Online supplement of Robust optimization in simulation: Taguchi and Krige combined

Gabriella Dellino
IMT Institute for Advanced Studies,
Piazza San Ponziano 6, 55100 Lucca, Italy, g.dellino@imtlucca.it
Jack P.C. Kleijnen
Department of Information Management / CentER,
Tilburg School of Economics and Management, Tilburg University, Postbox 90153, 5000 LE Tilburg, Netherlands, kleijnen@uvt.nl
Carlo Meloni
Department of Electrical Engineering and Electronics, Polytechnic of Bari,
Via E. Orabona 4, 70125 Bari, Italy, meloni@deemail.poliba.it

---

1. Introduction

This online publication supplements the article *Robust optimization in simulation: Taguchi and Krige combined* by G. Dellino, J.P.C. Kleijnen and C. Meloni. It provides a brief survey of robust simulation-optimization methods, introducing Taguchian and metamodel-based approaches. In particular, it summarizes Kriging metamodeling. It illustrates computational issues through an additional example based on Bertsimas et al. (2010), taking into account both environmental factors and implementation errors.

2. Robust simulation-optimization: an overview

The simulation-optimization process aims to identify the setting of input parameters leading to optimal system performance, evaluated through a simulation model of the system itself. The factors involved in the simulation model are often noisy and cannot be controlled or varied during the design process, due to measurement errors or other implementation issues; moreover, some factors are determined by the environment, not the managers. Therefore, the presumed optimal solution may turn out to be sub-optimal or even infeasible. Robust optimization tackles problems affected by uncertainty, providing solutions that are in some sense insensitive to perturbations in the model parameters.

Several alternative methods have been proposed for achieving robustness in simulation-based optimization problems, adopting different experimental designs and/or metamodeling techniques. The present section reviews the current state of the art on robust optimization approaches, focusing on simulated systems. First, we summarize robust Mathematical Programming. Then we discuss the approach introduced by Taguchi in the 1970s. Finally, we consider methods to tackle robustness using metamodels (Kriging, in particular).

2.1 Robust Mathematical Programming through Uncertainty Sets

The robust optimization methodology developed by Ben-Tal and Nemirovski (2008) investigates different choices of uncertainty sets to model data uncertainty, in order to characterize the structure of the resulting robust counterparts of optimization problems. In particular, their research focuses on robust formulations for Linear Programming (LP), Mixed Integer Programming (MIP), Second Order Cone Programming (SOCP), and Semidefinite Programming (SDP) problems. For this family of problems a fundamental issue is related to the feasibility of the solutions with respect to the classical optima; in particular, the challenge is to guarantee that the constraints will be satisfied for any possible value of the parameters in a given uncertainty set. The computational complexity of the deterministic problem and its robust counterpart is also investigated, to ensure that the problem remains tractable.

Although this approach has a strong theoretical background, there are several practical problems to which it cannot be applied, for many reasons (see Beyer and Sendhoff, 2007): the main disadvantage is the necessity of modeling a real-world problem through a linear model with (at most) conic or quadratic constraints. Moreover, in order to satisfy all the assumptions under which the method is applicable, the approximate model might become very complex and difficult to manage. Finally, if the objective function is not defined through a mathematical expression but can only be evaluated through simulations, the methodology cannot be applied.

Zhang (2004) deals with some of the aforementioned cases, proposing a mathematical formulation extending Ben-Tal's approach to parameterized nonlinear programming, with both equality and inequality constraints; the inequality constraints are supposed to be strictly satisfiable and are referred to as *safety constraints*. Zhang points out that his approach is especially suitable for applications where the satisfaction of safety constraints is of crucial importance. However, the formulation of the robust problem assumes that a reasonable estimate for the uncertain parameters is available, and the magnitude of the variations in

the uncertain parameters is relatively small. He proved that his formulation reduces to Ben-Tal's formulation when the objective function and the inequality constraints are linear and there is no uncertainty in the equality constraints. Anyway, further research is needed to develop algorithms able to effectively solve the proposed formulation.

Mainly based on Ben-Tal's approach, Bertsimas proposes a formulation for stochastic and dynamic optimization problems using uncertainty sets, in contrast to the stochastic programming approach which assumes full knowledge of the underlying probability distributions. Bertsimas and Sim (2004) propose a robust optimization methodology—based on linear and mixed-integer programming—to find an optimal supply chain control strategy, assuming stochastic demand. Their approach incorporates demand randomness in a deterministic manner, without making any specific assumption on the demand distribution. First, a robust formulation is given for the simple uncapacitated single-station case; then, capacity constraints are introduced, both on the orders and on the inventory level; finally, the network case is considered. The numerical experiments showed that, if only the mean and the variance of the demand distribution are known, the robust policy often outperforms the nominal policy, and also policies computed assuming full but erroneous knowledge of the demand distribution. The authors also prove that the nominal problem and its robust counterpart belong to the same complexity class, and that the robust formulation does not suffer from the curse of dimensionality. The method guarantees the robust solution to be feasible if less than a prespecified number of coefficients change; moreover, if the coefficient of variation affects a bigger number of factors, those authors provide a probabilistic guarantee that the solution will be feasible with high probability. The method has been applied by Bertsimas and Thiele (2004, 2006).

In a recent paper, Bertsimas et al. (2010) propose an approach to solve robust optimization problems in which the objective function is not explicitly available, but is derived from simulation models. They implement an iterative local search method, moving along descent directions of the worst-case cost function. The first step of the proposed algorithm consists of exploring a (properly defined) neighborhood of the current point; then, a descent direction can be found by solving a SOCP problem. The robust local search is designed to terminate at a robust local minimum, which is a point where no improving directions are available for the algorithm.

2.2 Taguchi's Approach

In the late 1970s, Genichi Taguchi, a Japanese textile engineer, introduced new ideas on quality improvement, resulting in an innovative parameter design approach for reducing variation in products and processes (see Taguchi, 1987). His methodology has been successfully applied in many important industries in the USA, such as Ford Motor Company and Xerox.

Taguchi identifies three stages in the design process:

- *System Design* is a general approach to design a process that includes defining the system's objectives and goals.
- *Parameter Design* involves defining responses of interest and optimizing them w.r.t. their mean and variation.
- *Tolerance Design* corresponds to fine-tuning the variables that have been optimized in the previous stage by controlling the factors that affect them.

Notice that the last two stages may appear quite similar to each other, so it may be difficult to keep them distinct; see Beyer and Sendhoff (2007). In fact, from a mathematical point of view, parameter and tolerance design differ only in the *granularity* by which design parameters are treated. On the other hand, from a practical point of view, it is important to distinguish between the two phases, because they can occur under very different constraints, e.g. design time versus operation time. Taguchi—focusing on Parameter Design—distinguishes between two different types of factors when designing a product or process:

- *control* or *decision* factors (which we denote by $d_j, j = 1, \ldots, n_d$) are under the control of the users; e.g., in inventory management, the order quantity may be controllable.
- *noise* or *environmental* factors (denoted by $e_k, k = 1, \ldots, n_e$) cannot be controlled while the process operates or the product is used; e.g. the demand rate in inventory problems.

Notice that, in practice, the controllability of a factor depends on the specific situation; e.g., in production or inventory management the decision makers may affect the demand rate through an advertising campaign.

Other authors distinguish between environmental uncertainty (e.g., demand uncertainty) and system uncertainty (e.g., yield uncertainty); see Mula et al. (2006) and also Beyer and Sendhoff (2007). Implementation errors may also be a source of uncertainty. These errors occur whenever recommended (optimal) values of control factors have to be realized in practice; see Stinstra and den Hertog (2008). Continuous values are hard to realize in practice, because only limited accuracy is then possible; e.g., the optimal solution in the Economic Order Quantity model (EOQ) turns out to be the square root of some expression, but in practice only a discrete number of units could be ordered. Besides implementation errors, there are validation errors of the simulation model—compared with the real system—and the metamodel—compared with the simulation model; see Kleijnen and Sargent (2000).

