Arch Computat Methods Eng (2019) 26:245–274
DOI 10.1007/s11831-017-9240-5

ORIGINAL PAPER

A Critical Review of Surrogate Assisted Robust Design Optimization

Tanmoy Chatterjee[1] · Souvik Chakraborty[2] · Rajib Chowdhury[1]

Received: 1 June 2017 / Accepted: 1 July 2017 / Published online: 13 July 2017

Abstract Robust design optimization (RDO) has been eminent, ascertaining optimal configuration of engineering systems in presence of uncertainties. However, computational aspect of conventional RDO can often get computationally intensive as neighborhood assessments of every solution are required to compute the performance variance and ensure feasibility. Surrogate assisted optimization is one of the efficient approaches in order to mitigate this issue of computational expense. However, the performance of a surrogate model plays a key factor in determining the optima in multi-modal and highly non-linear landscapes, in presence of uncertainties. In other words, the approximation accuracy of the model is principal in yielding the actual optima and thus, avoiding any misguide to the decision maker on the basis of false or, local optimum points. Therefore, an extensive survey has been carried out by employing most of the well-known surrogate models in the framework of RDO. It is worth mentioning that the numerical study has revealed consistent performance of a model out of all the surrogates utilized. Finally, the best performing model has been utilized in solving a large-scale practical RDO problem. All the results have been compared with that of Monte Carlo simulation results.

✉ Tanmoy Chatterjee
tanmoydce88@gmail.com

Souvik Chakraborty
csouvik41@gmail.com

Rajib Chowdhury
rajibfce@iitr.ac.in

[1] Department of Civil Engineering, Indian Institute of Technology Roorkee, Roorkee 247667, India

[2] Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

1 Introduction

The presence of uncertainty is inevitable in the life cycle analysis and design optimization of an industrial output. In order to illustrate its omnipresence, manufacturing variations can be considered as a classical example of one of the major sources of uncertainties inherently associated with a product [1]. Deterministic approaches may not always converge to the desired optima, especially when the design solution is highly sensitive to such variations. In such critical situations, it could lead to either unsafe or over safe outcomes. Therefore, incorporating uncertainty in product analysis and design is necessary to yield economically viable solutions.

Robust design optimization (RDO) is one of the most popular approaches which take into account the effect of uncertainties into the design optimization formulation [2–4]. RDO has been observed to improve product quality significantly, and yield insensitive solutions, even in real time industrial applications [5, 6]. Over the last two decades, it has gained much attention across various domains, such as, telecommunications and electronics [7–9], aerospace [10–12], automobile [13–15], ship design [16–18], structural mechanics [19–21], structural dynamics and vibration control [22, 23] and fatigue analysis [24–26].

RDO establishes a mathematical framework for optimization motivated to minimize the propagation of input uncertainty to output responses [27, 28]. A graphical representation presented in Fig. 1 illustrates the concept of RDO. Out of the two optimal solutions, x_2 is more robust as compared to x_1, since the former does not affect the objective function $f(x)$ much, and hence is less sensitive. Broadly, the problem formulations of RDO generally constitute of maximizing the performance, minimizing the performance variance, or both. For the details and insight

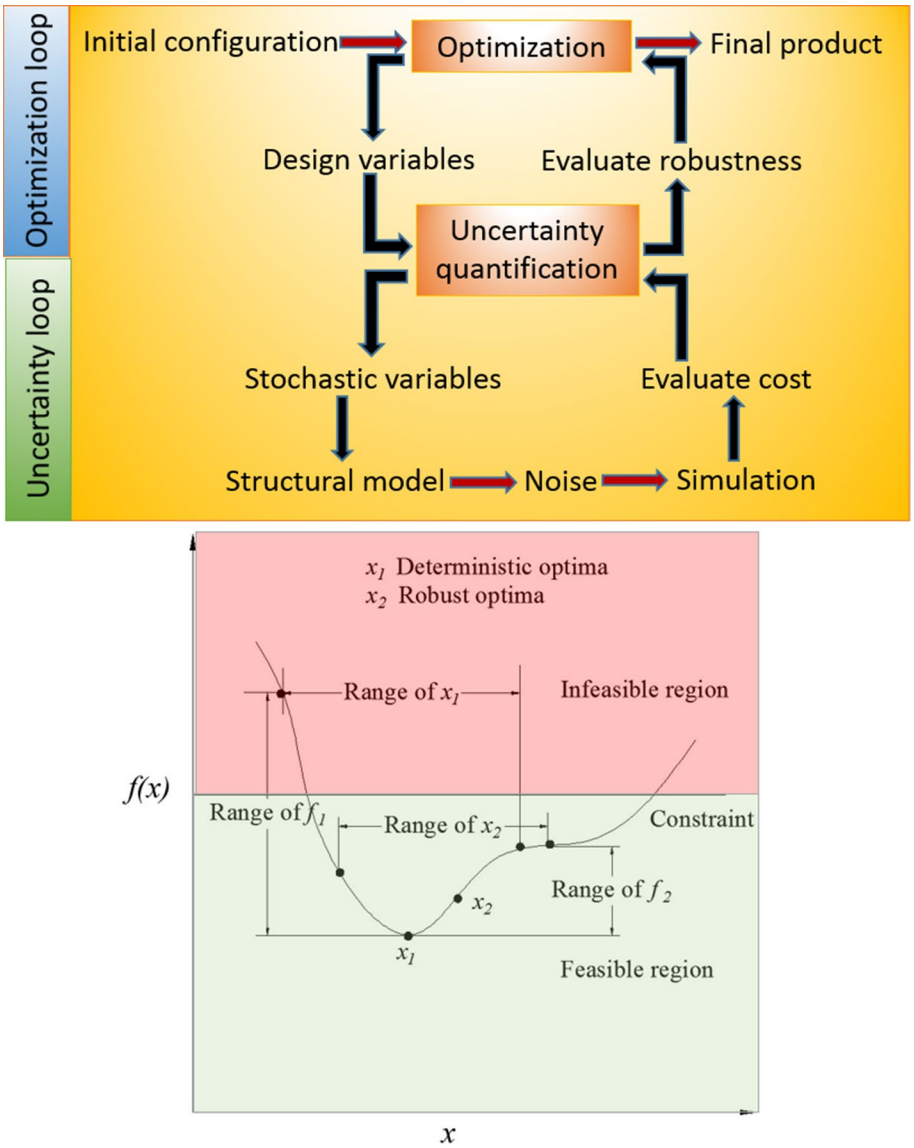

Fig. 1 Schematic diagram illustrating robust design optimization (RDO)

of RDO formulation, the reader is referred to the following literatures [2, 29, 30]. Since RDO attempts to resolve sensitive design solutions in a random environment, it suffers from computational issues as an obvious consequence. Despite the advances in computer configuration and speed, the mammoth computational costs of running such complex simulation codes have prohibited their usage.

Therefore, in order to avoid running such high fidelity simulations, various approximation schemes have emerged [31]. These techniques, also known as surrogate modelling, have been observed to resolve the computational expenses significantly by approximating the underlying model in a sample space [32, 33]. Various such techniques have been developed till date, such as least square approximation [34], moving least square [35], polynomial chaos expansion [36], anchored ANOVA decomposition [37], Kriging [38], radial basis function [39], artificial neural network [40], support vector machines [41] and multivariate adaptive regression splines [42].

Such techniques have been successfully utilized in design optimization for engineering applications. A comprehensive review of the use of approximation models in mechanical and aerospace systems, and multidisplinary design optimization can be found in [12, 43, 44]. However, in context to optimization, such efficient paradigms have found their use primarily in deterministic frontier. To further illustrate the gap in research in this particular area, it may be justified to reveal that application of approximation models in RDO is quite limited in number and content [1]. A possible explanation for the lack of investigation in this area can be derived from the fact that the performance output is likely to be vulnerable and unstable in presence of

uncertainties in an optimization framework. Additionally, solutions of constrained RDO problems can easily move to infeasible regions, if they are located close to constraint boundaries [45]. Thus, accuracy of the method to capture the original model is the governing factor to yield stable results satisfying feasible bounds, as the optima is highly sensitive to the convergence of each iteration.

Thus, the primary motivation of this paper lies in investigating most of the popular surrogate models and study their performance in the framework of RDO. Considering the lack of relevant literature in surrogate assisted RDO, an extensive review and thorough comparative assessment of various surrogate models in RDO has been felt as the need of the hour by the authors. This work is expected to serve as the guiding and selection application of the appropriate approximation models for the solution of high-fidelity computationally expensive stochastic optimization problems. Six benchmark RDO examples have been solved by utilizing as many as eleven popular surrogate models. Finally, a practical problem has been realistically modelled and solved by utilizing the most consistently performing model. To be specific, the dimensionality of the problems in terms of the number of stochastic variables vary from 2 to 48. The robust optimal solutions obtained have been validated with that of Monte Carlo simulation (MCS).

The paper has been organized in the following sequence. The theoretical development of the surrogate models employed for this study has been illustrated in Sect. 2. Section 3 explains the framework of surrogate assisted RDO technique. Numerical study has been carried out in Sect. 4 in order to illustrate the efficiency and accuracy of the various surrogate assisted RDO frameworks. A practical engineering RDO problem has been efficiently addressed in Sect. 5. Finally, conclusion has been drawn by discussing potentiality of the surrogate models in RDO framework on the basis of the results obtained from the study.

2 Surrogate Modelling

In this section, various surrogate models have been explained along with brief description of their mathematical formulations, which would provide complete insight to the readers.

2.1 Anchored ANOVA Decomposition

Various methods approximate multivariate functions in such a way that the component functions of the approximation are ordered starting from a constant and gradually approaching to multivariance as one proceeds along first order, second order and so on. An example of such method is ANOVA decomposition [37, 46–48] which is a general set of quantitative model assessment and analysis tool for mapping the high dimensional relationships between input and output model variables. It is an efficient formulation of the system response, if higher order co-operative effects are weak, allowing the physical model to be captured by the lower order terms. Practically for most well-defined physical systems, only relatively low order co-operative effects of the input variables are expected to have a significant effect on the overall response. ANOVA decomposition utilizes this property to fit an accurate hierarchical representation of the physical system. The fundamental concepts of generalized ANOVA decomposition has been discussed henceforth.

Let, $\mathbf{x} = \left(x_1, x_2, \ldots, x_P\right) \in \mathbb{R}^P$ and consider that $\eta = \mathbb{L}^2\left(\mathbb{R}^P, P_X\right)$. The generalized ANOVA decomposition expresses $\eta(\mathbf{x}) = \eta\left(x_1, x_2, \ldots, x_P\right)$ as the sum of the hierarchical correlated function expansion in terms of the variables [37, 48] as

$$\begin{aligned}\eta(\mathbf{x}) &= \eta_0 + \sum_{i=1}^{P} \eta_i\left(x_i\right) + \sum_{1\le i\le j\le P} \eta_{i,j}\left(x_i, x_j\right) + \ldots + \eta_{1,\ldots,P}(\mathbf{x}) \\ &= \sum_{u\subseteq\{1\ldots P\}} \eta_u\left(x_u\right)\end{aligned} \tag{1}$$

The expansion (1) exists and is unique under one of the hypothesis:

$$\begin{aligned}&\int \eta_u\left(x_u\right) dP_{X_i} = 0 \quad \forall i \in u, \forall u \subseteq \{1 \ldots P\} \\ &\int \eta_u\left(x_u\right)\eta_v\left(x_v\right) dP_X = 0 \quad \forall u, v \subseteq \{1 \ldots P\}, u \neq v\end{aligned} \tag{2}$$

where η_0 is constant term representing the zeroth order component function or the mean response. The function $\eta_i\left(x_i\right)$, referred to as first order component function representing the independent effect. Similarly, $\eta_{i,j}\left(x_i, x_j\right)$ is termed as second-order component function and represents co-operative effect of two-variables acting at a time. The higher-order terms indicate higher order co-operative effect with $\eta_{1,\ldots,P}(\mathbf{x})$ denoting the effect of all the variables acting together.

Moreover, each term η_u in the model function $Y = \eta(\mathbf{X})$ [in Eq. (1)] can be solved explicitly by integration and is given by [49],

$$\begin{aligned}&\eta_0 = \mathrm{E}(\mathbf{X}) \\ &\eta_i = \mathrm{E}\left(Y/X_i\right) - \mathrm{E}(Y),\ i = 1, \ldots, P \\ &\eta_u = \mathrm{E}\left(Y/X_u\right) - \sum_{v\subset u} \eta_v,\ |u| \ge 2\end{aligned} \tag{3}$$

The conditional expectation $\mathrm{E}\left(Y/X_u\right)$ effectively reduces a P dimensional function $\eta\left(x_1, x_2, \ldots, x_P\right)$ to a linear sum of following smaller $\alpha(<P)$ dimensional functions [50]:

$$\left[\prod_{1\le u\ne i,j,\ldots,i_\alpha\le P}\phi^{u}_{\mu_u}\left(x_u\right)\right]\times\eta\left(x_i,x_j,\ldots,x_{i_\alpha}:\left\{\prod_{1\le u\ne i,j,\ldots,i_\alpha\le P}x_u^{\mu_u}\right\}\right) \tag{4}$$

Over a $(P-\alpha)$ dimensional rectangular grid $\prod_{u\ne i,j,\ldots,i_\alpha} x_u^{\mu_u},\ \mu_u=1,2,\ldots,K_u$

Remark In this study, second order anchored ANOVA decomposition has been utilized for approximating actual functions. To generate the approximation of any function, initially a reference point $\bar{\mathbf{x}}=\left(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_P\right)$, has to be defined in the variable space. In practice, the choice of the reference point $\bar{\mathbf{x}}$ is essential, especially only if the first few terms, i.e., first and second order, in Eq. (1) are considered. The reference point $\bar{\mathbf{x}}$ at the middle of the input domain appears to be the ideal choice [51].

Based on the above formulation presented, a step-by-step procedure for approximation by utilizing anchored ANOVA decomposition has been provided in algorithm 1.

Table 1 The correspondence of the type of orthogonal polynomial with distribution pattern

Nature of distribution	Random variables	Type of orthogonal polynomial	Support
Continuous	Gaussian	Hermite	$(-\infty,\infty)$
	Gamma	Laguerre	$[0,\infty)$
	Beta	Jacobi	$[a,b]$
	Uniform	Legendre	$[a,b]$
Discrete	Poisson	Charlier	$\{0,1,\ldots\}$
	Binomial	Krawtchouk	$\{0,1,\ldots,N\}$
	Negative binomial	Meixner	$\{0,1,\ldots\}$
	Hypergeometric	Hahn	$\{0,1,\ldots,N\}$

$$\hat{g}(Z)=\sum_{|\mathbf{i}|=0}^{N}a_{\mathbf{i}}\Phi_{\mathbf{i}}(Z) \tag{5}$$

Algorithm 1. Illustration of the computational paradigm of anchored ANOVA decomposition

1. Identify the range of input variables x_i, $i=1\,\text{to}\,P$.
2. Select the center of the range of variables as the reference point $\bar{\mathbf{x}}$.
3. Determine η_0 which is a constant term, representing the response at $\bar{\mathbf{x}}$.
4. Compute first and second order component functions η_i and $\eta_{i,j}$ respectively, by utilizing Eq. (3)
5. Interpolate each of the low dimensional ANOVA expansion terms with respect to the input variables by utilizing Eq. (4). Specifically,
   a. In first order component functions and for $x_i=x_i^{i_1}$, n function values are given at $n=\left(3,5,7\,\text{or}\,9\right)$ regularly spaced sample points along variable axis x_i.
   b. Similarly, considering second order component functions for $x_i=x_i^{i_1}$ and $x_j=x_j^{i_2}$, n^2 function values are given on a grid formed by adopting $n=\left(3,5,7\,\text{or}\,9\right)$ regularly spaced sample points along x_i and x_j, axes.
6. Evaluate responses at these sample points.
7. Construct functional anchored ANOVA expansion.

2.2 Polynomial Chaos Expansion

The polynomial chaos expansion (PCE) is an efficient technique for obtaining the responses of stochastic systems. This has been introduced by Wiener [52] and hence, known as 'Wiener Chaos expansion'. The generalized results have been presented by Xiu and Karniadakis [53] for various continuous and discrete system from the so called Askey-scheme and further stated the $\mathcal{L}_2$ convergence in the corresponding Hilbert space.

Assuming $\mathbf{i}=\left(i_1,i_2,\ldots,i_n\right)\in\mathbb{N}_0^n$ be a multi-index with $|\mathbf{i}|=i_1+i_2+\cdots+i_n$, and let $N\ge 0$ be an integer. The Nth order PCE of $g(Z)$ can be stated as:

where $\{a_{\mathbf{i}}\}$ are unknown coefficients which are to be determined. $\Phi_{\mathbf{i}}(Z)$ are n-dimensional orthogonal polynomials with maximum order of N and satisfies the following relation

$$E\left(\Phi_{\mathbf{i}}(Z)\Phi_{\mathbf{j}}(Z)\right)=\int_{\Omega}\Phi_{\mathbf{i}}(Z)\Phi_{\mathbf{j}}(Z)\varpi(z)=\delta_{ij},0\le|\mathbf{i}|,|\mathbf{j}|\le N \tag{6}$$

here $\delta_{\mathbf{ij}}$ denotes the multivariate kronecker delta function. It is to be noted that if $\varpi(z)$ is Gaussian, the orthogonality relation in Eq. (6) yields Hermite polynomial as the optimal polynomial. The correspondence of the type of orthogonal

polynomial and type of random variable has been presented in Table 1 [53].

