Engineering Optimization
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/geno20

A new method for estimating missing values for a genetic algorithm used in robust design

E. Canessa [a], S. Vera [a] & H. Allende [b]
[a] Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, Balmaceda 1625, Recreo, CP, 3132386, Chile
[b] Departamento de Informática, Universidad Técnica Federico Santa María, CP, 110, Valparaíso, Chile

Available online: 07 Dec 2011

To cite this article: E. Canessa, S. Vera & H. Allende (2012): A new method for estimating missing values for a genetic algorithm used in robust design, Engineering Optimization, 44:7, 787-800

To link to this article: http://dx.doi.org/10.1080/0305215X.2011.613464

Engineering Optimization
Vol. 44, No. 7, July 2012, 787–800

A new method for estimating missing values for a genetic algorithm used in robust design

E. Canessa[a]*, S. Vera[a] and H. Allende[b]

[a]*Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, Balmaceda 1625, Recreo, CP 3132386, Chile;* [b]*Departamento de Informática, Universidad Técnica Federico Santa María, CP 110, Valparaíso, Chile*

(*Received 11 January 2011; final version received 21 July 2011*)

This article presents an improved genetic algorithm (GA), which finds solutions to problems of robust design in multivariate systems with many control and noise factors. Since some values of responses of the system might not have been obtained from the robust design experiment, but may be needed in the search process, the GA uses response surface methodology (RSM) to estimate those values. In all test cases, the GA delivered solutions that adequately adjusted the mean of the responses to their corresponding target values and with low variability. The GA found more solutions than the previous versions of the GA, which makes it easier to find a solution that may meet the trade-off among variance reduction, mean adjustment and economic considerations. Moreover, RSM is a good method for estimating the mean and variance of the outputs of highly non-linear systems, which makes the new GA appropriate for optimizing such systems.

Keywords: robust design; Taguchi methods; genetic algorithms; response surface methodology

1. Introduction

To deliver products and services of quality, a firm needs to achieve certain quality standards. Although there are many definitions of quality, and thus associated standards (see Hoyer and Hoyer (2001) for a good discussion), Taguchi's proposals and methods have had a profound impact on improving the production processes at many firms (Allende *et al.* 2005, de Mast 2004, Kackar and Shoemaker 1986). One of the key contributions of Taguchi has been robust design. Robust design is a two-stage method, which tries to set controllable input factors of a production system, so that the outputs of the system stay as stable as possible, and then it adjusts other control factors to bring the mean of the outputs as close as possible to their target values. This is done under different noise conditions (noise factors), so that the settings of the control factors found in the analysis are robust to such noise situations (Taguchi 1991). When an engineer needs to consider many control and noise factors and the system has many responses that must be simultaneously optimized, the typical techniques recommended by Taguchi become difficult to apply (Allende

*Corresponding author. Email: ecanessa@uai.cl

ISSN 0305-215X print/ISSN 1029-0273 online

http://dx.doi.org/10.1080/0305215X.2011.613464
http://www.tandfonline.com

et al. 2005, Maghsoodloo and Chang 2001). Therefore, to overcome such difficulties, a tool based on genetic algorithms (GA) (Holland 1974) was developed, which assists in applying robust design to multivariate systems (Allende *et al*. 2008, Canessa *et al*. 2011). Although this tool and subsequent improvements have worked rather well (Allende *et al.* 2008, Canessa *et al.* 2011), there is still room to refine them. Figure 1 shows a general diagram of the algorithm. In that diagram, note that there is a final adjustments stage in which the algorithm estimates the missing values of