Taguchi's basic idea in dealing with robust parameter design is to take account of the environmental (noise) factors in the experimental design and find the most insensitive, or robust, system configuration in the decision (controllable) factors with respect to the noise factors variation.

As far as the experimental strategy is concerned, Taguchi adopted crossed arrays, resulting from the product of two experimental designs; one design varies the decision factors $\mathbf{d}$ (obtaining the "inner array", as Taguchi calls it), and one design varies the environmental factors $\mathbf{e}$ (thus obtaining the "outer array"). Combining them together is equivalent to consider variations in the uncontrollable (environmental) factors at different locations in the space of the controllable factors. Taguchi further distinguishes between factors that have a *location* effect, changing the mean of the response or objective function, and factors that have a *dispersion* effect, since they affect the variance of the process. Therefore, in the optimization process, Taguchi takes into account the first two moments of the distribution of the objective function, and combines them using the signal-to-noise ratio (SNR). Taguchi suggests to consider three types of problems:

1. "Smaller the better": select the factor combination in the inner array that maximizes

$$SNR_S = -10 \log \sum_{i=1}^{n_0} \frac{w_i^2}{n_0} \tag{1}$$

   where $w_i = w(\mathbf{d}, e_i)$ and n_0 is the number of runs in the outer array.

2. "Larger the better": select the inner array point that maximizes

$$SNR_L = -10 \log \frac{1}{n_0} \sum_{i=1}^{n_0} \frac{1}{w_i^2} \tag{2}$$

3. "Target is best". Here Taguchi proposes a two-step approach, suggesting two cases:

   (a) μ_w (mean of w) is not related to σ_w (standard deviation of w). In this case the steps are the following:

      i - Select some control factors that maximize

      $$SNR_{T1} = -10 \log s^2 \tag{3}$$

      where s^2 is the sample variance of the outer array observations.

      ii - Select some other inner array factor (not varied before) to make

      $$\overline{w} \approx \mathcal{T} \tag{4}$$

      where $\overline{w}$ is the average of the outer array observations and $\mathcal{T}$ is the target of the quality characteristic.

   (b) If σ_w is proportional to μ_w—a case likely to occur in practice—then

      i - Select some control factors to maximize

      $$SNR_{T2} = -10 \log \frac{\overline{w}^2}{s^2} \tag{5}$$

      ii - Select some other control factors not varied before, to make

      $$\overline{w} \approx \mathcal{T}.$$

      Because the standard deviation is assumed to be proportional to the mean, the controllable factors will change the mean but will not change the ratio $\overline{w}^2/s^2$ much.

Some aspects of the Taguchian approach have been severely criticized; see Myers et al. (1992), Myers et al. (2009) and Del Castillo (2007). The mostly debated issues were the following:

- A data set with no outer array variability and one with considerable outer array variability may result in the same SNR; therefore, SNR would be ineffective in Robust Parameter Design.

- No attention is paid to the computational costs required by the experimental design: in fact, using a crossed array design often requires a large number of runs, which can be prohibitive in some industrial processes.

- The method does not enable flexible modeling of the design variables, not taking into account the interactions either among decision factors or between decision and environmental factors. Standard ANOVA techniques can be used to identify the control factors that impact SNR (see Myers et al., 2009; Robinson et al., 2004).

- Factors may have both location and dispersion effects, so the proposed two-step approach may be inadequate in practice. Moreover, the adoption of the SNR as performance characteristic appears to be too restrictive (Park et al., 2006) and may confound the mean and variance contributions; keeping them separately, instead, can provide further insight into the process behavior.

Although based on Taguchi's view of modeling uncertainty in a design process, some authors (Trosset, 1997) have suggested to directly model the response as a function of both decision and environmental factors, instead of using SNRs. Suppose we measure q performance indicators, $w_1, \ldots, w_q$; let $w_i(\mathbf{d}, \mathbf{e})$ denote the value of the i-th performance indicator when control and noise factors assume values $(\mathbf{d}, \mathbf{e})$ and let $l[w_1(\mathbf{d}, \mathbf{e}), \ldots, w_q(\mathbf{d}, \mathbf{e})]$ denote the corresponding loss. A robust design approach will seek a combination of control factors that minimizes the expected loss, computed with respect to the random vector $\mathbf{e}$. If the distribution of $\mathbf{e}$ does not depend on $\mathbf{d}$, then the objective function is

$$L(\mathbf{d}) = \int l[w_1(\mathbf{d}, \mathbf{e}), \ldots, w_q(\mathbf{d}, \mathbf{e})]\, p(\mathbf{e})\, d\mathbf{e} \tag{6}$$

where $p(\mathbf{e})$ denotes the probability density function of $\mathbf{e}$. The question arises: how have statisticians sought to minimize (6)? A numerical optimizer would answer this question in the following manner:

1. A design is chosen that specifies the (d_j, e_j) at which the w_i have to be evaluated; this approach results in a single "combined" array, instead of inner and outer arrays.

2. The $w_i(d_j, e_j)$ are used to estimate cheap-to-compute surrogate models $\widehat{y}_i$.

3. Optimization is carried out using the surrogate objective function

$$\hat{L}(\mathbf{d}) = \int l[\hat{y}_1(\mathbf{d}, \mathbf{e}), \ldots, \hat{y}_q(\mathbf{d}, \mathbf{e})]\, p(\mathbf{e})\, d\mathbf{e}. \tag{7}$$

A similar approach is suggested by Sanchez (2000), who proposes a robust methodology, starting from Taguchi's approach and combining it with metamodeling techniques. Focusing

on discrete-event simulation models, she identifies some performance characteristic, denoted by $w(\mathbf{d})$, $\mathbf{d}$ being the vector of decision factors, and an associated target value $\mathcal{T}$. The goal would be to select the decision factors to keep the objective function on target, with zero variance. However, this would be an ideal situation—hard to realize in practice. Therefore, to find a trade-off between performance mean and variability, Sanchez proposes to use a quadratic loss function, defined as follows: assuming that no loss occurs when $w(\mathbf{d})$ achieves the target $\mathcal{T}$, the quadratic loss function can be written as

$$l(w(\mathbf{d})) = c[w(\mathbf{d}) - \mathcal{T}]^2 \tag{8}$$

where c is a scaling factor, accounting for possible units conversions. It follows from (8) that the expected loss associated with configuration $\mathbf{d}$ is

$$E[l(w(\mathbf{d}))] = c[\sigma^2 + (\mu - \mathcal{T})^2] \tag{9}$$

where μ and σ^2 denote the true mean and variance of the output function w.

As far as the robust design is concerned, Sanchez tries to characterize the system behavior as a function of the control factors only. First, an appropriate experimental design is planned, for both decision and environmental factors. Then, for every combination of decision factor configuration i and environmental factor configuration j, the sample average $\overline{w}_{ij}$ and sample variance s_{ij}^2 are computed—after suitable truncation to remove initialization bias. Finally, summary measures across the environmental space for each decision factor configuration i are computed:

$$\overline{w}_{i\cdot} = \frac{1}{n_e} \sum_{j=1}^{n_e} \overline{w}_{ij} \tag{10}$$

$$\overline{V}_{i\cdot} = \frac{1}{n_e - 1} \sum_{j=1}^{n_e} (\overline{w}_{ij} - \overline{w}_{i\cdot})^2 + \frac{1}{n_e} \sum_{j=1}^{n_e} s_{ij}^2 \tag{11}$$

where n_e is the number of combinations in the environmental design.

