as liquid-liquid and liquid-gas processes. The model can be implemented as systems of differential equations (DEs) or of differential-algebraic equations (DAEs) depending on how the acid-base reactions are computationally expressed. Though complex, this model does not necessarily describe all the mechanisms involved in anaerobic digestion, and calibration and validation of model parameters are required to achieve sufficient accuracy for application to specific feedstocks and operating condition [22]. The ADM1 parameters are grouped into three categories: (1) stoichiometric parameters, (2) kinetic parameters, and (3) physico-chemical parameters. ADM1 has been successfully used to model and simulate digestion of sewage sludge [7, 23-25], agricultural biomass [26-28], food waste [29], and municipal solid wastes at industrial-scale co-digestion [30]. More works are being performed for the extension of model database via constants calibration on new substrates [31-33], and process conditions [30, 34]. For these reasons, ADM1 was an appropriate benchmark model for the proof-of-concept, i.e., the success of GSA-fPCA approach in a mathematically complex digestion model (ADM1) should be an indication of its capability to handle SA on simpler or similarly complex mathematical models on digestion.

The main objective of this study is to demonstrate a comprehensive approach of GSA for anaerobic digestion models through aggregation of time-dependent response patterns into time-independent coefficients of functional PCs combined with a traditional GSA (Morris' technique). The GSA is implemented on the 24 biochemical time-dependent outputs of ADM1 as influenced by 22 stoichiometric parameters through the transformation of the model outputs into functional principal components (PCs) and their associated scores.

The PC scores were then used as inputs for the GSA for the calculation of sensitivity indices. This work intends to broaden the knowledge area for SA methodologies on ADM1 and other digestion kinetics models by demonstrating a GSA-fPCA approach and its implications in computational modelling of anaerobic digestion systems.

2. METHODOLOGY

The implementation of GSA-fPCA on ADM1 is schematically depicted in Figure 1.



p_k	k-th model parameter under study	F	Set of differential and algebraic equations
r	Number of elementary effects per parameter	u	External inputs to model (e.g. feedstock)
Δ	Parameters sampling steps	$y_i(t)$	i-th model time-series response
$P_{1\ldots r}$	Sampled parameters design	$y(t)$	Array of model time-series responses
σ_{pd}	Standard deviation of sampled parameters	$\xi_j(t)$	Basis function for the j-th PC
μ^*	Mean of absolute of elementary effects	ω_{ij}	Score of i-th response on j-th PC
σ	Standard deviation of elementary effects	σ_ω	Standard deviation of PC scores

Figure 1. Overview of the GSA-fPCA approach implemented for the sensitivity analysis of stoichiometric parameters of ADM1.

2.1 The ADM1

ADM1 could be represented as a model of the form$\mathbf{y}(t) = F(u, \mathbf{P}, t)$, where$\mathbf{y}(t)$ are the N time-dependent response curves, u are external model inputs, $\mathbf{P}$ is the $k$parameters vector, i.e.,$\mathbf{P} = (p_1, p_2, \ldots, p_k)$, and Fconsists mainly of a system of differential and algebraic equations that are highly nonlinear in terms of $\mathbf{P}$. The details on the implemented DAE for ADM1 are summarized in the supplementary material. The default mass basis was kilograms chemical oxygen demand (kg-COD), the volume basis was metric (m^3), and the time basis was in days (d). All of the biochemical, physico-chemical, and mass transfer components in the ADM1 were included in the computational algorithm as all of these components have some dependencies on one another, but only the 24 biochemical responses were included in the GSA. Specifically, the target responses for GSA were:$\mathbf{y}^*(t) = [S_{\mathrm{su}}(t), S_{\mathrm{aa}}(t), S_{\mathrm{fa}}(t), S_{\mathrm{va}}(t), S_{\mathrm{bu}}(t), S_{\mathrm{pro}}(t), S_{\mathrm{ac}}(t), S_{h2}(t), S_{ch4}(t), S_{\mathrm{IC}}(t), S_{\mathrm{IN}}(t), S_I(t), X_c(t), X_{\mathrm{ch}}(t), X_{\mathrm{pr}}(t), X_{\mathrm{li}}(t), X_{\mathrm{su}}(t), X_{\mathrm{aa}}(t), X_{\mathrm{fa}}(t), X_{c4}(t), X_{\mathrm{pro}}(t), X_{\mathrm{ac}}(t), X_{h2}(t), X_I(t)]$. The S's are the soluble components and the X's are the particulate components with units expressed as kg-COD/m^3. See Nomenclature section for definitions of model outputs $\mathbf{y}^*(t)$. Among all model parameters, only the 22 stoichiometric parameters were used for this study. They were of two types: $f_{prod,subs}$, which are the product yield coefficients from specific substrates and are of no units (dimensionless); and Y_{subs}, which are the biomass yield coefficients from specific substrates and with units expressed as kg-COD/kg-COD. These were: $\mathbf{P}^* = [f_{fa,li}, f_{va,aa}, f_{bu,su}, f_{bu,aa}, f_{pro,su}, f_{pro,aa}, f_{ac,su}, f_{ac,aa}, f_{h2,su}, f_{h2,aa}, f_{xI,xc}, f_{sI,xc}, f_{pr,xc}, f_{ch,xc}, f_{li,xc}, Y_{\mathrm{aa}}, Y_{\mathrm{su}}, Y_{c4}, Y_{\mathrm{fa}}, Y_{\mathrm{pro}}, Y_{\mathrm{ac}}, Y_{h2}]$. The $\mathbf{P}^*$sampling band was set within $\pm 30\%$ of the suggested nominal levels (Table 1) to emulate the common parameter variation levels suggested in the ADM1 literature [22]. The other model parameters were set their nominal values (see supplementary material). The notations were chosen to be consistent with the nomenclature in the ADM1 literature. The ADM1 model implementation was based on the digestion conditions, hence modelling conditions, in the work of Demitry [35] which focused on the co-digestion of municipal sewage sludge and bakery waste, and on the use of ADM1 as modelling platform [36, 37]. Specifically, the levels of initial conditions for the state variables were used to emulate the feedstock characteristics in the said previous work [35]. This case study system was chosen with the intent of demonstrating the various aspects of the GSA-fPCA approach. The period for simulation runs was set to 30 days to capture the transitional behavior from initial conditions to steady state conditions in the system.

Table 1. ADM1 stoichiometric parameters used for the GSA implementation.

Parameter	**Coefficient Definition**	**Nominal**	**Nominal (-) 30%**	**Nominal (+) 30%**
$f_{fa,li}$	Fatty acids from lipids	0.95	0.910+	0.980+
$f_{va,aa}$	Valerate from amino acids	0.23	0.16	0.30
$f_{bu,su}$	Butyrate from sugars	0.13	0.09	0.17
$f_{bu,aa}$	Butyrate from amino acids	0.26	0.18	0.34
$f_{pro,su}$	Propionate from sugars	0.27	0.18	0.35
$f_{pro,aa}$	Propionate from amino acids	0.05	0.03	0.07
$f_{ac,su}$	Acetate from sugars	0.41	0.28	0.53
$f_{ac,aa}$	Acetate from amino acids	0.40	0.28	0.53
$f_{h2,su}$	Hydrogen from sugars	0.19	0.13	0.25
$f_{h2,aa}$	Hydrogen from amino acids	0.06	0.04	0.08
$f_{xI,xc}$	Particulate inerts from composites	0.25	0.18	0.32
$f_{sI,xc}$	Soluble inerts from composites	0.10	0.07	0.13
$f_{pr,xc}$	Proteins from composites	0.20	0.14	0.26
$f_{ch,xc}$	Carbohydrates from composites	0.20	0.14	0.26
$f_{li,xc}$	Lipids from composites	0.25	0.18	0.32
Y_{aa}	Biomass from amino acids	0.08	0.06	0.10
Y_{su}	Biomass from sugars	0.10	0.07	0.13
Y_{c4}	Biomass from butyrate and valerate	0.06	0.04	0.08
Y_{fa}	Biomass from fatty acids	0.06	0.04	0.08
Y_{pro}	Biomass from propionate	0.04	0.02	0.05

Parameter	Coefficient Definition	Nominal	Nominal (-) 30%	Nominal (+) 30%
Y_{ac}	Biomass from acetate	0.05	0.03	0.07
Y_{h2}	Biomass from hydrogen	0.06	0.04	0.08

+The ±30% variation constraint was not applied. Batstone, Keller, Angelidaki, Kalyuzhnyi, Pavaostathis, Rozzi, Sanders, Seigrist and Vavilin [22] noted that the typical range is 0.91-0.98 depending on the LCFA chain length.