Since the emergence of generalised PCE [53], discrete variants of PCE have been developed. It is worth mentioning that each of the variants of PCE are based on Eq. (5). However, the uniqueness resides in the algorithm utilized to determine the unknown coefficients associated with the bases. The Weiner-Askey PCE, proposed by Xiu and Karniadakis [53] is based on the Galerkin projection. In this method, Galerkin projection has been utilized to decompose the governing stochastic partial differential equation into a system of coupled differential equations. Furthermore, it has been demonstrated that PCE based on Galerkin approach yields excellent results and provide exponential convergence with increase in order of PCE. However, the method, being intrusive in nature, requires knowledge regarding the governing partial differential equation of the system. Consequently, it is not applicable to real-life problems with unknown governing differential equation. In order to address this issue, special attention has been provided to develop non-intrusive PCE. The most popular non-intrusive PCE is the one based on least square method [54, 55]. In this method, the least square technique is implemented to determine the unknown coefficients associated with the bases. Least square based PCE is easy to implement and applicable for systems with unknown governing differential equations. Other alternatives that have been investigated for determining the unknown coefficients associated with the bases are quadrature method [56, 57] and collocation approach [58, 59]. However, all the variants of PCE discussed above are only suitable for small scale problems. This is because, the number of unknown coefficients associated with PCE increases factorially with increase in number of variables. This renders application of PCE to large scale problems infeasible.

Blatman and Sudret [60, 61] proposed two adaptive sparse PCE for solving high-dimensional problems. The purpose of the methods is to determine, in an iterative manner, the components/variables that significantly contributes to the response of interest. The number of unknown coefficients associated with the bases are reduced by eliminating the components/variables having no or less effect on the output response. While the first approach [60] utilizes change in coefficient of determination (R^2) to identify the significant components, the second approach utilizes least angle regression scheme [61] to identify the less important components. Moreover, it has been demonstrated that the proposed adaptive sparse PCE is capable of treating systems having number of variables as large as 500. However, only problems governed by elliptical partial differential equations have been investigated as part of the above works.

Due to its superior performance, PCE has found wide applications is various domains. PCE has been utilized for solving the stochastic steady state diffusion problem by Xiu and Karniadakis [62] and the stochastic Navier–Stokes equation [63]. Further, PCE has been utilized for sensitivity analysis by Sudret [64]. A reduced PCE has been developed for stochastic finite element analysis by Pascual and Adhikari [65, 66]. In each of the above mentioned applications, PCE has been found to yield excellent results. However, there are few issues regarding PCE that are yet to be answered. Firstly, PCE is only applicable to systems involving independent random variables. If the system under consideration involves correlated variables, ad hoc transformations, such as, Nataf transformation, needs to be employed to transform the dependent variables into independent variables. Secondly, PCE involves determining orthogonal polynomials for a system. However, orthogonal polynomials are known only for a few random variables as shown in Table 1. Hence, if the system under consideration involves variable(s), orthogonal polynomial(s) for which is not known, implementation of PCE may become tedious.

2.3 Multivariate Adaptive Regression Splines (MARS)

MARS [42] is governed by a set of bases which are selected for approximating the output response through a forward or backward iterative approach. The functional form of MARS is represented as:

$$\hat{g}(\boldsymbol{X}) = \sum_{k=1}^{n} \alpha_k \, H_k^f(X_i) \tag{7}$$

with

$$H_k^f(X_1, X_2, X_3, \ldots, X_n) = 1 \tag{8}$$

where α_k and $H_k^f(x_i)$ are the coefficient of the expansion and the basis functions, respectively. The basis function can be represented as

$$H_k^f(X_i) = \prod_{i=1}^{i_k} \left[z_{i,k} \left(X_{j(i,k)} - t_{i,k} \right) \right]_{tr}^{q} \tag{9}$$

where i_k is the order of interaction in the kth basis function and $z_{i,k} = \pm 1$. $x_{j(i,k)}$ in Eq. (9) is the jth variable with $1 \leq j(i,k) \leq n$. $t_{i,k}$ represents the knot location on each of the corresponding variables. $H_k^f(x_i)$ in Eq. (9) represents the multivariate spline basis function and is represented as the product of univariate spline basis functions $z_{i,k}$, which is either of order one or cubic, depending on the degree of continuity of the approximation. The notation "tr" denotes the function is a truncated power function.

$$\begin{aligned} \left[z_{i,k} \left(X_{j(i,k)} - t_{i,k} \right) \right]_{tr}^{q} &= \left[z_{i,k} \left(X_{j(i,k)} - t_{i,k} \right) \right]^q \text{ for } \left[z_{i,k} \left(X_{j(i,k)} - t_{i,k} \right) \right] < 0 \\ &= 0, \text{ otherwise} \end{aligned} \tag{10}$$

Each function is piecewise linear with a knot *tr* at every $X_{(i,k)}$. MARS models the function by allowing the basis function to bend at the knots. The maximum number of knots considered, the minimum number of observations between knots, and the highest order of interaction terms are to be determined. Automated variable screening is performed within MARS by using the generalized cross-validation (GCV) model fit criterion, which has been developed by Craven and Wahba [67]. The location and number of spline bases needed is determined by a (a) over-fitting a spline function through each knot and (b) removing the knots that have least contribution to the overall fit of the model as determined by the modified GCV criterion. The following equation shows the lack-of-fit (LOF) criterion used by MARS:

$$L_{fc}\left(\hat{g}_k\right) = G_{cv}(k) = \frac{(1/n)\sum_{k=1}^{n}\left[g\left(X_i\right) - \hat{g}_k\left(X_i\right)\right]^2}{\left[1 - \{\tilde{c}(k)/n\}\right]^2} \tag{11}$$

where

$$\tilde{c}(k) = c(k) + d.k \tag{12}$$

MARS has been utilized by Sudjianto et al. [68] to emulate a conceptually intensive complex automotive shock tower model in fatigue life durability analysis. A comparative assessment of MARS as compared to linear, second-order and higher-order regression models has been carried out by Wang et al. [69]. MARS has been utilized by Friedman [42] to approximate behaviour of performance variables in a simple alternating current series circuit. The primary advantage of MARS is its computational efficiency. Moreover, MARS model is capable of handling large data and almost no data preparation is required for building it. However, accuracy of MARS model is lower as compared to other surrogate techniques. Additionally, MARS is incapable of predicting the confidence bound of prediction and additional sample points are required for its validation. This, in turn, reduces the computational efficiency of the MARS model.

2.4 Radial Basis Function

Radial basis function (RBF) is another surrogate model which is quite popular among researchers. RBF is often used to perform the interpolation of scattered multivariate data [70–72]. The metamodel appears in a linear combination of Euclidean distances, which may be expressed as

$$\hat{g}(\boldsymbol{X}) = \sum_{k=1}^{n} w_k \varphi_k\left(\boldsymbol{X}, x_k\right) \tag{13}$$

where n is the number of sampling points, w_k is the weight determined by the least-squares method and $\varphi_k(X, x_k)$is the k-th basis function determined at the sampling point x_k. Various symmetric radial functions are used as basis function. The radial functions for RBF model are illustrated as,

$$R_f(\boldsymbol{X}) = \exp\left(-\frac{(\boldsymbol{X}-c)^T(\boldsymbol{X}-c)}{r^2}\right) \quad \text{(For Gaussian)} \tag{14}$$

$$R_f(\boldsymbol{X}) = \sqrt{1 + \frac{(\boldsymbol{X}-c)^T(\boldsymbol{X}-c)}{r^2}} \quad \text{(For multi-quadratic)} \tag{15}$$

$$R_f(\boldsymbol{X}) = \frac{1}{\sqrt{1 + \frac{(\boldsymbol{X}-c)^T(\boldsymbol{X}-c)}{r^2}}} \quad \text{(For inverse multi-quadratic)} \tag{16}$$

$$R_f(\boldsymbol{X}) = \frac{1}{1 + \frac{(\boldsymbol{X}-c)^T(\boldsymbol{X}-c)}{r^2}} \quad \text{(For Cauchy)} \tag{17}$$

It is to be noted that unlike PCE and response surface method (RSM), RBF is not a regression technique. Rather, RBF may be broadly considered as an interpolation technique. As a result, RBF, unlike regression techniques, yields exact result at the sample points.

Till date, RBF has found wide application in the domain of structural reliability and uncertainty quantification. A dynamic surrogate model based on stochastic RBF has been developed by Volpi et al. [73] for uncertainty quantification. The method has been equipped with auto-tuning scheme based on curvature, adaptive sampling scheme, parallel infill and multi-response criterion. It has also been illustrated that the surrogate based on stochastic RBF outperforms popular surrogate such as Kriging. A hybridized RBF has been proposed by Dai et al. [74] for structural reliability analysis. To be specific, the hybridized RBF has been formulated by replacing the learning network of RBF with support vector algorithm. As a consequence, it has been possible to exploit the advantages of support vector algorithm, such as good generalization and global optimization. Comparative assessment illustrated that the method proposed outperform both RBF and support vector algorithm. Other works on RBF include, but are not limited to, development of performance measure approach based reliability analysis technique using RBF [75] and integration of RBF into the FORM algorithm [76]. However, RBF yields satisfactory results only for problems that are linear and/or, weakly non-linear.

2.5 Artificial Neural Network

Artificial neural networks (ANNs) are a family of surrogate model inspired by biological functioning of brain and nervous system. ANNs are generally presented as a system of interconnected 'neuron'. The neurons in ANN, often termed as nodes, consists of some primitive functions. All the connections have numeric weight that are tuned based on input data. A typical structure of ANN has been illustrated in Fig. 2. The network is represented as a function Φ obtained by combining the primitive functions f_1, f_2, f_3 and f_4. $\alpha_1, \alpha_2, \ldots, \alpha_5$ are termed as weights and determined by employing some learning algorithm. Based on the above points, it should be clear that three elements govern the formation of an ANN, namely the primitive function associated with the node, topology of the network (single layer or multilayer) and the learning algorithm used to determine the weights associated with the connections. Based on these criteria, multiple variants of ANN have evolved over the years.

Multilayer feed-forward neural network (MFFNN) [77] is the most popular and widely used ANN. Here, the neurons are arranged in three layers, namely input layer, hidden layer and output layer. It is worthwhile to mention that the number of hidden layers in feed-forward neural network may be more than one. Therefore, it is necessary to perform convergence study to determine the optimum number of layers. Moreover, the number of neurons/nodes in each hidden layer should also be determined by using some appropriate convergence criterion. Each neuron/node consists of a transfer function that expresses the internal activation level of the neuron. A transfer function may either of linear or nonlinear. An account of popular transfer functions has been provided in Table 2.

Another popular ANN scheme is the well-known back-propagation algorithm based ANN [78]. In this scheme, the errors in prediction are propagated backward to the inputs. Based on the errors received, the weights associated with the connections are further updated. The process is repeated until errors at the output layer is less than a specified threshold.

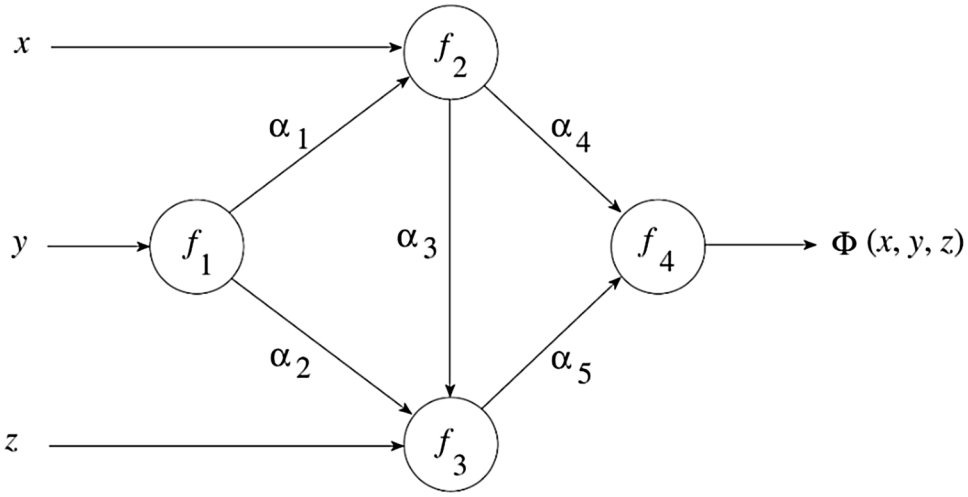

Fig. 2 Typical structure of ANN

Table 2 Popular transfer functions in ANN

Name	Input (x)/output (y) relation	Name	Input (x)/output (y) relation
Hard limit	$y = 0, x < 0$ $y = 1, x \geq 0$	Symmetric saturating linear	$y = -1,\ x < -1$ $y = n,\ -1 \leq x \leq 1$ $y = 1,\ x > 1$
Symmetrical hard limit	$y = -1, x < 0$ $y = 1, x \geq 0$	Log-sigmoid	$y = 1/(1 + \exp(-x))$
Linear	$y = x$	Hyperbolic tangent sigmoid	$y = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$
Saturating linear	$y = 0, x < 0$ $y = n, 0 \leq x \leq 1$ $y = 1, x > 1$	Positive linear	$y = 0, x < 0$ $y = n, x \geq 0$

Due to its high accuracy level, ANN has found wide application in uncertainty quantification and reliability analysis. ANN based response surface method has been presented by Shu and Gong [79] for reliability analyses of c-phi slopes. The soil properties have been assumed to be having spatial randomness and hence modelled as random field. It has been observed that the proposed approach yield accurate estimation of failure probability. A multi-wavelet neural network based response surface method has been proposed by Dai et al. [80] for structural reliability analysis. It has been illustrated that the proposed algorithm outperforms the well-known multilayer perceptron based response surface method. Other works which utilized neural network in the field of uncertainty quantification and reliability analysis include [40, 81, 82].

2.6 Support Vector Regression

Support vector regression (SVR) is a variant of the Support Vector Machine (SVM) utilized for regression analysis. SVR uses a subset of data samples, support vectors, in order to construct an approximation model that has a maximum deviation of ε from the function value corresponding to each training data. For a linear mapping, the SVR model may be represented as

$$\hat{g}(\boldsymbol{X}) = \langle W \cdot \boldsymbol{X} \rangle + b \tag{18}$$

where $\hat{g}(\boldsymbol{X})$ is the approximate value of the objective function at x, W represents a vector of weights, b is the bias term, and $\langle \cdot \rangle$ denotes the inner product. Equation (18) may be further expressed as a convex optimization problem as

$$\begin{aligned} &\arg\min \quad 0.5||W||^2 \\ &\text{s.t.} \quad \begin{cases} g_i(X_i) - \langle W.X_i \rangle - b \le \varepsilon \\ \langle W.X_i \rangle + b - g(X_i) \le \varepsilon \end{cases} \end{aligned} \tag{19}$$

where $g_i(X_i)$ represent the responses at the sample points. It should be noted that there might not be a function that satisfies the condition in Eq. (19). Hence, introducing slack variables ξ_i, ξ_i^*, Eq. (19) can be rewritten as:

$$\begin{aligned} &\arg\min \quad 0.5||W||^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \\ &\text{s.t.} \quad \begin{cases} g_i(X_i) - \langle W.X_i \rangle - b \le \varepsilon + \xi_i \\ \langle W.X_i \rangle + b - g_i \le \varepsilon + \xi_i^* \\ \xi, \xi_i^* \ge 0 \end{cases} \end{aligned} \tag{20}$$

where n is the number of sample points. The regularization parameter, C, determines the trade-off between the model complexity and the degree for which deviation larger than ε is allowed in Eq. (20). The formulation discussed corresponds to dealing with a ε-insensitive loss function, as proposed by [83]

$$G(x) = \begin{Bmatrix} 0 & |g(\boldsymbol{X}) - \hat{g}(\boldsymbol{X})| \le \varepsilon \\ |g(\boldsymbol{X}) - \hat{g}(\boldsymbol{X})| - \varepsilon & \text{otherwise} \end{Bmatrix} \tag{21}$$

Next introducing a Lagrange multiplier as:

$$\begin{aligned} L: = 0.5||W||^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \\ - \sum_{i=1}^{n}\alpha_i\left(\varepsilon + \xi_i - g(X_i) + \langle W.X_i \rangle + b\right) \\ - \sum_{i=1}^{n}\alpha_i^*\left(\varepsilon + \xi_i^* + g(X_i) - \langle W.X_i \rangle - b\right) \\ - \sum_{i=1}^{n}\left(\eta_i\xi_i^* + \eta_i^*\xi_i^*\right) \end{aligned} \tag{22}$$

It is to be noted that the dual variables in Eq. (22) should satisfy the positivity constraints, i.e., $\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \ge 0$. Furthermore, it can be shown that Eq. (22) has a saddle point with respect to the primary and dual variables at the optimal solution [84]. Hence,

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) = 0 \tag{23}$$

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right)X_i = 0 \tag{24}$$

$$\frac{\partial L}{\partial \xi_i^{(*)}} = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0 \tag{25}$$

where $\alpha_i^{(*)}$ includes both α_i and α_i^*. Similarly, $\eta_i^{(*)}$ also includes both η_i and η_i^*. Substituting Eqs. (23)–(25) into Eq. (22) yields

$$\begin{aligned} &\text{Maximise} \begin{cases} -0.5\sum_{i,j=1}^{n}\left(\alpha_i - \alpha_i^*\right)\left(\alpha_j - \alpha_j^*\right)\langle X_i, X_j \rangle \\ -\varepsilon\sum_{i=1}^{n}\left(\alpha_i + \alpha_i^*\right) + \sum_{i=1}^{n} g(X_i)\left(\alpha_i - \alpha_i^*\right) \end{cases} \\ &\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases} \end{aligned} \tag{26}$$

It should be noted that variables η_i and η_i^* are not present in Eq. (26). Additionally using Eq. (24),

$$w = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) X_i \tag{27}$$

In order to compute b, the so-called Karush–Kuhn–Tucker (KKT) condition has been utilized. According to the KKT condition,

$$\begin{aligned} \alpha_i(\varepsilon + \xi_i - g(X_i) + \langle W.X_i \rangle + b) = 0 \\ \alpha_i^*(\varepsilon + \xi_i^* + g(X_i) - \langle W.X_i \rangle - b) = 0 \end{aligned} \tag{28}$$

and

$$\begin{aligned} (C - \alpha_i)\xi_i = 0 \\ (C - \alpha_i^*)\xi_i^* = 0 \end{aligned} \tag{29}$$

From Eq. (29), important remarks can be presented. First, $\left(C - \alpha_i^{(*)}\right) = 0$ only if the samples lie outside the ε-insensitive zone. Secondly, $\alpha_i \alpha_i^* = 0$, i.e., either α_i or α_i^* should always be zero. Finally, for $\alpha_i^{(*)} \in (0, C)$, $\xi_i^{(*)} = 0$. Hence, the second factor in Eq. (28) vanishes. Therefore,

$$\begin{aligned} b = g(X_i) - \langle W.X_i \rangle - \varepsilon \\ b = g(X_i) - \langle W.X_i \rangle + \varepsilon \end{aligned} \tag{30}$$

Similar to the procedure described above, a non-linear regression can be achieved by replacing the $\langle \cdot \rangle$ in Eq. (18) with a kernel function, K, [83] as

$$\hat{g}(\boldsymbol{X}) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(X_i, \boldsymbol{X}) + b \tag{31}$$

All other operations will be applicable as discussed previously.

