doi.org/10.26398/IJAS.0034-004

SPLIT-PLOT DESIGNS AND MULTI-RESPONSE PROCESS OPTIMIZATION: A COMPARISON BETWEEN TWO APPROACHES

Rossella Berni[1], **Lorenzo Piattoli**
Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Italy

Christine M. Anderson-Cook
Los Alamos, New Mexico, USA

Lu Lu
Department of Mathematics and Statistics, University of South Florida, Florida, USA

Abstract: *Nowadays split-plot designs play a crucial role in technological fields, both for their flexibility when applying a robust design approach and in relation to the modelling step, by considering mixed Response Surface models and/or the class of Generalized Linear Mixed Models (GLMMs). In this paper, a split-plot design is studied in a process optimization scenario involving several response variables, a multi-response situation, where two optimization methods are compared. More precisely, by considering a real case study related to the improvement of a measurement process of a Numerical Control machine for measuring dental implants, the optimization is carried out with the Pareto front approach and then compared with an analytical optimization method obtained starting from the definition of a risk function. In the final discussion advantages and disadvantages of application for both methods are evaluated.*

Keywords: *split-plot designs, multi-response process optimization, Pareto front approach, analytical optimization methods*

1. INTRODUCTION

Process optimization is a key issue for statistical quality control, and its relevance has increased since the long and fruitful scientific debate related to Taguchi's two-step procedure for robust design (Nair, 1992). Currently, the robust design approach involves three key-steps: experimental design, modelling, and optimization. For a successful implementation, it is important to

[1] Corresponding author: Rossella Berni, email: rossella.berni@unifi.it

involve noise factors starting from the experimental planning, and then modelling them through a suitable analysis in the subsequent optimization phase. Therefore, the concept of process optimization is extended to robust process optimization, where control and noise variables are jointly studied to attain the best set of control factor levels that simultaneously achieves the target value and minimize the process variability with a robust configuration. In this context, the process optimization strictly depends on the designed experiment and on the class of statistical models applied. Specifically, Response Surface Methodology (RSM) approaches (Myers et al., 2016) may be used, or, alternatively, an experimental design may be planned outside the RSM context, modelling the experimental data with a more flexible class of statistical models, e.g. Generalized Linear Models-GLMs (McCullagh and Nelder, 1989; Nelder and Lee, 1991; Lee and Nelder, 2003) or Generalized Linear Mixed Models-GLMMs (Dror and Steinberg, 2006; Robinson et al., 2006; Berni and Bertocci, 2018).

Undoubtedly, the two choices of experimental design and modelling are related to what is known about the process to be studied and optimized; the same line of reasoning should be applied when deciding on which, how many, and the type of response variables have to be considered. In fact, the multi-response situation should be evaluated when the real scenario shows that several response variables are naturally involved, and they are important to the overall process under study. If this is not the case, then a simpler choice should be used, since collinearity, often present among responses (Box et al., 1973; Chiao and Hamada, 2001), can lead to serious complications both in analysis and optimization.

Moreover, in a multi-response situation, the optimization step is crucial, since it is generally not feasible in practical terms to reach an ideal optimum simultaneously for all the responses. In literature there are several methods for achieving optimal solutions as a compromise among several response variables, including Derringer and Suich (1980), Khuri and Conlon (1981), Del Castillo et al. (1996), Copeland and Nelson (1996).

A further issue concerns the conjunction of a multi-response case and the dual response approach. The simultaneous optimization of several variables, jointly with the consideration of two statistical models, e.g., location and dispersion, complicates the analysis and interpretation of results. Indeed, we should make a distinction when considering the differences between the dual response approach, or, alternatively, when a single response model, opportunely weighted with respect to the estimated dispersion values, is applied in a "true" multi-response case. In the latter case, the application of analytical methods for

optimizing can give fruitful and satisfactory results, particularly by considering recent developments involving noise factors.

To this end, several optimization measures are suggested in the literature (Lin and Tu, 1995; Tang and Xu, 2002). The Pareto front approach is a multi-response optimization method of the analytical-qualitative type, consisting of two sequential steps: a first step based on objective conditions identifies the dominant solutions, and a second step based on examining the identified alternatives and then selecting the best solution conditions that subjectively match the experimenters' priorities (Lu et al., 2011; Chapman et al., 2014a).

In this paper, data from a split-plot experiment (Berni, 2010) are optimized through the Pareto front approach (Chapman et al., 2014b), and following, the results obtained are compared with a proposal of the analytical optimization method. More specifically, both methods are compared and discussed through an empirical example in the orthodontic field, in order to improve the accuracy in the measurements of a Numerical Control (N/C) machine, which provides some automatic control of machining tools.

This paper is structured as follows: in Section 2, the basics of split-plot designs are reviewed and briefly illustrated. Section 3 provides a short description of both optimization methods, and Section 4 presents the case study including optimization results. The paper ends with a general comparison between the two approaches, and final remarks.

2. SPLIT-PLOT DESIGNS FOR STATISTICAL QUALITY CONTROL: A REVIEW

The split-plot design (Cochran and Cox, 1957) has been developed and characterized over the years, proving to be an experimental design widely used in industrial, technological, and environmental fields, because of the need to restrict randomization on expensive or hard-to-change factors in the experiment.

By considering the developments that the fractional factorial and RSM designs, and modelling have had since the 1980s, the split-plot design has experienced a particular renewal (Box and Jones, 1992), expounding its theoretical features, specific usefulness for the statistical quality control and robust design concepts, initially introduced by Genichi Taguchi, (Nair, 1992). In this context, the two seminal papers of Vining and Myers (1990) and Myers et al. (1992) extended the two-step procedure into the dual response approach, and the combined-array is considered as a milestone for recent developments and robust process optimization. Within this methodological framework, the split-plot design plays a central role, starting from the tutorial by Box and Jones

(1992), in which the authors proposed the split-plot design as an efficient alternative in many experiments, for example with hard-to-change factors, to Taguchi's product-array for a robust design approach, and also in a fractional factorial setting (Bisgaard, 2000).

Recently, split-plot designs have great relevance for the latest developments in robust process optimization, extending the initial concept of the robust design approach, with a focus on the design and modelling steps (Kowalski and Potcner, 2003; Kowalski et al., 2007; Jones and Nachtsheim, 2009). More recently, the split-plot design has been revised and included in the class of *crossed bi-randomized* experimental designs (Myers et al., 2016), given the possibility of including environmental/noise factors as Whole-Plot (WP) factors and process factors as Sub-Plot (SP) factors. The standard allocation of the environmental/noise factors as WP allows for the most accurate estimate of the factors of interest, as well as the estimate of the 1st order interactions, e.g., the 1st order interaction between a WP factor (for example a noise factor) with a SP (process) factor, in order to perform a robust design (Berni and Nikiforova, 2022). This structure is common in many applications, given the generally high cost of controlling the noise factors in production.

