REVISTA DE GESTÃO E SECRETARIADO
MANAGEMENT AND ADMINISTRATIVE PROFESSIONAL REVIEW
ISSN: 2178-9010

Revista GeSec
São Paulo, SP, Brasil
v. 14, n.6, p. 9461-9471, 2023

DOI: http://doi.org/10.7769/gesec.v14i6.2310

Simultaneous optimization response surface experiments with several response variables

Experimentos de superfície de resposta de otimização simultânea com várias variáveis de resposta

Larissa Barbosa de Santana[1]
Vinícius Fernandes Rodrigues[2]
Nilo Antonio de Souza Sampaio[3]

Abstract

A problem facing the product development community is developing simultaneous solutions of response variables (to several properties) that depends on a number of independent variables or sets of responses. Harrington, among others, addressed this problem and presented a desirability function with a functional approach. Derringer and Suich altered their approach and illustrated how multiple variables can be transformed into a convenience function. This work redid the calculation performed by them using another software and made a comparative discussion of the results found.

Keywords: Statistical Software. Optimizer. Modeling. Desirability Function.

Resumo

Um problema enfrentado pela comunidade de desenvolvimento de produtos é desenvolver soluções simultâneas de variáveis de resposta (para várias propriedades) que dependem de um número de variáveis independentes ou conjuntos de respostas. Harrington, entre outros,

[1]Technical Degree in Logistics, Universidade do Estado do Rio de Janeiro, Rod. Pres. Dutra, km 298, Polo Industrial, Resende - RJ, CEP: 27537-000. E-mail: ls.ppcp@gmail.com
Orcid: https://orcid.org/0009-0007-1073-8876
[2]Bachelor's Degree in Administration, Universidade do Estado do Rio de Janeiro, Rod. Pres. Dutra, km 298, Polo Industrial, Resende - RJ, CEP: 27537-000. E-mail: viniciusferrodri@gmail.com
Orcid: https://orcid.org/0009-0009-5220-6418
[3]Doctor in Mechanical Engineering, Universidade do Estado do Rio de Janeiro, Rod. Pres. Dutra, km 298, Polo Industrial, Resende - RJ, CEP: 27537-000. E-mail: nilo.samp@terra.com.br
Orcid: https://orcid.org/0000-0002-6168-785X

abordou esse problema e apresentou uma função de desejabilidade com uma abordagem funcional. Derringer e Suich alteraram sua abordagem e ilustraram como múltiplas variáveis podem ser transformadas em uma função de conveniência. Este trabalho refez o cálculo realizado por eles utilizando outro software e fez uma discussão comparativa dos resultados encontrados.

Palavras-chave: Software Estatístico. Otimizador. Modelagem. Função de Desejabilidade.

Introduction

A common problem in product development involves selecting a set of conditions, the Xs (input variables), that will result in a product with a combination of properties, the Ys (output variables). Essentially this becomes a problem of simultaneous optimization of the Ys, or response variables, each of which depends on a set of independent variables. As an example that occurs a plasma etching process for the fabrication of integrated circuits. During fabrication, a mask of the tracks is drawn on a printed circuit, and the factors influencing this process are Distance, which ranges from 1 cm to 1.4 cm, and Power, which ranges from 350 W to 400 W. The response variables are the Engraving Rate in (Å/m), which must be maximized to get the best response in the process, and the Uniformity in (Å/m), which must be minimized to get the best response in the process (Fonseca et al., 2023). This problem was described in the book by Montgomery; D. C. from 2004. This paper performed the optimization with the same data proposed by the author using Minitab 19 software and reached conclusions using the desirability function.

Theoretical Framework

In the last century, statistics has revolutionized science by presenting useful models that have modernized the research process toward better research results, making it possible to guide decision making in many areas. Statistical methods were developed as a mixture of science and logic for the solution and investigation of problems in various areas of human knowledge. Today, statistics has contributed significantly to the decision-making process because much of what is produced is based on quantitative methods, and statistics is one such area. In the information and knowledge age, statistics uses mathematics to support professionals in business, government, and researchers (Mazza et al., 2023; Rezende et al., 2023).

