Computers and Chemical Engineering 138 (2020) 106814

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Surrogate modeling for fast uncertainty quantification: Application to 2D population balance models

Georgios Makrygiorgos[a], Giovanni Maria Maggioni[b,c], Ali Mesbah[a,*]

[a] Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA
[b] School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c] Engineering and Technology, Environmental and Solids Processing, Bayer AG, 51368 Leverkusen, Germany

ARTICLE INFO

Article history:
Received 27 December 2019
Revised 16 February 2020
Accepted 12 March 2020
Available online 16 March 2020

Keywords:
Surrogate models
Uncertainty quantification
Probabilistic uncertainty
Sparse polynomial chaos
Kriging
Population balance models

ABSTRACT

Surrogate modeling is a useful tool for enabling uncertainty quantification (UQ) tasks that require many expensive model evaluations, as it replaces expensive high-fidelity models with cheap-to-evaluate surrogates. This paper investigates sparse polynomial chaos and Kriging methods for surrogate modeling of first-principles models with probabilistic uncertainty in parameters and initial conditions. The surrogate modeling methods are demonstrated on a 2-dimensional population balance (2D-PB) model for batch cooling crystallization of ibuprofen with 20 uncertain parameters. Our analysis indicates that not only sparse polynomial chaos expansions are powerful for probabilistic UQ, but also the approximation accuracy of Kriging surrogate models can be significantly improved when polynomial chaos expansions are used to describe their trend. A basis-adaptive least-angle-regression strategy is shown to be particularly useful for inducing sparsity in polynomial chaos expansions, allowing for dealing with problems with a relatively large number of uncertain inputs. The utility of sparse polynomial chaos- and Kriging-based surrogate models is illustrated for various forward and inverse UQ problems, including global sensitivity analysis as well as Bayesian and maximum a posteriori parameter estimation of the 2D-PB model, where massive savings in computational cost (up to 30,000-fold) are observed.

1. Introduction

High-fidelity models based on first principles are indispensable tools for describing and understanding the complex behavior of technical systems. Nonetheless, first-principles models are generally subject to various sources of uncertainty. These include: 1) model structure uncertainty due to, for example, incomplete system knowledge in describing chemical reaction kinetics or physicochemical phenomena, e.g., nucleation, growth, and agglomeration in particulate processes, 2) uncertainty in model parameters and/or initial conditions, e.g., uncertainty in kinetic parameters, and 3) experimental uncertainty, e.g., measurement noise. The existence of uncertainty in model predictions, along with the intricate interplay between the various sources of uncertainty, poses an important challenge when applying high-fidelity models in decision-support tasks such as parameter estimation Renardy et al. (2018); Hermanto et al., 2008, optimal experiment design Huan and Marzouk (2013); Streif et al. (2014), or optimal control Paulson and Mesbah (2019). Hence, it is crucial to quantify and analyze the impact of model uncertainty on the predictions of quantities of interest (QoIs), particularly to ensure compliance of processes and products with safety and regulatory requirements.

In recent years, computational methods for efficient uncertainty quantification (UQ) have been the subject of intense research in science and engineering (e.g., see Najm (2009); Russi et al. (2010)). UQ consists of the forward problem of propagating uncertain model inputs to predictions of QoIs as well as the inverse problem of estimating unknown model inputs from measurements under experimental uncertainties. There is a variety of approaches to forward UQ. In deterministic perturbation methods, a model is run using perturbed inputs around their assumed values to characterize how the QoIs are affected Elishakoff et al., 1994; Ma and Braatz (2001); Gunawan et al. (2002); Sudret (2008); Kim et al. (2013). Alternatively, probabilistic UQ methods look to obtain probability distributions of QoIs by propagating the distribution of uncertain inputs through a system model. To this end, Monte Carlo-based (MC-based) methods are widely used for sample-based approximation of the QoI distributions Nagy and Braatz (2007). A useful feature of MC-based methods is that their convergence rate is independent of the number of uncertain inputs. That is, the convergence rate is $O(1/\sqrt{N_s})$ as the number of

* Corresponding author.
E-mail address: mesbah@berkeley.edu (A. Mesbah).

https://doi.org/10.1016/j.compchemeng.2020.106814
0098-1354/

samples $N_s \to \infty$ Caflisch, 1998. This is particularly advantageous for UQ of complex dynamical systems with a high uncertainty dimension. However, the error convergence rate is slow, often requiring a large number of samples to achieve acceptable accuracy. As such, even for a moderate number of uncertain inputs (on the order of 10), a large number of samples may be required to achieve accurate uncertainty propagation, which can make the UQ tasks computationally prohibitive. In MC-based UQ, this computational challenge is further compounded with the cost of running expensive high-fidelity models.

This paper looks to demonstrate the utility of surrogate modeling (also known as metamodeling) for fast UQ of complex systems. Surrogate models are computationally cheap-to-evaluate surrogate representations of high-fidelity, expensive-to-run, models Maceiczyk and Demello (2014); del Rio-Chanona et al. (2019). The key notion is to construct surrogates for high-fidelity models using a limited amount of simulation training data, such that the high-fidelity simulations are used as "black-box" (i.e., without any modification of the governing equations) Schenkendorf et al. (2017); Luu Trung Duong et al. (2018). Surrogate modeling has been shown to be useful for performing expensive UQ tasks that require predicting QoIs for many random realizations of the uncertain model inputs Zhang and Sahinidis, 2013; Sudret et al. (2017). As such, surrogate models are particularly advantageous for sample-based estimation of probability distribution of QoIs, or their statistical moments.

In this work, we investigate sparse polynomial chaos Blatman, Sudret, 2011 and Kriging Hu and Mahadevan (2016) methods for surrogate modeling of first-principles models that are subject to time-invariant, probabilistic uncertainty in parameters and initial conditions. The surrogate modeling methods are demonstrated on a two-dimensional population balance (2D-PB) problem. PB models are widely used for describing disperse, multi-phase systems in physics (e.g., statistical physics Smoluchowski (1916)), industrial processes (e.g., crystallization, colloidal coalescence, reactive precipitation Hulburt and Katz (1964); Randolph (1964); Mesbah et al. (2011)), and more recently medical research (e.g., cancer or leukemia progression Ramkrishna and Singh (2014); Solsvik and Jakobsen (2015)). The fact that dynamic PB models become computationally excessively expensive in two or higher dimensions (e.g., when crystals are modeled as 2D objects, or when external spatial coordinates are considered Ramkrishna (2000); Marchisio and Fox (2013); Sherer et al. (2007); Jakobsen (2008); Marchisio and Fox (2013)) warrants the use of surrogate modeling for their fast UQ.

We consider a 2D-PB model of batch cooling crystallization of ibuprofen with probabilistic uncertainty in 20 parameters related to the kinetics of growth and dissolution, as well as to the physico-chemical properties of the crystals Maggioni et al. (2017); Crose et al. (2015); Ma and Roberts (2018). We first demonstrate the effectiveness of a basis-adaptive least-angle-regression strategy Blatman, Sudret, 2011 for building sparse polynomial chaos surrogate models that can handle the relatively large number of uncertain parameters in the 2D-PB model. We then explore polynomial chaos-Kriging surrogate models, where polynomial chaos expansions are used as the trend (or mean) of the Kriging surrogate model, leading to improved predictions of the global behavior of QoIs compared to using generic polynomial trends. At the same time, the variance of local Kriging predictions allows for quantifying the uncertainty of predicted QoIs Schöbi and Sudret (2014). Finally, the utility of surrogate models is illustrated for a variety of forward and inverse UQ problems, including global sensitivity analysis as well as Bayesian and maximum a posteriori parameter estimation of the 2D-PB model.

The remainder of the paper is structured as follows. Section 2 discusses the sparse polynomial chaos and Kriging surrogate modeling methods. Section 3 introduces the 2D-PB model of the batch cooling crystallization process, followed by performance analysis of the different surrogate modeling methods. Section 4 demonstrates the application of surrogate modeling for fast probabilistic UQ. The paper is concluded by discussing potential avenues for future research.

2. Overview of surrogate modeling methods

Consider a computationally expensive, first-principles model $\mathcal{M}$ that describes a physical system and contains M uncertain inputs (i.e., model parameters and/or initial conditions), denoted by $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_M\}, \boldsymbol{\xi} \in \mathbb{R}^M$. Uncertainty in model parameters and initial conditions can be represented by a random variable $\boldsymbol{\Xi}$ with some joint probability density function (pdf) $f_{\boldsymbol{\Xi}}$, from which random realizations $\boldsymbol{\xi}$ are drawn. The support of this joint pdf is $\mathcal{D}_{\boldsymbol{\Xi}}$. "a remove". The model $\mathcal{M}$ is a mapping in the form of $\mathcal{M} : \mathbb{R}^M \to \mathbb{R}^N$ from the space of uncertain inputs to the space of N quantities of interest (QoIs) predicted by the model. This mapping does not have to be known explicitly; hence, the model can also be a "black-box". Let $\hat{\mathcal{M}}$ denote an approximation of the mapping $\mathcal{M}$ and $\mathbf{Y}$ a vector representing the QoIs

$$\mathbf{Y} = \mathcal{M}(\boldsymbol{\Xi}) \approx \hat{\mathcal{M}}(\boldsymbol{\Xi}). \tag{1}$$

The approximation $\hat{\mathcal{M}}$ is a so-called surrogate model that can be used to replace the original model $\mathcal{M}$. We denote the predictions of the surrogate model as $\hat{\boldsymbol{Y}}$, such that $\hat{\boldsymbol{Y}} = \boldsymbol{Y} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ denotes the error on the predictions. All surrogate modeling methods discussed in this paper have a representational form of $\hat{\mathcal{M}}$ through continuous functions, scaled by some coefficients. In order to estimate the unknown coefficients of the surrogate models, it is imperative to use training data, referred to as experimental design (ED). There are many methods to generate the ED, such as Monte Carlo (MC) sampling, Latin Hypercube sampling (LHS) Shapiro (2003), or Sobol' sequences Sobol' And and Levitan (1999). The experimental design consists of N_{ed} samples $\widetilde{\boldsymbol{\Xi}} = \{\boldsymbol{\xi_1}, \boldsymbol{\xi_2}, \ldots, \boldsymbol{\xi_{N_{ed}}}\}$ with corresponding responses $\widetilde{\boldsymbol{Y}} = \{\mathcal{M}(\boldsymbol{\xi_1}), \mathcal{M}(\boldsymbol{\xi_2}), \ldots, \mathcal{M}(\boldsymbol{\xi_{N_{ed}}})\}$. Note that the quantity and quality of the training data directly affect the quality of the surrogate model approximation.

There are various approaches for estimating the coefficients of surrogate models. We primarily focus on non-intrusive methods that have attained substantial attention recently Jones et al. (2013), and usually involve solving some nonlinear regression problem. The advantage of non-intrusive approaches is that they do not rely on implicit equations for the QoIs or any manipulations of the equations, contrary to the intrusive Galerkin methods Ghanem and Spanos (1991). Once a surrogate model $\hat{\mathcal{M}}$ is established, its capability to reproduce the responses of the original computational model $\mathcal{M}$ should be verified. Whenever the data are limited and must be used both for training and validation, one can resort to the so-called "leave-k-out" cross-validation Goodfellow et al. (2016). The leave-one-out (LOO) cross validation error estimate, ϵ_{loo}, is known to prevent overfitting and is generally preferred in statistical learning models. The LOO error is defined as

$$\epsilon_{loo} = \frac{\sum_{i=1}^{N_{ed}} \left(\mathcal{M}\left(\boldsymbol{\xi}^i\right) - \hat{\mathcal{M}}^{\setminus i}\left(\boldsymbol{\xi}^i\right)\right)^2}{\sum_{i=1}^{N_{ed}} \left(\mathcal{M}\left(\boldsymbol{\xi}^i\right) - \mu_{\widetilde{\boldsymbol{Y}}}\right)^2}, \tag{2}$$

where $\hat{\mathcal{M}}^{\setminus i}$ denotes the surrogate model constructed without taking the i^{th} experimental design realization into account and $\mu_{\widetilde{\boldsymbol{Y}}}$ is the empirical mean of the training set.

The generalization capability of a surrogate model, however, may be overestimated when relying only on the LOO cross-validation error. Therefore, we introduce another crucial performance assessment metric, the validation error, ϵ_{val} , based on N_{val}

samples different from those used to generate the training set. The validation error is defined as

$$\epsilon_{val} = \left(\frac{N_{val}-1}{N_{val}}\right)\frac{\sum_{i=1}^{N_{val}}\left(\mathcal{M}\left(\boldsymbol{\xi}^i\right)-\hat{\mathcal{M}}\left(\boldsymbol{\xi}^i\right)\right)^2}{\sum_{i=1}^{N_{val}}\left(\mathcal{M}\left(\boldsymbol{\xi}^i\right)-\mu_{\tilde{\boldsymbol{Y}}_{\boldsymbol{val}}}\right)^2}, \tag{3}$$

where $\mu_{\tilde{\boldsymbol{Y}}_{\boldsymbol{val}}}$ is the empirical mean of the validation samples. Typically, N_{val} should be a few orders of magnitude larger than N_{ed} to ensure a reliable assessment of the generalization of the surrogate model. To summarize, the surrogate model coefficient estimation generally relies heavily on the LOO cross-validation error, as it is estimated based on the training data. If a validation set is available, it can be used to ensure the predictive capability of the surrogate model. In this work, we use random samples from the original, expensive-to-evaluate model to validate the results of the surrogate models. In a real application, however, cross-validation error could be used to estimate the accuracy of the surrogate model. Moreover, although we extensively sample the original model in this work one could also use a much lower number of samples to get similar information about the validation error at the expense of loss in the validation error accuracy. We now present the polynomial chaos- and Kriging-based surrogate modeling methods considered in this work.

