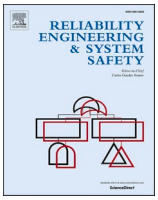
Reliability Engineering and System Safety 233 (2023) 109122

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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



Quantification and propagation of Aleatoric uncertainties in topological structures

Zihan Wang [a], Mohamad Daeipour [b], Hongyi Xu [a,*]

[a] Mechanical Engineering, University of Connecticut, Storrs, CT 06269, United States of America
[b] Materials Science & Engineering, University of Connecticut, Storrs, CT 06269, United States of America

ARTICLE INFO

Keywords:
Uncertainty quantification
Uncertainty propagation
Aleatoric uncertainty
Random field
Non-simply connected space
Topological structure

ABSTRACT

Quantification and propagation of aleatoric uncertainties distributed in complex topological structures remain a challenge. Existing uncertainty quantification and propagation approaches can only handle parametric uncertainties or high dimensional random quantities distributed in a simply connected spatial domain. There lacks a systematic method that captures the topological characteristics of the structural domain in uncertainty analysis. Therefore, this paper presents a new methodology that quantifies and propagates aleatoric uncertainties, such as the spatially varying local material properties and defects, distributed in a topological spatial domain. We propose a new random field-based uncertainty representation approach that captures the topological characteristics using the shortest interior path distance. Parameterization methods like PPCA and β-Variational Autoencoder (βVAE) are employed to convert the random field representation of uncertainty to a small set of independent random variables. Then non-intrusive uncertainties propagation methods such as polynomial chaos expansion and univariate dimension reduction are employed to propagate the parametric uncertainties to the output of the problem. The effectiveness of the proposed methodology is demonstrated by engineering case studies. The accuracy and computational efficiency of the proposed method is confirmed by comparing with the reference values of Monte Carlo simulations with a sufficiently large number of samples.

1. Introduction

Engineering structures with complex topological characteristics have excellent performances that cannot be achieved by traditional simple structure designs. The reliability and robustness of complex structural systems are influenced by uncertainties from various sources, such as the boundary conditions [1–3], assembly tolerance [4], dimensions and geometries [5,6], time dependencies [7–9], and the distribution of local material properties [6,10,11]. To represent those aleatoric uncertainties quantitatively, researchers proposed statistical distribution functions for parametric uncertainties [12–15] or random field models for high dimensional uncertainties distributed in a simply connected space [10]. However, quantifying and propagating the uncertainties in topological structure have remained a challenge. In literature, simplified yet inaccurate approaches have been proposed to represent aleatoric uncertainties distributed in a topological spatial domain. For example, one widely used approach is to first create a random field model of the uncertain quantity of interest (QoI) without considering the topological characteristics of the structure and then mask the random field realizations with the structure image ("dig holes" in the random realizations) [2,10]. Another approach is to divide the topological spatial domain into multiple simply connected sub-domains and create an independent random field model of the uncertain QoI for each sub-domain [16]. The main issue with this approach is the lack of continuity at the interface between sub-domains. In our previous work [17], we proposed a network distance-based method to represent the spatially correlated uncertainties in the lattice structure. This method captures the topological characteristics of the lattice structure while neglecting the variations in QoI along the diameter direction of each lattice bar.

In addition to capturing topological characteristics, another challenge is to reduce the dimensionality of the input uncertainties in order to improve the efficiency of uncertainty quantification. Lots of dimension reduction methods are proposed to parameterize the high-dimensional variables, like principal component analysis (PCA) and its variations [18–20], non-negative matrix factorization (NMF) [21], linear discriminate analysis (LDA) [22], and generalized discriminant analysis (GDA) [23]. With the advancements in neural networks, autoencoder (AE) draws increasing attention for its capability of

* Corresponding author.
E-mail address: hongyi.3.xu@uconn.edu (H. Xu).

https://doi.org/10.1016/j.ress.2023.109122
Received 9 June 2022; Received in revised form 18 January 2023; Accepted 21 January 2023
Available online 26 January 2023

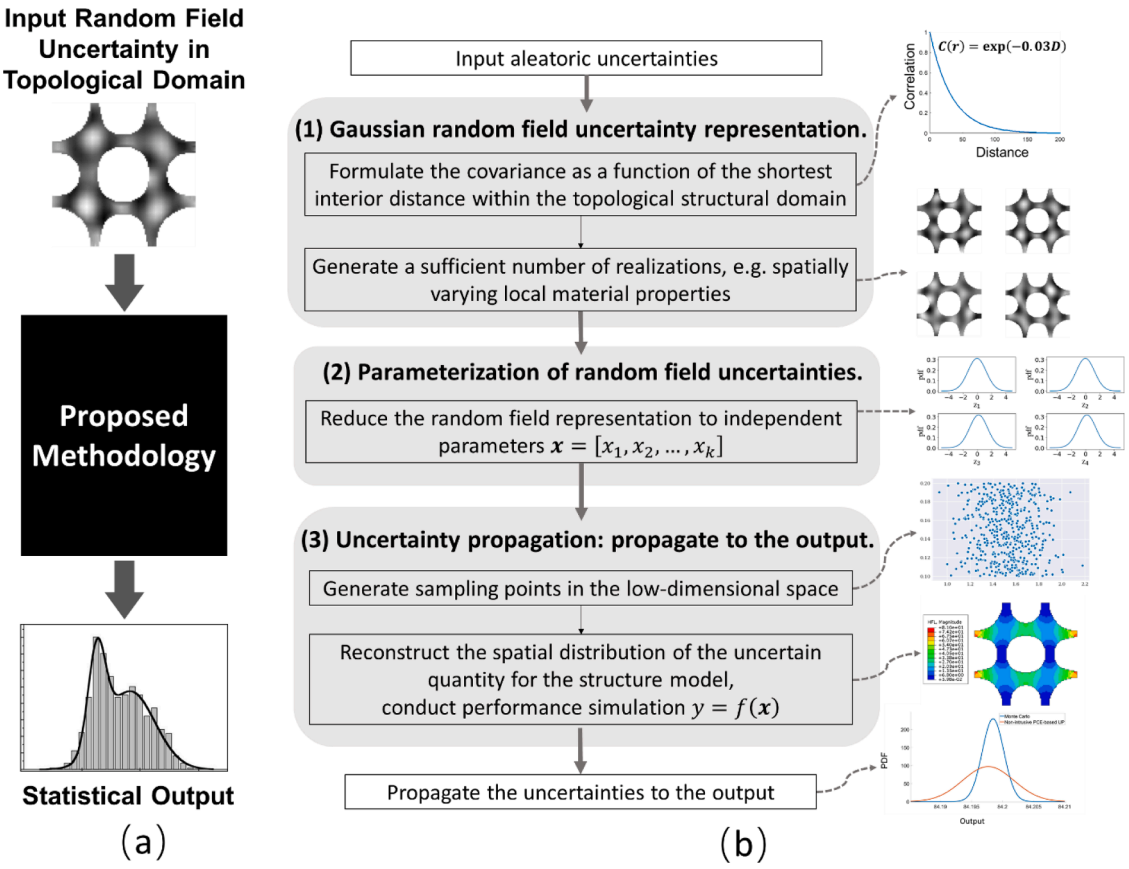


Fig. 1. (a) Problem Statement. (b) Overview of the proposed methodology for quantifying and propagating aleatoric uncertainties in a topological structural domain.

compressing data into a low-dimensional latent space [24]. Recently, various AE-based methods have been successfully implemented to reduce the dimensionality of high-dimensional data. Yant et al. [25] proposed a sparse AE-based deep neural network to perform fault detection in high-dimensionality multi-component systems. Health indicator built in the latent space has been proposed for condition monitoring of engineering systems [26]. Liu et al. [27] proposed a variational Autoencoder (VAE)-based dimensionality reduction method for the quantification of uncertainties in geophysical inverse problems.

Uncertainty propagation (UP) is a method for quantifying uncertainty in the system output based on random or noisy inputs. UP is of great importance for the reliability-based design and robust design. Major efforts have been made to develop UP methods for reliability and safety applications [28–30], including simulation-based methods [29–33], function expansion-based methods [29,31,34,35], most probable point-based methods [36,37], and numerical integration-based methods [38–40]. The simulation-based methods include Monte Carlo simulation (MCS), importance sampling, and adaptive sampling. MCS method, which is computationally expensive, is widely used to provide a reference value to validate other methods. Among function expansion-based methods, polynomial chaos expansion (PCE) is regarded as one of the most used methods due to its mathematically rigorous concept, strong theoretical basis, and ability to represent stochastic quantities [35]. PCE represents a random variable by a series of polynomial chaos basis, which allows the statistical moments of the system output to be estimated. Proposed by Ghanem and Spanos [35], the Hermite PCE has been applied to a variety of engineering fields [41–45]. In the most probable point-based method, first-order reliability method [46] and second-order reliability method [47] are the two most popular approaches. The drawback of first-order reliability method and second-order reliability method is that the information regarding variable distribution is completely ignored, so they work well only when the limit state is linear or quadratic and the variables are normally distributed. The numerical integration-based methods derive the probability distribution function (PDF) of performance function, and the probability is evaluated by a simple one-dimensional integration over PDF. Univariate dimension reduction (UDR) method is a numerical integration-based UP method, which approximates a multidimensional moment integral through multiple reduced-dimensional integrals based on additive decomposition of the performance function [40]. Due to its high efficiency, UDR has been widely used in reliability-based design and robust design [48–51]. The aforementioned UP methods propagate parametric uncertainties to the output.

The goal of this work is to establish a new methodology that quantifies and propagates the uncertainties distributed in a topological structural domain. We propose a new covariance function of Gaussian random field, which captures the topological characteristics using the shortest interior path distance between two sampling points in a structural domain. Dimension reduction methods, PPCA and βVAE, are employed to convert the high-dimensional random field uncertainties into independent, normally distributed random parameters. Then non-intrusive PCE and UDR are employed to propagate the parameterized uncertainties to the output. The effectiveness of the proposed method is validated by two case studies.

The main contributions of this work are summarized as follows:

(1) A new shortest interior path distance-based uncertainty representation method is proposed. Compared with the traditional Euclidean distance-based method, the proposed method can better represent the spatially correlated uncertainties distributed in a topological spatial domain.
(2) A comparative study is presented to provide new insights on the methods that reduce the high-dimensional random field representation of uncertainties in a topological spatial domain to a

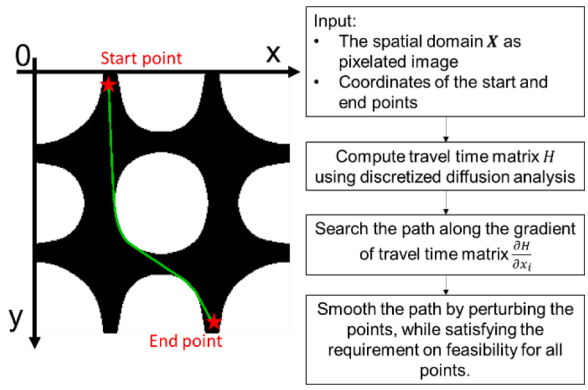


Fig. 2. Finding the shortest interior path in a topological feasible region.

low-dimensional representation of a few independent random variables.

(3) To the authors' best knowledge, it is the first UP methodology that propagates aleatoric uncertainties distributed in a topological spatial domain to quantify the uncertainties of the system responses.

The remainder of this paper is organized as follows. In Section 2, an overview of the proposed uncertainty quantification and propagation methodology is introduced. Details of each component of the framework are introduced in Section 3, 4, and 5. Section 3 introduces the statistical representation of high-dimensional uncertainties in a topological structural domain. In Section 4, we present two ways for parameterizing the high-dimensional uncertainties, PPCA and βVAE, and present a comparative study to understand their relative merits. In Section 5, we introduce the UP process using non-intrusive PCE and UDR. Section 6 includes two engineering case studies and a comparison of different methods. The conclusion of this work is provided in Section 7.

2. Overview of the proposed methodology

The overarching goal is to predict the statistical output for an engineering problem based on the statistical distribution of the input aleatoric uncertainties (Fig. 1(a)). Aleatoric uncertainty, also known as statistical uncertainty, is representative of unknowns that differ each time we run the same experiment/simulation. Examples of aleatoric uncertainties include manufacturing errors, inherent stochasticity in material microstructure and properties, uncertain working conditions, etc. Here we present a new methodology (Fig. 1(b)) that (i) provides a statistical representation of the high-dimensional aleatoric uncertainties distributed in a topological spatial domain, (ii) reduces the random field representation to independent random parameters, and (iii) predicts the statistics of the output. The steps for implementing the proposed methodology include a:

1 Gaussian random field uncertainty representation with the shortest interior path distance-based covariance function. When modeling the covariance between two sampling points, we propose to formulate the covariance as a function of the shortest interior path distance within the topological structural domain. In this way, the topological characteristics can be embodied in the random field model.

2 Parameterization of the random field uncertainties. Dimension reduction methods, PPCA and βVAE, are employed to convert the random field uncertainties to independent random parameters $\boldsymbol{x} = [x_1, x_2, \ldots, x_k]$, where k is the dimension of the parametric representation. Each parameter x_i is described by its PDF. A sufficiently large set of realizations (statistically representative) are generated using the Gaussian random field model established in the previous step and the parameterization is conducted on the realizations.

3 UP. Non-intrusive PCE and UDR are conducted to propagate the input uncertainties, which are in the form of independent random parameters, to the output of the problem. Compared to the MSC-based approaches [52,53], the proposed approach has the advantage of efficiency as it requires a smaller set of sample points, which are generated in the low-dimensional parametric space.

3. Statistical representation of Aleatoric uncertainties

A proper mathematical representation of the input aleatoric uncertainty is a stepping stone for the following uncertainty propagation studies. In this section, we firstly introduce the Gaussian random field for representing high-dimensional uncertainties, and then present a new formulation of the Gaussian random field covariance function for topological structural domains.

