PARAMETRIC AND DISTRIBUTION-FREE BOOTSTRAPPING
IN ROBUST SIMULATION-OPTIMIZATION

Gabriella Dellino

Department of Information Engineering
University of Siena
Via Roma 56
53100 Siena, ITALY

Jack P.C. Kleijnen

Department of Information Management
Tilburg University
Postbox 90153
5000 LE Tilburg, THE NETHERLANDS

Carlo Meloni

Department of Electrical Engineering and Electronics
Polytechnic of Bari
Via E. Orabona 4
70125 Bari, ITALY

ABSTRACT

Most methods in simulation-optimization assume known environments, whereas this research accounts for uncertain environments combining Taguchi's world view with either regression or Kriging (also called Gaussian Process) meta-models (emulators, response surfaces, surrogates). These metamodels are combined with Non-Linear Mathematical Programming (NLMP) to find robust solutions. Varying the constraint values in this NLMP gives an estimated Pareto frontier. To account for the variability of this estimated Pareto frontier, this contribution considers different bootstrap methods to obtain confidence regions for a given solution. This methodology is illustrated through some case studies selected from the literature.

1 INTRODUCTION

Simulation-optimization aims at optimizing one or more performance measures that characterize the behavior of a system described through a simulation model. A major part in simulation-optimization assumes known environments, which implies that all relevant simulation inputs are supposed to be known. Unfortunately, ignoring the uncertainty in some inputs of the simulation model may lead to a suboptimal or infeasible solution. The goal of robust optimization is to tackle uncertainty affecting a system and derive solutions that are relatively insensitive to perturbations in the simulation inputs.

In such an uncertain world, strategic decision-making may use Taguchi's approach, originally developed to help Toyota design 'robust' cars; i.e., cars that perform reasonably well in many circumstances; see, among others, Beyer and Sendhoff (2007), Kleijnen (2008) and Taguchi (1987).

Our contribution is structured as follows. Section 2 describes the framework we propose for robust simulation-optimization, focussing on three alternative metamodeling approaches. Section 3 discusses the analysis of uncertainty through bootstrapping and confidence intervals. Section 4 presents the results of our methodology applied to a robust variant of the classic Economic Order Quantity (EOQ) model. Section 5 summarizes our conclusions.

2 METHODOLOGY FOR ROBUST SIMULATION-OPTIMIZATION

We combine Taguchi's view of systems affected by uncertainty with two metamodeling techniques, namely regression and Kriging. Metamodeling treats the simulation model as a black box; i.e., metamodeling observes only the

Input/Output (I/O) of the simulation model. Metamodels run much faster than the underlying–possibly computationally expensive–simulation models, enabling fast approximation. We estimate one univariate metamodel per simulation output (response), and combine these metamodels in a nonlinear Mathematical Programming model; the latter model selects one of the multiple simulation outputs as the goal variable, while the remaining outputs must satisfy given constraints (thresholds). Changing these thresholds on the constrained outputs and solving the corresponding Mathematical Programming problems provides an estimate of the Pareto frontier, enabling the users (managers) to select a robust solution for the simulated system. In summary, our methodology combines the following three building blocks: 1. Taguchi's worldview; 2. Design Of Experiments (DOE) and Metamodeling; 3. Mathematical Programming.

Taguchi (1987) identifies two types of input variables; namely, *decision* or *control* variables, which we denote by $\mathbf{d} = (d_j)_{j=1,\ldots,k}$, and *environmental* or *noise* variables, which we denote by $\mathbf{e} = (e_g)_{g=1,\ldots,c}$. The former are under the control of the manager, whereas the latter cannot be controlled. Taguchi considers a single output (say) w, and accounts for the uncertainty due to the environmental factors through optimizing a mean-to-variance ratio of this output.

Taguchi focusses on physical systems, whereas we focus on simulated systems. Assuming that a single run of a simulation model is computationally expensive, we approximate the system's behavior through three alternative approaches detailed in the following three subsections. Notice that we do not use Taguchi's statistical techniques, which have been criticized by many statisticians; for a detail discussion, see the panel report by Nair (1992). Assuming that the environmental variables follow a given statistical distribution, we combine a space-filling design for the decision variables with Latin Hypercube Sampling (LHS) for the environmental factors, sampled from the corresponding distribution.

2.1 Regression Metamodeling

In our first approach, we use regression. Myers et al.(2009) combine Taguchi's worldview with Response Surface Methodology (RSM). RSM proceeds stepwise (multi-stage); i.e., it consists of a sequence of first-order polynomial regression metamodels to search for the minimum, using steepest descent; it ends with a second-order polynomial to estimate the extremum (maximum, minimum, saddle point).

