Moment-Independent Sensitivity Analysis of Enzyme-Catalyzed Reactions with Correlated Model Parameters ⋆

Xiangzhong Xie [*,***] **Rüdiger Ohs** [**] **Antje Spieß** [**,***] **Ulrike Krewer** [*,***] **René Schenkendorf** [*,***]

[*] *Institute of Energy and Process Systems Engineering, Braunschweig University of Technology, 38106 Braunschweig, Germany*
[**] *Institute of Biochemical Engineering, Braunschweig University of Technology, 38106 Braunschweig, Germany*
[***] *Center of Pharmaceutical Engineering (PVZ), Braunschweig University of Technology, 38106 Braunschweig, Germany*
e-mail:
{*x.xie,r.ohs,a.spiess,u.krewer,r.schenkendorf*}*@tu-braunschweig.de*

Abstract: The dynamic models used for biological and chemical process analysis and design usually include a significant number of uncertain model parameters. Sensitivity analysis is frequently applied to provide quantitative information regarding the influence of the parameters, as well as their uncertainties, on the model output. Various techniques are available in the literature to calculate parameter sensitivities based on local derivatives or changes in dedicated statistical moments of the model output. However, these methods may lead to an inevitable loss of information for a proper sensitivity analysis and are not directly available for problems with correlated model parameters. In this work, we demonstrate the use of a moment-independent sensitivity analysis concept in the presence and absence of parameter correlations and investigate the correlation effect in more detail. Moment-independent sensitivity analysis calculates parameter sensitivities based on changes in the entire probability density distribution of the model output and is formulated independently of whether the parameters are correlated or not. Technically, a single-loop Monte Carlo simulation method in combination with polynomial chaos expansion is implemented to reduce the computational cost significantly. A sampling procedure derived from Gaussian copula formalism is used to generate sample points for arbitrarily correlated uncertain parameters. The proposed concept is demonstrated with a case study of an enzyme-catalyzed reaction network. We observe evident differences in the parameter sensitivities for cases with independent and correlated model parameters.

Keywords: moment-independent, sensitivity analysis, parameter correlations, Gaussian copula, polynomial chaos expansion, enzyme-catalyzed reactions

1. INTRODUCTION

Increasing competition in the process engineering industry, stringent safety requirements and regulations for reliable operations necessitate model-based design and control for biological and chemical processes (Biegler, 2010). However, mathematical models that attempt to mimic the dynamic processes contain many uncertain model parameters. The parameter uncertainties may result from either inaccurate experiment measurements or the inherent randomness of dynamic systems. Sensitivity analysis aims at quantifying the intensity of the influence of uncertain parameters on the model output and has also been used for model-based experiment design (Telen et al., 2014). The existence of parameter correlations, however, might increase the difficulty of properly quantifying parameter sensitivities. The goal of this paper is to present a reliable method for quantifying the sensitivities of correlated parameters for dynamic systems focusing on efficient implementation.

Based on the considered amount of parameter information, the methods for sensitivity analysis can be categorized into two groups: 1) Local sensitivity analysis explores the parameter sensitivity close to a given parameter reference point (e.g., differential-based or derivative-based methods) and 2) global sensitivity analysis methods cover the entire parameter space (e.g., non-parametric or variance-based approaches). For more details, the interested reader is referred to the work of Borgonovo and Plischke (2016) and references therein. However, most sensitivity analysis methods are inappropriate for complex dynamic systems with correlated model parameters. To address also relevant parameter correlations, we present the moment-independent sensitivity analysis (MISA) method (Bor-

⋆ Funding of Promotionsprogramm "μ-Props" by MWK Niedersachsen is gratefully acknowledged. Xiangzhong Xie is also a member of the International Max Planck Research School for Advanced Methods in Process and Systems Engineering, 39106 Magdeburg, Germany

gonovo, 2007). Moreover, MISA evaluates the entire probability density function of the simulation result instead of a limited number of statistical moments providing more informative insights.

Although MISA has been applied in various studies (Borgonovo and Tarantola, 2008; Rajabi et al., 2015), its efficient numerical calculation is still challenging. In this paper, we present a single-loop Monte Carlo simulation framework (Wei et al., 2013) which renders the original nested problem of two Monte Carlos simulation loops in a more explicit form. Moreover, polynomial chaos expansion (PCE; Xiu and Karniadakis (2002)) is used to lower the computational cost. In detail, we present a sampling procedure to generate parameter samples from arbitrarily correlated uncertain model parameters.

We demonstrate the proposed concept with an enzyme-catalyzed reaction network for the synthesis of 2-hydroxy ketones which is of interest in the pharmaceutical and chemical industries (Demir et al., 2001). The remainder of the paper is organized as follows. In Sections 2 and 3, the general sensitivity analysis problem is defined, and the methods for MISA are presented. We discuss the results in Section 4 and end with conclusions in Section 5.

