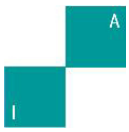
Inteligencia Artificial, 23(66) (2020), 9-25
doi: 10.4114/intartif.vol23iss66pp9-25



INTELIGENCIA ARTIFICIAL

http://journal.iberamia.org/

Forest-Genetic method to optimize parameter design of multiresponse experiment

Adriana Villa-Murillo, Andrés Carrión and Antonio Sozzi.
Universidad Viña del Mar. Life Sciences Department.
Viña del Mar. Chile. orcid.org/0000-0002-2645-4217.
adriana.villa@uvm.cl
Universitat Politécnica de Valencia. Dept of Applied Statistics, Operations Research and Quality.
Valencia. España. orcid.org/0000-0002-0953-2500.
acarrion@eio.upv.es
Universidad de la República. Soil and Water Department.
Montevideo. Uruguay.
asozzi@fagro.edu.uy

Abstract
We propose a methodology for the improvement of the parameter design that consists of the combination of Random Forest (RF) with Genetic Algorithms (GA) in 3 phases: normalization, modelling and optimization. The first phase corresponds to the previous preparation of the data set by using normalization functions. In the second phase, we designed a modelling scheme adjusted to multiple quality characteristics and we have called it Multivariate Random Forest (MRF) for the determination of the objective function. Finally, in the third phase, we obtained the optimal combination of parameter levels with the integration of properties of our modelling scheme and desirability functions in the establishment of the corresponding GA. Two illustrative cases allow us to compare and validate the virtues of our methodology versus other proposals involving Artificial Neural Networks (ANN) and Simulated Annealing (SA).
Keywords: Artificial Intelligence, Genetic Algorithm, Random Forest, Artificial Neural Networks, Multivariate Analysis.

Resumen
Proponemos una metodología para la mejora del diseño de párametros que consiste en la combinación de Random Forest (RF) con Algoritmos Genéticos (GA) en 3 fases: normalización, modelización y optimización. La primera fase corresponde a la preparacion previa del conjunto de datos mediante funciones de normalización. En la segunda fase, diseñamos un esquema de modelización ajustado a múltiples características de calidad, que hemos llamado Multivariante Random Forest (MRF) para la determinación de la función objetivo. Finalmente, en la tercera fase se obtiene la combinación óptima de los niveles de los parámetros mediante la integración de propiedades dadas por nuestro esquema de modelización y las desirabibity functions en el establecimiento del correspondiente GA. Dos casos ilustrativos nos permiten comparar y validar las virtudes de nuestra metodología versus otras propuestas que involucran Redes Neuronales Artificiales (ANN) y Simulated Annealing (SA).

Palabras Clave: Inteligencia Artificial, Algoritmos Genéticos, Random Forest, Redes Neuronales Artificiales, Analisis Multivariante.

ISSN: 1137-3601 (print), 1988-3064 (on-line)

1 Introduction

The robust design technique's proposed by Taguchi have to goal controllable input variables (control factor) of a system, so that its outputs (response variables) stay as close as possible to their corresponding target values and whit minimum variability, even in the presence of noise factor, which cannot be controlled [5]. Then, to accomplish those goal robust design use two methods: Tolerance design and parameter design. The first design seeks determine how much variability of the input parameter to the system will be allowed. Parameter design is a procedure which tries to reduce variability at a low cost [1]. All this is known as Quality Engineering with strengths but also limitations [21]

The demands of the current market involve more complex processes where multiple quality characteristics are usually considered. In these cases, the Taguchi methodology is limited [32]. In the literature some proposals for the improvement of the design of parameters for multi-response cases can be found, among the most recent ones, the following are found: Castillo *et al* [9] who proposed a modification to the function of desirability of Harrington E. [11] for optimize multiple answers but it is limited to non-differentiable cases. Su and Tong [26] proposed a method based on Principal Component Analysis (PCA), where the set of original responses is transformed into uncorrelated sets and subsequently optimized, Antony, J. [2] adapts the loss function of Taguchi to the case of multiple responses by assigning weights to each quality feature, Kim and Lin [18] present a modification of the desirability function (exponential) in the modelling of multiple responses and later Hsu *et al* [4] used those exponential functions in combination with ANNs for simultaneous optimization of the broadband tap coupler optical. Sarajit and Susanta [23] proposed to adjust multiple regression models to each quality characteristic separately to obtain the mean and variance estimate that will later allow to maximize the weighted signal-to-noise (SN) ratio; combining Taguchi's ideas with multiple regression techniques, Koksoy [22] considers each response as independent and uncorrelated for use the MSE criterion (Mean Squared of Error) in each response, but in real terms this is unlikely and Canessa *et al* [5] propose a Pareto genetic algorithm to finds the pareto frontiers of solutions to problems of robust design in multiobjetive systems.

All those previous proposals demonstrated great abilities in their respective illustrations and constitute great contributions for the improvement of the design of parameters. However, from a practical point of view, those approaches are complicated to apply and can only obtain the best solution from a specified set of levels of the control factors, ie they are unable to achieve an optimal combination of control factors when having continuous values [17].

Other proposals show the power of meta-heuristic and data mining techniques to solve problems involving continuous type factors through the work of Chang and Chen [16], Hou *et al* [12], Hsu *et al* [13], Huang and Hung [20], Su *et al* and [7]. More recently Chang[15] proposed the combination of ANN with SA for the case of multiple quality features as an extension of Su and Chang work's [6]. The ANNs are powerful techniques in the recognition of input/output patterns, but they are unstable in some cases and costly in the consumption of computational resources. Chang and Chen [17] modify the previous proposal, they recommend using GA instead of SA for the optimization.

Finally, Villa-Murillo *et al* [30] provides the Forest-Genetic method as a complete and statistically robust alternative for improving parameter's design, firstly considering experiment with a single quality characteristic as response variable [29] and later generalized to cases with multiple quality characteristic. This latter case is the objective of this article, which we called Multivariate Forest-Genetic Method, where we proposal an schema for design at Multivariate Random Forest to determine the objective function and using its properties in combination with desirability functions to drive the Genetic Algorithm into efficient optimization at the parameter levels.

This paper is organized as follows: in section 2 a review of the related works including brief description of the Classifications and Regression Trees as a base to the Multivariate Regression Tree and Random Forest are presented. This section also describes the operating process of the Genetic Algorithms and Desirability functions. Section 3 describes our proposal in 3 phases: normalization, modelling and optimization. Finally, in section 4 the effectiveness of the proposed approach with 2 illustrative examples is demonstrated and conclusions are provided in section 5.

2 Related works

2.1 Multivariate regression tree, Random Forest and Genetic Algorithms

The Classifications and Regression Trees (CART), defined for Breiman at 1984, consists of a recursive division of N cases on which a response variable Y and a set of predictors X are observed. Such partitioning procedure is known as regression tree when the response variable is numerical, and as classification tree when the response variable is categorical. [25]. For univariate responses, regression CART uses the exhaustive search method to select the splitting variable at each node. It considers binary split of the form $X \leq a$, so the method simultaneously finds covariate X and point a to minimize some node impurity criterion.

Multivariate Regression Tree (MRT) is a direct extension of CART. Its result are trees with the terminal nodes are composed of subgroups that minimize the sum of squares intra-group, by the next partition phase is defined by the threshold of the explanatory variables [8]. Each descendant node will depend directly on the threshold marked by the previous node, this makes it sensitive to the variations that can be created in the determination of the training and test sets. A determinant point in the formation of the nodes corresponds to the mean of the responses within each group; i.e., very different scales between the response variables will also noticeably affect the structure of the tree.

Dine *et al* [10] proposed a method for the construction of MRT by creating substitute variables resulting from the grouping of response variables of the same type, highlighting the benefits of this method only for descriptive purposes and leaving open the study for predictive cases because of the variability that entails the construction of a single tree. This variability significantly affects stability statistics of the modelling process and even create problems of over-adjustment.

