Mathematical Biosciences and Engineering

https://www.aimspress.com/journal/mbe

MBE, 21(9): 7139–7164.
DOI: 10.3934/mbe.2024316
Received: 25 March 2024
Revised: 02 August 2024
Accepted: 08 August 2024
Published: 25 September 2024

Research article

Global sensitivity analysis and uncertainty quantification for a mathematical model of dry anaerobic digestion in plug-flow reactors

Daniele Bernardo Panaro[1], Andrea Trucchia[2], Vincenzo Luongo[1,*], Maria Rosaria Mattei[1] and Luigi Frunzo[1]

[1] Department of Mathematics and Applications "Renato Caccioppoli", University of Naples Federico II, via Cintia Monte S. Angelo, Naples 80126, Italy

[2] CIMA Research Foundation, via A. Magliotto 2, Savona 17100, Italy

* **Correspondence:** Email: vincenzo.luongo@unina.it; Tel: +39081675637.

Abstract: In many applications, complex biological phenomena can be reproduced via structured mathematical models, which depend on numerous biotic and abiotic input parameters, whose effect on model outputs can be of paramount importance. The calibration of model parameters is crucial to obtain the best fit between simulated and experimental data. Sensitivity analysis and uncertainty quantification constitute essential tools in the field of biological systems modeling. Despite the significant number of applications of sensitivity analysis in wet anaerobic digestion, there are no examples of global sensitivity analysis for mathematical models describing the dry anaerobic digestion in plug-flow reactors. For the first time, the present study explores the global sensitivity analysis and uncertainty quantification for a plug-flow reactor model. The investigated model accounts for the mass/volume variation that takes place in these systems as a result of solid waste conversion in gaseous value-added compounds. A preliminary screening based on the Morris' method allowed for the definition of three different groups of parameters. A surrogate model was constructed to investigate the relation between input and output parameters without running demanding simulations from scratch. The obtained Sobol' indices allowed to perform the quantitative global sensitivity analysis. Finally, the uncertainty quantification results led to the definition of the probability density function related to the investigated quantity of interest. The study showed that the net methane production is mostly sensitive to the values of the conversion parameter related to the particulate biodegradable volatile solids in acetic acid k_1 and to the kinetic parameter describing the acetic acid uptake k_2. The application of these techniques led to helpful information for model calibration and validation.

Keywords: anaerobic digestion; global sensitivity analysis; uncertainty quantification; plug-flow reactor modeling; partial differential equations

1. Introduction

The role of partial differential equations (PDEs) is fundamental in modeling and analyzing engineering biological processes that evolve over time and space within bioreactors. PDEs provide a robust mathematical framework for describing the spatial and temporal distribution of critical variables, such as nutrient concentrations, microbial populations, and metabolic byproducts. By solving these equations, it is possible to predict the dynamic behavior of these variables, facilitating the optimization of bioreactor design and operational parameters in complex and specific applications. This predictive capability is essential for enhancing process efficiency, achieving precise control over bioprocesses, and maximizing the yield of desired bioproducts. Anaerobic digestion (AD) is a well known biotechnology for waste and wastewater treatment, in which the metabolic activities of specific microorganisms are used to convert complex organic substrates into a value-added gaseous product. During the AD process, complex organic materials are progressively degraded into simpler organic compounds, until a methane rich biogas is obtained as the final product. The biogas can be further utilized for clean energy generation in accordance to environmental policies. Consequently, AD can be considered one of the leading technologies in the context of biorefining systems. Moreover, it is important to account for additional by-products usually contained in the residual non-gaseous effluent, i.e., digestate of the AD process. The digestate is usually rich in nutrients and micro-elements, and it can be successfully used as a fertilizer for agricultural scopes [1, 2]. However, some digestate constituents are able to negatively affect the composition of soils and lead to undesired environmental impacts. In some cases, additional treatments are recommended, such as filtration procedures [3], to reduce the content of polluting components or to recover dissolved compounds from the liquid fraction of the digestate. The correct management of digestate, from its generation to the final disposal, represents a crucial point for a sustainable application of the AD process. Its correct use and utilization allow to fulfill the objectives of the Kyoto Protocol and EU Policies concerning organic waste disposal and renewable energy production.

Based on the solid content of the fed bio-waste, AD can be performed in dry, semi-dry, or wet conditions. In dry AD, the total solids (TS) content is higher than 15%, while in semi-dry AD, the TS content ranges between 10 and 15%. Indeed, the most frequently adopted wet AD is used for diluted bio-wastes with a TS content lower than 10% [4]. Over the last decades, high attention has been devoted to the dry AD as its application leads to several advantages compared to wet systems. Small reactor volume, low water addition, easy management of digested by-products, and reduced pre-treatment requirement and nutrient loss are crucial design aspects of dry AD [5, 6].

The most used configuration for AD processes are continuously stirred tank reactor (CSTR) and plug-flow reactor (PFR). CSTR is an ideal hydrodynamic model where perfect mixing is hypothesized. For this reason, the concentration of soluble and particulate compounds is supposed to be homogeneous in each point of the reactor [7]. In a real-scale AD plant, this ideal condition is warranted by continuous mixing systems; the more efficient is the mixing system, the closer is the real-scale reactor to the CSTR model. The PFR is characterized by a preferential axial direction in which a continuous flow of waste is forced to pass. The PFR scheme is constituted by a cylindrical/tubular shape reactor where the concentration of different components is a spatial variable in the axial direction. Conversely, soluble and particulate compound concentrations along the radial direction are supposed to be homogeneous. For this reason, the total volume of a PFR can be schematized as an infinite number of in series CSTR [8]. This leads to higher average substrate concentrations and related conversion rates, which could be

beneficial for bioreactor design and performances.

In real-scale AD plants, CSTR configuration is commonly used for wet AD [9], while PFR is adopted for dry AD due to the higher viscosity of the medium and the consequent complexity of mixing procedures [10]. The implementation of wet or dry systems has similar costs in terms of investment and management. On the other hand, the two systems are substantially different from an environmental point of view. Typically, wet AD leads to the consumption of around $1m^3$ of fresh water per ton of waste, e.g., the organic fraction of municipal solid waste (OFMSW), while the water demand of dry AD reactors is ten-fold lower. For this reason, the volume of digestate produced by dry AD processes is several-fold lower than that generated by wet AD systems [6].

1.1. Modeling of AD

Due to the complex biological interaction occurring in AD processes, the role of mathematical modeling is crucial for the development of management/optimization tools able to predict methane production in full-scale applications. The anaerobic digestion model 1 (ADM1) [11] is one of the most popular mathematical models for the prediction of AD dynamics in wet conditions. From its first publication in 2002, several authors were inspired by the ADM1 approach and proposed different modifications to account for specific mechanisms occurring in the AD process or to include relevant phenomena that were neglected in the original version. For instance, the variation of kinetics and stoichiometric parameters due to different pressure levels in AD bioreactors has been recently addressed in [12]. Mathematical models of wet AD are typically based on ordinary differential equation (ODE) systems [13]. Since each state variable only depends on time, they are homogeneous in the reactor domain. On the other hand, the mathematical problem of PFR is described by a PDE system, where the state variables are functions of space and time. This aspect is relevant as it highly increases the complexity of the mathematical modeling of dry AD processes. The effects of turbulence, dispersion, and accumulation of particles in bioreactors strongly affect dry AD performance, as these phenomena are naturally characterized by great uncertainty levels. For this reason, it is possible to state that the modeling of dry AD is a less explored field than wet AD. Some authors avoided to solve the complete PDE system by schematizing the PFR as a number of n in series CSTR [8, 14, 15]. Other attempts considered a fixed velocity inside the reactor [16] or coupled computational fluid dynamics analysis with a steady-state kinetic model described with ODE [17].

Recently, Panaro et al. [18] proposed an original one-dimensional mathematical model for dry AD in PFR systems. The model was based on mass balance principles for selected state variables resulting in a non-linear PDE system. The convective velocity of the substrate moving along the reactor represented a further unknown, and the model was able to consider the mass/volume variation of the waste along the reactor. Indeed, the solids conversion in gaseous compounds led to a relevant mass variation in dry AD, which in wet systems is negligible due to the high quantity of water in the system. In the same work, the velocity variation was obtained under two main hypotheses: i) the density of the waste substrate moving along the reactor was supposed constant over time, and ii) the volume fractions of the bio-components constituting the treated substrate were constrained to sum up to the unity. These statements implied that the mass of the substrate mixture was constant along the reactor, and the mass reduction of volatile solids (VS) due to the degradation processes was continuously balanced by the variation of velocity. Under these hypotheses, the crucial role of relatively low convective velocity as well as the contextual effect of diffusion and biological reactions in a PFR was addressed. It represents the first mathematical

model able to completely describe the bio-physics of the dry anaerobic conversion accounting for the reactor geometry, substrate characteristics, and microbial activities and interactions.

1.2. Quantification of uncertainty and sensitivity analysis

In many mathematical models of biosystems, global sensitivity analysis (GSA) plays a critical role by identifying key parameters that significantly impact model outputs. This analysis helps in simplifying complex models by focusing on the most influential variables, thereby making the models more manageable and interpretable. GSA also quantifies the impact of uncertainties in input parameters, providing insights into the robustness and reliability of model predictions. Additionally, GSA guides model validation and calibration by highlighting discrepancies between model outputs and experimental data, leading to more accurate and predictive models.

In mathematical models including several governing equations, the role of model parameters is crucial to capture specific real world dynamics [19]. To better fit the experimental data, adequate calibration procedures are required, and sensitivity analysis (SA) methods, e.g., GSA techniques, are considered powerful tools to obtain a meaningful screening of model parameters relevance [20]. Indeed, several works explored the input space along mono-dimensional corridors, e.g., local sensitivity analysis (LSA), with a consequent neglect of a significant part of the space of parameters. In their literature review, the authors noticed that 42% of the cited papers do not satisfy the requirements to properly explore the space of input parameters. It is clear that the introduction of high-quality practices and standard procedures is essential to the fields of SA and uncertainty quantification (UQ).