2. PROBLEM FORMULATION

Consider a non-linear dynamic model described by ordinary differential equations (ODEs):

$$\dot{\mathbf{x}}(t,\boldsymbol{\theta}) = \mathbf{f}(\mathbf{x}(t,\boldsymbol{\theta}),\mathbf{u}(t),\boldsymbol{\theta}), \qquad \mathbf{x}(0) = \mathbf{x}_0, \tag{1}$$

where $t \in [0, t_f]$ denotes the time, $\mathbf{x}(t,\boldsymbol{\theta}) \in \mathbb{R}^{n_x}$ represents the system states with the initial conditions $\mathbf{x}(0) = \mathbf{x}_0$, $\mathbf{u} \in \mathbb{R}^{n_u}$ represents the system inputs, and $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_{n_\theta}) \in \mathbb{R}^{n_\theta}$ represents the time-invariant parameters. The nonlinear function $\mathbf{f}$: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_x}$ represents the vector field of the dynamic system.

In what follows, we assume that the uncertainties of the model parameters can be described probabilistically. The probability space $(\Omega, \mathcal{F}, P)$ consists of the sample space Ω, the σ-algebra $\mathcal{F}$, and the probability measure P. The uncertain parameters $\boldsymbol{\theta}(\omega)$ are functions of $\omega \in \Omega$ on the probability space and are associated with probability density functions (PDFs) $\mathbf{p}(\boldsymbol{\theta}) = [p_1(\theta_1), \ldots, p_{n_\theta}(\theta_{n_\theta})]$. The parameter uncertainties $\mathbf{p}(\boldsymbol{\theta})$ are propagated through the dynamic model (1), and in turn, the system states $\mathbf{x}(t,\boldsymbol{\theta})$ are vectors of random variables.

The resulting variation in the system states $\mathbf{x}$ induced by individual model parameters and parameter combinations is different. The relations between the parameters and the states cannot be directly observed due to the complexity of the dynamic system and are counterintuitive in some cases. This paper aims at quantifying the influence of the model parameters on the system states by including the full parameter information to ensure credible parameter sensitivity studies. In the literature, various techniques are available for sensitivity analysis. Our focus is on a parameter sensitivity concept with the following features: 1) has a quantitative measure of parameter sensitivities, 2) is global for the entire parameter space, 3) is independent of the model structure, 4) is moment-independent, and 5) is available for independent and correlated model parameters which might be of vital importance (Rajabi et al., 2015). Based on this, the problem of global sensitivity analysis can be stated as below.

Problem 1 Measure the influence of correlated model parameters $\boldsymbol{\theta}(\omega)$ on the variation of the model states $\mathbf{x}(t,\boldsymbol{\theta})$ in which we are interested.

3. METHOD

MISA gives the solution to Problem 1, at the same time it fulfills all the desired features 1) to 5) of sensitivity analysis. The key issue in performing MISA is the efficient computation of the sensitivity indicators. In what follows, we introduce the single-loop Monte Carlo simulation method to reduce the computational cost. The computational demand is further decreased by substituting the original CPU-intensive model with a fast PCE model. Moreover, we include a sophisticated sampling strategy to generate samples for correlated model parameters of arbitrary distributions.

3.1 Moment-Independent Sensitivity Analysis

MISA expresses the influence of parameter uncertainties on the entire PDF of the model output (Borgonovo, 2007). Consider the dynamic model in (1). The model output of interest $y = h(\mathbf{x}(t,\boldsymbol{\theta}))$ is a function of state variables $\mathbf{x}$. The uncertainties of the parameters propagate through the dynamic model and function h, and lead to probabilistically distributed output y with PDF $p_y(y)$. MISA compares the difference between probability distribution $p_y(y)$ and conditional probability distributions $p_{y|\theta_i}(y) (i = 1 \cdots n_\theta)$ of the model output y to calculate the parameter sensitivities (Borgonovo, 2007), which is mathematically expressed as

$$s(\theta_i) = \int_{\mathcal{I}_y} |p_y(y) - p_{y|\theta_i}(y)| dy, \tag{2}$$

where $\mathcal{I}_y$ is the support domain of y and $s(\theta_i)$ is called the shift function. The average of the shift function on the entire distribution of θ_i is then given by

$$E_{\theta_i}[s(\theta_i)] = \int_{\mathcal{I}_{\theta_i}} \left[\int_{\mathcal{I}_y} |p_y(y) - p_{y|\theta_i}(y)| dy \right] p_i(\theta_i) d\theta_i, \tag{3}$$