Random Forest (RF) is based on the construction of trees of prediction through the use Bootstrap and Bagging, which makes the process stable [3]. No single tree is generated, a large number of trees are generated without pruning. These trees are constructed from bootstrap samples with replacement for a correct the prediction error and have an independent sample for each tree, called Out-of-Bag predictions. So, the observations that are not part of the bootstrap sample or subsample, respectively, are referred to as Out-of-Bag (OOB) observations and can be used for estimating the prediction error and performance of RF. This is justified since approximately one-third of the original sample is excluded from each sample generated by bootstrap. Finally, for each division of a node, a set of variables of a previously established size is selected and the selection of the division variable is restricted to that set. Remember that in CART, the best variable of the total set is selected. In this way, greater tree variability is included and the dependence of the result with past divisions is minimized. We note that, RF is a univariate methodology, but we have explained it in this section because it is the starting point for the design of our modelling strategy, defined as MRF.

The Genetic Algorithms (GA) are adaptive methods inspired by the biological theory of evolution formulated by Darwin in the middle 19th century. GA ideas are transferred to optimization problems quite naturally. The feasible solutions of a specific problem correspond to the members of a particular species, where the fitness of each member is measured by the value of the objective function. The current population in each iteration (generation) consists of a set of test solutions that are the living members of the species. Some of the younger members of the population (the fittest members) survive adulthood and become parents (randomly matched) who will have children (new test solutions) that would have some of the characteristics (genes) of the parents. As the fittest members of the population are more likely to become parents, GA tends to create improved populations in each generation. Sometimes mutations occur, so that new generations can adopt characteristics that are not possessed by parents. These mutations help the GA to explore a part of the feasible region, perhaps better, than previously considered. Finally, the survival of the fittest individual leads GA to a test solution (the best of all considered) closer to optimal.

2.2 Desirability functions

Commonly, when multiple responses are handled simultaneously as quality characteristics, quality measures may have different importance, different measurement scales or can be mutually opposed measures. All this hinders their simultaneous management. Harrington [11] provides a mathematical solution to these types of problems by transforming the estimated response vector $\hat{\mathbf{Y}}$ into a function d_i (e.g., the i-th estimaded response vector), defined as a desirability function. This is a value between 0 and 1, which increases as much as the desirability of the response increases, then form a composite function D defined as the geometric mean of the d_i. The objective of optimization for the parameter design is to find the set of independent variables that maximize the value of D.

Many proposals have arisen around the functions of exponential desirability among which are Kim and Lin [18], Hsieh *et al* [19], Wu and Yeh[31] and Chang [14] who makes a slight modification to Wu's proposal[31]. More recently Sujarit and Susanta[27] extend the concept of desirability function for ordinal response variables. Such proposals have as a common factor the classification of the functions of desirability according to the quality characteristic of the vector $\hat{\mathbf{Y}}$ as alternatives to the measures of performance proposed by Taguchi. The present work adopts the approach of Chang [14], who defines the functions of desirability based on three response types as follows.

Nominal is better (NTB) type with upper specifications limit (USL) and lower specification limit (LSL), the desirability function of the d value denoted by d_i^{NTB}

$$d_i^{NTB} = \exp\left(-\left|\frac{2\hat{y}_i - \left(y_i^{\max} - y_i^{\min}\right)}{y_i^{\max} + y_i^{\min}}\right|\right) \tag{1}$$

Small is better (STB) type with USL, the desirability function of the d value denoted by d_i^{STB}

$$d_i^{STB} = \exp\left(-\left(1 + \frac{\hat{y}_i - y_i^{\max}}{y_i^{\max}}\right)\right) \tag{2}$$

Larger is better (LTB) type with LSL, the desirability function of the d value denoted by d_i^{LTB}

$$d_i^{LTB} = \exp\left(-\exp\left(-\frac{\hat{y}_i - y_i^{\min}}{y_i^{\min}}\right)\right) \tag{3}$$

Where $y_i^{\min}$ and $y_i^{\max}$ correspond to the maximum and minimum limits in each case.

3 Proposed approach

The proposed methodology, which is called Multivariate Forest-Genetic Method, consists of 3 phases: normalization, modelling and optimization. We assume a dynamic multi-response problem defined as:

$$\mathbf{Y} = f_i\left(M_j, \mathbf{X}_k, Z_l\right) + \varepsilon_{ijkl} \tag{4}$$

Where $f_i\left(M_j, \mathbf{X}_k, Z_l\right)$ denotes the function between the ijkl-th vector response $\mathbf{Y}$ and the corresponding j-th combination of signal factor M, k-th level of control factor $\mathbf{X}$ and the l-th level of the vector noise Z. Finally ε_{ijkl} represents an error term.

Normalization phase
The initial phase comprises the preparation of the data set by the normalization function (5) recommended by Villa-Murillo[29], to minimize the present variability. In addition, our methodology may be an alternative in the use of ANNs as a technique for establishing the relation input/outputs, so for the purpose of numerical comparisons is important the equality of scales in the dataset.

Table 1: Metric of modelling for RFM

$P_1(Y) = \sum_{i=1}^{r} y_i$	$P_5(Y) = \frac{1}{3}\sum_{i=1}^{r} y_i^2$
$P_2(Y) = \left(\prod_{i=1}^{r} y_r\right)^{\frac{1}{r}}$	$P_6(Y) = \frac{1}{3}\sum_{i=1}^{r} \exp(y_i)$
$P_3(Y) = \left(\prod_{i=1}^{r} y_i^{-1}\right)^{\frac{1}{r}}$	$P_7(Y) = \frac{1}{3}\sum_{i<s}^{r} \frac{y_i-\mu_s}{\sigma_s}$
$P_4(Y) = \left(\prod_{i=1}^{r} y_i^2\right)^{\frac{1}{r}}$	

$$f_2(y) = \begin{cases} 0 & y = 0 \\ \frac{1}{1+\exp(-y')} & \text{in other cases} \end{cases} \qquad \text{with} \qquad y' = \frac{y - min}{max - min} \tag{5}$$

Modellig phase
The modelling phase comprises the MRF design, which starts with the unification of the response vector $\mathbf{Y}$, considering modelling metrics that will be defined by the multiple quality characteristics for each combination of parameters and will have as final product the estimation of the responses through a sequence of univariate regression trees and the allocation of weights. We propose the study of several metrics considering the possible quadratic effects or correlations existing in $\mathbf{Y}$. Table 1 presents such functions, assuming a response vector $\mathbf{Y} = (y_1, ..., y_r)$ for each i-th observation.

The study of these functions will lead to algorithms MRF_{Pp} $(p = 1, ..., 7)$, i.e the Multivariate Random Forest for each P_p metric, with $(p = 1, ..., 7)$, Table 1. Each study begins with the adjustment of the MRT and the calculation of the corresponding $RMSE_{MRT}$ (Root Mean Square Error of the adjusted Multivariate Random Tree), and finalizes with the choice of the algorithm that provides the lowest RMSE to guarantee the greatest robustness in the modelling phase. For the MRT setting the parameters $minsplit, minbucket$ y cp must be determined, as well as your RMSE that we will call $RMSE_{Tree_p}$. Where $minsplit$ is the minimum number of observations at each node, $minbucket$ is the minimum number of observations at the terminal nodes and cp is the complexity parameter. To adjust MRF_{Pp}, we must define the number of trees to be assembled (n_{tree}) and the number of variables present in each division $(mtry)$, this last parameter must be optimized by a preliminary study of the OOB error rate (out of bag error rate).

Optimization phase
The optimization phase, based on GA, consists of the integration of the desirability functions and the measures the importance of variables thrown by MRF for the design of the genetic operators. We consider that a dynamical system can be treated as a static multi-target system [28], which allows us to establish the desirability functions defined in (1), (2) and (3) as our performance (PMs) according to the respective response types, where $y_{ij}^{\min}$, $y_{ij}^{\max}$ represents the lower and upper limits of specification of the i-th response in the j-th level of signal factor, k-th leven of control factor and l-th of noise factor.(Table 2).