We adapt Myers et al.'s *robust* variant of RSM, accounting for the particularities of simulation described in Section 1. Like Myers et al., we assume

- a *second-order* polynomial with first-order effects $\boldsymbol{\beta}$ and second-order effects $\mathbf{B}$ for the decision factors $\mathbf{d}$;
- a *first-order* polynomial with first-order effects $\boldsymbol{\gamma}$ for the environmental factors $\mathbf{e}$;
- *decision-by-environmental two-factor interactions* $\boldsymbol{\Delta}$;

these assumptions result in the low-order polynomial metamodel (which is a regression model that is linear in its parameters)

$$y = \beta_0 + \boldsymbol{\beta}'\mathbf{d} + \mathbf{d}'\mathbf{B}\mathbf{d} + \boldsymbol{\gamma}'\mathbf{e} + \mathbf{d}'\boldsymbol{\Delta}\mathbf{e} + \varepsilon \tag{1}$$

where y denotes the metamodel's estimator (predictor) of the simulation output w, ε denotes the residual with $E(\varepsilon) = 0$ and constant variance σ_ε^2, $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_k)'$, $\mathbf{d} = (d_1,\ldots,d_k)'$, $\mathbf{B}$ denotes the $k \times k$ symmetric matrix with main-diagonal elements $\beta_{j;j}$ and off-diagonal elements $\beta_{j;j'}/2$, $\boldsymbol{\gamma} = (\gamma_1,\ldots,\gamma_c)'$, $\mathbf{e} = (e_1,\ldots,e_c)'$, and $\boldsymbol{\Delta} = (\delta_{j;g})$.

This metamodel (1) implies the regression predictor for the true mean $E(w)$

$$E(y) = \beta_0 + \boldsymbol{\beta}'\mathbf{d} + \mathbf{d}'\mathbf{B}\mathbf{d} + \boldsymbol{\gamma}'E(\mathbf{e}) + \mathbf{d}'\boldsymbol{\Delta}E(\mathbf{e}) \tag{2}$$

and the regression predictor for the true variance $var(w)$

$$var(y) = (\boldsymbol{\gamma}' + \mathbf{d}'\boldsymbol{\Delta})\boldsymbol{\Omega}_e(\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{d}) + \sigma_\varepsilon^2 \tag{3}$$

where the environmental factors are assumed to have a distribution with known $E(\mathbf{e}) = \boldsymbol{\mu}_{\mathbf{e}}$ and $cov(\mathbf{e}) = \boldsymbol{\Omega}_{\mathbf{e}}$, and σ_ε^2 is the (constant) variance of the residual; see Dellino, Kleijnen, and Meloni (2010a). To predict the standard deviation of the output, we take the square root of the right-hand side of (3).

2.2 One-Layer Kriging Metamodel (Kriging variant 1)

In our second approach, we replace the polynomial metamodel by a *Kriging* metamodel. Kriging provides more flexible metamodels than a low-order polynomial does; i.e., Kriging is better suited to global instead of local fitting; see del Castillo (2007), Kleijnen (2008). The polynomial metamodel (1) implies the mean and variance predictors (2) and (3); the Kriging metamodel, however, has no such analogue. We therefore estimate the response mean $E(w)$ and variance $var(w)$ from the crossed design (which is popular in the Taguchian approach). This design combines the (say) n_d combinations of the decision variables **d** and the n_e combinations of the environmental variables **e**. More specifically, we cross a space-filling design for the decision factors and a LHS design accounting for the distribution of the environmental factors. This design gives the simulation outputs $w_{i;j}$, so the mean and variance estimators become

$$\overline{w_i} = \frac{\sum_{j=1}^{n_e} w_{i;j}}{n_e} \qquad i = 1, \ldots, n_d, \tag{4}$$

$$s^2(w_i) = \frac{\sum_{j=1}^{n_e}(w_{i;j} - \overline{w_i})^2}{n_e - 1} \qquad i = 1, \ldots, n_d. \tag{5}$$

To the n_d estimates resulting from (4), we fit a Kriging metamodel for the mean; from the estimates in (5) we compute the n_d standard deviations, and fit a Kriging metamodel for the standard deviation.

2.3 Two-Layer Kriging Metamodel (Kriging variant 2)

Our third approach is still based on Kriging. In layer 1, however, our approach does not distinguish between decision and environmental variables yet, and no assumption is made for the distribution of the environmental factors; i.e., we select a space-filling design for the $k + c$ factors. In layer 2, we select a larger design by crossing a space-filling design for the decision factors, and a LHS design for the environmental factors accounting for the distribution of these environmental factors. For this large design we do not run expensive simulations but compute the Kriging predictions of the simulation output w. We use these predictions to estimate the conditional means and standard deviations; i.e., in the right-hand sides of (4) and (5) we replace $n_\mathbf{e}$ and $n_\mathbf{d}$ by $N_\mathbf{e}$ and $N_\mathbf{d}$—the large-sample analogues of the small-sample sizes $n_\mathbf{e}$ and $n_\mathbf{d}$—and w by $\widehat{y}$ where $\widehat{y}$ denotes Kriging prediction. From these I/O data we again build two Kriging metamodels for the mean and standard deviation of the output.

2.4 Mathematical Programming

Using the metamodels obtained through one of the approaches described in the preceding three subsections, we formulate the following Mathematical Programming (MP) problem:

$$\begin{aligned} &\min \overline{w} \\ &\text{s.t. } s_w \leq T \end{aligned} \tag{6}$$

for a given threshold T, quantifying the manager's risk attitude; i.e., lower T-values correspond to risk averse management, whereas higher T-values denote a risk seeking behavior. We use the standard deviation instead of the variance because both the mean and the standard deviation are expressed in the same units.

