Recent Advances in Surrogate Modeling Methods for Uncertainty Quantification and Propagation

Chong Wang 1,*, Xin Qiang 1, Menghui Xu 2 and Tao Wu 3

1 Institute of Solid Mechanics, Beihang University, Beijing 100191, China; xinqiang@buaa.edu.cn
2 Faculty of Mechanical Engineering & Mechanics, Ningbo University, Ningbo 315211, China; xumenghui@nbu.edu.cn
3 Department of Mechanical Science and Engineering, Technische Universität Dresden, Holbeinstraße 3, 01307 Dresden, Germany; wutao202@hotmail.com

* Correspondence: chongwang@buaa.edu.cn

Abstract: Surrogate-model-assisted uncertainty treatment practices have been the subject of increasing attention and investigations in recent decades for many symmetrical engineering systems. This paper delivers a review of surrogate modeling methods in both uncertainty quantification and propagation scenarios. To this end, the mathematical models for uncertainty quantification are firstly reviewed, and theories and advances on probabilistic, non-probabilistic and hybrid ones are discussed. Subsequently, numerical methods for uncertainty propagation are broadly reviewed under different computational strategies. Thirdly, several popular single surrogate models and novel hybrid techniques are reviewed, together with some general criteria for accuracy evaluation. In addition, sample generation techniques to improve the accuracy of surrogate models are discussed for both static sampling and its adaptive version. Finally, closing remarks are provided and future prospects are suggested.

Keywords: uncertainty quantification; uncertainty propagation; surrogate modeling; sampling strategy; symmetrical engineering systems

1. Introduction

Practical engineering structures are inevitably rife with diverse types of uncertainties related to model assumption, material property, loads, boundary conditions, etc. In general, these uncertainties can be classified into two categories: aleatory uncertainty and epistemic uncertainty [1]. In this context, aleatory uncertainty reveals the inherent variation in the system and is irreducible but can be described by a probability distribution. However, epistemic uncertainty reflects a lack of knowledge of the system and is reducible if more information is procured. Tackling these uncertainties in an effective manner has become a critical consideration for both practitioners and academicians.

Uncertainty quantification is a route of the quantitative representation of and reduction in uncertainties in both simulation and practical applications. With the perquisite of sufficient samples, the probabilistic model has always been the most popular scheme to tackle the aleatory uncertainty [2,3]. For those scenarios with scarce samples, various representative non-probabilistic theories have been developed for the treatment of epistemic uncertainty, including fuzzy theory, interval theory, convex models, evidence theory, and so on [4]. Apart from the traditional schemes, several recent studies have also suggested a more applicable hybrid framework to tackle coexisting uncertainties in engineering systems with increasing complexity [5–7].

As another important issue in uncertainty treatment, uncertainty propagation focuses on characterizing the impact of fluctuations in the input parameters on system responses. Compared with the probability distribution function, some easy-to-procure indicators, e.g., the statistical moments of response, are more applicable in engineering practices [8]. To
this end, various numerical methods have flourished and have been successfully applied in structural stochastic response analysis, reliability analysis, robust design, multidisciplinary optimization, etc. [9–11]. As the complexity of simulation models grows, however, the corresponding computational cost of these conventional numerical methods becomes increasingly unaffordable.

To alleviate the computing burden, the surrogate model (also known as the metamodel) has been an attractive alternative, wherein a cheap-to-run approximation model is constructed to replace the original time-consuming high-fidelity simulation. Until now, various surrogate models and auxiliary optimization algorithms have been developed to deliver better predictions [12,13]. In addition to pursuing a more accurate surrogate model, selecting appropriate samples is another way to help enhance the prediction accuracy, and thus the literature on sampling strategies of the surrogate model has also seen a rapid increase in recent decades [14].

This paper aims to provide a general review of the advances in both uncertainty treatment and the surrogate model in the past two decades. The remainder of this paper is organized as follows. Firstly, mathematical models for uncertainty quantification, including probabilistic, non-probabilistic and hybrid ones, are discussed in Section 2. In Section 3, numerical methods are divided into four categories to address their differences in uncertainty propagation. Section 4 presents several popular surrogate models and their hybrid strategies successively, together with a range of commonly used criteria for accuracy evaluation. For sampling strategies in surrogate modeling, Section 5 discusses the one-shot and sequential ones, respectively. Finally, Section 6 closes the paper by encapsulating the main points and concluding remarks.

2. Mathematical Models in Uncertainty Quantification

In this section, various uncertainty modeling techniques, including probabilistic, non-probabilistic and hybrid methods, are reviewed with a focus on their recent advances. Since probabilistic methods have been well studied and applied in engineering scenarios, the invariant/time-variant/space-variant characteristics of random parameters are summarized and discussed in this section. As a series of attractive tools in measuring the uncertainties with insufficient information, existing non-probabilistic methods are subsequently discussed, together with their fundamentals. Finally, hybrid strategies concerning multiple uncertain modeling techniques are classified with a closing review.

2.1. Probabilistic Models

Under the probabilistic framework, as shown in Figure 1, the statistical characteristic of uncertain parameters can be described by random variables, stochastic process or random fields [15]. The probability density functions (PDF) with different colors in Figure 1a illustrate the impact of Gaussian distribution parameters on random variables. Considering the time-/space-varying characteristics of the random variable, stochastic process and random fields are displayed in Figure 1b and 1c, respectively. According to the probability theory, a random variable \(X\) defined on the probability triple \((\Omega, \mathcal{F}, P)\) maps a random event \(\theta \in \Omega\) to real value in \(\mathbb{R}^n\) [15]. For brevity but without loss of generality, a collection of random variables \(\{H : H(\theta, w) = X_w(\theta), \theta \in \Omega\}\) can be exploited to represent the stochastic process or random field, where \(w\) is the index indicating time or position, respectively.

For many random parameters measured in engineering practices, such as geometry size, material properties and loads applied to structures, there are inevitably small fluctuations in their observed values. When sufficient samples are provided, the accurate PDF of these random variables can be easily procured to quantify their uncertainties. There are a variety of probability distributions widely adopted in engineering practices, including Gaussian, Poisson, and Weibull distributions [16–19]. For multidimensional random variables, the easy-to-procure incomplete information such as statistical moments and marginal probability distributions, instead of the precise joint probability distribution, is
often exploited to measure the randomness of parameters. To obtain the above statistical information, many probabilistic methods have been widely adopted, such as Nataf transformation and the Copula function [20,21]. Based on data-driven strategies, emerging machine learning-aided techniques also provide an attractive alternative for uncertainty quantification [22,23]. As an extension of deterministic finite element method (FEM), the stochastic FEM has been regarded as a powerful approach when solving the problems with random properties [24].