SVR, of late, has found wide application in uncertainty quantification and reliability analysis. Least square based SVR has been utilized [85] for reliability analysis. It has been illustrated that the structural risk associated with SVR is inherently minimized and therefore, suitable as a surrogate model for reliability analysis. Bootstrap technique has been integrated into the framework of SVR by Lins et al. [86] for obtaining confidence and prediction interval using SVR. The bootstrap based SVR has been utilized for a large scale problem involving component degradation of offshore oil industry. A novel dynamic-weighted probabilistic SVR has been proposed by Liu et al. [87]. The approach developed has been utilized for a real case study on reactor coolant pump governed by 20 failure scenarios. Other significant work on SVR include [88, 89].

Tuning the parameters for obtaining optimum performance is an important aspect associated with SVR. A hybrid method (APSO-SVR) based on the particle swarm optimization and analytical selection for tuning of parameters in SVR has been developed by Zhao et al. [90]. It has been demonstrated that APSO-SVR outperforms the conventional SVR is terms of convergence. Other significant contribution to parameter tuning of SVR include work by Zhao et al. [91] and Coen et al. [92].

2.7 Kriging

Kriging is a surrogate model which is based on Gaussian process modelling. The basic idea of Kriging is to incorporate interpolation, governed by prior covariances, in order to obtain responses at unknown points [93, 94]. In this method, the functional response characteristics is illustrated as:

$$\hat{\boldsymbol{g}}(\boldsymbol{X}) = \boldsymbol{y}_0(\boldsymbol{X}) + \boldsymbol{Z}(\boldsymbol{X}) \tag{32}$$

where $\hat{\boldsymbol{g}}(\boldsymbol{X})$ is the response function of interest, $\boldsymbol{X}$ is an N dimensional vector (N design variables), $y_0(\boldsymbol{X})$ is the known approximation (usually polynomial) function and $Z(x)$ represents is the realization of a stochastic process with mean zero, variance, and non-zero covariance. In the model, the local deviation at an unknown point ($\mathbf{X}$) is expressed using stochastic processes. The sample points are interpolated with the help of Gaussian as the correlation function to estimate the trend of the stochastic processes [95, 96].

Consider, $\mathbf{X} = \{X_1, X_2, \ldots, X_N\}^T \in \mathfrak{R}^N$ be the vector of basic random variables and $g(\mathbf{X})$ be the system response output. In universal Kriging, $y_0(\boldsymbol{X})$ is represented by using a multivariate polynomial as:

$$y_0(x) = \sum_{i=1}^{p} a_i b_i(\boldsymbol{X}) \tag{33}$$

where $b_i(\boldsymbol{X})$ represents the ith basis function and a_i denotes the coefficient associated with the ith basis function. The primary idea behind such a representation is that the regression function captures the variance in the data (the overall trend) and the Gaussian process interpolates the residuals. Suppose $X = \{X^1, X^2, \ldots, X^n\}$ represents a set of n samples. Also, assume $g = \{g_1, g_2, \ldots, g_n\}$ to be the responses at training points. Therefore, the regression part can be written as a $n \times p$ model matrix F,

$$F = \begin{pmatrix} b_1(X^1) & \cdots & b_p(X^1) \\ \vdots & \ddots & \vdots \\ b_1(X^n) & \cdots & b_p(X^n) \end{pmatrix} \tag{34}$$

whereas, the stochastic process is defined using a $n \times n$ correlation matrix Ψ

$$\Psi = \begin{pmatrix} \psi(X^1, X^1) \cdots \psi(X^1, X^n) \\ \vdots \ddots \vdots \\ \psi X(x^n, X^1) \cdots \psi(X^n, X^n) \end{pmatrix} \tag{35}$$

where $\psi(\cdot,\cdot)$ is a correlation function, parameterised by a set of hyperparameters θ. The hyperparameters are further identified by maximum likelihood estimation (MLE). A detailed account of MLE in the context of Kriging can be found in [38]. The prediction mean and variance can be obtained as:

$$\mu(\boldsymbol{X}) = M\alpha + r(\boldsymbol{X})\Psi^{-1}(g - F\alpha) \tag{36}$$

and

$$s^2(\boldsymbol{X}) = \sigma^2\left(1 - r(\boldsymbol{X})\Psi^{-1}r(\boldsymbol{X})^T + \frac{\left(1 - F^T\Psi^{-1}r(\boldsymbol{X})^T\right)}{F^T\Psi^{-1}F}\right) \tag{37}$$

where $M = \left(\, b_1\left(X_p\right) \;\ldots\; b_p\left(X_p\right) \,\right)$ is the modal matrix of the predicting point X_p,

$$\alpha = \left(F^T\Psi F\right)^{-1}F^T\Psi^{-1}g \tag{38}$$

is a $p \times 1$ vector consisting of the unknown coefficients determined by generalised least squares regression and

$$r(x) = \left(\, \psi\left(X_p, X^1\right) \;\cdots\; \psi\left(X_p, X^p\right) \,\right) \tag{39}$$

is an $1 \times n$ vector denoting the correlation between the prediction point and the sample points. The process variance σ^2 is given by

$$\sigma^2 = \frac{1}{n}(g - F\alpha)^T\Psi^{-1}(g - F\alpha) \tag{40}$$

It is worthwhile to mention that the universal Kriging, as formulated above, is an interpolation technique. This can be easily validated by substituting the ith sample point in Eq. (36) and considering that $r\left(X^i\right)$ is the ith column of Ψ:

$$\mu\left(X^i\right) = M\alpha + g^i - M\alpha = g^i \tag{41}$$

One issue associated with the universal Kriging is selection of the optimal polynomial order. Conventionally, the order of the polynomial is selected empirically. However, such non-adapted framework may render the modelling inefficient. Recent works [97–104] have addressed these issues.

An essential feature associated with Kriging is selection of appropriate covariance function [105–107]. Mostly, the covariance functions used with Kriging surrogate are stationary and can be expressed in the following form:

$$\psi\left(x, x'\right) = \prod_j \psi_j\left(\theta, x_i - x_i{}'\right) \tag{42}$$

The correlation function defined in Eq. (42) has two desirable properties. Firstly, the correlation function for multivariate functions can be represented as product of one dimensional correlations. Secondly, the correlation is stationary and depends only in the distance between two points. Few standard stationary covariance functions, which have been investigated are namely, (a) exponential correlation function, (b) generalised exponential correlation function (c) Gaussian correlation function (d) linear correlation function (e) spherical correlation function (f) cubic correlation function and (g) spline correlation function. The mathematical forms of the above correlation functions are provided below:

i. Exponential correlation function:

$$\psi_j\left(\theta; d_j\right) = \exp\left(-\theta_j\left|d_j\right|\right) \tag{43}$$

ii. Generalised exponential correlation function:

$$\psi_j\left(\theta; d_j\right) = \exp\left(-\theta_j\left|d_j\right|^{\theta_{n+1}}\right),\ 0 < \theta_{n+1} \le 2 \tag{44}$$

iii. Gaussian correlation function:

$$\psi_j\left(\theta; d_j\right) = \exp\left(-\theta_j {d_j}^2\right) \tag{45}$$

iv. Linear correlation function:

$$\psi_j\left(\theta; d_j\right) = \max\left\{0, 1 - \theta_j\left|d_j\right|\right\} \tag{46}$$

v. Spherical correlation function:

$$\psi_j\left(\theta; d_j\right) = 1 - 1.5\,\xi_j + 0.5\,\xi_j^2,\ \xi_j = \min\left\{1,\ \theta_j\left|d_j\right|\right\} \tag{47}$$

vi. Cubic correlation function:

$$\psi_j\left(\theta; d_j\right) = 1 - 3\,\xi_j^2 + 2\,\xi_j^3,\ \xi_j = \min\left\{1,\ \theta_j\left|d_j\right|\right\} \tag{48}$$

vii. Spline correlation function:

$$\psi_j\left(\theta; d_j\right) = \begin{cases} 1 - 5\xi_j^2 + 30\xi_j^3, & 0 \le \xi_j \le 0.2 \\ 1.25\left(1 - \xi_j^3\right), & 0.2 \le \xi_j \le 1 \\ 0, & \xi_j > 1 \end{cases} \tag{49}$$

where $\xi_j = \theta_j\left|d_j\right|$.

For all the correlation functions described above, $d_j = x_i - x_i{}'$.

2.8 Locally Weighted Polynomials

An approach for the pointwise estimation of the unknown function from known samples based upon Taylor's series expansion has been proposed by Cleveland [108]. This idea was further extended into a statistical framework for model

approximation. Later, these methods have been generalized into kernel regression approach [109, 110].

Locally weighted polynomial (LWP) regression is one such form of instance-based algorithm for learning continuous non-linear mappings [111]. Basically, it is a non-parametric regression approach that combine multiple models in a k-nearest-neighbor based metamodel [112]. A low-order weighted least square model is fitted at each training point. Let the input–output mapping be represented as,

$$Y_j = M\left(x_j\right) + \sigma\left(x_j\right)\varepsilon_j \tag{50}$$

where $\varepsilon_j \sim N(0,1)$ and $\sigma^2\left(x_j\right)$ is the variance of Y_j at x_j. For cases where homoscedastic variance is assumed, $\sigma^2(x) = \sigma^2$. $M(x)$ can be obtained by solving the followinsg for α

$$\sum_{j=1}^{n}\left(Y_j - \sum_{i=1}^{p}\alpha_i\left(x_j - x_0\right)^i\right)^2 K_b\left(x_j - x_0\right) \tag{51}$$

where $K_b(\cdot)$ controls the weights and b controls the size of neighbourhood around x_0. Equation (51) can be rewritten as

$$\arg\min_{\alpha}\left(\mathbf{y} - \mathbf{x}\boldsymbol{\alpha}\right)^T\mathbf{W}\left(\mathbf{y} - \mathbf{x}\boldsymbol{\alpha}\right) \tag{52}$$

where $\mathbf{W}$ is a diagonal matrix of weights, $W_{ii} = K_b\left(x_j - x_0\right)$. Coefficient vector $\boldsymbol{\alpha}$ can be obtained by the following expression

$$\boldsymbol{\alpha} = \left(\mathbf{x}^T\mathbf{W}\mathbf{x}\right)^{-1}\mathbf{x}^T\mathbf{W}\mathbf{y} \tag{53}$$

Thus, there are three principal parameters whose selection may have an effect on the approximation, which are bandwidth (b), the order of polynomial (p), and the kernel or weight function (K_b) [113]. A natural way to select the bandwidth and calibrate the tradeoff is to minimize the mean squared error [114]. In this context, a variable bandwidth selector for kernel regression which can be extended for local linear regression has been proposed in [115]. An adaptive method has been proposed by Fan and Gijbels [116] for selecting the appropriate order of polynomials based on local factors, allowing p to range through various points within the support. In context to choose K_b, it has been established in [113], that the constant for the Epanechnikov kernel is the smallest and hence optimal in terms of integrated mean squared error. However, the difference between the kernels is negligible, so the selection may depend upon the user's preference at large.

Expressions for the asymptotic bias and variance of an estimate have been presented in [117]. The local linear model using the Epanechnikov kernel has been proven to optimize the linear minimax risk [118], which is a criterion to benchmark the efficiency of an estimator in terms of sample size required for obtaining a certain level of accuracy in results. Later, these results were extended to LWP in [119]. LWP is observed to perform well near the boundary of support of the data, unlike most of non-parametric models in which rate of convergence is slow [110]. However, the basis of slow convergence has been explained as lower number of training points were utilized for estimators near the boundary. It has been illustrated in [120], that no linear estimator can prove to be superior on the boundary in a minimax sense in terms of mean squared error in comparison to LWP. Further details can be found in [109]. Few references in which computational efficiency of LWP has been improved upon are [121–123]. Some recent extensions of LWP include [124–126].

After an extensive literature review of popular surrogate models, the following section has been attributed to demonstrate the utilization of any of the above models in RDO framework.

3 Surrogate Assisted RDO Framework

This section discusses that how the surrogate models will be employed to address the issue of computational expense of robust optimization. The objective and/or constraint functions in RDO involve mean and standard deviation of stochastic responses, which is the main reason for making the computational platform cumbersome. This is due to large number of simulations required to approximate the statistical quantities of responses during each optimization iteration. It is worth mentioning that, generally there are two ways to introduce efficiency in an RDO approach, which are,

- To avoid expensive original function/FE evaluation within an optimization iteration,
- To reduce the number of optimization iterations by preserving the elite solutions.

The scope of the present study is limited so as to address the first point discussed above, in order to present a comparative assessment of efficient surrogate assisted RDO tools. Therefore, instead of simulating original objective and constraint functions, the functions are approximated by surrogate models, which have been utilized within the optimization routine. This efficient surrogate based framework of RDO assists the computation to be limited to nominal costs, especially in case of large scale finite-element based models. A flow diagram of the entire steps involved in the surrogate assisted RDO approach has been depicted in Fig. 3 for better understanding.

During the evaluation of objective and constraint functions in Fig. 3, it is to be noted that in order to compute the

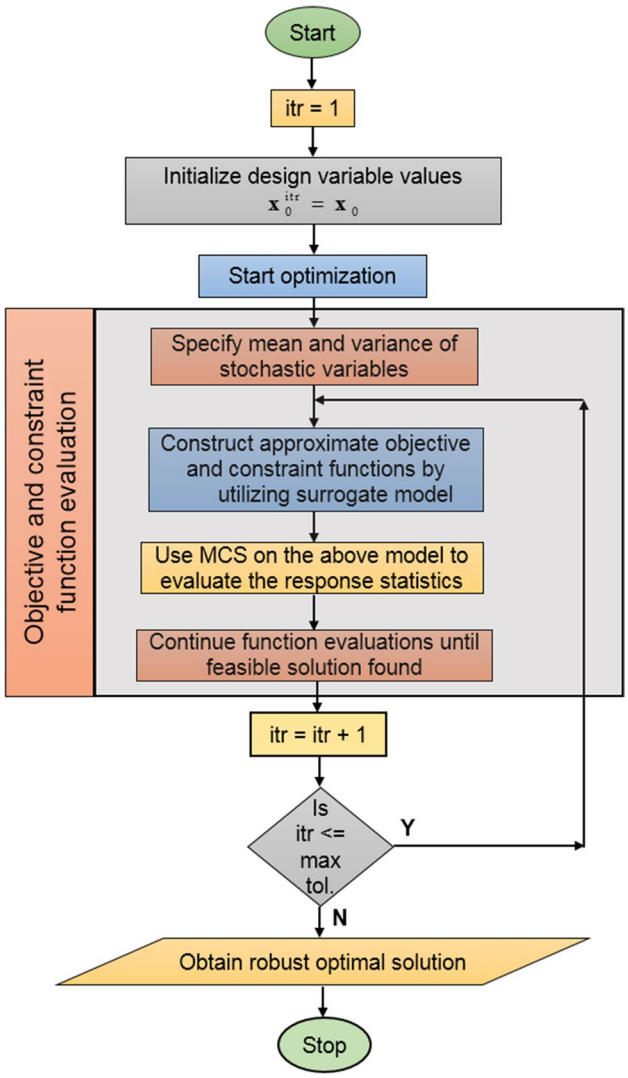

Fig. 3 Flowchart of surrogate assisted RDO framework utilized in this study

response statistics, simulations are carried out based upon the model generated by the surrogate. This renders significant level of computational efficiency in comparison to MCS performed on the actual FE model. Moreover, since simulations have to be carried out in each of the optimization iteration, computational savings can be achieved in each of such iterations, until convergence of the optima.

Thus, the computational efficiency of surrogates as compared to simulation based RDO framework can be realized as an obvious matter of fact, however, approximation accuracy of the former is a crucial factor yet to be investigated. Therefore, various surrogate models as described in Sect. 2 have been employed in solving few typical non-linear analytical examples in the following section.

4 Numerical Examples

In order to illustrate the efficiency and accuracy of the various surrogate models in RDO platform, six benchmark examples have been considered in this section. In example 1, a test function has been investigated. A two bar plane truss has been studied in example 2. Conceptual design of a bulk carrier is of great concern in shipping industry, which is considered in example 3. Design of a welded beam has been taken up as the fourth example. A speed reducer problem has been investigated in example 5. Side impact crashworthiness of a car has been considered in example 6. Each of the optimization problems deal with single objective and multiple constraint functions. The sequence of examples has been placed in accordance to the increasing number of stochastic variables and hence, increasing complexity. Results obtained have been compared with that of Monte Carlo simulations (MCS) based RDO solutions.

In this paper, the computational platform has been MATLAB® version 8.1 R2013a. MATLAB® toolbox *fmincon* has been utilized as the optimization search engine. Grid sampling has been utilized for generating training

Table 3 List of the surrogate models utilized and their abbreviations

S. no.	Surrogate model	Abbreviation used
1	Anchored ANOVA decomposition	ANOVA-D
2	Universal Kriging	UK
3	Kriging with first order polynomial bases	FK
4	Radial basis functions	RBF
5	Quadrature based polynomial chaos expansion	PCE-Q
6	Ordinary least squares based polynomial chaos expansion	PCE-OLS
7	Least angle regression based polynomial chaos expansion	PCE-LAR
8	Locally weighted polynomials	LWP
9	Artificial neural network	ANN
10	Multivariate adaptive regression splines	MARS
11	Support vector machine	SVM

points for constructing anchored ANOVA model. While the other surrogate models have been trained by utilizing latin-hypercube sampling [127]. For the comparison of computational effort, the number of original actual function evaluations is chosen as the primary index tool. This is due to the fact that the number of function evaluations indirectly indicates the CPU time usage.