We solve the MP problem (6) through Matlab's `fmincon`, but many alternative solvers are available in the MP software.

Then, we move the threshold within a given interval, and solve the problem for each T value. Collecting these solutions, we derive an estimate of the Pareto frontier balancing the mean and standard deviation of the output w.

3   BOOTSTRAPPING FOR UNCERTAINTY ANALYSIS

The metamodels for $\overline{w}$ and s_w derived in Section 2 are based on estimates; therefore we further investigate these metamodels through bootstrapping.

3.1 Parametric bootstrapping of the regression metamodel

Following Dellino, Kleijnen, and Meloni (2010a), p. 17, Eq. (15), we apply parametric bootstrapping to the vector of regression coefficients $\boldsymbol{\zeta}$ using (7):

$$\widehat{\boldsymbol{\zeta}}^* \sim N_q(\widehat{\boldsymbol{\zeta}}, \widehat{\mathbf{cov}(\widehat{\boldsymbol{\zeta}})}) \tag{7}$$

where (also see Dellino, Kleijnen, and Meloni (2010a), p. 6, Eq. (5), and p. 7, Eq. (6)):

$$\widehat{\boldsymbol{\zeta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{w} \tag{8}$$

is the least squares estimator of the regression parameters $\boldsymbol{\zeta} = (\beta_0, \boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\gamma}, \boldsymbol{\delta})$ of size q where $\mathbf{b}$ is the vector that follows from stapling the elements of the matrix $\mathbf{B}$, and $\boldsymbol{\delta}$ is the vector following from stapling the elements of $\boldsymbol{\Delta}$, and

$$\widehat{\mathbf{cov}(\widehat{\boldsymbol{\zeta}})} = (\mathbf{X}'\mathbf{X})^{-1}\widehat{\sigma_\varepsilon^2} \tag{9}$$

where $\mathbf{X}$ is the $n \times q$ matrix of explanatory variables, accounting for the (first-order and some second-order) effects of the decision factors, environmental factors and their interactions; $n = n_\mathbf{d} \cdot n_\mathbf{e}$ denotes the number of combinations of the k decision factors which are combined in $n_\mathbf{d}$ combinations, and c environmental factors which are combined in $n_\mathbf{e}$ combinations, so n denotes the number of combinations in the crossed design; $\mathbf{w}$ consists of the n simulation outputs corresponding with the n simulation input combinations of decision and environmental factors; $\widehat{\sigma_\varepsilon^2}$ is estimated by the mean squared residuals. We emphasize that the covariance matrix defined in (9) assumes that the explanatory variables are fixed; i.e., this covariance matrix is conditional on the values of the environmental variables that have been sampled in the $n_\mathbf{e}$ combinations.

From (7) we get $\widehat{\boldsymbol{\zeta}}^*$, which we plug-in into (2) and (3) to obtain the bootstrapped regression predictors for the mean and variance of the simulation output:

$$\widehat{y}^* = \widehat{\beta_0}^* + \widehat{\boldsymbol{\beta}}^{*\prime}\mathbf{d} + \mathbf{d}'\widehat{\mathbf{B}}^*\mathbf{d} + \widehat{\boldsymbol{\gamma}}^{*\prime}\boldsymbol{\mu}_\mathbf{e} + \mathbf{d}'\widehat{\boldsymbol{\Delta}}^*\boldsymbol{\mu}_\mathbf{e}, \tag{10}$$

$$\widehat{var(y)}^* = (\widehat{\boldsymbol{\gamma}}^{*\prime} + \mathbf{d}'\widehat{\boldsymbol{\Delta}}^*)\boldsymbol{\Omega}_\mathbf{e}(\widehat{\boldsymbol{\gamma}}^* + \widehat{\boldsymbol{\Delta}}^{*\prime}\mathbf{d}) + \widehat{\sigma_\varepsilon^2}. \tag{11}$$

Notice that Kriging does not enable a parametric bootstrap procedure analogous to the one described for regression modeling.

3.2 Distribution-free bootstrapping of the regression and Kriging models

Following Dellino, Kleijnen, and Meloni (2010b), p. 8, we apply distribution-free (non-parametric) bootstrapping to the simulation output data $\mathbf{w}$, by resampling—with replacement—the $n_\mathbf{e}$ vectors $\mathbf{w}_j$ ($j = 1, \ldots, n_\mathbf{e}$), each with $n_\mathbf{d}$ elements; we resample these vectors instead on the $n = n_\mathbf{d} \times n_\mathbf{e}$ scalars w because all elements within such a vector are correlated (they use the same values for the environmental factors: common random numbers). This resampling gives the $n_\mathbf{e}$ bootstrapped output observations $\mathbf{w}_j^*$, which together give the matrix $\mathbf{W}^*$ with n elements. To reduce the sampling error in this bootstrapping, we repeat this sampling (say) B times; B is called the bootstrap sample size (in our numerical example we use $B = 100$).

