On The Use of Surrogate Models in Engineering Design Optimization and Exploration: The Key Issues

Pramudita Satria Palar
Faculty of Mechanical and Aerospace Engineering
Institut Teknologi Bandung
Bandung, Indonesia
pramsp@ftmd.itb.ac.id

Rhea Patricia Liem
Department of Mechanical and Aerospace Engineering
The Hong Kong University of Science and Technology
Hong Kong, China
rpliem@ust.hk

Lavi Rizki Zuhal
Faculty of Mechanical and Aerospace Engineering
Institut Teknologi Bandung
Bandung, Indonesia
lavirz@ftmd.itb.ac.id

Koji Shimoyama
Institute of Fluid Science, Tohoku University
Sendai, Japan
shimoyama@tohoku.ac.jp

ABSTRACT

Surrogate models are invaluable tools that greatly assist the process of computationally expensive analyses and optimization. Engineering optimization reaps the benefit from surrogate models in order to perform expensive optimization that could potentially be computationally intractable in the pre-high-performance computing age. Moreover, surrogate models provide a means to allow engineering design exploration with high-fidelity computer simulations. Despite their wide use and substantial research progresses, there are still some key issues and challenges that need to be addressed by researchers. Most of these issues stem from the growing complexity of engineering design optimization and exploration in real-world problems. In other words, the sophistication of the problem that we have to tackle increases faster than that of computing power and technology. It is thus imperative to have accurate and yet computationally efficient surrogate models that are suitable for real-world engineering problems. In this paper, we discuss key issues and challenges of the application of surrogate models in engineering design optimization and exploration. This paper is directed toward general readers, in which we aim to present general discussions regarding the effectiveness, issues, and future of surrogate-based optimization and exploration in engineering.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability;

KEYWORDS

ACM proceedings, LaTeX, text tagging

GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic

ACM ISBN 978-1-4503-6748-6/19/07…$15.00
https://doi.org/10.1145/3319619.3326813

ACM Reference Format:
Pramudita Satria Palar, Rhea Patricia Liem, Lavi Rizki Zuhal, and Koji Shimoyama. 2019. On The Use of Surrogate Models in Engineering Design Optimization and Exploration: The Key Issues. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion), July 13–17, 2019, Prague, Czech Republic.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3319619.3326813

1 INTRODUCTION

Computational design optimization and exploration methodology have played an increasingly bigger role in engineering. An optimization aims to produce optimal designs that improve the performance over the baseline designs so as to gain benefits such as lower operating cost and higher efficiency. On the other hand, the aim of design exploration is to gain knowledge, insight into, and physical understanding of the system that eventually lead to a better design process. Optimization and exploration go hand-in-hand to produce both optimized design solutions and also important design insight. In many instances where exhaustive real experiments are not possible, computer simulations that can help predict design performance become an indispensable component in design optimization and exploration. However, an accurate optimization design and process requires high-fidelity computer simulations, which can make the implementation of population-based optimization methods such as genetic algorithms computationally impractical. To that end, it is now common to deploy surrogate models, i.e., mathematical models that can provide cheaper evaluations of the input-output relationship of a system, in such computationally expensive tasks.

Surrogate-based optimization (SBO) has been demonstrated in various engineering fields, ranging from aerospace, chemical, mechanical, and environmental engineering, to name a few. Research activities in its method and algorithmic development goes hand-in-hand with those in the engineering applications. The engineering community normally works by a problem-driven philosophy, i.e., the optimization method should be able to deal with various complexities of the current real-world problems at hand. On the other hand, other communities such as evolutionary computation and applied mathematics communities primarily focus on the method and algorithmic development with fewer concerns on actual real-world challenges faced by the engineering community. Both research directions are equally critical, and have seen significant development in recent years. We note, however, that there are still gaps between the two research communities. The algorithmic development might not be directly applicable to real-world problems. Likewise, not all complexities in real-world problems are adequately addressed by the existing methods. We strongly believe that identifying these gaps is the first key step to bridge these two different research communities, which can in turn further accelerate the advancement of the engineering optimization and exploration and their applications. As an example, researchers typically assume that all solutions will return objective function values, while in reality some of these solutions might fail due to convergence issues. To that end, it is important to raise awareness regarding the key issues, concerns, and challenges faced in typical engineering design optimization and exploration problems.

The main objective of this paper is to summarize the key issues and challenges concerning the applications of SBO techniques in real-world problems. Instead of focusing on a particular field of application, we look into the common pitfalls and challenges that are applicable to various engineering disciplines. We write this paper to address both research communities, that is, we hope that this paper will introduce the key issues in engineering design optimization and exploration to researchers who work in algorithmic development and also applications. We certainly hope that this paper can help researchers identify the key research focus in the field of SBO and its implementations, and to set the directions of the future research.

2 SURROGATE-BASED METHODS

In this section, we discuss the common applications of surrogate models in engineering design analyses. Although our main discussion focuses on engineering design optimization and exploration, it is also important to briefly examine other applications of surrogate models to give readers an awareness of the other functionalities of surrogate models in engineering. We then also discuss how surrogate models can be used within various optimization algorithm frameworks for aiding design optimization and exploration.

2.1 Applications of surrogate models

In general, surrogate models of lower complexity that are inexpensive to evaluate and approximate accurately a large-scale model can greatly facilitate computationally intensive analysis tasks at hand. Such analyses include optimizations, which are typically performed iteratively, or by relying on population-based algorithms. Other examples include analyses that involve Monte Carlo simulations that require many thousands of scenarios to characterize the effects of the uncertainty in the system. These analyses, when using the large-scale model, are intractable. The prediction of surrogate models aids designers in performing a wide range of analysis with an affordable number of computer simulations.

2.1.1 Design optimization and exploration. One of the most important applications of surrogate models is design optimization. In design optimization, surrogate models are used to model the relationship of the input variables and the objective and constraint functions, in which the exact relationship is difficult or even impossible to obtain. Often, real-world engineering optimization problems involve computationally expensive computer simulations or, in few cases, expensive physical experiments. The use of approximation models in engineering optimization is relatively old and can be traced back to structural optimization community. According to Viana et al. [107], the use of surrogate models in engineering optimization and exploration started to flourish after the publication of Sacks et al. seminal paper in the design and analysis of computer experiments [90].

Additionally, surrogate models are also invaluable tools in assisting design exploration, to gain the knowledge, insight, and physical interpretation that can be very useful in the design endeavor [30, 73]. Surrogate models can also be coupled with global sensitivity analysis tools such as the Sobol' method [98] to identify the impact of each design variable to the output. This method can provide information on the importance ranking of the inputs, based on how much their variations affect the variation of the output. Multi-objective optimization and design exploration can work hand-in-hand in the form of multi-objective design exploration (MODE) [73]. In this approach, design exploration is applied both in the global space and also among the non-dominated designs. Exploration of non-dominated designs sheds a light on the trade-off between optimal designs, which is a key information for engineers.

2.1.2 Uncertainty quantification and sensitivity analysis. Other important applications of surrogate models in engineering are uncertainty quantification (UQ) and global sensitivity analysis (GSA). Instead of treating the system deterministically, we treat the inputs and outputs as random variables. In UQ, we aim to quantify the uncertainty in the output as a function of the random inputs. On the other hand, the goal of GSA is to quantify the contribution of each random variable to the output variation, by taking into account the interactions between inputs. Monte Carlo simulations are commonly used to perform UQ and GSA. These simulations generate an ensemble of random realizations by running the model deterministically for different sets of random inputs. Using surrogate models can significantly reduce the number of required function evaluations, from the number of function evaluations required to perform the Monte Carlo simulation to the number of samples required to construct the surrogate models. When the uncertainty is described via probability theory, the surrogate model should be accurate in the entire domain of the random space. In some SBO practice, on the other hand, the surrogate modeling accuracy might only be emphasized in the vicinity of the optimal solution.

