Int. J. Production Economics 125 (2010) 52–59

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

Robust optimization in simulation: Taguchi and Response Surface Methodology

Gabriella Dellino [a,*], Jack P.C. Kleijnen [a], Carlo Meloni [b]

[a] Department of Information Management/CentER, Tilburg School of Economics and Management, Tilburg University, Postbox 90153, 5000 LE Tilburg, Netherlands
[b] Department of Electrical Engineering and Electronics, Polytechnic of Bari, Via E. Orabona 4, 70125 Bari, Italy

A R T I C L E I N F O

Article history:
Received 20 August 2008
Accepted 5 December 2009
Available online 16 December 2009

Keywords:
Pareto frontier
Bootstrap
Latin hypercube sampling

A B S T R A C T

Optimization of simulated systems is tackled by many methods, but most methods assume known environments. This article, however, develops a 'robust' methodology for uncertain environments. This methodology uses Taguchi's view of the uncertain world, but replaces his statistical techniques by Response Surface Methodology (RSM). George Box originated RSM, and Douglas Montgomery recently extended RSM to robust optimization of real (non-simulated) systems. We combine Taguchi's view with RSM for simulated systems. We illustrate the resulting methodology through classic Economic Order Quantity (EOQ) inventory models, which demonstrate that robust optimization may require order quantities that differ from the classic EOQ.

© 2010 Published by Elsevier B.V.

1. Introduction

The major purpose of this article is to present a *methodology* for *robust* simulation–optimization. Simulation is much applied in production economics. The goal of these simulations is often the optimization of the real production system being simulated. 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. Therefore we derive our methodology, which combines (i) Taguchi's view of the world, and (ii) RSM. The Taguchian worldview has been very successful in production engineering (see the references below). Nevertheless, statisticians have criticized Taguchi's statistical techniques (see Nair, 1992). (Our example shows that the classic EOQ and our robust EOQ differ nearly 25% if the managers are risk-averse.) Therefore Myers and Montgomery (1995) combine the Taguchian worldview with RSM; RSM has already built a track record in the 'classic' (non-robust) optimization of real-life (non-simulated) system. We adapt Myers and Montgomery's robust RSM to account for the particularities of simulation–optimization; i.e., whereas in real-life experiments it is hard to vary a factor over many values, this restriction does not apply in simulation experiments (we use Latin Hypercube Sampling (LHS), which has as many values per factor as it has combinations; see Section 4.2). So a change of mindset of the simulation experimenter is necessary (also see Kleijnen et al., 2005). Moreover, we further adapt Myers–Montgomery's RSM; e.g., we add bootstrapping (a statistical technique), Mathematical Programming, and Pareto frontiers.

More precisely, RSM uses low-order polynomial regression metamodels (metamodels are also called response surfaces, surrogates, emulators, auxiliary models, repromodels, etc.). These metamodels run much faster than the—possibly computationally expensive—simulation models. RSM was introduced by Box and Wilson (1951) as an iterative heuristic for optimizing real (non-simulated) systems. RSM was further developed for robust optimization of such systems by Myers and Montgomery (1995). We use a less restrictive assumption; i.e., we replace Myers and Montgomery's (1995, p. 493) assumption for the environmental variables $\mathbf{e}$ (namely $E(\mathbf{e}) = \mathbf{0}$ and $\mathbf{cov}(\mathbf{e}) = \sigma_e^2\mathbf{I}$) by a more general and realistic assumption (namely, $E(\mathbf{e}) = \boldsymbol{\mu}_\mathbf{e}$ and $\mathbf{cov}(\mathbf{e}) = \boldsymbol{\Omega}_\mathbf{e}$). When the noise factors do not have constant variances (so $\mathbf{cov}(\mathbf{e}) \neq \sigma_e^2\mathbf{I}$), then Myers and Montgomery (1995, p. 504) point out that confidence intervals for the robust optimum become complicated; to solve this problem, we use parametric bootstrapping. To find a 'robust' solution, Myers and Montgomery (1995, p. 504) superimpose contour plots for the mean and variance of the output, whereas we use more general and flexible Mathematical Programming, to minimize the mean output such that the output variance remains below a given threshold. Our Mathematical Programming approach, however, requires specification of threshold values that managers may find hard to quantify; we therefore try different values, and estimate the

* Corresponding author.
E-mail addresses: G.Dellino@uvt.nl, dellino@dii.unisi.it (G. Dellino), meloni@deemail.poliba.it (C. Meloni).
URL: http://center.uvt.nl/staff/kleijnen (J.P.C. Kleijnen).

0925-5273/$ - see front matter © 2010 Published by Elsevier B.V.
doi:10.1016/j.ijpe.2009.12.003

corresponding Pareto frontier showing a trade-off between the output and its variability. To estimate the variability of this Pareto frontier, we use bootstrapping. Our methodology has the following particularities. Simulation experiments (unlike real-life experiments) enable the exploration of many values per input and many combinations of these values; whereas Myers and Montgomery (1995, pp. 463–534) use only two values per environmental factor, we use LHS. Instead of the F lack-of-fit test used in classic and robust RSM, we use leave-one-out cross-validation. Notice that we assume that the decision variables are continuous (like Myers and Montgomery, 1995, p. 486 do).

In practical applications, simulation may be computationally expensive; e.g., a single run took 36 to 160 hours of computer time to simulate a crash model at Ford (see Simpson et al., 2004). In such applications, the extra computer time needed for fitting RSM metamodels, followed by bootstrapping and Mathematical Programming is negligible. We illustrate our methodology through a simple simulation model, namely the EOQ model. Our choice has several advantages: it saves much computer time; its results can be verified through the analytical EOQ solution (in practice, simulation is used because there is no known analytical solution); yet this EOQ simulation enables us to illustrate some details of our methodology. (The EOQ model is closely related to the Economic Production Quantity or EPQ (see Darwish, 2008). The EOQ model is also a building block for realistic supply chain simulations.)

The rest of this article is organized as follows. Section 2 summarizes Taguchi's worldview. Section 3 summarizes and extends Myers and Montgomery's (1995) approach that uses RSM for robust optimization. Section 4 illustrates our new methodology through the classic EOQ simulation. Section 5 presents our conclusions and possible topics for future research. Our online companion paper, Dellino et al. (2008), gives more details and additional references enabling further study of robust simulation–optimization.

2. Taguchi's worldview

Based on Kleijnen (2008, pp. 130–137), we summarize the extensive literature on *Taguchi*'s view as follows. Taguchi (1987) distinguishes between two types of variables: (i) decision (or control) factors, which we denote by d_j $(j=1,\ldots,k)$, and (ii) environmental (or noise) factors, e_g $(g=1,\ldots,c)$. Taguchi assumes a single output (say) w. He focusses on the mean and the variance of this output. By definition, the decision factors are under the control of the users; e.g., in inventory management, the order quantity is controllable. The environmental factors are not controlled by the users; e.g., in inventory management, the demand rate may not be controllable.