Two initial metamodels are then built, using regression polynomials: one for the performance mean, and one for the performance variability; for discrete-event simulation experiments, Sanchez recommends a design which allows for fitting at least a quadratic effect. Robust configurations are identified by combining information resulting from the mean and variance metamodels, using (9) where the true mean and variance are replaced by the estimate given in (10) and (11). If the configurations suggested by the robust design were not

among those initially tested, further experimentation could be needed: in this case, however, computational time could be saved, by screening the decision factors involved in the experiment.

Al-Aomar (2002) presents an iterative scheme to solve simulation-based optimization problems. His work considers a discrete-event simulation model; the (controllable) design parameters are $d_1, \ldots, d_n$, and its performances are evaluated through the metrics $w_1, \ldots, w_q$. Then, an overall utility function U is defined combining multiple performance measures into a single function. The general formulation of the system design problem can be defined as follows:

$$\begin{aligned} &\max\ U(w_1, \ldots, w_q) \\ &s.t.\ w_i = f_i(d_1, \ldots, d_n), \quad 1 \leq i \leq q \\ &d_j \in S, \quad 1 \leq j \leq n \end{aligned} \tag{12}$$

where S is the feasible space for the control variable $\mathbf{d}$. His methodology consists of four modules: i) the Simulation Modeling (SM) module uses a discrete-event simulation model to evaluate the set of performance metrics w_i associated with each solution alternative $\mathbf{d}$, in terms of means and variances; ii) the Robustness Module (RM) transforms the mean and variance of each performance measure into a Signal-to-Noise Ratio—thus adopting a Taguchian approach; iii) the Entropy Method (EM) module builds the utility function U by linearly combining the performance criteria, through a proper choice of the weights, dynamically updated at each iteration; iv) the Genetic Algorithm (GA) module is utilized as a global optimizer, working on a set of possible solutions that are selected basing on the overall utility function value at each point. A convergence test at the end of each step controls whether any stopping criterion is met (maximum number of generations reached or convergence rate achieved). For a detailed discussion we refer to El-Haik and Al-Aomar (2006).

Because of the criticism on SNRs, some authors like Myers et al. (2009) suggest to build separate models for the mean and variance of the system performance, adopting the so-called Dual Response Surface approach. This methodology has some advantages:

- It provides an estimate of the mean and standard deviation at any location in the space of control design variables.

- Some insight can be gained regarding the roles of these variables in controlling process mean and variance.

- It could be easily integrated into process optimization based on a squared error loss criterion, $\widehat{E}_{\mathbf{e}}(w-\mathcal{T})^2 = [\widehat{E}_{\mathbf{e}}(w)-\mathcal{T}]^2 + \widehat{\sigma}^2_{\mathbf{e}}(w)w$ or the maximization of an estimated quantile $\hat{E}_{\mathbf{e}}(w) - 2\hat{\sigma}_{\mathbf{e}}(w)$ in the Taguchian "larger the better" case, or the minimization of $\widehat{E}_{\mathbf{e}}(w) + 2\widehat{\sigma}_{\mathbf{e}}(w)$ in the Taguchian "smaller the better" case.

- It allows the use of constrained optimization; that is, choosing a target value of $\widehat{\mu}_{\mathbf{e}}[w(\mathbf{d},\mathbf{e})]$ or—better to say—a threshold T below which one cannot accept the solution. Therefore, the following problem has to be solved:

$$\min_{\mathbf{d}} \widehat{\sigma}^2_{\mathbf{e}}[w(\mathbf{d},\mathbf{e})] \quad s.t. \quad \widehat{\mu}_{\mathbf{e}}[w(\mathbf{d},\mathbf{e})] \leq T \tag{13}$$

  Several values of T may be used to consider different user's alternatives.

The Dual Response Surface approach has been successfully applied to robust process optimization; see Del Castillo (2007). Quite often, in fact, the purpose is to reach a desired performance for the process that manufactures some products—e.g. by minimizing the cost of operation in a production process, or the variability of a quality characteristic, or by maximizing the throughput of the manufacturing process. Evidently, multiple—and sometimes conflicting—responses are usually considered in practical problems. However, due to noisy data and/or to uncertainty affecting some parameters of the model, achieving robust performances is of interest.

Miró and del Castillo (2004) point out that the classical Dual Response Surface approach takes into account only the uncertainty due to the noise factors; they identify an additional component due to the uncertainty in the parameter estimates. Therefore, they propose an extension of the Dual Response Surface approach, introducing the additional variance of the parameters estimates into an objective function that combines it with the noise factor variance. Optimizing such a function will achieve a process that is robust with respect to both noise factor variation and uncertainty in the parameter estimates. One such function is the variance of the predicted response, where the variance is now taken with respect to both the parameter estimates of the model and the noise factors.

Robustness is also a central issue in design optimization. Many engineering applications have to deal with the uncertainty which affects the components of the system under design;

ignoring the source of uncertainty and assuming some parameters to be *exactly* known and constant might cause the designed system not to be adequate whenever the environmental setting changes.

Bates et al. (2006) compare different methods to perform robust optimization, applying them to solve robust design optimization of a mechanical component: the objective is to achieve a given mean cycle time while minimizing the standard deviation of the cycle time. The authors discuss the following issues: the Taguchian approach, using a crossed-array design and maximizing the SNR; the Response Model analysis, involving both decision and environmental factors, and accounting for factor interactions; and the Dual Response Surface approach. They propose a framework called Stochastic Emulator Strategy, consisting of the following building blocks: (i) DoE, using an array that includes both design and noise factors; they prefer space-filling designs (such as LHS) or lattice designs rather than orthogonal arrays or fractional factorials, to achieve more uniform coverage of the input space. (ii) Metamodel (or emulator, as they called it) fitting, to represent the relationship among all factors—disregarding whether they are decision or environmental factors—and the chosen response. (iii) Metamodel prediction, to estimate the mechanical component cycle time for a given set of factor values and evaluate the effect of noise on the output by studying how it behaves when subjected to small changes in factor values. (iv) Optimization process, minimizing the output variance with a target value for the mean cycle time.

Lee and Park (2006) present a methodology—based on Kriging metamodels—to tackle robust optimization in deterministic simulation-based systems. They use simulated annealing to solve the optimization problem. The approach is basically the one proposed by Taguchi, employing mean and variance as statistics to study the insensitivity of the response to possible variations in the noise factors.

The use of Kriging as an approximation technique is justified because Kriging provides reliable approximation models of highly nonlinear functions, and this feature is even more useful in robust optimization than it is in classical optimization because in general the nonlinearity of the response variance could be higher than that of the mean. Kriging is also recommended by Jin et al. (2003), who compare some metamodeling techniques and—based on the results of some tests performed on both mathematical functions and a more complex case study—they conclude that Kriging models provide higher accuracy than the other alternatives. However, Allen et al. (2003) notice that regression modeling should not be quickly discarded for cases in which the number of runs is particularly low.

Lee and Park aim at determining a design point $\mathbf{d}$ providing a target response value $\overline{\mu}_w$ with the smallest variation σ_w^2. Therefore they formulate the robust optimization problems as follows:

$$\min \quad \sigma_w^2 \tag{14}$$

$$s.t. \quad \mu_w \leq \overline{\mu}_w \tag{15}$$

Because the analytical computation of the mean and variance of a given response w is not always possible (too time expensive or too difficult), Lee and Park approximate these two statistics by means of the first-order Taylor expansion:

$$\mu_w \approx w(\mathbf{d}, \mathbf{e})_{\bar{\mathbf{d}},\bar{\mathbf{e}}} \tag{16}$$

$$\sigma_w \approx \sum_{i=1}^{n_d} \left(\frac{\partial w}{\partial d_i}\right)_{\bar{\mathbf{d}}}^2 \sigma_{d_i}^2 + \sum_{j=1}^{n_e} \left(\frac{\partial w}{\partial e_j}\right)_{\bar{\mathbf{e}}}^2 \sigma_{e_j}^2 \tag{17}$$

where $\bar{\mathbf{d}}$ and $\bar{\mathbf{e}}$ denote the mean vectors of the control and noise factors, and $\sigma_{d_i}^2$ and $\sigma_{e_j}^2$ represent the variance of the i-th control variable and the j-th noise variable. However, (16) and (17) are valid approximations only for monotonic functions, which is usually a property difficult to ascertain when working with black-box simulation models.

