ROBUST SIMULATION-OPTIMIZATION USING METAMODELS

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

Optimization of simulated systems is the goal of many methods, but most methods assume known environments. In this paper we present a methodology that does account for uncertain environments. Our methodology uses Taguchi's view of the uncertain world, but replaces his statistical techniques by either Response Surface Methodology or Kriging metamodeling. We illustrate the resulting methodology through the well-known Economic Order Quantity (EOQ) model.

1 INTRODUCTION

The main purpose of this contribution is to present a *methodology* for *robust* simulation-optimization. Simulation is much applied in Operations Research and Management Science (ORMS). The goal of these simulations is often the optimization of the real system being simulated (another goal is sensitivity analysis, which often precedes optimization; see Kleijnen 2009). The resulting problem domain is called *simulation-optimization*; this domain includes a variety of methodologies (see Fu 2007). Unfortunately, these methodologies ignore the fact that, in practice, some inputs of the given simulation model are *uncertain*, so the optimum solution that is derived—ignoring these uncertainties—may be completely wrong! Robust optimization accounts for this uncertainty during the optimization process, thus deriving solutions that are relatively insensitive to perturbations in the model parameters. We summarize a methodology that solves robustness issues in simulation-optimization through the use of a metamodel of the underlying simulation model; such a metamodel is used because simulation runs are often computationally expensive. More details on our methodology are presented in Dellino, Kleijnen, and Meloni (2008) and Dellino, Kleijnen, and Meloni (2009).

We organize this paper as follows. Section 2 details the main steps of the proposed methodology. Section 3 discusses the results of an inventory example, namely the Economic Order Quantity (EOQ) model. Section 4 draws some conclusions and proposes directions for future research.

2 PROPOSED METHODOLOGY

Strategic decision-making in an uncertain world may use Taguchi's approach, originally developed to help Toyota design 'robust' cars; i.e., cars that perform reasonably well in many circumstances—see Beyer and Sendhoff (2007), Kleijnen (2008), Park et al. (2006), Taguchi (1987), Wu and Hamada (2000).

We combine Taguchi's approach with either Response Surface Methodology (RSM) or Kriging metamodeling. *Metamodels* treat the simulation model as a black box; i.e., only the Input/Output (I/O) of the simulation model is observed (metamodels are also called response surfaces, surrogates, emulators, auxiliary models, repromodels, etc.). Metamodels run much faster than the underlying—possibly computationally expensive—simulation models, enabling fast approximation. We use one (univariate) metamodel per simulation output (response) combined with 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 the thresholds (right-hand sides) in the Mathematical Programming solver to the metamodels gives an estimate of the Pareto frontier, enabling the selection of a robust solution of the simulation model, as we shall show in Section 2.3.

We therefore derive a methodology that combines: 1. Taguchi's worldview; 2. Design Of Experiments (DOE) and Metamodeling; 3. Mathematical Programming. We detail each of the three scientific approaches in the next subsections.

2.1 Taguchi's Worldview

Taguchi (1987) distinguishes between two types of input factors:

- Decision (or control) factors, denoted by (say) d_j $(j = 1, \ldots, k)$
- Environmental (or noise) factors, e_g $(g = 1, \ldots, c)$.

By definition, the decision factors are under the control of the users; e.g., in inventory management, the order quantity (denoted by Q) is controllable. The environmental factors are not controlled by those users; e.g., in inventory management the demand rate may not be controllable. Taguchi assumes a single output (say) w, and focuses on its mean $E(w) = \mu_w$ and variance $var(w) = \sigma_w^2$.

The Taguchian worldview has been very successful in production engineering. Nevertheless, statisticians have criticized Taguchi's statistical techniques; see Nair et al. (1992). We agree with this criticism. Moreover, Taguchi limits his work to real-life experiments, whereas we focus on *simulation experiments*. The latter experiments enable the exploration of many values per input and many combinations of these values; also see Kleijnen et al. (2005). To select these combinations, risk and uncertainty analyses often use Latin Hypercube Sampling (LHS); see Kleijnen (2008) for a discussion including references and websites. In our robust simulation-optimization, we use LHS to select the (say) n_e combinations of the environmental factor values; we use a space-filling design (e.g., a uniform grid if there only a few decision factors) to select the n_d combinations of the decision factors.

Moreover, we do not use Taguchi's *scalar loss function* such as the signal-to-noise or mean-to-variance ratio; for a further discussion of these loss functions we refer to Park et al. (2006), and Wu and Hamada (2000). Instead, we assume a single random simulation output w, and formulate the problem as a constrained minimization problem where we minimize the mean simulation output, under the constraint of a fixed threshold for the standard deviation. To solve this problem, we apply Mathematical Programming to a metamodel for this mean and a metamodel for this standard deviation. Then, we change the upper bound for the standard deviation, which gives the Pareto-optimal efficiency frontier—briefly called the *Pareto frontier*—where we consider the mean and standard deviation as criteria to be balanced. This is a classical approach to solve optimization problems with multiple criteria (see Miettinen 1999).

