

Review Article

Basic Framework and Main Methods of Uncertainty Quantification

Juan Zhang ¹, Junping Yin ², and Ruili Wang ²

¹Institute of Artificial Intelligence, Beihang University, Beijing, China

²Institute of Applied Physics and Computational Mathematics, Beijing, China

Correspondence should be addressed to Ruili Wang; wang_ruili@iapcm.ac.cn

Received 30 April 2020; Revised 20 July 2020; Accepted 14 August 2020; Published 31 August 2020

Academic Editor: Xuping Zhang

Copyright © 2020 Juan Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since 2000, the research of uncertainty quantification (UQ) has been successfully applied in many fields and has been highly valued and strongly supported by academia and industry. This review firstly discusses the sources and the types of uncertainties and gives an overall discussion on the goal, practical significance, and basic framework of the research of UQ. Then, the core ideas and typical methods of several important UQ processes are introduced, including sensitivity analysis, uncertainty propagation, model calibration, Bayesian inference, experimental design, surrogate model, and model uncertainty analysis.

1. Introduction

In weapon manufacturing [1], aerodynamics modeling [2], detonation modeling [3], inertial confinement fusion [4], and other natural science and engineering fields, *test*, *modeling*, and *simulation* are three main approaches to understand a complex process. The basic element of scientific exploration is experiment and observation. For example, cylinder test is a reliable and representative measure of an explosive ability to accelerate metal [5]. However, due to the limitation of test cost, lack of test conditions, environmental damage caused by test, political constraints, and other reasons, sometimes we have to rely on other means, such as *modeling and simulation (M&S)* [6]. The process of mathematical and physical modeling is to decompose and refine the complex system and then attempt to reveal the basic principle behind a phenomenon. Often, complex mathematical models, such as partial differential equations (PDEs), have no closed-form solution [7]. Hence, we need to rely on some numerical simulation methods, such as finite element or finite difference schemes, to obtain the results. The importance of simulation is that it allows parameters to be changed in the models to understand the cause and effect of some complex phenomena which might be too expensive or dangerous for conventional experimental methods [8].

In these processes, many uncertainties are introduced. For the experiment, the main sources of uncertainty are observation error and random perturbation of some experimental conditions. Then, because of the complexity of reality, the incompleteness of information, and the limitation of cognition, the mathematical model will neglect some influence factors and can only approximate the real behavior at a certain level of accuracy, which naturally introduces discrepancies between the physical system and the mathematical model [7]. The discrepancies are the source of uncertainty, including multiple model forms describing the same physical process, random variables, unknown distribution parameters, and initial and boundary conditions. In addition, in the process of numerical solution, time-space discretization, truncation error, rounding error, and the inherent accuracy of digital system all introduce uncertainties to the system [9].

Generally, uncertainties are broadly classified into two categories: aleatory and epistemic uncertainty. Aleatory uncertainty describes the natural/intrinsic variability of a quantity of interest (QoI). Epistemic uncertainty, on the other hand, describes the lack of knowledge and is potentially reducible by acquiring more knowledge. This lack of knowledge comes from many sources [10], for example, inadequate understanding of the processes, incomplete

knowledge of the phenomena, and imprecise evaluation of the related characteristics.

The uncertainties are major obstacles for the predictive capability and the reliability of simulations [11]. Different types of uncertainties may present in a given problem and interact with each other, which will affect almost all aspects of engineering modeling and design. Hence, it is important to quantify the errors in order to be able to interpret the results [9], including identifying the main sources of uncertainty, analyzing how the uncertainty propagates in complex systems, and finding stable optimized solutions across a wide range of inputs, and then make better decisions at a known level of confidence, so as to reduce development time, prototype cost, and unexpected failure.

The research on the uncertainty in the deterministic engineering modeling of complex physical processes dates back to around 1980 [12]. After nearly four decades of development, uncertainty quantification (UQ) has played an important role and has been successfully applied in many fields. For example, NASA Ames Research Center has carried out a research on Mars atmospheric entry condition [13] to isolate the rate-limiting mechanisms and identify the chief sources of aeroheating uncertainty. Through Monte Carlo sensitivity analysis and uncertainty analysis, a total of 130 input parameters are statistically varied to shortlist a handful of parameters that essentially control the heat flux prediction. In the research of theoretical nuclear physics, Lawrence Livermore National Laboratory has carried out a research on the reliable UQ of nuclear forces [14], which is an urgent necessity in ab initio nuclear structure and nuclear reaction calculations. Based on the discrepancy between the systematic uncertainty of the Skyrme parameters and the sample standard deviation of the six interactions, they discussed and showed how a detailed statistical scrutiny of the nucleon-nucleon scattering data may provide valuable hints on the interplay between theory and experiment and their assumed uncertainties. In addition, UQ and aeroelasticity are important ingredients for the design optimization of more reliable aircraft [15]. Through designing under uncertainty, designers can seek to minimize structural weight while meeting probabilistic safety constraints. At the same time, a surrogate model with high accuracy and low computational complexity is also an important tool for aerodynamic optimization design [16]. Moreover, due to the uncertainty involved, the credibility of risk assessment results is still a major issue [17]. Hence, the uncertainty quantification in risk assessment is crucial in high-risk fields such as the nuclear and chemical industries. Parameter uncertainty characterization (representation) must be managed correctly so that it can then be propagated through risk models and produce a satisfactory output. Then, by considering all the uncertainties in the risk estimation, the decision with lower risk will be obtained.

Since 2000, the research of UQ has been highly valued and strongly supported by academia and industry and has become one of the important research directions of Applied Mathematics [18]. At the same time, different approaches of UQ have been developed, for instance, sensitivity analysis, Monte Carlo simulation, response surface approaches, evaluation of classical

statistical confidence bounds, Dempster–Shafer theory, and Bayesian inference. At present, the methods of UQ are relatively isolated and have not formed into a complete system. This paper will give the basic framework of research on UQ and systematically review the core ideas and typical methods of several important UQ processes.

2. The Framework of UQ

Figure 1 shows the connections between some of these essential components of UQ. It is obvious that the forward UQ process, such as sensitivity analysis and uncertainty propagation, always starts with the characterization of the input uncertainties. Unfortunately, this information is not always readily available. Such condition is known as the “lack of input uncertainty information” issue. Up to now, in the uncertainty, sensitivity, and validation studies of engineering, “expert opinion” or “user self-assessment” has been predominantly used [19]. Such ad hoc specifications of input uncertainty information have been considered reasonable for a long time. However, these approaches are subjective and lack mathematical rigor and can lead to inconsistencies.

The “lack of input uncertainty information” issue necessitates the research on inverse UQ (backward problem [20] or inverse uncertainty propagation [21]), including model calibration and Bayesian Inference. According to Oberkampf [20], “the backward problem asks whether we can reduce the output uncertainty by updating the statistical model using comparisons between computations and experiments.”

While the simulation-based uncertainty quantification seems straightforward, the slow convergence rate poses a major challenge in applications where the computational cost of each sample is high [11]. Experimental design is an effective method, which focuses on how to achieve the same effect as full factorial design with fewer samples [22]. Another approach for overcoming the difficulty of expensive model simulations is surrogate model or response surface method [11]. The surrogate models provide an approximate functional mapping $\widehat{\mathcal{M}}$ that replaces the true mapping \mathcal{M} . Once constructed, the surrogate models can be evaluated at negligible computational costs.

Model uncertainty also needs attention. This is mainly aimed at some phenomenological models or data-driven models with insufficient theoretical support coupled in complex physical systems or engineering problems, as well as the influence of numerical discretization techniques, such as the density and type of discrete mesh. Multiple models with similar effects but different structures can be evaluated and integrated using model averaging and model selection method combining the differences between the test and the simulations [11]. Note that, in some instances, it also depends on the optimization and calibration of model parameters, so as to eliminate the influence of other factors, such as parameter uncertainty.

3. Sensitivity Analysis

Sensitivity analysis (SA) is to study the impacts of the input parameters on the outputs of a mathematical model or

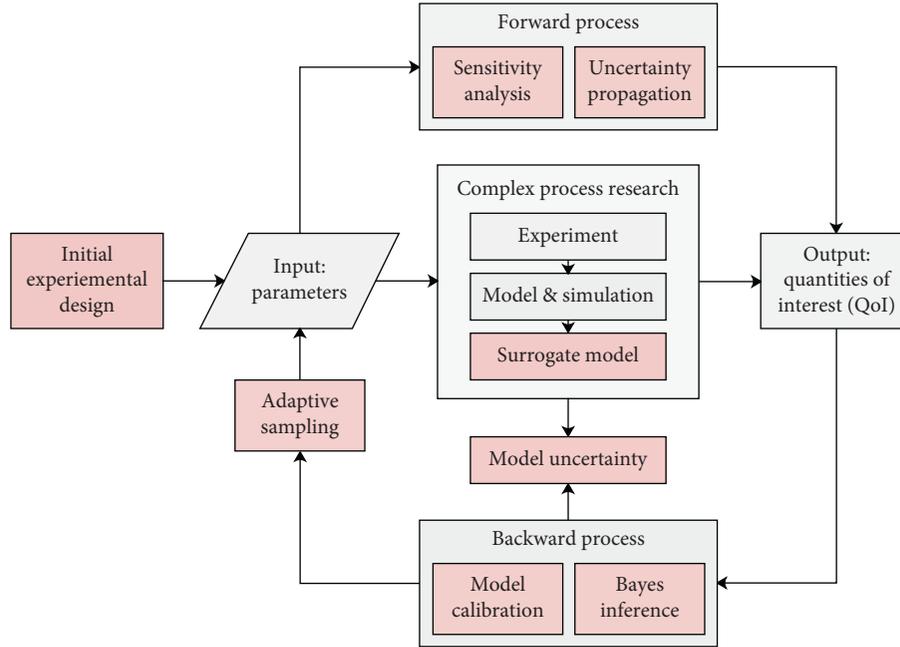


FIGURE 1: The framework and main process of UQ (boxes with pink).

system. Through SA, the importance ranking of the input parameters can be given according to their contribution to output uncertainties. This ranking can provide two main ideas for system optimization design.

First, by neglecting the uncertainty of input factors with small importance, we can effectively reduce the dimension of high-dimensional complex system to solve the “disaster of dimension,” and reduce the cost of trial and error to provide more efficient guidance information for system optimization design.

Second, by controlling the uncertainty of the important input factors, the designer can reduce the uncertainty of the model output with the minimum economic and time cost, so as to improve the robustness of the model prediction or reduce the failure probability of the structure system to the greatest extent and directly achieve the optimal design of the structure.

There are generally two types of sensitivity analysis: local sensitivity analysis and global sensitivity analysis [23]. Both statistical and deterministic methods are used for sensitivity analysis [24]. In principle, both types of procedures can be used for either local or global sensitivity and uncertainty analysis, but in practice, deterministic methods are used mostly for local analysis, while statistical methods are mostly used for global analysis.

3.1. Local Sensitivity Analysis. The local sensitivity analysis is to analyze the effects of local changes of a parameter in the system [25]. It can further gain more insight of the system for the local structure of the system. The main method for the local sensitivity analysis is the partial derivatives method [26], calculating the exact slopes provided by $\partial y / \partial x_i$, $i = 1, \dots, m$ of the model response $y = f(x)$ with respect to the i -th model parameters x_i . For some complex systems, the

partial derivatives can be determined indirectly by recalculations of the response using parameter values that deviate by a small amount, δx_i , 1% from their nominal values x_i^0 as follows:

$$\left\{ \frac{\partial y}{\partial x_i} \right\}_{\mathbf{x}^0} = \frac{f(x_1^0, \dots, x_i^0 + \delta x_i, \dots, x_m^0) - f(\mathbf{x}^0)}{\delta x_i}, \quad (1)$$

where $\mathbf{x}^0 = (x_1^0, \dots, x_m^0)$. Although this *indirect (or brute-force) method* [27] is conceptually simple to use and requires no additional model development, it is expensive for models based on nonlinear partial equations with many sensitivity coefficients because each solution of the equations is itself expensive to obtain. Additionally, it involves a trial-and-error process when selecting the parameter perturbations δx_i . Note that erroneous sensitivities will be obtained if (i) δx_i is chosen to be too small, in which case the computational round off errors will overwhelm the correct values and (ii) the parameter dependence is nonlinear and δx_i is chosen too large, in which case the assumption of local linearity is violated.

