Water Research 171 (2020) 115377

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Self-optimizing attainable regions of the anaerobic treatment process: Modeling performance targets under kinetic uncertainty

F. Abunde Neba [a, b, c, *], Hoese M. Tornyeviadzi [a, c], Stein W. Østerhus [b], Razak Seidu [c]

[a] Abunde Sustainable Engineering Group, AbundeSEG, Ghana
[b] Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, Trondheim, Norway
[c] Department of Marine Operations and Civil Engineering, Norwegian University of Science and Technology, Ålesund, Norway

ARTICLE INFO

Article history:
Received 2 August 2019
Received in revised form
5 December 2019
Accepted 6 December 2019
Available online 10 December 2019

Keywords:
Self-optimizing attainable regions
Practical identifiability
Uncertainty quantification
Anaerobic digester synthesis

ABSTRACT

Despite the advantage of model-based design, anaerobic digesters are seldom designed using biokinetic models due to lack of reliable kinetic coefficients and/or systematic approaches for incorporating kinetic models into digester design. This study presents a systematic framework, which couples practical identifiability, uncertainty quantification and attainable region (AR) concepts for defining process performance targets, especially when reliable kinetic coefficients are unavailable. Within the framework, we introduce the concept of self-optimizing ARs, which define performance targets that results in near optimal operation in spite of variations in kinetic coefficients. Using the case of modified Hill model, only 3 out of the 6 model parameters (unidentifiable set) are responsible for the model prediction uncertainty. The uncertainty bands (mean, 10th percentile and 90th percentile) on the model states has been computed using the Monte Carlo Simulation procedure and attainable regions for the different levels of uncertainty has been constructed and the boundaries interpreted into digester structures. The self-optimizing attainable regions have been defined as the intersection region of the attainable regions corresponding to the mean, 10th percentile and 90th percentile. Incorporating uncertainty significantly reduces performance targets of the process but increases self-optimality in defining performance targets. Unlike the attainable region, which represents the limits of achievability for defined kinetics, the self-optimizing attainable region represents the set of all possible states attainable by the system even in cases of kinetic uncertainty. In summary, the concept of self-optimizing ARs provides a systematic way of defining process performance targets and making economic decisions under conditions of uncertainty.

1. Introduction

Anaerobic digestion (AD) of animal manure is of great importance to the waste treatment and bioenergy industries, since the biomethane produced is a promising renewable energy alternative to fossil fuels. The modelling of anaerobic treatment process is a mature research area, now with a strong shift from model development towards application development, aimed at solving various design and operational challenges. Various models have been constructed to describe the anaerobic treatment process and the key motivations for model development have mainly been operational analysis, technology development, as well as digester design (Batstone, 2006). The model-based design is particularly important as the capital cost for anaerobic digesters determined from design is a key motivation for implementers. The kinetics captured by AD models is highly important for an optimal digester design since operating conditions, volumetric gas production, process stability (Finn et al., 2013; Yu et al., 2013; Batstone, 2006; Kythreotou et al., 2014), as well as effluent quality can be predicted (Kythreotou et al., 2014; Yu et al., 2013). Despite the advantage of model-based design, anaerobic digesters are seldom designed using biokinetic models but rather based on a combination of hydraulic and organic loading, where the digester capacity is determined for a given loading rate, temperature regime, mixing, etc. (Wang et al., 2007). This is because the use of biokinetic models is highly dependent upon availability of kinetic coefficients (Batstone, 2006; Wang et al., 2007), but it is often difficult to get reliable kinetic parameters in practical operation, which results in kinetic uncertainty (and hence uncertain process performance) if the models are used for digester

* Corresponding author. Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, Trondheim, Norway.
E-mail address: fabricen@stud.ntnu.no (F. Abunde Neba).

https://doi.org/10.1016/j.watres.2019.115377

design. Summarily, three main challenges can be mentioned with the use of models to design anaerobic digesters: (1) Lack of systematic approaches for incorporating process kinetics in digester design. (2) The reliability of some of the studies using kinetic models to guide design of anaerobic digesters is undermined by uncertainty existing in the in kinetic coefficients. (3) Existing model-based studies are limited to single stage digesters and operating the process as a single stage generally limits overall performance. This is supported by the fact that AD involves multiple bioreactions steps (each step catalyzed by a specific group of microorganisms) and when operated as a single stage, it limits the possible combination of pathways since the process conditions are only suitable for all microorganisms with no reaction being optimized (EPA, 2006). However, most of the modeling studies on AD have focused on the model development approach, techniques for parameter estimation, with less effort devoted to assessing model reliability (identifiability and uncertainty) or how to incorporate uncertainty in digester design and operation. A systematic approach for handling kinetic uncertainty in design of anaerobic digesters with focus on multi-stage anaerobic digesters networks as opposed to single stage systems will be a breakthrough in advancing model-based optimization of anaerobic digestion.

This study is therefore designed to develop a systematic model-based framework (Fig. 1) for performance targeting and synthesis of anaerobic digester networks, when reliable kinetic models are not available. The framework (Fig. 1) is realized in two phases, which may involve feedback checks at specific steps depending on the system performance after every step. In the first phase (model reliability assessment), one selects a kinetic model appropriate for a digested substrate of interest (e.g. solid waste, sludge, wastewater, etc.), assesses the model's reliability and quantifies the model prediction uncertainty resulting from kinetic uncertainty. In the second phase (self-optimizing design), one mainly defines the robust performance targets of the system considering the uncertainty bounds computed in phase 1. Phase 2 is based on the concept of attainable regions, which is a geometric optimization technique that is used for both performance targeting and reactor network synthesis (Hildebrandt et al., 1990). The AR is a collection of all possible output for all possible reactor designs by interpreting chemical processes as geometric objects that define a region of achievability without having to explicitly enumerate all possible design combinations (Hildebrandt and Glasser, 1990). Central to the AR concept is the availability of reliable kinetic models of all fundamental processes (e.g. biochemical, physicochemical, physical in the case of AD) occurring within the system. In particular, simplified kinetic process models are emphasized as the AR theory involves mixing and attainability of states through a relatively complex geometric and hydrodynamic analysis (Hildebrandt et al., 1990; Ming et al., 2016).

The novel idea presented in this study is that instead of using AR to define an optimal performance target, which can only be achieved some of the times (due to kinetic uncertainty), the authors define a near optimal performance target, which can be attained all the time. This however involves an acceptable loss in process performance resulting from the kinetic uncertainty. Anaerobic digester systems designed with such an acceptable loss in performance resulting from uncertainty in kinetic coefficients are referred to as self-optimizing. Self-optimizing operation also referred to as self-optimizing design/systems (Permin et al., 2016; Gausemeier et al., 2006) is when we can achieve an acceptable loss by using constant setpoint values for design/operation variables (e.g. temperature, kinetics, substrate characteristics, etc., for the case of anaerobic digestion) without the need to reoptimize when variations occur. In the case of this study, we define self-optimizing

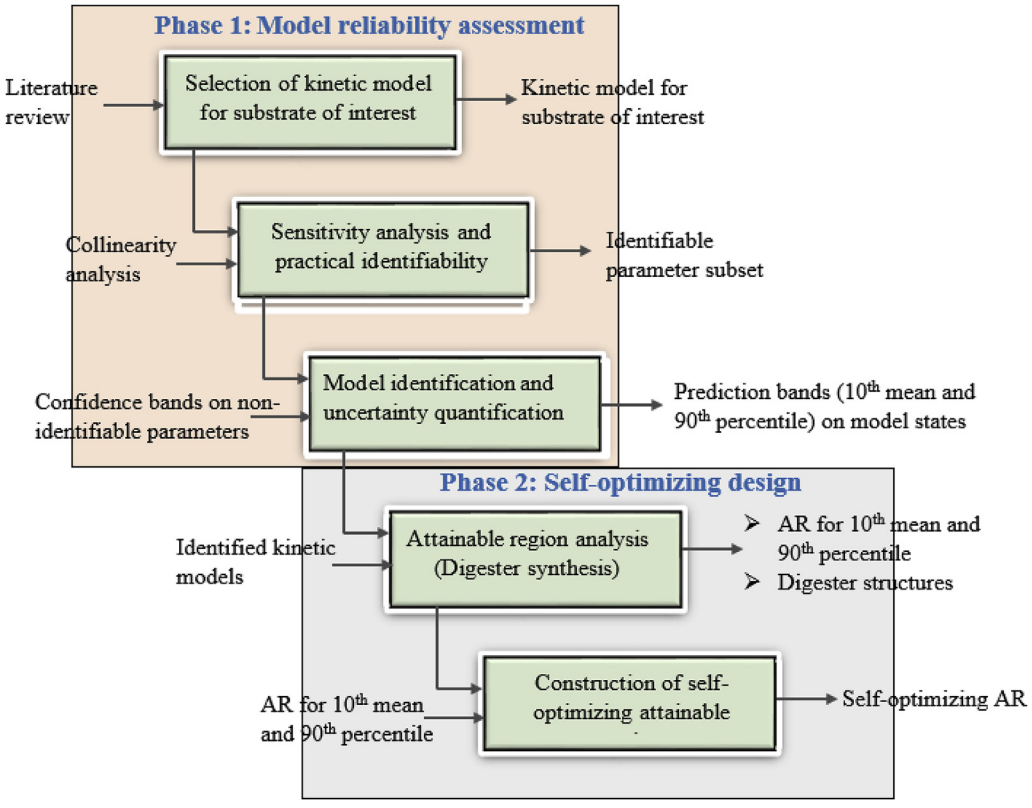

Fig. 1. Two-phase framework for model-based synthesis of methane bioreactors.

operation as an attainable region or performance target that results in near optimal operation despite variations in kinetic coefficients of the process. In the context of process engineering, a similar concept has been applied to plant wide control, known as self-optimizing control, characterized by the choice of self-optimizing controlled variables (Skogestad, 2000; Jäschke et al., 2017).

In order to illustrate the applicability of the framework presented in Fig. 1, the modified Hill model published by (Finn et al., 2013) was selected as a case study. The model considers substrate effects and applies to anaerobic digestion of animal manure (diary, poultry, beef or swine wastes) and predicts acidogens, methanogens, organic substrate and volatile acids. In addition, the model eliminates the need for factors such as alkalinity, concentration of cation, dissolved CO_2 and ammonia gas because their effect is already lumped into two important parameters found in the model, the biodegradability constant (B_0) and acidity factor (A_f).Even though the modified Hill model has been selected, it is important for readers to note that the major contribution of this study is the development of a systematic framework, which couples practical identifiability, uncertainty quantification and attainable region (AR) concepts for defining process performance targets and synthesizing anaerobic digester networks, especially when reliable kinetic coefficients are unavailable. The framework can be used for any other dynamic model selected to describe the kinetics of the anerobic treatment process.

Our recent studies have been first to illustrate the usefulness of AR to define performance targets and model digester configurations that optimize methane productivity and volatile solids reduction (Abunde Neba et al., 2019c), as well as stability of methanogenic archaea (Abunde Neba et al., 2019b). Both studies put together have illustrated that a change in the kinetic model structure or value of kinetic coefficients, induced by differences in substrate and inoculum characteristics significantly influences the performance target as well as the optimal digester configuration required to achieve the target. In another recent study by the authors, a framework was developed and embedded into a software for using simplified microbial kinetic models for AR analysis in cases where data requirements are limited (Abunde Neba et al., 2020a). The integration of economic feasibility indicators (such as payback period, benefit cost ratio, net present value and internal rate of returns) with attainable region analysis has also been presented by the authors, which is very interesting for synthesizing digester structures based on economic objectives (Abunde Neba et al., 2019a). Finally, another approach, which only relies on experimental data (no model required) is developed by coupling attainable regions and fuzzy multicriteria decision making for selection of digester subunits and synthesis of digester network configurations (Abunde Neba et al., 2020b). It is interesting to mention at this point that unlike other model-based studies on AR, which assume that the kinetic coefficients of a model are known before constructing the attainable regions, this study is novel in that it rather quantifies the uncertainty in the kinetic coefficients and propagates it onto the attainable regions. The authors call such regions 'self-optimizing attainable regions', because they will always be attained even if variations occur in kinetic coefficients. Once the AR is obtained, its boundary can always be interpreted into digester structures, which can be used for industrial operation in order to achieve the performance target defined by the region.

