Ingeniería e Investigación
ISSN: 0120-5609
revii_bog@unal.edu.co
Universidad Nacional de Colombia
Colombia

Canessa, E.; Chaigneau, S.
Response surface methodology for estimating missing values in a pareto genetic algorithm used in parameter design
Ingeniería e Investigación, vol. 37, núm. 2, agosto, 2017, pp. 89-98
Universidad Nacional de Colombia
Bogotá, Colombia

DOI: http://dx.doi.org/10.15446/ing.investig.v37n2.57152

Response surface methodology for estimating missing values in a pareto genetic algorithm used in parameter design

Metodología de superficie de respuesta para estimar valores faltantes en un algoritmo genético de pareto usado en diseño de parámetros

E. Canessa[1], and S. Chaigneau[2]

ABSTRACT

We present an improved Pareto Genetic Algorithm (PGA), which finds solutions to problems of robust design in multi-response systems with 4 responses and as many as 10 control and 5 noise factors. Because some response values might not have been obtained in the robust design experiment and are needed in the search process, the PGA uses Response Surface Methodology (RSM) to estimate them. Not only the PGA delivered solutions that adequately adjusted the response means to their target values, and with low variability, but also found more Pareto efficient solutions than a previous version of the PGA. This improvement makes it easier to find solutions that meet the trade-off among variance reduction, mean adjustment and economic considerations. Furthermore, RSM allows estimating outputs' means and variances in highly non-linear systems, making the new PGA appropriate for such systems.

Keywords: Robust design, parameter design, pareto genetic algorithm, response surface methodology.

RESUMEN

En este artículo se presenta un Algoritmo Genético de Pareto (AGP) mejorado que encuentra soluciones a problemas de diseño robusto en sistemas multi-respuesta con 4 respuestas y hasta 10 factores de control y 5 de ruido. Ya que algunas respuestas podrían no haber sido obtenidas en el experimento de diseño robusto y se necesitan en el proceso de búsqueda, el AGP usa metodología de superficie de respuesta (MSR) para estimarlas. El AGP no solo entregó soluciones que ajustan adecuadamente la media de las respuestas a sus valores meta y con poca variabilidad, sino que también encontró más soluciones Pareto eficientes que una versión previa del AGP. Esta mejora facilita encontrar soluciones que alcancen el balance entre reducción de variabilidad, ajuste de media y consideraciones económicas. Además, la MSR permite estimar las medias y varianzas de las respuestas de sistemas altamente no lineales, haciendo apropiado el uso del AGP en dichos sistemas.

Palabras clave: Diseño robusto, diseño de parámetros, algoritmo genético de pareto, metodología de superficie de respuesta.

Received: April 25th 2016
Accepted: March 8th 2017

Introduction

Parameter Design (PD) is a two-stage method for executing robust design (RD) interventions, which tries to set controllable input factors of a production or service system, so that the system's outputs stay as stable as possible. Then it adjusts other control factors to bring the mean of the outputs as close as possible to their target values, even under the presence of noise factors (Taguchi, 1991). To find solutions that achieve both objectives of PD and to be able to assess the relative merit of setting different control factors to reduce variability and obtain mean adjustment, the use of a Pareto Genetic Algorithm (PGA) has been proposed (Canessa, Bielenberg & Allende, 2014). The PGA is especially useful when an engineer needs to consider many control and noise factors and the system has many responses that must be simultaneously optimized (Canessa *et al*., 2014; Lin *et al*., 2014). The PGA delivers Pareto efficient solutions and thus, it exposes the trade-off between reducing variability and getting mean adjustment. Using such solutions, the engineer then incorporates other considerations (e.g. costs to implement each solution, importance of obtaining a product or service with a given level of variance and adjustment to target

[1] Electronics Engineer, APN, Chile. MBA, PhD, University of Michigan, USA. Affiliation: Associate Professor, Universidad Adolfo Ibáñez, Facultad Ingeniería y Ciencias, Chile. E-mail: ecanessa@uai.cl.

[2] Psychologist, Universidad de Chile, Chile. PhD, Emory University, USA. Affiliation: Full Professor, Universidad Adolfo Ibáñez, Escuela Psicología, Chile. E-mail: sergio.chaigneau@uai.cl

How to cite: Canessa, E., and Chaigneau, S. (2017). Response surface methodology for estimating missing values in a Pareto Genetic Algorithm. *Ingeniería e Investigación*, 37(2), 89-98.
DOI: 10.15446/ing.investig.v37n2.57152

values) to decide which of the solutions is the most cost-effective. For this analysis to be worthwhile, the solutions delivered by the PGA must be a good approximation to the Pareto frontier, both in terms of Pareto efficiency and number of solutions found. The original PGA developed by Canessa *et al*. (2014) has worked well under different scenarios, but results have shown that a bigger number of data points fed to the PGA may enhance the approximation to the Pareto frontier (Canessa *et al*., 2014). Thus, the current work applies Response Surface Methodology (RSM) (Myers & Montgomery, 2002) to estimate the outputs of the system, based on a small number of data points gathered in experiments. Then, using the real and estimated data, the PGA finds a better approximation to the Pareto frontier. The application of genetic algorithms to problems of RD and PD is not new (see e.g. Allende, Bravo & Canessa, 2010; Canessa, Droop & Allende , 2011; Wan & Birch, 2011). Also, some researchers have studied the use of RSM to optimize systems (e.g. Nair, Tam & Ye, 2002; Robinson *et al*., 2004; Vining & Myers, 1990; Wu & Hamada, 2009). However, to the best of our knowledge, the simultaneous application of PGA and RSM to problems of PD with many responses, control and noise factors, is still an open research topic. Moreover, the PGA presented here will automatically calculate the RSM and use it in the estimation of missing values of responses. This last point is important, given that to lower the cost of PD studies, generally, engineers use highly fractioned experimental designs (Maghsoodloo, Jordan & Huang, 2004; Roy, 2001; Taguchi, 1991). Thus, when the PGA is iterating, it frequently needs to know the values of 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, during the search process of the PGA, it needs to estimate those values. A good estimate of these values will in turn enhance the search capabilities of the PGA and allow the PGA to obtain a good approximation to the Pareto frontier.

