Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

An Empirical Comparison of Meta-Modeling Techniques for Robust Design Optimization

Sibghat Ullah, Hao Wang, Stefan Menzel, Bernhard Sendhoff, Thomas Bäck

2019

Preprint:

This is an accepted article published in IEEE Symposium Series on Computational Intelligence. The final authenticated version is available online at: https://doi.org/10.1109/SSCI44817.2019.9002805 Copyright 2019 IEEE

An Empirical Comparison of Meta-Modeling Techniques for Robust Design Optimization

Sibghat Ullah[1], Hao Wang[1], Stefan Menzel[2], Bernhard Sendhoff[2], and Thomas Bäck[1]

[1]Leiden Institute of Advanced Computer Science (LIACS), Leiden University, The Netherlands
Email: {s.ullah,h.wang,t.h.w.baeck}@liacs.leidenuniv.nl

[2]Honda Research Institute Europe GmbH (HRI-EU), Offenbach/Main, Germany
Email: {stefan.menzel,bernhard.sendhoff}@honda-ri.de

The article has been accepted for publication in the IEEE Symposium Series on Computational Intelligence.

Copyright notice:

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

An Empirical Comparison of Meta-Modeling Techniques for Robust Design Optimization

Sibghat Ullah[1], Hao Wang[1], Stefan Menzel[2], Bernhard Sendhoff[2], and Thomas Bäck[1]

[1]Leiden Institute of Advanced Computer Science (LIACS), Leiden University, The Netherlands
Email: {s.ullah,h.wang,t.h.w.baeck}@liacs.leidenuniv.nl

[2]Honda Research Institute Europe GmbH (HRI-EU), Offenbach/Main, Germany
Email: {stefan.menzel,bernhard.sendhoff}@honda-ri.de

***Abstract*—This research investigates the potential of using meta-modeling techniques in the context of robust optimization namely optimization under uncertainty/noise. A systematic empirical comparison is performed for evaluating and comparing different meta-modeling techniques for robust optimization. The experimental setup includes three noise levels, six meta-modeling algorithms, and six benchmark problems from the continuous optimization domain, each for three different dimensionalities. Two robustness definitions: robust regularization and robust composition, are used in the experiments. The meta-modeling techniques are evaluated and compared with respect to the modeling accuracy and the optimal function values. The results clearly show that Kriging, Support Vector Machine and Polynomial regression perform excellently as they achieve high accuracy and the optimal point on the model landscape is close to the true optimum of test functions in most cases.**

***Keywords*—meta-modeling, surrogate-assisted optimization, robust optimization, quality engineering, machine learning**

I. Introduction

Uncertainty is an imperative motif in design optimization. The classical view on black-box optimization does not account for these uncertainties. It is important to state that throughout this paper uncertainty and noise refer to the same concept, (i.e., unexpected drifts in the optimization setup). These unexpected drifts can be found in the design and environmental parameters, e.g., temperature, stiffness, structural rigidity etc. as well as in the constraints and objectives. Accounting for these uncertainties leads to the concept of robust design optimization aka quality engineering [1]. Recently, many methodologies have been investigated for robust optimization. Most of these approaches for robust optimization focus on using direct search methods, in particular evolution strategies and surrogate-assisted (aka meta-model based) optimization. We believe that while the former has been discussed at length in the context of robust optimization [2]–[7], the latter still needs further consideration. In particular, the practical areas of interest regarding the computational tractability, fidelity and flexibility of meta-model based optimization need to be investigated [8]–[10].

Uncertainties and noise comprise one of the most challenging areas in optimization literature. They are encountered frequently in real-world optimization problems. These notions are significant in engineering optimization such as automobile manufacturing, building construction and steel production due to the potentially serious impact in case of a failure. As such, achieving robustness in design optimization is important. Despite the significance, achieving robustness in design optimization is challenging. We believe this is due to the veracity of optimization landscapes, high dimensionality and the intrinsic complexity of product life cycle. On the other hand, computational statistics based supervised learning methods have been successfully employed in Artificial Intelligence applications in health-care, transportation, banking and finance, signal processing, multimedia and entertainment. As such, it is desirable to investigate the suitability of these methods to help designing robust systems.

For many reasons, incorporating robustness in design optimization is challenging. Firstly, uncertainty has many shapes and forms [7], [11]. The nature and structure of uncertainty is often unknown in advance. Secondly, uncertainty has many sources [7], [11] such as environmental and design parameters, objectives and constraints. Thus, accounting for all sources of uncertainty in an accurate fashion is often impossible. Thirdly, uncertainty modeling and representation is itself quite tedious and exhaustive. Finally, there is lack of cohesive literature to minimize the effects of all types of uncertainty and design a robust system. As such, it is desirable to obtain practically useful insights to be adopted by engineers to account for the unexpected changes during optimization.

Early approaches for robust design optimization are based on Taguchi methods [1], [12] to assess the quality of an engineering design. In particular, Taguchi postulated a three-stage robust design optimization which includes system design, parameter design and tolerance design. Taguchi used signal to noise ratio (SNR) and mean squared deviation (MSD) to formulate the objective of quality engineering and while this work is an important milestone in itself, it comes with an array of quibbles targeting the efficiency of the mechanism and the philosophical understanding of SNR and MSD [13]–[15]. Later theories for robust optimization hinge upon the so called robust-counterpart approach [16] which incorporates the robustness in the formulation of the optimization problem. Robust-counterpart approaches have been employed regularly in the literature with combinations of mathematical programming techniques and direct search methods. The limitations of

mathematical programming techniques are discernible [11] as they are not naturally associated with simulation which forms an important part of the optimization pipeline in engineering [17]. Also, this family of techniques are hardly flexible and analytically tractable in most real-world design optimization tasks. The combination of the robust-counterpart approach with direct search methods is usually focusing on evolutionary algorithms only. As such, we believe there is a dire need for combining the robust-counterpart approach with intelligent, experience based optimization schemes such as meta-model based optimization which are easy to model, appraise and practice.

In this paper, we study sensitivity robustness [18], [19], i.e., robustness related to the sensitivity w.r.t. noise on the design parameters. We use six optimization problems. We then construe three levels of additive noise characterizing small, medium and big uncertainty during optimization. Furthermore, to incorporate robustness for the corresponding noisy functions, we choose two robust-counterpart approaches, namely the robust regularization based on the worst case scenario and the robust composition, i.e., robustness achieved through combining the expectation and dispersion of a noisy function. As a result, we have thirty-six noisy optimization cases due to three levels of noise and two robust-counterpart approaches. We study six meta-modeling techniques: Kriging, Support Vector Machines (SVM), Radial Basis Function Network (RBFN), Random Forest (RF), K Nearest Neighbors (KNN) and Elastic net with second order polynomial regression function (ELN). We perform hyper-parameter optimization of the corresponding hyper-parameters for all meta-modeling algorithms, to obtain the best results for each scenario. Conclusively, we evaluate and compare these meta-modeling techniques on these test cases and discuss their generality, reliability, and computational tractability.

The rest of the paper is organized as follows. We present the idea of robustness in design optimization in section II. This includes robust counter-part approaches namely the robust regularization and robust composition discussed in section II-A and section II-B respectively. Section III provides a brief introduction to meta-modeling. In section IV, we present the experimental design for this work. This is followed by results in section V. Finally, we discuss the logical conclusion of the paper along-side the future research line in section VI.

II. Robustness in black-box optimization

In this work, we consider real-valued black-box optimization problems $f\colon \mathcal{S} \rightarrow \mathbb{R}$, with $\mathcal{S} \subseteq \mathbb{R}^d$, where the so-called feasible region $\mathcal{S}$ is specified by inequality constraints $g_j(\mathbf{x}) \leq 0$ ($j \in \{1, \ldots, J\}$) and equality ones $h_k(\mathbf{x}) = 0$ ($k \in \{1, \ldots, K\}$), $\mathbf{x} \in \mathcal{S}$. Without loss of generality, the objective function f is subject to minimization. In most real-world design optimization tasks, the objective and constraint functions are non-linear.

One of the key problems in solving the optimization is to handle the uncertainty, since in practice, the evaluation of inputs can introduce noises and even the objective function could be dynamic in time. As an example, environmental parameters such as pressure, humidity etc. vary unexpectedly at times. Likewise, design parameters can fluctuate because of floating point arithmetic, division by zero and quality compromises which are associated with high precision costs in manufacturing [11]. Furthermore, the systematic errors are unavoidable in the modeling of physical processes.