The surrogates employed in order to solve the examples have been illustrated in Table 3. The abbreviations of the surrogate models mentioned in Table 3, have been utilized throughout the remaining paper.

4.1 Example 1: Test Function [21]

The first example considered is RDO of a test function [21]. The description of the problem has been stated as:

$$\begin{aligned}
&\text{Minimize } F = \frac{\sigma_f}{\sigma_f^*}\\
&\text{Subjected to } G = \mu_g - k\sigma_g \geq 0\\
&\text{where } f(x_1,x_2) = (x_1-4)^3 + (x_1-3)^4 + (x_2-5)^2 + 10\\
&g(x_1,x_2) = x_1 + x_2 - 6.45\\
&1 \leq \mu_{x_1} \leq 10, 1 \leq \mu_{x_2} \leq 10
\end{aligned} \tag{54}$$

The objective is to minimize the standard deviation of f with a probabilistic constraint on g. σ_f^* and k have been adopted to be 15 and 3, respectively. The design variables x_1 and x_2 follow normal distribution with standard deviation 0.4.

The number of sample points utilized for comparison of various methods have been presented in Table 4. The corresponding robust optimal solutions obtained have been reported in Table 5. The number of iterations and function calls required for yielding the optimal solutions have been presented in Fig. 4.

4.2 Example 2: Two Bar Planar Truss [21]

RDO of a planar two bar truss [21] has been considered as the second example. The problem consists of two design variables, which are the cross sectional area x_1 and the horizontal span of each truss x_2. The density of bar material ρ, the magnitude of the applied load Q, and the material's tensile strength S are the other parameters of the problem. The objective is to minimize the volume of structure subject to constraints on axial strength of each of the members. The description of the deterministic optimization can be stated as:

$$\begin{aligned}
&\text{Minimize } f(x_1,x_2) = \rho x_1\sqrt{1+x_2^2}\\
&\text{Subjected to}\\
&g_1(x_1,x_2) = 1 - \frac{5Q}{\sqrt{65}S}\sqrt{1+x_2^2}\left(\frac{8}{x_1} + \frac{1}{x_1x_2}\right) \geq 0\\
&g_1(x_1,x_2) = 1 - \frac{5Q}{\sqrt{65}S}\sqrt{1+x_2^2}\left(\frac{8}{x_1} + \frac{1}{x_1x_2}\right) \geq 0\\
&0.2 \leq x_1 \leq 20, 0.1 \leq x_2 \leq 1.6
\end{aligned} \tag{55}$$

The values of the other parameters ρ, Q and S are 10^4 kg/m^3, 800 kN and 1050 MPa, respectively. The RDO formulation has been presented in Eq. (56).

$$\begin{aligned}
&\text{Minimize } F = w_1\frac{\mu_f}{\mu_f^*} + w_2\frac{\sigma_f}{\sigma_f^*}\\
&\text{Subjected to } G_1 = \mu_{g_1} - k\sigma_{g_1} \geq 0\\
&G_2 = \mu_{g_2} - k\sigma_{g_2} \geq 0\\
&0.2 \leq \mu_{x_1} \leq 20,\ 0.1 \leq \mu_{x_2} \leq 1.6
\end{aligned} \tag{56}$$

The weighing factors w_1 and w_2 have been adopted to be 0.5. μ_f^*, σ_f^* and k have been set as 10, 2 and 3, respectively. The description of the random variables has been presented in Table 6. The coefficient of variation of the two design variables is 0.02.

The number of sample points utilized for comparison of various methods have been presented in Table 7. The

Table 4 Number of sample points utilized for example 1 (Sect. 4.1)

Function	MCS	Anchored ANOVA decomposition	Other surrogate models[a]
f	10^5	25	32
g	10^5	25	32

[a]Except quadrature based PCE (PCE-Q)

Table 5 Robust optimal solutions for example 1 (Sect. 4.1)

Approach	Design variables		Objective function
	x_1	x_2	
MCS	3.3622	4.9941	0.0748
ANOVA-D	3.3629	4.9991	**0.0748**
UK	3.3384	4.9161	0.0834
FK	3.3541	4.9213	0.0832
RBF	3.3616	4.9533	0.0762
PCE-Q	3.3571	5.0008	0.0755
PCE-OLS	3.3572	5.001	0.0755
PCE-LAR	3.357	5.0009	0.0755
LWP	3.3535	4.9477	0.0764
ANN	3.448	5.0793	0.1113
MARS	3.2562	4.9152	0.0779
SVM	4.5073	5.3833	0.2643

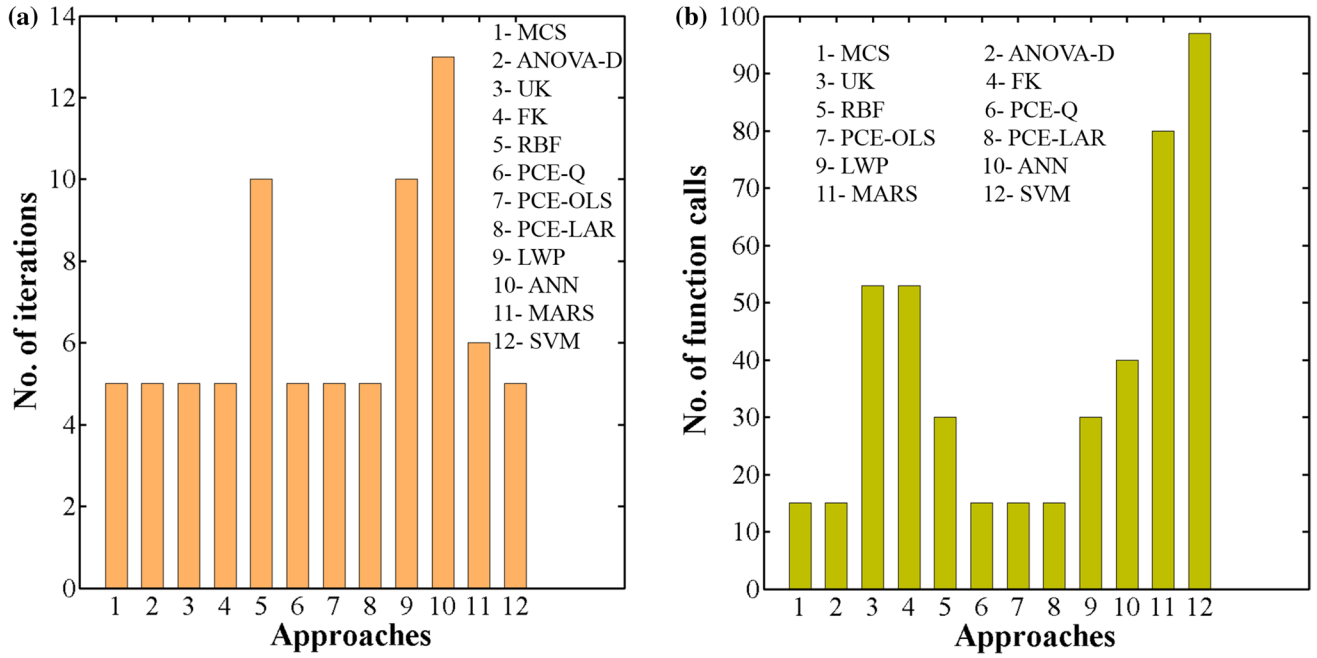

Fig. 4 Comparative assessment of the surrogate models for example 1 in terms of (**a**) number of iterations (**b**) number of function calls, required in yielding the optimal solutions. The number of function calls refer to the number of function evaluations in the optimization loop. The results obtained by MCS are also provided

Table 6 Description of the random variables

Random variable	Distribution	Mean	Standard deviation
x_1	Normal	μ_{x_1}	$0.02\mu_{x_1}$
x_2	Normal	μ_{x_2}	$0.02\mu_{x_2}$
ρ	Beta[a]	10,000	2000
Q	Gumbel	800	200
S	Lognormal	1050	250

[a]Parameters for beta distribution have been set as $\gamma = \eta = 5$

Table 7 Number of sample points utilized for solving the example 2 (Sect. 4.2)

Function	MCS	Anchored ANOVA decomposition	Other surrogate models[a]
f	10^5	25	64
g	10^5	49	128

[a]Except quadrature based PCE (PCE-Q)

corresponding robust optimal solutions obtained have been reported in Table 8. The number of iterations and function calls required for yielding the optimal solutions have been presented in Fig. 5.

Table 8 Robust optimal solutions for example 2 (Sect. 4.2)

Approach	Design variables		Objective function ($\times 10^4$)
	x_1	x_2	
MCS	11.6798	0.3771	1.2516
ANOVA-D	11.7409	0.3771	1.2582
UK	12.9081	0.3774	1.3834
FK	11.9496	0.3781	1.2813
RBF	11.7907	0.3783	1.2662
PCE-Q	11.6765	0.3771	**1.2513**
PCE-OLS	11.6964	0.377	1.2535
PCE-LAR	11.6946	0.3768	1.2533
LWP	12.3217	0.3764	1.3201
ANN	15.6275	0.6233	1.8465
MARS	10.0795	0.3246	1.0639
SVM	8.4505	0.1048	8.4673

4.3 Example 3: Bulk Carrier Design [16]

The third example considered is that of an RDO of a bulk carrier [16]. The basic cost function of the optimization problem has been considered to be the unit transportation cost. The six design variables have been described in Table 9. The formulation involves some design constraints and have been constructed based on geometry, stability and model validity.

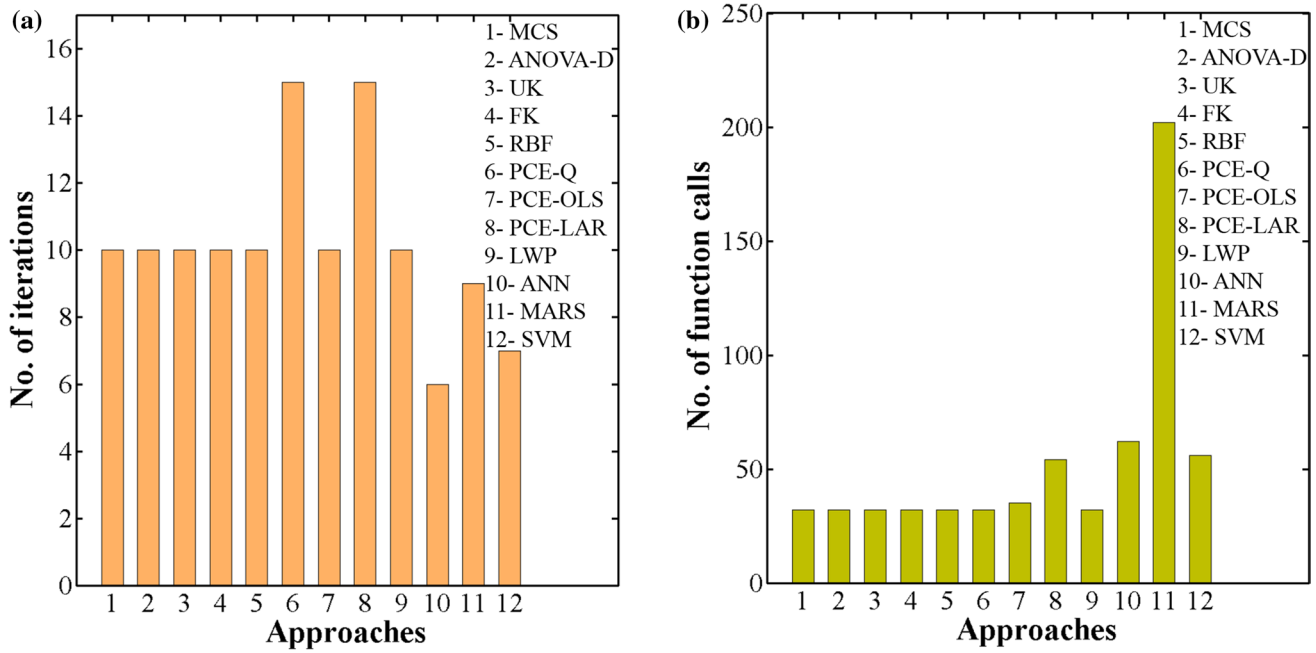

Fig. 5 Comparative assessment of the surrogate models for example 2 in terms of (**a**) number of iterations (**b**) number of function calls, required in yielding the optimal solutions. The number of function calls refer to the number of function evaluations in the optimization loop. The results obtained by MCS are also provided

Table 9 Description of design variables

S. no.	Variable	Symbol	Unit	Lower bound	Upper bound
1	Length	L	m	100	600
2	Beam	B	m	10	100
3	Depth	D	m	5	30
4	Draft	T	m	5	30
5	Block coefficient	C		0.63	0.75
6	Cruise speed	V	knots	14	18

The mathematical model of the cost function has been discussed briefly.

$$\text{Annual cost} = \text{capital costs} + \text{running costs} + \text{voyage costs} \tag{57}$$

$$\text{Capital costs} = 0.2\ (\text{ship costs}) \tag{58}$$

$$\text{Ship cost} = 1.3(2000W_S^{0.85} + 3500W_0 + 2400P^{0.8}) \tag{59}$$

$$\text{Steel weight} = W_S = 0.034L^{1.7}B^{0.7}D^{0.4}C^{0.5} \tag{60}$$

$$\text{Outfit weight} = W_0 = L^{0.8}B^{0.6}D^{0.3}C^{0.1} \tag{61}$$

$$\text{Machinery weight} = W_m = 0.17P^{0.9} \tag{62}$$

$$\text{Displacement} = 1.025LBTC \tag{63}$$

$$\text{Power} = P = \text{displacement}^{2/3}V^3/(a + bF_n) \tag{64}$$

$$\text{Froude number} = F_n = V_k/(gL)^{0.5} \tag{65}$$

$$V_k = 0.5144\ \text{V} \tag{66}$$

With the help of Eq. (66), V_k has units of m/s and $g = 9.8065\ \text{m/s}^2$ in Eq. (65).

$$a = 4977.06C^2 - 8105.61C + 4456.51 \tag{67}$$

$$b = -10847.2C^2 + 12817C - 6960.32 \tag{68}$$

$$\text{Running costs} = 40{,}000DWT^{0.3} \tag{69}$$

$$\text{Deadweight} = DWT = \text{displacement} - \text{light ship weight} \tag{70}$$

$$\text{Light ship weight} = W_S + W_0 + W_m \tag{71}$$

$$\text{Voyage costs} = (\text{fuel cost} + \text{port cost})RTPA \tag{72}$$

$$\text{Fuel cost} = 1.05\ \text{daily consumption} \times \text{sea days} \times \text{fuel price} \tag{73}$$

$$\text{Daily consumption} = 0.19P24/1000 + 0.2 \tag{74}$$

$$\text{Sea days} = \text{round trip miles}/24V \tag{75}$$

$$\text{Port cost} = 6.3DWT^{0.8} \tag{76}$$

$$\text{Round trips per year} = RTPA = 350/(\text{sea days} + \text{port days}) \tag{77}$$

Table 10 Description of uncertain variables

S. no.	Parameter	Unit	Distribution type	Lower bound	Upper bound
1	Port handling rate	Ton/day	Uniform	1000	11,000
2	Round trip	nm	Uniform	1000	5000
3	Fuel price	GBP/ton	Uniform	50	150

Table 11 Description of objective functions for example 3 defined in Sect. 4.3

S. no.	Objective function
1	Standard deviation
2	Mean + standard deviation

Table 12 Number of sample points utilized for solving the example 3 (Sect. 4.3)

Function	MCS	Anchored ANOVA decomposition	Other surrogate models[a]
Objective function	10^5	127	128

[a]Except quadrature based PCE (PCE-Q)

$$\text{Port days} = 2\left[(\text{cargo deadweight}/\text{handling rate}) + 0.5\right] \tag{78}$$

$$\text{Cargo deadweight} = DWT - \text{fuel carried} - \text{miscellaneous } DWT \tag{79}$$

$$\text{Fuel carried} = \text{daily consumption (sea days} + 5) \tag{80}$$

$$\text{Miscellaneous } DWT = 2DWT^{0.5} \tag{81}$$

$$\text{Annual cargo capacity} = DWT \times \text{round trips per year} \tag{82}$$

$$\text{Unit transportation cost} = \text{annual cost/annual cargo capacity} \tag{83}$$

The unit transportation cost has been adopted to be the objective function of the optimized conceptual design of a bulk carrier and can be evaluated using Eq. (83). Since the design problem incorporates several environmental factors and involves detailed modelling, large number of parameters have been involved. Therefore, to maintain disambiguity, all parameters required for evaluating Eq. (83) have been well defined in Eqs. (57)–(82). The constraints pertaining to the optimization problem have been defined in Eqs. (84)–(91).

$$L/B \geq 6 \tag{84}$$

$$L/D \leq 15 \tag{85}$$

$$L/T \leq 19 \tag{86}$$

$$T \leq 0.45DWT^{0.31} \tag{87}$$

$$T \leq 0.7D + 0.7 \tag{88}$$

$$25,000 \leq DWT \leq 5,00,000 \tag{89}$$

$$F_n \leq 0.32 \tag{90}$$

$$GMT = KB + BMT - KG \leq 0.07B \tag{91}$$

where KB, BMT and KG have been defined in Eqs. (92)–(94), respectively.

$$\text{Vertical center of buoyancy} = KB = 0.53T \tag{92}$$

$$\text{Metacentric radius} = BMT = (0.085C - 0.002)B^2/TC \tag{93}$$

Table 13 Robust optimal solutions corresponding to case 1 of Table 11 for example 3

Approach	Design variables						Objective function
	L	B	D	T	C	V	
MCS	151.1748	25.1958	14.9491	10.389	0.75	14	2.3704
ANOVA-D	151.1748	25.1958	14.9491	10.389	0.75	14	**2.3705**
UK	151.9329	25.3222	16.1029	10.389	0.7473	14	2.4395
FK	164.7326	27.4554	16.4	10.389	0.6487	14	1.9441
RBF	165.7109	27.6185	16.5099	10.3890	0.6424	14	2.0618
PCE-Q	180.0045	28.5101	17.0979	11.1554	0.6649	14	3.5237
PCE-OLS	180.0045	28.5101	17.0979	11.1554	0.6649	14	3.5822
PCE-LAR	180.0045	28.5101	17.0979	11.1554	0.6649	14	3.5909
LWP	162.6665	27.1111	16.1701	10.389	0.6624	14	2.1427
ANN	179.9694	26.7837	19.2892	10.389	0.63	14	2.7193
MARS	179.8739	26.7022	18.2304	10.3705	0.63	14	2.452
SVM	183.1758	28.9947	17.4435	11.0532	0.6807	15.3888	0.9631

Table 14 Robust optimal solutions corresponding to case 2 of Table 11 for example 3

Approach	Design variables						Objective function
	L	*B*	*D*	*T*	*C*	*V*	
MCS	153.5981	25.5997	15.1984	10.389	0.7298	14	10.9932
ANOVA-D	153.2929	25.5471	15.1657	10.389	0.7323	14	11.1545
UK	154.2596	25.7099	15.2672	10.389	0.7244	14	**10.9477**
FK	161.9953	23.6782	14.0228	10.389	0.7495	14	10.2549
RBF	154.0877	25.68128	15.24929	10.389	0.725811	14	10.5391
PCE-Q	180.0045	28.51008	17.0979	11.1554	0.66491	14	12.0307
PCE-OLS	180.0045	28.51008	17.0979	11.1554	0.66491	14	12.1421
PCE-LAR	180.0045	28.51008	17.0979	11.1554	0.66491	14	12.1441
LWP	154.5059	25.75098	15.29291	10.389	0.722457	14	10.526
ANN	166.8899	23.6186	13.9817	10.389	0.7334	14	10.9417
MARS	153.9085	25.65141	15.23064	10.389	0.727256	14	10.7451
SVM	184.08	28.35978	17.20125	11.1232	0.695016	15.0315	10.2736

$$\text{Vertical center of gravity} = KG = 1 + 0.52D \tag{94}$$

The description of random variables have been provided in Table 10. Two case studies have been performed for different configurations of objective function as presented in Table 11. The number of sample points utilized for comparison of various methods have been presented in Table 12. The robust optimal solutions corresponding to the case studies undertaken have been reported in Tables 13 and 14. The number of iterations and function calls required for yielding the optimal solutions have been presented in Fig. 6.