In our regression model we go back to (8), where we now replace the vector $\mathbf{w}$ by $\mathbf{w}^*$; this vector $\mathbf{w}^*$ results from stapling the $n_\mathbf{e}$ ($n_\mathbf{d}$-dimensional) vectors in the matrix $\mathbf{W}^*$; this gives the bootstrapped regression parameter $\widehat{\boldsymbol{\zeta}}^*$:

$$\widehat{\boldsymbol{\zeta}}^* = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{w}^*. \tag{12}$$

Like in parametric bootstrapping, we derive the bootstrapped regression metamodels for the mean and standard deviation of the output from (10) and (11), where $\widehat{\boldsymbol{\zeta}}^*$ is now given by (12).

Comparing (12) and (7), we see that parametric bootstrapping relies on the fixed set of simulation outputs $\mathbf{w}$, whereas distribution-free bootstrapping creates variability in the simulation outputs; this variability corresponds with resampling of the environmental variables.

In the first Kriging variant we estimate the $n_{\mathbf{d}}$ bootstrapped conditional means and variances, analogously to (4) and (5):

$$\overline{w_i^*} = \frac{\sum_{j=1}^{n_{\mathbf{e}}} w_{i,j}^*}{n_{\mathbf{e}}} \quad (i = 1, \ldots, n_{\mathbf{d}}), \tag{13}$$

$$s_i^{2*} = \frac{\sum_{j=1}^{n_{\mathbf{e}}} \left(w_{i,j}^* - \overline{w_i^*}\right)^2}{n_{\mathbf{e}} - 1} \quad (i = 1, \ldots, n_{\mathbf{d}}). \tag{14}$$

From the estimates computed through (13) and (14) respectively we estimate two Kriging metamodels.

Our second Kriging variant directly fits a Kriging metamodel to the bootstrapped output observations in $\mathbf{W}^*$. Then we follow the procedure outlined in Section 2 to derive the Kriging metamodels for the mean and the standard deviation of the output.

3.3 Bootstrapped Confidence Intervals for means and standard deviations

Based on the bootstrapped metamodels obtained through either parametric or distribution-free bootstrapping described in Sections 3.1 and 3.2, we compute confidence intervals (CIs) for the *expected output* through the following procedure (CIs for the standard deviation will follow).

Select (say) p combinations for the decision variables $\mathbf{d}$ that are spread out over the area of interest. For each $\mathbf{d}_i$ $(i = 1, \ldots, p)$ repeat the following steps:

1. Compute the predictor of the expected output using the bootstrapped predictor for $E(w)$ as discussed in Sections 3.1 and 3.2. We now use the symbol $\widehat{y}^*$ to denote either a bootstrapped estimated regression or Kriging predictor of the mean.

2. Because (10) is a linear transformation of the q-variate normal variable defined in (7), compute the *parametric CI*, following Efron and Tibshirani (1993), p. 160:

$$\left[\overline{\widehat{y}_i^*} - t_{\alpha/2;B-1} se_i, \overline{\widehat{y}_i^*} + t_{\alpha/2;B-1} se_i\right] \tag{15}$$

where

$$\overline{\widehat{y}_i^*} = \frac{\sum_{b=1}^{B} \widehat{y}_b^*(\mathbf{d}_i)}{B}; \tag{16}$$

α is based on Bonferroni's inequality; i.e., assuming we are interested in an 80% confidence interval (because Bonferroni's inequality is conservative), we take $\alpha = 0.20/p$; se_i is the estimated standard error of $\overline{\widehat{y}_i^*} = \overline{\widehat{y}^*}(\mathbf{d}_i)$ defined in (16), computed as follows:

$$se_i = \sqrt{\frac{\sum_{b=1}^{B} \left(\widehat{y}_b^*(\mathbf{d}_i) - \overline{\widehat{y}_i^*}\right)^2}{B(B-1)}}. \tag{17}$$

Note that the t statistic is quite insensitive to deviations from normality, so (15) may also be used to compute the CI for the expected output in the Kriging-based approaches.

Alternatively, when the data deviate from normality, we compute a *non-parametric CI* based on order statistics denoted by the subscript ( ), following Efron and Tibshirani (1993), p. 170:

$$\left[\widehat{y}_{(\lfloor B\alpha/2 \rfloor)}^*(\mathbf{d}_i), \widehat{y}_{(\lceil B(1-\alpha/2) \rceil)}^*(\mathbf{d}_i)\right] \tag{18}$$

To compute CIs for the estimated *standard deviation* of the output, we follow a similar procedure. For each $\mathbf{d}_i$ $(i = 1, \ldots, p)$ repeat the following steps:

1. Compute the predictor of the expected standard deviation of the output using the bootstrapped predictor for $s(w)$ derived in Sections 3.1 and 3.2. We now use the symbol s^* to denote either a bootstrapped estimated regression or Kriging predictor of the mean.

2. Obviously, the standard deviation estimated through one of the three metamodeling approaches is not normally distributed. However, the t-statistic is known to be not very sensitive to nonnormality, so compute the *parametric CI*; i.e., in (15) we replace $\widehat{y}$ by s.