2.2 Functional Principal Component Analysis (fPCA)

The N time-dependent model outputs $\mathbf{y}(t)$ may be projected into some familiar functional expansions, and this study used functional bases computed through fPCA. The $\mathbf{y}(t)$ components are usually smoothed as spline curves using n samples from the time-series results of model simulation. The fPCs are specifically adapted to maximize the variance of the data projection onto the first basis vector, then to the spanned first and second basis vector, and so on, resulting into information aggregation through the first few PCs [19]. This approach considers time-series data as a function rather than discrete measurements. The PCs may be interpreted as a set of curves of the form$\xi(t) = (\xi_1(t),\ \xi_2(t), \ldots,\ \xi_q(t))$such that each model output $y_i(t)$ can be expressed as:

$y_i(t) = \sum_{j=1}^{q} \omega_{ij} \xi_j(t);\ for\ i = 1, \ldots, N (1)$

where ω_{ij} is the score for the model output i on the j^{th} PC. The PC score (ω_{ij}) captures how much of the model response $y_i(t)$ is represented in the basis function $\xi_j(t)$[21]. The ω_{ij}'s, therefore, are satisfactory measures of the aggregate behavior of the model outputs, and are appropriate analysis inputs for global SA [19]. For simplicity of GSA, the units of $\mathbf{y}(t)$, which were in kg-COD/m^3, have been assumed to be associated with the$\xi(t)$, so the ω_{ij}'s were dimensionless.

2.3 Global Sensitivity Analysis (GSA)

The GSA was the start and the end of the SA methodology (Figure 1). At the beginning, GSA set the level combinations of the k^* target parameters ($\mathbf{P}^*$) , using an assignment approach, which is dependent on a particular GSA technique. This generated a design of experiment (DOE) for the$\mathbf{P}^*$ that were then used to solve the ADM1 DAEs and to generate the responses that were eventually transformed via fPCA to determine the ω_{ij}'s. After these intermediate steps, the ω_{ij}'s were passed back to the GSA technique to analyze the ω_{ij}'s dataset against the DOE of $\mathbf{P}^*$, hence, determining the SA indices. The GSA used in this study was the Morris' screening design.

Morris' technique works on the evaluation of the elementary effect ($EE_{d,\ ij}$) of the d^{th} parameter defined as follows [38, 39].

$$EE_{d,\ ij}(\mathbf{P}^*) = \frac{[\omega_{ij}(p_1,\ldots,p_{d-1},p_d+,p_{d+1},\ldots,p_{k*}) - \omega_{ij}(\mathbf{P}^*)]}{\left(\frac{\sigma_{pd}}{\sigma_\omega}\right)} (2)$$

For each input, a number of incremental ratios called Elementary Effects (EE's) were calculated from which statistics are calculated to be aggregate measures, hence, global sensitivities. The is selected such that $\mathbf{P}^* +$ is still in the set of the allowable values of the k^* parameters [38]. This $EE_{d,\ ij}$equation incorporates a scaling factor, which is the ratio of the standard deviations of the model output (in this case the PC scores),σ_ω, and of the d^{th} parameter,σ_{pd}. Scaling of the DOE levels of$\mathbf{P}^*$ after building the design and before computing the EE's is necessary to eliminate the influence of step size on the sensitivity results especially when the parameter values are very small [39]. Changing for each p_d by r times generates a DOE containing$\mathbf{P}^*_1, \ldots, \mathbf{P}^*_r$. The simulation of the ADM1 model using these randomly sampled parameter values emulate a Monte-Carlo simulation, which has been the main approach in generating time-dependent outputs for SA of several wastewater treatment models [16, 18, 20]. The ADM1 model simulation using the

$\mathbf{P}^*$ DOE followed by the transformation of$\mathbf{y}^*(t)$ to PCs and calculation of elementary effects consequently result to a finite distribution of$\mathrm{EE}_{d,\ ij}$ values that can be aggregated using statistics. Hence, Morris method requires only $r(k^*+1)$ model runs to generate r EE's for each of the k^*parameters.[40] The $r$trajectories for the various 's are randomly assigned; hence, the calculated EE's are random samples [38]. The mean of the absolute values of the elementary effects (μ^*), and standard deviation of the elementary effects (σ) are unbiased estimators, which are meaningful indices for SA [40]. The aggregation of local indices through these estimators makes the Morris technique a global SA. A high level of μ^* indicates high overall effect of the d^{th} parameter on thei^{th} model output as represented by the j^{th}PC. The relative values of μ^*'s for the k^* parameters are used to rank the parameters for their influence on theω_{ij} and hence on the $y_i(t)$. A relatively high level of σ indicates a parameter with significant interactions with other parameters. As demonstrated by Sin, Gernaey, Neumann, van Loosdrecht and Gujer [16], ranking of sensitivity indices has been a convenient way of reporting GSA results due to the need to identify the most sensitive and least sensitive parameters.

$$\mu^* = \frac{\sum_r |\mathrm{EE}_{d,\ ij}|}{r} \quad (3)$$

$$\mu = \frac{\sum_r \mathrm{EE}_{d,\ ij}}{r} \quad (4)$$

$$\sigma = \sqrt{\frac{1}{r}\sum_r \left(\mathrm{EE}_{d,\ ij} - \mu\right)^2} \quad (5)$$

2.4 Computational Implementation

All computational works were implemented using the open-source R statistical software (version 3.2). The Morris' screening design for GSA was performed using the 'morris()' function from the 'sensitivity'[41] R-package (version 1.15) with the following arguments: r (varied) = 20, 50, 100; levels = 6; grid jump = 3; scaled = TRUE. The GSA parameters were the 22 stoichiometric coefficients of ADM1 defined above as$\mathbf{P}^*$. Morris technique randomly generated the various parameter combinations (hence trajectories) to create the design of experiments (DOE) for the SA. Ten random number generator seeds were used through the 'set.seed()' function for each set of experiments to facilitate randomization of the r trajectories. After the DOE of the stoichiometric parameters, i.e., a set of $\mathbf{P}^*$'s, was generated, the ADM1 was simulated using the 'ode()' function from the 'deSolve'[42] R-package (version 1.20). The simulation was set from 0-30 days of digestion, and$n = 100$ equally-spaced data samples were taken from each response variable in $\mathbf{y}(t)$. The simulation results on the 24 biochemical responses, i.e.,$\mathbf{y}^*(t)$, were transformed via fPCA. fPCA was performed using the functions from the 'fda'[43] R-package (version 2.4.7). This fPCA step generated the PC scores (ω_{ij}'s) on all the target ADM1 outputs. Finally, the scores were analyzed against the DOE of $\mathbf{P}^*$ via analysis of the EE's through the 'morris()' function. The Morris GSA indices were the$\mu^*$and σ of the EE's. The program script written for the execution of these steps is in the supplementary material.