In this article, we use Taguchi's view, but not his statistical methods; instead we use RSM. Our main reason is that simulation experiments enable the exploration of many more factors, factor levels, and combinations of factor levels than real-life (physical) experiments do. Moreover, we do not use a scalar *Taguchian loss function* such as the signal-to-noise or mean-to-variance ratio; instead we allow each output to have a statistical distribution, which we characterize through its mean and standard deviation. We solve the resulting problem through the Pareto-optimal efficiency frontier—briefly called the *Pareto frontier*. Also see Beyer and Sendhoff (2007), Lee and Nelder (2003), Myers and Montgomery (1995, p. 491), Park et al. (2006), Wu et al. (2009); the latter authors focus on the mean–variance trade-off in the newsvendor's inventory problem.

3. RSM and robust optimization

In their robust RSM Myers and Montgomery (1995) use a polynomial of a degree as low as possible. They fit a *second-order* polynomial for the decision factors d_j, to estimate the optimal combination of these factors. To model possible effects of the environmental factors e_g, they fit a first-order polynomial for these factors. To estimate interactions between the two types of factors, they fit 'control-by-noise' two-factor interactions. Altogether, they fit:

$$\begin{aligned} y &= \beta_0 + \sum_{j=1}^{k}\beta_j d_j + \sum_{j=1}^{k}\sum_{j'\geq j}^{k}\beta_{j;j'}d_j d_{j'} + \sum_{g=1}^{c}\gamma_j e_j + \sum_{j=1}^{k}\sum_{g=1}^{c}\delta_{j;g}d_j e_g + \varepsilon \\ &= \beta_0 + \boldsymbol{\beta}'\mathbf{d} + \mathbf{d}'\mathbf{B}\mathbf{d} + \boldsymbol{\gamma}'\mathbf{e} + \mathbf{d}'\boldsymbol{\Delta}\mathbf{e} + \varepsilon, \end{aligned} \tag{1}$$

where y denotes the regression predictor of the output w, ε the residual with $E(\varepsilon)=0$ if this metamodel has no *lack of fit* and with constant variance σ^2_ε, $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_k)'$, $\mathbf{d}=(d_1,\ldots,d_k)'$, $\mathbf{B}$ 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 $\boldsymbol{\Delta}=(\delta_{j;g})$.

Design Of Experiments (DOE) uses *coded*—also called *standardized* or *scaled*—factor values (say) x_j. So the experiment consists of n factor combinations of the 'original' factors (say) z_j, which correspond with d_j and e_g in (1). Coding is also discussed in Kleijnen (2008, p. 29).

Myers and Montgomery (1995, p. 493) assume that the environmental variables satisfy $E(\mathbf{e})=\mathbf{0}$ and $\mathbf{cov}(\mathbf{e})=\boldsymbol{\sigma}^2_e\mathbf{I}$. We, however, use the more general and realistic assumption $E(\mathbf{e})=\boldsymbol{\mu}_\mathbf{e}$ and $\mathbf{cov}(\mathbf{e})=\boldsymbol{\Omega}_\mathbf{e}$. Analogous to Myers and Montgomery (1995), we then derive that this metamodel (1) implies the regression predictor for the true mean $E(w)$

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

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

$$var(y)=(\boldsymbol{\gamma}'+\mathbf{d}'\boldsymbol{\Delta})\boldsymbol{\Omega}_\mathbf{e}(\boldsymbol{\gamma}+\boldsymbol{\Delta}'\mathbf{d})+\sigma^2_\varepsilon=\mathbf{l}'\boldsymbol{\Omega}_\mathbf{e}\mathbf{l}+\sigma^2_\varepsilon \tag{3}$$

where in (3) $\mathbf{l}=(\boldsymbol{\gamma}+\boldsymbol{\Delta}'\mathbf{d})=(\partial y/\partial e_1,\ldots,\partial y/\partial e_c)'$; i.e., $\mathbf{l}$ is the gradient with respect to the environmental factors—which follows directly from (1). So, the larger the gradient's elements are, the larger the predicted variance of the simulation output is—which stands to reason. Furthermore, if $\boldsymbol{\Delta}=\mathbf{0}$ (no control-by-noise interactions), then $var(y)$ cannot be controlled through the control variables $\mathbf{d}$. Notice the difference between the predicted variance, $var(y)$, and the variance of the predictor, $var(\widehat{y})$ with $\widehat{y}=\widehat{\boldsymbol{\zeta}}'\mathbf{x}$; see (4) and (5).

Eq. (3) implies that the predicted output y has heterogeneous variances, because changing the control factors $\mathbf{d}$ changes $var(y)$ (heterogeneous variances arise even if $\mathbf{cov}(\mathbf{e})=\boldsymbol{\sigma}^2_e\mathbf{I}$). If a particular decision factor has no effects on the mean output but has important interactions with the noise factors, then these interactions can be utilized to decrease the output variance. If there are multiple decision factors (unlike our EOQ example), then we first select the values of some decision factors such that $\mathbf{l}=\mathbf{0}$, so $var(y)$ in (3) is minimized; next we select the remaining decision factors such that the predicted mean output $E(y)$ in (2) gets the desired threshold value. For more details we refer to Myers and Montgomery (1995, p. 494).

Myers and Montgomery (1995, p. 495) also discuss *constrained optimization*, which minimizes (e.g.) the variance in (3) subject to a constraint on the mean in (2). To select an appropriate compromise or 'robust' solution, those authors often superimpose contour plots for the mean and variance. We, however, shall use Mathematical Programming, because it is more general and flexible.