Alternatively, we compute a *non-parametric CI* based on order statistics:

$$\left[s^*_{(\lfloor B\alpha/2 \rfloor)}(\mathbf{d}_i), s^*_{(\lceil B(1-\alpha/2) \rceil)}(\mathbf{d}_i) \right] . \tag{19}$$

4 COMPUTATIONAL EXPERIMENTS WITH EOQ MODEL

For the classic EOQ model, Zipkin (2000), pp. 30-39, uses the following symbols and assumptions: (i) The demand is known and constant, say a units per time unit. (ii) The order quantity is Q units. (iii) Total costs consist of setup cost per order, K; cost per unit purchased or produced, c; and holding cost per inventory unit per time unit, h. Management's goal is to minimize the total costs per time unit C, over an infinite time horizon.

To study robustness issues, we consider a variant of the classic EOQ model which deals with parameters affected by uncertainty—like Dellino, Kleijnen, and Meloni (2010a) do. Note that the robustness of the EOQ model is also examined by Yu (1997), who uses other criteria and other methods than we do (he uses two minmax criteria and analytical methods instead of simulation).

Note that the EOQ model is usually a building block for more complex supply chain models.

4.1 EOQ model with uncertain demand rate

Following Dellino, Kleijnen, and Meloni (2010a), we assume that a (demand per time unit) is an *unknown* constant; i.e., a has a Normal distribution with mean μ_a and standard deviation σ_a: $a \sim N(\mu_a, \sigma_a)$. Furthermore, we assume $\mu_a = 8000$ (nominal value), and $\sigma_a = 0.10\mu_a$ (uncertainty about the true input parameter). Because this standard deviation can give negative values for a, we truncate the normal distribution at zero; we use this adjustment of the Normal distribution for the sampling phase but ignore this truncation in our further analysis. We apply the three metamodel approaches discussed in Section 2.

For the regression and Kriging variant 1, we select a DOE consisting of $n_\mathbf{d} = 10$ values for the order quantity Q, equally spaced within the interval $[15000, 45000]$ and almost centered around the classic EOQ, which is close to 30000, and $n_\mathbf{e} = 100$ values for the demand rate sampled through a normal LHS (we use Matlab's lhsnorm). For the Kriging variant 2, we first select $n_\mathbf{d} \times n_\mathbf{e} = 1000$ points from a uniform grid in the decision×environmental space; then, for the second layer, we cross $N_\mathbf{d} = 30$ equally spaced values for Q with $N_\mathbf{e} = 200$ values sampled from a normally distributed LHS.

Based on these input designs, we run the simulation model and then fit either the regression or the Kriging metamodels for the mean and standard deviation of the output. Then, we estimate the Pareto frontier through solving the problem in (6) for different threshold values T; we consider 100 values for T in the interval $[8200, 8600]$.

Next we compute CIs following the procedures sketched in Section 3.3. Because Bonferroni's inequality becomes more conservative as p (number of CIs) increases, we take $p = 3$ values for the threshold T ($T_1 = 8200$, $T_2 = 8300$, $T_3 = 8600$) and compute the CIs at the corresponding estimated optimal Q-values.

We report only the parametric CIs for the mean computed through (15) and the non-parametric CIs for the standard deviation, applying (19). In fact, our regression metamodel gives a predicted mean through (10) that passes the normality test because $\boldsymbol{\zeta}^*$ is normal (see (7)), and (10) is a linear transformation of $\boldsymbol{\zeta}^*$. We expect that our regression metamodel gives a predicted standard deviation that is not normal, because (11) is a nonlinear transformation of $\boldsymbol{\zeta}^*$. Fig. 1 shows the CIs for parametric bootstrapping, for the two extreme values of the threshold, $T = T_1$ and $T = T_3$; obviously these CIs do not cover the true values of the means and standard deviations.