3. RESULTS

The analysis results started with the transformation of model outputs$y^*(t)$ to PCs (Figure 2) followed by the calculation of the μ^* (Figure 3 & Figure 4) and σ (Figure 5 & Figure 6). Finally, the parameters were ranked for their influence on the model outputs based on μ^* (Figure 7) and σ (Figure 8). Several intermediate results (ADM1 time-dependent output series as graphs, fPCA curves as graphs, and tabulated values of graphical results presented below) are summarized in the supplementary material. The GSA experiments were completed in approximately two days using a desktop computer. Though the GSA-fPCA approach was implemented with varying$r$ (20, 50, 100), the results used for discussion are from$r = 100$, which produced improved convergence of values. The results from all three r levels are summarized in the supplementary material.

3.1 Model Outputs Transformation to PCs

The strategy for the use of the ω_{ij}'s as inputs for the elementary effects analysis in the Morris' technique is predicated on the assumption that the j^{th} PC captures most of the variations in the time-dependent model outputs. The reliability of this transformation must be examined prior to the analysis of the SA indices. Figure 2 shows the percentages of variations in the biochemical time-dependent outputs of ADM1

as captured by the first three principal components (PC1, PC2 & PC3). These components captured more than 99% of the total variations in all of the model outputs. Notably, the PC1 captured between 95 to 99% of the variations. This warranted the use of PC1 (hence ω_{i1}'s) as the fPCA component representing the variations of the target model outputs$\mathbf{y}^*(t)$ in the computation of the SA indices.

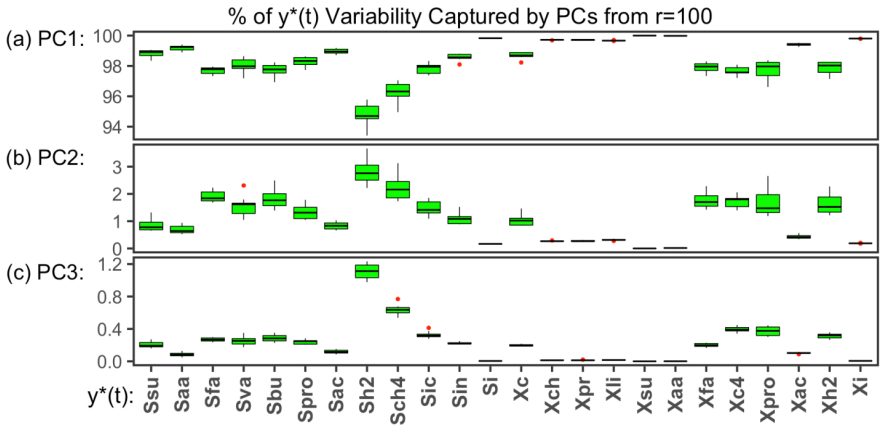


Figure 2. The first three functional principal components (PC1, PC2 & PC3) as transformations of the biochemical dynamical outputs in ADM1. The boxplots represent the values distribution from the analyses performed using ten different seeds for random number generator that perturbed the Morris screening design.

3.2 Sensitivity Indices on Stoichiometric Parameters

After verifying that the PC1 was sufficient enough to represent the variations in the target model outputs (Figure 2), the PC1 scoresω_{i1}'s were passed to the fPCA step for the calculation of the μ^* and σ for each parameter in$\mathbf{P}^*$ for every target model output in$\mathbf{y}^*(t)$. Figure 3 and Figure 4 show theμ^* results for the S and the X components, respectively. In general, some parameters consistently standout from the others across the model outputs (Figure 3a-3k; Figure 4a, 4g-4k) while some parameters become significant only in few model outputs (Figure 3l; Figure 4b-4f, 4l). The range of μ^* levels varies across the model outputs due to the direct relationship of the ω_{i1}'s with the levels of the time-dependent model outputs $\mathbf{y}^*(t)$. Theω_{i1}'s are high when the levels of $y_i(t)$are high; otherwise ω_{i1}'s are low. This behavior is unique to each model formulation. However, the ranking for the influence of parameters on the model outputs is implemented on the elementary effects of the parameters within each model output (S or X). The parameters ranking in a model output is compared with the parameters ranking associated with other model outputs. This approach results in a summary of the various rankings of the parameters on their influence on various model outputs. Figure 7 shows the rankings summary of the$\mathbf{P}^*$ parameters based on the PC1 μ^*.

Figure 5 and Figure 6 show the σ results for the S and the X components, respectively. Unlike the μ^* that measures significance of influence on model output, σ measures the non-linear effect of the parameter on the mode output or the interaction effect of the parameter with other model parameters. The σsignifies the possibility of significant interaction effects with other parameters, but it does not indicate which parameters are interacting. In general, most of the parameters exhibit consistent rank across the model outputs based on PC1 σ (Figure 8

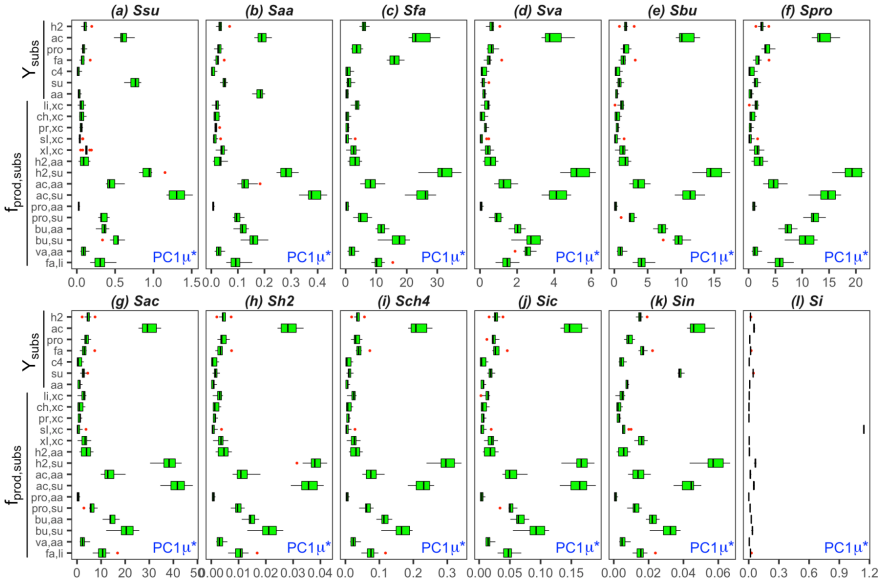


Figure 3. Mean absolute values (μ*) of the first principal component (PC1) elementary effects for the influence of the stoichiometric parameters on the solubles (S's) time-dependent outputs in ADM1. The boxplots represent the values distribution from the analyses performed using ten different seeds for random number generator that perturbed the Morris screening design

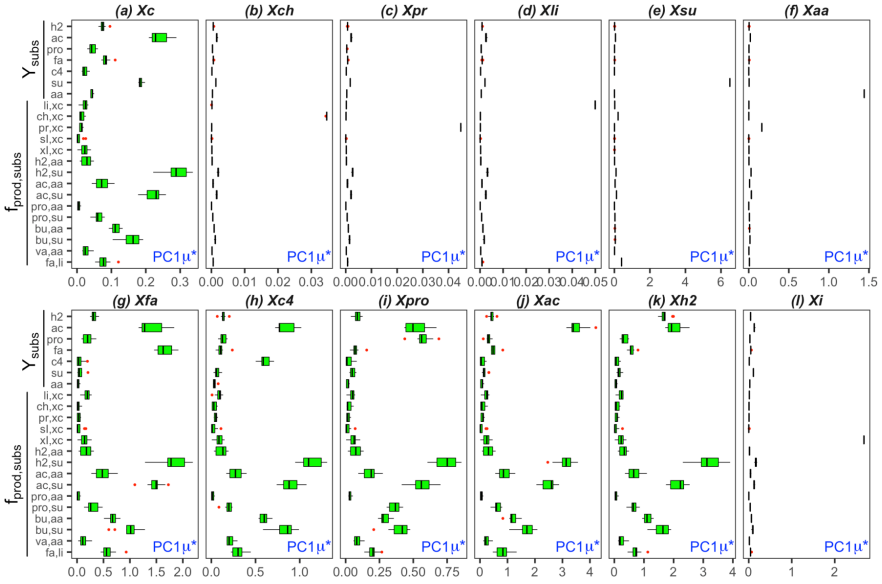


Figure 4. Mean absolute values (μ*) of the first principal component (PC1) elementary effects for the influence of the stoichiometric parameters on the particulates (X's) time-dependent outputs in ADM1. The boxplots represent the values distribution from the analyses performed using ten different seeds for random number generator that perturbed the Morris screening design.