To *estimate* the unknown (regression) parameters in (1) we reformulate (1) as the following *linear regression model*:

$$y = \boldsymbol{\zeta}'\mathbf{x} + \varepsilon \tag{4}$$

with $\boldsymbol{\zeta} = (\beta_0, \boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\gamma}, \boldsymbol{\delta})'$ where $\mathbf{b}$ denotes the vector with the $k \times (k-1)/2$ interactions between the decision factors plus their k purely quadratic effects, and $\boldsymbol{\delta}$ denotes the $k \times c$ control-by-noise interactions, $\mathbf{x}$ is defined in the obvious way, e.g., the element corresponding with the interaction effect $\beta_{1,2}$ is $d_1 d_2$. 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

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

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) determined by DOE that are actually simulated, and q denotes the number of parameters collected in $\boldsymbol{\zeta}$; $\mathbf{w}$ consists of the n simulation outputs. The covariance matrix of this estimator is

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

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

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

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

Furthermore, we assume that $\mathbf{y}$ is *normally* distributed; i.e., $\mathbf{e}$ and ε in (1) are normally distributed. We can then test the estimated regression parameters $\widehat{\zeta_j}$ through the following t statistic with $n-q$ degrees of freedom:

$$t_{n-q} = \frac{\widehat{\zeta_j} - \zeta_j}{s(\widehat{\zeta_j})} \text{ with } j = 1, \ldots, q \tag{8}$$

where $s(\widehat{\zeta_j})$ is the square root of the j^{th} element on the main diagonal of (6) with σ_w^2 estimated through (7). Myers and Montgomery (1995, p. 488) keep only the significant effects in their response model. It is well-known that this test is not very sensitive to nonnormality. We agree that when estimating the robust optimum, we should use the *reduced* metamodel, which eliminates all non-significant effects in the full model—except for those non-significant effects that involve factors that have significant higher-order effects; see the 'strong heredity' assumption in Wu and Hamada (2000). For example, if the estimated main effect $\widehat{\beta_1}$ is not significant but the estimated quadratic effect $\widehat{\beta_{1,1}}$ is, then $\widehat{\beta_1}$ is not set to zero.

To construct *confidence intervals* for the robust optimum, Myers and Montgomery (1995, p. 498) assume normality and derive an F statistic. Myers and Montgomery (1995, p. 504) notice that the analysis becomes complicated when the noise factors do not have constant variances. We shall therefore use *parametric bootstrapping* in the EOQ examples; by definition, parametric bootstrapping assumes that the distribution of the relevant random variable is known; e.g., in the EOQ examples, the distribution of the unknown demand rate is modified normal (to avoid negative demand rates). In general, bootstrapping is a simple numerical method for obtaining the estimated density function of a—possibly complicated—statistic for a—possibly non-Gaussian—parent distribution; e.g., our Mathematical Programming solution (output) of the EOQ problem is non-normal, and the input distribution is modified normal. Also see Efron and Tibshirani (1993) and Kleijnen (2008, p. 86).

Like Myers and Montgomery (1995, p. 495) we simply plug in the LS estimators (5) for β_0, $\boldsymbol{\beta}$, $\mathbf{B}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\Delta}$ in the right-hand side of (2); the factors $\mathbf{d}$ and $\boldsymbol{\mu}_\mathbf{e}$ are known. To estimate (3), we again plug in the estimators for $\boldsymbol{\gamma}$, $\boldsymbol{\Delta}$, and σ_ε^2; $\boldsymbol{\Omega}_\mathbf{e}$ is known. However, we point out that (3) has products of unknown parameters, so it implies a *nonlinear estimator* $\widehat{\sigma_y^2}$. Plugged-in estimators certainly create bias; this bias we ignore when estimating the Pareto frontier that balances $\widehat{y}$ and $\widehat{\sigma_y}$. To study the variability of this estimated Pareto frontier caused by estimating the regression parameters, we use bootstrapping.

Analogous to Myers and Montgomery (1995, pp. 41–54) we wonder whether the RSM model (1) *adequately* approximates the true Input/Output (I/O) function implicitly defined by the underlying simulation model. There are several methods for answering this question (see Kleijnen, 2008, p. 54). Following Kleijnen (2008, p. 57), we use *leave-one-out cross-validation*:

Step 1: Delete I/O combination i from the complete set of n combinations, to obtain $(\mathbf{X}_{-i}, \mathbf{w}_{-i})$. Assume that $\mathbf{X}_{-i}$ is not collinear; a necessary condition is $n > q$.

Step 2: Recompute the LS estimator from the I/O data in Step 1:

$$\widehat{\boldsymbol{\zeta}_{-i}} = (\mathbf{X}_{-i}'\mathbf{X}_{-i})^{-1}\mathbf{X}_{-i}'\mathbf{w}_{-i}. \tag{9}$$

Step 3: Use $\widehat{\boldsymbol{\zeta}_{-i}}$ from Step 2 to compute $\widehat{y_{-i}}$, the regression predictor of the simulation output generated by $\mathbf{x}_i$ which is the simulation input of the combination deleted in Step 1:

$$\widehat{y_{-i}} = \mathbf{x}_i'\widehat{\boldsymbol{\zeta}_{-i}}. \tag{10}$$

Step 4: Repeat the preceding three steps, until all n combinations have been processed. This gives $\widehat{y_{-i}}$ $(i = 1, \ldots, n)$.

Step 5: To judge whether the metamodel is adequate, use a scatterplot with the n pairs $(w_i, \widehat{y_{-i}})$.

Step 6: Because the scaling of this scatterplot may give the wrong impression, also evaluate the relative prediction errors $\widehat{y_{(-i)}}/w_i$.

Step 7: Examine the recomputed effects $\widehat{\boldsymbol{\zeta}_{-i}}$ $(i = 0, 1, \ldots, n)$ where $\widehat{\boldsymbol{\zeta}_{-0}}$ denotes the estimator when zero combinations are deleted (so $\widehat{\boldsymbol{\zeta}_{-0}} = \widehat{\boldsymbol{\zeta}}$); these effects should not change much if the regression model is adequate.

The final goal of our robust optimization is to minimize the estimated mean $\widehat{y}$ while keeping the estimated standard deviation $\widehat{\sigma_y}$ below a given threshold (say) T. We solve this constrained minimization problem through a Mathematical Programming solver. We decide to use Matlab's `fmincon`, but a different solver may be used (Gill et al., 2000). This gives the values of the 'estimated robust decision variables' (say) $\widehat{\mathbf{d}^+}$ and the corresponding estimated mean $\widehat{y}$ and standard deviation $\widehat{\sigma_y}$. Next, we vary the threshold value T (say) 100 times. These changes may give different solutions $\widehat{\mathbf{d}^+}$ with their corresponding $\widehat{y}$ and $\widehat{\sigma_y}$. These pairs $(\widehat{y}, \widehat{\sigma_y})$ enable us to estimate the Pareto frontier. We estimate the variability of this frontier through bootstrapping of the estimated regression estimates that gave $\widehat{y}$ and $\widehat{\sigma_y}$. This methodology is illustrated in the next section.