3.1. Gaussian random field for uncertainty representation

Gaussian-type random field is one of the most used uncertainty representation methods due to its simplicity, tractability, and flexibility. A Gaussian random field $R_G(\boldsymbol{x})$ can be completely characterized by its mean $E(\boldsymbol{x})$ and covariance function $\boldsymbol{C}(\boldsymbol{x})$, where $\boldsymbol{x}$ is an N-dimensional field variable:

$$R_G(\boldsymbol{x}) \sim \mathrm{MVN}(E(\boldsymbol{x}),\ \boldsymbol{C}(\boldsymbol{x})) \tag{1}$$

where "MVN" represents multivariate normal distribution. The covariance function $\boldsymbol{C}(\boldsymbol{x})$ gives the covariance evaluated at any two observation locations in the input space. In literature, exponential function-based kernel functions are widely used as the covariance function [10, 54-56]. Those functions capture the decaying correlation between two observation locations as the distance increases. In our previous work [57], oscillating function-based formulations are proposed to capture the negative correlations and fluctuating correlations along with distances. Several examples of covariance kernel functions are listed below:

A traditional exponential function-based kernel function:

$$C(\boldsymbol{x}_i,\ \boldsymbol{x}_j) = \sigma^2 \cdot \exp\left(-\frac{D(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta}\right) \tag{2}$$

Oscillating function-based kernel functions:

$$C(\boldsymbol{x}_i,\ \boldsymbol{x}_j) = \sigma^2 \cdot \exp\left(-\frac{D(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta_1}\right) \cdot \cos\left(\frac{D(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta_2}\right) \tag{3}$$

$$C(\boldsymbol{x}_i,\ \boldsymbol{x}_j) = \sigma^2 \cdot \left(1 + \frac{D(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta_1}\right)^{-1} \cdot \cos\left(\frac{D(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta_2}\right) \tag{4}$$

where σ is a calibration parameter that controls the variance; $D(\boldsymbol{x}_i,\ \boldsymbol{x}_j)$ represents the distance between the two sampling locations; θ, θ_1 and θ_2 are kernel function parameters. Traditionally, the distance is measured by the Euclidean distance $D(\boldsymbol{x}_i, \boldsymbol{x}_j) = \| \boldsymbol{x}_i, \boldsymbol{x}_j \|$.

The values of the unknown parameters in the covariance function are obtained by fitting to the random field realizations (observed data). First, the distance and covariance between any of the two sampling locations are evaluated. Then the covariance-distance dataset is sorted based on the distance value. For each distance value, the average covariance value is calculated. The distance-average covariance relationship is then used to fit the kernel function. Different types of kernel functions need to be compared to obtain the best fitting accuracy. It is also suggested to set a cutoff distance in order to capture the short-distance covariance better. The detailed process of fitting the covariance kernel function can be found in our previous work [57].

1.2. Shortest interior path distance between two sampling locations

Instead of constructing the covariance function based on the

Euclidean distance, we propose to establish the spatial relationship between two sampling locations based on the shortest interior path distance within the structural domain. As shown in Fig. 2, the topological structural domain is represented by the feasible region $\boldsymbol{X}$ in the 2D or 3D spatial domain. The shortest interior path distance $D_{SP}(\boldsymbol{x}_i,\ \boldsymbol{x}_j)$ between any two sampling locations $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ within the feasible region (within the structure) is evaluated for constructing the covariance function [58]. Taking the oscillating function-based covariance formulation shown in Eq. (5) as an example, the covariance between two sampling locations is calculated as:

$$C(\boldsymbol{x}_i,\ \boldsymbol{x}_j) = \sigma^2 \cdot \left(1 + \frac{D_{SP}(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta_1}\right)^{-1} \cdot \cos\left(\frac{D_{SP}(\boldsymbol{x}_i,\ \boldsymbol{x}_j)}{\theta_2}\right) \tag{5}$$

A three-step algorithm is employed to solve for the shortest interior path distance within a topological feasible region.

- First, the travel time matrix $\boldsymbol{H}$ is computed using discretized diffusion analysis [59].
- Second, the path is searched along the gradient of travel time matrix $\frac{\partial H}{\partial x_i}$. The algorithm rapidly propagates through all possible pathways to find a representative shortest route, which is recorded as a set of points in the feasible region.
- Third, the path is smoothed by optimization, which perturbs the locations of the points while satisfying the requirement of feasibility. A feasible route is one that is fully contained in the feasible region.

4. Parameterization of the high dimensionality input uncertainty

In this section, we introduce two dimension reduction methods that are employed to parameterize the high-dimensional random field uncertainties into low-dimensional parametric uncertainties. In addition to low dimensionality, orthogonality (statistical independence) is also desired, as it improves the data compression efficiency and facilitates identifying the key features from the data. Parameterization of random field uncertainties brings benefit to UP (Section 5), as the state-of-the-art methods work best on independent random variables as the input. Two methods, beta-variational autoencoder and probabilistic principal component analysis, are introduced and compared in this section.

4.1. Probabilistic principal component analysis (PPCA)

Probabilistic principal components analysis (PPCA) is a dimension reduction technique derived from principal component analysis (PCA) [60]. As a probabilistic version of PCA, PPCA is based on a Gaussian latent variable model that parameterizes the high dimensional input data to a low dimensional latent space [19]. PPCA overcomes some disadvantages of PCA, such as difficulty in assessing the fit quality [61], inability to reveal groups of subjects in the data [62,63], and inability to deal with noisy data and missing data [64].

Let $\boldsymbol{x}$ be a set of p-dimensional input variables (which is the input structure in our case) and $\boldsymbol{z}$ be the q-dimensional linear projection in the latent space that best represents the data in the least-squares sense. Noted that, each latent dimension of $\boldsymbol{z}$ is fully independent of others. For the traditional PCA, z_i can be seen as the principal score of subject i. PPCA can be written as:

$$\boldsymbol{x} = \boldsymbol{W}\boldsymbol{z} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_i \tag{6}$$

where $\boldsymbol{W}$ is $p \times q$ linear transformation matrix; $\boldsymbol{\mu}$ is the mean vector, where $\boldsymbol{\mu} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}$; $\boldsymbol{\varepsilon}_i$ is multivariate Gaussian for subject i. The dependency between $\boldsymbol{z}$ and $\boldsymbol{x}$ is linear. It is assumed that $p(\boldsymbol{\varepsilon}_i) = MVN_p(0, \sigma^2\boldsymbol{I})$, where $\boldsymbol{I}$ denotes the identity matrix. The latent variable $\boldsymbol{z}$ is assumed to be isotropic Gaussian, for example, $p(\boldsymbol{z}_i) = MVN_q(0,\boldsymbol{I})$. The conditional distribution of the input data given the latent variables can be written as:

$$p(\boldsymbol{x}_i|\boldsymbol{z}_i) = MVN_p\left(\boldsymbol{W}\boldsymbol{z}_i + \boldsymbol{\mu}, \sigma^2\boldsymbol{I}\right) \tag{7}$$

The distribution of $p(\boldsymbol{x}_i)$ is:

$$p(\boldsymbol{x}_i) = MVN_p\left(\boldsymbol{\mu}, \boldsymbol{W}\boldsymbol{W}^T + \sigma^2\boldsymbol{I}\right) \tag{8}$$

One can obtain the maximum-likelihood estimation of the parameters $\boldsymbol{\mu}$, $\boldsymbol{W}$ and σ^2 based on dataset $\boldsymbol{x}$. Any observed data point $\boldsymbol{x}_i$ can be represented in the latent space by its q-dimensional latent variable $\boldsymbol{z}$. Then the distribution of the latent variable on observed data can be expressed as:

$$p(\boldsymbol{z}_i|\boldsymbol{x}_i) = MVN_q\left(\boldsymbol{M}^{-1}\boldsymbol{W}^T(\boldsymbol{x}_i - \boldsymbol{\mu}), \sigma^2\boldsymbol{M}^{-1}\right) \tag{9}$$

where M is defined as $\boldsymbol{M} = \boldsymbol{W}\boldsymbol{W}^T + \sigma^2\boldsymbol{I}$. It is to be noted that, the output of PPCA is similar to that obtained by PCA [19], but PPCA offers an additional advantage of uncertainty assessment.

4.2. Beta-variational autoencoder (βVAE) and it with Gaussian mixture prior

4.2.1. Beta-variational autoencoder (βVAE)

A non-linear extension of the PPCA is the variational autoencoder (VAE) [65]. VAE is a generative machine learning model that aims at extracting the knowledge from the original space to construct a low dimensional latent space and also provides the capability of reconstructing the data in the original space [66].

βVAE, a variant of VAE, disentangles the latent features [67]. Disentanglement means that one latent feature is only sensitive to changes in one generative factor, while being relatively invariant to changes in other factors [68]. βVAE learns the mapping relationship between the input high dimensional data $\boldsymbol{x}$ (e.g. structure image) and the corresponding latent space $\boldsymbol{z}$.

βVAE consists of an encoder (recognition model) and a decoder (generative model). The encoder performs nonlinear dimensionality reduction and compresses the input data from the original high dimensional space into a low dimensional latent space. The encoder can be expressed as $Q_\phi(\boldsymbol{z}|\boldsymbol{x})$, which is the approximate posterior represented by a normal distribution. ϕ is the encoder model parameters. The decoder is used to reconstruct the high dimensional data by mapping the latent feature vector back to the original space. The generative model is expressed as:

$$P_\theta(\boldsymbol{x}, \boldsymbol{z}) = P_\theta(\boldsymbol{x}|\boldsymbol{z})P_\theta(\boldsymbol{z}) \tag{10}$$

where θ is the vector of parameters of the decoder, $P_\theta(\boldsymbol{z})$ is the prior distribution of $\boldsymbol{z}$, and $P_\theta(\boldsymbol{x}|\boldsymbol{z})$ is the conditional distribution of $\boldsymbol{x}$ on $\boldsymbol{z}$.

The loss of βVAE is written as:

$$L_{\beta VAE} = E_{Q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\log P_\theta(\boldsymbol{x}|\boldsymbol{z})] - \beta D_{KL}\left(Q_\phi(\boldsymbol{z}|\boldsymbol{x}) \parallel P_\theta(\boldsymbol{z})\right) \tag{11}$$

This loss function consists of two parts, the reconstruction loss and the Kullback-Leibler (KL) divergence loss. The hyperparameter β in the KL divergence loss controls the independency of the latent feature variables. When $\beta = 1$, the βVAE is the same as VAE. When $\beta > 1$, it applies a larger KL divergence loss, and the model is pushed to a more independent latent representation [67]. By tuning the coefficient β, we could obtain an approximately orthogonal latent feature representation of the data. In our engineering case studies (Section 6), the βVAE models generate latent features that are approximately orthogonal to each other, i.e. the correlation coefficient between two latent features is in the range of $[-0.1, 0.1]$ (the angles between two feature axes is in the range of $[85^\circ, 95^\circ]$).

4.2.2. Gaussian mixture beta variational autoencoder (GM-βVAE)

Different from the typical VAE and βVAE, the prior over each latent variable of GM-βVAE is modeled as a mixture of Gaussian, which

captures a complex multimodal statistical distribution. Each mode of the latent feature corresponds to a specific class of data. Each class obeys a Gaussian distribution. The distance between two classes of data is measured by the Mahalanobis distance [69].

GM-βVAE learns the relationship between the input image $\boldsymbol{x}$ and the corresponding latent space z. Moreover, a categorical variable $\boldsymbol{y}$ is introduced to identify which Gaussian distribution in the multimodal distribution of a latent feature each data point belongs to. The encoder (recognition model) of the GM-βVAE is expressed as:

$$Q_\phi(\boldsymbol{y},z|\boldsymbol{x}) = Q_\phi(\boldsymbol{y}|\boldsymbol{x})Q_\phi(z|\boldsymbol{x},\boldsymbol{y}) \tag{12}$$

where ϕ is the vector of the recognition model parameters and $Q_\phi(\boldsymbol{y},z|\boldsymbol{x})$ is the approximate posterior that is represented by a mixture of Gaussian distributions. Each latent feature consists of k distinct Gaussian distributions, i.e., $Q_\phi(z|\boldsymbol{x},\boldsymbol{y}_i)$, where $i \in 0,\ 1,\ 2,\ \ldots k-1$. Thus, the approximate posterior becomes a Gaussian mixture. The means and variances of the Gaussian distributions are learned by the inference model.

The decoder of GM-βVAE is expressed as:

$$P_\theta(\boldsymbol{x},\boldsymbol{y},z) = P_\theta(\boldsymbol{x}|z)P_\theta(\boldsymbol{y}|z)P(\boldsymbol{y}) \tag{13}$$

where θ is the network parameters of the generative model, $P_\theta(\boldsymbol{x}|z)$ is the approximated distribution of $\boldsymbol{x}$ conditioned on z, $P_\theta(\boldsymbol{y}|z)$ is the approximated distribution of $\boldsymbol{y}$ conditioned on z, and $P(\boldsymbol{y})$ is the initial prior on $\boldsymbol{y}$, which is selected to be a uniform multinomial distribution.

In the GM-βVAE model, we are pursuing both an orthogonal latent representation of data and the capability to capture multi-mode characteristics of the original input data. The evidence lower bound can be expressed as:

$$L_{GM-\beta VAE} = E_{Q_\phi(\boldsymbol{y},z|\boldsymbol{x})}\left[logP(\boldsymbol{y}) - logQ_\phi(\boldsymbol{y}|\boldsymbol{x}) + \beta log\frac{P_\theta(z|\boldsymbol{y})}{Q_\phi(z|\boldsymbol{x},\boldsymbol{y})} + logP_\theta(\boldsymbol{x}|z)\right] \tag{14}$$

where $E_{Q_\phi(\boldsymbol{y},z|\boldsymbol{x})}[logQ_\phi(\boldsymbol{y}|\boldsymbol{x})]$ is the conditional entropy that reflects how informative $\boldsymbol{x}$ is on $\boldsymbol{y}$. $log\frac{P_\theta(z|\boldsymbol{y})}{Q_\phi(z|\boldsymbol{x},\boldsymbol{y})}$ is the KL divergence loss term and $logP_\theta(\boldsymbol{x}|z)$ is the reconstruction loss, respectively.

5. Non-Intrusive uncertainty propagation

With the parameterized uncertainties, the last step of the proposed methodology is to propagate the uncertainty to the output of the problem and predict the associated statistical moments. Two UP methods, non-intrusive PCE and UDR, are reviewed and discussed in this section. A comparative study of these two methods will be presented in Section 6.