Despite many mathematical models focused on the AD process, only a few studies include GSA. The majority of available works focused on local procedures, thus neglecting more exhaustive global techniques [21]. Several examples of LSA procedures can be found in literature. Tartakovsky et al. [22] used relative sensitivity functions to reduce the number of parameters of an ADM1-based model of an up-flow anaerobic sludge blanket (UASB) reactor; Noykova and Gyllenberg [23] used a similar approach in the case of a mathematical model for an anaerobic process influenced by substrate inhibition; Bernard et al. [24] performed a LSA for an AD process model in an upflow fixed bed reactor (FBR) using sensitivity coefficients evaluated for the chosen output variable; and Vavilin et al. [25] adopted a similar approach for a distributed model of solid waste AD in batch and continuous reactors. They identified the key parameters by changing their baseline values by a certain percentage. Similarly, Lin and Wu [26] used the least-square (LS) method for the parameters of an AD model including the degradation of phenolic compounds. However, these works are based on analysis that overlook the effect of the possible interactions between parameters on model outputs.

On the other hand, only a few works focused on the GSA of AD models, and they were all related to wet AD in CSTR. Solon et al. [27] performed a GSA on the ADM1 implemented in the context of the benchmark simulation model no. 2 (BSM2) using the standardized regression coefficients method and the Morris' elementary effects screening method. They tested different reactor configurations and concluded that sensitivity results were more accurate when AD was performed at a low sludge retention time (SRT) and mesophilic conditions (around 35°C). In the open-loop version of the BSM2, a GSA study on the ADM1 using Monte Carlo (MC) tests and linear regression was proposed [28]. They focused on the reduction of computational time and on the optimal number of simulations to achieve an acceptable sensitivity analysis result. They highlighted the importance of the numerical solver optimization and stated that the optimal number of simulations to warrant an accurate GSA result

depends on the number of investigated parameters, that should be multiplied by 50. Zonta et al. [29] used a Bayesian SA tool to evaluate the variance-based SA index in a ADM1-based model accounting for inhibitory phenomena. Carrera-Chapela et al. [30] developed a simplified mathematical model for the AD process to account for hydrogen sulphide generation and related gaseous emissions. They carried out a GSA using ANOVA decomposition and determined the higher sensitive parameters on the model output by changing the inlet flow and the substrate concentration in the reactor. Based on the sensitivity profiles, they selected the less influenced parameters using a collinearity analysis.

To the best of our knowledge, there are no examples of GSA related to PDE models describing dry AD in PFR. Indeed, the objective of the present work is to apply GSA and UQ techniques to perform an exhaustive GSA of the model presented in [18], as it is uniquely able to account for the effective mass variation occurring during the anaerobic biological conversion in PFR. The screening of model parameters represents a helpful reference for further calibration and validation works related to mathematical modeling of dry AD, in particular for real-scale PFR applications. Moreover, the adopted GSA technique allowed to highlight the most influencing aspects affecting model outputs and results. The present work is organized into six main sections. The complete model of dry anaerobic digestion in PFR is recalled in Section 2, and the model parameters and QoI selection for SA and UQ studies are introduced in Section 3. Moreover, that section outlines the test case and the databases used for simulations. Section 4 describes the Morris' screening and the surrogate-based SA, which are the adopted methods for UQ and SA. Section 5 synthesizes the main results of the study, and Section 6 contains remarks and future perspectives.

2. Model of plug-flow reactor

The one-dimensional mathematical model proposed in [18] describes the evolution of soluble and particulate components in a PFR. The state variables represent the concentrations of these components and are functions of time t and space, where z represents the spatial coordinate, which is oriented along the reactor axis and directed from the inlet to the outlet section. For each state variable, a convection-diffusion-reaction equation describes the transport, along the reactor, and the related bio-chemical transformation. The convective and diffusive fluxes are characterized by the velocity $v(z,t)$ and by a constant diffusion coefficient D along the z direction, respectively. The feeding waste substrate was composed by different components X_i with the same density $\rho(z,t) = \rho$. Specifically, it was constituted by water X_1, particulate inert X_2, biodegradable X_3 and non-biodegradable X_4 VS, microbial biomass X_5, and additional dissolved compounds (i.e., soluble acetic acid S_1 and soluble methane S_2). The produced biogas was supposed to be constituted of methane gas G, and it was collected in the head-space of the reactor.

The kinetic scheme describing the bio-conversion processes occurring in dry AD is reported in Figure 1. The degradation of the biodegradable VS fraction X_3 in soluble acetic acid S_1 is regulated by the kinetic rate r_1, which considers the disintegration process as the rate-limiting step for acetic acid generation; S_1 is further consumed through the non-linear Monod-type kinetic rate r_2 producing both soluble methane S_2 and microbial biomass X_5; the microbial biomass decay r_4 leads to the production of further biodegradable X_3 and non-biodegradable X_4 VS; finally, soluble S_2 and gaseous G methane is regulated by a gas-transfer law with the kinetic rate r_3. Note that the only microbial biomass explicitly considered as a particulate component in the model is X_5, representing methanogenic bacteria using

acetate S_1 for methane production.

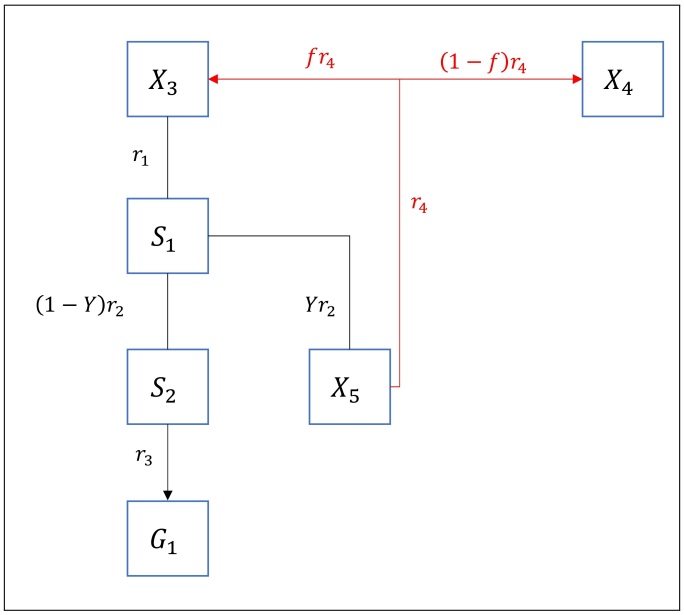


Figure 1. Kinetic scheme of AD conversion processes.

The convection-diffusion-reaction equations related to each compound constitute the non-linear PDE systems (2.1) and (2.2):

$$\frac{\partial X_h(z,t)}{\partial t} + \frac{\partial (v(z,t)X_h(z,t))}{\partial z} - D\frac{\partial^2 X_h(z,t)}{\partial z^2} = F_{X,h}\left(z,t,\mathbf{X},\mathbf{S},G\right), \quad 0 < z < L, \quad t > 0, \quad h = 1,...,5, \quad (2.1)$$

$$\frac{\partial S_l(z,t)}{\partial t} + \frac{\partial (v(z,t)S_l(z,t))}{\partial z} - D\frac{\partial^2 S_l(z,t)}{\partial z^2} = F_{S,l}\left(z,t,\mathbf{X},\mathbf{S},G\right), \quad 0 < z < L, \quad t > 0, \quad l = 1,2, \quad (2.2)$$

where:

- $\mathbf{X} = (X_1,...,X_5)$;
- $\mathbf{S} = (S_1, S_2)$;
- $F_{X,h}\left(z,t,\mathbf{X},\mathbf{S},G\right)$, $h = 1,...,5$, is the reaction term of the compound X_h;
- $F_{S,l}\left(z,t,\mathbf{X},\mathbf{S},G\right)$, $l = 1,2$, is the reaction term of the dissolved compound S_l.
- $F_{X,1}\left(z,t,\mathbf{X},\mathbf{S},G\right) = F_{X,2}\left(z,t,\mathbf{X},\mathbf{S},G\right) = 0$;
- $F_{X,3}\left(z,t,\mathbf{X},\mathbf{S},G\right) = fr_4 - r_1$;
- $F_{X,4}\left(z,t,\mathbf{X},\mathbf{S},G\right) = (1-f)r_4$;
- $F_{X,5}\left(z,t,\mathbf{X},\mathbf{S},G\right) = Yr_2 - r_4$;
- $F_{S,1}\left(z,t,\mathbf{X},\mathbf{S},G\right) = m(r_1 - r_2)$;

- $F_{S,2}(z,t,\mathbf{X},\mathbf{S},G) = m(1-Y)r_2 - r_3$;
- D is the diffusion coefficient;
- $r_1 = k_1 X_3$;
- $r_2 = k_2 X_5 S_1/(K_1 + S_1)$;
- $r_3 = k_3(S_2 - RTK_H G)$;
- $r_4 = k_4 X_5$;
- m is the conversion factor of VS in COD $\left[g_{COD}\, g_{VS}^{-1}\right]$;
- Y is the yield of biomass on substrate;
- f is the fraction of dead microbial biomass converted in biodegradable substrate;
- k_1 is the rate of volatile solids X_3 consumption, with the dimension of $\left[T^{-1}\right]$;
- k_2 is the Monod maximum specific uptake rate for the acetic acid $\left[T^{-1}\right]$;
- K_1 is the half-saturation constant $\left[M\, L^{-3}\right]$ for the acetic acid;
- k_3 is the gas-liquid transfer coefficient $\left[T^{-1}\right]$;
- R is the gas law constant $\left[L^2\, T^{-2}\, \Theta^{-1}\right]$;
- T is the operating temperature $[\Theta]$;
- K_H is the Henry's law coefficient $\left[L^2\, T^{-2}\right]$;
- k_4 is the first-order decay rate of the microbial biomass X_6 $\left[T^{-1}\right]$.

The velocity displacement in the preferential direction of the reactor constitutes a further unknown of the problem, and the equation describing its variation is derived by assuming, for any time and space:

$$\begin{cases} \sum_{h=1}^{5} X_h(z,t)/\rho_h = 1, \\ \rho_h = \rho, \quad h = 1,...,5 \end{cases} \quad \Longrightarrow \quad \sum_{h=1}^{5} X_h(z,t) = \rho. \tag{2.3}$$

Equation (2.3) implies that the total mass of the waste (composed of water, inert, VS, and microbial biomass) is constant over time. This results in varying the convection velocity along the reactor. Its variation depends on the kinetics of the mixture's compounds. To keep the mass constant, the velocity variation must balance the VS consumption.