Nevertheless, a split-plot design in a RSM context generally requires that all the variables included in the experimental plan (irrespective of whether WP or SP factors) must be quantitative in nature. In fact, in case of a qualitative process variable, the optimization step is conditioned to the levels of the categorical variable involved. To this end, the inclusion of a qualitative variable should be limited (Berni, 2010) or restricted to two levels where they can be treated as quantitative in standard models. Moreover, the presence of measurable noise factors, involved as random effects, is possible when the split-plot design is applied through mixed Response Surface (RS) models, or alternatively, through GLMMs.

Additional developments in the literature have contributed significantly to the inclusion of the split-plot design in RSM, showing the equivalence of *Ordinary Least Squares* (OLS) with *Generalized Least Squares* (GLS) for split-plot designs and mixed RS models (Vining et al., 2005); and improving inference issues (Vining and Kowalski, 2008).

2.1. THEORY ABOUT THE SPLIT-PLOT DESIGN

When implementing a split-plot design, it is essential to begin with a primary classification between WP factors and SP factors. It is therefore desirable to carefully evaluate what is possible in the experimental set-up for the specific process (industrial process, laboratory experiment) under study,

considering the exact definition of the response (quantitative) variables. Also, it is necessary to define the role of each variable in the study, in order to plan the split-plot design based on the most efficient arrangement for the specific scenario. This step plays a central role, not only in the attribution of factors as whole-units or sub-units, but also determines the subsequent model estimation, in which each variable plays a specific role, according to its nature (qualitative, discrete quantitative or continuous). It is also a crucial point to clarify the distinction between fixed and random effects.

In short, a Whole-Unit (WU) is defined by runs involving the set of WP factors, and the run order of all the WUs is randomized. Subsequently, the Sub-Units (SUs), defined through the combination of levels (runs) of the SP factors within the WUs, are associated with a particular WU and randomized separately.

In this study, we consider a split-plot design in a RSM context. More precisely, we are interested in applying a standard second-order polynomial model with random effects. The model with random effects for the single experimental observation y_u $(u = 1, \dots, n)$ and J variables $(x_1, \dots, x_j, \dots, x_J)$ is (Khuri, 1996):

$$y_u = \beta_0 + \boldsymbol{f}'(\boldsymbol{x}_u)\boldsymbol{\beta} + \boldsymbol{z}_u'\boldsymbol{\gamma} + \boldsymbol{g}'(\boldsymbol{x}_u)\boldsymbol{\Delta}\boldsymbol{z}_u + \varepsilon_u \qquad (1)$$

where β_0 is the intercept; $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ is the column vector $[p\times 1]$ of the unknown fixed parameters ($p \geq J$); $\boldsymbol{x}_u = (x_{u1}, \dots, x_{uj}, \dots, x_{uJ})'$ is the vector of design settings at the u-th experimental run; $\boldsymbol{f}(\boldsymbol{x}_u)$ is a vector function of dimension $[p\times 1]$ defined for each $\boldsymbol{x}_u$ and related to the p second-order effects for the J variables. Therefore, $\boldsymbol{F}$ is the so-called "extended" matrix of dimension $[n\times p]$, formed by the n rows $\boldsymbol{f}'(\boldsymbol{x}_u)$; ε_u is the residual error. For the random effects, $\boldsymbol{z}_u = (z_{u1}, \dots, z_{ub})'$ is the vector of binary values $(0,1)$ to describe the presence and structure of the block factors; $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_b)'$ is the column vector $[b\times 1]$ of the unknown coefficients relating to the random effects. The matrix $\boldsymbol{\Delta}$ characterizes the 1st order interactions between polynomial effects (fixed) and random effects. The maximum dimension of $\boldsymbol{\Delta}$ is achieved if the interactions of all fixed effects with random effects are included in model (1). Note that this matrix contains the key estimated coefficients for evaluating the robust design as the control-by-noise interactions that can be exploited to achieve robustness.

Starting from the model expression (1), the second-order polynomial model of response surfaces from a split-plot design is outlined, considering quantitative variables only. For further details, see Myers et al. (2016).

Let's consider two sets of factors: $\boldsymbol{z} = (\boldsymbol{z_1}, \ldots \boldsymbol{z_i}, \ldots, \boldsymbol{z_I})$ are WP random factors/variables and $\boldsymbol{x} = (\boldsymbol{x_1}, \ldots \boldsymbol{x_j}, \ldots, \boldsymbol{x_J})$ are SP variables/factors.[2] Furthermore, let y_u be the u-th observation of the k-th block, for the i-th WP factor and the j-th SP factor respectively $(i = 1, \ldots, I; j = 1, \ldots, J; k = 1, \ldots, K)$; therefore, the second-order mixed RS split-plot model for a single replicate $(K = 1)$ and a single observation u is defined as follows:

$$y_u(\boldsymbol{z}, \boldsymbol{x}) = \beta_0 + \boldsymbol{z_i'\gamma} + \boldsymbol{z_i'\Gamma z_i} + \boldsymbol{x_{ij}'\beta} + \boldsymbol{x_{ij}'Bx_{ij}} + \boldsymbol{z_i'\Delta x_{ij}} + \psi_u + \varepsilon_u \qquad (2)$$

where β_0 is the intercept; $\boldsymbol{z_i} = (z_{i1}, \ldots, z_{iu}, \ldots, z_{in})'$ is the vector of the i-th WP factor; $\boldsymbol{\gamma}$ contains the unknown coefficients of linear terms of the WP variables; $\boldsymbol{\Gamma}$ is the array related to the coefficients of the 1st order interaction and quadratic terms of the WP variables; $\boldsymbol{x_j} = (x_{j1}, \ldots, x_{ju}, \ldots, x_{jn})'$ is the vector of the j-th SP factor; $\boldsymbol{\beta}$ contains the unknown coefficients of linear terms of the SP variables; $\boldsymbol{B}$ is the array containing 2nd order coefficients for the fixed effects of SP variables; the matrix $\boldsymbol{\Delta}$, contains coefficients of the 1st order interaction effects between the WP and SP factors. The model terms in the matrix $\boldsymbol{\Delta}$ are of primary interest in the context of robust design evaluation. Regarding the two error components, ψ_u is the WP error component and ε_u is the SP error component, where in general the two error components are assumed to be independent and identically Normally distributed, i.e. $\psi \sim iid\ N(0, \sigma_\psi^2)$ and $\varepsilon \sim iid\ N(0, \sigma_\varepsilon^2)$, respectively. In addition we also assume that $Cov(\psi_u, \varepsilon_u) = 0\ \forall u$. The assumptions about the error variances and covariances are equivalent to assume constant covariance between two observations belonging to the same WU, across all its observations.