The following issues emerge from Lee and Park (2006):

- The authors fit one *single* metamodel over the control-by-noise factors space. They suggest that this metamodel be highly accurate, because it will be used to derive the approximation model for the variance.

- To derive a model for the mean of the response, they use the approximation provided by (16), applying it to the metamodel computed.

- To derive a model of the variance, they use Monte-Carlo simulations performed not on the simulation model but on the (inexpensive) metamodel obtained in the beginning.

- They point out that post-processing may be necessary because of the nonlinearity both of the mean response and (even more) of its variance, and the approximation errors coming from fitting the metamodel of the variance based on the metamodel of the mean response function. The post-processing consists in solving the following optimization

problem, restricting the search area to the neighborhood of the optimal solution found so far:

$$\min \quad \widehat{\sigma}_w^2 = \sum_{i=1}^{n_d} \left(\frac{\partial \hat{y}}{\partial d_i} \right)^2 \sigma_{d_i}^2 + \sum_{j=1}^{n_e} \left(\frac{\partial \hat{y}}{\partial e_j} \right)^2 \sigma_{e_j}^2 \,, \tag{18}$$

$$s.t. \quad \hat{y}(\mathbf{d}, \mathbf{e})_{\bar{\mathbf{d}}, \bar{\mathbf{e}}} \leq \bar{\mu}_w \tag{19}$$

This post-processing would aim to further refine the robust optimum, but experimental results usually show rather small improvements.

- As a further research topic, they suggest to adopt two *distinct* metamodels, approximating the true response and the true variance. They suggest to adopt this approach for strongly nonlinear models, especially the variance model.

3. Robust Optimization and Kriging metamodels

The purpose of the section is to integrate the description of Kriging within our framework for robust simulation-optimization discussed in the main paper. We still interpret the simulated system from the Taguchian viewpoint, but use Kriging as a metamodeling technique.

In this section we describe the characteristics of the Kriging metamodeling technique in more general terms than in the main paper, adopting a notation which goes beyond the one characterizing our robust methodology. We also refer to Sacks et al. (1989) and Santner et al. (2003) for a detailed exposition of both theory and implementation of Kriging.

A Kriging model is

$$y(\mathbf{x}) = f(\mathbf{x}) + Z(\mathbf{x}) \,, \tag{20}$$

where $f(\mathbf{x})$ is a function of the n-dimensional $\mathbf{x}$ that is a global model of the original function, and $Z(\mathbf{x})$ is a stochastic process with zero mean and non-zero variance that represents a local deviation from the global model. Usually, $f(\mathbf{x})$ is

$$f(\mathbf{x}) = \sum_{i=0}^{p} \beta_i \, f_i(\mathbf{x}) \tag{21}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, p$, are polynomial terms (typically of zero, first or second order). The coefficients β_i, $i = 1, \ldots, p$, are regression parameters. The $p + 1$ regression functions can be regarded as components of a vector

$$\mathbf{f}(\mathbf{x}) = [f_0(\mathbf{x}), \ldots, f_p(\mathbf{x})]^T \,. \tag{22}$$

Suppose the design sites are $(\mathbf{x}_1, \ldots, \mathbf{x}_{N_s})$, where $\mathbf{x}_i \in \mathbb{R}^n$, $i = 1, \ldots, N_s$. Then we can compute the matrix $\mathbf{F}$ by evaluating the vector $\mathbf{f}(\mathbf{x})$ at the design sites, thus obtaining:

$$\mathbf{F} = \begin{bmatrix} f^T(\mathbf{x}_1) \\ \vdots \\ f^T(\mathbf{x}_{N_s}) \end{bmatrix} = \begin{bmatrix} f_0(\mathbf{x}_1), & \ldots, & f_p(\mathbf{x}_{N_s}) \\ \vdots & & \vdots \\ f_0(\mathbf{x}_1), & \ldots, & f_p(\mathbf{x}_{N_s}) \end{bmatrix}. \tag{23}$$

The covariance of $Z(\mathbf{x})$ is:

$$Cov[Z(\mathbf{x}_j), Z(\mathbf{x}_k)] = \sigma^2 \mathbf{R}(\mathbf{x}_j, \mathbf{x}_k), \qquad j, k = 1, \ldots, N_s \tag{24}$$

where σ^2 is the so-called process variance and $\mathbf{R}$ is the correlation matrix with elements $R_{jk} = R_\theta(\mathbf{x}_j, \mathbf{x}_k)$, representing the correlation function between any two of the N_s samples $\mathbf{x}_j$ and $\mathbf{x}_k$, with unknown parameters θ. $\mathbf{R}$ is a symmetric matrix of dimension $N_s \times N_s$, with diagonal elements equal to 1. The form of the correlation function $R_\theta(\mathbf{x}_j, \mathbf{x}_k)$ can be chosen among a variety of functions proposed in the literature. Yet, the exponential family is used most frequently

$$R_{\theta,p}(\mathbf{x}_j, \mathbf{x}_k) = \prod_{i=1}^{n} exp\left(-\theta_i |x_{ji} - x_{ki}|^{p_i}\right) \tag{25}$$

where n is the dimension of the input variable. When $p_i = 2$, then (25) is called the Gaussian correlation function. The parameters p_i determine the smoothness of the correlation function; e.g., $p_i = 2$ implies an infinitely differentiable function.

The Kriging predictor can be written as a linear combination of the observed responses:

$$y(\mathbf{x}) = \mathbf{c}^T(\mathbf{x}) \mathbf{y}_s \,, \tag{26}$$

where $\mathbf{y}_s$ is the vector of the response function evaluated at the N_s design sites, $\mathbf{y}_s = [y(\mathbf{x}_1), \ldots, y(\mathbf{x}_{N_s})]^T$. The weights $\mathbf{c}(\mathbf{x})$ are obtained by minimizing the Mean Squared Error (MSE), which is given by

$$MSE[y(\mathbf{x})] = E\left[\left(\mathbf{c}^T(\mathbf{x})\mathbf{y}_s - y(\mathbf{x})\right)^2\right]. \tag{27}$$

In order to keep the predictor unbiased, the following constraint has to be satisfied:

$$\mathbf{F}^T \mathbf{c}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \,. \tag{28}$$

It can be proven that the MSE in (27) can be rewritten as

$$MSE[y(\mathbf{x})] = \sigma^2 [1 + \mathbf{c}^T(\mathbf{x}) \mathbf{R} \mathbf{c}(\mathbf{x}) - 2\mathbf{c}^T(\mathbf{x}) \mathbf{r}(\mathbf{x})] \tag{29}$$

where $\mathbf{r}(\mathbf{x}) = [R(\mathbf{x}_1, \mathbf{x}), \ldots, R(\mathbf{x}_{N_s}, \mathbf{x})]^T$ is the vector of the correlations between $Z(\mathbf{x}_i)$ and $Z(\mathbf{x})$. Minimizing the MSE in (29) with respect to $\mathbf{c}(\mathbf{x})$ under the constraint (28) gives the Kriging predictor

$$\widehat{y(\mathbf{x})} = \widehat{\mathbf{c}}^T(\mathbf{x})\mathbf{y}_s = \widehat{\mathbf{r}}^T\widehat{\mathbf{R}}^{-1}(\mathbf{y}_s - \mathbf{F}\widehat{\boldsymbol{\beta}}) + \mathbf{f}^T\widehat{\boldsymbol{\beta}}\,, \tag{30}$$

where

$$\widehat{\boldsymbol{\beta}} = (\mathbf{F}^T\widehat{\mathbf{R}}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\widehat{\mathbf{R}}^{-1}\mathbf{y}_s \tag{31}$$

follows from the Generalized Least-Squares (GLS) criterion or the Maximum Likelihood Estimation (MLE) criterion.

