Journal of Computational Physics 228 (2009) 8726–8751

Contents lists available at ScienceDirect

Journal of Computational Physics

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

Polynomial chaos representation of spatio-temporal random fields from experimental measurements

Sonjoy Das [a,*], Roger Ghanem [a], Steven Finette [b]

[a] University of Southern California, Kaprielian Hall 210, Los Angeles, CA 90089, USA
[b] Acoustics Division, Naval Research Laboratory, Washington, DC 20375, USA

ARTICLE INFO

Article history:
Received 4 September 2008
Received in revised form 5 June 2009
Accepted 25 August 2009
Available online 1 September 2009

Keywords:
Karhunen–Loève expansion
Non-Gaussian random process
Non-homogeneous/non-stationary random process
Polynomial chaos expansion
Probability density function estimation
Rosenblatt transformation
Spearman's rank correlation coefficient

ABSTRACT

Two numerical techniques are proposed to construct a polynomial chaos (PC) representation of an arbitrary second-order random vector. In the first approach, a PC representation is constructed by matching a target joint probability density function (pdf) based on sequential conditioning (a sequence of conditional probability relations) in conjunction with the Rosenblatt transformation. In the second approach, the PC representation is obtained by having recourse to the Rosenblatt transformation and simultaneously matching a set of target marginal pdfs and target Spearman's rank correlation coefficient (SRCC) matrix. Both techniques are applied to model an experimental spatio-temporal data set, exhibiting strong non-stationary and non-Gaussian features. The data consists of a set of oceanographic temperature records obtained from a shallow-water acoustics transmission experiment [1]. The measurement data, observed over a finite denumerable subset of the indexing set of the random process, is treated as a collection of observed samples of a second-order random vector that can be treated as a finite-dimensional approximation of the original random field. A set of properly ordered conditional pdfs, that uniquely characterizes the target joint pdf, in the first approach and a set of target marginal pdfs and a target SRCC matrix, in the second approach, are estimated from available experimental data. Digital realizations sampled from the constructed PC representations based on both schemes capture the observed statistical characteristics of the experimental data with sufficient accuracy. The relative advantages and disadvantages of the two proposed techniques are also highlighted.

1. Introduction

Many applications in science and engineering involve modeling spatio-temporal phenomena. Within the confines of the probabilistic framework, the Gaussian stochastic process has been the most common form for describing these phenomena. Non-Gaussian models, although clearly more realistic in most instances, have had to contend with the scarcity of consistent mathematical theories for describing general infinite-dimensional probability measures. Moreover, for most available mathematical models, algorithms may not be available for generating samples of the corresponding stochastic processes, that are consistent with target infinite-dimensional probability measures. In addition to these mathematical challenges, another crucial difficulty often stems from the paucity of data on which available models are to be based. This has limited the scope of non-Gaussian models to transformations of Gaussian vectors and processes, or to models that are completely characterized by their lower order statistics. Since Gaussian processes are characterized only by their mean and covariance functions, they

* Corresponding author.
E-mail addresses: sdas@usc.edu (S. Das), ghanem@usc.edu (R. Ghanem), steve.finette@nrl.navy.mil (S. Finette).

0021-9991/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2009.08.025

require a manageable amount of information and thus often provide a rational modeling alternative. These challenges notwithstanding, it remains a recognized fact that many processes representing physical phenomena rarely satisfy the assumptions and constraints associated with a Gaussian process. The current work focuses on the construction of a probability model of a non-stationary and non-Gaussian random process by using a set of measurement data and the associated simulation technique based on the constructed model. The resulting mathematical model readily lends itself to the generation of consistent samples of the process.

The most common approach in digitally generating realizations of a non-Gaussian process consists of specifying a set of target non-Gaussian marginal probability density functions (marpdfs) and a target correlation coefficient (corrcoef) function or spectral density function (sdf) [2,3]. The set of multidimensional statistics required to characterize these probabilistic models can be estimated from available data. In synthesizing realizations in this manner, the target stochastic process is assumed to be a nonlinear mapping of some underlying Gaussian process, where the corrcoef function of the Gaussian process is usually determined through a suitable (analytical or optimization based) root-finding technique. Realizations of this Gaussian process are then transformed, pointwise, into realizations of the target non-Gaussian random process using the Nataf transformation yielding a match of the target non-Gaussian marpdfs and the target correlation structure (the Nataf transformation is essentially a scalar-variate version of the Rosenblatt transformation [4]). This step involves an additional computational burden of constructing the inverse mappings of the target marginal probability distribution functions (marPDFs) if they are not readily available, which is often the case if these latter have been constructed using some nonparametric density estimation techniques [5,6]. In addition to computational challenges, it is well known that certain stochastic processes, specified through their marginal pdfs and their second-order correlation, cannot be characterized as Nataf transformation of a Gaussian process [7], thus limiting the applicability of these approaches.

It is clear that the construction of a probability model to describe observational data requires some definition of proximity between random variables of interest. The most traditional way to define the proximity, as alluded to above, is via the correlation coefficient, also known as Pearson's correlation coefficient (PCC). In many classical statistical tests, for instance, hypothesis test (such as to test if the correlation, i.e., PCC, between two random variables is zero), it is required that the random variables be normally distributed [8,9]. It can be inferred from the statistical literature that nonparametric correlation coefficients based on distribution-free approach are likely to be more useful when the assumption of Gaussianity violates. A number of other notions of distribution-free correlation have been proposed that might be more useful in the present context [8] as they provide alternative measures of statistical dependency, useful in characterizing non-Gaussian random processes. These include Spearman's rank correlation coefficient (SRCC) or Spearman's ρ and Kendall's τ. A recent simulation-based study [10, Section 12.5.2] investigates the admissibility (i.e., positive-definiteness) of the PCC matrix of an underlying Gaussian random vector when the statistical dependency of the non-Gaussian random vector is characterized by the correlation coefficient matrices based on SRCC and Kendall's τ. It was found that constraining the target non-Gaussian process through its SRCC, instead of the Kendall's τ, provides a characterization that is more amenable to be described as a nonlinear (e.g., Nataf) transformation of some underlying Gaussian process. Thus, in addition to describing a more meaningful statistical dependency structure, the SRCC also provides a characterization that is more amenable to digital synthesis and simulation when compared to Kendall's τ. In light of these observations, the SRCC will be considered in the ensuing discussion as the main correlation structure describing the stochastic processes at hand.

The simulation of non-Gaussian random processes by specifying a set of target non-Gaussian marpdfs and a target SRCC function has already been considered in the literature. It was found that whenever the underlying Gaussian process *exists*, *no* special search technique is required to determine its feasible PCC function [11] resulting in significant computational savings. However, efficient algorithms are still required for the computation of the inverse functions of the target marPDFs. Clearly, simulation techniques, based on a target PCC/SRCC function or sdf and an underlying Gaussian process, are not computationally efficient, particularly, in the case when the target set of marpdfs are estimated by employing nonparametric techniques. The characterization of non-Gaussian random processes continues to be an active field of research. Many practically appealing issues concerning the estimation of the underlying family of multivariate joint probability density functions (mjpdfs) from finite data sets have provided the impetus for much of the innovation in the recent past [5–7,12]. By making use of such techniques, the problem of non-existence of an underlying Gaussian random process can be mitigated at the expense of computational efficiency. Advanced simulation algorithms such as Markov chain Monte Carlo (MCMC) sampling are required to sample from the resulting family of mjpdfs thus placing a further demand on computational resources. This difficulty could be a major bottleneck particularly in the context of propagating the statistical characteristics of stochastic system parameters to model-based predictions if the stochastic system parameters are modeled as non-Gaussian random processes. A number of studies [13,14] have been carried out to circumvent this particular difficulty by representing the non-stationary and non-Gaussian random processes through polynomial chaos (PC) expansions [15]. The underlying concept of these studies is similar to the one introduced earlier in the literature [16], which again assumes the existence of an underlying Gaussian process.

The PC formalism, nevertheless, provides a theoretically sound backbone facilitating efficient construction of the probability model of the non-stationary and non-Gaussian second-order random process possibly representing some model parameters of a stochastic system [17–20] and systematic propagation of the statistical properties of these stochastic system parameters to the response of the model in diverse fields of application [15,21–24].

Within the purview of the PC framework, the probability model of the random process refers to a decomposition of the process itself constructed with respect to (w.r.t.) a set of random variable basis functions. The basis functions constitute a set

of orthogonal functions w.r.t. a suitably chosen probability measure [15,18,19]. The coordinates (often referred as PC coefficients in the literature) w.r.t. the basis functions are the representative statistics and must be estimated from data. The set of PC coefficients thus provides a parametrization of the probabilistic description of the stochastic process. Moreover, given the fact that this is a linear parametrization of the process itself, and not its multidimensional pdf, synthesizing digital realizations of the process is an immediate task involves virtually no computational overhead. We also note that the PC based procedures do not require an assumption of ergodicity, and the resulting PC representations of the stochastic process do no exhibit this characteristic either.

Recent attempts [25] at constructing PC representations from a given set of data have assumed the dominant Karhunen–Loève (KL) random variable components of the stochastic process to be statistically independent, and have relied on Bayesian inference to construct posterior distributions for the PC coefficients. In fact, it is common to assume the uncorrelated (statistically dependent) KL random variables associated with a non-Gaussian process to be mutually statistically independent. Due to the central limit theorem, the assumption of statistically independent KL random variables causes a fundamental difficulty in representing a non-Gaussian process. This is particularly true for a process with wide-banded evolutionary sdf [26] (for stationary process, it reduces to the usual sdf). Under certain mild condition, the central limit theorem guarantees then that the KL-based representation of the process tends to produce in distribution a process with Gaussian marpdfs rather than the original non-Gaussian process.[1]

Additional efforts [28], that do not assume the existence of an underlying Gaussian process and relax the condition of statistical independence among the KL components, propose to employ the principle of maximum likelihood. The likelihood function is, however, approximated by the product of one-dimensional marginal likelihood functions for computational expediency at the cost of accuracy loss. This approximation is essentially similar to the independence composite likelihood representation [29]. Further work based on the principle of maximum entropy and the Rosenblatt transformation has recently been proposed to identify the asymptotic mjpdf of the PC coefficients [30]. In that work, no assumption about the underlying Gaussian process is made, and the statistical dependencies among the dominant KL components are characterized by accurately capturing their higher order *joint* statistical features.

The current work considers the problem of constructing the probability model of a non-stationary and non-Gaussian random process *directly from* the experimental measurements. At the outset, let us set forth a clarification of terminology for the present work. When the indexing set of the stochastic process is multidimensional, reference is often made to a random field, and a stationary random process is then referred to as a statistically homogeneous random field. In the present work, and to emphasize the identical underlying mathematical structure, the term 'stochastic process' or 'random process' will be ubiquitously used and the equivalence of the concepts of statistical homogeneity and stationarity will be implied. Throughout this work, the bold face character will be used to indicate that the quantity under consideration is either random or multidimensional. The realizations of a multidimensional random quantity are, however, denoted by the respective normal characters for distinction purpose.

We will assume that a collection of measurement data is observed over a finite subset of the indexing set of the random process, and consequently, treated as a collection of observed samples of $\mathbb{R}^N$-valued random vector $\boldsymbol{\mathcal{Y}}$ with $\mathbb{R}^N$ denoting the N-dimensional Euclidean space. It should be clear that $\boldsymbol{\mathcal{Y}}$ represents a finite-dimensional representation of the original (infinite-dimensional) stochastic process under investigation. Let the set of random variable components of $\boldsymbol{\mathcal{Y}}$ be denoted by $\{\mathbf{y}_i\}_{i=1}^N$ and the multivariate joint probability distribution function (mjPDF) of $\boldsymbol{\mathcal{Y}} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]^T$ by $P_{\mathbf{y}_1,\ldots,\mathbf{y}_N}$. Here, T is the transpose operator. The probability measure of the underlying random process is then completely characterized by the family of mjPDF: $\{P_{\mathbf{y}_1,\ldots,\mathbf{y}_N}\},\ \forall N \in \mathbb{N} = \{0, 1, 2, \ldots\}$. Since N is always finite in an experimental or numerical context, characterizing the underlying stochastic process has to be performed, in some approximate sense, through a characterization of $\boldsymbol{\mathcal{Y}}$. The probability measure $P_{\boldsymbol{\mathcal{Y}}} \equiv P_{\mathbf{y}_1,\ldots,\mathbf{y}_N}$ estimated by using the measurement data set satisfies Kolmogorov's existence theorem [27, Section 36] implying the existence of a stochastic process compatible with the set $\{P_{\mathbf{y}_1,\ldots,\mathbf{y}_n}\},\ \forall n \in \{1, \ldots, N\}$. The value of N, required to achieve a certain fidelity in the finite-dimensional representation, depends on the characteristics of the stochastic fluctuations of the original stochastic process over its spatio-temporal domain. Then, the constructed stochastic process characterized by $P_{\boldsymbol{\mathcal{Y}}}$ and the underlying original stochastic process are equivalent in the sense that they have N-dimensional identical mjPDFs.

The current work presents two different computational techniques to estimate the probability model of a finite-dimensional approximation of the underlying non-stationary and non-Gaussian stochastic process that is assumed to be completely characterized by experimental measurements taken simultaneously over space and time. In the context of the present work, the term 'probability model' refers to 'PC representation'. The first approach constructs the PC representation based on a target mjpdf, and the other approach is based on a set of all the target marpdfs and a target SRCC matrix. The target mjpdf, marpdfs and SRCC matrix correspond, respectively, to the observed joint histogram density, observed marginal histogram densities and sample SRCC matrix estimated by using the available measurements. No assumption about the exis-

[1] It follows from Lindberg–Feller theorem and requires, for every $\varepsilon > 0$, that [27, Section 27]

$$\lim_{M\to\infty} \frac{1}{C_M^2} \sum_{i=1}^{M} \int_{|z-\mu_{\mathbf{z}_i}|>\varepsilon C_M} (z-\mu_i)^2 \, dP_{\mathbf{z}_i}(z_i) = 0,$$

in which $\{\mathbf{z}_i\}_{i=1}^M$ is the set of M dominant KL random variable components, $P_{\mathbf{z}_i}$ is the marginal probability distribution function of $\mathbf{z}_i$, $\mu_{\mathbf{z}_i}$ is the mean of $\mathbf{z}_i$, $\sigma^2_{\mathbf{z}_i}$ is the variance of $\mathbf{z}_i$, and $C_M = \left(\sum_{i=1}^M \sigma^2_{\mathbf{z}_i}\right)^{1/2}$.

tence of an underlying Gaussian vector is made for any of the approaches presented here, nonetheless the second approach can exploit the advantage of existence of such a vector whenever possible.

The paper starts with a discussion of the PC formalism in Section 2, following which the two approaches are presented in Sections 3.1 and 3.2. Since considerable use of the properties of SRCC is made in the second approach, the definition and relevant features of SRCC are highlighted before presenting the second approach. As an illustration of the two proposed techniques, a set of oceanographic data obtained from a shallow-water acoustics transmission experiment [1] is used to model the spatio-temporal random temperature field and the results are discussed in Section 4. Finally, conclusions inferred from the present contribution are presented in Section 5.

The present work also contains two appendices. An interpolation-based scheme to efficiently compute the required PC coefficients is described in Appendix A. In Appendix B, overall convergence, verification and validation issues are briefly discussed.

2. Polynomial chaos formalism

It is assumed in the present work that the stochastic processes of interest are second-order, guaranteeing the existence of mean-squared convergent series representations, including the PC representation, of $\mathcal{Y}$. Current developments in polynomial chaos methodologies consist of various adaptations of the Cameron–Martin theorem [31] according to which second-order nonlinear functionals of the Brownian motion, defined on the space, $\mathcal{C}$, of all real-valued continuous functions on a compact support, are approximated by a Hilbertian decomposition with respect to a set of multidimensional orthogonal Hermite polynomials. The set of Hermite polynomials is constructed with respect to a finite-dimensional set of statistically independent Gaussian random variables, obtained through suitable Hilbertian projections of the Brownian motion. The resulting representation is shown [31] to converge in mean-square to the nonlinear functional being approximated as the dimension and order of the Hermite polynomial tend to infinity. This mean-square error (MSE) is measured with respect to the Wiener measure [32] on $\mathcal{C}$.

While the Cameron–Martin theorem referred explicitly to the representation of functionals of infinite-dimensional Brownian motion, some of their recent application, motivated by practical considerations have adapted this result to functionals of Gaussian and non-Gaussian finite-dimensional vectors, using orthogonal projections in suitable measure spaces [15,17,18,33,19].

Let $\underline{\xi} \equiv (\xi_1, \ldots, \xi_d)$ be a $\mathbb{R}^d$-valued random vector with probability measure, $P_{\underline{\xi}}$, that is absolutely continuous w.r.t. the Lebesgue measure, $d\underline{\xi}$, on $\mathbb{R}^d$ implying that $P_{\underline{\xi}}(d\underline{\xi}) = p_{\underline{\xi}}(\underline{\xi})d\underline{\xi}$, in which $d\underline{\xi} = \prod_{i=1}^{d} d\xi_i$ with $d\xi_i$ being the Lebesgue measure on $\mathbb{R}$ and $p_{\underline{\xi}}$ is the mjpdf of $\underline{\xi}$. The set $\underline{\xi}$ plays the role of an underlying source of randomness inducing, through an arbitrary nonlinear transformation, the randomness in the observed phenomenon. The choice of the probability distribution, $P_{\underline{\xi}}$, is thus a modeling decision that should reflect some judgment regarding the essential sources of uncertainty. Based on the chosen $P_{\underline{\xi}}$, the PC representation of each component of $\mathcal{Y}$ can be expressed as,

$$\mathbf{y}_k \equiv \mathbf{y}_k\left(\underline{\xi}\right) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^d} y_{\boldsymbol{\alpha},k}\,\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi}), \quad k = 1, \ldots, N, \tag{1}$$

if $\mathbf{y}_k(\underline{\xi})$ is a second-order random variable, i.e., $E[|\mathbf{y}_k(\underline{\xi})|^2] < \infty$ with $|x|$ representing the absolute value of x and $E[\cdot]$ representing the expectation operator with respect to the chosen probability measure, $P_{\underline{\xi}}$ (this second-order condition is satisfied here since the underlying stochastic process is assumed to be second-order). Here, $y_{\boldsymbol{\alpha},k},\ \boldsymbol{\alpha} \equiv (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, represent the PC coefficients (to be determined) with respect to the basis functions, $\Upsilon_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \mathbb{N}^d$. If p_{ξ_i} denotes the marpdf of ξ_i induced by $p_{\underline{\xi}}$, then the basis functions in (1) are given by [19],

$$\begin{aligned} \Upsilon_{\mathbf{0}}(\underline{\xi}) &= 1, && \text{if } \boldsymbol{\alpha} = \mathbf{0} \in \mathbb{N}^d, \\ \Upsilon_{\boldsymbol{\alpha}}(\underline{\xi}) &= \left(\frac{\prod_{i=1}^{d} p_{\xi_i}(\xi_i)}{p_{\underline{\xi}}(\underline{\xi})}\right)^{1/2} \prod_{i=1}^{d} \Psi_{\alpha_i}(\xi_i), && \text{if } \boldsymbol{\alpha} \neq \mathbf{0}, \end{aligned} \tag{2}$$

in which Ψ_{α_i} are orthogonal polynomials of order α_i. The orthogonality of the basis functions $\Psi_{\alpha_i},\ \alpha_i \in \mathbb{N}$, also implies [19] the orthogonality of the set, $\{\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi}),\ \boldsymbol{\alpha} \in \mathbb{N}^{n_d}\}$, with respect to $P_{\underline{\xi}}$. In the case of statistically independent $\xi_1, \ldots, \xi_d$ random variables, (2) simplifies to,

$$\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi}) = \prod_{i=1}^{d} \Psi_{\alpha_i}(\xi_i). \tag{3}$$

The equality, '=', in (1) should be interpreted in the mean-square sense such that $E\left[\left\{\mathbf{y}_k(\underline{\xi}) - \sum_{\boldsymbol{\alpha}:|\boldsymbol{\alpha}|\leq n_o} y_{\boldsymbol{\alpha},k}\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi})\right\}^2\right] \to 0$ as $n_o \to \infty$, where the expectation operator is w.r.t. $P_{\underline{\xi}}$ [19], $|\boldsymbol{\alpha}| = \sum_{i=1}^{d} \alpha_i$, and n_o is the maximum order (i.e., order of the PC representation) of all the basic orthogonal polynomials, $\{\Psi_{\alpha_i},\ \alpha_i \in \mathbb{N},\ i \in (1, \ldots, d)\}$, included in (1). However, for computational purpose, this infinite series is truncated after a finite number of terms that is typically determined by the available computational budget and target accuracy (usually in the sense of MSE).