Harrington [11] defines one Overall Performance Index (OPI) to evaluate the performance of quality responses simultaneously. This index is defined as the geometric mean of the quality responses in their corresponding desirability function as presented in (6) and that we incorporate into our GA as our fitness evaluation function.

Table 2: PMs according to the type of response quality

DNTB: $d_i = \exp\left(-\frac{1}{sn}\sum_{j=1}^{s}\sum_{k=1}^{n}\left\lvert\frac{2\hat{y}_{ijkl}-\left(y_{ij}^{\max}+y_{ij}^{\min}\right)}{y_{ij}^{\max}+y_{ij}^{\min}}\right\rvert\right)$
DSTB: $d_i = \exp\left(-\left(1+\frac{1}{sn}\sum_{j=1}^{s}\sum_{k=1}^{n}\frac{\hat{y}_{ijkl}-y_{ij}^{\max}}{y_{ij}^{\max}}\right)\right)$
DLTB: $d_i = \exp\left(-\exp\left(-\frac{1}{sn}\sum_{j=1}^{s}\sum_{k=1}^{n}\frac{\hat{y}_{ijkl}-y_{ij}^{\min}}{y_{ij}^{\min}}\right)\right)$

$$OPI = \left(\prod_{j=1}^{r} d_j\right)^{1/r} \tag{6}$$

Where d_j represents the desirability value in the j-th vector $\mathbf{Y}$, with $(j = 1, ..., r)$.

The optimization phase is initiated by randomly forming n chromosomes. These chromosomes are formed by $(2 + p)$ gen, the signal factor $M_j, (j = 1, ..., s)$, the noise factor $Z_l, (l = 1, ..., n)$ and the vector of k control factor $X_k, (k = 1, ..., p)$ in normalized scale.

We used the crossing of a point but assigning weights to the genes of the vector of control factors $\mathbf{X}$ by the measure the importance of the variables given by the MRF_{Pp}. In this way, we increased the probability of crossing the chromosomes around the genes of greater importance in the design. The weight of each gene is defined by equation (7), where I_{x_k} represents the importance by k-th gen in $MRF_{Pp}, (k = 1, ..., p)$.

$$PC = \frac{I_{x_k}}{\sum_{k=1}^{p} I_{x_k}} \tag{7}$$

For the rate and type of mutation used, we propose a complete scheme consisting of 2 mutation rates, considered with and without gene replacement, in 4 different functions. We propose 2% and 5% as a mutation rate, which directly influences the number of mutated genes in each generation. We then studied the possibility that the same gene may or may not be mutated in different generations, which we have defined as *mutation with or without replacement*. Finally we define 4 mutation functions $m(g)$ that determine the search field of the algorithm: the function $m_1(g)$ and $m_2(g)$ are non-uniform mutations designed to alter the gene within the range of $\mathbf{X}$, $m_3(g)$ and $m_4(g)$ are uniform mutations that could allow, in some cases, for the algorithm to search outside of the range of $\mathbf{X}$. The functions are presented in table 3, where g is the gene to mutate in the corresponding chromosome and g_{min}, g_{max} the minimum and maximum values of the gene in the population.

MRF_{Pp} has its basis in CART, so any new observations (chromosome) will be adjusted in the limits of the corresponding node. This limits the GA in its exploratory task. So, considering that the crossing operation corresponds to a genetic exchange between chromosomes and that in MRF_{Pp} the genes are represented by nodes, we assign as response value to the new chromosomes the interpolation between those nodes in the gene of greater weight of the chromosome. The above is summarized in equation (8) where $\hat{y}_{m_i}$ y $\hat{y}_{p_i}$ are the prediction of the i-th mother and the i-th father coming from MRF_{Pp} . Finally, g_{p_i}, g_{m_i} and g_{h_i} are the values of the i-th father, i-th mother and the i-th children with the highest weight in the model MRF_{Pp}.

Table 3: Mutations functions

$m_1(g) = \begin{cases} g + g(0.10) & g < g_{max} \\ g & \text{other cases} \end{cases}$
$m_2(g) = \begin{cases} g - g(0.10) & g > g_{min} \\ g & \text{other cases} \end{cases}$
$m_3(g) = g + 1$
$m_4(g) = g - 1$

$$\hat{y}_{h_i} = \left| \frac{\hat{y}_{m_i} (g_{p_i} - g_{h_i}) - \hat{y}_{p_i} (g_{m_i} - g_{h_i})}{g_{p_i} - g_{m_i}} \right| \tag{8}$$

Figure 1 shows the schema of the our approach and the procedure of each phase is described below.

Normalization phase

Initialization: Normalize the data set with (5)

Modeling phase

Step 1: Adjust MRT and calculate $RMSE_{MRT}$

Step 2: Adjust and train MRF_{Pp} algorithms as follows:

- Do $p = i$ where $i = 1, 2, ..., 7$
- Apply the function Pp in the vector $\mathbf{Y}$
- Adjust $Tree_{Pp}$ using the set of data obtained in the previous step to determine $minsplit, minbucket$, cp and $MRSE_{Tree_p}$
- Adjust RFM_{Pp} by determining (n_{tree}) and $(mtry)$
- Randomly divide the sample into train and test set
- Train the MRF_{Pp} model with training set and obtain the predictions of the test responses using this model

Step 3: $\forall p$, calculate $RMSE_{MRF_{Pp}}$ from the predictions obtained in test by MRF_{Pp}. If $MRSE_{Tree_{Pp}} < RMSE_{MRF_{Pp}}$ go back to the step 2, Otherwise go to the step 4

Step 4: Calculate $f^* = argmin(RMSE_{MRF_{Pp}})$, this is, the algorithm MRF_{Pp} that gives the lowest RMSE.

Optimization phase

Step 5: Do z=b

- Use MRF_{Pp} to calculate the vector $\hat{\mathbf{Y}} = (\hat{y}_1, ..., \hat{y}_r)$
- Apply equation (7) for the assignment of the weight of importance to each gene

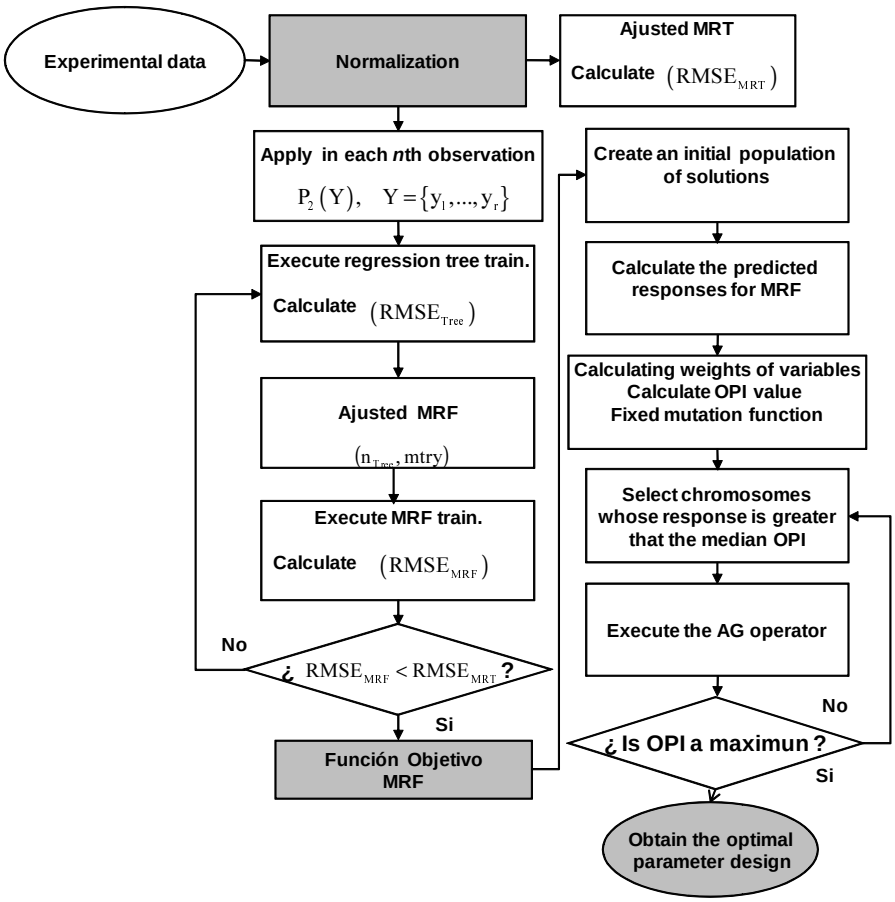


Figure 1: The schema of the Multivariate Forest-Genetic algorithm

- Transform the set of chromosomes with their corresponding predictions to the initial scale. So it is formed the initial population. We will call the population a_z , so the initial population $z = b$ is call $a_z = a_b$ where a_{bi} is the chromosome i-th of the b population