5.1. Non-intrusive PCE

PCE, originally proposed by Norbert Wiener [70], is a well-established functional expansion-based method based on the spectral representation of randomness (uncertainties). Hermite polynomial, Legendre polynomial, and Laguerre polynomial are optimal basis functions for Gaussian, uniform and, exponential distribution functions for the input uncertainties, respectively [70,71]. The details of polynomial chaos expansions can be found in literature [72–74].

The spectral representation of uncertainty decomposes a random variable (or function) into deterministic and stochastic components. For example, a metamodel for the stochastic response of a model $\alpha^*(\boldsymbol{x})$ is constructed on N independent random variables that construct a random vector: $\boldsymbol{x} = [x_1, \ldots\ldots x_N]$. Then $\alpha^*(\boldsymbol{x})$ is expanded into a basis of orthogonal polynomials to give a spectral representation of uncertainty:

$$\alpha^*(\boldsymbol{x}) = \sum_{i=0}^{P} \alpha_i \Psi_i(\boldsymbol{x}) \tag{15}$$

where $\Psi_i(\boldsymbol{x})$ is the random basis function corresponding to the i^{th} mode. We assume α^* to be a function of deterministic independent variable vector $\boldsymbol{c}$ and the N-dimensional random variable vector $\boldsymbol{x}$ that follows a specific probability distribution. The discrete sum can be expressed as a function of polynomial chaos order p and the number of random variables N, with $P+1 = \frac{(N+p)!}{N!p!}$. There are P linear combination terms and the multivariate orthogonal polynomial is expressed as:

$$\Psi_i(\boldsymbol{x}) = \prod_{k=1}^{N} \psi^{(k)}_{I_i^{(k)}}(x_k),\ i \in [1,P] \tag{16}$$

where the rows of the index matrix I contain the multi-index and the $\psi^{(k)}$ represents the input random variable.

To build a non-intrusive polynomial chaos model, approximations of the coefficients of polynomial will be obtained without making changes to the deterministic part [75]. Sampling based, collocation based, and quadrature methods are commonly used to obtain the polynomial coefficients in a non-intrusive polynomial chaos model.

Using the sampling based and the quadrature methods, the equation is projected on k^{th} basis:

$$\langle \alpha^*(\boldsymbol{c},\boldsymbol{x}), \Psi_k(\boldsymbol{x})\rangle = \left\langle \sum_{i=0}^{P} \alpha_i \Psi_i(\boldsymbol{x}), \Psi_k(\boldsymbol{x}) \right\rangle \tag{17}$$

where the inner product $<f(\mathbf{b}),g(\boldsymbol{x})>=\int_R f(\boldsymbol{x})g(\boldsymbol{x})p_N(\boldsymbol{x})d\boldsymbol{x}$, here the $p_N(\boldsymbol{x})$ is the probability density function of $\boldsymbol{b}$. We know that the basis functions are orthogonal, thus we have:

$$\langle \alpha^*(\boldsymbol{c},\boldsymbol{x}), \Psi_k(\boldsymbol{x})\rangle = \alpha_k \langle \Psi_k^2(\boldsymbol{x})\rangle \tag{18}$$

Then we can estimate the polynomial coefficients as:

$$\alpha_k = \frac{\langle \alpha^*(\boldsymbol{c},\boldsymbol{x}), \Psi_k(\boldsymbol{x})\rangle}{\langle \Psi_k^2(\boldsymbol{x})\rangle} \tag{19}$$

Using point-collocation based method [76], we start by replacing the uncertain variables with their polynomial expansions given by Eq. (20). Then $P+1$ vectors are chosen for a given PCE with $P+1$ modes and the deterministic solution is evaluated at these points. A system of equations is solved to obtain the polynomial coefficient (α_k) of the random variable α^*:

$$\begin{bmatrix} \Psi_0(x_0) & \Psi_1(x_0) & \ldots & \Psi_P(x_0) \\ \Psi_0(x_1) & \Psi_1(x_1) & \ldots & \Psi_P(x_1) \\ \ldots. & \ldots. & \ldots. & \ldots. \\ \Psi_0(x_p) & \Psi_0(x_p) & \ldots. & \Psi_p(x_p) \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ \ldots. \\ a_p \end{Bmatrix} = \begin{Bmatrix} a^*(x_0) \\ a^*(x_1) \\ \ldots. \\ a^*(x_p) \end{Bmatrix} \tag{20}$$

The mean and standard deviation of the solution are:

$$\mu = \alpha_0 \tag{21}$$

$$\sigma = \sqrt{\sum_{i=1}^{P} \alpha_i^2 \Psi_P^2(\boldsymbol{x})} \tag{22}$$

The high-order moments such as skewness (γ) and kurtosis (κ) can be obtained by computing the expectation products of three or four multivariate polynomials:

$$\gamma = \frac{1}{\sigma_{\alpha^*}^3} \sum_{i=1}^{P}\sum_{j=1}^{P}\sum_{k=1}^{P} E\left[\Psi_i(\boldsymbol{x})\Psi_j(\boldsymbol{x})\Psi_k(\boldsymbol{x})\right]\alpha_i\alpha_j\alpha_k \tag{23}$$

$$\kappa = \frac{1}{\sigma_{\alpha^*}^4} \sum_{i=1}^{P}\sum_{j=1}^{P}\sum_{k=1}^{P}\sum_{k=1}^{P} E\left[\Psi_i(\boldsymbol{x})\Psi_j(\boldsymbol{x})\Psi_k(\boldsymbol{x})\Psi_l(\boldsymbol{x})\right]\alpha_i\alpha_j\alpha_k\alpha_l \tag{24}$$

The solution of Eq. (20) requires $P+1$ function evaluations. If more than $P+1$ deterministic function values are used, the over-determined system of equations can be solved using the least square error. Hosder et al. [77] investigated this option by increasing the number of

collocation points in a systematic way through the introduction of a parameter n_P, which is defined as:

$$n_P = \frac{number\ of\ samples}{P+1} \tag{25}$$

They show that increasing the number of collocation points to twice or more of the minimum number required ($n_P \geq 2$) gives a better approximation to statistics at each polynomial order. In this study, we use $n_P = 2$, so the number of sampling points equals $2 \times (P+1)$ for different polynomial orders.

There are different sampling methods that can be used for the point-collocation non-intrusive PCE approach. For example, random sampling, Latin hypercube sampling, Hammersley sampling, Halton sampling, Sobol sampling, etc. These methods are also used as design of experiment (DOE) techniques in design and optimization [78]. Latin hypercube sampling is used in this work.

5.2. Univariate dimension reduction (UDR)

5.2.1. Evaluation of statistical moments by Gauss quadrature

The k^{th} order statistical moment of a function of random variable, $g(\boldsymbol{x})$, can be evaluated using Gauss quadrature:

$$E\left[g^k\right] = \int_{\Omega} [g(x)]^k f_X(x)dx \cong \sum_{i=1}^{m} \omega_i [g(x_i)]^k \tag{26}$$

where $f_X(\boldsymbol{x})$ represents the PDF of random variable $\boldsymbol{X}$ defined on Ω. m is the order of quadrature sum, ω_i and x_i are the i^{th} weight and Gauss node. The quadrature sum can achieve at most $2m-1$ polynomial quadrature order. In a Gauss-type quadrature formula, the Gauss node x_i is defined by the roots of the orthogonal polynomial; the weight ω_i is calculated by Lagrange interpolation polynomial [79]. When $\boldsymbol{X}$ follows the normal distribution, ω_i and x_i can be directly derived from the Gauss-Hermite quadrature formula [80].

5.2.2. Transformation to the standard normal distribution

A non-normally distributed variable can be transformed to a Gaussian distributed variable [81] using methods like Rosenblatt transformation [82] and Nataf transformation [83]. In this way, the Gauss nodes and weights of the Gauss-Hermite quadrature formula can be applied to compute the statistical moments if the original input variables are not normally distributed. The transformation T is expressed as:

$$w = T(\boldsymbol{x}) = P^{-1}(F(\boldsymbol{x})) \tag{27}$$

where w is the transformed variable, $P(\cdot)$ is the cumulative distribution function (CDF) of a transformed variable, and $F(\cdot)$ is the CDF of the original arbitrary distributed random variable, respectively.

Then the statistical moments of a function of input variable x can be written as:

$$\begin{aligned} E\left[g^k\right] &= \int_{\Omega} [g(x)]^k f_X(x)dx = \int_{-\infty}^{+\infty} P_w(w)\left[g\left(T^{-1}(x)\right)\right]^k dw \\ &\cong \sum_{i=1}^{m} \frac{\omega_i}{\sqrt{\pi}}\left[g\left(T^{-1}\sqrt{2}\,x_i\right)\right]^k \end{aligned} \tag{28}$$

5.2.3. Univariate dimension reduction

The UDR method [40,84,85] approximates a multivariate function by a summation of univariate functions. The statistical moments of each univariate function can be evaluated using the Gauss quadrature method introduced in Section 5.2.1.

Consider a mechanical system with the independent random input variables $\boldsymbol{x} = [x_1, \ldots\ldots, x_N]^T$, where N is the number of random variables. Let $g(\boldsymbol{x})$ be the response of the mechanical system, then the k^{th} order statistical moment is expressed as:

$$E\left[g^k(x)\right] = \int_R N[g(x)]^k f_X(x)dx \tag{29}$$

where $f_X(\boldsymbol{x}) = f_{X_1 \ldots X_N}(x_1, \ldots, x_N)$ is the joint probability density function of $\boldsymbol{X}$ and $E[\cdot]$ is the expectation operator. In the univariate approximation of $g(\boldsymbol{x})$, each term only depends on one variable and the other variables are fixed to their mean values:

$$\begin{aligned} g(\boldsymbol{x}) &= g(x_1, \ldots, x_N) \\ &= \sum_{i=1}^{N} g(\mu_1, \ldots, \mu_{i-1}, x_i, \mu_{i+1}, \ldots, \mu_N) - (N-1)g(\mu_1, \ldots, \mu_N) \end{aligned} \tag{30}$$

where N is the total dimension of variables and μ_i is the mean value of variable x_i.

Then the multi-dimensional integration is simplified as:

$$\begin{aligned} E\left[g^k(\boldsymbol{x})\right] &\cong E\left[\left(\sum_{i=1}^{N} g(\mu_1, \ldots, \mu_{i-1}, x_i, \mu_{i+1}, \ldots, \mu_N) - (N-1)g(\mu_1, \ldots, \mu_N)\right)^k\right] \\ &= \sum_{l=1}^{k} \binom{k}{l} \cdot E\left[\left(\sum_{i=1}^{N} g(\mu_1, \ldots, \mu_{i-1}, x_i, \mu_{i+1}, \ldots, \mu_N)\right)^k\right] \cdot [-(N-1)g(\mu_1, \ldots, \mu_N)]^{k-l} \end{aligned} \tag{31}$$

Noted that, the error of this approximation is mainly induced by the approximation of the interactions among variables [40].

The moments of g_i can be obtained using the one-dimensional Gauss quadrature-based numerical integration (as discussed in Section 5.2.1). If we use the same number (M) of Gauss nodes and weights for the one-dimensional integration of all g_i, the total number of simulations (evaluations of $g(\boldsymbol{x})$) will be at least $(M-1)N+1$ and at most $MN+1$, where M represents the number of Gauss nodes and weights, N represents the number of random variables [86].

Subsequently, the statistical moments are given by:

$$\mu = E[g(\boldsymbol{x})] \tag{32}$$

$$\sigma = \sqrt{E[g^2(\boldsymbol{x})] - \mu^2} \tag{33}$$

$$\gamma = \frac{E[g^3(\boldsymbol{x})] - 3E[g^2(\boldsymbol{x})]\mu + 2\mu^2}{\sigma^3} \tag{34}$$

$$\kappa = \frac{E[g^4(\boldsymbol{x})] - 4E[g^3(\boldsymbol{x})]\mu + 6E[g^2(\boldsymbol{x})]\mu^2 - 3\mu^4}{\sigma^4} \tag{35}$$

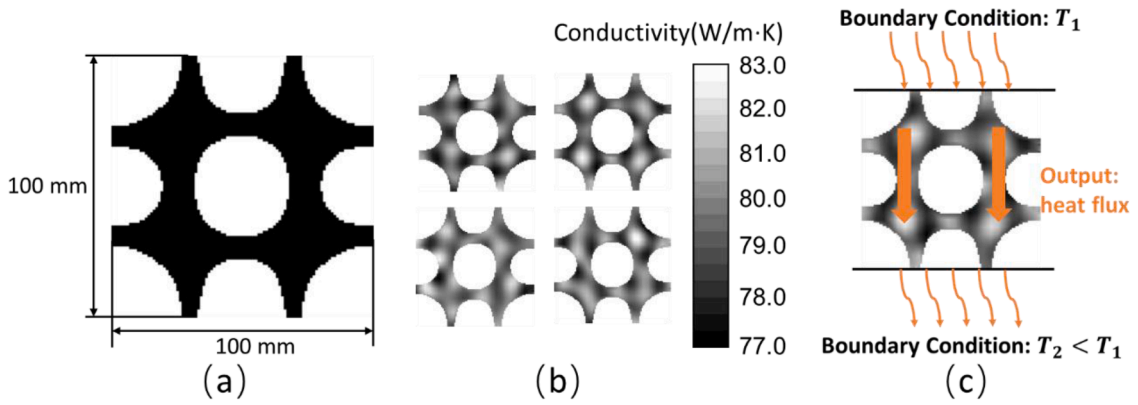


Fig. 3. (a) Pixelated topological structural domain. (b) Samples of the stochastic local material properties (thermal conductivity of each element) distributed in the topological structure. (c) Heat transfer simulation predicting the heat flux distribution throughout the structure domain.