The accuracy of a truncated approximation and the rate of convergence of the PC representation also depend on the choice of $P_{\underline{\xi}}$ and consequently, on the resulting set of orthogonal basis functions used in (1). The proper selection of the probability measure $P_{\underline{\xi}}$ should reflect a physical modeling process through which the significant and sufficient sources of uncertainty are identified. Once this choice for $P_{\underline{\xi}}$ has been made and the mapping, $\underline{\xi} \mapsto \mathbf{y}_k(\underline{\xi})$, is constructed, either explicitly or implicitly, the PC coefficients can be computed using the orthogonality property of $\Upsilon_{\boldsymbol{\alpha}}$'s,

$$y_{\boldsymbol{\alpha},k} = \frac{E\left[\mathbf{y}_k(\underline{\xi})\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi})\right]}{E\left[\Upsilon^2_{\boldsymbol{\alpha}}(\underline{\xi})\right]}, \quad \boldsymbol{\alpha} \in \mathbb{N}^d, \;\; k = 1, \ldots, N. \tag{4}$$

The denominator in (4) can be evaluated by using (2) or (3) as appropriate. When $\xi_1, \ldots, \xi_d$ are statistically independent, then $E[\Upsilon^2_{\boldsymbol{\alpha}}(\underline{\xi})]$ reduces to,

$$E\left[\Upsilon^2_{\boldsymbol{\alpha}}(\underline{\xi})\right] = \prod_{i=1}^{d} E\left[\Psi^2_{\alpha_i}(\xi_i)\right], \tag{5}$$

in which $E[\Psi^2_{\alpha_i}(\xi_i)]$ are readily available [19] for many commonly employed measure, P_{ξ_i}. The numerator in (4), on the other hand, is given by,

$$E\left[\mathbf{y}_k(\underline{\xi})\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi})\right] = \int_{S_{\underline{\xi}}} \mathbf{y}_k(\underline{\xi})\Upsilon_{\boldsymbol{\alpha}}(\underline{\xi}) p_{\underline{\xi}}(\underline{\xi}) d\underline{\xi}, \tag{6}$$

in which $S_{\underline{\xi}} \subseteq \mathbb{R}^d$ is the support of $\underline{\xi}$. Clearly, computation of this integral requires knowledge of the mapping, $\underline{\xi} \mapsto \mathbf{y}_k(\underline{\xi})$, which is not available in the present work because the information, that is assumed to be available here, is *only* the measurement data on $\mathcal{Y}$. Two schemes are presented next that define this mapping, enabling the computation of the integral in (6) to determine the PC coefficients in (4), and subsequently, facilitating construction of the required PC representation in (1).

3. Construction of PC representation from data

In this section, the two approaches for constructing PC representations are presented in detail. Each approach consists of two major steps. In the first step, a mjPDF is estimated based on available measurements. The estimated mjPDF is referred to as the target mjPDF. The second step involves developing a PC representation such that the associated mjPDF is within a desired tolerance to the target mjPDF.

The first approach borrows ideas from the literature of pdf estimation [12] and PC coefficient estimation [30]. This approach, which can be viewed as a supplement to our previous work [30], is based on the Rosenblatt transformation that makes use of a complete set of properly ordered conditional probability distribution functions (PDFs). The set of conditional PDFs uniquely defines the target mjPDF. The associated target mjpdf is simply taken as the observed joint normalized histogram estimated by using the available measurements.

The second approach, on the other hand, is founded on the properties of SRCC and the Rosenblatt transformation, applied individually on each marPDF of the involved scalar-variate random variable components. It borrows ideas from the literature of computer simulation of non-Gaussian random vectors when these are characterized by their marpdfs and SRCC. In this approach, the target mjPDF is assumed to be completely characterized by a set of marpdfs and a SRCC matrix. The set of target marpdfs and the target SRCC matrix are, respectively, taken as the observed marginal normalized histogram and the sample SRCC matrix that are estimated from the available measurements.

Both approaches rely on the Rosenblatt transformation to define a nonlinear mapping $\mathbf{f} : \underline{\xi} \to \mathcal{Y}$ such that $\mathbf{f}(\underline{\xi})$ is equal in distribution to $\mathcal{Y}$. The function $\mathbf{f}(\cdot)$ is then approximated by using PC representation constructed w.r.t. the chosen measure of $\underline{\xi}$. By construction (see details in Sections 3.1 and 3.2) and choice of $P_{\underline{\xi}}$, the function $\mathbf{f}(\cdot)$ is absolutely continuous on $S_{\underline{\xi}}$ in addition to being second-order, $E(\|\mathbf{f}(\underline{\xi})\|^2) < \infty$, where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^N$. This last condition guarantees the existence of the sought-after PC representation for $\mathbf{f}(\underline{\xi})$ [19].

3.1. Approach 1: based on conditional PDFs

The unknown mapping, $\underline{\xi} \mapsto \mathcal{Y}$, in this case is defined by using the Rosenblatt transformation. While any suitable density estimation technique could be applied to compute the target mjPDF, $P_{\mathcal{Y}}$, of $\mathcal{Y}$ by using the available measurement data, it is simply obtained, in the present work, from the normalized $(N+1)$-dimensional histogram of the available N-variate data of $\mathcal{Y}$. The normalized histogram can be used to determine the corresponding target mjPDF, $p_{\mathcal{Y}}$. The histogram is first estimated over a discrete array of a finite number of grid points spread over the support, $S_{\mathcal{Y}} \subset \mathbb{R}^N$, of $\mathcal{Y}$. This discrete array of grid points typically represents the center points of the histogram bins. An N-dimensional linear interpolation scheme is subsequently employed to determine the value of the histogram of $\mathcal{Y}$ at any other arbitrary point, $\mathcal{Y} \in S_{\mathcal{Y}}$, thus resulting in the target mjpdf, $p_{\mathcal{Y}}$, and therefore, the target mjPDF, $P_{\mathcal{Y}}$, over the entire $S_{\mathcal{Y}}$. Use of the normalized histogram to approximate $p_{\mathcal{Y}}$ is acceptable in the view of the fact that the density estimation techniques currently existing in the literature are founded on this primitive notion of normalized histogram. It should also be noted that the final objective of the present work is

not estimation of the mjpdf of $\mathcal{Y}$ but construction of the PC representation of $\mathcal{Y}$. The resulting $P_{\mathcal{Y}}$ is an absolutely continuous function on $S_{\mathcal{Y}}$ because of the use of a linear interpolation scheme. A requirement for using the Rosenblatt transformation is that $P_{\mathcal{Y}}$ be an absolutely continuous function on $S_{\mathcal{Y}}$.

The simplest formulation of Approach 1 can be explained by considering a random variable, say, $\mathbf{Z}$, which is a function of two *dependent* random variables, say, $\mathbf{X}$ and $\mathbf{Y}$. The random variable $\mathbf{Z}$ is first approximated in terms of a PC representation constructed w.r.t. a weight defined on the support of $\mathbf{X}$. The resulting PC representation captures the effects of $\mathbf{Y}$ into the PC coefficients which are random since they depend on $\mathbf{Y}$. Subsequently, these random PC coefficients are further approximated with another set of orthogonal expansions constructed w.r.t. another weight defined on the support of $\mathbf{Y}$ [34].

Let us illustrate the approach now by using a 2-dimensional random vector, say, $\mathcal{Y} = [\mathbf{y}_1, \mathbf{y}_2]^T$. The formulation can be readily extended to the random vector with more than two random variable components. Consider the 2-dimensional data set as shown in Fig. 1. The corresponding histogram is shown in Fig. 2. The target mjpdf, $p_{\mathcal{Y}}$, based on 2-dimensional linear interpolation of the histogram is shown in Fig. 3. The motive here is to pictorially describe the formulation; therefore, the specific values of the associated data or values of the resulting function and variables are not relevant.

The conditional pdf of $\mathbf{y}_1$, given $\mathbf{y}_2 = y_2$, induced by $p_{\mathbf{y}_1\mathbf{y}_2}$ is denoted by $p_{1|2}$ and shown in Fig. 4 for different values of $y_2 \in s_{\mathbf{y}_2}$, in which $s_{\mathbf{y}_2} = [l_2, m_2] \subset \mathbb{R}$ is the support of $\mathbf{y}_2$. The slices representing $p_{1|2}$ as shown in this figure are obtained from the corresponding slices of Fig. 3 by simply making the area under each slice unity because area under a pdf is always unity,

$$p_{1|2}(y_1|y_2) = \frac{p_{\mathbf{y}_1\mathbf{y}_2}(y_1, y_2)}{\int_{s_{\mathbf{y}_1}} p_{\mathbf{y}_1\mathbf{y}_2}(y_1, y_2)dy_1} = \frac{p_{\mathbf{y}_1\mathbf{y}_2}(y_1, y_2)}{p_{\mathbf{y}_2}(y_2)},$$

in which $s_{\mathbf{y}_1} = [l_1, m_1] \subset \mathbb{R}$ is the support of $\mathbf{y}_1$ and $p_{\mathbf{y}_2}$ is the marpdf of $\mathbf{y}_2$. Let the associated conditional PDF be denoted by $P_{1|2}$ given by,

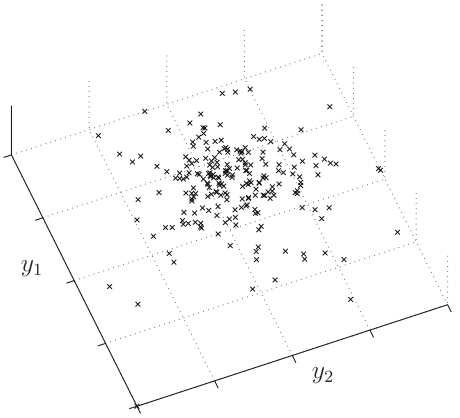


Fig. 1. 2-Dimensional illustration: data points.

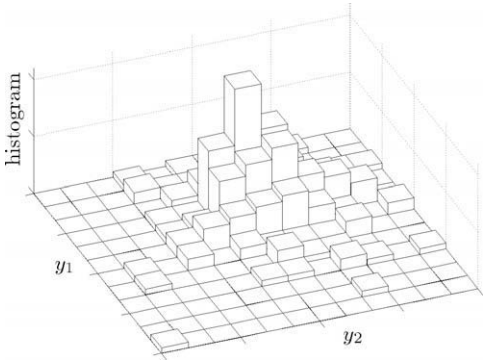


Fig. 2. 2-Dimensional illustration: histogram.

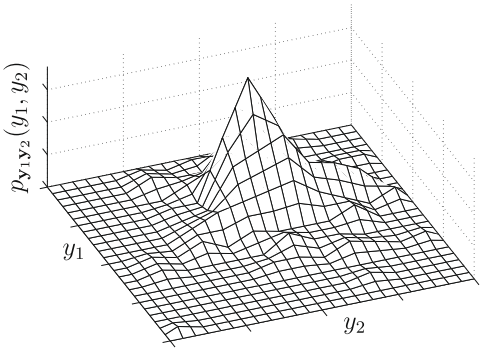


Fig. 3. 2-Dimensional illustration: the target mjpdf, $p_{\mathcal{Y}} \equiv p_{\mathbf{y}_1\mathbf{y}_2}$, of $\mathcal{Y} = [\mathbf{y}_1, \mathbf{y}_2]^T$.

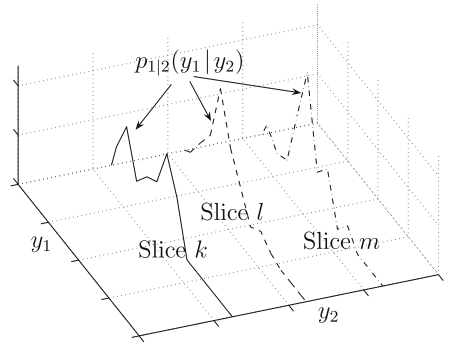


Fig. 4. 2-Dimensional illustration: three slices representing the conditional pdf of $\mathbf{y}_1$, given $\mathbf{y}_2 = y_2$, for three different y_2's.

$$P_{1|2}(y_1|y_2) = \frac{\int_{l_1}^{y_1} p_{\mathbf{y}_1\mathbf{y}_2}(y, y_2)dy}{p_{\mathbf{y}_2}(y_2)},$$

as depicted in Fig. 5. Consider $P_{1|2}(\mathbf{y}_1|y_2)$ and $P_{\xi_1}(\xi_1)$ as two random variables (functions of $\mathbf{y}_1$ and ξ_1, respectively). The PDF of both the random variables are uniform distribution supported over $[0,1]$ [10, Theorem 2.1]. Then, the mapping, $\mathbf{f} : \underline{\xi} \to \mathcal{Y}$, can be defined by employing the Rosenblatt transformation [4] as shown below,

$$P_{1|2}(\mathbf{y}_1|y_2) \stackrel{d}{=} P_{\xi_1}(\xi_1) \tag{7}$$

$$\Rightarrow \mathbf{y}_1 \stackrel{d}{=} (P_{1|2}^{-1} P_{\xi_1})(\xi_1|y_2) \tag{8}$$

$$= \lim_{K\to\infty} \sum_{j=0}^{K} a_j(y_2)\Psi_j(\xi_1). \tag{9}$$

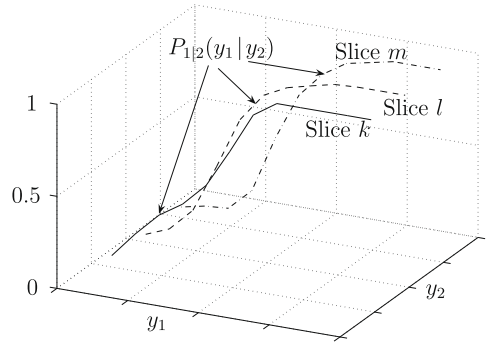


Fig. 5. 2-Dimensional illustration: three slices representing the conditional PDFs of $\mathbf{y}_1$, given $\mathbf{y}_2 = y_2$, for three different y_2's.

The equalities, "$\stackrel{d}{=}$", above should be interpreted in the sense of distribution implying that the PDFs of the random variables in the left-hand-side (lhs) and the right-hand-side (rhs) of each equality are identical. Eq. (7) ensures that the conditional PDF of $\mathbf{y}_1$, given $\mathbf{y}_2 = y_2$, is precisely given by $P_{1|2}$ [10, Theorem 2.1] as required.

It should be noted here that by construction (because of use of a linear interpolation scheme) and choice of PDF of ξ_1, $f_{1|2} \equiv P_{1|2}^{-1} P_{\xi_1}$ is piecewise smooth [35, p. 18] on the support, $s_{\xi_1} \subseteq \mathbb{R}$, of ξ_1, and also second-order, $\int_{s_{\xi_1}} f_{1|2}^2(\xi_1|y_2) p_{\xi_1}(\xi_1) d\xi_1 < \infty$. This results in the PC representation of $f_{1|2}$ as shown by the rhs of (8). It should be noted that while $\mathbf{y}_1$ and the rhs of (8) is equal in distribution,

$$\mathbf{y}_1 \stackrel{d}{=} f_{1|2}(\xi_1|y_2) = \lim_{K\to\infty} \sum_{j=0}^{K} a_j(y_2) \mathbf{\Psi}_j(\xi_1), \tag{10}$$

the equality, "=", above or in (8) follows from $f_{1|2}(\xi_1|y_2)$ (not from $\mathbf{y}_1$) and is valid in the mean-square sense. By the piecewise smoothness and second-order conditions, the equality is also valid at every continuity point of $f_{1|2}$ [36, Chapter 4] implying that the equality can as well be interpreted in almost sure (a.s.) sense w.r.t. P_{ξ_1}.

The deterministic (since, given y_2) PC coefficient, $\{a_j(y_2),\ j \in \mathbb{N}\}$, in (10) is given by,

$$a_j(y_2) = \frac{E\left[f_{1|2}(\xi_1|y_2)\mathbf{\Psi}_j(\xi_1)\right]}{E\left[\mathbf{\Psi}_j^2(\xi_1)\right]}, \quad j \in \mathbb{N}. \tag{11}$$

The determination of $a_j(y_2)$ requires computation of the following integral,

$$E\left[f_{1|2}(\xi_1|y_2)\mathbf{\Psi}_j(\xi_1)\right] = \int_{s_{\xi_1}} (P_{1|2}^{-1} P_{\xi_1})(\xi_1|y_2)\mathbf{\Psi}_j(\xi_1) p_{\xi_1}(\xi_1) d\xi_1.$$

The evaluation of this integral involves computation of the inverse of $P_{1|2}$. Since, in the current context, $P_{1|2}$ is based on observed histogram-based conditional PDF, no suitable analytical inverse function exists for such nonparametric PDF. Therefore, inverse of this function needs to evaluated numerically while evaluating the above integral. This might be computationally expensive or/and numerically instable. A computationally efficient scheme based on a surrogate function (instead of using $P_{1|2}^{-1} P_{\xi_1}$) is described in Appendix A.