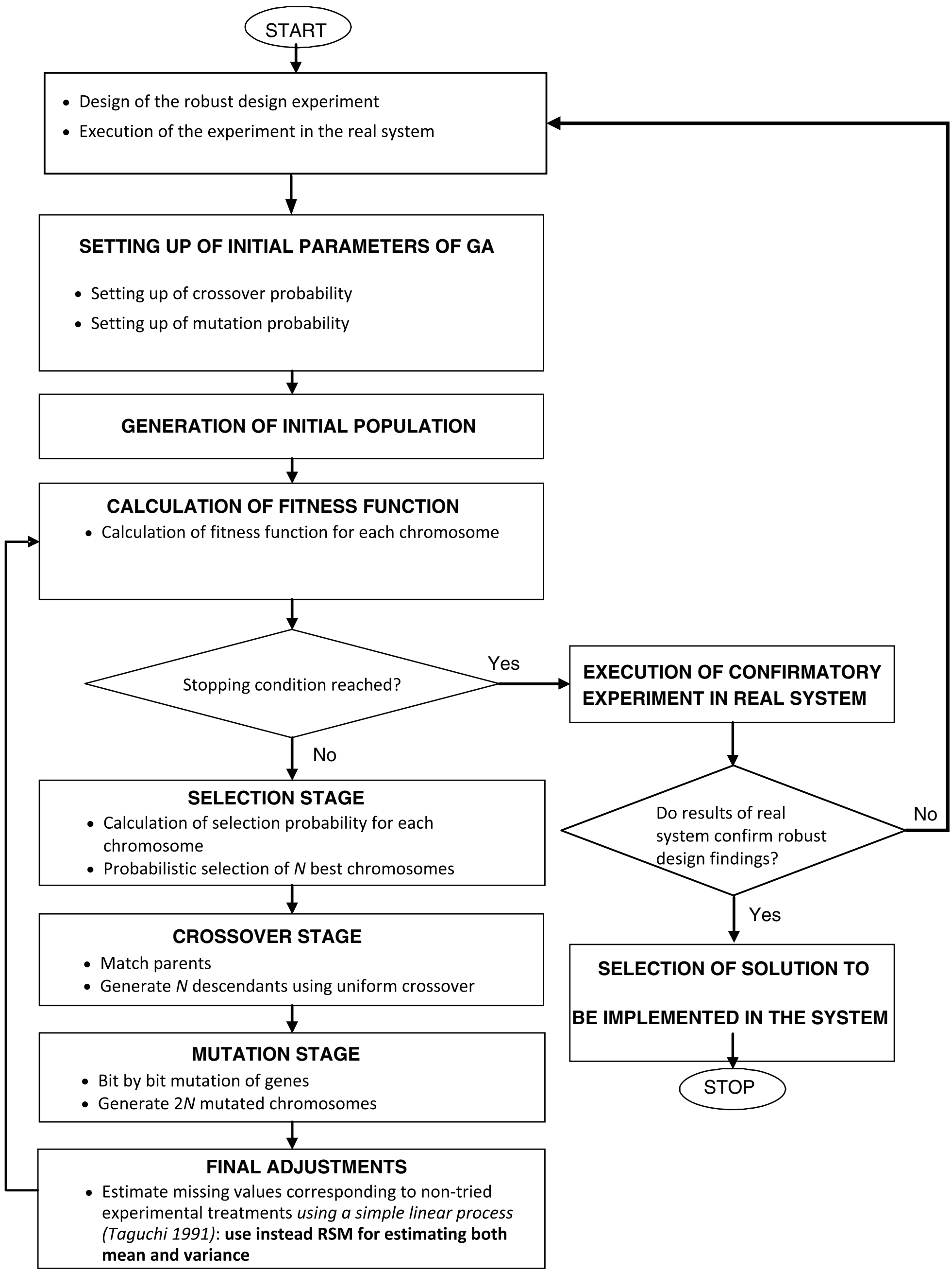

Figure 1. Application of the genetic algorithm to robust design in multivariate systems and proposed changes to the GA (modified from Canessa *et al*. 2011).

the responses corresponding to combinations of control factors (treatments) that might not have been part of the experiment that the engineer conducted to gather the data. Thus, if during the search process of the GA, the GA needs the values of those responses, they should be estimated.

In previous GAs, Allende *et al.* (2008) and Canessa *et al.* (2011) used a simple approximation to calculate the missing values. For estimating the mean and variance of the response for a non-tried treatment, the GAs basically calculated the main effect of each level of each control factor on both mean and variance. Then for estimating the mean of a response the GAs added to the mean of all the responses of the experiment that was carried out, the corresponding main effects. For the variance, the GAs computed a total variance considering all the replications of all the treatments tried in the original experiment and then added to it, the corresponding variance main effects. Both procedures are the usual form of estimating responses in robust design (Taguchi 1991, Yuan 2006). Although those two simple methods have adequately worked, previous studies indicate that response surface methodology (RSM) is a better procedure to accomplish such task (Vining and Myers 1990, Myers and Montgomery 2002, Nair *et al.* 2002). Thus, this article explores the benefits of changing the original method used in previous GAs for estimating missing values of responses, to a new procedure based on RSM.

The rest of the article is organized as follows: Section 2 presents some details of the GA and RSM, which will be needed in order to explain the changes made to the algorithm. Then, Section 3 shows the results obtained with the modified GA, both for univariate and multivariate systems. Additionally, Section 3 compares the results obtained by the previous GA and the outcomes generated by the modified GA. The article ends with a summary of the results and their implications for the use of the modified GA in robust design.

2. Changes to the genetic algorithm using RSM

The following subsections present some parts of the previous GA and concepts related to RSM necessary for understanding the changes made to the GA. More details of the previous GA may be found in Canessa *et al.* (2011) and Allende *et al.* (2008).

2.1. *Some details of the previous GA*

The previous GA developed in Canessa *et al.* (2011) represents the combinations of k control factors that may take s different levels (values) of a robust design experiment using an integer codification. One chromosome will be composed of a combination of different levels for each factor, which corresponds to a particular treatment of the experiment.

Let f_{lj} be the factor j of chromosome l, with $j = 1, 2, \ldots, k$ and $l = 1, 2, \ldots, N$. Each f_{lj} can take the value of a given level of the factor j, that is $1, 2, \ldots, s$. One chromosome (or solution) is expressed as a row vector (see Equation (1)). The matrix representing the total population of solutions X will be composed of N chromosomes (see Equation (2)).

$$x_l = [f_{l1}, f_{l2}, \ldots, f_{lk}] \tag{1}$$

$$X = [x_1, x_2, \ldots, x_N]^T \tag{2}$$

Each of the chromosomes (solutions) x_l will generate a different response y of the system when the control factors are set to the corresponding levels specified in the chromosome x_l. As can be seen in Figure 1, the GA searches through the space of possible treatment combinations, finding the combinations that minimize the variance of the responses and adjust their means as close as possible to their corresponding target values. The fitness function used to guide the GA for each response is the following:

$$Fitness\ function = \phi(x_l) = -\{s^2(x_l) + [T - \bar{y}(x_l)]^2\} \quad (3)$$

Where:

- T is the target value for nominal the better (NTB) quality characteristic;
- $\bar{y}(x_l) = \frac{1}{n}\sum_{i=1}^{n} y_i(x_l)$ mean of the response $y(x_l)$ of chromosome x_l;
- $s^2(x_l) = \frac{1}{n}\sum_{i=1}^{n}[y_{li} - \bar{y}(x_l)]^2$ variance of the response $y(x_l)$ of chromosome x_l.

Then, using the fitness value calculated by Equation (3), the GA computes a desirability and penalty function $D_l(\phi(x_l) - P_l(x_l))$, which normalizes the value of $\phi(x_l)$ to the $[0, 1]$ interval taking into account the maximum and minimum possible values and variation that the output of the system should have (for further details, see Allende *et al.* 2008). Additionally, since the system may have r responses corresponding to a multivariate system ($r = 1, 2, 3, \ldots, R$), the GA aggregates each response desirability, corresponding to $D_l(\phi_r(x_l)) - P_{lr}(x_l)$), into a single one using the work of Ortiz *et al.* (2004) and Del Castillo *et al.* (1996), which consists in employing an aggregated desirability function $D*_l(\phi(x_l))$:

$$D_l^*(\phi(x_l)) = D(D_l(\phi_r(x_l)) - P_{lr}(x_l)) \quad (4)$$

In Equation (4), $D*_l(x_l)$ represents the aggregated desirability of chromosome x_l over all the system's responses, which the GA uses as the aggregated fitness of x_l. $D*_l(\phi(x_l))$ is calculated as the aggregated desirability $D_l(\phi_r(x_l))$ of each of the r individual fitness functions of the responses of chromosome x_l, minus the penalty function $P_{lr}(x_l)$ of each of the penalties corresponding to the r responses of the chromosome x_l. More details regarding how the GA computes Equation (4) may be found in Allende *et al.* (2008).