Determining a process improvement is typically complex due to variations in customer demand and technological advances. Generally, several responses must be considered in order to achieve an overall process improvement (Maciel Gomes et al., 2019). The simultaneous optimization of multiple responses is necessarily a priority in many industries, and much of the effort has been directed to researching alternative methods for determining an efficient response that achieves a given goal. Multi-response optimization problems often involve conflicting objectives making it difficult to solve them, such as minimizing one response variable and maximizing another (Stojkovski, 2009).Currently the most used process optimization method in scientific work is the joint employment of the Desirability agglutination method with the mathematical Generalized Reduced Gradient search method (Corrêa et al., 2020; Yu et al., 2012).

DOE is a structured and organized method used to determine the relationship between different process input and output factors, involving the definition of the set of experiments, in which all relevant factors are systematically varied. By analyzing the results obtained, you can determine the degree of influence of each factor used on the response variable, as well as the interactions between the factors and the optimum conditions (Cardoso et al., 2023; G. Novaes et al., 2017; Rezende et al., 2023). In processes with multiple responses, you should model each of the responses you wish to optimize by means of a function that describes the so-called response surface, that is, which allows you to estimate the value of the response within the range of variation defined for the variables involved in the study. These functions (multiple regression equations) are usually obtained from the analysis of the results of experiments designed by the Box-Behnken, Central Composite or three level factorial designs, and are usually second order equations, highlighting that the Central Composite Design (CCD) model is the most widely used (G. Novaes et al., 2017; Kibria et al., 2014; Rehman et al., 2015).

A factor unknown by many studies using Design of Experiments (DOE) for process optimization, especially those involving multiple responses, is the individual quality of the models obtained (Derringer & Suich, 1980). In most cases, one or more models end up with a very low degree of fit. The success of the optimization is linked to the robustness of the models (Cardoso, Reis, Silva, Barros, et al., 2023; Gomes et al., 2019).One of the most used techniques for simultaneously optimizing multiple responses is to transform the equations that model each of these responses into individual utility functions and then proceed to optimize an overall utility function, known as Total Desirability (D), which is described in terms of the individual utility functions. The simultaneous optimization of multiple responses therefore becomes the optimization of a single function (Gomes et al., 2019; Vera Candioti et al., 2014).

Method

Scientific research can be classified as to its nature, approach, objectives, and procedures (Espuny et al., 2022; Kothari & Garg, 2019; Yin, 2017). Table 1 highlights the main classifications of this research. As for the nature, this work is characterized as applied research, because it has practical interest so that the results are applied and/or used in the solution of real problems (Provdanov & Freitas, 2013).

As to the objectives, this research is descriptive and exploratory. Descriptive because it allows to describe the characteristics of the phenomenon observed in relation to the delimitation made in this project and exploratory because it will provide greater familiarity of the researcher with the research problem in order to provide her with greater in loco contact/familiarity with the elements to be studied and the data were processed in May 2023 (Kothari & Garg, 2019).

NATURE OF RESEARCH	APPROACH	FIELD	OBJECTIVES/PURPOSES	RESEARCH PROCEDURES
BASIC	QUANTITATIVE	BIBLIOGRAPHICAL	EXPLORATORY	EXPERIMENT
APPLIED	QUALITATIVE	LABORATORY	EXPLANATORY	Survey
	COMBINED		DESCRIPTIVE	MODELING AND SIMULATION
			NORMATIVE	BIBLIOGRAPHIC REVIEW
				ACTION RESEARCH
				DOCUMENT SEARCH
				CASE STUDY

Table 1 - Research classification
Source: Adapted from Kothari and Garg (2019).

Results and Discussions

The data that appear in the book Introduction to Statistical Quality Control (Montgomery, 2004), in which a plasma etching process takes place for manufacturing integrated circuits. During fabrication a mask of the tracks is drawn on a printed circuit and the factors that influence this process are the Distance that varies from 1cm to 1.4 cm and the Power that varies from 350 W to 400 W. The response variables are the Engraving Rate in (Å/m) which must be maximized to get the best response in the process and the Uniformity in (Å/m) which must be minimized to get the best response in the process. Table 2 below presents the 13 non-randomized runs with the variables already decoded (because in the original table they were coded) performed using the Minitab 19 Software for the Response Surface Experiment, in this specific case using the Central Composite Design, with 5 replicates in the central point and a 45º rotation in the 4 star points, which is exactly in the vicinity of the optimal region.