One popular surrogate model in UQ and GSA is the polynomial chaos expansion (PCE) method [12, 19, 101, 109]. PCE is powerful, as it is theoretically convergent (under certain conditions), and its polynomial coefficients directly provide the statistical moments and sensitivity indices. Kriging is also frequently used, e.g., in UQ of airfoil [96] and composite material [71]. A combination of PCE and Kriging has also been explored, as was applied in the computational dosimetry [51]. As far as we know, other surrogate models besides PCE and Kriging have not been widely used in UQ and GSA, although more careful observation is needed to confirm this hypothesis.

2.1.3 Structural reliability Analysis. Surrogate models are also used in structural reliability analysis to compute the probability of failure. In this respect, the surrogate model is used to classify the problem space into a feasible and infeasible region. In structural reliability analysis, the surrogate models need to be accurate primarily near the limit state function region, i.e., the boundary between the feasible and infeasible domain. Neural-network was one popular surrogate model for such an application [40, 100]; however, recently, PCE [39, 66] and Kriging [26] (also its polynomial-chaos Kriging variant [92]) are paving their way in this field. Kriging is particularly attractive, owing to the availability of uncertainty structures. Lu et al. recently developed an improved kriging with extremum response surface method for structural dynamic reliability and sensitivity analyses [64].

2.2 Classification of surrogate-based optimization techniques

In the following discussion, we classify SBO into several categories based on how surrogate models are used in the optimization procedure. Each classification is described briefly below. Note that this discussion is limited to those where surrogate models are directly involved in optimization procedures. Other approaches, such as the use of surrogate models to assist in the derivation of multipoint objective function formulations [58, 59], are beyond the scope of this paper.

2.2.1 Decoupled methods. This is the simplest approach, where the surrogate models and optimization algorithms are fully decoupled. The surrogate models are typically used to approximate the objective function of the optimization, and the constraints when we deal with constrained optimizations. In this case, any optimization algorithms such as metaheuristics and local search can be applied to find the optimum of the surrogate models. The optimum of the surrogate model is then evaluated and when the sequential update is adopted, the loop of surrogate building and optimization is repeated again until one runs out of budget. In addition to its simplicity, the choice of the surrogate modeling method and optimization algorithm are independent of each other in this case. Some application examples that used decoupled techniques are axial compressor blade optimization [91] and composite laminate design [72].

2.2.2 Surrogate-driven methods. In this category, the surrogate models are intrinsically built in the optimization procedure. One example is the Bayesian optimization technique which relies on the sequential updates of probabilistic surrogate models to find the optimal solution. The popular expected improvement-based efficient global optimization (EGO) [49] is one form of Bayesian optimization technique. EGO is a popular algorithm in the engineering community and has found its usages in disciplines such as robotics [14], aerospace design [34, 47, 54], and petroleum engineering [18]. Trust-region methods can also be classified as a surrogate-driven method since they rely on approximation models to solve the non-linear programming (NLP) problems [3, 4, 88]. This approach relies on the use of a local surrogate model (typically a quadratic function) to approximate the objective function around the current best solution. The algorithm then computes the step size and the improvement direction, followed by constructing a local surrogate in the new "trust region." The level of trust is determined by assessing the improvement, i.e., whether it is to subtle or even produces a negative improvement. Note that the techniques mentioned above might involve some internal optimization procedures (e.g. optimization of acquisition function in Bayesian optimization). This internal optimization procedure is independent of the surrogate-driven approach discussed here.

Gaussian processes and Kriging models are the most commonly used surrogate models for the Bayesian optimization approach [55, 95]. Viana et al. proposed to import the uncertainty structure of Kriging so that other non-probabilistic surrogates can be used in EGO algorithms [106]. Radial basis function (RBF) is also widely used as a surrogate model for surrogate-driven techniques, sometimes as a cheaper alternative to kriging, thanks to its interpolating capability and also the ability to model highly non-linear functions. Gutmann's optimization method is one of the earliest surrogate-driven techniques that utilizes RBF [32]. Following Gutmann's work, many RBF-based methods that specifically handle expensive problems were developed [36, 86, 87].

2.2.3 Surrogate-assisted methods. Unlike surrogate-driven techniques, surrogate-assisted techniques depend on surrogate models to aid the operators in other optimization techniques (e.g., metaheuristics such as evolutionary algorithms). Surrogate models play an important role in such algorithms but their role is not as principal as in surrogate-driven techniques. One prominent example is memetic algorithms that are assisted by surrogate models in local search. The surrogate-assisted memetic algorithm has been applied in multi-objective coastal aquifer management [99], car design [57], and airfoil optimization [76], to name a few. Other metaheuristics also reap benefits from surrogate models, which include the evolutionary algorithm [114], differential evolution [69], evolution strategy [50], and particle swarm optimization [102]. Readers are referred to a paper by Jin [48] for a more comprehensive review of this topic. Some advantages of surrogate-assisted techniques, particularly the population-based algorithms, are the parallelizability of the function evaluations. In the multidisciplinary optimization (MDO) community, however, surrogate-assisted metaheuristics have not been commonly used, due to the scarcity of the open source solvers and commercial implementation of such techniques.

2.3 When should we use surrogate models?

In this section, we would like to discuss about the common design problems where using surrogate models is desirable. We also wish to highlight the conditions where surrogate models give an edge compared to other classes of optimization methods (e.g. model-free metaheuristics and gradient-based methods).

2.3.1 When the budget of function evaluations is low to moderate. The effectiveness of using surrogate models largely depend on the cost of running the full-model function evaluation, the cost of constructing the surrogate model, and the available computational budget (time and resources). In the case of inexpensive function evaluations, it is better to deploy non-surrogate methods such as metaheuristics or multi-start local search. Likewise, surrogate models are not required when the objective and constraint function are evaluated by cheap and very low-fidelity solvers or analytical models. Surrogate models are the most effective when we deal with computationally costly high-fidelity function evaluations with limited computational time and computing power.

2.3.2 When the dimensionality is low to moderate. Although there is a good progress in high-dimensional and gradient-enhanced surrogate models, the current use of conventional surrogate models is still limited to problems with low to moderate dimensionality. That is, surrogate model construction and usages still suffer from the "curse-of-dimensionality". Approximating very high-dimensional problems (with thousands or even millions of decision variables [1]) still imposes serious challenges in surrogate modeling construction. Nevertheless, research is progressing in this direction, as will be further discussed in Section 3.2.

The definitions of low, moderate, and high dimensionality depend on the problem and discipline. As a rule of thumb in a general surrogate modeling context; we define low dimensional problems when the number of variables is lower than 10, moderate from 10-50, and high when the dimensionality is higher than 50. Most surrogate models are adequate for low-dimensional problems but the efficiency decreases as we add more design variables [107]. Fortunately, such a weakness of conventional surrogate models has long been realized. Researchers are working on increasing the efficiency of surrogate models in moderate to high-dimension by introducing techniques such as gradient-enhanced surrogate models, high-dimensional model representation, and dimensionality reduction. It is worth noting that one advantage of surrogate-based optimization is that it can be used with and without gradient information.

2.3.3 When design exploration is one of the main goals. Design exploration, as the name suggests, involves the exploration of various configurations and designs [30, 73]. Frequently, at the preliminary design stage, fewer design variables are consider to gain insight into the problem. Surrogate models play a key role when exploring the design space is an issue. Using surrogate models that are constructed based on high-fidelity models enables an accurate design exploration even at the preliminary stage. The accuracy of the surrogate models, of course, needs to be validated prior to the usage. Although standard optimization can also assist the design exploration process (e.g., by investigating multiple local optimum designs) surrogate models can provide insight into the effect of changing variables or parameters without running extra computer simulations. This feature is not typically available with optimization alone. Surrogate models can also give access to a quick global sensitivity analysis, that helps designers in identifying the importance of each design variable [46, 47].

