Research in Engineering Design
https://doi.org/10.1007/s00163-020-00336-7

ORIGINAL PAPER

Managing computational complexity using surrogate models: a critical review

Reza Alizadeh[1] · **Janet K. Allen**[1] · **Farrokh Mistree**[2]

Received: 12 August 2018 / Revised: 4 March 2020 / Accepted: 9 March 2020

Abstract
In simulation-based realization of complex systems, we are forced to address the issue of computational complexity. One critical issue that must be addressed is the approximation of reality using surrogate models to replace expensive simulation models of engineering problems. In this paper, we critically review over 200 papers. We find that a framework for selecting appropriate surrogate modeling methods for a given function with specific requirements has been lacking. Having such a framework for surrogate model users, specifically practitioners in industry, is very important because there is very limited information about the performance of different models before applying them on the problem. Our contribution in this paper is to address this gap by creating practical guidance based on a trade-off among three main drivers, namely, size (how much information is necessary to compute the surrogate model), accuracy (how accurate the surrogate model must be) and computational time (how much time is required for the surrogate modeling process). Using the proposed guidance a huge amount of time is saved by avoiding time-consuming comparisons before selecting the appropriate surrogate model. To make this contribution, we review the state-of-the-art surrogate modeling literature to answer the following three questions: (1) What are the main classes of the design of experiment (DOE) methods, surrogate modeling methods and model-fitting methods based on the requirements of size, computational time, and accuracy? (2) Which surrogate modeling method is suitable based on the critical characteristics of the requirements of size, computational time and accuracy? (3) Which DOE is suitable based on the critical characteristics of the requirements of size, computational time and accuracy? Based on these three characteristics, we find six different qualitative categories for the surrogate models through a critical evaluation of the literature. These categories provide a framework for selecting an efficient surrogate modeling process to assist those who wish to select more appropriate surrogate modeling techniques for a given function. It is also summarized in Table 4 and Figs. 2, 3. MARS, response surface models, and kriging are more appropriate for large problems, acquiring less computation time and high accuracy, respectively. Also, Latin Hypercube**,** fractional factorial designs and D-Optimal designs are appropriate experimental designs. Our contribution is to propose a qualitative evaluation and a mental model which is based on quantitative results and findings of authors in the published literature. The value of such a framework is in providing practical guide for researchers and practitioners in industry to choose the most appropriate surrogate model based on incomplete information about an engineering design problem. Another contribution is to use three drivers, namely, computational time, accuracy, and problem size instead of using a single measure that authors generally use in the published literature.

Keywords Surrogate model · Model selection · Meta model · Computational complexity · Design · Response surface

Abbreviations

ANN	Artificial Neural Network
CCD	Central composite design
CFD	Computational fluid dynamics
CPU	Central processing unit
DOE	Design of experiments
DST	Dempster–Shafer theory
EA	Evolutionary algorithm
FD	Factorial design
FEA	Finite element analysis
FFD	Fractional factorial design
GSME	Generalized mean square error
KRG	Kriging
LS	Least squares
MAE	Mean absolute error
MAPE	Mean absolute percentage error

✉ Janet K. Allen
janet.allen@ou.edu

Extended author information available on the last page of the article

Published online: 06 April 2020

MARS	Multivariate Adaptive Regression Splines
MAXE-CV	Maximum absolute cross-validation error
MSEGO	Multiple surrogate efficient global optimization
MEMO	Multimodal-based evolutionary multiple-objective
NSGA-II	Non-dominated sorting genetic algorithm II
OA	Orthogonal array
PNN	Polynomial Neural Network
POF	Pareto optimal front
PRESS	Predicted residual error sum of squares
PRS	Polynomial response surface
PSO	Particle swarm optimization
RBDO	Reliability-based design optimization
RBF	Radial basis function
RMSE	Root Mean Square Error
RSM	Response surface models
SE	Square error
SVM	Support vector machines
WAS	Weighted average surrogate
WLSR	Weighted least square regression

1 Frame of reference

Over the last few decades, computer simulation models, which are often used to represent physical problems via mathematical models and computer code, have begun to play a crucial role in engineering problems. Many simulation models are applied in various types of engineering problems, including optimization design, uncertainty design, reliability analysis, reliability-based design optimization (RBDO), and robust design. However, these computer simulations have a strong tendency to be computationally expensive due to their excessively detailed representation of real-world systems. Furthermore, many model-based engineering problems require simulation models to be run thousands of times to develop an appropriate solution, which also demands a high computational budget.

Surrogate models (SMs) or metamodels are used to replace expensive simulation models of engineering problems. Although computers are becoming faster and more powerful every day, the increasing computational complexity demand is still much higher than the increasing power and speed of computers. Additionally, serious difficulties in using computer simulations can occur due to their high fidelity regardless of how fast and powerful the computers used are. Even methods, such as parallel computing, which is a computation in which many calculations or executions of processes are carried out simultaneously, are not very helpful (Grama 2013). The disadvantages of parallel computing are (a) designing shared-memory machines becomes difficult and expensive as the number of processors increases; (b) the lack of scalability between memory and CPUs because adding more CPUs can increase traffic on the shared-memory-CPU and catch-memory management paths (Geist 2000). Steuben studies one of the most promising parallel computing platforms which has been gaining wider usage over the last decade, namely, the Graphics Processing Unit, or GPU. He hypothesized that three factors must be considered in designing simulations utilizing General Purpose Graphics Processing Unit (GPGPU): parallelization, synchronization, and approximation (Steuben 2014).

In this review, we use three metrics: computational time, accuracy and problem size instead of the Big O notation because we are looking at problem-side and not algorithm-side complexity. So, computational time is problem specific and software-hardware dependency cannot be generalized. SMs are used in the initial phases of design to explore the solution space and develop a set of solutions. There are many types of surrogate modeling techniques for approximating reality. Response surface methodology (RSM) (Box et al. 1978) are methods used to create fast estimations of intricate mathematical models. Many people have used kriging instead of complicated mathematical models, such as (Sacks et al. 1989). Recently, other statistical methods, such as inductive learning (Simpson et al. 2001), adaptive learning (Picheny et al. 2010; Viana et al. 2010a, b, c; Wang et al. 2014), and different machine learning methods such as support vector machines (SVM) (Gunn 1998; Brereton and Lloyd 2010; Viana et al. 2013) have attracted attention.

In several papers, authors use one SM for a specific task. Simpson et al. (1998) compared kriging and RSM to address an optimization design problem with three design variables. In Giunta et al. (1994), another comparison between kriging and RSM was performed. Additionally, (Varadarajan et al. 2000) compared RSM to the ANN technique to address an engine design problem including nonlinear thermodynamic performance.

Multiple factors in a particular problem, including linearity/nonlinearity, the size of the problem, the desired level of accuracy, the speed of the process, the required amount of information, the size of the sample, and the availability of convenient computer code impact the appropriateness of a SM. Using only one measure of merit, such as accuracy, for comparing SMs in surrogate modeling is not efficient, and a set of multiple measures must be considered (Jin et al. 2000). Therefore, we contend that to select one SM for a task, knowledge about the characteristics of various SM using various modeling measures is needed.

In this paper, we review the literature regarding surrogate modeling and shed light on the state-of-the-art methods in this area of research. The intellectual questions that are posed and investigated in this paper are summarized in Table 1. We define *size,* as the number of input variables used in the surrogate modeling process. C*omputational time*

Table 1 Questions investigated in this paper

Research questions	Relevant sections
1. What are the main classes of the design of experiment (DOE) methods, surrogate modeling methods and model-fitting methods based on the requirements of size, computational time, and accuracy?	1, 2.1, 2.2
2. Which surrogate modeling method is suitable based on the critical characteristics of the requirements of size, computational time and accuracy?	1, 2.2
3. Which DOE is suitable based on the critical characteristics of the requirements of size, computational time and accuracy?	2.1, 2.2

is also defined as the time needed to create the data set, that is, simulation time. Accuracy is the measure used to evaluate the deviation of the prediction results from the actual values. It should be noted that accuracy and precision have been used interchangeably. However, these two measures are different and to achieve a good prediction performance, acceptable values for both should be obtained. Accuracy is defined as the closeness of a predicted value to a standard value while precision is defined as the closeness of two or more predicted values to each other. We use accuracy as the measure for the predictive performance of the SMs.

The abovementioned questions help us to identify a more appropriate surrogate modeling process, including the selection of a DOE to produce data and subsequent use of a model to represent the generated data. Then, we need to create the best fit between the model and the data. Then, section two is divided into three subsections based on this classification. In the first subsect. (1.1), DOE methods are introduced and described. Then, before describing the different DOE techniques, we explain the problem of size and how to address this issue with the most well-known dimensionality reduction methods, such as principal component analysis, variable screening, partitioning and Bayesian updating. A broad range of DOE techniques are then reviewed. In the second subsection (1.2), an overview of different methods that can be used to represent the data generated by DOEs is provided. In this subsection, in addition to methods that have been used by researchers for a long time, such as RSMs, kriging and MARS, more recently developed methods, such as reference-point-based Nondominated Sorting SVM, are reviewed. Also, methods used for fitting the selected models to represent the data are discussed. Finally, an approach for comparing different surrogate modeling methods is introduced based on these fitting indicators.

Next, an overview of the reviews of the surrogate modeling literature is presented. A summary of previous review papers is provided in Table 2. In Simpson et al. (2001), the authors perform a broad survey of metamodeling uses in aerospace and mechanical systems; in Barthelemy and Haftka (1993), a review of metamodeling applications in structural optimization is carried out. The use of surrogate modeling in multidisciplinary design optimization is discussed in Sobieszczanski-Sobieski and Haftka (1997). Barton presents developments in this area with a particular focus on experimental designs (Barton 1994). Also, they provide a review of statistical methods that are useful in conducting computer experiments with a focus on the task of metamodeling, which is driven by the goal of optimizing a complex system via a deterministic simulation model in case of stochastic simulation (Chen et al. 2006). A review (Viana 2008) advises against employing a single SM and alternatively proposes that MSs are appropriate and even can be used in association with one another. The MSs are useful in water resource applications (Chen et al. 2006). Kriging correlation parameters can be interpreted to some extent in that large values of a dimension indicate a highly nonlinear function in that dimension, while small values indicate a smooth function with limited variation (Maier and Dandy 2000). Simpson et al. (2008) report that the prevalent theme in six highly cited metamodeling (or design and analysis of computer experiments) review papers is indeed the high cost of computer simulations. Global optimization algorithms based on RSMs, such as EGO (Jones et al. 1998), GMSRBF, MLMSRBF (Regis and Shoemaker 2007), and Gutmann's method (Gutmann 2001), and uncertainty analysis algorithms, such as ACUARS (Mugunthan and Shoemaker 2006) and RBF-enabled MCMC (Bliznyuk et al. 2008), can help circumvent the computational budget limitations associated with computationally intensive simulation models. However, even though Jones (2001) and Simpson et al. (2008) both note that surrogate modeling has more advantages than simply reducing computational time, reviewing other possible motivations for surrogate modeling is beyond the scope of this paper.

Simpson et al. (2001) and Wang et al. (2014) also review the literature regarding RSM for engineering design optimization problems. Simpson et al. (2004) summarize a discussion panel regarding RSM held at the 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization. Simpson et al. (2008) review the literature regarding RSM and motivations from a historical perspective and emphasize the appeal of lower-fidelity physically based surrogate modeling. Forrester and Keane (2009) review recent advances in surrogate modeling, including

Table 2 Summary of previous review papers

Paper	Focus	Type of Problem	Types of metamodels
Fang et al. (2017)	Design optimization for structural crashworthiness	Engineering design	RSM radial basis function (RBF) model, kriging (KRG), ANN,, ensemble of surrogates (EoS)
Laurent et al. (2017)	Gradient-enhanced surrogate models	Numerical comparisons of metamodels carried out for approximating analytical test functions	Classical, weighted and moving least squares, Shepard weighting functions, and kernel-based methods like RBF, kriging and SVM
Chen et al. (2006)			RSM; Spatial correlation models; MARS; Regression trees; ANNs; RBFs; Least interpolating polynomials
Turner (2005)	Hyperdimensional performance models for engineering design	Engineering design	RSMs; Spline models, Rational B-splines (NURBs), MARS; KRG, Design and Analysis of Computer Experiments (DACE), SVM, ANN, Frequency Domain Methods
Dey et al. (2017)	Comparison of metamodel results with traditional Monte Carlo simulation for high fidelity uncertainty quantification	Composite plates	RSM, kriging, high dimensional model representation (HDMR), PCE, ANN, MLS, SVM, MARS, RBF and PNN
Pickett and Turner (2011); Barton (1994)	General purpose mathematical approximations to input—output functions	Engineering design	Splines, RBF, kernel smoothing, spatial correlation models, and frequency-domain approximations
Clarke et al. (2004)	The computationally efficient theory behind SVR is reviewed, and SVR approximations are compared	Engineering design	KRG, RSM, RBF, MARS and SVR
Kianifar and Campean (2019)	Framework for describing the typology of engineering problems, in terms of dimensionality and complexity, and the modelling conditions	Engineering problems	KRG, RSM, and RBF
Viana and Haftka (2008)	Multiple SMs	Engineering design	KRG, RBF, linear Shepard and SVR
Chen et al. (2006)	Wastewater treatment stochastic dynamic programming (SDP)	Aerospace engineering: engineering design and optimization; electrical engineering: CAD/CAM modeling and optimization; chemical engineering: optimization of a continuous-stirred tank reactor; continuous-state stochastic dynamic programming	RSM, RSM, MARS, spatial correlation models, OA designs, regression trees and related methods, least interpolating polynomials, ANNs, RBFs
Kleijnen (2009)	Extension of kriging to random simulation and discussion of bootstrapping to estimate the variance of the KRG predictor	Classic one-shot statistical designs, sequentialized and customized designs for sensitivity analyses and optimization	KRG, RSM, such as RBF models
Maier and Dandy (2000)	Review of modeling issues and applications	ANN for water resource forecast	ANN
Jones Schonlau and Welch (1998)	Efficient global optimization	Balancing the need to use the approximating surface with the need to enhance the estimation	RSM
Sacks et al. (1989)	Scientific phenomena are currently investigated by complex computer code	Design and analysis of computer experiments	RSM, neural networks, function bounds
Sobieszczanski-Sobieski and Haftka (1997)	Primary challenges in MDO, computational expense	Engineering design	Kriging, RBNN, linear Shepard