2. Theoretical concepts and methods description

2.1. Model reliability assessment

The reliability of a mechanistic model has to do with the degree of uncertainty (the confidence band) of its model parameters and it is influenced by three main factors (Sin et al., 2009, 2010a): (1) the mathematical structure of the model, (2) the nature of the experimental data used for identification, and (3) the set of model parameters used in the identification process. In this paper, the focus is on analyzing the relation amongst model structure (factor 1), identifiable set of parameters (factor 2) and reliability of anaerobic digestion model although the discussion of the results is extended to also reflect on the impact of the information content in the experimental data (factor 2).

Given a kinetic model for a process, we define the following three key steps needed to completely assess the reliability and usage of the model:

Step 1: Perform a sensitivity-based identifiability to determine the identifiable set of model parameters
Step 2: Estimate the identifiable set of model parameters and quantify the confidence band
Step 3: Quantify the model prediction (output) uncertainty using the unidentifiable parameter set as inputs

For studying the identifiability of the biokinetic models, the sensitivity and collinearity analysis are used. For parameter estimation, the method of first order gradients, with gradients computed using the discrete adjoint method, was used and 95% joint and marginal confidence regions were used to assess the identifiability following parameter estimation. For the input-output uncertainty analysis, the Monte Carlo simulation procedure was used.

The objective of this section is to analyze the aforementioned necessary steps with respect to its application to the anaerobic treatment process. However, the analysis requires that the process model is known, and we therefore begin by describing the model of the anaerobic treatment process.

2.1.1. Model selection and description

A number of simplified state-space dynamic models for the anaerobic digestion process have been reviewed by Finn et al. (2013). The modified Hill model which was developed for anaerobic digestion of animal manure (diary, poultry, beef and swine wastes) was selected for this study. The model lumps the effect of hydrolysis, alkalinity, cation concentration, dissolved carbon dioxide and ammonia into two important constants, the biodegradability constant (Bo) and acidity factor (AF) present in the modified Hill model. The Hill model is a mechanistic (model parameters have a physical meaning), which makes it interesting to understand the identifiability characteristics of the model. The identifiability characteristics of a model relates to set of parameters to be estimated in order to accurately describe the observed mechanisms described by the model (Donoso-Bravo et al., 2013) An "over-calibrated" model would accurately describe/fit experimental data but would lose its capability to predict(Donoso-Bravo et al., 2011), which weakens the model's reliability and hence applicability for design purposes.

Fig. 2 presents an illustration of the model by showing the flow of information between four compartments in the methane bioreactor, which include inoculum, substrate, liquid phase and gas phase.

The species conservation and biogas production equations for the modified Hills model is presented as follows

a) Biodegradable volatile solids (S_1) in the liquid phase of the bioreactor

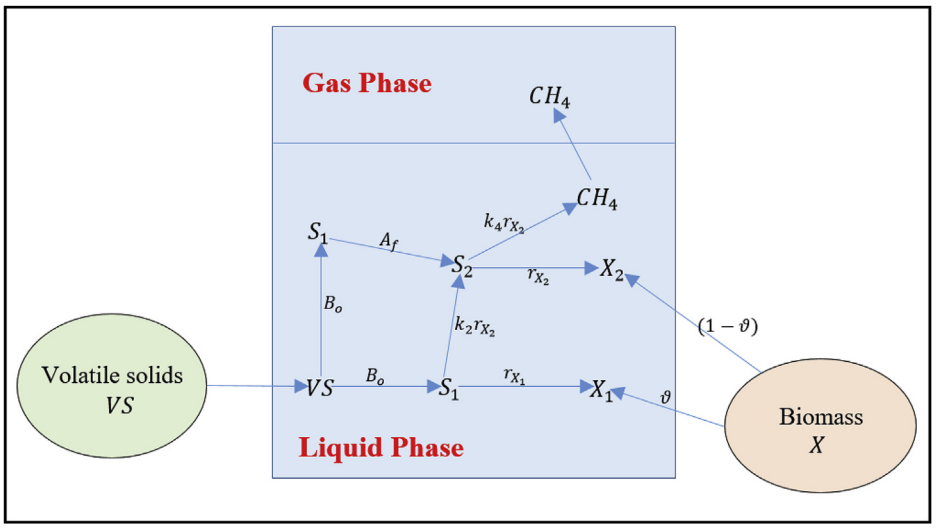

Fig. 2. Information flow of the modified Hill model.

$$\frac{dS_1}{dt} = (S_{1_{in}} - S_1)D - k_1\mu_1 X_1 \quad (1)$$

b) Volatile fatty acids (S_2) in the liquid phase of the bioreactor

$$\frac{dS_2}{dt} = (S_{2_{in}} - S_2)D + k_2\mu_1 X_1 - k_3\mu_2 X_2 \quad (2)$$

c) Acidogens (X_1) in the liquid phase of the bioreactor

$$\frac{dX_1}{dt} = (\mu_1 - K_{d_1} - D)X_1 \quad (3)$$

d) Methanogens (X_2) in the liquid phase of the bioreactor

$$\frac{dX_2}{dt} = (\mu_2 - K_{d_2} - D)X_2 \quad (4)$$

e) Methane gas flow rate

$$Q_{CH_4} = V\mu_2 k_4 X_2 \quad (5)$$

The organic waste is characterized by using the two parameters, which are biodegradability (B_o), Eq. (6) and acidity (A_f), Eq. (7). In the modified Hill model, B_o measures the ease with which the organic substrate can be broken down and stabilized by anaerobic bacteria while A_f of a substrate can be defined as the amount of volatile fatty acids contained in the substrate per unit mass of biodegradable volatile solids

$$S_{1_{in}} = B_0 S_{in} \quad (6)$$

$$S_{2_{in}} = A_f S_{1_{in}} \quad (7)$$

In the modified Hill model the anaerobic biodegradability can be computed via Eq. (8) while the acidity factor is computed using Eq. (9).

$$B_0 = \frac{g\ VS_{destroyed}}{g\ VS_{added}} \quad as\ HRT \to \infty \quad (8)$$

$$A_f = \frac{VFA_{in}}{B_0 \times VSL} \quad (9)$$

The modified Hill's model considers temperature dependence of the anaerobic treatment process through an empirical model, Eq. (10) and since the death rates are set to one tenth of the maximum reaction rates, Eq. (11) they are also show temperature dependent.

$$\mu_{1m}(T) = \mu_{2m}(T) = 0.012T - 0.086 \quad (10)$$

$$K_{d1} = K_{d2} = 0.1\mu_{1m} \quad (11)$$

$$10^\circ C < T < 60^\circ C$$

In the modified Hill's model, the Monod function, Eq. (12) is used to describe the growth rates of acidogenic and methanogenic microorganisms.

$$\mu_1 = \frac{\mu_{m1} S_1}{K_{s1} + S_1} \quad (12)$$

It is known that anaerobic digestion is sensitive to a wide-range of inhibitory conditions either from toxic substrates or by-products of microbial metabolism (Chen et al., 2014). Since the methanogenic archaea are most sensitive to inhibition than any other group of anaerobic microorganisms (Chen et al., 2008), the Monod function used to describe the growth rate of methanogenic archaea will be replaced by an inhibition counterpart, the Haldane model, Eq. (13). The Haldane model is suited for growth processes affected by the allosteric effectors present in the acidified substrate, non-competitive inhibition (Kythreotou et al., 2014).

$$\mu_2 = \frac{\mu_{m2} S_2}{(K_{s2} + S_2)\left(1 + S_2/K_i\right)} \quad (13)$$

After having defined the kinetic model, we now proceed with assessing the model's reliability for synthesis of anaerobic digesters.

2.1.2. Sensitivity-based identifiability

Model sensitivity analysis provides dynamic information on how the states of a process vary with changes in the model parameters. This information can be used to identify time intervals where experimental data points carry more or less importance for the parameter estimation process. For instance, if the sensitivity of a model state to a given parameter is zero or close to zero in some time interval, then variations in that parameter would have a little influence on that state variable. What this means in practical operation is that having a more accurate experimental measurement of the state variable at that insensitive time interval will not serve to improve the reliability of the parameter estimate. The sensitivity-based identifiability consist of analyzing the sensitivity of the model states to the model parameters, and using these sensitivities to screen for parameter significance ranking by calculating a sensitivity measure, δ_k^{msqr} and for analyzing the near-linear dependency between parameters by a measure called the collinearity index, K.

Given a model for a process, the following five key steps needs to be performed in order to completely assess the reliability and usage of the model (Brun et al., 2002).

Step 1: Compute the absolute sensitivity

Since we do not have an explicit solution to the differential equation model, the absolute sensitivities must be computed using the sensitivity equations. For an n-dimensional system given by Eq. (14)

$$\dot{Y} = f(t, Y; \beta), \quad Y(0) = Y_0 \tag{14}$$

With state variable $Y \in \mathbb{R}^n$, the parameter $\beta \in \mathbb{R}^p$ and Y_0 the initial condition, the matrix of sensitivities $\partial Y/\partial\beta$ satisfy

$$\frac{d}{dt}\frac{\partial Y}{\partial \beta} = \frac{\partial F}{\partial Y}\frac{\partial Y}{\partial \beta} + \frac{\partial F}{\partial \beta} \tag{15}$$

With initial conditions

$$\frac{\partial Y(0)}{\partial \beta} = 0_{n\times p} \tag{16}$$

$\partial Y/\partial\beta$ is the Jacobian of the system. The sensitivity equations are coupled with the original model differential equations and solved to obtain the parameter sensitivities for the necessary time points. The resulting matrix of absolute sensitivities at time point t $S_a(t) = \partial Y/\partial\beta$ will be of the form shown by Eq. (17).

$$S_a(t) = \begin{bmatrix} S_{a,11} & S_{a,12} & \cdots & S_{a,1p} \\ S_{a,21} & S_{a,22} & \cdots & S_{a,2p} \\ \vdots & \vdots & \vdots & \vdots \\ S_{a,n1} & S_{a,12} & \cdots & S_{a,np} \end{bmatrix} \tag{17}$$

Step 2: Compute the non-dimensional sensitivity

The sensitivities of the observables are scaled using the same weights as in Eq. (18), resulting in scaled sensitivities for an output j and a parameter i:

$$S_{nd} = S_a(t) \cdot W \tag{18}$$

The non-dimensional scaling/weighting matrix W is of the form shown by Eq. (19) while the resulting non-dimensional sensitivity of the form given by Eq. (20).

$$W = \begin{bmatrix} \beta_1/S_{c1} & \beta_2/S_{c1} & \cdots & \beta_p/S_{c1} \\ \beta_1/S_{c2} & \beta_2/S_{c2} & \cdots & \beta_p/S_{c2} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_1/S_{c3} & \beta_2/S_{c3} & \cdots & \beta_p/S_{cn} \end{bmatrix} \tag{19}$$

$$S_{nd}(t) = \begin{bmatrix} S_{nd,11} & S_{nd,12} & \cdots & S_{nd,1p} \\ S_{nd,21} & S_{nd,22} & \cdots & S_{nd,2p} \\ \vdots & \vdots & \vdots & \vdots \\ S_{nd,n1} & S_{nd,12} & \cdots & S_{nd,np} \end{bmatrix} \tag{20}$$

Step 3: Compute the sensitivity measure

From the matrix of non-dimensional sensitivities, we compute an overall coring for each parameter, called root mean squared sensitivity, δ_k^{msqr}, to consider changes in time or across experiments. The root mean squared sensitivity is computed using Eq. (21)

$$\delta_k^{msqr} = \sqrt{\frac{1}{N}\sum_{l=1}^{N}\left(S_{nd,lk}\right)^2} \tag{21}$$

N is the number of state variables and $k = 1, 2 \ldots p$where p is the number of model parameters. A vector of root mean squared sensitivities of the different model parameters is created of the form given by Eq. (22)

$$\delta^{msqr} = \begin{bmatrix} \delta_1^{msqr} & \delta_2^{msqr} & \cdots & \delta_p^{msqr} \end{bmatrix} \tag{22}$$

The sensitivity measure (δ^{msqr}) measures the relative importance of the parameters with respect to how the influence the model outputs (states). The higher the magnitude of the sensitivity measure the more important the influence of the parameter on the states.