Using Response Surface Methodology along with the Pareto Genetic Algorithm

The following subsections present some parts of the previous PGA (PGA1) and concepts related to RSM necessary for understanding the changes made to PGA1. More details of PGA1 may be found in Canessa *et al*.(2014).

Some details of the original Pareto GA (PGA1).

The previous PGA1 developed in Canessa et al. (2014) 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)).

$$x_l = \left[f_{l1}, f_{l2}, \cdots, f_{lk} \right] \tag{1}$$

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

Each of the chromosomes (solutions) x_l will generate different responses $y_r (r = 1,2,3,\ldots,R)$ of a multi-response system when the control factors are set to the corresponding levels specified in the chromosome x_l. The PGA 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 problem that the PGA needs to solve can be expressed by Equation (3):

$$\begin{aligned} Min\ & f(x_l) = \left(\left[\bar{y}_r(x_l) - t_r \right]^2, s_r^{\,2}(x_l) \right) \\ subject\ to: & L_r \le \bar{y}_r(x_l) \le H_r; \forall l, r \\ where: & x_l = \left(f_{l1}, f_{l2}, \ldots, f_{lk} \right) \in X \end{aligned} \tag{3}$$

Expression (3) means that the PGA will try to find solutions x_l, which must adjust the mean of each response y_r to its corresponding target value τ_r and simultaneously reduce the variance of each response. L_r and H_r are lower and upper limits within which, each response y_r is feasible and must be set by the engineer. Following the work presented in Del Castillo, Montgomery & McCarville (1996), a penalty function is used to enforce such limits. Finally, in the case of single-response systems ($R = 1$), the PGA directly uses expression (3). On the other hand, for multi-response systems ($R > 1$), the PGA aggregates $\left[\overline{y_r}(x_l) - \tau_r \right]^2$ over all responses y_r into φ_1 using a desirability function proposed in Ortiz *et al*. (2004). The same is done with all the s_r^2 to obtain φ_2. Given that the desirability functions are defined so that the smaller $\left[\overline{y_r}(x_l) - \tau_r \right]^2$ and s_r^2, the larger φ_1 and φ_2 (in the range [0,1]), the problem for multi-response systems is stated as a maximization of φ_1 and φ_2, as Equation (4) shows:

$$\begin{aligned} Max\ & f(x_l) = \left(\phi_1\left(\left[\bar{y}_r(x_l) - t_r \right]^2 \right), \phi_2\left(s_r^{\,2}(x_l) \right) \right) \\ subject\ to: & L_r \le \bar{y}_r(x_l) \le \mathrm{H}_r; \forall l, r \\ where: & x_l = \left(f_{l1}, f_{l2}, \ldots, f_{lk} \right) \in X \end{aligned} \tag{4}$$

More details of these mechanisms may be found in Canessa *et al*. (2014) and are not repeated here for the sake of length and because an exhaustive explanation is unnecessary for understanding the work shown here.

From Equations (3) and (4), one can see that in the calculation of the value of those expressions PGA1 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, PGA1 needs to estimate those responses. In PGA1, for estimating the mean of a response for a non-tried chromosome (treatment), PGA1 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, PGA1 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), PGA1 uses a similar procedure. PGA1 first computes a global variance considering all the replications of all the treatments tried in the original experiment. Then PGA1 calculates the main effect of each control factor on the variance. Finally, PGA1 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 (for a worked out numerical calculation, see example in Taguchi (1991, p. 16). In the search process, PGA1 uses roulette or probabilistic selection, a uniform crossover with $p_c = 0{,}3$ and a bit by bit (factor by factor) mutation operator with $p_m = 0{,}05$. PGA1's stopping criterion is to iterate 15 times. Larger values of iterations were tried, but the performance of PGA1 did not improve (Canessa *et al*., 2014). To implement PGA1, the Strength Pareto Evolutionary Algorithm 2 (SPEA2) (Laumans, 2008) is used, with minor adjustments that the interested reader may find in Canessa *et al*. (2014).

RSM and its application to the estimation of missing values in the new Pareto GA (PGA2)

Using techniques that are part of Response Surface Methodology (RSM), Vining & 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 optimization problems. Further analysis of that approach indicated that it provides many benefits (for a good discussion see Nair, Taam & Ye, 2002). Vining & 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*):

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

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

Expression (5) corresponds to the response surface that estimates the mean of the response and Equation (6) approximates its variance. 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 & Montgomery, 2002). Accordingly, Expressions (5) and (6) must be modified to accommodate the use of indicator variables as Equations (7) and (8) show:

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

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

where there are *k* quantitative control variables, and z_1, z_2, …, z_n are *n* indicator variables, which represent the qualitative control variables. Also, 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 & Montgomery, 2002), which agrees with Taguchi's suggestions for conducting parameter design experiments (Taguchi, 1991, cf. Roy, 2001).