Figure 1. Three types of probabilistic methods.

Due to the changing environmental factors, the uncertainty of parameters often has time-varying characteristic in many practical problems, such as the aerodynamic heat on hypersonic vehicles and wind excitations of buildings [25,26]. In this context, the stochastic process model can be employed to conduct uncertainty modeling for these time-varying parameters. To approximate the stochastic process, many methods have been developed to simulate Gaussian or non-Gaussian and stationary or non-stationary processes [27,28]. The Markov process, which describes the transitions between a sequence of states, plays an increasingly important role in system survival/failure behavior evaluation [29]. In addition, recent investigations have shown the potential of data-driven methods in simulating the statistical properties of stochastic processes [30].

In addition to the above two descriptions, some researchers have also noticed that many parameters vary with spatial locations in practical cases, such as the mechanical properties of geotechnical materials and composite structures [31,32]. In this context, this spatially distributed uncertainty can be statistically described by means of random fields. To simulate the spatial-varying uncertainty, various random field discretization techniques have been applied in geotechnical engineering and structural vibration analysis, such as local average subdivision, turning-band methods, and Karhunen–Loève (KL) expansion [31,33–35]. The stochastic spectral element method (SEM) has also been proven to be a powerful approach to tackle the spatial uncertainty in dynamic systems [36–38]. The success of machine learning techniques is also spreading their application to random fields [39].

2.2. Non-Probabilistic Models

Probabilistic approaches have offered a stable framework for uncertainty analysis, with the premise that sufficient experimental samples are available to construct the precise probability distribution of uncertain parameters. However, the experimental conditions or costs of many practical engineering problems often restrict the acquisition of adequate data, which further leads to the inapplicability of the probabilistic method [40,41]. To procure more credible results for engineering problems with insufficient data, many non-probabilistic methods have flourished in recent decades, providing an attractive framework for uncertainty analysis.

Fuzzy set theory. In a large number of real-world problems, the relationship between elements and sets is sometimes vague, and it is difficult to give a crisp partition. Conventional set theory exploits the binary \{0, 1\} to describe the degree of an element \(x\) with respect to the set \(A\). In contrast, fuzzy set theory leverages a real number \(\mu_A(x)\) in the
closed interval $[0, 1]$ to quantitatively measure this affiliation degree more precisely. On this basis, the fuzzy set $A^f$ can be defined as

$$A^f = \{(x, \mu_{A^f}(x)) | x \in A, \mu_{A^f}(x) \in [0, 1]\}$$

(1)

where the superscript ‘$f$’ is the symbol of fuzzy set; the reference set $A$ is called the universe of discourse; the real-valued function $\mu_{A^f}$ is called a membership function; $\mu_{A^f}(x)$ represents the membership degree of element $x$ to the fuzzy set $A^f$. The techniques used to generate the membership function have received extensive investigations, and there are many widely used membership functions such as Gaussian, triangular and trapezoidal functions [42]. In addition to the conventional fuzzy sets, many extended fuzzy sets such as intuitionistic and hesitant ones have emerged in recent decades [43,44]. Considering the time-variant properties in some fuzzy variables, many scholars have also begun to focus on the time-dependent fuzzy uncertainty [45,46]. With its high degree of maturity, fuzzy set theory has been extensively applied in structural response analysis, reliability assessment, parameter identification, etc. [45,47,48].

Interval theory. In many engineering problems, the values of external loads or material properties often fluctuate within a certain range. With interval theory, these uncertain-but-bounded parameters can be described as follows

$$x^I = (x^I_i)_m = ([x_i, x_i])_m = (x^I_i + \Delta x^I_i)_m i = 1, 2, \ldots, m$$

(2)

where $x^I$ denotes the interval vector composed of $m$ interval variables; $x_i$ and $x_i$ are called the lower and upper bound of the interval variable $x^I_i$, respectively; and $x^I_i = (x_i + x_i)/2$ and $\Delta x^I_i = (x_i - x_i)/2$ are the midpoint and the radius of $x^I_i$, respectively. Since the interval model only needs the bounds of the parameters to quantify the uncertainty, it has been widely applied in various engineering fields [49–51]. Similar to the probabilistic methods, existing interval-related studies can also be divided into those on interval variables [49], interval process [52] and interval fields [51].

Ellipsoid model. To remedy the deficiency that the interval model can only handle independent variables, the ellipsoid model has been applied to tackle various engineering problems with dependent variables [41,53,54]. The explicit mathematical formula to represent the ellipsoid model $\Omega_e$ is given as follows

$$\Omega_e = \left\{ x | (x - x^0)^T G_e (x - x^0) \leq 1 \right\}$$

(3)

where $x = (x_1, x_2, \ldots, x_m)^T$ stands for the variable vector in $m$-dimensional space; $G_e$ and $x^0$ are the characteristic matrix and the centroid of the ellipsoid, respectively; and the characteristic matrix $G_e$ measures the shape and orientation of the ellipsoid. When conducting the ellipsoid-based uncertainty quantification, the optimal ellipsoid model is considered to envelop all experimental samples with a minimal volume. In this context, many ellipsoid-modeling techniques have flourished, such as the rotation matrix method, the correlation approximation method and data-driven methods [55–58]. Theoretically, efficiently constructing a reasonable ellipsoid model under high-dimensional space remains a challenging issue.

Evidence theory. The evidence theory, also known as Dempster–Shafer theory, has seen increasing applications due to its advantages in flexibly dealing with imprecise and incomplete uncertain information from multiple sources [59–61]. Two important measures, i.e., belief and plausibility, are considered in the evidence theory for each proposition $A$ in the frame of discernment $\Theta$.

$$Bel(A) = \sum_{\xi | \Theta \subseteq A} m(\xi), Pl(A) = \sum_{\xi | \Theta \cap A \neq \emptyset} m(\xi)$$

(4)
where \( m(\xi) \) is interpreted as the basic probability assignment (BPA) of possible proposition \( A \); \( Bel(A) \) denotes the aggregate of \( m(\xi) \) values that totally support the proposition \( A \); \( Pl(A) \) represents the aggregate of \( m(\xi) \) values that totally or partially support the proposition \( A \). Generation techniques of BPA, which play a key role in practical applications, have also received extensive investigations [62]. Considering that existing evidence inevitably conflicts in multi-source information, ways of measuring and fusing these inconsistencies more reasonably have been a research focus in recent decades [63].