By summing Eq (2.1) on $h = 1,...,5$ and considering Eq (2.3),

$$\frac{\partial v(z,t)}{\partial z} = \frac{Yr_2 - r_1}{\rho}. \tag{2.4}$$

In addition, the mass balance on the head-space volume $V_{gas} = A_{gas}L$ leads to an additional differential equation describing the dynamics of the gaseous methane $G(t)$ (Eq (2.5)):

$$\frac{dG(t)}{dt} = \frac{A}{V_{gas}} \int_0^L r_3(z,t)\ dz, \tag{2.5}$$

where A is the constant cross-sectional area occupied by the treated substrate. Equation (2.5) accounts for the total gas-transfer rate in the PFR. The ratio between A and the volume of gas V_{gas} connects the gas-transfer kinetic rate to the volume of the treated substrate.

Finally, boundary and initial conditions are prescribed according to Eqs (2.6) to (2.13):

$$v(0,t) = v_0, \quad t \geq 0, \tag{2.6}$$

$$-D\frac{\partial X_h(0,t)}{\partial z} = v_0(X_{h,IN} - X_h(0,t)), \quad h = 1,...,5, \quad t > 0, \tag{2.7}$$

$$\frac{\partial X_h(L,t)}{\partial z} = 0, \quad h = 1,...,5, \quad t > 0, \tag{2.8}$$

$$-D\frac{\partial S_l(0,t)}{\partial z} = v_0(S_{l,IN} - S_l(0,t)), \quad l = 1,2, \quad t > 0, \tag{2.9}$$

$$\frac{\partial S_l(L,t)}{\partial z} = 0, \quad l = 1,2, \quad t > 0, \tag{2.10}$$

$$X_h(z,0) = X_{h,0}, \quad h = 1,...,5, \quad 0 \le z \le L, \tag{2.11}$$

$$S_l(z,0) = S_{l,0}, \quad l = 1,2 \quad 0 \le z \le L. \tag{2.12}$$

$$G(0) = G_0. \tag{2.13}$$

The value v_0 in Eq (2.6) can be obtained by fixing the hydraulic retention time (HRT) of the components moving along the reactor:

$$v_0 = \frac{L}{HRT}. \tag{2.14}$$

In general, HRT refers to the mean residence time of a certain substrate within a biological reactor. This concept, usually used in CSTR, is equally applicable to PFR, wherein the HRT is intrinsically linked to the convective velocity of all compounds traversing the reactor. As a consequence, the HRT in a PFR can be defined as a function of the physical dimensions (L) and flow dynamics within the reactor system. In Eqs (2.7) and (2.8), $X_{h,IN}$ and $S_{l,IN}$ are the concentrations of water or particulate components, and dissolved compounds in the incoming flow rate, respectively.

It is important to emphasize that, unlike conventional ADM1-based models, the only microbial biomass described in the present model is methanogenic bacteria that utilize acetate as the organic source for methane production. This approximation can appear highly limiting, as it fails to account for the various fermentation pathways and associated byproducts accumulating in an AD reactor. Specifically, the inability to predict the dynamics of all classical anaerobic species restricts the determination of crucial compounds, such as volatile fatty acids, which are known to cause pH inhibition processes affecting system dynamics. Nevertheless, when limited information is available about microbial species in an AD bioreactor, the model can still successfully reproduce real-case experimental data. An illustrative example is provided in [18], where model parameters were adjusted to replicate lab-scale experimental data, with Tables 1 and 2 listing all the parameters used for data fitting and related boundary conditions.

3. Uncertainty sources, quantities of interest, and databases

In the presented mathematical model, the state variables whose dynamics are described by the system of Eqs (2.1)–(2.3) and (2.5), coupled with boundary (2.6)–(2.10) and initial conditions (2.11)–(2.13), depend on a significant set of parameters. As recently reported [18], the knowledge about model responses to the variation of parameters is crucial for calibration and validation procedures based on real-world data. In the present investigation, seven different parameters have been selected as reported in Table 1. The first group includes i) the first-order kinetic constant for the conversion process of biodegradable VS in acetic acid k_1, ii) the Monod maximum specific uptake rate for acetate k_2, iii) the

first-order decay rate for the microbial biomass k_4, iv) the half-saturation constant K_1, and v) the yield of biomass on substrate Y linked to bio-chemical processes. The second group is constituted by vi) the gas-transfer coefficient k_3, and vii) the diffusion coefficient D, which was assumed constant for all the components. For the input parameters, uniform distributions were assumed in their physically meaningful literature range [11, 16].

Table 1. Parameters investigated and related variation ranges.

Parameter	Description	Range	Value	Reference
k_1	VS degradation rate	0.005-0.5	0.1 d^{-1}	[18]
k_2	Acetate maximum uptake rate	0.080-8.0	8.00 d^{-1}	[11]
k_3	Gas transfer coefficient	0.300-300	20 d^{-1}	[18]
k_4	Microbial biomass decay rate	0.001-0.05	0.02 d^{-1}	[11]
K_1	Half-saturation constant	0.015-1.5	0.15 $g_{COD}l^{-1}$	[11]
D	Diffusion coefficient	10^{-8}-10^{-6}	10^{-7} m^2s^{-1}	[18]
Y	Yield of biomass on substrate	0.04-0.1	0.05	[11]

3.1. Quantity of interest

The PFR model describes the AD process evolving in time $t \in (0, \tau)$ and space $x \in (0, L)$. It also accounts for the already mentioned key variable $G(t)$, which is the biologically produced methane described as a function of time, $X_h(z, t)$, $h = 1, ..., 5$, and $S_l(z, t)$, $l = 1, 2$, which are spatially distributed state variables. To gain a deeper understanding of how uncertain inputs influence the behavior of the PFR model, it is crucial to concentrate on a select set of scalar outputs.

A specific quantity of interest (QoI) has been considered in this study; denoting with y the time integral from $t = 0$ and $t = \tau$ of the total methane volume accumulated in the head-space of the reactor at normal temperature and pressure conditions, it is possible to obtain the equation

$$y = \int_0^\tau \frac{RT_0}{P_0} V_{gas} \frac{G(t)}{64} dt \qquad [l_{CH_4} d], \tag{3.1}$$

where:

- $G(t)$ is the gaseous methane concentration expressed in $g_{COD}\ l_{gas}^{-1}$;
- T_0 and P_0 represent the values of temperature and pressure at normal conditions;
- 64 is a stoichiometric coefficient expressing the grams of COD per mole of methane.

The selection of the QoI is strictly connected to the main objective AD process: to realize an efficient biological treatment for the waste organic compounds fed to the PFR by producing methane. Hence, the main goal of the present analysis is the reduction of the uncertainty related to this specific model output.

3.2. Description of test case

To achieve useful information and significant datasets for the SA, a test case was built for numerical simulations. The test case was based on experimental evidence and procedures aimed at maximizing the methane production from AD in a real PFR. Indeed, many experimental studies focused on the measured methane yield variation when different feeding or environmental conditions were applied to bioreactors,

such as different organic loading rates (OLR) or solid contents, different pretreatment on the organic substrate, and different working temperature [31]. These parameters can strongly affect bioreactor design and performances, as well as uncertainty quantification results. In the present study, the more frequent parameters adopted for real-scale PFR were selected. More precisely, OLR is a measure of the organic compounds fed to the bio-reactor, evaluated as the ratio between the concentration of VS in the inlet flow and the HRT, which represents the average treatment time of fed components in the AD process. All these aspects were fundamental to the test case definition.

Table 2 reports the value of the defined geometric and operational parameters adopted in this study. These parameters include the PFR length L, the cross-sectional area A, the reactor head-space volume V_{gas}, the HRT, the inlet convective velocity v_0, the OLR, and the temperature T. The conversion coefficient of VS in COD was fixed at $m = 1.5\ g_{COD}\ g_{VS}^{-1}$, while the fraction of dead microbial biomass converted in biodegradable substrate was set at $f = 0.2$. The composition of the feeding substrate was chosen in accordance to lab-scale experiments and is reported in Table 3. The initial substrate consisted of a mixture of biodegradable, non-biodegradable, and microbial biomass solids, the latter constituting the microbial biomass required as the inoculum for the anaerobic process. Initial and boundary conditions used for simulations are summarized in Table 4. The simulation time was fixed to 120 d, as it is a common practice to suppose that the reactor stabilization is achieved when three HRT have been performed by the PFR in the same working conditions ($HRT = 40\ d$).

Table 2. Operating parameters of the test case.

Parameter	Description	Dimension	Value
L	Reactor length	m	1.34
A	Reactor cross-section	m^2	0.0224
V_{gas}	Volume of the head-space	l	10.0
HRT	Hydraulic retention time	d	40.0
v_0	Incoming flow rate velocity	$cm\ d^{-1}$	3.35
OLR	Organic loading rate	$g_{VS}\ l^{-1}\ d^{-1}$	6.0
T	Reactor temperature regime	°C	37.0

Table 3. Initial mixture and inlet substrate characterization.

Parameter	Dimension	Initial mixture	Inlet substrate
Total solids	$g_{TS}\ g^{-1}$	0.15	0.30
Volatile solids	$g_{VS}\ g_{TS}^{-1}$	0.80	0.80
Biodegradable VS	$g_{VS}\ g_{VS}^{-1}$	0.07	0.396
Non-biodegradable VS	$g_{VS}\ g_{VS}^{-1}$	0.905	0.60
Microbial biomass	$g_{VS}\ g_{VS}^{-1}$	0.025	0.004

Table 4. Initial conditions and inlet flow concentrations used in model simulations.