where $p_i(\theta_i)$ is the marginal PDF of parameter θ_i. Based on (3), a sensitivity indicator for MISA was introduced by Borgonovo (2007) as

$$\delta_i = \frac{1}{2} E_{\theta_i}[s(\theta_i)]. \tag{4}$$

The indicator can also be directly extended to a group of parameters equal to

$$E_{\boldsymbol{\theta_u}}[s(\boldsymbol{\theta_u})] = \int_{\mathcal{I}_{\boldsymbol{\theta_u}}} \left[\int_{\mathcal{I}_y} |p_y(y) - p_{y|\boldsymbol{\theta_u}}(y)| dy \right] p_{\boldsymbol{\theta_u}}(\boldsymbol{\theta_u}) d\boldsymbol{\theta_u}, \tag{5}$$

$$\delta_{\boldsymbol{u}} = \frac{1}{2} E_{\boldsymbol{\theta_u}}[s(\boldsymbol{\theta_u})], \tag{6}$$

in which $\boldsymbol{\theta_u} = (\theta_{i1}, \cdots, \theta_{ir})$ is the subgroup of $\boldsymbol{\theta}$ with dimension r, and $f_{\theta_u}(\theta_u)$ is the joint probability distribution of the subgroup.

The sensitivity indicator δ for MISA has five properties (Borgonovo, 2007): 1) δ_i ($\delta_{\boldsymbol{u}}$) varies in the range [0,1],

Algorithm 1 Single-Loop Monte Carlo Simulation Method

1: Generate a sample set $\mathcal{A} = \left\{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^j, \ldots, \boldsymbol{\theta}^N\right\}$ from joint PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ of the uncertain parameters
2: Evaluate dynamic model (1) and $y(\boldsymbol{\theta}) = h(\mathbf{x}(t, \boldsymbol{\theta}))$ for the samples and obtain a vector of the model output $\mathrm{Y} = \left[y^1 \cdots y^N\right]^T$
3: Estimate PDF $p_y(y)$ of model output with Y using a kernel density estimator
4: **for** $i = 1$ **to** n_θ
5: Estimate joint PDF $p_{y,\theta_i}(y, \theta_i)$ with Y and the i^{th} column of sample matrix A using a kernel density estimator
6: Calculate the interpolated values of the PDFs $p_y(y), p_i(\theta_i), p_{y,\theta_i}(y, \theta_i)$ at points $[\theta_i^j, y^j]$, $j = 1, \cdots N$
6: Compute $\delta_i = \frac{1}{2N} \sum_{j=1}^{N} \left| \frac{p_y(y^j) p_i(\theta_i^j)}{p_{y,\theta_i}(y^j, \theta_i^j)} - 1 \right|$
7: **end for**

and 2) model output y is independent of parameter θ_i (subgroup $\boldsymbol{\theta}_{\boldsymbol{u}}$) if $\delta_i(\delta_{\boldsymbol{u}}) = 0$, 3) the indicator for all uncertain parameters ($\delta_{1,2,\ldots,n}$) equals unity, 4) δ_{ij} for input i and j is bounded as $\delta_i \leq \delta_{ij} \leq \delta_i + \delta_{i|j}$, in which $\delta_{i|j}$ denotes δ_i conditioned on θ_j, and 5) $\delta_j = 0 \iff \delta_{ij} = \delta_i$, i.e., the model output is independent of θ_j. For further descriptions and proofs, we refer to Borgonovo (2007). MISA can be directly used for dynamic models with correlated model parameters because independent parameters are not required for its definition (Borgonovo, 2007).

3.2 Single-Loop Monte Carlo Simulation Method

The key problem in performing MISA is to calculate the sensitivity indicator δ with (3) and (4). Two numerical methods, single-loop and double-loop Monte Carlo simulations, were introduced by Wei et al. (2013). The double-loop Monte Carlo simulation approach is straightforward in implementation but requires sampling from conditional distributions that leads to a high total sampling number which might be prohibitive for many practical applications (Wei et al., 2013). In contrast, the single-loop Monte Carlo simulation method is more efficient but requires some modifications of (3) and (4) as described below.