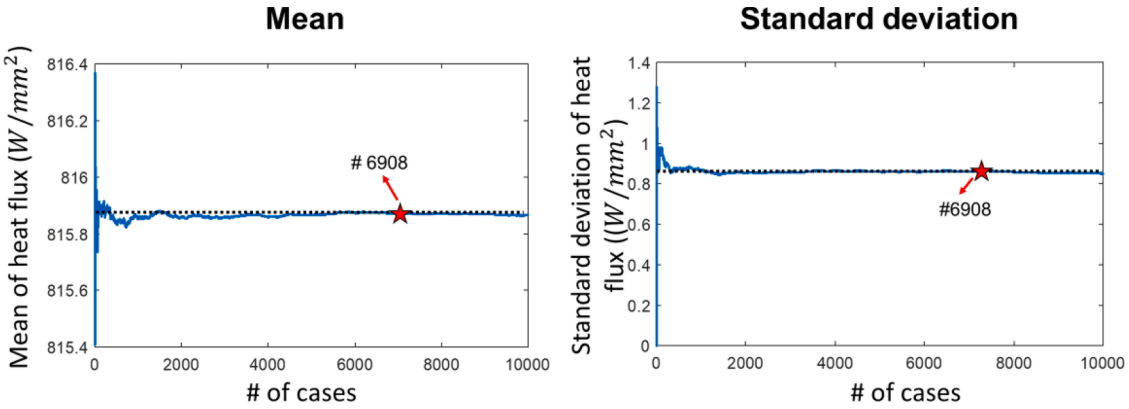


Fig. 4. Evaluation of the output's mean and the standard deviation by MCS.

5.2.4. Estimation of the probability density function

Once the first four order statistical moments are obtained, the probability density function (PDF) can be evaluated by empirical distribution systems, such as the Pearson system [87]. The Pearson system approximates the PDF $f(x)$ as solution of differential equation:

$$\frac{df(\boldsymbol{x})}{d\boldsymbol{x}} = -\frac{a+\boldsymbol{x}}{c_0 + c_1\boldsymbol{x} + c_2\boldsymbol{x}^2} f(\boldsymbol{x}) \tag{36}$$

The Pearson system includes the normal, Gamma, and Beta distribution among the families, which depend on the Pearson parameters (a, c_0, c_1, c_2). The Pearson parameters can be expressed in terms of the first four central moments of the distribution, which expressed as:

$$c_1 = a = -\frac{\mu_3(\mu_4 + 3\mu_2^2)}{10\mu_4\mu_2 - 18\mu_2^3 - 12\mu_3^2} \tag{37}$$

$$c_0 = -\frac{\mu_2(\mu_2\mu_4 - 3\mu_3^2)}{10\mu_4\mu_2 - 18\mu_2^3 - 12\mu_3^2} \tag{38}$$

$$c_2 = -\frac{(2\mu_2\mu_4 - 3\mu_3^2 - 6\mu_2^3)}{10\mu_4\mu_2 - 18\mu_2^3 - 12\mu_3^2} \tag{39}$$

where μ_1, μ_2, μ_3, μ_4 correspond to the first four central moments of the system. Therefore, once we know the first four central moments, we can construct a PDF that is consistent with these moments.

6. Engineering case studies

In this section, two engineering cases are presented to demonstrate the effectiveness of the proposed methodology. We also test and compare the different technical options in Step 2 Parameterization and Step 3 UP to understand their relative merits. The MCS method is applied to provide reference solutions for validating the proposed methodology.

6.1. Engineering case 1: Thermal conductivity analysis considering stochastic local material properties

For the first case study, the input uncertainties are the spatially varying local properties generated by the shortest interior path distance-based Gaussian random field model. The goal is to quantify the uncertainty associated with the output, which is the heat flux through a topological structure. The case study tests two components in the proposed method: uncertainty parameterization and UP.

6.1.1. Problem description, uncertainty representation, and parameterization

A finite element simulation model is established to predict the heat flux (output) with consideration of the stochastic local material properties (input uncertainty). The topological structural domain is discretized into pixelated elements. Each pixel in the structure is assumed to be 1 mm in length, the dimensions of the topological structure are shown in Fig. 3(a). The shortest interior path distance between any of the two elements $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ in the pixelated feasible region is evaluated for constructing the covariance matrix. The local material properties (thermal conductivity) are generated using the shortest interior path distance-based Gaussian random field model, which is expressed as:

$$C(\boldsymbol{x}_i, \boldsymbol{x}_j) = 80^2 \cdot \exp\left(-\frac{D_{SP}(\boldsymbol{x}_i, \boldsymbol{x}_j)}{30}\right) \tag{40}$$

By sampling the Gaussian random field model, we generate random

Table 1
Number of data points used in non-intrusive PCE and UDR for 8-dimensional latent space.

PCE	Polynomial order	1	2	3	4	5
	Number of data points	18	90	330	990	2574
UDR	Number of Gaussian nodes	2	4	6	8	10
	Number of data points	17	33	49	64	81

realizations of the local material property distribution in the structural domain. 10,000 realizations are generated (Fig. 3(b)). Each realization is one sample of the spatial distribution of element-wise local properties. It is to be noted that sampling the Gaussian random field model is efficient as no finite element simulation is involved at this stage.

This dataset is compressed by PPCA or βVAE to low-dimensional normally distributed, independent random variables. A convergence test is conducted to investigate the number of independent random variables, as shown in Appendix A.2. To achieve a balance between the accuracy and the dimensionality of the latent space, we select the 8-dimensional latent space for the following UP studies.

6.1.2. *Uncertainty propagation: from stochastic local properties to heat flux*

The commercial finite element software package, ABAQUS, was used for all case 1 heat transfer simulations. The 2D geometrical domain and set up used for the heat transfer simulations is shown in Fig. 3(c). A thermal flux is generated in the system by assigning two different temperature values (T_1 & T_2) to the upper and lower boundaries of the domain. All 10,000 samples are simulated to provide a large dataset for MCS. A convergence test is conducted to find the minimum number of samples required by the MCS method in order to obtain an accurate prediction of the mean and variance of the output (Fig. 4 and Appendix A.3).

Two UP methods, non-intrusive PCE and UDR are compared. Table 1 lists the number of data points (simulations) required by non-intrusive PCE with different polynomial orders and by UDR with different

Table 2
Results of UP and validation. The results of different combinations of parametrization and UP methods are compared with the reference values obtained by MC sampling.

		Mean (Difference)	STD (Difference)	Skewness (Difference)	Kurtosis (Difference)	# of sampling points
Monte Carlo		815.8638	0.8509	−0.0856	2.9254	6908
PCE	PPCA	815.8718 (0.0010%)	0.8312 (2.3152%)	−0.0926 (8.1776%)	3.0128 (2.9876%)	330
	βVAE	815.8726 (0.0011%)	0.8249 (3.0556%)	−0.0875 (2.2196%)	2.8564 (2.3586%)	330
UDR	PPCA	815.8719 (0.0010%)	0.8293 (2.5385%)	−0.0848 (0.9345%)	3.0465 (4.1396%)	49
	βVAE	815.8504 (0.0016%)	0.8465 (0.5171%)	−0.1233 (44.0420%)	4.2122 (43.9871%)	49

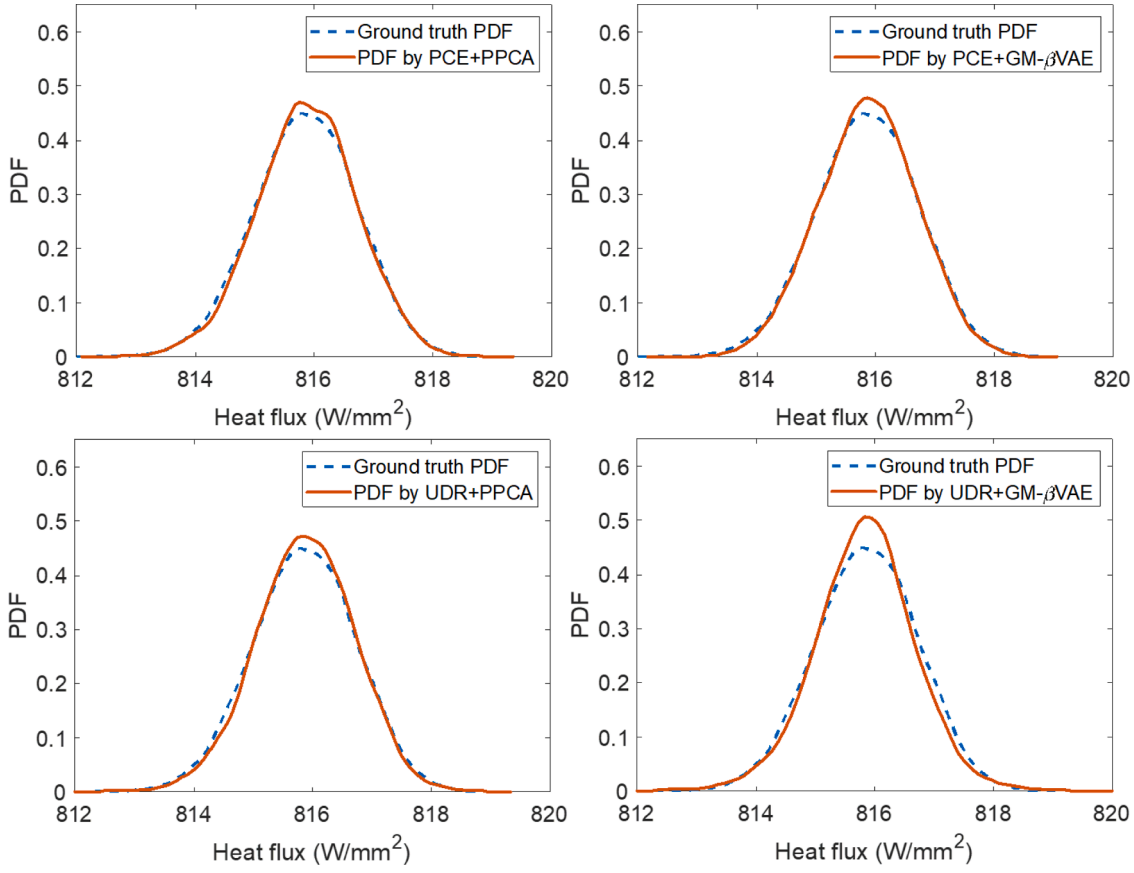


Fig. 5. PDF generated by different combinations of UP using Pearson system.

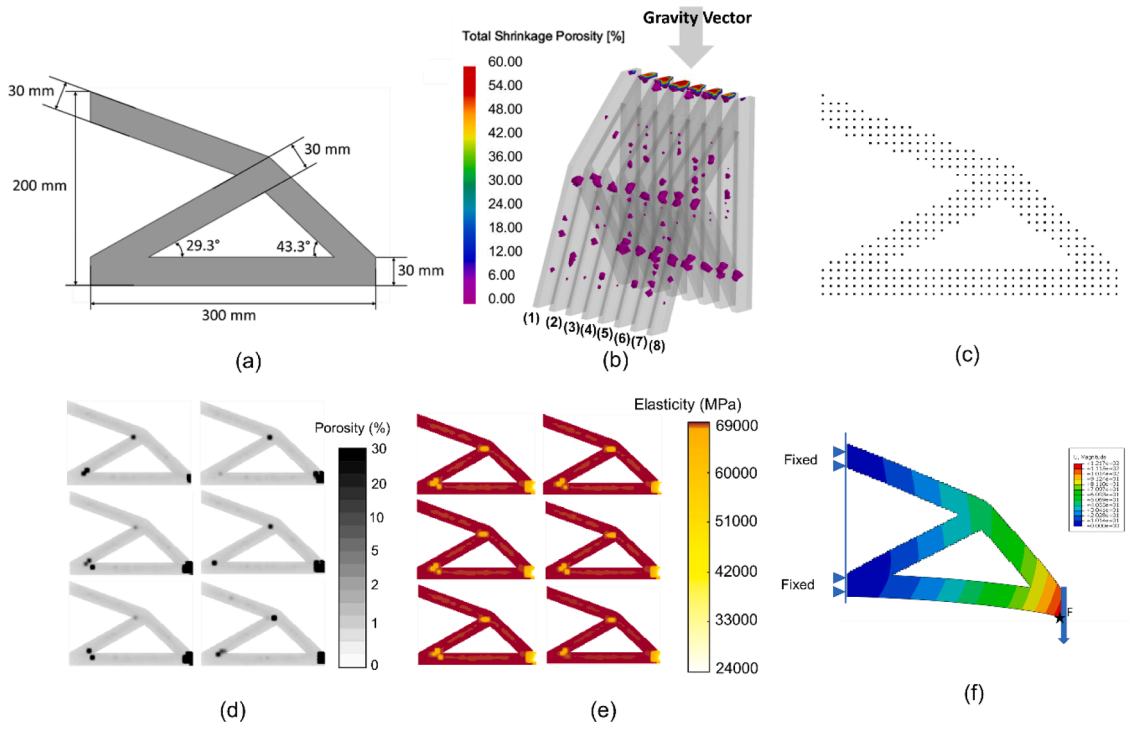


Fig. 6. (a) Geometry and dimensions of the topological truss structure. (b) 8 parts are cast per mold; Example shrinkage porosity predictions are shown. (c) Sampling locations for constructing kernel functions. (d) Examples of the spatially varying local shrinkage properties. (e) Examples of the spatially varying local mechanical properties (elastic moduli). (f) Simulation of the structure deflection under a point load.

Table 3
Process parameters, interface parameters and material properties used in the casting solidification simulation. The material properties marked with * is temperature-dependent, the temperature-property curves are shown in Appendix A.1, Fig. A.1.

Process Parameters	Mean	Range
Gravity (mg/s^2)	9.80	±0.15%
Ambient Temperature (°C)	20	±0.3%
Air Cooling HTC (W/m^{2}K)	10	±0.15%
Initial Temperature of Casting (°C)	720	±0.3%
Initial Temperature of Mold (°C)	100	±0.3%
Interface Parameters	Mean	Range
HTC while Casting is Liquid (W/m^{2}K)	500	±0.15%
HTC while Casting is Solid (W/m^{2}K)	300	±0.15%
Material Properties	Mean	Range
Latent Heat of AlSi7Mg03 (kJ/kg)	431	±0.3%
Density of AlSi7Mg03 (kg/m^3)	*	/
Fraction Solid of AlSi7Mg03	*	/
Conductivity of AlSi7Mg03 (W/mK)	*	/
Specific Heat of AlSi7Mg03 (kJ/kgK)	*	/
Density of Sand (kg/m^3)	1520	±0.15%
Specific Heat of Sand (kJ/kgK)	*	/
Conductivity of Sand (W/mK)	*	/
Liquidus Temperature of AlSi7Mg03 (°C)	613	±0.05%
Solidus Temperature of AlSi7Mg03 (°C)	548	±0.05%

numbers of Gaussian points using 8-dimensional latent space.