Step 4: Compute the normalized sensitivity

From the matrix of non-dimensional sensitivity, we compute the normalized sensitivity for each parameter using Eq. (23), re-written as Eq. (24).

$$S_{norm} = \frac{S_{nd,lk}}{S_{nd}(t)} \tag{23}$$

$$S_{norm} = \frac{S_{nd,(l,k)}}{\sqrt{\sum_{k=1}^{n} S_{nd,(l,k)}^2}} \quad l = 1, 2, \ldots., n; k = 1, 2, \ldots, p \tag{24}$$

Step 5: Compute the collinearity index

Finally, step five consist of computing the collinearity index γ_K using Eq. (25)

$$\gamma_K = \frac{1}{\sqrt{\min\lambda_K}} \tag{25}$$

$$\lambda_K = eigen\left(S_{norm,\,K}^T S_{norm,\,K}\right)$$

K stands for the index of the parameter subset, which is a combinatorial function of the parameter vector β.

If the sensitivity functions of two or more parameters are orthogonal (implying parameters are independent), the index of that parameter subset (K) is equal to unity, but if the parameters are

linearly dependent, the index approaches infinity. In order to find an identifiable parameter subset, a threshold value (1-15) is usually used (Brun et al., 2002; Sin and Vanrolleghem, 2007) where by any parameter subset having an index (K) greater than the threshold is said to be unidentifiable.

2.1.3. Parameter estimation: confidence bounds and correlation analysis

In this section, we describe the adjoint-based gradient method for parameter estimation. The method is selected rather than the standard finite difference method because it takes less computing time and is less sensitive to round-off and truncation errors, which becomes very attractive for optimization problems with large number of variables (Benítez et al., 2017). To facilitate mathematical developments in subsequent sections, we redefine the model states and parameters as follows;

$$Y_1 = S_1, \quad Y_2 = S_2, \quad Y_3 = X_1 \quad Y_4 = X_2\,, \quad Y_5 = Q_{CH_4}$$

$$\beta_1 = k_1, \quad \beta_2 = k_2, \quad \beta_3 = k_3, \quad \beta_4 = k_4, \quad \beta_5 = K_{i1}, \quad \beta_6 = K_{i2}$$

$$Minimize \quad J(\beta) = \frac{1}{2} Y_i^{obs} - HM\left(t, Y^i; \beta\right)^2 \tag{26}$$

subject to

$$\frac{dY_1}{dt} = (Y_{1_{in}} - Y_1)D - \beta_1 \mu_1 Y_3 \tag{26a}$$

$$\frac{dY_2}{dt} = (Y_{2_{in}} - Y_2)D + \beta_2 \mu_1 Y_3 - \beta_3 \mu_2 Y_4 \tag{26b}$$

$$\frac{dY_3}{dt} = (\mu_1 - K_{d_1} - D)Y_3 \tag{26c}$$

$$\frac{dY_4}{dt} = (\mu_2 - K_{d_2} - D)Y_4 \tag{26d}$$

$$Y_5 = V\mu_2\beta_4 Y_4 \tag{26e}$$

$$\beta_1 < \beta_3; \quad \beta_5 < \beta_6 \quad \beta_1, \beta_2, \beta_3, \beta_4 > 0 \quad \beta_5, \beta_6 \geq 0$$

$$\mu_i = \mu_i(\beta_{i+4}), \quad i = 1, 2$$

Eq. (26) presents a constraint nonlinear optimization problem, where the constraints are differential algebraic equations. In order to find the numerical solution of the problem there exist indirect and direct methods of minimization of the objective function. In the direct method, the state equations are influenced only by the parameters, and the minimization of the function is done by direct adjustment of the model parameters. The simplest approach to a direct method is that of first order gradients in which the state and co-state equations remain separated. The system of continuous equations is regarded as a limiting case of a system of discrete equations as the time of a subinterval approaches zero. The optimization problem is solved using the method of conjugate gradients with the gradient computed by the adjoint method. The conjugate gradient algorithm is illustrated as follows

Given $J : \mathbb{R}^n \to \mathbb{R}$ and $\nabla J(\beta)$. Let $\beta^{(0)}$ be the initial guess and set $w^{(0)} = -\nabla_\beta J(\beta^{(0)})$

For $k = 0, 1, 2, 3, \ldots$

Step 1: Perform a line search in the direction of to compute $\gamma = Arg\,min\Phi(\rho)$, which minimizes the scalar function $\Phi(\rho) = f(\beta^{(k)}, \rho w^{(k)})$
Step 2: Compute $\beta^{(k+1)} = \beta^{(k)} + \gamma^{(k)} w^{(k)}$
Step 3: Test for convergence. If satisfied Exit, else go to Step 4.
Step 4: Define $\tau^{(k)} = \frac{\nabla J(\beta^{(k)})^2}{\nabla J(\beta^{(k-1)})^2}$
Step 5: Compute $w^{(k+1)} = -\nabla_\beta J(\beta^{(k+1)}) + \tau^{(k)} w^{(k)}$ and go to step (1).

We notice that from the computational point of view a discrete adjoint approach is the one needed to accurately compute the gradient. The model equations are discretized using the Runge-Kutta 4th order scheme as shown by Eq. (27)

$$Y^{(k+1)} = M\left(Y^{(k)}, \beta\right) \tag{27}$$

$$f(t, Y(t);\ \beta) = \begin{bmatrix} (Y_{1_{in}} - Y_1)D - \beta_1\mu_1 Y_3 \\ (Y_{2_{in}} - Y_2)D + \beta_2\mu_1 Y_3 - \beta_3\mu_2 Y_4 \\ (\mu_1 - K_{d_1} - D)Y_3 \\ (\mu_2 - K_{d_2} - D)Y_4 \\ V\mu_2\beta_4 Y_4 \end{bmatrix} \tag{27a}$$

$$K_1 = f\left(t^{(k)}, Y^{(k)}; \beta\right) \tag{27b}$$

$$K_2 = f\left(t^{(k)} + 0.5h, Y^{(k)} + 0.5K_1; \beta\right) \tag{27c}$$

$$K_3 = f\left(t^{(k)} + 0.5h, Y^{(k)} + 0.5K_2; \beta\right) \tag{27d}$$

$$K_4 = f\left(t^{(k)} + h, Y^{(k)} + K_3; \beta\right) \tag{27e}$$

$$M\left(Y^{(k)}, \beta\right) = Y^{(k)} + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4) \tag{27f}$$

The optimization problem can then be simply written in a discrete and compressed form as in Eq. (28)

$$Minimize\ J(\beta) = \frac{1}{2}\sum_{k=0}^{N}\left(Y_k^{obs} - HY^{(k)}\right)^2 \tag{28}$$

subject to

$$Y^{(k+1)} - M\left(Y^{(k)}, \beta\right) = 0 \tag{28a}$$

The adjoint method consist of transforming a constraint optimization problem into an unconstrained problem by defining the Lagrangian, Eq. (29)

$$L(Y,\ \beta,\ \lambda) = J(\beta) + \sum_{k=0}^{N}\lambda_k\left[Y^{(k+1)} - M\left(Y^{(k)}, \beta\right)\right] \tag{29}$$

From the Lagrangian, we can then derive the state equations, Eq. (30) and gradient of the optimization problem, Eq. (31). The Adjoint model is given by Eq. (32) (Roulston, 1999).

$$\frac{\partial L(Y, \beta, \lambda)}{\partial Y^{(k)}} = \frac{\partial J_k(\beta)}{\partial Y^{(k)}} + \lambda_{k-1} - \lambda_k\left[\frac{\partial M\left(Y^{(k)}, \beta\right)}{\partial Y^{(k)}}\right]^T \tag{30}$$

$$\nabla_{\beta} J(\beta) = \frac{\partial L(Y,\beta,\lambda)}{\partial \beta} = -\sum_{k=0}^{N} \lambda_k M_{\beta}^{T}\left(Y^{(k)},\beta\right) \tag{31}$$

$$\lambda_{k-1} - M_{Y}^{T}\left(Y^{(k)},\beta\right)\lambda_k = e_k \quad \lambda_N = 0 \tag{32}$$

Step 1: Choose an initial guess $\beta^{(0)}$ and set counter $k = 0$
Step 2: Solve the forward model $Y^{(k+1)} = M\left(Y^{(k)},\beta\right)$ and compute the criterion $J(\beta)$
Step 3: Solve the Adjoint model $\lambda_{k-1} - M_{Y}^{T}(Y^{(k)},\beta)\lambda_k = e_k$ and compute the gradient

$$\nabla_{\beta} J(\beta) = -\sum_{k=0}^{N} \lambda_k M_{\beta}^{T}\left(Y^{(k)},\beta\right)$$

Step 4: Determine the descent direction

if $k = 0$

$$w^{(k)} = -\nabla_{\beta} J\left(\beta^{(k)}\right)$$

else

$$w^{(k)} = -\nabla_{\beta} J\left(\beta^{(k)}\right) + \frac{\nabla J\left(\beta^{(k)}\right)^{2}}{\nabla J\left(\beta^{(k-1)}\right)^{2}} w^{(k-1)}$$

Step 5: Perform a line search in the direction of to compute $\gamma = Arg\ min\Phi(\rho)$, which minimizes the scalar function $\Phi(\rho) = f(\beta^{(k)}, \rho w^{(k)})$
Step 6: Compute a new state vector estimate $\beta^{(k+1)} = \beta^{(k)} + \gamma^{(k)} w^{(k)}$
Step 7: Set $k =: k+1$ and return to step 2 until a termination condition is reached

All the work on the computer was carried out using Matlab R2017b (Mathworks Natick) using i7-6600U, 2.6 GHz CPU PC with 16 GB RAM and 64bits operating system.

2.1.4. Uncertainty quantification in model predictions

As mentioned in section 1, self-optimizing operation of anaerobic digesters is when we have an acceptable loss in performance as a result of kinetic uncertainty in the model. In order to therefore use the model to model the self-optimizing performance target, one needs to quantify the model prediction uncertainty resulting from uncertainty in kinetic coefficients. In order to quantify the model prediction uncertainty, the Monte Carlo simulation procedure, presented in Fig. 3 was applied in a similar way as in Sin et al. (2010b).

Input-output uncertainty analysis is highly dependent on the input uncertainty range (confidence bounds) as well as correlation coefficients. The variance metrics and correlation coefficients of the unidentifiable set of model parameters for the different biokinetic models were obtained by estimating the complete set of parameters (identifiable and unidentifiable) using the estimation procedure presented in section 2.1.3.