From expressions (3) and (4), one can see that in the calculation of the fitness the GA needs to know the responses corresponding to the experimental treatment, which each chromosome represents. However, some of those treatments might not have been part of the experiment that the engineer conducted to gather the data. Thus, the GA needs to estimate those responses. For estimating the mean of a response for a non-tried chromosome (treatment), the GA calculates the main effect of each of the treatment levels on the response and a grand mean using all the observations corresponding to the experiment that was carried out. Then, the GA adds to the grand mean, the corresponding main effects of the levels indicated by the chromosome. For estimating the variance of the response for a non-tried chromosome (treatment), the GA uses a similar procedure. The GA first computes a global variance considering all the replications of all the treatments tried in the original experiment. Then the GA calculates the main effect of each control factor on the variance. Finally, the GA sums the main effects of the levels indicated in the chromosome to the global variance. Those procedures correspond to a linear estimation usually applied in the Taguchi method (Taguchi 1991, Yuan 2006, for a worked out numerical calculation, see for example Taguchi 1991, p. 16).

In the search process, the GA uses roulette or probabilistic selection, a uniform crossover and a bit by bit (factor by factor) mutation operator. To be able to compare the performance of the original GA with the new one, the stopping criterion of reaching between 3000 and 4000 chromosomes in the population was kept. For the same reason, the new GA continued using a crossover probability of 0.3 and a mutation probability of 0.05.

2.2. *RSM and its application to the estimation of missing values in the GA*

Using techniques that are part of response surface methodology (RSM), Vining and Myers (1990) proposed the use of second order polynomials to adjust a surface that may represent the mean and variance of a system and employ that surface in robust design experiments. Further analysis of that approach indicated that it provides many benefits (for a good discussion see Nair *et al.* 2002 and Myers and Montgomery 2002). Vining and Myers (1990), recommended the following expressions for estimating the mean and variance of the output of a system, based on the vector of k control factors (X):

$$m(X) = \beta_0 + \sum_{p=1}^{k} \beta_p x_p + \sum_{p=1}^{k} \beta_{pp} x_p^2 + \sum_{p<q}^{k} \beta_{pq} x_p x_q + \varepsilon_m \tag{5}$$

$$v(X) = \gamma_0 + \sum_{p=1}^{k} \gamma_p x_p + \sum_{p=1}^{k} \gamma_{pp} x_p^2 + \sum_{p<q}^{k} \gamma_{pq} x_p x_q + \varepsilon_v \tag{6}$$

Expression (5) corresponds to the response surface that estimates the mean of the response and Equation (6) approximates the variance of it. Note that both expressions assume that each control factor x_p is a quantitative variable. However, it may be the case that some experiments might involve one or more qualitative variables. In such case, the researcher will need to represent those qualitative variables using indicator variables (Myers and Montgomery 2002). Accordingly, expressions (5) and (6) must be modified to accommodate the use of indicator variables as follows:

$$m(X) = \beta_0 + \sum_{p=1}^{k} \beta_p x_p + \sum_{i=1}^{n} \delta_i z_i + \sum_{p=1}^{k} \beta_{pp} x_p^2 + \varepsilon_m \tag{7}$$

$$v(X) = \gamma_0 + \sum_{p=1}^{k} \gamma_p x_p + \sum_{i=1}^{n} \eta_i z_i + \sum_{p=1}^{k} \gamma_{pp} x_p^2 + \varepsilon_v \tag{8}$$

where $z_1, z_2, \ldots, z_i$ represent n indicator variables and the response surface assumes that there are no interactions between indicator and quantitative variables and among quantitative variables. If that occurs, the experimenter must add some terms to Equations (7) and (8), but commonly the researcher tries to avoid including those interactions (Myers and Montgomery 2002), which agrees with Taguchi's suggestions for conducting robust design experiments (Taguchi 1991, Roy 2001). The Appendix shows the application of expressions (5) through (8) to the specific test cases used in this article, and thus helps to understand such equations.

To estimate the response surfaces characterized by expressions (5) and (6) or (7) and (8), the ordinary least squares (OLS) regression method is generally used (Myers and Montgomery 2002). Thus, the modified GA applies specific versions of expressions (5) through (8) and implements an OLS routine to estimate the response surfaces corresponding to the mean and variance of each output. Then, the GA uses those surfaces to approximate the mean and variance of an output, if such values are not directly available from the data collected from the robust design experiment that was carried out. Since the response surface for the variance is only an estimation of the true and unknown surface, it could happen that a missing value of variance estimated by the surface might be inappropriate, *i.e.* negative. That may occur because the GA is estimating a value outside the range of the values of the control factors used in calculating the regression, *i.e.* it is extrapolating, or simply because the estimated surface is not totally representative of the behaviour of the variance of the system. Therefore, the algorithm verifies whether the estimated variance is negative and, if that is the case, it estimates the variance again using the original approach.

Summarizing Section 2, the present proposal consists in changing the simple and linear estimation process for the mean and variance of responses corresponding to non-tried experimental treatments implemented in the original GA, by a new and more accurate method based on RSM.

3. Application of the modified GA

To evaluate the performance of the modified GA (GA2) and also compare it to the previous version (GA1) (Canessa *et al.* 2011), two case studies were used. The first one corresponds to a real application of robust design to adjust the automatic body paint process in a car manufacturing plant. The second case study uses a multivariate process simulator with four responses, ten control factors and five noise factors. This simulator is described in Canessa *et al.* (2011) and was used to test GA1.

3.1. *Results obtained for the univariate real system*

In this case, a robust design experiment was carried out to adjust the width of the painted strip of a car painting system to a nominal width of 40.0 cm. The design of the experiment consisted of an orthogonal array $L_9(3^4)$ for the four control factors and a $L_4(2^3)$ for the three noise factors. More details and the data may be found in Vandenbrande (2000) and Vandenbrande (1998). The Appendix presents the estimation of the response surfaces for the mean and variance of the system. 30 experimental runs using GA1 and GA2 were performed. The best solution obtained by GA2 is the same as the one generated by GA1 and corresponds to chromosome $x = [2\ 3\ 1\ 2]$, *i.e.* combination of control factors A = 2, B = 3, C = 1 and D = 2. This solution appeared in 100% of the 30 runs and since it was found by both algorithms it is highly reliable. For that combination, the width of the painted strip is 41.025 cm with a standard deviation of 1.439 cm. Note that since the system has one response to be optimized and only four control factors and three noise factors, the same solution can be manually computed. However, the simplicity of this test case allows the correct functioning of the algorithm to be verified.