The results of non-intrusive PCE and UDR are compared in Table 2 and Fig. 5. For non-intrusive PCE, the polynomial order is set as $p = 3$; for UDR method, we use $M = 6$ Gauss nodes and weights. Increasing the order or number of quadrature points beyond these chosen numbers does not noticeably improve the accuracy of predicting the statistical moments, while the computational cost will be increased significantly (Appendix A.5, Fig. A.3). Non-intrusive PCE-based UP methods and UDR-based UP methods both perform well in low-order moments estimations (means and STDs). However, when estimating the high-order moments, the combination of UDR and βVAE shows a large error (>40%) in skewness and kurtosis estimation, which is due to the effect of interactions between variables. Both PPCA and βVAE work well in non-intrusive PCE, while βVAE will have less error in skewness estimation. It shows that the non-intrusive PCE approximations are accurate within 3.1% of the MC results when calculating the low-order moments (means and STDs) with 330 sampling points, UDR approximations are accurate within 2.6% of the MC results with 49 sampling points when calculating the low-order moments (means and STDs). It is suggested to use PPCA as the parameterization method and UDR as the UP method in this study since it balances the number of sampling points as well as the accuracy of statistical moments.

6.2. *Engineering case study 2: quantification of manufacturing-induced uncertainties based on a process-property-performance simulation*

The second engineering case is a comprehensive study that tests all aspects of the proposed methodology: uncertainty representation, uncertainty parameterization, and UP. The goal is to quantify the uncertainty for structural performance based on manufacturing-induced uncertainties input within a process-property-performance model. This case study is challenging because the data does not follow the normal distribution.

6.2.1. *Problem description, data collection and uncertainty representation*

It is well established in literature that pouring temperatures, pouring speeds, cooling rates, and thermal gradients will directly influence a casting structure and its associated mechanical properties during solidification [88]. Casting defects, such as macrosegregation, shrinkage porosities, and coarse-grained zone, can be predicted through numerical simulation [89–91]. The stochastic geometry and spatial distribution of the casting defects are important sources of material uncertainties. Furthermore, variations in the processing parameters will lead to variations in thermal energy and deviation of initial parameters, which adds another source of uncertainty to the problem. Literature has revealed that the pouring temperature [92], solidification rate [93], thermal

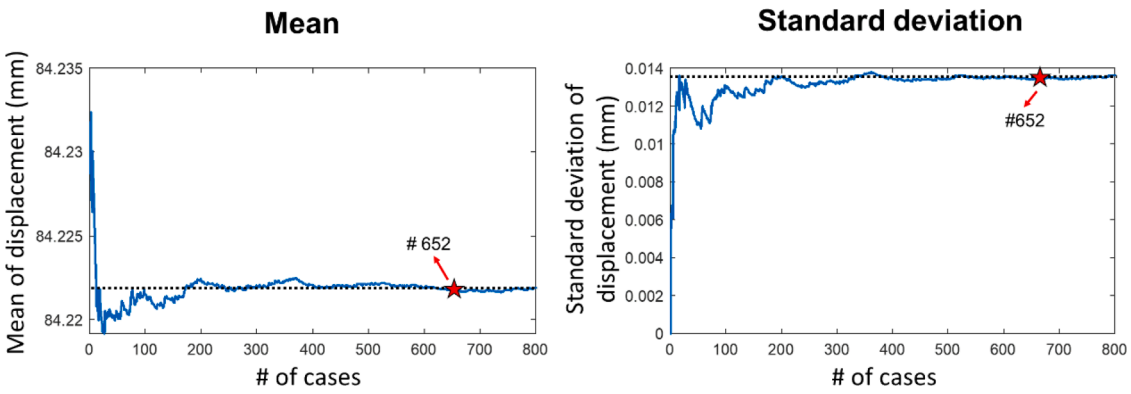


Fig. 7. Evaluation of the output's mean and the standard deviation by MCS.

Table 4
Prediction of the mean and standard deviation of the output. Different combinations of parametrization methods and UP methods are compared.

PCE		Mean (Difference)	STD (Difference)	Skewness (Difference)	Kurtosis (Difference)	# of sampling points
		84.2219	0.0136	1.0218	3.1626	652
	PPCA	84.2229 (0.0012%)	0.0115 (15.4412%)	0.4408 (56.8604%)	1.1204 (64.5734%)	168
	GM-βVAE	84.2222 (0.0004%)	0.0143 (5.1471%)	0.9529 (6.6451%)	3.4207 (8.1610%)	168
UDR	PPCA	84.2212 (0.0008%)	0.0101 (25.7353%)	1.0855 (6.2341%)	5.7453 (81.6638%)	73
	GM-βVAE	84.2347 (0.0151%)	0.0145 (6.6176%)	0.0660 (93.5408%)	6.4407 (103.6521%)	49

gradients [94], and other material properties and processing parameters will influence the mechanical properties of a cast part.

The geometry of a topological truss structure is created in Solidworks, and the solidification process is simulated using the commercial finite element software package, ProCAST. The size of each truss structure is 300 mm × 200 mm × 10 mm (Fig. 6(a)). A 5 mm tetrahedron mesh and an adaptive time stepper is used for all case 2 simulations. As shown in Fig. 6(b), each mold contains 8 truss structure castings of the same design. Uncertainties for the parameters used in the solidification simulations (process parameters, interface parameters, and material properties) are represented by random parameters with uniform distributions (Table 3). AlSi7Mg03 ternary alloy is used as the casting material and silica sand is assigned to the mold [95]. The castings are poured at 720 °C into a preheated mold at 100 °C. The heat transfer coefficient (HTC) between the casting and the mold was assigned to be constant at 500 W/m^{2}K while the casting is liquid and constant at 300 W/m^{2}K when the casting is solid. A linear deviation between the solid and liquid HTC values is used during solidification. Air cooling is assigned to the exterior surfaces of the mold by using an ambient temperature 20 °C and a HTC value of 10 W/m^{2}K. In total 800 structure samples (100 molds) were simulated and analyzed. The 800 truss structure samples follow the same geometrical design, but have varying performances due to the manufacturing and material uncertainties.

Shrinkage porosity values at each node of the casting are extracted and exported. The thickness of the structure is 1/20 of the width and 1/30 of the length of the structure, so variations along the thickness dimension are neglected. We focus on analyzing the spatially varying material properties in the 2D topological domain. At each location in the 2D domain, the average porosity value along the thickness direction is used. The casting is then re-meshed to a pixelated structure and each element is assigned the corresponding local mechanical properties. The element size of the pixelated structure is 3 mm. The sampling locations for observing the spatially varying local porosities are shown in Fig. 6(c). In order to achieve the best accuracy, we group all structural samples into four groups based on their location in the mold. The four groups correspond to locations (1) and (8), (2) and (7), (3) and (6), and (4) and (5) in the mold shown in Fig. 6(b). For each group, a Gaussian random field model with the kernel function of Eq. (4) is fitted to represent the spatial distribution of porosities. By sampling the Gaussian random field model, 10,000 random realizations of the local shrinkage porosity distribution in the topological structural domain are generated (Fig. 6(d)).

6.2.2. *Porosity-elasticity relation and parameterization*

The elasticity of each element is obtained with the assumption that the porosity-elasticity relation follows this equation:

$$E = E_0 e^{-B\varnothing_p} \tag{41}$$

where E represents Young's modulus of the porous pixelated element, E_0 represents Young's modulus of nonporous AlSi7Mg03, B is an empirical constant, which is selected to be $B = 2$ for this example, and $\varnothing_p$ is the fractional pore volume of the body predicted by the solidification simulation. Note, that a real-world casting would not behave the same as the part defined because of the simplified porosity-elasticity relationship. With Eq. (41), the random field realizations of local porosities are converted to the spatial distribution of local elastic moduli (Fig. 6(e)). It is noted that the local elastic modulus inside the topological structural domain is not normally distributed (Appendix A.4, Fig. A.2).

PPCA and βVAE are used to compress the dataset of 10,000 realizations of the local elastic moduli to a low-dimensional feature space that is formed by normally distributed, independent random variables. Because the local elastic modulus is not normally distributed, it's difficult for βVAE to obtain an efficient accuracy and get independent latent variables (Appendix A.2). Therefore, we propose to use the GM-βVAE, which would capture the complex statistical distribution with a multimodal statistical distribution function. To achieve a balance between the dimensionality of compressed feature space (computational cost) and the accuracy (Appendix A.2), we use 6 latent variables in PPCA and 6 latent variables in GM-βVAE for PCE, and use 12 latent variables in PPCA and 8 latent variables in GM-βVAE for UDR. The latent features, which follow the Gaussian mixture distribution, will be converted to the

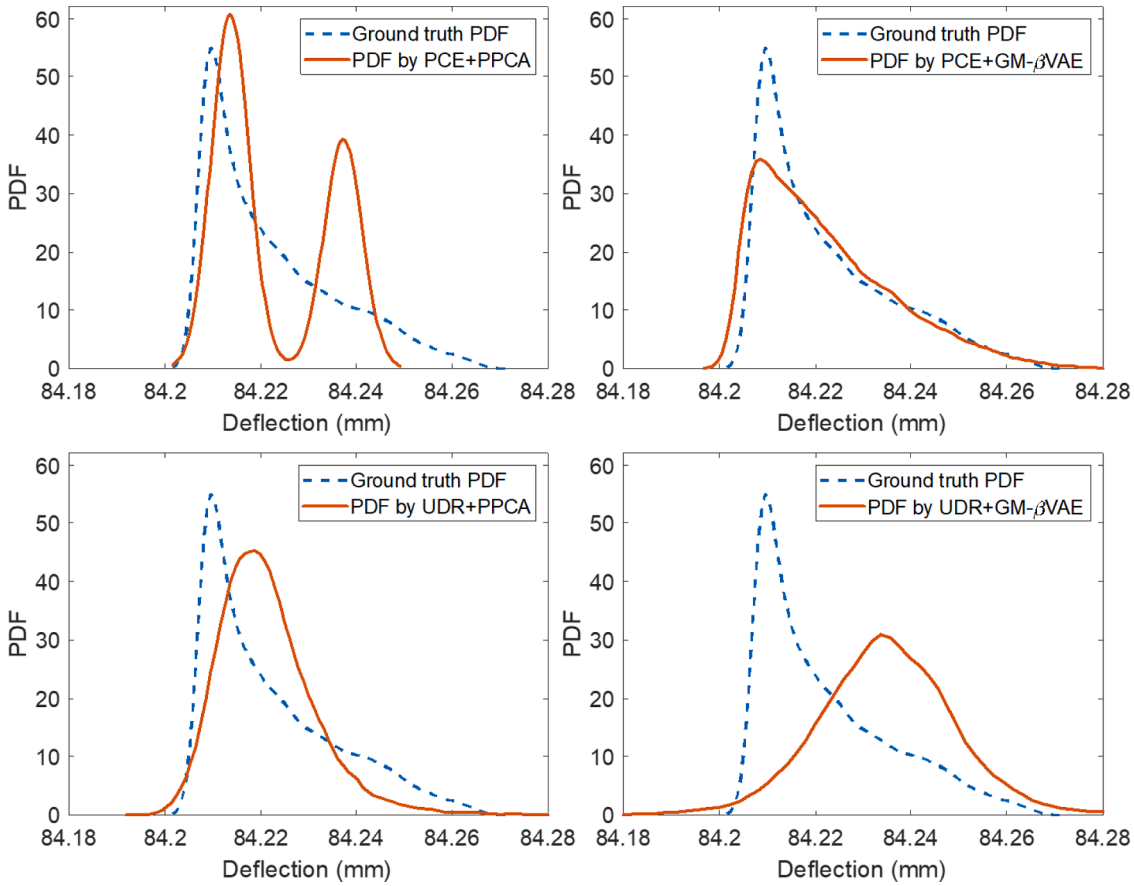


Fig. 8. The predicted PDF via the Pearson system by different combinations of the UP strategies.

Table 5
Comparison of PPCA, βVAE and GM-βVAE.

	PPCA	βVAE	GM-βVAE
Computational cost	~0.84 min	~541 min	~562 min
Linearity	Linear	Non-linear	Non-linear
Orthogonality	Fully orthogonal	Approximately orthogonal	Approximately orthogonal
Latent distribution	Gaussian	Gaussian	Mixture of Gaussians
Prediction accuracy	Satisfactory accuracy when compressing data with normally distributed marginal distributions; low accuracy when compressing data with non-normally distributed marginal distributions	Satisfactory accuracy when compressing data with normally distributed marginal distributions; low accuracy when compressing data with non-normally distributed marginal distributions	Highest accuracy when compressing data with normally distributed marginal distributions; satisfied accuracy when compressing data with non-normally distributed marginal distributions
Potential issues in implementation	Low reconstruction accuracy if the dataset has large variations	Hard to balance the data reconstruction accuracy and the orthogonality of latent variables	May induce errors when transforming from the mixture of Gaussians to Gaussian distribution

normal distribution using the transformation method introduced in Section 5.2.2 when performing UDR.

6.2.3. Uncertainty propagation: from stochastic local elastic moduli to structure deflection

The structural performance is measured by the deflection under a point load. The boundary condition of the problem is shown in Fig. 6(f). The deflection of the structure is obtained in ABAQUS and the displacement of the marked node is used as the output of the problem. The uncertainties associated with the local mechanical properties are propagated to predict the statistical moments of the structure deflection.

The results of the 800 process-property-performance simulations provide the reference values of the deflection statistical moments. A convergence test (Fig. 7 and Appendix A.3) is conducted to find the minimum number of samples required by the MCS method in order to obtain an accurate prediction of the mean and variance of the output. We find that a minimum of 652 samples are needed for MCS convergence.