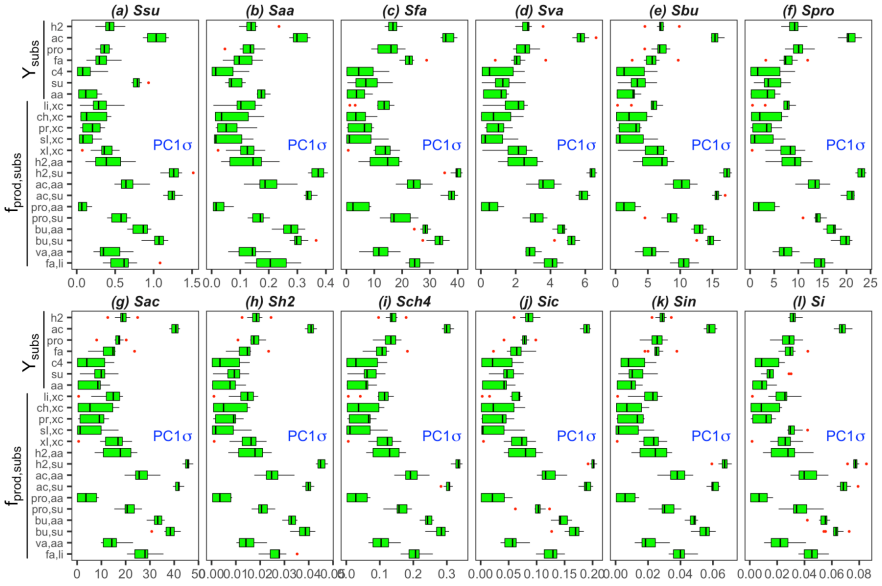


Figure 5. Standard deviations (σ) of the first principal component (PC1) elementary effects for the influence of the stoichiometric parameters on the solubles (S's) time-dependent outputs in ADM1. The boxplots represent the values distribution from the analyses performed using ten different seeds for random number generator that perturbed the Morris screening design.

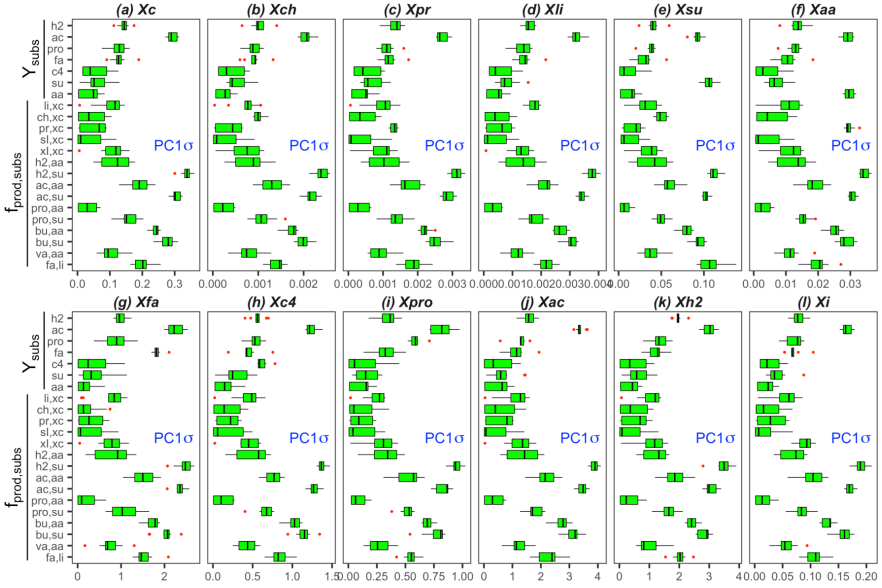


Figure 6. Standard deviations (σ) of the first principal component (PC1) elementary effects for the influence of the stoichiometric parameters on the particulates (X's) time-dependent outputs in ADM1. The boxplots represent the values distribution from the analyses performed using ten different seeds for random number generator that perturbed the Morris screening design.

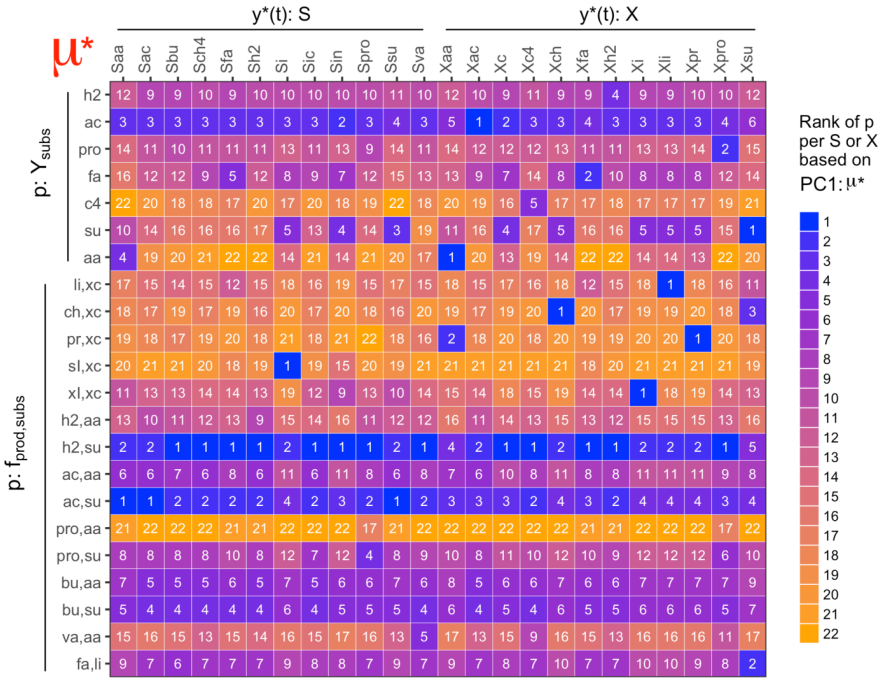


Figure 7. Ranking of the ADM1 stoichiometric parameters (p) for their influence on the time-dependent outputs (S or X) based on the overall mean absolute value (μ*) of elementary effects in the first principal component (PC1) scores. A high level of the mean (hence rank) indicates high overall effect of a parameter on a model output. High rank – low number in matrix grid.