Parameter	Symbol	Unit	Value
Density of the treated substrate	ρ	$g\ l^{-1}$	1000.0
Initial H_2O concentration	$X_{1,0}$	$g\ l^{-1}$	850.0
Initial inert concentration	$X_{2,0}$	$g\ l^{-1}$	30.0
Initial biodegradable VS concentration	$X_{3,0}$	$g_{VS}\ l^{-1}$	8.4
Initial non-biodegradable VS concentration	$X_{4,0}$	$g_{VS}\ l^{-1}$	108.6
Initial soluble acetic acid concentration	$S_{1,0}$	$g_{COD}\ l^{-1}$	0.0
Initial soluble methane concentration	$S_{2,0}$	$g_{COD}\ l^{-1}$	0.0
Initial microbial biomass concentration	$X_{5,0}$	$g_{VS}\ l^{-1}$	3.0
Initial gas-phase methane concentration	G_0	$g_{COD}\ l^{-1}$	0.0
Inlet H_2O concentration	$X_{1,IN}$	$g\ l^{-1}$	700.0
Inlet inert concentration	$X_{2,IN}$	$g\ l^{-1}$	60.0
Inlet biodegradable VS concentration	$X_{3,IN}$	$g_{VS}\ l^{-1}$	95.04
Inlet non-biodegradable VS concentration	$X_{4,IN}$	$g_{VS}\ l^{-1}$	144.0
Inlet soluble acetic acid concentration	$S_{1,IN}$	$g_{COD}\ l^{-1}$	0.0
Inlet soluble methane concentration	$S_{2,IN}$	$g_{COD}\ l^{-1}$	0.0
Inlet microbial biomass concentration	$X_{5,IN}$	$g_{VS}\ l^{-1}$	0.96

3.3. Experimental designs and databases

By using a design of experiments, the input space $Z_\Theta \in \mathbb{R}^d$ (hypercube) of uncertain parameters ($d = 7$ is the number of parameters investigated) was discretized. In this way, N realizations of the parameters $\boldsymbol{\theta}_i$ were defined. The PFR model was treated as a "black box" to generate N functional outputs $\mathbf{y}$. The functional outputs were further analyzed, and useful statistics were obtained. The ensemble was compiled into a database $\mathcal{D}_N$:

$$\mathcal{D}_N = \left\{\left(\boldsymbol{\theta}^{(l)}, \mathbf{y}^{(l)}\right)_{1\le l\le N}\right\}, \tag{3.2}$$

where $\mathbf{y}^{(l)} = \mathcal{F}\left(\boldsymbol{\theta}^{(l)}\right)$ is the integration of the PFR model $\mathcal{F}$ obtained by fixing the l^{th} set of input parameters $\boldsymbol{\theta}^{(l)}$.

Table 5. Datasets $\mathcal{D}_N$ of the PFR model [32] simulations for the Morris screening, the building of surrogates ("training"), and for the validation ("validation").

Method	Aim	Sample size
Randomized algorithm	Morris screening	400
Halton's sequence	Surrogate training	2^{10}
Faure's sequence	Surrogate validation	2^{10}

In this study, the parameter set used to carry out the screening $\boldsymbol{\theta}_{\text{Morris}}$ consisted of $N_M = 400$ samples. It was generated using the random algorithm proposed in [32]. For the surrogate-based study, quasi-

Monte Carlo sampling methods were selected to compile two different databases of size $N = 2^{10}$. Halton's sampling was used to generate the first database, which was utilized as a *training* set. Faure's sampling was used to generate the second database, or *validation* database. Indeed, the latter was used to investigate the precision of various surrogate techniques. Table 5 summarizes the compiled databases. Remarkably, the considered PFR model features significant non-linearities for the QoI y when $\boldsymbol{\theta}$ varies in Z_Θ. Figure 2 portrays 40 representative PFR model output snapshots sampled from the Morris database.

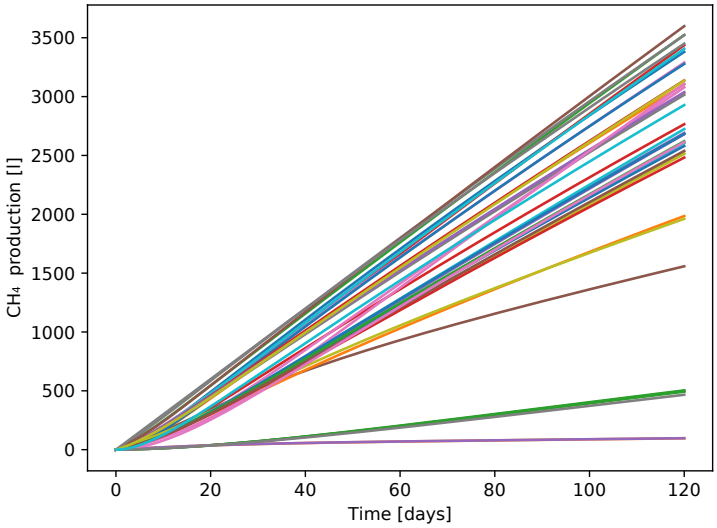


Figure 2. CH_4 profiles from the Morris' algorithm sampling database with different values of θ.

4. Surrogate-based SA

A two-phase SA was conducted on the PFR model: 1) the more rapid preliminary screening process using the *elementary effect test* (EET), proposed by Morris in [32]; and 2) the exhaustive surrogate-based GSA (4.2), aimed to achieve more accurate results. All GSA computations were pursued with the Python package *OpenTURNS* [33] (see www.openturns.org). Moreover, the PFR model was developed and integrated in the `MATLAB` platform. An original software, based on the finite difference upwind method, was used for numerical simulations. The method is conditionally stable and it was implemented according to [18].

4.1. Morris' scheme screening

According to Morris [32], an effective screening sensitivity measure for identifying a model's most significant parameters is the elementary effect (EE). Based on the computation of incremental ratios (the EEs), whose mean is used to measure global sensitivity, the method determines the overall significance of each input parameter on the QoI. The analysis was conducted by adopting randomized single-variable experiments, i.e., One-At-Time (OAT). These involve systematically changing one variable at a time while keeping others constant, with the order of changes randomized to reduce bias and better understand

the impact of each variable on the outcome. Initially, the input parameters were initially assumed to be uniformly distributed within [0, 1], and then their distributions were mapped from the unit hypercube to their corresponding distributions.

For a given value of $\theta_i \in \boldsymbol{\theta}$, the elementary effect EE_i can be expressed as:

$$EE_i = \frac{y(\theta_1^*, \ldots, \theta_i^* + \delta_i, \ldots, \theta_d^*) - y(\theta_1^*, \ldots \theta_i^*, \ldots, \theta_d^*)}{\delta_i}, \tag{4.1}$$

where $\delta \in \left\{\frac{1}{n_l - 1}, 1 - \frac{1}{n_l - 1}\right\}$, n_l is the number of levels, $\boldsymbol{\theta}^* = (\theta_1^*, \ldots, \theta_d^*)$ is a random value in the hypercube Z_θ such that the point $(\boldsymbol{\theta}^* + \boldsymbol{e}_i\delta)$ still maps to a point in Z_θ for each $i \in 1, \ldots, d$, and $\boldsymbol{e}_i$ is a zero valued vector except for its i^{th} component $e_i = 1$. For each input parameter θ_i, the empirical distribution of elementary effects EE_i was derived, with a random sampling of $\boldsymbol{\theta}$, s.t. $EE_i \sim F_i$, and its mean μ_i, and standard deviation σ_i were used as sensitivity measures. Moreover, the correction to the μ_i introduced in [34] was adopted. To avoid eliminating increments of different signs, their alternative measure employs the absolute value of the EEs instead of the mean μ_i. Their measure can be written as:

$$S_i = \mu^*_{\text{Morris}} = \frac{1}{n}\sum_{j=1}^{n} EE^j = \frac{1}{n}\sum_{j=1}^{n} \left| \frac{y(\theta_1^j, \ldots \theta_i^j + \delta_i^j, \ldots, \theta_d^j) - y(\theta_1^j, \ldots \theta_i^j, \ldots, \theta_d^j)}{\delta_i^j} \right| c_i. \tag{4.2}$$

To avoid large databases of experiments and/or simulations, it is a common practice to sample r elementary effects from each F_i and constructing r trajectories of $(d + 1)$ points in the input parameter space, which provides d elementary effects. The total cost of the experimental set is thus $r(d + 1)$ model evaluations. Apart from the mean (and its eventual corrections), standard deviation of the EEs provides useful information. It may constitute a proxy of the level of interaction between the parameters and it allows to understand if a certain factor has non-linear effects on the QoI.

4.2. Surrogate modeling

An emulator (also known as metamodel, or surrogate model) of the PFR model described in Section 2 was built by adopting two distinct algorithms: i) the generalized polynomial chaos (gPC) expansion, and ii) the Gaussian process (GP) model. Indeed, gPC-expansion and GP model are robust and widely spread techniques, which are well described in literature. In particular, the mathematical setting is detailed in [35, 36] and the previous work of Roy et al. [37]. Both approaches created a surrogate for the QoI y using a (finite) sum of basis functions:

$$y = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \gamma_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}. \tag{4.3}$$

The Halton's training database $\mathcal{D}_N$ with $N = 2^{10}$ (see Section 3.3) was used to determine the coefficients $\{\gamma_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}}$ and the basis functions $\{\Psi_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathcal{A}}$ in Eq (4.3). The coefficients can be determined using different methods.

Three different algorithms were tested in the present work: two variations of the gPC-expansion and an application of the GP. While performing a polynomial chaos, the basis functions of Eq (4.3) are multivariate orthonormal polynomial functions (see e.g., [38]). The two implementations of gPC differ by the rule that determines the finite sum of polynomial basis in Eq (4.3). One variant of gPC, referred to as standard least squares, used a linear truncation scheme to guide the choice of the polynomial basis,

while the second attempt, based on least angle regression, employs a sparse strategy for the truncation scheme [39]. Finally, the execution of the GP used the squared exponential (radial basis function) as the correlation kernel [40]. The algorithms to compute a gPC-expansion and GP surrogate are reported in [41] and adapted for the specific application. For the GP surrogate, the radial basis function (RBF) kernel was used according to [35].

5. Results

5.1. Morris' screening

Figure 3 and Table 6 summarize the results of Morris' screening procedure applied to the QoI y. The parameters were classified in three groups in terms of relevance. First, the most relevant parameters for the specific QoI are the kinetic constant related to the conversion process of the biodegradable VS in acetic acid k_1 and the Monod maximum uptake rate of acetic acid k_2. These results were expected from a biological point of view, as k_1 affects the kinetics of the disintegration process, which is one of the most limiting bio-process in AD, and determines the conversion rate of the substrate in intermediate products [42]. Moreover, k_2 directly affects the methane production rate and is recognized as one of the most influencing parameter in AD modeling [43]. However, the results suggest that it could be appropriate to focus future research on the disintegration kinetics modeling occurring in this case, and take into account the particle size distribution of biodegradable components as in Panaro et al. [44].