Another method of calculating the sensitivity coefficients is the *direct method*. In the direct method, the sensitivity coefficients are calculated from an auxiliary set of equations derived from the model equations. For some applications of the direct method in kinetics, the model equations and auxiliary equations have been coupled and solved together. This coupled solution procedure, however, has been found to be unstable or has failed completely for several stiff problems, and it is also quite inefficient [27]. For other applications, the auxiliary equations have been decoupled from the model equations and the two sets of equations solved separately. Among these, the most advanced and computationally economical method is the *decoupled direct method (DDM)* [23], which is originally proposed by Dunker

[27]. In this approach, the Jacobian matrix needed to solve the original system at a given time step is also used to solve the sensitivity equations at the respective time step before proceeding to solve both the original and sensitivity systems at the next time step. The decoupling procedure greatly increases the efficiency of the method by taking advantage of the fact that the auxiliary equations for different sensitivity coefficients are quite similar.

To overcome the difficulties connected with the early coupled direct method, the *Green's function method (GFM)* was developed [28]. The basis for the method is the well-known Green's function technique. By working with the same differential equations as the conventional direct method, this new approach reduces the number of differential equations to be solved and replaces them by a set of integrals. Sensitivity coefficients of all orders are expressed in integral form and evaluated in a recursive manner. Since evaluating well-behaved integrals is usually much easier than solving stiff differential equations, substantial savings can be achieved by the method when the number of system parameters is large. Afterwards, a modification of the GFM with the analytically integrated Magnus (GFM/AIM) [29] was presented, which dramatically reduces the computational effort required to determine linear sensitivity coefficients. The technique employs the piecewise Magnus method for more efficient calculation of Green's function kernels and treats the sensitivity integrals analytically.

For large-scale systems, in which the number of system parameters to be considered exceeds the number of responses, the *adjoint sensitivity analysis procedure (ASAP)* is, by far, the most efficient method even though it can only be implemented with an appropriately constructed adjoint sensitivity system [30]. The remarkable efficiency of the ASAP stems from the fact that the adjoint sensitivity system is linear in the adjoint function and is independent of any parameter variations. Hence, the adjoint sensitivity equation needs to be solved only once, for each response, in order to obtain the adjoint function. In particular, if the original model is linear in the state (i.e., dependent) variables, then the adjoint sensitivity equation can be solved independently of the original model. In turn, once the adjoint function has been calculated, it is used to obtain the sensitivities to all system parameters, by simple quadratures, without needing to solve repeatedly differential and/or integral equations.

In addition, there is another kind of sensitivity, namely, implicit sensitivity. In some cases, the implicit sensitivity was ignored because it is always difficult to be quantified. But this may lead to wrong conclusions in some cases. For example, in nuclear reactor physics calculation, the perturbation of micro cross section may impact the resonance calculation and consequently impact the macro cross section and then to the eigenvalue and neutron flux distribution. Response (eigenvalue, reaction rate, etc.) sensitivity with respect to cross sections can be divided into two parts, namely, explicit sensitivity and implicit sensitivity. The former is the direct impact of cross section perturbation on the responses through neutron transport equation, while the latter is the indirect impact of cross section perturbation on the responses through resonance self-shielding procedure [31]. As

an indirect impact related with resonance calculation, implicit sensitivity is often neglected in many sensitivity and uncertainty analysis, and many sensitivity and uncertainty analysis codes lack the ability to perform implicit sensitivity calculation. However, from the original research of Greenspan et al. [32] to the subsequent research of Williams et al. [31], the results indicated that the implicit sensitivity had a nonnegligible importance relative to the explicit sensitivity and the implicit effect had a magnitude that was more than 40% of the explicit effect in some cases. Up to now, however, most implicit sensitivity studies are mainly established for simple resonance-calculation methods such as Bondarenko method [31], generalized Stammeler method [33], and so on [34], which are not applicable for complex fuel and core designs. In order to expand the implicit sensitivity analysis method to wider application domain, Liu et al. [35] proposed a method based on the generalized perturbation theory (GPT) to calculate the implicit sensitivity coefficients by using the subgroup method in the resonance self-shielding calculation. The numerical results show that it is necessary to perform implicit sensitivity analysis in sensitivity and uncertainty analysis to obtain more rigorous results.

3.2. Global Sensitivity Analysis. To overcome the limitations of local methods (linearity and normality assumptions, local variations) [26], another class of methods has been developed in a statistical framework. In contrast to local SA, global SA aims to evaluate the entire parameter space [24] and measure the contribution of input variables to the output from the average point of view. In the global SA framework, the uncertainty of the inputs is modeled by random vectors. Common approaches for global sensitivity analysis include variance-based methods, moment-independent importance measures, and reliability-based SA.

The *variance-based methods*, such as the correlation ratio-based methods and the Sobol method [36], use variance as a measure of the importance of a parameter x_i in contributing to the overall uncertainty of the response y . For the correlation ratio-based method, the sensitivity measure (first-order effect) is written as follows:

$$S_i = \frac{\text{Var}[E(y|x_i)]}{\text{Var}(y)}, \quad (2)$$

where $E(y|x_i)$ denotes the conditional expectation of y while keeping x_i fixed, $\text{Var}(y)$ denotes the variance of y , and

$$\text{Var}[E(y|x_i)] = \int [E(y|x_i) - E(y)]^2 p(x_i) dx_i. \quad (3)$$

If x_i has great influence on y , then the fluctuation of y will be basically determined by that of x_i . It means that the variation of y at any given x_i , i.e., $\text{Var}(y|x_i)$, should be really small as well as $E[\text{Var}(y|x_i)]$. From the following decomposition of variance

$$\text{Var}(y) = \text{Var}[E(y|x_i)] + E[\text{Var}(y|x_i)], \quad (4)$$

we can see that $\text{Var}[E(y|x_i)]$ should be large as well as S_i . Hence, S_i can reflect the importance of the impact of x_i on y . The advantage of this index is that it avoids the specific form

of y and is applicable to all systems with quadratic integrability. The disadvantage is that it can only describe the importance of a single variable x_i to y , but cannot reflect the joint influence of a group of variables.

To supplement the insufficient, Sobol et al. proposed some more generalized indexes based on the orthogonal decomposition, Hoeffding decomposition [37], of the objective function [38]. Under the assumption that x_i is independent with each other and $f(\mathbf{x})$ is quadratic integral, the objective function $y = f(\mathbf{x}) = f(x_1, \dots, x_m)$ can be divided into the sum of several functions with different dimensions:

$$f = f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} + \dots + f_{12\dots m}, \quad (5)$$

where $f_i = f_i(x_i)$, $f_{ij} = f_{ij}(x_i, x_j)$, and so on for a total of 2^k terms, including f_0 . The unicity condition of the decomposition is granted by [39]

$$\int_0^1 f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) dx_k = 0, \quad (6)$$

where $k \in \{i_1, \dots, i_s\}$. The functions f_{i_1, \dots, i_s} are obtained from

$$\begin{aligned} f_0 &= E(y), \\ f_i &= E(y | x_i) - E(y), \\ f_{ij} &= E(y | x_i, x_j) - f_i - f_j - f_0, \\ &\dots \end{aligned} \quad (7)$$

and similarly for higher orders:

$$f_M(\mathbf{x}_M) = \sum_{N: N \subseteq M} (-1)^{|M|-|N|} E(y | \mathbf{x}_N), \quad (8)$$

where $M = \{k_1, \dots, k_s\} \subseteq \{1, 2, \dots, m\}$, $\mathbf{x}_M = (x_{k_1}, \dots, x_{k_s})$, and $E(y | \mathbf{x}_M)$ denotes the conditional expectation of y at given \mathbf{x}_M . Because of the orthogonality of the decomposition,

$$E(f_M(\mathbf{x}_M) f_N(\mathbf{x}_N)) = \text{Cov}(f_M(\mathbf{x}_M), f_N(\mathbf{x}_N)) = 0, \quad \forall M \neq N, \quad (9)$$

where $\text{Cov}(f_M(\mathbf{x}_M), f_N(\mathbf{x}_N))$ denotes the covariance of $f_M(\mathbf{x}_M)$ and $f_N(\mathbf{x}_N)$. Then, we can get the following conclusions:

$$D = \sum_i D_i + \sum_i \sum_{j>i} D_{ij} + \dots + D_{12\dots m}, \quad (10)$$

where $D_M = \text{Var}(f_M(\mathbf{x}_M))$. In this way, the fluctuation of output variable y is decomposed into the sum of the fluctuation of several items. Based on this variance decomposition, several indexes are proposed:

Sobol (1993) [39]:

$$S_M = \frac{D_M}{D}, \quad (11)$$

known as the Global Sensitivity Index, well reflects the interaction effect of several input variables \mathbf{x}_M on the response y .

Homma and Saltelli (1996) [40]:

$$S_{Ti} = \sum_{M_i} S_{M_i} = S_i + \sum_{j>i} S_{ij} + \dots + S_{12\dots m}, \quad (12)$$

where S_{Ti} denotes the total effect of x_i on y and M_i denotes all the subsets of $\{1, 2, \dots, m\}$ that including i .

These indexes can be used to measure the importance of the variables. And they can identify whether the interaction effect of a group of input variables on output variables exists. More importantly, they have no restrictions on the specific form of the model. The model can be nonlinear, non-monotonic, and nonadditive. However, due to the high-dimensional integration of complex functions involved in the calculation of expectation and variance, this method is sometimes difficult to achieve.

The class of *moment-independent importance measures* comprises sensitivity measures based on discrepancies between density functions, cumulative distribution functions, and value of information [41]. The name moment independent communicates the intuition that these sensitivity measures take into account the change in the entire distribution (density) of the model output, instead of the variation of one of its particular moments (e.g., variance). The common idea of these methods is to construct an index which can measure the difference between conditional distribution (by fixing a variable) and unconditional distribution to reflect the importance of the variable on the system. The following are some examples.

- (i) *Density-Based Importance Measure*: The first class of moment-independent sensitivity measures introduced is the class of density-based sensitivity measures:

$$\delta_i = \frac{1}{2} E \left[\int_{\mathcal{Y}} |p(y) - p(y | x_i)| dy \right], \quad (13)$$

where the factor 1/2 is inserted for normalization purpose, $p(y)$ denotes the density function of y , and $p(y | x_i)$ denotes the conditional density function by fixing x_i . In particular, δ_i assumes the null value if and only if y is independent of x_i . In fact, if y and x_i are independent, fixing x_i leaves the distribution of y unchanged.

The δ sensitivity measure is defined by the L^1 -norm. Similarly, several other distance measurements, such as Kullback–Leibler (K-L) divergence, can also be used to define the probabilistic sensitivity measures:

$$\theta_i = E[b_i(x_i)] = E \left[\int_{\mathcal{Y}} p(y | x_i) \log \frac{p(y | x_i)}{p(y)} dy \right], \quad (14)$$

where $b_i(x_i)$ represents the K-L divergence between the conditional and unconditional model output distributions.

Finally, while full details about estimation cannot be given due to space limitations, we observe that the estimation of any density-based sensitivity measure involves the problem of estimating the empirical density.