The results of non-intrusive PCE and UDR are compared in Table 4. The PDFs estimated by Eq. (36) are shown and compared in Fig. 8. For the non-intrusive PCE, we set the order of polynomial as $p = 3$. For UDR, we use $M = 6$ Gauss nodes. Further increasing the polynomial order or the number of Gauss nodes will not improve the prediction accuracy noticeably, while the computational cost will be increased significantly (Appendix A.6). We estimate the mean value, standard deviation, skewness, and kurtosis. The PDFs are estimated using Eq. (36) and compared with the ground truth PDF, as shown in Fig. 8. The combination of GM-βVAE and non-intrusive PCE provides the best prediction accuracy, 0.0004% for mean, 5.1471% for STD, 6.7430% for skewness, and 8.1610% for kurtosis. This combination requires 168 sampling points, which are 74% fewer than those required by MCS. The combination of GM-βVAE and UDR has the lowest prediction accuracies among all the combinations. The accuracy of estimating the high-order moments using the combination of GM-βVAE and UDR is impaired by

the interaction effects among the latent variables. Therefore, the combination of GM-βVAE and UDR is not suitable for estimating the high-order moments. For the estimation of low-order moments, we recommend the combination of GM-βVAE and UDR, which balance the prediction accuracy and the computational cost, 0.0151% error for mean, 6.6176% error for STD, and 49 sampling points.

Compared with the first case study, we see larger prediction errors in this case. We identify the following factors that may contribute to the prediction error:

1 Large variations are observed in the local porosity values. The extreme values and the associated irregular statistical distribution, increase the difficulty in parameterization. The variation is further increased by converting the porosity values to elasticities of each element in the structure using an exponential function, which is shown in Eq. (41).
2 Another source of error lies in the process of transforming the mixture of Gaussian to a normal distribution.

6.3. Comparison of the methods of uncertainty parameterization and uncertainty propagation

In this section, the uncertainty parameterization methods (PPCA and VAE-based) and the uncertainty propagation methods (non-intrusive PCE and UDR) are compared. Suggestions on selecting appropriate methods are provided.

6.3.1. Comparison of the parameterization methods

PPCA and βVAE-based methods (βVAE or GM- βVAE) are compared in the two engineering case studies presented in Section 6.1 and Section 6.2, as well as an additional engineering case in Appendix A.7. The comparison is conducted with respect to the computational cost, the linearity of the compression operation, orthogonality of the obtained parameters, ability of generating normally-distributed parameters, the overall prediction accuracies, and some potential issues in implementation are listed in Table 5.

The guidelines for selecting parameterization methods is summarized as follows:

1 When the marginal distributions of the QoIs are normally distributed, it is suggested to use PPCA since it balances the computational cost and the accuracy.
2 When the local values of QoIs (e.g. local properties in a structure) have large variations, it is suggested to use GM-βVAE, which can capture the multi-mode characteristics, in order to achieve a better reconstruction accuracy.
3 When calculating the high-order moments using UDR-based UP approaches, it is suggested to use PPCA since it can ensure the independency among the parameters.

6.3.2. Comparison of the uncertainty propagation methods

The non-intrusive PCE and UDR are implemented and compared in the aforementioned case studies. It might not be appropriate or complete to judge or rank methods with only a few examples. However, some characteristics, advantages, and disadvantages can be observed from those case studies:

1 Accuracy. Non-intrusive PCE and UDR show no great difference regarding the accuracy when calculating the low-order moments (mean and variance). However, the UDR is more sensitive to the interaction effect among the input parameters, especially when calculating the high-order moments (skewness and kurtosis). Using the UDR method will lead to large errors in calculating high-order moments when a relatively strong interactive effect exists among input parameters.
2 Efficiency. UDR is more efficient compared to non-intrusive PCE. The computational cost of the quadrature-based non-intrusive PCE depends highly on the order of polynomial expansion (refer to Eq. (25)), and it increases dramatically with the number of input parameters. By contrast, the computational cost of UDR increases linearly with the number of input parameters.
3 Other observations. UDR has advantages in estimating low-order moments. When applying UDR to estimate the high-order moments, it could be problematic as UDR is sensitive to the interactive effects among random variables. Non-intrusive PCE is more accurate when used to estimate the high-order moments or to predict the PDF.

Our recommendation of selecting UP methods are summarized as follows:

1 To estimate the low-order statistical moments, it is suggested to use UDR-based UP methods for a lower computational cost.
2 To estimate the high-order statistical moments or the complete PDF, it is suggested to use non-intrusive PCE-based methods for a higher accuracy.
3 In the proposed framework, the combination of UDR and βVAE (or GM-βVAE) is not recommended for calculating high-order moments. The deep learning-based parameterization methods cannot fully disentangle the interactions among latent variables (though the correlation values are low), thus it will result in huge errors in high-order moment estimation.

7. Conclusion

In this paper, a new methodology is presented for quantifying and propagating aleatoric uncertainties distributed in topology structure domains. Gaussian random field model based on the shortest interior path distance is proposed to capture the topological characteristics of the spatial domain. PPCA and variants of VAE are employed to convert the high dimensional random field uncertainties to independent, normally distributed random parameters. Non-intrusive PCE and UDR are employed to propagate the input uncertainties to obtain the statistical moments of the output. The proposed method is demonstrated on two engineering case studies, where the source of uncertainties is the spatially varying, stochastic local material properties distributed in the topological structural domain. Our major conclusions are summarized as follows:

- First, the proposed method is effective in quantifying and propagating the uncertainties distributed in complex topological structures. Compared with the MCS method, our method achieves a reasonable level of accuracy while significantly reducing the number of samples required for evaluating the statistical moments of the output.
- Second, we suggest using βVAE (GM-βVAE) as the parameterization method and UDR as the UP method since it balances the number of sampling points as well as the accuracy of statistical moments.

This work marks our first step toward quantifying uncertainties in complex mechanical systems. We also identify opportunities to improve the proposed method:

- First, the uncertainty representation method in our proposed methodology is not powerful enough to capture non-Gaussian data.
- Second, the proposed method only considers the characteristics of the spatial domain. In future works, we will improve the method in

order to capture both spatial (e.g. structure design) and temporal (e.g. toolpath) information of the input aleatoric uncertainties.
- Third, the proposed methodology can be integrated with parametric or topology structure design frameworks to enable the robustness and/or reliability-based design of topological structures (e.g. meta-material unit cell) under manufacturing-induced uncertainties.

APPENDIX

A.1 Material properties used in solidification simulation

A.2 Selection of the dimension of latent variables

PPCA and βVAE are used to compress the random field realizations of the local material properties to a low-dimensional feature space that is formed by normally distributed, independent random variables. The low-dimensional latent variables can be mapped back to the original high-dimensional structural space to reconstruct the spatial distribution of the local material properties. The reconstruction accuracy will affect the accuracy when performing UP. The reconstruction accuracy is measured by the R squared value of input data and the corresponding reconstruction.

Engineering case 1: For both PPCA and βVAE parameterization methods, we test the reconstruction accuracies with the number of latent variables of 6, 8, 10, and 12, respectively (Table A.1). When the number of latent variables equals or is larger than 8, the R squared values of both two methods are larger than 0.9. We select the 8-dimensional latent space for the following UP studies.

Engineering case 2: The reconstruction accuracies of PPCA, βVAE, and GM-βVAE are shown in Table A.2. While PPCA performs slightly better than βVAE, they both have low reconstruction qualities. GM-βVAE has a satisfactory reconstruction accuracy.

A.3 Criterion of convergence of MCS

The convergence criterion of MCS [96] in our engineering cases is defined as an approximate 99% probability that the mean value of n MCS samples μ_{MCSn} is within 0.0001 unit value of the true mean $\overline{\mu_{MCS}}$.

The mean value of n MCS samples is computed by the unbiased estimator $\overline{\mu_{MCS}}$:

$\mu_{MCS}=1/n\Sigma i = 1\ n\ \mu_{MCSi}$ (42) and the unbiased estimator of σ is written as:

$$S_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\mu_{MCSi} - \mu_{MCS})} \tag{43}$$

The confidence interval (CI) is obtained as:

$$CI = \mu_{MCS} \pm z^* \frac{\sigma}{\sqrt{n}} \tag{44}$$

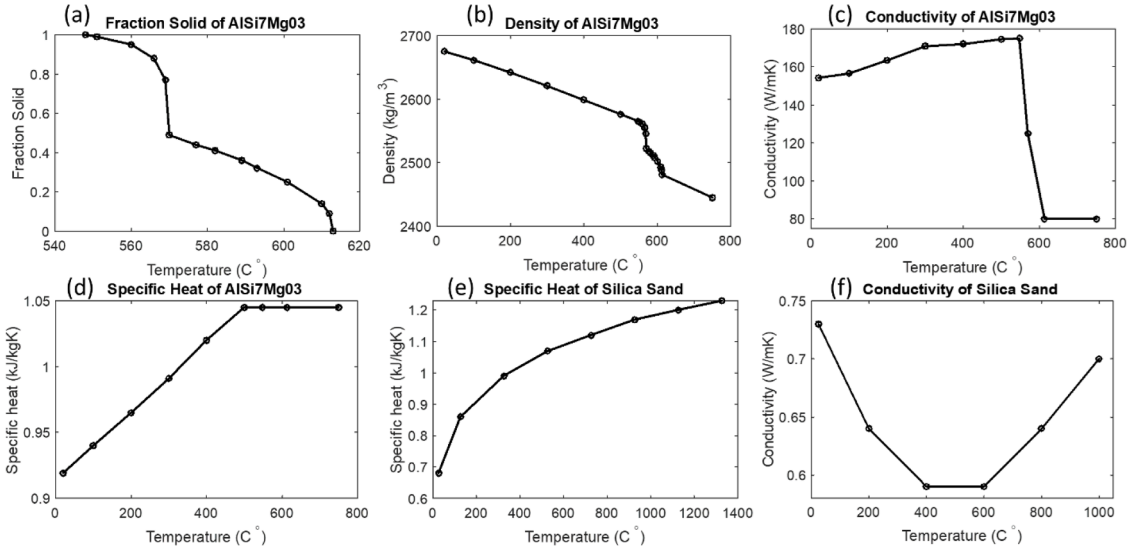


Fig. A.1. Temperature-dependent material properties used for AlSi7Mg03 and silica sand in the solidification simulations.

Table A.1
Reconstruction accuracies measured by the R squared values of PPCA and βVAE with different number of latent variables in engineering case 1.

# of latent variables	6	8	10	12
PPCA	0.8313	0.9252	0.9787	0.9999
βVAE	0.7825	0.9064	0.9212	0.9290

Table A.2
Reconstruction accuracies measured by the R squared values of PPCA, βVAE, and GM-βVAE with different number of latent variables in engineering case 2.

# of latent variables	6	8	10	12
PPCA	0.3677	0.3800	0.3923	0.4039
βVAE	−0.0304	0.0217	0.0578	0.1022
GM-βVAE	0.8072	0.8271	0.8755	0.8932

where a represents the confidence level value, for 99% CI, z^* equals 2.576. σ represents the standard deviation of the samples. n represents the number of MCS cases.

Using the central limit theorem, the probability of the sample lies in the 99% CI is written as:

$$P\left(\left|\frac{\mu_{MCS}n - \mu_{MCS}}{\frac{\sigma}{\sqrt{n}}}\right| \leq 2.576\right) = 99\% \tag{45}$$

It is to be noted that, Eq. (45) represents that there is a 99% probability that the sample mean $\mu_{MCS}n$ is within $2.576\frac{\sigma}{\sqrt{n}}$ units of the true mean μ_{true}.

Combining Eqs. (44 and 45), we get:

$$2.576\sqrt{\frac{\sum_{i=1}^{n}(\mu_{MCSi} - \mu_{MCS})}{n(n-1)}} < 0.0001 \tag{46}$$

When the number of MCS samples n satisfies Eq. (46), we recognize the MCS converges at n samples.

A.4 Marginal distribution of local properties

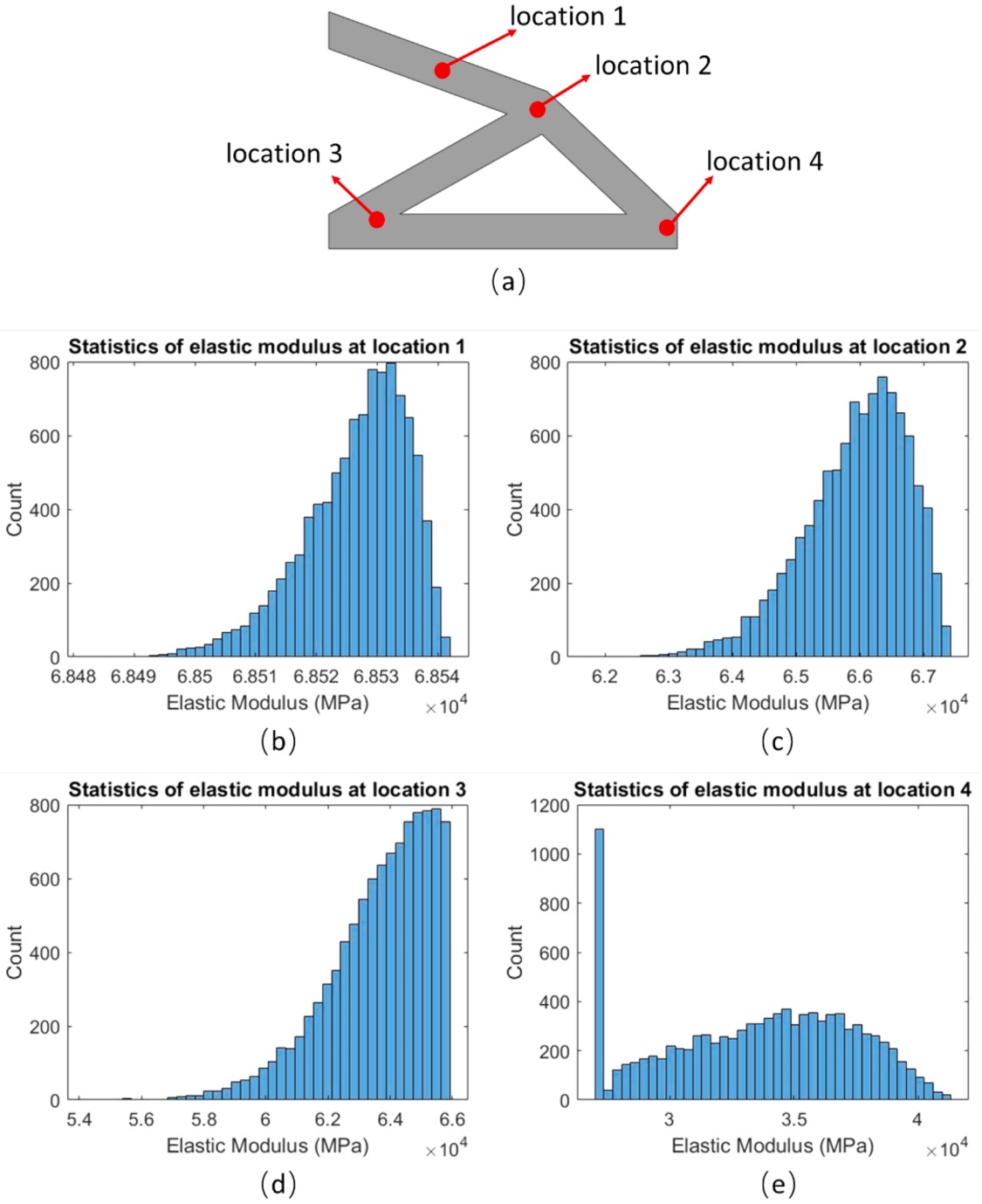


Fig. A.2. (a) Locations for observing the marginal distributions. (b) Marginal distribution at sampling location 1. (c) Marginal distribution at sampling location 2. (d) Marginal distribution at sampling location 3. (e) Marginal distribution at sampling location 4.