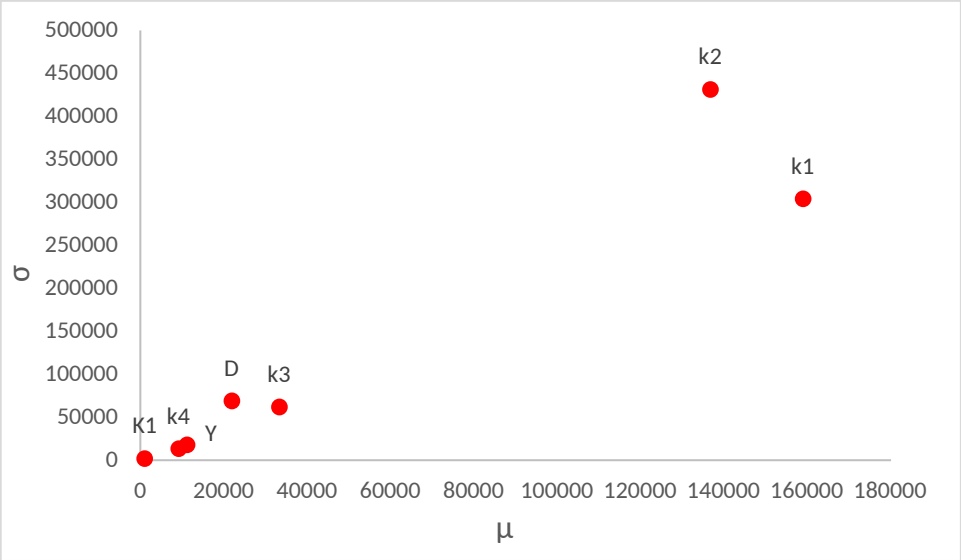


Figure 3. Morris algorithm with respect to y.

Table 6. Uniform distribution, Morris (M) rank, and Sobol' indices (SI) rank for the investigated parameters. Uniform marginal PDFs associated to the vector θ are reported in the third column. $\mathcal{U}(min, max)$ is the uniform distribution between the minimum and the maximum value of the specified parameter.

Parameter	Description	Uniform distribution	M rank	SI rank
k_1	VS degradation rate	$\mathcal{U}(0.005, 0.5)$	1	3
k_2	Acetate maximum uptake rate	$\mathcal{U}(0.08, 8.0)$	2	1
k_3	Gas-liquid transfer coefficient	$\mathcal{U}(0.3, 300)$	3	6
k_4	Microbial biomass decay rate	$\mathcal{U}(0.001, 0.05)$	6	5
K_1	Half-saturation constant	$\mathcal{U}(0.015, 1.5)$	7	2
D	Diffusion coefficient	$\mathcal{U}(1 \times 10^{-8}, 1 \times 10^{-6})$	4	7
Y	Yield of biomass on substrate	$\mathcal{U}(0.04, 0.1)$	5	4

A second group made of physical parameters includes k_3, and D, which are sufficiently relevant in determining the QoI value. The gas-transfer coefficient k_3 regulates the amount of soluble methane released in the gas-phase from the reactor environment. Of course, its low value leads to a small amount of soluble methane transferred in the gas-phase. When the value of k_3 exceeds a specific threshold, all the methane is released in the gas-phase and the value of the gas-transfer coefficient becomes negligible with respect to the QoI. It is clearly important to notice that this observation is due to the atmospheric pressure level considered for the PFR. Indeed, pressure and consequent gas-transfer conditions are critical factors influencing the overall efficiency of the biogas generation process. Elevated pressures can enhance the solubility of gases, thereby potentially increasing the contact between microorganisms and substrates, which can lead to higher methane yields. However, excessively high pressures may also inhibit microbial activity and disrupt the delicate balance of microbial communities, ultimately reducing methane production [12]. Therefore, optimizing pressure conditions is essential for maximizing methane output while maintaining the stability and performance of PFR. Regarding the diffusion coefficient, it was unexpected for D to be so less important than biological parameters for the specific QoI, as the PFR model reproduces convective-diffusive-reactive phenomena. This observation highlighted the crucial role of kinetic coefficients when modeling the effect of diffusion on the process development. Indeed, it has been demonstrated that diffusion strongly affects the hydrolysis of complex material in dry AD processes [45, 46]. The one-dimensional structure of the PFR model leads to neglecting the influence of different values of the diffusion coefficient along other directions, which can affect the determination of the QoI in 2D and 3D models. In these cases, a more relevant relationship between the diffusion and the rate of conversion of substrates in methane could emerge, as different diffusion processes in the other directions may influence acids and by-products distribution in the PFR [47]. Moreover, a different formulation of kinetics parameters including the effect of the diffusion coefficient in such particular environment could be proposed for future works.

The last qualitative information obtained through the Morris' screening was that the group of parameters consisting in the half-saturation constant K_1, the microbial biomass decay rate k_4, and the yield of biomass on substrate Y present very small sensitivities. The roles of the half-saturation constant and the yield of biomass on substrate are not emerging in the qualitative preliminary screening. Indeed,

when K_1 assumes very small values, the consumption kinetic of the acetic acid becomes linear, which means that the methane production process is more efficient. The yield of biomass on substrate Y determines the amount of acetic acid that becomes methane and, even if it varies among 0 and 1, it is expected to influence methane production. These aspects required a deeper investigation with the quantitative variance based GSA.

5.2. Estimation of the error

Metamodels of the PFR model for UQ and SA must be evaluated for their ability to reproduce model variability. Indeed, it is possible to compute the *a posteriori* approximation error resulting from the surrogate model by

$$\epsilon_{\text{emp}} = \frac{1}{N_{\text{halton}}} \sum_{l=1}^{N_{\text{halton}}} \left(y^{(l)} - \widehat{y}^{(l)} \right), \tag{5.1}$$

where $y^{(l)}$ is the l^{th} element of the training set, $\widehat{y}^{(l)}$ is the corresponding prediction (gPC or GP), and $N = 2^{10}$ (see Table 5). It is noteworthy that this estimator of the metamodel error is prone to issues of *overfitting*, which could result in a severe underestimation of the mean square error [39]. Furthermore, since the GP model interpolates the training set points, it is guaranteed to achieve $\epsilon_{\text{emp}} = 0$ when noise-free kernels are used. For any surrogate model, algorithm, and configuration that were tested, the ratio of the empirical error, ϵ_{emp}, to the empirical mean of the QoI, $\bar{y}$, calculated over the Halton dataset, remains consistently below 4.0×10^{-3}.

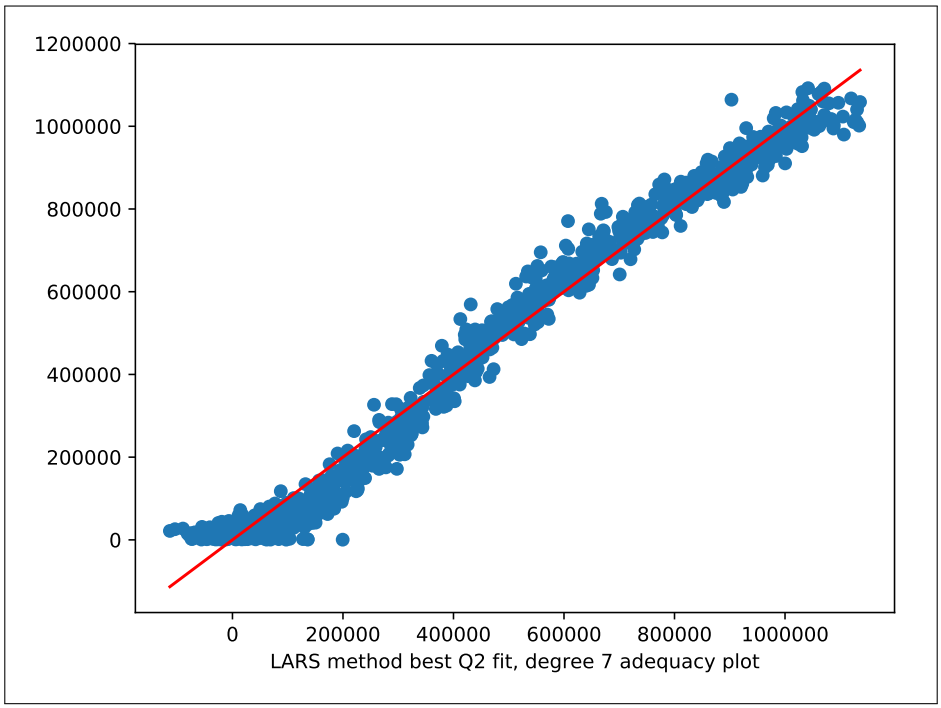
To avoid the aforementioned limitations, the surrogates were validated by employing the Q_2 predictive coefficient, a cross-validation error metric that utilizes an independent dataset derived from Faure's low discrepancy sequence, as reported in Table 5. In Eq (5.2), this coefficient is described as

$$Q_2 = 1 - \frac{\sum_{l=1}^{N_{\text{Faure}}} \left(y^{(l)} - \widehat{y}^{(l)} \right)^2}{\sum_{l=1}^{N_{\text{Faure}}} \left(y^{(l)} - \bar{y} \right)^2}, \tag{5.2}$$