An example of an optimization problem with various sources of uncertainty is illustrated in figure 1. Uncertainties in parameters are related with sensitivity robustness [18], [19] whereas uncertainty in objectives and constraints is strongly linked with reliability [20], [21] and feasibility robustness [22]. Parameters in a design optimization problem can be classified into two types, namely design and environmental ones. Incorporating robustness in design optimization with respect to these parameters leads to the concept of sensitivity robustness. The effect of additive noise $\delta_{\mathbf{x}} \in \mathbb{R}^d$ in design parameters of an objective function is formulated in Eq. (1) where $\tilde{f}(\mathbf{x})$ is the noisy counter-part of an arbitrary objective function $f(\mathbf{x})$.

$$\tilde{f}(\mathbf{x}) = f(\mathbf{x} + \delta_{\mathbf{x}}) \tag{1}$$

Additionally, the operational or environmental conditions fluctuate or are known only to a certain extent. This means that our existing knowledge about the system is limited. In the classical optimization setting, environmental conditions are treated as constants. However, considering the different phases of the product life cycle that is hardly the case in a practical application. We thus extend the output dependency to the environmental variables set $\mathcal{C}$ where $\alpha \in \mathcal{C}$ represents an individual environmental variable. The effect of such additive noise $\delta_\alpha \in \mathbb{R}$ on the output of an objective function is represented in Eq. (2).

$$\tilde{f}(\mathbf{x}; \alpha) = f(\mathbf{x}; \alpha + \delta_\alpha) \tag{2}$$

In this work, we employ the robust counter-part approach [16] to achieve robustness. The idea behind the robust counter-part approach is to reformulate the optimization problem in a way so as to minimize the effect of parametric uncertainties $\delta_{\mathbf{x}}$ and δ_α in Eqs. (1) and (2). Hence, it is important to model $\delta_{\mathbf{x}}$ and δ_α to design a robust system. The robust counter-part approach also requires to reformulate the optimization goal, i.e., whether to optimize the expected or worst case behavior of the objective function. Further, the constraint functions are revised to design a robust counter-part which by definition includes the robustness w.r.t. parametric imprecision. Finally, the resulting robust counter-part is substituted in place of the original problem and solved using any optimization methodology of interest. We study two robust counter-part approaches, namely the robust regularization and robust composition.

A. Robust Regularization

It can be stated without loss of generality that robust regularization works on the "minimax" principle [23] as it minimizes the worst output of the noisy function $\tilde{f}(\mathbf{x})$. Given an objective function $f(\mathbf{x})$ to be minimized, this robust

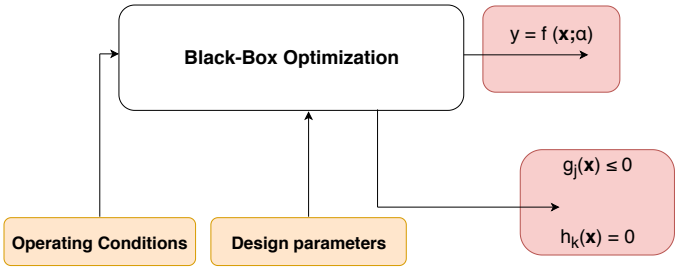

Fig. 1. An example of a constrained, single-objective design optimization problem with four common sources of uncertainty (i.e., design and environmental parameters in orange shade rectangles, objective and constraints in orchid shade rectangles).

counter-part approach first defines a neighbourhood $N_\epsilon(\mathbf{x})$ of design $\mathbf{x}$ whose size is determined by the parameter ϵ [13]. Then, an upper-bound of $f(\mathbf{x})$ is defined by taking into account the worst value of design $\mathbf{x}$ including neighborhood $N_\epsilon(\mathbf{x})$. Finally, the optimization goal is characterized so as to minimize this upper bound. The minimum returned by using this strategy is called "least upper-bound". Formally, the robust counter-part $R(\mathbf{x})$ based on robust regularization is defined in Eq. (3).

$$R(\mathbf{x}) = \sup_{\xi \in N_\epsilon(\mathbf{x})} f(\boldsymbol{\xi}) \tag{3}$$

The optimization goal then becomes to solve Eq. (3). This approach is called robust regularization because ϵ in Eq. (3) acts as a regularization parameter (i.e., defines the size of the neighbourhood).

B. Robust Composition

Robust regularization is a theoretically stable strategy to design systems with sensitivity robustness and has been discussed in operations research [16]. Different from robust regularization, engineers can also optimize the expected output of a noisy function $E[\tilde{f}(\mathbf{x})]$ while minimizing the dispersion $(V[\tilde{f}(\mathbf{x})])^{0.5}$ simultaneously. We refer to it as robust composition similar to [8]. Robust composition requires the uncertainty $\delta_\mathbf{x}$ be specified in the form of a probability distribution. The expectation $E[\tilde{f}(\mathbf{x})]$ and dispersion (i.e., standard deviation) $(V[\tilde{f}(\mathbf{x})])^{0.5}$ of the noisy function are combined at each point $\mathbf{x}$ in $\mathcal{S}$ to produce a robust scalar output. The optimization goal thus becomes to find a point $\mathbf{x}^*$ in $\mathcal{S}$ which minimizes this scalar. The robust counter-part with robust composition is defined in Eq. (4). In this paper, we assume $\delta_\mathbf{x} \sim \mathcal{N}(0, \sigma^2)$ where σ^2 depends on the range of parameters and noise level.

$$\begin{aligned} \text{minimize} \quad & R(\mathbf{x}) = E[\tilde{f}(\mathbf{x})] + \sqrt{V[\tilde{f}(\mathbf{x})]} \\ & \tilde{f}(\mathbf{x}) = f(\mathbf{x} + \delta_\mathbf{x}) \\ & \delta_\mathbf{x} \sim \mathcal{N}(0, \sigma^2) \end{aligned} \tag{4}$$

III. Meta-Modeling in Optimization

The idea of using a simple empirical approximation to substitute a complex model was proposed as early as 1974 [24]. This line of research was particularly used in structural optimization [25]–[29] with the name of Response Surface Approximation. The first attempt to classify such methodologies based on their accuracy in the context of structural engineering was made in 1993 [30]. On the other hand, the first significant attempt to replace a deterministic computer simulation with a meta-model was presented in [31]. Meta-modeling for robust design optimization has been investigated in [8], [9], [32]–[35]. For a detailed review of meta-modeling and its applications in structural engineering, the reader is directed to [8].

The idea behind meta-modeling is to build an empirical approximation model $\hat{f}(\mathbf{x})$ of an objective function $f(\mathbf{x})$. The approximation $\hat{f}(\mathbf{x})$ then acts as the meta-model (aka surrogate-model) of $f(\mathbf{x})$. This abstraction is useful in a variety of situations. Firstly, it simplifies the task to a great extent in simulation based modeling and optimization. Secondly, it provides the engineer with an opportunity to evaluate $f(\mathbf{x})$ indirectly if the exact computation of $f(\mathbf{x})$ is too costly or complex. Additionally, it provides the engineer with practically useful insights. Meta-modeling first evaluates $f(\mathbf{x})$ at several points of interest $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, e.g., using Latin hyper-cube sampling, Plackett-Burman [36], Box-Behnken [37] design etc., and generates the data set of input $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ and output $\{f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_N)\}$ pairs, i.e., sample points and resulting function values. The data generated can then be used to build a nonlinear regression model for approximating the objective function $f(\mathbf{x})$. In principle, any regression based machine learning methodology (e.g., Kriging, neural networks, polynomials etc.) may be used. Since no regression model can perfectly approximate the original function $f(\mathbf{x})$ by means of a limited number of evaluations, the resulting meta-model $\hat{f}(\mathbf{x})$ will be relatively inaccurate.

Although meta-modeling is relatively easy to use compared to classical optimization, it must be carefully designed. Firstly, selecting the set of points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ to evaluate $f(\mathbf{x})$ is not straightforward and requires some experience. To this end, the engineer can take help from design of experiment methodologies such as Latin hyper-cube sampling, factorial designs etc. Secondly, the sample size N to evaluate $f(\mathbf{x})$ is important as well. This is since the engineer is interested to find a meta-model $\hat{f}(\mathbf{x})$ as good as the original function $f(\mathbf{x})$ with minimum data i.e., function evaluations. Consequently, in most applications, the designer must come to a compromise on the accuracy of meta-model and computational budget. Additionally, evaluating the accuracy of a meta-model might have many aspects such as modeling accuracy and quality of an optimal solution of the problem etc. As such, choosing the proper criteria to evaluate the meta-model is important.