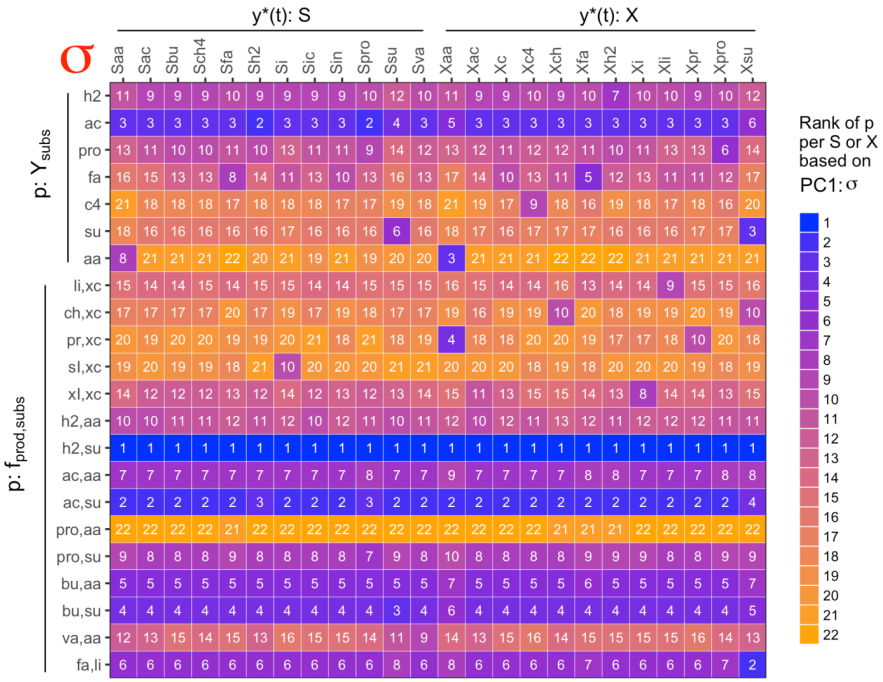


Figure 8. Ranking of the pooled standard deviations (σ) of the elementary effects of each parameter (p)

influencing the ADM1 time-dependent outputs (S or X). A relatively high standard deviation value (hence rank) indicates a parameter with significant interactions with other parameters. High rank – low number in matrix grid.

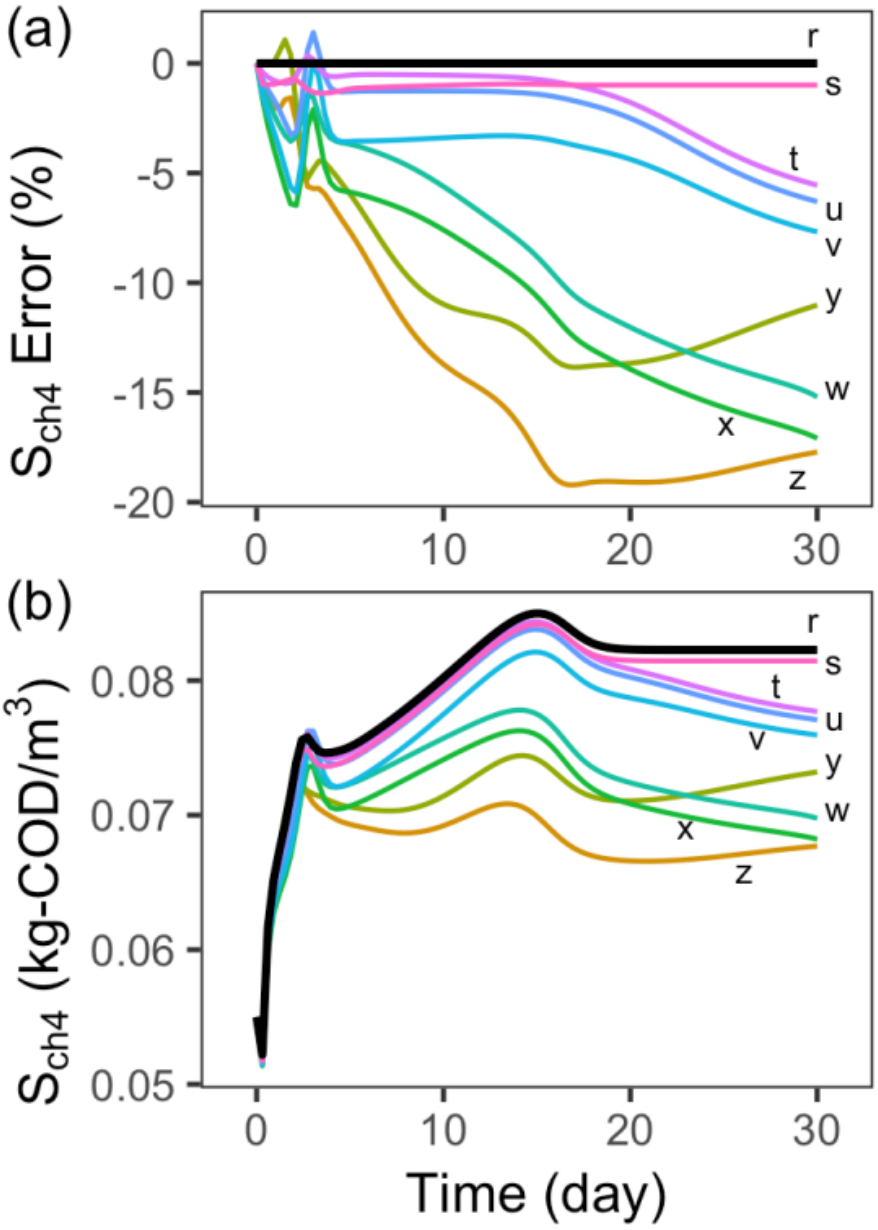
4. DISCUSSION

In this work, a data projection approach for the sensitivity analysis of time-dependent outputs in anaerobic digestion models has been demonstrated with the goal of establishing a response pattern aggregation technique to address the issue of time-varying character of SA indices. The approach integrates basis function representation via fPCA with traditional GSA (Morris' technique) to remove the time-varying character of SA indices. Removing the time-varying character of the sensitivity indices has been a key step in the interpretation of GSA results, and this has been typically achieved by calculating the grand average of the time-series indices [20]. An alternative approach proposed in this work is GSA-fPCA, which accounts for the amount of variation being captured via time-independent coefficients (PC scores) that aggregately measure the curvatures of the original time-series outputs. It is shown in this work has shown that the use of functional PCs as the projection functions for the transformation of the time-dependent variables into time-independent coefficients (PC scores) is valid for a complex digestion kinetics model such as ADM1. This projection via fPCA essentially reduces the dimensionality of the model output $y_i\,(t)$ from the n time-series samples to a single j^{th} PC score ω_{ij} in each of the$N$ responses for each simulation. The PC1 has captured most of the variations (95-99%) in the target model outputs (Figure 2) indicating that fPCA can successfully transform model outputs and aggregate curve patterns into a simple measure (ω_{ij}). This warrants the claim that fPCA may successfully reduce the dimensionality of a dataset prior to other data analysis steps such as GSA.

The ADM1 stoichiometric parameters, among all model parameters, have been chosen for this work due to their significance to the mass conversion kinetics in digestion, i.e., they indicate the mass flow profile of various organics in the digestion process. The$f_{prod,subs}$ coefficients measure the tendencies of the organics to be converted to extracellular products. The Y_{subs}coefficients, on the other hand, measure the tendencies of the organics to be assimilated by the anaerobes and be metabolically incorporated as intracellular structure. Findings about the sensitivities of variables of interest, such as methane production, to stoichiometric coefficients may show digestion pathways that significantly influence the process variables of interest. For example, consider the parameters ranking under the state variable for soluble methane (S_{ch4}) based onμ^* as shown in Figure 7. The results indicate that S_{ch4}kinetics is greatly influenced by the conversions of (ranks 1 to 5, respectively) sugars to hydrogen ($f_{h2,su}$), sugars to acetic acid ($f_{ac,su}$), acetic acid to biomass (Y_{ac}), sugars to butyric acid ($f_{bu,su}$), and amino acids to butyric acid ($f_{bu,aa}$). Interestingly, these parameters have been ranked on the same order under S_{ch4} based on the σ (Figure 8), which indicates their degree of interaction with other parameters in the model. This means that the stoichiometric parameters on whichS_{ch4} is highly sensitive are the parameters that have strong interaction with other stoichiometric parameters in the model. In a model-formulation perspective, the ranking results based on $\mu^*$are supported by the ranking results based on σ. That is, include these stoichiometric parameters (digestion pathways) in the model not only because they significantly influence the target state variable (S_{ch4}), but also because they exhibit strong interaction effects with other model parameters. The high-rank (low-numbered) parameters also imply the biological mechanisms that drive the digestion system.