- (ii) *Cumulative Distribution-Based Sensitivity Measures:* A second class of moment-independent sensitivity measures is represented by taking the separation between cumulative distribution functions into account. In general, we write

$$\beta_i = E[h(F(y), F(y|x_i))], \quad (15)$$

where $F(y)$ is the unconditional model output cumulative distribution function, $F(y|x_i)$ is the cumulative distribution function conditional on fixing model input x_i , and $h(\cdot, \cdot)$ is a distance or divergence between cumulative distribution function, such as generic L_p -distance metric, Kolmogorov–Smirnov distance, and Kuiper’s distance. These metrics hold different properties, including scale invariance and transformation invariance. In Baucells and Borgonovo [42], the choice of the metric is thoroughly discussed.

In many practical problems, the primary interest of the analyst may be focused on a particular mode of failure of the system under consideration, while the detailed spectrum of probabilistic outcomes may be of secondary concern [43]. For such problems, the so-called *reliability-based algorithms* [41] provide much faster and more economical solutions regarding the particular mode of failure. The concept of “failure” is characterized by a threshold level that specifies mathematically. The reliability algorithms most often used are first-order reliability methods and second-order reliability methods [44]. Both of these methods use optimization algorithms to seek “the most likely failure point” in the space of uncertain parameters, which is defined by the mathematical model and the response function. Once this most likely failure point (referred to as the “design point”) has been determined, the probability of failure is approximately evaluated by fitting a first- (or second-) order surface at that point. Reliability algorithms have been applied to a variety of problems including structural safety, offshore oil field design and operation, and multiphase flow and transport in subsurface hydrology.

4. Uncertainty Propagation

Uncertainty propagation (UP, or error propagation) [45] aims to measure the impact of disturbances in the input variables on the system output. This information is critical when the system needs to quantify the confidence of the outputs [9]. Through

appropriate methods, we can make horizontal comparisons of the uncertainties generated by several different models or analyze which steps are the key factors of the diffusion of uncertainty in a complex multiphysical process coupled system. Unlike SA, which focuses on the importance ranking of input, UP pays more attention to the result of error propagation in the system and its influence on system stability. In general, probabilistic method is a mature methodology for uncertain problems when sufficient sample information is available to construct the accurate probability distributions of random inputs [46]. However, due to the expensive experimental cost, the sample data in many engineering practices are always scarce. In this case, there is not enough information to make accurate probabilistic representation of the different types of uncertainties existing in the system. Therefore, some other promising uncertainty theories have emerged, for example, nonprobabilistic interval process [47] and Dempster–Shafer theory [48]. Compared with the traditional probability theory, these new representations of uncertainty are able to represent the epistemic uncertainty more accurately.

4.1. Probability Boxes. In many cases, knowledge on the input is incomplete and probability theory is not sufficient to describe the uncertainty. This motivates the introduction of so-called probability boxes (p-boxes) which account for aleatory as well as epistemic uncertainty in the description of a variable [49]. There are two distinguished types of p-boxes, namely, free and parametric p-boxes.

Mathematically speaking, a *free p-box* is defined by lower and upper bounds denoted by $\underline{F}(x)$ and $\overline{F}(x)$ on the cumulative distribution function (CDF) of a variable. A free p-box can be constructed by Kolmogorov–Smirnov confidence bound [50], Chebyshev’s inequalities [51], and robust Bayes’ method [52]. This implies that the true CDF can have an arbitrary shape as long as it fulfills the characteristics of a generic CDF and lies within the bounds of the p-box. Because the shape of the true CDF is not specified, different types of curves are possible, including nonsmooth ones [49].

Parametric p-boxes (or distributional p-boxes) are defined as distribution function families, the parameters of which are known in intervals:

$$F(x|\theta), \quad \theta \in \mathcal{D}_\theta, \quad (16)$$

where \mathcal{D}_θ is the interval domain of the distribution parameters. And the lower and upper boundary curves of the parametric p-box are obtained by

$$\begin{aligned} \underline{F}(x) &= \min_{\theta \in \mathcal{D}_\theta} F(x|\theta), \\ \overline{F}(x) &= \max_{\theta \in \mathcal{D}_\theta} F(x|\theta). \end{aligned} \quad (17)$$

A parametric p-box can be generated by determining confidence bounds onto the distribution parameters. Parametric p-boxes allow for a clear separation of aleatory and epistemic uncertainty [7]: aleatory uncertainty is represented by the distribution function family, whereas epistemic uncertainty is represented by the intervals in the distribution parameters. However, parametric p-boxes are

more restrictive than free p-boxes because they require knowledge on the distribution family.

4.2. Dempster–Shafer Theory. *Dempster–Shafer theory (D-S theory)* is also known as evidence theory. In particular, evidence theory considers two measures, *i.e.*, belief and plausibility, for each event \mathcal{E} in the event space Ω :

$$\begin{aligned} \text{Bel}(\mathcal{E}) &= \sum_{\mathcal{F} \subset \mathcal{E}} m(\mathcal{F}), \\ \text{Pl}(\mathcal{E}) &= \sum_{\mathcal{F} \cap \mathcal{E} \neq \emptyset} m(\mathcal{F}), \end{aligned} \quad (18)$$

where the belief measure $\text{Bel}(\mathcal{E})$ is defined as the minimum amount that must be associated with an event \mathcal{E} , whereas the plausibility is defined as the maximum amount that could be associated with \mathcal{E} . And $m(\mathcal{F})$ is interpreted as the amount of likelihood that is associated with event \mathcal{F} but without any specifications of how this likelihood might be appointed. Note that

$$\sum_{\mathcal{F}} m(\mathcal{F}) = 1. \quad (19)$$

The advantage of using evidence theory lies in the fact that it can be successfully used to quantify the degree of uncertainty when the amount of information available is small. And the evidence theory loosens the strict assumption of a single probability measure \mathbb{P} in probability theory. Apart from that, since there is uncertainty in the given information, the evidential measure for the occurrence of an event and that for its negation do not have to sum to unity. However, Dempster–Shafer theory is based on the assumption that these sources are independent. This rule has been subject to some criticism in the sense that it tends to completely ignore the conflicts that exist between the available evidence from different sources.

4.3. Interval Process. *Interval process* is considered as a flexible method for uncertainty propagation because only the variational ranges need to be well defined without any detailed statistical characteristics of uncertain parameters [47]. Along with the widespread concern in the recent two decades, many interval approaches have been successfully developed [53]. By treating the response bounds as two extreme value models, theoretically the optimization-based methods could obtain accurate results [54]. It means that when the system response was monotonic with respect to the uncertain parameters, the accurate response bounds could be derived by the interval vertex methods, where only the endpoint combinations of interval parameters were simulated [55]. When the system does not satisfy monotonicity, the response bounds can be directly simulated via substantial sample processes, but sometimes also obtained via monotonic interval determination with the help of sensitivity analysis to avoid the huge computational cost.

In practice, p-boxes and D-S theory provide measures of the uncertainty of the variables. However, in order to

perform the uncertainty propagation, we need to combine the interval process. To obtain more details about the response on each of the variables with different domains and weaken the presupposition of monotonicity, slicing algorithm needs to be introduced. The slicing algorithm transforms the propagation of p-boxes into the propagation of a large number of intervals [7].

First, each p-box is discretized into a number of intervals and associated probability masses. For variable x_i , the interval $[0, 1]$ is divided into n_{x_i} subintervals with corresponding thickness $m_i^{(j)}$ where $j = 1, \dots, n_{x_i}$ and $\sum_j m_i^{(j)} = 1$. Let us denote the lower and upper boundaries of these intervals by $\underline{c}_i^{(j)}$ and $\bar{c}_i^{(j)}$, respectively. Given the bounds of the free p-box $[\underline{F}_i, \bar{F}_i]$ of x_i , the corresponding intervals in \mathcal{D}_i are as follows:

$$\underline{x}_i^{(j)} = \bar{F}_i^{-1}(\underline{c}_i^{(j)}), \bar{x}_i^{(j)} = \underline{F}_i^{-1}(\bar{c}_i^{(j)}). \quad (20)$$

Then, let \mathcal{K} be a set of multi-indices defining a combination of intervals of each input parameter x_i :

$$\mathcal{K} = \{\mathbf{k} = (k_1, \dots, k_m), k_i \in \{1, \dots, n_{x_i}\}, i = 1, \dots, M\}. \quad (21)$$

Let $\mathcal{D}_{\mathbf{k}}$ be the hyperrectangle defined by

$$\mathcal{D}_{\mathbf{k}} = \left[\underline{x}_1^{(k_1)}, \bar{x}_1^{(k_1)} \right] \times \dots \times \left[\underline{x}_m^{(k_m)}, \bar{x}_m^{(k_m)} \right]. \quad (22)$$

For each $\mathcal{D}_{\mathbf{k}}$, two optimization problems are solved to define the associated the bounds of y :

$$\begin{aligned} \underline{y}^{\mathbf{k}} &= \min_{\mathbf{x} \in \mathcal{D}_{\mathbf{k}}} f(\mathbf{x}), \\ \bar{y}^{\mathbf{k}} &= \max_{\mathbf{x} \in \mathcal{D}_{\mathbf{k}}} f(\mathbf{x}). \end{aligned} \quad (23)$$

When f and n_{x_i} become large, this quickly becomes intractable due to the large number of optimizations. Hence, sometimes we need surrogate models to simplify the process.

5. Model Calibration

Here we focus on the difference between computer simulations and physical experiments. Model calibration aims to reduce the parameter uncertainty and improve consistency between the computer models and physical experiments by adjusting some selected tunable parameters through estimation, optimization approaches, and so on. Tunable parameters refer to the modeling or calculation parameters that have no physical meaning or have physical meaning but have great cognitive defect and have obvious influence on Model and Simulation (M&S) results. Calibration can be classified as deterministic and statistical calibration [56]. Deterministic calibration merely determines the point estimates of best-fit input parameters to minimize the discrepancies between code output and experimental data [57].

The earliest calibration work is mostly manual, or based on the modeler's experience. For example, for the parameter calibration of the equation of state in detonation test, Lee et al. [58] put forward an initial guess for the parameters and varied their values in a hydrocode and then repeatedly calculated

until a satisfactory agreement between calculation and experimental result is obtained. In these research studies, parameters were “guessed” [58] or “fiddled” [59] based on experience. Current deterministic calibration models range from manually tuning parameters [60], to brute-force search methods [61]; all rely on repeated simulation. An underlying challenge in simulation-based calibration is the curse of dimensionality, which characterizes problems in which the search space (and therefore computation time) increases drastically—often exponentially—with the number of parameters being simultaneously estimated. Hence, simulation-based calibration method is more suitable for small-scale problems. As for large-scale applications, we need to rely on high-precision surrogate models [62].

Moreover, the calibration problem requires the location of the minimum of an objective function that measures the fit of the model results to the data (e.g., the sum of squares of errors). Unfortunately, the objective function may have many local minima and the global minimum may reside in a part of the parameter space that the modeler considers undesirable [60]. However, the statistical method can get a more global calibration result. We will review the statistical inverse UQ in detail in the next section.

6. Bayesian Inference

The inference method based on Bayesian is a typical UQ method under probabilistic representation. It can be used to analyze how the uncertainty is transferred from input to output in a complex multilevel system. It can also be used to quantify the uncertainty of the input parameters through a posteriori probability analysis to reduce the discrepancy between experimental and numerical simulation.

6.1. Hierarchical Bayesian Approach. Simulation of a complex physical system involves multiple levels of modeling, such as material (lowest level) to component to subsystem to system (highest level). Different interdependent physics-based predictive models and their implementation in computer simulation codes are developed at each level. The uncertainties in the input variables are propagated through the simulation codes from one level to next level. The challenges are to identify the relationships between the models at various levels and uncertainty propagation between the computational models.