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 & Montgomery, 2002). Thus, the modified PGA (PGA2) 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, PGA2 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 PGA2 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 behavior 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 of PGA1. As a reviewer noted, we acknowledge that a better approach to dealing with this problem could be to calculate the response surface of variance using a logarithmic transformation (Wu & Hamada, 2009), which we will implement in future work. However, as results will show, our simple method works rather well.

Summarizing this section, 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 PGA1, by a new and more accurate method based on RSM in PGA2.

Application of the new PGA2

To evaluate the performance of the new PGA (PGA2) and also compare it to the previous version (PGA1) (Canessa *et al*., 2014), two case studies were used. The first one corresponds to a real application of PD to adjust the automatic body paint process in a car manufacturing plant. The second case study uses a multi-response process simulator with four responses, ten control factors and five noise factors. This simulator is described in Canessa, Droop & Allende (2011) and was used to test PGA1. To be able to compare the performance of PGA1 with the one of PGA2, both PGAs were set with the same parameters described before (e.g. crossover probability $p_c = 0{,}3$, mutation probability $p_m = 0{,}05$, etc.). Here we present the most important aspects of the calculation of the response surfaces and the assessment of the performance of PGA2. For further details, see the appendix.

Results obtained for the single-response real system

In this case, a parameter 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. Table 1 shows the control factors and their levels. More details and the data may be found in Vandenbrande (2000, 1998), and also see the appendix.

Table 1. Control factors and levels for the single-response 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

Source: Authors

As can be seen from Table 1, 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 & Montgomery, 2002). Given that factor A has three levels, the model must use two dummy variables for representing it. Since the design of the experiment consisted of an orthogonal array L9(34) for the four control factors, 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, 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. 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:

$$m(X) = \beta_0 + \sum_{p=2}^{4} \beta_p x_p + \sum_{i=1}^{2} \delta_i z_i + \sum_{p=2}^{3} \beta_{pp} x_p^{\ 2} + \varepsilon_m \qquad (9)$$

$$v(X) = \gamma_0 + \sum_{p=2}^{4} \gamma_p x_p + \sum_{i=1}^{2} \eta_i z_i + \sum_{p=2}^{3} \gamma_{pp} x_p^{\ 2} + \varepsilon_v \qquad (10)$$

Table 2 shows the values of the coefficients of Expressions (9) and (10). β_0 and γ_0 are the position parameters. 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 variables x_2, x_3, x_4 were rescaled and centered on their means, so that the values of them lie between -1 and 1 (per the procedure suggested in Myers & Montgomery, 2002). 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.

Table 2. Regression coefficients for the mean and variance response surfaces

Coefficients for the Mean Response Surface ($R^2 = 0{,}899$)	Value	Coefficients for the Variance Response Surface ($R^2 = 0{,}987$)	Value
β_0	37,400	γ_0	44,506
δ_1 (z_1)	1,017	η_1 (z_1)	-40,629
δ_2 (z_2)	-2,567	η_2 (z_2)	-40,384
β_2 (x_2)	-3,838	γ_2 (x_2)	0,325
β_3 (x_3)	-3,642	γ_3 (x_3)	-11,378
β_4 (x_4)	2,992	γ_4 (x_4)	16,637
β_{22} (x_2^2)	-1,813	γ_{22} (x_2^2)	17,536
β_{33} (x_3^2)	5,075	γ_{33} (x_3^2)	-12,251

Source: Authors

The two response surfaces were used in PGA2 to find an approximation to the Pareto frontier and 30 experimental runs were performed. Figure 1 shows the Pareto efficient solutions that were found by PGA2 (marked with a square), along with the ones delivered by PGA1 (marked with a rhomb). In Figure 1, the values shown in square brackets indicate the levels of factors set in each solution. For example, solution [1222], represents control factor A = 1 (spray gun Type 1), B = 2 (paint flow of 440 [cc / min]), C = 2 (fan air flow of 220 [Nl / min] and D = 2 (atomizing air flow of 330 [Nl / min]). The horizontal axis shows the mean adjustment as $|y - T|$ attained by each solution

in [cm]. The vertical axis shows variability reduction as *s* (standard deviation) of the response in [cm]. Note that the solutions delivered by PGA1 and PGA2 define similar Pareto frontiers. All of the solutions were consistently found in the 30 runs. Thus, for this case, PGA1 and PGA2 have a similar performance. Given that this single-response system has a few control and noise factors and is quite simple, a more refined missing value estimation mechanism does not impact too much the performance of the PGAs. Also, and as expected, since the surfaces that represent the mean and variance of the system are a good approximation to the real system, PGA2 has a good performance. In the next subsection, we test PGA2 under less favorable conditions, so that we can show how much PGA2's performance is reduced.