2.2 DOE and Metamodeling

ORMS and engineering often use expensive computer simulation models to optimize complex systems. To reduce the high computational costs of running such simulation models, metamodels are used. However, integrating metamodels in a classical, evolutionary or meta-heuristic optimization scheme is not straightforward; i.e., various issues emerge in their construction and management in order to coordinate optimization strategies and approximation efforts, pursuing the overall quality of the process (Dellino et al. 2009). In this paper we consider two types of metamodels; namely, regression and Kriging. We derive and manage such approximations through the three alternative approaches summarized in Figure 1 and detailed in what follows.

In our first approach, we use RSM. Myers and Montgomery (2002) combine Taguchi's worldview with RSM. RSM proceeds stepwise (multi-stage); i.e., it consists of a sequence of first-order polynomial metamodels to search for the minimum, using steepest descent; it ends with a *second-order polynomial* to estimate the minimum (or saddle point). Originally, Box and Wilson (1951) developed RSM for 'classic' optimization of real-life systems; i.e., they ignored environmental uncertainty.

Robust Optimization combining Taguchi & Metamodels

Input: design matrix $\mathbf{D}$ in the decision-by-environmental space $(\mathbf{d}, \mathbf{e})$;
  probability distribution of environmental factors;

begin
  for each design point (d_j, e_j) in $\mathbf{D}$, $j = 1, \ldots, n_\mathbf{d} \times n_\mathbf{e}$
    1. Run the simulation model f: $w_j = f(d_j, e_j)$
  switch sel_approach
    case Approach 1
      2. Fit a regression metamodel for the output, y
      3. Validate the metamodel, using leave-one-out cross validation
      4. From the fitted metamodel derive a metamodel for the mean of the response, $\bar{y}$, and one for its standard deviation, $s(y)$
    case Approach 2
      2. Fit one Kriging metamodel for the average response, $\bar{y}$, and one for its standard deviation, $s(y)$
      3. Validate the metamodels, using leave-one-out cross validation
    case Approach 3
      2. Fit a Kriging metamodel for the output, y
      3. Validate the metamodel, using leave-one-out cross validation
      4. Generate a new bigger design in the decision-by-environmental space
      5. Fit one Kriging metamodel for the mean of the predicted response, $\widehat{\bar{y}}$, and one for its standard deviation, $s(\widehat{y})$
      6. Validate the metamodels, using leave-one-out cross validation
  7. Estimate the Pareto frontier, solving constrained optimization problems with varying threshold values
  switch sel_approach
    case Approach 1
      8. Apply parametric bootstrap to the regression coefficients, B times
    case Approach 2, 3
      8. Apply distribution-free bootstrap to the simulation output data w, B times
  9. Derive confidence regions using the bootstrapped metamodels **end**

Figure 1: Robust optimization through metamodels: three approaches

This type of RSM has already built a track record; see Myers, Khuri, and Carter (1989) and again Myers and Montgomery (2002).

  We adapt Myers and Montgomery's *robust* variant of RSM, and account for the particularities of simulation as described in Section 2.1, whereas they use the classic Central Composite Design (CCD). Like Myers and Montgomery, 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 effects $\boldsymbol{\gamma}$ for the environmental factors $\mathbf{e}$;
- *decision-by-environmental two-factor interactions* $\boldsymbol{\Delta}$,

which results 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 $\boldsymbol{\Delta} = (\delta_{j;g})$. (The constant-variance assumption is unrealistic in many simulation applications. Relaxation of this assumption is discussed by Kleijnen 2008, pp. 87–93. Its consequences for our RSM approach deserve more research.)

  This metamodel 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 $var(w)$

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

where $\mathbf{l} = (\boldsymbol{\gamma} + \boldsymbol{\Delta}'\mathbf{d}) = (\partial y/\partial e_1, \ldots, \partial y/\partial e_c)'$; i.e., $\mathbf{l}$ is the gradient with respect to the environmental factors. So, the larger the gradient's components are, the larger the predicted variance of the simulation output is. Furthermore, if $\boldsymbol{\Delta} = \mathbf{0}$ (no decision-by-environmental interactions), then $var(y)$ cannot be controlled through the decision variables $\mathbf{d}$. Notice the difference between the predicted variance, $\widehat{var(y)}$, and the variance of the predictor, $var(\widehat{y})$ with $\widehat{y} = \widehat{\beta}_0 + \widehat{\boldsymbol{\beta}}'\mathbf{d} + \mathbf{d}'\widehat{\mathbf{B}}\mathbf{d} + \widehat{\boldsymbol{\gamma}}'\mathbf{e} + \mathbf{d}'\widehat{\boldsymbol{\Delta}}\mathbf{e}$.

To estimate the regression parameters in the RSM metamodel (1) (which implies (2) and (3)), we rewrite (1) as

$$y = \boldsymbol{\zeta}'\mathbf{x} + \varepsilon \qquad (4)$$

with $\boldsymbol{\zeta} = (\beta_0, \boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\gamma}, \boldsymbol{\delta})'$ where its components were defined below (1). Note that (4) is linear in the regression parameters $\boldsymbol{\zeta}$, whereas (1) is not linear in the decision variables $\mathbf{d}$. Then (4) gives the Least Squares (LS) estimator

$$\widehat{\boldsymbol{\zeta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{w}, \qquad (5)$$

where $\mathbf{X}$ is the $n \times q$ matrix of explanatory variables with n denoting the number of scenarios (combinations of decision and environmental factors) that are actually simulated, and q denoting the number of parameters collected in $\boldsymbol{\zeta}$; $\mathbf{w}$ consists of the n simulation outputs. The covariance matrix of $\widehat{\boldsymbol{\zeta}}$ is

$$\mathbf{cov}(\widehat{\boldsymbol{\zeta}}) = (\mathbf{X}'\mathbf{X})^{-1}\sigma_w^2. \qquad (6)$$