To quantify the uncertainties in multilevel models, a *hierarchical Bayesian approach* is proposed [63], which is a structural equation modeling (SEM) with latent variables supplemented with Bayesian regression and inference methods. The variabilities and uncertainties are quantified at individual levels, and the probability of meeting requirements may be assessed by model extrapolation from the subsystem to the full system. By using Bayesian estimation, unbiased estimates are derived for the relationships between latent variables at different levels and the measurement and prediction errors are modeled explicitly, while the variabilities of input variables in the computational model are updated effectively [64].

Multilevel Bayesian methods allow using more flexible assumptions regarding both the model parameters and prediction error probability distributions [65]. Development of hierarchical Bayesian models has brought about many successful applications in different scientific disciplines [66]. In molecular dynamics, hierarchical models have recently been developed for calibrating parametric models and fusing heterogeneous experimental data from different system operating conditions [67]. In structural dynamics, Behmanesh et al. [68] have developed a hierarchical framework to model and consider the variability of modal parameters over dissimilar experiments. This framework has found extensive applications in uncertainty quantification and propagation of dynamical models. Nagel et al. [69] have proposed a unified multilevel Bayesian framework for calibrating dynamical models for the special case of having noise-free vibration measurements. Sedehi et al. [70] introduce a novel Bayesian hierarchical setting, which breaks time-history vibrational responses into several segments so as to capture and identify the variability of inferred parameters over multiple segments.

6.2. Bayesian Inference for Inverse UQ. Inverse UQ aims to quantify the uncertainty in input parameters such that the discrepancies between code output and observed experimental data can be reduced [71]. In this sense, it is similar to model calibration or parameter estimation. However, unlike the deterministic calibration methods, the statistical calibration methods also capture the uncertainty of the estimates rather than merely determining point of the best-fit input parameters [19].

The inverse UQ mainly employs the Bayesian inference theory and explores the posterior PDF with Monte Carlo sampling. According to the Bayesian inference theory, the information, posterior PDF, for the input parameters \mathbf{x} can be obtained given the observation of the data:

$$p(\mathbf{x} | y) = \frac{p(y | \mathbf{x})\pi(\mathbf{x})}{\int p(y | \mathbf{x})\pi(\mathbf{x})d\mathbf{x}}, \quad (24)$$

where $\pi(\mathbf{x})$ is the prior, $p(y | \mathbf{x})$ is the likelihood function, and $p(\mathbf{x} | y)$ is the posterior. In brief, prior and posterior probabilities represent degrees of belief about possible values of \mathbf{x} , before and after observing the data y . A simple model for the likelihood assumes that independent additive errors account for the deviation between predicted and observed values of y :

$$y = f(\mathbf{x}) + \varepsilon, \quad (25)$$

where the components of ε are *i.i.d.* (independent and identically distributed) random variables. The posterior PDF $p(\mathbf{x} | y)$ is the Bayesian solution to the inverse problem. Compared with most of the deterministic inverse methods, it results in not just a single value but a PDF. Various moments and marginal densities can be computed from the posterior PDF. However, such defined posterior PDF is nonstandard and implicit, not normalized, and we need a numerical sampling method to explore this posterior PDF.

Monte Carlo sampling is commonly used to explore the posterior PDF by approximating the integrals

$$I(f) = \int f(x)p(x)dx, \quad (26)$$

with the numerical form

$$\hat{I}_N(f) = \frac{1}{N} \sum_{i=1}^N f(x_i). \quad (27)$$

Plain Monte Carlo methods are rarely used due to its difficulty to randomly sample from complex or high-dimensional distributions. Instead, *Markov chain Monte Carlo* (MCMC) methods are commonly used, which are a class of sequential sampling strategies in which the next sampled state only depends on the current state [11]. The MCMC algorithm samples from a given distribution by constructing a Markov chain whose stationary distribution coincides with the target distribution, such as Metropolis–Hastings sampling and Gibbs sampling. A major challenge of MCMC is that it requires a large number of samples to achieve statistical convergence. Typically the required number of samples range from $O(10^5)$ to $O(10^6)$, with the specific number depending on the shape of the posteriori distribution and the effectiveness of the sampling. In CFD applications, each evaluation involves a simulation that takes hours or even weeks to run. Clearly, it is impractical to perform a full simulation for each evaluation of likelihood in the MCMC sampling. Therefore, surrogate models are commonly used for likelihood evaluation in MCMC-based model uncertainty quantification to alleviate the high computational cost of simulations [72].

Besides, when the exact probability is not critical and only the low-order moments such as the mean and the variance are important, various approximate Bayesian inference methods can be used. These methods use *maximum a posteriori* (MAP) probability estimate to obtain the mode (peak) of the posterior and not the full posterior distribution. The MAP point can be obtained by finding the optimal values for \mathbf{x} that maximize the log posterior (minimize the negative log posterior):

$$\mathbf{x}_{\text{MAP}} = \arg\min_{\mathbf{x}} [-\log(p(\mathbf{x} | y))]. \quad (28)$$

This effectively eliminates the burn-in procedure for an MCMC chain where some initial portion of the Markov chain is discarded, as the MCMC chain can instead be initiated from a high probability starting point: the MAP solution.

Further, the MAP estimate can be computed in several other ways, among which the most commonly used are variational methods [73] and ensemble methods [74]. In variational methods, the minimization problem is often solved by using gradient descent methods, with the gradient obtained with adjoint methods. In contrast, ensemble methods use samples to estimate the covariance of the state vector, which is further used to solve the optimization problem. Variational methods have been the standard in data assimilation and still dominate the field, while ensemble

methods such as ensemble Kalman filtering have matured in the past decades [75]. Hybrid approaches combining both approaches are an area of intense research and have been explored in CFD applications. Moreover, mathematicians have performed analyses to shed light on why they have worked well in practice even with theoretical limitations [76].

7. Experimental Design

In M&S procedure, the choice of the experimental design, *i.e.*, the set of input samples, is crucial for an accurate representation of the computational model [77]. Various approaches are available, from purely deterministic to fully stochastic sampling techniques. The performance of a sampling strategy and the quality of its resulting samples directly control the efficiency and robustness of any associated sampling-based analysis. The key point is how to estimate the behavior of the computational model with a few representative samples, which should be carefully chosen so that the experimental design covers the entire space of input parameters.

7.1. Initial Experimental Design. Intuitively, the regular grid is a good choice to cover the whole space of input variables in a deterministic way [77]. It is a full factorial design, *i.e.*, all regions are covered regularly with the same density of samples in each subdomain. A drawback of the regular grid is that the design is full factorial. This implies that it involves a large number of samples, *i.e.*, $N = l^m$, where l is the number of levels of each dimension and m is the number of variables (dimensions). Therefore, we need some more efficient methods to generate samples.

In theory, *Monte Carlo sampling* is a purely stochastic design of experiments. The samples are generated randomly according to an assumed probability density function. For more efficient sampling strategies, semirandom designs were created, which include randomness and a deterministic part [77].

Orthogonal experimental design is one method of the deterministic part. When there exists interaction between variables, the workload of the experiment will become very large, even difficult to implement. To solve this problem, orthogonal experimental design is undoubtedly a better choice [78]. A highly efficient tool of orthogonal experimental design is orthogonal table [79]. The experimenter can find out the corresponding orthogonal table according to the number of factors, the level number of factors, and whether there is interaction and then select some representative points from the comprehensive experiment. The orthogonality makes it possible to achieve the equivalent result with the least number of experiments as a large number of comprehensive experiments. Therefore, orthogonal table design is an efficient, fast, and economic multifactor test design method [80].

Latin hypercube sampling (LHS) is the special case of an orthogonal array [77], which is the most widely used method at present. The property of this experimental design is that

the projection onto any axis in the p -dimensional space results in a uniform distribution [81]. The space of input parameters is defined by a regular grid. The samples are then arranged so that there is one sample in each dimension of the grid. Inside each square, the sample coordinates are chosen randomly [82], such as random Latin hypercube (RLHS) and median Latin hypercube (MLHS). LHS is typically used to save computer processing time when running Monte Carlo simulations.

Due to the multiple dimensions of the input parameters in reactor-physics M&S, the sample size required by the conventional sampling methods is very huge and it is impossible to directly estimate the exact sample size which can provide the converged UQ results. Therefore, it is an issue to notably reduce and determine the sample size on the basis of ensuring the consistent UQ results with those under infinite sample size and infinitesimal statistical fluctuations. In this context, Sui et al. [83] have proposed a covariance-oriented sample transformation (COST) to generate multivariate normal distribution samples for uncertainty analysis. In this method, samples from the standard normal distribution are transformed linearly in which the mean and covariance of the transformed samples are ensured equal to that of the input parameter population, respectively. In this way, the transformed samples can fully describe the uncertainty information of the input parameters. From the numerical result comparisons, it can be observed that the consistent uncertainty analysis results can be provided by COST with very small sample size, compared with the conventional sampling methods with very huge sample size.

7.2. Adaptive Experimental Design. In practice, we hope to have a “proper sample size” from the initial sampling process for a given simulation model and sampling-based analysis [84]. First, it refers to a sufficiently large number of sample points that can ensure the convergence or robustness of the analysis results. Then, on the premise of achieving the previous goal, we hope the sample size can be as small as possible. However, the “proper sample size” is not typically known a priori. It means that we may have to supplement some samples in the subsequent analysis process according to the information we get from the initial samples and calculations.

One major drawback of traditional LHS and many other sampling strategies is that they generate the entire sample points at once [84], which is referred to as one-stage or one-shot sampling. This requires users to specify the sample size prior to the associated sampling-based analysis. Also, it is often the case that the user is not satisfied with the resulting sampling-based analysis (e.g., convergence criteria are not met), and needs to enlarge the sample size and resumes the sampling-based analysis with the updated/new sample. The need that maintains the desired distributional proprieties while the sample size grows progressively warrants the development and application of multistage or sequential sampling. This way, sequential sampling will allow the user to monitor the performance of the sampling-based analysis and assess the stopping criteria (e.g., convergence criteria) in an online manner. To address these, Sheikholeslami et al.

[84] introduced a novel strategy, called *PLHS* (*progressive Latin hypercube sampling*), for sequentially sampling the input space while progressively maintaining the Latin hypercube properties. The proposed PLHS is composed of a series of smaller slices generated in a way that the union of these slices from the beginning to the current stage optimally preserves the desired distributional properties and at the same time achieves maximum space-filling.

In addition, unlike space-filling DoE techniques (e.g., LHS), there is another special kind of adaptive design technique generating more design points in areas of interest, for example, the areas with high gradients of the response function. It starts with generating initial design by one of the space-filling techniques and builds an approximation based on this initial sample. Adaptive design technique then enriches sample iteratively by adding the next best design point to the most interesting regions, minimizing the uncertainty of the approximation [85]. The question then becomes one of determining the meaning of best. In information theory, the mutual information is a measure of the reduction in the uncertainty of one random variable due to the knowledge of another [86]. Recast into the context of experimental design, the mutual information represents how much the proposed experiment and resulting observation would reduce the uncertainties in the model parameters. Therefore, given a set of experimental design conditions, that which maximizes the mutual information is the most desirable. This is the premise that motivates the Bayesian experimental design algorithm implemented in Dakota [85]. Similarly, Sudret [7] introduced an efficient adaptive experimental design strategy to add multiple samples at each iteration to increase the accuracy in the estimation of the QoI. Further, this algorithm is equipped with a new stopping criterion which monitors the convergence of the QoI better than existing ones. This further reduces the total computational resources needed for a more accurate estimation.