2.3.4 When the design process relies on physical experiments. Surrogate models can also be conveniently deployed in physical-experiment based optimization and design exploration. This is particularly relevant when running computer simulations is too expensive due to the complex physical phenomenon, e.g., as in unsteady flow in flapping wing [16, 17]. In fact, we believe that SBO is the only efficient method to solve physical-experiment based problems. In such cases, computing the gradient information is almost impossible due to the noise and high cost. Surrogate models enable physical experiment-based optimization and exploration by providing approximation models that filter the experimental noise. However, we still need to be careful on choosing the SBO algorithm for physical-experiment based optimization. This is mostly because the majority of SBO algorithms are developed to solve computer simulation-based problems, with different sets of characteristics than those of the physical-experiment based case (e.g., deterministic outputs).

3 KEY ISSUES AND CHALLENGES

In this section, we discuss some key issues and challenges that are typically encountered by engineering optimization community when tackling real-world problems using surrogate models.

3.1 Computational budget

The selection of the suitable optimization method largely depends on the computational budget and the problem dimension. The computational budget itself depends on many factors, including the computing power and the model fidelity. Here, we assume that the computational budget is analogous with the number of available function evaluations.

Some SBO algorithms are specifically designed for a low budget (e.g. EGO) or moderate budget (e.g. surrogate-assisted memetic algorithms). Having more budget is beneficial especially in design exploration and multi-/many-objective cases since we can have a more accurate representation of the input-output relationship and the design space. The use of the standard Bayesian optimization procedure with Kriging surrogate model is prohibitive when a large number of function evaluations is required (e.g., more than 1000). This is in particular due to the exponential training time of Kriging, and the problem is further exacerbated when more than one kriging model is required. With some modifications, however, the surrogate-driven SBO can also be used with a large budget (e.g., by clustering [108]).

Surrogate-assisted metaheuristics are more suitable for moderate budget problems. The evolutionary operators of surrogate-assisted metaheuristic are also powerful in the sense that they enhance diversity and exploration, which are important features for solving multi- and many-objective optimization problems. With a moderate budget, it is possible to cover more region in the objective space with more points, allowing for a better design exploration process through data mining technique.

3.2 Dimensionality and non-linearity

The non-linearity of the problem notably affects the effectiveness of optimization algorithms. For example, the COBYLA algorithm, which uses a linear approximation, is not suitable for highly non-linear problems. If, say, we falsely assume that the function is linear, the optimization progress would fail to find the optimum point. Researchers typically assume that the problem is nonlinear and deploy methods that can specifically handle nonlinear functions. In this respect, modern surrogate models such as Kriging, RBF and support vector regression (SVR) [25] are more suitable. Note that there are also some real-world situations where the input-output relationship cannot be precisely captured by a single global surrogate model (e.g. discontinuous response); in such situations, we can use techniques such as the mixture of experts [8, 60].

However, algorithms that rely on global surrogate models will encounter significant challenges if the black-box function are extremely non-linear (e.g. Rastrigin function). On the other hand, some techniques which include surrogate-assisted memetic algorithms rely on local surrogate modeling, thus are more capable of handling extremely non-linear problems by capturing the locality of the optimization problems. Interestingly, although we typically make an assumption that the problem is non-linear, some real-world problems are indeed linear or near-linear and can be approximated well with less complex models (e.g. polynomial regression). It is suggested that, if possible, one performs preliminary analyses regarding this issue [22]. One can also use cross-validation to determine the most suitable surrogate model, especially when a global surrogate model is used.

It is worth noting that most of the current state-of-the-art surrogate models are designed to handle problems with low to moderate dimensionality. Nevertheless, we can exploit the optimization problem's hidden structures so as to make surrogate models more effective in high-dimension. In most cases, some variables will have little or no contributions to the optimum solution. It is therefore imperative to prune such variables out of the optimization process, to eliminate the unnecessary complexity of the computation. The idea of dimensionality reduction stems from this issue, which led to the development of dimensionality reduction assisted surrogate models [13, 104]. When the problem is expensive, one can employ tools such as the active subspace which can determine the directions that contribute most to the problem [22].

3.3 Paralellization

With the advent of parallel computing, parallelization is also one of the key issues in SBO. There are basically two parallelization approaches in computer simulations: the first one is to parallelize the solvers, e.g. parallel computational fluid dynamics (CFD) computation, and the second one is to run the multiple function evaluations parallely. With parallelization in one function evaluation, a single-update method such as the conventional Bayesian optimization can be directly used. However, it is worth noting that a linear speedup due to the parallelization is impossible to achieve. Metaheuristics that are assisted by surrogate models are easier for parallelization since they primarily work with the population-based principle. On the other hand, sequential-based SBO strategies such as Bayesian optimization and RBF-based methods need to be modified to enable evaluating multiple function evaluations at the same time (see [31, 53, 106, 112] for examples).

One potential challenge with parallel optimization is when the calculation times for different designs vary widely. This is typical in cases that involve unsteadiness, e.g., unsteady flow simulation, or fluid-structure interaction. The parallel optimization algorithms need to be modified when handling such problems to maximize the use of parallel computational resources. For a more in-depth discussion, readers are referred to a good review of parallel function evaluations for optimization algorithms by Haftka [33].

3.4 Multi-fidelity surrogate modeling

Fidelity can be interpreted as the extent to which one simulation scheme is closer to the truth. The different levels of fidelity are typically defined based on how closely the models represent the system's physics [111]. Although there might be various levels of fidelity, for simplicity, here we classify the fidelity into high- and low-fidelity. Low-fidelity simulations can be obtained by reducing the number of mesh elements, using simpler governing equations, and partially converged simulation. In computational optimization, high-fidelity simulations are typically more expensive but more accurate than low-fidelity simulations. A multi-fidelity surrogate model aims to find the right balance between the two levels of fidelity, by combining a large number of low-fidelity function simulations with fewer high-fidelity simulations. The main aim is to find the balance between accuracy and computational cost. Some multi-fidelity surrogate models that have been applied in engineering are multi-fidelity Kriging [35], PCE [76], and support vector regression [44]. A multi-fidelity surrogate model has also been combined with evolutionary algorithms and applied to antenna design [61]. A good review of multi-fidelity models is given by Fernandez et al. [27].

However, it is worth noting that one issue of multi-fidelity surrogate modeling is the proper selection of the low-fidelity simulation. It is possible that the low-fidelity simulations are not helpful in aiding multi-fidelity surrogate modeling, for example, if the trend of the low-fidelity simulation is inconsistent with that of the high-fidelity simulation. Therefore, it is important to perform preliminary analysis on the correlation between the low- and high-fidelity simulations prior to formulating the multi-fidelity approach [76, 103]

3.5 Robust and reliability based design optimization

The aim of robust optimization is to find optimal solutions that are robust in the presence of uncertainty. In real-world engineering design, uncertainties could come from the variations in design condition (e.g., disturbances in velocity and temperature), manufacturing error, and even uncertainties in the computer simulations themselves (i.e., model form uncertainties), to name a few [9]. Robust optimization can also be defined as an optimization process that aims to design a product or system that performs well in the design and off-design conditions (e.g. different flight conditions for aircraft design). Notice that the design variables could also be uncertain.

On the other hand, reliability-based design optimization aims to find solutions that are optimal and also satisfy the constraint in the probability of failure [10, 11]. Note that both frameworks can be combined into robust and reliability-based design optimization (RBDO). In such tasks, there are two parts where surrogate models can be useful. First, surrogate models can be used as the tool for evaluating the robustness (i.e., UQ) and the reliability of a design (i.e., structural reliability analysis) to replace the expensive Monte Carlo simulation (MCS). One popular surrogate model-based technique for reliability computation is the active Kriging-MCS (AK-MCS) [26]. This method uses Kriging to sequentially generate new samples in the limit state region by exploiting the uncertainty structure of Kriging. In order to reduce the overall computational cost for both robust optimization and RBDO, it is also possible to combine the design and random variables into one space. In such a case, only one surrogate model is needed [15, 70]. The outputs from UQ and reliability computation are typically the statistical moments (i.e., mean and standard deviation) and the probability of failure, respectively, that are used as the merit indicator of a design in the optimization process. Second, an outer loop consisting of SBO can be used in a manner similar to that of a standard deterministic optimization to optimize the statistical moments while satisfying the probability of failure.