Fig. 2 displays the CIs for distribution-free bootstrapping of the output data (as described in Section 3.2). Comparing Figs. 1 and 2 shows that non-parametric bootstrapping gives longer CIs that always cover the true values. Our conclusion is that parametric bootstrapping assuming fixed environmental variables may give misleading results.
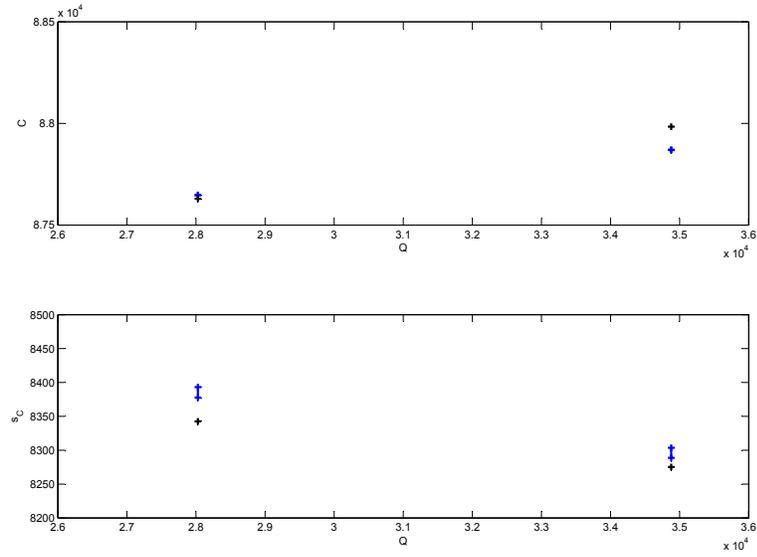
Figure 1: CIs for parametric bootstrapping of regression metamodels for the EOQ model with uncertain demand rate; true values are displayed in black.
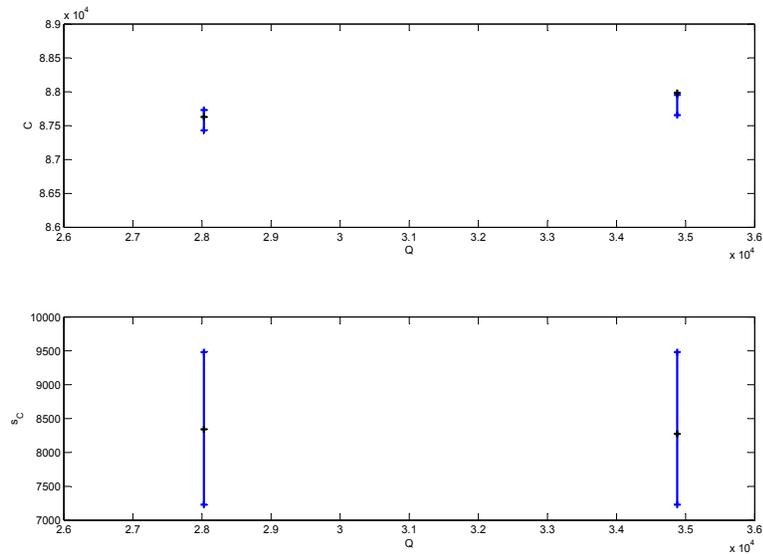
Figure 2: CIs for non-parametric bootstrapping of regression metamodels for the EOQ model with uncertain demand rate; true values are displayed in black.
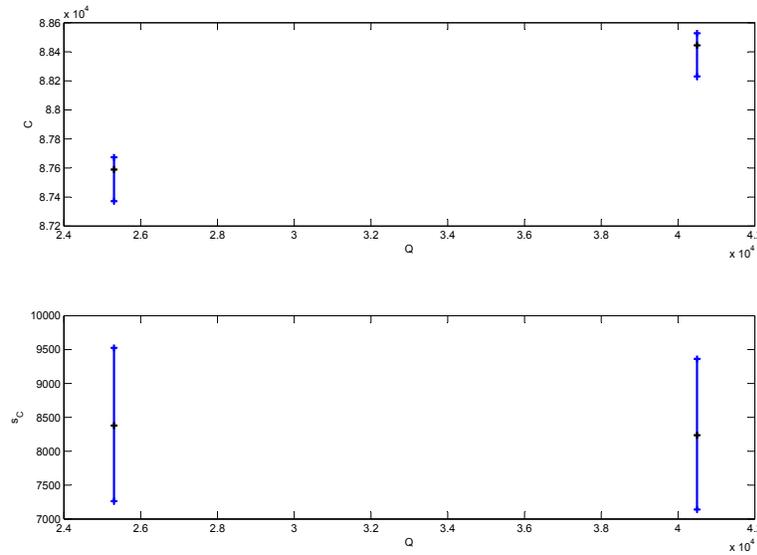
Figure 3: CIs for non-parametric bootstrapping of Kriging metamodels through variant 1 for the EOQ model with uncertain demand rate; true values are displayed in black.

Note that we also investigate whether the number of bootstrap samples B could change the results of the comparison. Eq. (17) prove that the CI length decreases with $\sqrt{B}$; see the factor B in the denominator (the factor $B-1$ is needed because the numerator adds B terms). For the distribution-free CI it is not so simple to apriori assess the behavior of the CI length as a function of B. We increase B from 100 to 1000, and observe that the CIs do not change significantly, with a maximum difference in length around 5%.

For the two Kriging variants we apply the distribution-free bootstrapping described in Section 3.2. Although Kriging models with estimated weights are nonlinear so they give nonnormal predictors, the B bootstrapped Kriging predictions of the mean for a given $Q = Q_i$ ($i = 1, \ldots, p$) pass Lilliefors test; i.e., the null hypothesis of a normal distribution is not rejected, meaning that the actual distribution is "close" to normal. Therefore we compute the parametric CIs for the mean through (15). For the standard deviation we do not expect normality, so we use (19) to derive the non-parametric CI.

Figures 3 and 4 show that both Kriging variants provide comparable CIs, always covering the true values.

4.2 EOQ with uncertain cost parameters and demand rate

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

Figure 5 shows that in the regression-based approach the CI for the mean cost never covers the true value, while the CI for the standard deviation covers the true value in one case only. So the regression metamodel appears a bad approximation of the simulation model's I/O function.