8. Surrogate Model

State-of-the-art numerical simulations are often characterized by a vast number of input parameters [87]. That means these applications always require thousands to millions of runs of the high-fidelity (parameterized) computational models, which is not affordable in many practical cases even with high-performance, parallel computing architectures [7]. Hence, the idea of using a simpler surrogate model to represent a complex phenomenon has gained increasing popularity over past three decades [88]. Surrogate model, also known as response surface or meta-model, is a set of easy-to-evaluate mathematical functions that approximates the actual simulation model based on pairs of input-output samples [89]. Here, we focus on surrogate models which treat the computational model as a black box and just need the input values and the corresponding output values of the QoI. Among them, there are several types of models that are widely used and have received sustained attention from researchers: polynomial chaos expansions (PCEs), kriging, reduced-order model (ROM), and artificial neural network (ANN).

8.1. Polynomial Chaos Expansions. PCE models approximate the computational model by the linear combination of a series of multivariate polynomials:

$$y = \sum_{\alpha \in \mathbb{N}^m} a_{\alpha} \psi_{\alpha}(\mathbf{x}), \quad (29)$$

where $\alpha = (\alpha_1, \dots, \alpha_m)$ are the multi-indices with $\alpha_i, i = 1, \dots, m$ denoting the degree of the univariate polynomial in x_i ,

$$\psi_{\alpha}(\mathbf{x}) = \prod_{i=1}^m \psi_{\alpha_i}^i(x_i), \quad (30)$$

are multivariate orthonormal polynomials which are built in coherency with the distribution of the input random vector \mathbf{x} , and $\{a_{\alpha}, \alpha \in \mathbb{N}^m\}$ are the associated deterministic coefficients.

If the simulation response function can be assumed as a continuous and well-behaved function of the input variables, then one may represent the function as a Taylor series about some particular environment condition [90]. However, since the computation model is not explicitly expressed, the coefficients of a Taylor series cannot be obtained by derivation, but rather by regression methods such as least square (LS) and maximum likelihood (ML) [91]. In practice, it is not tractable to use an infinite series expansion. In the early days, people focused on how to find the orthonormal basis [92, 93]. The so-called projection methods were developed [94] for a comprehensive review. Their findings open the possibility of representing stochastic processes with different orthogonal polynomials according to the property of the processes.

In practice, the use of infinite-dimensional PCEs is not tractable. And typically for smooth functions, a small number of polynomials are able to represent accurately the output of the computational model. Hence, some other researchers pay attention to obtain approximate representations by means of truncation. They introduced some truncation schemes to reduce the number of candidate polynomials [95]. The truncated representation is given by

$$y = \sum_{\alpha \in \mathcal{A}} a_{\alpha} \psi_{\alpha}(\mathbf{x}) + \epsilon, \quad (31)$$

in which \mathcal{A} is a truncation set and ϵ is the truncation-induced error. A classical truncation scheme consists in selecting all polynomials of the total degree less than or equal to p , when the truncation set reads

$$\mathcal{A}_q^{m,p} = \{\alpha \in \mathbb{N}^m : \|\alpha\|_q \leq p\}, \quad (32)$$

where

$$\|\alpha\|_q = \left(\sum_{i=1}^m \alpha_i^q \right)^{1/q}, \quad (33)$$

where $0 < q \leq 1$ is a hyperparameter and a decreasing q leads to a smaller number of interactive polynomials. Later, some variable selection methods in the field of statistics are gradually introduced into the construction process of a PCE model to screen out variables that have an important impact

on response. By means of submodel testing, adding penalty term to the likelihood function or to the least square objective function [96], the variables are reduced and a sparse model is constructed.

8.2. Kriging. Kriging [97], also known as Gaussian process modeling, assumes that the computational model is a realization of a Gaussian random process. It is a regression algorithm for spatial modeling and prediction of stochastic process or random field based on covariance function [98]. The property of the best linear unbiased predictor makes it widely applied in various fields related to spatial statistics, such as geostatistics, environmental science, and atmospheric science. The general form of a kriging surrogate can be formulated as the summation of two terms: a trend of mean prediction defined by some known independent basis functions at specified location and a random error with zero mean distribution implied the correlation of two distinct samples:

$$y = \boldsymbol{\beta}^T \mathbf{g}(\mathbf{x}) + \sigma^2 z(\mathbf{x}, \omega), \quad (34)$$

where $\boldsymbol{\beta}^T \mathbf{g}(\mathbf{x}) = \sum_{j=1}^{n_r} \beta_j g_j(\mathbf{x})$ is the mean value (a.k.a trend) of the Gaussian process, and β_j is the trend coefficient corresponding to the j -th function g_j . σ^2 is the process variance and $z(\mathbf{x}, \omega)$ is a zero-mean, unit-variance stationary Gaussian process. $z(\mathbf{x}, \omega)$ is characterized by an autocorrelation function $R(\mathbf{x}, \mathbf{x}') = R(\|\mathbf{x} - \mathbf{x}'\|; \rho)$ with the hyperparameter ρ .

Various correlation functions can be found in the literature [99], including linear, exponential, Gaussian (also called squared exponential), and Matérn autocorrelation functions. According to the complexity of the trend function, there are three popular kriging models [100], namely, simple, ordinary, and universal kriging. *Simple kriging* assumes that the trend has a known constant value, i.e., $\boldsymbol{\beta}^T \mathbf{g}(\mathbf{x}) = \beta_0$. *Ordinary kriging* assumes that the trend has a constant but unknown value, i.e., $n_r = 1$, $f_1(\mathbf{x}) = 1$ and β_1 is unknown. The most general and flexible is *universal kriging*, which assumes that the trend is composed of a sum pre-selected functions $g_j(\mathbf{x})$.

However, specifying a trend or a value for the mean when the underlying function is unknown may lead to inaccuracy in prediction. To avoid this problem, blind kriging is proposed [101] using a Bayesian technique to select models that have maximum posterior probability. Overall, kriging focuses on the local behavior of the computational model, resulting in high prediction accuracy close to sample points of the experimental design, but its global behavior can be poor. To make up for this, Schöbi et al. [77] proposed a novel PC-kriging approach combining kriging with the classical PCE:

$$y = \sum_{\alpha \in \mathcal{A}} a_{\alpha} \psi_{\alpha}(\mathbf{x}) + \sigma^2 z(\mathbf{x}, \omega), \quad (35)$$

which holds a good global approximation property. And numerical results show that PC-kriging performs better than the two traditional meta-modeling techniques taken separately [7].

8.3. Reduced-Order Model. In the study of many complex physical processes, the solution is usually high-dimensional, which poses many mathematical challenges for traditional statistical methods. For example, despite tremendous progress seen in the computational fluid dynamics community for the past few decades, numerical tools are still too slow for the simulation of practical flow problems, consuming thousands or even millions of computational core-hours. To enable feasible multidisciplinary analysis and design, the numerical techniques need to be accelerated by orders of magnitude. Reduced-order modeling (ROM) has been considered one promising approach for such purposes.

From a mathematical perspective, the problem of dimensionality/order reduction can be formulated as follows: given the m -dimensional vector $\mathbf{q} = \{q_1, \dots, q_m\}^\top$, find a reduced representation $\mathbf{q}^r = \{q_1^r, \dots, q_{m^r}^r\}^\top$ which can retain the geometry of the data as much as possible. Here, m^r is the intrinsic dimensionality, and $m^r \ll m$ is expected. Moreover, intrinsic dimensionality indicates that the original data \mathbf{q} is lying on or near a manifold with dimensionality m^r which is embedded in the M -dimensional space. In general, dimensionality reduction techniques can be divided into two groups: i.e., linear and nonlinear methods. Linear techniques assume that the data lies on or near a linear subspace of the high-dimensional space, which can be written as $\mathbf{q} = \mathbf{W}\mathbf{q}^r$, where $\mathbf{W} \in \mathbb{R}^{m \times m^r}$ is a linear transformation matrix. For nonlinear methods, the linear transformations are replaced with $\mathbf{q} = \mathbf{G}(\mathbf{q}^r)$, where \mathbf{G} denotes the nonlinear mapping function. Here we introduce two popular ROM technologies; for a detailed review of other ROMs, the readers are referred to work by Benner et al. [102] and Yu et al. [103].

One of the most popular linear techniques is the proper orthogonal decomposition (POD) [103], which is also known as Karhunen–Loève procedure, principal component analysis (PCA), Hotelling analysis, empirical component analysis, quasiharmonic modes, and empirical eigenfunction decomposition in different fields. POD was originally introduced to fluid problems to help analyze coherent structures of turbulence. Through singular value decomposition (SVD), this yields an orthonormal basis, the POD basis. It is optimal in the sense that, for an orthonormal basis of size r , it minimizes the least squares error of snapshot reconstruction. Due to its broad applicability to linear and nonlinear systems, as well as to parametrically varying systems, the proper orthogonal decomposition (POD) has become widely used in many different application domains as a method for computing the reduced basis.

Dynamic mode decomposition (DMD) has been designed to decompose time-resolved data into modes, each one of which corresponds to a single characteristic frequency and growth/decay rate. In principle, DMD is the eigendecomposition of a best-fit linear operator that approximates the underlying dynamics embedded in the datasets. A number of improved variants of DMD have occurred in the literature recently. Hemati et al. [104] proposed an efficient DMD method for large datasets. Also, a parallel version of DMD was developed by Belson et al. [105]. Moreover, it should be noted that the original DMD modes are not orthogonal, which is especially undesirable for a reduced-

order model. Orthogonalized DMD modes can be obtained with a recursive method [106]. For problems with highly intermittent dynamics, the original DMD method shows poor performance, for which the multiresolution [107] and time-delay DMD [108] methods are promising.

8.4. Artificial Neural Network. ANN is also a good choice for surrogate model due to its proved universal approximation, that is to say, a standard multilayer feedforward network with a locally bounded piecewise continuous activation function can approximate any continuous function to any degree of accuracy if and only if the network's activation function is not a polynomial [109, 110]. ANN has been widely applied in modern surrogate modeling with its advantage of consuming trivial computational effort; thus, it can be used as the assistant for CFD calculation [111, 112] for a large number of designed geometries, which greatly increases the optimization efficiency. Its accurate generalization and parallel computation capabilities in complex engineering design problems are helpful in the rapid investigation of design space and searching for the optimal solution [113]. For example, ANN has been used to expedite decision-making process in early stages of aircraft design process and to select proper combination of engine thrust, wing area, and the aircraft weight without going through elaborate details of other direct approaches [114].

However, more often than not, in the course of analyzing complex physical, the cost of data acquisition is prohibitive and we are inevitably faced with the challenge of drawing conclusions and making decisions under partial information. In this small data regime, the vast majority of state-of-the-art ANN techniques are lacking robustness and fail to provide any guarantees of convergence [115]. Therefore, unsupervised learning methods for complex mathematical model (PDEs) will become the focus and difficulty of future research. Recently, Raissi et al. [111] introduced physics-informed neural networks, which are data-driven algorithms for inferring solutions to general nonlinear partial differential equations. Results showcase a series of promising results for a diverse collection of problems in computational science, which open a new path for constructing surrogate models for mathematical physics under zero sample circumstance.

9. Model Uncertainty Analysis

Nowadays, model uncertainty has become one of the most important problems in both academia and industry [116]. The modeling of complex environmental stochastic systems is a difficult task. Because of the complexity of reality, the incompleteness of information, and the limitation of cognition, the mathematical model will neglect some influence factors and can only approximate the real behavior at a certain level of accuracy [7]. While scientists have made substantial progress in getting a better insight into environmental relationships and changes, model uncertainties are still a major obstacle for the predictive capability of simulations [11]. There are many ways in which model uncertainty may occur: the structure of the

model may be misspecified a priori, or the model may be identified incorrectly from the data, etc. [117]. All these situations could cause serious problems. For example, in the simulation of turbulent flow, since it is often not possible to know beforehand if one or more of such flow features will be present in a new flow configuration, predictions based on the Reynolds averaged Navier–Stokes equations are flawed by a structural (i.e., model-form) uncertainty [11]. Here we introduce three main approaches to remove or reduce structural uncertainty.