Robust optimization and RBDO are computationally intensive processes since they couple optimization, uncertainty analysis, and reliability computation. Therefore, when the problem is not sensitive to uncertainties, performing them might not be well-justified computationally. If necessary, the performance of the deterministically optimized design can be checked post-optimization. However, RBDO is suggested if the design will encounter significant disturbances from the nominal condition.

3.6 Constraint handling

Most real-world problems are formulated as a constrained optimization problem. To that end, it is necessary to equip SBO with the suitable constraint handling capabilities. The presence of constraints might also change the Pareto front of multi- and many-objective optimization problems. For Bayesian optimization, constraints are typically handled with special techniques such as the probability of feasibility [93] or the modified version that gives a chance for solutions near the boundary of the feasible domain to be selected [6]. Basudhar et al. proposed the use of support vector machines (SVM) to approximate the boundary of the feasible domain. This method is also capable of reducing the number of constraints surrogates to just one [7]. There also exist several RBF-based methods that are equipped with constraint handling capabilities [83, 84].

The complexity of constrained problems increases as the number of constraints increase; furthermore, there is a high chance that the percentage of the feasible region will also decrease. Consequently, the complexity of surrogate-driven methods that need extra surrogate models to model the constraints (e.g., the probability of feasibility) also grows with the increasing number of constraints. To handle this, Zhang et al. investigated several constraint aggregation methods for SBO of an aircraft wing [113]. Note that the method proposed by Basudhar et al. can also handle this problem by SVM [7].

For surrogate-assisted techniques, the constraint handling should be facilitated in both the primary algorithm (see Mezura and Coello for a review of constraint handling on metaheuristics methods [68]) and the corresponding operators that use surrogate models (e.g. local search in memetic algorithms). For decoupled techniques, the constraint handling purely depends on the selected optimization framework. Surrogate models can be used to approximate the constraint functions too in this case.

3.7 Noisy optimization

Computer simulations are deterministic, where the same inputs will always produce the same outputs. The noise that is typically present in physical experiments is therefore irrelevant in such a setting. However, numerical "noise" might still exist as artifacts of the numerical scheme. Such numerical noises might yield an undesirable effect on interpolating surrogate models (e.g. overfitting). In Bayesian optimization literature, the surrogate models and infill criteria should be modified to take into account the effect of noise [43, 81, 94]. Some specialized mechanism have also been developed to enhance surrogate-based optimization techniques so as to handle noisy problems [37, 52]. When decoupled techniques are used, the constructed surrogate models should also consider the impact of noise. Thus, regression, instead of interpolating, models should be used in the presence of noise.

It is worth noting that noisy optimization is not equivalent to robust optimization (a.k.a optimization under uncertainty). In robust optimization, uncertainty in design variables or design parameters are considered but the simulation itself might be deterministic (robustness is evaluated by using several deterministic simulations). However, in noisy optimization, the function evaluation might be corrupted by noise but the design variables or design parameters are treated deterministically. Considering both cases simultaneously would significantly increase the optimization problem complexity, since it is difficult to distinguish the impact of noise and uncertain variables/parameters.

3.8 Failed and untrustworthy simulations

Failed simulations can return no valid function values, which imposes serious problems when they are used in an optimization procedure. In surrogate-assisted methods, particularly those that are based on metaheuristics, the presence of failed simulations is not a big problem since such solutions can be excluded from the selection process (e.g., by giving a very large value for minimization). When surrogate models are not used in a sequential process, e.g., for the global sensitivity analysis and design exploration purposes, the failed simulations can be discarded once they are identified. However, the presence of failed simulations imposes a serious problem in sequential surrogate-driven optimization since it might terminate the optimization process prior to reaching its optimum value.

One simple solution to tackle this problem is by performing a random simulation to continue the optimization process; however, there is still no guarantee that we will not revisit the failed region. Forrester suggested to impute the values of the prediction added by the standard deviation when using kriging as the surrogate model [29]. Such an approach is inapplicable for non-probabilistic surrogate models (e.g., SVR). Moreover, when users are unable to modify the optimization code, dealing with failed simulations becomes more challenging. In that case, users have to stop the optimization, put a random sample, and then continue the optimization process. This is inconvenient especially if the users want the optimization process to be fully automatic. When dealing with a black-box optimization code, the best that the users can do is to impute some artificial values to the failed simulations. We then suggest that more research should be dedicated to the imputation of failed simulations in surrogate-based optimization. Treatment of missing data is a specific subject of study itself and researchers can borrow the idea from the statistical research community to handle missing data in surrogate-based optimization [38, 105].

Another possible scenario is the existence of sampling points that return untrustworthy outputs. Such untrustworthy outputs might occur due to errors in meshing, where the simulation can still return a value, albeit invalid. Detecting this kind of outputs is more difficult than identifying failed simulations, so great care needs to be taken. Experts in the problem can identify such untrustworthy simulations by observing the outputs, for example, if the output is too high or too low and seems to be unrealistic. Untrustworthy outputs can also be identified by outlier detection schemes as done by Martinez et al. [67] in the context of Bayesian optimization.

3.9 Multi- and many-objective optimization

In contrast to single-objective optimization, multi- and many-objective optimization are particularly useful when engineers/researchers want to perform exploration and trade-off between competing objective functions. Typically, multi-objective refers to two or three objective functions, while many-objective considers more than three objective functions [28, 42]. One simple way to use surrogate models in multi-/many-objective optimization is to construct surrogate models for each objective function and then optimize the surrogate models using algorithms such as NSGA-II (i.e. decoupled techniques); in this case, we fully trust the model. Such a framework has been applied to the optimization of composite laminate [72], cyclone separator [97], and microchannel heat sink [41], to name a few. Another advantage of this method is that the function evaluations are easily parallelizable. However, when the input-output relationship is highly non-linear, it is not reliable to fully trust the model. In that case, the multi-objective Bayesian optimization (MOBO) is one potential sequential approach that carefully adds new solutions to construct the Pareto front. Similar to its single-objective counterpart, MOBO typically uses the Kriging model and its uncertainty structure. MOBO, with various infill criteria, has been applied to engineering cases such as aerodynamic design [77] and chemical reactor design [78], to name a few. Other iterative surrogate-based methods for multi-objective optimization are the RBF-based Gap Optimized Multi-objective Optimization using Response Surfaces (GOMORS) [2] and the Multi-Objective Constrained Stochastic optimization using Response Surfaces (MOCS-RS) [85].

In contrast to multi-objective optimization, many-objective optimization with surrogate models has not yet been widely explored and used in engineering design optimizations. This is mainly due to the increased problem complexity as a consequence of the increase in the number of objectives. There have been some recent developments on methods that can specifically handle many-objective optimization problems. As an example, the Kriging-assisted reference vector evolutionary algorithm (K-RVEA) can solve many-objective problems with computationally expensive simulations [21]. This method has been demonstrated to optimize a blast furnace with eight objectives [20].

3.10 Utilizing derivatives information

Some numerical codes provide direct information of derivatives that can be used to enhance the prediction accuracy of surrogate models. The first-order derivative (gradient) information is frequently used while higher-order derivatives are rarely used. Some approaches use the second-order derivative (Hessian) [89, 110]. Some CFD codes are equipped with a routine to evaluate the gradient for an arbitrary number of variables, primarily via the adjoint method [45]. Typically, the evaluation of gradient via adjoint takes only one additional CFD simulation for an arbitrary number of variables, which explains its popularity in the aerodynamic optimization community. However, the presence of noise in the gradient information can significantly affect the performance of a gradient-based optimization. Another alternative is to utilize gradient-enhanced surrogate models that can filter out the noise in both function response and gradient. The effectiveness of gradient-enhanced surrogate models have been demonstrated in the optimization of turbomachinery components [5].