In the case study illustrated in Section 4, the multi-response case is related to the optimization involving three split-plot models, one for each response, estimated applying the RSM model above.

3. OPTIMIZATION METHODS

This Section includes a short description of both optimization methods considered. The Pareto front approach (Lu et al., 2011; Chapman et al., 2014a; Chapman et al., 2014b) is also illustrated within the case study, considering the application (Subsections 4.2 and 4.3); a brief introduction of the analytical

[2] Please note that we are referring to a factor/variable considering that the experimental region χ is defined by the factor ranges; a finite number of experimental points, forming the experimental design, is then selected by the experimental region. Following, the model estimation is performed within the whole experimental region, by inferring from a discrete set of points, e.g., the experimental points, to a continuous one.

method is illustrated in Subsection 3.2. For further details see (Berni and Gonnelli, 2006; Berni, 2010; Berni and Burbui, 2014).

3.1. THE PARETO FRONT APPROACH

The Pareto front approach is a multi-response analytical-qualitative optimization method, which allows the search for optimum to take subjective priorities and constraints into account, such as those due to a company's requirements (for example, costs or technical/engineering specifications). It consists of two sequential steps (Chapman et al., 2014a; Myers et al., 2016; Anderson-Cook, 2017), as outlined below.

Suppose that χ defines the entire experimental region; within this region a finite set, possibly a grid, of points, is selected and used to estimate the responses of interest and used to define a Pareto-optimal set. A possible solution is called non-inferior (or Pareto-optimal), if and only if, there is no other combination within the set for which the values of all the responses are at least as good, and the value of at least one response is strictly better; otherwise, the solution is called inferior or dominated. The set of non-inferior (or Pareto-optimal) input combinations is called the Pareto-optimal set, and the corresponding set of vectors for the responses under consideration is known as the Pareto front or frontier. Since the inferior solutions are not rational choices conditional on the choice of responses under consideration, they are not considered further and definitively discarded (Zitzler, 1999; Marler and Arora, 2004; Coello Coello et al., 2007). This leads to a reduced number of alternative solutions to be considered further in later stages of the optimization.

The Pareto front approach can be summarized with the following two steps:

1. An objective step, where the Pareto-optimal set is identified from the initial set of choices, based on the corresponding estimated response values;
2. A subjective step, in which the points belonging to the Pareto-optimal set are examined and then compared. Only points that provide the best combination of responses are considered as a compromise among all the estimated response values (quantitative considerations). This choice is based on evaluation and incorporation of the priorities/preferences of the company.

It must be noted that several optimal points corresponding to input combinations could be considered and compared for selection, by considering the priorities of different teams (decision-makers) involved in the study. Therefore, the best optimal solution combines the quantitative results with the

decision-makers' priorities. Moreover, graphical methods are a useful tool for discussion, comparison and achieving a consensus among all stakeholders (Anderson-Cook and Lu, 2018).

3.2. THE ANALYTICAL METHODS FOR A ROBUST PROCESS OPTIMIZATION

When dealing with several response variables, it is generally not feasible in practical terms to simultaneously achieve the optimum for each of them with a single input combination. To this end, many authors, starting from the methods suggested by Derringer and Suich (1980) and Khuri and Conlon (1981), have proposed methods to synthetize and optimize the responses, such as Ames et al. (1997), Del Castillo et al. (1996), Rajagopal et al. (2005).

In addition, a further issue emerges when considering the multi-response case and the dual response approach. Here, the simultaneous optimization of several variables jointly with the consideration of two statistical models, e.g., location and dispersion models, increases the complexity and dimensionality of the problem.

In order to solve the latter issue, which could imply a notable computational burden, analytical optimization methods can be defined and simplified starting from the dual approach theory and the building of a suitable performance measure (Leon et al., 1987). To this end, we consider a multiplicative relationship between the expected value $\big(E(y) = \mu(\boldsymbol{x})\big)$ and the process variance $\big(Var(y) = \sigma^2(\boldsymbol{z}, \boldsymbol{x})\big)$ defined as the variance of the response variable. Moreover, the expected value of the response could be identified in relative to the target value (e.g., $E(y) = \tau$), according to the Nominal the Best (NTB), Smaller the Better (STB), or Larger the Better (LTB) situations. At the beginning a general risk function is expressed as follows:

$$R(\boldsymbol{z}, \boldsymbol{x}) = (\mu(\boldsymbol{x}) - \tau)^2 + f\big(\mu(\boldsymbol{x})\big)\sigma^2(\boldsymbol{z}, \boldsymbol{x}) \tag{3}$$

Formula (3) explicitly involves two terms: i) $(\mu(\boldsymbol{x}) - \tau)^2$ which expresses the adjustment to the target value, while ii) $f\big(\mu(\boldsymbol{x})\big)\sigma^2(\boldsymbol{z}, \boldsymbol{x})$ is related to the multiplicative relation between location (adjustment) and process variance (dispersion). Therefore, formula (3) allows for defining specific objective functions in a dual response approach perspective, (Berni and Gonnelli, 2006), and particularly, to optimize several response variables without separately estimating the two statistical models for each response. This approach provides a simplification as well as a weighting of the responses according to their relative importance.

More recently, split-plot designs and modelling have been optimized by explicitly involving one model only for each response in a robust process optimization context, in which random effects are also evaluated (Berni and Bertocci, 2018; Berni and Nikiforova, 2022).

Let's start by defining a general response surface model, y_t $(t = 1, \dots, T)$, for each of the T responses. The simultaneous optimization may be performed considering the T estimated surfaces; each estimated model is evaluated as a single function to be included in the objective function to be optimized. Starting by formula (3) and considering the concept of a dual response approach, a general objective function can be defined as the distance between the estimated surface $\hat{y}_t$ and the corresponding desired target value τ_t:

$$S_t(\boldsymbol{z}, \boldsymbol{x}) = (\hat{y}_t(\boldsymbol{z}, \boldsymbol{x}) - \tau_t)^2 \quad \forall t$$

The approach can be easily adapted for responses where the goal is to achieve a maximum or minimum value. Subsequently, the minimization on the coded experimental region χ is performed by minimizing the sum of all the distances, as follows:

$$\min_{\chi} \left\{ \sum_t S_t(\boldsymbol{z}, \boldsymbol{x}) \right\} \tag{4}$$

The objective function (formula (4)) is optimized conditional on the whole experimental region χ defined by the process variable ranges (and potentially any limiting constraints for other technological issues), as well as involving the estimated confidence interval for each random coefficient when random noise factors are present.

In the following section, we compare the two optimization methods, the Pareto front approach (Subsection 3.1), and the objective function of formula (4), where the goal is to improve the accuracy of the measurement process of a N/C machine, used in the orthodontic field; the measurements of which are analyzed for a generic dental implant.