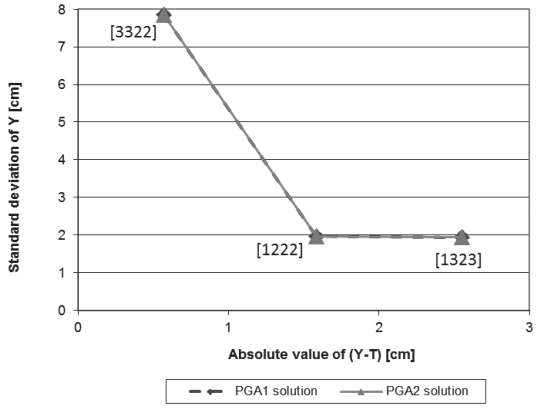

Figure 1. Pareto frontier for the solutions corresponding to the real single-response system.
Source: Authors

Results obtained for the single-response complex systems

To test both PGAs under a more complex situation, a simulator was built, which is described in detail in Canessa, Droop & Allende (2011). The robust design for this situation considers using an inner array $L_{64}(4^{10})$ for the ten control factors and an outer array $L_{16}(4^5)$ for the five noise factors, which represents a hard to treat system (Taguchi, 1991; Roy, 2001). For the following case studies, the four responses of the simulator are optimized independent from each other, so that both PGAs deal with four single-response systems.

Since there are four single-response systems with ten 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 four response surfaces (see the appendix for further important details). Thus, here we present only the most relevant aspects of that procedure. Readers interested in further details may contact the corresponding author. The 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). Given that 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 models for the mean and variance of each of the four responses shown by Equations (11) and (12), which are versions of Expressions (5) and (6) since all the ten control factors are quantitative variables:

$$m(X) = \beta_0 + \sum_{p=1}^{10} \beta_p x_p + \sum_{p=1}^{10} \beta_{pp} x_p^{\ 2} + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \varepsilon_m \quad (11)$$

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

Expressions (11) and (12) 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 (see the appendix for further details). Later, for corroborating whether those models were a good approximation to the true response surfaces, the estimated models and coefficients were compared to the expressions that comprise the simulator. The expressions and coefficients of the simulator may be seen in Allende, Bravo & Canessa (2010) and the comparison showed a modest approximation. Table 3 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 deliberately 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. To be on the safe side, we added a large noise because we wanted to assess the performance of RSM and PGA2 under unfavorable conditions. Given that for the real system the R^2 values are high (0,899 and 0.987) and PGA2 works very well, we thought that it would be appropriate to test PGA2 under less favorable conditions. Hence, if the method works under those adverse conditions, it should perform even better in more favorable situations, as the results for the real system show.

Table 3. R^2 for each of the estimated mean and variance response surfaces

Response surface representing	R^2 for the Mean Response Surface	R^2 for the Variance Response Surface
y1	0,779	0,481
y2	0,382	0,399
y3	0,517	0,405
y4	0,350	0,391

Source: Authors

After estimating the response surfaces, these were used in PGA2 and 30 runs were executed for finding solutions. The next paragraphs present the results for each of the four responses.

Figure 2 presents the solutions corresponding to the Pareto frontier found by PGA2 for the four single-response systems (marked with a square). Those solutions were consistently found in the 30 runs performed. Also, Figure 2 shows the solutions delivered by PGA1 in a previous study (Canessa *et al*., 2014), marked by a rhomb. The horizontal axis shows the mean adjustment as $|y_r - T_r|$ attained by each solution for response *r*. The vertical axis shows variability reduction as s_r (standard deviation) for each response *r*.