IV. Experimental Setup

In this section, we first report the optimization problems studied for this work. We then move forward to outline three levels of noise and their effect on the robust counter-part approach. This refers to the integration of each noise level with robustness schemes (i.e., robust regularization and robust composition). Lastly, we describe the meta-modeling

TABLE I
ALL SIX OPTIMIZATION PROBLEMS WITH TEST FUNCTIONS, KEY LANDSCAPE CHARACTERISTICS, DIMENSIONS, AND BOX CONSTRAINTS.

Function	Landscape	Dim.	Bounds
Ackley	Multi-Modal	2	$x_i \in [-32.768, 32.768]$
Branin	Multi-Global	2	$x_1 \in [-5, 10], x_2 \in [0, 15]$
Ackley	Multi-Modal	5	$x_i \in [-32.768, 32.768]$
Sphere	Isotropic	5	$x_i \in [-5, 5]$
Ackley	Multi-Modal	10	$x_i \in [-32.768, 32.768]$
Rastrigin	Multi-Modal	10	$x_i \in [-5.12, 5.12]$

techniques and the evaluation criteria to compare the meta-models.

We choose six optimization problems. Each problem is uniquely identified based on the choice of test function and dimensionality $D \in \{2, 5, 10\}$. We select the test functions known as Ackley, Branin, Sphere and Rastrigin. Among the test functions, Branin is only defined for $2D$, Sphere for $5D$, Rastrigin for $10D$ and Ackley is tested for all of these dimensions. This results in six optimization problems. Additionally, each one of these problems is investigated on three levels of additive noise (i.e., 5, 10 and 20% noise perturbation) and two robustness strategies (i.e., robust regularization according to Eq. (3) and robust composition according to Eq. (4)). All six optimization problems are presented in Table I, including the box constraints and key landscape characteristics.

We employ three levels of additive noise $\delta_{\mathbf{x}}$. The effect of additive noise in design and environmental parameters has already been presented in Eq. (1) and Eq. (2) respectively. Let $R = |u - l|$ be the absolute range of parameters where u and l serve as the upper and lower limits of design parameters. Further, let Z be the additive noise level. In the case of robust regularization, this means having a neighbourhood of design $\mathbf{x}$ whose scale ϵ is defined by the parameters range R and noise level Z. As an example, the Ackley function is defined from $l = -32.768$ to $u = 32.768$, having an absolute range of $R = 65.536$. Considering the first noise level, i.e., $Z = 5\%$ in Eq. (3), this means the regularization parameter $\epsilon = ZR = 0.05 \cdot 65.536 = 3.2768$. Throughout this paper, we assume the noise is symmetric[1], hence a neighbourhood $[-3.2768, 3.2768]$ of design $\mathbf{x}$ is constructed and Eq. (3) can be solved. In this work, we classify all parameters as design, however, it is important to note that real-world engineering applications have additional environmental parameters which can be modeled in exactly the same way since Eqs. (1) and (2) prescribe the similar additive noise.

For robust composition in this paper, we employ a normal distribution $\mathcal{N}(0, \sigma^2)$ where the standard deviation $\sigma^2 = ZR/6$ and Z and R serve as the noise level and absolute range of parameters, same as above. Once the noise is specified in the form of a probability distribution, the robust composition in Eq. (4) can be solved. Robust composition can be labelled as probabilistic treatment of uncertainty as opposed to robust regularization which is classified as deterministic treatment of uncertainty.

[1]This assumption is not always realistic and would be focused upon in the future research with asymmetric noise cases.

We evaluate and compare six machine learning based meta-modeling techniques: Kriging, Support Vector Machines (SVM), Radial Basis Function Network (RBFN), Random Forest (RF), K-Nearest Neighbors (KNN) and Elastic-net with second order polynomial regression function (ELN) to build a meta-model of $R(\mathbf{x})$ in Eqs. (3) and (4) for each optimization problem with three levels of noise. We evaluate the meta-models on the basis of modeling accuracy and optimal values of $R(\mathbf{x})$, taking the so-called relative mean absolute error:

$$\text{RMAE} = \frac{1}{M} \sum_{i=1}^{M} 100 \cdot \left(\frac{|y_i - \hat{y}_i|}{|y_i|} \right) \tag{5}$$

More specifically, in all cases, we train these meta-modeling techniques on ten different training sample sizes $N \in \{5D, 10D, 15D, 20D, 25D, 30D, 35D, 40D, 45D, 50D\}$ and evaluate the resulting meta-model each time on a test data set with size $M = 75D$. Here, D are the dimensions of the problem and N and M serve as the training and test sample sizes respectively. The sample points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\}$ to train and test the meta-model are generated using Latin hyper-cube sampling. Additionally, to achieve the best results in model training, we perform a detailed hyper-parameter optimization with cross validation. Finally, to evaluate the meta-models for accuracy, we choose % relative mean absolute error (RMAE) as presented in Eq. (5). In Eq. (5), y and $\hat{y}$ are arbitrary target and predicted values respectively whereas M denotes the size of the test data set. This criterion helps us understand the accuracy and computational tractability (i.e., modeling accuracy vs computational budget) of meta-modeling techniques.

The second criterion to evaluate the meta-models is based on the optimal values of $R(\mathbf{x})$ in Eqs. (3) and (4) found by each meta-model. To find the optimal values on the meta-models $\hat{R}(\mathbf{x})$ and the original model $R(\mathbf{x})$, a benchmark optimization algorithm is run. To this end, the Sequential Least Square Programming (SLSQP) [38] is chosen as benchmark. This criterion helps us understand the reliability of meta-models in practical situations. To evaluate the meta-models on this criterion, each meta-model is first trained using hyper-parameter optimization on a training sample of $N = 50D$ where D denotes the dimensions of the problem. An optimization run with SLSQP is then performed on the trained meta-models to find the optimal values of $R(\mathbf{x})$. This process is repeated for 100 times, and the mode of the group is chosen as the final optimal value of the $R(\mathbf{x})$ using the meta-model.

We now move on to discuss the results obtained from this experimental design.

V. RESULTS

Graphs showing the accuracy of meta-models by varying the training size, evaluated on the basis of RMAE are presented in figures 2-7. Standard error (SE) for each RMAE computation is also presented in the graphs. In particular, figure 2 shows the

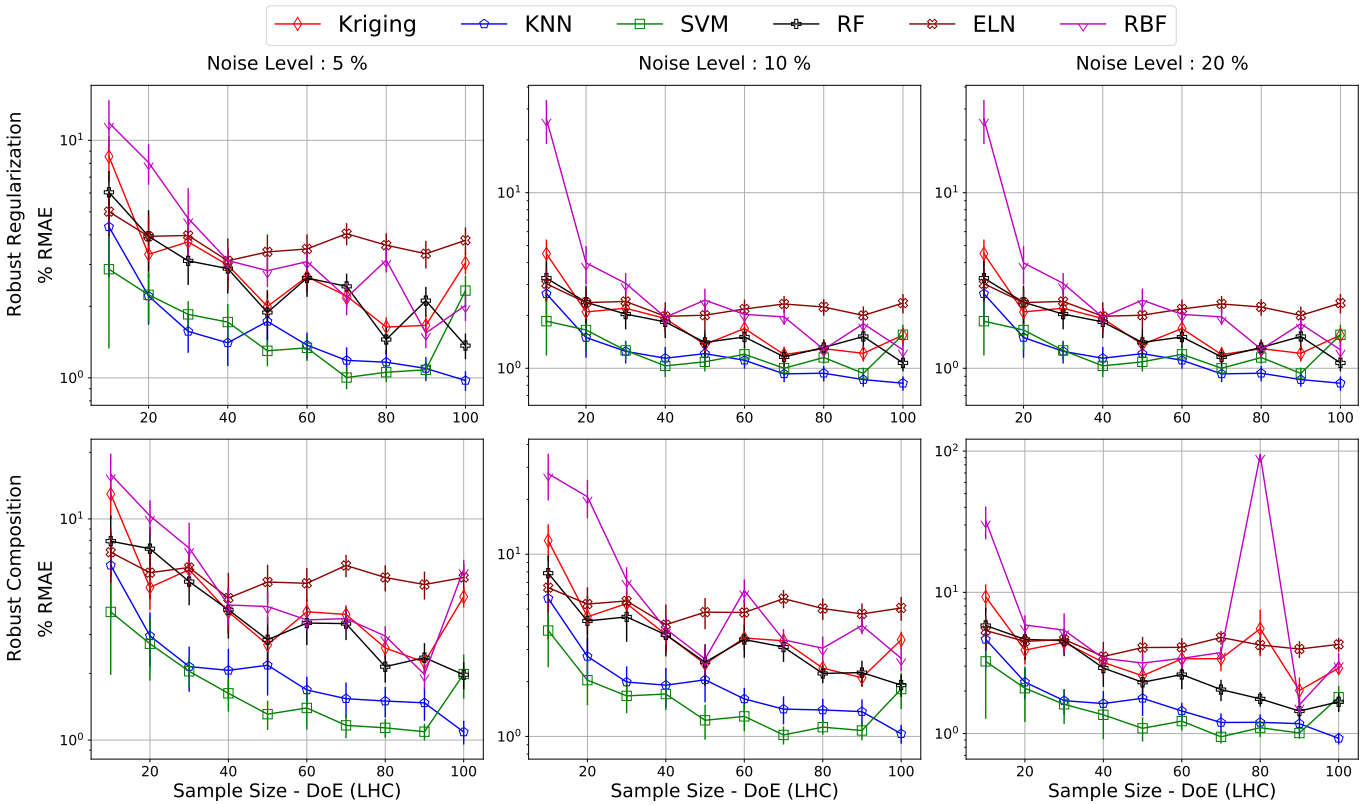

Fig. 2. Modeling accuracy of meta-model techniques on Ackley $2D$ with three noise levels, two robustness strategies and ten training sample sizes.