4. THE CASE STUDY: DATA DESCRIPTION AND PROCESS OPTIMIZATION

In this Section, comparison of multi-response optimization methods is made, after a short description of the experimental planning and data. For further details see Berni (2010).

4.1. SPLIT-PLOT DESIGN AND DATA DESCRIPTION

The aim of the study is to improve the accuracy in measurements for a N/C machine, jointly with a reduction of the measuring time. The machine uses a feeler pin with a movable bridge framework to facilitate the positioning of the measured piece (dental implant). The machine needs specific environmental conditions to function properly, all of which were ensured previously (see Berni and Gonnelli, 2006).

In Berni (2010), five response variables, $T = 5$, were optimized simultaneously applying formula (4) and related to the different positionings of the feeler pin on the dental implant during the process measurement steps. In this paper we focus on the optimization comparison by looking at a subset of three response variables.

By considering the dental implant used to set the measurement process, the three responses are (with their respective targets in brackets): maximum circle diameter-$crmax$ $(\tau_1: 3.000)$, minimum circle diameter-$crmin$ $(\tau_2: 2.790)$, and eccentricity-$eccen$ $(\tau_5: 0.000)$. There is no problem with correlation among the three dependent variables, since each type of measurement is carried out as a distinct step; moreover, each response variable is independent from the others during the measurement of the piece. In order to reduce the measuring time, it is possible to intervene on the process phase related to the identification of the cone frustum, identified by three circles, at three different distances. In Figure 1, the frustum of cone is shown by highlighting the three circles used to locate it.

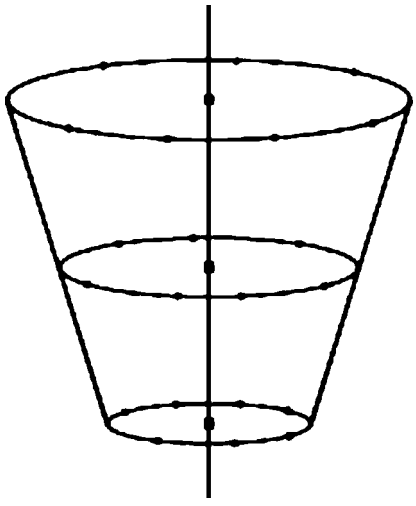

Fig. 1: Location of the frustum of cone by three circles and definition of the "circle-point" factor

In the initial setting, the N/C machine measures 7 points on each circle (7,7,7), each point is denoted with a dot in Figure 1. A categorical input factor "circle-point-cp" is then defined at four levels with each level corresponding to a different number of points measured on each circle: (1) 7,7,7; (2) 7,5,7; (3) 5,7,5; (4) 5,5,5.

Two other variables are involved in the split-plot design: measurement speed-$mspeed$ (mm/sec), and drift speed-$dspeed$ (mm/sec). Therefore, a split-plot design with three factors is planned: two WP process factors, both at two levels (measurement and drift speeds), and only one SP control factor, the cp categorical factor at four levels. The final split-plot has 112 runs with seven replicates.

Standardization of the responses was carried out (Berni, 2010) to compensate for differences in magnitude among responses, even though both responses and WP factors are expressed with the same unit of measurement.

4.2. THE PARETO FRONT APPROACH: OBJECTIVE PHASE

In order to identify the Pareto-optimal set, a series of 1764 combinations of the factor $mspeed$, $dspeed$ and cp levels were identified, from which, the predicted response values $\widehat{crmax}$, $\widehat{crmin}$ and $\widehat{eccen}$ were estimated using the model form described in formula (2). The set of possible input combinations (Figure 2) was formed by constructing a grid of points based on discrete levels of $mspeed$ and $dspeed$ for each level of the factor cp. The fineness of the mesh of each grid is 0.1, since this choice balances the complexity of calculation and valid coverage of the two-dimensional region, (ranges of $mspeed$ and $dspeed$ factors). The combinations of the possible solutions are labelled from 1 to 1764 according to the approach described in Chapman et al. (2014a): i) from the first grid on the upper left to the lower right grid; ii) inside each grid starting from the bottom row and moving from left to right, then starting at the end of each row, from the leftmost point of the next row.

The obtained Pareto-optimal set consists of 61 combinations, highlighted by the solid circles in Figure 2. These all combinations involve $cp = 4$, which requires the smallest number of measured points. Therefore, irrespective of the choice in the subjective phase, the Pareto front results ensure that an improvement in the measurement time is always obtained.

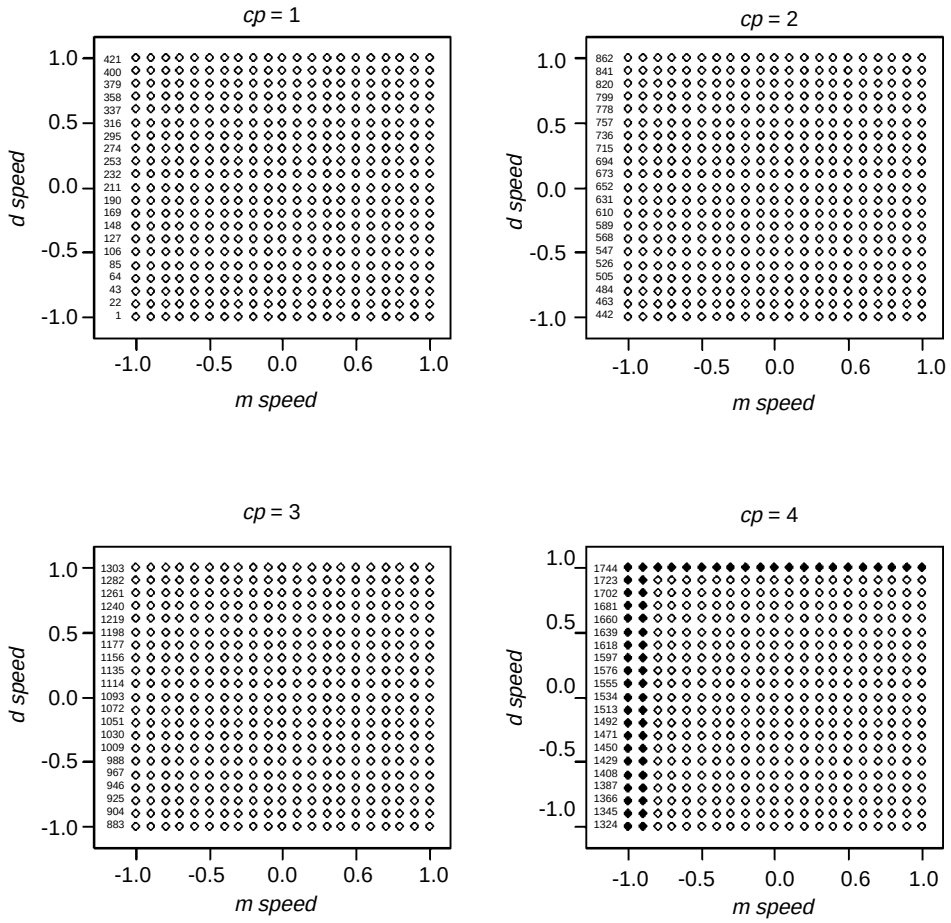

Fig. 2: Input grid plot

Figure 3 shows the pairwise scatterplots of the points belonging to the Pareto front (Chapman et al., 2014a; Anderson-Cook, 2017). The analysis of this set of points shows that there is a strong trade-off between the maximum circle diameter and eccentricity. The trade-offs between the other two pairs of response variables appear less important. Finally, by observing the relatively small ranges of the predicted responses, we note that all 61 combinations of the Pareto-optimal set lead to values of $\widehat{crmax}$, $\widehat{crmin}$ and $\widehat{eccen}$, close to the respective desired targets.