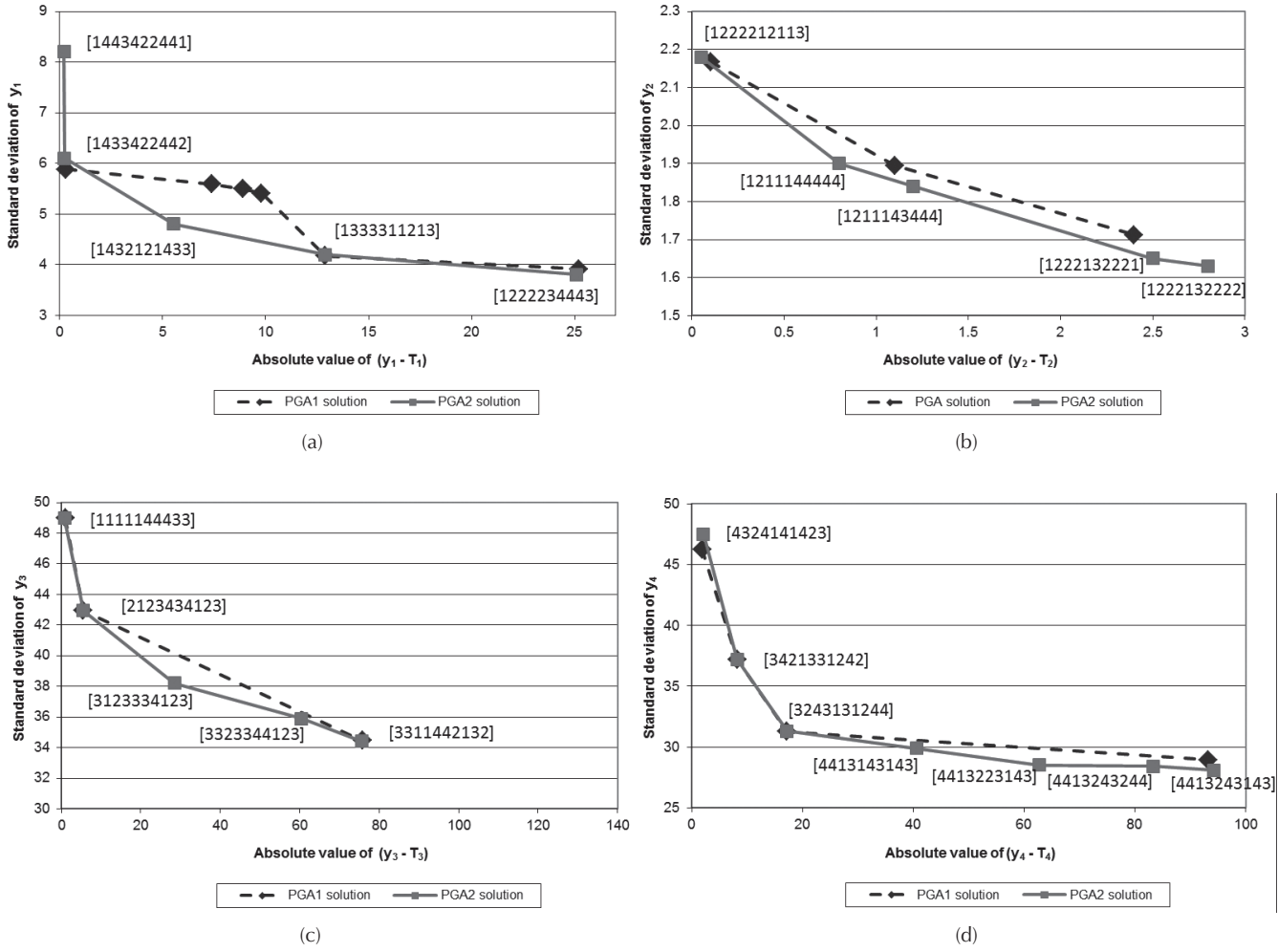

Figure 2. Pareto frontier for the solutions corresponding to the single-response system: (a) y_1 (b) y_2 (c) y_3 (d) y_4.
Source: Authors

For response y_1, Figure 2 (a) indicates that PGA2 delivered five solutions, whereas PGA1 found 6 points. However, note that the Pareto frontier defined by PGA2 is better than the one obtained by PGA1. PGA2's Pareto frontier is totally convex, while the one of PGA1 is not. Also, the Pareto frontier delivered by PGA2 is a better approximation to the real unknown Pareto frontier, given that PGA2's solutions define a Pareto frontier that is closer to the optimal point of mean adjustment and variability reduction, i.e. $|y_1 - T_1| = 0$ and $s_1 = 0$, than the one obtained by PGA1. Thus, we can say that the missing value RSM estimation method improved the performance of PGA2 for response y_1. For further details regarding the assessment of performance see the appendix.

For response y_2, Figure 2 (b) shows that PGA2 found five efficient solutions, whereas PGA1 found only three. As before, PGA2's solutions define a better approximation to the Pareto frontier and PGA2 also found two more extreme solutions ([1222132221] and [1222132222]).

Regarding response y_3, PGA2 and PGA1 found almost the same solutions in terms of mean adjustment and variability reduction. However, PGA2 delivered one additional solution [3123334123] that helps better define the Pareto frontier. Note that solution [3123334123] shifts down-left the Pareto frontier, which strictly speaking, removes solution [3323344123] from the frontier. Given that PGA2 delivered solution [3323344123] as a Pareto efficient one, the Pareto

frontier is not totally convex. However, the approximation is good.

Finally, for response y_4, Figure 2 (d) shows that two solutions obtained by PGA2 are exactly the same, in terms of $|y_4 - T_4|$ and s_4, as the ones found by PGA1 ([3421331242] and [3243131244]). Also, two other solutions are very similar ([4324141423] and [4413243143]). Additionally, PGA2 found three new solutions that define a somewhat better Pareto frontier than the one delineated by PGA1. Although PGA2's frontier is not totally convex, the approximation is very good.

In summary, for the four single-response systems, the RSM-based missing value estimation method improves the performance of PGA2 compared with PGA1. Note that although the response surfaces are modest approximations to the real outputs of each system, still the new method attains good results. Recall that we deliberately added a large noise to the responses. Given that the method works rather well, we believe it should perform even better in more favorable situations.

Results obtained for the multi-response 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 PGAs are optimizing a four–dimensional multi-response system. The estimation of the response surfaces for PGA2 is the same as the one already presented for the four single-response systems, i.e. the estimation procedure is the same for univariate and multivariate systems.

As in the previous analysis, 30 runs for PGA1 (see Canessa *et al*., 2014) and PGA2 were carried out. Figure 3 presents the solutions found by both PGAs. Note that in this case, the horizontal axis of Figure 3 shows the desirability function value φ_1 corresponding to mean adjustment and the vertical axis presents the desirability function φ_2 representing variance reduction. Remember that a larger value of φ_1 means a better mean adjustment and a larger value of φ_2 represents a smaller variance, so that the PGAs must maximize φ_1 and φ_2 (see Equation (4) and the appendix).

From Figure 3, one can see that both PGAs found three solutions. These solutions were consistently found in all 30 runs performed. Of them two are similar (solutions [3312442132] and [3421331242]) and one is a different solution ([2122334124] of PGA2). This solution of PGA2 defines a better approximation to the Pareto frontier. Given that the PGAs are dealing with a maximization problem, the frontiers should be concave, which is the case for the frontier defined by PGA2's solutions, whereas the one for PGA1 is not totally concave. Furthermore, since the Pareto frontier delivered by PGA2 is closer to the upper-right corner of the graph, which corresponds to the optimal point $\varphi_1 = \varphi_2 = 1{,}0$, this frontier is a closer approximation to the real unknown Pareto frontier.