There exist various gradient-enhanced extensions of surrogate models. Arguably, one of the most popular gradient enhanced surrogate models is the gradient-enhanced Kriging (GEK) [63]. However, the gradient-enhanced version of PCE [79], and RBF [74], can also be found in literature. Gradient-enhanced surrogate models are highly useful for high-dimensional modeling to leverage the curse-of-dimensionality. However, the gradient information itself should be effectively computed, which can be done by employing the adjoint or automatic differentiation method [82] (the latter typically costs about 4-5 times of the original function evaluation for an arbitrary number of variables).

When gradient information is utilized, it is worth noting that noise could also exist in the gradient values. Similar to that of function response, the accuracy of gradient-enhanced surrogate models also depends on the presence of gradient noise and how this noise is taken into account during the construction of surrogates. To the best of our knowledge, there are only few methods explicitly designed to handle gradient noise. In this respect, GEK can be further extended to handle the noise in gradient by introducing a second regression factor that alters the diagonal matrix [24]. De Baar et al. shows that the extended GEK is capable of reducing approximation error as compared to the standard GEK that only uses one regression factor applied to the response.

3.11 Benchmarking of optimization algorithms

Benchmarking is important to examine the strengths and weaknesses of various optimization algorithms. It can help users decide the most suitable algorithm for the problem at hand. Since not all characteristics of real-world problems can be fully replicated by artificial problems, benchmarking with real-world problems is important to give us insight into the performance of optimization algorithms when handling real-world complexities. Some studies that proposed new optimization algorithms also performed such a study after benchmarking with artificial problems, e.g., the comparison of K-RVEA, ParEGO, MOEA/D, and RVEA on free-radical polymerization [21]. Although there are some existing studies on the comparison of surrogate models for various real-world problems, e.g., ground water remediation process [65], wing design [75], and turbomachinery design [80], these studies compared the approximation capability of surrogate models and not the surrogate-based optimization strategies. Some studies that focus on comparison of various surrogate-based optimization algorithms or strategies in real-world problems can be found in Han et al., [62], Zuhal et al. [115].

Unfortunately, despite its importance, studies to compare various optimization algorithms on real-world problems are still limited, mainly because such problems are typically not publicly available. It is therefore imperative to establish a library of benchmarking problems based on real-world problems that are accessible to researchers. The comparison does not really need to be performed with computationally expensive simulations. Instead, benchmarking can be done with low-fidelity computer simulations (which are non-algebraic problems) so that multiple independent runs can be executed. It is worth noting that the use of high-fidelity simulations in optimization problem does not necessarily translates to complex optimization problems. One interesting publicly available test suite is the set of CFD-based problems developed by Daniels et al. [23]. These test problems are suitable for comparing various surrogate-based optimization algorithms since they are based on solving PDEs (e.g. fluid flow equation).

4 THE FUTURE OF SURROGATE-BASED METHODS

Considering their usefulness and wide use in many engineering disciplines, we project that surrogate models will continue to be a key element in advancing engineering design optimization and exploration. We expect to see further growth and development in terms of the form and types and surrogate models. In the past, linear and polynomial models were the most popular forms. However, at present, there is a great interest in surrogate models such as RBF, SVR, and Kriging [107] that are more suitable for non-linear functions. We also need to note that polynomial regressions in the form of PCE have found its popularity in the UQ and SA communities. Despite its origin from UQ and SA community, PCE (in particular its non-intrusive version) is a surrogate model that can be virtually used in many disciplines and applications.

As noted by Viana et al. [107], the complexity of the problems that the engineering community has to tackle grows faster than that of surrogate modeling technology. We expect that SBO will still hold an important role in the far future. As the computing power grows, we become more capable of tackling problems that were previously too expensive to perform. For example, CFD with Euler solver was considered as a high-fidelity technique back in the 1970s. Today, however, solving the Euler equation of a three-dimensional model can be executed with a personal computer within one day. This basically implies that our definitions of high-fidelity and low-fidelity simulations change over time. Regardless of the changes in the scope and definition of what constitutes as a high-fidelity model, we are confident that surrogate models will continue to lend themselves to help make running computationally intensive tasks that involve high-fidelity models intractable.

The curse-of-dimensionality in surrogate modeling remains an open issue. While gradient information has been shown to be helpful in leveraging this problem, it is not always readily available in all computational cases. Researchers still work towards tackling the curse-of-dimensionality without relying too much on the gradient information, and works are still underway. There have also been some attempts to couple surrogate models with methods that could reduce dimensionality such as the partial least squares [13] and the principal component analysis [56]. However, high-dimensionality problems are typically only tackled in the later stage of design and optimization. In design exploration and preliminary optimization, engineers typically focus on a manageable number of design variables for an easier interpretation of various designs. Therefore, obtaining higher accuracy surrogate models and expensive optimization in low to moderate dimensions are still topics that are highly relevant in the near and far future.

5 CONCLUSIONS

In this paper, we review and discuss some key issues in engineering design optimization and exploration using surrogate models. The main objective of this paper is to introduce the common issues and challenges that are typically faced by engineering optimization community to general readers. The classification of SBO based on the role of the surrogate model in the optimization process was also discussed. Discussions on particular situations and conditions where the use of surrogate models was more beneficial than other optimization techniques were also given. We also discussed some particular aspects that were not so widely addressed outside the engineering optimization community, such as the failed simulations, gradient-enhanced surrogate models, and multi-fidelity simulations. From the discussion, we argue that surrogate models will continue to play an important role in design and optimization. We are progressively tackling more complex problems (e.g. RBDO) by still relying on the surrogate model as the key technology.

We also argue that interaction between communities who are working in algorithmic development (e.g., evolutionary computation and machine learning) and those in engineering and applications should be encouraged. This is necessary to accelerate the research pace and also for the development of better problem-solving procedures in the context of design optimization and exploration.

ACKNOWLEDGEMENT

Part of this work was funded through Riset KK ITB program, which we gratefully acknowledge. Part of the work was also carried out under the Collaborative Research Project of the Institute of Fluid Science, Tohoku University.

REFERENCES

[1] Niels Aage, Erik Andreassen, Boyan S Lazarov, and Ole Sigmund. 2017. Giga-voxel computational morphogenesis for structural design. *Nature* 550, 7674 (2017), 84.
[2] Taimoor Akhtar and Christine A Shoemaker. 2016. Multi objective optimization of computationally expensive multi-modal functions with RBF surrogates and multi-rule selection. *Journal of Global Optimization* 64, 1 (2016), 17–32.
[3] Natalia M Alexandrov, JE Dennis, Robert Michael Lewis, and Virginia Torczon. 1998. A trust-region framework for managing the use of approximation models in optimization. *Structural optimization* 15, 1 (1998), 16–23.
[4] Natalia M Alexandrov and Robert Michael Lewis. 2001. An overview of first-order model management for engineering optimization. *Optimization and Engineering* 2, 4 (2001), 413–430.
[5] Jan Backhaus, Andreas Schmitz, Christian Frey, Max Sagebaum, Sebastian Mann, Marc Nagel, and Nicolas R Gauger. 2017. Application of an algorithmically differentiated turbomachinery flow solver to the optimization of a fan stage. In *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. 3997.
[6] Samineh Bagheri, Wolfgang Konen, Richard Allmendinger, Jürgen Branke, Kalyanmoy Deb, Jonathan Fieldsend, Domenico Quagliarella, and Karthik Sindhya. 2017. Constraint handling in efficient global optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 673–680.
[7] Anirban Basudhar, Christoph Dribusch, Sylvain Lacaze, and Samy Missoum. 2012. Constrained efficient global optimization with support vector machines. *Structural and Multidisciplinary Optimization* 46, 2 (2012), 201–221.