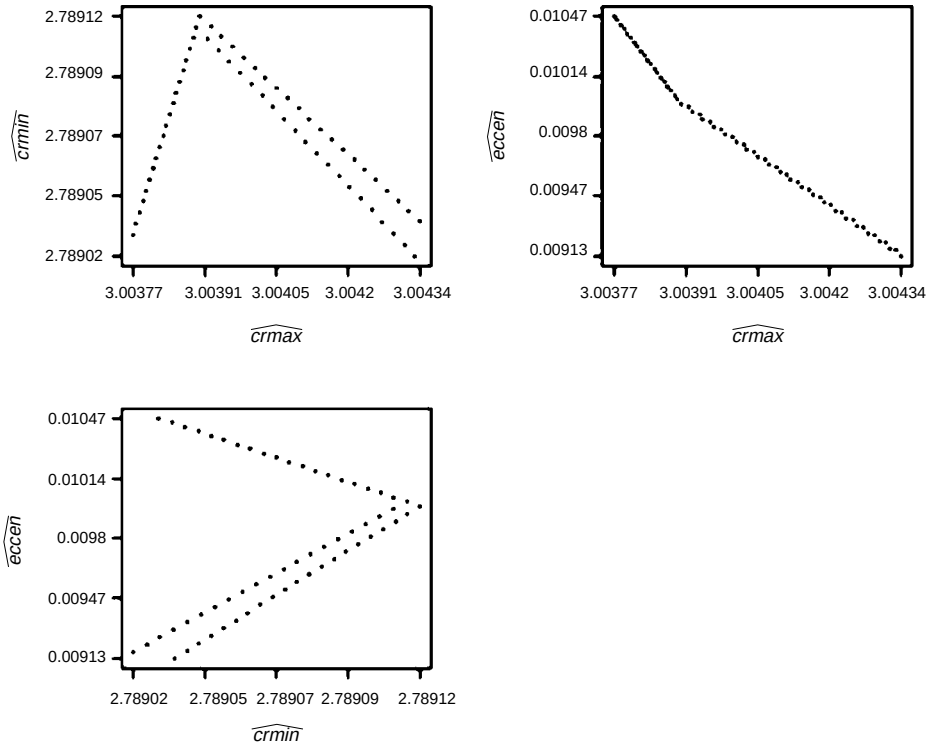

Fig. 3: Pairwise scatterplots of the points belonging to the Pareto front

4.3. THE PARETO FRONT APPROACH: SUBJECTIVE PHASE

In order to compare the 61 combinations of the Pareto-optimal set, the following procedure is carried out (Chapman et al., 2014a; Myers et al., 2016; Anderson-Cook, 2017): i) the Pareto front values of each predicted response are transformed into desirability values, so that the best value obtained (from the set of solutions comprising the front) for each response is scaled to one, while the worst value is scaled to zero; ii) for each combination of the Pareto-optimal set, the respective desirability values are combined in a single global desirability function. Since it was considered appropriate to heavily penalize undesirable predicted response values, we choose the standard multiplicative desirability form, based on the geometric mean expression, as follows:

$$D(\boldsymbol{x}_{\mathcal{P}}, \boldsymbol{w}) = d_{crmax}(\boldsymbol{x}_{\mathcal{P}})^{w_{crmax}} \, x \, d_{crmin}(\boldsymbol{x}_{\mathcal{P}})^{w_{crmin}} \, x \, d_{eccen}(\boldsymbol{x}_{\mathcal{P}})^{w_{eccen}}$$

where $\boldsymbol{x}_{\mathcal{P}}$ is a combination of the Pareto-optimal set; $d_{crmax}(\boldsymbol{x}_{\mathcal{P}})$, $d_{crmin}(\boldsymbol{x}_{\mathcal{P}})$,

$d_{eccen}(\boldsymbol{x}_{\mathcal{P}})$ the single desirability values related to the three predicted responses; $\boldsymbol{w} = (w_{crmax}, w_{crmin}, w_{eccen})'$ a weight vector, with $w_{crmax}, w_{crmin}, w_{eccen} \geq 0$ representing the weights assigned to the three response variables and $w_{crmax} + w_{crmin} + w_{eccen} = 1$. We note here that the small deviations of the response values from their desired targets mean that even small misses from the target are being strongly penalized, because one minimal error in measurements can lead to a serious risk for a patient.

Figure 4 shows the mixture plot, which identifies the best combination (i.e., the optimum point for achieving the highest value of the global desirability function) for each possible weighting of the response variables. Each point of the mixture plot represents a weight vector (e.g., the left bottom vertex represents $\boldsymbol{w} = (1,0,0)'$, the centroid marked with a black cross represents $\boldsymbol{w} = (1/3, 1/3, 1/3)'$, and the bottom edge represents the weight vectors with $w_{crmax}, w_{crmin} > 0$ and $w_{eccen} = 0$). For further details see Cornell (2002).