The RSM metamodel (1) implies that σ_w^2 equals σ_ε^2. This variance is estimated by the Mean Squared Residuals (MSR)

$$MSR = \frac{(\widehat{\mathbf{y}} - \mathbf{w})'(\widehat{\mathbf{y}} - \mathbf{w})}{n - q} \qquad (7)$$

where $\widehat{\mathbf{y}} = \widehat{\boldsymbol{\zeta}}'\mathbf{x}$; also see Kleijnen 2008, p. 23.

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 (del Castillo 2007, Kleijnen 2008). The simple 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 μ_w and variance σ_w^2 from the crossed design (which is popular in the Taguchian approach), which combines the (say) n_d combinations of the decision variables and the n_e combinations of the environmental variables and gives the simulation outputs $w_{i;j}$ so the mean and variance estimators become

$$\overline{w}_i = \frac{\sum_{j=1}^{n_e} w_{i;j}}{n_e} \qquad i = 1, \ldots, n_d, \qquad (8)$$

$$s^2(w_i) = \frac{\sum_{j=1}^{n_e} (w_{i;j} - \overline{w}_i)^2}{n_e - 1} \qquad i = 1, \ldots, n_d. \qquad (9)$$

To the n_d estimates in (8), we fit a Kriging metamodel for the mean; to the estimates in (9) we fit a Kriging metamodel for the standard deviation. To analyze these two metamodels, we use bootstrapping (see Section 2.3 below). We cross the following specific designs for the selection of the n_d and n_e combinations: a space-filling for the decision factors and LHS accounting for the distribution of the environmental factors .

We also implement a third approach still based on Kriging. First, we fit a Kriging metamodel for the simulation output, w. At this stage, however, the procedure does not distinguish between decision and environmental factors yet, and no assumption is made for the distribution of the environmental factors; i.e., we select a space-filling design for the $n_d + n_e$ factors. Next, we build a larger design by crossing a uniform space-filling design for the decision factors, and a LHS design for the environmental factors, this time accounting for their distribution. For this large input set we do not run expensive simulations but compute the Kriging predictions of the simulation output w and use these data to derive the conditional

means and standard deviations (w.r.t. the environmental variables); i.e., in the right-hand sides of (8) and (9) we replace n_e and n_d by N_e and N_d—the large-sample analogues of the small-sample n_e and n_d—and w by $\widehat{y}$ where the 'hat' denotes Kriging prediction.

In our three approaches we validate all the metamodels by applying the leave-one-out cross validation; see Kleijnen (2008).

2.3 Mathematical Programming and Bootstrapping

To solve the compromise between mean and variance, we use a Mathematical Programming formulation minimizing the mean subject to the constraint that the standard deviation remain below a given threshold (say) T. By using an appropriate Mathematical Programming solver, we estimate the robust optimal decision variables (say) $\widehat{\mathbf{d}}^+$ and the corresponding predicted mean $\widehat{y}$ and standard deviation $\widehat{\sigma}_y$. Next, we vary the threshold value T (say) S times (e.g., $S = 100$) to obtain different solutions $\widehat{\mathbf{d}}_t^+$ with corresponding $\widehat{y}_t$ and $\widehat{\sigma}_{y;t}$, $(t = 1, \ldots, S)$. The S pairs $(\widehat{y}_t, \widehat{\sigma}_{y;t})$ give the estimated Pareto frontier; see Figure 4, to be discussed later.

Figure 2(a) displays the average cost versus the order quantity in the upper part, and the estimated standard deviation versus the order quantity in the lower part, for the EOQ example (detailed in the next section). The upper part resembles the classic EOQ with a known demand rate; the lower part illustrates that the standard deviation decreases as the order quantity increases. The lower part also displays two *threshold* values for the standard deviation, called T_1 and T_2. If the threshold is high (T_1), the forbidden area (displayed in red) is small and the permitted area is large (not displayed explicitly); we may then select the order quantity that minimizes the cost, ignoring this constraint: see Q_1 in the upper part. If the threshold is small (T_2), the permitted area (displayed in green), is smaller and the forbidden area is larger (not displayed); the optimal order size must be higher than Q_1, namely Q_2; the constraint is binding (active).

The two threshold values T_1 and T_2 in Figure 2(a) give two points on the *Pareto* curve; see Figure 2(b). The point corresponding with the low threshold T_2 is a binding constraint, which we denote in Figure 2(b) by displaying the symbol T_2 in the color red; the point corresponding with the high threshold T_1 is not binding, which we denote by green. The binding constraint implies that the standard deviation equals the threshold T_2; the nonbinding constraint implies that the standard deviation is lower than the threshold T_1.

We let management make a choice based on the best estimate we have of the Pareto curve (namely, the 'original' curve, not the bootstrapped curves; see Figure 3). Management makes such a choice based on its *risk attitude*; e.g., if they are risk averse, then they impose the low threshold T_2 for the standard deviation of the cost. Hence they should select the solution that corresponds with T_2 in Figure 2(b); i.e., they are prepared to accept a relative high expected cost, because this choice implies a relative low standard deviation.