Table 2 (continued)

Paper	Focus	Type of Problem	Types of metamodels
Barthelemy and Haftka (1993)	Applications of nonlinear programming methods to large structural design problems	Design and analysis of computer experiments	RSM, ANN, function bounds
Regis and Shoemaker (2007)	Efficient global optimization	Groundwater bioremediation models	RSM, GMSRBF and MLMSRBF
Mugunthan and Shoemaker (2006)	Uncertainty analysis algorithms	Assessing the impacts of parameter uncertainty in computationally expensive groundwater models	ACUARS and RBF-enabled MCMC
Bliznyuk et al. (2008)	Basic sequential framework	Automatic calibration and Bayesian uncertainty analysis of an environmental model	RBF
Jones (2001)	Adaptive–recursive framework	Local optimization	RSM, kriging
Simpson et al. (2008)	Multi-level/multi-fidelity approximations and ensembles of metamodels	Availability of metamodels in commercial software for design space exploration and visualization	Two-dimensional response, RSM, MARS, spatial correlation models, OA designs
Simpson et al. (2001)	Computer-based engineering design	Approximation of deterministic computer analysis codes	Two-dimensional response, surfaces and one ten-dimensional surface, MARS, spatial correlation models, OA designs, regression trees and related methods, least interpolating polynomials, ANNs, RBFs
Wang et al. (2016)	Model approximation, design space exploration, problem formulation, optimization support	Finite element analysis (FEA) and computational fluid dynamics (CFD)	Polynomial (linear, quadratic, or higher), splines (linear, cubic, and NURBS), MARS, Gaussian process, kriging, RBF, least interpolating polynomials, ANN, knowledge base or decision tree, SVM, hybrid models
Simpson et al. (2004)	Approximation methods in multidisciplinary analyses and optimization	Design of experiments versus design and analysis of simulations or computer experiments, reflecting experimental results, and data from approximation models, capturing uncertainty	Surfaces and one ten-dimensional surface, MARS, spatial correlation models, RBFs
Forrester and Keane (2009)	Efficient global optimization	Advances in lower-fidelity physically based surrogates in the field of optimization	Polynomials, moving least squares, RBF, RBFF of noisy data, kriging, universal kriging, blind kriging with noisy data, SVM
Bandler and Madsen (2001); Bandler et al. (2008)	Editorial-surrogate modeling and space mapping	Engineering optimization	ANN, kriging, RSM, and approaches based on splines
Ratto et al. (2012)	Tools involved in surrogate modeling	Reduction and sensitivity analyses of complex environmental models	RSM, kriging, RBF, ANN
Jin (2005)	Comprehensive survey of fitness approximation in evolutionary computation	RSM if used with evolutionary optimization algorithms	RSM, EAs
Chen et al. (2003)	Primarily on the task of metamodeling, which is driven by the goal of optimizing a complex system via a deterministic simulation model	Applications in electrical engineering, chemical engineering, mechanical engineering, and dynamic programming	RSM, kriging, regression splines, regression trees, ANN, orthogonal arrays, Latin hypercubes, number-theoretic methods

advances in lower-fidelity physics-based surrogates in the field of optimization. Special journal issues focusing on surrogate modeling summarize the first and second International Workshops on Surrogate Modeling and Space Mapping for Engineering Optimization (Bandler and Madsen 2001; Bandler et al. 2008). Another special issue publication regarding surrogate modeling is a thematic journal issue focusing on surrogate modeling for sensitivity analyses and the reduction of complex environmental models (Ratto et al. 2012). In addition, there are more specific review papers focusing on specific tools/strategies for surrogate modeling. Kleijnen (2009) reviews kriging and its applications in RSM. Jin et al. (2001) and Chen et al. (2006) review and compare multiple function approximation models acting as RSM. Jin (2005) focuses on RSM used with evolutionary optimization algorithms. Laurent et al. (2017) review gradient-enhanced metamodels, which involve function gradients that use common auxiliary information and are useful for predicting functions based on locally changing behaviors. Turner (2005) review a wide range of SMs by grouping them into three main families including geometric models, spline models, stochastic and heuristic models. He compares these surrogate modeling methods focusing on Non-Uniform Rational B-splines (NURBs) methods (Turner and Crawford 2005). Dey et al. (2017) present a state-of-the-art review along with a comparative analysis on SMs for evaluation of uncertainty in frequencies of composite plates based on computational efficiency and accuracy. Clarke et al. (2004) review support vector regression (SVR) as an alternative technique for approximating complex engineering analyses. In this survey, the computationally efficient theory behind SVR is reviewed, and SVR approximations are compared to the aforementioned four metamodeling techniques using a test bed of 26 engineering analysis functions. Kianifar and Campean (2019) introduce a framework for describing types of engineering problems, in terms of dimensionality and complexity, and the modeling conditions, reflecting the noisiness of signals and the affordability of sample sizes, and on this basis present a systematic evaluation of the performance of frequently used metamodeling techniques. Additionally, Fang et al. (2017) review structural crashworthiness in design optimization. Surrogate modeling has become increasingly popular over the last decade within the water resources community, which is consistent with the increasing utilization of metamodels in the scientific literature since 1990 as documented by Viana and Haftka (2008).

As shown in Table 2, (Barton 1994; Clarke et al. 2004; Turner 2005; Chen et al. 2006; Fang et al. 2017) review the most commonly used SMs, compare them, and provide recommendations for the choice of SMs. However, no survey analyzes the characteristics of more appropriate SM for a design problem to create a classification and generalize it for guidance for industrial practitioners in engineering design. In the published review papers, a comprehensive set of criteria for identifying appropriate SMs has been lacking. Furthermore, no authors have discussed the automation of the surrogate modeling process. Our contribution is to address these gaps.

2 Design of experiments (DOEs)

In this section, we explain the DOEs, discuss the different types of DOEs and perform a comparison of the types; for any comparison, some indices or metrics are needed; thus, we provide an overview of existing and most frequently used evaluation methods used as metrics for assessing DOEs.

In this section, we review the various types of DOEs and measures of merit used to evaluate the DOEs.

2.1 The problem of size

In the surrogate modeling process, addressing the problem of size, which is related to the number of variables, is among the greatest challenges. Hence, addressing this issue first is essential as it directly influences the entire process of surrogate modeling and particularly the DOE. Because the computational time is directly related to the number of factors, an effective strategy for managing the computational time is to reduce the problem dimension. Determining which variables or phenomena ("factors" in statistical terminology) in an experiment are the "important" ones. This has two aspects: first, whether an effect is large enough that we can be sure it is real and not due simply to noise (or error) alone (i.e., "statistically significant")? We have treated this question to some extent in our paper. The second aspect is if the effect of a factor is indeed real; is it of sufficiently large magnitude to be of practical importance? While the answer to this question is important to understanding the outcome of the experiment, it is not a statistical question, so, we did not address it in our paper. In our review, different design options, including complete factorial and fractional factorial designs are discussed and compared to address this limitation and difficulty of DOEs.

2.1.1 Principal component analysis (PCA)

In this technique, we use an orthogonal conversion through to transform an observed variable set which is correlated with an independent linear set of variables called principal components (PCs). The number of PCs is always smaller than the number of initially observed design variables. In

PCA eigenvalues are used to determine lower-dimensional sets of variables which retain the maximum possible information (Gao et al. 2018).

2.1.2 Variable screening

Screening the variables refers to discovering significant variables in a set of random variables. In this definition, the word "significant" has various implications based on the problem investigated (He et al. 2013). There are several types of variable screening methods; for example, Kochet al. (1999) use a Pareto method to examine the outcomes of an experimental design in grading the significance of the factors in each response. They use a 2-level fractional factorial experiment to screen 26 control and noise factors this includes 64 designs and one center point. Additionally, in Cho et al. (2014), the authors introduce an approach for variable screening in surrogate modeling with fractional output variance based on univariate dimension reduction. In this variable selection process, the authors chose variables that yield higher output variance as vital factors.

2.1.3 Sensitivity analysis

"Sensitivity analysis screening methods aim to isolate the most important factors in experiments involving a large number of significant factors and interactions (Campolongo and Braddock 1999). Sensitivity testing is an important component of building mathematical and simulation models. The values of model parameters and the initial (input) values of variables are subject to many sources of uncertainty. An understanding of the sensitivity of the model outputs to the uncertainty in the values of the input variables and parameters, is necessary in developing confidence in the model and its predictions. Sensitivity analysis (SA) experiments may be performed on mathematical and computational models to determine the relative contribution of input variables and parameter values to the observed variations in model outputs. These computational experiments can determine, within reasonable limits, which parameters or initial variable values may have effects on the model outputs which are negligible, significant, linear or non-linear.

Box et al. (1978) discuss and define the main effects (first order), as well as the higher order effects (second order, third order, etc.). They relate these effects to the Taylor series expansion of the response (or output) function of the model, as it depends on the values of the model parameters and initial values of model variables. In particular, the first derivative terms in the Taylor series correspond to the main effects (i.e., the effect on the response function due to perturbation of a single variable); the second derivative terms correspond to two-factor interactions (including the quadratic effects which are two-factor interactions of a factor with itself); etc.

Several sensitivity techniques are available in the literature to estimate main and higher order effects (Helton 1999). Those techniques can be divided into three main classes, according to the problem setting to which they are applied: (1) screening methods: where the aim is to isolate the most important factors from amongst a large number that may affect a particular response (Bettonvil and Kleijnen 1997); (2) global sensitivity methods: where the emphasis is on apportioning the output uncertainty to the uncertainty in the input parameters (Sobol 1993); (3) local sensitivity methods: where by computing the partial derivatives, the local response of the output function to variation in its input parameters is investigated (Rbitz 1989).

2.1.4 Partitioning

Partitioning the problem and multiple-dimensional or multiple-level optimization represent other approaches for addressing dimensionality. A comprehensive survey of methods used to decompose a problem into more manageable partitions is available in Lewis and Mistree (1998). In addition, in Sobieszczanski-Sobieski and Haftka (1997), several multidisciplinary design optimization techniques for framing and solving partitioned problems concurrently are introduced. Moreover, Balling and Wilkinson (1997) classify hierarchical partitioning methods into the following three classes: (i) single-stage problems, (ii) concurrent subproblems, and (iii) collaborative techniques. The most distinctive feature of the single-stage methods is the single discipline feasible exploration and design. In the 2nd type of problem partitioning, we address partitions independently and concurrently during the solution process. Then, the compatibility of the solutions for each subproblem is assessed (Koch et al. 1999).

2.1.5 Bayesian updating

If several simulations can be performed and many design variables must be included in the SM, Bayesian updating is useful and can overcome the difficulties causing classical stochastic optimization to fail (Cheng and Currie 2004). Thus, in the case of numerous design variables and a complex simulation model, only limited runs can be implemented, and approximating all coefficients of the SM using classical tools, such as maximum likelihood, may be impossible (Liang and Mahadevan 2016). In such cases, Bayesian updating has an enormous benefit.

In Bayesian updating, the initial probability distribution $f(x)$ is the first estimate of the probable value of the parameters that can achieve a feasible solution for starting the simulation runs. Then, the start point value is gradually improved, and estimates of the distribution of x become more accurate as the simulation runs proceed (Beck and Au 2002; Hasselman et al. 2005). Therefore, the prior lack of information

regarding the parameters is addressed by allocating initial distributions to the parameters and using experimental data (Kennedy and O'Hagan 2001; Ghanem and Doostan 2006).

2.2 The main types of experimental designs

Numerous DoE methods are available, including factorial designs, fractional factorial designs (Gunst and Mason 2009; Montgomery 2017), central composite designs (Montgomery 2017; Box and Behnken 1960), orthogonal arrays (Hedayat et al. 2012; Dey 1985; Montgomery 2017), D-optimal (Myers et al. 2016), Placket-Burman (Gustafsson et al. 2013), hexagon (Montgomery 2017), hybrid (Myers et al. 2016), Latin Hypercube (McKay et al. 1979; Montgomery 2017), random selection, grid search (Barthelemy and Haftka 1993; Güntert et al. 1998; Hsu et al. 2003; Corchado et al. 2007), etc. We review some of these methods.

Factorial designs are the most fundamental experimental designs. In this experimental design method, the number of factor levels identifies the number of design points. Two main factorial designs (FD) are 2^k (for two levels and k factors) (Box and Hunter 1961) and 3^k. See also Box et al. (1978) for more details.

The number of runs needed for a full factorial design increases exponentially and rapidly exceeds the available resources of most experimenters by increasing the factors through a 2^k factorial design. Under such circumstances, the engineering designer can obtain knowledge regarding the critical impacts using only a small portion of the full design (Montgomery 2017). Fractional factorial designs (FFD) utilize the specifications of the design to decrease the size of an experiment and limit the loss of critical knowledge (Gunst and Mason 2009).

Central composite designs (CCDs) are typical second-order designs used to explore higher-order impacts, mainly quadratic effects that require many design points. CCDs are 2-level FDs with 2^k star points for a design with k factors. Each star point is placed at a position with $\alpha \pm \beta$ distance from each factor (Montgomery 2017).