In a first step, we obtain a new expression for the sensitivity indicator δ as

$$\begin{aligned}
\delta_i &= \frac{1}{2} \int_{\mathcal{I}_{\theta_i}} \left[\int_{\mathcal{I}_y} |p_y(y) - p_{y|\theta_i}(y)| dy \right] p_i(\theta_i) d\theta_i \\
&= \frac{1}{2} \int_{\mathcal{I}_{\theta_i}} \int_{\mathcal{I}_y} |p_y(y) p_i(\theta_i) - p_{y|\theta_i}(y) p_i(\theta_i)| dy d\theta_i \\
&= \frac{1}{2} \int_{\mathcal{I}_{\theta_i}} \int_{\mathcal{I}_y} |p_y(y) p_i(\theta_i) - p_{y,\theta_i}(y, \theta_i)| dy d\theta_i \\
&= \frac{1}{2} \int_{\mathcal{I}_{\theta_i}} \int_{\mathcal{I}_y} \left| \frac{p_y(y) p_i(\theta_i)}{p_{y,\theta_i}(y, \theta_i)} - 1 \right| p_{y,\theta_i}(y, \theta_i) dy d\theta_i,
\end{aligned} \tag{7}$$

where $p_{y,\theta_i}(y, \theta_i) = p_{y|\theta_i}(y) p_i(\theta_i)$ is the joint PDF of y and θ_i. With the expression in (7), Monte Carlo simulations can be directly used to solve the integral term over the domain of y and θ_i with the joint PDF $p_{y,\theta_i}(y, \theta_i)$ as summarized in Algorithm 1. Note that only one for-loop is required in the algorithm. In the algorithm, the unknown PDFs $p_y(y)$ and $p_{y,\theta_i}(y, \theta_i)$ are estimated with the multivariate kernel density estimator toolbox (Botev et al., 2010). Equation (7) and Algorithm 1 can be directly extended to $\delta_{\boldsymbol{u}}$ since the kernel density estimator can also compute joint PDF of higher dimensions $p_{y,\boldsymbol{\theta}_{\boldsymbol{u}}}(y, \boldsymbol{\theta}_{\boldsymbol{u}})$.

3.3 Sampling strategy for Correlated Model Parameters with Arbitrary Distributions

Although the marginal PDFs $\mathbf{p}(\boldsymbol{\theta})$ and the correlation matrix Σ of the uncertain model parameters can be derived from experimental data, the joint PDF $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ to generate sample set $\mathcal{A}$ is still missing (Step 1 in Algorithm 1). For independent uncertain parameters, the joint PDF is simply given by the product of the marginal distributions as

$$p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \prod_{i=1}^{n_\theta} p_i(\theta_i). \tag{8}$$

In the case of model parameter correlations, this simplifying assumption of (8) no longer holds. Alternatively, a Gaussian copula-based approach (Nelsen, 2007) is used to calculate the joint PDF as follows

$$p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \frac{\partial^{n_\theta} F_{n_\theta}[F^{-1}(\mu_1), \cdots, F^{-1}(\mu_{n_\theta}); \Sigma]}{\partial \theta_1 \cdots \theta_{n_\theta}}, \tag{9}$$

where F^{-1} denotes the inverse cumulative distribution function (CDF) of the standard Gaussian distribution, and F_{n_θ} denotes the joint CDF of multivariate standard Gaussian distributions with the correlation matrix Σ. The CDF of $\boldsymbol{\theta}$, which is $[\mu_1, \cdots, \mu_{n_\theta}] = [F_1(\theta_1), \ldots, F_{n_\theta}(\theta_{n_\theta})]$, can be arbitrary.

Practically, it is non-trivial to directly calculate $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ from (9) for parameter sampling. Therefore, we present a sampling procedure according to (9) in combination with the inverse Nataf transformation (Nelsen, 2007) as summarized in Algorithm 2. Within Algorithm 2, sample set $\mathcal{A}$ is derived from the sampling of a multivariate standard normal distribution which can be directly generated from the *randn* function in MATLAB®.

Algorithm 2 Sampling for correlated random variables, adapted from Lataniotis et al. (2015)

1: Generate samples $\mathbf{G} = [\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^N]$ from $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: Perform Cholesky decomposition for the correlation matrix $\Sigma = \mathbf{L}\mathbf{L}^T$, where L is a lower triangular matrix;
3: Add correlations to the samples $\mathbf{G}$, $\mathbf{V} = \mathbf{L}\mathbf{G}$;
4: Transform $\mathbf{V}$ into samples of the Gaussian copula, $\mathbf{W} = [F(V_1), \cdots, F(V_\theta)]$;
5: Transform $\mathbf{W}$ into samples of uncertain parameters $\boldsymbol{\theta}$, $\mathbf{A} = [F_1^{-1}(W_1), \cdots, F_\theta^{-1}(W_\theta)]$.

3.4 Polynomial Chaos Expansion

The main computational burden for solving (7) is the repeated evaluation of the dynamic system (1) as indicated in Step 2 of Algorithm 1. The single-loop Monte Carlo simulation approach has reduced the required sample number significantly compare to the double-loop approach,

but at least 10,000 Monte Carlo simulations are required to approximate the integral term in (7) accurately which might be still prohibitive in the case of computationally expensive model evaluations. To confront this problem, an easy-to-evaluate PCE model is derived first.