4.4 Example 4: Welded Beam Design [128]

The fourth example considered is that of a welded beam design [128]. The objective is to minimize the cost of the beam subject to constraints on shear stress, bending stress, buckling load, and end deflection. There are four continuous design variables, namely, beam thickness x_1, beam width x_2, weld length x_3, and weld thickness x_4.

The problem description can be stated as follows:

$$\text{Minimize} \quad f(\mathbf{x}) = 1.10471x_1^2x_2 + 0.04811x_3x_4\left(14 + x_2\right) \tag{95}$$

s.t.

$$\begin{aligned} g_1(\mathbf{x}) &= t - t_{\max} \le 0 \\ g_2(\mathbf{x}) &= s - s_{\max} \le 0 \\ g_3(\mathbf{x}) &= x_1 - x_4 \le 0 \\ g_4(\mathbf{x}) &= d - d_{\max} \le 0 \\ g_5(\mathbf{x}) &= P - P_c \le 0 \end{aligned} \tag{96}$$

where

$$M = P\left(L + x_2/2\right) \tag{97}$$

$$R = \sqrt{0.25\left(x_2^2 + \left(x_1 + x_3\right)^2\right)} \tag{98}$$

$$J = \sqrt{2}x_1x_2\left(x_2^2\Big/12 + 0.25\left(x_1 + x_3\right)^2\right) \tag{99}$$

$$P_c = 64746.022\left(1 - 0.0282346x_3\right)x_3x_4^3 \tag{100}$$

$$t_1 = P\Big/\left(\sqrt{2}x_1x_2\right) \tag{101}$$

$$t_2 = MR/J \tag{102}$$

$$t = \sqrt{t_1^2 + t_1t_2x_2/R + t_2^2} \tag{103}$$

$$S = 6PL\Big/\left(x_4x_3^2\right) \tag{104}$$

$$d = 2.1952\Big/\left(x_4x_3^3\right) \tag{105}$$

$$\begin{aligned} &P = 6000,\ L = 14, E = 30 \times 10^6,\ G = 12 \times 10^6, \\ &t_{\max} = 13,600,\ s_{\max} = 30,000,\ x_{\max} = 10,\ d_{\max} = 0.25 \\ &0.125 \le x_1 \le 10,\ 0.1 \le x_i \le 10, \text{for } i = 2, 3, 4. \end{aligned} \tag{106}$$

For the RDO formulation of the problem, each of the design variables have been assumed to be normally distributed with standard deviation to be 5%. Case study has been performed considering objective function to be: $\text{mean}(f(\mathbf{x})) + \text{SD}(f(\mathbf{x}))$, where SD is standard deviation.

The number of sample points utilized for comparison of various methods have been presented in Table 15. The corresponding robust optimal solutions obtained have been reported in Table 16. The number of iterations and function calls required for yielding the optimal solutions have been presented in Fig. 7.

4.5 Example 5: Speed Reducer [129]

The fifth example considered is that of speed reducer, which is a standard optimization problem. The details of theoretical

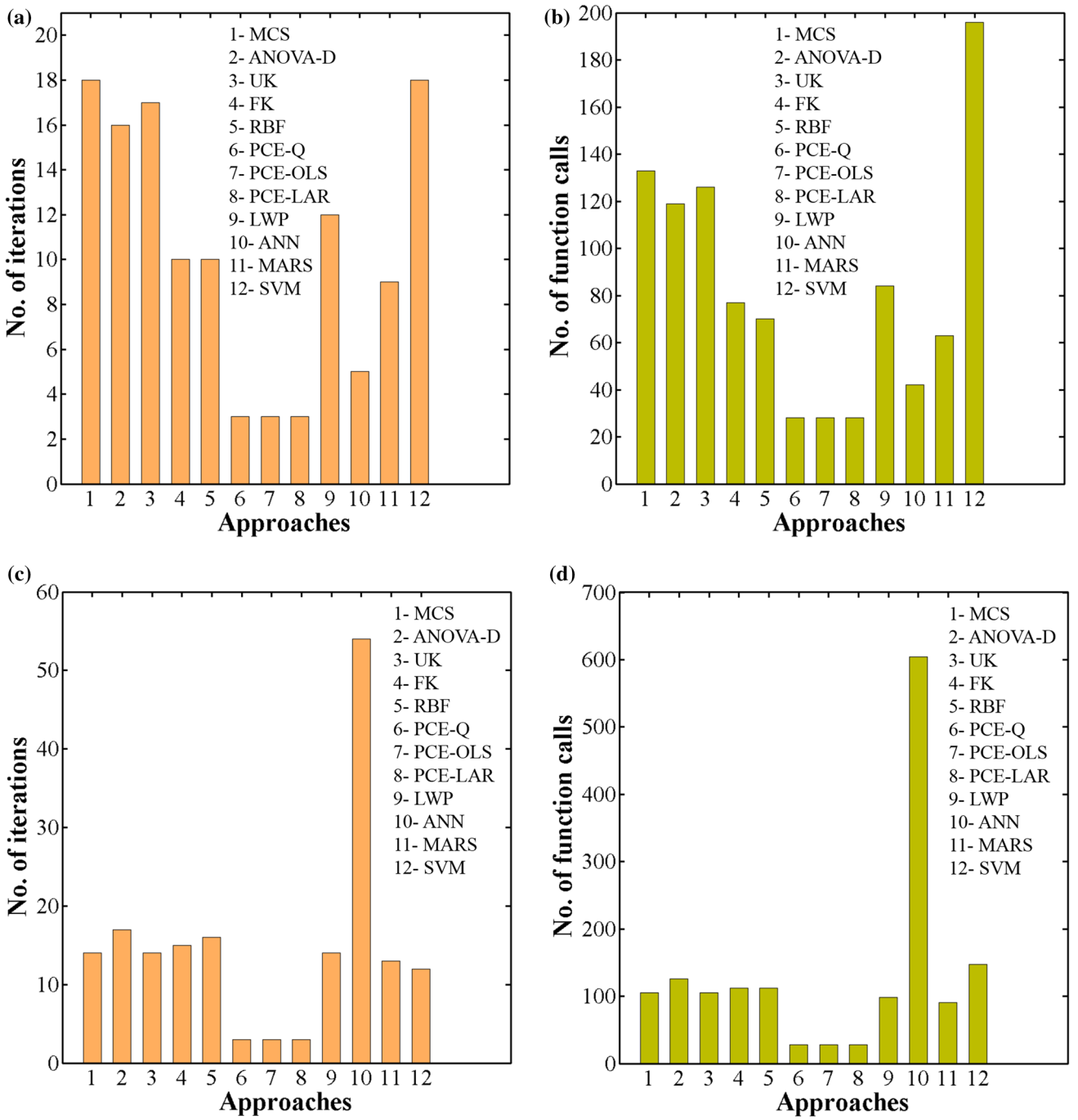

Fig. 6 Comparative assessment of the surrogate models for example 3 in terms of (**a**) number of iterations (**b**) number of function calls, required in yielding the optimal solutions, corresponding to case 1 of Table 11, (**c**) number of iterations (**d**) number of function calls, required in yielding the optimal solutions, corresponding to case 2 of Table 11. The number of function calls refer to the number of function evaluations in the optimization loop. The results obtained by MCS are also provided

formulation can be found elsewhere [129]. The problem consists of seven design variables and eleven constraint functions. The mathematical description of the problem has been presented below.

$$\text{Minimize } f(\mathbf{x}) = 0.7854x_1x_2^2A - 1.508x_1B + 7.477C + 0.7854D \tag{107}$$

where, $A = 3.3333x_3^2 + 14.9334x_3 - 43.0934$

$B = x_6^2 + x_7^2$

$C = x_6^3 + x_7^3$ (108)

$D = x_4x_6^2 + x_5x_7^2$

Table 15 Number of sample points utilized for solving the example 4 (Sect. 4.4)

Function	MCS	Anchored ANOVA decomposition	Other surrogate models[a]
f	10^4	113	128
g_1	10^4	61	128
g_2	10^4	25	128
g_3	10^4	25	128
g_4	10^4	25	128
g_{5ss}	10^4	25	128

[a]Except quadrature based PCE (PCE-Q)

Table 16 Robust optimal solutions for example 4 (Sect. 4.4)

Approach	Design variables				Objective function
	x_1	x_2	x_3	x_4	
MCS	0.2312	6.7505	8.7636	0.2299	2.9059
ANOVA-D	0.2312	6.7479	8.7617	0.2299	**2.905**
UK	0.5852	1.3633	0.4902	6.143	2.9887
FK	0.125	0.8514	2.413	2.4641	4.395
RBF	0.3103	0.1	0.1	10	1.0259
PCE-Q	0.2296	0.1410	8.7935	0.2307	1.6875
PCE-OLS	0.2309	6.7715	8.7635	0.231	2.9114
PCE-LAR	0.2309	6.7614	8.7634	0.231	2.9095
LWP	0.6002	0.3459	0.4499	0.5999	0.3651
ANN	0.1502	0.3122	0.1609	1.3808	0.2094
MARS	3.5638	0.8112	0.3011	3.5635	12.9351
SVM	0.1999	6.0000	7.9999	0.1500	1.8194

$$
\begin{aligned}
\text{s.t.,}\quad & g_1(\mathbf{x}) = \left(27 - x_1x_2^2x_3\right)/27 \le 0\\
& g_2(\mathbf{x}) = \left(397.5 - x_1x_2^2x_3^2\right)/397.5 \le 0\\
& g_3(\mathbf{x}) = \left(1.93 - \left(x_2x_6^4x_3\right)/x_4^3\right)/1.93 \le 0\\
& g_4(\mathbf{x}) = \left(1.93 - \left(x_2x_7^4x_3\right)/x_5^3\right)/1.93 \le 0\\
& g_5(\mathbf{x}) = \left(\left(A_1/B_1\right) - 1100\right)/1100 \le 0\\
& g_6(\mathbf{x}) = \left(\left(A_2/B_2\right) - 850\right)/850 \le 0\\
& g_7(\mathbf{x}) = \left(x_2x_3 - 40\right)/40 \le 0\\
& g_8(\mathbf{x}) = \left(5 - \left(x_1/x_2\right)\right)/5 \le 0\\
& g_9(\mathbf{x}) = \left(\left(x_1/x_2\right) - 12\right)/12 \le 0\\
& g_{10}(\mathbf{x}) = \left(1.9 + 1.5x_6 - x_4\right)/1.9 \le 0\\
& g_{11}(\mathbf{x}) = \left(1.9 + 1.1x_7 - x_5\right)/1.9 \le 0
\end{aligned}
\tag{109}
$$

$$
\begin{aligned}
\text{where, } A_1 &= \left[\left(745x_4/\left(x_2x_3\right)\right)^2 + \left(16.91\times10^6\right)\right]^{0.5}\\
B_1 &= 0.1x_6^3\\
A_2 &= \left[\left(745x_5/\left(x_2x_3\right)\right)^2 + \left(157.5\times10^6\right)\right]^{0.5}\\
B_2 &= 0.1x_7^3
\end{aligned}
\tag{110}
$$

The variable bounds have been presented in Eq. (111).

$$
\begin{aligned}
&2.6 \le x_1 \le 3.6\\
&0.7 \le x_2 \le 0.8\\
&17 \le x_3 \le 28\\
&7.3 \le x_4, x_5 \le 8.3\\
&2.9 \le x_6 \le 3.9\\
&5 \le x_7 \le 5.5
\end{aligned}
\tag{111}
$$

For the RDO formulation of the problem, each of the design variables have been assumed to be normally distributed with standard deviation to be 5%. Case studies have been performed considering various objective functions as shown in Table 17.

The number of sample points utilized for comparison of various methods have been presented in Table 18. The corresponding robust optimal solutions obtained have been reported in Tables 19 and 20. The number of iterations and function calls required for yielding the optimal solutions have been presented in Fig. 8.

4.6 Example 6: Side Impact Crashworthiness of Car [130]

The sixth example considered is that of an RDO of side impact crashworthiness of car [130]. This example has been reformulated as a robust optimization problem. The problem consists of eleven stochastic variables and nine variables out of them are design variables. The description of the stochastic and design variables have been provided in Tables 21 and 22.

The optimization problem formulation can be stated as:

$$
\text{Minimize } f(\mathbf{x}) = \text{Weight} \tag{112}
$$

$$
\begin{aligned}
\text{s.t.,}\quad & g_1(\mathbf{x}) = \text{Abdomen load} \le 1 \text{ kN}\\
& g_2(\mathbf{x}) = V \times C_u \le 0.32 \text{ m/s}\\
& g_3(\mathbf{x}) = V \times C_m \le 0.32 \text{ m/s}\\
& g_4(\mathbf{x}) = V \times C_l \le 0.8 \text{ m/s}\\
& g_5(\mathbf{x}) = \text{upper rib deflection} \le 32 \text{ mm}\\
& g_6(\mathbf{x}) = \text{middle rib deflection} \le 32 \text{ mm}\\
& g_7(\mathbf{x}) = \text{lower rib deflection} \le 32 \text{ mm}\\
& g_8(\mathbf{x}) = \text{pubic force} \le 4 \text{ kN}\\
& g_9(\mathbf{x}) = \text{vel. of V-pillar at midpoint} \le 9.9 \text{ mm/ms}\\
& g_8(\mathbf{x}) = \text{front door vel. at B - Pillar} \le 15.7 \text{ mm/ms}
\end{aligned}
\tag{113}
$$

The functional forms of the objective and constraint functions have been provided in Eqs. (114)–(124).

$$
\begin{aligned}
f(\mathbf{x}) = {} & 1.98 + 4.9x_1 + 6.67x_2 + 6.98x_3 + 4.01x_4\\
& + 1.78x_5 + 0.00001x_6 + 2.73x_7
\end{aligned}
\tag{114}
$$

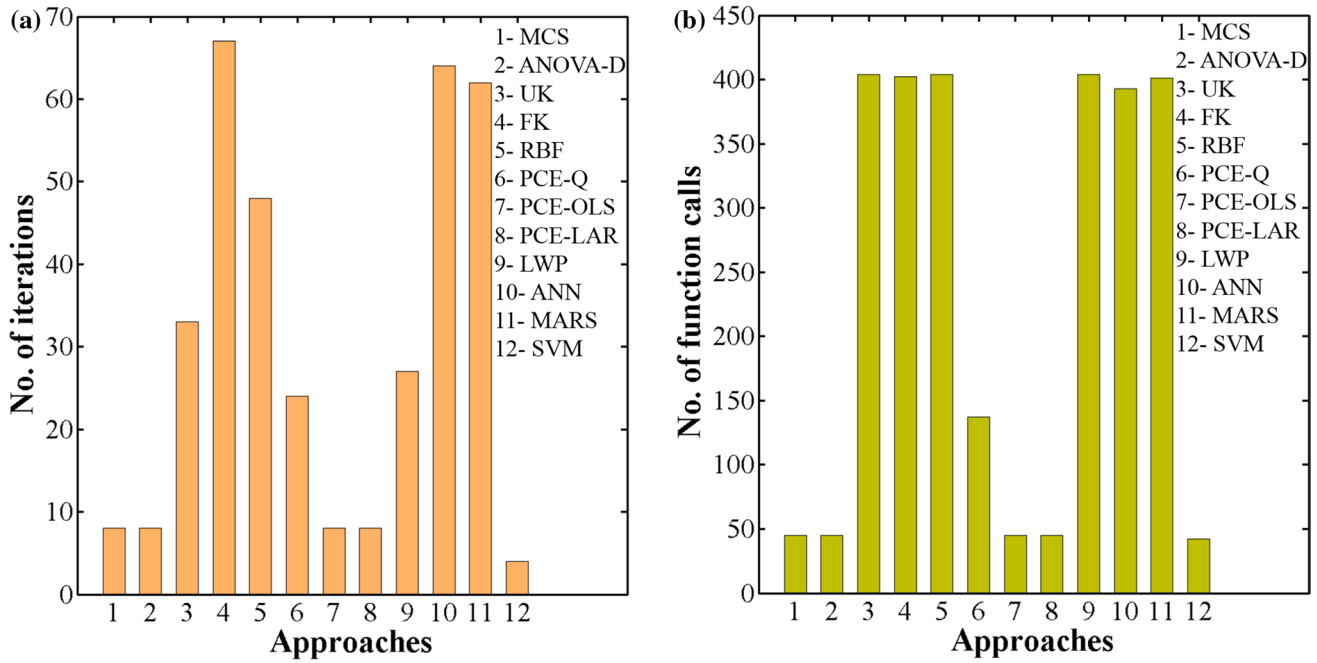

Fig. 7 Comparative assessment of the surrogate models for example 4 in terms of (**a**) number of iterations (**b**) number of function calls, required in yielding the optimal solutions. The number of function calls refer to the number of function evaluations in the optimization loop. The results obtained by MCS are also provided