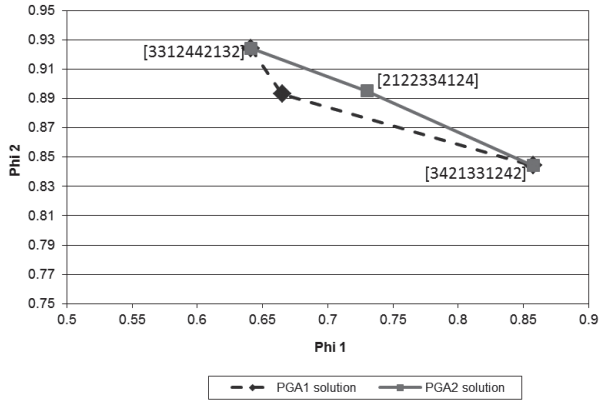

Figure 3. Pareto frontier for the solutions corresponding to the multi-response system.
Source: Authors

Note also that although the response surfaces are a modest approximation to the real mean and variance surfaces of the outputs of the system, given the large noise deliberately added to the responses (see Table 3, R^2 of surfaces between 0,5 and 0,35, except for the surface of the mean of y_1 with R^2 equal to 0,8, to which we added a small noise), the missing value estimation process based on RSM works rather well, enhancing the capability of PGA2 of finding a good approximation to the Pareto frontier. Hence, we should expect that under more favorable noise conditions, PGA2 should work even better.

Conclusions

The results of the application of the modified PGA (PGA2) to the multi-response system and single-response systems, showed that PGA2 outperforms PGA1. Most notably, PGA2 found better approximations to the Pareto frontier, even if the estimated response surfaces were a modest approximation to the real response surfaces of the outputs of the systems.

Additionally, for three out of four single-response systems, PGA2 found new Pareto efficient 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. 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 PGA. Thus, if the PGA 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 PGA2, 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) trough (8) and the number of data points that he or she will collect in the robust design experiment. Those considerations are important in the design of the experiment and will have an impact on the ability of the response surfaces to adequately represent the true surfaces that characterize the mean and variance of the outputs of the system (Myers & Montgomery, 2002). 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 (Myers & Montgomery, 2002). However, the results suggest that even a modest approximation to the true response surfaces, may enhance the performance of PGA2.

On the other hand, given that RSM is a good technique for estimating mean and variance of highly non-linear systems (Myers & Montgomery, 2002; Nair, Taam & Ye, 2002) the engineer should use PGA2 instead of PGA1 when he/she suspects that is the case. Remember that PGA1's missing value estimation procedure is totally linear, whereas the RSM method implemented in PGA2 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 PGA2 from the beginning would be sensible. Of course, using PGA2 comes to the expense of having to define the form of the response surfaces, per Expressions (5) to (8), but we think that is not a too high price to pay.

References

Allende, H., Bravo, D., Canessa, E., (2010). *Robust Design in Multivariate Systems using Genetic Algorithms*. Quality & Quantity Journal, Vol. 44, 315-332.

Canessa E., Droop, C., Allende, H., (2011). *An improved genetic algorithm for robust design in multivariate systems*. Quality & Quantity Journal, Vol. 42, 665-678.

Canessa E., Bielenberg, G., Allende, H., Robust (2014). *Design in Multiobjective Systems using Taguchi's Parameter Design Approach and a Pareto Genetic Algorithm*. Revista Facultad Ingeniería Universidad de Antioquia, Vol. 72, 73-86.

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

Laumans, M., *SPEA2-The Strength Pareto Evolutionary Algorithm 2*. Swiss Federal Institute of Technology (ETH) Zurich, 2008. Retrieved from: http://www.tik.ee.ethz.ch/~sop/pisa/selectors/spea2/?page=spea2.php. Accessed 10 October 2015.

Lin, C.D., Anderson-Cook, C.M., Hamada, M.S., Moore, L.M., Sitter, R.R., (2014). *Using Genetic Algorithms to Design Experiments: A Review*. Quality Reliability Engineering International, doi: 10.1002/qre.1591.

Maghsoodloo, S., Ozdemir, G., Jordan, G., Huang, CH., (2004). *Strengths and limitations of Taguchi's contributions to quality, manufacturing, and process engineering*. Journal of Manufacturing Systems, Vol. 23, Nr. 2, 73-126.

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

Nair V.N., Taam, W., & Ye K.Q., (2002). *Analysis of functional responses from robust design studies*. Journal of Quality Technology, Vol. 34, Nr. 4, 355-370.

Ortiz, F., Simpson, J., Pigniatello J. Jr., Heredia – Langner, A., (2004). *A Genetic Algorithm Approach to Multiple – Response Optimization*. Journal of Quality Technology, Vol. 36, Nr. 4 432– 449.

Roy, R.K., (2001). D*esign of Experiments Using the Taguchi Approach*. New York: J. Wiley & Sons.

Robinson, T.J., Borror, C.M., & Myers, R.H., (2004). *Robust Parameter Design: A Review*. Quality and Reliability Engineering International, Vol. 20, 81-101.

Taguchi, G., (1991) *System of experimental design*. Dearborn: American Supplier Institute.