Step 6: Evaluation Calculate the by PMs each response type of vector $\hat{\mathbf{Y}}$ according to the table 2 and OPI measure with equation (6)

Step 7: Selection Calculate the median OPI (We will call $\widetilde{OPI}$), and Select all chromosomes with values greater than $\widetilde{OPI}$. This selection will be called a'_{bi} with $a'_{bi} = \left\{ a_{bi} \,;\, OPI\,(a_{bi}) > \widetilde{OPI} \right\}$

- Crossing: To determine randomly the factor of vector $\mathbf{X}$ as crossing point and make cross $\forall a'_{bi}$
- Mutation: in all a'_{bi} use the function $m(g)$ (table 3) or the mutation of genes with t rate mutation and the q replacement factor
- Prediction for the new generation: $\forall a'_{bi}$ calculate the responses vector $\hat{\mathbf{Y}}$ with equation (8)

Step 8: Replacement The union of the parent chromosomes a_{bi} and the children a'_{bi} with his $\hat{\mathbf{Y}}$ will form the new generation. This is $a_{Bi} = a_{bi} \bigcup a'_{bi}$

Step 9: Convergence Calculate $OPI_{a_{Bi}}$. If $\max[OPI\,(a_{Bi})] \leq OPI\,(a_{bi})$, do $z = b + 1$ and go back the step 7. Otherwise go at step 10

Step 10: Obtain the optimal combination of levels of the control factor $\mathbf{X}$ and its corresponding multiresponse.

4 Numerical examples

Two case studies are considered in order to perform an exhaustive numerical analysis. We considered the first case as a pilot, then the results obtained are extend for the second case study and thus, to verify the virtues and deficiencies of the proposed methodology. These cases correspond to studies adopted by Chang in the year 2008[15] and 2001[17]. This cases consist of 3 variable responses (y_1, y_2, y_3), obtained by simulating Monte Carlo and considering the method of Park and Yum[24], where the objective is the simultaneous optimization of the 3 responses under the quality characteristics DLB, DNB and DSB respectively.

4.1 Parameter control

All the procedures and routines used in the design of our algorithms have been programmed with R Language (version 2.13.1) where it has been necessary to use some libraries and adjust the control parameters of the modelling algorithms. All these are presented bellow.

- Adjusted MRT :minsplit= 10, minbucket= 2, cp=0.0001, xval= 10, library mvpart, version 3.1-46
- Adjusted CART: minsplit= 10, minbucket= 2, cp=0.0001, xval= 10, library rpart, version 1.4-0
- Adjusted MRF: library randomForest version 4.5-33, $n_{Tree} = 1000$, estimation of RMSE through OOB no pruning criterion is established between trees.

These parameters are common for the two case studies, with the exception of *mtry* parameter because this depends on the data set.

4.2 Example 1

The data set are presented under an orthogonal matrix L_{18} with 6 control factors $\mathbf{X} = (x_1, x_2, x_3, x_4, x_5, x_6)$. The data set are show in table 4. Signal factor has 3 levels $M_1 = 10$, $M_2 = 20$ and $M_3 = 30$, Noise factor has 2 levels N_1 y N_2. The specification values of the 3 responses are shown in table 5.

In order to make the comparison with the work present by Chang [15], the data is randomly divided into two sets. The training set consists of 185 observations and the test set with 31 observations.

4.2.1 Normalization and modelling phase

Function (5) is used for the normalization of the data. Figure 2 shows the RMSE values obtained by the initial MRT, MRF and ANN, where Pp, $(p = 1, ..., 7)$ represents the 7 metric of the table 1 and (7-13-3), (8-12-3) the best ANN architecture presented in the works of Chang (2006) [14] and Chang (2008)[15].

It can be noted in figure 2(a), that P_7 is totally inefficient in obtaining the RMSE, with values higher than given by MRT. For this reason it is excluded as a modelling metric and of the figure 2(b), where P_2 is observed with the lowest RMSE values in training and test set. Notice how the highest RMSE values correspond the 2 ANN architectures compared to all the metrics used.

Table 6 also shows the RMSE obtained from MRT and MRF, where the objective is the determination of the algorithm that provides the lowest RMSE, ie $f^* = argmin(RMSE_{RFM_{Pp}})$. For this we defined in the table the measure $\mathbf{P}\%$ as the progress of the values obtained in each MRF algorithm versus the initially adjusted MRT, observing the training and test sets. A possible quadratic effect on the data leads a lower RMSE values in the P_2 ,P_4 and P_5 functions, and consequently a better performance of MRF: 70.60%, 63.19% and 63.43% in training set and 42.81% 32.53% and 35.27% in test set. In contrast, P_1 obtains a RMSE close to P_2, 69.68% in training but less than P_4 and P_5 in test, This translates to an over-adjustment of the model RFM_{P1}. Then RFM_{P1} is chosen as the modelling scheme, since it presents the smaller RMSE and the smaller discrepancies between the training and test sets.

Table 7 compares the RMSE of RFM_{P2} with the obtained by Chang (2006)[14] and Chang (2008)[15]. Note that our modelling proposal gets a RMSE 5 times lower in 1000 interactions versus 3000 and 4500