Rough set theory. Rough set theory is recognized as a promising technique for uncertainty management, especially for those uncertainties with incomplete and inconsistent information. By exploiting a boundary region of a set, the classical rough set theory defines the following two operations to express vagueness:

\[
R(X) = \{ x \in U : R(x) \subseteq X \}, \quad \overline{R}(X) = \{ x \in U : R(x) \cap X \neq \emptyset \}
\]  

where \( R(\cdot) \) is an equivalence relation (also known as indiscernibility relation) on the universe \( U \); \( x \) is an arbitrary element in a subset \( X \subseteq U \); \( R(X) \) and \( \overline{R}(X) \) are named the \( R \)-lower and \( R \)-upper approximations of \( X \), respectively. As the above equivalence relations are too stringent, many scholars have proposed more general rough sets (e.g., probabilistic-rough, fuzzy-rough and rough-soft sets) for engineering applications [64–66].

The rough set model has been widely applied in decision making, attribute reduction and fault diagnosis [65,67,68].

In addition to the above models, other types of non-probabilistic methods have also been developed in terms of various requirements for uncertainty modeling, such as possibility theory (see [69]), the information gap model (see [70]), the parallelepiped model (see [71]), the super ellipsoid model (see [72]) and the convex polyhedral model (see [73]). To intuitively exhibit the characteristic of different non-probabilistic convex modeling techniques in uncertainty quantification, the envelope results of experimental samples under five types of convex models are shown in Figure 2.

![Figure 2. Envelopes of experimental samples using five types of convex models.](image)

### 2.3. Hybrid Models

Overall, significant studies have investigated uncertainty quantification based on both probabilistic and non-probabilistic schemes. In some engineering practices, however, diverse types of uncertainty often coexist in a unified system, and using conventional single-type models becomes incompetent [2,40,74]. In this context, a variety of hybrid uncertainty modeling strategies have been increasingly attended and investigated in recent years. Theoretically, they can be divided into two categories: parallel and embedded hybrid methods. As a common hybrid case, parallel-type hybrid methods enable different uncertain parameters exist in systems simultaneously and independently. In contrast, the embedded-type ones are more general when dealing with various coexisting uncertainties...
but also remain technically challenging at the same time [40,75]. For a better understanding, Table 1 summarizes related investigations on the above hybrid strategies in past five years.

Table 1. Hybrid strategies for dual-uncertainty modeling in past five years.

<table>
<thead>
<tr>
<th>Hybrid Strategy</th>
<th>Uncertainty Involved</th>
<th>Descriptions</th>
<th>Reference Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel-type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probabilistic interval</td>
<td>Mutually independent</td>
<td>[2,76]</td>
<td></td>
</tr>
<tr>
<td>Probabilistic fuzzy</td>
<td>Mutually independent</td>
<td>[77]</td>
<td></td>
</tr>
<tr>
<td>Fuzzy interval</td>
<td>Mutually independent</td>
<td>[6]</td>
<td></td>
</tr>
<tr>
<td>Probabilistic convex</td>
<td>Mutually independent</td>
<td>[78]</td>
<td></td>
</tr>
<tr>
<td>Fuzzy evidence</td>
<td>Mutually independent</td>
<td>[7,79]</td>
<td></td>
</tr>
</tbody>
</table>

| Embedded-type   |                      |              |                 |
|-----------------|                      |              |                 |
| Probabilistic interval | Distribution parameters of random variables are interval values | [2,80] |
| Probabilistic fuzzy | Distribution parameters of random variables are fuzzy numbers | [40,81] |
| Fuzzy interval  | Interval variables with fuzzy bounds | [75,82] |
| Fuzzy evidence | Evidence variables with fuzzy focal elements | [83] |

Take the probabilistic-interval hybrid model as an example; the parallel type is suitable for those problems with determined probability distribution types but inaccurate parameter values. In contrast, the embedded type is suitable for solving problems where only the fluctuation intervals can be obtained for some parameters due to the lack of available samples or expert experience. In practice, it is usually strenuous to choose the most reasonable hybrid strategy for a certain problem and the selection criteria of these two types of hybrid models have not been reported. With the increasing complexity of uncertainties in multidisciplinary practices, the study of engineering problems concerning more than two kinds of uncertainties under a unified framework is promising but mostly unexplored [3,5].


After the results of uncertainty quantification are obtained, the next priority is to measure the impact of disturbances in the input parameters on the system responses, i.e., uncertainty propagation. In this section, several popular numerical methods in uncertainty propagation are reviewed, including sampling-based, expansion-based, optimization-based and integration-based ones. They are schematically summarized in the following subsection.

3.1. Sampling-Based Method

Sampling-based techniques predominantly include Monte Carlo simulation (MCS) and its variants. As shown in Figure 3, MCS generates random variables from probability density functions to estimate unknown parameters and then calculate their expected value and variances. Based on the law of large numbers and the central limit theorem, one usually works with the multivariable normal distribution. In practice, the MCS procedure remains in essence identical regardless of the complexity or the computational burden of the simulation model. Theoretically, MCS requires extensive samplings (usually $10^6$ runs or greater) to procure reliable statistics, which implies that it is often computationally expensive for complex engineering problems [84]. By updating a Markov chain with desired distribution as the stationary distribution, the Markov Chain Monte Carlo (MCMC) method aims to recover the properties of an unknown probability distribution and is deemed an
important complement to the problems with known probability distribution. Both MCS and MCMC have been widely applied to obtain reliable results in various domains [85–87].

Figure 3. Diagram of Monte Carlo simulation [88].

Another important family of Monte Carlo techniques is importance sampling (IS) and its adaptive version (AIS). As a variance reduction technique, IS utilizes a targeted sampling strategy to reduce the number of model evaluations. In contrast, AIS focuses on employing the information of previously drawn samples to adjust proposals to further reduce the variance in the desired estimators. For applications of importance sampling in reliability analysis, one can consult [89–91].