As much as results of high-sensitivity parameters are important, the results on low-sensitivity parameters may also imply some key refinements on theoretical and practical aspects of digestion modelling. Consider again the parameters ranking under S_{ch4} (Figure 7). Starting from the bottom of the parameters ranking, S_{ch4} is least sensitive to the conversions of (ranks 22 up to 18, respectively) amino acids to propionic acid ($f_{pro,aa}$), amino acids to biomass (Y_{aa}), composites to soluble inerts ($f_{xI,xc}$), composites to proteins ($f_{pr,xc}$), and combined butyric and valeric acids to biomass (Y_{c4}). In modelling perspective, these least sensitive parameters may not need calibration during data fitting into the mathematical model. These can then be (1) fixed to the nominal values suggested in literature such as those in the standard ADM1 [22], or (2) discarded from the model since the underlying mechanism they describe are not so influential to the behavior

of the system. Model fitting through regression is not within the scope of this study, so approach (1) is not discussed in this work. Nonetheless, a consequence of approach (2) has been evaluated through a simple modification of the computational algorithm from the GSA work, and the results are shown in Figure 9, which shows the changes on S_{ch4} as ADM1 model parameters that are least sensitive are zeroed starting from the 22^{nd} up to the 15^{th} rank. Zeroing the model parameters means excluding the mass conversion mechanism represented by those parameters. As the parameters are eliminated, S_{ch4} is gradually being underestimated (Figure 9b). The simulation error (Figure 9a) can be interpreted as the error the regression algorithm must minimize as the remaining parameters are calibrated. This approach (2) is how model reduction is implemented, i.e., the mathematical model is modified by removing (in this case, zeroing) the factors that may be considered non-influential and the remaining model parameters are adjusted to minimize the error between the fitted values and the measurement data. The reduced model may be a more accurate representation of the system mechanisms for further study such as estimation of optimum operating conditions targeted for methane or hydrogen production. In practical aspect, model reduction has several advantages including reduction of the cost of monitoring in a process control system by eliminating unnecessary measurements. That is, if a conversion mechanism is insensitive to critical variables such as organic acids (short-chain or long-chain) concentration in the digester, the variables related to that mechanism may be ignored, hence, not measured. For example in Figure 7, all of the product yield coefficients using composites as substrate ($f_{xI,xc}$, $f_{sI,xc}$, $f_{pr,xc}$, $f_{ch,xc}$,$f_{li,xc}$, ,) are low-rank parameters when considering the organic acids concentration (S_{ac}, S_{pro},S_{bu}, S_{va}, S_{fa}), and the organic acids degraders (X_{ac}, X_{pro},X_{c4}, X_{fa}). Hence, the composites (X_c), which are the decayed anaerobic microorganisms is not critical variable for control of organic acids concentration. Approach (2) shall complement other techniques being developed towards a systematic way of simplifying complex digestion mathematical models.[8]



Notation:

r - ADM1 Full Model

s - Zeroed $f_{pro,aa}$

t - Zeroed $f_{pro,aa}$, Y_{aa}

u - Zeroed $f_{pro,aa}$, Y_{aa}, $f_{sl,xc}$

v - Zeroed $f_{pro,aa}$, Y_{aa}, $f_{sl,xc}$, $f_{pr,xc}$

w - Zeroed $f_{pro,aa}$, Y_{aa}, $f_{sl,xc}$, $f_{pr,xc}$, Y_{c4}

x - Zeroed $f_{pro,aa}$, Y_{aa}, $f_{sl,xc}$, $f_{pr,xc}$, Y_{c4}, $f_{ch,xc}$

y - Zeroed $f_{pro,aa}$, Y_{aa}, $f_{sl,xc}$, $f_{pr,xc}$, Y_{c4}, $f_{ch,xc}$, Y_{su}

z - Zeroed $f_{pro,aa}$, Y_{aa}, $f_{sl,xc}$, $f_{pr,xc}$, Y_{c4}, $f_{ch,xc}$, Y_{su}, $f_{li,xc}$

$$\text{Error} = 100\% * \frac{S_{ch4,s...z} - S_{ch4,r}}{S_{ch4,r}}$$

Figure 9. Sample results on the ADM1 model reduction by successive elimination (zeroing) of the stoichiometric parameters under soluble methane state variable starting from the least sensitive parameter going up the ranking (s:22$^{\text{nd}}$ up to z:15$^{\text{th}}$rank).

An interesting comparative view of parameter influences can be made when the sensitivities of all the target biochemical state variables on the stoichiometric coefficients are calculated, ranked, and summarized according to Figure 7 and Figure 8. There are parameters that dominantly influence most of the state variables (Figure 7): $f_{h2,su}$,$f_{ac,su}$, Y_{ac}, $f_{bu,su}$, $f_{bu,aa}$,$f_{fa,li}$, etc. These same parameters are significantly interacting with other stoichiometric parameters in almost all of the state variables based on the rankings in Figure 8. There are some parameters that are very influential on the certain state variables. Take for example the parameter for conversion of amino acids to biomass (Y_{aa}), which is high-rank parameter under soluble amino acids S_{aa} and under amino acid degraders (X_{aa}), but is a low-rank parameter in other state variables. This is reasonable since they are directly related. That is, if the Y_{aa} is low (allowed to fluctuate from lower bound to upper bound levels during GSA) then there are lower amounts of amino acids degraded and assimilated as biomass resulting to lowerX_{aa}, and higher S_{aa}. The reverse of this relation is also true, i.e., higher Y_{aa} means higherX_{aa} and lower S_{aa}. This apparent diversion and conservation of mass is due to the imposed balancing of COD (kg-COD/m^3) in the ADM1, which has been standardized on the mass balance of organics (as COD) in digestion. The above discussions on how the results may be interpreted and used aim to put into perspective the potential of the GSA-fPCA methodology in broadening the techniques on SA for anaerobic digestion models. SA has always been integral in digestion kinetics modelling and parameter estimation.

The GSA method (Morris') used in this work is one of many that can be used for ADM1 and other bioprocessing kinetics models. Morris' technique, however, has been tested as a practical approach for parameters SA of bioprocess models that include many parameters that are nonlinearly related to the variables of interest [21]. It has also been used for GSA of models on chemical reaction for dimethylsulphide (DMS), which is a gas involved in climate change [40]; on human insulin signaling pathway [21]; on urban water supply yield [44]; and on whole-year dynamics of grasslands [45]. These studies have shown that Morris' technique is at par with other GSA methods. An aspect of the Morris' technique that has been emphasized in this work is the randomization of the parameter sampling trajectories by setting unique random number generator seed for each set of r elementary effects. This randomization improves the coverage of the trajectories in the sampling space of all the parameters under study. This ensures that many combinations of parameter levels have been accounted, hence, model responses have exhibited multitudes of possible patterns. Though this work focused only on Morris' technique as the GSA method, others may be amenable to integration with fPCA as demonstrated by Sumner, Shephard and Bogle [21] on integrating fPCA with the Sobol method. Though Morris' requires fewer model evaluations, it is still at par with Sobol method which requires significantly more model evaluations (In the work by Sumner, Shephard and Bogle [21]: 15min for Morris' and 1.06 days for Sobol to achieve the same level analysis). Integrating fPCA with previously explored GSA for biological wastewater treatment models such as Extended-FAST and SRC may also worth considering.