Table 4: The experimental data. Example 1

No.	N	Responses																	
		y_1						y_2						y_3					
		$M_1=10$		$M_2=20$		$M_3=30$		$M_1=10$		$M_2=20$		$M_3=30$		$M_1=10$		$M_2=20$		$M_3=30$	
1	$N1$	61.6	78.2	128.0	106.0	230.6	226.9	7.4	7.2	16.7	13.2	23.7	24.1	1.9	1.9	4.6	3.7	7.6	4.6
	$N2$	70.8	57.1	137.3	160.3	282.2	252.5	9.1	10.2	22.8	17.8	26.2	26.7	2.0	2.1	3.9	4.8	4.7	4.3
2	$N1$	88.3	93.6	175.2	181.5	259.7	304.5	10.1	8.8	23.4	22.6	29.6	30.3	1.8	2.0	4.0	2.8	6.0	3.3
	$N2$	72.9	72.7	174.0	145.5	258.4	214.6	8.7	9.2	19.1	24.1	31.4	30.5	1.9	2.2	3.8	3.5	6.7	6.1
3	$N1$	80.8	81.1	154.3	157.4	238.1	237.8	10.8	10.5	20.1	20.9	30.6	32.4	1.0	3.2	4.4	5.4	8.1	6.7
	$N2$	77.2	83.3	167.1	159.0	251.8	257.9	10.6	10.9	21.8	23.7	30.5	32.7	1.4	2.6	3.8	3.9	8.0	2.7
4	$N1$	65.9	71.3	179.2	151.5	196.1	221.6	7.6	7.2	15.3	14.8	22.5	22.2	1.7	2.2	4.1	4.9	5.9	6.0
	$N2$	83.7	78.4	135.6	177.0	246.9	291.8	8.1	7.6	14.1	14.7	21.9	21.8	2.1	2.6	3.6	3.3	7.7	6.9
5	$N1$	79.4	88.6	121.9	151.6	248.8	245.1	11.9	12.7	25.6	25.7	36.6	35.7	2.0	2.1	2.7	3.8	4.8	5.8
	$N2$	67.8	87.3	113.6	141.3	171.5	244.7	10.5	11.8	25.8	26.2	39.1	33.1	2.6	1.7	4.1	3.6	5.5	5.2
6	$N1$	90.5	87.0	161.8	169.4	286.9	236.5	10.2	10.2	23.7	21.8	32.0	32.8	1.8	2.2	2.5	4.1	4.9	5.4
	$N2$	87.6	87.8	160.7	163.9	231.4	288.7	10.7	11.2	22.5	20.0	34.2	28.0	1.2	2.0	4.0	4.4	6.6	3.1
7	$N1$	80.9	74.7	165.9	163.4	232.2	246.4	11.7	12.1	23.3	23.3	33.5	32.7	1.9	1.6	5.2	5.1	6.6	5.3
	$N2$	69.9	78.7	141.7	159.1	260.4	239.7	11.6	11.6	22.0	22.5	33.8	34.0	1.9	1.7	3.6	5.1	5.8	5.1
8	$N1$	92.3	71.7	185.7	154.3	233.1	240.5	8.3	8.4	16.9	18.5	28.1	27.3	2.8	1.5	3.4	3.4	6.2	3.7
	$N2$	104.8	89.4	216.1	173.0	340.6	308.9	8.2	6.1	18.5	15.1	29.1	21.1	1.8	2.2	4.1	4.3	9.0	5.0
9	$N1$	92.8	59.8	130.7	142.2	257.3	266.0	8.6	9.7	17.3	18.4	30.9	31.5	2.3	1.7	4.5	5.1	7.4	6.5
	$N2$	82.1	87.0	175.1	138.2	161.4	274.4	8.2	8.4	18.6	18.0	31.3	30.9	2.2	1.2	4.0	4.0	7.2	6.7
10	$N1$	86.0	100.0	179.1	175.4	246.9	244.9	6.9	6.6	15.8	16.4	22.2	24.9	1.7	1.7	3.9	5.8	5.2	8.4
	$N2$	81.9	91.2	190.8	137.8	293.5	227.1	8.3	7.2	14.3	16.6	23.1	24.2	2.1	1.8	3.9	4.8	4.5	8.0
11	$N1$	76.3	78.2	140.0	154.7	264.1	260.1	9.5	9.2	27.9	25.4	32.3	33.8	1.0	1.7	4.1	4.0	3.9	5.6
	$N2$	67.1	76.0	169.5	175.4	239.2	251.8	10.8	12.2	23.6	20.4	26.8	29.2	2.1	2.3	5.0	4.6	6.7	4.5
12	$N1$	91.4	81.8	160.1	167.2	238.5	197.2	11.8	10.5	22.9	25.9	36.9	36.8	2.1	1.9	4.0	3.7	7.4	6.2
	$N2$	85.1	63.8	123.8	166.4	233.5	242.0	11.6	11.5	22.1	22.6	32.5	31.9	1.6	0.9	4.2	4.1	2.1	6.9
13	$N1$	87.9	82.5	146.5	167.0	212.4	222.2	10.1	10.3	19.9	22.2	27.1	27.7	2.0	2.3	5.0	4.7	6.7	7.3
	$N2$	57.4	78.5	91.6	182.3	250.3	207.3	10.6	10.1	20.1	19.2	27.8	24.6	2.2	2.1	4.5	4.7	7.4	6.2
14	$N1$	88.1	78.1	156.7	170.1	239.1	215.2	12.1	10.0	24.1	20.2	28.0	35.6	1.7	2.1	3.8	3.3	7.2	4.9
	$N2$	81.7	75.7	140.2	127.8	241.4	211.7	11.3	11.8	23.3	24.4	32.1	38.9	1.8	2.7	3.1	3.8	4.9	4.7
15	$N1$	101.8	78.2	168.3	180.8	240.6	235.2	10.2	7.6	14.9	19.4	26.9	19.5	1.7	2.2	5.3	3.6	5.0	7.5
	$N2$	80.4	76.5	206.7	222.6	325.1	285.2	8.7	7.2	16.8	14.1	26.6	25.1	2.2	1.9	4.8	2.6	3.8	5.5
16	$N1$	77.4	75.4	171.7	159.0	201.3	219.7	10.4	10.6	20.7	22.2	34.4	30.0	1.9	2.3	3.1	4.2	5.8	3.3
	$N2$	72.0	69.5	189.1	168.6	254.3	237.3	11.1	10.9	20.6	21.4	30.5	31.4	2.1	1.6	4.1	4.8	5.4	6.9
17	$N1$	71.4	69.2	145.0	152.5	223.8	218.7	8.8	8.4	19.0	13.8	26.2	24.1	1.5	2.1	3.8	4.4	6.9	4.2
	$N2$	77.0	70.5	158.4	154.0	218.4	224.1	9.2	9.0	16.7	17.4	27.1	26.2	1.6	1.8	3.7	4.4	4.6	4.2
18	$N1$	82.8	67.8	183.7	175.5	276.1	254.4	10.7	9.2	19.8	20.1	27.3	31.0	2.5	2.2	3.1	3.4	7.7	7.4
	$N2$	85.2	92.0	154.4	157.6	249.3	286.1	11.3	7.7	19.3	22.6	29.4	26.6	1.6	1.6	3.9	4.7	6.0	7.3

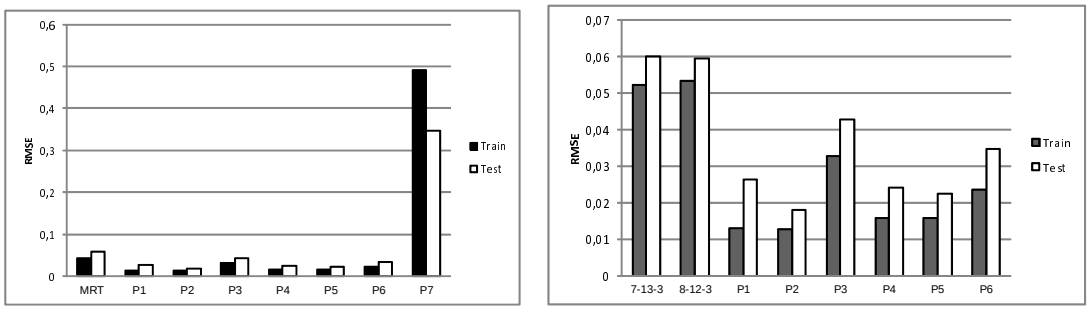
Table 5: The specifications for the responses. Example 1

Responses	y_1	y_2		y_3
Bounds	$y_{jk}^{\min}$	$y_{jk}^{\min}$	$y_{jk}^{\max}$	$y_{jk}^{\max}$
M_1	55	7	13	3
M_2	110	14	26	6
M_3	165	21	39	9

Table 6: RMSE progress of MRF_{Pp} front MRT. Example 1

Function	RMSE		**P**%	
	Train	Test	Train	Test
MRT	0.0432	0.0584	–	–
RFM_{P1}	0.0131	0.0264	69.68	28.77
RFM_{P2}	0.0127	0.0182	**70.60**	**42.81**
RFM_{P3}	0.0327	0.0429	24.31	0.51
RFM_{P4}	0.0159	0.0242	63.19	32.53
RFM_{P5}	0.0158	0.0226	63.43	35.27
RFM_{P6}	0.0236	0.0347	45.37	14.55

required for ANN in the training and in test set. In conclusion, our proposal allows a considerable saving



(a) Metric modelling (b) Metric modelling and the ANN

Figure 2: Compared the RMSE in the metric and the ANN

of computational resources.