Assuming the stochastic process $Z(x)$ to be Gaussian, MLE maximizes a likelihood function using numerical optimization techniques to determine an estimate $\widehat{\theta}$ (Sacks et al., 1989). The likelihood function depends on the coefficients $\boldsymbol{\beta}$ in the regression model, the process variance σ^2, and the correlation parameters $\boldsymbol{\theta}$. Given the correlation parameters θ and hence R, the MLE of $\boldsymbol{\beta}$ is given by (31), and the MLE of σ^2 is given by

$$\widehat{\sigma}^2 = \frac{1}{N_s}(\mathbf{y}_s - \mathbf{F}\widehat{\boldsymbol{\beta}})^T\mathbf{R}^{-1}(\mathbf{y}_s - \mathbf{F}\widehat{\boldsymbol{\beta}})\,. \tag{32}$$

Therefore, the MLE of R follows from

$$\min_{\boldsymbol{\theta}} (\det\mathbf{R})^{1/N_s}\widehat{\sigma}^2\,; \tag{33}$$

this is a global nonlinear optimization problem, which requires a heuristic procedure for its solution.

In our study, we estimate Kriging models using algorithms and functions of the MATLAB DACE Toolbox (see Lophaven et al., 2002), which includes some heuristics to solve the problem in (33), and offers some support for experimental designing.

In our own experiments we use so-called *ordinary Kriging*

$$y(\mathbf{x}) = \mu + Z(\mathbf{x}) \tag{34}$$

where $\mu = \beta_0$ and $f_0 \equiv 1$.

4. Examples

A set of inventory problems based on the Economic Order Quantity (EOQ) model provides a first group of examples we use to illustrate and test our methodology (as documented in the

main paper). In spite of its simplicity—and maybe even thanks to that simplicity—it helps to describe each step of the heuristic procedure, and the computational results are easy to understand. The main characteristics of these models are described in the main paper. In subsection 4.1 we detail the computational results for the robust EOQ model with uncertain demand rate and cost coefficients. Subsection 4.2 introduces another example taken from Bertsimas et al. (2010).

4.1 Robust optimization with uncertain demand rate and cost coefficients

Inspired by Borgonovo and Peccati (2007), we extend our robust formulation of the EOQ model such that it accounts for fixed but uncertain cost parameters. So the number of environmental factors increases from one (demand rate) to three (holding and set-up costs, besides demand rate). We assume that all the environmental factors follow normal distributions, with mean equal to the nominal value (the value when no uncertainty is assumed; namely, $\mu_h = 0.3$ for the holding cost, $\mu_K = 12000$ for the set-up cost, and $\mu_a = 8000$ for the demand rate); the standard deviation of each factor is equal to 10% of the nominal value; we resample negative values; we assume that the three factors are independent.

4.1.1 1L-KM approach

We again adopt a crossed design, combining a uniform space filling design of size $n_Q = 10$ for the decision factor Q and a LHS design of size $n_{\mathbf{e}} = 120$ for the three-dimensional space of the environmental factors a, K, and h.

Next we average over the environmental factors to derive a set of n_Q output values for both the mean and the standard deviation, using (18) and (19) in the main paper with n_a replaced by $n_{\mathbf{e}}$. Based on these I/O data, we fit one Kriging metamodel for each of the two outputs; see the solid curve in Figures 1 and 2, which also displays the true cost function obtained from

$$E(C) = \left(\frac{\mu_K}{Q} + c\right)\mu_a + \frac{\mu_h Q}{2}\,, \tag{35}$$

and the true standard deviation computed through

$$\sigma_C = \sqrt{\sigma_a^2\left(c^2 + \frac{2c\mu_K}{Q} + \frac{\sigma_K^2 + \mu_K^2}{Q^2}\right) + \frac{\mu_a^2\sigma_K^2}{Q^2} + \frac{\sigma_h^2 Q^2}{4}}\,. \tag{36}$$

We use leave-one-out cross-validation to validate our two metamodels. Based on the relative prediction errors which are smaller than 1%, we decide to accept both metamodels.

We let the threshold T vary in the interval $[8200, 8600]$ (the same as for the EOQ model with uncertain demand rate only), which gives the estimated Pareto frontier in Figure 3.

4.1.2 2L-KM approach

Following the 2L-KM approach, we first build a LHS design for the four input factors, without distinguishing between decision and environmental factors; we choose a design of size $n = 1200$ (we keep the same sample size as in the 1L-KM approach).

We run the simulation model over these n input combinations, and collect the corresponding output values C_i $(i = 1, \ldots, n)$. Next, we fit a Kriging metamodel based on these n I/O combinations. We produce a bigger DOE by crossing a uniform space filling design for the order quantity Q and a LHS design for the three environmental factors, accounting for their distribution; the overall design size is $N = N_Q \times N_{\mathbf{e}} = 30 \times 200 = 6000$. Notice that the bigger value for N does not imply a computationally expensive task, because we use this bigger design to compute Kriging predictions instead of running the simulation model.

Next we compute the sample average of cost predictions through (22) in the main paper and the sample standard deviation of cost through (23) in the main paper, and derive two Kriging metamodels; namely, one for the expected cost and one for the cost standard deviation; see again Figures 1 and 2. We validate these metamodels through leave-one-out cross-validation. As noticed in the EOQ example with uncertain demand rate, both metamodels give small relative prediction errors; namely, around 10^{-6}.

Finally, we take 100 equally spaced values of the threshold T in the interval $[8200, 8600]$, and collect the corresponding optimal solutions to estimate the Pareto frontier; see again the dashed curve in Figure 3.

4.1.3 Bootstrapped confidence regions

To analyze the variability in the optimal solutions of the Pareto frontier, we apply bootstrapping, as discussed in Section 4.2 in the main paper. Therefore, we derive the bootstrapped output data and use them to fit B pairs of bootstrapped Kriging metamodels. We use these bootstrapped metamodels to derive the rectangular *confidence regions* for two points on the original estimated Pareto curve. Part (a) of Figure 4 corresponds with the relatively small threshold value $T = 8250$ so $\widehat{Q^+} = 25531.91$ for the 1L-KM approach and $\widehat{Q^+} = 29943.98$ for

the 2L-KM approach; part (b) corresponds with the larger threshold value $T = 8600$ so we obtain the smaller $\widehat{Q^+} = 25291.39$ for 1L-KM and $\widehat{Q^+} = 24001.84$ for 2L-KM. Both Kriging approaches give confidence regions that cover the true point, although the confidence region associated with 2L-KM is smaller—which confirms what we have already observed in Section 5.2 in the main paper.

Again, we notice that there is still a probability of violating the threshold on the standard deviation for some elements of the Pareto frontier. Therefore, the manager may prefer to implement a solution providing a relatively small probability of becoming infeasible (estimated through the confidence region), accepting higher costs.

4.2 Bertsimas et al.'s example

Besides uncertainties in coefficients or parameters, uncertainties can result from implementation errors; i.e. models cannot be implemented to infinite precision (Stinstra and den Hertog (2008)). Therefore, in the optimization process, the decision maker could account for both these aspects. In this section, we consider an example proposed by Bertsimas et al. (2010) including both sources of uncertainty.