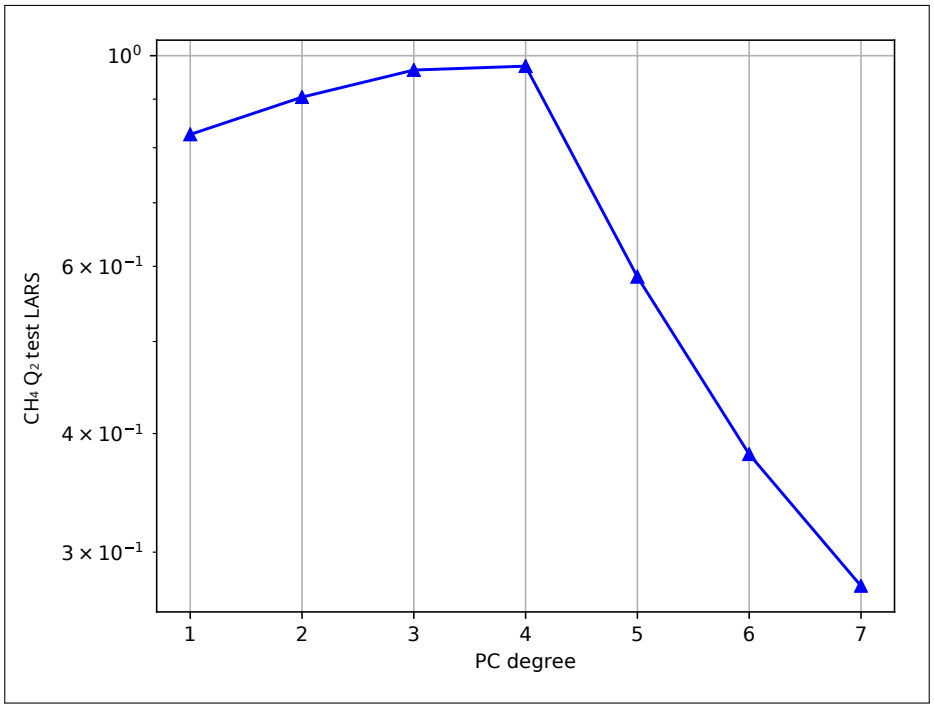
where $\bar{y}$ denotes the empirical mean calculated from Faure's validation set ($N_{\text{faure}} = 2^{10}$). The Q_2 predictor coefficient (with 1 as the target value) provides a normalized estimation of the generalization error, which quantifies the surrogate model's error when evaluating points that are not included in Halton's training set, thereby offering a measure of the model's predictive performance on unseen data [48]. The closer Q_2 approaches 1, the better the surrogate replicates the PFR dynamical system. By applying the Q_2 indicator, it is possible to rank surrogate models based on their ability to reproduce the PFR model dynamics. When the g-PC technique is applied, the surrogate model characterized by a total polynomial order P demonstrates superior accuracy, yielding the most precise results compared to other adopted models with different polynomial orders. In this study, the polynomial order P ranged from 1 to 7.

Figure 4(a) presents the adequacy plots, which illustrate the performance of the metamodel by comparing it with the actual forward model $\mathcal{F}$ runs, using points from the design of experiment (DoE) to assess how well the metamodel approximates the forward model. By plotting the Q_2 values in Figure 4(b), the robustness of the LAR-gPC algorithm is illustrated. When $P > 4$, the dimensionality

of the underlying statistical model hinders the convergence. As a final note, Table 7 provides a comprehensive overview of the error estimators associated with the various surrogate techniques that have been applied, detailing the measures used to evaluate the accuracy and reliability of each technique.



(a) Adequacy plot for the LAR gPC algorithm



(b) Q_2 test for varying maximum order of gPC from $p_0 = 1$ to $p_1 = 7$

Figure 4. Adequacy plots (a) and Q_2 test (b) related to the LAR-based gPC algorithm.

Table 7. Values of observed errors related to the different surrogate models built in this study. The most favorable results for the LAR-gPC and the SLS-gPC, corresponding to the explored range of values for P, are reported.

Metamodel	Q_2	$\epsilon_{emp}/\overline{y}$
SLS-gPC ($P = 4$)	0.972	$3.40e-03$
LAR-gPC ($P = 4$)	0.975	$3.79e-03$
GP (RBF Kernel)	0.993	0

5.3. Quantitative GSA

Variance decomposition-based analysis is usually performed using Sobol' indices [49, 50]. These allow to quantify the weight of the uncertainty of a single input parameter, considered as an independent variable with respect to the other parameters, based on the variance of the QoI y. By indicating the variance of the output random variable y and the variance linked to the variability of the i^{th} parameter as $\mathbb{V}(y)$ and $\mathbb{V}_i(y)$, respectively, the first-order Sobol' index S_i, referred to the i^{th} parameter of $\mathbf{\Theta}$, is defined as:

$$S_i = \frac{\mathbb{V}_i(Y)}{\mathbb{V}(Y)}. \tag{5.3}$$

Indeed, S_i values ranges between 0 and 1. To account for parameter interactions while defining the measure of the i^{th} input parameter contribution on the output variance, the total Sobol' index S_{T_i} is defined as:

$$S_{T_i} = \sum_{\substack{I \subset \{1,\dots,d\} \\ I \ni i}} S_I. \tag{5.4}$$

Consequently, $S_{T_i} \geq S_i$ by definition. By observing the difference between the first-order and the total indices, it is possible to detect the presence of these interactions. In the case of the GP-surrogate approach, the Sobol' indices were obtained through a stochastic process, with Martinez's formulation employed as a stable and reliable estimator to ensure precision and robustness in the calculation of these indices [51]. The first-order and total Sobol' indices are derived directly from the gPC coefficients for the LAR gPC expansion. As an example, the first-order Sobol index is expressed as:

$$S_{i,\text{pc}} = \frac{1}{\sigma_y^2} \sum_{\substack{\boldsymbol{\alpha} \in \mathcal{A}, \\ \alpha_i > 0 \text{ and } \alpha_{k \neq i} = 0}} \gamma_{\boldsymbol{\alpha}}^2, \tag{5.5}$$

where σ_y is the empirical output sample STD (see [35, 36]).

The first-order and total Sobol' indices obtained by the three algorithms are shown in Figure 5. Based on the Q_2 error, the GP surrogate was the best-performing algorithm, so the SA and UQ results related to this algorithm were discussed. A comparison of first-order and total Sobol' indices reveals interactions between factors. It is noteworthy that the range of parameters utilized in the investigation can significantly influence the analysis outcomes. To mitigate errors, the parameters were selected based on a comprehensive review of various anaerobic digestion models that examine the bioconversion

of different substrates under anaerobic conditions [52–54]. The specific ranges of these parameters are detailed in Table 6. Considering the specific QoI (methane production), Sobol' indices analysis confirmed the Morris screening results for k_2. Both the Sobol' index and the total Sobol' index of the Monod maximum uptake rate k_2 were higher than all the other parameters. As expected, the kinetic constant k_1 maintained its relevance, while the half-saturation constant K_1 and the yield Y gained positions on k_3, k_4, and D. As already remarked, the half-saturation constant K_1 and the yield Y directly affect the acetic acid maximum consumption rate and the consumed acetic acid, respectively. As a result of the Morris screening, the QoI was not affected by the variation in these parameters. However, quantitative SA highlighted their effects. Finally, the small sensitivities of the microbial decay rate kinetic constant k_4, and physical parameters k_3 and D were confirmed by the quantitative SA with Sobol' indices.

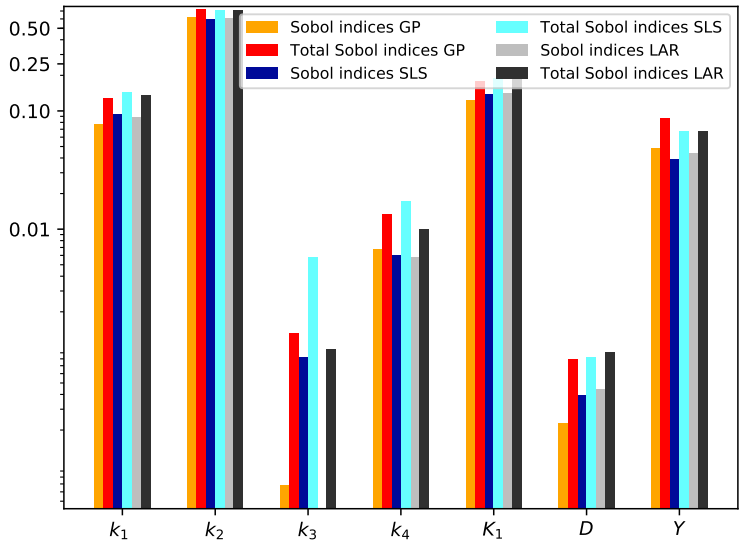


Figure 5. First-order and total Sobol' indices (logarithmic scale) related to uncertain parameters $\boldsymbol{\theta}$ and their effect on y. For the GP, orange and red are first-order and total Sobol' indices, respectively. For the SLS-gPC, light blue and dark blue are first-order and total Sobol' indices, respectively. For the LAR-gPC, gray and dark gray are first-order and total Sobol' indices, respectively.

5.4. Uncertainty quantification

Based on the Q_2 error estimator, the analysis was restricted to the surrogate model that provides the highest level of accuracy, such as the Gaussian process metamodel. To achieve a comprehensive statistical characterization of the QoI, including its statistical moments and PDF, a Monte Carlo random sampling method was employed to sample the uncertain input space Z_Θ with a sample size of $10,000$ members, followed by the evaluation of the GP surrogate model to analyze the resulting data. The mean, standard deviation, skewness, and kurtosis of the QoI, computed from such Monte Carlo samples are given in Table 8.

Table 8. Statistical moments for the PDF.

Moment	y ($l_{CH_4}d$)
Mean	398514.91
Standard deviation	353356
Skewness	0.275
Kurtosis	1.5638

Figure 6 presents the QoI PDF related to the CH_4 net production. Relevant information can be inferred from both the shape of the distribution and its higher-order moments, as these elements collectively provide insight into the underlying characteristics and behavior of the data. Depending on the model parameter values and the initial conditions, the two different peaks revealed a high probability of having a small amount of produced methane (corresponding to the failure of the system) and a medium-high probability of reaching high methane production. Indeed, when the input parameter vector is characterized by unfavorable values, with slow conversion of the biodegradable COD in acetic acid or slow conversion of the acetic acid, it is very difficult to reach a valuable methane production. On the other hand, when the vector of parameters is characterized by a favorable combination of input factors with respect to the reactor working conditions, there is a good chance to achieve relevant methane productions. The PDF drops to zero, after the value of approximately 1×10^6 $[l_{CH_4} \cdot d]$. This result highlights that it is impossible to produce more methane in the corresponding period of time with the used values of HRT, v_0, and OLR. This revealed the significant limitation due to the operational conditions influencing the performance of the PFR.