For several different values of $y_2 \in s_{\mathbf{y}_2}$, the PC coefficients, $\{a_j(y_2)\}_{j\in\mathbb{N}}$, need to be computed. Let the support, $s_{\mathbf{y}_2} = [l_2, m_2] \subset \mathbb{R}$, be divided equally into $n_2 \in \mathbb{N}$ intervals. Then, coordinates of the points defining these intervals are given by $y_2^{(k)} = l_2 + k[(m_2 - l_2)/n_2], k = 0, \ldots, n_2$. For each slice defined by $P_{1|2}(\mathbf{y}_1|y_2^{(k)})$, compute the PC coefficients, $\{a_j(y_2^{(k)})\}_{j\in\mathbb{N}}$, by using (11). A few typical profiles of the mapping, $\mathbb{N} \ni j \mapsto a_j(y_2) \in \mathbb{R}$, for given y_2 are depicted in Fig. 6.

For any given $j \in \mathbb{N}$, the set of pairs, $\{y_2^{(k)}, a_j(y_2^k)\}_{k=0}^{n_2}$, as just determined is next used to construct the mapping, $s_{\mathbf{y}_2} \ni y_2 \mapsto a_j(y_2) \in \mathbb{R}$, by simply employing a linear interpolation scheme (note that this is an 1-dimensional version of a similar problem for estimating pdf from the histogram defined over a discrete array of points). A few profiles of this mapping are sketched in Fig. 7.

Since $n_2 \in \mathbb{N}$ is a *finite* (but large) number, the mapping, $y_2 \mapsto a_j(y_2)$, for any given $j \in \mathbb{N}$, defined via linear interpolation with $\{y_2^{(k)}, a_j(y_2^k)\}_{k=0}^{n_2}$, is piecewise smooth. By the second-order condition on $f_{1|2}$, it also implies that $|a_j(y_2)| < \infty$ for any given $j \in \mathbb{N}$. It is, therefore, straightforward to select a suitable weight, say, defined by $s_{\mathbf{y}_2} \ni y_2 \mapsto w_2(y_2) \in (0, \infty)$, such that $\int_{s_{\mathbf{y}_2}} a_j^2(y_2) w_2(y_2) dy_2 < \infty$. Then, a set of basis functions, $\{\psi_k\}_{k\in\mathbb{N}}$, orthogonal w.r.t. the weight $w_2(\cdot)$, $\int_{s_{\mathbf{y}_2}} \psi_m(y_2)\psi_n(y_2) w_2(y_2) dy_2 = 0,\ m \neq n$, can be employed to expand the function, $y_2 \mapsto a_j(y_2)$, in the following series [36, Chapter 4],

$$a_j(y_2) = \lim_{K\to\infty} \sum_{k=0}^{K} b_{jk}\psi_k(y_2). \tag{12}$$

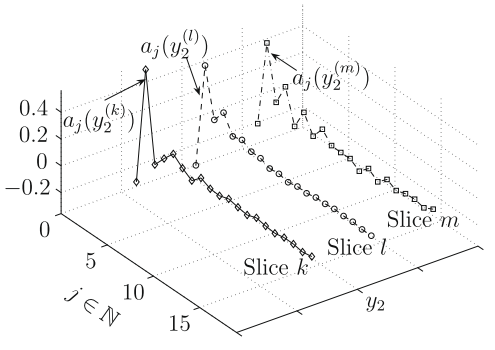


Fig. 6. 2-Dimensional illustration: $j \mapsto a_j(y_2)$ for given y_2.

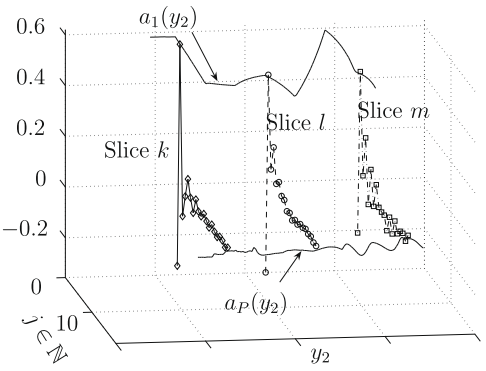


Fig. 7. 2-Dimensional illustration: $y_2 \mapsto a_j(y_2)$ for given $j \in \mathbb{N}$.

This series expansion is valid at every continuity point of a_j with b_{jk} computed from,

$$b_{jk} = \frac{\int_{s_{\mathbf{y}_2}} a_j(y_2)\psi_k(y_2)w_2(y_2)dy_2}{\int_{s_{\mathbf{y}_2}} \psi_k^2(y_2)w_2(y_2)dy_2}. \tag{13}$$

The denominator is readily available in the literature for many commonly used orthogonal polynomials [19]. The numerator can be evaluated by using any standard numerical integration scheme.

Use of (12) in (10) results in,

$$\mathbf{y}_1 \overset{d}{=} f_{1|2}(\xi_1|y_2) = \lim_{\substack{K_1\to\infty \\ K_2\to\infty}} \sum_{j=0}^{K_1}\sum_{k=0}^{K_2} b_{jk}\psi_k(y_2)\mathbf{\Psi}_j(\xi_1). \tag{14}$$

The marPDF P_2 of $\mathbf{y}_2$ can also be similarly (consider 1-dimensional cases of the series of Figs. 1–6) employed to obtain the following PC expansion for $\mathbf{y}_2$,

$$\mathbf{y}_2 \overset{d}{=} f_2(\xi_2) = \lim_{K\to\infty} \sum_{j=0}^{K} c_j\mathbf{\Psi}_j(\xi_2), \tag{15}$$

in which $f_2 \equiv P_2^{-1}P_{\xi_2}$ and c_j is given by,

$$c_j = \frac{E\left[f_2(\xi_2)\mathbf{\Psi}_j(\xi_2)\right]}{E\left[\mathbf{\Psi}_j^2(\xi_2)\right]}, \quad j \in \mathbb{N}, \tag{16}$$

and can be efficiently computed by using the simple scheme described in Appendix A.

The PC expansions, (14) and (15), constructed from the available measurement data, completely characterize the random vector $\mathcal{Y} = [\mathbf{y}_1, \mathbf{y}_2]^T$. In a computational setting, the series in (14) and (15) are truncated after a suitable large number of terms.

Sampling of $\mathcal{Y}$ is straightforward. The random variables, ξ_1 and ξ_2, are statistically independent. First, use (15) to generate a sample, y_2, of $\mathbf{y}_2$ and then use the realized value, y_2, in (14) to get y_1. Repeat the process until the desired number of samples of $\mathcal{Y} = [\mathbf{y}_1, \mathbf{y}_2]^T$ is generated.

Extension of the above 2-dimensional formulation to the N-variate $\mathcal{Y}$ is now summarized below,

$$\begin{aligned}
\mathbf{y}_1 &\overset{d}{=} P_{1|2:N}^{-1}P_{\xi_1}(\xi_1|y_2,\ldots,y_N) = \sum_{i_1=0}^{K_1^{(1)}}\cdots\sum_{i_N=0}^{K_N^{(1)}} b^{(1)}_{i_1 i_2\cdots i_N}\psi_{i_N}(y_N)\cdots\psi_{i_2}(y_2)\mathbf{\Psi}_{i_1}(\xi_1),\\
\mathbf{y}_2 &\overset{d}{=} P_{2|3:N}^{-1}P_{\xi_2}(\xi_2|y_3,\ldots,y_N) = \sum_{i_2=0}^{K_2^{(2)}}\cdots\sum_{i_N=0}^{K_N^{(2)}} b^{(2)}_{i_2\cdots i_N}\psi_{i_N}(y_N)\cdots\psi_{i_3}(y_3)\mathbf{\Psi}_{i_2}(\xi_2),\\
&\vdots\\
\mathbf{y}_N &\overset{d}{=} P_N^{-1}P_{\xi_N}(\xi_N) = \sum_{i_N=0}^{K_N^{(N)}} b^{(N)}_{i_N}\mathbf{\Psi}_{i_N}(\xi_N).
\end{aligned}$$

Here, $P_{i|(i+1):N}$ is the conditional PDF of $\mathbf{y}_i$, given $\mathbf{y}_{i+1} = y_{i+1},\ldots,\mathbf{y}_N = y_N$, induced by $P_{\mathcal{Y}}$ and $b^{(i)}_{j_i j_{i+1}\cdots j_N}$ represents $(N-(i-1))$-dimensional matrix of PC coefficients of size $K_i^{(i)}\times\cdots\times K_N^{(i)}$ with $K_i^{(i)},\ldots,K_N^{(i)}$ being the suitable large integers retained in

the corresponding series expansion. The random variables, $\xi_1,\ldots,\xi_N$, are statistically independent. Each digital sample of $\mathcal{Y}$ is generated starting with sampling $\mathbf{y}_N$ and successively proceeding towards sampling $\mathbf{y}_{N-1},\mathbf{y}_{N-2},\ldots,\mathbf{y}_1$.

Finally, let us conclude this section by emphasizing that $P_{i|(i+1):N}$ should not be computed by integrating $p_{\mathcal{Y}}$ since it would then involve a substantial computational effort to perform several multidimensional integrations while approximating the corresponding function, $P^{-1}_{i|(i+1):N}P_{\xi_i}$ (see Appendix A). Instead, $P_{i|(i+1):N}$ should be computed from the estimate of mjpdf of $(\mathbf{y}_i,\ldots,\mathbf{y}_N)$ determined by considering only the measurement data associated with $\mathbf{y}_i,\ldots,\mathbf{y}_N$, and ignoring the data associated with $\mathbf{y}_1,\ldots,\mathbf{y}_{i-1}$. This would always involve 1-dimensional integration in computation of $P_{i|(i+1):N}$,

$$P_{i|(i+1):N}(y_i|y_{i+1},\ldots,y_N)=\frac{\int_{l_i}^{y_i}p_{\mathbf{y}_i,\ldots,\mathbf{y}_N}(y_i,\ldots,y_N)dy_i}{p_{\mathbf{y}_{i+1},\ldots,\mathbf{y}_N}(y_{i+1},\ldots,y_N)}.$$

Here, the integration is carried over the domain, $[l_i,y_i]\subseteq s_{\mathbf{y}_i}=[l_i,m_i]\subset\mathbb{R}$, where $s_{\mathbf{y}_i}$ is the support of $\mathbf{y}_i$. This scheme would be relatively computationally inexpensive even after the additional computational overhead required to estimate the set of pdfs, $p_{\mathbf{y}_2,\ldots,\mathbf{y}_N},p_{\mathbf{y}_3,\ldots,\mathbf{y}_N},\ldots,p_{\mathbf{y}_N}$, (from the corresponding data) that need to determined only once at the outset.

3.2. Approach 2: based on marginal PDFs and SRCC

In this approach, the unknown relationship between $\underline{\xi}$ and $\mathcal{Y}$ is defined again by having recourse to the Rosenblatt transformation establishing a set of N mappings, each of which is similar to (15), between the corresponding kth components, $\mathbf{y}_k$ and ξ_k, $k=1,\ldots,N$. Only marPDF of $\mathbf{y}_k$ is used in this approach. Unlike ξ_k's in Approach 1, the random variables, $\xi_1,\ldots,\xi_N$, here are, however, statistically dependent enforcing the required statistical dependencies among $\mathbf{y}_k$'s. The statistical dependency is characterized via SRCC. As indicated earlier, $P_{\mathcal{Y}}$ in this approach is assumed to be completely characterized by the marPDFs and the SRCC matrix of $\mathcal{Y}$. In the following, the definition and the relevant properties of SRCC are briefly reviewed before describing Approach 2.

3.2.1. Spearman's rank correlation coefficient

The rank correlation coefficient or Spearman's ρ is named after Charles Edward Spearman who first introduced it in 1904. The rank correlation coefficient between random variables, $\mathbf{y}_i$ and $\mathbf{y}_j$, is simply the PCC applied to the rank of the observed samples of $\mathbf{y}_i$ and $\mathbf{y}_j$ rather than to their observed or measured values. When there is no tie in the observed values of the data, a simple formula exists for the calculation of SRCC Further theoretical treatment and calculation procedure of SRCC including the case of tied data values can be found in the literature [9]. The statistical toolbox of MATLAB provides the function, `corr`, that can be used to calculate SRCC.

Definition 1. The *Spearman's rank correlation coefficient* between two random variables, $\mathbf{y}_i$ and $\mathbf{y}_j$, with marginal probability distribution functions, respectively, being given by $P_{\mathbf{y}_i}$ and $P_{\mathbf{y}_j}$, is defined as,

$$\rho_s(\mathbf{y}_i,\mathbf{y}_j)=\rho(P_{\mathbf{y}_i}(\mathbf{y}_i),P_{\mathbf{y}_j}(\mathbf{y}_j))=12\mathrm{cov}(P_{\mathbf{y}_i}(\mathbf{y}_i),P_{\mathbf{y}_j}(\mathbf{y}_j)). \tag{17}$$

Here, ρ is the Pearson's correlation coefficient (or the usual product–moment correlation coefficient), cov is the covariance and the multiplying factor, 12, emanates from variance of $P_{\mathbf{y}_k}(\mathbf{y}_k),k=i,j$, since $P_{\mathbf{y}_k}(\mathbf{y}_k)\sim U(0,1)$ with $U(0,1)$ being uniform distribution on [0,1]; see e.g., [10, Theorem 2.1].

It must be noted from the above definition that SRCC and PCC coincide if PDFs of $\mathbf{y}_i$ and $\mathbf{y}_j$ are $U(0,1)$. In general, they are, however, different.

A collection of a few salient properties of ρ_s is enlisted below [8,37]. It

- always exists and is symmetric;
- is independent of marpdf of $\mathbf{y}_i$ and $\mathbf{y}_j$;
- is invariant under strictly monotone transformation of $\mathbf{y}_i$ and $\mathbf{y}_j$;
- can take any values from the closed interval, $[-1,1]$;
- is zero if $\mathbf{y}_i$ and $\mathbf{y}_j$ are statistically independent, the converse is not true.

The most important property to be used in the present work is invariance under monotone transformation property of SRCC.

Now that the relevant information on SRCC is set forth, Approach 2 is described below by introducing the mapping, $\xi_k\mapsto\mathbf{y}_k$, $k=1,\ldots,N$, through the use of the Rosenblatt transformation [4] applied on each ξ_k separately,

$$\mathbf{y}_k\overset{d}{=}q_k(\xi_k)=\lim_{K_k\to\infty}\sum_{j=0}^{K_k}c_{jk}\mathbf{\Psi}_j(\xi_k),\quad q_k\equiv P^{-1}_{\mathbf{y}_k}P_{\xi_k}. \tag{18}$$

This PC representation is similar to (15). The marPDF, $P_{\mathbf{y}_k}$, is estimated from the normalized and linearly interpolated 1-dimensional histogram of the measurement data on each random variable component, $\mathbf{y}_k$, separately. This can be readily per-

formed as already discussed in Section 3.1. The PC representation of q_k in (18) is, therefore, valid at every continuity of q_k implying that the equality, '=', can also be interpreted in a.s. sense w.r.t. P_{ξ_k}. The PC coefficient, c_{jk}, is given by,

$$c_{jk} = \frac{E\left[q_k(\xi_k)\Psi_j(\xi_k)\right]}{E\left[\Psi_j^2(\xi_k)\right]}, \quad j \in \mathbb{N}. \tag{19}$$

A simple and computationally efficient scheme based on 1-dimensional interpolated surrogate function, approximating $P_{\mathbf{y}_k}^{-1}P_{\xi_k}$, is described in Appendix A to determine $\{c_{jk}\}_{j\in\mathbf{N}},\ k = 1,\ldots,N$. The series in (18) is truncated after a large number of terms K_k in a computational setting.

Since SRCC is preserved under monotone transformation, the SRCC matrices of $\underline{\xi} = [\xi_1,\ldots,\xi_N]^T$ and $\mathcal{Y}$ are identical. The target $N \times N$ SRCC matrix, $[\rho_s]$, is simply estimated from the available measurement data on $\mathcal{Y}$. For $i,j = 1,\ldots,N$, the (i,j)th element of $[\rho_s]$ is denoted by $(\rho_s)_{ij}$, where $(\rho_s)_{ij} = \rho_s(\mathbf{y}_i,\mathbf{y}_j)$. The samples of $\underline{\xi}$, with SRCC matrix $[\rho_s]$, are first generated. Subsequently, samples of each ξ_k are substituted in the corresponding PC expansion of $\mathbf{y}_k$ to obtain the realizations of $\mathbf{y}_k$. The resulting samples of $\mathcal{Y}$ are consistent with the target set, $\{p_{\mathbf{y}_k}\}_{k=1}^N$, of marpdfs and the target SRCC matrix $[\rho_s]$.

The commonly used PC random variables, $\xi_1,\ldots,\xi_N$, that are often chosen to construct PC representation, are standard Gaussian random variables, uniform random variables on $[-1,1]$, beta type I random variables on $[-1,1]$ or gamma random variables. The generation of samples of such statistically independent random variables, as required in Approach 1, is straightforward. The samples of statistically dependent random variables, particularly when the statistical dependency is characterized by a specified SRCC matrix $[\rho_s]$, as required in Approach 2, can also be generated by using the existing simulation techniques. For the sake of completeness of the present work, two such useful and easily implementable techniques are summarized in the next two subsections. These two techniques are directly related to concept of copula [38,39], knowledge of which, though useful, is not required here.

3.2.2. Normal copula technique

This technique assumes existence of an underlying correlated N-dimensional standard Gaussian random vector, $\mathbf{X} = [\mathbf{x}_1,\ldots,\mathbf{x}_N]^T$, in which each component $\mathbf{x}_i$ is a standard Gaussian random variable. Then, the correlation (also, covariance) matrix $[\rho]$ of $\mathbf{X}$ is determined as follows.

It was shown by Pearson in 1904 that [37] the PCC and SRCC are related according to,

$$\rho(\mathbf{x}_i,\mathbf{x}_j) = 2\sin\left(\frac{\pi}{6}\rho_s(\mathbf{u}_i,\mathbf{u}_j)\right), \tag{20}$$

in which $\mathbf{u}_i \sim U(0,1)$ and $\mathbf{u}_j \sim U(0,1)$ are uniform random variables. If the PDF of the standard Gaussian random variable is denoted by $\Phi(\cdot)$, then $\Phi(\mathbf{x}_i) \sim U(0,1),\ \forall i$ [10, Theorem 2.1]. Let us then select $\mathbf{u}_i$'s in (20) as $\mathbf{u}_i \equiv \Phi(\mathbf{x}_i)$. Consider now the following mapping based on the Rosenblatt transformation,

$$\mathbf{u}_i \equiv \Phi(\mathbf{x}_i) \stackrel{d}{=} P_{\mathbf{y}_i}(\mathbf{y}_i), \quad i = 1,\ldots,N, \tag{21}$$

since $P_{\mathbf{y}_i}(\mathbf{y}_i) \sim U(0,1)$ [10, Theorem 2.1]. By the invariance under monotone transformation property of the SRCC, we have $\rho_s(\mathbf{y}_i,\mathbf{y}_j) = \rho_s(P_{\mathbf{y}_i}(\mathbf{y}_i),P_{\mathbf{y}_j}(\mathbf{y}_j))$. Then, by (17) and "$\stackrel{d}{=}$" in (21), the SRCC matrix of $\mathbf{U} = [\mathbf{u}_1,\ldots,\mathbf{u}_N]^T$ is given by $[\rho_s]$ with its (i,j)th element being given by $\rho_s(\mathbf{y}_i,\mathbf{y}_j)$ estimated based on the measurement data on $\mathcal{Y}$. The correlation (or covariance) matrix $[\rho]$ of $\mathbf{X}$ then follows from (20), with the (i,j)th, $i,j = 1,\ldots,N$, element ρ_{ij} of $[\rho]$ being given by $\rho(\mathbf{x}_i,\mathbf{x}_j)$. Simulation of the standard Gaussian random vector $\mathbf{X}$ with covariance matrix $[\rho]$ is then straightforward. In the literature, PDF of $\mathbf{U}$ is usually referred as normal copula.