Let $f(\mathbf{x}, \widehat{\mathbf{p}})$ be the planned cost of design vector $\mathbf{x}$ where $\widehat{\mathbf{p}}$ is an estimator of the true problem coefficient $\mathbf{p}$. Because $\widehat{\mathbf{p}}$ is an estimate, the true coefficient $\mathbf{p}$ may turn out to be $\widehat{\mathbf{p}} + \mathbf{\Delta p}$ where $\mathbf{\Delta p}$ is the additive parameter uncertainty. Often, the *design* optimization problem

$$\min_{x} f(\mathbf{x}, \widehat{\mathbf{p}}) \tag{37}$$

is solved ignoring the presence of uncertainties.

Bertsimas et al. consider problem (37) where both $\mathbf{\Delta p} \in \mathbb{R}^m$ and implementation errors $\mathbf{x} \in \mathbb{R}^n$ are present, while further assuming that $\mathbf{\Delta z} = \begin{pmatrix} \mathbf{\Delta x} \\ \mathbf{\Delta p} \end{pmatrix}$ lies within the uncertainty set

$$\mathcal{U} = \{\mathbf{\Delta z} \in \mathbb{R}^{n+m} \mid \|\mathbf{\Delta z}\|_2 \leq \Gamma\} \tag{38}$$

where $\Gamma > 0$ is a scalar describing the size of the errors or perturbations. Those authors seek a design $\mathbf{x}$ that minimizes the worst-case cost given a perturbation in $\mathcal{U}$ whereas, according to Section 2 in our main paper, we seek a design that minimizes expected (mean) cost under a constraint on the variability of this cost.

The Bertsimas et al.' example consists of the following polynomial function:

$$\begin{aligned} f(x_1, x_2) = 2x_1^6 - 12.2x_1^5 + 21.2x_1^4 + 6.2x_1 - 6.4x_1^3 - 4.7x_1^2 + x_2^6 - 11x_2^5 \\ +43.3x_2^4 - 10x_2 - 74.8x_2^3 + 56.9x_2^2 - 4.1x_1x_2 - 0.1x_2^2x_1^2 + 0.4x_2^2x_1 + 0.4x_1^2x_2 = \\ = \sum_{r>0,s>0}^{r+s\leq 6} c_{rs}x_1^r x_2^s \,. \end{aligned} \tag{39}$$

The case study considers the uncertainty in each of the 16 coefficients of the objective function; i.e. the objective function is

$$\tilde{f}(x_1, x_2) = \sum_{r>0,s>0}^{r+s\leq 6} c_{rs}(1 + 0.05\Delta p_{rs})x_1^r x_2^s \,. \tag{40}$$

where $\mathbf{\Delta p}$ is the vector of uncertain parameters; the robust optimization problem considers also the implementation errors and has to minimize the following objective function

$$\tilde{f}(x_1 + \Delta x_1, x_2 + \Delta x_2) \,. \tag{41}$$

The vector $\mathbf{\Delta z} = \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \mathbf{\Delta p} \end{pmatrix}$ is used to define the following uncertainty set

$$\mathcal{U}_f = \{\mathbf{\Delta z} \in \mathbb{R}^{18} \,|\, \|\mathbf{\Delta z}\|_2 \leq 0.5\} \tag{42}$$

The problem has two decision variables, x_1 and x_2, and 18 environmental factors; namely, 2 implementation errors, Δx_1 and Δx_2, and 16 coefficients uncertainties $\mathbf{\Delta p}$. Therefore, the uncertainty set $\mathcal{U}_f$ defined in (42) is a hypersphere of dimension 18.

Whereas Bertsimas et al. focus on worst-case robust optimization, we account for the probability of each specific realization of the vector $\mathbf{\Delta z}$ and apply our robust optimization approach. In order to conduct a computational experiment giving results comparable with those obtained by Bertsimas et al., we consider the same region for uncertainties, and choose a uniform distribution for them.

As we discussed in the main paper, we apply 1L-KM and 2L-KM to the Bertsimas et al.'s example. The 1L-KM approach crosses a space-filling design for the decision factors and a LHS design for the environmental factors; more specifically, we take $n_{\mathbf{d}} = 100$ equally spaced points in the rectangle $[0, 5] \times [0, 8]$—which is the experimental area Bertsimas et al. choose—, and we sample $n_{\mathbf{e}} = 100$ points from a uniform distribution in the hypercube

$[-0.5, 0.5]^{18}$. The 2L-KM approach starts from a uniform LHS design of size $n = 1600$ in the entire space of decision-by-environmental factors, to fit the metamodel of the output f. In the second stage we use a crossed design obtained through a grid of $N_{\mathbf{d}} = 100$ points for the decision factors and a uniform LHS design of $N_{\mathbf{e}} = 100$ points for the environmental factors.

After deriving the metamodels for the mean and the standard deviation of the output through 1L-KM and 2L-KM, we solve the optimization problem for different values of T in the interval $[3, 20]$; this provides an estimated Pareto frontier. Then, we apply the bootstrap procedure to derive a confidence region for the optimal solution resulting from the original estimated Pareto frontier. Figure 5 shows the results obtained for an intermediate threshold value; namely, $T = 8.58$. Like the EOQ examples in the main paper, 2L-KM provides a smaller confidence region; both approaches ensure confidence regions which cover the original (i.e., non-bootstrapped) optimal solution and guarantee the optimal solutions to remain feasible despite the metamodel variability.

Bertsimas et al. also provide the global optimum for the nominal problem (when there is no uncertainty on the parameters) $\mathbf{x}_{opt} = (2.8, 4)$, with corresponding output $f(\mathbf{x}_{opt}) = -20.8$. Computing the prediction of the expected output and its standard deviation on the Kriging metamodels in our approaches gives the following results: $\widehat{\overline{y}}_{1L-KM}(\mathbf{x}_{opt}) = -2.225$, $\widehat{s}_{1L-KM}(\mathbf{x}_{opt}) = 246.3$; and $\widehat{\overline{y}}_{2L-KM}(\mathbf{x}_{opt}) = 7.436$, $\widehat{s}_{2L-KM}(\mathbf{x}_{opt}) = 111.97$. Notice that Bertsimas et al. associate a worst-case cost of 450 to the nominal solution $\mathbf{x}_{opt}$, which appears compatible with the high standard deviation we obtained from the metamodel approximations.

References

Al-Aomar, R. 2002. A Robust Simulation-based Multicriteria Optimization Methodology, *Proceedings of the Winter Simulation Conference*, Yücesan, E. and Snowdon, C.-H. and Charnes, J. M., Eds. 1931–1939.

Allen, T.T., M.A. Bernshteyn, and K. Kabiri-Bamoradian. 2003. Constructing Meta-Models for Computer Experiments, *J. of Quality Tech.* **35** 264–274.

Bates, R. A., R. S. Kenett, D. M. Steinberg, and H. P. Wynn. 2006. Achieving Robust Design from Computer Simulations. *Quality Tech. & Quant. Management* **3** 161–177.

Ben-Tal, A. and A. Nemirovski. 2008. Selected topics in robust convex optimization. *Mathematical Programming* **112** 125–158.

Bertsimas, D., O. Nohadani, K.M. Teo. 2010. Robust optimization for unconstrained simulation-based problems. *Oper. Res.* **58** 161–178.

Bertsimas, D. and M. Sim 2004. The Price of Robustness. *Oper. Res.* **52** 35–53.

Bertsimas, D. and A. Thiele. 2004. A Robust Optimization Approach to Supply Chain Management, *Integer Programming and Combinatorial Optimization*, Springer, Lecture Notes in Computer Science **3064** 86–100.

Bertsimas, D. and A. Thiele. 2006. A Robust Optimization Approach to Inventory Theory. *Operations Research* **54** 150–168.

Beyer, H., B. Sendhoff. 2007. Robust optimization—a comprehensive survey. *Comput. Methods in Appl. Mech. and Engrg.* **196** 3190–3218.