2.2. Self-optimizing performance targeting

Given a set of reactions and associated kinetics, the following five key steps needs to be performed in order to define the performance target of a process using attainable region analysis (Ming et al., 2016):

- Define the reaction, dimension and feed set
- Define the fundamental processes occurring in the system
- Generate the AR using combinations of the fundamental processes
- Interpret the AR boundary in terms of reactor equipment
- Define the objective function and overlay this onto the AR to determine point of intersection with the AR boundary
- Determine the specific reactor configuration required to achieve the intersection point

The previous two bullet points are important if the attainable region is to be used to answer a specific design or optimization question.

Some necessary conditions for AR can be summarized as follows (Hildebrandt and Glasser, 1990; Hildebrandt et al., 1990):

- The AR includes all feeds to the system.
- The AR is convex.
- No process vector point out of the AR boundary.
- No rate vectors in the complement of the AR when extended backward intersects the AR.

The objective of this section is to analyze the aforementioned necessary requirements with respect to its application to the anaerobic treatment process.

2.2.1. Reaction scheme and process kinetics

Using the information flow diagram of the kinetic model presented in Fig. 2, a stoichiometric scheme of the bioreaction occurring in the anaerobic digester consist of two main reactions catalyzed by acid-forming bacteria, Eq. (33) and methane-forming bacteria Eq. (34)

$$k_1 S_1 \xrightarrow{r_{X_1}} X_1 + k_2 S_2 \tag{33}$$

$$k_3 S_2 \xrightarrow{r_{X_2}} X_2 + k_4 CH_4 \tag{34}$$

If we assume the specific death rate to be negligible compared to the specific growth rate of both microbial populations, the rate expressions for the different reaction species is defined by Eq. (35) – (38)

$$r_{X_1} = \mu_1 X_1 \tag{35}$$

$$r_{X_2} = \mu_2 X_2 \tag{36}$$

$$r_{S_1} = -k_1 \mu_1 X_1 \tag{37}$$

$$r_{S_2} = k_2 \mu_1 X_1 - k_3 \mu_2 X_2 \tag{38}$$

2.2.2. Fundamental processes

Various fundamental processes can occur within a system, which for bioreactors may include: mass transfer, mixing, bioreaction (biodegradation, bioconversion), adsorption, heat transfer, etc. The AR approach requires the fundamental processes taking place in the system be identified. The following two main fundamental processes are identified to be associated with the anaerobic treatment process: Biodegradation and mixing. The attainable

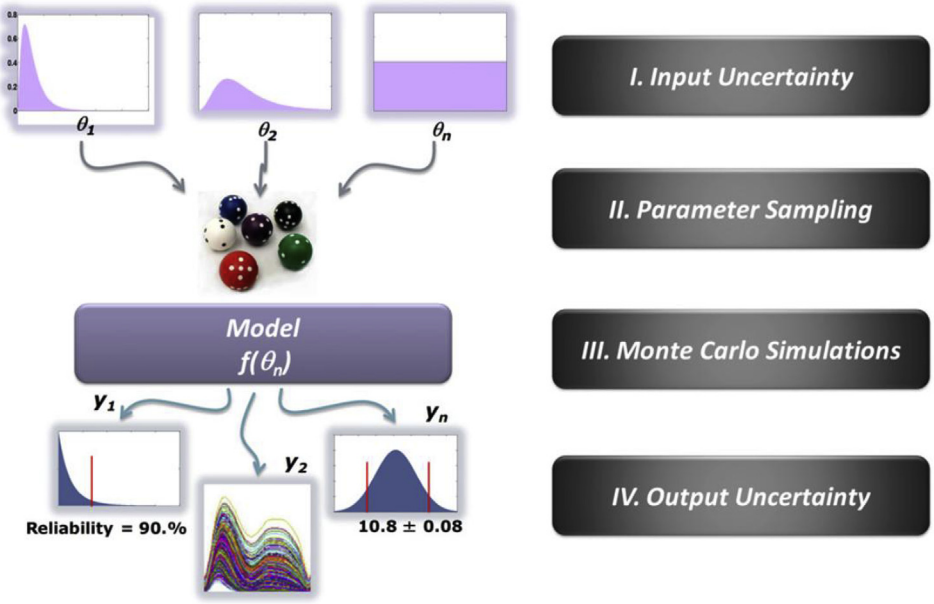

Fig. 3. Monte-Carlo simulation method for uncertainty propagation (Morales-Rodriguez et al., 2012).

region (AR) for the anaerobic treatment process therefore represents the set of all possible states that can be achieved by a combination the two fundamental processes, biodegradation and mixing. In AR theory, mixing is performed by a continuous stirred tank reactor (CSTR) while reaction (biodegradation) is achieved in a plug flow reactor (PFR), since the operation of both reactors respectively mimic the two fundamental processes. At steady state operation, the general mathematical representation of a CSTR and PFR are given by Eqs. (39) and (40) respectively.

$$C = C_f + \tau r(C) \tag{39}$$

$$\frac{dC}{d\tau} = r(C) \tag{40}$$

C is the state vector while $r(C)$ is the reaction rate vector as shown by Eqs. (41) and (42) respectively.

$$\boldsymbol{C} = [X_1 \quad X_2 \quad S_1 \quad S_2]^T \tag{41}$$

$$\boldsymbol{r}(\boldsymbol{C}) = [r_{X_1} \quad r_{X_2} \quad r_{S_1} \quad r_{S_2}]^T \tag{42}$$

2.2.3. Dimensionality analysis and model reduction

The reaction stoichiometry of the system can be used to determine the dimension of the system. The dimension of the AR is determined from the number of independent reactions occurring in the reactor system, which defines the dimension of the stoichiometric subspace (the rank of the stoichiometric coefficient matrix A), in which the AR must reside. Since there are two independent reactions occurring in the system, the set of points generated by the anaerobic treatment process must reside in a two-dimensional subspace in $\mathbb{R}^5$ (Ming et al., 2016). The reduced state and reaction rate vectors are therefore presented by Eqs. (43) and (44).

$$\boldsymbol{C} = [S_2 \quad X_2]^T \tag{43}$$

$$\boldsymbol{r}(\boldsymbol{C}) = [\, r_{S_2} \quad r_{X_2}]^T \tag{44}$$

The reduction involved expressing S_1 and X_1, as a function of S_2 and X_2, as shown by Eqs. (45) and (46).

$$S_1 = S_{1_{in}} - k_1(X_1 - X_{1_{in}}) \tag{45}$$

$$X_1 = X_{1_{in}} + \frac{1}{k_2}[S_2 - S_{2_{in}} + k_3(X_2 - X_{2_{in}})] \tag{46}$$

This reduction in the dimensions of the state and rate vectors was done using the approach illustrate in our recent study using attainable regions for synthesis and optimization of methane bioreactors (Ref). The model reduction assumes that the specific death rates of acidogens and methanogens is negligible compared to their respective specific growth rates.

2.2.4. AR construction and defining performance target for the system

After stating the process kinetics, the AR construction process is initiated by defining feed point and process conditions that influence the system. In anaerobic treatment, the digester is normally maintained at constant temperature (isothermal process) throughout retention time, which makes the AR dependent on the particular temperature in the system. The anaerobic digestion was carried out under mesophilic conditions at a temperature of 35°C. Using the specified feed, kinetics and temperature conditions, the set of points generated by solving the PFR equation are called the PFR trajectory and those generated by solving the CSTR equation are called the CSTR locus. The convex hull of the set of points generated by the system defines the attainable region, which represents the limits of achievability by the system.

3. Results and discussion

In this section, data from a real experiment is utilized to illustrate the theories presented in the previous sections. The case study is based on a batch methane bioreactor operated with diary manure, where experimental measurements of volatile fatty acids

and methane gas flowrate (which, can be used to get the concentration of methanogenic archae) were obtained (Zaher et al., 2009).

3.1. Parameter identifiability measures

3.1.1. Sensitivity analysis

In this section, the objective was to determine, which set of parameters should be estimated to accurately describe the mechanisms of the anaerobic digestion process. This depends on analyzing the sensitivity function of the model parameters with respect to with respect to the states. Fig. 4 presents the sensitivity functions (dynamic sensitivities) of the states for the parameters of the biokinetic model. From the shape of the sensitivity functions, the authors made the following remarks: (1) All the states show some sensitivity to the model parameters, which can either be a negative or positive sensitivity. (2) The anaerobic microorganisms mostly show negative sensitivity while the substrates show both negative and positive sensitivities to the model parameters. Table 1 presents the numerical characteristics sensitivity analysis, which include: the nominal values and scale of the model parameters; the mean, minimum and maximum values of the dynamic sensitivities; the sensitivity measures (L1 and L2) as well as the number of data points (N). The 130 data points corresponds to the small time step of 0.0769 that was used to integrate the sensitivity equations from 0 to 10days.

Particularly, it is worth mentioning that for the substrates, the biodegradable volatile solids is most sensitive to the acid yield coefficient (k_2) while volatile fatty acids are most sensitivity to the Monod saturation constant for volatile acids (K_{s2}) and inhibition constant (K_i). For the anaerobic microorganisms, the acidogenic bacteria is highly sensitive to the Monod saturation constant (K_{s2}), while the methanogenic archae are highly sensitive to inhibition constant (K_i). These outcome accurately describe the underlying theories of the anaerobic treatment process, which include: breakdown of volatile solids into volatile fatty acids by acidogenic bacteria, utilization of volatile fatty acids for growth of methanogenic archaea as well as high sensitivity of methanogenic archaea to inhibitions (Henze et al., 2008; Wang et al., 2007). Hence the results clearly illustrate the ability of the model to describe the anaerobic digestion process.

Fig. 5 presents the use of the sensitivity measure (sum of sensitivity functions of the available measurements with respect to