Finally, we warn management that their choice may turn out to result in a mean and standard deviation that differs from the values displayed in Figure 2(b). The cause is that the Pareto curve in Figure 2(b) is not known with certainty; i.e., this curve is *estimated* from a simulation with sampled values for the environmental factors (namely, the demand rate in the EOQ example). We can quantify this uncertainty through a *confidence region* for the mean cost and the standard deviation of the cost, given the order quantity Q_2 (Q_2 was displayed in Figure 2(a)). Technically, we derive such a confidence region through *bootstrapping*; see Efron and Tibshirani (1993), Kleijnen (2008). We use *parametric* bootstrapping for our RSM approach and *nonparametric distribution-free* bootstrapping for our two Kriging approaches. For further details on bootstrapping, we refer to Efron and Tibshirani (1993) and Chernick (2008).

In our RSM approach, we apply *parametric* bootstrapping of the estimated regression parameters $\widehat{\boldsymbol{\zeta}}$, in order to estimate the variability of the RSM metamodel. More specifically, we sample—via the Monte Carlo method, using pseudo-random numbers—(say) B times from a multivariate normal distribution with mean vector $\widehat{\boldsymbol{\zeta}}$ and covariance matrix $\widehat{\mathbf{cov}(\boldsymbol{\zeta})}$ defined in Eqs. (5-7). This sampling gives $\widehat{\boldsymbol{\zeta}}_b^* = (\widehat{\beta}_{0b}^*, \widehat{\boldsymbol{\beta}}_b^*, \widehat{\mathbf{b}}_b^*, \widehat{\boldsymbol{\gamma}}_b^*, \widehat{\boldsymbol{\delta}}_b^*)'$ $(b = 1, \ldots, B)$, from which we derive $\widehat{y}_b^*$; see (2) with β_0, $\boldsymbol{\beta}$, $\mathbf{B}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\Delta}$ replaced by their bootstrapped estimates. It also gives $\widehat{\sigma}_{y_b^*}$; see (3) where σ_ε^2 is replaced by the estimate computed from the bootstrapped parameters using 7.

In our second approach we use Kriging metamodels instead of linear regression RSM models, and we apply *nonparametric* bootstrapping to the output data $w_{i;j}$ $(i = 1, \ldots, n_d; j = 1, \ldots, n_e)$; i.e., we resample—with replacement—the n_e vectors $\mathbf{w}_j$, which gives $\mathbf{w}_j^*$ (this bootstrapping is analogous to the bootstrapping in case of Common Random Numbers; see Kleijnen, van Beers, and van Nieuwenhuyse (2009)). Analogous to (8) and (9) we compute the n_d bootstrapped averages and variances:

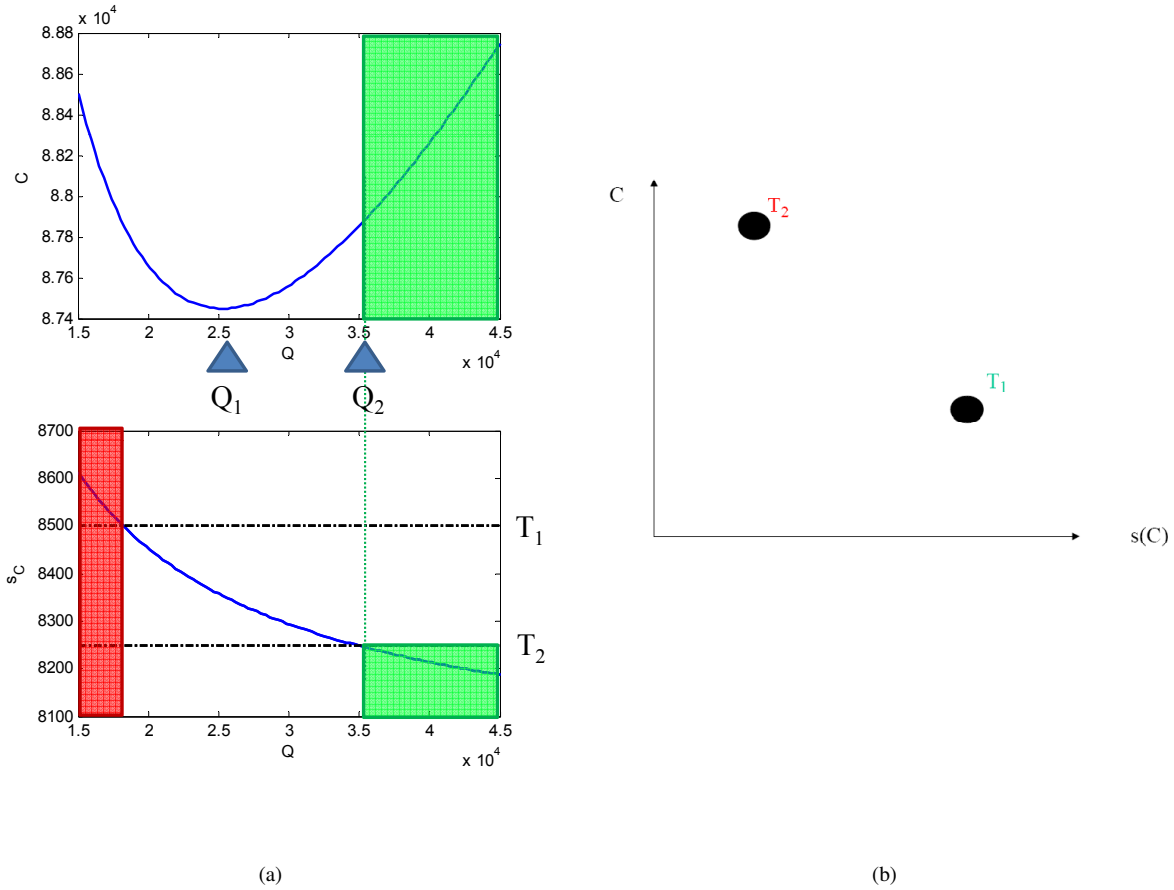$$\overline{w_i^*} = \frac{\sum_{j=1}^{n_e} w_{ij}^*}{n_e} \qquad i = 1, \ldots, n_d, \tag{10}$$