2.1. Polynomial chaos expansions

Polynomial chaos (PC) has been established as a popular surrogate modeling method for estimation and optimal control of uncertain nonlinear systems (e.g., Nagy and Braatz (2007); Sudret (2008); Chaffart and Ricardez-Sandoval (2017); Wang et al. (2018)). Polynomial chaos expansions (PCEs) are based on the assumption that a finite-variance QoI can be represented by an infinite series Xiu and Em Karniadakis (2002)

$$\boldsymbol{Y} = \sum_{\boldsymbol{a}\in\mathbb{N}^M} y_{\boldsymbol{a}}\Psi_{\boldsymbol{a}}(\boldsymbol{\Xi}). \tag{4}$$

The basis functions $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\Xi})$ in the expansion of $\boldsymbol{Y}$ are multivariate polynomials (orthogonal with respect to $f_{\boldsymbol{\Xi}}$) and $\boldsymbol{y_a}$ are the coefficients of the basis functions. $\mathbb{N}$ is the set of Natural numbers and the multi-index $\boldsymbol{\alpha}$ is an M-dimensional vector in $\mathbb{N}$. In this form, i.e., an infinite series representation, an infinite number of coefficients is retrieved. The orthogonality condition holds over $\mathcal{D}_{\boldsymbol{\Xi}}$, which is the support of the joint distribution $f_{\boldsymbol{\Xi}}$, and is defined as

$$\mathbb{E}\{\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\Xi})\Psi_{\boldsymbol{\beta}}(\boldsymbol{\Xi})\} = \int_{\mathcal{D}_{\boldsymbol{\Xi}}}\Psi_{\boldsymbol{\alpha}}\left(\boldsymbol{\xi}\right)\Psi_{\boldsymbol{\beta}}\left(\boldsymbol{\xi}\right)f_{\boldsymbol{\Xi}}\left(\boldsymbol{\xi}\right)d\boldsymbol{\xi} = \delta_{\alpha\beta}, \quad \forall\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^M, \tag{5}$$

where $\delta_{\alpha\beta}$ denotes the Kronecker delta. Constructing the orthogonal basis $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\Xi})$ relies on the form of the joint pdf of the uncertain inputs $f_{\boldsymbol{\Xi}}$. Expression (5) represents the tensor product of M univariate polynomials that are orthonormal with respect to their corresponding marginal density f_{ξ_i}. For several continuous and discrete probability distributions, optimal series convergence is established based on the Wiener-Askey scheme Xiu and Em Karniadakis (2002). Note that one may choose an arbitrary family of orthogonal basis functions for representing a given uncertain input and the PCE will still be convergent. The Wiener-Askey scheme, however, provides the optimal L_2-convergence rate with increasing polynomial order by the Martin-Cameron theorem Cameron, Martin, 1947; Xiu and Karniadakis (2002). For arbitrary probability distributions, the orthogonal polynomials can be constructed using the three-term recurrence procedure Gautschi (1982). When the joint probability density $f_{\boldsymbol{\Xi}}$ does not have an independent copula (i.e., the uncertain inputs are statistically dependent) an iso-probabilistic transformation, such as the Rosenblatt transformation, that preserves the pdf of the uncertainty elements can be used Paulson and Mesbah (2018a). In this work, we focus on uncertain parameters whose joint pdf has an independent copula and whose optimal univariate polynomial representation possesses an analytical expression based on the Wiener-Askey scheme.

For practical reasons, the expansion (4) must be truncated up to a finite order. The truncated polynomial chaos expansion can be cast as follows

$$\hat{\boldsymbol{Y}} = \hat{\mathcal{M}}^{PC} = \sum_{\boldsymbol{a}\in\mathcal{A}} y_{\boldsymbol{a}}\Psi_{\boldsymbol{a}}(\boldsymbol{\Xi}) = \boldsymbol{y}^T\Psi(\boldsymbol{\Xi}). \tag{6}$$

The order of the expansion is described by the multi-index $\boldsymbol{a}\in\mathcal{A}$, where $\mathcal{A}\subset\mathbb{N}^M$ represents the set of the multi-indices kept in the truncated expansion. The truncation scheme aims to limit the possible infinite expansion to a series of maximum order p, so that $\mathcal{A}^{M,p} = \{\boldsymbol{a} : |\boldsymbol{a}| \le p\}$ and thus the cardinality of $\mathcal{A}$ is equal to $P = \binom{M+p}{p}$. The coefficients $\boldsymbol{y}$ of (6) can be determined by solving a least-squares problem

$$\widehat{\boldsymbol{y}} = \underset{y\in\mathbb{R}^P}{\operatorname{argmin}}\,\mathbb{E}\left[\left(\mathcal{M}(\boldsymbol{\Xi}) - \mathbf{y}^{\mathrm{T}}\Psi(\boldsymbol{\Xi})\right)^2\right], \tag{7}$$

which seeks to minimize the discrepancy between predictions of the surrogate model $\widehat{\mathcal{M}}^{PC}$ and the model $\mathcal{M}$. When an ED is available, the coefficient estimation problem can be solved analytically by ordinary least squares (OLS) as

$$\hat{\boldsymbol{y}} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\tilde{\boldsymbol{Y}}, \tag{8}$$

where matrix $\boldsymbol{A}$ is the so-called experimental or design matrix whose elements are defined as $A_{ij} = \Psi_j\left(\boldsymbol{\xi}_i\right)$, $i = 1,\ldots,N_{ed}, j = 1,\ldots,P$, i.e., the basis functions evaluated on the ED points. Accounting for more uncertain inputs in (6) drastically increases the cardinality of $\mathcal{A}$ and, therefore, the number of coefficients to be estimated. A similar challenge arises when increasing the surrogate model complexity by adding higher order terms to the expansion. To address these challenges, we look to introduce sparsity in the expansion (6). First, the basic truncation scheme can be replaced by the hyperbolic truncation scheme, also known as the q-norm scheme,

$$\mathcal{A}^{M,p,q} = \{\boldsymbol{a}\in\mathcal{A}^{M,p} : ||\boldsymbol{a}||_q \le p\}, ||\boldsymbol{a}||_q = \left(\sum_{i=1}^{M} a_i^q\right)^{\frac{1}{q}}. \tag{9}$$

The main idea behind the hyperbolic truncation scheme is that the high-order univariate polynomials for each single variable should appear in the expansion, but high-order interaction terms are omitted. This follows the sparsity-of-effects principle, indicating that only low-order interactions in the uncertain inputs are relevant . For $q = 1$ the full basis is retained, while terms disappear from the expansion as q decreases. Second, once a basis has been constructed using the q-norm scheme, further sparsity effects can be introduced by modifying the coefficient estimation problem. This is achieved by adding a regularization term $||\boldsymbol{y}||_1 = \sum_{\boldsymbol{a}\in\mathcal{A}}\boldsymbol{y_a}$, which yields low-rank solutions and leads to the regularized estimation problem Hastie et al. (2015)

$$\widehat{\boldsymbol{y}} = \underset{y\in\mathbb{R}^P}{\operatorname{argmin}}\,\mathbb{E}\left[\left(\mathcal{M}(\boldsymbol{\Xi}) - \mathbf{y}^{\mathrm{T}}\Psi(\boldsymbol{\Xi})\right)^2\right] + \lambda\,\|\,\boldsymbol{y}\|_1. \tag{10}$$

A key aspect in solving (10) is the proper choice of the regularization parameter $\lambda \ge 0$, which affects the number of non-zero coefficients retained in the expansion (6). The penalized problem can be efficiently solved with the least-angle-regression (LAR) algorithm Efron et al. (2004). Briefly, LAR seeks to identify the polynomials $\Psi(\boldsymbol{\Xi})$ (known as "regressors" in the statistical literature) with the greatest impact on the model responses among a large set of candidates based on a q-norm adaptive scheme Blatman, Sudret, 2011. The convergence criterion of the LAR algorithm relies on ϵ_{loo} (see

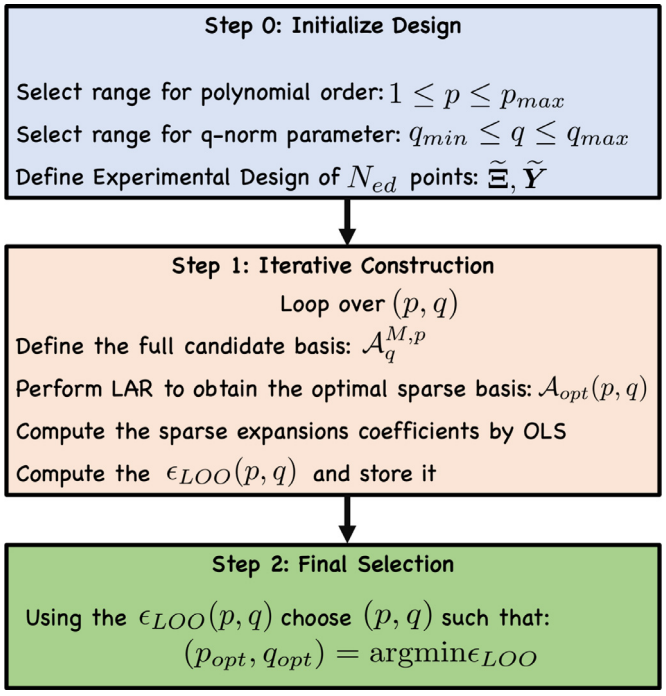

Fig. 1. Sparse polynomial chaos with adaptive order and basis truncation.

(3)). We refer to PCEs constructed by using a hyperbolic truncation scheme and the LAR algorithm as sparse PC (sPC) expansions. Fig. 1 shows the main procedure through which the sPC expansion is constructed.

2.2. Kriging

Kriging, also known as Gaussian process Rasmussen and Williams (2006), is a stochastic interpolation method that was initially adopted in spatial statistics Cressie (1990). In Kriging, the predictions of the model $\mathcal{M}$ are approximated as realizations of a Gaussian process, so that the surrogate model structure takes the form

$$\widehat{\boldsymbol{Y}} = \widehat{\mathcal{M}}^{KG}(\boldsymbol{\xi}) = \boldsymbol{\beta}^T \boldsymbol{q}(\boldsymbol{\xi}) + \sigma^2 Z(\boldsymbol{\xi}). \tag{11}$$

The first term in (11) describes the trend of the Gaussian process, which can assume different functional forms. The trend is analogous to the multivariate basis functions in polynomial chaos, as it attempts to characterize the global behavior of the model response as a function of the uncertain inputs. A general choice for the trend consists of polynomial functions. The second term in (11), $Z(\boldsymbol{\xi})$, is a stationary Gaussian process (zero mean, unit variance) that describes the variance of the surrogate model predictions. Similar to polynomial chaos, the polynomial order in the trend can be defined through the multi-indices that belong to the set $\mathcal{A}^{M,p} = \{\boldsymbol{a} : |\boldsymbol{a}| \leq p\}$. Thus, the general form of the polynomial trend can be given as

$$\boldsymbol{\beta}^T \boldsymbol{q}(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}^{M,p}} \beta_a q_a(\boldsymbol{\xi}), \quad q_a(\boldsymbol{\xi}) = \prod_{i=1}^{M} \xi^{a_i}. \tag{12}$$

The coefficients $\boldsymbol{\beta}$ as well as the variance σ^2 should be estimated using the available experimental design. The zero-mean Gaussian process $Z(\boldsymbol{\xi})$ is determined by the so-called kernel function that defines a pairwise correlation between input samples based on their mutual distance $R(\boldsymbol{\xi}, \boldsymbol{\xi}') = R(|\boldsymbol{\xi} - \boldsymbol{\xi}'|; \boldsymbol{\theta})$. The idea behind the use of a kernel (or auto-correlation) function is that sample points that are "close" (in the absolute distance sense) should produce close outputs. Hence, the kernel function provides a measure of similarity. The hyperparameters $\boldsymbol{\theta}$ of the kernel function should be estimated along with $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. In the present study, we have chosen a kernel function belonging to the class of separable correlation functions Lataniotis et al. (2015)

$$R(\boldsymbol{\xi}, \boldsymbol{\xi}'; \boldsymbol{\theta}) = \prod_{i=1}^{M} R(\xi_i, \xi_i'; \theta_i) \tag{13}$$

and particularly the Matern correlation function

$$R(\boldsymbol{\xi}, \boldsymbol{\xi}'; \boldsymbol{\theta}) = \prod_{i=1}^{M} \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\sqrt{2\nu} \frac{|\xi_i - \xi_i'|}{\theta_i} \right)^{\nu} \mathcal{K}_{\nu} \left(\sqrt{2\nu} \frac{|\xi_i - \xi_i'|}{\theta_i} \right), \tag{14}$$

where ν denotes the degree or shape parameter, Γ is the Euler Gamma function, and $\mathcal{K}$ is the modified Bessel function of second kind Rasmussen and Williams (2006). Although the Matern correlation function is a popular choice, other alternatives exist, such as linear, exponential, and Gaussian kernels Moustapha et al., 2018.

We now discuss the calibration of a Kriging surrogate model, i.e., the procedure of estimating the hyperparameters. Since Kriging is a stochastic interpolation method, the underlying Gaussian process assumption is fundamentally related to the data used to train the surrogate model. Consider an ED of N_{ed} points $\widetilde{\Xi} = \{\boldsymbol{\xi_1}, \boldsymbol{\xi_2}, \ldots, \boldsymbol{\xi_{N_{ed}}}\}$ with responses $\widetilde{\boldsymbol{Y}} = \{\mathcal{M}(\boldsymbol{\xi_1}), \mathcal{M}(\boldsymbol{\xi_2}), ..., \mathcal{M}(\boldsymbol{\xi_{N_{ed}}})\}$. Let $\hat{\boldsymbol{Y}}$ denote the predictions of a Kriging surrogate model on random "unseen" uncertainty realizations $\hat{\boldsymbol{\xi}}$. Then, the vector $\{\hat{\boldsymbol{Y}}, \widetilde{\boldsymbol{Y}}\}$ has a joint Gaussian probability density function. We define $\boldsymbol{r}(\boldsymbol{\xi})$ as the cross-correlation vector that relates unseen realizations $\boldsymbol{\xi}$ to the observations from the ED through the kernel function $\boldsymbol{r}_i(\boldsymbol{\xi}) = R(\boldsymbol{\xi}, \boldsymbol{\xi}_i; \boldsymbol{\theta}), i = 1, 2, \ldots, N_{ed}$. Let $\boldsymbol{R}(\boldsymbol{\xi}, \boldsymbol{\theta})$ be the correlation matrix between the ED samples whose terms are defined as $R_{ij} = R(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j; \boldsymbol{\theta}), i, j = 1, 2, \ldots, N_{ed}$. The trend coefficients and variance explicitly depend on the hyperparameters $\hat{\boldsymbol{\theta}}$ through the following expressions

$$\boldsymbol{\beta}(\widehat{\boldsymbol{\theta}}) = \left(\boldsymbol{Q}^T \boldsymbol{R}^{-1} \boldsymbol{Q}\right)^{-1} \boldsymbol{Q} \boldsymbol{R}^{-1} \widetilde{\boldsymbol{Y}}, \tag{15}$$

$$\sigma^2(\widehat{\boldsymbol{\theta}}) = \frac{1}{N} \left(\widetilde{\boldsymbol{Y}} - \boldsymbol{Q}\boldsymbol{\beta}\right)^T \boldsymbol{R}^{-1} \left(\widetilde{\boldsymbol{Y}} - \boldsymbol{Q}\boldsymbol{\beta}\right), \tag{16}$$

where $\boldsymbol{Q}$ denotes the information matrix whose elements correspond to the Kriging trend evaluated at the training data so that $Q_{ij} = q_j(\boldsymbol{\xi}_i), i = 1, \ldots, N_{ed}, j = 1, \ldots, p$. The hyperparameters $\hat{\boldsymbol{\theta}}$ can be estimated through a maximum likelihood (ML) problem of the form Marrel et al. (2008)

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \left[\frac{1}{N} \left(\widetilde{\boldsymbol{Y}} - \boldsymbol{Q}\boldsymbol{\beta}\right)^T \boldsymbol{R}^{-1} \left(\widetilde{\boldsymbol{Y}} - \boldsymbol{Q}\boldsymbol{\beta}\right) (\det(\boldsymbol{R}))^{1/N_{ed}} \right]. \tag{17}$$

Alternatively, the hyperparameters can be estimated by solving a cross-validation problem Bachoc, 2013. Various approaches exist for solving the ML problem (17). An effective approach is the covariance matrix adaptation-evolution Scheme (CMAES), which belongs to the class of global optimization methods Hansen and Ostermeier (2001).