Figure 6 shows that in the first Kriging variant the CIs for both the mean and the standard deviation of the inventory cost do cover the true values for all the selected optimal Q_i values. So Kriging is flexible enough to provide a valid metamodel, and the distribution-free bootstrap gives CIs with proper coverage.
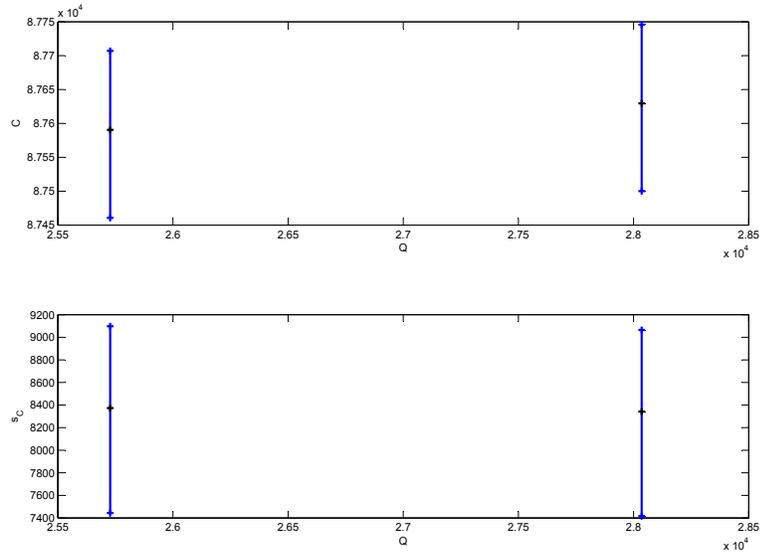
Figure 4: CIs for non-parametric bootstrapping of Kriging metamodels through variant 2 for the EOQ model with uncertain demand rate; true values are displayed in black.
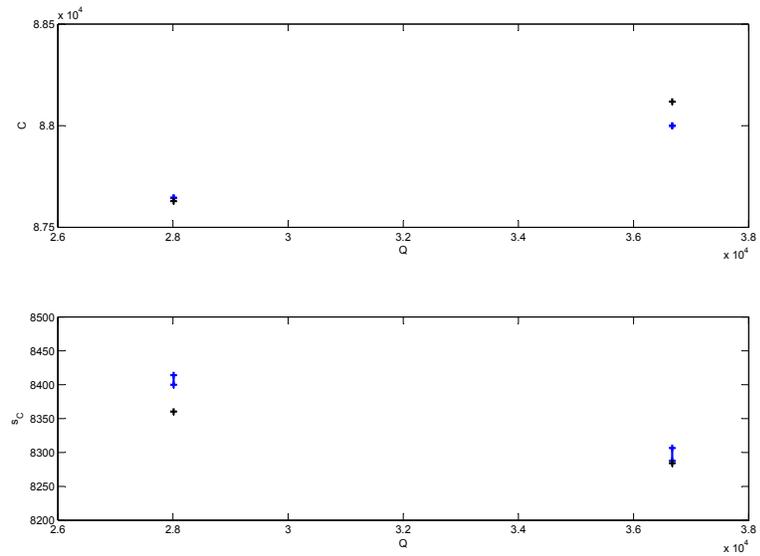
Figure 5: CIs for parametric bootstrapping of regression metamodels for the EOQ model with uncertain demand rate and costs; true values are displayed in black.
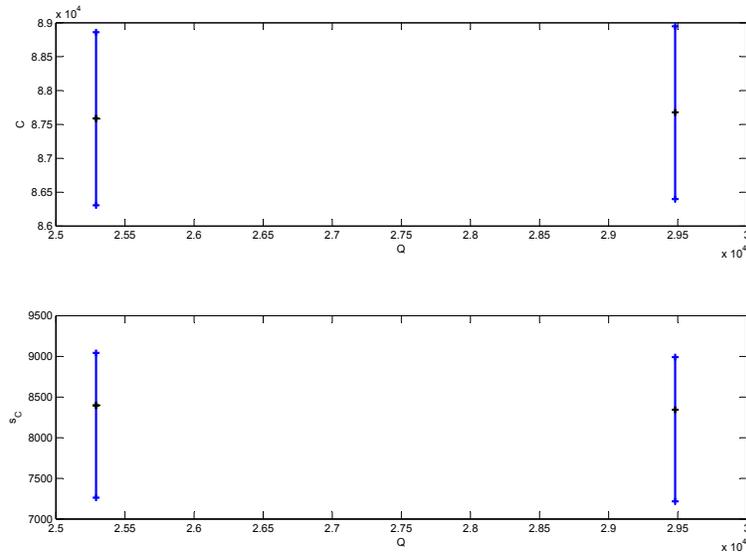
Figure 6: CIs for distribution-free bootstrapping of Kriging metamodels through variant 1 approach for the EOQ model with uncertain demand rate and costs; true values are displayed in black.

Figure 7 shows that in the second Kriging variant the CIs computed through (18) and (19) always cover the true values for both the mean and the standard deviation of the cost. The *t*-statistic (not displayed) gives much tighter CIs for the standard deviation than the order statistics do, but sometimes the former CIs miss the true standard deviations. A smaller type-I error rate would give longer CIs that do cover the true values.

5   CONCLUSIONS

We presented a framework for robust optimization of simulated systems based on three metamodeling approaches; namely, one regression and two Kriging models. To further analyze the effects of the uncertainty in the environmental variables, we proposed to use bootstrapping; more specifically, we compared parametric and distribution-free bootstrapping and derived confidence intervals for the optimal solutions belonging to the Pareto frontier.

We tested our new methodology through a popular inventory model; namely the EOQ model. We introduced a robust formulation of the EOQ model when the demand rate and the cost coefficients are uncertain. Because we can derive the true expected cost and its standard deviation, we could verify the performance of our heuristic.