Figure 2: (a) The average and the standard deviation of cost versus the order quantity, and the role of the threshold. (b) Two points on the Pareto curve.

$$s^2(w_i^*) = \frac{\sum_{j=1}^{n_e} \left(w_{ij}^* - \overline{w_i^*}\right)^2}{n_e - 1} \qquad i = 1, \ldots, n_d. \tag{11}$$

Then we fit two Kriging metamodels to these bootstrapped output data and their corresponding input data, $\mathbf{d}_i$. To reduce the sampling error, we repeat this sampling B times, which gives the bootstrapped averages and variances $\overline{w_{i;b}}^*$ and $s^2(w_{i;b}^*)$ and their Kriging metamodels ($b = 1, \ldots, B$). Figure 3 gives results for the EOQ example, discussed in the following section. In this figure we select the B points for the predicted mean that correspond with Q_2; see left part of this figure. We do the same for the standard deviation in the right part of this figure. More precisely, we derive a conservative confidence interval using Bonferroni's inequality and the classic bootstrap confidence interval per output. This interval is $\left[\widehat{\overline{w}}_{(\lfloor B(\alpha/2)/2 \rfloor)}^{+*}, \widehat{\overline{w}}_{(\lceil B(1-(\alpha/2))/2 \rceil)}^{+*}\right]$ where the subscript () denotes the order statistic (i.e., the observations are ranked or sorted from smallest to largest), $\lfloor \rfloor$ denotes the floor function (which gives the integer part), $\lceil \rceil$ denotes the ceiling function (rounding upwards), $\alpha/2$ gives a two-sided confidence interval, Bonferroni's inequality implies that the type-I error rate for the interval per output is divided by the number of outputs (which is two, namely, the mean and standard deviation). For the EOQ example, this gives Figure 5, discussed below.

In Approach 3, we replace the simulation output values $w_{i;j}$ by the corresponding Kriging predictions $\hat{y}_{i;j}$, and bootstrap these observations as described before.

Note that our bootstrap procedure is a fast alternative to macroreplicates, which can be obtained as follows. Repeat the LHS sampling (say) L times; i.e., sample new values for the environmental factors from their distribution, while keeping the
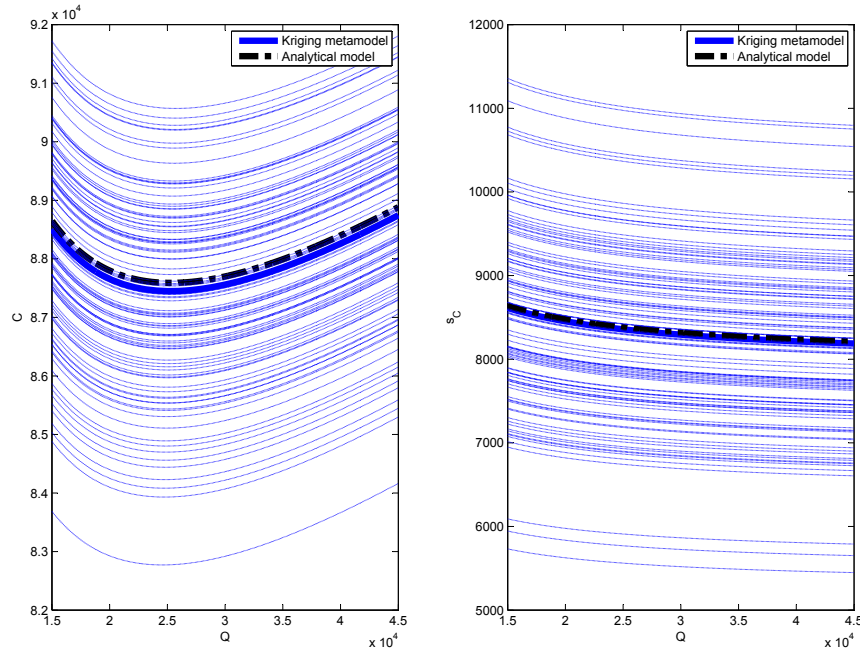
Figure 3: Bootstrapped metamodel predictions of the mean and standard deviation of the cost for Approach 2 (i.e., Kriging variant 1)

values of the decision variables fixed. Then, apply RSM or Kriging to the simulation I/O data corresponding with each of the *L* designs.