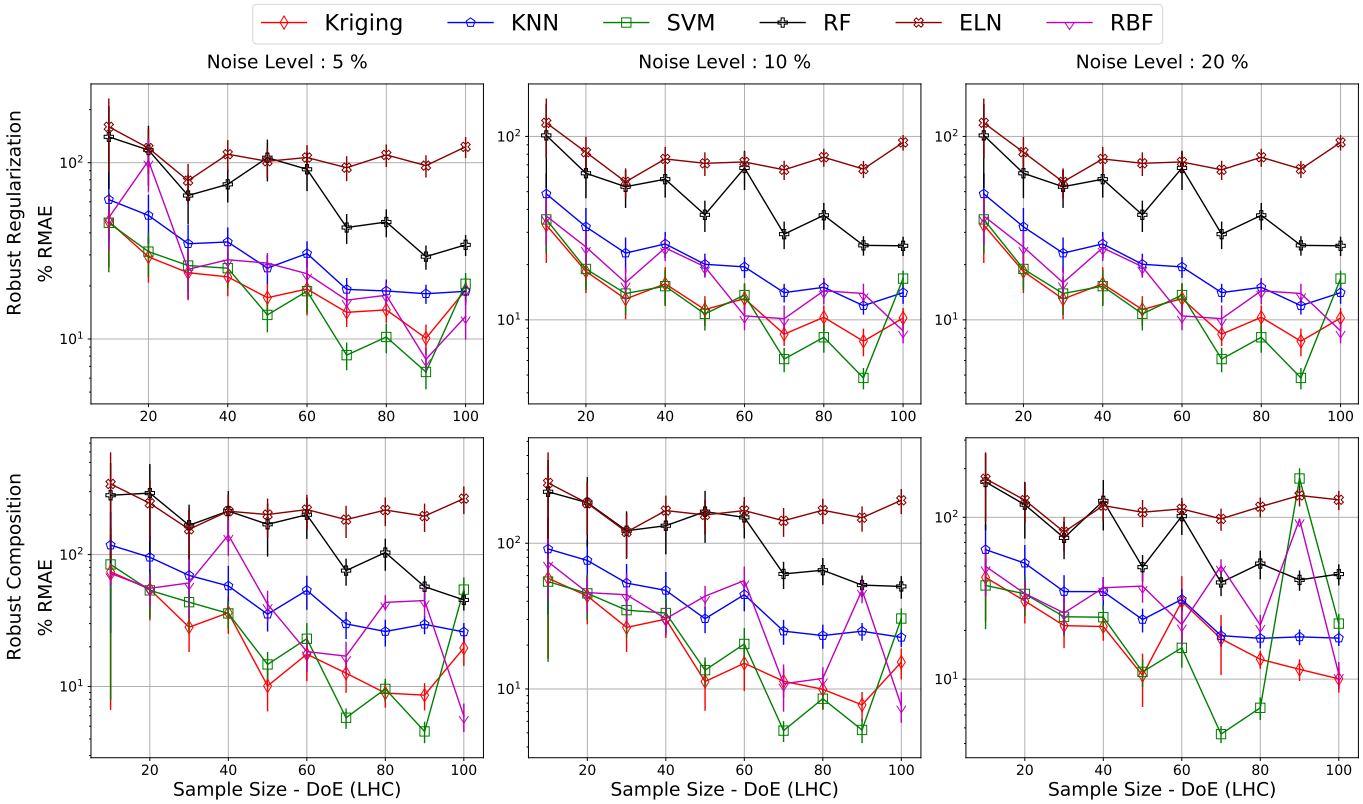

Fig. 3. Modeling accuracy of meta-model techniques on Branin $2D$ with three noise levels, two robustness strategies and ten training sample sizes.

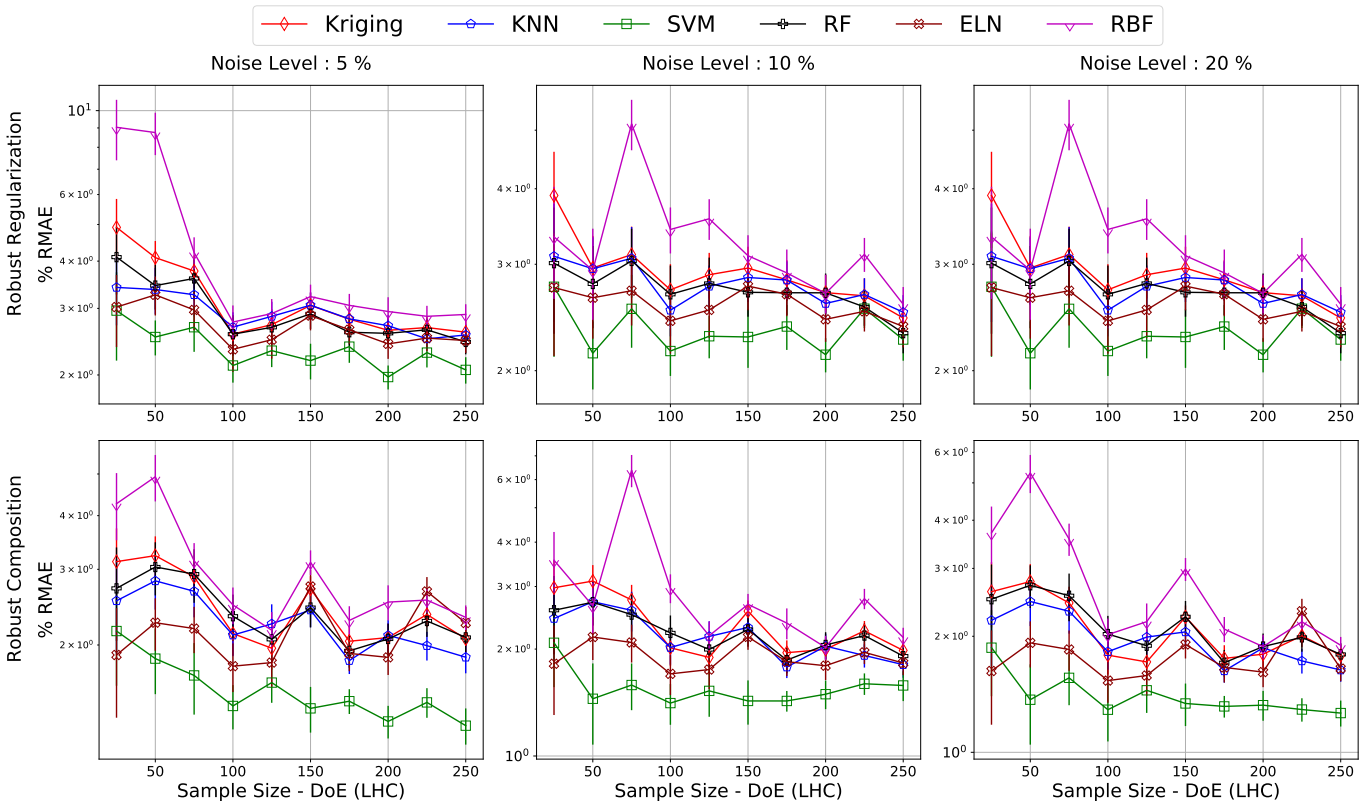

Fig. 4. Modeling accuracy of meta-model techniques on Ackley $5D$ with three noise levels, two robustness strategies and ten training sample sizes.