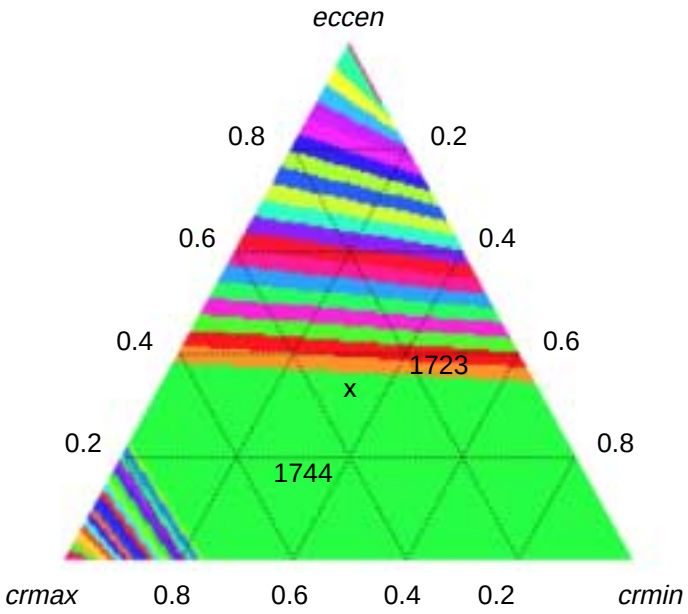

Fig. 4: Mixture plot

In the case study, 41 of the 61 combinations belonging to the Pareto-optimal set appear in the mixture plot (each colored area identifies a different combination), that is, they are best for at least one weight vector. Assuming that the three response variables are thought to be of equal importance, the weights reflecting company priorities/preferences are those around the centroid of the triangle. The two best points for these weight combinations are 1723 and 1744. In particular, 1744 is better for most of these weights, including the one directly at the centroid of the triangle as well. Table 1 shows the detailed results obtained for these two points, (1723 and 1744), differing only in the *dspeed* level value, and providing similar predicted responses.

Tab. 1: Factor levels and predicted response values for the combinations 1723 and 1744

Combination	Factors			Predicted responses		
	mspeed	*dspeed*	*cp*	$\widehat{crmax}$	$\widehat{crmin}$	$\widehat{eccen}$
1723	-1	0.9	4	3.00392	2.78911	0.00994
1744	-1	1	4	3.00390	2.78912	0.00998

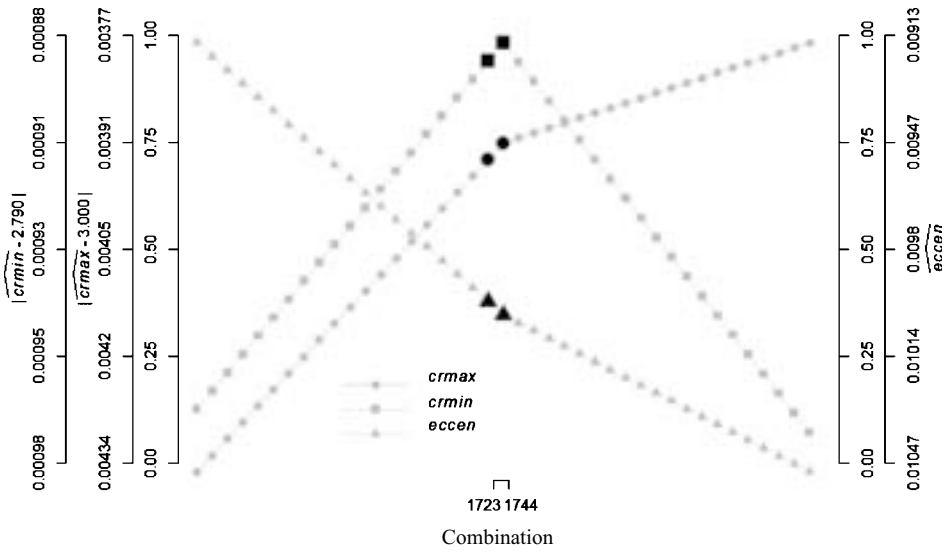

Fig. 5: Trade-off plot of the 41 best point combinations for at least one weight combination

Figure 5 contains the trade-off plot illustrating the desirability values (internal vertical axes) and absolute value differences between the predicted response values and the respective targets (external vertical axes) considering the 41 solutions that are best for at least one weight combination. In Figure 5, the trade-offs between the pairs of responses are similar to those highlighted by the pairwise scatterplots in Figure 3. Moreover, as shown in the mixture plot (Figure 4); the point combinations 1723 and 1744 provide an ideal balance among the three responses when they are all prioritized as being equally important.

In order to better analyze and compare these two combinations of interest, Figure 6 shows the synthesized efficiency plots (Lu and Anderson-Cook, 2012) which allow comparison of the relative efficiency of individual solutions with the best available across all the possible weight vectors.[3] The synthesized efficiency of a point combination (belonging to the Pareto-optimal set) $\boldsymbol{x}_{\mathcal{P}}$, with weight vector $\boldsymbol{w}$, is defined as follows:

$$\frac{D(\boldsymbol{x}_{\mathcal{P}}, \boldsymbol{w})}{max_{\boldsymbol{x}_{\mathcal{P}}}[D(\boldsymbol{x}_{\mathcal{P}}, \boldsymbol{w})]}$$

The shading, from white to black, represents the transition from high to low values of the synthesized efficiency. Each of the 19 shades of grey, starting from the lightest, corresponds to a decrease in the synthesized efficiency of 0.05.

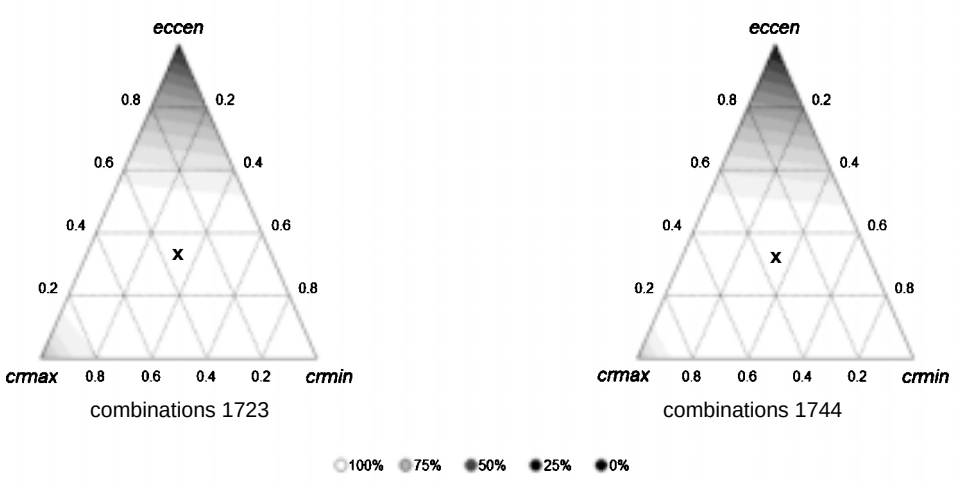

Fig. 6: Synthesized efficiency plot for the point combinations 1723 and 1744

[3] It should be noted that for the construction of the synthesized efficiency plots (Figure 6) and the mixture plot (Figure 4), a set of 20301 weight combinations has been defined, where adjacent weights related to a same response variable are separated by a distance equal to 0.005.

The large white region characterizing the two graphs, represents approximately 75% and 74% of the triangles, respectively, and indicates that both points have a synthesized efficiency of at least 0.95 for a substantial number of weight combinations. In particular, the white region around the centroid of the triangle, indicated in Figure 6 with a black cross, shows that both point combinations give excellent performance at the weighting region, which reflects the company's priorities/preferences. The graphical plots provide detailed information about the relative performance of the different contending solutions and allow the experimenter to understand what alternatives are available.