3 EXAMPLE: THE ECONOMIC ORDER QUANTITY MODEL

We illustrate our three robust simulation-optimization approaches through a deterministic EOQ model (Zipkin 2000). Starting from the classic model assumptions, we add the assumption that the demand rate is *uncertain*. Now the goal is to estimate an optimal order quantity, which minimizes the total cost while satisfying a given constraint (threshold) for the standard deviation of those costs.

More specifically, we keep the focus on a single-period inventory model; i.e., we have to select the order quantity value such that it faces the demand before the process begins, and keep it constant throughout the observed period. Note that such a model is useful to describe several service and manufacturing applications; e.g., ordering of fashion items (Chase, Jacobs, and Aquilano 2006).

In Taguchian terminology, we consider the order quantity *Q* as the decision factor, while the demand rate *a* is the environmental factor; all the other parameters—shown in column 1 of Table 1—are supposed to be fixed and known constants. To select these parameters, we use a famous textbook, namely Hillier and Lieberman (2001); see the column denoted by EOQ_HL in Table 1. Because this ratio of the setup cost and the unit holding cost is very large in this column, we also consider a different ratio, denoted by EOQ_CJA, taken from Chase et al. (2006).

Table 1: Nominal values of the parameters in the EOQ example

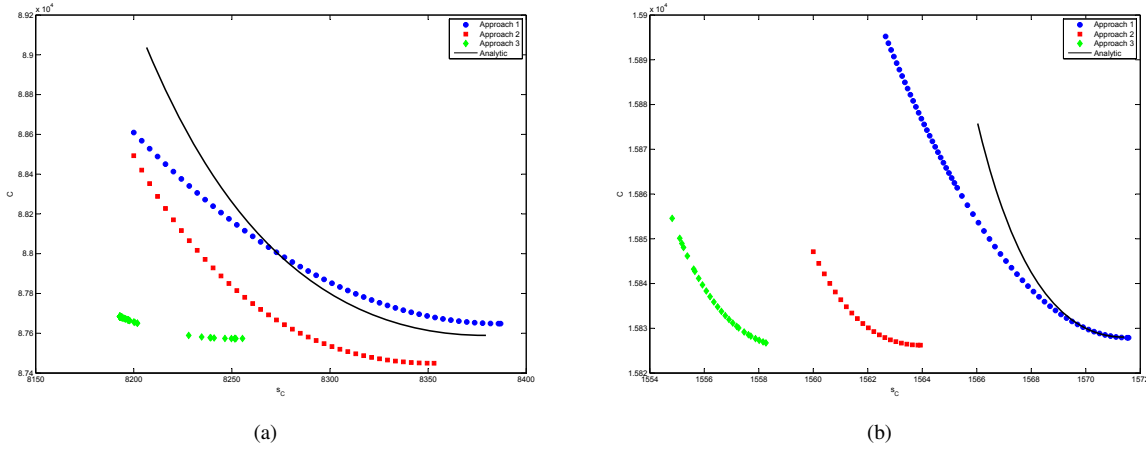| Parameter	Symbol	EOQ_HL	EOQ_CJA
demand rate	a	8000	1040
setup cost	K	12000	10
unit ordering cost	c	10	15
unit holding cost	h	0.30	2.50

Figure 4: Pareto frontiers based on three approaches, and analytical solution: (a) EOQ_HL in Table 1; (b) EOQ_CJA in Table 1.

We assume—without loss of generality—that the demand rate a follows a Normal (Gaussian) distribution with mean μ_a and standard deviation σ_a:

$$a \sim \mathcal{N}(\mu_a, \sigma_a),\tag{12}$$

where μ_a denotes the nominal value used in the classic EOQ model without uncertainty (see Table 1) and σ_a quantifies the uncertainty about the true input parameter. We experiment with three alternative environmental settings, each representing a different degree of volatility: $\sigma_a = k_a \mu_a$, where $k_a \in \{0.50; 0.10; 0.01\}$. Because the qualitative results turn out to be very similar, we will focus on the intermediate case, $k_a = 0.10$.

We apply the steps sketched in Figure 1 for our three approaches to both EOQ_HL and EOQ_CJA. After deriving the metamodels for the mean of the costs C and the cost's standard deviation $s(C)$, we solve the following optimization problem using the Matlab solver `fmincon`:

$$\min \overline{C}$$
$$\text{s.t. } s(C) \leq T\tag{13}$$

where the threshold T is given. Next we change it S times within a range of values properly chosen to avoid infeasible problems (in practice, management would give the T values). The set of S optimal solutions provides an estimate of the Pareto frontier; see Figure 4.

We compare the optimal solution estimated through the three metamodels with the optimal solution found through mathematical analysis, given the same threshold value T; namely, $T = 8600$ for EOQ_HL and $T = 1580$ for EOQ_CJA (such an analysis is possible, because the EOQ model is so simple; the analytical solution is given in Dellino et al. 2008). The comparison concerns the decision variable Q^+, the objective function $\overline{C}^+$ and constrained $s^+(C)$, using the Average Absolute Relative Error (AARE); see Table 2. This table implies that in our EOQ example Approach 1 gives the minimum errors—except for the objective function $\overline{C}^+$ in the EOQ_CJA case.