3.2. Expansion-Based Method

The perturbation method (also known as the small-parameter expansion method) expresses the desired solution in terms of a formal power series (i.e., perturbation series) in small parameter $\epsilon$ that measures the deviation from the fully solvable problem [92]. A pivotal idea of this method is an intermediate operation that breaks the problem into ‘solvable’ and ‘perturbative’ parts. On this basis, the full solution $A$ can be subsequently represented by a series in $\epsilon$ with the first-order perturbative correction:

$$A = A_0 + \epsilon A_1 + o(\epsilon^2)\quad (\epsilon \to 0)$$

(6)

where $A_0$ and $A_1$ indicate the know solution and the first-order terms, respectively, and $o(\epsilon^2)$ denotes the truncation error of high-order terms. As the basis of other expansion methods, parameter and subinterval perturbation strategy have been widely applied in heat transfer, structural-acoustic systems, etc. [6,93,94].

The Taylor series expansion method expands the system function $y = f(x)$ at a certain point $x_0$ and constructs a polynomial using each order’s partial derivative to approximately replace the original system function [95]. Due to its implementation-friendly characteristic, the first-order Taylor series expansion has been extensively used, in which the high-order terms are truncated:

$$y(x) \approx f(x_0) + \sum_{i=1}^{m} \frac{\partial f(x_0)}{\partial x_i} (x_i - x_{i0})$$

(7)

where $m$ indicates the dimension of the uncertain variable vector and $x_i$ and $x_{i0}$ are the $i$th components of $x$ and $x_0$, respectively. In practice, the first-order Taylor series is widely used to reduce the computational cost in heat conduction, structural vibration analysis, etc. [96,97].

The truncated Neumann series can be employed for approximate matrix inversion by introducing a linear operator $T(x)$ [98]. The inverse of a matrix $A$ can be approximately written as:

$$A^{-1} \approx \sum_{i=0}^{n} (I - A)^i$$

(8)
where $T(x) = (I - A)x$ and $T$ satisfies the norm condition; $A^{-1}$ and $I$ are the inverse of $A$ and the identity matrix, respectively; and the Neumann series is truncated with $n$ terms. Neumann series have been employed to approximate the interval matrix inverse in acoustic field, coupled structural-acoustic field prediction, etc. [99,100].

Karhunen–Loève (KL) expansion represents the stochastic process as an infinite linear combination of orthogonal functions, similar to the Fourier series representation of functions on a bounded interval [27,101]. Assume that $X_t$ is a zero-mean square-integrable stochastic process defined over the probability space on a closed interval $[a, b]$ and admits the following decomposition:

$$X_t = \sum_{k=1}^{\infty} Z_k e_k(t)$$

where $Z_k$ are pairwise uncorrelated random variables and the real-valued functions $e_k$ are continuous on $[a, b]$ that are pairwise orthogonal in $L^2([a, b])$. KL expansion has been widely applied in dynamic uncertainty analysis, including stochastic process, random field models and interval process [101–103].

3.3. Optimization-Based Method

According to the type of uncertain variables involved, existing optimization-based methods in uncertainty propagation can be broadly classified into three main categories: stochastic programming, fuzzy programming and interval optimization [104].

Stochastic programming combines conventional deterministic optimization with random variables and probabilistic constraints, which often require time-consuming simulations [105]. Such problems can be calculated by either classical methods, such as nonlinear programming or quadratic programming, or other advanced methods, e.g., Nondominated Sorting Genetic Algorithm II (NAGA-II) or simulated annealing (SA) optimization [106–108].

Fuzzy programming broadly includes two main concerns: the possibilistic programming approach, using the possibility or necessity measures to convert the fuzzy mathematical programming problems into conventional ones, and the ordering-based approach, considering non-dominated solutions based on the ordering of fuzzy sets [109]. Many heuristic algorithms have been introduced in a variety of non-deterministic problems with fuzzy numbers, such as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Genetic Algorithm (GA) [110–113].

Interval optimization transforms the non-deterministic optimization problem into a deterministic, double-loop nested optimization problem. The outer optimizer is adopted to search the optimal design variable, while the inner optimizer is employed to compute the bounds of uncertain objective functions and constraints [114]. Though such processing has high precision, it suffers from expensive computational cost. For the consideration of efficiency, various approaches have been introduced to decouple the nested optimization into a single-layer one, including the degree of interval constraint violation (DICV) [115], the Karush–Kuhn–Tucker (KKT) condition [116], lightning attachment procedure optimization (LAPO) [117], affine arithmetic [118], etc.

In addition to the above uncertainty, optimization methods can also be incorporated into ellipsoid models [119], evidence theory [120], etc. Moreover, optimization schemes have received increasing applications in hybrid uncertainties, e.g., probabilistic-interval [121], probabilistic-fuzzy [122], and interval-fuzzy uncertainties [123].

3.4. Integration-Based Method

When determining the statistical moments of performance function, analytical solutions are often strenuous to obtain. In this context, various numerical integration methods are introduced to estimate the probability distribution of the system response. In general, the statistical moments of system response can be measured in two numerical manners: the point estimation method (PEM) and the dimension-reduction method (DRM).
For the first type, the statistical moments are evaluated by the weighted sum of values of the response functions at a set of collection points in random space. Appropriate strategies for selecting the computational nodes have been increasingly investigated to enhance computational efficiency. The simplest scheme is the full factorial numerical integration (FFNI) method, which utilizes a tensor product based on the one-dimensional quadrature rule. As the number of dimensions grows, however, it suffers from the well-recognized ‘curse of dimensionality’ [124]. In this context, the sparse grid numerical integration (SGNI) method constructed by the Smolyak algorithm is more popular in engineering problems [125]. In addition, several innovative techniques, e.g., adaptive SGNI [126], high-order unscented transformation [127], cubature formulation [128] and the quasi-symmetric point method [129], have also been developed for efficiency consideration. In such cases, considerable computing efforts are usually required on high-dimensional problems.

For the second type, the multiple-dimension integral is decomposed into a sum of several low-dimensional integrals. The univariate dimension-reduction method (UDRM) [130], the eigenvector UDRM [131] and the multiplicative UDRM (M-UDRM) [132] are considered the most popular ones due to their simplicity and efficiency in moderate nonlinear problems where few performance function calls are required. For the system with large random variations and high nonlinearity, multivariate DRM (e.g., the bivariate [133] and the trivariate DRM [134]) and adaptive DRM [135] can be employed to enhance the accuracy of the evaluation of moments. Although the above significant improvements have been achieved, the scheme that can strike a trade-off between accuracy and efficiency is still of great interest in the assessment of statistical moments.