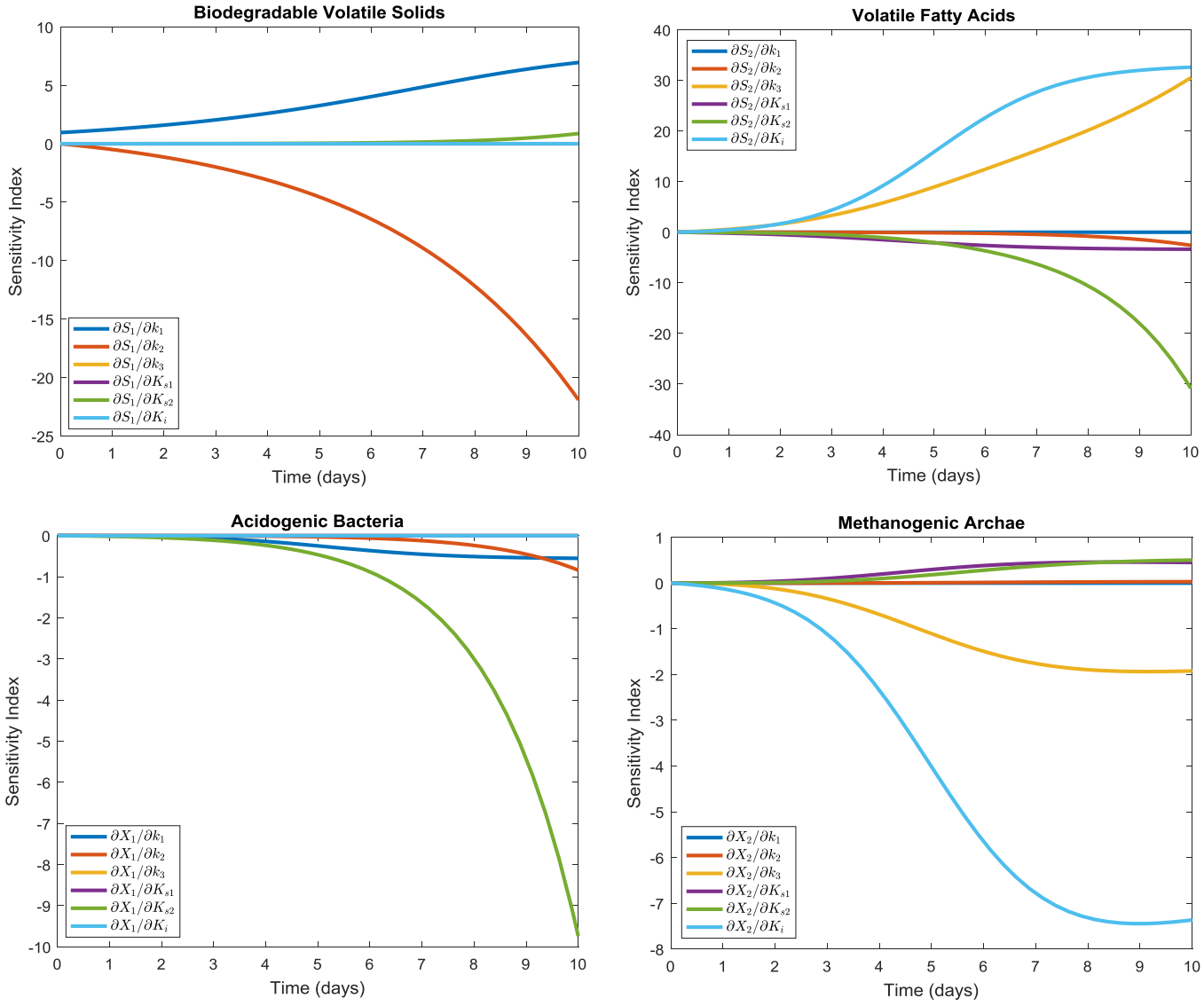

Fig. 4. Sensitivity of model states to model parameters.

Table 1
Sensitivity measures of the observable states to the model parameters.

Parameters	Value	Scale	L1	L2	Mean	Min	Max	N
$\boldsymbol{k_1}$	0.1920	0.1920	0.0000	0.0000	0.0000	0.0000	0.0000	130
$\boldsymbol{k_2}$	0.5029	0.5029	0.2131	0.6086	−0.2042	−2.5915	0.0302	130
$\boldsymbol{k_3}$	0.1920	0.1920	4.6606	9.4970	3.9058	−1.9351	30.5131	130
$\boldsymbol{K_{s1}}$	25.0687	25.0687	0.7843	1.3894	−0.5965	−3.3613	0.4591	130
$\boldsymbol{K_{s2}}$	0.0899	0.0899	2.7457	7.4232	−2.5820	−30.8306	0.5022	130
$\boldsymbol{K_i}$	1088.8324	1088.8324	7.3896	12.9875	4.5331	−7.4461	32.5866	130

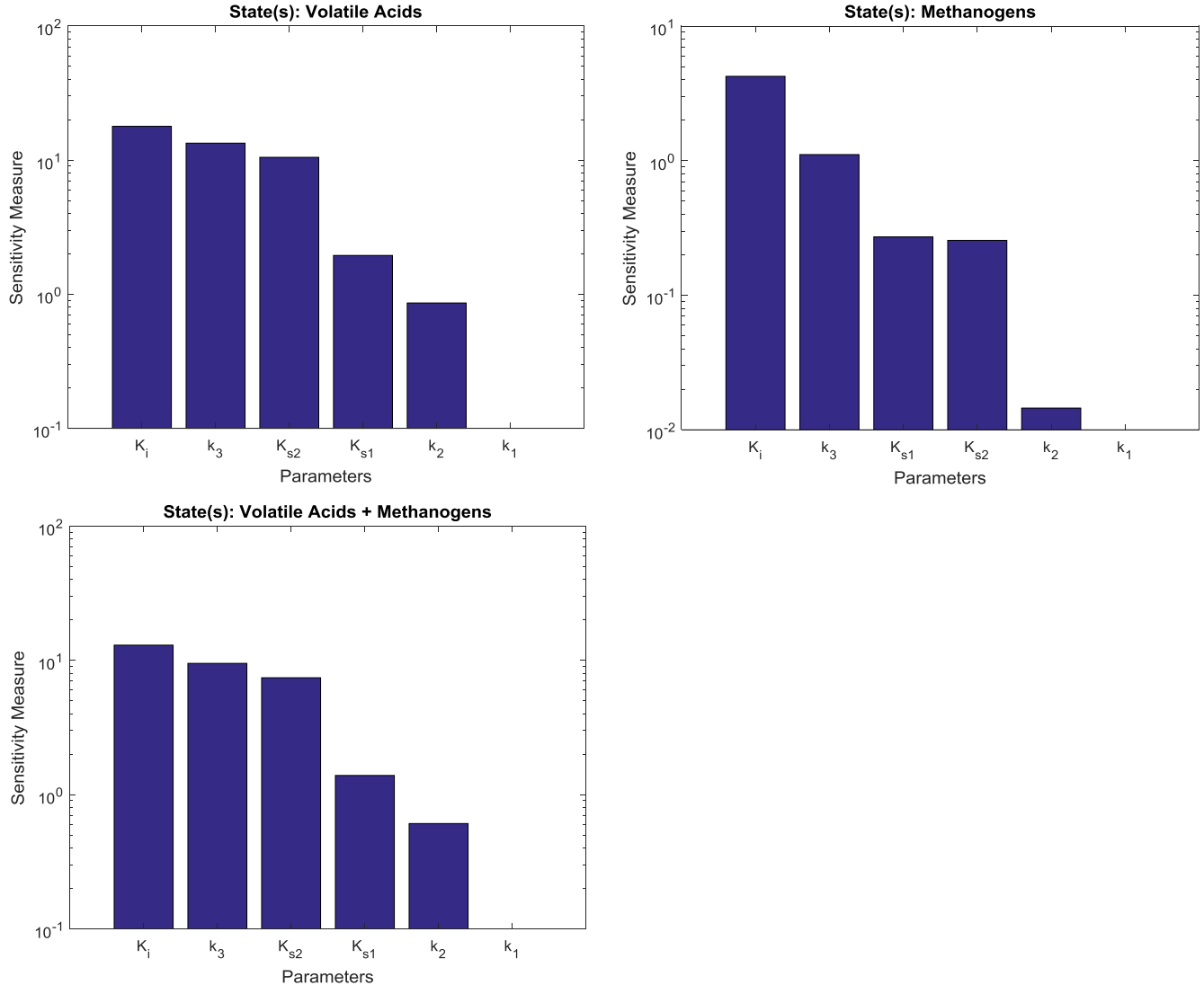

Fig. 5. Parameter significance ranking for observable states based on sensitivity measure.

parameters) to rank the model parameters, which reveals the relative significance of the parameters with respect to the observable states. Only the relative ranking for volatile acids, methanogenic archaea as well as volatile acids + methanogenic archae as model as model outputs are considered. This is because the experimental data, which is used as our case study only contains measurements for volatile acids methanogenic archae (calculated form methane flowrate) and we therefore determine the set of model parameters that are identifiable considering these two observable states.

It is noteworthy that the yield coefficient (k_1) is found completely not sensitive to any of the observable states. The inhibition constant (K_i) was found most significant for all the observable states followed by the yield coefficient (k_3). The Monod constants (K_{s1} and K_{s2}) and the yield coefficient (k_2) were found significant albeit to a relatively lower degree. The practical relevance of the parameter significance ranking is that only those parameters with significant sensitivity measure with respect to the observable states can be identified. Therefore, considering volatile fatty acids and methanogenic archaea as the only observable states in the system, the yield coefficient k_1 cannot be identified from the available data since it has a sensitivity measure of zero. This implies

only 5 out of the 6 model parameters can be candidates for parameter estimation.

3.1.2. Collinearity analysis

This section of the identifiability analysis only considers those set of parameters (5 out of the 6 model parameters were significant), which have a significant effect on the observable states. The collinearity analysis screens all possible subsets of the potential candidate parameters to determine the identifiable subsets using a collinearity index. The five potential candidates for parameter estimation gives a total of approximately 31 parameter subset combinations, with a maximum subset size of five parameters

Fig. 6a presents the collinearity analysis for all possible combinations of parameter subsets while Fig. 6b presents collinearity analysis for the potentially identifiable subsets. From Fig. 6, it can be observed that of the 31 possible subset combinations, only 18 are potentially identifiable (those having collinearity index less than 15) and with a maximum identifiable subset size of three parameters

These findings suggest that for a given set of observable states (experimental measurements), there exist many identifiable subset combinations of model parameters having a maximum number of parameters that can be estimated uniquely. Unique estimation means that by using an identifiable subset, the estimated parameters should have a relatively lower correlation values and/or confidence intervals. The results corroborate the theoretical premise that subjecting an overparameterized model to limited quality/quantity of data limits the number of parameter that can be estimated to uniquely and accurately describe the system (Brun et al., 2002; Sin and Vanrolleghem, 2007).

3.2. Model fits and parameter uncertainty

Two cases of parameter estimation were considered: one with an identifiable subset (specifically k_2, K_{s2} and K_i) and one with all the model parameters (known as the nominal case) so that the effect of identifiability analysis can be ascertained. The model fits for both cases are presented in Fig. 7 while parameter estimates together with their 95% marginal confidence intervals are shown in Table 2. Visually, both cases show a good fit between the experimental measurements and model predictions with no observable difference in both cases. However, from a numerical perspective (see Table 2), the parameter estimates from the nominal case shows a much higher degree of uncertainty (given as the standard deviation, which relates to the 95% marginal confidence interval) than that of the identifiable case. Put it in another way, the identifiability analysis has served to reduce the degree of uncertainty in model parameter estimates.

The results indicate that despite variation in parameter uncertainty, the quality of the model fit to experimental data is not compromised and using an identifiable subset of model parameters serves to improve the quality of the model parameters. It is worth mentioning that the identifiable parameter subset utilized for parameter estimation is just one of the three identifiable subsets available with size of 3. Other three-parameter combinations of identifiable subsets can still be selected as candidates for the parameter estimation. The focus of this study is not to consider all the identifiable subset, but to illustrate how these identifiability issues should be incorporated in digester synthesis.

3.3. Uncertainty quantification on model states

Recall from step 3 of section 2, which stated the need to quantify the model prediction (state) uncertainty using the unidentifiable parameter set as inputs. From section 3.2, we have demonstrated the use of an identifiable parameter subset (k_2, K_{s2} and K_i) to reduce uncertainty in model parameters. Even though the use of an identifiable subset reduces parameter uncertainty, it causes another problem, which is that of model uncertainty. This is because those parameters that are not identifiable (k_1, k_2 and K_{s1}) need to be kept constant (probably using values estimated from previous studies or independent experiments), which influences the reliability of the model. Since the geometric optimization technique of attainable regions presented in this study for synthesis of anaerobic digesters is unique for a given kinetic model, an unreliable model will therefore result in an unreliable digester system, which can easily lead to operational failure. Hence before using the model to construct the attainable regions (which can be interpreted into digester structures), we quantify and incorporate the model prediction uncertainty into the limits of achievability of the system, which is defined by the attainable regions.