Table 17 Description of objective functions for example 5 (Sect. 4.5)

S. no.	Objective function
1	Standard deviation
2	Mean + Standard deviation

Table 18 Number of sample points utilized for solving the example 5 (Sect. 4.5)

Function	MCS	Anchored ANOVA decomposition	Other surrogate models[a]
f	10^4	99	128
g_1	10^4	19	128
g_2	10^4	19	128
g_3	10^4	33	128
g_4	10^4	33	128
g_5	10^4	33	128
g_6	10^4	33	128
g_7	10^4	9	128
g_8	10^4	9	128
g_9	10^4	9	128
g_{10}	10^4	9	128
g_{11}	10^4	9	128

[a]Except quadrature based PCE (PCE-Q)

$$g_1(\mathbf{x}) = 1.16 - 0.3717x_2x_4 - 0.00931x_2x_{10} - 0.484x_3x_9 + 0.01343x_6x_{10} \quad (115)$$

$$g_2(\mathbf{x}) = 0.261 - 0.0159x_1x_2 - 0.188x_1x_8 - 0.019x_2x_7 + 0.0144x_3x_5 + 0.0008757x_5x_{10} + 0.08045x_6x_9 + 0.00139x_8x_{11} + 0.00001575x_{10}x_{11} \quad (116)$$

$$g_3(\mathbf{x}) = 0.214 + 0.00817x_5 - 0.131x_1x_8 - 0.0704x_1x_9 + 0.03099x_2x_6 - 0.018x_2x_7 + 0.0208x_3x_8 + 0.121x_3x_9 - 0.00364x_5x_6 + 0.0007715x_5x_{10} - 0.0005354x_6x_{10} + 0.00121x_8x_{11} + 0.00184x_9x_{10} - 0.018x_2^2 \quad (117)$$

$$g_4(\mathbf{x}) = 0.74 - 0.61x_2 - 0.163x_3x_8 + 0.001232x_3x_{10} - 0.166x_7x_9 + 0.227x_2^2 \quad (118)$$

$$g_5(\mathbf{x}) = 28.98 + 3.818x_3 - 4.2x_1x_2 + 0.0207x_5x_{10} + 6.63x_6x_9 - 7.77x_7x_8 + 0.32x_9x_{10} \quad (119)$$

$$g_6(\mathbf{x}) = 33.86 + 2.95x_3 + 0.1792x_{10} - 5.057x_1x_2 - 11x_2x_8 - 0.0215x_5x_{10} - 9.98x_7x_8 + 22x_8x_9 \quad (120)$$

$$g_7(\mathbf{x}) = 46.36 - 9.9x_2 - 12.9x_1x_8 + 0.1107x_3x_{10} \quad (121)$$

Table 19 Robust optimal solutions corresponding to case 1 of Table 17 for example 5 (Sect. 4.5)

Approach	Design variables							Objective function
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
MCS	3.4863	0.7	17	7.3	7.7158	3.3515	5.2867	230.027
ANOVA-D	3.4859	0.7	17	7.3	7.7158	3.3515	5.2867	**230.0031**
UK	3.0026	0.7503	22.9953	7.7987	7.7951	3.3983	5.2507	370.3597
FK	3.4899	0.7002	17.3942	8.2563	7.8465	3.3575	5.3175	237.7963
RBF	3.0099	0.7	23.2772	7.7445	7.8474	3.4012	5.3505	352.9711
PCE-OLS	3.4818	0.7	17	7.3	7.7165	3.3522	5.2877	228.5885
PCE-LAR	3.4818	0.7	17	7.3	7.7165	3.3522	5.2877	228.5851
LWP	3.0630	0.7	22.23509	7.5397	7.9276	3.5978	5.4790	332.1858
ANN	3.0231	0.7491	21.8341	7.8765	7.8096	3.3523	5.4985	338.3447
MARS	3.1384	0.7444	21.4819	7.6380	7.9358	3.3257	5.4794	338.2782
SVM	3	0.75	23	7.8	7.8	3.4	5.25	0.4839

Quadrature based PCE (PCE-Q) did not converge

Table 20 Robust optimal solutions corresponding to case 2 of Table 17 for example 5 (Sect. 4.5)

Approach	Design variables							Objective function
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
MCS	3.4863	0.7	17	7.3	7.7158	3.3515	5.2867	3232.5
ANOVA-D	3.4859	0.7	17	7.3	7.7158	3.3515	5.2867	**3232.4**
UK	2.4807	0.7982	19.9598	7.719	8.2962	3.4142	5.374	4090.3
FK	2.6243	0.7	18.8062	7.9017	7.5009	3.836	5.0891	3156.1
RBF	2.8906	0.7	25.0018	7.7546	7.6379	3.5215	5.2159	4543.7
PCE-OLS	3.4818	0.7	17	7.3	7.7165	3.3522	5.2877	3225.7
PCE-LAR	3.4818	0.7	17	7.3	7.7165	3.3522	5.2877	3225.7
LWP	3.553	0.7625	18.9213	7.4558	7.9566	3.5702	5.4763	4253.3
ANN	3.5576	0.7034	23.1254	7.9142	7.6379	3.3776	5.2197	4685.1
MARS	2.9308	0.7467	23.5264	7.7401	8.2991	3.3556	5.3752	4673.2
SVM	3	0.75	23	7.8	7.8	3.3999	5.2499	4185.5

Quadrature based PCE (PCE-Q) did not converge

$$g_8(\mathbf{x}) = 4.72 - 0.5x_4 - 0.19x_2x_3 - 0.0122x_4x_{10} + 0.009325x_6x_{10} + 0.000191x_{11}^2 \quad (122)$$

$$g_9(\mathbf{x}) = 10.58 - 0.674x_1x_2 - 1.95x_2x_8 + 0.02054x_3x_{10} - 0.0198x_4x_{10} + 0.028x_6x_{10} \quad (123)$$

$$g_{10}(\mathbf{x}) = 16.45 - 0.489x_3x_7 - 0.843x_5x_6 + 0.0432x_9x_{10} - 0.0556x_9x_{11} - 0.000786x_{11}^2 \quad (124)$$

Case studies have been performed considering various objective functions as shown in Table 23. The number of sample points utilized for comparison of various methods have been presented in Table 24. The corresponding robust optimal solutions obtained have been reported in Tables 25 and 26. The number of iterations and function calls required for yielding the optimal solutions have been presented in Fig. 9.

4.7 Results and Discussion

The results obtained for each of the six problems by utilizing the surrogate models have been discussed in this section. This discussion has been presented in order to provide guidance to the users about the appropriateness of a particular surrogate model to be utilized in a specific problem. For each of the case studies performed, the best performing surrogate model (in terms of closeness to MCS based solutions) has been marked in bold.

Firstly, in case of example 1, ANOVA-D performs excellently, almost exactly matching the MC based optimal solutions. Apart from ANN and SVM, all other models exhibit acceptable performance in terms of approximation accuracy. ANOVA-D and the PCE based models have been observed to converge in less number of iterations and function calls, in comparison to the other surrogate models. In example 2, ANOVA-D, FK, RBF and the PCE based approaches yield strikingly similar results as compared to

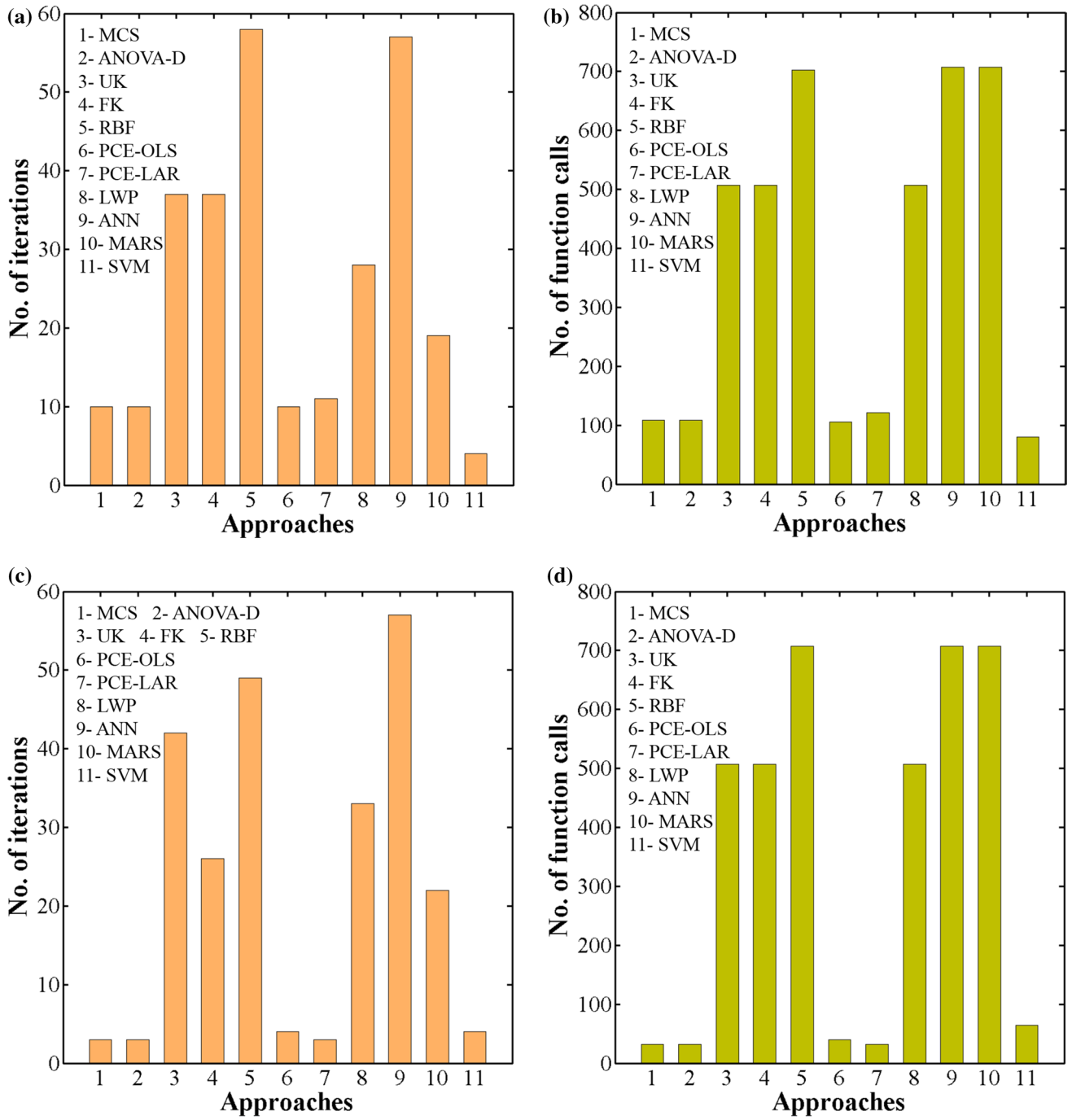

Fig. 8 Comparative assessment of the surrogate models for example 5 in terms of (**a**) number of iterations (**b**) number of function calls, required in yielding the optimal solutions, corresponding to case 1 of Table 17, (**c**) number of iterations (**d**) number of function calls, required in yielding the optimal solutions, corresponding to case 2 of Table 17. The number of function calls refer to the number of function evaluations in the optimization loop. The results obtained by MCS are also provided

MCS based RDO. Also, the above models yield results in relatively less number of iterations as compared to the other models, which probably experience delayed convergence due to some false optima at intermediate iterations. In example 3, excellent similar results have been obtained by ANOVA-D and UK as compared to MC based solutions, both in terms of accuracy and rate of convergence. Models such as, RBF, LWP, ANN and MARS yield slightly inaccurate results, however, in relatively less number of iterations as compared to the above ones. In example 4, ANOVA-D, PCE-OLS and PCE-LAR achieve almost exact results as that of MCS, in relatively less number of iterations and

Table 21 General description of the stochastic and design variables

S. no.	Variable	Notation
1	Thickness of B-Pillar inner	x_1
2	Thickness of B-Pillar reinforcement	x_2
3	Thickness of floor side inner	x_3
4	Thickness of cross members	x_4
5	Thickness of door beam	x_5
6	Thickness of door beltline reinforcement	x_6
7	Thickness of roof rail	x_7
8	Material of B-Pillar inner	x_8
9	Material of floor side inner	x_9
10	Barrier height	x_{10}
11	Barrier hitting position	x_{11}

Table 22 Statistical description of stochastic variables

Variable	Distribution	Standard deviation	Lower limit	Upper limit
x_1	Normal	0.03	0.5	1.5
x_2	Normal	0.03	0.5	1.35
x_3	Normal	0.03	0.5	1.5
x_4	Normal	0.03	0.5	1.5
x_5	Normal	0.03	0.5	2.625
x_6	Normal	0.03	0.5	1.2
x_7	Normal	0.03	0.5	1.2
x_8	Normal	0.006	0.192	0.345
x_9	Normal	0.006	0.192	0.345
x_{10}	Normal	10	$x_{10}$and x_{11} are not considered as design variables	
x_{11}	Normal	10		

Table 23 Description of objective functions for example 6 (Sect. 4.6)

S. no.	Objective function
1	Standard deviation
2	Mean + Standard deviation

Table 24 Number of sample points utilized for solving the example 6 (Sect. 4.6)

Function	MCS	Anchored ANOVA decomposition	Other surrogate models[a]
f	10^4	73	75
g_1	10^4	73	205
g_2	10^4	201	205
g_3	10^4	201	205
g_4	10^4	73	205
g_5	10^4	163	205
g_6	10^4	129	205
g_7	10^4	51	205
g_8	10^4	73	205
g_9	10^4	99	205
g_{10}	10^4	99	205

[a]Except quadrature based PCE (PCE-Q)

function calls as compared to other surrogate models. Excellent results in terms of similarity to MCS have been obtained by ANOVA-D, PCE-OLS and PCE-LAR in example 5. The above models outperform the other surrogates not only in response approximation but also in terms of rate of convergence. FK has also achieved decent and satisfactory results, however, it requires significantly higher number of iterations to converge. It is also worth mentioning that PCE-Q did not achieve convergence due to the lack of ability to accurately approximate the response functions in presence of multiple constraints. Exactly same results have been achieved by ANOVA-D as compared to MC based solutions in example 6. However, other models such as, UK, PCE-LAR, LWP, ANN and MARS have performed very well both in terms of approximation accuracy and rate of convergence. It is also worth mentioning that PCE-Q and PCE-OLS did not achieve convergence due to the lack of ability to accurately approximate the response functions in multiple constrained environment.

The results of each of the examples as discussed above illustrate that performance of a surrogate model is very sensitive in an RDO framework, and may easily lead to an incorrect optima. In all of the above examples carried out, only one surrogate model i.e., ANOVA-D, has been consistent in accurately capturing the non-linearity and multi-modal landscapes in presence of constraints. It is also worth mentioning that PCE-LAR, being an adaptive sparse model, has achieved good results in most of the problems. Rest of the models have been observed to perform well in few problems and found unsuitable for other complex non-linear problems. Thus, on the basis of the above results, it is recommended to employ ANOVA-D as a surrogate model for highly complex problems whose response landscapes are difficult to capture and additionally comprising of multiple non-linear constraints. Since ANOVA-D has been observed to perform in a superior manner not only in terms of accuracy but also in convergence rate, therefore, it has been employed to solve a large-scale practical engineering problem in the next section.

5 Practical Problem: RDO of a Hydroelectric Dam Model

Electricity generation using a hydroelectric dam is primarily governed by the hourly water supplied through the

Table 25 Robust optimal solutions corresponding to case 1 of Table 23 for example 6 (Sect. 4.6)

Approach	Design variables									Objective function
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	
MCS	1.006	1.008	1.036	1.081	1	1	1.005	0.337	0.316	0.3591
ANOVA-D	1.006	1.008	1.036	1.081	1	1	1.005	0.337	0.316	**0.3591**
UK	1.006	1.008	1.036	1.081	1	0.999	1.005	0.338	0.317	0.3568
FK	1	1	1	1	1	1	1	0.3	0.3	0.3568
RBF	1.007	1.007	1.035	1.079	1	0.999	1.005	0.338	0.317	0.5232
PCE-LAR	1.006	1.008	1.036	1.083	1	1	1.005	0.337	0.317	0.3578
LWP	1.006	1.008	1.035	1.081	1	0.999	1.005	0.338	0.317	0.3568
ANN	1.006	1.008	1.037	1.082	1	1.001	1.005	0.337	0.317	0.3561
MARS	1.006	1.008	1.035	1.081	1	0.999	1.005	0.338	0.316	0.3568
SVM	1	1	0.999	0.999	0.999	0.999	0.999	0.299	0.3	0.1007

PCE-Q and PCE-OLS did not converge

Table 26 Robust optimal solutions corresponding to case 2 of Table 23 for example 6 (Sect. 4.6)

Approach	Design variables									Objective function
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	
MCS	1.306	0.863	0.5	1.314	0.543	1.2	0.5	0.345	0.345	25.5875
ANOVA-D	1.306	0.863	0.5	1.314	0.543	1.2	0.5	0.345	0.345	**25.5875**
UK	1.306	0.863	0.5	1.313	0.541	1.2	0.5	0.345	0.345	25.5821
FK	0.5	0.506	0.518	1.024	0.943	1.199	0.970	0.345	0.278	20.2156
RBF	1.059	0.974	0.5	1.293	1	0.999	1.2	0.345	0.345	21.9368
PCE-LAR	1.306	0.863	0.5	1.316	0.544	1.2	0.5	0.345	0.345	25.5924
LWP	1.306	0.863	0.5	1.313	0.541	1.2	0.5	0.345	0.345	25.5821
ANN	1.306	0.863	0.5	1.313	0.541	1.2	0.5	0.345	0.345	25.5801
MARS	1.306	0.863	0.5	1.313	0.541	1.2	0.5	0.345	0.345	25.5837
SVM	1.003	0.998	1	1	1	1	0.999	0.298	0.3	29.170

PCE-Q and PCE-OLS did not converge

turbine and the water level in the reservoir. It is quite obvious that due to environmental variations, large amount of uncertainties are associated with a hydroelectric dam. Moreover, cost of energy is also influenced by various factors. Hence, it is of utter importance to consider the presence of uncertainties while optimizing (maximizing) the overall revenue of a hydroelectric dam.