Moreover, we select the optimal solution as represented by combination 1744, since it is slightly better for a large number of weight combinations, and in particular, for the weighting giving equal importance to the three response variables. However, input combination 1723 provides a similar performance, thus representing a valid competitive alternative.

4.4. RESULT COMPARISON BETWEEN THE TWO APPROACHES

Table 2 shows the results from both multi-response optimization methods: the Pareto front approach and the analytical method applied in Berni (2010). By comparing (Table 2) the optimum point combination 1744 (Piattoli, 2020) and the optimal analytical solution, we can observe how only one process factor, *cp*, shows the same optimal level; nevertheless, it must be noted that the circle-point variable is the main process variable that we are interested in optimizing. Although both combinations provide similar predicted response values, the analytical method allows for obtaining a better value for the response $\widehat{eccen}$. This is an important result in view of the relevance that this response variable has in the actual process. The constructed Pareto front contains similar solutions to those identified by the optimal analytical solution, but corresponding to different weight combinations. Hence, with a more thorough exploration of the solution set identified with the Pareto front, a similar solution could be selected relaxing the assumption that all the responses were of equal importance.

Tab. 2: Optimization results: the comparison

Method	Factors			Predicted responses		
	mspeed	***dspeed***	***cp***	$\widehat{crmax}$	$\widehat{crmin}$	$\widehat{eccen}$
Pareto front	-1	1	4	3.00390	2.78912	0.00998
Analytical	0.710	0.362	4	3.00300	2.78500	0.00100

It is important to note, however, that although a Pareto front can be constructed for any number of responses of interest, the graphical tools considered here only used three response variables, unlike the five considered in Berni (2010). For this reason, it was only possible to make a partial comparison between the results obtained through the two different methods. Nevertheless, the Pareto front approach offers the possibility of using additional graphical tools (Lu et al., 2017), which enable multi-response optimization of more than three response variables.

Moreover, through the analytical approach, the optimization was carried out considering both non-standardized and standardized data, where the latter gave the best optimization results.

5. GENERAL COMPARISON AND FINAL REMARKS

The results obtained through the case study allows us to perform an empirical comparison between the two approaches, where some specific differentiations could be viewed in a theoretical perspective, as outlined in the following scheme (Table 3). Both methods use the same experimental plan, data and analysis, but then differ in how choose to optimize the settings of inputs.

Tab. 3: Theoretical step comparison between the Pareto front approach and the analytical method

Step	***Method***	
	Pareto front	**Analytical**
1	DoE: planning and trials	DoE: planning and trials
2	Statistical modelling	Statistical modelling
3	Optimization: A) objective phase - identification of Pareto-optimal set; B) subjective phase - choice of the optimal solution among the points belonging to the Pareto-optimal set, taking the quantitative results and the decision-makers' priorities into account; C) validation of the results obtained at step (B), intrinsic in the subjective phase.	Optimization: a) definition of objective function (formula (4)); b) minimization (or maximization) of the objective function, and identification of the optimal solution (optimal process variable levels); c) validation of the obtained results at step (b) by: 1. the objective function value, (results also checked through: convergency, gradient estimates, determinant of the Hessian matrix); 2. application of the optimal solution (obtained through step (b)) in the real (actual) production process, by involving the stakeholders (engineers).

Undoubtedly, the Pareto front approach offers the advantage of using graphical tools in a simple and intuitive way, enabling straightforward identification of leading solutions with discussion allowing for consensus of the optimal solution among the various company teams involved. The elimination of non-competitive choices streamlines where to focus further discussion. Moreover, a subjective evaluation (Table 3, step B) can also be performed, with the possible achievement of a unanimous decision among different stakeholders, and considering different weightings of how important the performance on each response is to overall results. It is possible to compare different identified solutions, and see their relative strengths and weaknesses for each of the responses of interest. Indeed, it allows for accurate comparison among several input combinations of interest. A further advantage is the flexibility in response weighting to handle multiple combinations of business priorities and to examine the impact of these choices on the identified results. The transparent nature of the Pareto front presents the experimenter with different alternatives that can be explored and compared. Nevertheless, this is also possible by performing analytical optimization methods (see Lin and Tu, 1995). In addition, response weighting and analytical methods assign relative importance to each response according to the estimated corresponding weight (Berni, 2010). This is particularly helpful for solving technological issues and constraints, that can be measured and evaluated in a wide and general context.

An advantage of the analytical method is the ability to include random effects, within both the modelling and the optimization steps. Therefore, the fixed as well as random effects are wholly involved, and as a result a robust process optimization can be carried out, and the final validation (Table 3, step c.2) specifically verified in the firm, allows for checking the validity in the actual context. It is straightforward to use the same optimization function based on the inclusion of random effects for each of the responses as the basis for constructing the Pareto front.

The aforementioned advantages and disadvantages highlight the significant relevance of both methods, as each has specific strengths and weaknesses that would be relevant for a wide range of empirical situations (real industrial processes, technological contexts) where they can be effectively applied.

REFERENCES

Ames, A.E., Mattucci, N., Macdonald, S., Szonyi, G. and Hawkins, D. M. (1997). "Quality loss functions for optimization across multiple response surfaces". *Journal of Quality Technology*, 29 (3): 339-346.

Anderson-Cook, C.M. (2017). "Optimizing in a complex world: a statistician's role in decision making". *Quality Engineering*, 29 (1): 27-41.

Anderson-Cook, C.M. and Lu,L. (2018). "Graphics to facilitate informative discussion and team decision making". *Applied Stochastic Models in Business and Industry*, 34 (6): 963-980.

Berni, R. (2010). "Split-plot and robust designs: weighting and optimization in the multiple response case". In: Giovagnoli, A., Atkinson, A. C., Torsney, B., and May, C., eds. *mODa 9 – advances in model-oriented design and analysis: proceedings of the 9th international workshop in model-oriented design and analysis held in Bertinoro, Italy, June 14-18, 2010*. Heidelberg: Physica-Verlag, 25-32.

Berni, R. and Bertocci, F. (2018). "Optimization of gas sensors based on advanced nanomaterials through split-plot designs and GLMMs". *Sensors*, 18 (11): 3858.

Berni, R., Bertocci, F., Nikiforova, N.D. and Vining, G. G. (2020). "A tutorial on randomizing versus not randomizing split-plot experiments". *Quality Engineering*, 32 (1): 25-45.

Berni, R. and Burbui, M. (2014). "Process optimization of a superfinishing machine through experimental design and mixed response surface models". *Quality Engineering*, 26 (4): 404-415.

Berni, R. and Gonnelli, C. (2006). "Planning and optimization of a numerical control machine in a multiple response case". *Quality and Reliability Engineering International*, 22 (5): 517-526.