Table 2: AARE of estimated optimal solutions based on three approaches for EOQ_HL and EOQ_CJA

	EOQ_HL				EOQ_CJA		
	Approach 1	Approach 2	Approach 3		Approach 1	Approach 2	Approach 3
$AARE(Q^+)$	0.6426	0.6805	0.6660	$AARE(Q^+)$	0.9921	0.9940	0.9933
$AARE(\overline{C}^+)$	0.0010	0.0025	0.0025	$AARE(\overline{C}^+)$	0.0011	1.73e-7	0.0001
$AARE(s^+(C))$	0.9050	0.9052	0.9066	$AARE(s^+(C))$	0.9009	0.9012	0.9016

Finally, we apply bootstrapping to derive confidence regions, as we discussed in Section 2.3; see Figure 5. This figure which demonstrates that the bigger sample in Approach 3 gives smaller regions than Approach 2; of course, this approach
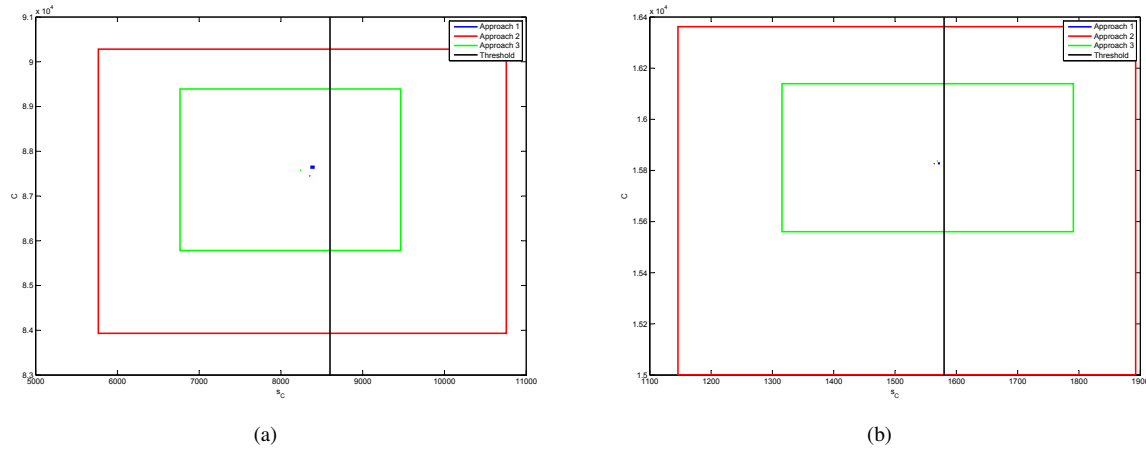
(a)

(b)

Figure 5: Confidence regions for $s(C)$ (on x-axis) and $\overline{C}$ (on y-axis) based on bootstrapped metamodels for (a) EOQ_HL, at $T = 8600$ (see the vertical line) and (b) EOQ_CJA at $T = 1580$ (see the vertical line).

assumes that the Kriging model does not give biased predictions. From the figure, we conclude that RSM (approach 1) gives much smaller variability than Kriging (approaches 2 and 3). We think that the reason is that the RSM metamodel imposes a specific type of model, namely, the low-order polynomial defined by equation (1); so the spread arises only from perturbations of the estimated parameters in (1). Kriging gives a better metamodel when (1) does not give a valid metamodel.

4   CONCLUSIONS AND FURTHER RESEARCH

Whereas *optimization* is a 'hot' topic in the simulation literature, *robust* optimization is neglected in that literature—except for a few publications; see the five references in Kleijnen 2008, p. 131, which in turn reference several more simulation studies using Taguchi's methods.

To solve this problem, we integrate the following scientific approaches:

- *Taguchi*'s view of the uncertain world; i.e., we assume that the inputs of the simulation model that represent the environmental factors do not have fixed values but have a distribution.
- *DOE and metamodeling*; i.e., we replace Taguchi's statistical techniques by techniques that have become popular in simulation; e.g., LHS for the environmental factors and either RSM's low-order polynomial metamodels or Kriging metamodels.
- *Mathematical Programming*; i.e., we replace Taguchi's scalar loss function by a Mathematical Programming model that minimizes the expected simulation output such that the standard deviation of the simulation output does not exceed a given threshold. Varying this threshold gives the *Pareto* frontier, which enables management to select a solution (combination of decision factors) that accounts for their risk attitude.
- *Bootstrapping*; i.e., we derive a (conservative, rectangular)) confidence region for the estimated mean and standard deviation of the simulation output, given management's choice of a particular combination of decision factors. The corresponding confidence interval for the standard deviation quantifies the probability of exceeding the threshold.

We illustrated our new methodology through the well-known *EOQ* model. The following topics deserve *further research*.

- Is *RSM*'s simple low-order polynomial metamodel an adequate approximation, or is more flexible *Kriging* needed? Which Kriging variant is better? Which *designs* are best; e.g., the crossed design is popular among Taguchians, but *split plot* designs are proposed by Dehlendorff, Kulahci, and Andersen (2008).
- The *Mathematical Programming model* may need revision; e.g., the uncertainty of the simulation output may be quantified through a specific *quantile* (instead of the mean and standard deviation). Kleijnen, Pierreval, and Zhang (2009) use a lower bound for the 5% quantile of the simulation output. Instead of a single simulation output (characterized through its mean, standard deviation, quantiles) practical simulations have *multiple* simulation outputs; e.g., cost and service rate in inventory simulations.