Since $P_{\xi_i}(\xi_i) \sim U(0,1)$ [10, Theorem 2.1], use of the following transformation (again based on the Rosenblatt transformation),

$$\left.\begin{aligned} P_{\xi_i}(\xi_i) \stackrel{d}{=} \Phi(\mathbf{x}_i) \equiv \mathbf{u}_i \\ \Rightarrow \xi_i \stackrel{d}{=} P_{\xi_i}^{-1}\Phi(\mathbf{x}_i), \end{aligned}\right\}, \quad i = 1,\ldots,N, \tag{22}$$

yields the samples of $\underline{\xi} = [\xi_1,\ldots,\xi_N]^T$. The SRCC matrix of $\underline{\xi}$ again turns out to be $[\rho_s]$ by the invariance under monotone transformation property of the SRCC. The closed form expression of, or the efficient algorithm to compute the inverse function, $P_{\xi_i}^{-1}$, associated with the commonly used PC random variables can be readily extracted from the standard textbook on MC simulation [40,10].

The MATLAB statistical toolbox provides many such useful functions. Clearly, the simulation of $\underline{\xi}$ with SRCC matrix $[\rho_s]$ essentially reduces to the simulation of an N-dimensional standard Gaussian random vector with covariance matrix $[\rho]$ (if it exists).

Let us consider the last remark about the existence of feasible covariance matrix of $\mathbf{X}$ more carefully. Denote the set of symmetric $N \times N$ positive-definite real matrices by $\mathbb{M}_N^+(\mathbb{R})$ and $\mathbb{S}_N(\mathbb{R}) = \{A : A \in \mathbb{M}_N^+(\mathbb{R}),\ A_{ii} = 1\}$, in which A_{ij} is (i,j)th element of A. Then, for any $[\rho_s] \in \mathbb{S}_M(\mathbb{R})$, there always exists a random vector with uniform marPDFs and SRCC matrix $[\rho_s]$ [37, Theorem 4.4, pp. 100, 124–125]. It does not, however, necessarily mean that its uniform random variable components can be given by $\Phi(\mathbf{x}_i)$'s. Counterexamples exist in the literature ([7]; [37, Section 4.2]; [10, Section 12.5.2]; [37, Section 4.2]) showing that application of the mapping defined by (20) on each element $(\rho_s)_{ij}$ of $[\rho_s]$ may produce a matrix $[\rho^{(1)}]$ that is not a

positive-definite matrix, thus limiting the applicability of the normal copula technique. This problem becomes more common as the dimension N of the random vector increases. An adaptation of the normal copula technique has been developed to address this difficulty; it is described next. However, if $\mathbf{X}$ exists, i.e., if a feasible (positive-definite) covariance matrix is found, then the normal copula technique is the fastest method among all the currently existing methods.

3.2.3. Augmented normal copula techniques

Application of these techniques ensures that the samples of $\mathbf{U}$ follow the uniform marPDFs but the SRCC or PCC matrix (identical by definition for uniform distribution) is approximate in the sense that the target correlation matrix is modified to a 'new' correlation matrix that is close, in some sense, to the originally estimated correlation matrix. Let us denote the original matrix by $[\rho_s^{(1)}]$ and the modified positive-definite correlation matrix by $[\rho_s]$. With this new target correlation matrix $[\rho_s]$, the use of the normal copula technique, as described in the previous subsection, becomes feasible.

One such technique [13, Section 5] suggests to adapt $[\rho_s^{(1)}]$ and $[\rho^{(1)}]$, to new positive-definite correlation matrices, $[\rho_s]$ and $[\rho]$, by using a simple iterative scheme based on the spectral decomposition of real Hermitian matrices. While this scheme might work in practice, it is likely to be little unwieldy, particularly in high dimension, requiring too many iterations often resulting in relatively large error between the old and modified matrices.

Another technique [11] is a constrained minimization problem in the space of $\mathbf{X}$ and relatively more robust. The norms $L_1 = \sum_{i<j}|\rho_{ij} - \rho_{ij}^{(1)}|$and $L_\infty = \max_{i<j}|\rho_{ij} - \rho_{ij}^{(1)}|$ are minimized subject to $[\rho] \in \mathbb{S}_N(\mathbb{R})$ [11]. It is, however, not guaranteed that the resulting 'new' correlation matrix $[\rho_s]$ of $\mathbf{U}$ (by applying the inverse transformation of (20) on $[\rho]$) would be positive-definite and close to the originally specified target correlation matrix $[\rho_s^{(1)}]$ of $\mathbf{U}$. In such a situation, an iterative scheme like the one proposed earlier in the literature [13, Section 5] might be adopted.

In the present work, the following constrained minimization problem, similar to the works presented in earlier literature [11], is recommended,

$$\begin{aligned} &\text{minimize} && \|[\rho] - [\rho^{(1)}]\|_F \\ &\text{subject to} && [\rho] \in \mathbb{S}_N(\mathbb{R}), \end{aligned} \tag{23}$$

or/and other meaningful constraints; see e.g., [11, Section 5]. Here, $\|\cdot\|_F$ is the Frobenius (matrix) norm defined by $\|C\|_F = (\sum_{ij}|c_{ij}|^2)^{1/2}$, in which c_{ij} is the (i,j)th element of C. The Frobenius norm is preferred (over L_1 and L_∞ norms) since it shows relatively much smaller error (even in high dimension). The above optimization problem can be efficiently solved by employing the semidefinite program (SDP) [41]. Many efficient freely available softwares[2] exist to solve such SDP. In the present work, a public domain MATLAB toolbox, YALMIP, developed by Löfberg [42], is used.

The techniques as discussed above should be applied only if the new correlation matrix $[\rho_s]$ of $\mathbf{U}$ is positive-definite and is close, in the appropriate sense, to the originally specified target correlation matrix, $[\rho_s^{(1)}]$. Otherwise, alternative techniques [7,37] at the expense of additional computational time and resource might be investigated. In many practical applications, the two recommended techniques—normal copula technique and augmented normal copula technique—are, nevertheless, likely to be satisfactory.

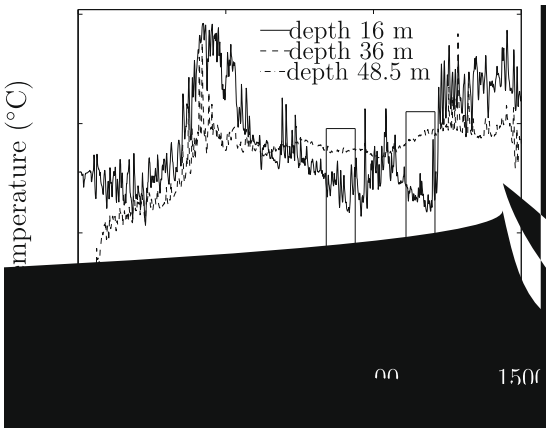
4. Practical illustration and discussion

The proposed techniques are employed here to construct the PC representation of a spatio-temporal random temperature field by using a set of oceanographic data obtained from a shallow-water acoustics transmission experiment. This experiment will be referred to in the sequel as SWARM95 (Shallow Water Acoustics in Random Medium). It was conducted during the month of July–August in 1995 in the Mid-Atlantic Bight continental shelf region off the coast of New Jersey [1].

The primary objective of the SWARM95 experiment was to investigate the effects of random variations of the oceanographic parameters, for example, temperature and salinity fields, on the statistical properties of an acoustic field propagating through the water column bounding the ocean surface and the ocean bottom. The acoustic field is perturbed significantly by a small change in water column sound speed distribution. The sound speed variation is a function of an internal gravity wave field that, in turn, depends on water temperature and salinity distributions. Both temperature and salinity typically have a small random component. This internal wave field is also governed by partial differential equations with random coefficients depending on the oceanographic parameters. Further details and precise objectives of the experiment are documented and discussed in other papers [1,43]. In this section, modeling of the spatio-temporal random temperature field from the oceanographic measurements of SWARM95 experiment is considered. The PC representation of the spatio-temporal random field modeling the oceanographic parameters would be useful in propagating the uncertainty in a rational manner in order to predict the statistical properties of the acoustic field and in estimating the confidence interval of the associated statistical parameters by employing the techniques available elsewhere [44,45].

There are three vertically oriented thermister strings located within a water column of maximum depth 72 m. Each string contains 11 temperature sensors measuring the temperature histories. These temperature senors are located at depth $h \in D = \{16, 21, 26, 31, 36, 38.5, 41, 46, 48.5, 51, 56\}$ m. The temperature data are sampled every minute for a total of 17,281 samples from each sensor. The three strings will hereafter be referred to as tav309, tav307 and tav598. A few typical

[2] http://www-user.tu-chemnitz.de/~helmberg/semidef.html.



time histories obtained from tav309 are shown in Fig. 8. Both linear (background) and nonlinear internal wave fields are present in the data. The later appears in the form of solitary wave trains, while the former are spatially diffuse and typically treated as a random field. It is necessary to separate the background internal wave field from the solitary wave contribution while computing some intermediate oceanographic parameters, for example, buoyancy frequency, that is required to compute the sound speed fluctuation [43]. Therefore, only the "quiescent" part of the measurement data excluding the solitary waves must be used while computing such intermediate parameters. The most active solitary wave region is in the upper half of the water column.

4.1. Selecting the regions of low internal solitary wave activity

There is some subjectivity in choosing the quiescent part of the temperature data because it is next to impossible to completely separate the background internal wave field from the solitary wave contribution. The mathematical decomposition of the sound speed distribution into deterministic, time-dependent field and a random fluctuation about this deterministic field, as discussed in previous work [43], is an idealization. In a real ocean experiment, the situation is much more complicated. In order to estimate the oceanographic parameters, e.g., buoyancy frequency, it is important to try to stay away from regions containing the obvious large fluctuations that often start with a jump discontinuity. These regions are usually associated with the main components of the solitary wave train. Therefore, the highly variable regions, containing the strong solitary wave activity, are not used in the following analysis.

By visual inspection, the regions in the boxes, for example, as shown in Fig. 8, are examples of "low" internal solitary wave activity and suitable for reliable estimation of the buoyancy frequency, and consequently, selected for further analysis. A total of 8 time-segments each with 99 temperature measurements at any $h \in D$ are selected from the whole span of the experimentally measured time history. Out of 17,281 samples available from each sensor, only 8×99 samples are deemed to be useful in constructing the PC representation of the spatio-temporal random temperature field. The resulting PC representation would be useful for other analysis involving (stochastic) oceanographic parameters that depend on the random temperature field.

More detailed features of a typical quiescent segment showing the time histories collected from a few sensors (at different depths) attached to one of the three strings (tav309) are shown in Fig. 9. Each quiescent segment with 99 samples is further divided into nine smaller segments with each of these smaller segments containing 11 samples as shown in this figure.

At any given time instant, all the 11 sensors located at $h \in D$ are measuring the temperature (at 1 min sampling rate) simultaneously. Consider a spatio-temporal domain defined by one *smaller* segment associated with the quiescent zone and the 72 m depth of water column within which SWARM95 experiment was conducted. Let us assume that the random temperature field is statistically independent and identically distributed (i.i.d.) both across the smaller segments with 11 samples as shown in Fig. 9 within a given quiescent zone as well as across the different quiescent zones as shown in Fig. 8. Without any further loss of generality, time can be reset to $t = 0$ at the beginning of each of these smaller segments as illustrated in Fig. 10. Denote the spatio-temporal domain thus described by $(\mathcal{T} \times \mathcal{D})$ in which $\mathcal{T} = (0, 11)$ min and $\mathcal{D} = (0, 72)$ m. Denote the random temperature field evolving over $(\mathcal{T} \times \mathcal{D})$ by $(\mathcal{T} \times \mathcal{D}) \ni (t, h) \mapsto \mathbf{\Gamma}(t, h) \in \mathbb{R}$.

4.2. Detrending the data

The average trends of the oscillatory time histories are obtained by fitting the data linearly within each smaller segment as shown as dotted lines in Fig. 10. Within a given segment, suppose that the experimentally measured data, for any given

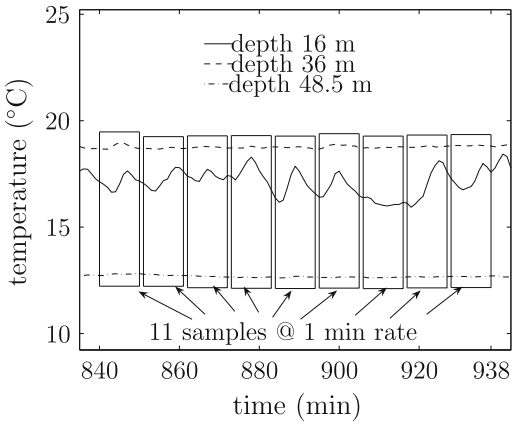


Fig. 9. A typical quiescent zone divided into nine smaller segments with 11 samples (shown for a few sensors).

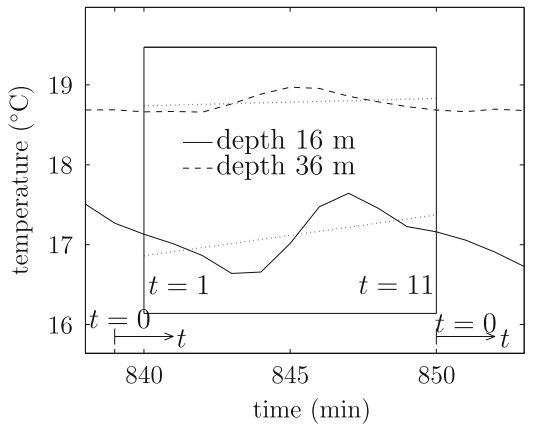


Fig. 10. A typical subset of $(\mathcal{T} \times \mathcal{D})$ with two time histories collected from tav309; dotted lines indicate linear fit to the experimental data.

$h \in D$, is represented by $\Gamma^{(\text{meas})}(t,h)$ and the linear trend of the measurement by $\overline{\Gamma}(t,h)$. Then, define a normalized spatio-temporal random temperature field, $\mathbf{\Gamma}^{(n)}(t,h)$, as,

$$\mathbf{\Gamma}^{(n)}(t,h) = \frac{\mathbf{\Gamma}(t,h) - \overline{\Gamma}(t,h)}{\overline{\Gamma}(t,h)}. \tag{24}$$

The experimental samples of $\mathbf{\Gamma}^{(n)}(t,h)$ can be readily deduced by substituting $\mathbf{\Gamma}(t,h)$ with $\Gamma^{(\text{meas})}(t,h)$ in (24). A few such typical experimental samples of $\mathbf{\Gamma}^{(n)}(t,h)$ are shown in Fig. 11.

In the following, $\mathbf{\Gamma}^{(n)}(t,h)$ is modeled by employing the approaches as proposed in the present work based on the resulting experimental samples. Once the PC representation of $\mathbf{\Gamma}^{(n)}(t,h)$ is available, the PC representation of the original random temperature field, $\mathbf{\Gamma}(t,h)$, immediately follows from $\mathbf{\Gamma}(t,h) = \overline{\Gamma}(t,h)\mathbf{\Gamma}^{(n)}(t,h) + \overline{\Gamma}(t,h)$. The linear fit, $\overline{\Gamma}(t,h)$, has already been deduced by using the experimental samples of $\mathbf{\Gamma}(t,h)$. The separation of this average trend from $\mathbf{\Gamma}(t,h)$ essentially adds a certain flexibility to the scheme of modeling $\mathbf{\Gamma}(t,h)$ as adopted in this numerical illustration. This, in particular, facilitates in inferring the PC coefficients of $\mathbf{\Gamma}(t,h), (t,h) \notin (\mathcal{T} \times \mathcal{D})$ (assuming that the corresponding $\overline{\Gamma}(t,h)$ can be reliably estimated from the experiment or is available from other sources/experiments). The normalization by $\overline{\Gamma}(t,h)$ as shown in (24) also facilitates in achieving certain numerical stability to the ensuing analysis since values of the experimental measurements collected from sensors at different depths show significant variations (see Fig. 12). This variation should be compared with the variation after the normalization as shown in Fig. 13.

4.3. Stochastic modeling of $\mathbf{\Gamma}^{(n)}(t,h)$

For any given $(t,h) \in (\mathcal{T} \times \mathcal{D})$, $\mathbf{\Gamma}^{(n)}(t,h)$ represent a random variable. Clearly, the experimental measurements essentially represent the samples of a finite set of these random variables. Recall that D represents the set of depth coordinates of the

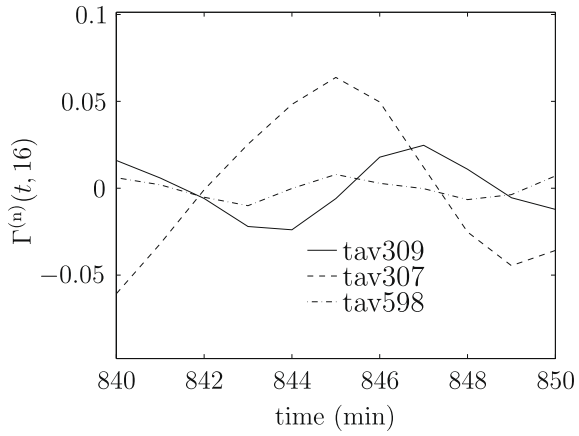


Fig. 11. A few typical profiles of experimental samples of $\mathbf{\Gamma}^{(n)}(t,h)$; $(t,h)\mapsto\Gamma^{(n)}(t,h)$ at $h = 16$ m depth.