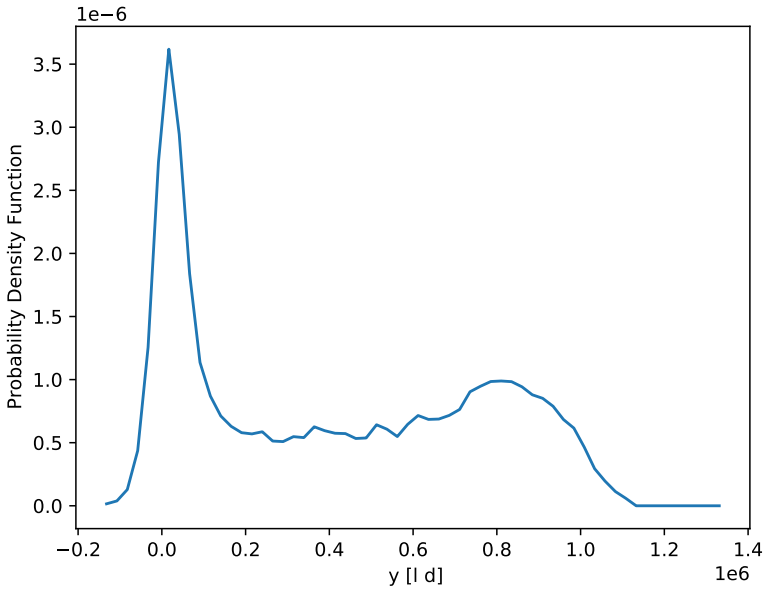


Figure 6. Probability density function for the quantity of interest y with the GP metamodel.

6. Conclusions

Using non-intrusive techniques, UQ and GSA techniques were adopted on a novel dry AD model in a PFR. The model depends on a conspicuous set of input parameters and, up to now, had lacked rigorous studies serving as bases for the application of effective calibration an validation procedures. GSA application leads to identifying which parameters most significantly influence model outputs, thereby guiding targeted adjustments and ensuring that model predictions are robust and reliable. On the other hand, UQ addresses the inherent uncertainties in model inputs and parameters, providing a comprehensive understanding of the range and likelihood of possible outcomes, which is essential for accurate model validation and for making informed decisions in any biotechnological application. The use of these techniques was aimed at predicting the overall effect of uncertainty of the input parameters on CH_4 production, which represents the value-added product achievable from this bio-tecnology. The identification of the most sensitive parameters upon which the dry AD process depends, and the identification of the uncertainty associated with their variation within reasonable ranges, represent an initial step in reducing the process-related uncertainty.

The Morris screening test and the quantitative sensitivity analysis with Sobol' indices showed, with a good accordance, that the net methane concentration is mostly sensitive to the values of the conversion parameter related to the particulate biodegradable volatile solids in acetic acid k_1, and to the kinetic parameter describing the acetic acid uptake k_2. The relevance of the half-saturation constant K_1 and the yield of biomass on substrate Y were highlighted by the quantitative sensitivity analysis through the Sobol' indices. The selected QoI was less sensitive to other biochemical and physical parameters. According to the presented results, it is possible to claim that in 1-D PFR models, the role of the diffusion coefficient is negligible with respect to more relevant kinetic parameters. It is important to emphasize that the primary objective of the present study was to investigate the bioconversion of acetic acid into methane in a PFR. Consequently, microbial species responsible for acetic acid accumulation, along with other common metabolic species such as sulfate-reducing bacteria, have been excluded from consideration. A natural progression of this research would involve developing an extended mathematical model that includes a broader range of microbial species and associated parameters. An analytical study of this extended model, combined with the application of SA and UQ techniques, could provide valuable insights into the most influential metabolic processes under these more comprehensive conditions.

Finally, the present analysis revealed the crucial role of UQ for adequate calibration procedures and further validation of such complex mathematical models. These routines, ranging from simple screening analysis to more complex and computationally demanding variance-based metamodel analysis, constitute a robust asset for modelers, designers, and managers of anaerobic reactors. For the performance evaluation of a dry AD process, the group of more sensitive parameters are those connected to the kinetic description of the biological process, leading to the contextual waste substrate degradation and methane production in anaerobic reactors.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work has been partially developed in the context of D.D. n. 1377 on June 5, 2017, additional PhD fellowships for 2017/2018 academic year, course XXXIII within the framework of the "Programma Operativo Nazionale Ricerca e Innovazione (PON RI 2014/2020) Action I.1 - Innovative PhDs with industrial characterization".

VL acknowledges the project "FBIOMA-Fermentative BIOhydrogen production MAthematical modeling" (CUP: E63C22004130001).

MRM acknowledges support from the National Research Center for Agricultural Technologies - Agritech, funded by the European Union NextGenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA PNRR, Missione 4 Componente 2 Investimento 1.4 - D.D. 1032 17/06/2022, Project code CN00000022) The present work has been performed under the auspices of the GNFM of INdAM.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. G. Esposito, L. Frunzo, F. Liotta, A. Panico, F. Pirozzi, Bio-methane potential tests to measure the biogas production from the digestion and co-digestion of complex organic substrates, *Open Environ. Eng. J.*, **5** (2012), 1–8. https://doi.org/10.2174/1874829501205010001
2. G. Esposito, L. Frunzo, A. Giordano, F. Liotta, A. Panico, F. Pirozzi, Anaerobic co-digestion of organic wastes, *Rev. Environ. Sci. Bio/Technol.*, **11** (2012), 325–341. https://doi.org/10.1007/s11157-012-9277-8
3. V. Luongo, M. R. Mattei, L. Frunzo, B. D'Acunto, K. Gupta, S. Chellam, et al., A transient biological fouling model for constant flux microfiltration, *Math. Biosci. Eng.*, **20** (2023), 1274–1296. https://doi.org/10.3934/mbe.2023058
4. Y. Li, S. Y. Park, J. Zhu, Solid-state anaerobic digestion for methane production from organic waste, *Renewable Sustainable Energy Rev.*, **15** (2011), 821–826. https://doi.org/10.1016/j.rser.2010.07.042
5. O. Karthikeyan, C. Visvanathan, Bio-energy recovery from high-solid organic substrates by dry anaerobic bio-conversion processes: A review, *Rev. Environ. Sci. Bio/Technol.*, **12** (2013), 257–284. https://doi.org/10.1007/s11157-012-9304-9
6. P. Vandevivere, New and broader applications of anaerobic digestion, *Critical Rev. Env. Sci. Technol.*, **29** (1999), 151–173. https://doi.org/10.1080/10643389991259191
7. G. Policastro, V. Luongo, L. Frunzo, N. Cogan, M. Fabbricino, A mechanistic mathematical model for photo fermentative hydrogen and polyhydroxybutyrate production, *Math. Biosci. Eng.*, **20** (2023), 7407–7428. https://doi.org/10.3934/mbe.2023321
8. A. Donoso-Bravo, C. Sadino-Riquelme, D. Gómez, C. Segura, E. Valdebenito, F. Hansen, Modelling of an anaerobic plug-flow reactor. process analysis and evaluation approaches with non-ideal mixing considerations, *Bioresour. Technol.*, **260** (2018), 95–104. https://doi.org/10.1016/j.biortech.2018.03.082

9. I. Białobrzewski, K. Waszkielis, K. Bułkowska, The application of Anaerobic Digestion Model No. 1 for the optimization of biogas production from maize silage, pig manure, cattle manure, and digestate in a full-scale biogas plant, *Fuel*, **357** (2024), 129789. https://doi.org/10.1016/j.fuel.2023.129789
10. R. Kothari, A. Pandey, S. Kumar, V. Tyagi, S. Tyagi, Different aspects of dry anaerobic digestion for bio-energy: An overview, *Renewable Sustainable Energy Rev.*, **39** (2014), 174–195. https://doi.org/10.1016/j.rser.2014.07.011
11. D. J. Batstone, J. Keller, I. Angelidaki, S. Kalyuzhnyi, S. Pavlostathis, A. Rozzi, et al., The iwa anaerobic digestion model no 1 (adm1), *Water Sci. Technol.*, **45** (2002), 65–73. https://doi.org/10.2166/wst.2002.0292
12. C. De Crescenzo, A. Marzocchella, D. Karatza, S. Chianese, D. Musmarra, Autogenerative high-pressure anaerobic digestion modelling of volatile fatty acids: Effect of pressure variation and substrate composition on volumetric mass transfer coefficients, kinetic parameters, and process performance, *Fuel*, **358** (2024), 130144, https://doi.org/10.1016/j.fuel.2023.130144
13. C. De Crescenzo, A. Marzocchella, D. Karatza, A. Molino, P. Ceron-Chafla, R. E. Lindeboom, et al., Modelling of autogenerative high-pressure anaerobic digestion in a batch reactor for the production of pressurised biogas, *Biotechnol. Biofuels Bioprod.*, **15** (2022). https://doi.org/10.1186/s13068-022-02117-x
14. C. Fall, J. Loaiza-Navía, Design of a tracer test experience and dynamic calibration of the hydraulic model for a full-scale wastewater treatment plant by use of AQUASIM, *Water Environ. Res.*, **79** (2007), 893–900. https://doi.org/10.2175/106143007X176068
15. Y. Muslu, Numerical approach to plug-flow activated sludge reactor kinetics, *Comput. Biol. Med.*, **30**, (2000), 207–223. https://doi.org/10.1016/S0010-4825(00)00009-3
16. V. Vavilin, L. Lokshina, X. Flotats, I. Angelidaki, Anaerobic digestion of solid material: Multidimensional modeling of continuous-flow reactor with non-uniform influent concentration distributions, *Biotechnol. Bioeng.*, **97** (2007), 354–366. https://doi.org/10.1002/bit.21239
17. B. Wu, Integration of mixing, heat transfer, and biochemical reaction kinetics in anaerobic methane fermentation, *Biotechnol. Bioeng.*, **109** (2012), 2864–2874. https://doi.org/10.1002/bit.24551
18. D. B. Panaro, M. R. Mattei, G. Esposito, J. P. Steyer, F. Capone, L. Frunzo, A modelling and simulation study of anaerobic digestion in plug-flow reactors, *Commun. Nonlinear Sci. Numer. Simul.*, **105** (2022), 106062. https://doi.org/10.1016/j.cnsns.2021.106062
19. Y. Han, Z. Du, X. Hu, Y. Li, D. Cai, J. Fan, et al., Production prediction modeling of food waste anaerobic digestion for resources saving based on SMOTE-LSTM, *Appl. Energy*, **352** (2023), 122024. https://doi.org/10.1016/j.apenergy.2023.122024
20. A. Saltelli, K. Aleksankina, W. Becker, P. Fennell, F. Ferretti, N. Holst, et al., Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices, *Environ. Modell. Software*, **114** (2019), 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012
21. A. Donoso-Bravo, J. Mailier, C. Martin, J. Rodríguez, C. A. Aceves-Lara, A. V. Wouwer, Model selection, identification and validation in anaerobic digestion: A review, *Water Res.*, **45** (2011), 5347–5364. https://doi.org/10.1016/j.watres.2011.08.059