Consider $y(\boldsymbol{\theta})$ as a random variable of finite variance, which can be represented as (Sudret, 2008)

$$y(\boldsymbol{\theta}) = \sum_{k=0}^{\infty} \alpha_k \Psi_k(\boldsymbol{\theta}), \tag{10}$$

where $\{\Psi_k(\boldsymbol{\theta})\}_{k=0}^{\infty}$ and $\{a_k\}_{k=0}^{\infty}$ are multidimensional orthogonal polynomials and polynomial coefficients, respectively. Assuming independent model parameters, the multivariate polynomials $\Psi_k(\boldsymbol{\theta})$ can be constructed as a product of univariate polynomials (Xiu and Karniadakis, 2002) according to

$$\Psi_k(\boldsymbol{\theta}) = \Phi^1_{k_1}(\theta_1)\Phi^2_{k_2}(\theta_2)\cdots\Phi^{n_\theta}_{k_{n_\theta}}(\theta_{n_\theta}), \tag{11}$$

where $\Phi^i_{k_i}(\theta_i)$ denotes the univariate polynomial of order k_i for model parameter θ_i. The univariate polynomial type is chosen based on the PDF of parameter θ_i. Algorithms for constructing orthogonal univariate polynomials for standard and arbitrary PDFs can be found in Xiu and Karniadakis (2002) and Oladyshkin and Nowak (2012). For practical reasons, the infinite expansion in (10) is truncated and substituted by

$$y(\boldsymbol{\theta}) = \sum_{k=0}^{P-1} \alpha_k \Psi_k(\boldsymbol{\theta}), \tag{12}$$

where P is the number of the retained polynomials and determined by the expansion order and the hyperbolic index (Marelli and Sudret, 2014). The coefficients α_k in (12) can be estimated by using a non-intrusive simulation-based approach. More details about the truncation strategy and coefficient estimation are referred to Blatman and Sudret (2011). Note that the PCE model is determined by ignoring the parameter correlations first. Only the sample set $\mathcal{A}$ used to calculate the parameter sensitivities reflects the parameter correlation subsequently.

4. CASE STUDY: ENZYME-CATALYZED REACTION NETWORK

Typically, enzymatic-catalyzed processes operate under mild conditions while at the same time they ensure a high selectivity (Stillger et al., 2006; Nakamura et al., 2003). In this work, we are interested in the synthesis of enantiopure hydroxy ketones catalyzed by benzaldehyde lyase from *Pseudomonas fluorescens* (*Pf*BAL), which are important building blocks in the chemical and pharmaceutical industries (Demir et al., 2001). The biochemical reaction network shown in Fig. 1 includes the substrates benzaldehyde (B) and propanal (A), the side product (*R*)-benzoin (BB), and the desired product (*R*)-2-hydroxy-1-phenylbutan-1-one (BA). The five reaction steps are determined by 10 kinetic parameters $\theta_i, i = 1, \cdots, 10$.

Model-based process optimization tools might be applied to obtain a higher selectivity and yield of the target product BA (Ploch, 2014). However, the kinetic parameters estimated from experiments are uncertain and correlated. These parameter imperfections might affect the performance of the model-based design negatively; i.e., the derived operating condition is sub-optimal or biased . Therefore, we investigate the effect of parameter variation and parameter correlation on the simulation results.

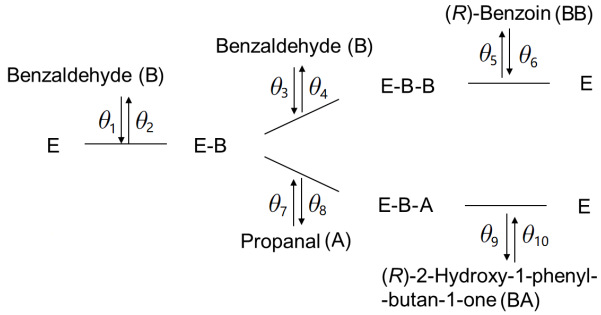


Fig. 1. Reaction network of the synthesis of desired product (*R*)-2-hydroxy-1-phenylbutan-1-one (BA) and side product (*R*)-benzoin (BB) from benzaldehyde (B) and propanal (A) catalyzed by *Pf*BAL (E) (Ploch, 2014)