Berni, R. and Nikiforova, N. D. (2022). "A joint multiresponse split-plot design modeling and optimization including fixed and random effects". *Austrian Journal of Statistics*, 51: 83-101. https://doi.org/10.17713/ajs.v51i1.1211

Bisgaard, S. (2000). "The design and analysis of $2^{k-p} \times 2^{q-r}$ split plot experiments". *Journal of Quality Technology*, 32 (1): 39-56.

Box, G.E.P., Hunter, W. G., MacGregor, J. F. and Erjavec, J. (1973). "Some problems associated with the analysis of multiresponse data". *Technometrics*, 15 (1): 33-51.

Box, G.E.P. and Jones, S. (1992). "Split-Plot designs for robust product experimentation". *Journal of Applied Statistics*, 19 (1): 3-26.

Chapman, J.L., Lu,L., and Anderson-Cook, C.M. (2014a). "Process optimization for multiple responses utilizing the Pareto front approach". *Quality Engineering*, 26 (3): 253-268.

Chapman, J. L., Lu, L. and Anderson-Cook, C. M. (2014b). "Incorporating response variability and estimation uncertainty into Pareto front optimization". *Computers & Industrial Engineering*, 76: 253-267.

Chiao, C.-H. and Hamada, M. (2001). "Analyzing experiments with correlated multiple responses". *Journal of Quality Technology*, 33 (4): 451-465.

Cochran, W.G. and Cox, G.M. (1957). *Experimental designs*. 2° ed. New York: Wiley.

Coello Coello, C.A., Lamont, G.B. and Van Veldhuizen, D.A. (2007). *Evolutionary algorithms for solving multi-objective problems*. 2º ed. New York: Springer, 1-13.

Copeland, K.A.F. and Nelson, P.R. (1996). "Dual response optimization via direct function minimization". *Journal of Quality Technology*, 28 (3): 331-336.

Cornell, J.A. (2002). *Experiments with mixtures: designs, models, and the analysis of mixture data*. 3° ed. New York: Wiley.

Del Castillo, E., Montgomery, D.C. and McCarville, D.R. (1996). "Modified desirability functions for multiple response optimization". *Journal of Quality Technology*, 28 (3): 337-345.

Derringer, G. and Suich, R. (1980). "Simultaneous optimization of several response variables". *Journal of Quality Technology*, 12 (4): 214-219.

Dror, H. A. and Steinberg, D.M. (2006). "Robust experimental design for multivariate generalized linear models". *Technometrics*, 48 (4): 520-529.

Jones, B. and Nachtsheim, C.J. (2009). "Split-plot designs: what, why, and how". *Journal of Quality Technology*, 41 (4): 340-361.

Khuri A. I. (1996). "Response surface models with mixed effects". *Journal of Quality Technology*, 28 (2): 177-186.

Khuri, A.I. and Conlon, M. (1981). "Simultaneous optimization of multiple responses represented by polynomial regression functions". *Technometrics*, 23 (4): 363-375.

Kowalski, S.M., Parker, P.A. and Vining, G.G. (2007). "Tutorial: industrial split-plot experiments". *Quality Engineering*, 19 (1): 1-15.

Kowalski, S.M. and Potcner, K.J. (2003). "How to recognize a split-plot experiment". *Quality Progress*, 36 (11): 60-66.

Lee, Y. and Nelder, J.A. (2003). "Robust design via generalized linear models". *Journal of Quality Technology*, 35 (1): 2-12.

León, R.V., Shoemaker, A.C. and Kacker, R. N. (1987). "Performance measures independent of adjustment". *Technometrics*, 29 (3): 253-285.

Lin, D.K.J. and Tu, W. (1995). "Dual response surface optimization". *Journal of Quality Technology*, 27 (1): 34-39.

Lu, L. and Anderson-Cook, C.M. (2012). "Rethinking the optimal response surface design for a first-order model with two-factor interactions, when protecting against curvature". *Quality Engineering*, 24 (3): 404-422.

Lu, L., Anderson-Cook, C.M. and Robinson, T. J. (2011). "Optimization of designed experiments based on multiple criteria utilizing a Pareto frontier". *Technometrics*, 53 (4): 353-365.

Lu, L., Chapman, J.L. and Anderson-Cook, C.M. (2017). "Multiple response optimization for higher dimensions in factors and responses". *Quality and Reliability Engineering International*, 33 (4): 727-744.

Marler, R.T., and Arora, J. S. (2004). "Survey of multi-objective optimization methods for engineering". *Structural and Multidisciplinary Optimization*, 26 (6): 369-395.

McCullagh, P. and Nelder, J.A. (1989). *Generalized linear models*. 2° ed. London: Chapman & Hall.

Myers, R.H., Khuri, A.I. and Vining, G.G. (1992). "Response surface alternatives to the Taguchi robust parameter design approach". *The American Statistician*, 46 (2): 131-139.

Myers, R.H., Montgomery, D.C. and Anderson-Cook, C.M. (2016). *Response surface methodology: process and product optimization using designed experiments*. 4° ed. Hoboken, New Jersey: Wiley.

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

Nelder, J.A. and Lee, Y. (1991). "Generalized linear models for the analysis of Taguchi-type experiments". *Applied Stochastic Models and Data Analysis*, 7 (1): 107-120.

Piattoli, L. (2020). *Disegno split-plot per l'ottimizzazione multiresponse in ambito tecnologico* [MD Thesis]. Florence: University of Florence, pp. 48-60.

Rajagopal, R., Del Castillo, E. and Peterson, J.J. (2005). "Model and distribution-robust process optimization with noise factors". *Journal of Quality Technology*, 37 (3): 210-222.

Robinson, T.J., Wulff, S.S., Montgomery, D.C. and Khuri, A.I. (2006). "Robust parameter design using generalized linear mixed models". *Journal of Quality Technology*, 38 (1): 65-75.

Tang, L.C. and Xu, K. (2002). "A unified approach for dual response surface optimization". *Journal of Quality Technology*, 34 (4): 437-447.

Vining, G.G. and Kowalski, S.M. (2008). "Exact inference for response surface designs within a split-plot structure". *Journal of Quality Technology*, 40 (4): 394-406.

Vining, G.G., Kowalski, S.M. and Montgomery, D. C. (2005). "Response surface designs within a split-plot structure". *Journal of Quality Technology*, 37 (2): 115-129.

Vining, G.G. and Myers, R.H. (1990). "Combining Taguchi and response surface philosophies: a dual response approach". *Journal of Quality Technology*, 22 (1): 38-45.

Zitzler, E. (1999). *Evolutionary algorithms for multiobjective optimization: methods and applications* [dissertation]. Zurich: Swiss Federal Institute of Technology Zurich, 5-10.