Once the hyperparameters and, subsequently, the Kriging surrogate model parameters $\boldsymbol{\beta}$ and σ^2 are estimated, the prediction of model responses, assumed to obey a Gaussian distribution, attains a mean and variance

$$\mu_{\hat{\boldsymbol{Y}}} = \boldsymbol{q}(\boldsymbol{\xi})^T \boldsymbol{\beta} + \boldsymbol{r}(\boldsymbol{\xi})^T \boldsymbol{R}^{-1} \left(\widetilde{\boldsymbol{Y}} - \boldsymbol{Q}\boldsymbol{\beta}\right) \tag{18}$$

$$\sigma^2_{\hat{\boldsymbol{Y}}} = \sigma^2_{\hat{\boldsymbol{Y}}} \left(1 - \langle \boldsymbol{q}(\boldsymbol{\xi})^T \boldsymbol{r}(\boldsymbol{\xi})^T \rangle \begin{bmatrix} \boldsymbol{0}\boldsymbol{Q}^T \\ \boldsymbol{Q}\boldsymbol{R} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{q}(\boldsymbol{\xi}) \\ \boldsymbol{r}(\boldsymbol{\xi}) \end{bmatrix} \right). \tag{19}$$

2.3. Polynomial Chaos-Kriging

Recently, Schöbi and Sudret (2014) proposed combining polynomial chaos and Kriging to improve the global approximation capability of PCEs by leveraging the local interpolation scheme of Kriging. The sparse polynomial chaos-Kriging (sPCK) surrogate model has a similar form to that of Kriging, with the notable difference being the generic polynomial trend is substituted by a PCE

$$\hat{\mathbf{Y}} = \hat{\mathcal{M}}^{sPCK}(\boldsymbol{\xi}) = \sum_{\boldsymbol{a}\in\mathcal{A}} y_{\boldsymbol{a}}\Psi_{\boldsymbol{a}}(\boldsymbol{\xi}) + \sigma^2 Z(\boldsymbol{\xi}). \tag{20}$$

The model coefficients estimation procedure is very similar to that of Kriging. There are two strategies for estimating the polynomial chaos trend coefficients as well as the Kriging kernel hyperparameters the sequential and optimal methods Schöbi and Sudret (2014). In both methods, the set of multi-indices $\mathcal{A}$ is determined by (9) through applying the LAR algorithm to increase the sparsity of the expansion. Sequential PCK commences by estimating the optimal basis set for the trend independently. Then, the PCE coefficients are treated as the $\boldsymbol{\beta}$ parameters in Kriging, which are estimated by solving the ML problem (17). On the other hand, optimal PCK solves (17) for the hyperparameters at each step of the LAR algorithm, i.e., each time new regressors are entered into the expansion. sPCK has received less attention than sPC in the literature, but examples of application of sPCK can be found in the works of Weinmeister et al. (2018) and Makrygiorgos et al. (2020).

3. Case study: Surrogate modeling of a 2D population balance model

3.1. 2D Population balance model

3.1.1. General formulation of population balances

Consider a system consisting of many individual entities that interact with one another and with the surrounding environment, usually a continuous medium. Describing the evolution of each entity individually is computationally prohibitive, whereas modeling the entities collectively as an ensemble is more convenient. Additionally, such a representation often provides a deeper insight into the main physical features and properties of the system as a whole. In population balance (PB) modeling, the constitutive entities are called *particles*, while the ensemble itself *population*. Here, we assume that not only the medium is continuous, but also the population itself, f, is mathematically described by a *continuous* function. f represents the number distribution (or density) of particles and depends on time, t, on the position in the physical space (i.e., the external coordinates **x**), and on a set of relevant properties (i.e., the internal coordinates **L**). The choice of the internal coordinates depends on what the particles represent; in crystallization problems they usually represent the crystal characteristic lengths, while in cell population modeling the age or a specific genetic trait. Each type of coordinates is defined over a sub-domain, Ω_t, $\Omega_{\mathbf{x}}$, and $\Omega_{\mathbf{L}}$, respectively. The domain Ω of f is given by the joint space obtained from the Cartesian product $\Omega = \Omega_t \times \Omega_{\mathbf{x}} \times \Omega_{\mathbf{L}}$. In this work, all coordinates are continuous variables. Note that the population density f is not a proper probability distribution because its integral over Ω is not unit, but yields the total number of particles per unit mass (or volume). However, f can be given an intuitive probabilistic interpretation upon normalization. The evolution of f can be thought of as a motion in Ω. During this motion, the number of particles may change due to sinks / sources that consume / produce new entities, respectively. If the motion is purely convective with velocity **U** in $\Omega_{\mathbf{x}}$ and **G** in $\Omega_{\mathbf{L}}$, the conservation equation, usually known as population balance equation (PBE) Ramkrishna (2000), takes the form

$$\frac{\partial f(t,\mathbf{x},\mathbf{L})}{\partial t} + \nabla_{\mathbf{x}}\cdot(\mathbf{U}(t,\mathbf{x},\mathbf{L},\mathbf{z})f(t,\mathbf{x},\mathbf{L})) + \nabla_{\mathbf{L}}\cdot(\mathbf{G}(t,\mathbf{x},\mathbf{L},\mathbf{z})f(t,\mathbf{x},\mathbf{L})) = \mathcal{S}(t,\mathbf{x},\mathbf{L},\mathbf{z}), \tag{21}$$

where **z** is the vector of the relevant system states (e.g., temperature, concentration, pH) and $\mathcal{S}$ is the function representing the sink and source terms. The solution of (21) requires appropriate initial and boundary conditions. It is apparent that it is akin to the conservation equations of mass, thermal energy, momentum, or electric charge.

3.1.2. Benchmark case: 2D-PB model of batch cooling crystallization

We now present the PB model of batch cooling crystallization of ibuprofen (i.e., solute) in ethanol (i.e., solvent). We assume that the suspension within the vessel is well-mixed at any time and that breakage and agglomeration do not occur in the system. Without loss of generality, we also assume that the process starts at $t_0 = 0$, i.e. $\Omega_t = \mathbb{R}^+$, and that at t_0 the suspension is at the thermodynamic equilibrium for the given initial conditions.

The needle-like crystals are modelled as square-based prisms, whose geometry is fully described by two characteristic lengths, $L_1 \in \mathbb{R}^+$ for the base and $L_2 \in \mathbb{R}^+$ for the height. The characteristic lengths are used as internal coordinates, i.e., $\mathbf{L} = [L_1, L_2]^T$, hence $\Omega_{\mathbf{L}} = \mathbb{R}^+ \times \mathbb{R}^+$. The crystal population constitutes a 2D distribution, called *particle size and shape distribution* (PSSD). Note that the above definition of $\Omega_{\mathbf{L}}$ assumes that the smallest possible crystal is a (virtual) particle of null size in both L_1 and L_2. For each characteristic length, the rate of change along the internal coordinates, **G**, is given by the growth rates (for supersaturated systems) or by the dissolution rates (for undersaturated systems). We assume that **G** depends on thermodynamic variables, but not on **L**. Because the crystals do not agglomerate or break, the only source term is represented by nucleation, J. This total nucleation rate is given by the additive contribution of primary and secondary nucleation, with rates J_p and J_s, respectively. Under the above assumption, the PBE (21) simplifies to

$$\frac{\partial f(t,L_1,L_2)}{\partial t} + G_1(\mathbf{z})\frac{\partial f(t,L_1,L_2)}{\partial L_1} + G_2(\mathbf{z})\frac{\partial f(t,L_1,L_2)}{\partial L_2} = (J_p(\mathbf{z}) + J_s(\mathbf{z}))\delta(L_1)\delta(L_2), \tag{22}$$

with initial and boundary conditions

$$f(0,L_1,L_2) = f_0(L_1,L_2) \tag{23}$$

$$f(t,0,L_2) = 0 \tag{24}$$

$$f(t,L_1,0) = 0. \tag{25}$$

The initial PSSD, f_0, is modeled either as a (truncated) normal distribution (seeded processes), or as an identically null distribution (unseeded processes). The PBE (22) is coupled with a material balance for the solute present in the solution

$$\frac{dc}{dt} = -\rho_c k_v \frac{d\phi_{21}}{dt}, \quad c(0) = c_0, \tag{26}$$

where c is the concentration, c_0 is the initial concentration, ρ_c is the crystal density, k_v is the volume shape factor (equal to 1 for a square-based prism), and ϕ_{21} is the cross-moment of the PSSD. The general definition of the cross-moment (ij) is

$$\phi_{ij}(t) = \int_0^{+\infty}\int_0^{+\infty} L_1^i L_2^j f(t,L_1,L_2)\,dL_1\,dL_2. \tag{27}$$

The elements of the state vector **z** in (22) are the temperature, T, the supersaturation, S, and the system volume, V. The thermodynamic driving force of the crystallization phenomena is the supersaturation, S. For mildly non-ideal solutions and for compounds whose activity coefficients are independent of the concentration,

the supersaturation is given by the ratio between the solute concentration and its solubility, c_s, at the temperature T

$$S(T) = \frac{c}{c_s(T)}, \tag{28}$$

where c_s is given by the van't Hoff expression

$$c_s(T) = k_1 \exp\left(-\frac{k_2}{T}\right). \tag{29}$$

The expression for each characteristic length $j = 1, 2$ of the growth rate ($S \geq 1$) follows the birth and spread mechanisms Lindenberg and Mazzotti (2009), while that of the dissolution rate is empirical of the form Mersmann (2001)

$$G_j(S,T) = \begin{cases} g_{j,0} \exp\left(-\frac{g_{j,1}}{T}\right)(S-1)^{1/6} \exp\left(-\frac{g_{j,2}}{T^2 \ln S}\right) & S > 1 \\ d_{j,0} \exp\left(-\frac{d_{j,1}}{T}\right)(1-S)^{d_{j,2}} & S < 1. \end{cases} \tag{30}$$

The primary nucleation is defined according to the classical nucleation theory Debenedetti (1996); Kashchiev (2000)

$$J_p(S,T) = a_0 \exp\left(-\frac{a_1}{T}\right) S \exp\left(-\frac{a_2}{T^3 \ln^2 S}\right), \tag{31}$$

which holds for $S > 1$, otherwise $J_p = 0$. Similarly, for $S > 1$, the secondary nucleation is given by the empirical law Mersmann (2001)

$$J_s(S,T) = b_0 \epsilon^{b_1} m_s^{b_2} \bar{G}^{b_3} = b m_s^{b_2} \bar{G}^{b_3}, \tag{32}$$

where $\bar{G}$ is an average growth rate Salvatori and Mazzotti (2017), and ϵ is the turbulence intensity. For $S < 1$, $J_s = 0$. The last equality on the right hand side of (32), with $b = b_0 \epsilon^{b_1}$, is obtained by assuming a constant ϵ.

There is no analytical solution for the PBE (22). Among the various numerical methods for solving PBEs Mesbah et al. (2009), high-order finite volume methods are particularly suited for high-resolution solution of population balances, are computationally efficient in the case of pure growth and nucleation phenomena, and easily generalize to n-dimensional PBEs LeVeque (2002); Gunawan et al. (2004); Qamar et al. (2009). Thus, we used high-resolution finite volume methods in this work to discretize the PSSD along the 2D internal coordinates. The numerical oscillations induced by this method are suppressed by using a Van Leer flux-limiting function LeVeque (2002). The discretization along the internal coordinates transforms the PBE into a set of ordinary differential equations along the temporal coordinate, solved with explicit time integration and subject to the CFL condition Gunawan et al. (2004). Numerical solution of the 2D-PB model (22)-(32) is relatively expensive. A typical run of the model for a simulation time of one hour takes approximately 15 to 20 s. This can lead to excessive computation times in UQ tasks that require many (i.e., on the order of 10^5 to 10^6) model evaluations.

Here, we consider 20 uncertain parameters in the 2D-PB model (22)–(32), as listed in Table 1 (i.e., $M = 20$). The uncertain parameters are assumed to follow independent uniform distributions $\boldsymbol{\xi} \sim \mathcal{U}(\boldsymbol{l_b}, \boldsymbol{u_b})$, where $\boldsymbol{l_b}$ and $\boldsymbol{u_b}$ denote the lower and upper bounds of the distribution support. The nominal values of each parameter are $\boldsymbol{\xi_{mean}} = \frac{\boldsymbol{u_b}+\boldsymbol{l_b}}{2}$. The joint probability density has an independent copula. The uncertain parameters and QoIs are scaled so that they attain values on the order of $\mathcal{O}(10^{-1} - 1)$. Let $\hat{\xi}_i$ denote the unscaled uncertainties, which can be written as $\hat{\xi}_i = o_i \xi_i$. The supports $[\boldsymbol{l_b}, \boldsymbol{u_b}]$ of the marginals are defined symmetrically for each uncertain parameter and set to $\pm$ 20% or $\pm$ 30% of the corresponding mean value.

Table 1
The list of 20 uncertain parameters in the 2D-PB model (22)–(32), which correspond to the kinetics of growth and dissolution and the physicochemical properties of crystals. The uncertain parameters are assumed to follow a uniform distribution on the support $[\boldsymbol{l_b}, \boldsymbol{u_b}]$.

ξ	p	l_b	u_b
ξ_1	α_0	1.16	1.74
ξ_2	α_1	1.2	1.8
ξ_3	α_2	2.416	3.624
ξ_4	b_0	0.8	1.2
ξ_5	b_1	1.6	2.4
ξ_6	b_2	1.2	1.8
ξ_7	$g_{1,0}$	0.07	0.13
ξ_8	$g_{1,1}$	1.04	1.56
ξ_9	$g_{1,2}$	3.04	4.56
ξ_{10}	$g_{2,0}$	4.2	7.8
ξ_{11}	$g_{2,1}$	0.96	1.44
ξ_{12}	$g_{2,2}$	4.8	7.2
ξ_{13}	$d_{1,0}$	4.8	7.2
ξ_{14}	$d_{1,1}$	1.28	1.92
ξ_{15}	$d_{1,2}$	1.6	2.4
ξ_{16}	$d_{2,0}$	4.8	7.2
ξ_{17}	$d_{2,1}$	1.2	1.8
ξ_{18}	$d_{2,2}$	1.6	2.4
ξ_{19}	ρ_c	0.91	1.69
ξ_{20}	k_v	0.5497	1.021

The QoIs consist of the solute concentration c, the cross-moment ϕ_{21}, and the average crystal lengths $\widetilde{L}_1$ and $\widetilde{L}_2$ (i.e., $N = 4$). The characteristic quantities that non-dimensionalize and scale the QoI are defined as the nominal solution values at the final time step. The nominal solution is obtained for the realization that corresponds to the mean values of the parameters. The QoIs are sampled at $n_t = 15$ discrete time steps over a fixed simulation time horizon. The simulation time is set to $t_f = 3600$ s and the solution is available on 1000 equally-spaced time steps through interpolation from the solution with a time-adaptive integrator. The n_t sampling times are

$$t = \{396, 648, 936, 1080, 1368, 1440, 1872, 2160, 2340, 2556, 2700, 2880, 3060, 3276, 3600\}\ \text{s}.$$

For each QoI at each sampling time, we construct a separate surrogate model, so the dynamics are captured through $n_t \times N$ mappings of the form $\hat{\mathcal{M}} : \mathbb{R}^M \to \mathbb{R}$.