Orthogonal arrays (OAs) are the experimental designs recommended by Taguchi. These designs represent a type of FFD that is used primarily for 2 or 3 levels (Owen 1992). These designs are built to decrease the number of essential design points and are also known as a Plackett–Burman design (Hedayat et al. 2012).

Space filling design allow users to explore all regions of the design space equally (Montgomery 2017).

2.3 Comparison of different DOEs: evaluation metrics

Selecting the appropriate design is critical for efficient experimentation, i.e., balancing between obtaining knowledge about the relationships among factors, the response variables and computational time. Thus, we review the criteria used to evaluate and compare DOEs in this section.

The most significant challenge choosing an experimental design is the size of the problem. Unsaturated designs are the most common types of designs in which the number of design points is twice the number of factors. In saturated DOEs, the number of design points and number of factors are equal (Chen and Lin 1998). OAs, and two-level Placket–Burman designs are the most frequently used designs (Simpson et al. 2001).

Supersaturated designs represent a particular type of FD and are typically used if the factors outnumber the observations (Nguyen 1996). Building this sort of design by a random process is recommended (Kathleen and Cox 1962; Lin 1993). We use this type of design in a screening process to identify principal and sparse factors through a process which requires limited computation (Phoa et al. 2009). Frederick and co-authors in Phoa et al. (2009) studied supersaturated designs using the Dantzig selector and found that this variable selection method is highly effective in predicting the size of the model. If the essential runs of the simulations are not extremely expensive, using unsaturated designs for predictive functions is recommended (Simpson et al. 2001).

All approximators must be unbiased. Thus, the long-term expected value or average should be the same as the estimated parameter (Karson et al. 1969).

The variance of the coefficient predictions in a first-order model is minimal if we build it as follows:

$$\sum_{u=1}^{N} x_{iu}^2 = N, \tag{1}$$

where the approximation variance ($\hat{y}$) is an unmoving variance, and the distance from the design center is constant (Simpson et al. 2001).

Rotatability is an evaluation metric used for second-order models. For a 2nd-order model, we create appropriate approximations for the entire solution space (Khuri 1988). Box and Hunter recommend that a 2nd-order RSM must have rotatability (Box and Hunter 1957). This feature in a design problem means that the variance remains constant

as the design is rotated around the center (rotatable design) (Draper and Guttman 1988).

If the summation of the products of the N design points for each x_i and x_j is zero (as shown below), the design is orthogonal.

$$\sum_{u=1}^{N} x_{iu}x_{ij} = 0. \tag{2}$$

If the impacts of any factor balance out the impacts of other factors (sum to zero), the experimental design is orthogonal (Gunst 1996).

By performing a comprehensive review of different DOEs and their evaluation metrics, we found that the critical characteristics that should be considered in identifying the most appropriate surrogate modeling method are the problem of size, the accuracy level and the computational time In the following section, we compare the utility and appropriateness of the presented DOEs based on the required numbers of factors and design points, the computational time and the accuracy level.

2.4 Model choice

2.4.1 Response surfaces

If we assume that x is an independent vector of factors and y is the vector of response, then the impact of x on y and their relationship is:

$$y = f(x) + \varepsilon, \tag{3}$$

where ε denotes the normally distributed random error with a mean of zero and a standard deviation of 0 (Box and Draper 1987). Because the actual RSM function is indefinite, instead of $f(x)$, a new RSM $g(x)$ is built and used as a surrogate for $f(x)$. Thus, $\hat{y} = g(x)$ yields the estimated values of $f(x)$ (Khuri and Mukhopadhyay 2010).

The approximation function is the 1st-order model if we can demonstrate the actual model by a linear function of the factors as in:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_i x_i + e. \tag{4}$$

However, sometimes, we need to use a higher-order polynomial, such as a 2nd-order model (see Eq. 5.), if there is curvature in the problem as in

$$y = \alpha_0 + \sum_{i=1}^{k} \alpha_i x_i + \sum_{i=1}^{k} \alpha_{ii} x_i^2 + \sum \sum_{i<j} \alpha_{ii} x_i x_j + e. \tag{5}$$

RSM is very easy to use and very convenient (Deaton and Grandhi 2014).

2.4.2 Kriging

Kriging is a surrogate modeling technique introduced as a mixture of a polynomial model of x, i.e., $f(x)$, and localized deviations of x, i.e., $Z(x)$, as shown in Eq. 6.

$$y(x) = f(x) + Z(x). \tag{6}$$

where $Z(x)$ refers to the concept of a normally distributed Gaussian random process with a mean of zero, variance, and non-zero covariance. In Eq. 1, $f(x)$ is a polynomial function of an RSM that delivers a 'global' model and is often a fixed function (Stein 2012). Kriging is a technique of interpolation and interpolated values are modeled by a Gaussian process using information about prior covariances (Gano et al. 2006).

2.4.3 Radial basis functions (RBF)

This technique is a mathematical function that accepts real values, and its value is calculated based on the distance between the origin and each point as shown in Eq. 7 (Shan and Wang 2010).

$$\begin{aligned} Q\,(x) &= Q\,(\|x\|)\ \text{(distance from the origin)};\ Q(x,\ c) \\ &= Q(\|x - c\|)\ \text{(distance from the center)}. \end{aligned} \tag{7}$$

We can use these functions to construct SMs, as shown in Eq. 8.

$$\hat{y}(x) = \sum_{i=1}^{M} r_i Q(\|x - x_i\|), \tag{8}$$

where the surrogate function $\hat{y}$ refers to the integration of M radial basis functions, and each function is linked to a distinct x_i and has a weight of w_i (Broomhead and Lowe 1988).

2.4.4 Inductive learning

Inductive learning is a machine learning paradigm that is similar to metamodeling and regression (Dumais et al. 1998). In this method, the necessary rules are derived from examples; these rules are used to divide the data into distinct classes and are the principal modeling concepts (Michalski 1983). Both the features and results are assumed to be real values but discrete values can also be used (Witten et al. 2016).

2.4.5 Boosted trees and random forests

Friedman introduced boosted trees (BTs) or a gradient boosting machine as a method of supervised learning (Friedman 2001). BTs are used in supervised learning problems where multiple features x_i are available in a training dataset to estimate the response variable y_i (Madala and Ivakhnenko 1994). BTs produce a prediction model in the form of an ensemble of weak prediction models that are typically decision trees (De'ath 2007).

They also inherit the desirable features of trees while eliminating many undesirable features. The most desirable feature is robustness (La Fuente and Andres 2016; MacCalman et al. 2016). Sensitivity to long-tailed distributions and outliers is also eliminated (Jagadeesh et al. 2016; Lemercier et al. 2012). Internal feature selection and retaining robustness after addition of input variables are other advantages of BTs (Sim et al. 2018). BTs address missing values in an elegant manner (Song et al. 2018).

The disadvantage of a single tree is inaccuracy (Friedman 2001) as a result of the coarse characteristics of their piecewise continuous estimates, particularly for smaller trees. For larger trees, a major disadvantage is uncertainty, and the fact that they include predominately high-order interactions (Amit and Geman 1997; Breiman 2001; Cutler et al. 2012; Kleinberg 1990, 1996, 2000). These disadvantages are mitigated by boosting (Burnham and Anderson 2003; Han et al. 2011; Vapnik 2013; Witten et al. 2016).

The advantage of a single tree is interpretability, whereas boosted trees are thought to lack this feature (Duda et al. 2012). Small trees can be easily interpreted with caution while the interpretability of more massive trees is questionable (Ripley 2007). Overfitting the training sets and high variance are other disadvantages of BTs.

Random forests are ensemble learning techniques for regression, classification, and other tasks that are performed by building an aggregation of decision trees at the training stage and outputting the group that is the mode of the groups (classification) or mean estimation (regression) of single trees (Ho 1995, 1998). Random forests correct for decision trees' habit of overfitting to their training set (Friedman et al. 2001).

2.4.6 Adaptive learning/active learning

Adaptive learning is a particular machine learning technique that is semi-supervised, and its algorithm can refer to the information source interactively to achieve the favored outcomes for a newly generated data set (Carbonell 1970). It is instrumental in collecting data; although the computation is very expensive, identifying the precise distribution of the data is nontrivial in the initial stages of design (Gorissen et al. 2009). In active learning, there is a function for minimizing the number of samples chosen during the iterations and maximizing the knowledge attained with each step (Gorissen et al. 2010).

2.4.7 Hyperdimensional performance models

Hyperdimensional performance models are another type of SMs which are introduced by Turner in (2005). The distinct feature of Turner's proposed method is that SMs are constructed from a wide variety of mathematical basis functions (Yurner and Crawford 2005a; Turner and Crawford 2005b; Turner et al. 2011). However, Hyperdimensional performance models (HyPerModels) derived from Non-Uniform Rational B-splines (NURBs). NURBs are defined by a set of control points, knot vectors and the NURBs orders, resulting in a highly robust and flexible curve definition. The defining components of a NURBs HyPerModel can be used to define adaptive sequential sampling algorithms that allow a designer to efficiently survey the design space for interesting regions (Steuben and Turner 2014). The data collected from design space surveys can be represented with a HyPerModel by adapting NURBs fitting algorithms, originally developed for computer graphics, to address the unique challenges of representing a hyperdimensional design space (Pickett and Turner 2011). HyPerModels support design space analysis for adaptive sequential sampling algorithms, to detect robust design space regions or for fault detection by comparing multiple HyPerModels obtained from the same system (Turner et al. 2003). With HyPerMaps, an engineering designer has a window into the hyperdimensional design space, allowing a designer to explore the design space for undiscovered design variable combinations with superior performance capabilities (Turner 2005). NURBs-based SMs have been used for various types of problems, including, performance forecasting in manufacturing systems (Anderson and Turner 2015), robotic navigation (Steuben and Turner 2011a, b, 2015), total knee replacement (Adams et al. 2015), real-time composite material characterization (Steuben et al. 2015a, b), robust and global optimization (Turner et al. 2007a, b, c; Ajetunmobi et al. 2008; Steuben and Turner 2010, 2012; Steuben et al. 2013; Audoux et al. 2018), mixed integer optimization (Turner et al. 2007a, b, c), design space analysis (Turner et al. 2004, 2006, 2007a, b, c, 2011; Turner and Crawford 2005), product and process design (Turner 2011), fault detection (Turner et al. 2006), ice detection on wind turbine blades (Malavé and Turner 2011), in vivo tissue condition diagnosis (Turner 2010), data modeling (Turner and Crawford 2009; Hammond and Turner 2010), inverse characterization of composite materials (Steuben et al. 2015a, b), additive manufacturing modeling and simulation (Steuben et al. 2019), and stochastic analysis of composite laminates (Dey et al. 2017).

The main difference between this review and the review and Turner's (2005) review is that the focus in his survey is on hyperdimensional performance models under the assumption that all the models are complete and accurate while we recognize that models are not necessarily complete and accurate. As SMs are used to approximate the design space and then using them to find a global optimum solution is not a logical. What we propose in this survey is the qualitative practical guidance for industrial designers who are looking for good enough solutions. Also, in Turner's study, using hyperdimensional performance models to create robust solutions again is not possible when one is looking for single solutions on the boundary of the feasible design space under the assumption that models are complete and accurate. However, in our review, SMs are defined and critically evaluated as methods to approximate the reality to find solutions which are relatively insensitive to variations.

2.4.8 Ensembles of surrogates

Often an individual surrogate is chosen to model a specific problem and develop an approximation function based on prior knowledge from past experience. Ensemble of surrogates (EoS) have been used to overcome the drawbacks of other approaches (Acar 2015; Babaei and Pan 2016; Badhurshah and Samad 2015; Basudhar 2012; Bhattacharjee et al. 2018; Ezhilsabareesh et al. 2018; Glaz et al. 2009; Goel et al. 2006; Goel et al. 2007; Karwan and Rardin 1980; Liu et al. 2015; Luo et al. 2019; Lv et al. 2019; Shankar Bhattacharjee et al. 2016; Viana and Haftka 2008a, b). For example, adaptive sampling algorithms that add one point per cycle are readily available. The addition of one point at a time may be inefficient if running simulations in parallel is possible. To address this problem, (Viana et al. 2010a, b, c) proposed an algorithm for adding several points per optimization cycle based on the simultaneous use of EOS. In another study, (Song et al. 2018) analyzed the effectiveness of EOS in providing accuracy, robustness, and efficiency requirements for many specific problems. Xu and Zeger (2001) proposed two simple, complementary ways to address the relative benefits of EOS over a single surrogate.

EOS are also extensively used in surrogate-assisted optimization (Viana et al. 2013; Villanueva et al. 2013; Chaudhuri and Haftka 2014; Adhav et al. 2015; Wang et al. 2016; Song et al. 2018). In engineering design, first, (Mack et al. 2005) explored the complex design space for shape optimization of a bluff body. Then, Samad et al. (2006) used EOS in shape optimization of a turbomachinery blade. Yin et al. (2018) proposed a new EOS with multiple regional optimized weight factors. Alizadeh et al. (2019) used cross validation to create EoS with small data sets. They implemented EoS to predict banded microstructure in a hot rod rolling problem. They found that using EoS improves the predictions of SMs even with the smaller sets of data points and shorter simulation times.

Some studies analyze the SM-based design optimization concern of the modeling fidelity of the approximation functions (Bellary and Samad 2017; Habib et al. 2017; Shi et al. 2016; Ye et al. 2018; Song et al. 2018) (E-AHF) (Xing et al. 2019; Al-Juboori and Datta 2019; Hou et al. 2019). Many studies have been conducted using EOS as evolutionary computation methods (Lim et al. 2007; Samad et al. 2007; Bellary et al. 2016; Bhat et al. 2010; Arias-Montano et al. 2012; Hong et al. 2019; Bhattacharjee et al. 2016, 2017; Lv et al. 2019). Cross-validation has been used to assign weights to surrogates by generating EOS and facilitate the measurement of the accuracy (Viana et al. 2009). Also, Viana et al. (2008) discuss whether to use the best PRESS solution or a weighted surrogate if a single surrogate is needed.