As can be seen in Figure 1, and following Taguchi's recommendation, it is suggested that after the engineer completes the robust design study, he/she carry out a confirmatory experiment in the real system (Taguchi 1991). This experiment allows the engineer to assess whether the conclusions of the robust design analyses hold in the real system, *i.e.* whether the combination of values for the control factors actually produce a mean of the output close enough to the target value and with an acceptable low variation. If that is the case, the engineer can then select the solution that will be implemented in the system, taking into account mean adjustment, variance reduction and other considerations, for example, necessary cost and time for implementing each solution. In this particular case study, the description of results allows concluding that the best solution found by the GA2 indeed holds for the real system (Vandenbrande 2000). However, the engineers did not select the best solution ([2 3 1 2]), but solution [2 3 1 3]. This latter solution is close to the best one, but is less expensive to implement (Vandenbrande 2000). Another point that the engineers considered when making that decision was that they wanted to focus more on variance reduction than on mean adjustment, which the solution [2 3 1 3] better helped to accomplish. For more details, see Vandenbrande (2000).

For this univariate simple system, note that the surfaces that represent the mean and variance of the system have a high R^2 (all above 0.89, see last paragraph of section A1.2), which indicates a very good approximation of the estimated surface to the real system. For that system, GA2 delivers very good solutions; in fact it finds the best solution one can manually compute. Thus, it can be said that when the surfaces calculated by RSM are a good approximation to the real system, it is worth using the new missing value estimation method implemented in GA2.

3.2. *Results obtained for the univariate complex systems*

To test both GAs under a more complex situation, a simulator was built, which is described in detail in Canessa *et al.* (2011). The robust design for this situation considers using an inner array

Table 1. Solutions obtained by GA1 and GA2 for the univariate complex systems.

Responses	Chromosome (solution)	Response mean: Target value	Response mean: Obtained	Response mean: Difference (%)	Standard deviation: Minimum possible (s_{min})	Standard deviation: Obtained (s_{obt})	Standard deviation: s_{obt}/s_{min} (times)
		Previous GA (GA1)					
y_1	3-4-2-1-3-3-1-2-4-2	200	190.50	−4.75	2.1	2.10	1.00
y_1	3-3-1-2-4-3-1-2-4-1	200	197.40	−1.30	2.1	2.10	1.00
y_2	4-3-2-4-1-4-1-3-2-2	50	52.10	4.20	0.8	0.80	1.00
y_2	2-3-4-1-2-1-2-3-4-3	50	47.26	−5.48	0.8	0.80	1.00
y_3	2-2-1-4-3-3-4-1-2-4	1000	1028.80	2.88	36.0	37.10	1.03
y_3	3-3-1-2-4-4-2-1-3-2	1000	981.00	−1.90	36.0	38.10	1.06
y_4	1-4-4-4-4-1-1-1-1-4	500	521.90	4.38	28.0	28.00	1.00
y_4	3-2-4-3-1-2-4-3-1-2	500	474.90	−5.02	28.0	28.00	1.00
		Modified GA (GA2) (new solutions)					
y_1	1-3-3-3-1-1-1-3-4-3	200	196.64	−1.68	2.1	2.10	1.00
y_1	1-1-1-4-4-1-1-1-1-4	200	202.74	1.37	2.1	2.10	1.00
y_2	3-2-3-3-1-3-4-3-1-3	50	50.82	1.64	0.8	0.80	1.00
y_2	1-4-4-4-4-3-3-4-3-3	50	47.25	−5.50	0.8	0.86	1.08
y_3	4-4-4-3-4-1-2-2-1-4	1000	1019.03	1.90	36.0	36.00	1.00
y_3	1-1-4-4-4-2-2-4-4-3	1000	841.47	−15.85	36.0	36.00	1.00
y_4	2-3-3-3-3-3-2-1-2-4	500	478.85	−4.23	28.0	28.00	1.00
y_4	1-1-4-3-1-1-2-4-1-1	500	381.09	−23.78	28.0	28.00	1.00

Note: Solutions listed for GA1, excluding [1-4-4-4-4-1-1-1-1-4], were also delivered by GA2 and are not repeated in the rows that list the new solutions found by GA2.

$L_{64}(4^{10})$ for the 10 control factors and an outer array $L_{16}(4^5)$ for the five noise factors. For the following case studies, the four responses of the simulator are optimized independent from each other, so that both GAs deal with four univariate systems. For each response, both algorithms were run 30 times. The Appendix presents the estimation of the response surfaces for the mean and variance of each of these four univariate systems.

Table 1 presents the mean and standard deviation of the responses of the four univariate systems for the best solutions found by GA1 and GA2, along with their corresponding target values and minimum possible variability that can be achieved. These best solutions appeared in all the 30 runs performed for each response. Additionally, the last column of Table 4 labelled s_{obt}/s_{min}, shows the ratio of the standard deviation of each response attained by each solution to the minimum standard deviation possible to achieve according to the settings of the system simulator. Note that the modified GA (GA2) found all the solutions listed in that table for the previous GA (GA1) and also generated new solutions. The solutions found by GA1 and GA2 are highly reliable. The only solution that was not found by GA2, but was discovered by GA1 corresponds to chromosome [1 4 4 4 4 1 1 1 1 4], which is one of the solutions that optimizes response y_4. In general, note that the new solutions found by GA2 achieve an adjustment of the mean and variance reduction similar to the ones attained by GA1. To further see whether the solutions found by GA1 and GA2 are comparable, the mean fitness value of the best solutions was calculated (see Equations (3), (4) and related explanation in subsection 2.1). For GA1 the mean fitness is 0.92 (standard deviation = 0.05) and for GA2 it is 0.96 (standard deviation = 0.01). The p-value corresponding to the t-statistic to test whether those two means are different is 0.02. That suggests that GA2 mean fitness for its solutions is larger than the mean for GA1's best solutions. Thus, it can be said that GA2 performed better than GA1. Since the study carried out 30 runs for GA1 and GA2 for each response, and the solutions were consistent among runs, the analysis can adequately rule out that the better performance of GA2 over GA1 is due just to chance. Regarding the solution missed by GA2 (chromosome [1 4 4 4 4 1 1 1 1 4]), it must be treated with caution, since the study cannot explain why GA2 did not find it, whereas GA1 did. However, note that GA2 reported a better solution (chromosome [2 3 3 3 3 3 2 1 2 4]), with a closer fit to the target value than the missed chromosome and similar reduction in variance. Thus, the missed solution does not pose a big problem for the practical application of GA2.