548

- Besides the simple EOQ example, *more examples and case studies* are needed to derive general conclusion about our robust simulation-optimization methodology.
- Besides *deterministic* simulation with uncertain inputs (like the EOQ simulation), we are investigating *discrete-event* simulation (like the (s, S) inventory and the $M/M/1$ queueing simulations). By definition, the latter simulation type has *aleatory* uncertainty (e.g., random demand per day, exponential interarrival and service times) besides the *epistemic* uncertainty (e.g., unknown mean demand per day, and unknown arrival and service rates).

REFERENCES

Beyer, H. G., and B. Sendhoff. 2007. Robust optimization — a comprehensive survey. *Computer Methods in Applied Mechanics and Engineering* 196 (33–34): 3190–3218.

Box, G. E. P., and K. B. Wilson. 1951. On the experimental attainment of optimum conditions. *Journal Royal Statistical Society, Series B* 13 (1): 1–38.

Chase, R. B., F. R. Jacobs, and N. J. Aquilano. 2006. *Operations management for competitive advantage*. 11th ed. McGraw Hill.

Chernick, M. R. 2008. *Bootstrap methods: A guide for practitioners and researchers*. 2nd ed. John Wiley & Sons.

Dehlendorff, C., M. Kulahci, and K. Andersen. 2008. Designing simulation experiments with controllable and uncontrollable factors. In *Proceedings of the 2008 Winter Simulation Conference*, ed. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 2909–2915. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

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

Dellino, G., J. P. C. Kleijnen, and C. Meloni. 2008, July. Robust optimization in simulation: Taguchi and response surface methodology. CentER Discussion Paper 2008-69, CentER - Tilburg University. Under review.

Dellino, G., J. P. C. Kleijnen, and C. Meloni. 2009. Robust optimization in simulation: Taguchi and Krige combined. Working paper, Tilburg University. In preparation.

Dellino, G., P. Lino, C. Meloni, and A. Rizzo. 2009, April. Kriging metamodel management in the design optimization of a CNG injection system. *Mathematics and Computers in Simulation* 79 (8): 2345–2360.

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

Fu, M. C. 2007. Are we there yet? the marriage between simulation & optimization. *OR/MS Today* 34 (3): 16–17.

Hillier, F. S., and G. J. Lieberman. 2001. *Introduction to operations research*. 7th ed. McGraw Hill.

Kleijnen, J. P. C. 2008. *Design and analysis of simulation experiments*. New York: Springer Science + Business Media.

Kleijnen, J. P. C. 2009. Sensitivity analysis of simulation models. In *Encyclopedia of Operations Research and Management Science*, ed. J. Cochran, L. Cox, P. Keskinocak, J. Kharoufeh, and J. Smith. New York: Wiley.

Kleijnen, J. P. C., H. Pierreval, and J. Zhang. 2009. Methodology for determining the acceptability of given design in uncertain environments. CentER Discussion Paper 2009-03, Tilburg University.

Kleijnen, J. P. C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa. 2005. State-of-the-art review: a user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing* 17 (3): 263–289.

Kleijnen, J. P. C., W. C. M. van Beers, and I. van Nieuwenhuyse. 2009. Constrained optimization in simulation: novel approach. *European Journal of Operational Research*. In press.

Miettinen, K. M. 1999. *Nonlinear multiobjective optimization*, Volume 12 of *International Series in Operations Research & Management Science*. Boston: Kluwer Academic Publishers.

Myers, R. H., A. I. Khuri, and W. H. Carter. 1989. Response surface methodology: 1966-1988. *Technometrics* 31 (2): 137–157.

Myers, R. H., and D. C. Montgomery. 2002. *Response surface methodology: Process and product optimization using design experiments*. 2nd ed. New York: John Wiley & Sons.

Nair, V. N., B. Abraham, J. MacKay, J. A. Nelder, G. Box, M. S. Phadke, R. N. Kacker, J. Sacks, W. J. Welch, T. J. Lorenzen, A. C. Shoemaker, K. L. Tsui, J. M. Lucas, S. Taguchi, R. H. Myers, G. G. Vining, and C. F. J. Wu. 1992, May. Taguchi's parameter design: A panel discussion. *Technometrics* 34 (2): 127–161.

Park, G. J., T. H. Lee, K. H. Lee, and K. H. Hwang. 2006, January. Robust design: An overview. *AIAA Journal* 44 (1): 181–191.

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

Wu, C. F. J., and M. Hamada. 2000. *Experiments: planning, analysis, and parameter design optimization*. New York: John Wiley & Sons.

Zipkin, P. H. 2000. *Foundations of inventory management*. McGrawHill.

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 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, computer simulation methodologies, and Multidisciplinary Design Optimization. She is a member of INFORMS. In 2008 she has been visiting the CentER Graduate School, at Tilburg University.

JACK P.C. KLEIJNEN is Professor of 'Simulation and Information Systems' at Tilburg University. He is a member of the 'Department of Information Management' and the 'Operations Research' group of CentER. His research concerns simulation, especially its statistical design and analysis. He has been a consultant for several organizations in Europe and the US. He serves on many international editorial boards and scientific committees. He spent some years in the US, at universities and private companies. 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. 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.