[8] Dimitri Bettebghor, Nathalie Bartoli, Stéphane Grihon, Joseph Morlier, and Manuel Samuelides. 2011. Surrogate modeling approximation using a mixture of experts based on EM joint estimation. *Structural and multidisciplinary optimization* 43, 2 (2011), 243–259.

[9] Hans-Georg Beyer and Bernhard Sendhoff. 2007. Robust optimization–a comprehensive survey. *Computer methods in applied mechanics and engineering* 196, 33-34 (2007), 3190–3218.

[10] Barron J Bichon, Michael S Eldred, Sankaran Mahadevan, and John M McFarland. 2013. Efficient global surrogate modeling for reliability-based design optimization. *Journal of Mechanical Design* 135, 1 (2013), 011009.

[11] Barron J Bichon, Michael S Eldred, Laura Painton Swiler, Sandaran Mahadevan, and John M McFarland. 2008. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal* 46, 10 (2008), 2459–2468.

[12] Géraud Blatman and Bruno Sudret. 2011. Adaptive sparse polynomial chaos expansion based on least angle regression. *J. Comput. Phys.* 230, 6 (2011), 2345–2367.

[13] Mohamed Amine Bouhlel, Nathalie Bartoli, Abdelkader Otsmane, and Joseph Morlier. 2016. Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction. *Structural and Multidisciplinary Optimization* 53, 5 (2016), 935–952.

[14] Roberto Calandra, André Seyfarth, Jan Peters, and Marc Peter Deisenroth. 2016. Bayesian optimization for learning gaits under uncertainty. *Annals of Mathematics and Artificial Intelligence* 76, 1-2 (2016), 5–23.

[15] Tanmoy Chatterjee, Rajib Chowdhury, and Palaniappan Ramu. 2019. Decoupling uncertainty quantification from robust design optimization. *Structural and Multidisciplinary Optimization* (2019), 1–22.

[16] Anirban Chaudhuri, Raphael T Haftka, Kelvin T Chang, Jordan Van Hall, and Peter Ifju. 2015. Multi-objective experimental optimization with multiple simultaneous sampling for flapping wings. In *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference.* 1586.

[17] Anirban Chaudhuri, Raphael T Haftka, Peter Ifju, Kelvin Chang, Christopher Tyler, and Tony Schmitz. 2015. Experimental flapping wing optimization and uncertainty quantification using limited samples. *Structural and Multidisciplinary Optimization* 51, 4 (2015), 957–970.

[18] Mingjie Chen, Andrew FB Tompson, Robert J Mellors, and Osman Abdalla. 2015. An efficient optimization of well placement and control for a geothermal prospect under geological uncertainty. *Applied energy* 137 (2015), 352–363.

[19] Seung-Kyum Choi, Ramana V Grandhi, Robert A Canfield, and Chris L Pettit. 2004. Polynomial chaos expansion with latin hypercube sampling for estimating response variability. *AIAA journal* 42, 6 (2004), 1191–1198.

[20] Tinkle Chugh, Nirupam Chakraborti, Karthik Sindhya, and Yaochu Jin. 2017. A data-driven surrogate-assisted evolutionary algorithm applied to a many-objective blast furnace optimization problem. *Materials and Manufacturing Processes* 32, 10 (2017), 1172–1178.

[21] Tinkle Chugh, Yaochu Jin, Kaisa Miettinen, Jussi Hakanen, and Karthik Sindhya. 2018. A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization. *IEEE Transactions on Evolutionary Computation* 22, 1 (2018), 129–142.

[22] Paul G Constantine, Eric Dow, and Qiqi Wang. 2014. Active subspace methods in theory and practice: applications to kriging surfaces. *SIAM Journal on Scientific Computing* 36, 4 (2014), A1500–A1524.

[23] Steven J Daniels, Alma AM Rahat, Richard M Everson, Gavin R Tabor, and Jonathan E Fieldsend. 2018. A suite of computationally expensive shape optimisation problems using computational fluid dynamics. In *International Conference on Parallel Problem Solving from Nature.* Springer, 296–307.

[24] JHS de Baar, TP Scholcz, and RP Dwight. 2015. Exploiting adjoint derivatives in high-dimensional metamodels. *AIAA Journal* 53, 5 (2015), 1391–1395.

[25] Harris Drucker, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. 1997. Support vector regression machines. In *Advances in neural information processing systems.* 155–161.

[26] B Echard, N Gayton, and M Lemaire. 2011. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Structural Safety* 33, 2 (2011), 145–154.

[27] M Giselle Fernández-Godino, Chanyoung Park, Nam-Ho Kim, and Raphael T Haftka. 2016. Review of multi-fidelity models. *arXiv preprint arXiv:1609.07196* (2016).

[28] Peter J Fleming, Robin C Purshouse, and Robert J Lygoe. 2005. Many-objective optimization: An engineering design perspective. In *International conference on evolutionary multi-criterion optimization.* Springer, 14–32.

[29] Alexander IJ Forrester, András Sóbester, and Andy J Keane. 2006. Optimization with missing data. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 462, 2067 (2006), 935–945.

[30] Adam Gaier, Alexander Asteroth, and Jean-Baptiste Mouret. 2017. Data-efficient exploration, optimization, and modeling of diverse designs through surrogate-assisted illumination. In *Proceedings of the Genetic and Evolutionary Computation Conference.* ACM, 99–106.

[31] Javier González, Zhenwen Dai, Philipp Hennig, and Neil Lawrence. 2016. Batch bayesian optimization via local penalization. In *Artificial Intelligence and Statistics.* 648–657.

[32] H-M Gutmann. 2001. A radial basis function method for global optimization. *Journal of global optimization* 19, 3 (2001), 201–227.

[33] Raphael T Haftka, Diane Villanueva, and Anirban Chaudhuri. 2016. Parallel surrogate-assisted global optimization with expensive functions–a survey. *Structural and Multidisciplinary Optimization* 54, 1 (2016), 3–13.

[34] Zhong-Hua Han, Jing Chen, Ke-Shi Zhang, Zhen-Ming Xu, Zhen Zhu, and Wen-Ping Song. 2018. Aerodynamic shape optimization of natural-laminar-flow wing using surrogate-based approach. *AIAA Journal* 56, 7 (2018), 2579–2593.

[35] Zhong-Hua Han and Stefan Görtz. 2012. Hierarchical kriging model for variable-fidelity surrogate modeling. *AIAA journal* 50, 9 (2012), 1885–1896.

[36] Kenneth Holmström. 2008. An adaptive radial basis algorithm (ARBF) for expensive black-box global optimization. *Journal of Global Optimization* 41, 3 (2008), 447–464.

[37] Daniel Horn, Melanie Dagge, Xudong Sun, and Bernd Bischl. 2017. First Investigations on Noisy Model-Based Multi-Objective Optimization. In *International Conference on Evolutionary Multi-Criterion Optimization.* Springer, 298–313.

[38] Nicholas J Horton and Ken P Kleinman. 2007. Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician* 61, 1 (2007), 79–90.

[39] Chao Hu and Byeng D Youn. 2011. Adaptive-sparse polynomial chaos expansion for reliability analysis and design of complex engineering systems. *Structural and Multidisciplinary Optimization* 43, 3 (2011), 419–442.

[40] Jorge E Hurtado and Diego A Alvarez. 2001. Neural-network-based reliability analysis: a comparative study. *Computer methods in applied mechanics and engineering* 191, 1-2 (2001), 113–132.

[41] Afzal Husain and Kwang-Yong Kim. 2010. Enhanced multi-objective optimization of a microchannel heat sink through evolutionary algorithm coupled with multiple surrogate models. *Applied Thermal Engineering* 30, 13 (2010), 1683–1691.

[42] Hisao Ishibuchi, Noritaka Tsukamoto, and Yusuke Nojima. 2008. Evolutionary many-objective optimization: A short review. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence).* IEEE, 2419–2426.

[43] Alexander I J. Forrester, Andy J Keane, and Neil W Bressloff. 2006. Design and analysis of" Noisy" computer experiments. *AIAA journal* 44, 10 (2006), 2331–2339.