The EOQ example gives encouraging results, so in future research we might apply the methodology to more complex and realistic models.

ACKNOWLEDGMENTS

REFERENCES

Beyer, H., and B. Sendhoff. 2007. Robust optimization—A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering* 196:3190–3218.

Borgonovo, E., and L. Peccati. 2007. Global sensitivity analysis in inventory management. *International Journal of Production Economics* 108:302–313.

del Castillo, E. 2007. *Process optimization: a statistical approach*. New York: Springer.
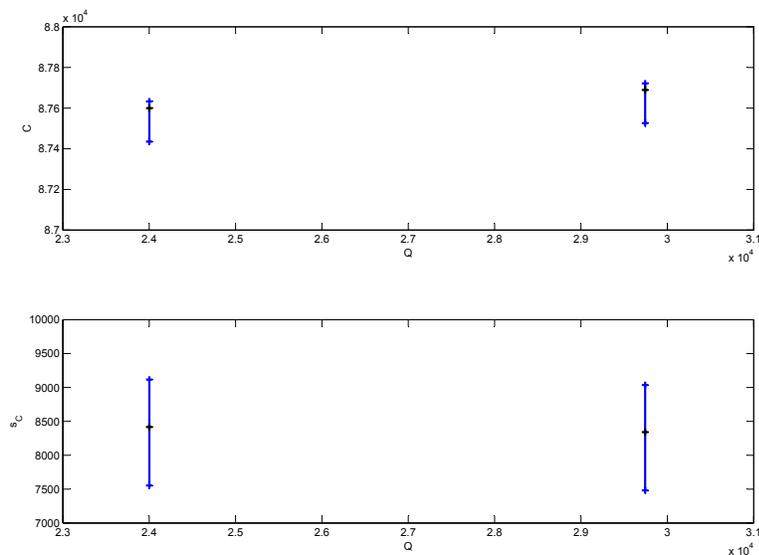
Figure 7: CIs for distribution-free bootstrapping of Kriging metamodels through variant 2 for the EOQ model with uncertain demand rate and costs; true values are displayed in black.

Dellino, G., J.P.C. Kleijnen, and C. Meloni. 2010. Robust optimization in simulation: Taguchi and Response Surface Methodology. *International Journal of Production Economics* 125:52–59.

Dellino, G., J.P.C. Kleijnen, and C. Meloni. 2009. Robust optimization in simulation: Taguchi and Krige combined. Working Paper Tilburg University, Tilburg, The Netherlands.

Efron, B., and R.J. Tibshirani. 1993. *An introduction to the bootstrap*. New York: Chapman & Hall.

Kleijnen, J.P.C. 2008. *Design and analysis of simulation experiments*. New York: Springer Science + Business Media. (Chinese translation will be published by Publishing House of Electronics Industry)

Myers, R.H., D.C. Montgomery, and C.M. Anderson-Cook. 2009. *Response Surface Methodology: Process and Product Optimization using Designed Experiments*. New York: John Wiley & Sons.

Nair, V.N. (ed.). 1992. Taguchi's parameter design: a panel discussion. *Technometrics* 34:127–161.

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

Yu, G. 1997. Robust economic order quantity models. *European Journal of Operational Research* 100:482–493.

Zipkin, P.H. 2000. *Foundations of inventory management*. McGraw-Hill.

AUTHOR BIOGRAPHIES

GABRIELLA DELLINO graduated cum laude in Computer Science Engineering at the Politecnico di Bari, in 2005. In 2009 she got a Ph.D. in Applied Mathematics from the University of Bari. She is currently working at the University of Siena, as a research associate in Operations Research. She took part in research projects funded by academic organizations and companies. Her main research interests are optimization models and methods, and computer simulation methodologies, with applications to Multidisciplinary Design Optimization, Health Care Operations Management, and energy systems.. She is a member of INFORMS. In 2008 she was visiting the CentER Graduate School, at Tilburg University. Her email address is <dellino@dii.unisi.it>.

JACK P.C. KLEIJNEN is Professor of 'Simulation and Information Systems' at Tilburg University, where he is a member of the 'Department of Information Management' and the 'Operations Research' group of CentER of the Tilburg School of Economics and Management. His research concerns simulation, especially its statistical design and analysis. He received several national and international awards, including a Knighthood in the Order of

the Netherlands Lion in 2008 and the INFORMS Simulation Society's Lifetime Professional Achievement Award (LPAA) in 2005. He serves on many international editorial boards and scientific committees. He spent some years in the US, at universities and private companies, and has been a consultant for several organizations in Europe and the US. More details are given on his website `<center.uvt.nl/staff/kleijnen>`.

CARLO MELONI is Assistant Professor of Systems Engineering and Optimization at the Politecnico di Bari. He participated in several research projects funded by various academic organizations and companies. His main research and professional interests concern applications of optimization and other OR/MS methodologies to Multidisciplinary Systems Design, Supply Chain Management, Advanced Planning & Scheduling, and Health Care Service Operations Management. He is a member of the Italian Society of Operations Research (AIRO) and INFORMS. His email address is `<meloni@deemail.poliba.it>`.