2.4.9 Multi-fidelity models

Multi-fidelity models are models which are used to support different dimensions of computational science and engineering, from discovery to design to decision-making and more (Xiong et al. 2009, 2013; Xue et al. 2013; Easterling and Berger 2002).

In uncertainty propagation, model input is described by a random variable and we are interested in statistics of the model output. Using Monte Carlo simulation to estimate statistics of the model output often requires many model evaluations to achieve accurate approximations of the output statistics. A multi-fidelity method that combines outputs from computationally cheap low-fidelity models with outputs from high-fidelity models can lead to significant reductions in runtime and provide unbiased estimators of the statistics of the high-fidelity model outputs. Other approaches to uncertainty propagation include fuzzy set approaches and worst-case scenario analysis (Sadaghiani et al. 2014; Alizadeh et al. 2016a, b; Peherstorfer et al. 2018).

Multi-fidelity models may also be used for optimization. In optimization, we seek an input that leads to an optimal model output with respect to a given objective function. Optimization is typically solved using an iterative process that requires evaluations of the model in each iteration. Multi-fidelity optimization can reduce the runtime of the optimization process using low-fidelity models to accelerate the search or using a low-fidelity model in conjunction with adaptive corrections and a trust-region model management scheme. For example, efficient global optimization is a multi-fidelity optimization method that adaptively constructs a low-fidelity model by interpolating the objective function corresponding to the high-fidelity model with kriging (Peherstorfer et al. 2018).

2.5 Model fitting

2.5.1 Methods to train a surrogate model

After using DoEs to create a sample data set, SM can be created using different fitting methods such as the method of least squares, mean absolute percentage error, mean absolute error, weighted least square regression, or log likelihood for different polynomial approximations and backpropagation for neural networks.

The least square (LS) method is used when the selection of the design points through DOE is completed, then an RSM is built. In LS random errors are identically scattered (with a "0" mean and unknown variance) and are independent of each other (Vinzi et al. 2010). For *i*th observation, the difference between the fitted value ($\bar{y}_i$) and the observed (y) is the residual and is $e_i = y_i - \bar{y}_i$. And β_i approximations in the equation $y = X\beta + e$ must minimize the residuals' sum of squares, this is usually referred as the sum of squares of the errors (SSE) and presented as,

$$\text{Minimize } S = \sum_{i=1}^{n} e_i^2. \tag{9}$$

The residuals can be expressed as $e = y - X\beta$. Then S becomes $S = e^{\mathrm{T}}\, e = (y - X\beta)^{\mathrm{T}}\,(y - X\beta)$.

The mean absolute percentage of error is very frequently used to test the accuracy of an approximation because it is interpretable and scale-independent (Dobler and Anderson-Cook 2005). To address the issue of infinite values MAAPE[1] has been developed which is another version of MAPE (Kim and Kim 2016).

The mean absolute error (MAE) is another indicator used to measure two variables' variance. If we generate a scatter chart of m points, where the point k has a location (a_k, b_k). (See Eq. 10)

$$MAE = \frac{\sum_{k=1}^{m} |b_k - a_k|}{m} \tag{10}$$

If the errors have no static variation in typical LS, weighted least squares regression WLSR is used (Hansen et al. 2012). WLSR at a given point involves identifying the best multivariate polynomial approximation (Bettebghor et al. 2011).

Likelihood is a function of combinations of coefficients and parameters in a model that creates a statistical collection of data. Usually, using log-likelihood as the likelihood is more straightforward (Montgomery 2017).

Backpropagation is a technique used to compute a loss function's gradient in an ANN's weights and for training the data (Demuth et al. 2014). It is an iterative two-step process of propagation and weight adjustment. Once a given vector enters a space, it is propagated throughout the process until it arrives at the final layer. Then, using a loss function, the outcome is analyzed against the ideal result, and we can compute the error value of each neuron in the last layer (Chauvin and Rumelhart 2013).

The performance of a model is evaluated by comparing predicted values with the actual values. The most important measures used to evaluate the accuracy of the SMs are discussed in the following sub-section.

2.5.2 Methods for estimating model accuracy

Accuracy, or the ability of a model to predict as closely to the actual value as possible, is evaluated using different measures which some of these are discussed here.

R-Squared and Adjusted R-Squared: an accurate SM is built by adding/eliminating input variables by backward removal, forward addition or stepwise addition/elimination. It includes the computation of the *P*-value (the probability which shows the risk of incorrectly rejecting a given hypothesis) and the *F*-value < Prob. (which shows the percentage of time we may expect to obtain a specified *F*-value if there is no significant effective factor). The SM built must be checked against several criteria such as R-Squared an indicator which shows the variation around the average expressed by the model and $\mathrm{R-Squared_{adj}}$ which is an indicator of the variation around the average explained by the model, which is adjusted for the number of terms in the model, and $\mathrm{R-Squared_{pred}}$ which is an indicator of the estimation ability of the RSM) written as follows:

$$\mathrm{R\text{-}Squared_{pred}} = \frac{S_R}{S_T} = 1 - \frac{S_E}{S_T}\ (0 \le \mathrm{R - Squared_{pred}} \le 1)$$

$$\mathrm{R\text{-}Squared_{adj}} = 1 - \frac{\frac{S_E}{n-k-1}}{\frac{S_T}{n-1}} = 1 - \frac{(n-1)}{(n-k-1)}\left(1 - R^2\right)$$

$$\left(0 \le \mathrm{R\text{-}Squared_{adj}} \le 1\right).$$

where $S_T = S_E + S_R$ is the total sum of square. The values of R–Squared, $\mathrm{R\text{-}Squared_{adj}}$ and $\mathrm{R\text{-}Squared_{pred}}$ must be close to 1 (Venter et al. 1997).

Mean Absolute Error (MAE): For a specific block of observations, the mean absolute error for model A at estimation horizon h is defined as $\mathrm{MAE}_{A,h} = \frac{1}{N_{k^*}} \sum_{t \in k^*} \left| y_t - \hat{\mu}_{t,h}^{A} \right|$, where N_{k^*} is the number of observations in block k^*.

Maximum Mean Square Error (MMSE) is

$$MMSE = max. \left[\frac{1}{k} \sum_{i=1}^{k} (\bar{y}_i - y_i)^2 \right]$$

[1] Mean arctangent absolute percentage error.

The Root Mean Square Error (RMSE) is, $RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\bar{y}_i - y_i)^2}{n}}$, where y_i and $\bar{y}_i$ are the vector of the true values and the vector corresponding to ith prediction, respectively.

Cross validation is used to evaluate the outputs of a statistical study and generalize the results to other sets of data (Qian 2009; Qian et al. 2008, 2009a, b; Higdon et al. 2004; IT 2002; Jaeckel and Rebonato 1999; Reese et al. 2004; Shankar Bhattacharjee et al. 2016; Trucano et al. 2006; Kennedy and O'Hagan 2000). In this method the data is divided into test and training data sets and the process is repeated several times (Geisser 1993). In surrogate modeling, if we wish to fit a model to a given known data set, we train the model by running it with the training data set. Then, an SM is validated against the test data. Using this method a researcher gains insight into whether the SM can be generalized for a new independent set of data (Lucas 1994).

Besides accuracy measures like R-Squared, and Adjusted R-Squared, MAE, MMSE, and RMSE, many surrogate modeling users use various model checking plots to evaluate the accuracy of the created SMs. In the following section, some of these model checking plots are evaluated.

Adequate Model Checking Plots offer an alternative method of checking adequate accuracy, compares the range of the estimated values at the design points to the mean estimation error. Generally, a value larger than 4 shows an adequate model. Moreover, additional plots should also be checked, for example, the normal plot of residuals (which shows whether the residuals follow a normal distribution, if this is true, the points will follow a straight line), an estimated versus residual plot (plot of the residuals against ascending estimated response values), an estimated versus actual plot (a plot of estimated responses against actual response values for the design points used for SM construction. It is used to help to distinguish a value which is not simply estimated by the model) and Box-Cox plot (to identify the most suitable power transformation to be implemented) (Dey et al. 2017).

All the methods and measures described in Sect. 2.5.2 are used to evaluate the accuracy of SMs. However, to compare SMs, the number of variables used in creating the model is also a very important consideration as it defines the computational time of the process, so, in the following sub-section we discuss the two drivers, time and size.

2.5.3 Comparison of chosen models

The appropriateness of a SM depends on the trade-off between the model accuracy and the computational expense of the model. Because no SM is perfect, any SM used for a system's physical behavior can be refined further to increase its accuracy at an increased computational expense (Paiva et al. 2010). On the other hand, an alternative approach to reduce the computational expense is to replace detailed simulations with simplified approximate simulations, thereby sacrificing accuracy for reduced computational time. In this case, a strategy to maintain accuracy along with reducing the computation time can be to integrate data from approximate and detailed simulations to build a SM that describes the relationship between output and input parameters (Qian et al. 2006). Therefore, the question asked by a designer is, 'How much refinement of a SM is appropriate for a particular design problem?' Value-of-information has been addressed in the engineering design literature to determine whether to make a decision using the available information or collect more information before reaching a decision (Messer et al. 2010). However, the main drawback of applying existing value-of-information-based metrics for model refinement problems is that the existing metrics only account for time, size and accuracy separately; these metrics do not account for all three combined in SMs, and this consideration has an impact on design decisions (Panchal et al. 2009). To address the lack of current metrics in the context of model refinement, in this paper a value-of-information-based method for identifying the appropriate amount of refinement of SMs is proposed.

Our contribution in this paper is to introduce a SM selection guidance focused on designers in industry and practitioners who need a practical guide to select an acceptable SM. This is critical since most of the industrial designers use only one SM for different engineering design problems while our proposed guidance can help them to choose the suitable SM for the specific problem. Other contribution is that there is very limited information about the performance of SMs before applying and examining them on the problem. In the following section, we compare the chosen SMs by balancing size of the problem (the number of factors or design variables), simulation time and accuracy level.

3 Results

In this section, we outline the primary applications of each DOE, selecting a model to represent the data and model-fitting approaches to evaluate the fitness of the chosen models. We first discuss how we review the surrogate modeling process, and then, we recommend more appropriate approaches based on the evaluation metrics, computational time and accuracy.

3.1 Review of the applications of DOE, data representation models, and model-fitting methods

In this subsection, we summarize the different DOE techniques, various models used for representing data and methods used to evaluate the fitness of each model to the data generated by each DOE method. Here, we also discuss several surrogate modeling methods. For example, we conclude that if we use the fractional factorial design as a DOE and a linear or quadratic polynomial as the model representing the data, which are generated by a fractional factorial design, we can use LS regression as the model-fitting method; this process is a type of RSM. Table 3 provides an initial summary of different surrogate modeling procedures.

3.2 Suggestions for DOEs, model selection and application, and model-fitting methods

In this subsection, we develop scenarios that engineering designers may face while addressing problems. The computing time, number of design variables, and accuracy level for achieving an efficient SM process are the main critical factors that should be considered. Additionally, we consider accuracy in the explanation of the following nine different states: a-i, a-ii, a-iii, b-i, b-ii, b-iii, c-i, c-ii, and c-iii. Then, we explain the time-size-accuracy triangle (see Fig. 1) in more detail.

As shown in Fig. 2, we always should balance "the needed speed of a surrogate modeling process", "the required accuracy" and "the number of dimensions or degree of complexity we aim to maintain within the model". To achieve this trade-off, information about the available tools and their characteristics, capabilities, pros, and cons and problem-specific knowledge are needed; we provide the reader with this knowledge.

First, we review some state-of-the-art approaches to addressing the time-size-accuracy triangle. Li et al. (2017) classified the various metamodeling methodologies for multiple-objective optimization and propose a comparative analysis by explaining the pros and cons of each method.

In addition, Tutum and Deb (2015) and Li et al. (2017) applied two distinct selecting functions based on the following two recent ideas: (i) the KKT proximity measure function and (ii) the multimodal-based evolutionary multiple-objective (MEMO) selection function. Subsequently, they apply outcome surrogate modeling approaches to several standard two- and three-objective constrained and unconstrained test problems; the authors compared the efficiency of two surrogate modeling approaches, i.e., ANN and kriging, by applying these selection function methods to these problems. According to their results, the performance of the MEMO-based approach is high, and the ANN surrogate modeling technique is generally better able to approximate the multimodal selection function landscape portrayed by the overall approach. Additionally, the MEMO selection function enables the process to be completed using only a fraction of the solution evaluations (limited to 500–2000) compared to the hundreds of thousands of solution assessments.

Therefore, Deb and co-authors conclude that the MEMO method and ANN are more appropriate if the goal is to reduce the number of runs and, accordingly, the required time of the surrogate modeling process. Based on their recommendation, the use of different architectures and deeper ANNs for higher accuracy models, the use of other selection function modeling after successful EMO approaches and extending the approach to many-objective problems need to be further studied and are suggested as future research topics.

The authors also increased the understanding of the behavior of EMO approaches when confronting a substantial number of objectives and proposed some initial empirical analyses for multiple problems with multiple objectives (Li et al. 2017).

In addition to the extensive work investigating massive objectives and constraints, in Deb and Myburgh (2016), the authors proposed a method to simultaneously meet the time and size requirements. EAs can meet the size requirements suitably, but the run time is not efficient using these methods. However, in this study, they propose a method that can identify the optimum solution very quickly in problems with 300 to 2000 integer variables. Additionally, the method can manage computational complexity by addressing 50,000 to 10^9 variables (Table 3).