3.2. Performance of surrogate models

In this section, we compare the performance of the different surrogate modeling methods presented in Section 2. The specific choices for constructing the surrogate models are as follows. To build the polynomial chaos (PC) surrogate model, we choose an adaptive order scheme with varying orders from 1 to 3, where the coefficients are estimated using ordinary least squares (OLS). Note that using fourth order polynomials in the full PCE would increase the cardinality of $\mathcal{A}$ by an order of magnitude, leading to out-of-memory computations. A similar basis-adaptive approach is taken for building the Kriging (KG) surrogate models. The order of the trend p in (11) varies from 1 to 2, the surrogate models of both orders are computed and the surrogate model with the least leave-one-out error is stored. Usual choices for the shape factor of the Matern kernel function are $\nu = 3/2$ or $\nu = 5/2$; we choose the latter. Sparse polynomial chaos (sPC) and sparse PC-Kriging (sPCK) offer greater flexibility in selecting the polynomial basis order, which is varied from 1 to 15. In fact, due to sparsity, a plethora of terms are dropped. For the sPCK model, we employ the sequential PCK procedure Schöbi and Sudret (2014). Moreover, in both sPCK and sPC, the truncation norm q is varied from 0.7 to 0.8. Note that in the PC-based surrogate models, the univari-

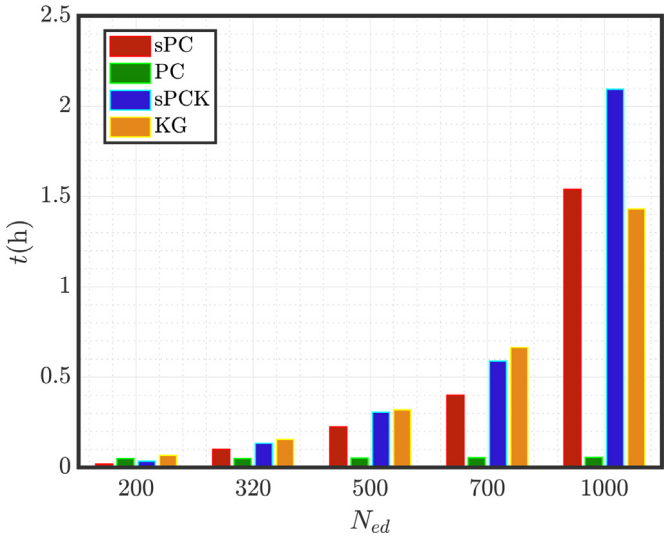

Fig. 2. Time required to train the sPC, PC, sPCK, and KG surrogate models in relation to the experimental design size N_{ed}. The total time t corresponds to the time needed to construct the surrogate models for the four quantities of interest at the desired times (i.e., $n_t \times N$).

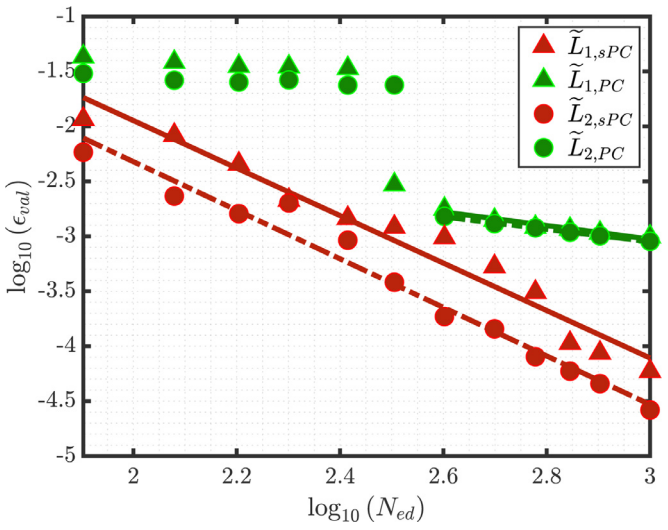

Fig. 3. Validation errors for the polynomial chaos surrogate models for $\widetilde{L}_1$ and $\widetilde{L}_2$ at sampling time $t = 2556$ s. Symbols correspond to the errors computed for each ED size while the lines represent the leading order fit. Triangles/solid lines correspond to $\widetilde{L}_1$, while circles/dashed lines correspond to $\widetilde{L}_2$. The estimations of sPC are denoted with red, while PC with green. A leading order fit of the error is superimposed for the validation errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ate basis functions are chosen as Legendre polynomials, which are the optimal choice for the uniform distributions of the uncertain parameters. The surrogate model construction is performed using `UQLab` in `MATLAB` Marelli and Sudret (2014). For the error convergence study and performance assessment of each surrogate modeling method, 12 experimental designs are used with sizes $N_{ed} = \{80, 120, 160, 200, 260, 320, 400, 500, 600, 700, 800, 1000\}$. The experimental designs are constructed using the Latin Hypercube Sampling (LHS) method by defining an initial set of points and sequentially enriching it until the maximum design size is reached.

First, we compare the computation time required to train each surrogate model. Fig. 2 shows the total time to construct the PC, sPC, sPCK, and KG surrogate models for all QoIs at the desired times (i.e., $n_t \times N$ surrogate models in total for each method) as a function of ED size. We compare all methods by allocating approximately the same time to construct the surrogate models. The PC expansions are the least time consuming to build. This can be attributed to the fact that the coefficient estimation via OLS has an analytical solution, where the crucial step is the inversion of the experimental design matrix. Generally, it is more expensive to train sPCK than sPC surrogate models: sPCK requires both the LAR basis adaption step and the hyperparameter optimization, while sPC relies only on the former. Kriging requires approximately the same amount of time as sPC, given that linear and quadratic trends are used in (12). We observed that adding a cubic trend would dramatically increase the time required for refining the Kriging hyperparamters, but did not improve significantly the accuracy of the surrogate model.

Next, we discuss the error convergence analysis. For brevity, only the QoIs $\widetilde{L}_1$ and $\widetilde{L}_2$ (i.e., the average crystal lengths in 2D) at time 2556 s are considered for this analysis. We begin by inspecting the validation errors computed by (3) as a function of the ED size. The validation set consists of $N_{val} = 20000$ points. To standardize the error analysis, let us assume that the error scales as $\epsilon_{val} = cN_{ed}^n$, where n is the leading scaling order LeVeque (2007). In the logarithmic scale, the scaling equation takes the form

$$log_{10}(\epsilon_{val}) = log_{10}(c) + nlog_{10}(N_{ed}), \tag{33}$$

which shows a linear relationship between the logarithm of the error and that of the ED size.

Fig. 3 shows the validation errors for the PC and sPC surrogate models as a function of the ED size for the two QoIs in the logarithmic scale. As expected, Fig. 3 indicates that the PC surrogate models perform much worse than sPC, with validation errors that are approximately two orders of magnitude higher. This suggests that the model responses cannot be captured reliably with cubic polynomials in the PC models. On the other hand, sPC have a larger representational capability due to sparsity effects, as higher order terms appear in the expansions. In addition, the cardinality of the basis becomes excessively large in the PC models, even for quadratic or cubic expansions, so that the available training data are less than the number of the coefficients to be estimated. LAR alleviates this issue in sPC by entering at most N_{ed} terms in the expansion Blatman, Sudret (2013).

On the other hand, Fig. 4 suggests that, as the ED size increases, the log10-validation error of sPCK and KG keeps decreasing and for large ED it is well below $\mathcal{O}(-3)$. Nevertheless, using orthogonal polynomials as the trend improves the predictions by approximately an order of magnitude for both QoIs ($\widetilde{L}_1$ and $\widetilde{L}_2$). This is due to not only the optimal basis representation in the trend, but also the fact that the polynomial trend in the sPCK surrogate model (20) allows for a higher order approximation.

Overall, sPC and sPCK surrogate models exhibit comparable performance in terms of their predictive capability. This implies that there do not exist significant local variations in the 2D-PB model under consideration, which a high-order polynomial approximation cannot capture. Generally, it is expected that the use of a larger training set would result in more accurate surrogate models. However, the error convergence results (Figs. 3 and 4) reveal that for some experimental designs, especially for small sized EDs, there is an irregular non-monotonic behavior of the validation error as a function of the ED size, especially for the Kriging model. This non-monotonic behavior can be due to the specific training procedure in each case. In fact, the validation errors are implicit functions of the quality of the ED points, the quality of the validation set, and the hyperparameter optimization of the surrogate models. A similar non-monotonic behavior in error convergence was also observed in Yamazaki (2015); Harenberg et al. (2019).

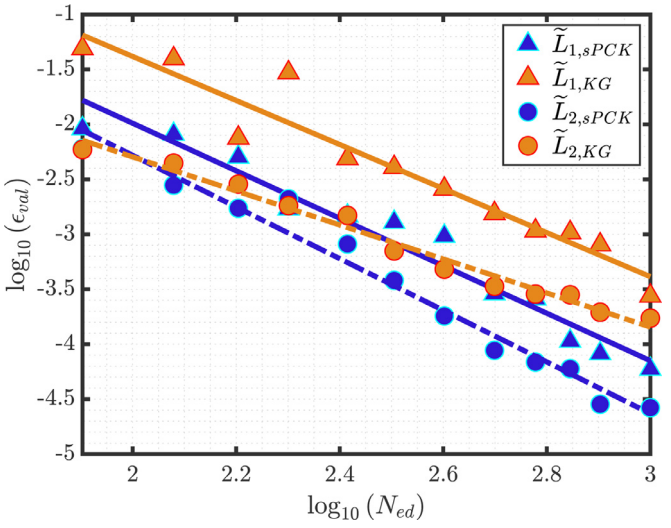

Fig. 4. Validation errors for the Kriging-based surrogate models for $\widetilde{L}_1$ and $\widetilde{L}_2$ at sampling time $t = 2556$ s. Symbols correspond to the errors computed for each ED size, while the lines represent the leading order fit. Triangles/solid lines correspond to $\widetilde{L}_1$, while circles/dashed lines correspond to $\widetilde{L}_2$. The estimations of KG are denoted with orange, while sPCK with blue. A leading order fit is superimposed for the validation errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

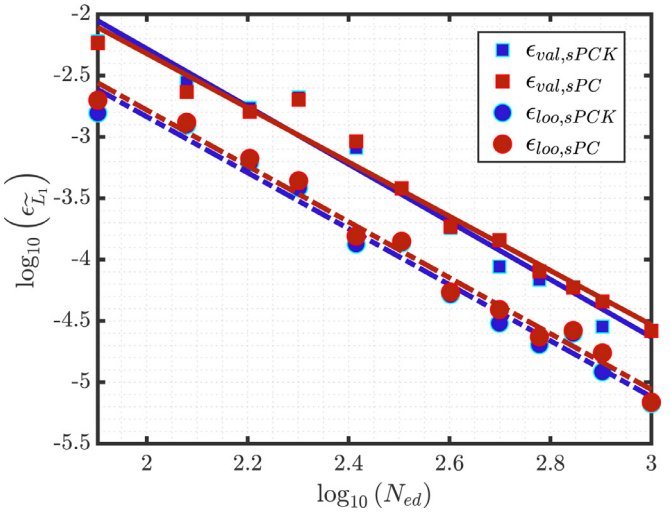

Fig. 5. Validation and LOO error convergence plot for $\widetilde{L}_1$ at sampling time $t = 2556$ s. PCK errors are denoted by blue color and sPCE errors by red color. Square symbols correspond to the validation error and circles to the LOO error. A leading order fit of the error is superimposed for both validation and LOO errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Furthermore, we perform an estimation of the leading scale order of the validation errors as a function of the ED size. The computed leading orders are the average of the predictions of the two QoIs. It is evident from Fig. 3 that the PC models exhibit discontinuity in the error, so that the convergence rate cannot be estimated by a single line over all considered ED sizes. Nevertheless, PC models exhibit a convergence rate of approximately $\mathcal{O}(-0.6)$ for larger EDs, which is the lowest among all the surrogate models. The KG surrogate model shows the second slowest convergence with a rate of $\mathcal{O}(-1.7)$. The leading orders for the sPC and sPCK are $\mathcal{O}(-2.18)$ and $\mathcal{O}(-2.25)$, respectively. This suggests that these surrogate models show a nearly quadratic convergence rate in relation to the ED size. The previous error analysis was based on the validation error, assuming that there is enough data in order to construct a validation set. The ϵ_{loo} error, which is available directly from the training data, is another potential candidate for estimating the predictive capabilities of the various surrogate models. Fig. 5 shows the cross-validation, leave-one-out error ϵ_{loo}, and the validation error ϵ_{val} for the QoI $\widetilde{L}_1$ for the sPC and sPCK surrogate models. As can be seen, the LOO error follows the same trend as the validation error, so that in the absence of a validation set it can potentially be used to estimate the accuracy of the surrogate models. However, the LOO errors tend to be smaller than the validation errors.

The performance of surrogate models is also assessed by examining the parity plots, which can provide insights into the error sources, i.e., identifying where inconsistencies between the original model $\mathcal{M}$ and the surrogate model $\hat{\mathcal{M}}$ lie. For brevity, we only present the parity plots for one QoI, that is, the average length L_1 at $t = 2556$ s. Fig. 6 shows the parity plots for the KG and PC surrogate models. As expected from the error convergence results, when the size of the ED increases, the surrogate model predictions will become more accurate. In addition, for surrogate models with worse predictive capability, the surrogate model predictions deviate from the $y = x$ line. Given a limited amount of training data, this can be attributed to the fact that the adopted training set may be richer in samples that are closer to the nominal values of the uncertain inputs and, thus, lack samples for model responses that are less likely to occur.