Borgonovo, E., L. Peccati. 2007. Global sensitivity analysis in inventory management. *International J. of Prod. Econom.* **108** 302–313.

Del Castillo, E. 2007. *Process optimization: a statistical approach*. Springer, New York.

Dellino, G., J.P.C. Kleijnen, C. Meloni. 2010. Robust optimization in simulation: Taguchi and Response Surface Methodology. *Internat. J. of Prod. Econom.* **125** 52–59.

El-Haik, B., and R. Al-Aomar. 2006. *Simulation-based lean Six-Sigma and design for Six-Sigma*. John Wiley & Sons, Inc., New York.

Jin, R., Du, X. and W. Chen. 2003. The use of metamodeling techniques for optimization under uncertainty. *Structural Multidisciplinary Optim.* **25** 99–116.

Kleijnen, J.P.C. and R.G. Sargent. 2000. A methodology for the fitting and validation of metamodels in simulation. *Eur. J. of Oper. Res.* **120** 14–29.

Lee, K.H., G.J. Park. 2006. A global robust optimization using Kriging based approximation model. *J. of the Japan Soc. of Mech. Engrg.* **49** 779–788.

Lophaven, S.N., H.B. Nielsen, J. Søndergaard. 2002. DACE: a MATLAB Kriging toolbox (version 2.0). Technical Report IMM-TR-2002-12, Technical University of Denmark, Lyngby.

Miró-Quesada, G. and E. del Castillo. 2004. Two Approached for Improving the Dual Response Method in Robust Parameter Design. *J. of Quality Tech.* **36** 154–168.

Mula, J., Poler, R. and J.P. García-Sabater, and F.C. Lario. 2006. Models for production planning under uncertainty : a review. *Internat. J. of Production Econom.* **103** 271–285.

Myers, R. H., A. I. Khuri, and G. Vining. 1992. Response Surface Alternatives to the Taguchi Robust Parameter Design Approach. *The Amer. Statist.* **46** 131–139.

Myers, R.H., D.C. Montgomery, C.M. Anderson-Cook. 2009. *Response surface methodology: process and product optimization using designed experiments; third edition.* John Wiley & Sons, New York.

Park, G-J., T-H. Lee, K.H. Lee, K-H. Hwang. 2006. Robust design: an overview. *AIAA J.* **44** 181–191.

Robinson, T.J., C.M. Borror, and R.H. Myers. 2004. Robust Parameter Design: a Review. *Quality and Reliability Engrg. Internat. J.* **20** 81–101.

Sacks, J., W.J. Welch, T.J. Mitchell, H.P. Wynn. 1989. Design and analysis of computer experiments. *Statist. Sci.* **4** 409–435.

Sanchez, S. M. 2000. Robust Design: Seeking the Best of All Possible Worlds. *Proceedings of the Winter Simulation Conference*, J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick (eds.), 69–76.

Santner, T.J., B.J. Williams, W.I. Notz. 2003. *The design and analysis of computer experiments.* Springer-Verlag, New York.

Stinstra, E., D. den Hertog. 2008. Robust optimization using computer experiments. *Eur. J. of Oper. Res.* **191** 816–837.

Taguchi, G. 1987. *System of experimental designs, volumes 1 and 2.* UNIPUB/ Krauss International, White Plains, New York.

Trosset, M. W. 1997. Taguchi and Robust Optimization, Dept. of Computational and Applied Mathematics, Rice University, Technical Report 96-31, Houston, Texas, USA.

Zhang, Y. 2004. General Robust-Optimization Formulation for Nonlinear Programming, Dept. of Computational and Applied Mathematics, Rice University, Technical Report 04-13, Houston, Texas, USA.

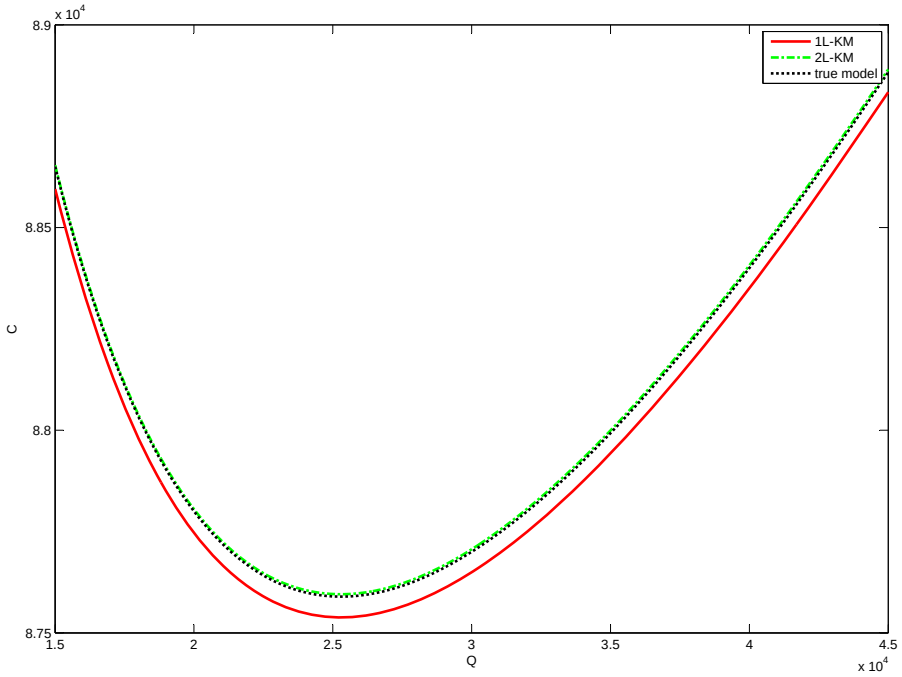


Figure 1: Kriging metamodels for the expected total cost of the EOQ model with uncertain demand rate and cost coefficients, derived through 1L-KM (solid curve) and 2L-KM (dashed curve). The dotted curve represents the true model.

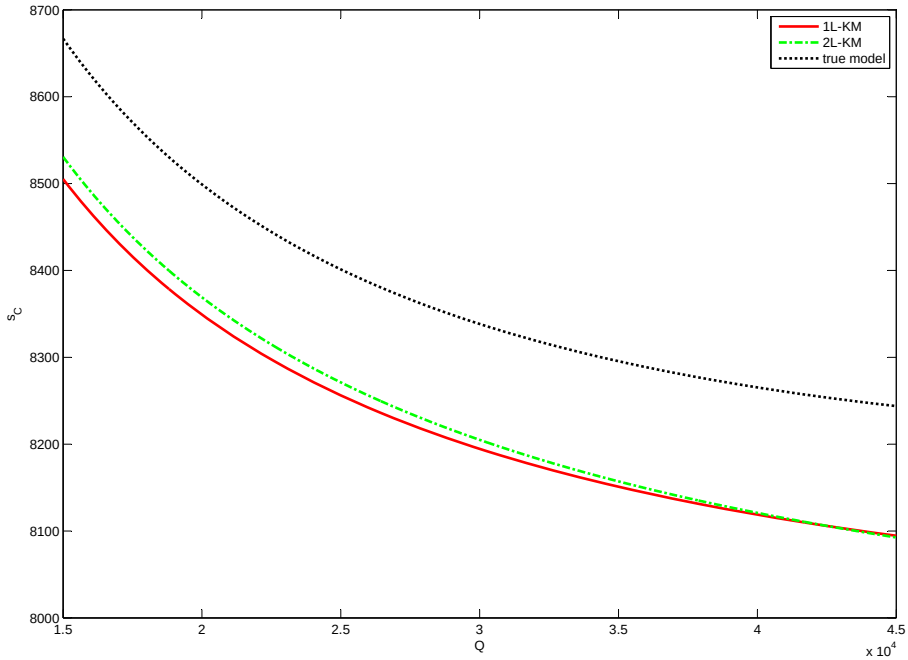


Figure 2: Kriging metamodels for the expected standard deviation of cost of the EOQ model with uncertain demand rate and cost coefficients, derived through 1L-KM (solid curve) and 2L-KM (dashed curve). The dotted curve represents the true model.