Fig. 8 presents the results obtained from the Monte Carlo simulations. From a general perspective, the results indicate that each of the model states have a time varying uncertainty band defined by the 10th and the 90th percentile. The methanogenic archaea shows insignificant uncertainty band to the model inputs at certain times instants, where the mean, 10th and 90th percentile are equal. The width of the band (difference between the 10th and the 90th

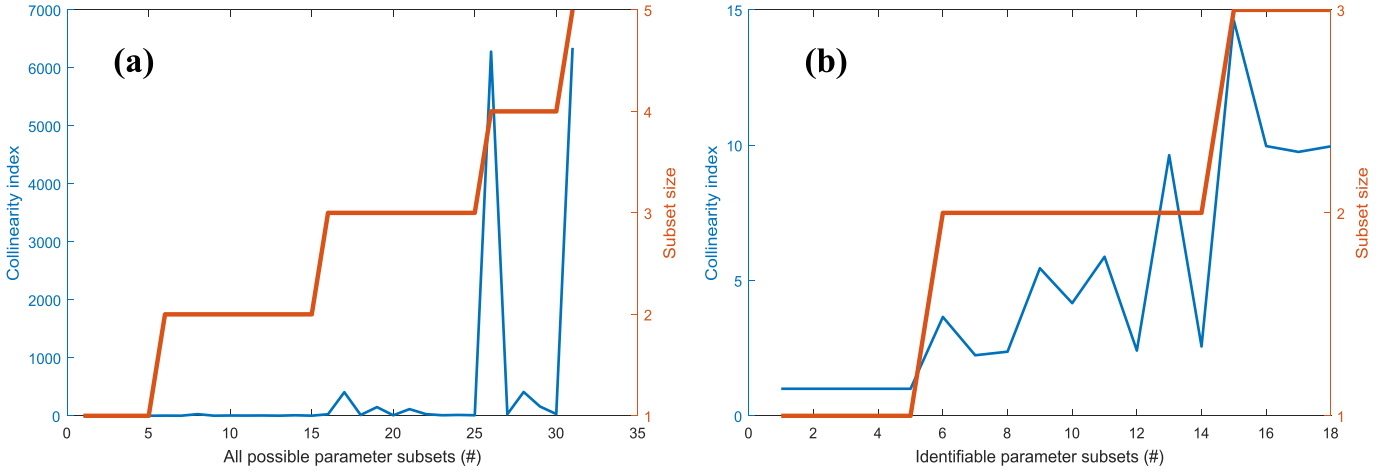

Fig. 6. Collinearity analysis for: (a) all possible combination of parameter subsets (b) potentially identifiable parameter subsets.

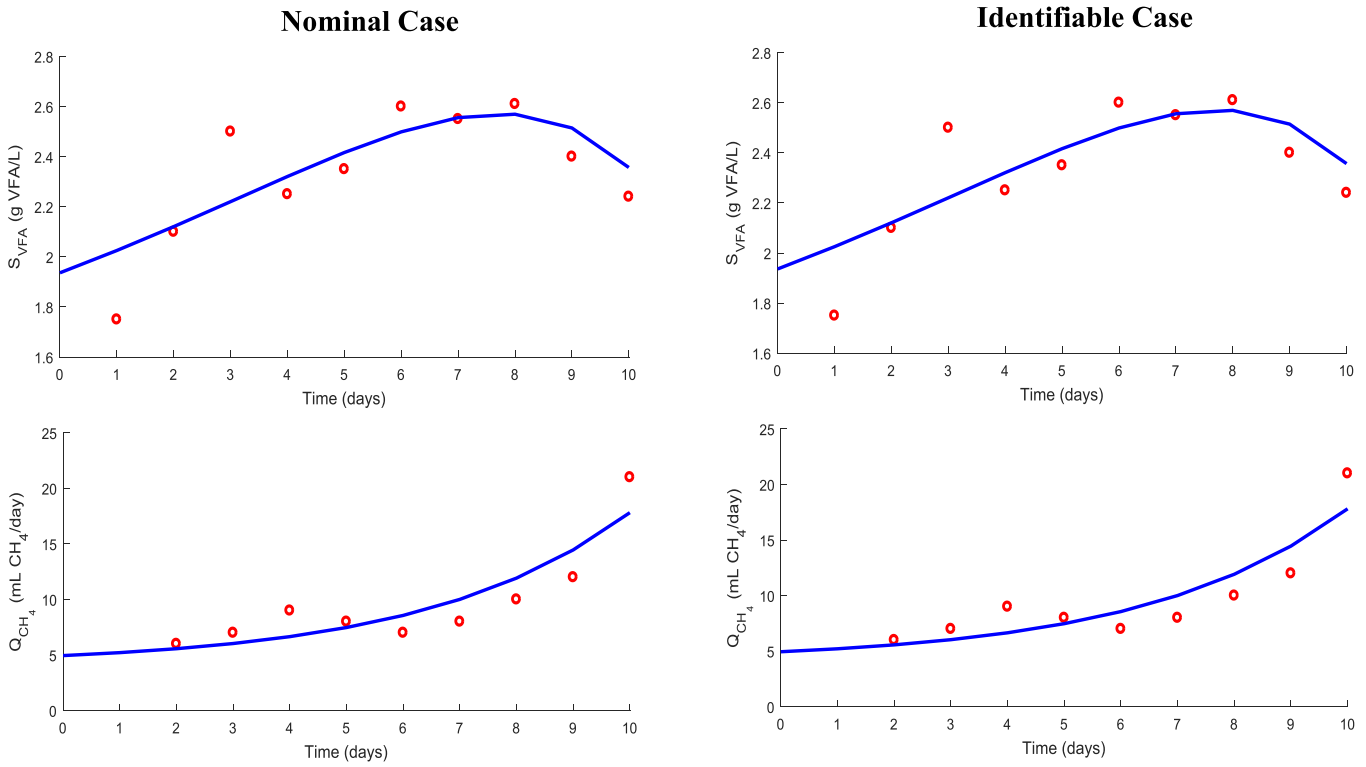

Fig. 7. Model fits to experimental data for both nominal and identifiable cases.

Table 2
Parameter estimates and uncertainty given as the standard deviation.

Model Parameters	Parameter Estimates		Standard Deviation	
	Nominal case	Identifiable case	Nominal case	Identifiable case
k_1	0.1920	0.1920	45.8498	N/A
k_2	0.5029	0.514889	3.2951	0.62322
k_3	0.1920	0.1920	3.6133	N/A
K_{s1}	25.0687	25.0687	386.8622	N/A
K_{s2}	0.0899	0.852436	0.2822	0.78416
K_i	1088.8324	1088.5435	22887.6440	12.1010

percentile) describes the spread of the distribution of the model states resulting from parameter uncertainty and the larger the width, the higher the degree of model output uncertainty. This is often called mapping/propagating parameter (input) uncertainty onto states (output) uncertainty.

The state uncertainty bands presented in Fig. 8 are highly dependent on the uncertainty range of the model parameters. The study used the joint confidence region (Fig. 9) of the sampled parameters, which takes into consideration the correlation amongst model parameters and eliminates the need to define the correlation amongst model parameters during the Monte Carlo procedure.

The interpretation of these results is based on the relationship between uncertainty band and model quality: the higher the uncertainty band, the lower the model quality. Hence model predictions for biodegradable volatile solids followed by volatile fatty acids and acidogenic bacteria are deemed of low quality (large uncertainty bands) while that of methanogenic archaea can be deemed acceptable.

3.4. Self-optimizing attainable regions

Surely, whether of acceptable quality or not, the prediction uncertainty around the model states affects the limits of achievability of the anaerobic digestion process and hence the nature of the optimal digester structures. This is because for synthesis of methane bioreactors using attainable region analysis, the predicted performance target or limits of achievability by the system is computed by the area of the convex hull for the set of states (outputs) achievable by the system.

Hence, when using attainable regions for performance targeting and digester network synthesis, we suggest that it should be mandatory to incorporate uncertainty of model prediction during construction of the attainable regions. The approach here relies on constructing the attainable regions using the three key points of the state's prediction (mean, 10th and 90th percentile) and superposing the regions to obtain a robust region which considers the effect of uncertainty. Fig. 10 presents the AR for the 10th percentile, mean and 90th percentile state predictions on to which the digester structures required to attain points on the AR boundary has been overlaid. A detailed explanation of how the AR boundary has been interpreted into digester structures is presented in our recent publication (Abunde Neba et al., 2019c). It can be observed that for all the cases where the AR boundary is convex, the optimal digester structure involves a plug flow digester in order to attain points on

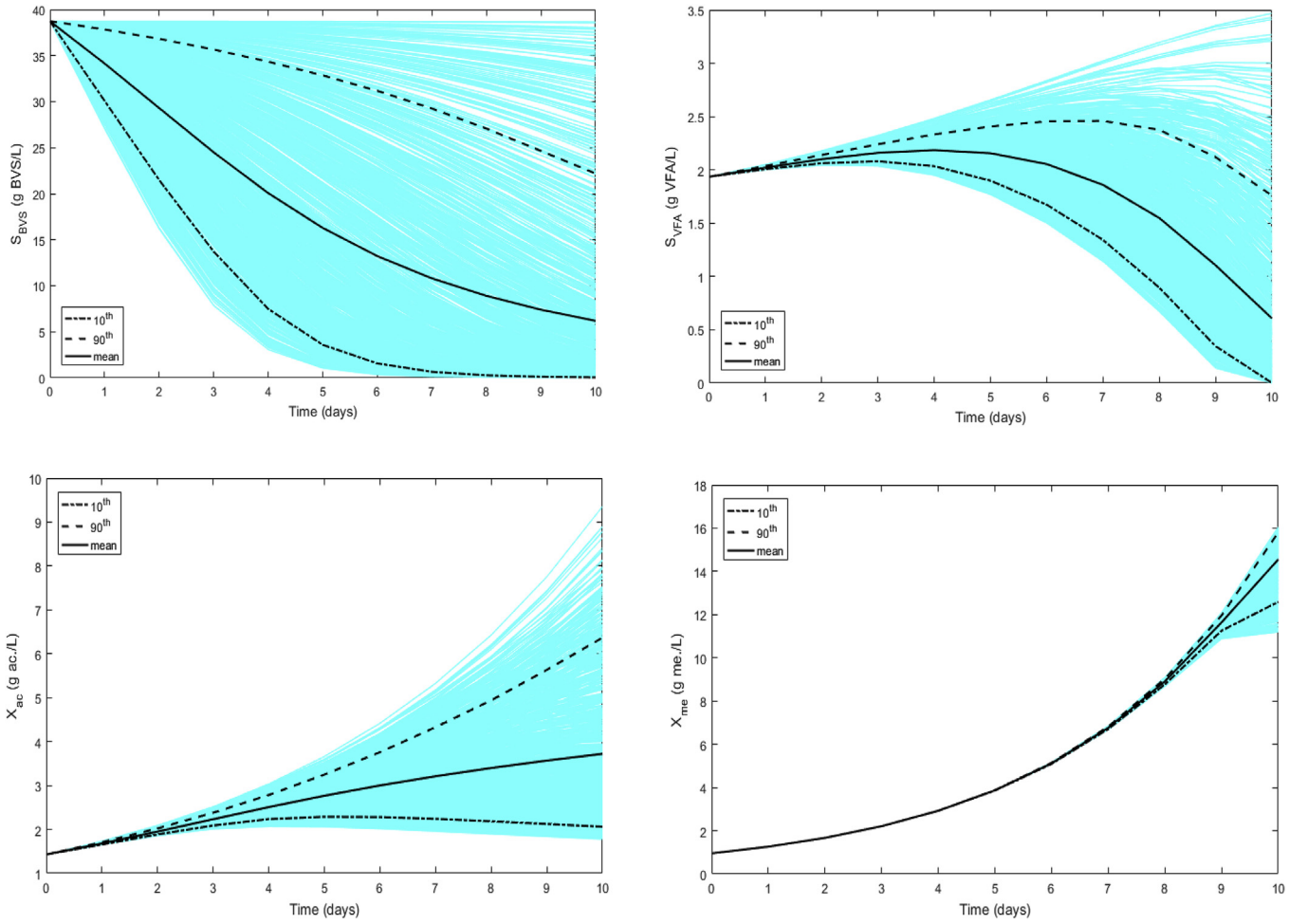

Fig. 8. Model uncertainty quantification (the mean, the 10th and 90th) using 1000 Monte Carlo simulations.