OrderPad	OrderEns	KindPt	Blocks	Potency	Distance	Recording Rate	Uniformity
1	1	1	1	350	1	1054	79,7
2	2	1	1	400	1	926	81,3
3	3	1	1	350	1,4	1179	78,5
4	4	1	1	400	1,4	1417	97,7
5	5	-1	1	339,644661	1,2	1049	76,4
6	6	-1	1	410,355339	1,2	1187	88,1
7	7	-1	1	375	0,917157288	927	78,5
8	8	-1	1	375	1,482842712	1345	92,3
9	9	0	1	375	1,2	1151	90,1
10	10	0	1	375	1,2	1150	88,3
11	11	0	1	375	1,2	1177	89,6
12	12	0	1	375	1,2	1196	90,1
13	13	0	1	375	1,2	1180	90,3

Table 2: Experimental Design
Source: Authors 2023

In the sequence it was performed the analysis of the experiment taking into account the two response variables Recording Rate and Uniformity in function of Power called (A) for the purpose of mounting the equation of the model and Distance called (B), these variables squared and the interaction between both because as it is in the region of inclination it is necessary to have a complete quadratic model to represent it. Table 3, Table 4 and Table 5 below show the Coded Coefficients Table, Summarized Model and ANOVA Table for the Regression Rate where the p-value of the terms appear, all of them significant (< 5%) with the exception of the Distance*Distance interaction, the Coefficient of Determination (R2) of 98.66% which shows that 98.66% of the dependent variables are explained by the independent variables. In ANOVA Table 5 shows that both Linear and Quadratic models are suitable as both are significant and except for the Distance*Distance interaction, everything else is significant and the Lack of Fit is not significant (pvalue = 38.6%), which shows that the model fits the data very well.

Term	Coefficients	EP de Coeff	TValue	PValue	VIF
Constant	1170,80	9,48	123,52	0,000	
Pot	38,15	7,49	5,09	0,001	1,00
Dist	150,89	7,49	20,14	0,000	1,00
Pot*Pot	-22,15	8,04	-2,76	0,028	1,02
Dist*Dist	-13,15	8,04	-1,64	0,146	1,02
Pot*Dist	91,5	10,6	8,63	0,000	1,00

Table 3: Coded Coefficients
Source: Authors 2023

S	R^2	R^2(aj)	R^2(pred)
21,1950	98,66%	97,70%	94,21%

Table 4: Model Summary
Source: Authors 2023

Source	DF	SQ (Aj.)	QM (Aj.)	F Value	PValue
Model	5	231436	46287	103,04	0,000
Linear	2	193789	96895	215,69	0,000
Pot	1	11640	11640	25,91	0,001
Dist	1	182149	182149	405,47	0,000
Square	2	4158	2079	4,63	0,052
Pot*Pot	1	3413	3413	7,60	0,028
Dist*Dist	1	1203	1203	2,68	0,146
Interaction with 2 Factors	1	33489	33489	74,55	0,000
Pot*Dist	1	33489	33489	74,55	0,000
Error	7	3145	449		
Lack of Fit	3	1562	521	1,32	0,386
Pure error	4	1583	396	*	*
Total	12	234581			

Table 5: Analysis of Variance
Source: Authors 2023

Equation (1) shows the Regression Equation of the Model:

(1) Regression Equation in Uncoded Units

Recording Rate = 2471 + 6,15 A - 5319 B + 18,30 A*A - 329 B*B - 0,0354 A*B

In Figure-1 you can see the Residuals Graph, this residual is the difference from the expected value to the calculated value and it is expected that this residual has a Normal Distribution around zero. That the observed error versus the calculated error is normally distributed around zero. And that is exactly what appears in this figure.