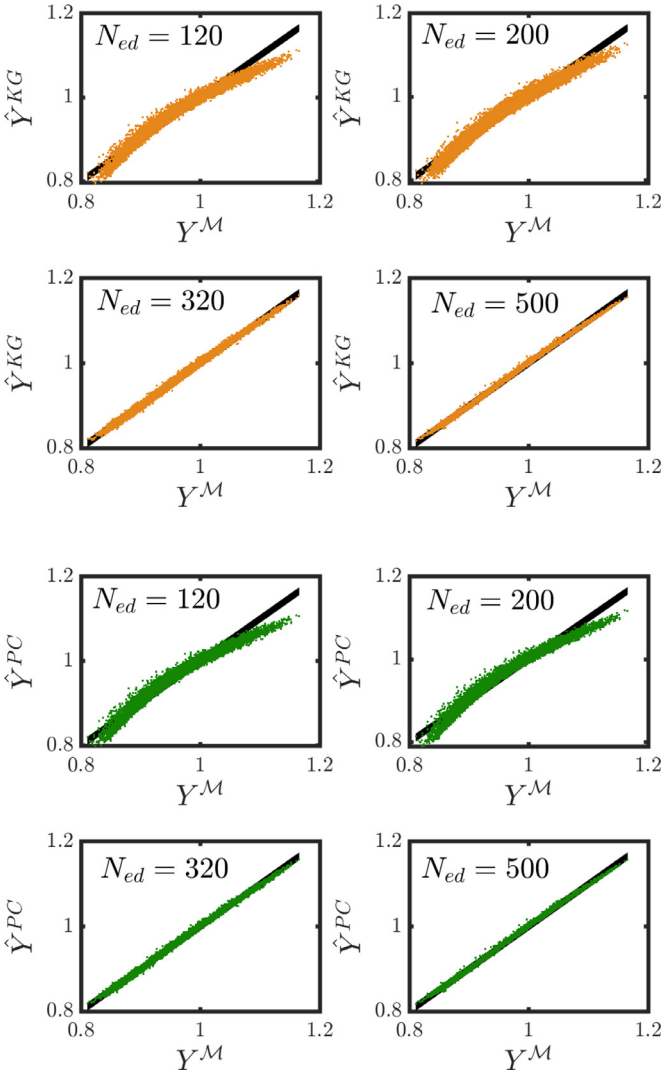

Fig. 6. Parity plots for the KG (orange symbols) and PC (green symbols) surrogate models for different experimental design sizes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

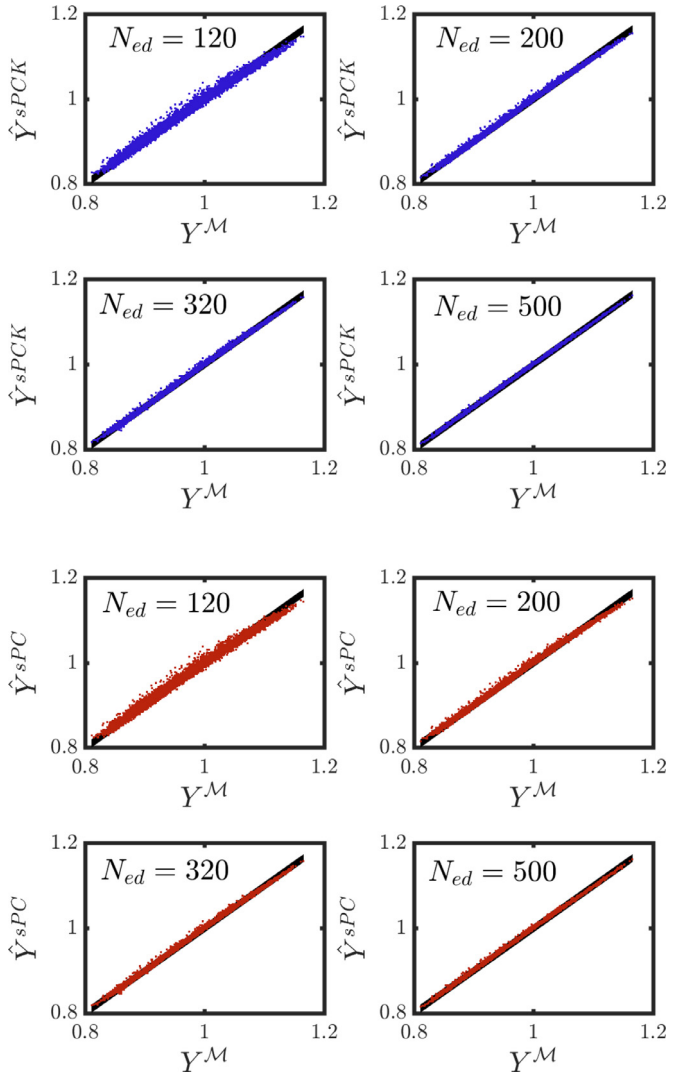

Fig. 7. Parity plots for the sPCK (blue symbols) and sPC (red symbols) surrogate models for different experimental design sizes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 7 shows the parity plots for the sPC and sPCK surrogate models. Both surrogate models provide accurate predictions even for the smallest ED, which is expected from their low validation. These results highlight the usefulness of leveraging sparsity in dealing with the relatively high uncertainty dimension of the 2D-PB model. Generally, sPCK and sPC surrogate models are expected to have comparable performance, as was also reported in Schöbi and Sudret (2014). Whether the auto-correlation function improves the sPC accuracy is a matter of the complexity of the uncertainty propagation due to the model at hand, which is generally not known *a priori*. Moreover, sPCK provides the confidence intervals of the surrogate model predictions that can be a useful feature in many UQ applications.

4. Surrogate models for uncertainty quantification

In this section, the application of the sparse polynomial chaos and polynomial chaos-Kriging surrogate models is demonstrated for fast solution of various forward and inverse uncertainty quantification problems for the 2D-PB model. This includes global sensitivity analysis as well as Bayesian and maximum a posteriori parameter estimation.

4.1. Global sensitivity analysis

Surrogate models are invaluable tools for UQ problems that rely on Monte Carlo (MC) sampling where a significant number of model evaluations is often necessary. One such application is global sensitivity analysis (GSA). In contrast to local sensitivity analysis that typically involves computing gradients of QoIs around a nominal value, GSA accounts for the entire domain of uncertain inputs to estimate the sensitivity of QoIs.

First, we introduce the Sobol' indices Sobol (1993); Sudret (2008). For simplicity of notation, we assume uniform uncertain inputs with support $\mathcal{D}_{\boldsymbol{\Xi}} = [0,1]^M$. The model response is assumed to take the following form, which is known as the analysis of variance (ANOVA) representation

$$\mathcal{M}(\boldsymbol{\xi}) = \mathcal{M}_0 + \sum_{i=1}^{M} \mathcal{M}_i(\xi_i) + \sum_{1\le i\le j\le M}^{M} \mathcal{M}_{ij}(\xi_{ij}) + \ldots + \mathcal{M}_{1,2,\ldots,M}(\boldsymbol{\xi}), \tag{34}$$

where $\mathcal{M}_0$ is the expected value of the response $\mathcal{M}(\boldsymbol{\xi})$ over the input space, while the integrals of the summands over their variables obey

$$\int_0^1 \mathcal{M}_{i_1,\ldots,i_s}(\xi_{i_1},\ldots,\xi_{i_s})\, d\xi_{ik} = 0, \quad 1 \le k \le s. \tag{35}$$

The ordered set $1 \le i_1 < \ldots < i_s \le M, s = 1,\ldots,M$ represents a subset of the uncertain inputs. For functions integrable over $\mathcal{D}_{\boldsymbol{\Xi}}$, whose input variables are independent, the expansion (34) can be shown to be unique Sobol (1993). The constant and summation terms in (34) are computed recursively as

$$\mathcal{M}_0 = \int_{\mathcal{D}_{\boldsymbol{\Xi}}} \mathcal{M}(\boldsymbol{\xi})\, d\boldsymbol{\xi} \tag{36}$$

$$\mathcal{M}_i(\xi_i) = \int_0^1 \cdots \int_0^1 \mathcal{M}(\boldsymbol{\xi})\, d\boldsymbol{\xi}_{\sim i} - \mathcal{M}_0 \tag{37}$$

$$\mathcal{M}_{ij}(\xi_i,\xi_j) = \int_0^1 \cdots \int_0^1 \mathcal{M}(\boldsymbol{\xi})\, d\boldsymbol{\xi}_{\sim(ij)} - \mathcal{M}_0 - \mathcal{M}_i(\xi_i) - \mathcal{M}_j(\xi_j), \tag{38}$$

where $\boldsymbol{\xi}_{\sim i} = \{\xi_1, \xi_2, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_M\}$, i.e., it represents the set where the i^{th} variable is excluded. Accordingly, the total variance of $\mathcal{M}(\boldsymbol{\xi})$, denoted by D, is given by

$$D = \int_{\mathcal{D}_{\boldsymbol{\Xi}}} \mathcal{M}^2(\boldsymbol{\xi})\, d\boldsymbol{\xi} - \mathcal{M}_0^2. \tag{39}$$

We also define the partial variance that represents the variance of the response of $\mathcal{M}$ when a subset of the uncertainties is perturbed. The partial variance $D_{i_1,\ldots,i_s}$ is computed as

$$D_{i_1,\ldots,i_s} = \int_0^1 \cdots \int_0^1 \mathcal{M}^2_{i_1,\ldots,i_s}(\xi_{i_1},\ldots,\xi_{i_s})\, d\xi_{i_1} \ldots d\xi_{i_s},$$
$$1 \le i_1 \le \ldots \le i_s \le M, \quad s = 1,\ldots,M. \tag{40}$$

Using the definitions of partial and total sensitivities, the sensitivity measures can now be defined as

$$S_{i_1,\ldots,i_s} = \frac{D_{i_1,\ldots,i_s}}{D}. \tag{41}$$

The sensitivity measures represent the contribution of a subset of uncertain parameters $\{\xi_{i_1},\ldots,\xi_{i_s}\}$ to the total variance of a QoI. Here, we focus on the total Sobol' indices. The total Sobol' index of an uncertain input ξ_i is denoted by S_i^T and defined as

$$S_i^T = \sum_{\{i_1,\ldots,i_s\}\supset i} S_{i_1,\ldots,i_s}. \tag{42}$$

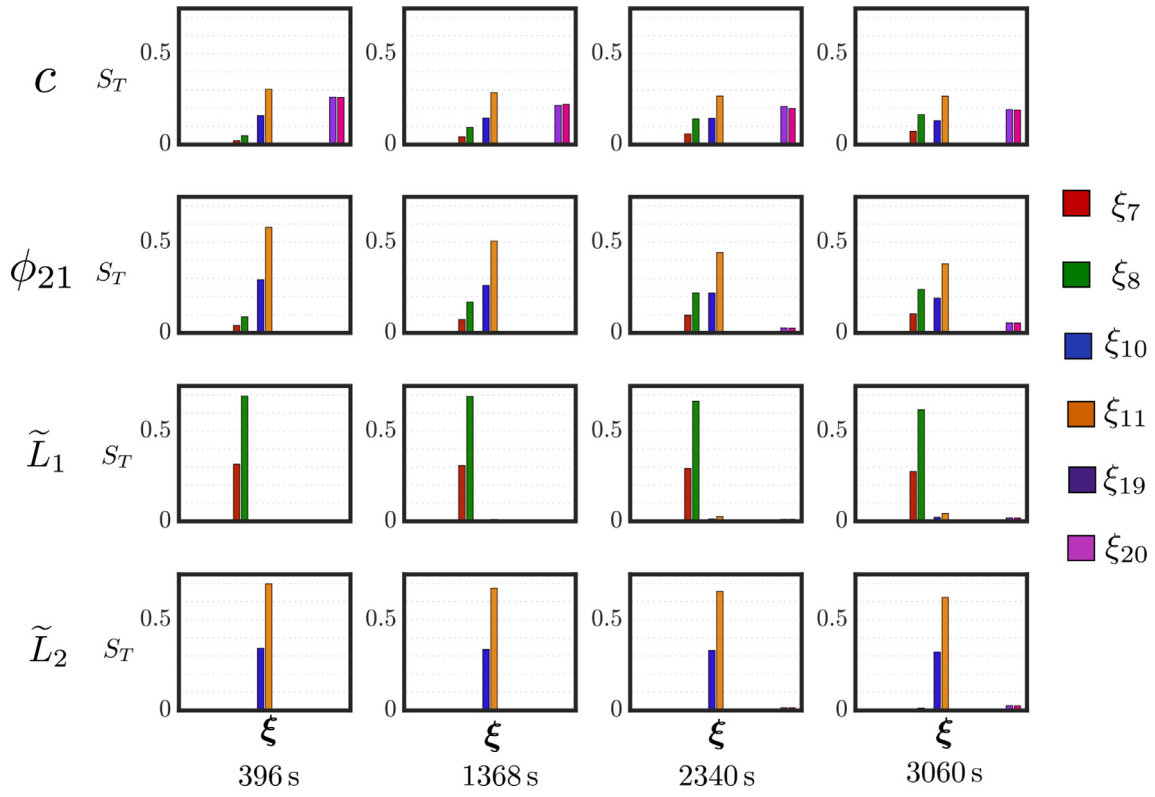

Fig. 8. Dynamic global sensitivity analysis of the four quantities of interest at four different sampling times. Each row represents a quantity of interest and each column a different sampling time. The sensitivity indices of the uncertain parameters are indicated with different colors. The index for ξ_7 is denoted by red, ξ_8 by green, ξ_{10} by blue, ξ_{11} by orange, ξ_{19} by purple, and ξ_{20} by magenta. Starting from top, the rows represent the concentration c, the cross moment ϕ_{21}, the average length on coordinate 1 (L_1), and the average length on coordinate 2 (L_2). The y-axis values represent the total Sobol' index for each uncertain parameter. The Sobol' indices for the other 14 uncertain parameters are practically zero, so they are not visible in any subplot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Thus, the estimation of the Sobol index S_i^T requires the computation of all partial sensitivites for the subsets that involve the input ξ_i. Given that the Sobol' indices of all parameters sum to 1, the Sobol index can be obtained from

$$S_i^T = 1 - S_{\sim i} = 1 - S_{\boldsymbol{v}}, \quad \boldsymbol{v} = \{1, \ldots, i-1, i+1, \ldots, M\}. \tag{43}$$

In general, there is no closed-form solution for the multivariate integrals in (39) and (40). Instead, these expressions are usually approximated by MC simulations with a large number of samples. In this work, we use (43) for the computation of the total Sobol' indices, where the $S_{\boldsymbol{v}}$ term is approximated using the Janon estimator Janon et al. (2014), which is also a sample-based method. Here, we use the sPCK surrogate model trained with 1000 samples for fast sampling of the system responses for all QoIs at the different sampling times. The dynamic Sobol' indices for the 20 uncertain parameters in the 2D-PB model are shown in Fig. 8. As can be seen, only 6 out of 20 parameters have a significant impact on the QoIs. The most influential parameters include the two kinetic parameters associated with the growth along L_1, which are $\xi_7 = g_{1,0}$ and $\xi_8 = g_{1,1}$, two kinetic parameters associated with the growth along L_2, which are $\xi_{10} = g_{2,0}$ and $\xi_{11} = g_{2,1}$, the crystal density, $\xi_{19} = \rho_c$, and the shape factor, $\xi_{20} = k_v$. At each time instant, the different QoIs have a unique sensitivity to the uncertain parameters. In addition, the relative importance of the parameters for a given QoI changes with time. For example, the solute concentration c is sensitive primarily to the parameters of growth along L_1 at the early stages of crystallization, but becomes increasingly sensitive to the growth parameters along L_2 at later times. Moreover, c is significantly sensitive to variations in the shape factor, k_v, and crystal density, ρ_c. The cross-moment ϕ_{21} exhibits a similar trend as the solute concentration; however, it is not sensitive to the shape factor and the crystal density until later stages of crystallization. Lastly, the parameters pertaining to the dissolution have negligible Sobol' indices. This is expected since the crystallization occurs in the domain where $S > 1$, thus no dissolution occurs. As for the average crystal lengths, $\widetilde{L}_1$ and $\widetilde{L}_2$ are affected by different mechanisms. As expected, the former is more sensitive to perturbations of parameters related to growth along L_1, while $\widetilde{L}_2$ to perturbations of parameters related to growth along L_2. We note that the outcomes of the GSA are dependent on the process input variables. For example, changing the cooling rate of crystallization could affect the impact of the uncertain parameters on the QoIs.