22. B. Tartakovsky, S. Mu, Y. Zeng, S. Lou, S. Guiot, P. Wu, Anaerobic digestion model No. 1-based distributed parameter model of an anaerobic reactor: II. Model validation, *Bioresour. Technol.*, **99** (2008), 3676–3684. https://doi.org/10.1016/j.biortech.2007.07.061

23. N. Noykova, M. Gyllenberg, Sensitivity analysis and parameter estimation in a model of anaerobic waste water treatment processes with substrate inhibition, *Bioprocess. Eng.*, **23** (2000), 343–349. https://doi.org/10.1007/s004499900169

24. O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, J. P. Steyer, Dynamical model development and parameter identification for an anaerobic wastewater treatment process, *Biotechnol. Bioeng.*, **75** (2001), 424–438. https://doi.org/10.1002/bit.10036

25. V. Vavilin, S. Rytov, S. Pavlostathis, J. Jokela, J. Rintala, A distributed model of solid waste anaerobic digestion: Sensitivity analysis, *Water Sci. Technol.*, **48** (2003), 147–154. https://doi.org/10.2166/wst.2003.0241

26. Y. Lin, C. Wu, Sensitivity analysis of phenol degradation with sulfate reduction under anaerobic conditions, *Environ. Model. Assess.*, **16** (2011), 213–225. https://doi.org/10.1007/s10666-010-9243-1

27. K. Solon, X. Flores-Alsina, K. V. Gernaey, U. Jeppsson, Effects of influent fractionation, kinetics, stoichiometry and mass transfer on CH_4, H_2 and CO_2 production for (plant-wide) modeling of anaerobic digesters, *Water Sci. Technol.*, **71** (2015), 870–877. https://doi.org/10.2166/wst.2015.029

28. L. Benedetti, D. J. Batstone, B. De Baets, I. Nopens, P. A. Vanrolleghem, Global sensitivity analysis of biochemical, design and operational parameters of the Benchmark Simulation Model no. 2, in *Proceedings of the 4th International Congress on Environmental Modelling and Software (iEMSs 2008)*, (2008), 1322–1330.

29. Ž. Zonta, M. Alves, X. Flotats, J. Palatsi, Modelling inhibitory effects of long chain fatty acids in the anaerobic digestion process, *Water Res.*, **47** (2013), 1369–1380. https://doi.org/10.1016/j.watres.2012.12.007

30. F. Carrera-Chapela, A. Donoso-Bravo, D. Jeison, I. D´ıaz, J. Gonzalez, G. Ruiz-Filippi, Development, identification and validation of a mathematical model of anaerobic digestion of sewage sludge focusing on H_2S formation and transfer, *Biochem. Eng. J.*, **112** (2016), 13–19. https://doi.org/10.1016/j.bej.2016.03.008

31. I. M. Nasir, T. I. Mohd Ghazi, R. Omar, Anaerobic digestion technology in livestock manure treatment for biogas production: A review, *Eng. Life Sci.*, **12** (2012), 258–269. https://doi.org/10.1002/elsc.201100150

32. M. D. Morris, Factorial sampling plans for preliminary computational experiments, *Technometrics*, **33** (1991), 161–174. https://doi.org/10.1080/00401706.1991.10484804

33. M. Baudin, A. Dutfoy, B. Iooss, A. L. Popelin, Open TURNS: An industrial software for uncertainty quantification in simulation, preprint, arXiv:1501.05242. https://doi.org/10.48550/arXiv.1501.05242

34. F. Campolongo, A. Saltelli, Sensitivity analysis of an environmental model: An application of different analysis methods, *Reliab. Eng. Syst. Saf.*, **57** (1997), 49–69. https://doi.org/10.1016/S0951-8320(97)00021-5

35. A. Trucchia, M. R. Mattei, V. Luongo, L. Frunzo, M. C. Rochoux, Surrogate-based uncertainty and sensitivity analysis for bacterial invasion in multi-species biofilm modeling, *Commun. Nonlinear Sci. Numer. Simul.*, **73** (2019), 403–424. https://doi.org/10.1016/j.cnsns.2019.02.024

36. A. Trucchia, V. Egorova, G. Pagnini, M. C. Rochoux, On the merits of sparse surrogates for global sensitivity analysis of multi-scale nonlinear problems: Application to turbulence and fire-spotting model in wildland fire simulators, *Commun. Nonlinear Sci. Numer. Simul.*, **73** (2019), 120–145. https://doi.org/10.1016/j.cnsns.2019.02.002

37. P. T. Roy, N. El Moçayd, S. Ricci, J. C. Jouhaud, N. Goutal, M. De Lozzo, et al., Comparison of polynomial chaos and Gaussian process surrogates for uncertainty quantification and correlation estimation of spatially distributed open-channel steady flows, *Stochastic Environ. Res. Risk Assess.*, **32** (2018), 1723–1741. https://doi.org/10.1007/s00477-017-1470-4

38. D. Xiu, G. E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, **24** (2002), 619–644. https://doi.org/10.1137/S1064827501387826

39. G. Blatman, B. Sudret, Efficient computation of global sensitivity indices using sparse polynomial chaos expansions, *Reliab. Eng. Syst. Saf.*, **95** (2010), 1216–1229. https://doi.org/10.1016/j.ress.2010.06.015

40. C. K. Williams, C. E. Rasmussen, *Gaussian Processes for Machine Learning*, MIT press Cambridge, MA, **2** (2006).

41. A. Trucchia, L. Frunzo, Surrogate based global sensitivity analysis of ADM1-based anaerobic digestion model, *J. Environ. Manage.*, **282** (2021), 111456. https://doi.org/10.1016/j.jenvman.2020.111456

42. F. Charte, I. Romero, M. D. Pérez-Godoy, A. J. Rivera, E. Castro, Comparative analysis of data mining and response surface methodology predictive models for enzymatic hydrolysis of pretreated olive tree biomass, *Comput. Chem. Eng.*, **101** (2017), 23–30. https://doi.org/10.1016/j.compchemeng.2017.02.008

43. E. Ficara, S. Hassam, A. Allegrini, A. Leva, F. Malpei, G. Ferretti, Anaerobic digestion models: A comparative study, *IFAC Proc. Vol.*, **45** (2012), 1052–1057. https://doi.org/10.3182/20120215-3-AT-3016.00186

44. D. Panaro, L. Frunzo, M. Mattei, V. Luongo, G. Esposito, Calibration, validation and sensitivity analysis of a surface-based ADM1 model, *Ecol. Modell.*, **460** (2021), 109726.

45. C. Veluchamy, A. S. Kalamdhad, A mass diffusion model on the effect of moisture content for solid-state anaerobic digestion, *J. Cleaner Prod.*, **162** (2017), 371–379 https://doi.org/10.1016/j.jclepro.2017.06.099

46. F. Xu, Z. W. Wang, L. Tang, Y. Li, A mass diffusion-based interpretation of the effect of total solids content on solid-state anaerobic digestion of cellulosic biomass, *Bioresour. Technol.*, **167** (2014), 178–185. https://doi.org/10.1016/j.biortech.2014.05.114

47. F. Liotta, P. Chatellier, G. Esposito, M. Fabbricino, E. D. Van Hullebusch, P. N. Lens, Hydrodynamic mathematical modelling of aerobic plug flow and nonideal flow reactors: A critical and historical review, *Crit. Rev. Environ. Sci. Technol.*, **44** (2014), 2642–2673. https://doi.org/10.1080/10643389.2013.829768

48. A. Marrel, B. Iooss, B. Laurent, O. Roustant, Calculations of Sobol indices for the Gaussian process metamodel, *Reliab. Eng. Syst. Saf.*, **94** (2009), 742–751. https://doi.org/10.1016/j.ress.2008.07.008

49. I. Sobolprime, Sensitivity analysis for nonlinear mathematical models, *Math. Model. Comput. Exp.*, **1** (1993), 407–414.

50. A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, et al., *Global Sensitivity Analysis: The Primer*, John Wiley & Sons, 2008.

51. M. Baudin, K. Boumhaout, T. Delage, B. Iooss, J. M. Martinez, Numerical stability of Sobol'indices estimation formula, in *Proceedings of the 8th International Conference on Sensitivity Analysis of Model Output (SAMO 2016)*, **30** (2016), 50–51.

52. J. Yang, L. Lu, W. Ouyang, Y. Gou, Y. Chen, H. Ma, et al., Estimation of kinetic parameters of an anaerobic digestion model using particle swarm optimization, *Biochem. Eng. J.*, **120** (2017), 25–32. https://doi.org/10.1016/j.bej.2016.12.022

53. N. Kythreotou, G. Florides, S. A. Tassou, A review of simple to scientific models for anaerobic digestion, *Renewable Energy*, **71** (2014), 701–714. https://doi.org/10.1016/j.renene.2014.05.055

54. J. Y. X. Ling, Y. J. Chan, J. W. Chen, D. J. S. Chong, A. L. L. Tan, S. K. Arumugasamy, et al., Machine learning methods for the modelling and optimisation of biogas production from anaerobic digestion: A review, *Environ. Sci. Pollut. Res.*, **31** (2024), 19085–19104. https://doi.org/10.1007/s11356-024-32435-6

AIMS Press