The hydroelectric dam considered in this study as presented in Fig. 10, is such that the water in the reservoir may either leave through the spillway or turbine. It is obvious that the water leaving through the turbine will be utilized in producing electricity. The primary objective of designing a hydroelectric dam is to maximize the revenue generated by selling the electricity. Conventional optimization of the above mentioned hydroelectric dam can be found in [131].

Various uncertainties are associated with any hydroelectric dam. For instance, the flow through spillway and turbine are generally controlled by some machine operated gates. However, it is not possible to exactly control the flow with such machineries and this results in some uncertainties. On the other hand, the in-flow to the reservoir is uncontrolled and hence large sources of uncertainties is associated with this. Moreover, market price of electricity depends on various factors and is highly uncertain. It is to be noted that flow through spillway, flow through turbine, in-flow and market price are generally monitored on an hourly basis. In the present study, the simulation is run for 12 h and hence, the system under consideration involves 48 random variables. A detailed account of the involved uncertain variables have been provided in Table 27.

The electricity produced in a hydroelectric dam depends on two primary parameters, namely amount of water owing through the turbine and the reservoir storage level. The storage of reservoir again depends on the three factors: (a) inflow, (b) flow through turbine and (c) flow through spillway. As the flow through turbine increases, the water in the reservoir decreases. Therefore, it is necessary to compute the optimum flow through the turbine and spillway that maximizes the electricity production. Moreover, certain constraints needs to be considered while solving the optimization problem.

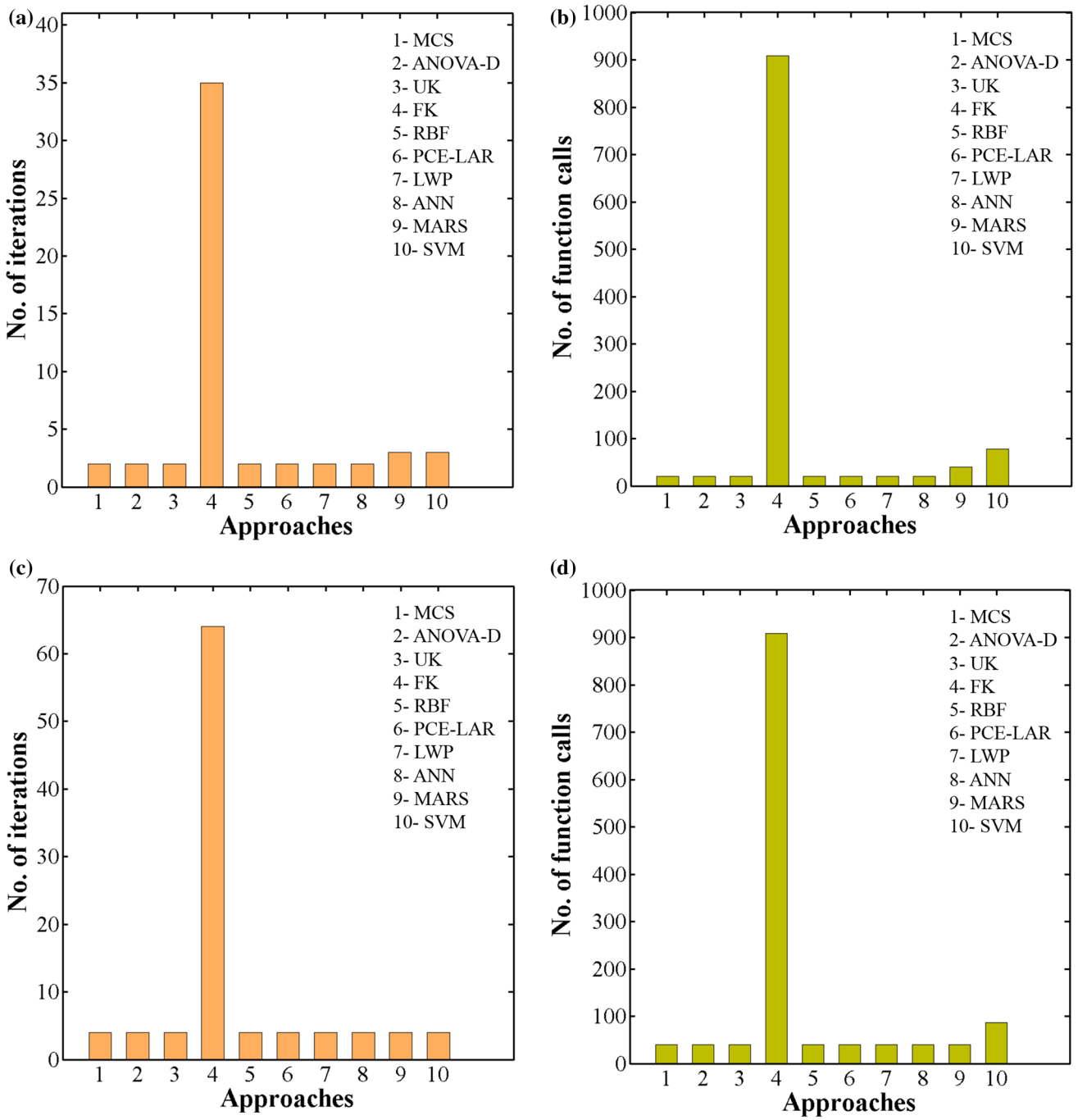

Fig. 9 Comparative assessment of the surrogate models for example 6 in terms of (**a**) number of iterations (**b**) number of function calls, required in yielding the optimal solutions, corresponding to case 1 of Table 23, (**c**) number of iterations (**d**) number of function calls, required in yielding the optimal solutions, corresponding to case 2 of Table 23. The number of function calls refer to the number of function evaluations in the optimization loop. The results obtained by MCS are also provided

First, both reservoir level and downstream flow rates should be within some specified limit. Secondly, maximum flow through the turbine should not exceed the turbine capacity. Finally, the mean reservoir level at the end of the simulation should be same as that at the beginning. This ensures that the reservoir is not emptied at the end of the optimization cycle. The RDO problem has been stated as:

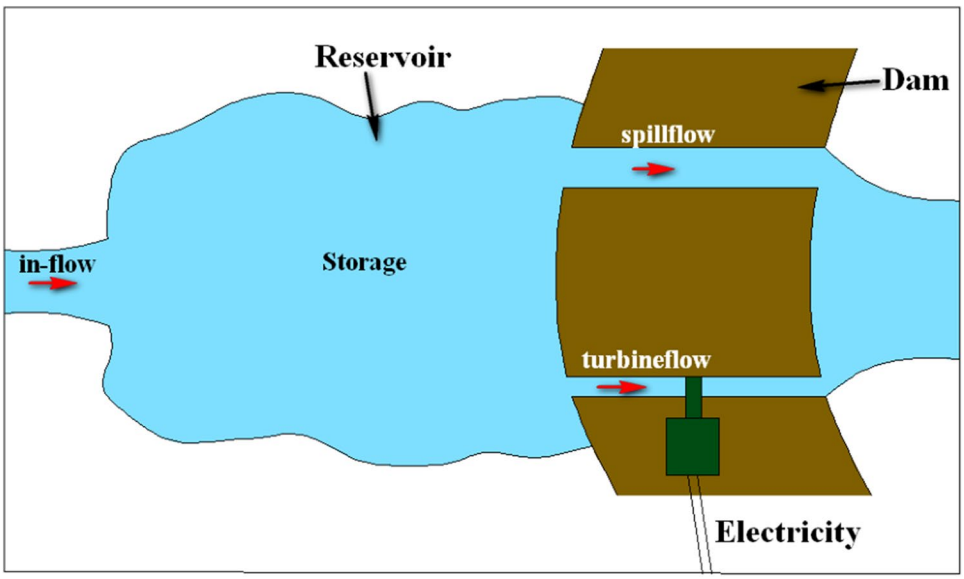

Fig. 10 Schematic diagram of hydroelectric dam

Table 27 Description of the random variables for hydroelectric dam model (Sect. 5)

S. no.	Variable	Distribution	Mean	COV/SD
1–12	Hourly in-flow	Normal	1070 CFS	0.05
13–24	Hourly electricity price	Normal	45 CFS	0.3
25–36	Hourly flow through turbine	Lognormal	–	100[a] CFS
37–48	Hourly flow through spillway	Lognormal	–	0.02

[a]Indicates standard deviation

standard deviation of the revenue generated. The number of sample points required for training the anchored ANOVA decomposition (ANOVA-D) model is 4609. The results obtained by ANOVA-D have been validated with that of MCS (10^4 samples) for each optimization iteration. The robust optimal solutions obtained by utilizing ANOVA-D as presented in Table 28 and Fig. 11 has achieved excellent similarity with benchmark MCS solutions. This illustrates high approximation accuracy of ANOVA-D. Additionally, ANOVA-D has utilized significantly less number of sample points as compared to MCS, which illustrates its computational efficiency. Overall, the performance of ANOVA-D is

$$
\begin{aligned}
&\arg\min \; -\left[\beta\mu_R + (1-\beta)\sigma_R\right] \\
&\text{s.t.} \quad \mu_{f_t(i)} - 3\sigma_{f_t(i)} \geq 0, \forall i \\
&\qquad \mu_{f_t(i)} + 3\sigma_{f_t(i)} \geq 25000, \forall i \\
&\qquad \mu_{f_t(i)} - 3\sigma_{f_t(i)} + \mu_{f_s(i)} - 3\sigma_{f_s(i)} \geq 500, \forall i \\
&\qquad \left|\mu_{f_t(i)} + 3\sigma_{f_t(i)} + \mu_{f_s(i)} + 3\sigma_{f_s(i)} - \mu_{f_t(i-1)} + 3\sigma_{f_t(i-1)} - \mu_{f_s(i-1)} + 3\sigma_{f_s(i-1)}\right| \leq 500, \forall i \\
&\qquad \mu_{S(i)} - 3\sigma_{S(i)} \geq 50000, \forall i \\
&\qquad \mu_{S(i)} + 3\sigma_{S(i)} \geq 100000, \forall i \\
&\qquad \mu_{S(\text{end})} = 90000
\end{aligned}
\tag{125}
$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ denote mean and standard deviation, respectively. R denotes the revenue generated and S denotes the storage of the reservoir. f_t and f_s in Eq. (125) represent the flow through turbine and spillway, respectively. β is the weightage factor. The objective is to determine f_t and f_s that minimizes the objective function as defined in Eq. (125).

The RDO results have been obtained for $\beta = 0.5$, i.e., equal weightage has been assigned to the mean and

acknowledgeable in such high-dimensional problem as this present one.

After carrying out an extensive numerical study by utilizing few surrogate models in RDO framework (Sect. 4) and also, addressing a practical engineering problem in an efficient manner (Sect. 5), the study has been summarized briefly in the next section.

Table 28 Robust optimal solutions for the hydroelectric dam problem (Sect. 5)

Design variables	MCS (10^4 samples)	ANOVA-D (4609 samples)
$f_t(1)$	800.0016	800
$f_t(2)$	800.0019	800
$f_t(3)$	800.002	800
$f_t(4)$	800.0028	800
$f_t(5)$	800.003	800
$f_t(6)$	800.0058	800
$f_t(7)$	840.7067	840.6930
$f_t(8)$	1.0407×10^{3}	1.0407×10^{3}
$f_t(9)$	1.2407×10^{3}	1.2407×10^{3}
$f_t(10)$	1.4407×10^{3}	1.4407×10^{3}
$f_t(11)$	1.6407×10^{3}	1.6407×10^{3}
$f_t(12)$	1.8407×10^{3}	1.8407×10^{3}
$f_s(1)$	9.7467×10^{-8}	4.1097×10^{-11}
$f_s(2)$	5.4615×10^{-8}	3.7066×10^{-11}
$f_s(3)$	3.1662×10^{-7}	1.7349×10^{-10}
$f_s(4)$	1.9007×10^{-9}	1.2200×10^{-12}
$f_s(5)$	9.2755×10^{-8}	1.2338×10^{-10}
$f_s(6)$	6.4504×10^{-9}	6.1031×10^{-12}
$f_s(7)$	1.0177×10^{-11}	6.9300×10^{-15}
$f_s(8)$	6.1147×10^{-9}	2.4160×10^{-12}
$f_s(9)$	4.4513×10^{-9}	2.2503×10^{-12}
$f_s(10)$	6.6375×10^{-8}	2.8331×10^{-11}
$f_s(11)$	1.2683×10^{-7}	7.6531×10^{-12}
$f_s(12)$	1.4256×10^{-7}	1.9602×10^{-11}
Objective function ($\beta = 0.5$)	224.2756	224.0604

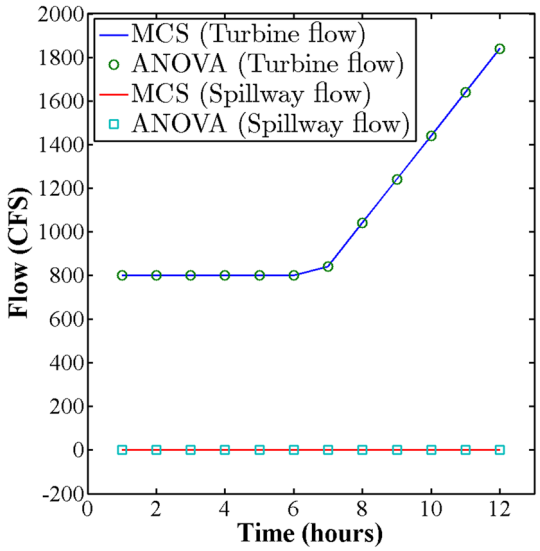

Fig. 11 Comparison of the flow through turbine and spillway (i.e., design variables) as obtained by utilizing ANOVA-D and MCS

6 Summary and Recommendations

An extensive survey has been carried out in this study illustrating the performance of surrogate models in RDO framework. As previously illustrated, the approximation accuracy of a surrogate model is a crucial factor in stochastic optimization as a slight deviation from the results of any intermediate iterations may easily deviate to yield a false or local optima. Therefore, the motivation of the study has been to access the performance of available surrogate models in terms of their approximation potential while solving typical non-linear RDO problems. This study would also serve as a guiding handbook for the selection of a suitable surrogate model for addressing a problem of a particular level of complexity. In this context, few salient points have been highlighted on the basis of the results achieved by various surrogate models:

- First and foremost, ANOVA-D has outperformed the other models investigated for solving typical non-linear RDO examples. It has proven its consistency and robustness in accurately approximating the response functions in all examples, unlike the other models. It is highly recommended for use in future applications of stochastic optimization.
- Secondly, performance of least angle regression based PCE is noteworthy both in terms of yielding accurate solutions and rate of convergence, except in the third example. After ANOVA-D, PCE-LAR has achieved the second ranking among the models utilized and thus, recommended to be utilized and further improved.
- Thirdly, ordinary least square and quadrature based PCE do not achieve convergence in examples 5 and 6. Thus, they are not recommended to be utilized for relatively high-dimensional non-linear problems with multiple constraints. However, they may be considered to be suitable for low-dimensional problems.
- Fourthly, performance of models such as, Kriging, RBF, LWP, ANN and MARS have been observed to vary in different problems and thus, found to be inconsistent. It is highly recommended to validate the results obtained by using these above models with benchmark solutions, if available.
- Lastly, SVM has yielded unacceptable results due to its lack of capability to accurately capture the non-linearity in functional space and thus, gets deviated from the true optima in presence of constraints.

Thus, after identifying high reliability of ANOVA-D in complex landscapes, it has been employed to solve a practical large-scale hydroelectric dam model. As expected,

ANOVA-D has yielded similar results as compared to MCS based RDO solutions, by utilizing limited number of training points, considering the scale of the problem. Thus, the study illustrates the resilience of anchored ANOVA decomposition is note-worthy and encouraging for applications in further complex engineering systems.

Acknowledgements TC acknowledges the support of MHRD, Government of India and RC acknowledges the support of CSIR via Grant No. 22(0712)/16/EMR-II.

Compliance with Ethical Standards

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest.