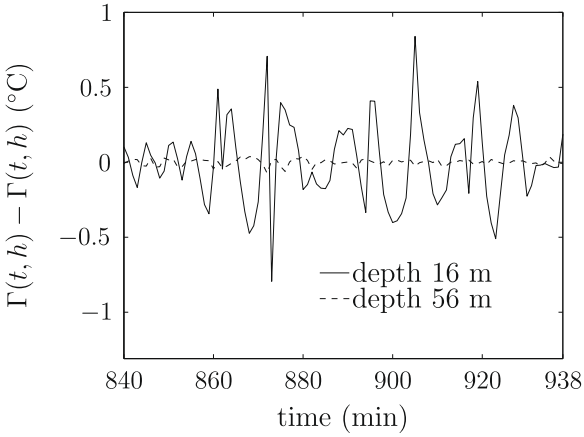


Fig. 12. Experimental variation of temperature measurements after removing the linear trends and before normalization (shown for two time histories and over a quiescent zone).

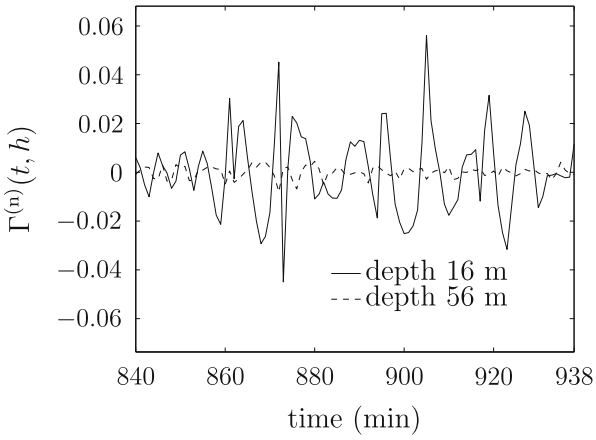


Fig. 13. Variation of the normalized temperature measurements (shown for two time histories and over a quiescent zone).

sensors attached to the thermister strings along 72 m water column. Let us now also denote the set of time instants (per convention of Fig. 10) of the collection of experimental samples by $T = \{1, 2, \ldots, 10, 11\}$ (since sampling rate is 1 min). Note the difference between the continuous space, $(\mathcal{T} \times \mathcal{D})$, over which $\mathbf{\Gamma}^{(n)}(t,h)$ is evolving and the discrete space, $(T \times D)$, consisting of only a finite set of points at which the experimental samples are available.

Let us denote the set of 11×11 random variables, $\{\mathbf{\Gamma}^{(n)}(t,h)\}_{(t,h)\in(T\times D)}$, collectively by $\mathcal{Y}$, i.e., $N = \dim(\mathcal{Y}) = 121$. Since each quiescent zone is divided into nine smaller segments (see Fig. 9) and eight quiescent zones are selected (see Section 4.1) across the whole span of the experimental time histories, there are $8 \times 9 = 72$ statistically independent samples of $\mathcal{Y}$ from each string. In the present work, a space and time separability condition of statistical dependency of the original random temperature field is assumed for the time and spatial extent spanning the sea surface. However, no such space–time separability is assumed for the time and depth, i.e., for $(\mathcal{T} \times \mathcal{D})$. The random variable components of $\mathcal{Y}$ are, therefore, statistically dependent. Furthermore, we note again that the underlying spatio-temporal stochastic process is modeled here as a non-stationary random field evolving over both time and depth (i.e., non-stationary in time and non-homogeneous in depth). From three vertical strings, tav309, tav307 and tav598 (about 10 km away from each other), a total of $n = 3 \times 72 = 216$ samples of $\mathcal{Y}$ are available.

The task is now to construct PC representations of $\mathcal{Y}$ by using the approaches as proposed in the present work with 216 samples of $\mathcal{Y}$. The PC representations would be consistent with the information extracted from these 216 experimental samples. Further details and results are discussed in the next subsections.

4.4. Modeling of $\mathcal{Y}$ via Approach 1

The Karhunen–Loève (KL) decomposition is first employed to construct a reduced order model of the non-Gaussian random vector, $\mathcal{Y}$. Though the resulting non-Gaussian KL random variable components are uncorrelated, they are, in general, statistically dependent. Approach 1 is subsequently used to characterize this reduced order model of $\mathcal{Y}$.

4.4.1. KL decomposition of $\mathcal{Y}$

The KL expansion [46, Chapter 11] of a stochastic process represents the *uncountably* infinite number of random variables, that characterize the stochastic process, as a linear decomposition of *countably* infinite number of random variables, often referred as KL random variables. The orthogonal basis functions of this linear decomposition are given by eigenfunctions of the covariance function of the stochastic process. The coordinates of the decomposition constitute the set of countably infinite number of random variables, and are global variables (like Fourier coefficients). The non-stationarity of the stochastic process is captured by the basis functions of the linear decomposition, i.e., eigenfunctions of the covariance function. For a stationary process, the KL expansion reduces to the usual spectral representation [3,47]. A finite number of dominant KL random variables is typically retained in the linear decomposition to obtain a KL approximation. The number of dominant KL random variables depends on the stochastic fluctuations of the stochastic process over its spatio-temporal domain. The KL approximation converges in mean-square sense to the stochastic process [46, Section 37.5]. It can be shown that the KL approximation is optimal in the sense of minimum MSE [15]. Furthermore, the mean-square convergence can be shown to be uniform on the spatio-temporal domain over which the stochastic process evolves [46]. The mean-square convergence criteria remains valid irrespective of the non-stationarity of the process.

For a finite-dimensional representation $\mathcal{Y}$ of the stochastic process, the KL approximation is typically employed as a device to further reduce the stochastic dimension (i.e., N) as described next. Let us denote the $N \times N$ covariance matrix of $\mathcal{Y}$ by $\mathbf{\Sigma}_{\mathcal{Y}}$. If n experimental samples of $\mathcal{Y}$ are denoted by $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, then the covariance matrix $\mathbf{\Sigma}_{\mathcal{Y}}$ can be estimated by using the samples as $(1/(n-1))Y_o Y_o^T$. Here, $Y_o = [\mathcal{Y}_{1o}, \ldots, \mathcal{Y}_{no}]$ represents an $N \times n$ matrix and $\mathcal{Y}_{ko} \equiv \mathcal{Y}_k - \overline{\mathcal{Y}},\ k = 1, \ldots, n$, with $\overline{\mathcal{Y}}$ being unbiased estimate of the mean vector of $\mathcal{Y}$, i.e., $\overline{\mathcal{Y}} = (1/n)\sum_{k=1}^{n} \mathcal{Y}_k$. Investigation of the effects of n and N on the estimate of $\mathbf{\Sigma}_{\mathcal{Y}}$ is beyond the scope of the present work. It is assumed here that these effects are negligible, and this estimate is used in the KL approximation as described next.

Following the KL expansion procedure [46, Chapter 11], let us collect the dominant KL random variable components, $\{\mathbf{z}'_1, \ldots, \mathbf{z}'_M\},\ M \leqslant N$, in an M-dimensional random vector, $\mathcal{Z}' = [\mathbf{z}'_1, \ldots, \mathbf{z}'_M]^T$. If V_i is the eigenvector of $\mathbf{\Sigma}_{\mathcal{Y}}$ associated with the ith largest eigenvalue of $\mathbf{\Sigma}_{\mathcal{Y}}$, then $\mathcal{Z}'$ is related to $\mathcal{Y}$ by,

$$\mathcal{Z}' = \mathcal{V}_M^T(\mathcal{Y} - \overline{\mathcal{Y}}), \tag{25}$$

in which $\mathcal{V}_M = [V_1, \ldots, V_M]$ is the $N \times M$ matrix of eigenvectors $\{V_1, \ldots, V_M\}$ associated with the M largest eigenvalues. The eigenvectors are chosen such that $\{V_i\}_{i=1}^N$ is a set of orthonormal eigenvectors. The variance of the ith KL random variable $\mathbf{z}'_i$ is given by the ith largest eigenvalue of $\mathbf{\Sigma}_{\mathcal{Y}}$ to be denoted by $\sigma^2_{\mathbf{z}'_i}$. The value of M is selected such that $\mathrm{tr}(\mathbf{\Sigma}_{\mathcal{Y}}) = \sum_{i=1}^{N}\mathrm{var}(\mathbf{y}_i) \approx \sum_{i=1}^{M}\mathrm{var}(\mathbf{z}'_i) = \sum_{i=1}^{M}\sigma^2_{\mathbf{z}'_i}$ with var and tr, respectively, representing variance and trace operators. The KL approximation $\mathcal{Y}^{(M)}$ of $\mathcal{Y}$ is then given by $\mathcal{Y}^{(M)} = \mathcal{V}_M\mathcal{Z}' + \overline{\mathcal{Y}}$, where $\mathcal{Z}'$ is defined by (25). If the error associated with KL approximation is measured by $\epsilon_M = \mathcal{Y} - \mathcal{Y}^{(M)} = \sum_{i=M+1}^{N} V_i\mathbf{z}'_i$, then the KL approximation $\mathcal{Y}^{(M)}$ converges in mean-square sense to $\mathcal{Y}$ implying that $E(\|\mathcal{Y} - \mathcal{Y}^{(M)}\|^2) \to 0$ as $M \to N$ [46,19]. Here, $\|\cdot\|$ represents Euclidean norm and $E(\cdot)$ represents expectation w.r.t. the measure of $\mathcal{Y}$. The associated minimum (optimal) MSE is given by $E(\|\mathcal{Y} - \mathcal{Y}^{(M)}\|^2) = E(\epsilon_M^T\epsilon_M) = \sum_{i=M+1}^{N}\mathrm{var}(\mathbf{z}'_i)$ [19].

The set of experimental samples of $\mathcal{Z}'$ can be immediately obtained by replacing $\mathcal{Y}$ with $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$ in (25) resulting in $\mathcal{Z}'_1, \ldots, \mathcal{Z}'_n$. Let $s_{\mathbf{z}'_i} = [\alpha_i, \beta_i] \subset \mathbb{R}$ be the support of $\mathbf{z}'_i$, and let $\underline{a} = [\alpha_1, \ldots, \alpha_M]^T$ and $\underline{b} = [\beta_1, \ldots, \beta_M]^T$. To enhance the regularity of the ensuing numerical problem and to improve the efficiency of the associated computation, the data set is further scaled to obtain another data set as shown below,

$$\mathcal{Z}_k = 2\left[(\mathcal{Z}'_k - \underline{a}) \circ \left(\frac{1}{\underline{b} - \underline{a}}\right)\right] - \mathbf{1}_M, \quad k = 1, \ldots, n. \tag{26}$$

Here, the symbol $\circ$ represents element-wise product operator or the Hadamard product operator, $\mathbf{1}_M$ is an M-dimensional column vector of 1's, α_i and β_i can be estimated, respectively, as $\min(z_i'^{(1)},\ldots,z_i'^{(n)})$ and $\max(z_i'^{(1)},\ldots,z_i'^{(n)})$ with $z_i'^{(k)}$ being the ith component of the kth sample $\mathcal{Z}_k' = [z_1'^{(k)},\ldots,z_M'^{(k)}]$, and finally, $1/(\underline{b}-\underline{a})$ needs to be interpreted as $M\times 1$ column vector with its ith, $i=1,\ldots,M$, element being given by reciprocal of the ith element of $(\underline{b}-\underline{a})$. Denote this M-dimensional normalized KL random vector associated with the samples, $\{\mathcal{Z}_k\}_{k=1}^n$, by $\boldsymbol{\mathcal{Z}} = [\mathbf{z}_1,\ldots,\mathbf{z}_M]^T$. The scaling in (26) is particularly chosen so that $\boldsymbol{\mathcal{Z}}$ is supported on $[-1,1]^M$, which would be in concordance, in some sense, of the support, $[-1,1]$, of the uniform distributions used as measures of the PC random variables in Section 4.4.2. The following relation between $\boldsymbol{\mathcal{Z}}$ and $\boldsymbol{\mathcal{Y}}$ then holds,

$$\boldsymbol{\mathcal{Y}} \approx \boldsymbol{\mathcal{Y}}^{(M)} = \mathcal{V}_M\left[\underline{a} + \left\{(\underline{b}-\underline{a})\circ\frac{1}{2}(\boldsymbol{\mathcal{Z}}+\mathbf{1}_M)\right\}\right] + \overline{\mathcal{Y}}. \tag{27}$$

The approximation sign, '$\approx$', in (27) emphasizes that $\boldsymbol{\mathcal{Y}}$ is projected into the space spanned only by the largest M dominant eigenvectors of $\boldsymbol{\Sigma}_{\mathcal{Y}}$ to obtain the reduced order representation, $\boldsymbol{\mathcal{Z}}$.

The covariance matrix $\boldsymbol{\Sigma}_{\mathcal{Y}}$ is first estimated based on $n=216$ samples. Here, M is chosen such that $\sum_{i=1}^M \hat{\sigma}^2_{\mathbf{z}_i'} \approx \sum_{i=1}^M \sigma^2_{\mathbf{z}_i'} = 0.999\sum_{i=1}^N \mathrm{var}(\mathbf{y}_i)$ dictating that $M=78$ dominant KL random variables should be considered (recall that $N=\dim(\boldsymbol{\mathcal{Y}})=121$). Here, $\hat{\sigma}_{\mathbf{z}_i'}$ is an estimate of $\sigma_{\mathbf{z}_i'}$ based on $n=216$ samples. Use of the dominant M eigenvectors, along with the samples of $\boldsymbol{\mathcal{Y}}$, in (25) yields the set of samples of $\boldsymbol{\mathcal{Z}}'$ which, in turn, yields the samples of $\boldsymbol{\mathcal{Z}}$ through (26). At this stage, a crosscheck is performed to ensure that enough information is propagated from $\boldsymbol{\mathcal{Y}}$ to $\boldsymbol{\mathcal{Z}}$ as the dimension is reduced from $N=121$ to $M=78$. The samples of $\boldsymbol{\mathcal{Y}}$ are reconstructed back from the samples, $\{\mathcal{Z}_k\}_{k=1}^n$, by using (27), i.e., $\mathcal{Y}_k^{(M)} = \widehat{\mathcal{V}}_M\left[\hat{\underline{a}} + \left\{(\hat{\underline{b}}-\hat{\underline{a}})\circ\frac{1}{2}(\mathcal{Z}_k+\mathbf{1}_M)\right\}\right] + \overline{\mathcal{Y}}$. The presence of hat symbols is a reminder that the quantities are estimated based on $n=216$ samples. The MSE of the relevant statistics computed from $\{\mathcal{Y}_k^{(M)}\}_{k=1}^n$ are compared with the corresponding statistics computed from the original experimental samples, $\{\mathcal{Y}_k\}_{k=1}^n$. The results are shown in Table 1.

Remark 1. Let us denote the ith eigenvalue of the true (unknown) continuous covariance function of the spatio-temporal stochastic process by ${}^{\mathrm{true}}\sigma^2_{\mathbf{z}_i'}, i\geqslant 1$. Then, an estimate of ${}^{\mathrm{true}}\sigma^2_{\mathbf{z}_i'}$ is given by $\hat{\sigma}^2_{\mathbf{z}_i'}$. The estimate $\hat{\sigma}_{\mathbf{z}_i'}$ depends on the number of experimental samples (i.e., n), the number of discrete spatio-temporal locations of the experimental measurements (i.e., N), and the interpolation functions (e.g., finite element type shape functions) that one employs to infer values of the experimental samples of the stochastic process at any other locations where experimental measurements are not available. It can be shown that the MSE associated with the KL approximation is given by [48, Theorem 2.10],

$$\sum_{i=1}^{M}({}^{\mathrm{true}}\sigma^2_{\mathbf{z}_i'} - \hat{\sigma}^2_{\mathbf{z}_i'}) + \sum_{i=M+1}^{\infty} {}^{\mathrm{true}}\sigma^2_{\mathbf{z}_i'}.$$

As already indicated earlier, M is typically less than N. The first error term can be reduced by increasing n, N and better interpolation operator (see, for example, [48, Proposiiton 3.3]). If the true continuous covariance function is compact and piecewise analytic, then it can be shown that the eigenvalues ${}^{\mathrm{true}}\sigma^2_{\mathbf{z}_i'},\ i\geqslant 1$, exponentially converge to 0 [48, Proposition 2.18]. The compact and piecewise analytic condition on the true (unknown) continuous covariance function can be assumed to hold for many practical applications including our work. Under more general conditions, other slow convergence, e.g., algebraic convergence, criteria might be suitable [48]. Such convergence criteria imply that the second error term can be made negligibly small for a sufficiently large M. For that M, if the first error term is also small enough, then we have achieved a good KL approximation of the stochastic process. However, it cannot be verified in practice since the first error term depends on the eigenvalues of the true continuous covariance function which will never be known in reality regardless of the approach used for constructing a probabilistic model (a PC based or non-PC based model) from experimental measurements.

Approach 1 as proposed in Section 3.1 is now employed to construct the PC representation of $\boldsymbol{\mathcal{Z}}$ based on $n=216$ experimental samples of $\boldsymbol{\mathcal{Z}}$. While the non-stationarity of the stochastic process is captured in the KL representation by the eigenfunctions of the covariance function of the stochastic process, the departure of the probability measure of the KL random variables from the chosen probability measure $P_{\underline{\xi}}$ will now be captured by the PC representation.

4.4.2. Construction of PC representation of $\boldsymbol{\mathcal{Z}}$ via Approach 1

In order to gain computational advantage, it is assumed here that 78 random variable components of $\boldsymbol{\mathcal{Z}}$ are pairwise statistically independent; particularly, the mjpdf of $\boldsymbol{\mathcal{Z}}$ has the following form,

Table 1
Comparison of statistics between experimental samples of the KL approximation $\boldsymbol{\mathcal{Y}}^{(M)}$ and experimental samples of $\boldsymbol{\mathcal{Y}}$; relative MSE as shown below is computed by using $\mathrm{relMSE}(\mathbf{S}^{(M)},\mathbf{S}) = 100(\|\mathbf{S}^{(M)}-\mathbf{S}\|_F^2)/\|\mathbf{S}\|_F^2$, in which $\mathbf{S}$ represents sample statistic of experimental samples $\{\mathcal{Y}_k\}_{k=1}^n$, and $\mathbf{S}^{(M)}$ represents the corresponding sample statistic of samples $\{\mathcal{Y}_k^{(M)}\}_{k=1}^n$.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
0	0	0.1220

$$p_{\mathcal{Z}}(\mathcal{Z}) = p_{\mathbf{z}_1\mathbf{z}_2}(z_1,z_2)p_{\mathbf{z}_3\mathbf{z}_4}(z_3,z_4)\cdots p_{\mathbf{z}_{77}\mathbf{z}_{78}}(z_{77},z_{78}). \tag{28}$$

In the present work, this form is found to be capable of accurately capturing the practically relevant and important information, as demonstrated later at the end of this section while discussing the results. Note that the random variable components, $\mathbf{z}_1,\ldots,\mathbf{z}_{78}$, are ordered in the descending order of values of the associated eigenvalues, $\hat{\sigma}^2_{\mathbf{z}'_1} > \cdots > \hat{\sigma}^2_{\mathbf{z}'_{78}}$, obtained in Section 4.4.1 (this, however, does not imply that $\{\hat{\sigma}_{\mathbf{z}_i}\}_{i=1}^{78}$ is also similarly ordered).