EOS is more appropriate than kriging if we use more than 100 experimental designs (Viana et al. 2010a, b, c). This method is more appropriate if we wish to obtain the results within 5 hours. Additionally, the authors apply the proposed method to problems with less than ten variables; thus, this approach is appropriate for the (a-i) type of problems in our taxonomy shown in Fig. 2.

Furthermore, the comprehensive literature review shown in Table 4 confirms that our results regarding the classification of methods based on time-size-accuracy triangle criteria are justifiable (Bettonvil and Kleijnen 1997).

After reviewing the surrogate modeling literature, the results of this analysis using the proposed taxonomy are summarized in Fig. 1.

(a-i) When the desired experimental time is 1–5 hours and the number of variables is less than 10, RSM performs well; however, this method is not appropriate if the functions are nonlinear or we have discrete and binary variables along with continuous variables. Under such circumstances, we can use kriging, but we should consider that utilizing this method leads

Table 3 Surrogate modeling techniques

Design of experiments	Surrogate model choice	Model fitting
Classical methods	Multivariate Adaptive Regression Splines (MARS)	Log-likelihood
Factorial design	Polynomial regression (Linear, quadratic)	Best Linear Unbiased Predictor (BLUP)
Optimal Designs	Kriging	(Weighted) Least Squares Regression
Central Composite	Splines (Linear, cubic, or NURBS)	Best Linear Predictor
Box-Behnken	Radial Basis Functions (RBF)	Sequential or Adaptive Metamodeling
Plackett–Burman	Artificial Neural Network (ANN)	Multipoint Approximation (MPA)
Space-filling methods	Least Interpolating Polynomials	Back Propagation (for ANN)
Hammersley Sequence	Support Vector Machine (SVM)	Entropy (inf.-theoretic, for inductive learning on decision tree)
Simple Grids	Knowledge Base or Decision Tree	
Minimax and Maximin	Gaussian Process	
Orthogonal Arrays (Taguchi)	High dimensional model representation (HDMR) [cut-HDMR, RS-HDMR, GHDMR]	
Latin Hypercube	Polynomial Chaos Expansion (PCE)	
Sobol Sequence	Group Method of Data Handling—Polynomial Neural Network (GMDH—PNN)	
Hammersley sequence	Weighted Least squares regression	
Hybrid Methods	Best Linear Unbiased Predictor (BLUP)	
Importance Sampling	Multipoint approximation (MPA)	
Discriminative Sampling	Sequential or adaptive metamodeling	
Directional Simulation	Hybrid Models	
Sequential or Adaptive Methods		
Random/Human Selection		

Table 4 Critical Literature Review Table: Survey of Engineering Application of DOE, Data Representation Models, and Model Fitting

Papers/criteria	Simulation time(hour)		Number of variables			Method-accuracy level
	$1 < \text{time} < 5$	$\text{time} > 5$	$v < 10$	$10 < v < 100$	$100 < v$	
Viana et al. (2010)	✓		✓			RSM—medium
Koch et al. (1999)	✓			✓		Kriging—high
Li et al. (2017)		✓			✓	SVM—medium
Li et al. (2017)	✓				✓	SVM—high
Panchal et al. (2008); Jiang et al. (2016)	✓			✓		Kriging—high
Yang et al. (2006)		✓			✓	MARS—medium
Geisser (1993); Jin et al. (2001); Viana et al. (2010); Jiang et al. (2016)	✓		✓			RSM—medium
Maier and Dandy (2000)		✓			✓	MARS—medium
Myers et al. (2016)	✓		✓			RSM—high
Seni and Elder (2010; Schmidhuber (2015)	✓				✓	SVM—high
Krauss et al. (2017)	✓				✓	MARS—high/SVM—high

to more complexity and that we can only use this method for deterministic applications.

(a-ii) If the desired experimental time is 1–5 hours and the number of variables is between 10 and 100, kriging is a more appropriate method, and the RSM loses efficiency. In addition, we can use SVM instead of RSM; however, if we need more accuracy, kriging is the better choice. Meanwhile, we can use variable screening and other dimensionality reduction methods if the number of factors is greater than 10 (Koch et al. 1999).

(a-iii) If the desired experimental time is 1–5 hours and the number of variables is very high, greater than 100, RSMs are not useful, and we can use MARS, SVM, and other EAs. Additionally, we can use kriging and even RSM if we perform dimensionality reduction steps using variable screening and PCA.

In this category, extensive analysis have been done on the time performance by Deb and Myburgh (2016) and Li et al. (2017) to demonstrate EAs efficiency. However, we need to integrate other more accurate methods such as kriging (Viana et al. 2013). Using EOS can address this challenge

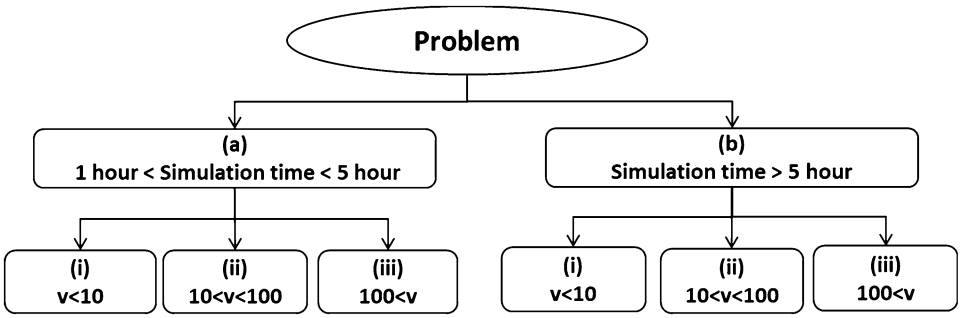

Fig. 1 Problem characterization based on the two aspects of time and size (v indicates the number of variables)

since we can combine more accurate surrogates, with faster surrogates, to compensate the weakness of each surrogate with the advantages of other surrogates.

(c-i) In this category, if the computational time can be more than 5 min and we do not have more than one variable, we can use RSM. However, the time can be reduced using EAs, such as GA and MARS.

(c-ii) If we require more than 5 min and the number of variables is high, for example between 10 and 100, we can use kriging (for high accuracy) and EA (for a more rapid process). In addition, we can use RSM after implementing some dimensionality reduction methods.

(c-iii) Evolutionary algorithms, e.g., SVM and MARS, are always appropriate when we have many variables. Therefore, in this category of problems, it is better to use EA. Of course, kriging can be used if we use some dimensionality reduction methods beforehand.

This six-type classification, which contains a wide range of SMs, demonstrates the differences among the models and allows us to adopt a more appropriate model for a specified variable size and level of accuracy in a given amount of time. In Table 5, we summarize the general characteristics of the SMs.

Based on our survey of the literature and the results reported in multiple studies, various methods are recommended for various situations in Fig. 3. Here, we demonstrated the interaction between three problem characterizations by high/low qualifiers in the context of different SMs. For instance, when the problem size is high, the accuracy level is low as long as the computation time remains the same, RSM and radial basis function are the appropriate SMs to use. While, when the size of the problem is high, MARS delivers higher accuracy in relatively shorter time comparing to the kriging. Another example is kriging, a method which results in high levels of accuracy and adequately performs for large problem sizes but not for problems with more than approximately 50 variables. Additionally, some authors mention in their outcomes that due to the complexity of the modeling, more time is needed to run kriging models (Simpson et al. 1998; Koch et al. 1999; Jin et al. 2000, 2001; Simpson et al. 2001; Deng et al. 2012a, b). Thus, we placed the kriging models in the triangle relatively close to the desired accuracy and relatively high for medium problem sizes, which need much more computational time than other methods. Kriging is placed close to the desired accuracy because it has a high level of accuracy and is at a medium distance from the problem size vertex because it is good enough for high-dimensional problems, but we placed it far from the computational time vertex because it is time-consuming.

SVM performs well for high-dimensional problems with high levels of nonlinearity. SVM has also been developed in recent years, and although the initial versions were very time-consuming, more recent versions are rapid. However, SVM is not very accurate as discussed in the literature (Broomhead and Lowe 1988; Cheng and Titterington 1994; Simpson et al. 2001; Yang et al. 2006; Demuth et al. 2014). Also, in most of the problems that SVM is used the number of variables is very high. Therefore, the high dimensionality is preferred over accuracy.

RSM has been used for a long time with high efficiency for small-sized problems with acceptable levels of accuracy. However, it loses its efficiency as the size of the problem grows, and if the functions are nonlinear (Simpson et al. 1997; Venter et al. 1997; Simpson et al. 1998; Khuri and Mukhopadhyay 2010; Myers et al. 2016). Therefore, we placed RSM close to the accuracy and computational time vertices because it is better in these criteria than in the size criteria.

Finally, RBF is very fast in modeling and better in solving models, but it's most powerful aspect is its performance in high-dimensional problems (Lin et al. 2015). Additionally, although it was not accurate enough in the past, some recent studies (Yang et al. 2006; Cho et al. 2014; Deb and Myburgh 2016; Li et al. 2017; Deb et al. 2017) have shown that the accuracy of the updated versions have been enhanced significantly. Thus, we decided to place this method at the center of the triangle but somewhat close to the size and time criteria.

Table 5 General Characteristics of SMs

Model selection	Time	Dimensionality	Accuracy	Complexity	Linearity	Deterministic/ stochastic
RSM	Rapid	Useful if the # of factors is less than 10	Not very high but acceptable	Easy to apply	Appropriate for linear functions	Appropriate if errors are random
Kriging	Slow but accurate	Applicable if the # of factors is more than 50	Very high	Complex but very flexible	Proper for linear and nonlinear functions	Appropriate for deterministic uses
EA (MARS, SVM)	Quick but not very accurate	Appropriate for problems with many parameters (~ 10,000)	Low level of accuracy	Useful if the complexity is very high	Proper for very nonlinear problems	Best performed in deterministic uses

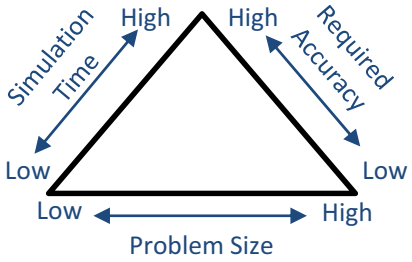

Fig. 2 Time-size-accuracy triangle

As the results have shown, conventional methods, such as RSM and kriging, are generally more appropriate if we wish to achieve more accurate SMs, and EAs are more appropriate if we wish to address high dimensionality and high linearity and obtain faster results. However, we find that EAs have achieved higher accuracy than they did previously; specifically, MARS is currently faster and more accurate. In the future, studying deep neural networks and integrating them with other EA methods through EOS enables SM users to compensate the weaknesses of various SMs; thus, this concept is an excellent area for future research. Finally, the time of the simulation is strictly associated with the circumstances under which the models run, e.g., the power of the computer used. Therefore, these circumstances should always be considered for comparison.

4 Closing remarks

In the simulation-based realization of complex systems we are forced to address the issue of computational complexity. SMs are used to replace expensive simulation models of engineering problems. We observe that a framework for selecting a more appropriate SM for a given function with

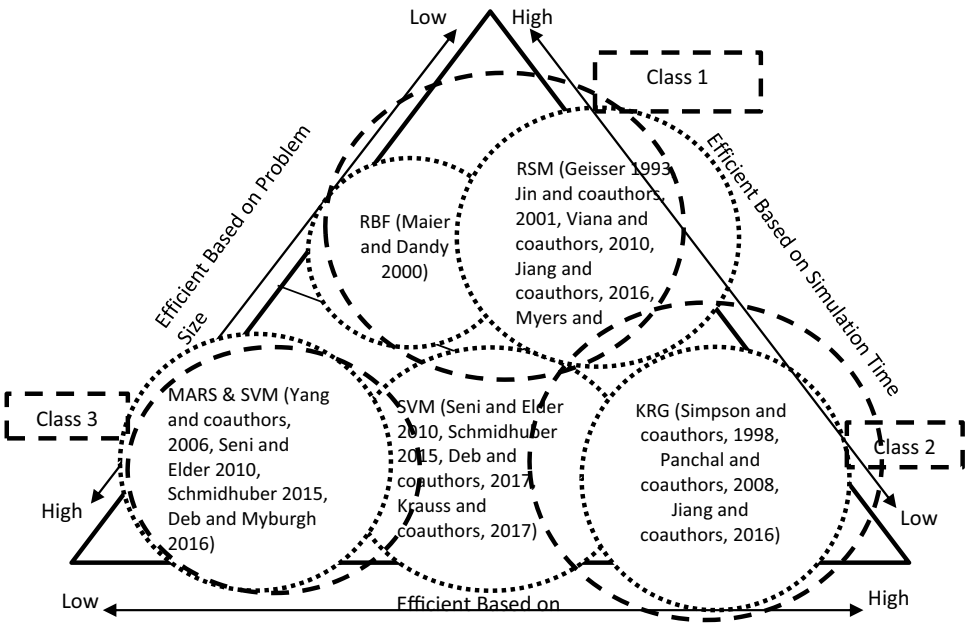

Fig. 3 The classification of some surrogate modeling techniques in a time-size-accuracy triangle

specific requirements has been lacking. To address this gap, we hypothesize that a trade-off among three main drivers, namely, size (how much information is necessary), accuracy (how accurate the model must be) and computational time (how much time is required for the surrogate modeling process) is needed. Our contribution is to create a framework, which is an easy to understand and use, for qualitative classification based on results in published literature, for selecting a more appropriate SM for a given function with specific requirements. As a result of testing the proposed hypotheses and our critical review of the state-of-the-art surrogate modeling literature we summarize our findings to three questions in Table 1:

Question 1: What are the main classes of design of experiment (DOE) methods, SMs and model-fitting methods based on the requirements of size, time and accuracy?

Finding 1: As shown in Table 4, Figs. 2, 3, based on three critical characteristics identified through the critical evaluation of the literature, six different categories of the SMs are introduced**.** This classification is a framework for selecting an efficient surrogate modeling process to assist those who aim to select a more appropriate surrogate modeling technique for a given function.