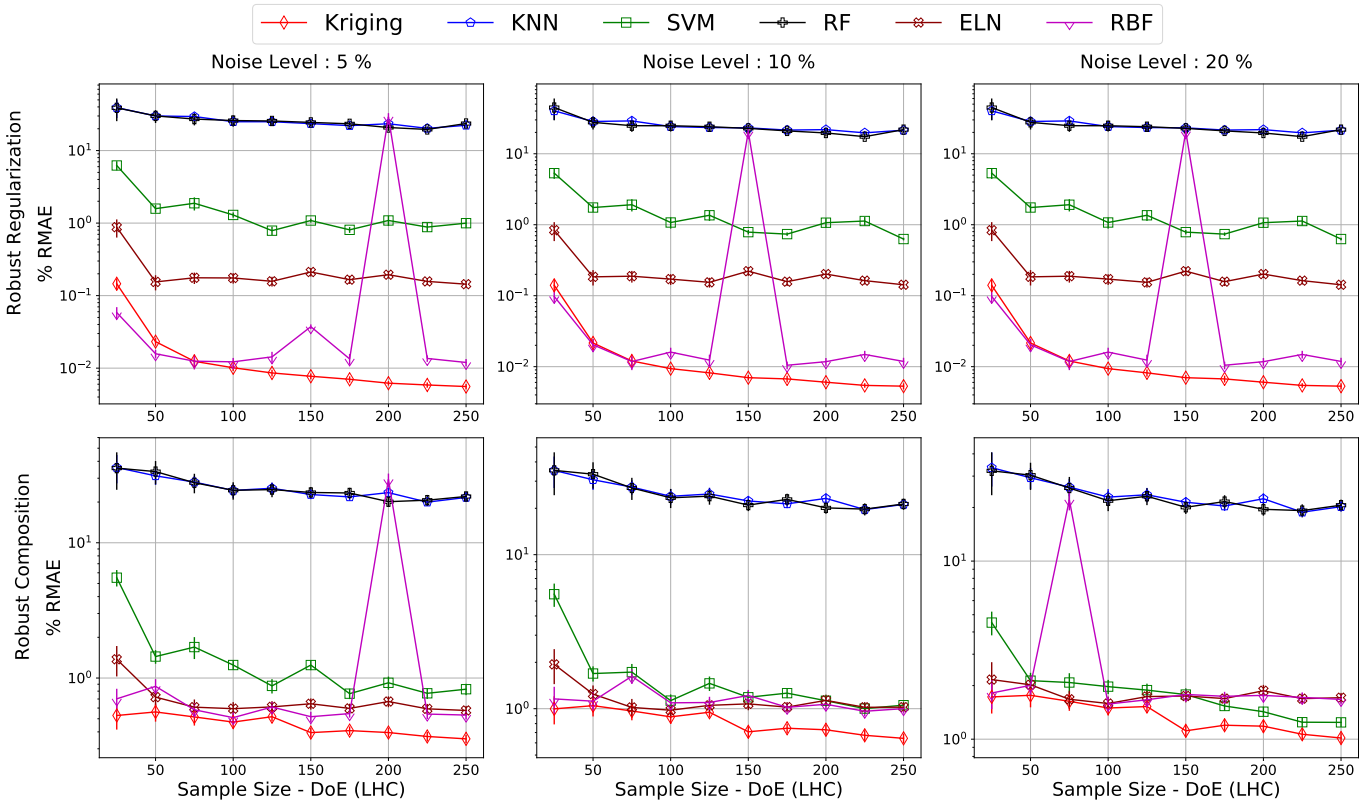

Fig. 5. Modeling accuracy of meta-model techniques on Sphere $5D$ with three noise levels, two robustness strategies and ten training sample sizes.

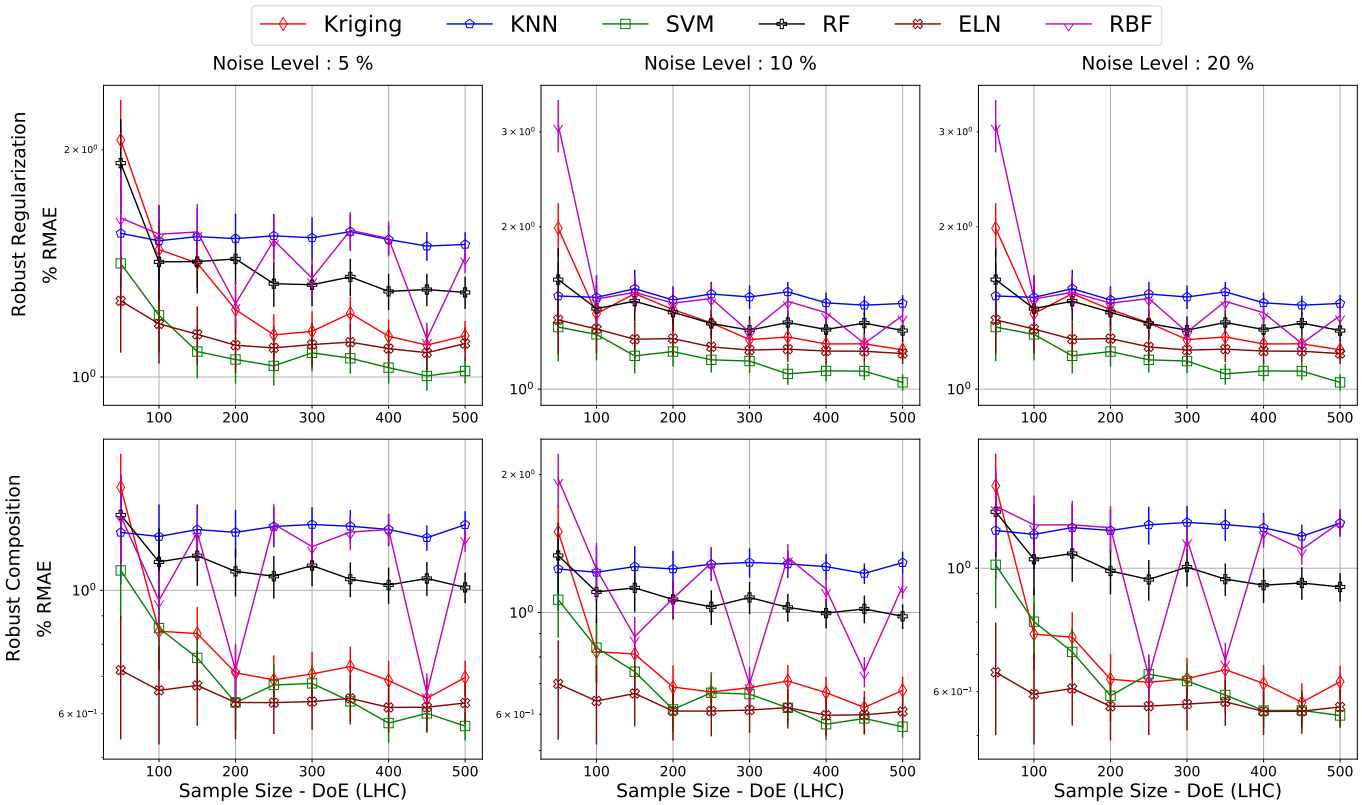

Fig. 6. Modeling accuracy of meta-model techniques on Ackley $10D$ with three noise levels, two robustness strategies and ten training sample sizes.