9.1. Model Correction. Here we first consider the fact that the observed quantities differ from the true ones by the experimental (observational) noise, which may be expressed through the relation:

$$z = \zeta + \epsilon, \quad (36)$$

where ζ is the true value for z and ϵ a random vector representative of the experimental noise. The experimental data noise ϵ is often assumed to be independently distributed without spatial correlation, and it is modeled as a Gaussian process with diagonal covariance matrix.

Theoretically, the true value for ζ could be obtained as an output of the model y , once a suitable set of parameters \mathbf{x} has been identified, i.e., $\zeta = y(\mathbf{x})$. In practice, however, few models are perfect [11], even if there is no parameter uncertainty, i.e., the predicted value will not equal the true value of the process [118]. The discrepancy is due to model inadequacy. A general framework to include the model inadequacy term in the stochastic model was first proposed in Kennedy and O’Hagan [119]. Model discrepancy can be taken into account by introducing an additional error term to the statistical model as in equation (36), which could be of additive nature, i.e.,

$$z = \zeta + \epsilon = y(\mathbf{x}) + \eta + \epsilon, \quad (37)$$

or of multiplicative nature:

$$z = \zeta + \epsilon = \eta \circ y(\mathbf{x}) + \epsilon. \quad (38)$$

The symbol \circ denotes the Hadamard (element-wise) multiplication. The choice of model inadequacy formulation largely depends on the nature and prior knowledge about the observed quantity. η is a random field representative of the model inadequacy. In several cases, η may involve additional parameters introduced for describing the error behavior, referred to as hyperparameters, which can be estimated based on likelihood maximization criteria or calibrated from the data along with the physical model parameters. When an additive model inadequacy term is used, it becomes difficult to separate its effect from that of the observational error. As for the multiplicative one, if the competing multiplicative statistical models describe the inadequacy term as Gaussian process, the observations can also be modeled as a Gaussian process based on the δ -method [120].

Although the use of model inadequacy terms is helpful in alleviating parameter overfitting problems, the approach suffers from several limitations: the correction terms are specific to

the observed QoI and depend on the spatial distribution of the observed data for the specific scenario [11, 121].

9.2. Model Selection. Model selection is to choose the best model in some statistical sense among a class of competing models. This consists in providing estimates of the posterior probability of each model in the considered set of models $\{\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_I\}$ given the observed data. The “model” here should be interpreted in a broader sense, including not only physical models with associated coefficients but also statistical models. Each model \mathcal{M}_k consists of a family of distributions $p(\mathcal{D} | \mathbf{x}_k, \mathcal{M}_k)$ indexed by \mathbf{x}_k . For such setups, the Bayesian approach provides a natural and general probabilistic framework that simultaneously treats both model and parameter uncertainty. Coupled with the advent of MCMC methods for posterior computation, the development and application of Bayesian methods for model uncertainty have seen remarkable evolution over the past decade [122].

The comprehensive Bayesian approach for multiple model setups proceeds by assigning a prior probability distribution $p(\mathbf{x}_k | \mathcal{M}_k)$ to the parameters of each model and a prior probability $p(\mathcal{M}_k)$ to each model. Margining out the parameters \mathbf{x} and conditioning on the data \mathcal{D} yields the posterior conditional model probabilities:

$$p(\mathcal{M}_k | \mathcal{D}) = \frac{p(\mathcal{D} | \mathcal{M}_k)p(\mathcal{M}_k)}{\sum_k p(\mathcal{D} | \mathcal{M}_k)p(\mathcal{M}_k)}, \quad (39)$$

where

$$p(\mathcal{D} | \mathcal{M}_k) = \int p(\mathcal{D} | \mathbf{x}_k, \mathcal{M}_k)p(\mathbf{x}_k | \mathcal{M}_k)d\mathbf{x}_k, \quad (40)$$

is the marginal likelihood of \mathcal{M}_k . Based on these posterior probabilities, pairwise comparison of models is summarized by the posterior odds

$$\frac{p(\mathcal{M}_k | \mathcal{D})}{p(\mathcal{M}_j | \mathcal{D})} = \frac{p(\mathcal{D} | \mathcal{M}_k)}{p(\mathcal{D} | \mathcal{M}_j)} \times \frac{p(\mathcal{M}_k)}{p(\mathcal{M}_j)}. \quad (41)$$

Insofar as the priors provide an initial representation of model uncertainty, the model posterior $p(\mathcal{M}_k | \mathcal{D})$ provides a complete representation of postdata model uncertainty that can be used for a variety of inferences and decisions. By treating $p(\mathcal{M}_k | \mathcal{D})$ as a measure of the “truth” of model \mathcal{M}_k , a natural and simple strategy for model selection is to choose the most probable \mathcal{M}_k , the modal model for which $p(\mathcal{M}_k | \mathcal{D})$ is largest.

However, the drawbacks of this approach exist. The selection of one particular model may lead to riskier decisions. In other words, if we choose a wrong model, the consequence will be disastrous. Moral-Benito [123] already pointed out the concern. From a pure empirical viewpoint, model uncertainty represents a concern because estimates may well depend on the particular model considered. Therefore, combining multiple models to reduce the model uncertainty is very desirable.

9.3. Model Averaging. The difficulty in making predictions with a single calibrated model clearly calls for a framework based on multimodel ensembles. Multimodel approaches have been used in aerodynamics [124] and many other applications [125, 126]. Bayesian modeling averaging is among the most widely used multimodel approaches, which combines multiple models enabling researchers to draw conclusions based on the whole universe of candidate models. Rather than choosing a single best model, the analyst may wish to utilize several models that are thought to be plausible a priori, or that seem to provide a reasonable approximation to the given data for the required objective. Prior knowledge is used to select a set of plausible models, and prior probabilities are attached to them. The data are then used to evaluate posterior probabilities for the different models, after which models with low posterior probabilities may be discarded to keep the problem manageable. Finally, a weighted sum of the predictions from the remaining competing models is calculated. There are two different approaches to model averaging in the literature, including frequentist model averaging (FMA) and Bayesian model averaging (BMA) [116].

Frequentist approaches focus on improving prediction and use weighted mean of estimates from different models while Bayesian approaches focus on the probability that a model is true and consider priors and posteriors for different models. The FMA approach does not consider priors, so the corresponding estimators depend solely on data. For its simplicity, the FMA approach has received some attention over the last decade [127].

The BMA approach includes a posterior of the predicted quantity ψ :

$$p(\psi | \mathcal{D}, \mathcal{M}) = \sum_{k=1}^I p(\psi | \mathcal{M}_k) p(\mathcal{M}_k | \mathcal{D}), \quad (42)$$

given calibration data \mathcal{D} and a set of models. In this framework, the posterior of ψ is an average of I posterior predictive distributions corresponding to I competing models weighted by their respective model posterior. The Bayesian approaches have the advantage of using arbitrary domain knowledge through a proper prior. However, how to set prior probabilities and how to deal with the priors when they are in conflict with each other are still problems [128]. The probably approximately correct- (PAC-) Bayes theory, first proposed by McAllester [129], combines the performance measurement of PAC, and hence, it can make full use of prior information, provide the tightest generalization error boundary for various learning algorithms, and evaluate the generalization performance of learning algorithms [130]. Generally speaking, compared to model selection methods, the result of an averaged model will not be as good as the (a priori unknown) best model but will not be as bad as the worst one [116].

10. Conclusion

At present, in the complex physical and engineering problems, M&S research has gradually developed a mature system, but also encountered new challenges. The

uncertainties introduced by many different sources become a major obstacle for the predictive capability and the reliability of simulations. The research of UQ aims to make better decisions, reduce the cost of trial and error during code development, and improve the reliability of simulation through identifying the main source of uncertainty, analyzing how the uncertainty propagates, searching for stable optimized solutions, and so on.

This paper gives a comprehensive review on the goals, ideas, and principle methods for each of the UQ processes. According to the data flow, the research of UQ can be divided into two categories: forward analysis concerning about the uncertainty propagation from input to output and backward analysis concerning about how to obtain the inference of input from experiment data and simulation output. In addition, as an aid, experimental design and surrogate models are introduced to improve the efficiency of UQ. In practice, most of the processes of UQ are related to and dependent on each other. For example, the good analysis or approximation requires representative samples, while the process of adaptive sampling depends on a good surrogate model and a sufficient understanding of the input, which needs to be obtained through sensitivity analysis, Bayes inference, etc. Last but not least, there comes a new trend that machine learning methods, which are suitable for the analysis of big data (e.g., parameter selection and high-dimensional approximation), are found to have great potential in the UQ research for complex physical processes, and will probably become an important research direction in the future.

Finally, it should be noted that this paper aims to highlight the main conceptual ideas of the whole system of UQ research and focus on innovative concepts, and hence, some details of the methods are intentionally/unintentionally ignored to avoid the impact on the integrity and coherence of the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by National Numerical Wind Tunnel Project (Grant no. NNW2019ZT7-A13), Science Challenge Project (Grant no. TZ2018001), and the Development Program for Defense Ministry of China (Grant no. C1520110002).

References

- [1] A. Gsponer, "Fourth generation nuclear weapons: military effectiveness and collateral effects," 2005, <http://arxiv.org/abs/0510071>.
- [2] M. Ghoreyshi and R. M. Cummings, "Unsteady aerodynamics modeling for aircraft maneuvers: a new approach using time-dependent surrogate modeling," *Aerospace Science and Technology*, vol. 39, pp. 222–242, 2014.