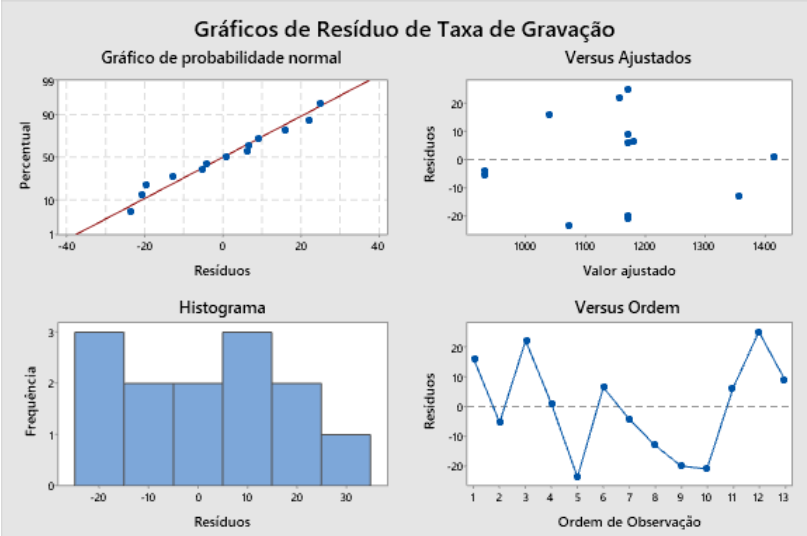

Figure-1: Residuals Plots for Recording Rate
Source: Authors 2023

The same analysis was performed for Uniformity generating the Tables 6,7 and 8 below, in which appear the Table of Coded Coefficients, Summarized Model and ANOVA Table for Uniformity in which appear the p-value of the terms, all of them significant (< 5%), the Coefficient of Determination (R2) of 98.51% which shows that 98.51% of the dependent variables are explained by the independent variables. In ANOVA Table 8 shows that both Linear and Quadratic models are adequate, because both are significant and all coefficients and interactions are significant and the Lack of Fit is not significant (pvalue = 19.5%), which shows that the model fits the data very well..

Termo	Coef	EP de Coef	T Value	P Value	VIF
Constant	89,680	0,469	191,33	0,000	
Pot	4,668	0,371	12,60	0,000	1,00
Dist	4,340	0,371	11,71	0,000	1,00
Pot*Pot	-3,596	0,397	-9,05	0,000	1,02
Dist*Dist	-2,021	0,397	-5,09	0,001	1,02
Pot*Dist	4,400	0,524	8,40	0,000	1,00

Table 6: Coded Coefficients
Source: Authors 2023

S	R2	R2(aj)	R2(pred)
1,04810	98,51%	97,45%	92,27%

Table 7: Model Summary
Source: Authors 2023

Source	DF	SQ (Aj.)	QM (Aj.)	F Value	PValue
Model	5	509,453	101,891	92,75	0,000
Linear	2	324,995	162,497	147,92	0,000
Pot	1	174,343	174,343	158,71	0,000
Dist	1	150,651	150,651	137,14	0,000

Square	2	107,019	53,509	48,71	0,000
Pot*Pot	1	89,969	89,969	81,90	0,000
Dist*Dist	1	28,421	28,421	25,87	0,001
Interaction with 2 Factors	1	77,440	77,440	70,49	0,000
Pot*Dist	1	77,440	77,440	70,49	0,000
Error	7	7,690	1,099		
Lack of Fit	3	5,042	1,681	2,54	0,195
Pure error	4	2,648	0,662	*	*
Total	12	517,143			

Table 8: Analysis of Variance
Source: Authors 2023

Equation (2) shows the Regression Equation of the Model:

(2) Regression Equation in Uncoded Units

Uniformity = -492 + 3,446 A - 187,0 B - 0,005754 A*A - 50,53 B*B + 0,880 A*B

In Figure 2 you can also see the Residuals Graph that also presents the Normalized Data.