4.2. Bayesian estimation of model parameters

we focus on the inverse UQ problem that aims at estimating the parameters that affect the QoIs. Bayesian parameter inference attempts to narrow down the uncertainty in these parameters, under some assumption on both prior knowledge about them and given available experimental measurements. We aim to estimate a subset of the uncertain inputs, i.e., the "important" parameters based on the GSA, which define a new vector denoted by $\boldsymbol{\xi}^n = \{\xi_7, \xi_8, \xi_{10}, \xi_{11}, \xi_{19}, \xi_{20}\}$.

A new sPCK surrogate model that has $\boldsymbol{\xi}^n$ as the stochastic input is trained using an ED of 250 samples. A validation error below $\epsilon_{target} = 5 \times 10^{-3}$ is reached for all QoIs, so no further enrichment of the ED is needed. For this demonstration, due to the fact that we do not have actual experimental data, we generate data for parameter estimation using a "true" computational model whose responses, $\boldsymbol{\xi}^n_{true}$, are corrupted by state-dependent Gaussian mea-

Algorithm 1 The algorithm for sequential Monte Carlo Liu and Chen, 1998; Chopin (2002).

0. **Initialization**: SMC algorithm is initialized by selecting N_p particles. N. Set $j = 1$ and use the prior $f_{\Xi}(\boldsymbol{\xi})$ to generate samples $\{\boldsymbol{\xi}\}_{i=1}^{N_p}$ and set uniform weights $w_i = \frac{1}{N_p}$
1. **Reweighting**: Based on the likelihood for each realization from the N_p particles, update the weights as $w_i \leftarrow w_i \cdot w_j(\boldsymbol{\xi}_i)$, where $w_j(\boldsymbol{\xi}_i) \propto f_{\boldsymbol{D}_j|\Xi}(\boldsymbol{d}_j|\boldsymbol{\xi}_i)$
2. **Resampling**: To retain a threshold of effective particles, resample $\{\boldsymbol{\xi}_i, w_i\}_{i=1}^{N_p}$ to obtain particles with equal weights $\{\boldsymbol{\xi}_i^r, \frac{1}{N_p}\}_{i=1}^{N_p}$
3. **Iteration** : If $j < j_f$ continue to the next sampling time by setting $\{\boldsymbol{\xi}_i, w_i\}_{i=1}^{N_p} \leftarrow \{\boldsymbol{\xi}_i^r, \frac{1}{N_p}\}_{i=1}^{N_p}$, set $j \leftarrow j+1$ and return to step 1.

surement noise. We denote as d_i^w the measurement of quantity w at time step t_i. The available data for the analysis are given,

$$\begin{aligned} d_i^c &= c\left(t_i, \boldsymbol{\xi}_{true}^n\right) + E_i^c, \quad i \in \{1, \dots, n_t\} \\ d_i^{\phi_{21}} &= \phi_{21}\left(t_i, \boldsymbol{\xi}_{true}^n\right) + E_i^{\phi_{21}}, \quad i \in \{1, \dots, n_t\} \\ d_i^{\widetilde{L}_1} &= \widetilde{L}_1\left(t_i, \boldsymbol{\xi}_{true}^n\right) + E_i^{\widetilde{L}_1}, \quad i \in \{1, \dots, n_t\} \\ d_i^{\widetilde{L}_2} &= \widetilde{L}_2\left(t_i, \boldsymbol{\xi}_{true}^n\right) + E_i^{\widetilde{L}_2}, \quad i \in \{1, \dots, n_t\}. \end{aligned} \tag{44}$$

Thus, at every sampling point, the available data is denoted by $\boldsymbol{D}_i = \{d_i^c, d_i^{\phi_{21}}, d_i^{\widetilde{L}_1}, d_i^{\widetilde{L}_2}\}$ and the corresponding error vector as $\boldsymbol{E}_i = \{E_i^c, E_i^{\phi_{21}}, E_i^{\widetilde{L}_1}, E_i^{\widetilde{L}_2}\}$. We further define $\boldsymbol{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, \dots, \boldsymbol{D}_{15}\}$ as the matrix containing all the measurements and $\boldsymbol{E} = \{\boldsymbol{E}_1, \boldsymbol{E}_2, \dots, \boldsymbol{E}_{15}\}$ as the errors matrix. We model the measurement error as a Gaussian variable with zero mean and state-dependent standard deviation, which is 5% of the measured value; $E_i^w \sim \mathcal{N}(0, \sigma_{w,i}^2)$ with $\sigma_{w,i}(\Xi) = 0.05|w(t_i, \Xi)|$. At the core of Bayesian inference problem lies Bayes theorem that describes how the posterior density function is estimated given the experimental data. Bayes theorem is formulated as Kennedy and O'hagan (2001)

$$f_{\Xi|\boldsymbol{D}}(\boldsymbol{\xi}|\boldsymbol{d}) = \frac{f_{\boldsymbol{D}|\Xi}(\boldsymbol{d}|\boldsymbol{\xi}) f_{\Xi}(\boldsymbol{\xi})}{f_{\boldsymbol{D}}(\boldsymbol{d})}, \tag{45}$$

where $f_{\Xi|\boldsymbol{D}}(\boldsymbol{\xi}|\boldsymbol{d})$ denotes the posterior density of the uncertain parameters after observing the experimental data; The prior density is denoted by $f_{\Xi}(\boldsymbol{\xi})$; the term $f_{\boldsymbol{D}|\Xi}(\boldsymbol{d}|\boldsymbol{\xi})$ is the likelihood function; and $f_{\boldsymbol{D}}(\boldsymbol{d})$ is the so-called evidence or marginal likelihood. The priors of the uncertain imputs for this problem follow the same uniform distributions as those presented in Section 3 The likelihood function is represented by a Gaussian distribution Raue et al. (2013)

$$f_{\boldsymbol{D}|\Xi}(\boldsymbol{d}|\boldsymbol{\xi}) = \prod_{i=1}^{n_t} \prod_{w=1}^{N} \frac{1}{\sqrt{2\pi\sigma_{w,i}^2(\boldsymbol{\xi})}} exp\left(-\frac{\left(d_i^w - w(t_i;\boldsymbol{\xi})\right)^2}{2\sigma_{w,i}^2(\boldsymbol{\xi})}\right). \tag{46}$$

The computation of the posterior can be expensive, as closed-form solutions exist only in few specific cases. Therefore, we resort to sequental Monte Carlo (SMC), a sample-based method that seeks to generate samples from the unknown posterior function Liu and Chen, 1998; Chopin (2002) In dynamic estimation problems. the SMC method works in an iterative fashion, as summarized in Algorithm 1. We approximate the distributions using N_p particles. At each iteration, the likelihood function has to be sampled for each particle in the simulation to obtain the updated weight factors. Thus, the computational model $\mathcal{M}$ has to be evaluated N_p times, which adds a vast computational challenge to the problem. In fact, the first-principles model $\mathcal{M}$ needs on average 15 s for a complete time integration For $N_p = 10^6$, the model evaluation would require about 173 days or roughly 2.5×10^5 min. However, using the sPCK surrogate model, this computation is vastly accelerated. We sample the likelihood function at each j_t iteration. The N_p computations at each iteration, including the rest of the SMC procedure, result in an approximately total time of 10 min for the posterior density to be estimated. The time needed to generate training data for the 250 samples is approximately an hour. The training of the sPCK surrogate model using 250 training data points lasts approximately 65 min. This translates to appriximately 30,000-fold savings in computation time.

The subplots in the diagonal of Fig. 9 depict the estimates of the posterior densities (blue), where the prior densities (green, flat) are superimposed. Within the same plots, the red lines represent the true value of the parameters while the black lines the mean of the posteriors. As can be seen, the posterior densities are, in general, narrower than the priors. Thus, taking the measurements into account greatly reduces the uncertainty of the prior beliefs. However, this Bayesian inference cannot provide the same level of accuracy for estimation of all uncertain inputs. For instance, the support of parameters ξ_7, ξ_8,ξ_{10}, ξ_{11} is less than half of the corresponding support of the priors, while this is not true for ξ_{19} and ξ_{20}. The most identifiable parameter is ξ_8 and the least identifiable one is ξ_{20} This observation is consistent with the Sobol' indices presented earlier. Overall, ξ_7, ξ_8,ξ_{10}, ξ_{11} attain high Sobol' index values across the measured QoIs. while ξ_{19}, ξ_{20} mainly affect only one QoI, the concentration c. This suggests the QoI distributions are mainly affected by the parameters that have the greatest influence on them, reflected by the Sobol' indices. Hence, it is expected that it is those parameters that are more identifiable.

In addition, obtaining the posterior density for the uncertain parameters allows for a full characterization of the predicted distributions. Fig. 10 shows the approximate distribution of the QoIs marginalized over the parameters uncertainty. The prior distributions $f_{\boldsymbol{Y}}$ (green) represent the probability density functions (pdfs) of the model QoIs marginalized over the prior parameter distributions. The posterior distributions (blue), $f_{\boldsymbol{Y}|\boldsymbol{D}}$. are updates of the prior predictive distributions based on the available data. The posterior distributions are obtained by marginalization over the posterior parameters distributions. Similar to the prior and posterior behavior for the uncertain parameters, the posterior distribution of QoIs are much narrower their prior distributions. predictive. In fact, when the true QoIs and the mean of the posterior distributions are superimposed on the plot, they may become indistinguishable.

4.3. Maximum a posteriori parameter estimation

For this application, we focus on the solution of an optimization problem where surrogate models are employed for accelerating the objective function evaluations. The problem involves estimation of the 20 uncertain parameters via estimating the mean of the posterior distribution as the solution of a maximum *a posteriori* (MAP) problem. The MAP problem that takes the form

$$\hat{\boldsymbol{\xi}}_{MAP}(\boldsymbol{d}) = \underset{\boldsymbol{\xi} \in \mathcal{D}_{\Xi}}{\operatorname{argmax}} f_{\Xi|\boldsymbol{D}}(\boldsymbol{\xi}|\boldsymbol{d}) = \underset{\boldsymbol{\xi} \in \mathcal{D}_{\Xi}}{\operatorname{argmax}} f_{\boldsymbol{D}|\Xi}(\boldsymbol{d}|\boldsymbol{\xi}) f_{\Xi}(\boldsymbol{\xi}). \tag{47}$$

This yields the MAP point estimates of the parameters. We re-visit the original uncertainty inputs space with the 20 parameters and we assume that the uncertainties follow a Gaussian distribution so that $\boldsymbol{\xi^G} \sim \mathcal{N}(\boldsymbol{\mu_G}, \boldsymbol{\sigma_G^2})$. The mean of the Gaussian distribution is the same as in the case of uniform priors, i.e., the midpoint of the each respective range, and the standard deviation $\sigma_{G,i}$ is assumed to be 15% of the mean for all parameters except ξ_7^G, ξ_{10}^G, ξ_{19}^G, ξ_{20}^G for which $\sigma_{G,i}$ is 25% of the corresponding mean. The covariance matrix for the uncertain inputs is considered to be diagonal and is denoted by $\boldsymbol{C_\xi}$. The prior in (47) is utilized as a regularization term that adds stability to the solution. The likelihood is assumed to be

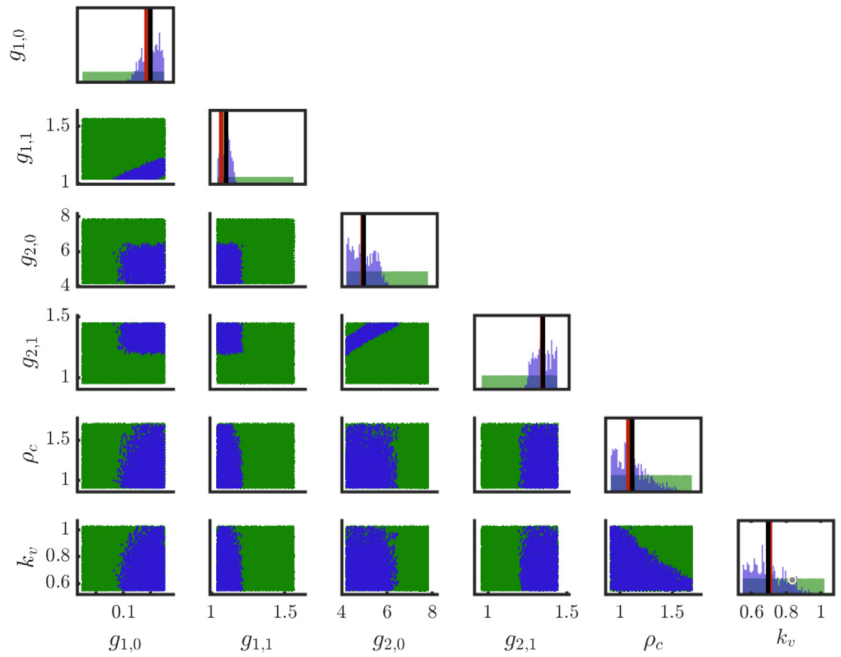

Fig. 9. Prior and posterior distribution of the uncertain parameters. The posteriors are denoted with blue, while the priors with green. The diagonal elements represent the prior (flat) and posterior marginal densities. The red lines represent the true value of the parameters, while the black lines the mean of the posterior. The off-diagonal elements represent the 2D projections of the joint probability density. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

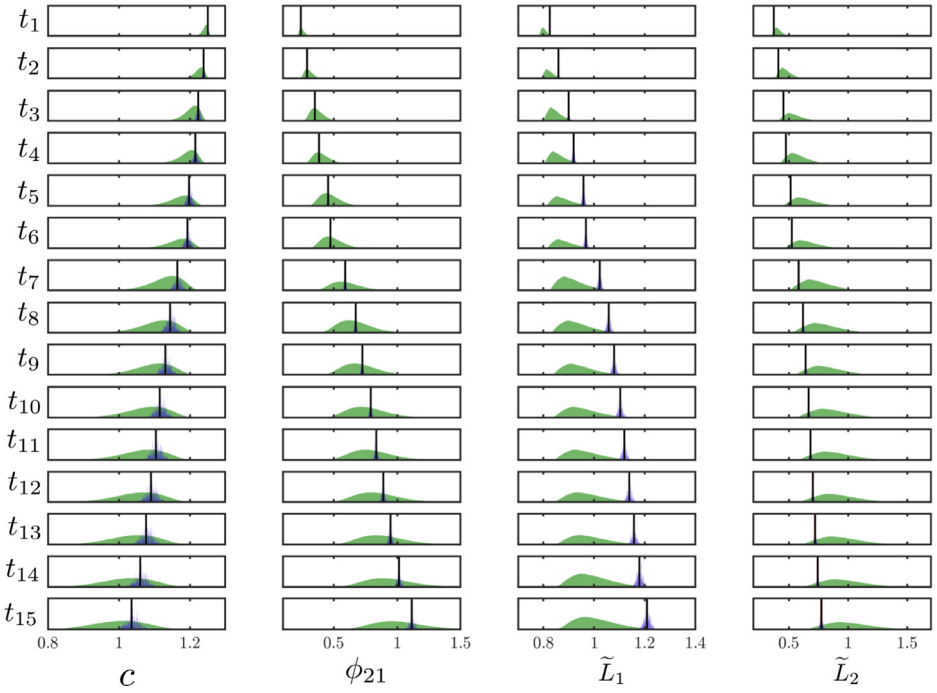

Fig. 10. Time variation of prior and posterior distribution of QoIs i.e., the model responses marginalized over the prior and posterior parameter distributions, respectively. Each column represents a different QoI and each row a different sampling time step, with descending order. The prior distributions are denoted by green and the posterior by blue. The black line represents the mean of the posterior distribution (point estimate). Some marginals posteriors are so narrow that are practically indistinguishable from the point estimates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Gaussian with a time-averaged state dependent variance, so that each diagonal element in the covariance matrix for the responses is $C_{\boldsymbol{Y}}(w,w) = 0.05|\frac{\sum_{i=1}^{n_t} w_i}{n_t}|$, where $w = 1,\ldots,N$ and corresponds to the QoIs.