State-of-the-art numerical simulations often involve complex mechanisms with a vast number of input parameters, where extensive runs of computational models incur unaffordable computing effort in many practical cases. In view of this issue, the surrogate model has been recognized as an attractive alternative to reduce the computing budget and has received sustained attention and wide applications in recent decades. Herein, two main concerns of the surrogate model—theoretical basis and sampling strategy—are discussed in the following two sections.

4. Theoretical Basis of Surrogate Model

Surrogate models are a series of easy-to-evaluate mathematical models that approximate the original time-consuming simulation models based on paired input–output experimental samples [41]. In this section, commonly used approaches for surrogate modeling are discussed with an emphasis on their recent advances. As one of the hotspots in surrogate models, state-of-the-art hybrid strategies are subsequently discussed. Finally, several popular accuracy evaluation criteria for the surrogate model are reviewed.

4.1. Commonly Used Surrogate Model

Polynomial response surface (PRS) model. This popular model is trained by the least-square method, which minimizes the variance of unbiased estimators of the coefficients by means of the conditions of the Gauss–Markov theorem. A typical second-order PRS model can be expressed as:

\[ \hat{g}(x) = \alpha_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} x_i x_j \]  

(10)

where \( x_i \) and \( x_j \) denote the \( i \)th and the \( j \)th components of the \( n \)-dimensional design variable, respectively; \( \alpha_0 \) is the constant term; \( \beta_{ii} \), \( \beta_{ij} \) stands for the coefficient of first-order term, the second-order term and the cross term, respectively.

When establishing the PRS model, the coefficients are interpreted as the significance of different terms. The remarkable smoothing capability of PRS model enables the fast convergence of noisy functions. Although the PRS model is simple and implementation-
friendly, a main drawback lies in its applications for high nonlinear problems. In such cases, a vast number of samples are usually required to estimate the coefficients of PRS model, and high-order polynomials may cause instabilities. In practice, linear and second-order PRS model are commonly used ones.

Radial basis function (RBF) model. The RBF model exploits linear combinations of radial symmetric kernel functions to approximate the system function. A general RBF model can be expressed as:

$$
\tilde{g}(x) = \sum_{i=1}^{n} \omega_i \psi_i(\|x, x_i\|)
$$

(11)

where \( n \) is the number of sample points; \( \|x, x_i\| \) denotes the Euclidean distance between predicted point \( x \) and observed point \( x_i \); \( \psi_i(\cdot) \) represents the radial basis function on \( x_i \); \( \omega_i \) stands for the weight coefficient to be determined. A typical structure of RBF neural network is displayed in Figure 4. Here, \( \Gamma^* \) and \( \Theta^* \) denote the input and output sample respectively.

![Figure 4. Schematic diagram of RBF neural network [136].](image)

By substituting all observed sample points \( x_i, i = 1, 2, \ldots, n \) into Equation (11), a group of equations related to the unknown weight coefficients \( \omega_i \) can be obtained

$$
\begin{pmatrix}
\psi(\|x_1, x_1\|) & \psi(\|x_2, x_1\|) & \cdots & \psi(\|x_n, x_1\|) \\
\psi(\|x_1, x_2\|) & \psi(\|x_2, x_2\|) & \cdots & \psi(\|x_n, x_2\|) \\
\vdots & \vdots & \ddots & \vdots \\
\psi(\|x_1, x_n\|) & \psi(\|x_2, x_n\|) & \cdots & \psi(\|x_n, x_n\|)
\end{pmatrix}
\begin{pmatrix}
\omega_1 \\
\omega_2 \\
\vdots \\
\omega_n
\end{pmatrix} =
\begin{pmatrix}
\tilde{g}(x_1) \\
\tilde{g}(x_2) \\
\vdots \\
\tilde{g}(x_n)
\end{pmatrix}
$$

(12)

where \( \tilde{g}(x_i) \), \( i = 1, 2, \ldots, n \) denotes the response at the observed point \( x_i \), calculated by the original system function. RBF model is usually employed to interpolate scattered multivariate data and has shown satisfactory approximations for arbitrary forms of response functions. Various radial basis functions, e.g., linear, Gaussian and multi-quadratic [71], can be flexibly determined for diverse practical requirements.

Polynomial chaos expansion (PCE) model. The PCE model aims to project the random variable onto a stochastic space spanned by a set of orthogonal polynomial basis. Then, a prototypical \( p \)-order PCE for an \( m \)-dimensional random variable is represented as [60]:

$$
\tilde{g}(x) = \sum_{|\alpha| \leq p} b_\alpha \Psi_{\alpha}(x)
$$

(13)

where \( \Psi_{\alpha}(x) \) denotes the polynomial basis function; \( b_\alpha \) is the unknown expansion coefficient vector; \( |\alpha| \) is calculated by \( |\alpha| = a_1 + a_2 + \cdots + a_m \); the total number of expansion terms is \( C_m^{m+p} = (m + p)! / (m! p!) \).

Classical families of orthogonal polynomials have been developed and extensively used, such as Hermite, Legendre, Laguerre and Jacobi polynomials [137]. Different types
of orthogonal polynomials are shown in Table 2. In practice, both generalized polynomial chaos (gPC) and arbitrary polynomial chaos (aPC) frameworks have also seen a promising future in various engineering practices with different probability measures [138].

Table 2. Different types of orthogonal polynomials with distribution.

<table>
<thead>
<tr>
<th>Polynomial Type</th>
<th>Distribution</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hermite</td>
<td>Gaussian</td>
<td>(–∞, ∞)</td>
</tr>
<tr>
<td>Laguerre</td>
<td>Gamma</td>
<td>[0, ∞)</td>
</tr>
<tr>
<td>Jacobi</td>
<td>Beta</td>
<td>[a, b]</td>
</tr>
<tr>
<td>Legendre</td>
<td>Uniform</td>
<td>[a, b]</td>
</tr>
</tbody>
</table>

Kriging model. The Kriging, also known as Gaussian process regression, is an interpolation methodology based on Gaussian processes governed by prior covariance, as shown in Figure 5. A general form of Kriging can be formulated as a summation of two components: a trend of mean prediction determined by several basis functions at known locations and a random error with zero-mean distribution [139]:

\[
g(x) = \sum_{i=1}^{n} \beta_i f_i(x) + Z(x) \tag{14}
\]

where \(f_i(x)\) and \(\beta_i\) are the ith basis function and its corresponding coefficient, respectively; \(n\) is the number of sample points; and \(Z(x)\) denotes a Gaussian process with a zero-mean and covariance function formulated as

\[
\text{Cov}(Z(x_i), Z(x_j)) = \sigma^2 R(x_i, x_j, \theta) \tag{15}
\]

where \(\sigma^2\) stands for the variance of \(Z(x)\) and \(R(x_i, x_j, \theta)\) is the correlation function between \(Z(x_i)\) and \(Z(x_j)\) with hyper-parameters \(\theta\).