- [3] R.-L. Wang, Q. Liu, and W.-Z. Wen, "Uncertainty quantification for JWL EOS parameters using non-intrusive polynomial chaos (in Chinese)," *Explosion and Shock Waves*, vol. 35, no. 1, pp. 9–15, 2015.
- [4] J. A. Gaffney, S. T. Brandon, K. D. Humbird et al., "Making inertial confinement fusion models more predictive," *Physics of Plasmas*, vol. 26, no. 8, Article ID 082704, 2019.
- [5] C. M. Lindsay, G. C. Butler, C. G. Rumchik, B. Schulze, R. Gustafson, and W. R. Maines, "Increasing the utility of the copper cylinder expansion test," *Propellants, Explosives, Pyrotechnics*, vol. 35, no. 5, pp. 433–439, 2010.
- [6] J. S. Carson, "Introduction to modeling and simulation," in *Proceedings of the 2005 Winter Simulation Conference*, pp. 8–19, Orlando, FL, USA, December 2005.
- [7] R. Schöbi, *Surrogate models for uncertainty quantification in the context of imprecise probability modelling*, PhD Thesis, ETH Zürich, Zürich, Switzerland, 2017.
- [8] R. Allan, "Creating and using research data," in *Virtual Research Environments*, R. Allan, Ed., pp. 39–54, Chandos Publishing, Cambridge, UK, 2009.
- [9] C. Ortega Absil, A. Riccardi, M. Vasile, and C. Tardioli, "SMART-UQ: uncertainty quantification toolbox for generalized intrusive and non-intrusive polynomial algebra," in *Proceedings of the 6th International Conference on Astrodynamics Tools and Techniques*, Darmstadt, Germany, March 2016.
- [10] S. Basu, "Evaluation of hazard and risk analysis," in *Plant Hazard Analysis and Safety Instrumentation Systems*, S. Basu, Ed., Academic Press, Cambridge, MA, USA, 2017.
- [11] H. Xiao and P. Cinnella, "Quantification of model uncertainty in RANS simulations: a review," *Progress in Aerospace Sciences*, vol. 108, pp. 1–31, 2019.
- [12] D. C. Cox and P. Baybutt, "Methods for uncertainty analysis: a comparative survey," *Risk Analysis*, vol. 1, no. 4, pp. 251–258, 1981.
- [13] D. Bose, M. J. Wright, and G. E. Palmer, "Uncertainty analysis of laminar aeroheating predictions for mars entries," *Journal of Thermophysics and Heat Transfer*, vol. 20, no. 4, pp. 652–662, 2006.
- [14] R. N. Pérez, J. E. Amaro, and E. R. Arriola, "Uncertainty quantification of effective nuclear interactions," *International Journal of Modern Physics E*, vol. 25, no. 5, Article ID 1641009, 2016.
- [15] P. Beran, B. Stanford, and C. Schrock, "Uncertainty quantification in aeroelasticity," *Annual Review of Fluid Mechanics*, vol. 49, no. 1, pp. 361–386, 2017.
- [16] B. Yuepeng, S. Wenping, H. Zhonghua, Y. Zhang, and L. Zhang, "Aerodynamic/aeroacoustic variable-fidelity optimization of helicopter rotor based on hierarchical kriging model," *Chinese Journal of Aeronautics*, vol. 33, no. 2, pp. 476–492, 2020.
- [17] H. Abdo, J.-M. Flaus, and F. Masse, "Uncertainty quantification in risk assessment-representation, propagation and treatment approaches: application to atmospheric dispersion modeling," *Journal of Loss Prevention in the Process Industries*, vol. 49, pp. 551–571, 2017.
- [18] R. Wang and S. Jiang, "Mathematical methods for uncertainty quantification in non linear multi-physics systems and their numerical simulations," *SCIENTIA SINICA Mathematica*, vol. 45, no. 6, pp. 723–738, 2015.
- [19] X. Wu and T. Kozlowski, "Inverse uncertainty quantification of reactor simulations under the Bayesian framework using surrogate models constructed by polynomial chaos expansion," *Nuclear Engineering and Design*, vol. 313, pp. 29–52, 2017.
- [20] W. L. Oberkampf and T. G. Trucano, "Verification and validation in computational fluid dynamics," *Progress in Aerospace Sciences*, vol. 38, no. 3, pp. 209–272, 2002.
- [21] C. Unal, B. Williams, F. Hemez, S. H. Atamturktur, and P. McClure, "Improved best estimate plus uncertainty methodology, including advanced validation concepts, to license evolving nuclear reactors," *Nuclear Engineering and Design*, vol. 241, no. 5, pp. 1813–1833, 2011.
- [22] C. R. Rao, "Some combinatorial problems of arrays and applications to design of experiments," in *A Survey of Combinatorial Theory*, J. N. Srivastava, Ed., pp. 349–359, North-Holland, Amsterdam, Netherlands, 1973.
- [23] D. Cacuci, *Sensitivity and Uncertainty Analysis, Volume I: Theory*, CRC Press, Boca Raton, FL, USA, 2003.
- [24] M. Alam, V. Abedi, J. Bassaganya-Riera et al., "Agent-based modeling and high performance computing," in *Computational Immunology*, J. B. Riera, Ed., Academic Press, Cambridge, MA, USA, 2016.
- [25] H. Christopher Frey and S. R. Patil, "Identification and review of sensitivity analysis methods," *Risk Analysis*, vol. 22, no. 3, pp. 553–578, 2002.
- [26] B. Iooss and A. Saltelli, *Introduction to Sensitivity Analysis*, Springer, Berlin, Germany, 2017.
- [27] A. M. Dunker, "The decoupled direct method for calculating sensitivity coefficients in chemical kinetics," *The Journal of Chemical Physics*, vol. 81, no. 5, pp. 2385–2393, 1984.
- [28] J. T. Hwang, E. P. Dougherty, S. Rabitz, and H. Rabitz, "The Green's function method of sensitivity analysis in chemical kinetics," *The Journal of Chemical Physics*, vol. 69, no. 11, pp. 5180–5191, 1978.
- [29] M. A. Kramer, J. M. Calo, and H. Rabitz, "An improved computational method for sensitivity analysis: Green's function method with 'AIM'," *Applied Mathematical Modelling*, vol. 5, no. 6, pp. 432–441, 1981.
- [30] M. Ionescu-Bujor and D. G. Cacuci, "A comparative review of sensitivity and uncertainty analysis of large-scale systems-I: deterministic methods," *Nuclear Science and Engineering*, vol. 147, no. 3, pp. 189–203, 2004.
- [31] M. L. Williams, B. L. Broadhead, and C. V. Parks, "Eigenvalue sensitivity theory for resonance-shielded cross sections," *Nuclear Science and Engineering*, vol. 138, no. 2, pp. 177–191, 2001.
- [32] E. Greenspan, Y. Karni, and D. Gilai, "High order effects in cross section sensitivity analysis," Technical report no. 780858-3, Ben-Gurion Univ. of the Negev, Dept. of Nuclear Engineering, Beersheba, Israel, 1978.
- [33] M. Dion and G. Marleau, "Resonance self-shielding effects on eigenvalue sensitivity," in *Proceedings of the 2013 International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering-M and C*, Sun Valley, Idaho, May 2013.
- [34] M. Kimura and T. Kitada, "Evaluation of indirect effect in sensitivity coefficients by SAINT-II," *Transactions of the American Nuclear Society*, vol. 106, no. 1, pp. 757–758, 2012.
- [35] Y. Liu, L. Cao, H. Wu, T. Zu, and W. Shen, "Eigenvalue implicit sensitivity and uncertainty analysis with the subgroup resonance-calculation method," *Annals of Nuclear Energy*, vol. 79, pp. 18–26, 2015.
- [36] X.-Y. Zhang, M. Trame, L. Lesko, and S. Schmidt, "Sobol sensitivity analysis: a tool to guide the development and evaluation of systems pharmacology models," *CPT*:

- Pharmacometrics & Systems Pharmacology*, vol. 4, no. 2, pp. 69–79, 2015.
- [37] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, “Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index,” *Computer Physics Communications*, vol. 181, no. 2, pp. 259–270, 2010.
- [38] X. Wang, *Global Sensitivity Analysis for High-Dimensional Complex Models (In Chinese)*, PhD Thesis, East China Normal University, Shanghai, China, 2012.
- [39] I. M. Sobol, “Sensitivity analysis for non linear mathematical model,” *Mathematical and Computer Modelling*, vol. 1, pp. 407–414, 1993.
- [40] T. Homma and A. Saltelli, “Importance measures in global sensitivity analysis of nonlinear models,” *Reliability Engineering & System Safety*, vol. 52, no. 1, pp. 1–17, 1996.
- [41] E. Borgonovo and B. Iooss, *Moment-Independent and Reliability-Based Importance Measures*, Springer, Cham, Switzerland, 2016.
- [42] M. Baucells and E. Borgonovo, “Invariant probabilistic sensitivity analysis,” *Management Science*, vol. 59, no. 11, pp. 2536–2549, 2013.
- [43] D. Cacuci, M. Ionescu-Bujor, and I. Navon, *Sensitivity and Uncertainty Analysis, Volume II: Applications to Large-Scale Systems*, CRC Press, Boca Raton, FL, USA, 2005.
- [44] A. D. Kiureghian and M. D. Stefano, “Efficient algorithm for second-order reliability analysis,” *Journal of Engineering Mechanics*, vol. 117, no. 12, pp. 2904–2923, 1991.
- [45] I. Hughes and T. Hase, *Measurements and Their Uncertainties: A Practical Guide to Modern Error Analysis*, Oxford University Press, Oxford, UK, 2010.
- [46] J. Liu, X. Meng, C. Xu, D. Zhang, and C. Jiang, “Forward and inverse structural uncertainty propagations under stochastic variables with arbitrary probability distributions,” *Computer Methods in Applied Mechanics and Engineering*, vol. 342, pp. 287–320, 2018.
- [47] C. Wang and H. G. Matthies, “Non-probabilistic interval process model and method for uncertainty analysis of transient heat transfer problem,” *International Journal of Thermal Sciences*, vol. 144, pp. 147–157, 2019.
- [48] H. Agarwal, J. E. Renaud, E. L. Preston, and D. Padmanabhan, “Uncertainty quantification using evidence theory in multidisciplinary design optimization,” *Reliability Engineering & System Safety*, vol. 85, no. 1–3, pp. 281–294, 2004.
- [49] R. Schöbi and B. Sudret, “Uncertainty propagation of p-boxes using sparse polynomial chaos expansions,” *Journal of Computational Physics*, vol. 339, pp. 307–327, 2017.
- [50] H. Zhang, H. Dai, M. Beer, and W. Wang, “Structural reliability analysis on the basis of small samples: an interval quasi-monte carlo method,” *Mechanical Systems and Signal Processing*, vol. 37, no. 1–2, pp. 137–151, 2013.
- [51] M. Oberguggenberger and W. Fellin, “Reliability bounds through random sets: non-parametric methods and geotechnical applications,” *Computers & Structures*, vol. 86, no. 10, pp. 1093–1101, 2008.
- [52] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer Science & Business Media, Berlin, Germany, 2013.
- [53] C. Wang, Z. Qiu, X. Wang, and D. Wu, “Interval finite element analysis and reliability-based optimization of coupled structural-acoustic system with uncertain parameters,” *Finite Elements in Analysis and Design*, vol. 91, pp. 108–114, 2014.
- [54] S. H. Chen and J. Wu, “Interval optimization of dynamic response for structures with interval parameters,” *Computers & Structures*, vol. 82, no. 1, pp. 1–11, 2004.
- [55] Z. Qiu, Y. Xia, and J. Yang, “The static displacement and the stress analysis of structures with bounded uncertainties using the vertex solution theorem,” *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 49–52, pp. 4965–4984, 2007.
- [56] K. Campbell, “Statistical calibration of computer simulations,” *Reliability Engineering & System Safety*, vol. 91, no. 10–11, pp. 1358–1363, 2006.
- [57] X. Wu, T. Kozłowski, H. Meidani, and K. Shirvan, “Inverse uncertainty quantification using the modular Bayesian approach based on Gaussian process, part 1: theory,” *Nuclear Engineering and Design*, vol. 335, pp. 339–355, 2018.
- [58] E. L. Lee, H. C. Hornig, and J. W. Kury, “Adiabatic expansion of high explosive detonation products,” Tech. Rep. UCRL-505422, Lawrence Radiation Laboratory, Berkeley, CA, USA, 1968.
- [59] P. C. Souers, B. Wu, and L. C. J. Haselman, “Detonation equation of state at LLNL,” Tech. Rep. UCRL-ID119262 Rev 3, Lawrence Livermore National Laboratory, Livermore, CA, USA, 1995.
- [60] M. J. O’Sullivan and J. P. O’Sullivan, “Reservoir modeling and simulation for geothermal resource characterization and evaluation,” in *Geothermal Power Generation* Elsevier, Amsterdam, Netherlands, 2016.
- [61] D. K. Hale, *System and Method for Automated Model Calibration, Sensitivity Analysis, and Optimization*, U.S. Patent no. 10509873, 2019.
- [62] X. Wu, T. Kozłowski, H. Meidani, and K. Shirvan, “Inverse uncertainty quantification using the modular Bayesian approach based on Gaussian process, part 2: application to trace,” *Nuclear Engineering and Design*, vol. 335, pp. 417–431, 2018.
- [63] X. Jiang and S. Mahadevan, “Bayesian hierarchical uncertainty quantification by structural equation modeling,” *International Journal for Numerical Methods in Engineering*, vol. 80, no. 6–7, pp. 717–737, 2009.
- [64] X. Jiang and S. Mahadevan, “Bayesian structural equation modeling method for hierarchical model validation,” *Reliability Engineering & System Safety*, vol. 94, no. 4, pp. 796–809, 2009.
- [65] P. D. Congdon, *Applied Bayesian Hierarchical Methods*, CRC Press, Boca Raton, FL, USA, 2010.
- [66] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, CRC Press, Boca Raton, FL, USA, 2013.
- [67] S. Wu, P. Angelikopoulos, C. Papadimitriou, R. Moser, and P. Koumoutsakos, “A hierarchical Bayesian framework for force field selection in molecular dynamics simulations,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2060, Article ID 20150032, 2016.
- [68] I. Behmanesh, B. Moaveni, G. Lombaert, and C. Papadimitriou, “Hierarchical Bayesian model updating for structural identification,” *Mechanical Systems and Signal Processing*, vol. 64–65, pp. 360–376, 2015.
- [69] J. B. Nagel and B. Sudret, “A unified framework for multilevel uncertainty quantification in Bayesian inverse problems,” *Probabilistic Engineering Mechanics*, vol. 43, pp. 68–84, 2016.
- [70] O. Sedehi, C. Papadimitriou, and L. S. Katafygiotis, “Data-driven uncertainty quantification and propagation in structural dynamics through a hierarchical Bayesian