Let us generically indicate any of the pairs in (28) by $(\mathbf{z}_l,\mathbf{z}_u),\ l\in\mathcal{L}\equiv\{1,3,\ldots,77\}$ and $u\in\mathcal{U}\equiv\{2,4,\ldots,78\}$. For any given $l\in\mathcal{L}$ and $u\in\mathcal{U}$, the target bivariate pdf, $p_{\mathbf{z}_l\mathbf{z}_u}$, is estimated by using a normalized histogram of the corresponding experimental samples appropriately collected from $\{\mathcal{Z}_k\}_{k=1}^n$. Each bivariate histogram is constructed with 12×12 bins on equally spaced grids on the support, $s_{\mathbf{z}_l\mathbf{z}_u}\equiv[-1,1]^2$, of $(\mathbf{z}_l,\mathbf{z}_u)$.

By using the set of pdfs, $\{p_{\mathbf{z}_l\mathbf{z}_u}\}_{l\in\mathcal{L},u\in\mathcal{U}}$, PC representations (similar to (14) and (15)) of all the pairs are constructed. The set of PC random variables, $\{\xi_i\}_{i=1}^{78}$, is assumed here to be a set of statistically independent uniform random variables, all of which are supported on $[-1,1]$. For such ξ_i's, the orthogonal polynomials are Legendre polynomials given by,

$$\left.\begin{aligned}&\Psi_0(\xi_i)=1,\quad \Psi_1(\xi_i)=\xi_i,\\ &\Psi_j(\xi_i)=\tfrac{1}{j}(2j-1)\xi_i\Psi_{j-1}(\xi_i)-\tfrac{1}{j}(j-1)\Psi_{j-2}(\xi_i),\quad \text{if } j\geqslant 2,\end{aligned}\right\} \tag{29}$$

and the variance of $\Psi_j(\xi_i)$ is given by,

$$E\left[\Psi_j^2(\xi_i)\right]=\frac{1}{2j+1}. \tag{30}$$

While computing the PC coefficients (see (11) and (16)), the proxy function, $\tilde{q}$, for $f_{l|u}\equiv P^{-1}_{l|u}P_{\xi_l}$ or $f_u\equiv P_u^{-1}P_{\xi_u}$, as appropriate, is based on dividing the support, $s_{\mathbf{z}_l}\equiv[-1,1]$, of $\mathbf{z}_l$ (when approximating $f_{l|u}$) or $s_{\mathbf{z}_u}\equiv[-1,1]$ of $\mathbf{z}_u$ (when approximating f_u), into 199 equal intervals (see Appendix A). In determining the series expansion (similar to (12)) of $z_u\mapsto a_j(z_u)$, the basis functions ψ_k's are also selected as Legendre polynomials that are orthogonal w.r.t. the weight $w(z)=1/2$ on $[-1,1]$ implying that the denominator of (13) is given by $1/(2k+1)$. The set of PC coefficients of $f_{l|u}(\xi_l|z_u)$ is computed for 200 slices, that are equally spaced along the support $s_{\mathbf{z}_u}$ resulting in $\{z_u^{(k)},a_j(z_u^k)\}_{k=0}^{199}$ (see Figs. 5–7). The function, $z_u\mapsto a_j(z_u)$, based on this set is first formed by employing linear interpolation scheme and later used to compute the PC coefficients b_{jk}'s. The resulting PC representation of $(\mathbf{z}_l,\mathbf{z}_u)$ given by expressions similar to (14) and (15) is truncated at $K_1=K_2=K=19,\ \forall l\in\mathcal{L}$ and $\forall u\in\mathcal{U}$.

The constructed PC representation can be used to analytically derive bivariate pdf $p^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u}$ for all the pairs $(\mathbf{z}_l,\mathbf{z}_u),\ l\in\mathcal{L},\ u\in\mathcal{U}$. In the following, the set of bivariate pdfs $\{p^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u}\}_{l\in\mathcal{L},u\in\mathcal{U}}$ is estimated by using a set of $n_{\mathrm{PC}}=50{,}000$ samples to test the quality of the constructed PC representations. Use of $n_{\mathrm{PC}}=50{,}000$ samples of statistically independent uniform random variables $\{\xi_i\}_{i=1}^{78}$ in the constructed PC representations yields a set, $\{\mathcal{Z}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, of 50,000 samples of $\mathcal{Z}$. Let the estimate of the bivariate pdf of each pair, $(\mathbf{z}_l,\mathbf{z}_u)$, based on 50,000 PC samples be denoted by $\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u}$. This bivariate pdf is simply determined by employing linear interpolation scheme on a normalized histogram of PC samples with 25×25 bins constructed on equally spaced grids on the support $s_{\mathbf{z}_l\mathbf{z}_u}$. Introduce the following relative MSE for pdf,

$$\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u},\hat{p}_{\mathbf{z}_l\mathbf{z}_u})=100\frac{\int_{[-1,1]^2}\left\{\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u}(z_l,z_u)-\hat{p}_{\mathbf{z}_l\mathbf{z}_u}(z_l,z_u)\right\}^2dz_l\,dz_u}{\int_{[-1,1]^2}\hat{p}^2_{\mathbf{z}_l\mathbf{z}_u}(z_l,z_u)dz_l\,dz_u},$$

in which $\hat{p}_{\mathbf{z}_l\mathbf{z}_u}$ represents an estimate of $p_{\mathbf{z}_l\mathbf{z}_u}$ based on $n=216$ samples.

It is found that $\max_{l\in\mathcal{L},u\in\mathcal{U}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u},\hat{p}_{\mathbf{z}_l\mathbf{z}_u})]=2.4136\%$ and $\min_{l\in\mathcal{L},u\in\mathcal{U}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u},\hat{p}_{\mathbf{z}_l\mathbf{z}_u})]=0.1217\%$. Bivariate pdf based on 216 experimental samples and 50,000 PC realizations are plotted in Fig. 14 corresponding to $\max_{l\in\mathcal{L},u\in\mathcal{U}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u},\hat{p}_{\mathbf{z}_l\mathbf{z}_u})]=2.4136\%$. The associated contour plots are shown in Fig. 15.

In Table 2, a few practically significant statistics of experimental samples, $\{\mathcal{Z}_k\}_{k=1}^n$, and PC samples, $\{\mathcal{Z}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, are compared.

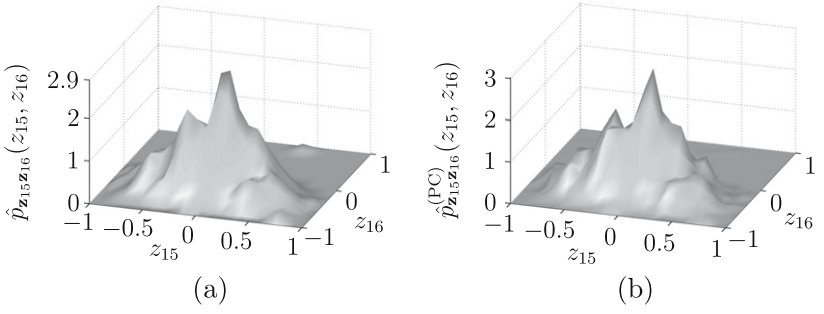


Fig. 14. Estimate of bivariate pdf of $(\mathbf{z}_l,\mathbf{z}_u)$ corresponding to $\max_{l\in\mathcal{L},u\in\mathcal{U}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u},\hat{p}_{\mathbf{z}_l\mathbf{z}_u})]=2.4136\%$: (a) $\hat{p}_{\mathbf{z}_l\mathbf{z}_u}$ and (b) $\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u}$.

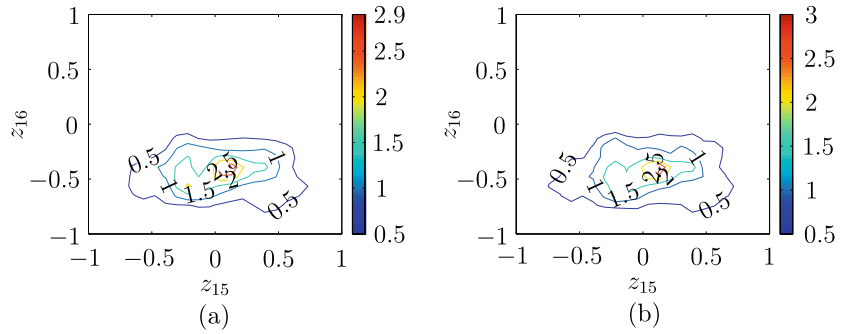


Fig. 15. Contour plots associated with the bivariate pdfs shown in Fig. 14: (a) $\hat{p}_{\mathbf{z}_l\mathbf{z}_u}$ and (b) $\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_l\mathbf{z}_u}$.

Table 2
Comparison of statistics between experimental samples and PC samples of the normalized KL random vector $\mathcal{Z}$; relative MSE as shown below is computed by using $\mathrm{relMSE}(\mathbf{S}^{(\mathrm{PC})}, \mathbf{S}) = 100(\|\mathbf{S}^{(\mathrm{PC})} - \mathbf{S}\|_F^2)/\|\mathbf{S}\|_F^2$, in which $\mathbf{S}$ represents the appropriate sample statistic of experimental samples $\{\mathcal{Z}_k\}_{k=1}^{n}$, and $\mathbf{S}^{(\mathrm{PC})}$ represents the corresponding sample statistic of PC realizations $\{\mathcal{Z}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
0.1103	0.1581	6.5934

Table 3
Comparison of statistics between experimental samples and PC samples of $\mathcal{Y}$; relative MSE is similarly computed as explained in the caption of Table 2.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
2.8525	0.0124	6.3700

While random variable components, $\{\mathbf{z}_i\}_{i=1}^{78}$, of normalized KL vector, $\mathcal{Z}$, are uncorrelated by construction resulting in zero off-diagonal elements of the covariance matrix of $\mathcal{Z}$, the SRCC matrix of $\mathcal{Z}$ would be fully populated (since $\{\mathbf{z}_i\}_{i=1}^{78}$ are statistically dependent in the present work). The covariance matrix and SRCC matrix estimated from the experimental samples, $\{\mathcal{Z}_k\}_{k=1}^{n}$, that contain the information about this statistical dependency indeed display the respective characteristics. The PC samples $\{\mathcal{Z}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, however, cannot capture the effect of this statistical dependency among the pairs $\{\mathbf{z}_l, \mathbf{z}_u\}_{l\in\mathcal{L},u\in\mathcal{U}}$ because of the assumption of pairwise statistical independence in (28). The effect of this assumption can be assessed, in some sense, by the deviation of the SRCC matrix based on PC samples from the SRCC matrix based on experimental samples. The value of relative MSE for SRCC matrix as shown in the third column of Table 2 implies that the assumption of pairwise statistical independence might be practically acceptable.

Finally, the set, $\{\mathcal{Z}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, are used to obtain the set, $\{\mathcal{Y}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, of PC samples of $\mathcal{Y}$ by taking recourse to (27). The statistics of the resulting samples, $\{\mathcal{Y}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, are compared to that of the experimental samples, $\{\mathcal{Y}_k\}_{k=1}^{n}$, and the results are shown in Table 3.

Remark 2. Computations of $a_j(z_u)$'s and b_{jk}'s involve bivariate histogram estimation and several numerical integrations. For any given $l \in \mathcal{L}$ and $u \in \mathcal{U}$, there are 200 slices representing 200 functions $f_{l|u}(\cdot|z_u)$'s for 200 discrete points along z_u. For each such slice, there are 20 PC coefficients $\{a_j(z_u)\}_{j=0}^{19}$ implying that we first need to compute 200×20 such PC coefficients $a_j(z_u)$'s. Subsequently, for any given $j \in \{0, 1, \ldots, 19\}$, use of 200 PC coefficients $a_j(z_u)$'s available at 200 discrete points along z_u results in 20 PC coefficients $\{b_{jk}\}_{k=0}^{19}$. For $\forall j$, it thus leads to 20×20 PC coefficients b_{jk}'s. Computations of 20×20 PC coefficients b_{jk}'s via 200×20 PC coefficients $a_j(z_u)$'s are the most expensive computational overhead in the proposed algorithm because of several numerical integrations involved in the computations. We need to repeat these computations for all the 39 pairs of $(\mathbf{z}_l, \mathbf{z}_u)$'s. These computations are executed solely in a single processor machine (Intel Xeon 3.2 GHz 2 GB Memory). It requires less than 72 h to compute these $39 \times 20 \times 20$ PC coefficients b_{jk}'s via $39 \times 200 \times 20$ PC coefficients $a_j(z_u)$'s thus highlighting the affordability of the computational cost. The rest of the computational burden required for Algorithm 1 is almost negligible compared to the above computational overhead. In this context, it is worth mentioning that the maximum likelihood based approach in the context of PC coefficients estimation [28] is theoretically more appealing since the dimension of the resulting PC representation of $\mathcal{Y}$ need not be same as the number, M, of dominant KL random variable components (in the present case, $M = 78$). Nevertheless, this maximum likelihood based approach, in its current state, is computationally prohibitive. In the authors' experience and opinion, it is believed that it would have taken several

months to estimate the PC representation of $\mathcal{Y}$ in a similar single processor machine. Finally, we note that several modules of the numerical tasks for both the approaches can be readily parallelized to substantially reduce the computational time.

Remark 3. It must be noted that computations of b_{jk}'s and c_j's are required *only once* as a set-up process. Once we have computed these PC coefficients, generation of PC samples, $\{\mathcal{Y}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, of $\mathcal{Y}$ is almost immediate. For example, generation of 50,000 PC realizations of each pair of $(\mathbf{z}_l, \mathbf{z}_u)$'s requires only about 1.1 second on average, and consequently, 50,000 PC realizations of $\mathcal{Y}$ (finite-dimensional representation of the stochastic process) are generated in about 45 s on average. As indicated earlier, most of the other simulation techniques involve either advanced MCMC simulation techniques or a suitable root-finding technique (to determine the corrcoef function of the underlying Gaussian process, if it exists) and numerical computation of the inverse mappings of the target marPDFs (for Nataf transformation) for generation of *each and every* realization of the stochastic process. Thus, these other techniques require substantial amounts of simulation time for generation of thousands of realizations (typically at the scale of hours). This essentially implies that the higher the number of PC realizations, the more computational savings achieved by employing the PC based simulation techniques (even after including the initial set-up time required for estimating the PC coefficients as highlighted in Remark 2).

In the next section, modeling of $\mathcal{Y}$ via Approach 2 is considered.

4.5. Modeling of $\mathcal{Y}$ via Approach 2

The experimental samples, $\{\mathcal{Y}_k\}_{k=1}^{n}$, of $\mathcal{Y}$ as obtained in Section 4.3 are used here again to deduce PC representation of $\mathcal{Y}$ by employing Approach 2. In this case, application of KL decomposition in order to obtain a reduced order representation of $\mathcal{Y}$ is not plausible since statistical dependency here would be characterized by SRCC not by PCC. However, for the sake of improved efficiency and regularity of the following numerical task, the samples of $\mathcal{Y}$ is scaled to obtain a set of samples of another N-dimensional random vector $\mathcal{Z} = [\mathbf{z}_1, \ldots, \mathbf{z}_N]^T$ supported on $[-1, 1]^N$ by employing a transformation similar to (26). In this case, $\mathcal{Y}$ is related to $\mathcal{Z}$ by,

$$\mathcal{Y} = \underline{a} + \left[(\underline{b} - \underline{a}) \circ \frac{1}{2}(\mathcal{Z} + \mathbf{1}_N) \right], \tag{31}$$

and the experimental samples, $\{\mathcal{Z}_k\}_{k=1}^{n}$, of $\mathcal{Z}$ follow from,

$$\mathcal{Z}_k = 2\left[(\mathcal{Y}_k - \underline{a}) \circ \left(\frac{1}{\underline{b} - \underline{a}} \right) \right] - \mathbf{1}_N, \quad k = 1, \ldots, n. \tag{32}$$

In (31) and (32), $\underline{a}$ and $\underline{b}$ are now, respectively, given by $\underline{a} = [\alpha_1, \ldots, \alpha_N]^T$ and $\underline{b} = [\beta_1, \ldots, \beta_N]^T$ with α_i and β_i, respectively, estimated as $\min(y_i^{(1)}, \ldots, y_i^{(n)})$ and $\max(y_i^{(1)}, \ldots, y_i^{(n)})$. Here, $y_i^{(k)}$ is the ith component, $i = 1, \ldots, N$, of the kth sample, $\mathcal{Y}_k = [y_1^{(k)}, \ldots, y_N^{(k)}]$.

The normalized marginal histogram of each random variable component, $\mathbf{z}_i$, $i \in \mathcal{I} = \{1, 2, \ldots, 121\}$ (recall $N = 121$), is constructed based on corresponding $n = 216$ experimental samples appropriately collected from $\{\mathcal{Z}_k\}_{k=1}^{n}$. Marginal histogram is based on 12 equal-sized bins on the support, $s_{\mathbf{z}_i} \equiv [-1, 1]$, of $\mathbf{z}_i$. Similar to previous approach, subsequent use of 1-dimensional linear interpolation scheme on this normalized histogram results in an estimate of the target marPDF of $\mathbf{z}_i$.

Based on the target marPDF $P_{\mathbf{z}_i}$, PC representation of each $\mathbf{z}_i$ (see (18)) is determined. In constructing these PC representations, orthogonal polynomials are again chosen as Legendre polynomials, given by (29), in terms of a set of uniform random variables, $\{\xi_i\}_{i=1}^{121}$, each of which is supported on $[-1, 1]$. In computing the corresponding PC coefficients, the approximate function, $\tilde{q}_i$, to be used in lieu of $q_i \equiv P_{\mathbf{z}_i}^{-1} P_{\xi_i}$ in (19) is based on dividing $s_{\mathbf{z}_i}$ into 199 equal intervals (see Appendix). The resulting PC representation, $\mathbf{z}_i \overset{d}{=} \lim_{K_i \to \infty} \sum_{j=0}^{K_i} c_{ji} \boldsymbol{\Psi}_j(\xi_i)$, is truncated at $K_i = 14, \forall\, i \in \mathcal{I}$.