Figure 5. Schematic diagram of Kriging [140].

Kriging starts with a prior distribution over functions, and a set of spatial-related observation values are then obtained. By combining the Gaussian prior with a Gaussian likelihood function for each observed value, unknown value can be predicted at new spatial locations, together with their means and covariance. The correlation function \(R(x_i, x_j, \theta)\) can be specified with various forms, including linear, exponential, Gaussian, etc. According to the various stochastic process assumed, there are different types of Kriging models, namely, ordinary, simple, universal Kriging, their adaptive versions, etc. [141].
Support vector regression (SVR) model. The SVR model is a supervised machine learning model and is regarded as a special version of the support vector machine for regression. SVR utilizes the kernel function to map the original data onto a high-dimensional feature space and then searches the optimal regression function in a linear feature space. A general form of the SVR model is a sum of basis function $\psi(\cdot)$, with weighting coefficient $\omega_i$, added to a constant term $b$, which can be written as:

$$\hat{g}(x) = \sum_{i=1}^{n} \omega_i \psi_i(x, x_i) + b$$

(16)

This form of SVR is similar to that of the RBF and the Kriging model. However, the way to calculate the unknown parameters in SVR model differs significantly from them. The purpose of SVR is to find a function that can estimate the output value with a deviation less than $\epsilon$ from the real value. The corresponding band of deviation is called the $\epsilon$-tube. The optimal regression function is determined by formulating a mathematical optimization problem:

$$\min \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*)$$

s.t.  
$$-\epsilon - \xi_i \leq y_i - (\omega^T x_i + b) \leq \epsilon + \xi_i^* , \quad i = 1, 2, \ldots, n$$

$$\xi_i, \xi_i^* \geq 0$$

(17)

where $\omega^T x_i + b$ is a linear version of SVR, $y_i$ denotes the predicted value, slack variables $\xi_i$ and $\xi_i^*$ allow the existence of outliers outside the $\epsilon$-tube, and the regularization constant (also known as penalty coefficient) $C$ here achieves a trade-off between the model complexity and the empirical risk. For better understanding, a typical SVR model is shown in Figure 6.

![Figure 6. Schematic diagram of SVR.](image)

Artificial neural network (ANN) model. ANNs are computational systems inspired by biological central neural networks and are gaining increasing popularity for surrogate modeling. In accordance with the universal approximation theorem, a three-layer ANN with a non-linear activation function is able to approximate any complex non-linear functions with satisfactory accuracy [142]. A typical three-layer feedforward ANN, which is trained by error back propagation algorithm, consists of an input layer, a hidden layer and an output layer, as shown in Figure 7.
where $T_i b + \omega x_i$ is a linear version of SVR, $y_i$ denotes the predicted value, slack variables $\xi_i$ and $\xi^*_i$ allow the existence of outliers outside the $\varepsilon$-tube, and the regularization constant (also known as penalty coefficient) $C$ here achieves a trade-off between the model complexity and the empirical risk. For better understanding, a typical SVR model is shown in Figure 6.

![Figure 6. Schematic diagram of SVR.](image)

Artificial neural network (ANN) model. ANNs are computational systems inspired by biological central neural networks and are gaining increasing popularity for surrogate modeling. In accordance with the universal approximation theorem, a three-layer ANN with a non-linear activation function is able to approximate any complex non-linear functions with satisfactory accuracy [142]. A typical three-layer feedforward ANN, which is trained by error back propagation algorithm, consists of an input layer, a hidden layer and an output layer, as shown in Figure 7.

![Figure 7. Structure of a typical three-layer ANN [41].](image)

The training process of an ANN mainly includes two stages: (1) signal forward propagation, in which each neuron collects incoming signals and the weighted summation of inputs is processed and transferred to the next layer of neurons by means of the activation function, and (2) error backward propagation, in which the deviation between the actual outputs and the forecasting outputs is calculated and back-propagated and then connection weights are updated by the gradient-descent strategy. Since machine learning techniques are increasingly important in many engineering fields, more complex ANNs have also received increasing applications in surrogate modeling, such as long short-term memory (LSTM), convolutional neural networks (CNNs) and deep neural networks (DNNs) [143–145].

4.2. Hybrid Strategies of Surrogate Model

Although an individual surrogate model can achieve good performance for certain problems, it is well recognized that no single surrogate model always performs the best for all types of engineering applications [146]. This motivates the idea of using a hybrid surrogate model that takes full advantage of the individual surrogate model to guarantee the accuracy and robustness of the predictions for diverse low-/high-dimensional problems. The basic principle of the hybrid surrogate model is to utilize a linearly weighted summation of the individual surrogate model as follows:

$$\hat{g}_h(x) = \sum_{i=1}^{m} \omega_i \hat{g}_i(x), \quad \sum_{i=1}^{m} \omega_i = 1$$

(18)

where $\hat{g}_h(x)$ is the response predicted by the hybrid surrogate model at point $x$, $m$ is the number of surrogate models involved and $\omega_i$ is the weight associated with the $i$th surrogate model $\hat{g}_i(x)$. Theoretically, the adjustable weights provide a flexibility to place more emphasis on the good surrogate model and less emphasis on the bad surrogate model as per the need [13].

According to the schemes for determining the weights, existing hybrid strategies can generally be classified as average measures (or global ensemble) and pointwise ones (or local ensembles) [147]. The weights evaluated by average measures remain constant in the whole design space [148,149]. However, the precision of the individual surrogate model may change significantly in the design space; accordingly, the hybrid surrogate model with fixed weights inevitably encounters a precision fluctuation. In contrast, pointwise measure-based schemes have shown more satisfactory precision, and various approaches have flour-
ished to reasonably determine the weights, such as minimal prediction error-based, cross validation-based, optimization-based and trust region-based approaches [146,150–154]. In general, auxiliary optimization procedures that are used to search for the weights also inevitably increase the computing effort, especially for high-dimensional problems.