Table 7: Compared the RMSE in the MRF and ANN

Proposal	Iteractions	RMSE	
		Training	Test
Chang (2008)[14]	30000	0.0524	0.0601
Chang (2008)[15]	4500	0.0573	0.0595
RFM_{P2}	1000	0.0127	0.0182

4.2.2 Optimization phase

400 chromosomes are randomly generated, whose structure is made up of a total of 8 genes: 6 corresponds to the vector of control factors $\mathbf{X} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, 1 corresponds to the signal factor $M = \{10, 20, 30\}$ and 1 corresponds to the noise factor $Z = \{N_1, N_2\}$. These chromosomes will be in a normalized scale in order to calculate their corresponding responses with our model MRF_{P2}. The objective of the present case is the simultaneous optimization of the response vector $\mathbf{Y} = \{y_1, y_2, y_3\}$ under the quality characteristics DLTB, DNTB and DSTB correspondingly. Desirability functions (table 2) and OPI (equation (6)) are determined by the expression (9), which represented the *fitness* function of the algorithm.

$$\begin{aligned} OPI \quad = \quad & \left\{ \exp\left[-\exp\left(-\frac{1}{6}\sum_{j=1}^{3}\sum_{l=1}^{2} \frac{\hat{y}_{1jkl} - \hat{y}_{1j}^{\min}}{\hat{y}_{1j}^{\min}} \right) \right] \cdot \exp\left(-\frac{1}{6}\sum_{j=1}^{3}\sum_{l=1}^{2} \left| \frac{2\hat{y}_{2jkl} - \left(y_{2j}^{\max} + y_{2j}^{\min}\right)}{y_{2j}^{\max} - y_{2j}^{\min}} \right| \right) \right. \\ \cdot \quad & \left. \exp\left[-\left(1 + \frac{1}{6}\sum_{j=1}^{3}\sum_{l=1}^{2} \frac{\hat{y}_{3jkl} + y_{3j}^{\max}}{y_{3j}^{\max}} \right) \right] \right\}^{\frac{1}{3}} \end{aligned} \tag{9}$$

Table 8 shows the important values (I_{xk}, $k = 1, ..., 6$) produced by MRF_{P2} and the weights assigned to the genes for crosses are calculated by equation (7).

Table 8: Measure of importance and weight of the genes by MRF_{P2}. Example 1

Gen	x_1	x_2	x_3	x_4	x_5	x_6
I_{xk}	0.403	**0.602**	0.398	0.464	0.436	0.459
Weight	0.146	**0.218**	0.144	0.168	0.158	0.166

Note that x_2 is the gen with greater value of importance which will later determine the function (10) of the present case, this is show in (10).

$$\hat{y}_{h_i} = \left| \frac{\hat{y}_{m_i}\left(x2_{p_i} - x2_{h_i}\right) - \hat{y}_{p_i}\left(x2_{m_i} - x2_{h_i}\right)}{x2_{p_i} - x2_{m_i}} \right| \tag{10}$$

where $\hat{y}_{m_i}$ y $\hat{y}_{p_i}$ represent the la prediction of i-th mother and i-th father coming of MRF_{Pp}; $x2_{p_i}, x2_{m_i}$ and $x2_{h_i}$ correspond of the gen value x_2 of i-th father, i-th mother and i-th children.

We propose as a mutation scheme the 2% and 5% rates, allowing the same gene to be mutated repeatedly with 4 functions (table 3). All this produces 16 algorithms that were processed for 15 generations. As all algorithms converge at the same OPI value after the first generation, then we have extracted this chromosome with the aim of avoiding a local optimum and thus to explore further the behaviour of the algorithms. So, 16 algorithms are processed again in 15 generations. Now all the algorithms found the optimum after the fifth generation. We have taken the results of the fifth generation of each algorithm and performed in respective analysis of variance (ANOVA), where the quadratic transformation of the OPI response was necessary to fulfil the assumptions of the ANOVA. table 9 shows these result.

Table 9: ANOVA for mutations functions. Example 1

Main effects	SS	Gl	MS	F	P-value
A: Reemp	0,0359505	1	0,0359505	2,20	0,1378
B: Rate	0,302677	1	0,302677	18,54	0,0000
C: Type	0,861467	3	0,287156	17,59	0,0000
Interactions					
AB	1,44034	1	1,44034	88,23	0,0000
AC	0,282535	3	0,0941782	5,77	0,0006
BC	0,986153	3	0,328718	20,14	0,0000
Error	101,46	6215	0,0163251		
Total	105,413	6227			

Table 9 shows high significance in the rate of mutation, type of mutation and their interactions. So, we used the Tukey HSD test at 95%. Figure 3 shows such test for the interaction type and mutation rate, where the numbers $1, 2, 3, 4$ represent the mutation functions $m_1(g)$, $m_2(g)$, $m_3(g)$ y $m_4(g)$ respectively. The maximum OPI observed is $m_1(g)$ with 5%. Said scheme was adopted in the establishment of GA, in addition we have decided to allow mutation with replacement in order to expand the GA search.

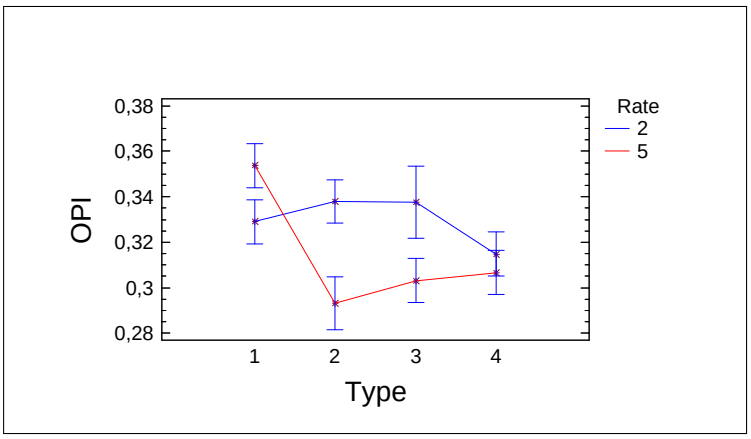


Figure 3: HSD Tukey for interactions Type-Rate

Chang[15] shows in his work the result obtained in comparison with the 5 best combinations obtained in his previous work at 2006[14], where the proposal is based in ANN. In the table 10 we have decided to present these results and add ours to determine numerically the validity of our proposal. As well as

the Chang [15] proposal at 2008, we were able to explore intermediate values in the ranges of control factors, which in cases of real applications could mean material savings. We exceeded the OPI value in both works of Chang, which allows us to verify the efficiency of our proposal.

Table 10: The results of the implementation. Example 1

Approach	Control factor values						OPI
	x_1	x_2	x_3	x_4	x_5	x_6	
Forest-Genetic	1.18	2.77	2.60	1.64	1.53	1.33	**0.673122**
Chang (2008)[15]	1.98	2.98	1.06	2.44	2.96	1.06	0.668877
Chang (2006)[14]	1	3	1	3	2	1	0.626443
	1	3	1	2	1	1	0.617955
	2	3	1	3	2	1	0.483897
	2	3	1	2	1	1	0.482461
	3	3	1	3	2	1	0.409331

4.3 Example 2

The data set are presented under an orthogonal matrix L_{18}, the control factor vector is defined as $X = \{A, B, C, D, E, F\}$ and the signal factor is defined with the levels $M_1 = 0.1$, $M_2 = 0.2$ y $M_3 = 0.3$. The experiment was conduced with 2 replicates. The data set are show in table 12 and the specification values for the 3 responses in the table 11. In order to make the comparisons with the work presented by Chang at 2011[17], the data set are randomly divided into two: the training set consisting of 92 observations and the test set with 16 observations.