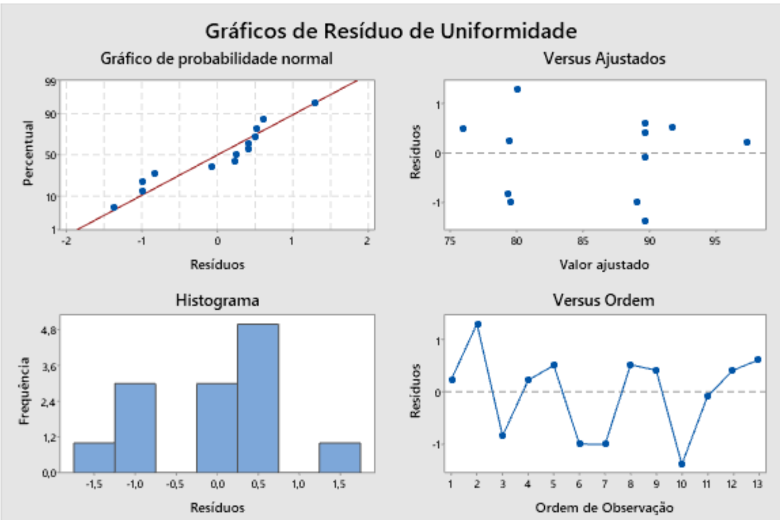

Figure-2: Residuals Plots for Recording Rate
Source: Authors 2023

Next, the optimization of the experiment was performed. The big problem is that, while one variable is optimized to meet the conditions required by the problem, the others are outside the ideal conditions required, that is, a conflicting behavior occurs and it is difficult to achieve the goals proposed for the problem. For this, a link function called Desirability was used. In Minitab software, you use the Optimize Response function and enter the desired targets to obtain the desired optimal response.

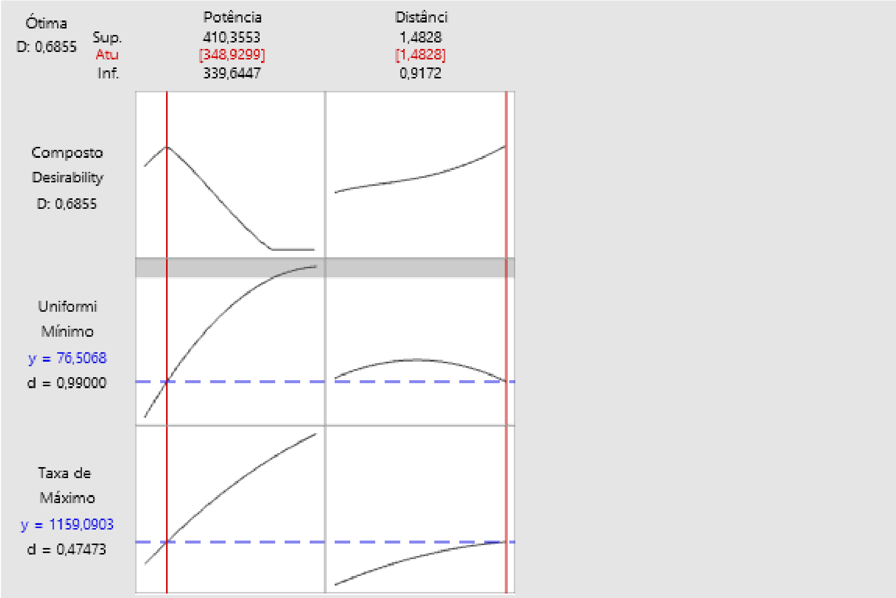

Figure-3: Optimized Response Graphs
Source: Authors 2023

In Figure-3 above can be observed the Optimized Response Graph, with the independent variables not coded the optimal values of Power (348.9299 W) and Distance (1.4828 Cm) were the ones found to reach the maximum value of Recording Rate (1159, 0903 Å/m) with an individual desirability of (d=0.47473) and the minimum value of Uniformity (76.5068 Å/m) with an individual desirability of (d=0.99) making a total desirability of d= 0.6855. Thus achieving the Process Optimization.