Using GSA-fPCA approach on ADM1 as demonstrated in this work, other SA problems related to digestion kinetics may be addressed via similar computational technique. For example, the kinetic and physicochemical parameters such as rate constants, inhibition constants, temperature, and pH may be subject to GSA for their influence on the time-dependent outputs. Moreover, though the focus of the work has been on parameter sensitivities, the GSA-fPCA approach may also be used for sensitivities on initial conditions, which are properties of the incoming feedstock stream in digestion. That is, allow the initial values of the S's and the X's to fluctuate within specified ranges and evaluate their elementary effects on their time-series values (transformed via fPCA). This is a reasonable analysis problem when testing for the robustness of a digestion setup with high chances of variations in the incoming organic feedstock. Examples of highly-varied feedstocks are organic wastes that are not homogenized in bulk amounts prior to digestion, and feedstocks that temporally vary due to external factors. Digestion stability analysis has been the concern in some applications of digestion models and actual setups [35]. The GSA-fPCA methodology may also be implemented to systems that include digestion-related processes such as activated sludge process in wastewater treatment.

A series of standard mathematical models as rigorous as the ADM1 related to wastewater are the Activated Sludge Models (ASM1, ASM2, ASM2d, and ASM3). These have been implemented as ADM1-coupled mathematical models to simulate and study aerobic sludge-asnaerobic digestion systems [46-48]. Evaluating them within the context of SA by subjecting them under GSA-fPCA approach may elucidate patterns only obvious under a systematic GSA methodology.

5. CONCLUSIONS

A comprehensive GSA on a complex and computationally demanding mathematical model can systematically elucidate pertinent model behaviors. In this work, the combination of the conventional GSA Morris' screening technique with functional PCA has been implemented to evaluate the influences of stoichiometric parameters on the state variables in the standard model ADM1 describing anaerobic digestion kinetics. The GSA-fPCA eliminates the time-varying character of sensitivity indices allowing improved interpretability of GSA indices and accounting for the amount of variations being captured for indices calculations. Functional PCA can capture most of the variabilities of the time-dependent model responses (95-99%) through the first few PCs resulting into a reliable responses projection into time-independent PC scores. Using these PC scores for elementary effects evaluation that are eventually aggregated through statistics essentially determines global sensitivities of the PC scores, hence model responses, to the fluctuations of stoichiometric parameters. The ranking of stoichiometric sensitivities can be used for various purposes including driving mechanisms identification, and mathematical model modification. This GSA-fPCA approach may be extended to other bioprocesses modeled as time-dependent dynamical kinetics to perform comprehensive numerical sensitivity analysis of model parameters. Depending on the application, these models are manipulated to be fast responding models, which usually are the simple models, preferred in process control; or to be complex models preferred during process and system design; or to be at varying model complexities such as those typically encountered during parameter calibration with empirical data. In any of these cases, GSA-fPCA offers the advantage of calculating time-independent sensitivity indices amenable to ranking of model parameters while being able to account for the amount of variations of the original time-dependent model outputs being captured by the first few PCs as basis function projections.

ACKNOWLEDGEMENTS

Partial funding for the research presented in this paper was provided by NASA-ESPCoR, the Louisiana Board of Regents, Louisiana Space Grant/LaSPACE (Baton Rouge, LA 70803, USA), and Cleco (Pineville, LA 71361, USA). Appreciation is felt towards the staff and students of the Energy Institute of Louisiana.

Conflict of Interests : Authors have no conflict of interest relevant to this article.

APPENDIX

See supporting information (SI) for details of the GSA-fPCA on ADM1 model and for samples of intermediate results.

NOMENCLATURE

S_{su} Monosaccharides

S_{aa} Amino acids

S_{fa} Long chain fatty acids

S_{va} Total valerate

S_{bu} Total butyrate

S_{pro} Total propionate

S_{ac} Total acetate

S_{h2} Soluble hydrogen

S_{ch4} Soluble methane

S_{IC} Inorganic carbon

S_{IN} Inorganic nitrogen

S_I Soluble inerts

X_c Composites

X_{ch} Carbohydrates

X_{pr} Proteins

X_{li} Lipids

X_{su} Sugar degraders

X_{aa} Amino acid degraders

X_{fa} LCFA degraders

X_{c4} Valerate & butyrate degradersX_{pro} Propionate degraders

X_{ac} Acetate degraders

X_{h2} Hydrogen degraders

X_I Particulate inerts

$f_{prod,subs}$ product yield coefficients from specific substrates ('dimensionless)Y_{subs}, biomass yield coefficients from specific substrates (kg-COD/kg-COD)

REFERENCES

[1] R.E. Speece, *Anaerobic biotechnology for industrial wastewater treatment* , Environmental Science & Technology 17 (1983), pp. 416A-427A.

[2] K.A. Taconi, M.E. Zappi, W. Todd French, and L.R. Brown, *Feasibility of methanogenic digestion applied to a low pH acetic acid solution* , Bioresource Technology 98 (2007), pp. 1579-1585.

[3] M.G. Mengistu, B. Simane, G. Eshete, and T.S. Workneh, *A review on biogas technology and its contributions to sustainable rural livelihood in Ethiopia* , Renewable and Sustainable Energy Reviews 48 (2015), pp. 306-316.

[4] K.V. Rajeshwari, M. Balakrishnan, A. Kansal, K. Lata, and V.V.N. Kishore,*State-of-the-art of anaerobic digestion technology for industrial wastewater treatment* , Renewable and Sustainable Energy Reviews 4 (2000), pp. 135-156.

[5] N. Kythreotou, G. Florides, and S.A. Tassou, *A review of simple to scientific models for anaerobic digestion* , Renewable Energy 71 (2014), pp. 701-714.

[6] *Control Strategy Based on Sensitivity Analysis. Application for Anaerobic Digestion Bioprocess* .

[7] D.J. Batstone,*Mathematical Modelling of Anaerobic Reactors Treating Domestic Wastewater: Rational Criteria for Model Use* , Reviews in Environmental Science and Bio/Technology 5 (2006), pp. 57-71.

[8] S. Hassam, E. Ficara, A. Leva, and J. Harmand, *A generic and systematic procedure to derive a simplified model from the anaerobic digestion model No. 1 (ADM1)* , Biochemical Engineering Journal 99 (2015), pp. 193-203.

[9] N.M. Atallah, M. El-Fadel, S. Ghanimeh, P. Saikaly, and M. Abou-Najm, *Performance optimization and validation of ADM1 simulations under anaerobic thermophilic conditions* , Bioresource Technology 174 (2014), pp. 243-255.

[10] F. Pianosi, K. Beven, J. Freer, J.W. Hall, J. Rougier, D.B. Stephenson, and T. Wagener,*Sensitivity analysis of environmental models: A systematic review with practical workflow* , Environmental Modelling & Software 79 (2016), pp. 214-232.

[11] A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo, *Sensitivity Analysis for Chemical Models* , Chemical Reviews 105 (2005), pp. 2811-2828.

[12] I. Havlik, J. Votruba, M. Sobotka, and B. Volesky, *Parametric sensitivity in modeling of anaerobic digestion* , Biotechnology Letters 6 (1984), pp. 607-610.

[13] N.A. Noykova, and M. Gyllenberg, *Sensitivity analysis and parameter estimation in a model of anaerobic waste water treatment processes with substrate inhibition* , Bioprocess Engineering 23 (2000), pp. 343-349.

[14] O. Lidholm, and E. Ossiansson, *Modeling anaerobic digestion: Validation and calibration of the Siegrist model with uncertainty and sensitivity analysis* . Lund University, 2008.

[15] A. Abusam, and K.J. Keesman,*Sensitivity analysis of the secondary settling tank double-exponential function model* , European Water Management Online (2002), pp. 1-9.

[16] G. Sin, K.V. Gernaey, M.B. Neumann, M.C.M. van Loosdrecht, and W. Gujer, *Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty* , Water Research 45 (2011), pp. 639-651.

[17] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis - The Primer* , John Wiley & Sons, Ltd, 2008.

[18] A. Cosenza, G. Mannina, P.A. Vanrolleghem, and M.B. Neumann, *Global sensitivity analysis in wastewater applications: A comprehensive comparison of different methods* , Environmental Modelling & Software 49 (2013), pp. 40-52.