4. EOQ inventory simulation

The EOQ inventory model is often used in practical supply chain management. First, we define this model, including symbols and assumptions (also see Pentico et al., 2009; Teng, 2008). We use the following assumptions, following Zipkin (2000, pp. 30–39): (i) The demand is a constant a units per time unit. (ii) The order quantity is Q units. (iii) No shortages are allowed. (iv) Delivery lead time is zero. (v) Review is continuous. (vi) Total costs consist of setup cost per order, K; cost per unit purchased or produced, c; holding cost per inventory unit per time unit, h. The goal is to minimize the costs per time unit (say) C, over an infinite time horizon. It is easy to derive that this problem has the following *true I/O function*, which we shall use to check our simulation results:

$$C = \frac{aK}{Q} + ac + \frac{hQ}{2}. \tag{11}$$

Differentiation of (11) shows that the true EOQ is

$$Q_o = \sqrt{\frac{2aK}{h}}, \tag{12}$$

and the corresponding cost is

$$C_o = C(Q_o) = \sqrt{2aKh} + ac. \tag{13}$$

In our illustration we use the parameter values in a famous textbook; namely, Hillier and Lieberman (2001, pp. 936–937, 942–943): $a = 8000$, $K = 12\,000$, $c = 10$, and $h = 0.3$; such a high cost K argues for buying products in large batches, leading to orders approximately placed once every three periods (see for further comments, Hillier and Lieberman, 2001). So (12) gives $Q_o = 25\,298$ and (13) gives $C_o = 87589$. (We also run another numerical example with a smaller ratio between the ordering cost coefficient K and the unit holding cost coefficient h; namely, an example from Chase et al. (2006) with $a = 1040$, $K = 10$, $c = 15$, $h = 2.50$. For this example the benefits of the robust optimal solution are even more evident; e.g., if the manager is risk averse, then the robust optimum differs approximately 25% from the estimated classic optimum, and the corresponding mean cost differs approximately 57% from the estimated classic cost.)

4.1. Classic simulation optimization of the EOQ model

We start the simulation run with an inventory of Q units. Our simulation experiment consists of the following four steps.

Step 1 (design): To select the experimental area, we start with the interval $[0.5\,Q_o, 1.5\,Q_o]$. This selection, however, would imply that the midpoint coincides with the true optimum input $Q_o = 25298$. We therefore shift the interval a little bit (namely, by < 5000 units) to the right so that it is centered at $Q = 30\,000$. Furthermore, we pick five equally spaced points, including the extreme points; namely, the lowest point $0.5 \times 30\,000 = 15\,000$ and the highest point $1.5 \times 30\,000 = 45\,000$; see row 1 of Table 1. The input parameters are fixed to their base (nominal) values ($a = 8000$, etc.). Note that a Central Composite Design (CCD), which is popular in RSM, would also have five points—albeit not equally spaced; see Dellino et al. (2008) and Myers and Montgomery (1995, p. 55).

Step 2 (simulation model): Simulation gives $C(Q_i) = C_i$, the cost corresponding with input value i ($i = 1, \ldots, 5$) selected in Step 1; see the I/O combinations (Q_i, C_i) displayed in Table 1.

Step 3 (RSM metamodel): Based on the I/O data resulting from Steps 1 and 2, we estimate a second-order polynomial. We focus on the effects of the <u>coded</u> decision variable, because the resulting effects (say) $\widehat{\boldsymbol{\beta}} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_{1,1}})'$ show the relative importance of the intercept, the main effect, and the quadratic effect; moreover, their numerical accuracy is better. This $\widehat{\boldsymbol{\beta}}$ is displayed in the row with $i = 0$ in Table 2.

Step 4 (cross-validation): The remaining rows of Table 2 display the re-estimated regression parameters following from (9), and the re-estimated regression prediction following from (10). This table also presents the relative prediction errors $\widehat{y_{(-i)}}/C_i$, which supplement the scatterplot in Fig. 1. The first four columns of Table 2 show that the estimated regression coefficients remain more or less the same.

When we apply the t statistic defined in (8), we see that the estimated main effect is <u>not</u> significantly different from zero; the estimated quadratic effect $\widehat{\beta_{1,1}}$ is. Because of the 'strong heredity' assumption discussed below (8), we do not replace the estimated main effect by zero. Notice that the estimated effects are not independent, because $(\mathbf{X}'\mathbf{X})^{-1}$ in (6) is not diagonal.

The estimated optimum (say) $\widehat{Q_o}$ follows from the first-order optimality condition $\partial\widehat{C}/\partial Q = \widehat{\beta_1} + 2\widehat{\beta_{1,1}}x_1 = 0$, where x_1 is the coded variable corresponding with Q. This condition gives $\widehat{Q_o} = 28\,636$. This $\widehat{Q_o}$ gives the estimated minimal cost $\widehat{C_o} = 87\,654$. In this simple example, we can verify the estimated optimum: $\widehat{Q_o}/Q_o = 28\,636/25\,298 = 1.13$ and $\widehat{C_o}/C_o = 87\,654/87\,589 = 1.001$ so the cost virtually equals the true minimum, even though the input is 13% off. This illustrates the well-known insensitivity property of the classic EOQ formula.

We also experiment with a smaller experimental area; i.e., a smaller Q range. The Taylor series suggests that this smaller area gives a better approximation. The smaller Q range indeed gives a more accurate metamodel; the resulting estimated optimum is only 1% below the true EOQ and the corresponding cost virtually equals the true cost (see Dellino et al., 2008).

4.2. Robust optimization of the EOQ model

In this subsection, we still assume that the demand per time unit is a constant a but this constant is unknown. Uncertainty in parameters of inventory management is addressed in Borgonovo and Peccati (2007) and Yu (1997). We assume—without loss of generality—that a has a Normal (Gaussian) distribution with mean μ_a and standard deviation σ_a: $a \sim N(\mu_a, \sigma_a)$. Furthermore, we assume that μ_a denotes the 'base' value used in the classic simulation–optimization in Section 4.1 (so $\mu a = 8000$), and σ_a quantifies the uncertainty about the true value of this input

Table 1
I/O data of EOQ simulation.