Table 11: The specifications for the responses. Example 2

Responses	y_1	y_2		y_3
Bounds	$y_{jk}^{\min}$	$y_{jk}^{\min}$	$y_{jk}^{\max}$	$y_{jk}^{\max}$
M_1	4.8	0.6	1.4	28
M_2	9.6	1.2	2.8	56
M_3	14.4	1.6	4.2	84

Table 12: Experimental data. Example 2

No.	Responses																	
	y_1						y_2						y_3					
	$M_1 = 0.1$		$M_2 = 0.2$		$M_3 = 0.3$		$M_1 = 0.1$		$M_2 = 0.2$		$M_3 = 0.3$		$M_1 = 0.1$		$M_2 = 0.2$		$M_3 = 0.3$	
1	7.80	8.13	14.22	14.92	25.96	28.84	0.98	1.09	1.63	1.42	2.79	5.53	15.00	16.53	32.24	41.81	48.25	83.92
2	8.63	7.53	17.01	16.52	27.13	31.76	1.02	1.05	2.22	1.73	3.14	3.44	16.00	16.30	42.92	29.17	57.08	41.51
3	8.12	7.28	16.65	15.84	25.98	26.05	1.05	0.94	2.17	2.15	2.90	2.92	23.00	23.29	40.63	48.37	34.52	57.33
4	8.18	8.07	18.29	15.92	25.34	20.76	0.68	0.72	1.46	1.50	2.19	2.26	25.00	15.98	37.74	41.75	59.79	47.14
5	7.04	7.58	13.11	16.53	27.66	22.89	1.14	1.23	2.64	2.27	3.44	3.98	18.00	14.40	23.80	44.36	41.62	43.45
6	8.32	9.79	16.80	14.74	26.55	26.82	1.00	0.96	2.49	1.97	3.36	2.95	26.00	10.28	40.45	30.69	23.84	67.64
7	8.02	8.30	14.46	15.42	25.74	23.10	1.22	1.20	2.29	2.39	3.18	3.29	28.00	19.68	40.57	50.66	61.05	72.99
8	6.36	8.24	18.23	17.48	20.24	28.28	0.73	0.86	1.43	2.13	2.11	2.18	12.00	26.70	31.01	32.74	82.76	66.55
9	5.93	8.65	16.51	13.43	22.36	19.92	1.12	0.91	1.92	1.77	2.52	2.98	17.00	19.78	49.92	28.39	56.18	52.64
10	8.56	8.88	17.57	19.17	25.73	23.20	0.80	0.75	1.45	1.62	2.36	2.40	21.00	28.16	39.08	47.59	71.62	83.71
11	7.61	9.85	17.34	16.31	27.06	28.60	0.92	1.23	2.55	2.54	3.95	3.47	20.00	16.24	43.19	28.68	60.13	70.66
12	7.88	8.07	16.89	12.55	22.98	24.26	1.08	1.05	2.28	2.22	3.32	3.23	18.00	11.38	46.14	22.51	66.97	65.73
13	8.73	6.82	18.22	15.64	25.64	20.26	0.95	0.99	2.00	2.00	2.94	2.93	26.00	22.32	64.67	40.40	94.98	58.26
14	7.97	9.72	16.72	11.98	23.27	23.10	1.17	1.14	1.95	2.35	3.91	3.58	16.00	23.16	24.82	44.13	51.38	63.52
15	9.16	8.77	16.72	15.86	24.97	30.30	0.85	0.79	1.42	1.75	2.33	2.34	14.00	12.88	40.57	33.27	33.99	60.82
16	9.32	8.71	14.86	15.67	21.87	28.43	1.05	1.10	2.01	2.26	3.29	3.02	22.00	15.90	51.58	43.90	75.55	86.55
17	8.32	6.91	16.03	14.10	22.70	18.87	0.80	0.85	2.07	1.99	2.71	2.46	23.00	20.34	42.91	32.95	36.92	64.79
18	8.71	6.37	14.87	18.74	31.61	22.69	1.14	0.98	1.92	1.58	3.57	2.97	19.00	12.43	37.70	38.89	69.16	55.98

4.3.1 Normalization and modeling phase

The present case study adopts the results obtained in case 1, i.e., we performed the numerical analysis considering the metric P_2 in table 1, the normalization function f_2 and the parameter control defined in 4.0.1 subsection.

Table 13 compared the RMSE obtained of MRF_{P2} with the best result presented by Chang at 2011[17] based on ANN. Note that the differences between the RMSE for the training sets are not very large, but in the test sets the differences are large. Our algorithm present less discrepancies between the RMSE of the training and test sets compared to the results of Chang[17] on this sets, which translates as a good performance and statistical robustness. Finally, our yields were obtained with only 1000 interactions in contrast to the 10000 required iterations in ANNs interactions, which gives us a clear advantage in terms of the use of computational resources.

Table 13: Compared the RMSE in the MRF and ANN

Approach	Iterations	RMSE	
		Train	Test
Chang (2011)[17]	10000	0.0479	0.0776
RFM_{P2}	1000	0.0347	0.0346

4.3.2 Optimization phase

200 chromosomes are generated randomly, whose structure is formed by 7 genes, 6 genes correspond to the vector of control factors $\mathbf{X} = \{A, B, C, D, E, F\}$ and 1 gen correspond to the signal factor $M = \{0.1, 0.2, 0.3\}$. These chromosomes will be on a normalized scale in order to calculate their corresponding responses with our predictive model MRF_{P2}. The objective of the present case study is the simultaneous optimization of the response vector $\mathbf{Y} = \{y_1, y_2, y_3\}$ under the quality characteristic DLTB, DNTB and DSTB respectively. The desirability functions and measure OPI are determinate by equation (11) and represent the *fitness* function of the algorithm.

$$\begin{aligned} OPI &= \left\{ \exp\left[-\exp\left(-\frac{1}{3}\sum_{j=1}^{3} \frac{\hat{y}_{1jk} - \hat{y}_{1j}^{\min}}{\hat{y}_{1j}^{\min}} \right) \right] \cdot \exp\left(-\frac{1}{3}\sum_{j=1}^{3} \left| \frac{2\hat{y}_{2jk} - \left(y_{2j}^{\max} + y_{2j}^{\min}\right)}{y_{2j}^{\max} - y_{2j}^{\min}} \right| \right) \right. \\ &\cdot \left. \exp\left[-\left(1 + \frac{1}{3}\sum_{j=1}^{3} \frac{\hat{y}_{3jk} + y_{3j}^{\max}}{y_{3j}^{\max}} \right) \right] \right\}^{\frac{1}{3}} \end{aligned} \tag{11}$$

Note that, the difference between equations (9) and (11) is that the present case study has no noise factor, then a summation is suppressed.

Table 14 shows the measure of importance from the algorithm MRF_{P2} adjusted to the present case study. B is observed as the most important gene and therefore with greater weight, so the prediction of new chromosomes is established as shown in equation (12).

Table 14: Measure of importance and weight of the genes by MRF_{P2}. Example 2

Gen	A	B	C	D	E	F
I_{xk}	0.006	**0.011**	0.008	0.009	0.006	0.009
Weight	0.013	**0.024**	0.017	0.020	0.013	0.011

$$\hat{y}_{h_i} = \left| \frac{\hat{y}_{m_i}\left(B_{p_i} - B_{h_i}\right) - \hat{y}_{p_i}\left(B_{m_i} - B_{h_i}\right)}{B_{p_i} - B_{m_i}} \right| \tag{12}$$

where $\hat{y}_{m_i}$ and $\hat{y}_{p_i}$ represent the prediction of the i-th mother and the i-th father from the RFM_{P2}. B_{p_i}, B_{m_i} and B_{h_i} correspond to the values for the gene B of i-th father, i-th mother and i-th children.

In consideration of the good results obtained in our pilot test, the mutation function $m(1)$ is used (table 3)with the 5% rate and the replacement in mutation of the gene is allowed.