In order to digitally generate realizations of $\mathcal{Z}$ (and $\mathcal{Y}$), a set of $n_{\mathrm{PC}} = 50{,}000$ samples of random vector, $\underline{\xi} = [\xi_1, \ldots, \xi_{121}]^T$, is simulated first as follows. Unlike Approach 1, the random variables, $\xi_1, \ldots, \xi_{121}$, here are statistically dependent. The statistical dependency of $\underline{\xi}$ is characterized by the SRCC matrix estimated based on experimental samples, $\{\mathcal{Z}_k\}_{k=1}^{216}$, of $\mathcal{Z}$. Application of the mapping defined by (20) on the resulting sample SRCC matrix of $\mathcal{Z}$, however, yields an estimate of $[\rho^{(1)}]$ which turns out to be a non-positive-definite matrix, thus rendering the normal copula technique inapplicable. Samples of the associated Gaussian random vector, $\mathbf{X}$, consisting of correlated standard normal random variables, $\mathbf{x}_1, \ldots, \mathbf{x}_{121}$, therefore, need to be generated by using the augmented normal copula technique as highlighted in Section 3.2.3. The constrained optimization problem defined by (23) is solved to determine an estimate of the feasible positive-definite covariance (or correlation) matrix $[\rho]$ of $\mathbf{X}$. It is found that $\mathrm{relMSE}([\hat{\rho}], [\hat{\rho}^{(1)}]) = 0.0006\%$ and $\mathrm{relMSE}([\hat{\rho}_s], [\hat{\rho}_s^{(1)}]) = 0.0006\%$, in which $[\hat{\rho}_s^{(1)}]$ is the sample (positive-definite) SRCC matrix based on $\{\mathcal{Z}_k\}_{k=1}^{216}$ and $[\hat{\rho}_s]$ is (again) a positive-definite matrix resulting from the element-wise application of the inverse mapping of (20) on $[\hat{\rho}]$, i.e., $(\hat{\rho}_s)_{ij} = (6/\pi)\arcsin(\hat{\rho}_{ij}/2)$. The hat symbols are used to indicate that the quantities are estimated based on $n = 216$ samples. Then, 50,000 samples of $\underline{\xi}$ consisting of statistically dependent uniform random variables, $\{\xi_i\}_{i=1}^{121}$, supported on $[-1, 1]^{121}$, with an estimate of its SRCC or PCC matrix being given by $[\hat{\rho}_s]$, can be readily generated by using the augmented normal copula technique. Use of these samples in the constructed PC representations for $\{\mathbf{z}_i\}_{i \in \mathcal{I}}$ yields a set, $\{\mathcal{Z}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, of 50,000 samples of $\mathcal{Z}$, and subsequently, the set, $\{\mathcal{Y}_k^{(\mathrm{PC})}\}_{k=1}^{n_{\mathrm{PC}}}$, of samples of $\mathcal{Y}$ follows from (31).

Let the estimate of marpdf of $\mathbf{z}_i, i \in \mathcal{I}$, be denoted by $\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}$ that is again determined from the corresponding marginal normalized linearly interpolated histogram. The histogram is based on 25 equal-sized bins on the corresponding support, $s_{\mathbf{z}_i}$. A comparison between two marpdfs based on 50,000 PC realizations and 216 experimental samples is shown in Fig. 16 for $\mathbf{z}_i$ corresponding to $\max_{i\in\mathcal{I}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}, \hat{p}_{\mathbf{z}_i})] = 2.1833\%$, in which $\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}, \hat{p}_{\mathbf{z}_i})$ is now defined by,

$$\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}, \hat{p}_{\mathbf{z}_i}) = 100\frac{\int_{s_{\mathbf{z}_i}} \left\{\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}(z_i) - \hat{p}_{\mathbf{z}_i}(z_i)\right\}^2 dz_i}{\int_{s_{\mathbf{z}_i}} \hat{p}^2_{\mathbf{z}_i}(z_i) dz_i},$$

with $\hat{p}_{\mathbf{z}_i}$ being an estimate of the marpdf based on 216 experimental samples of $\mathbf{z}_i$. Let us also report the minimum value of relMSE_p, $\min_{i\in\mathcal{I}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}, \hat{p}_{\mathbf{z}_i})] = 0.0729\%$.

Finally, summaries of practically significant statistics based on PC realizations are compared with that based on experimental samples for $\mathcal{Z}$ and $\mathcal{Y}$, respectively, in Tables 4 and 5. It must remarked here that the covariance matrix is not used as a measure of statistical dependency in Approach 2, the corresponding results are still shown in the second columns of these tables for the sake of completeness.

Remark 4. We remark here that the computational efforts required for estimating the PC coefficients via Approach 2 is less than 1/18 of that of the first approach (i.e., less than 4 h) since it involves only a set of marginal pdfs and SRCC unlike Approach 1 that involves a set of bivariate pdfs. The computation is carried out in the same single processor machine as indicated earlier in Remark 2. This computational cost also includes the cost required to solve the constrained optimization problem defined by (23). It requires less than 3 h to carry out this augmented normal copula technique associated with the resulting 121×121 SRCC matrix. In Approach 2, most of the computational time may likely be devoted in determining an estimate $[\hat{\rho}]$ of the feasible positive-definite covariance (or correlation) matrix of the standard Gaussian random vector $\mathbf{X}$ from the associated sample SRCC matrix. Simulation time for generating samples of $\mathcal{Y}$ from the constructed PC representation is, however, comparable with that of Approach 1.

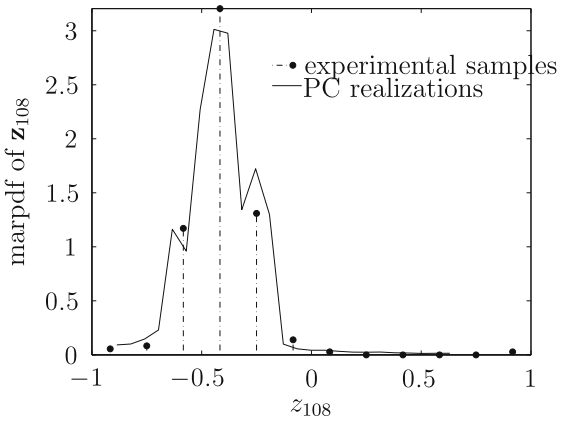


Fig. 16. Estimates of marginal pdf of $\mathbf{z}_i$ corresponding to $\max_{i\in\mathcal{I}}[\mathrm{relMSE}_p(\hat{p}^{(\mathrm{PC})}_{\mathbf{z}_i}, \hat{p}_{\mathbf{z}_i})] = 2.1833\%$.

Table 4
Comparison of statistics between experimental samples and PC samples of the N-dimensional normalized random vector $\mathcal{Z}$; relative MSE is similarly computed as explained in the caption of Table 2.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
0.0339	5.4139	0.0040

Table 5
Comparison of statistics between experimental samples and PC samples of $\mathcal{Y}$; relative MSE is similarly computed as explained in the caption of Table 2.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
2.5569	1.3123	0.0040

4.6. Reconstructing the original random temperature field

Construct the PC representation of $\mathcal{Z}$ either by using Approach 1 or Approach 2 as appropriate. The PC coefficients of the random variable components of $\mathcal{Z}$ and those of $\mathcal{Y}$ are related by linear mappings as can be readily verified by using (27) and (31) (see also the authors' other works [30] for further details). Since the set of 11×11 random variables, $\{\mathbf{\Gamma}^{(n)}(t,h)\}_{(t,h)\in(T\times D)}$, constitute $\mathcal{Y}$, PC coefficients of $\mathbf{\Gamma}(t,h),\ (t,h) \in (T \times D)$, immediately follow from the PC coefficients of $\mathcal{Y}$ by using the relation, $\mathbf{\Gamma}(t,h) = \overline{\Gamma}(t,h)\mathbf{\Gamma}^{(n)}(t,h) + \overline{\Gamma}(t,h)$. Inference of PC coefficients of the original random process, when $(t,h) \notin (T \times D)$, from those of $\mathbf{\Gamma}(t,h), (t,h) \in (T \times D)$, is essentially a task of interpolation or/and approximation technique as shown in numerous other occasions in the present work. Digital generation of realizations of the original random process similarly needs no further explanation.

5. Conclusion

Two approaches for constructing PC representations from experimental measurements are presented. These representations yield, in general, non-Gaussian models for a second-order random vector $\mathcal{Y}$, that has been experimentally observed. The random vector $\mathcal{Y}$ can be viewed as a finite-dimensional representation of a non-stationary, non-Gaussian, second-order random field evolving over space or space–time. The experimental data is measured on a finite countable subset of the indexing set. In many practical applications, e.g., prediction of acoustic field statistics involving oceanographic parameters as indicated in the previous section, use of a spatio-temporal random field to describe variability in model parameters, yields a more appropriate representation of reality. The PC representation of the random field representing such random system parameters has been proven to be an efficient tool in systematically propagating the uncertainty to the model-based response predictions of the stochastic system.

Approach 1 attempts to capture the complete information of a target mjpdf of $\mathcal{Y}$. This approach uses the knowledge of a complete set of properly ordered target conditional PDFs estimated from the experimental measurements and the Rosenblatt transformation. The set of target conditional pdfs, that uniquely defines the target mjpdf of $\mathcal{Y}$, are approximations, based on linear interpolation, of the corresponding set of normalized histograms of the appropriate set of experimental samples. Approach 2, on the other hand, satisfies the target marPDFs and the target SRCC matrix of $\mathcal{Y}$. The set of target marPDFs and the target SRCC matrix are similarly estimated by using the experimental samples. The second approach is also founded on the Rosenblatt transformation. In both approaches, appropriate functions based on the Rosenblatt transformation are first defined in terms of the selected PC variables, ξ_k's. The defined functions are equal to $\mathcal{Y}$ in the sense of distribution. Subsequently, construction of the PC expansion of these functions results in appropriate PC representations that can be readily employed within the PC framework to propagate the associated uncertainty. It should, however, be realized that the existence and uniqueness of these functions can rarely be established in practice due to ignorance of the true sources of uncertainty, let alone the manner in which they map into the system-level parameters. Nevertheless, the proposed approaches guarantee [10, Theorem 2.1] that such functions can always be constructed, using distributional equivalence. For efficient and fast computation of the PC coefficients, the Rosenblatt transformation based functions are further substituted by the appropriate interpolated functions.

One important distinction between the two proposed approaches is that while PC random variables, ξ_k's, are statistically independent in Approach 1, the corresponding set of PC random variables are statistically dependent in Approach 2. Moreover, Approach 1 will typically be more computationally demanding than Approach 2. Additional model reduction techniques, e.g., use of KL decomposition as discussed in the context of numerical illustration in Section 4.4.1, are recommended in order to manage this computational effort. Further probabilistic assumptions as made in Section 4.4.2, while modeling the spatio-temporal random temperature field, would also alleviate the computational burden albeit at the expense of accuracy. This practice is certainly recommended if the achieved accuracy is within acceptable bounds. The accuracy level of Approach 2 is expected to be higher than that of Approach 1 if additional probabilistic assumptions and model reduction schemes, as just indicated, are incorporated into it.

Acknowledgments

This work was supported by the Office of Naval Research, Air Force Office for Scientific Research, and the National Science Foundation.

Appendix A. Computation of PC coefficients

This appendix describes an 1-dimensional interpolation based scheme that would be useful for efficient computation of

1. $\{a_j(y_2)\}_{j\in\mathbb{N}}$ in (10), or
2. $\{c_j\}_{j\in\mathbb{N}}$ in (15), or
3. $\{c_{jk}\}_{j\in\mathbb{N}}$ for any given $k \in \{1,\ldots,N\}$ in (18).

Since all the cases above are similar, only the last case involving $\{c_{jk}\}_{j\in\mathbb{N}},\ k\in\{1,\ldots,N\}$, will be considered. Any other case can be readily tackled by replacing c_{jk}, $P_{\mathbf{y}_k}$ and P_{ξ_k} with the appropriate PC coefficients and PDFs.

Let the pdf and support of $\mathbf{y}_k$ be denoted, respectively, by $p_{\mathbf{y}_k}$ and $s_k=[l_k,m_k]\subset\mathbb{R}$. The computation of c_{jk} in (19) based on $q_k\equiv P_{\mathbf{y}_k}^{-1}P_{\xi_k}$ involves solving an integral equation. For some given ξ_k, the integral equation, $P_{\xi_k}(\xi_k)=\int_{l_k}^{y_k}p_{\mathbf{y}_k}(y)dy$, is to be solved for y_k. Solving this integral equation several times *within* the numerical integration algorithm, that is employed to compute c_{jk}, significantly increases the computational burden, and might also lead to certain numerical instability. To overcome these difficulties and to increase the computational expediency and efficiency, a surrogate function, $\tilde{q}_k$, determined based on 1-dimensional interpolation technique, is used for q_k in (19) to compute the PC coefficients, c_{jk}. The approximate function, $\tilde{q}_k$, needs to be determined only once $\forall j\in\mathbb{N}$.

Consider $u_k\equiv P_{\xi_k}(\xi_k)\overset{d}{=}P_{\mathbf{y}_k}(y_k)$ (u_k here should not be confused with the components of $\mathbf{U}$ in sub Sections 3.2.2 and 3.2.3). For a given $y_k\in s_{\mathbf{y}_k}=[l_k,m_k]$, finding u_k from $u_k=P_{\mathbf{y}_k}(y_k)$ is, in general, much cheaper than finding y_k from $y_k=P_{\mathbf{y}_k}^{-1}(u_k)$ for a given $u_k\in[0,1]$.

For each $k\in\{1,\ldots,N\}$, let the support, s_k, be divided equally into $n_k\in\mathbb{N}$ intervals. Then, coordinates of the points defining these intervals are given by $y_k^{(j)}=l_k+j[(m_k-l_k)/n_k],\ j=0,\ldots,n_k$. For each of these points, first compute $u_k^{(j)}=P_{\mathbf{y}_k}(y_k^{(j)})$, and then compute $\xi_k^{(j)}=P_{\xi_k}^{-1}(u_k^{(j)})$. Since P_{ξ_k}'s are suitably chosen standard measures associated with the commonly used PC random variables, computation of $P_{\xi_k}^{-1}$ via closed form expression or efficient algorithms is available in the statistical literature (see e.g., [40, Section 3.2]; [10, Section 2.1]). As already indicated, the statistical toolbox of MATLAB provides functions to evaluate the inverse of PDF for many such standard PC random variables. Since $P_{\mathbf{y}_k}$ and P_{ξ_k} are monotonically increasing function, the set of values in $\{\xi_k^{(j)}\}_{j=0}^{n_k}$ would be in the increasing order, $\xi_k^{(0)}<\cdots<\xi_k^{(n_k)}$. The set, $\{\xi_k^{(j)},y_k^{(j)}\}_{j=0}^{n_k}$, thus determined is now used to construct the approximate function, $s_{\xi_k}\ni\xi_k\mapsto\tilde{q}_k(\xi_k)\in s_{\mathbf{y}_k}$, by using standard interpolation technique [49]. The basic MATLAB package offers a function, `interp1`, use of which should be sufficient enough for deduction of $\tilde{q}_k$ for many practical purposes. The approximate function, $\tilde{q}_k$, is used as a proxy for q_k in (19) to compute the PC coefficients, c_{jk}'s.

The error in approximating $q_k(\xi_k)$ by the PC representation $\tilde{q}_k^{(K_k)}(\xi_k)=\sum_{j=0}^{K_k}c_{jk}\mathbf{\Psi}_j(\xi_k)$, for some large $K_k\in\mathbb{N}$, is bounded above by the following relation,

$$|q_k(\xi_k)-\tilde{q}_k^{(K_k)}(\xi_k)|\leqslant|q_k(\xi_k)-\tilde{q}_k(\xi_k)|+|\tilde{q}_k(\xi_k)-\tilde{q}_k^{(K_k)}(\xi_k)|\quad\text{a.s.}\tag{A.1}$$

The second error term is bounded above by some $e_{K_k}(\xi_k)$ satisfying $\lim_{K_k\to\infty}e_{K_k}(\xi_k)=0$ [36, Chapter 4] a.s. When a linear interpolation scheme is employed, the interpolated function, $\tilde{q}_k$, is piecewise linear in ξ_k and the first error term is then bounded above by $O(h_k^2)$, in which $h_k=\max_{1\leqslant i\leqslant n_k}(\xi_k^{(i)}-\xi_k^{(i-1)})$, [49, Example 1.1.4] a.s. In establishing this error bound, $O(h_k^2)$, it is necessary for the second derivative of q_k to be piecewise bounded by some finite $\mathcal{K}$, i.e., $|\partial^2q_k(\xi_k)/\partial\xi_k|\leqslant\mathcal{K}$ on s_{ξ_k} except possibly a finite number of points. As it is already mentioned that an assumption of piecewise smoothness is required to arrive at the PC representation in a.s. sense, the piecewise linear function, $\tilde{q}_k$, that is actually being represented by PC formalism, automatically satisfies the assumption of piecewise smoothness. Therefore, in order to bound both the error bounds in a.s. sense, the assumption of piecewise smoothness of the original function, q_k, needs to be replaced by a relatively stronger assumption of the piecewise boundedness of the second derivative of the function q_k.

Appendix B. Convergence, verification and validation analysis

The two approaches described in this paper can each be broken down into the following two essential steps. In the first step, a *target* mjPDF is constructed that purports to describe available information. The second step involves developing a PC representation of a set of random variables characterized by a known mjPDF. The PC representation is constructed such that the associated mjPDF is within specified tolerance to the known mjPDF of the random variables being described via PC representation. An analysis of the second step provides the verification task associated with our approaches. When the known mjPDF is taken as the target mjPDF constructed in the first step, then the suitability of the developed PC representation can be construed to define a "model validation" task associated with our approaches.

We will first consider the verification analysis along with the associated convergence issues followed by the validation process.

B.1. Convergence issues and verification analysis

Denote the PC based estimate of $P_{\mathcal{Y}}$ by $P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}$. The subscript $\mathbf{f}$ is used to indicate that the Rosenblatt transformation is employed in defining the unknown mapping $\mathbf{f}:\underline{\xi}\to\mathcal{Y}$ in the process of determining $P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}$. Several superscripts are used to show the dependence on them. Let us recall here that n_k refers to the number of slices as required in Section 3.1 (see, e.g., Fig. 4), and K refers to the number of terms retained in the constructed PC representation. It should be noted that K, in general, could be different along different random variable components of $\mathcal{Y}$ or its KL approximation as appropriate. While K is considered to be identical for all the components for the sake of simplified notation, the ensuing discussion can be readily

adapted to include the situation of different K's. We remind the readers that the superscript, (PC), as used in Sections 4.4.2 and 4.5 is a simplistic notation of the dependence as alluded to above.

The MSE of $P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}$, which measures the goodness of $P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}$ locally at $\mathcal{Y} \in S_{\mathcal{Y}}$, can be given by $|P_{\mathcal{Y}}(\mathcal{Y}) - P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}(\mathcal{Y})|^2$. For the verification purpose, the behavior of this MSE, as M, n_k, K and n_{PC} change, can be used to determine if the applications of the proposed approaches accurately yield consistent PC representations. The MSE is bounded by the following five error terms,

$$\begin{aligned}|P_{\mathcal{Y}}(\mathcal{Y}) - P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}(\mathcal{Y})|^2 \leqslant\ & |P_{\mathcal{Y}}(\mathcal{Y}) - P_{\mathcal{Y}}^{(M)}(\mathcal{Y})|^2 + |P_{\mathcal{Y}}^{(M)}(\mathcal{Y}) - P_{\mathbf{f}}^{(M)}(\mathcal{Y})|^2 + |P_{\mathbf{f}}^{(M)}(\mathcal{Y}) - P_{\mathbf{f}}^{(M,n_k)}(\mathcal{Y})|^2 + |P_{\mathbf{f}}^{(M,n_k)}(\mathcal{Y}) \\ & - P_{\mathbf{f}}^{(M,n_k,K)}(\mathcal{Y})|^2 + |P_{\mathbf{f}}^{(M,n_k,K)}(\mathcal{Y}) - P_{\mathbf{f}}^{(M,n_k,K,n_{\mathrm{PC}})}(\mathcal{Y})|^2, \end{aligned} \tag{B.1}$$

in which several superscripts represent the appropriate dependence.