Q	15 000	22 500	30 000	37 500	45 000
C	88 650	87 641.66	87 700	88 185	88 883.34

Table 2
Cross-validation of EOQ regression metamodel.

i	$\widehat{\beta_{0(-i)}}$	$\widehat{\beta_{1(-i)}}$	$\widehat{\beta_{1,1(-i)}}$	$\widehat{y_{(-i)}}$	$\widehat{y_{(-i)}}/C_i$
0	87 663.4257	202.004	1097.15		
1	87 731.998	522.008	640	87 849.94	0.991
2	87 769.82	139.94	1008.49	87 952.11	1.004
3	87 628.88	202.004	1137.79	87 628.92	0.999
4	87 583.63	155.46	1163.64	87 951.95	0.997
5	87 603.997	479.34	1493.34	89 576.98	1.008

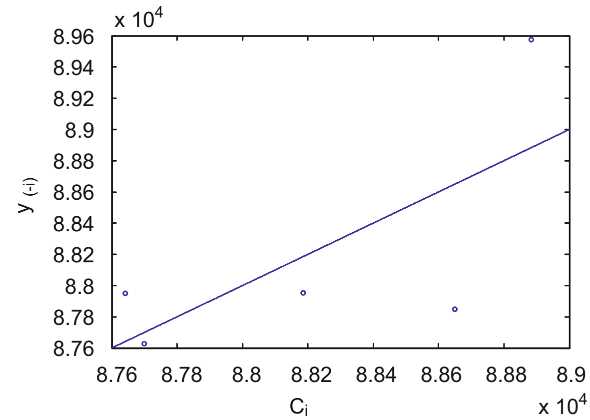

Fig. 1. Scatterplot for regression metamodel of EOQ simulation model.

parameter. We experiment with a 'low' and 'high' uncertainty: $\sigma_a = 0.10$ and $0.50\,\mu_a$. Because these standard deviations can give a negative value for a, we resample until we get non-negative values only; i.e., we adjust the normal distribution slightly. This adjustment, however, is ignored in our further analysis.

In their robust RSM, Myers and Montgomery (1995, pp. 463–534) use only two values per environmental factor; this suffices to estimate its main effect and its interactions with the decision factors. We, however, use LHS to select (say) $n_e = 5$ values for the environmental factor because LHS is popular in risk and uncertainty analysis (see Kleijnen, 2008). In our EOQ example, we split the range of possible a values ($0 < a < \infty$) into five equally likely subranges. We use `lhsnorm` from the Matlab Statistics Toolbox to select these five values from $N(\mu_a, \sigma_a)$; see The MathWorks Inc. (2005) and the first row in Table 3, which uses the relatively high uncertainty $\sigma_a = 0.50\mu_a$. Results for the smaller uncertainty $\sigma_a = 0.10\mu_a$ are presented in Dellino et al. (2008).

For the decision variable Q we select the five values that we also used in Table 1; see the first column in Table 3. We *cross* the two designs for a and Q respectively, as is usual in a Taguchian approach. (Nevertheless, we could also have used LHS to get a combined design for a and Q. Dellino et al. (2008) use a CCD instead of LHS to get a combined design. Myers and Montgomery (1995, p. 487) also discuss designs that are more efficient than crossed designs.)

We run the EOQ simulation model for all 5×5 combinations of the inputs (decision and environmental inputs), which gives the other entries of Table 3.

To analyze the I/O data of Table 3, we might compute the estimated *conditional* variance $\widehat{var(C|Q_i)}$ from the row with Q_i ($i = 1, \ldots, 5$) (also see Lee and Nelder, 2003). Instead we follow Myers and Montgomery (1995) and estimate the variance from all the elements in this table, using (3). The latter approach gives a better estimator, provided the RSM metamodel (1) is correct.

To compute the LS estimates in the RSM model, we must re-arrange the 5×5 elements of Table 3 into the $n \times q$ $\mathbf{X}$- matrix of (5) with $n = 25$ and $q = 5$; $\mathbf{w}$ is the vector with the 25 simulation outputs C. This gives the estimated intercept $\widehat{\beta_0} = 88\,150.40$, the estimated first-order effect $\widehat{\beta_1} = 190.56$ and the second-order effect $\widehat{\beta_{1,1}} = 1058.33$ of Q, the estimated first-order effect $\widehat{\gamma_1} = 36\,774.03$ of a, and the interaction $\widehat{\delta_{1,1}} = -899.67$. Dellino et al. (2008) also display the cross-validation results (analogous to Table 2) and the scatterplot (analogous to Fig. 1). Their results suggest that this metamodel is adequate for robust optimization through RSM. To verify the negative sign of $\widehat{\delta_{1,1}}$ (this estimate is -899.67), we use the analytical solution (11) to derive $\partial^2 C/\partial Q \partial a = -K/Q^2$, which is indeed negative.

Using a similar RSM metamodel for their example, Myers and Montgomery (1995, p. 501) derive *contour plots* for the mean and variance. Because our EOQ example has a single decision variable, we do not superimpose contour plots but present Figs. 2 and 3. Fig. 2 shows the $(Q, \widehat{C})$ plot; see (2) with the regression parameters replaced by their estimates. Fig. 3 shows the $(Q, \widehat{\sigma_C})$ plot. We prefer to use the standard deviation instead of the variance, because the former uses the same scale as the simulated cost C and its regression estimate $\widehat{C}$. We use (3) with $\boldsymbol{\gamma}$, $\boldsymbol{\Delta}$, and σ_ε^2 replaced by their estimates, including (7). This figure shows a second-order polynomial that resembles a linearly decreasing function in the relatively small domain of Q. For this simple example we know the true I/O function of the simulation model—namely (11)—so we derive the true standard error of the cost C:

$$\sigma_C = \sigma\left(\frac{aK}{Q} + ac + \frac{hQ}{2}\right) = \sigma\left(\frac{hQ}{2} + \left[\frac{K}{Q} + c\right]a\right) = \left(\frac{K}{Q} + c\right)\sigma_a = c\sigma_a + \frac{K\sigma_a}{Q}. \tag{14}$$

We also plot this σ_C against Q in Fig. 3. Comparing the two curves in this figure, we conclude that the estimated curve is an adequate approximation.

From Figs. 2 and 3 we derive the 'estimated robust optimal' order quantity (say) $\widehat{Q^+}$, which we define as the quantity that minimizes the estimated mean $\widehat{C}$ while keeping the estimated standard deviation $\widehat{\sigma_C}$ below a given threshold T. We solve this constrained minimization problem through Matlab's fmincon''. For example, if $T = 42\,500$, then Fig. 3 implies $\widehat{Q^+} = 28\,568$. However, when T becomes smaller (e.g., $T = 41\,500$) then Fig. 3 implies $\widehat{Q^+} = 35\,222$; see Fig. 4, in which the curve becomes a horizontal line with 'height' $\widehat{Q^+}$ if the threshold is high enough.

Section 4.1 gave the classic estimated EOQ—namely $\widehat{Q_o} = 28\,636$—assuming that the demand rate equals the nominal value. Now we assume different demand rates. The corresponding

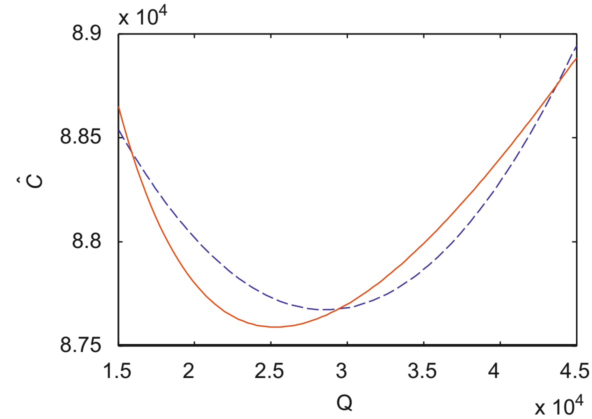

Fig. 2. Estimated (dashed) curve and true (solid) curve for mean cost versus order quantity.