To mitigate the computational burden in complex engineering practices, building a multi-fidelity (also known as variable-fidelity) surrogate model that combines a cheaper, low-fidelity model and a more expensive, high-fidelity model has gradually become another attractive alternative [155,156]. In such cases, the global trend of the system function is captured by the low-fidelity model, and the local accuracy is guaranteed by the high-fidelity model, as shown in Figure 8. Two main concerns, including the sampling strategy and the precision combination, have always been the work emphasis under various multi-fidelity modeling frameworks.

![Figure 8. Schematic diagram of multi-fidelity surrogate model.](image)

To address the uncertainties in multi-level systems, a series of multi-level surrogate modeling strategies have also emerged to simultaneously tackle different models. When constructing a multi-level surrogate model, some scholars prefer to employ a local explanation to modify the global surrogate model [157,158], while others focus on tackling the challenge of co-existing uncertainties in surrogate modeling [159–161].

### 4.3. Accuracy Evaluation of a Surrogate Model

To assess the accuracy of a surrogate model, different metrics can be used to measure the deviation between the predicted value and the actual value from the following perspectives.

**Coefficient of determination $R^2$.** The $R^2$ is used to gauge the overall reliability of the surrogate model, which can be written as:

$$R^2 = 1 - \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$  \hspace{1cm} (19)

where $y$, $\hat{y}$, $\bar{y}$ and $n$ denote the actual value, the predicted value, the mean of the actual value and the number of test samples, respectively. In general, the larger the $R^2$ is, the higher the accuracy of surrogate model is.

**Mean square error (MSE).** The MSE utilizes the square of Euclidean distance to measure the prediction error, and it is defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$  \hspace{1cm} (20)
Since the MSE does not have the same unit of measurement as the actual value, its square root version (root mean square error, RMSE) provides a more intuitive measurement:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}
\]  

(21)

Theoretically, both the MSE and the RMSE stay at positive values and decrease as the prediction error approaches zero.

Mean absolute error (MAE). The MAE describes the average deviation between the predicted value and actual value and is defined as:

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|
\]  

(22)

Relative average absolute error (RAAE). The RAAE is utilized to measure the global relative error, and it can be expressed as:

\[
RAAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| / \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2}
\]  

(23)

5. Sampling Strategy of Surrogate Model

Sampling (also known as Design of Experiments, DoE), the process of generating a good set of data points in the design space, has become a pivotal issue in experiments or simulations, with the purpose of maximizing the information gained from a limited number of samples. To guarantee the quality of surrogate models without incurring excessive samples, studying sampling techniques is of immense importance [13,162]. In general, sampling techniques can be classified into two categories: one-shot sampling and sequential sampling. Design considerations and progresses of different sampling strategies will be discussed in what follows.

5.1. One-Shot Sampling

One-shot sampling (or static sampling) determines the sample size and points in a single stage. Widely used one-shot sampling approaches include Monte Carlo sampling (MCS), Full/Fractional factorial design (FFD), Central composite design (CCD), Orthogonal array sampling (OAS), Latin hypercube sampling (LHS), etc.

Monte Carlo sampling utilizes pseudo-random numbers to generate a large number of samples, hoping to achieve space-filling by its random actions. To reduce possible unrepresented regions caused by randomness, Stratified Monte Carlo sampling is proposed to achieve space-filling by dividing several non-random strata [163]. Quasi-Monte Carlo sampling employs a quasi-random low-discrepancy sequence to generate samples, where several popular low-discrepancy sequences (e.g., Halton, Hammersley and Sobol sequences) are attractive for sampling [164].

Full factorial design takes into account all possible combinations of design variable levels, filling the whole design space regularly with the same density of samples in each sub-domain [165]. A main drawback of this method is that the computational budget explodes exponentially as the number of design variables (dimensions) grows. To overcome this disadvantage, Fractional factorial design has been introduced to neglect certain high-order interaction effects to reduce the number of experiments [166].

Central composite design is regarded as a full/fractional factorial-embedded design, augmented with a group of center and axial points (two axial points for each axis) [162]. It is a popular second-order design due to its unique feature of adding center and axial points. Another similar approach named Box–Behnken design requires fewer runs than the CCD, despite its poor coverage at the corner of the cube enclosing the design space.
Orthogonal array sampling utilizes the orthogonal table to generate some representative samples, in which the interaction between factors is considered. This method can procure uniform-dispersed and representative samples without executing expensive simulations under multi-factors and levels.

Latin hypercube sampling, a special case of OAS, has gained much popularity in various fields. In an \( m \)-dimensional design space, each dimension is divided into \( n \) equal bins and thus results in \( n^m \) hypercubes. Subsequently, \( n \) samples are arranged randomly such that there exists only one sample in each dimension of the hypercube (also known as the non-collapsing property). The LHS configuration is typically used as an improvement over MCS but may not guarantee adequate space-filling [167]. To generate regularly dispersed samples, the so-called optimal LHS has emerged, using different optimal criteria, e.g., the centered \( L_2 \)-discrepancy criterion, the \( \phi_p \) criterion, the max–min distance criterion and the entropy criterion [14]. For more details about search algorithms in optimal LHS, one can consult [168].

5.2. Sequential Sampling

Although the sampling strategies discussed so far are popular, some can lead to over/under-sampling and thus poor system approximations [167]. In this context, the sequential sampling (also known as adaptive sampling), as depicted in Figure 9, has been developed to augment new informative points on the basis of initial samples. Two major benefits of sequential sampling methodologies over one-shot ones are their low computational budget and better approximations.

![General flowchart of sequential sampling.](image)

Overall, sequential sampling includes two basic concepts: exploration and exploitation [14,167]. For sample settlement, the exploration of the design space tends to cover the poorly represented/unexplored regions in a global sense. In contrast, exploitation focuses on placing samples in poor-precision/nonlinear regions under local consideration. Theoretically, most sequential sampling techniques rely on designating an appropriate criterion to strike a trade-off between global exploration and local exploitation [13].