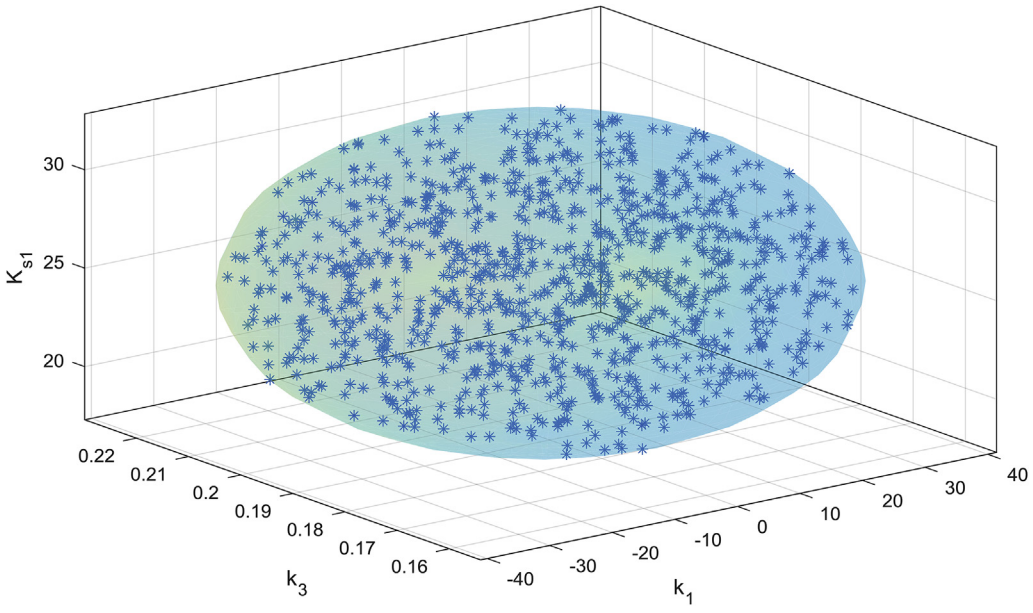

Fig. 9. Joint confidence region of unidentifiable parameters showing parameters sample for 1000 Monte Carlo simulations.

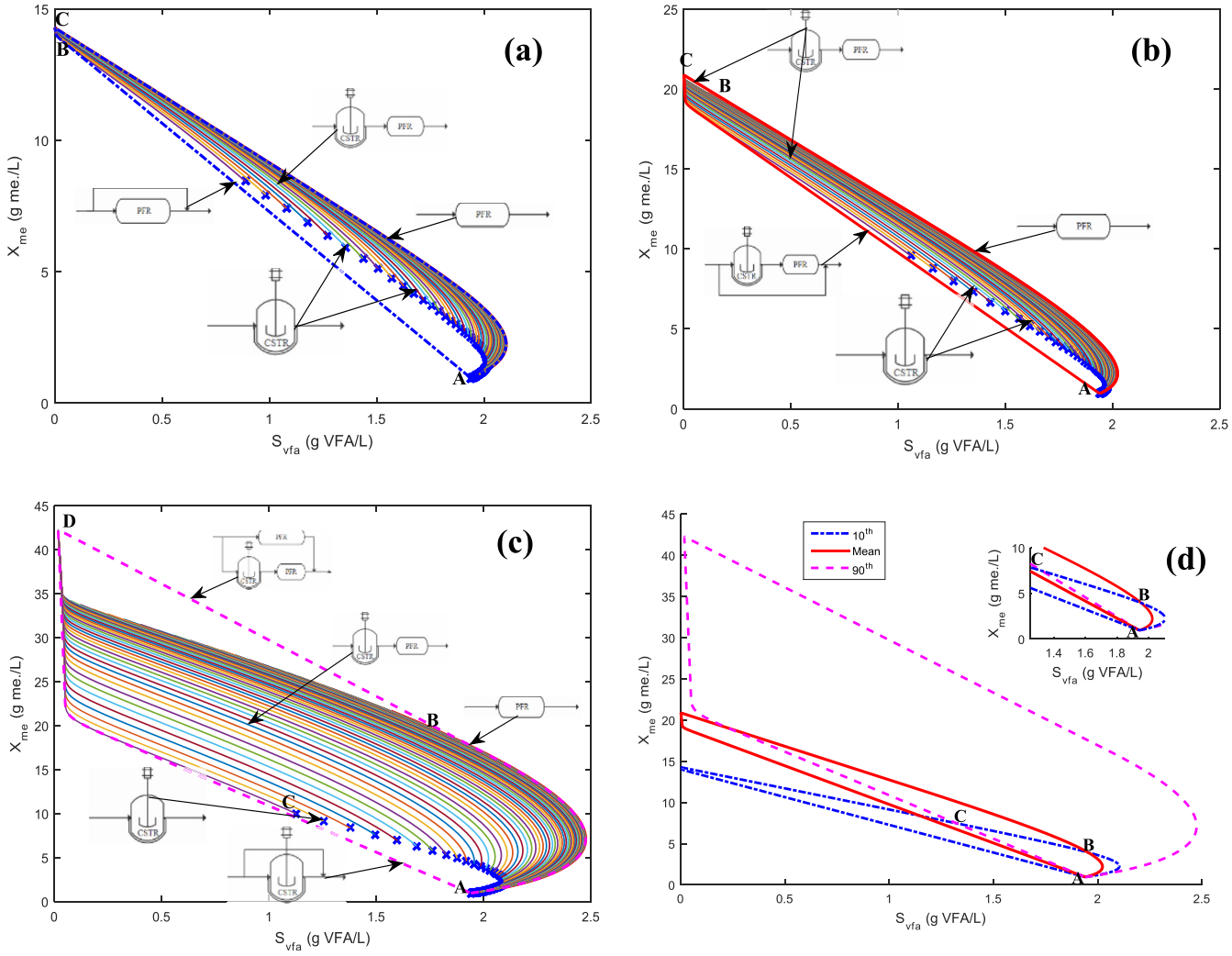

Fig. 10. Illustration of the self-optimizing attainable region: (a) Attainable region for 10th percentile uncertainty bound. (b) Attainable region for mean parameter values. (c) Attainable region for 90th percentile uncertainty bound. (d) Self-optimizing attainable regions of the anaerobic treatment process.

the AR boundary. This can be explained by the intrinsic geometric characteristics of this type of reactor in relation to the properties of the AR boundary. The AR boundary is composed entirely of reaction and mixing surfaces only. Reaction surfaces are always convex and the points that form convex sections of the AR boundary arise specifically from points on PFR trajectories (Ming et al., 2016). This is so because governing equations of a PFR is a system of first order ordinary differential equations Eq. (17), where a phase plane presentation of the solution of the system for a given organic load and digestion time is called PFR trajectory. Geometrically, the rate vector evaluated at points on the PFR trajectory is tangent to all points on the trajectory (Hildebrandt and Glasser, 1990; Ming et al., 2016). This implies the boundary of a true AR will always contain points originating from PFR trajectory, otherwise it becomes a candidate AR. Fig. 10d present the intersection of the three regions to define the self-optimizing attainable region. It can be observed that the region looks smaller than any of the three individual regions (10th percentile, mean and 90th percentile). This illustrates the accept loss in process operation mentioned in section 1. It is necessary here to re-clarify exactly what is meant by self-optimizing attainable regions. Unlike the attainable region, which represents the set of all possible states that is attainable by the system for a defined kinetics and initial condition (feed point), the self-optimizing attainable region represents the set of all possible states attainable by the system even in cases of kinetic uncertainty. The size of the self-optimizing attainable region is related to the domain of uncertainty defined for the unidentifiable set of model parameters used for the uncertainty propagation. The presence of uncertainty reduces the size of the self-optimizing AR and if the domain of uncertainty is reduced, the size increases. As mentioned in section 1, the attainable region defines the limits of achievability (performance targets) by a system. This implies that considering uncertainty has greatly reduced the limits of achievability by the system even though we have benefited from increased robustness. The authors will also like to clarify at this point that by performance targets, the authors refer to the totality concentration of microorganisms and substrates that can be output by the different digester combinations using the fundamental processes occurring in the system. This is defined by the attainable region of the system for a given kinetics and by the self-optimizing attainable region for different kinetic variations within a defined domain.

The findings from this study are therefore highly important in making economic feasibility decisions about the performance of biogas plants especially in cases where accuracy is very necessary. Put it in another way, when assessing the economic feasibility of the anaerobic treatment process, one can now consider the

economic performance of the process even in cases of uncertainty and reliably compare it with other process alternatives. In summary coupling uncertainty analysis and attainable region theory provides a systematic methodological framework for dealing with kinetic uncertainty during design of biogas digesters and hence allows biogas engineers to benefit from the advantages of model-based design. These advantages include easy digester scale-up, less experimental runs (hence less cost), as well as obtain optimal design parameters and digester configurations. This approach is therefore recommended as a reliable strategy for design of biogas plants in cases of kinetic uncertainty, which is very common with biokinetic models for anaerobic digestion.

4. Conclusion

A systematic model-based framework for the synthesis of biogas reactors under cases of kinetic uncertainty has been developed. Using the case of the modified Hill model for anaerobic digestion, the following conclusions are made:

- ➢ Identifiability analysis reveals that only 5 out of the 6 model parameters can be candidates for parameter estimation. The 5 potential candidates for parameter estimation gives a total of approximately 31 parameter subset combinations, with a maximum subset size of 5. Of the 31 possible subset combinations, only 18 are potentially identifiable and with a maximum identifiable subset size of 3.
- ➢ Parameter estimation indicates that despite variation in parameter uncertainty, the quality of the model fit to experimental data is not compromised and using an identifiable subset of model parameters serves reduce the degree of uncertainty (confidence interval) in model parameter estimates.
- ➢ Following sensitivity analysis, the biodegradable volatile solids are most sensitive to the acid yield coefficient (k_2) while volatile fatty acids are most sensitivity to the Monod saturation constant for volatile acids (K_{s2}) and inhibition constant (K_i). For the anaerobic microorganisms, the acidogenic bacteria is highly sensitive to the Monod saturation constant (K_{s2}), while the methanogenic archae are highly sensitive to inhibition constant (K_i).
- ➢ Uncertainty quantification reveals that of the four model states, the methanogenic archaea, shows an insignificant uncertainty band to the model inputs at certain times instants, while all the other sates show a degree of significant uncertainty to the model inputs at all times instants
- ➢ The systematic model-based framework proposed in this study has been based on the concept of attainable regions. Hence, when using attainable regions for performance targeting and digester network synthesis, we suggest that it should be mandatory to incorporate uncertainty of model prediction during construction of the attainable regions. The attainable region obtained in such cases is referred to as a self-optimizing attainable region, which is generally smaller than the attainable region. It is concluded that incorporating kinetic uncertainty onto attainable regions has greatly reduces the limits of achievability by the system even though we have benefited from increased robustness. When the AR is obtained, the boundary of the AR can be interpreted into digester structures, whereby the optimal digester structure always involves a plug flow digester in combination with either a CSTR and/or bypass streams.