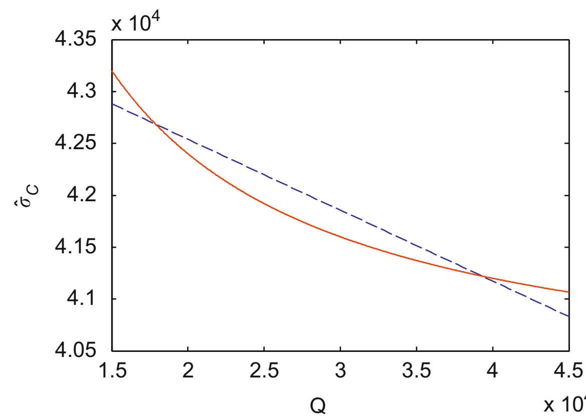

Fig. 3. Estimated (dashed) curve and true (solid) curve for standard deviation of cost versus order quantity.

Table 3
I/O simulation data for EOQ model with uncertain demand rate.

Q	a				
	4530.34	5478.85	7687.37	9329.26	11 559.02
15 000	51 177.72	61 421.54	85 273.65	103 006	127 087.4
22 500	51 094.63	61 085.52	84 348.68	101 643.2	125 130
30 000	51 615.59	61 480	84 448.7	101 524.3	124 713.8
37 500	52 378.16	62 166.7	84 958.71	101 902.9	124 914.1
45 000	53 261.54	62 999.49	85 673.71	102 530.4	125 422.6

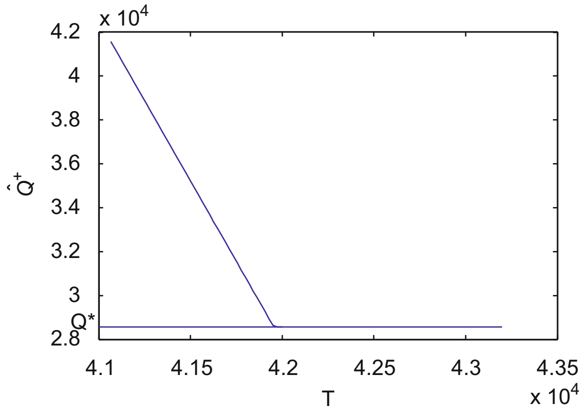

Fig. 4. Estimated robust optimal value for EOQ against threshold for standard deviation of cost.

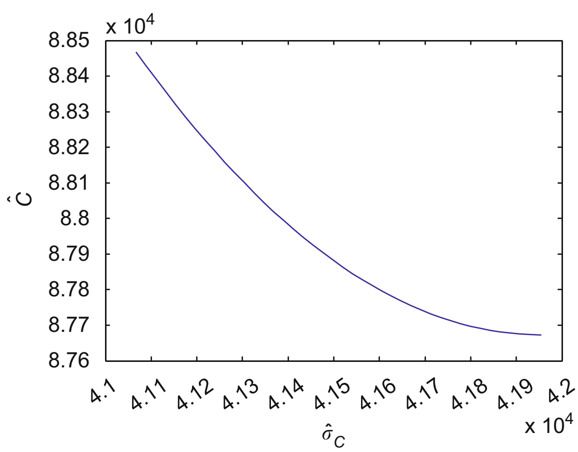

Fig. 5. Estimated Pareto frontier for EOQ simulation with threshold for standard deviation of cost.

model gives an estimated optimal order quantity $\widehat{Q^+}$ that differs from $\widehat{Q_o}$. This difference is nearly 25% if the managers are risk-averse (low threshold T).

We assume that management cannot give a single, fixed value for the threshold. Therefore we vary the threshold over the interval [41 067, 43 200]. This interval gives the estimated *Pareto frontier* in Fig. 5. This figure demonstrates that if management prefers low costs variability, then they must pay a price; i.e., the expected cost increases.

We repeat the experiment with a smaller σ_a, which implies a less volatile environment. Some reflection shows that we cannot keep the threshold values T the same in environments with different magnitudes of *volatility*. The new threshold values give a new estimated Pareto frontier. Dellino et al. (2008) report that a less volatile world gives lower mean cost. Their result quantifies the benefits of obtaining more information on the uncertain demand rate; e.g., such information may be provided by a marketing survey that decreases σ_a.

The estimated Pareto frontier is built on estimates only; namely $\widehat{\zeta}$ using (5). We therefore analyze this frontier further. Whereas Myers and Montgomery (1995, pp. 496–503) use rather complicated confidence intervals, we use *parametric bootstrapping*. So we sample (say) B times from the q-variate normal distribution with mean vector and covariance matrix given by (5) and (6), where we use the superscript * for bootstrapped values:

$$\widehat{\zeta}^* \sim N_q(\widehat{\zeta}, (\mathbf{X}'\mathbf{X})^{-1}\widehat{\sigma_w^2}). \tag{15}$$

This sampling gives $\widehat{\zeta}_b^*$ with $b = 1, \ldots, B$. This $\widehat{\zeta}_b^*$ gives $\widehat{C_b^*}$; see (2) with the regression parameters replaced by their bootstrapped estimates computed from the bootstrapped $\widehat{C_b^*}$. It also gives $\widehat{\sigma_{C_b^*}}$; see (3) where σ_ε^2 is replaced by the estimate computed from the bootstrapped parameters. These two bootstrapped variables $\widehat{C_b^*}$ and $\widehat{\sigma_{C_b^*}}$ give the bootstrapped optimal decision variable $\widehat{Q_b^+}^*$, computed through Matlab's `fmincon`. This bootstrap sample gives the B estimated Pareto frontiers of Fig. 6, with $B = 50$ and the true Pareto frontier derived from the analytical costs (11) and its standard deviation (14), and the original estimated frontier of