4.1 Mathematical Model

A batch reactor model is assumed for the sensitivity analysis (Ploch, 2014). Based on mass conservation, the governing equations read as

$$\frac{dc_A}{dt} = -\frac{N_{BA}}{D}c_E, \tag{13}$$

$$\frac{dc_B}{dt} = -\frac{2N_{BB} + N_{BA}}{D}c_E, \tag{14}$$

$$\frac{dc_{BA}}{dt} = -\frac{N_{BA}}{D}c_E, \tag{15}$$

$$\frac{dc_{BB}}{dt} = -\frac{N_{BB}}{D}c_E, \tag{16}$$

$$\frac{dc_E}{dt} = (-\theta_{11}\cdot c_A - \theta_{12}\cdot c_B - \theta_{13})c_E, \tag{17}$$

where c_i denotes the concentration of the i-th substrate; θ_{11}, θ_{12}, and θ_{13} are kinetic parameters that describe the deactivation in the enzymes; and N_{BA}, N_{BB}, and D are abbreviations for the nominators and denominators of the rate equations, which consist of the kinetic parameters of the reaction steps in Fig. 1 and concentration of components. For more details, we refer to Ploch (2014). Equations (13) to (16) describe the concentrations of the reactants and products, while (17) describes the concentration of the enzyme that remains active in the reactor. Uncertainties of all the kinetic parameters are described by a multivariate normal distribution with $\theta_i \sim \mathcal{N}(m_i, m_i \times 10\%)$, $\forall i = 1, \cdots, 13$ and the correlation matrix Σ as given in Fig. 2. The correlation matrix and parameter values are obtained from experimental data (Ohs et al., 2017). The mean values (m_i) of the kinetic parameters are summarized in Table 1.

Independent and correlated parameter samples are generated with Algorithm 2 according to their marginal distributions with the identity matrix I and the correlation matrix Σ, respectively. For the PCE setting, we use an expansion order of 6 and hyperbolic index of 0.5 (Marelli and Sudret, 2014). To calibrate the PCE model, 500 reference simulations were sufficient to guarantee a credible approximation; i.e., a dramatic reduction of computational

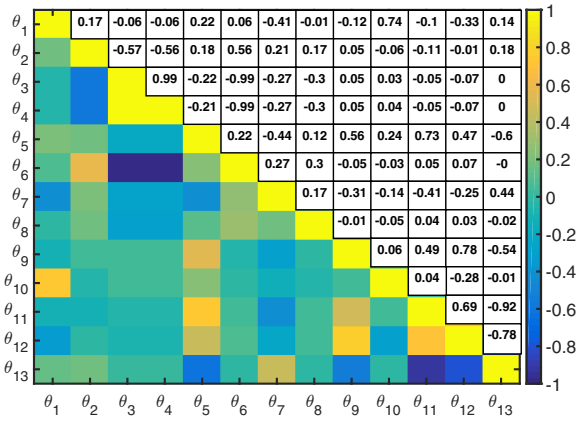


Fig. 2. Correlation values of all kinetic parameters

Table 1. Mean values of all kinetic parameters

θ_1	6.0×10^5	$(mM \cdot min)^{-1}$	θ_8	14	$(min)^{-1}$
θ_2	5.3×10^5	$(min)^{-1}$	θ_9	5.7×10^2	$(min)^{-1}$
θ_3	1.4×10^5	$(mM \cdot min)^{-1}$	θ_{10}	5.7×10^4	$(mM \cdot min)^{-1}$
θ_4	2.3×10^4	$(min)^{-1}$	θ_{11}	2.4×10^{-3}	$(mM \cdot min)^{-1}$
θ_5	1.3×10^6	$(min)^{-1}$	θ_{12}	1.7×10^{-3}	$(mM \cdot min)^{-1}$
θ_6	2.0×10^6	$(mM \cdot min)^{-1}$	θ_{13}	1.1×10^{-4}	$(min)^{-1}$
θ_7	3.3×10^3	$(mM \cdot min)^{-1}$			

costs compared to direct Monte Carlo simulations with 10,000 simulations. In parallel, Algorithm 1 is carried out for the correlated samples to perform MISA in MATLAB® using the UQLAB toolbox (Marelli and Sudret, 2014).

4.2 Results

The effects of the 13 kinetic parameters on different substrate concentrations are presented at time points 1.5, 40, 100 and 145 min. Results for enzyme E are displayed in Fig. 3. As we can see, the sensitivities at different time points have only slight differences, and θ_{11} always dominates the uncertainty of c_E. By comparing the results for the independent and correlated cases, we observe a significant increase in the sensitivities of θ_5, θ_{12}, and θ_{13} that is based on the strong correlations between them and θ_{11}. In Fig. 4, we show the results of the side product BB. Here, c_{BB} is dominated by the first six parameters at the start of the reaction. The impact of θ_9, θ_{10}, θ_{11} increases over time. Evident differences between the results for the independent and correlated case are also observed in Fig. 4. In Fig. 5, we summarize the sensitivities for the desired product BA. The impact of θ_9 on c_{BA} dominates at the start of the reaction and becomes similar to θ_7, θ_{10} and θ_{11} over time. The correlations among the parameters affect the sensitivity results here as well.