Question 2: Which SM is suitable based on the critical characteristics of size, time and accuracy?

Finding 2: As shown in Sect. 2.2, Fig. 3, SVM, RSMs, and kriging are more appropriate when the following characteristics are considered: large problem size, less computation time and high accuracy, respectively.

Question 3: Which DOE is more suitable based on the critical characteristics of the requirements of size, time accuracy?

Finding 3: As shown in the Sect. 2.2, and based on the answer to Question 2, Latin hypercube, fractional factorial design, and D-Optimal are appropriate DOEs for large problem sizes, less computation time and high accuracy, respectively.

4.1 Closing statement

We offer a designer or SM user the opportunity of particularizing the choice of SM for specific situations. However, sometimes, deciding which SM to use is difficult. As shown in Fig. 3, if only one or two priorities are to be considered, the decision of which method to use is not difficult. For instance, if computational time is the only concern, SVM and RSM are good. If accuracy and problem size are the most important issues, kriging is the appropriate method. The central question here is which of the SMs or which combination of models should be used if we wish to achieve a balance among these three priorities. A possible answer is to use EOS and cross-validation approaches. Our contribution is to create a practical framework which is a mental map and a critical qualitative evaluation based on quantitative results of authors in the published literature. The value of such a framework is to provide practical guidance for researchers and practitioners in industry to choose the most appropriate SM when there is incomplete information about the engineering design problem. Another of our contributions is to use three drivers, namely, computational time, accuracy, and problem size instead of using one single measure that authors generally have been using in the published literature.

Evidently by increasing the computer power, designers have been able to conduct high fidelity simulations and fitting process. Despite advances in computing power, the computational cost of sophisticated high-fidelity computer experiments often makes it impractical to rely exclusively on simulation for dealing with design problems (Viana et al. 2014). Moreover, the computational burden increases dramatically when the simulation models are applied to probabilistic design problems, for which a nested solving process is usually required (Zhao et al. 2012). In fact, implementing this framework entails complex and expensive probability calculation such as numerous samplings (Zhu and Du 2016). To reduce the computational cost, surrogate modeling techniques have been widely used for replacing sophisticated simulation models in many engineering applications (Haftka et al. 2016). In our review, we defined the computational cost as the cost of simulation and as computers become faster, the cost of simulations decreases. Although this transition may affect our results in terms of the computation time aspect, the accuracy and problem size criteria would be still effective in surrogate model selection process. So, practitioners would still be interested in choosing more expensive surrogate models to achieve higher accuracy for high dimensional problems.

In addition, the study presented here can help us develop a rule base for automating the process of SM selection. This is becoming increasingly important as computational platforms help us to move closer to automating design analysis processes.

Funding This study was supported by Tata Consultancy Services (Grant no. 105-373200) and The University of Oklahoma (Grant nos. 122-794800, 122-763300).

References

Acar E (2015) Effect of error metrics on optimum weight factor selection for ensemble of metamodels. Expert Syst Appl 42(5):2703–2709

Adams, MJ, Silverman AK, Turner CJ (2015) Pseudo elimination of geometry dependence in surrogate models of distributed knee loads from an explicit dynamic finite element analysis. International design engineering technical conferences. DETC2015-47850: V01AT02A055

Adhav R, Samad A, Kenyery F (2015). Design optimization of electric centrifugal pump by multiple surrogate models. SPE middle east oil and gas show and conference. Manama, Bahrain, Society of Petroleum Engineers

Ajetunmobi AM, Turner CJ, Crawford RH (2008). Robust optimization With NURBS HyPerModels. International design enginggering technical conferences. DETC2008-49961: 317–327

Al-Juboori M, Datta B (2019) Reliability-based optimum design of hydraulic water retaining structure constructed on heterogeneous porous media: utilizing stochastic ensemble surrogate model-based linked simulation optimization model. Life Cycl Reliab Saf Eng 8(1):65–84

Alizadeh R, Lund PD, Beynaghi A, Abolghasemi M, Maknoon R (2016a) An integrated scenario-based robust planning approach for foresight and strategic management with application to energy industry. Technol Forecast Soc Chang 104:162–171

Alizadeh R, Khodaei R, Mknoon R (2016b) A combined model of scenario planning and assumption-based planning for futurology, and robust decision making in the energy sector. J Energy Plan Policy Res 2(2):7–32

Alizadeh R, Jia L, Nellippallil AB, Wang G, Hao J, Allen JK, Mistree F (2019) Ensemble of surrogates and cross validation for rapid and accurate predictions using small data sets. Artif Intell Eng Des Anal Manuf 33(4):484–501

Amit Y, Geman D (1997) Shape quantization and recognition with randomized trees. Neural Comput 9(7):1545–1588

Anderson D, Turner CJ (2015). Applying NURBs-based surrogate models for performance forecasting in manufacturing systems. International mechanical engnieering congress and expo. IMECE2015-51862: V011T014A042

Arias-Montano A, Coello CAC, Mezura-Montes E (2012) Multi-objective airfoil shape optimization using a multiple-surrogate approach. Evolutionary computation (CEC), 2012 IEEE congress on, IEEE

Audoux Y, Montemurro M, Pailhes J (2018) A surrogate model based on non-uniform rational B-splines hypersurfaces. Procedia CIRP 70:463–468

Babaei M, Pan I (2016) Performance comparison of several response surface surrogate models and ensemble methods for water injection optimization under uncertainty. Comput Geosci 91:19–32

Badhurshah R, Samad A (2015) Multiple surrogate based optimization of a bidirectional impulse turbine for wave energy conversion. Renew Energy 74:749–760

Balling R, Wilkinson C (1997) Execution of multidisciplinary design optimization approaches on common test problems. AIAA J 35(1):178–186

Bandler JW, Madsen K (2001) Surrogate modelling and space mapping for engineering optimization. Optim Eng 2(4):367–368

Bandler WJ, Koziel S, Madsen K (2008) Editorial—Surrogate modeling and space mapping for engineering optimization. Optim Eng 9:307–310

Barthelemy J-FM, Haftka RT (1993) Approximation concepts for optimum structural design—a review. Struct Optim 5(3):129–144

Barton RR (1994). Metamodeling: a state of the art review. Proceedings of winter simulation conference, IEEE

Basudhar A (2012) Selection of anisotropic kernel parameters using multiple surrogate information. 12th AIAA aviation technology, integration, and operations (ATIO) conference and 14th AIAA/ISSMO multidisciplinary analysis and optimization conference

Beck JL, Au S-K (2002) Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. J Eng Mech 128(4):380–391

Bellary SAI, Samad A (2017) An alternative approach to surrogate averaging for a centrifugal impeller shape optimisation. Int J Comput Aided Eng Technol 9(1):62–83

Bellary SAI, Adhav R, Siddique MH, Chon B-H, Kenyery F, Samad A (2016) Application of computational fluid dynamics and surrogate-coupled evolutionary computing to enhance centrifugal-pump performance. Eng Appl Comput Fluid Mech 10(1):171–181

Bettebghor D, Bartoli N, Grihon S, Morlier J, Samuelides M (2011) Surrogate modeling approximation using a mixture of experts based on EM joint estimation. Struct Multidiscip Optim 43(2):243–259

Bettonvil B, Kleijnen JPC (1997) Searching for important factors in simulation models with many factors: sequential bifurcation. Eur J Oper Res 96(1):180–194

Bhat S, Viana FAC, Lind R, Haftka R (2010) Control-oriented design using H-infinity synthesis and multiple surrogates. 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, American Institute of Aeronautics and Astronautics

Bhattacharjee KS, Singh HK, Ray T, Branke J (2016) Multiple surrogate assisted multiobjective optimization using improved pre-selection. 2016 IEEE Congress on Evolutionary Computation (CEC)

Bhattacharjee KS, Isaacs A, Ray T (2017) Multi-objective optimization using an evolutionary algorithm embedded with multiple spatially distributed surrogates. MULTI-OBJECTIVE OPTIMIZATION: techniques and application in chemical engineering. World Scientific, Singapore, pp 135–155

Bhattacharjee KS, Singh HK, Ray T (2018) Multiple surrogate-assisted many-objective optimization for computationally expensive engineering design. J Mech Des 140(5):051403

Bliznyuk N, Ruppert D, Shoemaker C, Regis R, Wild S, Mugunthan P (2008) Bayesian calibration and uncertainty analysis for computationally expensive models using optimization and radial basis function approximation. J Comput Graphical Stat 17(2):270–294

Box GEP, Hunter JS (1957) Multi-factor experimental designs for exploring response surfaces. Ann Math Stat 28(1):195–241

Box GEP, Behnken DW (1960) Some new three level designs for the study of quantitative variables. Technometrics 2(4):455–475

Box GE, Hunter JS (1961) The 2 k—p fractional factorial designs. Technometrics 3(3):311–351

Box GE, Draper NR (1987) Empirical model-building and response surfaces. Wiley, Hoboken

Box GE, Hunter WG, Hunter JS (1978) Statistics for experimenters. Wiley, New York

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Brereton RG, Lloyd GR (2010) Support vector machines for classification and regression. Analyst 135(2):230–267

Broomhead DS, Lowe D (1988) Multivariable functional interpolation and adaptive networks. J Complex Syst 2:321–355

Burnham KP, Anderson DR (2003) Model selection and multimodel inference: a practical information-theoretic approach. Springer, Berlin

Campolongo F, Braddock R (1999) The use of graph theory in the sensitivity analysis of the model output: a second order screening method. Reliab Eng Syst Saf 64(1):1–12

Carbonell JR (1970) AI in CAI: an artificial-intelligence approach to computer-assisted instruction. IEEE Trans Man-Mach Syst 11(4):190–202
Chaudhuri A, Haftka RT (2014) Efficient global optimization with adaptive target setting. AIAA J 52(7):1573–1578
Chauvin Y, Rumelhart DE (2013) Backpropagation: theory, architectures, and applications. Psychology Press, London
Chen J, Lin DK (1998) On the identifiability of a supersaturated design. J Statis Plan Inference 72(1–2):99–107
Chen VC, Tsui K-L, Barton RR, Allen JK (2003) Ch. 7. A review of design and modeling in computer experiments. Handb Stat 22:231–261
Chen VC, Tsui K-L, Barton RR, Meckesheimer M (2006) A review on design, modeling and applications of computer experiments. IIE Trans 38(4):273–291
Cheng B, Titterington DM (1994) Neural networks: a review from a statistical perspective. Stat Sci 9(1):2–54
Cheng RC, Currie CS (2004) Optimization by simulation metamodelling methods. Proceedings of the 36th conference on Winter simulation, Winter Simulation Conference
Cho H, Bae S, Choi KK, Lamb D, Yang R-J (2014) An efficient variable screening method for effective surrogate models for reliability-based design optimization. Struct Multidiscip Optim 50(5):717–738
Clarke SM, Griebsch JH, Simpson TW (2004) Analysis of support vector regression for approximation of complex engineering analyses. J Mech Des 127(6):1077–1087
Corchado E, Corchado E, Corchado JM, Abraham A (2007) Innovations in hybrid intelligent systems (Advances in Soft Computing). Springer, Berlin
Cutler A, Cutler DR, Stevens JR (2012) Random forests. Ensemble machine learning. Springer, Berlin, pp 157–175
De'ath G (2007) Boosted trees for ecological modeling and prediction. Ecology 88(1):243–251
Deaton JD, Grandhi RV (2014) A survey of structural and multidisciplinary continuum topology optimization: post 2000. Struct Multidiscip Optim 49(1):1–38
Deb K, Myburgh C (2016) Breaking the billion-variable barrier in real-world optimization using a customized evolutionary algorithm. Proceedings of the 2016 on Genetic and Evolutionary Computation Conference, ACM
Deb K, Hussein R, Roy P, Toscano G (2017) Classifying metamodeling methods for evolutionary multi-objective optimization: first results. Evolutionary multi-criterion optimization: 9th International Conference, EMO 2017, Münster, Germany, March 19–22, 2017, Proceedings. Trautmann H, Rudolph G, Klamroth K, Schütze O, Wiecek M, Jin Y, Grimme C Springer, Cham 160–175
Demuth HB, Beale MH, De Jess O, Hagan MT (2014) Neural network design. Pws Pub, Boston
Deng H, Shao W, Ma Y (2012a) A Bayesian approach to kriging metamodeling for computer experiments. J Adv Mech Des Syst Manuf 6(4):404–419
Deng H, Shao W, Ma Y, Wei Z (2012b) Bayesian metamodeling for computer experiments using the Gaussian Kriging models. Qual Reliab Eng Int 28(4):455–466
Dey A (1985) Orthogonal fractional factorial designs. Wiley, NY
Dey S, Mukhopadhyay T, Adhikari S (2017) Metamodel based high-fidelity stochastic analysis of composite laminates: a concise review with critical comparative assessment. Compos Struct 171:227–250
Dobler CP, Anderson-Cook CM (2005) Forecasting, time series, and regression: an applied approach. Am Stat 59(3):278
Draper NR, Guttman I (1988) An index of rotatability. Technometrics 30(1):105–111
Duda RO, Hart PE, Stork DG (2012) Pattern classification. Wiley, NY
Dumais S, Platt J, Heckerman D, Sahami M (1998) Inductive learning algorithms and representations for text categorization. Proceedings of the seventh international conference on Information and knowledge management, ACM
Easterling RG, Berger JO (2002) Statistical foundations for the validation of computer models. Presented at computer model verification and validation in the 21st century workshop. Johns Hopkins University, Baltimore
Ezhilsabareesh K, Rhee SH, Samad A (2018) Shape optimization of a bidirectional impulse turbine via surrogate models. Eng Appl Comput Fluid Mech 12(1):1–12
Fang J, Sun G, Qiu N, Kim NH, Li Q (2017) On design optimization for structural crashworthiness and its state of the art. Struct Multidiscip Optim 55(3):1091–1119
Forrester AIJ, Keane AJ (2009) Recent advances in surrogate-based optimization. Prog Aerosp Sci 45(1):50–79
Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232
Friedman J, Hastie T, Tibshirani R (2001) The elements of statistical learning. Springer, New York
Gano SE, Renaud JE, Martin JD, Simpson TW (2006) Update strategies for kriging models used in variable fidelity optimization. Struct Multidiscip Optim 32(4):287–298
Gao H, Breitkopf P, Coelho RF, Xiao M (2018) Categorical structural optimization using discrete manifold learning approach and custom-built evolutionary operators. Struct Multidiscip Optim 58:215–228
Geisser S (1993) Predictive inference. CRC Press, Boca Raton
Geist A (2000) PVM–parallel virtual machine : a users' guide and tutorial for networked parallel computing. MIT Press, Cambridge
Ghanem RG, Doostan A (2006) On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data. J Comput Phys 217(1):63–81
Giunta AA, Dudley JM, Narducci R, Grossman B, Haftka RT, Mason WH, Watson LT (1994) Noisy aerodynamic response and smooth approximations in HSCT design. 5th Symposium on Multidisciplinary Analysis and Optimization, American Institute of Aeronautics and Astronautics
Glaz B, Goel T, Liu L, Friedmann PP, Haftka RT (2009) Multiple-surrogate approach to helicopter rotor blade vibration reduction. AIAA J 47(1):271–282
Goel T, Haftka R, Queipo N, Shyy W (2006). Performance estimate and simultaneous application of multiple surrogates. 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference
Goel T, Haftka RT, Shyy W, Queipo NV (2007) Ensemble of surrogates. Struct Multidiscip Optim 33(3):199–216
Gorissen D, Dhaene T, Turck FD (2009) Evolutionary model type selection for global surrogate modeling. J Mach Learn Res 10:2039–2078
Gorissen D, Couckuyt I, Demeester P, Dhaene T, Crombecq K (2010) A surrogate modeling and adaptive sampling toolbox for computer based design. J Mach Learn Res 11:2051–2055
Grama A (2013) Introduction to parallel computing. Addison-Wesley, Harlow
Gunn SR (1998) Support vector machines for classification and regression. ISIS Tech Rep 14:85–86
Gunst RF (1996) Response surface methodology: process and product optimization using designed experiments. Taylor & Francis Group, Routledge
Gunst RF, Mason RL (2009) Fractional factorial design. Wiley Interdiscip Rev 1(2):234–244