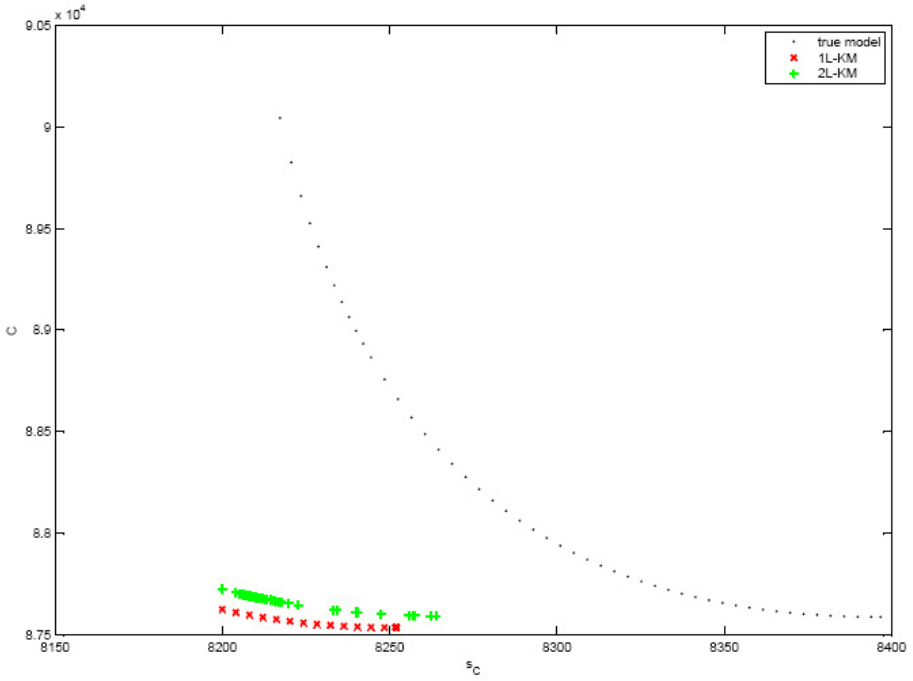
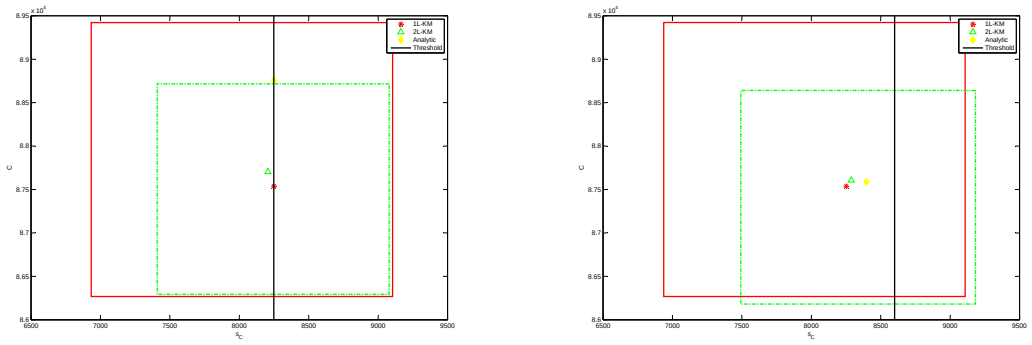


Figure 3: Pareto frontiers for the EOQ model with uncertain demand rate and cost coefficients, derived through 1L-KM approach (crosses) and 2L-KM (pluses). True model is displayed through dots.



(a) (b)

Figure 4: Confidence regions for σ_C (on x-axis) and $E(C)$ (on y-axis) in the extended EOQ example, based on bootstrapped Kriging in 1L-KM (solid rectangle) and 2L-KM (dash-dotted rectangle) at (a) $T = 8250$ and (b) $T = 8600$ (see the vertical line). '$*$' and '$\triangle$' denote the 'true' solutions based on (20) and (21) in the main paper, in 1L-KM and 2L-KM respectively; namely, (a) $(\widehat{\overline{C}}_{1L-KM} = 87538.62, \widehat{s}_{1L-KM} = 8248.49)$ at $\widehat{Q^+}_{1L-KM} = 25531.91$, and $(\widehat{\overline{C}}_{2L-KM} = 87703.95, \widehat{s}_{2L-KM} = 8205.61)$ at $\widehat{Q^+}_{2L-KM} = 29943.98$, (b) $(\widehat{\overline{C}}_{1L-KM} = 87538.28, \widehat{s}_{1L-KM} = 8251.94)$ at $\widehat{Q^+}_{1L-KM} = 25291.39$, and $(\widehat{\overline{C}}_{2L-KM} = 87605.84, \widehat{s}_{2L-KM} = 8287.48)$ at $\widehat{Q^+}_{2L-KM} = 24001.84$

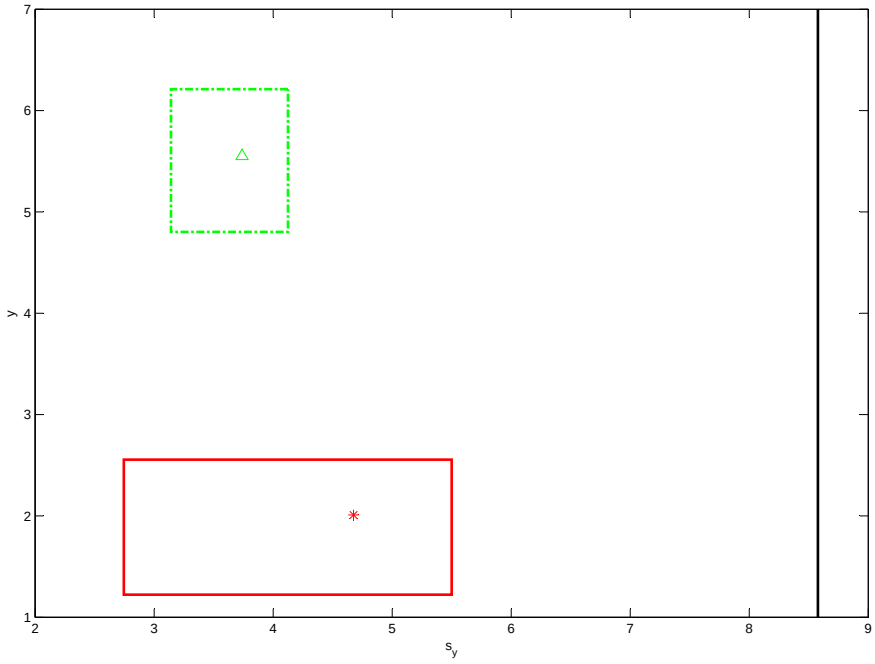


Figure 5: Confidence regions for σ_y (on x-axis) and $E(y)$ (on y-axis) based on bootstrapped Kriging in 1L-KM (solid rectangle) and 2L-KM (dash-dotted rectangle) at (a) $T = 8.58$ (see the vertical line). '$*$' and '$\triangle$' denote the 'true' solutions in 1L-KM and 2L-KM respectively; namely, $(\widehat{\overline{y}}_{1L-KM} = 2.01, \widehat{s}_{1L-KM} = 4.676)$ at $\widehat{\mathbf{x}^+}_{1L-KM} = (0,\ 0.16)$, and $(\widehat{\overline{y}}_{2L-KM} = 5.551, \widehat{s}_{2L-KM} = 3.739)$ at $\widehat{\mathbf{x}^+}_{2L-KM} = (0.089,\ 0.344)$