In addition, the savage score correlation coefficient (SSCC; Iman and Conover (1987)), which quantifies the agreement between two rankings, is used to compare the sensitivity results between independent and correlated parameter cases. It has a value between -1 and 1, where -1 and 1 indicate identical and opposite results, respectively. The SSCC for each analysis is shown in the corresponding figures, which also reveals a distinct effect of parameter correlations on the sensitivity results. The effect of parameter correlations on the entire PDF of the model output,

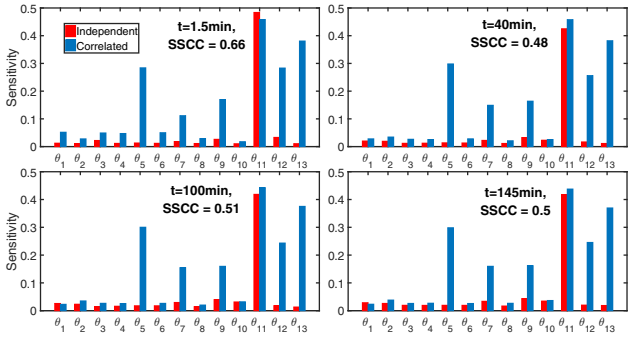


Fig. 3. MISA-based sensitivities of 13 parameters for enzyme E in the absence and presence of correlations at time points 1.5, 40, 100, and 145 min

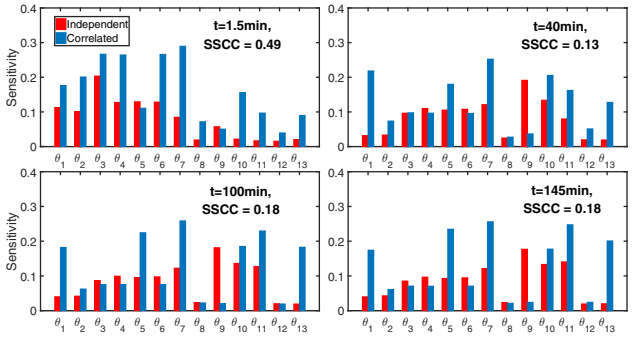


Fig. 4. MISA-based sensitivities of 13 parameters for substrate BB in the absence and presence of correlations at time points 1.5, 40, 100, and 145 min

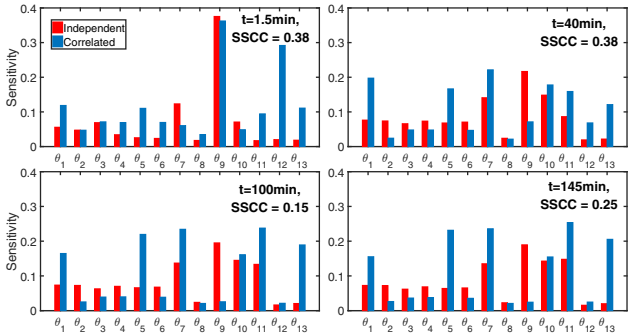
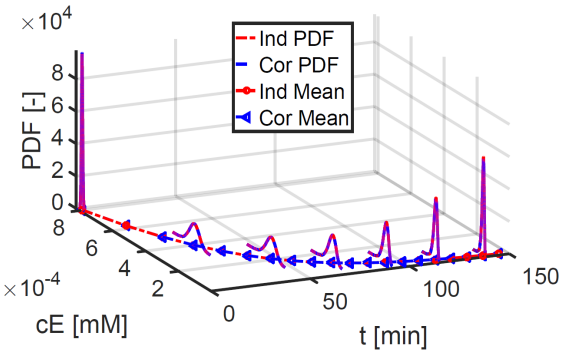


Fig. 5. MISA-based sensitivities of 13 parameters for substrate BA in the absence and presence of correlations at time points 1.5, 40, 100, and 145 min

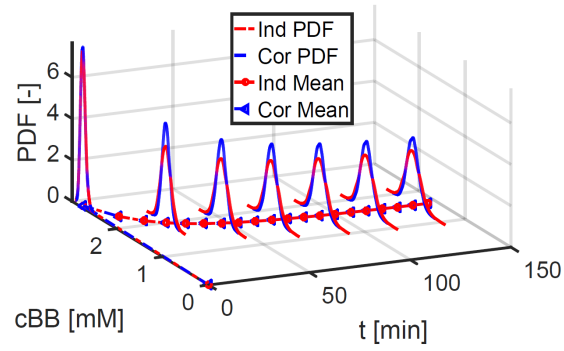
in turn, is minor as shown in Fig. 6. Only for c_{BB} and c_{BA} the PDFs in the presence and absence of parameter correlations are slightly distinct.