Moreover, an additional benefit of having the additional solutions delivered by GA2 is that the manager has more alternatives from which to choose and then implement the solution that better suits his/her other needs, *i.e.* the solution that is more economically feasible.

For the univariate complex systems, note that the R^2 obtained in the estimation of the mean and variance surfaces for each response (between 0.35 and 0.78, with a mean R^2 of 0.46, see Table 6), indicate a more modest fit of the response surfaces than in the case of the real univariate system. Since the corresponding simulator of the univariate complex system was set with high noise, that result is expected. However, GA2 still finds good solutions. Thus, the RSM estimation method works rather well even under noisy conditions.

Additionally, in the case of GA2, the R^2 for the y_1 surface is 0.78 for the mean and 0.48 for the variance. For the y_4 surface the R^2 is 0.35 and 0.39 respectively. Thus, the surface represents much better y_1 than y_4. If one compares the solutions for y_1 and y_4 delivered by GA2 (see Table 4), it is possible to see that mean adjustment is much better for y_1 than for y_4, which suggests a better performance of GA2 as the surface becomes more representative of the system.

The study also performed 30 new runs with GA1 and GA2 for responses y_1 and y_4 of the univariate complex system. During each run, the study calculated the residuals for each of the chromosomes whose values for the responses had to be estimated. Then, for the GA2 runs, the study compared those residuals with the ones obtained in the OLS regression procedure for estimating the response surfaces. A plot of the two types of residuals showed that the former residuals lie inside the envelope of the later ones. This suggests that the estimation error of the surfaces that were calculated during the OLS regression, and the corresponding R^2, adequately represent the approximation error of the missing value estimation procedure of GA2 during a run. Thus, it can be said that using the R^2 of y_1 and y_4 to characterize the impact of the approximation error of the RSM procedure on the performance of GA2 is adequate.

The study also plotted the residuals corresponding to the 30 new runs of both algorithms in one graph. The graph showed that the GA2 residuals lie inside the envelope of the GA1 residuals for each response, which suggests that the missing value estimation method of GA2 is better than the original procedure implemented in GA1. For a more reliable confirmation of this conclusion, the RMSEA for each GA and response was calculated. For GA1 response y_1, the RMSEA is 81.5 and for y_4 it is 284.1. For GA2 response y_1, the RMSEA is 59.8 and for y_4 it is 227.3. It can be seen that the RMSEA for GA2 are smaller than the ones for GA1, which confirms the visual analyses. Also note that, as expected, the RMSEA corresponding to response y_1 is smaller than the RMSEA for response y_4, since the simulator was set with a smaller variance for y_1 than for y_4.

Using the same new runs, the study also counted the number of iterations that GA1 and GA2 had to execute until delivering a chromosome with a fitness value equal to or above 0.9. The mean number of iterations executed by GA1 for y_1 was 81.7 (standard deviation = 11.3) and for y_4 was 91.3 (standard deviation = 7.4). For GA2 the mean number of iterations for y_1 was 28.3 (standard deviation = 6.0) and for y_4 was 42.4 (standard deviation = 9.8). Since GA1's number of iterations for both responses is larger than that for GA2, it may be concluded that GA2 converges faster than GA1. An ANOVA showed that the difference is statistically significant (F-stat $(3, 116) = 351.\ 2$ with p-value = 0.000). This faster convergence of GA2 may be attributed to its superior searching capability, due to the new RSM estimation procedure, caused by a more valid estimation of the real system outputs, *i.e.* an estimated value closer to the value of the outputs of the real system.

Finally, using the same count of iterations, the study compared the mean number of iterations for GA2 for responses y_1 (mean = 28.3, standard deviation = 6.0) and y_4 (mean = 42.4, standard deviation = 9.8). Since the mean for y_4 is significantly larger than the mean for y_1 (p-value of corresponding t-statistic $\ll 0.01$), it can be said that as the response surface becomes a more modest approximation of the system (remember that R^2 for response surface y_4 is smaller than for y_1), GA2 converges slower to a good solution. This may show that as the approximation of the response surface to the real system becomes more modest, the searching capability of GA2 deteriorates.

3.3. *Results obtained for the multivariate complex system*

This case study used the same simulator and the same experimental design as before, but it optimized the four responses at the same time. This means that the GAs are optimizing a four-dimensional multivariate system. As in the previous analysis, 30 runs for GA1 and GA2 were carried out. The Appendix shows the estimation of the response surfaces for this system, which are the same surfaces that were estimated for the univariate systems. Table 2 presents the solutions for this case.

From the three best solutions found by GA1, the GA2 obtained two of them, missing chromosome [2 2 1 4 3 3 4 1 2 4]. However, GA2 delivered a new solution [3 3 1 3 4 4 2 1 3 2] that is quite similar to the missed solution. That happened in all 30 runs, which indicates a consistent behaviour. Regarding mean adjustment, the new solution is much better than the missed one for response y_1, similar for response y_3 and worse for responses y_2 and y_4. In the case of variance reduction, the figures in Table 2 show that the new solution is slightly worse for responses y_1 and y_3, and as good as the missed solution for responses y_2 and y_4. The study also compared the fitness value of both chromosomes. Solution [2 2 1 4 3 3 4 1 2 4] has a fitness of 0.78 and solution [3 3 1 3 4 4 2 1 3 2] a fitness of 0.81. Hence, those figures confirm that both solutions are rather similar and for this reason, one can think that the missed solution does not hinder the practical application of GA2. Here again, the missed solution must be treated with caution, since it was found only by GA1.

Table 2. Solutions obtained by GA1 and GA2 for the multivariate complex system.

		Response mean			Standard deviation		
Responses	Chromosome (solution)	Target value	Obtained	Difference (%)	Minimum possible (s_{min})	Obtained (s_{obt})	s_{obt}/s_{min} (times)
		Previous GA (GA1)					
y_1	3-3-1-2-4-4-2-1-3-2	200	189.00	−5.53	2.10	2.16	1.03
y_2		50	54.00	7.93	0.80	0.81	1.01
y_3		1000	981.00	−1.90	36.00	38.10	1.06
y_4		500	570.00	14.00	28.00	28.50	1.02
y_1	2-2-1-4-3-3-4-1-2-4	200	174.90	−12.55	2.10	2.10	1.00
y_2		50	48.10	−3.80	0.80	0.80	1.00
y_3		1000	1028.80	2.88	36.00	37.10	1.03
y_4		500	553.800	10.76	28.00	28.30	1.01
y_1	1-3-3-3-3-2-2-2-2-4	200	200.30	0.17	2.10	2.10	1.00
y_2		50	43.00	−14.10	0.80	0.82	1.03
y_3		1000	791.20	−20.10	36.00	36.60	1.02
y_4		500	491.20	−1.76	28.00	28.00	1.00
		Modified GA (GA2) (new solution)					
y_1	3-3-1-3-4-4-2-1-3-2	200	192.56	−3.72	2.10	2.40	1.14
y_2		50	54.40	8.80	0.80	0.80	1.00
y_3		1000	971.04	−2.90	36.00	38.20	1.06
y_4		500	580.00	16.00	28.00	28.00	1.00

Note: Solutions listed for GA1, excluding [2-2-1-4-3-3-4-1-2-4], were also delivered by GA2 and are not repeated in the rows that list the new solutions found by GA2.