Vandenbrande, W., (2000). *Make love, not war: Combining DOE and Taguchi*. In ASQ´s 54th Annual Quality Congress Proceedings, 450– 456.

Vandenbrande, W., (1998). S*PC in paint application: Mission Impossible?* In ASQ´s 52nd Annual Quality Congress Proceedings, 708-715.

Vining, G.G., & Myers, R.H., (1990). *Combining Taguchi and response surface philosophies: a dual response approach*. Journal of Quality Technology, Vol. 22, 38–45.

Wan, W., & Birch, J.B., (2011). *Using a modified genetic algorithm to find feasible regions of a desirability function*. Quality Reliability Engineering International, Vol. 27, 1173–1182.

Wu, C.F.J., & Hamada, M.S., (2009). *Experiments: Planning, Analysis, and Optimization*, 2nd ed. New York: J. Wiley & Sons.

Appendix A: Most important details regarding the building of the Response Surfaces and assessment of PGA2 performance

This appendix presents some important details related to the building of the response surfaces corresponding to the mean and variance of each of the responses of a system and the assessment of PGA2 performance. The first subsection shows how the response surfaces for the first test case were built, which corresponds to a single-response real system, but including qualitative and quantitative variables. Since this system has a conveniently small number of variables, we present many of the details of the modelling process and assessment of PGA2's performance. Then, subsection three presents only the most important steps carried out to

build the response surfaces for the complex multi-response system and the corresponding performance evaluation.

Calculations for the single-response real system

This system is a robotic automatic car painting system. To improve the process, first a manager, operators and quality control personnel of the painting line, along with technical support people from the paint suppliers brain-stormed the possible control and noise factors that might affect the width of a painted strip. Then, simple 2^n screening experiments were done and analyzed to narrow down the control factors and possible levels to the ones shown in Table 1. Those experiments were also carried out for the noise factors, which were the color of the paint, inlet air pressure and paint viscosity at two levels (Vandenbrande, 1998, 2000). Then, the RD 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. The collected data was then used to calculate a regression equation for mean and variance of the width of the painted strip using the functional forms stated in Equations (7) and (8). Hence, we started with a regression equation for mean and another for variance with the position parameter, two dummy variables representing qualitative factor A at three levels, three linear terms for quantitative factors B, C and D, and three quadratic terms for those same factors. Given that the multicollinearity statistics of the quadratic term of factor D showed that a problem existed, we removed that term from the regression equations. Thus, we arrived at Equations (9) and (10), with all coefficients statistically significant at least at the 0,05 level. Additionally, the models exhibited a high R^2 (see Table 2), which indicates a good fit of the data to the model, and hence we finished the process. Table A.1 shows the corresponding data used in the regression analyses.

Table A.1. Single-response real system data for calculating R. Surfaces

z_1	z_2	x_2	x_3	x_4	x_2^2	x_3^2	y_{mean}	$y_{variance}$
1	0	0	0	0	0	0	41,150	8,243
1	0	-1	-1	-1	1	1	44,800	1,393
1	0	1	1	1	1	1	35,825	12,563
0	1	1	0	-1	1	0	24,825	3,163
0	1	0	-1	1	0	1	45,175	17,703
0	1	-1	1	0	1	1	41,025	2,069
0	0	1	-1	0	1	1	43,200	65,860
0	0	0	1	-1	0	1	34,475	2,056
0	0	-1	0	1	1	0	41,050	76,170

Source: Authors

Table A.1 shows the independent variables corresponding to Equations (9) and (10), and the observations, which are the mean of 4 replications. The variables z_1 and z_2 are the two dummy variables representing factor A: 00 = level 3, 01 = level 2, 10 = level 1, x_2, x_3 and x_4 represent the first order coefficient for factors B, C and D, where 1 = level 1, 0 = level 2 and -1 = level 3, and x_2^2 and x_3^2 represent the quadratic coefficients for factors B and C. Recall that the independent variables were rescaled and centered on their means, so that the values of them lie between -1 and 1 (per the procedure suggested in Myers & Montgomery, 2002). The treatment combinations of independent variables correspond to the ones of the inner orthogonal array $L_9(3^4)$. The coefficients obtained in the regressions and corresponding to Equations (9) and (10) are shown in Table 2.

Those response surfaces were then inputted to PGA2 and the algorithm delivered the solutions presented in Figure 1. For calculating the corresponding mean adjustment and variance of each solution, the corresponding levels of each factor were introduced in the regression Equations (9) and (10). Then, using those values the corresponding $|y - T|$ and std. deviation were calculated. For example, for solution [1222], y = 37,4 + 1 x 1,017 + 0 x -2,567 + 0 x -3,838 + 0 x -3,642 + 0 x 2,992 + 0 x 0 x -1,813 + 0 x 0 x 5,075 = 38,417 and $|y - 40| = 1{,}583$. For variance, v = 44,506 + 1 x -40,629 + 0 x -40,384 + 0 x 0,325 + 0 x -11,378 + 0 x 16,637 + 0 x 0 x 17,536 + 0 x 0 x -12,251 = 3,877 and $s = 1{,}969$. The same procedure was done for calculating the mean adjustment and std. deviation of the other solutions presented in Figure 1.