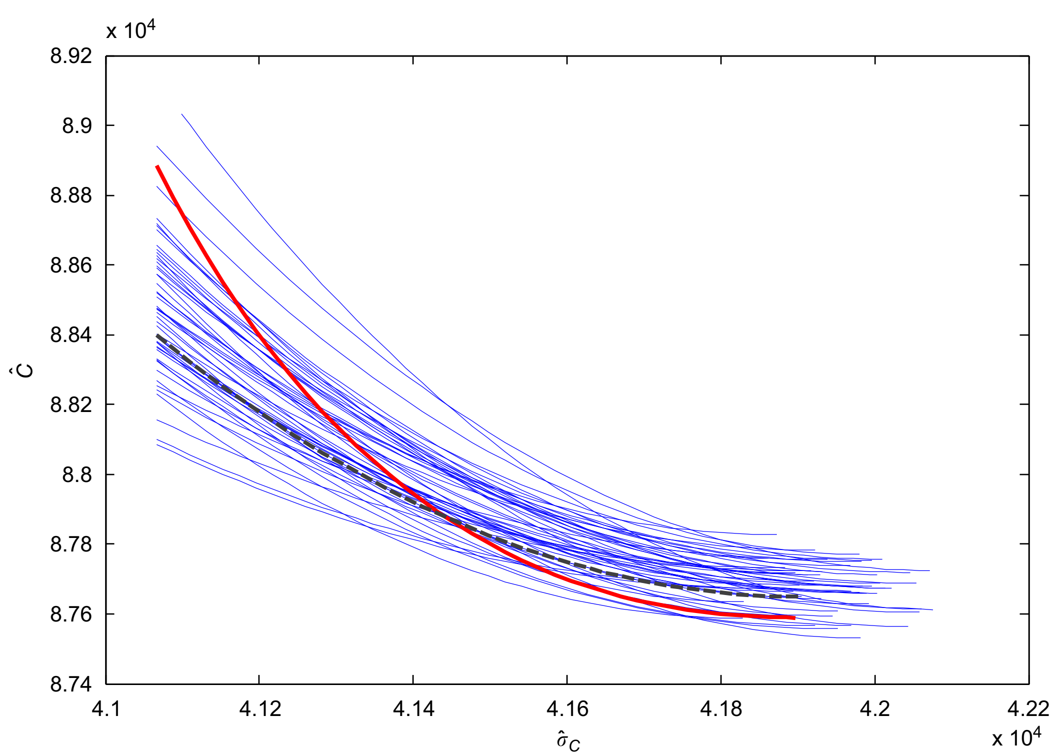

Fig. 6. Bootstrapped Pareto frontiers, original estimated frontier (dashed curve), and true frontier (heavy curve).

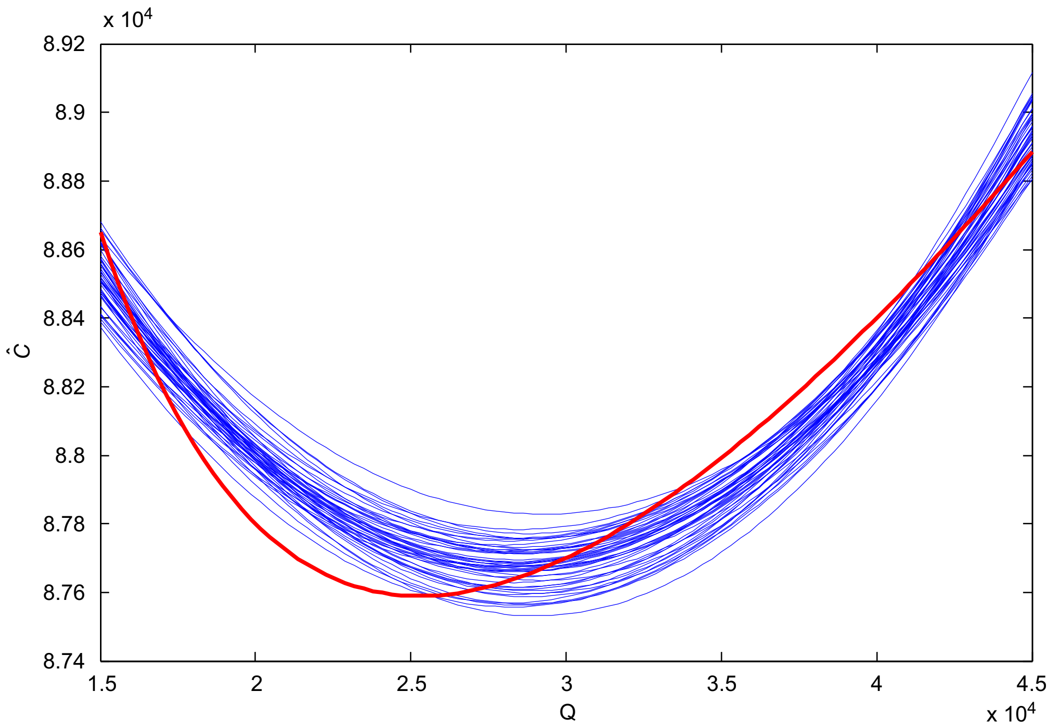

Fig. 7. Bootstrapped estimated costs, and true cost (heavy curve).

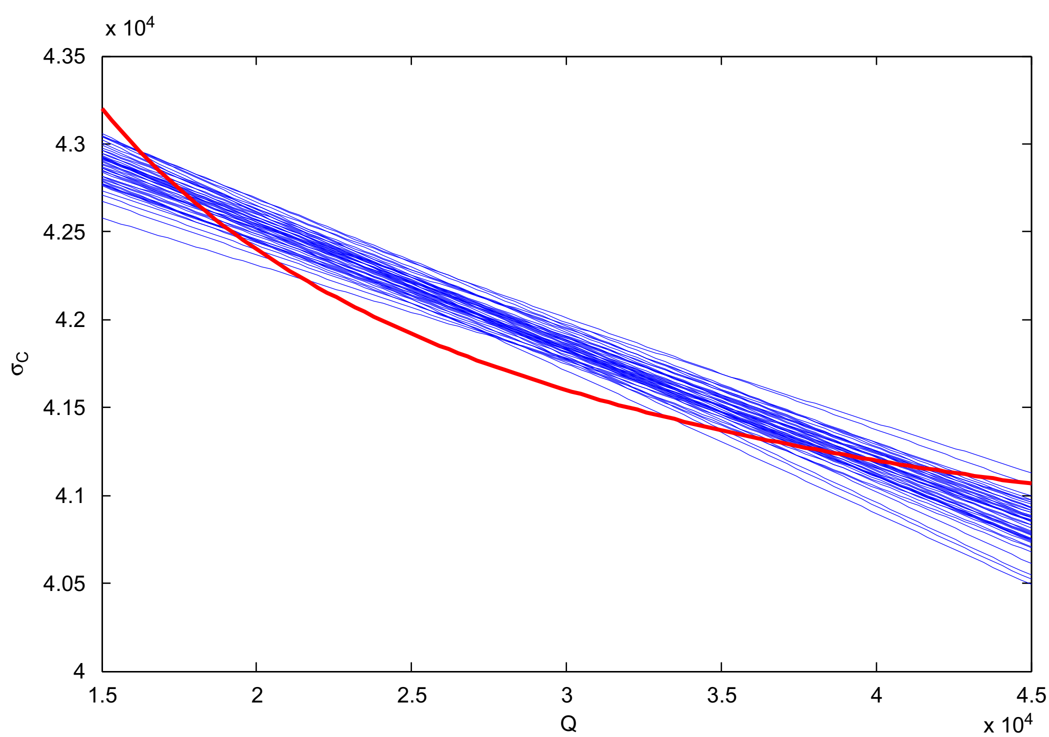

Fig. 8. Bootstrapped standard deviations of the cost, and true standard deviation of the cost (heavy curve).