Finally, the study compared the general performance of GA1 and GA2 for the multivariate system by computing a mean fitness value (see Equations (3), (4) and related explanation in subsection 2.1) for the best solutions delivered by both GAs. For GA1 the mean fitness is 0.81 (standard deviation = 0.02) and for GA2 is 0.91 (standard deviation = 0.08). The p-value for the corresponding t-statistic is 0.06, which suggests that there exists a partial statistically significant difference between both means and that GA2 performance is somewhat better than that of GA1.

For the multivariate complex system, it can also be seen that although the response surfaces are a modest approximation to the simulated system, GA2 delivers good solutions. Now, a comparison of the general performance of GA2 for the univariate simple system with the one for the univariate and multivariate complex systems shows that as the approximation of the response surfaces to the outputs of the system is better, the RSM estimating procedure of missing values works better. This is an expected result and highlights the importance of observing the rules stated in the Appendix in order to estimate adequate response surfaces.

4. Conclusions

The results of the application of the modified GA (GA2) to the multivariate complex system showed that GA2 outperforms GA1. In the case of the univariate complex systems, the performance of GA2 is better than that of GA1, as indicated by a significantly larger mean fitness of the best solutions delivered by GA2 than the mean fitness of the solutions provided by GA1. Also, evidence suggests that GA2 converges faster to a good solution than GA1, which may be a crucial feature of GA2 if it must be applied to optimize a system in real time.

Additionally, for the univariate systems, GA2 found almost all of the solutions generated by GA1, except one, but also found eight new solutions. This is important, since the manager of a production process generally needs to consider economic factors when selecting the solution to be implemented. The manager will need to assess the trade off between variance reduction and mean adjustment and see which of the two objectives is more important and cost-effective in delivering quality products to the customers. In order to be able to do that, the manager will need to take into account the cost that the firm must incur for implementing each of the solutions found by the GA. Thus, if the GA delivers a larger number of solutions, the manager will have more alternatives to choose from, which will make it easier to find a solution that may meet the trade-off among variance reduction, mean adjustment and economic considerations.

Regarding the practical application of the modified GA (GA2), note that the experimenter must decide the model that he/she wants to use for estimating the mean and variance response surfaces, based on expressions (5) to (8) and the number of data points that he or she will collect in the robust design experiment. As the Appendix explains, those considerations are important in the design of the experiment and will have an impact on the ability of the response surfaces for adequately representing the true surfaces that characterize the mean and variance of the outputs of the system. In general, the more levels each considered factor has, the more combinations of factors the experimental design considers and the larger the number of replications that the experimenter carries out, the more representative the response surfaces will be. Thus, it can be said that for robust design studies that use very highly fractioned (partitioned) experimental designs, with very few levels for each factor and a small number of replications, the difference in the benefits of applying GA2 instead of GA1 might not be too large. Additionally, note that the use of GA1 is totally automated, *i.e.* the manager conducts the robust design experiment and inputs the data into GA1, which will automatically find the solutions. Instead, if the manager uses GA2, he/she must decide the response surface expressions that GA2 will employ and input them into the GA2, along with the corresponding data. Then, the GA2 automatically computes the response surfaces using OLS regression. Although the additional effort to do the extra work is not too large, maybe a good approach would be to first use GA1 to find the solutions and only then employ GA2 if those solutions do not satisfy the requirements for variance reduction, mean adjustment and the other economic considerations. On the other hand, RSM is a good technique for estimating mean and variance of a system, especially when one suspects that the system is highly non-linear (Vining and Myers 1990, Nair *et al.* 2002, Myers and Montgomery 2002). Remember that GA1's missing value estimation procedure is totally linear, whereas the RSM method implemented in GA2 can accommodate non-linearity. Thus, when an engineer has evidence that the system under analysis is non-linear, or at least he/she cannot rule out that possibility, the use of GA2 from the beginning would be sensible. In that case, note that the engineer should consider an experimental design that meets the above mentioned requirements for estimating representative response surfaces, which are detailed in the Appendix. Moreover, since the present analyses suggest that the performance of GA2 improves as the response surfaces become a better approximation to the outputs of the real system, the experimenter should put every effort in estimating good surfaces.

Finally, it is important to note that there is still further work to be done to fully analyse the functioning of GA2. Although the study concluded that the performance of GA2 was improved by implementing a new missing value estimation procedure based on RSM, it is necessary to investigate exactly why that is happening. As pointed out by one reviewer, at this stage, the study cannot say whether the improved performance of GA2 stems from the fact that GA2 explores a different part of the parameter space or that GA2 explores a similar part of the parameter space, but is able to calculate a more adequate (true) fitness for the chromosomes. That is an interesting issue, but does not hinder the practical application of GA2. Along with investigating that subject, future research will analyse the impact of different experimental designs (*e.g.* different levels of fractioning), noise conditions and, several values for some of the parameters of the GAs (*e.g.* crossover and mutation probabilities) on their performance.

Acknowledgements

The authors would like to thank Rick L. Riolo, Center for the Study of Complex Systems, The University of Michigan, and Claudio Moraga, European Centre for Soft Computing and FB Informatik Universität, Dortmund, for their valuable comments regarding the original version of this article. This work was supported in part by Research Grants 1110854 Fondecyt and by FB081 Centro Tecnológico de Valparaíso.

References

Allende, H., Bravo, D. and Canessa, E., 2008. Robust design in multivariate systems using genetic algorithms. *Quality & Quantity Journal*, 44, 315–332.

Allende, H., Canessa, E. and Galbiati, J., 2005. *Diseño de Experimentos Industriales*. Valparaíso: Universidad Técnica Federico Santa María.