Calculations for the multi-response complex system

The procedures for calculating the response surfaces and assess PGA2's performance are similar to the ones already described for the single-response system. However, note that here we have at our disposal the system's simulator (Canessa *et al.*, 2011), and thus we can perform experiments with more treatments and also verify the goodness of the solutions found by PGA2 directly calculating the system's responses using such simulator. Given that the multi-response system has four responses, 10 control factors each at four levels, and 1024 design points, we don't have enough space to present the data and the calculated regression coefficients. That is why we present only the final model in Equations (11) and (12). Those equations give the refined model we arrive at by beginning with a full model, i.e. a model containing a position parameter, all 10 linear and 10 quadratic terms, and all the 45 possible double interactions, which corresponds to Equations (5) and (6). The values of the independent variables of the regression analysis were rescaled and centered on their means, so that the values of them lie between -1 and 1 (per the procedure suggested in Myers & Montgomery, 2002). Given that we used 4 levels for each control factor, we had 40 values for the independent variables of the regression equations. The first regression equation showed that many of the coefficients exhibited problems of high multicollinearity, as suggested by a VIF above 10,0. Thus, looking at the multicollinearity diagnostic statistics, we began an elimination process of quadratic and double interaction terms that might have caused that problem. We mainly focused on those terms, because generally higher order terms are the ones that might cause multicollinearity (Myers & Montgomery, 2002). Finally, we arrived at the

models shown in Equations (11) and (12), which indicate that the average of the responses y_1, y_2, y_3 and y_4 and their variances were modelled by all the linear and quadratic terms of the 10 factors A,B,C,D,E,F,G,H,I and J, and double interactions AB, AC, AD and BC.

After we obtained the response surfaces corresponding to mean and variance of each of the four responses, these were inputted into PGA2 and 30 runs were done for each case study (i.e., for responses y_1, y_2, y_3 and y_4 of the one-response complex system and for the four-response system). PGA2 then delivered solutions for responses y_1, y_2, y_3 and y_4 , independent from each other (i.e., corresponding to four single-response complex systems) and solutions for the four-response system (i.e., the one in which we must simultaneously optimize y_1, y_2, y_3 and y_4). For each of those solutions, the corresponding levels of the 10 control factors, corresponding to the solutions, were inputted into the system's simulator (Canessa, Droop & Allende, 2011) and 30 runs were executed to obtain the value of the corresponding responses. Remember that the simulator is stochastic and was set with Gaussian noise, so that to adequately characterize each response, we must use the mean of each response over the 30 runs. Additionally, we calculated the standard deviation of the 30 runs for each response. For example, for response y_1 of the single-response system and solution [1432121433], we inputted into the simulator factor A at level 1, B at level 4, etc. and obtained 30 responses y_1. Then we calculated the mean of those values $\overline{y}_1 = 205{,}66$ and the std. deviation $s_1 = 4{,}79$, so that $|y_1 - T_1| = 5{,}66$, where $T_1 = 200$. That point (5,66, 4,79) was graphed in Figure 2 (a) for solution [1432121433]. The same calculations were done for the rest of the solutions for y_1 in Figure 2 (a) and also for the solutions delivered by PGA2 for the rest of the responses shown in Figures 2 (b), (c) and (d).

For the four-response system, remember that we used two desirability functions, φ_1 related to mean adjustment of the four responses and φ_2 associated with variance reduction (see Equation (4)). We explain some details of the process by which the value of φ_1 and φ_2 is calculated for solution [2122334124] shown in Figure 3. First, the corresponding values of the levels indicated by the solution were inputted into the system's simulator (i.e., A = level 2, B = level 1, etc.) and 30 runs performed. Then, using the corresponding 30 values obtained from the simulator, a mean and variance was calculated for y_1, y_2, y_3 and y_4. With those values, we calculated $(y_r - T_r)^2$ for each response and the corresponding values of φ_1 and φ_2. Table A.2 shows those values.

Table A.2. Calculation of desirability functions φ_1 and φ_2

Response	Mean	Variance	*Tr*	$(y_r - T_r)^2$
y_1	204,6	9,40	200	21,2
y_2	45,2	0,341	50	23,0
y_3	789,9	41,12	1000	44,1 x 103
y_4	465,2	20,92	500	1211,0
		$\varphi_2 = 0{,}895$		$\varphi_1 = 0{,}729$

Source: Authors

Unfortunately, we don't have enough space to present and explain all the 10 expressions and parameters necessary to calculate φ_1 and φ_2 beginning with the $(y_r - T_r)^2$ and the variance of each response. The reader can find the details in Canessa *et al*. (2014). In brief, the process entails calculating an individual desirability D_{1r} for each response by a linear interpolation between a maximum *a1r* and minimum b_{ir} value of $(y_r - T_r)^2$ for each response, which can be set by the experimenter. The same is also done for calculating individual desirability values D_{2r} for the variance of each response, with limits a_{2r} and b_{2r}. Given that we will have a desirability for each response, those are aggregated by simply taking their mean, both for $(y_r - T_r)^2$ and variance, obtaining $\overline{D_1}$ and $\overline{D_2}$. Then, a penalty function P_r is calculated for each response, which penalizes solutions with the mean response outside some upper H_r and lower L_r limits set up by the experimenter (see Equation (4)). This penalty function assigns a linearly calculated penalty to responses, whose means are outside those two limits. To aggregate the penalties over all the responses, the geometric mean of them is calculated $\overline{P}$. Finally, $\varphi_1 = \overline{D_1} - \overline{P}$ and $\varphi_2 = \overline{D_2} - \overline{P}$.