References

1. Jin R, Du X, Chen W (2003) The use of metamodeling techniques for optimization under uncertainty. Struct Multidiscip Optim 25:99–116
2. Zang C, Friswell MI, Mottershead JE (2005) A review of robust optimal design and its application in dynamics. Comput Struct 83:315–326
3. Beyer H-G, Sendhoff B (2007) Robust optimization: a comprehensive survey. Comput Methods Appl Mech Eng 196:3190–3218
4. Chen W, Allen J, Tsui K, Mistree F (1996) Procedure for robust design: minimizing variations caused by noise factors and control factors. J Mech Des Trans ASME 118:478–485
5. Du X, Chen W (2000) Towards a better understanding of modeling feasibility robustness in engineering design. J Mech Des Trans ASME 122:385–394
6. Huang B, Du X (2007) Analytical robustness assessment for robust design. Struct Multidiscip Optim 34:123–137
7. Phadke M (1989) Quality engineering using robust design. Prentice Hall, Englewood Cliffs, NJ
8. Taguchi G (1987) System of experimental design: engineering methods to optimize quality and minimize costs, vol 1, UNIPUB/Kraus International Publications, White Plains, NY
9. Taguchi G (1986) Quality engineering through design optimization. Krauss International Publications, White Plains, NY
10. Alexandrov N, Lewis R (2002) Analytical and computational aspects of collaborative optimization for multidisciplinary design. AIAA J 40:301–309
11. Hicks R, Henne PA (1978) Wing design by numerical optimization. J Aircr 15:407–412
12. Sobieszczanski-Sobieski J, Haftka R (1997) Multidisciplinary aerospace design optimization: survey of recent developments. Struct Optim 14:1–23
13. Fang J, Gao Y, Sun G, Li Q (2013) Multiobjective reliability-based optimization for design of a vehicle door. Finite Elem Anal Des 67:13–21
14. Hwang K, Lee K, Park G (2001) Robust optimization of an automobile rearview mirror for vibration reduction. Struct Multidiscip Optim 21:300–308
15. Sun G, Li G, Zhou S, Li H, Hou S, Li Q (2010) Crashworthiness design of vehicle by using multiobjective robust optimization. Struct Multidiscip Optim 44:99–110
16. Diez M, Peri D (2010) Robust optimization for ship conceptual design. Ocean Eng 37:966–977
17. Hart CG, Vlahopoulos N (2009) An integrated multidisciplinary particle swarm optimization approach to conceptual ship design. Struct Multidiscip Optim 41:481–494
18. Parsons M, Scott R (2004) Formulation of multicriterion design optimization problems for solution with scalar numerical optimization methods. J Sh Res 48:61–76
19. Doltsinis I, Kang Z (2004) Robust design of structures using optimization methods. Comput Methods Appl Mech Eng 193:2221–2237
20. Lagaros ND, Plevris V, Papadrakakis M (2007) Reliability based robust design optimization of steel structures. Int J Simul Multidiscip Des Optim 1:19–29
21. Lee SH, Chen W, Kwak BM (2009) Robust design with arbitrary distributions using Gauss-type quadrature formula. Struct Multidiscip Optim 39:227–243
22. Cheng J, Liu Z, Wu Z, Li X, Tan J (2014) Robust optimization of structural dynamic characteristics based on adaptive Kriging model and CNSGA. Struct Multidiscip Optim 51:423–437
23. Roy BK, Chakraborty S (2015) Robust optimum design of base isolation system in seismic vibration control of structures under random system parameters. Struct Saf 55:49–59
24. Lee T, Jung J (2006) Metamodel-based shape optimization of connecting rod considering fatigue life. Key Eng Mater 211:306–308
25. Li F, Meng G, Sha L, Zhou L (2011) Robust optimization design for fatigue life. Finite Elem Anal Des 47:1186–1190
26. McDonald M, Heller M (2004) Robust shape optimization of notches for fatigue-life extension. Struct Multidiscip Optim 28:55–68
27. Lee I, Choi KK, Du L, Gorsich D (2008) Dimension reduction method for reliability-based robust design optimization. Comput Struct 86:1550–1562
28. Ramakrishnan B, Rao S (1996) A general loss function based optimization procedure for robust design. Eng Optim 25:255–276
29. Schuëller GI, Jensen HA (2008) Computational methods in optimization considering uncertainties—an overview. Comput Methods Appl Mech Eng 198:2–13
30. Bhattacharjya S (2010) Robust optimization of structures under uncertainty, Ph.D. Thesis, Department of Civil Engineering. Bengal Engineering and Science University, Shibpur
31. Kleijnen JPC (1987) Statistical tools for simulation practitioners. Marcel Dekker Inc., New York
32. Jin R, Chen W, Simpson T (2001) Comparative studies of metamodeling techniques under multiple modeling criteria. Struct Multidiscip Optim 23:1–13
33. Sudret B (2012) Meta-models for structural reliability and uncertainty quantification. In: Proc. 5th Asian-Pacific Symp. Stuctural Reliab. Its Appl. (APSSRA, 2012), Singapore, pp 53–76
34. Kim S-H, Na S-W (1997) Response surface method using vector projected sampling points. Struct Saf 19:3–19
35. Kang S-C, Koh H-M, Choo JF (2010) An efficient response surface method using moving least squares approximation for structural reliability analysis. Probabilistic Eng Mech 25:365–371
36. Jacquelin E, Adhikari S, Sinou J, Friswell MI (2014) Polynomial chaos expansion and steady-state response of a class of random dynamical systems. J Eng Mech 141:04014145
37. Rabitz H, Aliş ÖF (1999) General foundations of high-dimensional model representations. J Math Chem 25:197–233
38. Kaymaz I (2005) Application of kriging method to structural reliability problems. Struct Saf 27:133–151
39. Deng J (2006) Structural reliability analysis for implicit performance function using radial basis function network. Int J Solids Struct 43:3255–3291

40. Mareš T, Janouchová E, Kučerová A (2016) Artificial neural networks in the calibration of nonlinear mechanical models. Adv Eng Softw 95:68–81
41. Richard B, Cremona C, Adelaide L (2012) A response surface method based on support vector machines trained with an adaptive experimental design. Struct Saf 39:14–21
42. Friedman JH (1991) Multivariate adaptive regression splines. Ann Stat 19:1–67
43. Simpson T, Mauery T, Korte J, Mistree F (1998) Comparison of response surface and kriging models for multidisciplinary design optimization. In: 7th AIAA/USAF/NASA/ISSMO Symp. Multidiscip. Anal. Optim, St. Louis, MO, pp 381–391
44. Giunta A, Watson L, Koehler J (1998) A comparison of approximation modeling techniques: polynomial versus interpolating models. In: Proc. Seventh AIAA/USAF/NASA/ISSMO Symp. Multidiscip. Anal. Optim. AIAA-98-4758. pp 1–13
45. Ray T, Smith W (2006) A surrogate assisted parallel multiobjective evolutionary algorithm for robust engineering design. Eng Optim 38:997–1011
46. Chowdhury R, Rao BN, Prasad AM (2008) High dimensional model representation for piece-wise continuous function approximation. Commun Numer Methods Eng 24:1587–1609
47. Chowdhury R, Rao BN, Prasad AM (2009) High-dimensional model representation for structural reliability analysis. Commun Numer Methods Eng 25:301–337
48. Alis ÖF, Rabitz H (2001) Efficient implementation of high dimensional model representations. J Math Chem 29:127–142
49. Chastaing G, Gamboa F, Prieur C (2011) Generalized hoeffding-sobol decomposition for dependent variables-application to sensitivity analysis. Electron J Stat 6:2420–2448
50. Ho T-S, Rabitz H (2003) Reproducing kernel Hilbert space interpolation methods as a paradigm of high dimensional model representations: application to multidimensional potential energy surface construction. J Chem Phys 119:6433
51. Sobol IM (1993) Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1:407–414
52. Wiener N (1938) The homogeneous Chaos. Am J Math 60:897. doi:10.2307/2371268
53. Xiu D, Karniadakis GE (2002) The Wiener–Askey polynomial Chaos for stochastic differential equations. SIAM J Sci Comput 24:619–644
54. Hampton J, Doostan A (2015) Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression. Comput Methods Appl Mech Eng 290:73–97
55. Filomeno Coelho R, Lebon J, Bouillard P (2011) Hierarchical stochastic metamodels based on moving least squares and polynomial chaos expansion. Struct Multidiscip Optim 43:707–729
56. Madankan R, Singla P, Patra A, Bursik M, Dehn J, Jones M et al (2012) Polynomial chaos quadrature-based minimum variance approach for source parameters estimation. Procedia Comput Sci 9:1129–1138
57. Zheng Zhang TA, El-Moselhy IM, Elfadel L, Daniel (2014) Calculation of generalized polynomial-chaos basis functions and Gauss quadrature rules in hierarchical uncertainty quantification. IEEE Trans Comput Des Integr Circuits Syst 33:728–740
58. Zhao F, Tian Z (2012) Gear remaining useful life prediction using generalized polynomial chaos collocation method. In: Pham H (ed), Proc. 18TH ISSAT Int. Conf. Reliab. Qual. Des. pp 217–221
59. Hosder S, Walters RW, Balch M (2012) Point-collocation nonintrusive polynomial chaos method for stochastic computational fluid dynamics. AIAA J 48:2721–2730
60. Blatman G, Sudret B (2010) An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. Probabilistic Eng Mech 25:183–197
61. Blatman G, Sudret B (2011) Adaptive sparse polynomial chaos expansion based on least angle regression. J Comput Phys 230:2345–2367
62. Xiu D, Karniadakis GE (2002) Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. Comput Methods Appl Mech Eng 191:4927–4948
63. Xiu D, Karniadakis GE (2003) Modeling uncertainty in flow simulations via generalized polynomial chaos. J Comput Phys 187:137–167
64. Sudret B (2008) Global sensitivity analysis using polynomial chaos expansions. Reliab Eng Syst Saf 93:964–979
65. Pascual B, Adhikari S (2012) A reduced polynomial chaos expansion method for the stochastic finite element analysis. Sadhana-Acad Proc Eng Sci 37:319–340
66. Pascual B, Adhikari S (2012) Combined parametric-nonparametric uncertainty quantification using random matrix theory and polynomial chaos expansion. Comput Struct 112:364–379
67. Craven P, Wahba G (1978) Smoothing noisy data with spline functions. Numer Math 31:377–403
68. Sudjianto A, Juneja L, Agrawal H, Vora M (1998) Computer aided reliability and robustness assessment. Int J Reliab Qual Saf Eng 05:181–193
69. Wang X, Liu Y, Antonsson EK (1999) Fitting functions to data in high dimensionsal design space. In: ASME Des. Eng. Tech. Conf., Las Vegas, p DETC99/DAC-8622
70. Krishnamurthy T (2003) Response surface approximation with augmented and compactly supported radial basis functions. In: 44th AIAA/ASME/ASCE/AHS/ASC Struct. Struct. Dyn. Mater. Conf., American Institute of Aeronautics and Astronautics, Reston, Virigina
71. Hardy RL (1971) Multiquadric equations of topography and other irregular surfaces. J Geophys Res 76:1905–1915
72. Buhmann MD (2000) Radial basis functions. Acta Numer 9:1–38
73. Volpi S, Diez M, Gaul NJ, Song H, Iemma U, Choi KK et al (2015) Development and validation of a dynamic metamodel based on stochastic radial basis functions and uncertainty quantification. Struct Multidiscip Optim 51:347–368
74. Dai HZ, Zhao W, Wang W, Cao ZG (2011) An improved radial basis function network for structural reliability analysis. J Mech Sci Technol 25:2151–2159
75. Chau MQ, Han X, Jiang C, Bai YC, Tran TN, Truong VH (2014) An efficient PMA-based reliability analysis technique using radial basis function. Eng Comput 31:1098–1115
76. Chau MQ, Han X, Bai YC, Jiang C (2012) A structural reliability analysis method based on radial basis function. C Comput Mater Contin 27:128–142
77. Kriesel D (2010) A brief introduction to neural network
78. Hagan MT, Demuth HB, Beale MH, Jesús OD (2008) Neural network design, 2nd edn. Cengage Learning, New Delhi
79. Shu S, Gong W (2016) An artificial neural network-based response surface method for reliability analyses of c-φ slopes with spatially variable soil, China. Ocean Eng 30:113–122
80. Dai H, Zhang H, Wang W (2015) A multiwavelet neural network-based response surface method for structural reliability analysis. Comput Civ Infrastruct Eng 30:151–162
81. Peng W, Zhang J, You L (2015) The hybrid uncertain neural network method for mechanical reliability analysis. Int J Aeronaut Sp Sci 16:510–519
82. Zio E (2006) A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes. IEEE Trans Nucl Sci 53:1460–1478
83. Vapnik VN (1998) Statistical learning theory. Wiley, New York
84. Goldstein H (1986) Classical mechanics. Addison-Wesley, Reading, MA

85. Z. Guo, G. Bai, Application of least qquares support vector machine for regression to reliability analysis. Chin J Aeronaut 22 (2009) 160–166
86. Lins ID, Droguett EL, das Chagas Moura M, Zio E, Jacinto CM (2015) Computing confidence and prediction intervals of industrial equipment degradation by bootstrapped support vector regression. Reliab Eng Syst Saf 137:120–128
87. Liu J, Vitelli V, Zio E, Seraoui R (2015) A novel dynamic-weighted probabilistic support vector regression-based ensemble for prognostics of time series fata. IEEE Trans Reliab 64:1203–1213
88. Gunn SR (1997) Support vector machines for classification and regression, Technical report, image speech and intelligent systems research group. Southampton, UK
89. Xiao M, Gao L, Xiong H, Luo Z (2015) An efficient method for reliability analysis under epistemic uncertainty based on evidence theory and support vector regression. J Eng Des 26:340–364
90. Zhao W, Tao T, Zio E, Wang W (2016) A novel hybrid method of parameters tuning in support vector regression for reliability prediction: particle Swarm optimization combined With analytical selection. IEEE Trans Reliab 1–13
91. Zhao W, Tao T, Zio E (2013) Parameters tuning in support vector regression for reliability forecasting. Chem Eng Trans 33:523–528
92. Coen T, Saeys W, Ramon H, De Baerdemaeker J (2006) Optimizing the tuning parameters of least squares support vector machines regression for NIR spectra. J Chemom 20:184–192
93. Olea RA (2011) Optimal contour mapping using Kriging. J Geophys Res 79:695–702
94. Warnes JJ (1986) A sensitivity analysis for universal kriging. Math Geol 18:653–676
95. Krige DG (1951) A Statistical approach to some basic mine valuation problems on the witwatersrand. J Chem Metall Min Soc South Africa 52:119–139
96. Krige DG (1951) A statisitcal approach to some mine valuations and allied problems at the Witwatersrand. University of Witwatersrand
97. Joseph VR, Hung Y, Sudjianto A (2008) Blind Kriging: a new method for developing metamodels. J Mech Des 130:031102
98. Hung Y (2011) Penalized blind kriging in computer experiments. Stat Sin 21:1171–1190
99. Couckuyt I, Forrester A, Gorissen D, De Turck F, Dhaene T (2012) Blind Kriging: implementation and performance analysis. Adv Eng Softw 49:1–13
100. Kennedy M, O'Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. Biometrika 87:1–13
101. Wang B, Bai J, Gea HC (2013) Stochastic Kriging for random simulation metamodeling with finite sampling. In: Vol. 3B 39th Des. Autom. Conf., ASME, p V03BT03A056. doi:10.1115/DETC2013-13361
102. Qu H, Fu MC (2014) Gradient extrapolated Stochastic Kriging. ACM Trans Model Comput Simul 24:1–25
103. Kamiński B (2015) A method for the updating of stochastic kriging metamodels. Eur J Oper Res 247:859–866
104. Bhattacharyya B (2017) A critical appraisal of design of experiments for uncertainty quantification. Arch Comput Methods Eng. doi:10.1007/s11831-017-9211-x
105. Rivest M, Marcotte D (2012) Kriging groundwater solute concentrations using flow coordinates and nonstationary covariance functions. J Hydrol 472–473:238–253
106. Putter H, Young GA (2001) On the effect of covariance function estimation on the accuracy of Kriging predictors. Bernoulli 7:421–438
107. Biscay Lirio R, Camejo DG, Loubes J-M, Muñiz Alvarez L (2013) Estimation of covariance functions by a fully data-driven model selection procedure and its application to Kriging spatial interpolation of real rainfall data. Stat Methods Appt 23:149–174
108. Cleveland W (1979) Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 74:829–836
109. Hastie T, Loader C (1993) Local regression: automatic kernel carpentry (with discussion). Stat Sci 8:120–143
110. Fan J, Gijbels I (1992) Variable bandwidth and local linear regression smoothers. Ann Stat 20:196–216
111. Cleveland W, Devlin S (1988) Locally weighted regression: an approach to regression analysis by local fitting. J Am Stat Assoc 83:596–610
112. Cleveland W, Devlin S, Grosse E (1988) Regression by local fitting. J Econom 37:87–114
113. M. Wand, M. Jones (1995) Kernel smoothing, CRC Press, Boca Raton, FL
114. Fan J, Gijbels I (1996) Local polynomial modelling and its applications, Chapman and Hall
115. Schucany W (1995) Adaptive bandwidth choice for kernel regression. J Am Stat Assoc 90:535–540
116. Fan J, Gijbels I (1995) Adaptive order polynomial fitting: bandwidth robustification and bias reduction. J Comput Graph Stat 4:213–227
117. Lall U, Moon Y, Kwon H, Bosworth K (2006) Locally weighted polynomial regression: parameter choice and application to forecasts of the Great Salt Lake. Water Resour Res 42
118. Fan J (1993) Local linear regression smoothers and their minimax efficiencies. Ann Stat 21:196–216
119. Fan J, Gasser T, Gijbels I, Brockmann M, Engel J (1997) Local polynomial regression: optimal kernels and asymptotic minimax efficiency. Ann Inst Stat Math 49:79–99
120. Cheng M, Fan J, Marron J (1997) On automatic boundary corrections. Ann Stat 25:1691–1708
121. Fan J, Marron J (1994) Fast implementations of nonparametric curve estimators, J Comput Graph Stat 3:35–56
122. Hall P, Wand M (1996) On the accuracy of binned kernel density estimators. J Multivar Anal 56:165–184
123. Seifert B, Brockmann M, Engel J, Gasser T (1994) Fast algorithms for nonparametric curve estimation. J Comput Graph Stat 3:192–213
124. Seifert B, Gasser T (2000) Data adaptive ridging in local polynomial regression. J Comput Graph Stat 9:338–360
125. Li Q, Lu X, Ullah A (2003) Multivariate local polynomial regression for estimating average derivatives. J Nonparametr Stat 15:607–624
126. Kai B, Li R, Zou H (2010) Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression, J R Stat Soc Ser B-Stat Methodol 72:49–69
127. McKay MD, Beckman RJ, Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 2:239–245
128. Deb K (2001) Multi-objective optimization using evolutionary algorithms. Wiley, Chichester
129. Regis RG (2014) Evolutionary programming for high-dimensional constrained expensive black-box optimization using radial basis functions. IEEE Trans Evol Comput 18:326–347
130. Deb K, Gupta S, Daum D, Branke J, Mall AK, Padmanabhan D (2009) Reliability-Based Optimization Using Evolutionary Algorithms. IEEE Trans Evol Comput 13:1054–1074
131. DeLand S (2012) Solving large-scale optimization problems with MATLAB: A hydroelectric flow example