Canessa, E., Droop, C. and Allende, H., 2011. *An improved genetic algorithm for robust design in multivariate systems* [online]. *Quality & Quantity Journal*, DOI 10.1007/s11135-010-9420-y.

De Mast, J., 2004. A methodological comparison of three strategies for quality improvement. *Int. J. Quality & Reliability Management*, 21 (2), 198–213.

Del Castillo, E., Montgomery, D.C. and McCarville, D.R., 1996. Modified desirability functions for multiple response optimization. *Journal of Quality Technology*, 28, 337–345.

Hair, J., Anderson, R., Tatham, R. and Black, W., 1992. *Multivariate data analysis*. New York: Macmillan.

Holland, J., 1974. *Adaptation in natural and artificial systems*. Ann Arbor: The University of Michigan Press.

Hoyer, R.W. and Hoyer, B.Y., 2001. What is quality? *Quality Progress,* July, pp. 52–62.

Kackar, R.N. and Shoemaker, A.C., 1986. Robust design: A cost effective method for improving manufacturing processes. *AT&T Technology Journal*, 65 (2), 39–50.

Maghsoodloo, S. and Chang, C., 2001. Quadratic loss functions and SNR for a bivariate response. *Journal of Manufacturing Systems*, 20 (1), 1–12.

Myers, R. and Montgomery, D., 2002. *Response surface methodology: process and product optimization using designed experiments*. New York: John Wiley & Sons.

Nair, V.N., Taam, W. and Ye, K.Q., 2002. Analysis of functional responses from robust design studies. *Journal of Quality Technology*, 34 (4), 355–370.
Ortiz, F., Simpson, J., Pigniatello, J. and Heredia-Langner, A., 2004. A genetic algorithm approach to multiple-response optimization. *Journal of Quality Technology*, 36 (4), 432–449.
Roy, R.K., 2001. *Design of experiments using the Taguchi approach*. New York: John Wiley & Sons.
Taguchi, G., 1991. *Systems of experimental design*. Dearborn: American Supplier Institute.
Vandenbrande, W., 1998. SPC in paint application: Mission Impossible? *ASQ's 52nd Annual Quality Congress Proceedings*, 4–6 May, Philadelphia, PA. Philadelphia: American Society for Quality, 708–715.
Vandenbrande, W., 2000. Make love, not war: Combining DOE and Taguchi. *ASQ's 54th Annual Quality Congress Proceedings*, 8–10 May, Indianapolis, IN. Indianapolis: American Society for Quality, 450–456.
Vining, G.G. and Myers, R.H., 1990. Combining Taguchi and response surface philosophies: a dual response approach. *Journal of Quality Technology*, 22, 38–45.
Yuan, Y., 2006. *Multiple imputation for missing data: Concepts and new development*. Rockville, MD: SAS Institute Inc.

Appendix

This Appendix discusses some details related to the building of the response surfaces corresponding to the mean and variance of each of the responses of a system. Section A1 presents some general considerations that are important for designing and collecting the data in a robust design experiment, when using RSM. Then, Section A2 shows how the response surfaces for the first test case were built, which corresponds to a univariate simple system, but including qualitative and quantitative variables. Since this system has a conveniently small number of variables, this subsection presents all the details of the modelling process. Finally, Section A3 presents only the most important steps carried out to build the response surfaces for the complex multivariate system.

A1. *General considerations for defining a robust design experiment when using RSM*

If a manager wants to apply RSM to a robust design study, he/she needs to have some minimum number of data points, so that the estimated response surfaces for the mean and variance of the system adequately represent the outputs of the system. Moreover, since the system might be non-linear, the manager must be on the safe side, and thus assume non-linearity of the responses. Additionally, if the manager wants to test the statistical significance of the adjustment of the model to the data by using the F-ratio and Student's $t-$ test statistics, he/she needs to have enough degrees of freedom for the error term. To satisfy all those conditions, Myers and Montgomery (2002) suggest that the experimental design adhere to the following conditions:

(1) Consider at least three levels (values) for each of the control factors.
(2) Distinguish at least $1 + 2k + k(k-1)/2$ design points, for k control factors.
(3) Consider leaving at least one degree of freedom for the error term, where the number of d.f. for the error term are equal to $n - p - 1$, for n treatment combinations and p parameters to be estimated.

Another concern relates to the possible multi-collinearity problems that the model and data might exhibit. If high multi-collinearity exists, that does not imply that the corresponding response surface might be useless. However, since OLS estimates of the β coefficients are computed using the matrix equation $\beta = (X^TX)^{-1}X^TY$, the matrix X^TX might become singular or nearly singular and thus, OLS might not be able to estimate the value of the β coefficients or their values might be unstable (Myers and Montgomery 2002). Thus, when estimating the model, the manager should look at the collinearity statistics and see whether that problem exists and modify the model accordingly. Generally, that is done analyzing the tolerance or variance inflation factor (VIF) of the estimate of each coefficient, where a tolerance below 0.1, or equivalently a VIF above 10.0, indicates that a multi-collinearity problem might exist with the corresponding coefficient (Hair *et al.* 1992). In that case, Hair *et al.* (1992) recommend that the given variable be removed from the model.

Finally, Myers and Montgomery (2002) suggest that the quantitative variables be rescaled and centred on their means, so that the values of them lie between -1 and 1. For doing that, let ξ be the natural value of a variable, then x will be the codified and standardized variable calculated using the following expression:

$$x = \frac{\xi - \frac{\max(\xi)+\min(\xi)}{2}}{\frac{\max(\xi)-\min(\xi)}{2}} \tag{9}$$

A2. *Building of the response surfaces for the simple univariate system*

In this case, the experiment considers four control factors with three levels each (Vandenbrande 1998, 2000). Table 3 presents the control factors and their levels.

Since the experiment has four control factors with three levels each, the experimental design must consider at least 15 design points (1+8+4×3/2, according to condition (2) stated in A1). The robust design considered 36 design points (nine

Table 3. Control factors and levels for the simple univariate real system.

Control Factor	Level 1	Level 2	Level 3
A: Type of spray gun used	Type 1	Type 2	Type 3
B: Paint flow [cc/min]	490	440	390
C: Fan air flow [Nl/min]	260	220	180
D: Atomizing air flow [Nl/min]	390	330	270

Table 4. ***X*** and ***Y*** matrices for estimating the response surfaces.