In summary coupling identifiability analysis, uncertainty quantification and the attainable region theory provides a systematic methodological framework for defining the performance targets of the anaerobic treatment process under conditions of uncertainty. It is also worth mentioning that even though the study is based on the anaerobic treatment process, the framework can be applied to optimally design other environmental chemical processes, which can be described with a kinetic model.

More research is needed to extend the concept of self-optimizing attainable regions in the field of anaerobic digestion. This study has focused on kinetic uncertainty and it would be interesting to assess the effects of other potential sources of uncertainty (such as substrate characteristics, presence of inhibitions or temperature variations) on the performance targets (defined by the self-optimizing attainable regions) of the anaerobic treatment process.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Our team expresses gratitude to the following institutions; The Brew-Hammond Energy Centre, KNUST Ghana, The Water and Environmental Engineering Group, NTNU Ålesund and the Abunde Sustainable Engineering Group (AbundeSEG) for its immense technical support.

Nomenclature

$X_{1\,in}$	Influent concentration of acidogenic bacteria ($g\ ac./L$)
A_f	Acidity factor ($g\ VFA/L$)/($g\ BVS/L$)
B_0	Biodegradability constant ($g\ BVS/L$)/($g\ VS/L$)
K_{d_1}	Specific death rate of acidogenic bacteria (d^{-1})
K_{d_2}	Specific death rate of methanogenic bacteria (d^{-1})
K_i	VFA inhibition constant ($g\ VFA/L$)
K_{i1}	VFA inhibition constant for acidogenic bacteria ($g\ VFA/L$)
K_{i2}	VFA inhibition constant for methanogenic bacteria ($g\ VFA/L$)
K_s	Monod half-saturation constant (g/L)
K_{s1}	Monod half-saturation constant for acidogenic bacteria ($g\ BVS/L$)
K_{s2}	Monod half-saturation constant for acidogenic bacteria ($g\ VFA/L$)
Q_{CH_4}	Methane gas flow rate ($L\ CH_4/d$)
$S_{1_{in}}$	Influent concentration of biodegradable volatile solids ($g\ BVS/L$)
$S_{2_{in}}$	Influent concentration of volatile fatty acids ($g\ VFA/L$)
S_1	Concentration of biodegradable volatile solids in bioreactor ($g\ BVS/L$)
S_2	Concentration of volatile fatty acids in bioreactor ($g\ VFA/L$)
$S_a(t)$	Matrix of absolute sensitivities
S_{in}	Influent concentration of volatile solids ($g\ VS/L$)
S_{nd}	Non-dimensional sensitivity
$X_{2_{in}}$	Influent concentration of methanogenic bacteria ($g\ me./L$)
X_1	Concentration of acidogenic bacteria in bioreactor ($g\ ac./L$)
X_2	Concentration of methanogenic bacteria in bioreactor ($g\ me./L$)
X_{in}	Influent biomass concentration (g/L)
k_1	Yield constant ($g\ BVS/g\ ac./L$)
k_2	Yield constant ($g\ VFA/g\ ac./L$)
k_3	Yield constant ($gVFA/g\ me./L$)

$t_{v,\alpha/2}$	Student t-distribution parameter
δ_k^{msqr}	Root mean squared sensitivity
μ_1	Specific growth rate of acidogenic bacteria (d^{-1})
μ_{1m}	Maximum specific growth rate of acidogenic bacteria (d^{-1})
μ_2	Specific growth rate of methanogenic bacteria (d^{-1})
μ_{2m}	Maximum specific growth rate of methanogenic bacteria (d^{-1})
μ_m	Specific growth rate of bacteria (d^{-1})
D	Dilution rate (d^{-1})
HRT	Hydraulic retention time (d)
S	Substrate concentration (g/L)
T	Reactor temperature ($^{\circ}C$)
V	Volume of methane bioreactor (L)
VFA_{in}	Influent concentration of volatile fatty acids ($g\ VFA/\ L$)
VS	Volatile solids
VSL	Volatile solids loading ($g\ VS/L$)
W	non-dimensional scaling/weighting matrix
f	Inhibition factor
ϑ	Acidogenic fraction
Y	Model states
β	Parameter set
Y_0	Initial Condition
N	Number of state variables
n	Number of data points
p	Number of parameters
S_{norm}	Normalized sensitivity measure
γ_k	Collinearity index
λ_k	Eigen values of normalized sensitivity matrix
$J(\beta)$	Least Square Criterion
$\nabla J(\beta)$	Gradient of Least Square Criterion
Y_i^{obs}	Observable states
H	Observation matrix
$M(t, Y, \beta)$	Discretized model
$L(Y, \beta, \lambda)$	Lagrangian
w	Descent direction
ρ	Step length

Funding

This work was supported by EnPe-NORAD under the project Upgrading Education and Research Capacity in Renewable Energy Technologies (UPERC-RET).

References

Abunde Neba, F., Asiedu, N.Y., Addo, A., Morken, J., Østerhus, S.W., Seidu, R., 2019a. A coupled modeling of design and investment parameters for optimal operation of methane bioreactors: attainable region concept approach. Renew. Energy.

Abunde Neba, F., Asiedu, N.Y., Addo, A., Morken, J., Østerhus, S.W., Seidu, R., 2019b. Simulation of Two-Dimensional Attainable Regions and its Application to Model Digester Structures for Maximum Stability of Anaerobic Treatment Process. Water Research, p. 114891.

Abunde Neba, F., Asiedu, N.Y., Addo, A., Morken, J., Østerhus, S.W., Seidu, R., 2019c. Use of attainable regions for synthesis and optimization of multistage anaerobic digesters. Appl. Energy 242, 334–350.

Abunde Neba, F., Asiedu, N.Y., Addo, A., Morken, J., Østerhus, S.W., Seidu, R., 2020a. Biodigester rapid analysis and design system (B-Rades): a candidate attainable region-based simulator for the synthesis of biogas reactor structures. Comput. Chem. Eng. 132, 106607.

Abunde Neba, F., Asiedu, N.Y., Addo, A., Seidu, R., 2020b. Attainable regions and fuzzy multi-criteria decisions: modeling a novel configuration of methane bioreactor using experimental limits of operation. Bioresour. Technol. 295, 122273.

Batstone, D.J., 2006. Mathematical modelling of anaerobic reactors treating domestic wastewater: rational criteria for model use. Rev. Environ. Sci. Biotechnol. 5, 57–71.

Benítez, M., Bermúdez, A., Rodríguez-Calo, J.F., 2017. Adjoint method for parameter identification problems in models of stirred tank chemical reactors. Chem. Eng. Res. Des. 123, 214–229.

Brun, R., Kühni, M., Siegrist, H., Gujer, W., Reichert, P., 2002. Practical identifiability of Asm2d parameters—systematic selection and tuning of parameter subsets. Water Res. 36, 4113–4127.

Chen, J.L., Ortiz, R., Steele, T.W.J., Stuckey, D.C., 2014. Toxicants inhibiting anaerobic digestion: a review. Biotechnol. Adv. 32, 1523–1534.

Chen, Y., Cheng, J.J., Creamer, K.S., 2008. Inhibition of anaerobic digestion process: a review. Bioresour. Technol. 99, 4044–4064.

Donoso-Bravo, A., Mailier, J., Martin, C., Rodríguez, J., Aceves-Lara, C.A., Wouwer, A.V., 2011. Model selection, identification and validation in anaerobic digestion: a review. Water Res. 45, 5347–5364.

Donoso-Bravo, A., Mailier, J., Ruiz-Filippi, G., Wouwer, A.V., 2013. Identification in an anaerobic batch system: global sensitivity analysis, multi-start strategy and optimization criterion selection. Bioproc. Biosyst. Eng. 36, 35–43.

EPA, U. S. E. P. A, 2006. Biosolids Technology Fact Sheet, Multi-Stage Anaerobic Digestion. National Service Center or Environmental Publications (NSCEP).

Finn, H., Rune, B., Bernt, L., 2013. Adapting dynamic mathematical models to a pilot anaerobic digestion reactor. Model. Identif. Control 34, 35–54.

Gausemeier, J., Frank, U., Schmidt, A., Steffen, D., 2006. Towards a design methodology for self-optimizing systems. In: Elmaraghy, H.A., Elmaraghy, W.H. (Eds.), Advances in Design. Springer London, London.

Henze, M., Van Loosdrecht, M.C.M., Ekama, G.A., Brdjanovic, D., 2008. Biological Wastewater Treatment. Iwa Publishing.

Hildebrandt, D., Glasser, D., 1990. The attainable region and optimal reactor structures. Chem. Eng. Sci. 45, 2161–2168.

Hildebrandt, D., Glasser, D., Crowe, C.M., 1990. Geometry of the attainable region generated by reaction and mixing: with and without constraints. Ind. Eng. Chem. Res. 29, 49–58.

Jäschke, J., Cao, Y., Kariwala, V., 2017. Self-optimizing control – a survey. Annu. Rev. Contr. 43, 199–223.

Kythreotou, N., Florides, G., Tassou, S.A., 2014. A review of simple to scientific models for anaerobic digestion. Renew. Energy 71, 701–714.

Ming, D., Glasser, D., Hildebrandt, D., Glasser, B., Metzer, M., 2016. Attainable Region Theory: an Introduction to Choosing an Optimal Reactor. John Wiley & Sons, Inc, Hoboken, New Jersey.

Morales-Rodriguez, R., Meyer, A.S., Gernaey, K.V., Sin, G., 2012. A framework for model-based optimization of bioprocesses under uncertainty: lignocellulosic ethanol production case. Comput. Chem. Eng. 42, 115–129.

Permin, E., Bertelsmeier, F., Blum, M., Bützler, J., Haag, S., Kuz, S., Özdemir, D., Stemmler, S., Thombansen, U., Schmitt, R., Brecher, C., Schlick, C., Abel, D., Poprawe, R., Loosen, P., Schulz, W., Schuh, G., 2016. Self-optimizing production systems. Procedia Cirp 41, 417–422.

Roulston, M.S., 1999. Data Assimilation in Chaotic systems,Astrophysical and Geophysical Flows as Dynamical Systems 1998 Summer Study Program in Geophysical Fluid Dynamics.

Sin, G., Gernaey, K.V., Lantz, A.E., 2009. Good Modeling Practice for Pat Applications: Propagation of Input Uncertainty and Sensitivity Analysis, vol 25, pp. 1043–1053.

Sin, G., Meyer, A.S., Gernaey, K.V., 2010a. Assessing reliability of cellulose hydrolysis models to support biofuel process design—identifiability and uncertainty analysis. Comput. Chem. Eng. 34, 1385–1392.

Sin, G., Meyer, A.S., Gernaey, K.V., 2010b. Assessing reliability of cellulose hydrolysis models to support biofuel process design—identifiability and uncertainty analysis. Comput. Chem. Eng. 34, 1385–1392.

Sin, G., Vanrolleghem, P.A., 2007. Extensions to modeling aerobic carbon degradation using combined respirometric–titrimetric measurements in view of activated sludge model calibration. Water Res. 41, 3345–3358.

Skogestad, S., 2000. Plantwide control: the search for the self-optimizing control structure. J. Process Control 10, 487–507.

Wang, L.K., Shammas, N.K., Hung, Y.-T., 2007. Biosolids Treatment Processes. Humana Press Inc, New Jersey.

Yu, L., Wensel, P.C., Ma, J., Chen, S., 2013. Mathematical modeling in anaerobic digestion (Ad). Bioremediation & Biodegradation, pp. 2155–6199.

Zaher, U., Pandey, P., Chen, S., 2009. A simple elemental continuity based model application to study the anaerobic microbial activity for the treatment of dairy manure. Appl. Math. Model. 33, 3553–3564.