Under the Gaussian assumptions, the MAP optimization problem is converted to a regularized weighted least-squares problem, which is defined as

$$\widetilde{\boldsymbol{\xi}}_{MAP}(\boldsymbol{d}) = \underset{\boldsymbol{\xi}\in\mathcal{D}_{\Xi}}{\operatorname{argmin}} - \log\left(f_{\boldsymbol{D}|\Xi}\left(\boldsymbol{d}|\boldsymbol{\xi}\right)\right) - \log\left(f_{\Xi}\left(\boldsymbol{\xi}\right)\right). \tag{48}$$

The log-likelihood function in (48) includes the inverse of both $\boldsymbol{C}_{\boldsymbol{\xi}}$ and $\boldsymbol{C}_{\boldsymbol{Y}}$. If the diagonal of those matrices contains extremely small values, then they are ill-conditioned and their inversion may

be inaccurate. This can happen if the unscaled parameters and model outputs are used. Since the priors have been considered to be Gaussian here, we re-construct the surrogate models. As sPCK and sPC surrogate models perform similarly (as shown by the error convergence plots Figs. 3 and 4), we employ the latter for this problem. The main motivation behind this choice is that the QoI evaluations with sPC is faster than with sPCK, which involves the inversion of the correlation matrix at each evaluation. The estimation problem is solved using the CMAES method of \texttt{UQLab}. The solution of the optimization problem is performed in approximately 3.2 min. We also look to analyze the reliability and errors of these results. This is achieved by computing confidence inter-

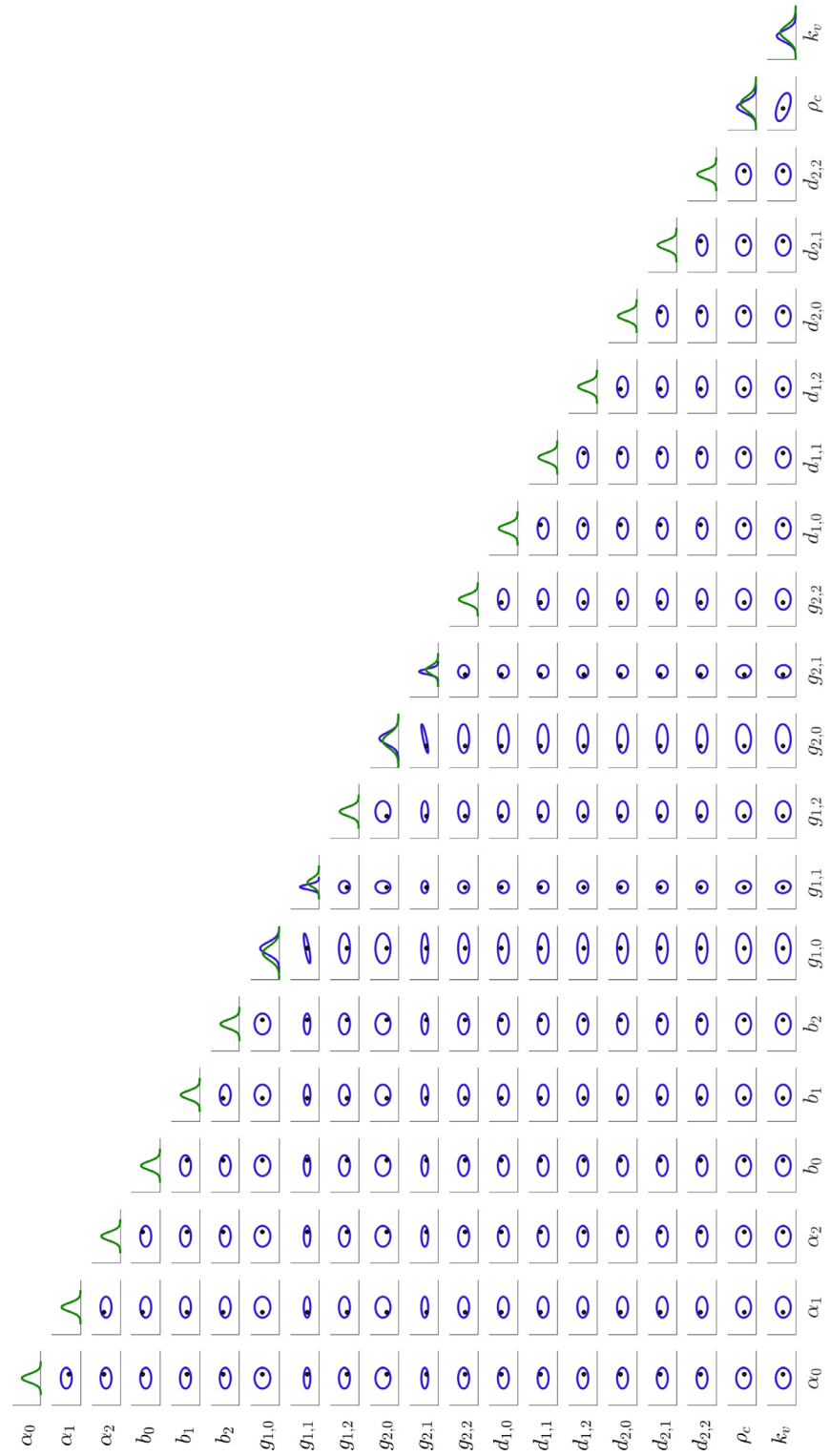

Fig. 11. Estimates of the posterior Gaussian distributions of the 20 uncertain parameters estimated via MAP. The diagonal elements represent the marginal densities of the priors (green) and posteriors (blue), while the off-diagonal sub-plots show the projected MAP estimates (pair-wise) along with the 95% confidence intervals. At the diagonal elements, where only the green Gaussian distribution is shown, the posterior coincides exactly with the prior and this it is not visible. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

vals around the estimations, as shown in Fig. 11. We employ the Laplace approximation of the log posterior $-\log\left(f_{\boldsymbol{D}|\boldsymbol{\Xi}}\left(\boldsymbol{d}|\boldsymbol{\xi}\right)\right)$ with a quadratic function from a second order Taylor expansion around the predicted mean. The ellipses drawn around the predicted mean are the 95% confidence intervals based on the chi-squared distribution. Press et al. (1992). Overall, the predicted MAP estimates are within the desired confidence intervals. The most important conclusion that can be drawn based on Fig. 11 is that the sensitivities of QoIs once again dictate which parameters are expected to be identifiable under the observation of the specific data. The juxtaposed prior and posterior distributions reveal that the optimizer has differentiated the prior and posterior distributions of the parameters with non-negligible Sobol' indices, which are the 6 parameters used in the Bayesian estimation problem. The other parameters have the same priors and posteriors, as expected from the prior regularization term that exists in the objective function. Furthermore, the MAP estimates do not reduce the uncertainty of the priors significantly, even on the identifiable parameters, especially when compared to the results of Bayesian inference. Nevertheless, the MAP estimation can be useful as a first step towards more sophisticated Bayesian inference.

5. Conclusions and future work

Surrogate modeling has shown great promise for substituting expensive, high-fidelity models with cheap-to-evaluate surrogates. Surrogate models can be especially useful for performing simulation and decision-making tasks that, for example, hinge on the solution of models consisting of partial differential equations Tripathy and Bilionis (2018), large-scale biochemical reaction network models Renardy et al. (2018), genome-scale models Kumar and Budman (2017); Paulson et al. (2019a), and closed-loop control simulations Paulson and Mesbah (2018b). As such, surrogate models facilitate fast uncertainty quantification in a variety of forward propagation and estimation problems. This paper focused on polynomial chaos-based and Kriging-based surrogate modeling methods, investigating UQ of a two-dimensional population balance model for batch cooling crystallization of ibuprofen under parameter uncertainty.

It is well-known that the truncation order of polynomial chaos expansions can grow very quickly with the increasing order of basis functions and a larger number of uncertain inputs, severely limiting the utility of polynomial chaos for UQ. Our analysis showed that inducing sparsity through basis-adaptive least-angles-regression is particularly effective for circumventing the curse of dimensionality of polynomial chaos expansions, while systematically accounting for the availability of a limited set of training data. Furthermore, it was shown that representing the trend term of Kriging surrogate models by sparse polynomial chaos expansions (i.e., sparse polynomial chaos-Kriging) not only improves the global approximation accuracy of Kriging, but also enables quantification of the local uncertainty of surrogate model predictions. Finally, we demonstrated that sparse polynomial chaos- and Kriging-based surrogate models can lead to significant computational speedups in UQ tasks such as Bayesian parameter estimation. Thus, the methods investigated in this work hold promise for creating new opportunities for probabilistic uncertainty quantification of complex dynamical systems.

Nonetheless, there are several important open challenges in polynomial chaos-based surrogate modeling that warrant further research. In this work, we adopted a time-frozen surrogate modeling approach Mai and Sudret (2017), where a separate surrogate model is constructed independently at each time instant for each QoI. This time-frozen approach can however fail in surrogate modeling of more complex dynamical systems, for example, systems with oscillatory response Mai and Sudret (2017) Gerritsma et al. (2010). To this end, several approaches have recently been investigated for constructing time-dependent surrogate models, including stochastic time wrapping Mai and Sudret (2017) and nonlinear autoregressive exogenous models Spiridonakos and Chatzi (2015).

Finally, improved sampling for training of surrogate models is an area of active research. Sampling methods such as Monte Carlo or Latin Hypercube are commonly used to generate training data, but do not follow any optimality measure. Optimal sampling methods can minimize the size of experimental design needed for surrogate model training, thus significantly enhancing the computational efficiency of building surrogate models (e.g., Hampton and Doostan (2015); Paulson et al., 2017, 2019; Sinsbeck and Nowak (2015)). Most optimal sampling methods are however tailored to specific surrogate modeling techniques, or suffer from limited scalability to large uncertainty dimensions, warranting further research in this area.

Declaration of Competing Interest

We declare no conflict of interest.

Acknowledgments

This material is based upon work supported by the National Aeronautics and Space Administration (NASA) under grant number NNX17AJ31G. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NASA. G. Makrygiorgos is grateful to the "A. Mentzelopoulos" Institution from the University of Patras, Greece.

References

Bachoc, F., 2013. Cross validation and maximum likelihood estimations of hyper-parameters of gaussian processes with model misspecification. Comput. Stat. Data Anal. 66, 55–69. doi:10.1016/j.csda.2013.03.016.

Blatman, G., Sudret, B., 2011. Adaptive sparse polynomial chaos expansion based on least angle regression. J Comput Phys 230 (6), 2345–2367. doi:10.1016/j.jcp.2010.12.021.

Blatman, G., Sudret, B., 2013. Sparse polynomial chaos expansions of vector-valued response quantities. In: Proc. 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR2013).

Caflisch, R.E., 1998. Monte carlo and quasi-Monte carlo methods. Acta Numerica doi:10.1017/S0962492900002804.

Cameron, A.R.H., Martin, W.T., 1947. Annals of mathematics the orthogonal development of non-linear functionals in series of fourier-hermite functionals. Ann. Math. 48 (2), 385–392.

Chaffart, D., Ricardez-Sandoval, L.A., 2017. Robust dynamic optimization in heterogeneous multiscale catalytic flow reactors using polynomial chaos expansion. J. Process Control 60, 128–140. doi:10.1016/j.jprocont.2017.07.002.

Chopin, N., 2002. A sequential particle filter method for static models. Biometrika 89 (3), 539–552. doi:10.1093/biomet/89.3.539.

Cressie, N., 1990. The origins of Kriging. Math. Geol. 22 (3), 47–55.

Crose, M., Kwon, J.S.-I., Tran, A., Orkoulas, G., Christofides, P.D., Nayhouse, M., Tran, A., Kwon, J.S.-I., Crose, M., Orkoulas, G., Christofides, P.D., 2015. Modeling and control of ibuprofen crystal growth and size distribution. Chem. Eng. Sci. 134, 414–422. doi:10.1016/j.ces.2015.05.033.

Debenedetti, P.G., 1996. Metastable Liquids Concepts and Principles. Princeton University Press doi:10.1088/0953-8984/15/1/308.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., Ishwaran, H., Knight, K., Loubes, J.M., Massart, P., Madigan, D., Ridgeway, G., Rosset, S., Zhu, J.I., Stine, R.A., Turlopti-ach, B.A., Weisberg, S., Johnstone, I., Tibshirani, R., 2004. Least angle regression. Ann. Stat. 32 (2), 407–499. doi:10.1214/009053604000000067.

Elishakoff, I., Haftka, R.T., Fang, J., 1994. Structural design under bounded uncertainty-optimization with anti-optimization. Comput. Struct. 53 (6), 1401–1405. doi:10.1016/0045-7949(94)90405-7.

Gautschi, W., 1982. On generating orthogonal polynomials. SIAM J. Sci. Stat. Comput. 3 (3), 289–317. doi:10.1137/0903018.

Gerritsma, M., van der Steen, J.B., Vos, P., Karniadakis, G., 2010. Time-dependent generalized polynomial chaos. J. Comput. Phys. 229 (22), 8333–8363. doi:10.1016/j.jcp.2010.07.020.

Ghanem, R.G., Spanos, P.D., 1991. Stochastic Finite Elements: A Spectral Approach. Springer-Verlag, Berlin, Heidelberg.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Gunawan, R., Fusman, I., Braatz, R.D., 2004. High resolution algorithms for multidimensional population balance equations. AIChE J. 50 (11), 2738–2749. doi:10.1002/aic.10228.

Gunawan, R., Ma, D.L., Fujiwara, M., Braatz, R.D., 2002. Identification of kinetic parameters in multidimensional crystallization processes. Int. J. Mod. Phys. B 16 (01n02), 367–374. doi:10.1142/S0217979202009883.

Hampton, J., Doostan, A., 2015. Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression. Comput Methods Appl Mech Eng 290, 73–97. doi:10.1016/j.cma.2015.02.006.

Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies. Evol. Comput. 9 (2), 159–195. doi:10.1162/106365601750190398.