A.5 convergence test for selecting the polynomial order and the number of gauss nodes and weights in engineering case 1

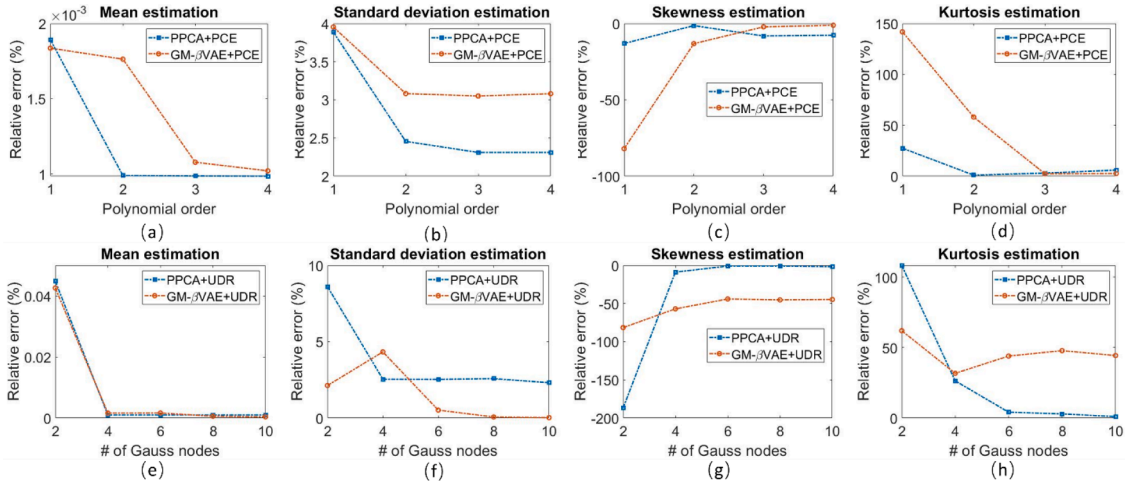


Fig. A.3. Case 1: Relative errors of (a) the mean and (b) the standard deviation (c) the skewness (d) the kurtosis estimations using different order of polynomials. Relative errors of (e) the mean and (f) the standard deviation (g) the skewness (f) the kurtosis estimations using different number of Gauss nodes and weights.

A.6 Convergence test for selecting the polynomial order and the number of Gauss nodes and weights in engineering case 2

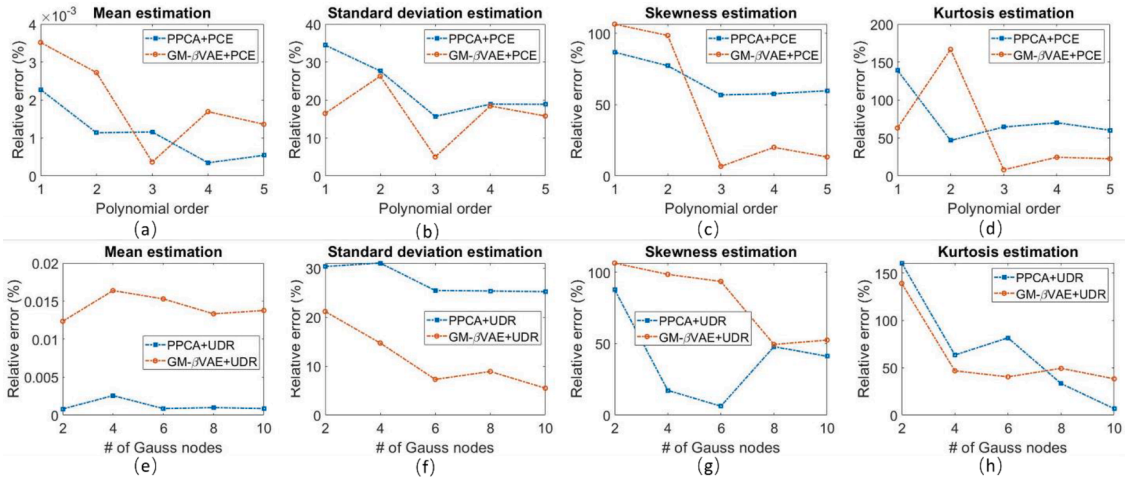


Fig. A.4. Case 2: Relative errors of (a) the mean and (b) the standard deviation (c) the skewness (d) the kurtosis estimations using different order of polynomials. Relative errors of (e) the mean and (f) the standard deviation (g) the skewness (f) the kurtosis estimations using different number of Gauss nodes and weights.

A.7 Engineering case study for selecting dimension reduction methods

The purpose of this supplementary case study is to compare the dimension reduction methods for parameterizing the random field uncertainties in a topological spatial domain. A simulation model is established to predict the structural distortions in lattice structures printed in batch by powder bed fusion (PBF). Multiple structures of the same design are printed on one building plate. The material is Ti6Al4V. The simulation model is implemented in software ABAQUS, which simulates the process of powder fusion and solidification, and result in the residual stress and distortion in the final printed parts. The uncertainties of the printing process come from two resources: (1) the uncertainties in material properties and process parameters; (2) the uncertainties in local thermal conditions. Detailed information of the simulation parameters can be found in our previous work [17].

As shown in Fig. A.5(a), six lattice structures are printed together on one building plate, therefore, one simulation yields six samples. In total, we generate 54 samples for training the uncertainty representation models. We group all structural samples into three groups based on their location on the building plate: the outer layers, the intermediate layers, and the inner layers (Fig. A.5(a)). The Y-distortions of the structure are selected as the quantity of interest. The dimensions of the lattice structures are 102 mm × 51 mm × 2 mm (Fig. A.5(b)). The Y-distortion values at each element of the cellular structure are extracted from the simulation results. At each element location in the 2D topological domain, the average Y-distortion value along the thickness direction is

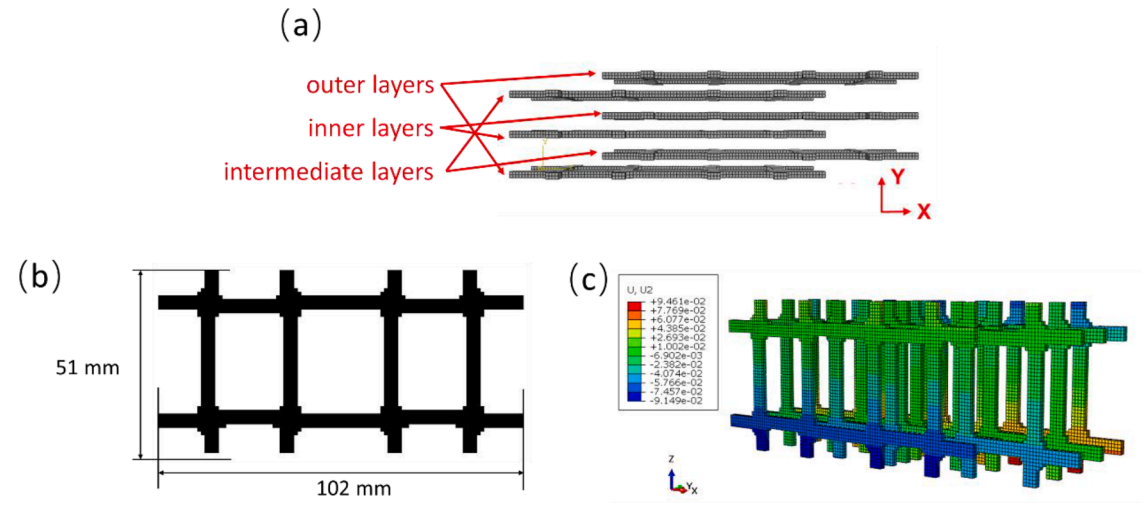


Fig. A.5. (a) Arrangement of six cellular structures on the building plate. (b) Dimensions of the 2D topological structural domain. (c) Y-axis distortions in the lattice structure.

Table A.3
Reconstruction accuracies measured by the R squared values. PPCA, βVAE, and GM-βVAE with different number of latent variables are compared in this table.

# of latent variables	2	3	4	5
PPCA	0.9669	0.9835	0.9901	0.9941
βVAE	0.0032	0.0822	0.1244	0.1596
GM-βVAE	0.9440	0.9526	0.9621	0.9684

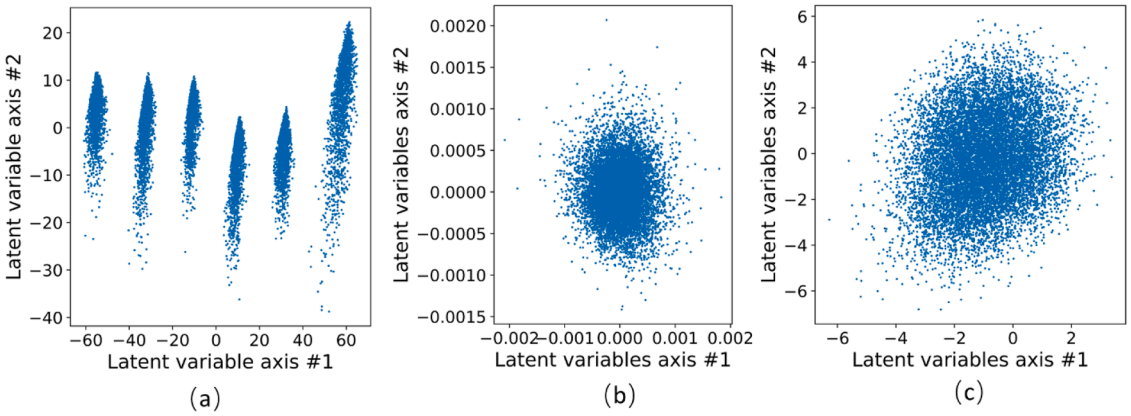


Fig. A.6. Latent variable distributions along latent axis #1 and #2 using (a) PPCA, (b) βVAE, (c) GM-βVAE.

Table A.4
Reconstruction accuracies measured by the R squared values of PPCA, βVAE, and GM-βVAE with different number of latent variables.

# of latent variables	2	4	6	8
PPCA	0.5837	0.7906	0.8918	0.9407
βVAE	0.5379	0.7424	0.8612	0.9079
GM-βVAE	0.9852	0.9911	0.9951	0.9987

used.

Two comparative studies are conducted using different subsets of the data. In the first case, we use the Y-distortion data for all the samples (multimodal distribution) to test the parameterization methods. In the second case, we only take the samples in the outer layers (unimodal distribution).

Parameterization based on multimodal data (all samples)

For each group (outer layers, intermediate layers, and inner layers), a Gaussian random field model with the kernel function of Eq. (3) is fitted to describe the spatial distribution of Y-distortions within the 2D lattice domain. By sampling the three Gaussian random fields each with 4000 random realizations, a data set of 12,000 random realizations, which is sufficiently large to be statistically invariant, are obtained for the following step of parameterization.

Three parameterization methods, PPCA, βVAE, and GM-βVAE, are implemented and compared. PPCA and βVAE compress the random field data to a low-dimensional feature space formed by normally distributed, independent random variables. GM-βVAE compresses the random field

data to a low-dimensional feature space formed by variables of the multi-variate Gaussian distribution. The performances of the three parameterization methods are evaluated based on the information loss, which can be quantified by the reconstruction accuracies, as shown in Table A.3. The PPCA has the highest reconstruction accuracy. However, the distribution of the latent vectors compressed by PPCA is not normally distributed, which indicate the failure in PPCA training (Fig. A.6). Therefore, PPCA is not suitable for this case study. βVAE has normally distributed latent variables, while the reconstruction accuracy is not acceptable. GM-βVAE has a satisfied reconstruction accuracy.

Parameterization based on unimodal data (outer layer samples only)

In total, we have 18 outer layer samples. A Gaussian random field model with the kernel function of Eq. (3) is fitted on the 18 samples to represent the spatial distribution of Y-distortion values in the outer layer lattice samples. By sampling the Gaussian random field model, we obtain a sample set with 12,000 random realizations, which is sufficiently large to be statistically invariant.

The random field realization data are compressed to the low dimensional feature space using PPCA, βVAE, and GM-βVAE, respectively. The reconstruction accuracies are shown in Table A.4. PPCA has relatively better accuracy compared with βVAE. Among all three methods, GM-βVAE works the best.

CRediT authorship contribution statement

Zihan Wang: Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Mohamad Daeipour:** Resources, Software, Writing – review & editing. **Hongyi Xu:** Conceptualization, Methodology, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

We gratefully acknowledge the financial support from the National Science Foundation (CAREER Award CMMI-2142290). We are grateful to ESI-Group for making an academic license for ProCast available to the UConn Casting and Solidification Research Group. We also sincerely thank Dr. Harold D. Brody for his insights and advices on this work.