The first error term in (B.1) is only relevant for Approach 1 since the KL approximation is only employed in Approach 1 to obtain a reduced order representation of $\mathcal{Y}$. For Approach 2, $P_{\mathcal{Y}}^{(M)}(\mathcal{Y}) \equiv P_{\mathcal{Y}}(\mathcal{Y})$ since the KL representation is not required, and, therefore, this error term is identically zero. For Approach 1, by the mean-square convergence criterion of the KL approximation as discussed in Section 4.4.1 and Cramér–Wold theorem [27, pp. 383], it can be concluded that $\lim_{M\to N} P_{\mathcal{Y}}^{(M)}(\mathcal{Y}) := \lim_{M\to N} P_{\mathcal{Y}^{(M)}}(\mathcal{Y}) = P_{\mathcal{Y}}(\mathcal{Y})$. Thus, the first error term can be made negligibly small by choosing an appropriate large value of M, which depends on the scales of stochastic fluctuations across the components of $\mathcal{Y}$.

Use of the Rosenblatt transformation guarantees that $\mathbf{f}(\underline{\xi})$ is equal in distribution to $\mathcal{Y}$. Equality in distribution, in turn, implies that the second error term in (B.1) is identically zero for Approach 1. Use of this equality in distribution along with the invariance under monotone transformation property of the SRCC, the second error term also turns out to be identically zero for Approach 2 since it is assumed that $P_{\mathcal{Y}}$ is completely characterized by the marPDFs and the SRCC matrix of $\mathcal{Y}$. To investigate this effect more rigorously, it should be noted that since $P_{\mathcal{Y}}^{(M)}(\mathcal{Y}) \equiv P_{\mathcal{Y}}(\mathcal{Y})$ is valid for Approach 2, the second error term is bounded above by $|P_{\mathcal{Y}}(\mathcal{Y}) - \widetilde{P}_{\mathcal{Y}}(\mathcal{Y})|^2 + |\widetilde{P}_{\mathcal{Y}}(\mathcal{Y}) - P_{\mathbf{f}}(\mathcal{Y})|^2$, where $\widetilde{P}_{\mathcal{Y}}$ is characterized by the marPDFs and SRCC matrix of $\mathcal{Y}$ but otherwise least committal to unavailable information. The assumption made in Approach 2 essentially implies that $\widetilde{P}_{\mathcal{Y}}(\mathcal{Y}) \equiv P_{\mathcal{Y}}(\mathcal{Y})$. Use of any numerical techniques in determining the correlation matrix of the Gaussian random vector $\mathbf{X}$ as required in Approach 2 is likely to introduce a small, but often practically acceptable, error given by $|\widetilde{P}_{\mathcal{Y}}(\mathcal{Y}) - P_{\mathbf{f}}(\mathcal{Y})|^2$. Finally, we note for Approach 1 that if any assumptions are made in the numerical procedure for constructing the PC representations, e.g., as introduced by the assumption of pairwise statistical independence in Section 4.4.2, then it implies that the second error term is bounded above by $|P_{\mathcal{Y}}^{(M)}(\mathcal{Y}) - \widetilde{P}_{\mathcal{Y}}^{(M)}(\mathcal{Y})|^2 + |\widetilde{P}_{\mathcal{Y}}^{(M)}(\mathcal{Y}) - P_{\mathbf{f}}^{(M)}(\mathcal{Y})|^2$. Here, the associated mjpdf of $\widetilde{P}_{\mathcal{Y}}^{(M)} := \widetilde{P}_{\mathcal{Y}^{(M)}} \equiv P_{\mathcal{Z}}$, e.g., in the context of the discussion in Section 4.4.2, is now explicitly given by the rhs of (28).

Let us consider the third error term now. It is only relevant for Approach 1. At any arbitrary point $\mathcal{Y} \in S_{\mathcal{Y}}$ through which no slice passes, $P_{\mathbf{f}}^{(M,n_k)}$ can be simply defined by employing any standard interpolation scheme. It is straightforward to see that the third error term $|P_{\mathbf{f}}^{(M)}(\mathcal{Y}) - P_{\mathbf{f}}^{(M,n_k)}(\mathcal{Y})|^2$ can be made negligibly small by choosing a large number n_k of slices [49]. Since the concept of slice is irrelevant for Approach 2 and $P_{\mathcal{Y}}$ is assumed to be completely characterized by the marPDFs and the SRCC matrix of $\mathcal{Y}$, this error term becomes identically zero for Approach 2 implying that $P_{\mathbf{f}}^{(M,n_k)}(\mathcal{Y}) \equiv P_{\mathbf{f}}^{(M)}(\mathcal{Y})$.

Before investigating the fourth error term, let us recall that the set of target conditional PDFs, $\{P_{i|(i+1):N}\}_{i=1}^{N}$, as required for Approach 1, uniquely defines $P_{\mathcal{Y}} \equiv P_{\mathbf{f}}$. Use of this condition along with the mean-square convergence criterion of the PC representation of $f_{i|(i+1):N} \equiv P_{i|(i+1):N}^{-1} P_{\xi_i}$ and the validity of the orthogonal series expansion of the interpolated function of each PC coefficient (associated with the PC representation of $f_{i|(i+1):N}$) at every continuity point of this interpolated function guarantees that $\lim_{K\to\infty} P_{\mathbf{f}}^{(M,n_k,K)}(\mathcal{Y}) = P_{\mathbf{f}}^{(M,n_k)}(\mathcal{Y})$ for Approach 1. For Approach 2, the mean-square convergence criterion of the PC representation of $q_i \equiv P_{\mathbf{y}_i}^{-1} P_{\xi_i}$ along with the invariance under monotone transformation property of the SRCC is relevant to conclude that $\lim_{K\to\infty} P_{\mathbf{f}}^{(M,K)}(\mathcal{Y}) = P_{\mathbf{f}}^{(M)}(\mathcal{Y})$. The accuracy of the resulting PC representations and orthogonal series expansions can be estimated by using the existing results on convergence analysis in the literature [19]. As already explained in Section 3, the above mean-square convergence can also be interpreted in a.s. sense for both approaches. Clearly, choosing a large value of K guarantees that the fourth error term is negligibly small for both approaches.

Finally, it is well known that the last error term in (B.1) can also be made negligibly small by choosing a large PC sample size n_{PC} [6,40,10].

Tables 6 and 7 show results similar to those in Tables 2 and 3 with $n_k = 25$ slices and $K = 14$ instead of $n_k = 200$ slices and $K = 19$, respectively. These results indicate, as expected, a decrease in relative MSE for various statistics with increase in n_k and K.

For Approach 2, while higher level of errors are observed for several statistics computed with lower values of K, say, $K = 12$, and other different parameters, these errors are not appreciably different from those shown in Tables 4 and 5.

Table 6
Comparison of statistics between experimental samples and PC samples of the normalized KL random vector $\mathcal{Z}$; relative MSE is similarly computed as explained in the caption of Table 2; $\max_{l\in\mathcal{L},u\in\mathcal{U}}[\mathrm{relMSE}_p(\hat{p}_{\mathbf{z}_l\mathbf{z}_u}^{(\mathrm{PC})}, \hat{p}_{\mathbf{z}_l\mathbf{z}_u})] = 24.9520\%$ and $\min_{l\in\mathcal{L},u\in\mathcal{U}}[\mathrm{relMSE}_p(\hat{p}_{\mathbf{z}_l\mathbf{z}_u}^{(\mathrm{PC})}, \hat{p}_{\mathbf{z}_l\mathbf{z}_u})] = 2.6312\%$.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
6.4140	1.9194	6.7806

Table 7
Comparison of statistics between experimental samples and PC samples of $\mathcal{Y}$; relative MSE is similarly computed as explained in the caption of Table 2.

Relative MSE in percentage (%) for		
Mean vector	Covariance matrix	SRCC matrix
>25	1.5434	6.7728

The convergence analysis as explained above provides a practically appealing and essential verification tool guaranteeing that the implementation of the proposed approaches results in PC representations consistent with the conceptual description of the target mjPDF $P_{\mathcal{Y}}$. By the use of Kolmogorov's existence theorem [27, Section 36], it can be concluded then that the constructed second-order, non-stationary and non-Gaussian stochastic process characterized by $P_{\mathcal{Y}}$ and the underlying original stochastic process are equivalent in the sense that they have identical N-dimensional mjPDF.

B.2. Validation formalities

For Approach 1, several statistics as shown in Tables 1–3 and the maximum and minimum values of $\text{relMSE}_p(\cdot,\cdot)$ as indicated in Section 4.4.2 are taken as the statistical features of interest. For Approach 2, similar statistical features are indicated in Tables 4, 5 and through a few other metrics as defined by $\text{relMSE}([\rho],[\rho^{(1)}])$ and $\text{relMSE}([\rho_s],[\rho_s^{(1)}])$ along with the maximum and minimum values of $\text{relMSE}_p(\cdot,\cdot)$ as discussed in Section 4.5.

The results presented in Section 4.4 for Approach 1 and in Section 4.5 for Approach 2 validate the respective PC representation. This validation must be judged within the degree of accuracy measured w.r.t. the relative MSE as reported for the corresponding statistical features of interest. Also, the importance of this validation must solely be interpreted from the perspective of the objective of our work, i.e., constructing the PC representation directly from the *available* experimental measurements by capturing the statistical information of interest as intended.

As more data become available modifying the dimension of $\boldsymbol{\mathcal{Y}}$ (i.e., N) and the experimental sample size (i.e., n), the degree of accuracy will change, which will call for the steps to re-validate the constructed PC representation. The relative MSE of several statistics should be recomputed to check if they are within the desired degree of accuracy. If they are within the desired degree of accuracy, then there will be no need to reconstruct the PC representation. Otherwise, new PC representation must be constructed to obtain a validated probability model.

References

[1] J.R. Apel, B. Badiey, C.S. Chiu, S. Finette, J. Headrick, J.N. Kemp, J.F. Lynch, A. Newhall, M.H. Orr, B.H. Pasewark, D. Tielbuerger, A. Turgut, K. von der Heydt, S. Wolf, An overview of the 1995 SWARM shallow-water internal wave acoustic scattering experiment, IEEE Journal of Oceanic Engineering 22 (3) (1997) 465–500.
[2] M. Cario, B. Nelson, Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix, Tech. rep., Department of Industrial Engineering and 35 Management Sciences, Northwestern University, Evanston, Illinois, 1997. URL: <http://citeseer.ist.psu.edu/cario97modeling.html>.
[3] G. Deodatis, R.C. Micaletti, Simulation of highly skewed non-Gaussian stochastic processes, Journal of Engineering Mechanics – ASCE 127 (12) (2001) 1284–1295.
[4] M. Rosenblatt, Remarks on multivariate transformation, The Annals of Mathematical Statistics 23 (3) (1952) 470–472.
[5] A. Izenman, Recent developments in nonparametric density estimation, Journal of the American Statistical Association 86 (413) (1991) 205–224.
[6] D. Scott, Multivariate Density Estimation: Theory, Practice and Visualization, Wiley-Interscience, 1992.
[7] S. Ghosh, S. Henderson, Chessboard distributions and random vectors with specified marginals and covariance matrix, Operations Research 50 (5) (2002) 820–834.
[8] P. Embrechts, A. McNeil, D. Straumann, Correlation and dependence in risk management: properties and pitfalls, in: M. Dempster (Ed.), Risk Management: Value at Risk and Beyond, Cambridge University Press, 2001, pp. 176–223.
[9] E. Lehmann, Nonparametrics: Statistical Methods Based on Ranks, McGraw-Hill International Book Company, New York, 1975.
[10] W. Hörmann, J. Leydold, G. Derflinger, Automatic Nonuniform Random Variate Generation, Springer, 2004.
[11] S. Ghosh, S. Henderson, Behavior of the NORTA method for correlated random vector generation as the dimension increases, ACM Transactions on Modeling and Computer Simulation 13 (3) (2003) 276–294. also see http://www.informs-cs.org/wsc02papers/034.pdf.
[12] Y. Kouskoulas, L. Pierce, F. Ulaby, A computationally efficient multivariate maximum-entropy density estimation (MEDE) technique, IEEE Transactions on Geoscience and Remote Sensing 42 (2) (2004) 457–468.
[13] P. van der Geest, An algorithm to generate samples of multi-variate distributions with correlated marginals, Computational Statistics and Data Analysis 27 (3) (1998) 271–289.
[14] S. Sakamoto, R. Ghanem, Simulation of multi-dimensional non-Gaussian non-stationary random fields, Probabilistic Engineering Mechanics 17 (2) (2002) 167–176.
[15] R. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, New York, USA, 1991 (revised edition published by Dover in 2003).
[16] H. Lancaster, Some properties of the bivariate normal distribution considered in the form of a contingency table, Biometrika 44 (1–2) (1957) 289–292.
[17] R. Ghanem, Ingredients for a general purpose stochastic finite elements implementation, Computer Methods in Applied Mechanics and Engineering 168 (1–4) (1999) 19–34.
[18] D. Xiu, G. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM Journal on Scientific Computing 24 (2) (2002) 619–644.
[19] C. Soize, R. Ghanem, Physical systems with random uncertainties: Chaos representations with arbitrary probability measure, SIAM Journal on Scientific Computing 26 (2) (2004) 395–410.
[20] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, Journal of Computational Physics 209 (2) (2005) 617–642.

[21] M. Pellissetti, R. Ghanem, Iterative solution of systems of linear equations arising in the context of stochastic finite elements, Advances in Engineering Software 31 (2000) 607–616.
[22] M. Reagan, H. Najm, R. Ghanem, O. Knio, Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection, Combustion and Flame 132 (3) (2003) 545–555.
[23] D. Ghosh, R. Ghanem, J. Red-Horse, Analysis of eigenvalues and modal interaction of stochastic systems, AIAA Journal 43 (10) (2005) 2196–2201.
[24] J. Witteveen, S. Sarkar, H. Bijl, Modeling physical uncertainties in dynamic stall induced fluid structure interaction of turbine blades using arbitrary polynomial chaos, Computers and Structures 85 (11–14) (2007) 866–878.
[25] R. Ghanem, A. Doostan, On the construction and analysis of stochastic models: Characterization and propagation of the errors associated with limited data, Journal of Computational Physics 217 (1) (2006) 63–81.
[26] M.B. Priestley, Spectral Analysis and Time Series, vols. I and II, Academic Press, New York, 1981.
[27] P. Billingsley, Probability and Measure, John Wiley & Sons, New York, 1995.
[28] C. Desceliers, R. Ghanem, C. Soize, Maximum likelihood estimation of stochastic chaos representation from experimental data, International Journal for Numerical Methods in Engineering 66 (6) (2006) 978–1001.
[29] B. Lindsay, Composite likelihood methods, Contemporary Mathematics 80 (1988) 221–239.
[30] S. Das, R. Ghanem, J.C. Spall, Asymptotic sampling distribution for polynomial chaos representation from data: a maximum entropy and Fisher information approach, SIAM Journal on Scientific Computing 30(5) (2008) 2207–2234. doi:10.1137/060652105.
[31] R. Cameron, W. Martin, The orthogonal development of non-linear functionals in series of Fourier–Hermite functionals, The Annals of Mathematics, Second Series 48 (2) (1947) 385–392.
[32] N. Wiener, The homogeneous chaos, American Journal of Mathematics 60 (4) (1938) 897–936.
[33] O. Le Maître, H. Najm, R. Ghanem, O. Knio, Multi-resolution analysis of Wiener-type uncertainty propagation schemes, Journal of Computational Physics 197 (2) (2004) 502–531.
[34] S. Das, Model, identification and analysis of complex stochastic systems: applications in stochastic partial differential equations and multiscale mechanics, Ph.D. Thesis, Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, USA, 2008. <http://digarc.usc.edu/assetserver/controller/view/search/etd-Das-20080513>.
[35] G. Tolstov, Fourier Series, Prentice-Hall, 1962.
[36] N. Lebedev, Special Functions and Their Applications, Dover, New York, 1972.
[37] D. Kurowicka, R. Cooke, Uncertainty Analysis with High Dimensional Dependence Modelling, John Wiley & Sons Ltd., 2006.
[38] V. Schmitz, Copulas and stochastic processes, Ph.D. Thesis, Rheinisch-Westfälische Technische Hochschule Aachen, Fakultät für Mathematik, Informatik und Naturwissenschaften, 2003. <http://deposit.ddb.de/cgi-bin/dokserv?idn=972691669>.
[39] R. Nelsen, An Introduction to Copulas, Springer, 2006.
[40] G. Fishman, Monte Carlo: Concepts, Algorithms and Applications, Springer, New York, 1996.
[41] L. Vandenberghe, S. Boyd, Semidefinite programming, SIAM Review 38 (1) (1996) 49–95.
[42] J. Löfberg, YALMIPA toolbox for modeling and optimization in MATLAB, in: Proceedings of the IEEE Conference on Computer Aided Control Systems Design (CACSD), Taipei, Taiwan, 2004. <http://control.ee.ethz.ch/joloef/yalmip.php>.
[43] S. Finette, M. Orr, A. Turgut, J. Apel, C. Badiey, B. Chiu, J. Headrick, J. Kemp, J. Lynch, A. Newhall, K.D. Heydt, B. Pasewark, S. Wolf, D. Tielbuerger, Acoustic field variability induced by time evolving internal wave fields, Journal of the Acoustical Society of America 108 (3) (2000) 957–972.
[44] M. Pellissetti, R. Ghanem, A method for validation of the predictive computations using a stochastic approach, Journal of Offshore Mechanics and Arctic Engineering 126 (2004) 227–234.
[45] S. Das, R. Ghanem, Uncertainty analysis for surface ship subjected to underwater detonation, in: Proceedings of the IMAC-XXII Conference and Exposition on Structural Dynamics, Dearborn, Michigan, USA, 2004.
[46] M. Loève, Probability Theory II, Springer-Verlag, New York, 1978.
[47] J. Hwang, A. Ding, Convergence study of the truncated Karhunen-Loève expansion for simulation of stochastic processes, International Journal for Numerical Methods in Engineering 52 (2001) 1029–1043.
[48] C. Schwab, R.A. Todor, Karhunen-Loève approximation of random fields by generalized fast multipole methods, Journal of Computational Physics 217 (2006) 100–122.
[49] G. Phillips, Interpolation and Approximation by Polynomials, Springer, 2003.