Harenberg, D., Marelli, S., Sudret, B., Winschel, V., 2019. Uncertainty quantification and global sensitivity analysis for economic models. Quant. Econom. 10 (1), 1–41. doi:10.3982/QE866.

Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman & Hall/CRC.

Hermanto, M.W., Kee, N.C., Tan, R.B., Chiu, M.-S., Braatz, R.D., 2008. Robust Bayesian estimation of kinetics for the polymorphic transformation of l-glutamic acid crystals. AIChE J. 54 (12), 3248–3259.

Hu, Z., Mahadevan, S., 2016. A Single-Loop Kriging Surrogate Modeling for Time-Dependent Reliability Analysis. Journal of Mechanical Design 138 (6). doi:10.1115/1.4033428. 061406

Huan, X., Marzouk, Y.M., 2013. Simulation-based optimal bayesian experimental design for nonlinear systems. J Comput Phys 232 (1), 288–317. doi:10.1016/j.jcp.2012.08.013.

Hulburt, H., Katz, S., 1964. Some problems in particle technology. Chem. Eng. Sci. 19 (8), 555–574. doi:10.1016/0009-2509(64)85047-8.

Jakobsen, H.A., 2008. Chemical reactor modeling: Multiphase reactive flows. Springer, Berlin Heidelberg doi:10.1007/978-3-540-68622-4.

Janon, A., Klein, T., Lagnoux, A., Nodet, M., Prieur, C., 2014. Asymptotic normality and efficiency of two sobol index estimators. ESAIM 18, 342364. doi:10.1051/ps/2013040.

Jones, B.A., Doostan, A., Born, G.H., 2013. Nonlinear propagation of orbit uncertainty using non-intrusive polynomial chaos. J. Guid. Control Dyn. 36 (2), 430–444. doi:10.2514/1.57599.

Kashchiev, D., 2000. Nucleation, basic theory with applications. Butterworth Heinemann.

Kennedy, M.C., O'hagan, A., 2001. Bayesian calibration of computer models. J. R. Statist. Soc. B 63, 425–464.

Kim, K.-K.K., Shen, D.E., Nagy, Z.K., Braatz, R.D., 2013. Wiener'S polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties [Historical Perspectives]. IEEE Control Syst. 33 (5), 58–67. doi:10.1109/MCS.2013.2270410.

Kumar, D., Budman, H., 2017. Applications of polynomial chaos expansions in optimization and control of bioreactors based on dynamic metabolic flux balance models. Chem. Eng. Sci. 167, 18–28. doi:10.1016/j.ces.2017.03.035.

Lataniotis, C., Marelli, S., Sudret, B., 2015. UQLab user manual Kriging (Gaussian process modelling). Technical Report. Chair of Risk, Safety& Uncertainty Quantification, ETH Zurich. Report UQLab-V0.9-105

LeVeque, R., 2002. Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, Cambridge doi:10.1017/CBO9780511791253.

LeVeque, R.J., 2007. Finite Difference Methods for Ordinary and Partial Differential Equations. Society for Industrial and Applied Mathematics doi:10.1137/1.9780898717839.

Lindenberg, C., Mazzotti, M., 2009. Experimental characterization and multi-scale modeling of mixing in static mixers. Part 2. Effect of viscosity and scale-up. Chem. Eng. Sci. 64 (20), 4286–4294. doi:10.1016/j.ces.2009.06.067.

Liu, J.S., Chen, R., 1998. Sequential Monte Carlo methods for dynamic systems. J Am Stat Assoc 93 (443), 1032–1044. doi:10.1080/01621459.1998.10473765.

Luu Trung Duong, P., Quang Minh, L., Abdul Qyyum, M., Lee, M., 2018. Sparse Bayesian learning for data driven polynomial chaos expansion with application to chemical processes. Chem. Eng. Res. Des. 137 (21), 553–565. doi:10.1016/j.cherd.2018.08.006.

Ma, C.Y., Roberts, K.J., 2018. Combining morphological population balances with face-Specific growth kinetics data to model and predict the crystallization processes for ibuprofen. Ind. Eng. Chem. Res. 57 (48), 16379–16394. doi:10.1021/acs.iecr.8b02140.

Ma, D.L., Braatz, R.D., 2001. Worst-case analysis of finite-time control policies. IEEE Trans. Control Syst. Technol. 9 (5), 766–774. doi:10.1109/87.944471.

Maceiczyk, R.M., Demello, A.J., 2014. Fast and reliable metamodeling of complex reaction spaces using universal kriging. J. Phys. Chem. C 118 (34), 20026–20033. doi:10.1021/jp506259k.

Maggioni, G.M., Bezinge, L., Mazzotti, M., 2017. Stochastic nucleation of polymorphs: experimental evidence and mathematical modeling. Cryst. Growth Des. 17 (12), 6703–6711. doi:10.1021/acs.cgd.7b01313.

Makrygiorgos, G., Sen Gupta, S., Menezes, A., Mesbah, A., 2020. Fast probabilistic uncertainty quantification and sensitivity analysis of a Mars life support system model. IFAC-PapersOnLine.

Mai, C.V., Sudret, B., 2017. Surrogate models for oscillatory systems using sparse polynomial chaos expansions and stochastic time warping. SIAM/ASA J. Uncertain. Quant. 5 (1), 540–571. doi:10.1137/16M1083621.

Marchisio, D.L., Fox, R.O., 2013. Computational Models for Polydisperse Particulate and Multiphase Systems. Cambridge University Press, Cambridge doi:10.1017/CBO9781139016599.

Marelli, S., Sudret, B., 2014. UQLab: A Framework for Uncertainty Quantification in Matlab. In: Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management - Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, ICVRAM 2014 and the 6th International Symposium on Uncertainty Modeling a, pp. 2554–2563. doi:10.1061/9780784413609.257.

Marrel, A., Iooss, B., Van Dorpe, F., Volkova, E., 2008. An efficient methodology for modeling complex computer codes with Gaussian processes. Comput. Stat. Data Anal. 52 (10), 4731–4744. doi:10.1016/j.csda.2008.03.026.

Mersmann, A., 2001. Crystallization Technology Handbook - Second Edition Revised and Expanded, 2nd Marcel Dekker.

Mesbah, A., Kramer, H.J., Huesman, A.E., Van den Hof, P.M., 2009. A control oriented study on the numerical solution of the population balance equation for crystallization processes. Chem Eng Sci 64 (20), 4262–4277.

Mesbah, A., Nagy, Z.K., Huesman, A.E., Kramer, H.J., Van den Hof, P.M., 2011. Nonlinear model-based control of a semi-industrial batch crystallizer using a population balance modeling framework. IEEE Trans. Control Syst. Technol. 20 (5), 1188–1201.

Moustapha, M., Bourinet, J.-M., Guillaume, B., Sudret, B., 2018. Comparative study of Kriging and support vector regression for structural engineering applications. ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A 4 (2), 4018005.

Nagy, Z.K., Braatz, R.D., 2007. Distributional uncertainty analysis using power series and polynomial chaos expansions. J. Process Control 17 (3), 229–240. doi:10.1016/j.jprocont.2006.10.008.

Najm, H.N., 2009. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. Annu Rev Fluid Mech 41 (1), 35–52. doi:10.1146/annurev.fluid.010908.165248.

Paulson, J.A., Buehler, E.A., Mesbah, A., 2017. Arbitrary polynomial chaos for uncertainty propagation of correlated random variables in dynamic systems. IFAC-PapersOnLine 50 (1), 3548–3553. doi:10.1016/j.ifacol.2017.08.954.

Paulson, J.A., Martin-Casas, M., Mesbah, A., 2019. Fast uncertainty quantification for dynamic flux balance analysis using non-smooth polynomial chaos expansions. PLoS Comput. Biol. 15 (8), e1007308. doi:10.1371/journal.pcbi.1007308.

Paulson, J.A., Martin-Casas, M., Mesbah, A., 2019. Optimal Bayesian experiment design for nonlinear dynamic systems with chance constraints. J. Process Control 1–17. doi:10.1016/j.jprocont.2019.01.010.

Paulson, J.A., Mesbah, A., 2019. An efficient method for stochastic optimal control with joint chance constraints for nonlinear systems. Int. J. Robust Nonlinear Control 29 (15), 5017–5037.

Paulson, J.A., Mesbah, A., 2018. Nonlinear model predictive control with explicit backoffs for stochastic systems under arbitrary uncertainty. IFAC-PapersOnLine 51 (20), 523–534. doi:10.1016/j.ifacol.2018.11.036.

Paulson, J.A., Mesbah, A., 2018. Shaping the Closed-Loop Behavior of Nonlinear Systems Under Probabilistic Uncertainty Using Arbitrary Polynomial Chaos. In: 2018 IEEE Conf. Decis. Control. IEEE, pp. 6307–6313. doi:10.1109/CDC.2018.8619328.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in Fortran 77: The Art of Scientific Computing. Cambridge University Press doi:10.2307/2153422.

Qamar, S., Warnecke, G., Elsner, M.P., 2009. On the solution of population balances for nucleation, growth, aggregation and breakage processes. Chem. Eng. Sci. 64 (9), 2088–2095. doi:10.1016/j.ces.2009.01.040.

Ramkrishna, D., 2000. Population Balances: Theory and Applications to Particulate Systems in Engineering. Academic Press.

Ramkrishna, D., Singh, M.R., 2014. Population balance modeling: current status and future prospects. Annu. Rev. Chem. Biomol. Eng. 5 (1), 123–146. doi:10.1146/annurev-chembioeng-060713-040241.

Randolph, A.D., 1964. A population balance for countable entities. Can. J. Chem. Eng. 42 (6), 280–281. doi:10.1002/cjce.5450420612.

Rasmussen, K., Williams, C., 2006. Gaussian processes for machine learning. MIT Press doi:10.1142/S0129065704001899.

Raue, A., Kreutz, C., Theis, F.J., Timmer, J., 2013. Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. Philosoph. Trans. Roy. Soc. A 371 (1984). doi:10.1098/rsta.2011.0544.

Renardy, M., Yi, T.M., Xiu, D., Chou, C.S., 2018. Parameter uncertainty quantification using surrogate models applied to a spatial model of yeast mating polarization. PLoS Comput. Biol. 14 (5), 1–26. doi:10.1371/journal.pcbi.1006181.

del Rio-Chanona, E.A., Wagner, J.L., Ali, H., Fiorelli, F., Zhang, D., Hellgardt, K., 2019. Deep learning-based surrogate modeling and optimization for microalgal biofuel production and photobioreactor design. AIChE J. 65 (3), 915–923. doi:10.1002/aic.16473.

Russi, T., Packard, A., Frenklach, M., 2010. Uncertainty quantification: making predictions of complex reaction systems reliable. Chem. Phys. Lett. 499 (1), 1–8. doi:10.1016/j.cplett.2010.09.009.

Salvatori, F., Mazzotti, M., 2017. Manipulation of particle morphology by crystallization, milling, and heating cycles a mathematical modeling approach. Ind. Eng. Chem. Res. 56 (32), 9188–9201. doi:10.1021/acs.iecr.7b02070.

Schenkendorf, R., Xie, X., Krewer, U., 2017. An efficient polynomial chaos expansion strategy for active fault identification of chemical processes, 40. Elsevier Masson SAS doi:10.1016/B978-0-444-63965-3.50281-6.

Schöbi, R., Sudret, B., 2014. PC-Kriging: A new metamodelling method combining polynomial chaos expansions and kriging. Proc. 2nd Int. Symp. Uncertain. Quantif. Stoch. Model..

Shapiro, A., 2003. Stochastic Programming, 10 doi:10.1016/S0927-0507(03)10006-0.

Sherer, E., Hannemann, R.E., Rundell, A., Ramkrishna, D., 2007. Estimation of likely cancer cure using first- and second-order product densities of population balance models. Ann. Biomed. Eng. 35 (6), 903–915. doi:10.1007/s10439-007-9310-1.

Sinsbeck, M., Nowak, W., 2015. An optimal sampling rule for nonintrusive polynomial chaos expansions of expensive models. Int. J. Uncertain. Quantif. 5 (3).
Smoluchowski, M., 1916. Drei vortrage uber diffusion. Brownsche Bewegung und Koagulation von Kolloidteilchen. Z. Phys. 17, 557–585.
Sobol, I.M., 1993. Sensitivity analysis for nonlinear mathematical models. Math. Model. Comput. 1 (4), 407–414. arXiv:1305.4373v1., doi: 10.18287/0134-2452-2015-39-4-459-461.
Sobol' And, I.M., Levitan, Y.L., 1999. A pseudo-Random number generator for personal computers. Computers and Mathematics with Applications 37, 33–40.
Solsvik, J., Jakobsen, H.A., 2015. The foundation of the population balance equation: areview. J. Dispers. Sci. Technol. 36 (4), 510–520. doi:10.1080/01932691.2014.909318.
Spiridonakos, M.D., Chatzi, E.N., 2015. Metamodeling of dynamic nonlinear structural systems through polynomial chaos NARX models. Comput. Struct. 157, 99–113. doi:10.1016/j.compstruc.2015.05.002.
Streif, S., Petzke, F., Mesbah, A., Findeisen, R., Braatz, R.D., 2014. Optimal experimental design for probabilistic model discrimination using polynomial chaos. IFAC Proceedings Volumes 47 (3), 4103–4109. doi:10.3182/20140824-6-ZA-1003.01562. 19th IFAC World Congress
Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. 93 (7), 964–979. doi:10.1016/j.ress.2007.04.002.
Sudret, B., Marelli, S., Wiart, J., 2017. Surrogate models for uncertainty quantification: an overview. In: 2017 11th European Conference on Antennas and Propagation (EUCAP), pp. 793–797. doi:10.23919/EuCAP.2017.7928679.
Tripathy, R.K., Bilionis, I., 2018. Deep uq: learning deep neural network surrogate models for high dimensional uncertainty quantification. J Comput Phys 375, 565–588. doi:10.1016/j.jcp.2018.08.036.
Wang, F., Xiong, F., Jiang, H., Song, J., 2018. An enhanced data-driven polynomial chaos method for uncertainty propagation. Eng. Optim. 50 (2), 273–292. doi:10.1080/0305215X.2017.1323890.
Weinmeister, J., Xie, N., Gao, X., Prasad, A.K., Roy, S., 2018. Analysis of a polynomial chaos-kriging metamodel for uncertainty quantification in aerospace applications. In: AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2018. American Institute of Aeronautics and Astronautics Inc, AIAA doi:10.2514/6.2018-0911.
Xiu, D., Em Karniadakis, G., 2002. The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. 24 (2), 619–644. doi:10.1137/S1064827501387826.
Xiu, D., Karniadakis, G.E., 2002. The wiener–Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. 24 (2), 619–644. doi:10.1137/S1064827501387826.
Yamazaki, W., 2015. Stochastic drag analysis via polynomial chaos uncertainty quantification. Trans. Jpn. Soc. Aeronaut. Space Sci. 58 (2), 89–99. doi:10.2322/tjsass.58.89.
Zhang, Y., Sahinidis, N.V., 2013. Uncertainty quantification in CO2 sequestration using surrogate models from polynomial chaos expansion. Indu. Eng. Chem. Res. 52 (9), 3121–3132. doi:10.1021/ie300856p.