Güntert P, Billeter M, Ohlenschläger O, Brown LR, Wüthrich K (1998) Conformational analysis of protein and nucleic acid fragments with the new grid search algorithm FOUND. J Biomol NMR 12(4):543–548
Gustafsson A, Herrmann A, Huber F (2013) Conjoint measurement: methods and applications. Springer, Berlin
Gutmann H-M (2001) A radial basis function method for global optimization. J Global Optim 19(3):201–227
Habib A, Kumar Singh H, Ray T (2017) A multiple surrogate assisted evolutionary algorithm for optimization involving iterative solvers. Eng Optim 50:1–20
Haftka RT, Villanueva D, Chaudhuri A (2016) Parallel surrogate-assisted global optimization with expensive functions—a survey. Struct Multidiscip Optim 54(1):3–13
Hammond C, Turner CJ (2010) Data Modeling using NURBS curves and modified genetic algorithms. Proceedings of the ASME 2010 International Mechanical Engineering Congress and Exposition. vol 3. Design and Manufacturing, Parts A and B, Vancouver, British Columbia, Canada, November 12–18, pp 1089–1098
Han J, Pei J, Kamber M (2011) Data mining: concepts and techniques. Elsevier, Amsterdam
Hansen PC, Pereyra V, Scherer G (2012) Least squares data fitting with applications. JHU Press, Baltimore
Hasselman T, Yap K, Lin C, Cafeo J (2005) A case study in model improvement for vehicle crashworthiness simulation. 23rd International Modal Analysis Conference
He X, Wang L, Hong HG (2013) Quantile-adaptive model-free variable screening for high-dimensional heterogeneous data. Ann Statist 41(1):342–369
Hedayat AS, Sloane NJA, Stufken J (2012) Orthogonal arrays: theory and applications. Springer, Berlin
Helton JC (1999) Uncertainty and sensitivity analysis in performance assessment for the Waste Isolation Pilot Plant. Comput Phys Commun 117(1):156–180
Higdon D, Kennedy M, Cavendish JC, Cafeo JA, Ryne RD (2004) Combining field data and computer simulations for calibration and prediction. SIAM J Sci Comput 26(2):448–466
Ho TK (1995) Random decision forests. Document analysis and recognition, 1995, proceedings of the third international conference on, IEEE
Ho TK (1998) The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 20(8):832–844
Hong C, Liao KP, Cai T (2019) Semi-supervised validation of multiple surrogate outcomes with application to electronic medical records phenotyping. Biometrics 75(1):78–89
Hou Z, Dai Z, Lao W, Wang Y, Lu W (2019) Application of mixed-integer nonlinear optimization programming based on ensemble surrogate model for dense nonaqueous phase liquid source identification in groundwater. Environ Eng Sci. https://doi.org/10.1089/ees.2018.0366
Hsu C-W, Chang C-C, Lin C-J (2003) A practical guide to support vector classification. Department of Computer Science and Information Engineering, National Taiwan University, Taipei
IT J (2002) Principal component analysis. Springer, New York
Jaeckel P, Rebonato R (1999) The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes. J Risk 2(2):17–28
Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, Bernstein JA, Bejerano G (2016) M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet 48(12):1581
Jiang Z, Chen S, Apley DW, Chen W (2016) Reduction of epistemic model uncertainty in simulation-based multidisciplinary design. J Mech Des 138(8):081403–081413
Jin Y (2005) A comprehensive survey of fitness approximation in evolutionary computation. Soft Comput 9(1):3–12
Jin R, Chen W, Simpson T (2000) Comparative studies of metamodeling techniques under multiple modeling criteria. 8th Symposium on Multidisciplinary Analysis and Optimization, American Institute of Aeronautics and Astronautics
Jin R, Chen W, Simpson TW (2001) Comparative studies of metamodelling techniques under multiple modelling criteria. Struct Multidiscip Optim 23(1):1–13
Jones DR (2001) A taxonomy of global optimization methods based on response surfaces. J Global Optim 21(4):345–383
Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. J Global Optim 13(4):455–492
Karson MJ, Manson AR, Hader RJ (1969) Minimum bias estimation and experimental design for response surfaces. Technometrics 11(3):461–475
Karwan MH, Rardin RL (1980) Searchability of the composite and multiple surrogate dual functions. Operations Res 28(5):1251–1257
Kathleen HVB, Cox DR (1962) Some systematic supersaturated designs. Technometrics 4(4):489–495
Kennedy MC, O'Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. Biometrika 87(1):1–13
Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. J R Stat Soc Ser B (Statistical Methodology) 63(3):425–464
Khuri A (1988) A measure of rotatability for response-surface designs. Technometrics 30(1):95–104
Khuri AI, Mukhopadhyay S (2010) Response surface methodology. Wiley Interdiscip Rev 2(2):128–149
Kianifar MR, Campean F (2019) Performance evaluation of metamodelling methods for engineering problems: towards a practitioner guide. Struct Multidiscip Optim 61:1–28
Kim S, Kim H (2016) A new metric of absolute percentage error for intermittent demand forecasts. Int J Forecast 32(3):669–679
Kleinberg E (1990) Stochastic discrimination. Ann Math Artif Intell 1(1):207–239
Kleinberg E (1996) An overtraining-resistant stochastic modeling method for pattern recognition. Ann Stat 24(6):2319–2349
Kleinberg EM (2000) On the algorithmic implementation of stochastic discrimination. IEEE Trans Pattern Anal Mach Intell 22(5):473–490
Kleijnen JP (2009) Kriging metamodeling in simulation: A review. Eur J Oper Res 192(3):707–716
Koch PN, Simpson TW, Allen JK, Mistree F (1999) Statistical approximations for multidisciplinary design optimization: the problem of size. J Aircr 36(1):275–286
Krauss C, Do XA, Huck N (2017) Deep neural networks, gradient-boosted trees, random forests: statistical arbitrage on the S&P 500. Eur J Oper Res 259(2):689–702
La Fuente D, Andres R (2016) Simulation metamodeling with gaussian process: a numerical study. Industrial Engineering, North Carolina State University, Raleigh
Laurent L, Le Riche R, Soulier B, Boucard P-A (2017) An overview of gradient-enhanced metamodels with applications. Arch Comput Methods Eng 26:61–106
Lemercier B, Lacoste M, Loum M, Walter C (2012) Extrapolation at regional scale of local soil knowledge using boosted classification trees: a two-step approach. Geoderma 171:75–84
Lewis K, Mistree F (1998) The other side of multidisciplinary design optimization: accommodating a multi-objective, uncertain and non-deterministic world. Eng Optim 31(2):161–189
Li K, Deb K, Altinoz T, Yao X (2017) Empirical investigations of reference point based methods when facing a massively large number of objectives: First results. International Conference on Evolutionary Multi-Criterion Optimization, Springer

Liang C, Mahadevan S (2016) Stochastic multidisciplinary analysis with high-dimensional coupling. AIAA J 54(2):1209–1219

Lim D, Ong Y-S, Jin Y, Sendhoff B (2007) A study on metamodeling techniques, ensembles, and multi-surrogates in evolutionary computation. Proceedings of the 9th annual conference on Genetic and evolutionary computation. London, England, ACM: 1288–1295

Lin DK (1993) A new class of supersaturated designs. Technometrics 35(1):28–31

Lin CD, Anderson-Cook CM, Hamada MS, Moore LM, Sitter RR (2015) Using genetic algorithms to design experiments: a review. Qual Reliab Eng Int 31(2):155–167

Liu K, Tovar A, Nutwell E, Detwiler D (2015) Thin-walled compliant mechanism component design assisted by machine learning and multiple surrogates, SAE World Congress and Exposition

Lucas JM (1994) Using response surface methodology to achieve a robust process. J Qual Technol 26(4):248–260

Luo P, Cai Z, Geng Z (2019) Criteria for multiple surrogates. Stat Sinica 29(3):1343–1366

Lv Z, Zhao J, Wang W, Liu Q (2019) A multiple surrogates based PSO algorithm. Artif Intell Rev 52:2169–2190

MacCalman A, Lesinski G, Goerger S (2016) Integrating external simulations within the model-based systems engineering approach using statistical metamodels. Proced Comput Sci 95:436–441

Mack Y, Goel T, Shyy W, Haftka R, Queipo N (2005) Multiple surrogates for the shape optimization of bluff body-facilitated mixing. 43rd AIAA Aerospace Sciences Meeting and Exhibit, American Institute of Aeronautics and Astronautics

Madala HR, Ivakhnenko AG (1994) Inductive learning algorithms for complex systems modeling. CRC Press, Boca Raton

Maier HR, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ Model Softw 15(1):101–124

Malavé V, Turner CJ (2011) Metamodel-assisted ice detection for wind turbine blades. International Mechanical Engineering Congress and Exposition. IMECE2011–62480: 565–572.

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 21(2):239–245

Messer M, Panchal JH, Krishnamurthy V, Klein B, Yoder PD, Allen JK, Mistree F (2010) Model selection under limited information using a value-of-information-based indicator. J Mech Des 132(12):121008

Michalski RS (1983) A theory and methodology of inductive learning. Mach Learn, I edn. Elsevier, Amsterdam, p 83

Montgomery DC (2017) Design and analysis of experiments. Wiley, NY

Mugunthan P, Shoemaker CA (2006) Assessing the impacts of parameter uncertainty for computationally expensive groundwater models. Water Resour Res 42(10):1–15

Myers RH, Montgomery DC, Anderson-Cook CM (2016) Response surface methodology: process and product optimization using designed experiments. Wiley, NY

Nguyen N-K (1996) An algorithmic approach to constructing supersaturated designs. Technometrics 38(1):69–73

Owen AB (1992) Orthogonal arrays for computer experiments, integration and visualization. Stat Sinica 2(2):439–452

Paiva RM, Carvalho ARD, Crawford C, Suleman A (2010) Comparison of surrogate models in a multidisciplinary optimization framework for wing design. AIAA J 48(5):995–1006

Panchal JH, Paredis CJJ, Allen JK, Mistree F (2008) A value-of-information based approach to simulation model refinement. Eng Optim 40(3):223–251

Panchal JH, Paredis CJ, Allen JK, Mistree F (2009) Managing design-process complexity: a value-of-information based approach for scale and decision decoupling. J Comput Inf Sci Eng 9(2):021005

Peherstorfer B, Willcox K, Gunzburger M (2018) Survey of multifidelity methods in uncertainty propagation, inference, and optimization. Siam Rev 60(3):550–591

Phoa FKH, Pan Y-H, Xu H (2009) Analysis of supersaturated designs via the Dantzig selector. J Stat Plan Inference 139(7):2362–2372

Picheny V, Ginsbourger D, Roustant O, Haftka RT, Kim N-H (2010) Adaptive designs of experiments for accurate approximation of a target region. J Mech Des 132(7):071008–071009

Pickett B, Turner CJ (2011) A review and evaluation of existing adaptive sampling criteria and methods for the creation of nurbs-based metamodels. International Design Engineering Technical Conferences, DETC2011-47288

Qian PZ (2009) Nested Latin hypercube designs. Biometrika 96(4):957–970

Qian Z, Seepersad CC, Joseph VR, Allen JK, Wu CJ (2006) Building surrogate models based on detailed and approximate simulations. J Mech Des 128(4):668–677