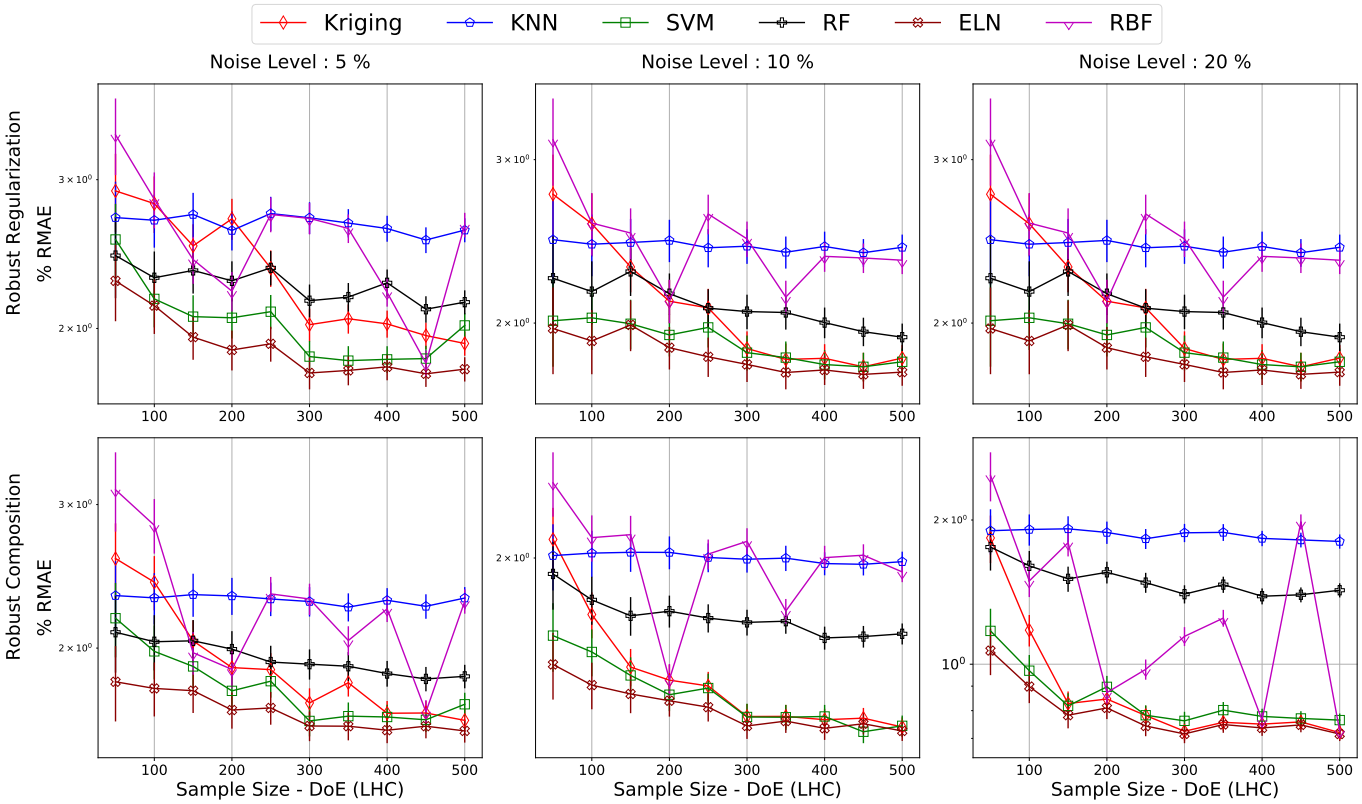

Fig. 7. Modeling accuracy of meta-model techniques on Rastrigin $10D$ with three noise levels, two robustness strategies and ten training sample sizes.

TABLE II

ALL 36 TEST CASES RESULTED FROM THE COMBINATION OF TWO META-MODELS, THREE LEVELS (NL), TWO ROBUST DEFINITIONS (ROBUST) AND FOUR FUNCTIONS. IN EACH TEST CASE, WE PICK THE BEST TWO META-MODELS IN TERMS OF AVERAGE RMAE. GIVEN THE ALTERNATIVE HYPOTHESIS (H_a), THE MANN WHITNEY U TEST IS PERFORMED TO CHECK IF THE META-MODEL WITH HIGHEST AVERAGE ACCURACY IS SIGNIFICANTLY BETTER THAN THE SECOND BEST USING $\alpha = 0.05$. THE RESULTING p-VALUES ARE PRESENTED.

Problem	Robust	NL	H_a	p-value
Ackley$2D$	RR	5%	KNN<SVM	0.425
Ackley$2D$	RR	10%	KNN<SVM	0.192
Ackley$2D$	RR	20%	KNN<SVM	**0.012**
Ackley$2D$	RC	5%	SVM<KNN	0.153
Ackley$2D$	RC	10%	SVM<KNN	0.136
Ackley$2D$	RC	20%	SVM<KNN	0.285
Branin$2D$	RR	5%	SVM<Kriging	0.395
Branin$2D$	RR	10%	Kriging<SVM	0.425
Branin$2D$	RR	20%	Kriging<SVM	0.153
Branin$2D$	RC	5%	Kriging<SVM	0.425
Branin$2D$	RC	10%	Kriging<SVM	0.425
Branin$2D$	RC	20%	Kriging<KNN	**0.037**
Ackley$5D$	RR	5%	SVM<ELN	**0.010**
Ackley$5D$	RR	10%	SVM<ELN	**0.010**
Ackley$5D$	RR	20%	SVM<KNN	**0.010**
Ackley$5D$	RC	5%	SVM<ELN	**0.0001**
Ackley$5D$	RC	10%	SVM<ELN	**0.0001**
Ackley$5D$	RC	20%	SVM<ELN	**0.0001**
Sphere$5D$	RR	5%	Kriging<ELN	**0.0001**
Sphere$5D$	RR	10%	Kriging<ELN	**0.0001**
Sphere$5D$	RR	20%	Kriging<ELN	**0.0001**
Sphere$5D$	RC	5%	Kriging<ELN	**0.0001**
Sphere$5D$	RC	10%	Kriging<RBFN	**0.0001**
Sphere$5D$	RC	20%	Kriging<ELN	**0.002**
Ackley$10D$	RR	5%	SVM<ELN	**0.012**
Ackley$10D$	RR	10%	SVM<ELN	**0.008**
Ackley$10D$	RR	20%	SVM<ELN	**0.015**
Ackley$10D$	RC	5%	ELN<SVM	0.285
Ackley$10D$	RC	10%	ELN<SVM	0.260
Ackley$10D$	RC	20%	ELN<SVM	0.120
Rastrigin$10D$	RR	5%	ELN<SVM	0.060
Rastrigin$10D$	RR	10%	ELN<SVM	0.060
Rastrigin$10D$	RR	20%	ELN<SVM	0.236
Rastrigin$10D$	RC	5%	ELN<SVM	**0.037**
Rastrigin$10D$	RC	10%	ELN<SVM	0.192
Rastrigin$10D$	RC	20%	ELN<SVM	0.052

results on Ackley $2D$. Similarly, figure 3 shows the accuracy on Branin $2D$, figure 4 shows the accuracy on Ackley $5D$, figure 5 presents the results for Sphere $5D$, figure 6 presents the results concerning Ackley $10D$ and lastly, figure 7 shows the results on Rastrigin $10D$. It is important to state, that for each optimization case, we select the two best meta-models based on the measure of lowest RMAE, which is averaged over all values of training size N. We further perform Mann-Whitney U test to find if the meta-model with the lowest average RMAE is significantly better than the other. The resulting p-values are presented in Table II. In that table, the first column reads the optimization problem under consideration, the second column presents the robustness strategies (i.e., RR for robust regularization and RC for robust composition), the third column reports the noise-level, the fourth column describes the two best meta-modeling techniques based on the lowest RMAE (i.e., alternative hypothesis) while the last column shows the p-value resulting from the Mann-Whitney U test. The frequencies of meta-modeling techniques achieving the highest accuracy for all optimization problems are presented in Table III. This Table follows the similar evaluation criteria i.e., lowest average RMAE.

TABLE III

THE FREQUENCY OF META-MODELING TECHNIQUES ACHIEVING HIGHEST ACCURACY (I.E., BASED ON THE VALUE OF LOWEST AVERAGE RMAE) FOR ALL SIX OPTIMIZATION PROBLEMS IS PRESENTED.

Opt.Problem	Kriging	SVM	RBFN	KNN	RF	ELN
Ackley$2D$	0/6	4/6	0/6	2/6	0/6	0/6
Branin$2D$	5/6	1/6	0/6	0/6	0/6	0/6
Ackley$5D$	0/6	6/6	0/6	0/6	0/6	0/6
Sphere$5D$	6/6	0/6	0/6	0/6	0/6	0/6
Ackley$10D$	0/6	3/6	0/6	0/6	0/6	3/6
Rastrigin$10D$	0/6	0/6	0/6	0/6	0/6	6/6

The results concerning the optimality of $R(\mathbf{x})$ for all thirty six cases are presented in Table IV. In this table, the first column reports the optimization problem, the second column reads the robustness scheme (i.e., RR for robust regularization and RC for robust composition), the third column presents the noise level, the fourth column shows the optimal value of $R(\mathbf{x})$, and all the next columns depict the optimal values of $R(\mathbf{x})$ proposed by all the meta-models.

Results from figures 2-7 suggest that robustness schemes and noise levels do not have much effect on the accuracy of meta-models as we mostly observe similar patterns (i.e., RMAE curves) across rows and columns. Additionally, these figures depict that the Radial Basis Function Network (RBFN) has a lot of variance in prediction in all cases. Furthermore, these figures characterize that setting training size $N = KD$ generally results in good modeling accuracy. Here $K \in \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$ is a scalar and D denotes the dimensions of the test problem. Hence, it can be observed that computational complexity of meta-models in this paper is a linear function of D. Table II shows that Kriging, Support Vector Machines (SVM) and Polynomial regression (ELN) achieve high accuracy in most test cases. In particular, Kriging performs well on Branin $2D$ and Sphere $5D$. Polynomial regression (ELN) performs well on all $5D$ and $10D$ cases whereas SVM performs excellently in most cases for all of these dimensions. Table III shows that K-Nearest Neighbors (KNN) and Random Forest (RF) generally do not achieve high accuracy compared to other meta-models. Finally, Table IV shows that meta-models were able to find an optimal or near optimal solution in most cases except Rastrigin$10D$.