- framework,” *Probabilistic Engineering Mechanics*, vol. 60, Article ID 103047, 2020.
- [71] R. Shrestha and T. Kozłowski, “Inverse uncertainty quantification of input model parameters for thermal-hydraulics simulations using expectation-maximization under Bayesian framework,” *Journal of Applied Statistics*, vol. 43, no. 6, pp. 1011–1026, 2016.
- [72] J. Ray, S. Lefantzi, S. Arunajatesan, and L. Dechant, “Bayesian parameter estimation of a k - ϵ model for accurate jet-in-crossflow simulations,” *AIAA Journal*, vol. 54, no. 8, pp. 2432–2448, 2016.
- [73] V. Mons, J.-C. Chassaing, T. Gomez, and P. Sagaut, “Reconstruction of unsteady viscous flows using data assimilation schemes,” *Journal of Computational Physics*, vol. 316, pp. 255–280, 2016.
- [74] M. A. Iglesias, K. J. Law, and A. M. Stuart, “Ensemble Kalman methods for inverse problems,” *Inverse Problems*, vol. 29, no. 4, Article ID 045001, 2013.
- [75] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter*, Springer Science & Business Media, Berlin, Germany, 2009.
- [76] C. Schillings and A. M. Stuart, “Analysis of the ensemble Kalman filter for inverse problems,” *SIAM Journal on Numerical Analysis*, vol. 55, no. 3, pp. 1264–1290, 2017.
- [77] R. Schöbi, P. Kersaudy, B. Sudret, and J. Wiart, “Combining polynomial chaos expansions and kriging,” Tech. Rep. RSUQ-2014-001, Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich, Zurich, Switzerland, 2014.
- [78] A. B. Owen, “Orthogonal arrays for computer experiments, integration and visualization,” *Statistica Sinica*, pp. 439–452, 1992.
- [79] L. Ji, Y. Si, H. Liu, X. Song, W. Zhu, and A. Zhu, “Application of orthogonal experimental design in synthesis of mesoporous bioactive glass,” *Microporous and Mesoporous Materials*, vol. 184, pp. 122–126, 2014.
- [80] L. Song, H.-Z. Liu, and Z.-X. Yang, “Orthogonal analysis based performance optimization for vertical axis wind turbine,” *Mathematical Problems in Engineering*, vol. 2016, Article ID 6241360, 2016.
- [81] M. D. McKay, R. J. Beckman, and W. J. Conover, “Comparison of three methods for selecting values of input variables in the analysis of output from a computer code,” *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.
- [82] L. Chrisman, *Latin Hypercube vs. Monte Carlo Sampling*, 2014, <https://lumina.com/latin-hypercube-vs-monte-carlo-sampling/>.
- [83] Z. Sui, L. Cao, C. Wan, and X. Zou, “Covariance-oriented sample transformation: a new sampling method for reactor-physics uncertainty analysis,” *Annals of Nuclear Energy*, vol. 134, pp. 452–463, 2019.
- [84] R. Sheikholeslami and S. Razavi, “Progressive Latin hypercube sampling: an efficient approach for robust sampling-based analysis of environmental models,” *Environmental Modelling & Software*, vol. 93, pp. 109–126, 2017.
- [85] B. M. Adams, W. Bohnhoff, K. Dalbey et al., “Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis,” Tech. Rep. SAND2014-4253, Sandia National Laboratories, Livermore, CA, USA, 2014.
- [86] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Hoboken, NJ, USA, 2012.
- [87] R. K. Tripathy and I. Billionis, “Deep UQ: learning deep neural network surrogate models for high dimensional uncertainty quantification,” *Journal of Computational Physics*, vol. 375, pp. 565–588, 2018.
- [88] A. Bhosekar and M. Ierapetritou, “Advances in surrogate based modeling, feasibility analysis, and optimization: a review,” *Computers & Chemical Engineering*, vol. 108, pp. 250–267, 2018.
- [89] D. Lu and D. Ricciuto, “Efficient surrogate modeling methods for large-scale earth system models based on machine learning techniques,” 2019, <http://arxiv.org/abs/1901.05125>.
- [90] G. A. Mihram, “An efficient procedure for locating the optimal similar response,” Tech. Rep., 1970.
- [91] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, Princeton, NJ, USA, 2010.
- [92] R. Askey and J. A. Wilson, *Some Basic Hypergeometric Orthogonal Polynomials that Generalize Jacobi Polynomials*, American Mathematical Society, Providence, RI, USA, 1985.
- [93] D. Xiu and G. E. Karniadakis, “The Wiener-Askey polynomial chaos for stochastic differential equations,” *SIAM Journal on Scientific Computing*, vol. 24, no. 2, pp. 619–644, 2002.
- [94] D. Xiu and J. S. Hesthaven, “High-order collocation methods for differential equations with random inputs,” *SIAM Journal on Scientific Computing*, vol. 27, no. 3, pp. 1118–1139, 2005.
- [95] G. Blatman, *Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis*, PhD Thesis, Université De Clermont-Ferrand, Clermont-Ferrand, France, 2009.
- [96] G. Heinze, C. Wallisch, and D. Dunkler, “Variable selection—a review and recommendations for the practicing statistician,” *Biometrical Journal*, vol. 60, no. 3, pp. 431–449, 2018.
- [97] D. Krige, “A statistical approach to some basic mine valuation problems on the Witwatersrand,” *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, no. 6, pp. 119–139, 1951.
- [98] N. Le and J. Zidek, *Statistical Analysis of Environmental Space-Time Processes*, Springer Science & Business Media, Berlin, Germany, 2006.
- [99] C. Rasmussen and C. Williams, *Gaussian Process for Machine Learning*, MIT Press, Cambridge, MA, USA, 2005.
- [100] M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*, Springer Science & Business Media, Berlin, Germany, 2012.
- [101] V. R. Joseph, Y. Hung, and A. Sudjianto, “Blind kriging: a new method for developing metamodels,” *Journal of Mechanical Design*, vol. 130, pp. 03–11, 2008.
- [102] P. Benner, S. Gugercin, and K. Willcox, “A survey of projection-based model reduction methods for parametric dynamical systems,” *SIAM Review*, vol. 57, no. 4, pp. 483–531, 2015.
- [103] J. Yu, C. Yan, and M. Guo, “Non-intrusive reduced-order modeling for fluid problems: a brief review,” *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 233, no. 16, pp. 5896–5912, 2019.
- [104] M. S. Hemati, M. O. Williams, and C. W. Rowley, “Dynamic mode decomposition for large and streaming datasets,” *Physics of Fluids*, vol. 26, no. 11, Article ID 111701, 2014.
- [105] B. A. Belson, J. H. Tu, and C. W. Rowley, “Algorithm 945,” *ACM Transactions on Mathematical Software*, vol. 40, no. 4, pp. 1–23, 2014.
- [106] B. R. Noack, W. Stankiewicz, M. Morzyński, and P. J. Schmid, “Recursive dynamic mode decomposition of transient and post-transient wake flows,” *Journal of Fluid Mechanics*, vol. 809, pp. 843–872, 2016.

- [107] J. N. Kutz, X. Fu, and S. L. Brunton, "Multiresolution dynamic mode decomposition," *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 2, pp. 713–735, 2016.
- [108] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, "Chaos as an intermittently forced linear system," *Nature Communications*, vol. 8, no. 1, pp. 1–9, 2017.
- [109] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [110] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, "Multi-layer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Networks*, vol. 6, no. 6, pp. 861–867, 1993.
- [111] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
- [112] Y. Zhu, N. Zabararas, P.-S. Koutsourelakis, and P. Perdikaris, "Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data," *Journal of Computational Physics*, vol. 394, pp. 56–81, 2019.
- [113] N. Van Nguyen, J.-W. Lee, M. Tyan, and S. Kim, "Repetitively enhanced neural networks method for complex engineering design optimisation problems," *The Aeronautical Journal*, vol. 119, no. 1220, pp. 1253–1270, 2015.
- [114] M. A. A. Oroumieh, S. M. B. Malaek, M. Ashrafizaadeh, and S. M. Taheri, "Aircraft design cycle time reduction using artificial intelligence," *Aerospace Science and Technology*, vol. 26, no. 1, pp. 244–258, 2013.
- [115] G. Sun and S. Wang, "A review of the artificial neural network surrogate modeling in aerodynamic design," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 233, no. 16, pp. 5863–5872, 2019.
- [116] Y. Huang, W. Huang, L. Li, and Z. Li, *Meta-Learning PAC-Bayes Priors in Model Averaging*, AAAI, Menlo Park, CA, USA, 2019.
- [117] C. Chatfield, *Model Uncertainty*, American Cancer Society, Atlanta, GA, USA, 2006.
- [118] J. Brynjarsdóttir and A. O'Hagan, "Learning about physical parameters: the importance of model discrepancy," *Inverse Problems*, vol. 30, no. 11, Article ID 114007, 2014.
- [119] M. C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 3, pp. 425–464, 2001.
- [120] G. Casella and R. L. Berger, *Taylor Approximation and the Delta Method*, pp. 240–245, Duxbury Press, Pacific Grove, CA, USA, 2001.
- [121] W. N. Edeling, P. Cinnella, R. P. Dwight, and H. Bijl, "Bayesian estimates of parameter variability in the k - ϵ turbulence model," *Journal of Computational Physics*, vol. 258, pp. 73–94, 2014.
- [122] M. Clyde and E. I. George, "Model uncertainty," *Statistical Science*, pp. 81–94, 2004.
- [123] E. Moral-Benito, "Model averaging in economics: an overview," *Journal of Economic Surveys*, vol. 29, no. 1, pp. 46–75, 2015.
- [124] S. V. Poroseva, M. Y. Hussaini, and S. L. Woodruff, "Improving the predictive capability of turbulence models using evidence theory," *AIAA Journal*, vol. 44, no. 6, pp. 1220–1228, 2006.
- [125] T. Diomede, S. Davolio, C. Marsigli et al., "Discharge prediction based on multi-model precipitation forecasts," *Metereology and Atmospheric Physics*, vol. 101, no. 3–4, pp. 245–265, 2008.
- [126] Q. Duan, N. K. Ajami, X. Gao, and S. Sorooshian, "Multi-model ensemble hydrologic prediction using Bayesian model averaging," *Advances in Water Resources*, vol. 30, no. 5, pp. 1371–1386, 2007.
- [127] H. Wang, X. Zhang, and G. Zou, "Frequentist model averaging estimation: a review," *Journal of Systems Science and Complexity*, vol. 22, no. 4, pp. 732–748, 2009.
- [128] N. L. Hjort and G. Claeskens, "Frequentist model average estimators," *Journal of the American Statistical Association*, vol. 98, no. 464, pp. 879–899, 2003.
- [129] D. A. McAllester, "PAC-Bayesian model averaging," in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, pp. 164–170, Santa Cruz, CA, USA, July 1999.
- [130] T. Li, G. Xiujun, and H. Li, "Survey on PAC-Bayes theory and application research (in Chinese)," *Journal of Frontiers of Computer Science and Technology*, vol. 9, no. 1, pp. 1–13, 2015.