[44] Jan Pieter Jacobs, S Koziel, and S Ogurtsov. 2013. Computationally efficient multi-fidelity Bayesian support vector regression modeling of planar antenna input characteristics. *IEEE Transactions on Antennas and Propagation* 61, 2 (2013), 980–984.

[45] Antony Jameson. 1988. Aerodynamic design via control theory. *Journal of scientific computing* 3, 3 (1988), 233–260.

[46] Shinkyu Jeong, Kazuhisa Chiba, and Shigeru Obayashi. 2005. Data mining for aerodynamic design space. *Journal of aerospace computing, information, and communication* 2, 11 (2005), 452–469.

[47] Shinkyu Jeong, Mitsuhiro Murayama, and Kazuomi Yamamoto. 2005. Efficient optimization design method using kriging model. *Journal of aircraft* 42, 2 (2005), 413–420.

[48] Yaochu Jin. 2011. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation* 1, 2 (2011), 61–70.

[49] Donald R Jones, Matthias Schonlau, and William J Welch. 1998. Efficient global optimization of expensive black-box functions. *Journal of Global optimization* 13, 4 (1998), 455–492.

[50] Stefan Kern, Nikolaus Hansen, and Petros Koumoutsakos. 2006. Local metamodels for optimization using evolution strategies. In *Parallel Problem Solving from Nature-PPSN IX.* Springer, 939–948.

[51] Pierric Kersaudy, Bruno Sudret, Nadège Varsier, Odile Picon, and Joe Wiart. 2015. A new surrogate modeling technique combining Kriging and polynomial chaos expansions–Application to uncertainty analysis in computational dosimetry. *J. Comput. Phys.* 286 (2015), 103–117.

[52] Joshua Knowles, David Corne, and Alan Reynolds. 2009. Noisy multiobjective optimization on a budget of 250 evaluations. In *International Conference on Evolutionary Multi-Criterion Optimization.* Springer, 36–50.

[53] Tipaluck Krityakierne, Taimoor Akhtar, and Christine A Shoemaker. 2016. SOP: parallel surrogate global optimization with Pareto center selection for computationally expensive single objective problems. *Journal of Global Optimization* 66, 3 (2016), 417–437.

[54] Rémi Lam, Matthias Poloczek, Peter Frazier, and Karen E Willcox. 2018. Advances in bayesian optimization with applications in aerospace engineering. In *2018 AIAA Non-Deterministic Approaches Conference.* 1656.

[55] Rémi Lam, Matthias Poloczek, Peter Frazier, and Karen E. Willcox. 2018. Advances in Bayesian Optimization with Applications in Aerospace Engineering. In *2018 AIAA Non-Deterministic Approaches Conference.*