VI. CONCLUSIONS AND OUTLOOK

In this paper, we evaluate and compare the meta-modeling techniques for the scenario of robust optimization on a variety of test cases. The findings suggest the usefulness of meta-models for robust design optimization. This is because in 8/36 cases reported in Table IV, at least one meta-model technique achieves the optimal function value i.e., based on mode and

TABLE IV

FINAL OPTIMAL FUNCTION VALUES FOR ALL THIRTY SIX CASES BY THE ORIGINAL MODEL AND ALL THE META-MODELS IS PRESENTED. FOR OPTIMIZATION, SLSQP [38] IS EMPLOYED ON THE ORIGINAL MODEL AND ALL THE META-MODELS. THE FUNCTION VALUES IN THE TABLE REPRESENT THE MODE OF 100 RUNS ALONGSIDE THE STANDARD DEVIATION, BOTH ROUNDED OFF TO THE INTEGER REPRESENTATION. THE STARTING VALUE FOR SLSQP IS RANDOMLY SELECTED FROM LATIN HYPER-CUBE SAMPLING FOR EACH SINGLE RUN. IN EACH CASE, META-MODELS WITH MOST OPTIMAL FUNCTION VALUES (I.E., BASED ON MODE AND IF TIED, BASED ON STANDARD DEVIATION) ARE HIGHLIGHTED.

Opt.Problem	Robustness	Noise Level	Original-Model	Kriging	SVM	RBFN	KNN	RF	ELN
Ackley2D	RR	5%	11 ± 3	14 ± 2	22 ± 3	13 ± 2	$\mathbf{12 \pm 3}$	14 ± 2	$\mathbf{12 \pm 3}$
Ackley2D	RR	10%	16 ± 1	17 ± 1	22 ± 1	18 ± 1	17 ± 1	17 ± 1	$\mathbf{16 \pm 0}$
Ackley2D	RR	20%	20 ± 0	$\mathbf{21 \pm 0}$	22 ± 0	$\mathbf{21 \pm 0}$	$\mathbf{21 \pm 0}$	$\mathbf{21 \pm 0}$	$\mathbf{21 \pm 0}$
Ackley2D	RC	5%	5 ± 7	10 ± 2	22 ± 7	10 ± 2	6 ± 4	10 ± 2	$\mathbf{5 \pm 5}$
Ackley2D	RC	10%	7 ± 6	10 ± 2	22 ± 7	11 ± 2	7 ± 4	11 ± 2	$\mathbf{7 \pm 0}$
Ackley2D	RC	20%	10 ± 5	12 ± 2	22 ± 5	12 ± 2	11 ± 3	12 ± 2	$\mathbf{10 \pm 0}$
Branin2D	RR	5%	3 ± 3	6 ± 64	9 ± 3	6 ± 64	$\mathbf{6 \pm 52}$	6 ± 64	17 ± 0
Branin2D	RR	10%	9 ± 4	13 ± 83	16 ± 3	13 ± 83	$\mathbf{13 \pm 56}$	13 ± 83	25 ± 0
Branin2D	RR	20%	20 ± 0	20 ± 129	$\mathbf{20 \pm 0}$	20 ± 146	20 ± 103	20 ± 146	106 ± 0
Branin2D	RC	5%	1 ± 0	1 ± 52	3 ± 7	$\mathbf{1 \pm 51}$	2 ± 43	2 ± 52	12 ± 0
Branin2D	RC	10%	1 ± 1	2 ± 54	4 ± 6	2 ± 54	$\mathbf{2 \pm 45}$	2 ± 54	13 ± 0
Branin2D	RC	20%	3 ± 3	4 ± 59	6 ± 4	4 ± 45	$\mathbf{3 \pm 48}$	4 ± 59	15 ± 0
Ackley5D	RR	5%	16 ± 1	16 ± 1	$\mathbf{12 \pm 1}$	17 ± 1	12 ± 2	17 ± 1	22 ± 0
Ackley5D	RR	10%	14 ± 1	16 ± 1	20 ± 0	16 ± 1	$\mathbf{14 \pm 1}$	16 ± 1	22 ± 0
Ackley5D	RR	20%	16 ± 1	18 ± 1	21 ± 0	18 ± 1	$\mathbf{14 \pm 1}$	18 ± 1	22 ± 0
Ackley5D	RC	5%	5 ± 5	19 ± 1	$\mathbf{12 \pm 2}$	18 ± 1	14 ± 2	19 ± 1	22 ± 0
Ackley5D	RC	10%	7 ± 5	19 ± 1	$\mathbf{9 \pm 2}$	18 ± 1	14 ± 2	19 ± 1	22 ± 0
Ackley5D	RC	20%	10 ± 4	19 ± 1	16 ± 1	22 ± 3	$\mathbf{15 \pm 2}$	19 ± 1	22 ± 0
Sphere5D	RR	5%	0 ± 0	0 ± 7	$\mathbf{0 \pm 0}$	1 ± 8	0 ± 12	7 ± 19	$\mathbf{0 \pm 0}$
Sphere5D	RR	10%	0 ± 0	0 ± 8	$\mathbf{0 \pm 0}$	0 ± 7	1 ± 12	5 ± 22	$\mathbf{0 \pm 0}$
Sphere5D	RR	20%	0 ± 0	0 ± 9	$\mathbf{0 \pm 0}$	1 ± 16	1 ± 12	8 ± 31	$\mathbf{0 \pm 0}$
Sphere5D	RC	5%	0 ± 0	5 ± 16	$\mathbf{0 \pm 0}$	9 ± 15	1 ± 7	9 ± 18	$\mathbf{0 \pm 0}$
Sphere5D	RC	10%	0 ± 0	9 ± 16	$\mathbf{0 \pm 0}$	7 ± 17	1 ± 8	10 ± 18	$\mathbf{0 \pm 0}$
Sphere5D	RC	20%	1 ± 0	10 ± 18	$\mathbf{1 \pm 0}$	10 ± 17	2 ± 11	11 ± 18	$\mathbf{1 \pm 0}$
Ackley10D	RR	5%	18 ± 0	19 ± 0	$\mathbf{8 \pm 0}$	19 ± 0	19 ± 0	19 ± 0	$\mathbf{8 \pm 0}$
Ackley10D	RR	10%	18 ± 0	20 ± 0	8 ± 0	19 ± 0	15 ± 1	20 ± 0	$\mathbf{7 \pm 0}$
Ackley10D	RR	20%	19 ± 0	20 ± 0	$\mathbf{12 \pm 0}$	19 ± 0	17 ± 1	20 ± 0	22 ± 0
Ackley10D	RC	5%	5 ± 5	20 ± 0	8 ± 0	20 ± 0	20 ± 0	20 ± 0	$\mathbf{5 \pm 0}$
Ackley10D	RC	10%	7 ± 5	20 ± 0	9 ± 0	20 ± 0	20 ± 0	20 ± 0	$\mathbf{7 \pm 0}$
Ackley10D	RC	20%	10 ± 5	20 ± 0	11 ± 0	21 ± 0	21 ± 0	21 ± 0	$\mathbf{10 \pm 0}$
Rastrigin10D	RR	5%	918 ± 27	1021 ± 33	1013 ± 0	1154 ± 28	$\mathbf{975 \pm 41}$	1020 ± 34	993 ± 0
Rastrigin10D	RR	10%	923 ± 28	1024 ± 40	987 ± 0	1083 ± 29	$\mathbf{985 \pm 31}$	1024 ± 38	996 ± 0
Rastrigin10D	RR	20%	919 ± 40	1038 ± 54	$\mathbf{952 \pm 0}$	1033 ± 50	975 ± 31	1047 ± 53	1032 ± 0
Rastrigin10D	RC	5%	946 ± 23	1020 ± 29	1034 ± 0	1164 ± 14	988 ± 31	1020 ± 29	$\mathbf{961 \pm 0}$
Rastrigin10D	RC	10%	995 ± 9	1042 ± 26	1005 ± 0	1134 ± 27	1013 ± 20	1041 ± 27	$\mathbf{996 \pm 0}$
Rastrigin10D	RC	20%	1003 ± 12	1058 ± 24	1045 ± 0	1041 ± 23	$\mathbf{1023 \pm 25}$	1064 ± 25	1055 ± 0

if tied, based on standard deviation when compared with the original model. Additionally, in 12/36 cases, at least one meta-model achieves a better optimal value of the function than the original model following the similar evaluation criteria. In most of the remaining cases, at least one technique achieves near optimal function value when compared with the original model. From the results, in particular Tables II, III and figures 2-7, the authors conclude that Kriging, SVM and ELN provide a high modeling accuracy with limited training data in most cases. Additionally, these techniques find optimal or near optimal function values compared with the original model in most cases as presented in Table IV. The findings are a further validation of Kriging and Polynomial regression (i.e., in response surface approximation in structural engineering) as two of the most accurate meta-modeling techniques. The study also suggests SVM as a promising and competitive meta-model technique since it provides the highest modeling accuracy in most cases i.e., lowest average RMAE in 14/36 as presented in Table III and proposes near optimal solutions in most $5D$ and $10D$ cases in as shown in Table IV.