Qian PZG, Wu H, Wu CJ (2008) Gaussian process models for computer experiments with qualitative and quantitative factors. Technometrics 50(3):383–396

Qian PZ, Ai M, Wu CJ (2009a) Construction of nested space-filling designs. Ann Stat 37(6A):3616–3643

Qian PZ, Tang B, Wu CJ (2009b) Nested space-filling designs for computer experiments with two levels of accuracy. Stat Sinica 19(1):287–300

Rabitz H (1989) Systems analysis at the molecular scale. Science 246(4927):221–226

Ratto M, Castelletti A, Pagano A (2012) Emulation techniques for the reduction and sensitivity analysis of complex environmental models. Elsevier, Amsterdam

Reese CS, Wilson AG, Hamada M, Martz HF, Ryan KJ (2004) Integrated analysis of computer and physical experiments. Technometrics 46(2):153–164

Regis RG, Shoemaker CA (2007) A stochastic radial basis function method for the global optimization of expensive functions. INFORMS J Comput 19(4):497–509

Ripley BD (2007) Pattern recognition and neural networks. Cambridge University Press, Cambridge

Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. Stat Sci 4(4):409–435

Sadaghiani M, Alizadeh R, Bahrami M (2014) Scenario-based planning for energy foresight case study: Iran's transportation industry. The 10th international energy conference (IEC 2014) Tehran, Iran, Ministry of Energy

Samad A, Kim K-Y, Goel T, Haftka RT, Shyy W (2006) Shape optimization of turbomachinery blade using multiple surrogate models. Fluids Engineering Division Summer Meeting: FEDSM2006–98368

Samad A, Lee K-D, Kim K-Y, Haftka R (2007) Application of multiple-surrogate model to optimization of a dimpled channel. 7th World Congress on Structural and Multidisciplinary Optimization. pp. 2276–2282

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117

Seni G, Elder JF (2010) Ensemble methods in data mining: improving accuracy through combining predictions. Synth Lect Data Min Knowl Discov 2(1):1–126

Shan S, Wang GG (2010) Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. Struct Multidiscip Optim 41(2):219–241

Shankar Bhattacharjee K, Kumar Singh H, Ray T (2016) Multi-objective optimization with multiple spatially distributed surrogates. J Mech Des 138(9):091401–091410

Shi R, Liu L, Long T, Liu J (2016) An efficient ensemble of radial basis functions method based on quadratic programming. Eng Optim 48(7):1202–1225

Sim DYY, Teh CS, Ismail AI (2018) Improved boosted decision tree algorithms by adaptive a priori and post-pruning for predicting obstructive sleep apnea. Adv Sci Lett 24(3):1680–1684

Simpson TW, Poplinski JD, Koch PN, Allen JK (1997) On the use of statistics in design and the implications for deterministic computer experiments. ASME Design Engineering Technical Conferences, Sacramento, California

Simpson T, Mistree F, Korte, Mauery T (1998) Comparison of response surface and kriging models for multidisciplinary design optimization. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, American Institute of Aeronautics and Astronautics

Simpson TW, Poplinski JD, Koch PN, Allen JK (2001) Metamodels for computer-based engineering design: survey and recommendations. Eng Comput 17(2):129–150

Simpson TW, Booker AJ, Ghosh D, Giunta AA, Koch PN, Yang R-J (2004) Approximation methods in multidisciplinary analysis and optimization: a panel discussion. Struct Multidiscip Optim 27(5):302–313

Simpson T, Toropov V, Balabanov V, Viana F (2008) Design and analysis of computer experiments in multidisciplinary design optimization: A review of how far we have come-or not. 12th AIAA/ISSMO multidisciplinary analysis and optimization conference

Sobieszczanski-Sobieski J, Haftka RT (1997) Multidisciplinary aerospace design optimization: survey of recent developments. Struct Optim 14(1):1–23

Sobol IM (1993) Sensitivity analysis for nonlinear mathematical models, mathematical modelling and computational experiments. Math Model Comput Exp 1(4):407–414

Song X, Lv L, Li J, Sun W, Zhang J (2018) An advanced and robust ensemble surrogate model: extended adaptive hybrid functions. J Mech Des 140(4):041402

Stein ML (2012) Interpolation of spatial data: some theory for kriging. Springer, Berlin

Steuben JC (2014) Massively parallel engineering simulations on graphics processors: parallelization, synchronization, and approximation. Colorado School of Mines, Golden

Steuben J, Turner CJ (2010) Robust optimization exploration using nurbs-based metamodeling techniques. International Design Engineering Technical Conferences International Design Engineering Technical Conferences, DETC2010–28226

Steuben J, Turner CJ (2011a). Robust optimization and analysis of nurbs-based metamodels using graph theory. International Design Engineering Technical Conferences DETC2011–47217: 587–598

Steuben J, Turner CJ (2011b) Waypoint-based robot navigation using NURBs-based metamodels. International Mechanical Engeineering Congress and Exposition. IMECE2011–62450: 1123–1133

Steuben JC, Turner CJ (2012) Robust optimization of mixed-integer problems using NURBS-based metamodels. J Comput Inf Sci Eng 12(4):041010–041017

Steuben J, Turner C (2014) Adaptive surrogate-model fitting using error monotonicity. International Design Engineering Technical Conferences, DETC2014–35037

Steuben JC, Turner CJ (2015) Graph analysis of non-uniform rational B-spline-based metamodels. Eng Optim 47(9):1157–1176

Steuben JC, Turner CJ, Crawford RH (2013) Robust engineering design optimization with non-uniform rational B-splines-based metamodels. Eng Optim 45(7):767–786

Steuben J, Michopoulos J, Iliopoulos A, Turner C (2015a) Inverse characterization of composite materials via surrogate modeling. Compos Struct 132:694–708

Steuben J, Michopoulos JG, Iliopoulos A, Turner C (2015b) Towards real-time composite material characterization using surrogate models and GPGPU computing. International Design Engineering Technical Conferences, DETC2015–46476

Steuben JC, Birnbaum AJ, Michopoulos JG, Iliopoulos AP (2019) Enriched analytical solutions for additive manufacturing modeling and simulation. Addit Manuf 25:437–447

Trucano TG, Swiler LP, Igusa T, Oberkampf WL, Pilch M (2006) Calibration, validation, and sensitivity analysis: what's what. Reliab Eng Syst Saf 91(10):1331–1357

Turner CJ (2005) HyPerModels: hyperdimensional performance models for engineering design. The University of Texas at Austin, Texas

Turner CJ (2010) Diagnosis via NURBs metamodel. International Mechanical Engineering Congress And Exposition IMECE2010–3823: 341–346

Turner CJ (2011) Metamodeling in product and process design. International Design Engineering Technical Conferences, DETC2011–47833: 639–651

Turner CJ, Crawford RH (2005a) Adapting non-uniform rational B-spline fitting approaches to metamodeling. International Design Engineering Technical Conferences, DETC2005:85544: 165–178

Turner CJ, Crawford RH (2005b) Selecting an appropriate metamodel: the case for NURBs metamodels. International Design Engineering Technical Conferences, DETC2005–85043: 759–771

Turner CJ, Crawford RH (2009) N-Dimensional Nonuniform Rational B-Splines for Metamodeling. J Comput Inf Sci Eng 9(3):031002–031013

Turner CJ, Campbell MI, Crawford RH (2003) Generic sequential sampling for metamodel approximations. International Design Engineering Technical Conferences, DETC2004–48230: 555–564

Turner CJ, Campbell MI, Crawford RH (2004) Metamodel defined multidimensional embedded sequential sampling criteria. International Design Engineering Technical Conferences, DETC2004–57722: 703–714

Turner CJ, Ajetunmobi AM, Crawford RH (2006) Fault detection with NURBs-based metamodels. International Design Engineering Technical Conferences, DETC2004–99637: 439–450

Turner CJ, Crawford RH, Campbell MI (2007a) Global optimization of NURBs-based metamodels. Eng Optim 39(3):245–269

Turner CJ, Crawford RH, Campbell MI (2007b) Mixed integer optimization with NURBs HyPerModels. Internationa Design Engineering Technical Conferences. DETC2007–35852: 371–379

Turner CJ, Crawford RH, Campbell MI (2007c) Multidimensional sequential sampling for NURBs-based metamodel development. Eng Comput 23(3):155–174

Turner C, Pickett B, Steuben J (2011) Design space analysis with hyperdimensional metamodels. Proceedings of the 2011 NSF-CMMI Grantees Conference, Atlanta, GA

Tutum CC, Deb K (2015) A multimodal approach for evolutionary multi-objective optimization (MEMO): proof-of-principle results. International Conference on Evolutionary Multi-Criterion Optimization, Springer

Vapnik V (2013) The nature of statistical learning theory, Springer, Berlin

Varadarajan S, Chen WEI, Pelka CJ (2000) Robust concept exploration of propulsion systems with enhanced model approximation capabilities. Eng Optim 32(3):309–334

Venter G, Haftka R, Chirehdast M, Venter G, Haftka R, Chirehdast M (1997) Response surface approximations for fatigue life prediction. 38th Structures, Structural Dynamics, and Materials Conference, American Institute of Aeronautics and Astronautics

Viana FAC, Haftka RT (2008a). Using multiple surrogates for metamodeling. Proceedings of the 7th ASMO-UK/ISSMO International Conference on Engineering Design Optimization

Viana FAC, RT Haftka (2008b) Using multiple surrogates for minimization of the rms error in meta-modeling. International Design Engineering Technical Conferences, DETC2008–49240, 851–860
Viana F, Haftka R, Steffen JV, Butkewitsch S, Leal MF (2008) Optimal use of multiple surrogate for reduced RMS error in meta-model. Proceedings of 2008 NSF Engineering Research and Innovation Conference, Knoxville, Tennessee
Viana FA, Picheny V, Haftka RT (2009) Conservative prediction via safety margin: design through cross-validation and benefits of multiple surrogates. ASME 2009 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers
Viana F, Haftka R, Watson L (2010a) Why not run the efficient global optimization algorithm with multiple surrogates? In 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 18th AIAA/ASME/AHS Adaptive Structures Conference 12th, p 3090
Viana FA, Gogu C, Haftka RT (2010b) Making the most out of surrogate models: tricks of the trade. In ASME 2010 International Design Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers Digital Collection, pp 587–598
Viana FAC, Picheny V, Haftka RT (2010c) Using cross validation to design conservative surrogates. AIAA J 48(10):2286–2298
Viana FA, Haftka RT, Watson LT (2013) Efficient global optimization algorithm assisted by multiple surrogate techniques. J Global Optim 56(2):669–689
Viana FAC, Simpson TW, Balabanov V, Toropov V (2014) Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come? AIAA J 52(4):670–690
Villanueva D, Haftka RT, Le Riche R, Picard G (2013) Locating multiple candidate designs with surrogate-based optimization. 10th World Congress on structural and multidisciplinary optimization, Orlando, USA, May, Citeseer
Vinzi V, Chin WW, Henseler J, Wang H (2010) Handbook of partial least squares. Springer, Berlin
Wang C, Duan Q, Gong W, Ye A, Di Z, Miao C (2014) An evaluation of adaptive surrogate modeling based optimization with two benchmark problems. Environ Model Softw 60:167–179
Wang H, Ye F, Li E, Li G (2016) A comparative study of expected improvement-assisted global optimization with different surrogates. Eng Optim 48(8):1432–1458
Witten IH, Frank E, Hall MA, Pal CJ (2016) Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, Burlington
Xing Z, Qu R, Zhao Y, Fu Q, Ji Y, Lu W (2019) Identifying the release history of a groundwater contaminant source based on an ensemble surrogate model. J Hydrol 572:501–516
Xiong Y, Chen W, Tsui K-L, Apley DW (2009) A better understanding of model updating strategies in validating engineering models. Comput Methods Appl Mech Eng 198(15):1327–1337
Xiong S, Qian PZ, Wu CJ (2013) Sequential design and analysis of high-accuracy and low-accuracy computer codes. Technometrics 55(1):37–46
Xu J, Zeger SL (2001) The evaluation of multiple surrogate endpoints. Biometrics 57(1):81–87
Xue Z, Lietz R, Rigoni E, Parashar S, Kansara S (2013) RSM improvement methods for computationally expensive industrial CAE analysis. 10th World Congress on Structural and Multidisciplinary Optimization
Yang T, Lin H-C, Chen M-L (2006) Metamodeling approach in solving the machine parameters optimization problem using neural network and genetic algorithms: a case study. Robot Comput Integr Manuf 22(4):322–331
Ye P, Pan G, Dong Z (2018) Ensemble of surrogate based global optimization methods using hierarchical design space reduction. Struct Multidiscip Optim 58(2):537–554
Yin H, Fang H, Wen G, Gutowski M, Xiao Y (2018) On the ensemble of metamodels with multiple regional optimized weight factors. Struct Multidiscip Optim 58(1):245–263
Zhao J, Liu Q, Pedrycz W, Li D (2012) Effective noise estimation-based online prediction for byproduct gas system in steel industry. IEEE Trans Industr Inf 8(4):953–963
Zhu Z, Du X (2016) Reliability analysis with Monte Carlo simulation and dependent Kriging predictions. J Mech Des. https://doi.org/10.1115/1.4034219

Affiliations

Reza Alizadeh[1] · Janet K. Allen[1] · Farrokh Mistree[2]

Reza Alizadeh
reza.alizadeh@ou.edu

Farrokh Mistree
farrokh.mistree@ou.edu

[1] The Systems Realization Laboratory, Industrial and Systems Engineering, University of Oklahoma, Norman, Oklahoma, USA

[2] The Systems Realization Laboratory, Aerospace and Mechanical Engineering, University of Oklahoma, Norman, Oklahoma, USA