[56] Christos Lataniotis, Stefano Marelli, and Bruno Sudret. 2018. Extending classical surrogate modelling to ultrahigh dimensional problems through supervised dimensionality reduction: a data-driven approach. *arXiv preprint arXiv:1812.06309* (2018).
[57] Minh Nghia Le, Yew Soon Ong, Stefan Menzel, Yaochu Jin, and Bernhard Sendhoff. 2013. Evolution by adapting surrogates. *Evolutionary computation* 21, 2 (2013), 313–340.
[58] Rhea P. Liem, Gaetan K. W. Kenway, and Joaquim R. R. A. Martins. 2015. Multimission Aircraft Fuel Burn Minimization via Multipoint Aerostructural Optimization. *AIAA Journal* 53, 1 (2015), 104–122. https://doi.org/10.2514/1.J052940
[59] Rhea P. Liem, G. K. W. Kenway, and Joaquim R. R. A. Martins. 2017. Expected Drag Minimization for Aerodynamic Design Optimization Based on Aircraft Operational Data. *Aerospace Science and Technology* 63 (2017), 344–362.
[60] Rhea P Liem, Charles A Mader, and Joaquim RRA Martins. 2015. Surrogate models and mixtures of experts in aerodynamic performance prediction for aircraft mission analysis. *Aerospace Science and Technology* 43 (2015), 126–151.
[61] Bo Liu, Slawomir Koziel, and Qingfu Zhang. 2016. A multi-fidelity surrogate-model-assisted evolutionary algorithm for computationally expensive optimization problems. *Journal of computational science* 12 (2016), 28–37.
[62] Jun Liu, ZH Han, and WP Song. 2012. Comparison of infill sampling criteria in kriging-based aerodynamic optimization. In *28th Congress of the International Council of the Aeronautical Sciences*. 23–28.
[63] Weiyu Liu and Stephen Batill. 2002. Gradient-enhanced response surface approximations using kriging models. In *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*. 5456.
[64] Cheng Lu, Yun-Wen Feng, Rhea P. Liem, and Cheng-Wei Fei. 2018. Improved Kriging with extremum response surface method for structural dynamic reliability and sensitivity analyses. *Aerospace Science and Technology* 76 (2018), 164–175. https://doi.org/10.1016/j.ast.2018.02.012
[65] Jiannan Luo and Wenxi Lu. 2014. Comparison of surrogate models with different methods in groundwater remediation process. *Journal of Earth System Science* 123, 7 (2014), 1579–1589.
[66] Stefano Marelli and Bruno Sudret. 2018. An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis. *Structural Safety* 75 (2018), 67–74.
[67] Ruben Martinez-Cantin, Kevin Tee, and Michael McCourt. 2017. Practical Bayesian optimization in the presence of outliers. *arXiv preprint arXiv:1712.04567* (2017).
[68] Efrén Mezura-Montes and Carlos A Coello Coello. 2011. Constraint-handling in nature-inspired numerical optimization: past, present and future. *Swarm and Evolutionary Computation* 1, 4 (2011), 173–194.
[69] Miha Mlakar, Dejan Petelin, Tea Tušar, and Bogdan Filipič. 2015. GP-DEMO: differential evolution for multiobjective optimization based on Gaussian process models. *European Journal of Operational Research* 243, 2 (2015), 347–361.
[70] Maliki Moustapha, Bruno Sudret, Jean-Marc Bourinet, and Benoît Guillaume. 2016. Quantile-based optimization under uncertainties using adaptive Kriging surrogate models. *Structural and multidisciplinary optimization* 54, 6 (2016), 1403–1421.
[71] Tanmoy Mukhopadhyay, S Chakraborty, S Dey, S Adhikari, and R Chowdhury. 2017. A critical assessment of Kriging model variants for high-fidelity uncertainty quantification in dynamics of composite shells. *Archives of Computational Methods in Engineering* 24, 3 (2017), 495–518.
[72] Mahdi Arian Nik, Kazem Fayazbakhsh, Damiano Pasini, and Larry Lessard. 2012. Surrogate-based multi-objective optimization of a composite laminate with curvilinear fibers. *Composite Structures* 94, 8 (2012), 2306–2313.
[73] Shigeru Obayashi, Shin-Kyu Jeong, Koji Shimoyama, Kazuhisa Chiba, and Hiroyuki Morino. 2010. Multi-objective design exploration and its applications. *International Journal of Aeronautical and Space Sciences* 11, 4 (2010), 247–265.
[74] Yew-Soon Ong, Kai Yew Lum, and Prasanth B Nair. 2008. Hybrid evolutionary algorithm with Hermite radial basis function interpolants for computationally expensive adjoint solvers. *Computational Optimization and Applications* 39, 1 (2008), 97–119.
[75] Ricardo M Paiva, André R D. Carvalho, Curran Crawford, and Afzal Suleman. 2010. Comparison of surrogate models in a multidisciplinary optimization framework for wing design. *AIAA journal* 48, 5 (2010), 995–1006.
[76] Pramudita Satria Palar, Takeshi Tsuchiya, and Geoffrey Thomas Parks. 2016. Multi-fidelity non-intrusive polynomial chaos based on regression. *Computer Methods in Applied Mechanics and Engineering* 305 (2016), 579–606.
[77] Pramudita Satria Palar, Kaifeng Yang, Koji Shimoyama, Michael Emmerich, and Thomas Bäck. 2018. Multi-objective aerodynamic design with user preference using truncated expected hypervolume improvement. In *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 1333–1340.
[78] Seongeon Park, Jonggeol Na, Minjun Kim, and Jong Min Lee. 2018. Multi-objective Bayesian optimization of chemical reactor design using computational fluid dynamics. *Computers & Chemical Engineering* 119 (2018), 25–37.
[79] Ji Peng, Jerrad Hampton, and Alireza Doostan. 2016. On polynomial chaos expansion via gradient-enhanced âĎŞ1-minimization. *J. Comput. Phys.* 310 (2016), 440–458.
[80] J Peter and M Marcelet. 2008. Comparison of surrogate models for turbomachinery design. *Wseas transactions on fluid mechanics* 3, 1 (2008), 10–17.
[81] Victor Picheny, Tobias Wagner, and David Ginsbourger. 2013. A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization* 48, 3 (2013), 607–626.
[82] Louis B Rall. 1981. Automatic differentiation: Techniques and applications. (1981).
[83] Rommel G Regis. 2011. Stochastic radial basis function algorithms for large-scale optimization involving expensive black-box objective and constraint functions. *Computers & Operations Research* 38, 5 (2011), 837–853.
[84] Rommel G Regis. 2014. Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points. *Engineering Optimization* 46, 2 (2014), 218–243.
[85] Rommel G Regis. 2016. Multi-objective constrained black-box optimization using radial basis function surrogates. *Journal of computational science* 16 (2016), 140–155.
[86] Rommel G Regis and Christine A Shoemaker. 2005. Constrained global optimization of expensive black box functions using radial basis functions. *Journal of Global optimization* 31, 1 (2005), 153–171.
[87] Rommel G Regis and Christine A Shoemaker. 2007. A stochastic radial basis function method for the global optimization of expensive functions. *INFORMS Journal on Computing* 19, 4 (2007), 497–509.
[88] Jose F Rodriguez, John E Renaud, and Layne T Watson. 1998. Trust region augmented Lagrangian methods for sequential response surface approximation and optimization. *Journal of mechanical design* 120, 1 (1998), 58–66.
[89] Markus P Rumpfkeil. 2012. Optimizations under uncertainty using gradients, Hessians, and surrogate models. *AIAA journal* 51, 2 (2012), 444–451.
[90] Jerome Sacks, William J Welch, Toby J Mitchell, and Henry P Wynn. 1989. Design and analysis of computer experiments. *Statistical science* (1989), 409–423.
[91] Abdus Samad, Kwang-Yong Kim, Tushar Goel, Raphael T Haftka, and Wei Shyy. 2008. Multiple surrogate modeling for axial compressor blade shape optimization. *Journal of Propulsion and Power* 24, 2 (2008), 301–310.
[92] Roland Schöbi, Bruno Sudret, and Stefano Marelli. 2016. Rare event estimation using polynomial-chaos kriging. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* 3, 2 (2016), D4016002.
[93] Matthias Schonlau, William J Welch, and Donald R Jones. 1998. Global versus local search in constrained optimization of computer models. *Lecture Notes-Monograph Series* (1998), 11–25.
[94] Warren Scott, Peter Frazier, and Warren Powell. 2011. The correlated knowledge gradient for simulation optimization of continuous parameters using gaussian process regression. *SIAM Journal on Optimization* 21, 3 (2011), 996–1026.
[95] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. 2016. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* 104, 1 (January 2016), 148–175. https://doi.org/10.1109/JPROC.2015.2494218
[96] Koji Shimoyama, Soshi Kawai, and Juan J Alonso. 2013. Dynamic adaptive sampling based on Kriging surrogate models for efficient uncertainty quantification. In *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. 1470.
[97] Prashant Singh, Ivo Couckuyt, Khairy Elsayed, Dirk Deschrijver, and Tom Dhaene. 2016. Shape optimization of a cyclone separator using multi-objective surrogate-based optimization. *Applied Mathematical Modelling* 40, 5-6 (2016), 4248–4259.
[98] Ilya M Sobol. 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation* 55, 1-3 (2001), 271–280.
[99] Jian Song, Yun Yang, Jianfeng Wu, Jichun Wu, Xiaomin Sun, and Jin Lin. 2018. Adaptive surrogate model based multiobjective optimization for coastal aquifer management. *Journal of hydrology* 561 (2018), 98–111.
[100] Lu-Kai Song, Guang-Chen Bai, Cheng-Wei Fei, and Rhea P. Liem. 2018. Transient probabilistic design of a flexible multibody system using a dynamic fuzzy neural network method with distributed collaborative strategy. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*. (2018). https://doi.org/10.1177/0954410018813213
[101] Bruno Sudret. 2008. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety* 93, 7 (2008), 964–979.
[102] Chaoli Sun, Yaochu Jin, Jianchao Zeng, and Yang Yu. 2015. A two-layer surrogate-assisted particle swarm optimization algorithm. *Soft computing* 19, 6 (2015), 1461–1475.
[103] David JJ Toal. 2015. Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models. *Structural and Multidisciplinary Optimization* 51, 6 (2015), 1223–1245.
[104] Rohit Tripathy, Ilias Bilionis, and Marcial Gonzalez. 2016. Gaussian processes with built-in dimensionality reduction: Applications to high-dimensional uncertainty propagation. *J. Comput. Phys.* 321 (2016), 191–223.
[105] Stef Van Buuren. 2018. *Flexible imputation of missing data*. Chapman and Hall/CRC.

[106] Felipe AC Viana, Raphael T Haftka, and Layne T Watson. 2013. Efficient global optimization algorithm assisted by multiple surrogate techniques. *Journal of Global Optimization* 56, 2 (2013), 669–689.

[107] Felipe AC Viana, Timothy W Simpson, Vladimir Balabanov, and Vasilli Toropov. 2014. Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come? *AIAA journal* 52, 4 (2014), 670–690.

[108] Hao Wang, Bas van Stein, Michael Emmerich, and Thomas Bäck. 2017. Time complexity reduction in efficient global optimization using cluster kriging. In *Proceedings of the Genetic and Evolutionary Computation Conference*. ACM, 889–896.

[109] Dongbin Xiu and George Em Karniadakis. 2002. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing* 24, 2 (2002), 619–644.

[110] Wataru Yamazaki, Markus Rumpfkeil, and Dimitri Mavriplis. 2010. Design optimization utilizing Gradient/Hessian enhanced surrogate model. In *28th AIAA Applied Aerodynamics Conference*. 4363.

[111] Jefry Yanto and Rhea P. Liem. 2018. Aircraft fuel burn performance study: a data-enhanced modeling approach. *Transportation Research Part D: Transport and Environment* 65 (2018), 574–595. https://doi.org/10.1016/j.trd.2018.09.014

[112] Dawei Zhan, Jiachang Qian, and Yuansheng Cheng. 2017. Balancing global and local search in parallel efficient global optimization algorithms. *Journal of Global Optimization* 67, 4 (2017), 873–892.

[113] Ke-Shi Zhang, Zhong-Hua Han, Zhong-Jian Gao, and Yuan Wang. 2019. Constraint aggregation for large number of constraints in wing surrogate-based optimization. *Structural and Multidisciplinary Optimization* 59, 2 (2019), 421–438.

[114] Qingfu Zhang, Wudong Liu, Edward Tsang, and Botond Virginas. 2010. Expensive multiobjective optimization by MOEA/D with Gaussian process model. *IEEE Transactions on Evolutionary Computation* 14, 3 (2010), 456–474.

[115] Lavi R Zuhal, Cahya Amalinadhi, Yohanes B Dwianto, Pramudita S Palar, and Koji Shimoyama. 2018. Benchmarking Multi-Objective Bayesian Global Optimization Strategies for Aerodynamic Design. In *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. 0914.