Chang[17] present the optimum OPI found through his optimization proposal in contrast to the maximum OPI found exclusively with the PMs functions under the 18 combinations established in the matrix L_{18}. They highlight the advantages of their proposal regarding the maximization of the OPI value and the possibility of finding continuous values for the range of the control factors. Forest-Genetic adopts the virtues of our MRF, resulting in a broader heuristic search (within and outside the established ranges) with a special emphasis on the most important design at control factors, producing in this way an improvement in the maximum value OPI. All this is show in the table 15.

Table 15: The results of the implementation. Example 2

Method	Combination of parameter						OPI
	A	B	C	D	E	F	
Forest-Genetic	1.66	3.06	2.42	2.21	1.48	2.86	**0.8840978**
Chang (2011)[17]	1.85	2.99	2.97	1.00	1.86	1.00	0.753553
OPI through PMs	2	3	3	1	2	1	0.745636

The table above shows the OPI values found and their combinations in the vector $\mathbf{X}$; note that between others proposal and ours, the biggest differences are found in the OPI values and the levels at factors D and F (D and F are the factors with greater weight in MRF according to the table 14).

5 Conclusions

Forest-Genetic has been schematic under the premises of statistical robustness and consumption of computational resources. The use of normalization functions and RMSE as a statistic test allowed us to minimize the variability present in the established models, we could also save computational resources by minimizing the number of iterations necessary in the modelling and optimization phases. The main virtues of the methods are summarized as follows.

- Forest-Genetic doesn't suppose to have independence between factors, nor linearity of the factors with the answers, therefore it is perfectly adjustable to cases with non-linear relations and presence of correlations between factors
- Forest-Genetic was designed considering noise and signal factors, so it can be adjusted in cases of absences or presence of such factors
- Forest-Genetic perfectly adjusts to continuous case and discrete case parameters, extending its real application
- The modelling phase of Forest-Genetic is based on RF which minimizes the risk of model overloading
- Forest-Genetic allows the mutation scheme to be optimized
- The crossover scheme designed in the optimization phase allows Forest-Genetic the most efficient direction in the search hyperplanes.

References

[1] H. Allende, D. Bravo, and E. Canessa. Robust design in multi multivariate systems using genetic algorithms. *Qual Quant*, 44:315–332, 2010.

[2] J Antony. Simultaneous optimization of multiple quality characteristics in manufacturing processes using taguchi's quality loss function. *Int. Adv. Manuf. Technol*, 17(2):134–138, 2012.

[3] L. Breiman. Random forest. *Machine Learning*, 45(1):5–32, 2001.

[4] Hsu C-M, Su C*T, and D. Liao. Simultaneous optimisation of the broadband tap coupler optical performance based on neural networks and exponential desirability functions. *Int. Adv Manuf Technol*, 23:896–902, 2004.

[5] E. Canessa, G. Bielenberg, and E. Allende. Robust design in multi objetive systems using taguchi's prameter design approach and pareto genetic algorithm. *Rev. Fac. Ing. Univ. Antioquia*, (72):73–85, 2014.

[6] Su Chao-Ton and Chang Hsu-Hwa. Optimization of parameter design: An intelligent approach using neural network and simulated annealing. *International Journal of Systems Science*, 312(12):1543–1549, 2000.

[7] Su Chao-Ton, Chen Mu-Chen, and Chan Hsiao-Ling. Applying neural network and scatter search to optimize parameter design with dynamic characteristics. *Journal of the Operational Research Society*, 56(10):1132–1140, 2005.

[8] Glenn De'ath. Multivariate regression tree: a new technique for modelingspecies-environment relationships. *Ecology*, 83(4):1105–1117, 2002.

[9] E Del Castillo, D Montgomery, and D McCarville. Modified desirability functions for multiple response optimization. *Journal of Quality Technology*, 28(3):337–345, 1996.

[10] A. Dine, D. Larocque, and F Bellavance. Multivariate trees for mixed outcomes. *Computational statistics and data analysis*, 53(11):3795–3804, 2009.

[11] E. Harrington. The desirability function. *Industrial Quality Control*, 21(10):494–498, 1965.

[12] Shang Hsu, Feng-Liang Lee, and Muh-Cherng Wu. An integrated approach to achieving optimal design of computer games. *Expert Syst. Appl.*, 31(1):145–149, 2006.

[13] Shang Hsu, Feng-Liang Lee, and Muh-Cherng Wu. An integrated approach to achieving optimal design of computer games. *Expert Syst. Appl.*, 31:145–149, 2006.

[14] Chang Hsu-Hwa. Dynamic multi-response experiments by backpropagation networks and desirability functions. *Journal of the Chinese Institute of Industrial Engineers*, 23(4):280–288, 2006.

[15] Chang Hsu-Hwa. A data mining approach to dynamic multiple responses in taguchi experimental design. *Expert System with Applications*, 35(3):1095–1103, 2008.

[16] Chang Hsu-Hwa and Chen Yan-Kwang. Neuro-genetic approach to optimize parameter design of dynamic multiresponse experiments. *Applied Soft Compunting*, 11(1):436–422, 2001.

[17] Chang Hsu-Hwa and Chen Yan-Kwang. Neuro-genetic approach to optimize parameter design of dynamic multiresponse experiments. *Applied Soft Computing*, 11:436–442, 2011.

[18] K. Kim, K. Dennis, and Lin J. Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 49(3):311–325, 2000.

[19] Hsieh Kun-Lin, Tong Lee-Ing, Chiu Hung-Pin, and Hsin-Ya Yeh. Optimization of a multi-response problem in taguchi's dynamic system. *Computers Industrial Engineering*, 49(4):556 – 571, 2005.

[20] Huang M-L and Hung Y. H. Combining radial basis function neural network and genetic algorithm to improve hdd driver ic chip scale package assembly yield. *Expert Syst. Appl.*, 34:588–595, 2008.

[21] S Maghsoodloo, G Ozdemir, V Jordan, and Ch Huang. Strengths and limitations of taguchi's contributions to quality, manufacturing and process engineering. *Journal of manufacturing systems*, 23(2):73–126, 2004.

[22] Koksoy Onur. A nonlinear programing solution to robust multi-response quality problem. *Applied Mathematics and Computation*, 196(2):603–612, 2008.

[23] Surajit Pal and Susanta Kumar Gauri. Multi-response optimization using multiple regressionâbased weighted signal-to-noise ratio (mrwsn). *Quality Engineering*, 22(4):336–350, 2010.

[24] Yang-Gil Park and Bong-Jin Yum. Development of performance measures of dynamic parameter design problems. *IJMTM*, 5:91–104, 2003.

[25] Roberta Siciliano and Francesco Mola. Multivariate data analysis and modeling through classification and regression trees. *Computational Statistics Data Analysis*, 32(3):285 – 301, 2000.

[26] Chao-Ton Su and Lee-Ing Tong. Multi-response robust design by principal component analysis. 1997.

[27] Pal Sujarit and Kumar G. Susanta. A desirability functions-based approach for misultaneos optimization of quantitative and ordinal response variables in industrial process. *The International Journal of Engineering, Science and Technology*, 10(1):76–87, 2018.

[28] Joseph V. Roshan and Wu C-F Jeff. Robust parameter design of multiple-target systems. *Technometric*, 44(4):338–346, 2002.

[29] Adriana Villa-Murillo, Andrés Carrión, and Antonio Sozzi. Optimización del diseño de parámetros: Métodos Forest-Genetic univariante. *Publicaciones en Ciencias y Tecnología*, 10(1):12–24, 2016.

[30] Adriana Villa-Murillo, Susana. San Matías, and Andrés Carrión. *Optimización del diseño de parámetros: Métodos Forest-Genetic*. Universidad Politécnica de Valencia. España, 2012. Tesis doctoral.

[31] Ful-Chiang Wu and Chi-Hao Yeh. Robust design of multiple dynamic quality characteristics. *The International Journal of Advanced Manufacturing Technology*, 25(5):579–588, 2005.

[32] C. Zang, M. Friswell, and JE. Monttershead. A review of robust ptimal design and its applications in dynamic. *Computer and Structures*, 83(4):315–326, 2005.