[19] K. Campbell, M.D. McKay, and B.J. Williams, *Sensitivity analysis when model outputs are functions* , Reliability Engineering & System Safety 91 (2006), pp. 1468-1472.

[20] A. Zhu, J. Guo, B.-J. Ni, S. Wang, Q. Yang, and Y. Peng, *A Novel Protocol for Model Calibration in Biological Wastewater Treatment* , Scientific Reports 5 (2015), p. 8493.

[21] T. Sumner, E. Shephard, and I.D.L. Bogle, *A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modelling* , Journal of the Royal Society Interface 9 (2012), pp. 2156-2166.

[22] D.J. Batstone, J. Keller, I. Angelidaki, S.V. Kalyuzhnyi, S.G. Pavaostathis, A. Rozzi, W.T.M. Sanders, H. Seigrist, and V.A. Vavilin, *The IWA Anaerobic Digestion Model No. 1 (ADM1)* , IWA Publishing, London, UK, 2002.

[23] F. Blumensaat, and J. Keller, *Modelling of two-stage anaerobic digestion using the IWA Anaerobic Digestion Model No. 1 (ADM1)* , Water Research 39 (2005), pp. 171-183.

[24] C. Mendes, K. Esquerre, and L. Matos Queiroz, *Application of Anaerobic Digestion Model No. 1 for simulating anaerobic mesophilic sludge digestion* , Waste Management 35 (2015), pp. 89-95.

[25] J. Bai, H. Liu, B. Yin, H. Ma, and X. Chen, *Modified ADM1 for modeling free ammonia inhibition in anaerobic acidogenic fermentation with high-solid sludge* , Journal of Environmental Sciences 52 (2017), pp. 58-65.

[26] A. Galí, T. Benabdallah, S. Astals, and J. Mata-Alvarez, *Modified version of ADM1 model for agro-waste application* , Bioresource Technology 100 (2009), pp. 2783-2790.

[27] S. Weinrich, and M. Nelles,*Critical comparison of different model structures for the applied simulation of the anaerobic digestion of agricultural energy crops* , Bioresource Technology 178 (2015), pp. 306-312.

[28] X. Chen, Z. Chen, X. Wang, C. Huo, Z. Hu, B. Xiao, and M. Hu, *Application of ADM1 for modeling of biogas production from anaerobic digestion of Hydrilla verticillata* , Bioresource Technology 211 (2016), pp. 101-107.

[29] D. Poggio, M. Walker, W. Nimmo, L. Ma, and M. Pourkashanian, *Modelling the anaerobic digestion of solid organic waste – Substrate characterisation method for ADM1 using a combined biochemical and kinetic parameter estimation approach* , Waste Management 53 (2016), pp. 40-54.

[30] E. Nordlander, E. Thorin, and J. Yan, *Investigating the possibility of applying an ADM1 based model to a full-scale co-digestion plant* , Biochemical Engineering Journal 120 (2017), pp. 73-83.

[31] S. García-Gen, J.M. Lema, and J. Rodríguez, *Generalised modelling approach for anaerobic co-digestion of fermentable substrates* , Bioresource Technology 147 (2013), pp. 525-533.

[32] S. García-Gen, P. Sousbie, G. Rangaraj, J.M. Lema, J. Rodríguez, J.-P. Steyer, and M. Torrijos,*Kinetic modelling of anaerobic hydrolysis of solid wastes, including disintegration processes* , Waste Management 35 (2015), pp. 96-104.

[33] E. Lee, J. Cumberbatch, M. Wang, and Q. Zhang, *Kinetic parameter estimation model for anaerobic co-digestion of waste activated sludge and microalgae* , Bioresource Technology 228 (2017), pp. 9-17.

[34] V. Razaviarani, and I.D. Buchanan, *Calibration of the Anaerobic Digestion Model No. 1 (ADM1) for steady-state anaerobic co-digestion of municipal wastewater sludge with restaurant grease trap waste* , Chemical Engineering Journal 266 (2015), pp. 91-99.

[35] M.E. Demitry,*Anaerobic Digestion Process Stability and the Extension of the ADM1 for Municipal Sludge Co- Digested with Bakery Waste* . Doctor of Philosophy, Utah State University, 2016.

[36] M.E. Demitry, C. Hansen, D. Stevens, and M. McFarland, *Extending the Applications of the ADM1 to Predict Performance of the Induced Bed Reactor (IBR) Co-Digesting Municipal Sludge with Bakery Waste* , Environment and Pollution 5 (2016), pp. 41-56.

[37] M.E. Demitry, J. Zhong, C. Hansen, and M. McFarland, *Modifying the ADM1 Model to Predict the Operation of an Anaerobic Digester Co-digesting Municipal Sludge with Bakery Waste* , Environment and Pollution 4 (2015), pp. 38-57.

[38] M.D. Morris, *Factorial Sampling Plans for Preliminary Computational Experiments* , Technometrics 33 (1991), pp. 161-174.

[39] G. Sin, and K.V. Gernaey,*Improving the Morris method for sensitivity analysis by scaling the elementary effects* , in *Computer Aided Chemical Engineering* , J. Jeżowski and J. Thullie eds., Elsevier, 2009, pp. 925-930.

[40] F. Campolongo, J. Cariboni, and A. Saltelli, *An effective screening design for sensitivity analysis of large models* , Environmental Modelling & Software 22 (2007), pp. 1509-1518.

[41] G. Pujol, B. Iooss, A. Janon, K. Boumhaout, S.D. Veiga, T. Delage, J. Fruth, L. Gilquin, J. Guil-, laume, L.L. Gratiet, P. Lemaitre, B.L. Nelson, F. Monari, R. Oomen, B. Ramos, O. Roustant, E. Song, J. Staum, T. Touati, and F. Weber, *'sensitivity': Global Sensitivity Analysis of Model Outputs* . Repository: cran.r-project.org; software available at: https://CRAN.R-project.org/package=sensitivity

[42] K. Soetaert, T. Petzoldt, and R.W. Setzer, *'deSolve': Solvers for Initial Value Problems of Differential Equations ('ODE', 'DAE', 'DDE')* . Repository: cran.r-project.org; software available at: https://cran.r-project.org/web/packages/deSolve/index.html

[43] J.O. Ramsay, H. Wickham, S. Graves, and G. Hooker, *'fda': Functional Data Analysis* . Repository: cran.r-project.org; software available at: https://cran.r-project.org/web/packages/fda/index.html

[44] D.M. King, and B.J.C. Perera, *Morris method of sensitivity analysis applied to assess the importance of input variables on urban water supply yield – A case study* , Journal of Hydrology 477 (2013), pp. 17-32.

[45] S.M. Thorsen,*Screening groups of parameters using the Morris method to select parameters for variance based sensitivity analysis* , Procedia - Social and Behavioral Sciences 2 (2010), pp. 7755-7756.

[46] J. Alex, L. Benedetti, J. Copp, K.V. Gernaey, U. Jeppsson, I. Nopens, M.-N. Pons, L. Rieger, C. Rosen, J.P. Steyer, P. Vanrolleghem, and S. Winkler, *Benchmark Simulation Model No. 1(BSM1)* , Lund University, Lund, Sweden, 2008.

[47] D. Vrecko, K.V. Gernaey, C. Rosen, and U. Jeppsson, *Benchmark Simulation Model No 2 in Matlab-Simulink: towards plant-wide WWTP control strategy evaluation* , Water Science and Technology 54 (2006), pp. 65-72.

[48] U. Jeppsson, C. Rosen, J. Alex, J. Copp, K.V. Gernaey, M.-N. Pons, and P.A. Vanrolleghem, *Towards a benchmark simulation model for plant-wide control strategy performance evaluation of WWTPs* , Water Science and Technology 53 (2006), pp. 287-295.