References

[1] Luo Y, et al. Non-probabilistic uncertainty quantification and response analysis of structures with a bounded field model. Comput Methods Appl Mech Eng 2019;347:663–78.
[2] Chen S, Chen W, Lee S. Level set based robust shape and topology optimization under random field uncertainties. Struct Multidiscip Optim 2010;41(4):507–24.
[3] Schueller G. On the treatment of uncertainties in structural mechanics and analysis. Comput Struct 2007;85(5–6):235–43.
[4] Xi Z, et al. Random field modeling with insufficient field data for probability analysis and design. Struct Multidiscip Optim 2015;51(3):599–611.
[5] Hess PE, et al. Uncertainties in material and geometric strength and load variables. Nav Eng J 2002;114(2):139–66.
[6] Holický M, Retief JV, Sýkora M. Assessment of model uncertainties for structural resistance. Probab Eng Mech 2016;45:188–97.
[7] Nguyen KT, Medjaher K, Gogu C. Probabilistic deep learning methodology for uncertainty quantification of remaining useful lifetime of multi-component systems. Reliab Eng Syst Saf 2022;222:108383.
[8] Lin M, et al. Battery health prognosis with gated recurrent unit neural networks and hidden Markov model considering uncertainty quantification. Reliab Eng Syst Saf 2023;230:108978.
[9] Zhang K, et al. An efficient reliability analysis method for structures with hybrid time-dependent uncertainty. Reliab Eng Syst Saf 2022;228:108794.
[10] Greene MS, et al. A generalized uncertainty propagation criterion from benchmark studies of microstructured material systems. Comput Methods Appl Mech Eng 2013;254:271–91.
[11] Bostanabad R, et al. Computational microstructure characterization and reconstruction: review of the state-of-the-art techniques. Prog Mater Sci 2018;95:1–41.
[12] Wei X, Du X. Uncertainty analysis for time-and space-dependent responses with random variables. J Mech Des 2019;141(2).
[13] Xi Z. Model-based reliability analysis with both model uncertainty and parameter uncertainty. J Mech Des 2019;141(5):051404.
[14] Asadpoure A, Tootkaboni M, Guest JK. Robust topology optimization of structures with uncertainties in stiffness–Application to truss structures. Comput Struct 2011;89(11–12):1131–41.
[15] Yuan X, et al. Efficient imprecise reliability analysis using the Augmented Space Integral. Reliab Eng Syst Saf 2021;210:107477.
[16] Richardson J, Coelho RF, Adriaenssens S. A unified stochastic framework for robust topology optimization of continuum and truss-like structures. Eng Optim 2016;48(2):334–50.
[17] Wang Z, Xu H. Quantitative Representation of Aleatoric Uncertainties in Network-Like Topological Structural Systems. J Mech Des 2021;143(3).
[18] Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat 2010;2(4):433–59.
[19] Tipping ME, Bishop CM. Probabilistic principal component analysis. J R Stat Soc Series B Stat Methodol 1999;61(3):611–22.
[20] Liu Y, et al. Uncertainty quantification for Multiphase-CFD simulations of bubbly flows: a machine learning-based Bayesian approach supported by high-resolution experiments. Reliab Eng Syst Saf 2021;212:107636.
[21] Lee D, Seung HS. Algorithms for non-negative matrix factorization. Adv Neural Inf Process Syst 2000:13.
[22] Izenman, A.J., *Linear discriminant analysis*, in *Modern multivariate statistical techniques*. 2013, Springer. p. 237–80.
[23] Baudat G, Anouar F. Generalized discriminant analysis using a kernel approach. Neural Comput 2000;12(10):2385–404.
[24] Zhai, J., et al. *Autoencoder and its various variants*. in *2018 IEEE international conference on systems, man, and cybernetics (SMC)*. 2018. IEEE.
[25] Yang Z, Baraldi P, Zio E. A method for fault detection in multi-component systems based on sparse autoencoder-based deep neural networks. Reliab Eng Syst Saf 2022;220:108278.
[26] González-Muñiz A, et al. Health indicator for machine condition monitoring built in the latent space of a deep autoencoder. Reliab Eng Syst Saf 2022;224:108482.
[27] Liu M, Grana D, de Figueiredo LP. Uncertainty quantification in stochastic inversion with dimensionality reduction using variational autoencoder. Geophysics 2022;87(2):M43–58.
[28] Kokkolaras, M., Z.P. Mourelatos, and P.Y. Papalambros. *Design optimization of hierarchically decomposed multilevel systems under uncertainty*. in *International design engineering technical conferences and computers and information in engineering conference*. 2004.
[29] Thoft-Cristensen, P. and M.J. Baker, *Structural reliability theory and its applications*. 2012: Springer Science & Business Media.
[30] Der Kiureghian A. Structural reliability methods for seismic safety assessment: a review. Eng Struct 1996;18(6):412–24.
[31] Madsen, H.O., S. Krenk, and N.C. Lind, *Methods of structural safety*. 2006: Courier Corporation.
[32] Bucher CG. Adaptive sampling—An iterative fast Monte Carlo procedure. Struct Saf 1988;5(2):119–26.
[33] Engelund S, Rackwitz R. A benchmark study on importance sampling techniques in structural reliability. Struct Saf 1993;12(4):255–76.
[34] Melchers R. Importance sampling in structural systems. Struct Saf 1989;6(1):3–10.
[35] Ghanem, R.G. and P.D. Spanos, *Stochastic finite elements: a spectral approach*. 2003: Courier Corporation.
[36] Hasofer AM, Lind NC. Exact and invariant second-moment code format. J Eng Mech Div, Am Soc Civ Eng 1974;100(1):111–21.
[37] Fiessler B, Neumann H-J, Rackwitz R. Quadratic limit states in structural reliability. J Eng Mech Div, Am Soc Civ Eng 1979;105(4):661–76.
[38] Evans DH. An application of numerical integration techniques to statistical tolerancing, III—General distributions. Technometrics 1972;14(1):23–35.
[39] Seo HS, Kwak BM. Efficient statistical tolerance analysis for general distributions using three-point information. Int J Prod Res 2002;40(4):931–44.
[40] Rahman S, Xu H. A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics. Probab Eng Mech 2004;19(4):393–408.
[41] Delgado J, et al. Hill chart modelling using the Hermite polynomial chaos expansion for the performance prediction of pumps running as turbines. Energy Convers Manage 2019;187:578–92.
[42] Su Q, Strunz K. Stochastic circuit modelling with Hermite polynomial chaos. Electron Lett 2005;41(21):1163–5.
[43] Xiu D, Karniadakis GE. Modeling uncertainty in flow simulations via generalized polynomial chaos. J Comput Phys 2003;187(1):137–67.

[44] Mara TA, Becker WE. Polynomial chaos expansion for sensitivity analysis of model output with dependent inputs. Reliab Eng Syst Saf 2021;214:107795.
[45] Zhang J, et al. Efficient reliability analysis using prediction-oriented active sparse polynomial chaos expansion. Reliab Eng Syst Saf 2022;228:108749.
[46] Maier HR, et al. First-order reliability method for estimating reliability, vulnerability, and resilience. Water Resour Res 2001;37(3):779–90.
[47] Der Kiureghian A, Lin H-Z, Hwang S-J. Second-order reliability approximations. J Eng Mech 1987;113(8):1208–25.
[48] Lee I, et al. Dimension reduction method for reliability-based robust design optimization. Comput Struct 2008;86(13–14):1550–62.
[49] Acar E, Rais-Rohani M, Eamon CD. Reliability estimation using univariate dimension reduction and extended generalised lambda distribution. Int J Reliab Saf 2010;4(2–3):166–87.
[50] Zou B, Xiao Q. Probabilistic load flow computation using univariate dimension reduction method. Int Trans Electric Energy Syst 2014;24(12):1700–14.
[51] Xiao Q, et al. Point estimate method based on univariate dimension reduction model for probabilistic power flow computation. IET Generat Transm Distribut 2017;11(14):3522–31.
[52] Kalos, M.H. and P.A. Whitlock, *Monte carlo methods*. 2009: John Wiley & Sons.
[53] Zhang J. Modern Monte Carlo methods for efficient uncertainty quantification and propagation: a survey. Wiley Interdiscip Rev Comput Stat 2021;13(5):e1539.
[54] Lu J, et al. Uncertainty propagation of frequency response functions using a multi-output Gaussian Process model. Comput Struct 2019;217:1–17.
[55] Oliver, T.A. and R.D. Moser. *Bayesian uncertainty quantification applied to RANS turbulence models*. in *Journal of physics: conference series*. 2011. IOP Publishing.
[56] Azzimonti D, et al. Quantifying uncertainties on excursion sets under a Gaussian random field prior. SIAM/ASA J Uncertain Quantificat 2016;4(1):850–74.
[57] Xu H. Constructing Oscillating Function-Based Covariance Matrix to Allow Negative Correlations in Gaussian Random Field Models for Uncertainty Quantification. J Mech Des 2020;142(7):074501.
[58] Kleder, M. *Shortest Path with Obstacle Avoidance (ver 1.3)* 2022 [cited 2022 February 15].
[59] Morton K, Sobey I. Discretization of a convection-diffusion equation. IMA J Numer Anal 1993;13(1):141–60.
[60] Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst 1987;2(1–3):37–52.
[61] Nyamundanda G, Brennan L, Gormley IC. Probabilistic principal component analysis for metabolomic data. BMC Bioinf 2010;11(1):1–11.
[62] Chang WC. On using principal components before separating a mixture of two multivariate normal distributions. J R Stat Soc Ser C Appl Stat 1983;32(3):267–75.
[63] McLachlan GJ, Lee SX, Rathnayake SI. Finite mixture models. Annu Rev Stat Appl 2019;6:355–78.
[64] Roweis S. EM algorithms for PCA and SPCA. Adv Neural Inf Process Syst 1997:10.
[65] Dai B, et al. Connections with robust PCA and the role of emergent sparsity in variational autoencoder models. J Mach Learn Res 2018;19(1):1573–614.
[66] Kingma, D.P. and M.J.a.p.a. Welling, *Auto-encoding variational bayes*. 2013.
[67] Higgins, I., et al., *beta-vae: learning basic visual concepts with a constrained variational framework*. 2016.
[68] Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 2013;35(8):1798–828.
[69] De Maesschalck R, Jouan-Rimbaud D, Massart DL. The mahalanobis distance. Chemom. Intell Lab Syst 2000;50(1):1–18.
[70] Wiener N. The homogeneous chaos. Am J Math 1938;60(4):897–936.
[71] Wiener N, Wintner A. The discrete chaos. Am J Math 1943;65(2):279–98.
[72] Walters, R.W. and L. Huyse, *Uncertainty analysis for fluid mechanics with applications*. 2002.
[73] Hosder S, Walters R, Perez R. A non-intrusive polynomial chaos method for uncertainty propagation in CFD simulations. In: 44th AIAA aerospace sciences meeting and exhibit; 2006.
[74] Crestaux T, Maıtre OLe, Martinez J-M. Polynomial chaos expansion for sensitivity analysis. Reliab Eng Syst Saf 2009;94(7):1161–72.
[75] Feinberg J, Langtangen HP. Chaospy: an open source tool for designing methods of uncertainty quantification. J Comput Sci 2015;11:46–57.
[76] Lacor, C. and É. Savin, *General introduction to polynomial chaos and collocation methods*, in *uncertainty management for robust industrial design in aeronautics*. 2019, Springer. p. 109–22.
[77] Hosder S, Walters R, Balch M. Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables. In: 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference; 2007.
[78] Giunta A, Wojtkiewicz S, Eldred M. Overview of modern design of experiments methods for computational simulations. In: 41st Aerospace Sciences Meeting and Exhibit; 2003.
[79] Krylov, V.I. and A.H. Stroud, *Approximate calculation of integrals*. 2006: Courier Corporation.
[80] Li G, Zhang K. A combined reliability analysis approach with dimension reduction method and maximum entropy method. Struct Multidiscip Optim 2011;43(1):121–34.
[81] Ditlevsen O, Madsen HO. Structural reliability methods, 178. New York: Wiley; 1996.
[82] Rosenblatt M. Remarks on a multivariate transformation. Annal Math Stat 1952;23(3):470–2.
[83] Der Kiureghian A, Liu P-L. Structural reliability under incomplete probability information. J Eng Mech 1986;112(1):85–104.
[84] Youn BD, Choi KK, Park YH. Hybrid analysis method for reliability-based design optimization. Int J Mech Mater Des 2003;125(2):221–32.
[85] Xu H, Rahman S. A generalized dimension-reduction method for multidimensional integration in stochastic mechanics. Int J Numer Methods Eng 2004;61(12):1992–2019.
[86] Lee SH, Chen W. A comparative study of uncertainty propagation methods for black-box-type problems. Struct Multidiscip Optim 2009;37(3):239.
[87] Pearson K. VII. Note on regression and inheritance in the case of two parents. Proc R Soc Lond 1895;58(347–352):240–2.
[88] Flemings MC. Solidification processing. Metallurgic Mater Trans B 1974;5(10):2121–34.
[89] Wang J, et al. Shrinkage porosity criteria and optimized design of a 100-ton 30Cr2Ni4MoV forging ingot. Mater Des 2012;35:446–56.
[90] Dabade UA, Bhedasgaonkar RC. Casting defect analysis using design of experiments (DoE) and computer aided casting simulation technique. Procedia CIRP 2013;7:616–21.
[91] Jafari H, et al. In situ melting and solidification assessment of AZ91D granules by computer-aided thermal analysis during investment casting process. Mater Des 2013;50:181–90.
[92] Malhotra V, Kumar Y. Casting defects: an literature. Technology (Singap World Sci) 2016;7(1):60–2.
[93] Luan Y, et al. Effect of solidification rate on the morphology and distribution of eutectic carbides in centrifugal casting high-speed steel rolls. J Mater Process Technol 2010;210(3):536–41.
[94] Kosec B, Kosec G, Sokovic M. Temperature field and failure analysis of die-casting die. Arch Comput Mater Sci Surf Eng 2007;28(3):182–7.
[95] Pehlke RD, Jeyarajan A, Wada H. Summary of thermal properties for casting alloys and mold materials. NASA STI/Recon Technical Report N 1982;83:36293.
[96] Gilman MJ. A brief survey of stopping rules in Monte Carlo simulations. Inst Electric Electron Eng (IEEE) 1968.