To enhance the global approximation accuracy of the surrogate model, it is pivotal to develop an effective local exploitation criterion for error prediction over the domain wherein additional samples are designed in these poorly estimated regions. Variance-based approaches embrace the assumption that regions with large prediction variances contain more estimation errors in the whole design space [169]. Entropy-based methods search
new points by maximizing the determinant of prior covariance matrix under the Bayesian framework [170]. Cross-validation (CV)-based techniques utilize leave-one-out CV errors to estimate prediction errors, evaluating the credibility of the surrogate model to some extent [171]. Gradient-based methods select the point with the maximal gradient for updating, aiming to improve the accuracy in these high-gradient regions [172]. Query-by-committee-based strategies assert that the point at which all committee members (several competing surrogate models) have maximal divergence are selected for supplement points [173].

In practical terms, if a global optimization is pursued, one can employ the following criteria to search the global optimum. Expected improvement (EI)-based techniques aim to generate the sample point that can maximize the EI function as new candidate points. Due to its easy-to-complement and robustness characteristic, this popular criterion has received increasing applications in recent years and is elaborately explained in [174]. Statistical-lower-bound-based methods choose the point at which the surrogate model yields the statistical lower bound as the new point. Though this criterion is simple, it is not easy to strike a balance between exploration and exploitation [175]. Probability of improvement-based approaches hope to find the point with the maximal probability that the system response is smaller than a threshold [156,176]. This method is sensitive to the above user-defined value, and arbitrary assignment may influence the performance of the criterion.

6. Conclusions

Despite the fact that surrogate modeling techniques have received increasing attention and investigations in the last few decades, the existing studies lack a comprehensive overview of surrogate modeling tricks in uncertainty-treatment practices. Firstly, this paper provides a thorough overview of two pivotal arms in uncertainty management, i.e., uncertainty quantification and propagation, together with their theoretical bases and recent applications. Subsequently, a comprehensive discussion lies in two main concerns in surrogate modeling: the theoretical basis and sampling strategy. The following remarks emphasize existing challenges and provide suggestions for future studies.

1) The probability framework is well-established and is still the most useful tool for uncertainty handling, and its incorporation of emerging machine learning techniques has become a research hotspot. Complementing this, non-probabilistic techniques have initially shown superiority in tackling the epistemic uncertainty, but there are still many unsolved issues in theoretical construction. In contrast, hybrid strategies that tackle different uncertainties simultaneously are flexible in theory but remain in the exploratory stage.

2) Existing numerical methods for uncertainty propagation are mostly developed to tackle aleatory uncertainty, and their potential in non-probabilistic frameworks should be emphasized. An open research area is investigations developing new tricks of uncertainty propagation to strike a trade-off between efficiency and accuracy.

3) Though various surrogate models have received wide applications in different scenarios, using a weighted mixture of different surrogate models rather than focusing performance improvements on single ones has been suggested as an easier way to deliver better predictions. In addition, finding ways to extend the combination of certain types of surrogate models to two arbitrary ones remains a challenging issue. Considering the well-recognized ‘curse of dimensionality’, future developments should place more emphasis on dimension-reduction-assisted techniques for surrogate modeling.

4) Sequential sampling relies heavily on sampling criteria to update the prior information on the dataset, and thus designating appropriate algorithms remains an active area of research. In addition, as the complexity of simulation-based engineering practices increases, developing an effective sampling strategy that can deliver multiple-source information or conduct multi-point updates has gradually become an important research topic.

5) In addition to developing new algorithms for both uncertainty treatment and the surrogate model, building software platforms or packages should also be given
attention. Many useful platforms have been successfully developed (e.g., Isight for uncertainty optimization [177] and UQLab for uncertainty treatment [178]), but further development is still needed to improve their functionality.

(6) The investigation of surrogate modeling in uncertainty treatment practices should be extended to more complex scenarios. With necessary extensions, more investigations should emphasize multidisciplinary design, multi-scale and other engineering practices, especial in symmetrical systems.

Author Contributions: Conceptualization, C.W. and M.X.; methodology, C.W. and X.Q.; data curation, X.Q.; writing—original draft preparation, C.W. and X.Q.; writing—review and editing, T.W.; supervision, C.W.; project administration, C.W.; funding acquisition, C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of PR China (No.12002015, No. 12132001), the Overseas High-Level Talents Plan of PR China, and the Young Talent Support Plan of Beihang University. And the APC was funded by the National Natural Science Foundation of PR China (No.12002015).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author thank the reviewers for their valuable suggestions, which were very helpful to improve the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

\[ \Omega \] sample space

\[ (\Omega, F, P) \] probability triple

\[ m(\cdot) \] basic probability assignment

\[ R(\cdot) \] equivalence relation

\[ \mathbb{R}^n \] \( n \)-dimensional real number field

\[ \varepsilon \] perturbation parameter

\[ X \] input variable

\[ y(\cdot), f(\cdot), g(\cdot) \] system function

\[ \mu_A(\cdot) \] membership function

\[ a, b, \omega, \beta \] coefficients

\[ \Delta(\cdot) \] width of interval number

\[ \varepsilon \] deviation threshold

\[ G_e \] characteristic matrix of ellipsoid

\[ \psi(\cdot) \] basis function

\[ x^0 \] centroid of ellipsoid

\[ \xi, \xi^* \] slack variable

\[ \Theta \] frame of discernment

\[ \sigma^2 \] variance

\[ Bel(\cdot), Pl(\cdot) \] belief and plausibility measure

\[ Cov(\cdot) \] covariance function

\[ C \] penalty coefficient

\[ \Delta \] width of interval number

\[ \varepsilon \] deviation threshold

Subscripts

\[ I \] Interval number

\[ e \] ellipsoid

\[ c \] midpoint

\[ \gamma \] Lower and upper bound of interval

\[ f \] Fuzzy set

\[ T \] transposition

References


82. Lü, H.; Shangguan, W.; Yu, D. Uncertainty quantification of squeal instability under two fuzzy-interval cases. *Fuzzy Sets Syst.* 2017, 328, 70–82. [CrossRef]


142. Leshno, M.; Lin, V.Y.; Pinkus, A.; Schöcken, S. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural Netw. 1993, 6, 861–867. [CrossRef]

143. Deng, Y.; Buccchianico, A.D.; Pechenizkiy, M. Controlling the accuracy and uncertainty trade-off in RUL prediction with a surrogate Wiener propagation model. Reliab. Eng. Syst. Saf. 2020, 196, 106727. [CrossRef]