5. CONCLUSIONS

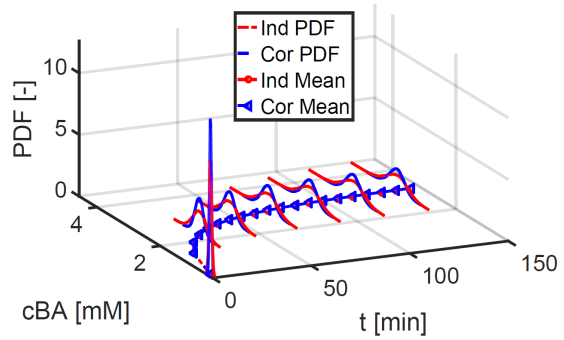
This work reveals the potential of MISA and its efficient computational implementation for analyzing realistic and complex dynamic systems in the absence and presence of parameter correlations. In detail, we demonstrated the use of MISA for enzyme-catalyzed reactions with independent and correlated model parameters, respectively. A single-loop Monte Carlo simulation method combined with PCE was used to compute the sensitivity indicator δ for all the kinetic parameters at low computational cost. According to the derived results, we observed that parameter corre-



(a)



(b)



(c)

Fig. 6. The time evolution of the probability density function (PDF) of concentrations c_E, c_{BB}, and c_{BA}, determined with 10,000 Monte Carlo simulations

lations can strongly affect the sensitivity ranking and the PDF of the model output. Therefore, it is necessary to consider parameter correlations in process analysis and design. In future work, we will transfer the proposed framework to model-based optimal experimental design problems.

REFERENCES

Biegler, L.T. (2010). *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. SIAM.

Blatman, G. and Sudret, B. (2011). Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6), 2345–2367.

Borgonovo, E. and Tarantola, S. (2008). Moment independent and variance-based sensitivity analysis with correlations: An application to the stability of a chemical reactor. *International Journal of Chemical Kinetics*, 40(11), 687–698.

Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6), 771–784.

Borgonovo, E. and Plischke, E. (2016). Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, 248(3), 869–887.

Botev, Z.I., Grotowski, J.F., Kroese, D.P., et al. (2010). Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5), 2916–2957.

Demir, A.S., Hamamci, H., Sesenoglu, O., Aydogan, F., Capanoglu, D., and Neslihanoglu, R. (2001). Simple chemoenzymatic access to enantiopure pharmacologically interesting (r)-2-hydroxypropiophenones. *Tetrahedron: Asymmetry*, 12(13), 1953–1956.

Iman, R.L. and Conover, W. (1987). A measure of top–down correlation. *Technometrics*, 29(3), 351–357.

Lataniotis, C., Marelli, S., and Sudret, B. (2015). Uqlab user manual–the input module.

Marelli, S. and Sudret, B. (2014). Uqlab: A framework for uncertainty quantification in matlab. In *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, 2554–2563.

Nakamura, K., Yamanaka, R., Matsuda, T., and Harada, T. (2003). Recent developments in asymmetric reduction of ketones with biocatalysts. *Tetrahedron: Asymmetry*, 14(18), 2659–2681.

Nelsen, R.B. (2007). *An introduction to copulas*. Springer Science & Business Media.

Ohs, R., Fischer, K., Schöpping, M., and Spiess, A.C. (2017). Model-based experimental analysis of an enzyme-catalyzed cross-ligation. *In preparation*.

Oladyshkin, S. and Nowak, W. (2012). Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, 106, 179–190.

Ploch, T. (2014). Model-based optimization of an enzyme-catalyzed reaction network. *Master thesis, RWTH Aachen University*.

Rajabi, M.M., Ataie-Ashtiani, B., and Simmons, C.T. (2015). Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations. *Journal of Hydrology*, 520, 101–122.

Stillger, T., Pohl, M., Wandrey, C., and Liese, A. (2006). Reaction engineering of benzaldehyde lyase from pseudomonas fluorescens catalyzing enantioselective c- c bond formation. *Organic process research & development*, 10(6), 1172–1177.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7), 964–979.

Telen, D., Vercammen, D., Logist, F., and Van Impe, J. (2014). Robustifying optimal experiment design for nonlinear, dynamic (bio) chemical systems. *Computers & Chemical Engineering*, 71, 415–425.

Wei, P., Lu, Z., and Yuan, X. (2013). Monte carlo simulation for moment-independent sensitivity analysis. *Reliability Engineering & System Safety*, 110, 60–67.

Xiu, D. and Karniadakis, G.E. (2002). The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2), 619–644.