Future investigations are necessary to validate these techniques on more complex problems, i.e., multi-objective cases, higher dimensions, asymmetric noise and real-world engineering case studies etc. Further, it is important to model multiplicative noise and define other robustness strategies characterizing the practical nature of robust design optimization. Finally, there is a dire need to produce tutorial based cohesive work for engineers to adopt the meta-models in practical cases.

ACKNOWLEDGMENT

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 766186.

REFERENCES

[1] G. Taguchi and M. S. Phadke, "Quality engineering through design optimization," in *Quality Control, Robust Design, and the Taguchi Method*, pp. 77–96, Springer, 1989.

[2] M. McIlhagga, P. Husbands, and R. Ives, "A comparison of search techniques on a wing-box optimisation problem," in *International Conference on Parallel Problem Solving from Nature*, pp. 614–623, Springer, 1996.

[3] D. Wiesmann, U. Hammel, and T. Bäck, "Robust design of multilayer optical coatings by means of evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 2, no. 4, pp. 162–167, 1998.

[4] J. W. Herrmann, "A genetic algorithm for minimax optimization problems," in *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, vol. 2, pp. 1099–1103, IEEE, 1999.

[5] E. Kazancioglu, G. Wu, J. Ko, S. Bohac, Z. Filipi, S. J. Hu, D. Assanis, and K. Saitou, "Robust optimization of an automobile valvetrain using a multiobjective genetic algorithm," in *Proceedings of DETC*, vol. 3, pp. 2–6, 2003.

[6] H.-G. Beyer and B. Sendhoff, "Evolution strategies for robust optimization," in *2006 IEEE International Conference on Evolutionary Computation*, pp. 1346–1353, IEEE, 2006.

[7] J. W. Kruisselbrink *et al.*, *Evolution strategies for robust optimization*. Leiden Institute of Advanced Computer Science (LIACS), Faculty of Science , 2012.

[8] F. Jurecka, *Robust design optimization based on metamodeling techniques*. PhD thesis, Technische Universität München, 2007.

[9] R. Jin, X. Du, and W. Chen, "The use of metamodeling techniques for optimization under uncertainty," *Structural and Multidisciplinary Optimization*, vol. 25, no. 2, pp. 99–116, 2003.

[10] X. Du and W. Chen, "Methodology for managing the effect of uncertainty in simulation-based design," *AIAA journal*, vol. 38, no. 8, pp. 1471–1478, 2000.

[11] H.-G. Beyer and B. Sendhoff, "Robust optimization–a comprehensive survey," *Computer methods in applied mechanics and engineering*, vol. 196, no. 33-34, pp. 3190–3218, 2007.

[12] G. Taguchi, S. Konishi, and S. Konishi, *Taguchi Methods: Orthogonal Arrays and Linear Graphs. Tools for Quality Engineering*. American Supplier Institute Dearborn, MI, 1987.

[13] M. W. Trosset, "Taguchi and robust optimization," tech. rep., 1996.

[14] V. N. Nair, B. Abraham, J. MacKay, G. Box, R. N. Kacker, T. J. Lorenzen, J. M. Lucas, R. H. Myers, G. G. Vining, J. A. Nelder, *et al.*, "Taguchi's parameter design: a panel discussion," *Technometrics*, vol. 34, no. 2, pp. 127–161, 1992.

[15] G. Box, S. Bisgaard, and C. Fung, "An explanation and critique of taguchi's contributions to quality engineering," *Quality and reliability engineering international*, vol. 4, no. 2, pp. 123–131, 1988.

[16] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*, vol. 28. Princeton University Press, 2009.

[17] Y. Carson and A. Maria, "Simulation optimization: methods and applications," in *Proceedings of the 29th conference on Winter simulation*, pp. 118–126, IEEE Computer Society, 1997.

[18] D. H. Jung and B. C. Lee, "Development of a simple and efficient method for robust optimization," *International Journal for Numerical Methods in Engineering*, vol. 53, no. 9, pp. 2201–2215, 2002.

[19] W. Chen, J. K. Allen, K.-L. Tsui, and F. Mistree, "A procedure for robust design: minimizing variations caused by noise factors and control factors," *Journal of mechanical design*, vol. 118, no. 4, pp. 478–485, 1996.

[20] H. Agarwal, *Reliability based design optimization: formulations and methodologies*. University of Notre Dame, 2004.

[21] M. Papadrakakis, N. D. Lagaros, and V. Plevris, "Design optimization of steel structures considering uncertainties," *Engineering Structures*, vol. 27, no. 9, pp. 1408–1418, 2005.

[22] X. Du and W. Chen, "Towards a better understanding of modeling feasibility robustness in engineering design," *Journal of Mechanical Design*, vol. 122, no. 4, pp. 385–394, 2000.

[23] L. El Ghaoui and H. Lebret, "Robust solutions to least-squares problems with uncertain data," *SIAM Journal on matrix analysis and applications*, vol. 18, no. 4, pp. 1035–1064, 1997.

[24] L. A. Schmit and B. Farshi, "Some approximation concepts for structural synthesis," *AIAA journal*, vol. 12, no. 5, pp. 692–699, 1974.

[25] K. Svanberg, "The method of moving asymptotesa new method for structural optimization," *International journal for numerical methods in engineering*, vol. 24, no. 2, pp. 359–373, 1987.

[26] V. Toropov, A. Filatov, and A. Polynkin, "Multiparameter structural optimization using fem and multipoint explicit approximations," *Structural optimization*, vol. 6, no. 1, pp. 7–14, 1993.

[27] W. Roux, N. Stander, and R. T. Haftka, "Response surface approximations for structural optimization," *International Journal for Numerical Methods in Engineering*, vol. 42, no. 3, pp. 517–534, 1998.

[28] G. E. Box and N. R. Draper, *Empirical model-building and response surfaces*. John Wiley & Sons, 1987.

[29] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, *Response surface methodology: process and product optimization using designed experiments*. John Wiley & Sons, 2016.

[30] J.-F. Barthelemy and R. T. Haftka, "Approximation concepts for optimum structural designa review," *Structural optimization*, vol. 5, no. 3, pp. 129–144, 1993.

[31] T. J. Santner, B. J. Williams, W. Notz, and B. J. Williams, *The design and analysis of computer experiments*, vol. 1. Springer, 2003.

[32] D. Lönn, Ø. Fyllingen, and L. Nilssona, "An approach to robust optimization of impact problems using random samples and metamodelling," *International Journal of Impact Engineering*, vol. 37, no. 6, pp. 723–734, 2010.

[33] V. Picheny and D. Ginsbourger, "Noisy kriging-based optimization methods: a unified implementation within the diceoptim package," *Computational Statistics & Data Analysis*, vol. 71, pp. 1035–1053, 2014.

[34] K.-H. Lee and G.-J. Park, "A global robust optimization using kriging based approximation model," *JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing*, vol. 49, no. 3, pp. 779–788, 2006.

[35] T. W. Simpson, T. M. Mauery, J. J. Korte, and F. Mistree, "Kriging models for global approximation in simulation-based multidisciplinary design optimization," *AIAA journal*, vol. 39, no. 12, pp. 2233–2241, 2001.

[36] K. Vanaja and R. Shobha Rani, "Design of experiments: concept and applications of plackett burman design," *Clinical research and regulatory affairs*, vol. 24, no. 1, pp. 1–23, 2007.

[37] S. C. Ferreira, R. Bruns, H. Ferreira, G. Matos, J. David, G. Brandao, E. P. da Silva, L. Portugal, P. Dos Reis, A. Souza, *et al.*, "Box-behnken design: an alternative for the optimization of analytical methods," *Analytica chimica acta*, vol. 597, no. 2, pp. 179–186, 2007.

[38] D. Kraft, "A software package for sequential quadratic programming," *Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt*, 1988.