Fig. 5. Fig. 6 demonstrates that bootstrapping gives a good idea of the variability of the estimated Pareto frontier; the bundle of bootstrapped curves 'envelop' the original estimated curve and the true curve. The bundle of bootstrapped estimated costs does not completely envelop the true curve; neither does the bundle for the bootstrapped standard deviations; see Figs. 7 and 8.

Besides the crossed design for Q and a Dellino et al. (2008) also use a CCD. They report that the CCD with its $n=9$ combinations gives a better estimate of the true frontier than the 5×5 crossed-design does. We conjecture that the bigger (crossed) design gives a more accurate estimator $\widehat{\zeta}$ of the *wrong* (misspecified) metamodel (namely, a second-order polynomial) for the true I/O function implied by the EOQ simulation model; see (11).

5. Conclusions and future research

This article leads to the following conclusions. Robust optimization of simulated systems may use *Taguchi*'s worldview, which distinguishes between decision variables to be optimized and environmental variables that remain uncertain but do affect the optimum. However, Taguchi's statistical techniques may be replaced by *RSM*. Myers and Montgomery's (1995) RSM for Taguchian optimization may be further adapted through bootstrapping, which better enable management to make the final compromise decision. Application of this new methodology to the classic *EOQ* model shows that—for a known environment—the methodology gives a good estimate of the true EOQ, and—for an environment with a demand rate that has a known distribution—the classic EOQ and the robust EOQ differ.

Future research may address the following issues. We conjectured that the bigger crossed design gave a more accurate LS estimator $\widehat{\zeta}$ of a misspecified metamodel. This is a good reason for using a different type of metamodel, namely *Kriging* models (Generalized Linear Models or GLMs are proposed by Lee and Nelder (2003)) as alternatives for RSM models. In a next article we shall present Kriging for robust optimization. Furthermore, we shall adjust our methodology for *discrete-event* simulation models that have so-called aleatory uncertainty; we focus on (s, S) models, with either explicit out-of-stock costs resulting in scalar output or a service constraint resulting in vector output (the difference $S-s$ is often based on the EOQ model)); this adjustment includes a switch from parametric bootstrapping to nonparametric (distribution-free) bootstrapping. Finally, we hope to apply our methodology to complex *supply chain* models (also see Hassini, 2008).

Acknowledgements

Carlo Meloni wishes to thank CentER for the financial support when he visited Tilburg University. We thank Wim van Beers (Tilburg University) for helpful discussions on our research.

We thank the two anonymous reviewers for their valuable comments on two preceding versions, which lead to the present paper. This research has been supported by the PRIN grant of the Italian Ministry of Education, 2007ZMZK5T.

References

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

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

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

Chase, R.B., Jacobs, F.R., Aquilano, N.J., 2006. Operations Management for Competitive Advantage, eleventh ed. McGraw-Hill.

Darwish, M.A., 2008. EPQ models with varying setup cost. International Journal of Production Economics 113 (1), 297–306.

Dellino, G., Kleijnen, J.P.C., Meloni, C., 2008. Robust optimization in simulation: Taguchi and Response Surface Methodology. CentER Discussion Paper, Tilburg University, Tilburg, the Netherlands. Available online via 〈http://center.uvt.nl/staff/kleijnen/papers.html〉 [accessed August 19, 2008].

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

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

Gill, P.E., Murray, W., Wright, M.H., 2000. Practical Optimization, twelfth ed. Academic Press, London.

Hassini, E., 2008. Editorial: supply chain optimization: current practices and overview of emerging research opportunities. INFOR, Information Systems and Operational Research 46 (2), 93–96.

Hillier, F.S., Lieberman, G.J., 2001. Introduction to Operations Research, seventh edition McGraw Hill, Boston.

Kleijnen, J.P.C., 2008. Design and Analysis of Simulation Experiments. Springer, New York.

Kleijnen, J.P.C., Sanchez, S.M., Lucas, T.W., Cioppa, T.M., 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.

Lee, Y., Nelder, J.A., 2003. Robust design. Journal of Quality Technology 35 (1), 2–12.

Myers, R.H., Montgomery, D.C., 1995. Response Surface Methodology: Process and Product Optimization Using Designed experiments. Wiley, New York.

Nair, V.N. (Ed.), 1992. Taguchi's parameter design: a panel discussion. Technometrics 34(2), 127–161.

Park, G-J., Lee, T-H., Lee, K.H., Hwnag, K-H., 2006. Robust design: an overview. AIAA Journal 44 (1), 181–191.

Pentico, D.W., Drake, M.J., Toews, C., 2009. The deterministic EPQ with partial backordering: a new approach. Omega 37 (3), 624–636.

Simpson, T.W., Booker, A.J., Ghosh, D., Giunta, A.A., Koch, P.N., Yang, R.-J., 2004. Approximation methods in multidisciplinary analysis and optimization: a panel discussion. Structural and Multidisciplinary Optimization 27 (5), 302–313.

Taguchi, G., 1987. System of Experimental Designs, vols. 1 and 2. UNIPUB/ Krauss International, White Plains, New York.

Teng, J.-T., 2009. A simple method to compute economic order quantities. European Journal of Operational Research, 198 (1), 351–353.

The MathWorks Inc., 2005. Statistics Toolbox, Natick, Massachusetts.

Wu, C.F.J., Hamada, M., 2000. Experiments Planning; Analysis, and Parameter Design Optimization. Wiley, New York.

Wu, J., Li, J., Wang, S., Cheng, T.C.E., 2009. Mean-variance analysis of the newsvendor model with stockout cost. Omega 37 (3), 724–730.

Yu, G., 1997. Robust economic order quantity models. European Journal of Operational Research 100 (3), 482–493.

Zipkin, P.H., 2000. Foundations of Inventory Management. McGraw-Hill.