1	z_1	z_2	x_2	x_3	x_4	x_2^2	x_3^2	y_{mean}	$y_{variance}$
1	1	0	0	0	0	0	0	41.150	8.243
1	1	0	−1	−1	−1	1	1	44.800	1.393
1	1	0	1	1	1	1	1	35.825	12.563
1	0	1	1	0	−1	1	0	24.825	3.163
1	0	1	0	−1	1	0	1	45.175	17.703
1	0	1	−1	1	0	1	1	41.025	2.069
1	0	0	1	−1	0	1	1	43.200	65.860
1	0	0	0	1	−1	0	1	34.475	2.056
1	0	0	−1	0	1	1	0	41.050	76.170

Table 5. Regression coefficients for the mean and variance response surfaces.

Coefficients for the mean RS	Value	Coefficient for the variance RS	Value
β_0	37.400	γ_0	44.506
δ_1	1.017	η_1	−40.629
δ_2	−2.567	η_2	−40.384
β_2	−3.838	γ_2	0.325
β_3	−3.642	γ_3	−11.378
β_4	2.992	γ_4	16.637
β_{22}	−1.813	γ_{22}	17.536
β_{33}	5.075	γ_{33}	−12.251

treatment combinations for control factors times four treatment combinations for the noise factors) (Vandenbrande 2000) and, thus conditions (1) and (2) are met. As can be seen, factor A is a qualitative variable, whereas the rest are quantitative. Thus, Equations (7) and (8) must be used for estimating the model. In general, the model needs $l-1$ dummy variables for representing a qualitative variable that has l levels (Myers and Montgomery 2002). Given that factor A has three levels, the model must use two dummy variables for representing it. Finally, the design of the experiment consisted of an orthogonal array $L_9(3^4)$ for the four control factors, so that the study has nine data points for the mean and variance of the response. Thus, in order to leave one degree of freedom for the error term and satisfy condition (3), the experimenter must choose only eight parameters to estimate. In this case, it was decided to estimate the position parameter and the four first order coefficients. Note that since factor A is represented by two dummy variables, the model consumes two degrees of freedom for factor A, and three more for quantitative factors B, C and D, and one for the position parameter, leaving two degrees of freedom for estimating other parameters. Then, by examining the fit of the model to the data, the researchers began to include some quadratic terms, concluding that a good fit was obtained when considering the quadratic terms for factors B and C. With those considerations, the final models are the following:

$$m(X) = \beta_0 + \sum_{p=2}^{4} \beta_p x_p + \sum_{i=1}^{2} \delta_i z_i + \sum_{p=2}^{3} \beta_{pp} x_p^2 + \varepsilon_m \tag{10}$$

$$v(X) = \gamma_0 + \sum_{p=2}^{4} \gamma_p x_p + \sum_{i=1}^{2} \eta_i z_i + \sum_{p=2}^{3} \gamma_{pp} x_p^2 + \varepsilon_v \tag{11}$$

Table 4 shows the extended X matrix of independent variables and the Y matrix of observations, corresponding to Equations (10) and (11). Note that the X matrix has a column of ones, which represents the position parameter β_0 and γ_0 respectively. The variables z_1 and z_2 are the two dummy variables representing factor A, x_2, x_3 and x_4 represent the first order coefficient for factors B, C and D and, x_2^2 and x_3^2 represent the quadratic coefficients for factors B and C. Note that the independent variables were codified and standardized according to expression (9). The treatment combinations of matrix X correspond to the ones of the inner orthogonal array $L_9(3^4)$ and the mean and variance of each row were calculated using the data provided in Vandenbrande (2000).

The OLS estimates of the response surfaces are shown in Table 5. The models as well as the coefficients are all statistically significant at least at the 0.05 level. The R^2 for the mean response surface is 0.899 and 0.987 for the variance surface. Thus, it can be said that the response surfaces are a good representation of the true system.

A3. *Building of the response surfaces for the complex univariate and multivariate system*

Since the univariate and multivariate system simulator has four responses, 10 control factors and five noise factors, it would take many pages to describe all the details and present all the figures of the building of the response surfaces. Thus, this subsection will present only the most relevant aspects of that procedure. Readers interested in further details may contact the corresponding author.

The robust experimental design considered factors with four levels each and a total of 1024 design points (64 treatment combinations for control factors times 16 treatment combinations for the noise factors), which met conditions (1) and (2) (minimum number of design points=1+2×10+10×9/2=66, see subsection A1). Since the inner array has 64 treatment combinations, the experimenter can estimate up to 63 parameters, leaving one degree of freedom for the error term. However, when the experimenters were estimating the parameters, the multicollinearity statistics showed that such a problem existed. Thus, they looked at the VIF of each coefficient and began an elimination process for the variables that might have caused that problem, arriving at the following models for the mean and variance of each of the four responses, where equations (5) and (6) were used since all the ten control factors are quantitative variables:

$$m(X) = \beta_0 + \sum_{p=1}^{10} \beta_p x_p + \sum_{p=1}^{10} \beta_{pp} x_p^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \varepsilon_m \tag{12}$$

$$v(X) = \gamma_0 + \sum_{p=1}^{10} \gamma_p x_p + \sum_{p=1}^{10} \gamma_{pp} x_p^2 + \gamma_{12} x_1 x_2 + \gamma_{13} x_1 x_3 + \gamma_{14} x_1 x_4 + \gamma_{23} x_2 x_3 + \varepsilon_v \tag{13}$$

Expressions (12) and (13) show that the experimenters estimated the position parameter, all the linear and quadratic coefficients and the interaction terms corresponding to double interactions AB, AC, AD and BC. Later, for corroborating whether those models were a good approximation to the true response surfaces, the estimated models and coefficient were compared to the expressions that comprise the simulator. The expressions and coefficients of the simulator may be seen in Canessa *et al.* (2010) and the comparison showed a modest approximation. Table 6 presents the R^2 for each of the four response surfaces corresponding to the mean and variance of each of the system's outputs. The R^2 are rather good, considering that the simulator was set with a large Gaussian noise, especially for response y_4, which causes that response surface y_4 to have the smallest R^2 among the four response surfaces. In this case, the R^2 values suggest that the response surfaces are a modest representation of the outputs of the true system.

Table 6. R^2 for each of the estimated mean and variance response surfaces.

Response surface	R^2 for the mean RS	R^2 for the variance RS
y_1	0.779	0.481
y_2	0.382	0.399
y_3	0.517	0.405
y_4	0.350	0.391