Conclusions

The main objective of this study was to perform an optimization using Response Surface Experiment with Central Composite with two response variables in the process described in the book by Montgomery; D. C. from 2004, in this sense, the objective was successfully achieved. The main academic contribution of this work was to show this optimization using Minitab 19 software and to perfectly identify the values of Power and Distance to maximize Recording Rate and minimize Uniformity. As a contribution for future works, the authors suggest the use of this type of optimization in other types of experiments, notably in the environmental area.

References

Cardoso, R. P., Reis, J. S. D. M., Silva, D. E. W., Almeida, M. da gloria diniz de, Barros, J. G. M. de, & Sampaio, N. A. de S. (2023). Scientific Research Trends About Metaheuristics in Process Optimization and Case Study Using the Desirability Function. *Revista de Gestão e Secretariado*, *14*(3), 3348–3367. https://doi.org/10.7769/gesec.v14i3.1809

Cardoso, R. P., Reis, J. S. da M., Silva, D. E. W., Barros, J. G. M. de, & Sampaio, N. A. de S. (2023). How to Perform a Simultaneous Optimization with Several Response Variables. *Revista de Gestão e Secretariado*, *14*(1), 564–578. https://doi.org/10.7769/gesec.v14i1.1536

Corrêa, J. M., dos Santos, E. L., Simões, M. R., Kadowaki, M. K., Gandra, R. F., & Simão, R. de C. G. (2020). Optimization of C. crescentus β-Xylosidases and Expression of xynB1–5 Genes in Response to Agro-Industrial Waste. *Waste and Biomass Valorization*, *11*(11), 6169–6178. https://doi.org/10.1007/s12649-019-00881-w

Derringer, G., & Suich, R. (1980). Simultaneous Optimization of Several Response Variables. *Journal of Quality Technology*, *12*(4), 214–219. https://doi.org/10.1080/00224065.1980.11980968

Espuny, M., Costa, A. C. F., Reis, J. S. da M., Barbosa, L. C. F. M., Carvalho, R., Santos, G., & Oliveira, O. J. de. (2022). Identification of the Elements and Systematisation of the Pillars of Solid Waste Management. *Quality Innovation Prosperity*, *26*(2), 147–169. https://doi.org/10.12776/qip.v26i2.1717

Fonseca, D., Correa, M. P. de O., Santos, R. R., Cardoso, R. P., Reis, J. S. da M., & Sampaio, N. A. de S. (2023). Effect of Pollution on Physical and Chemical Water Data: A Multivariate Statistical Analysis. *Revista de Gestão e Secretariado*, *14*(5), 7353–7366. https://doi.org/10.7769/gesec.v14i5.2125

G. Novaes, C., T. Yamaki, R., F. de Paula, V., B. do Nascimento Júnior, B., A. Barreto, J., S. Valasques, G., & A. Bezerra, M. (2017). Optimization of Analytical Methods Using Response Surface Methodology - Part I: Process Variables. *Revista Virtual de Química*, *9*(3), 1184–1215. https://doi.org/10.21577/1984-6835.20170070

Gomes, F. M., Pereira, F. M., Silva, A. F., & Silva, M. B. (2019). Multiple response optimization: Analysis of genetic programming for symbolic regression and assessment of desirability functions. *Knowledge-Based Systems*, *179*, 21–33. https://doi.org/10.1016/j.knosys.2019.05.002

Kibria, G., Doloi, B., & Bhattacharyya, B. (2014). Modelling and optimization of Nd:YAG laser micro-turning process during machining of aluminum oxide (Al 2 O 3) ceramics using response surface methodology and artificial neural network. *Manufacturing Review*, *1*, 12. https://doi.org/10.1051/mfreview/2014011

Kothari, C. R., & Garg, G. (2019). Research methodology methods and techniques. In *New Age International* (4º). New Age International.

Maciel Gomes, F., Monteiro Pereira, F., Augusto Silva Marins, F., & Borges Silva, M. (2019). Comparative study between different methods of agglutination in multiple response optimization. *Revista Gestão Da Produção Operações e Sistemas*, *14*(1), 95–113. https://doi.org/10.15675/gepros.v14i1.2080

Mazza, F. C., de Souza Sampaio, N. A., & von Mühlen, C. (2023). Hyperspeed method for analyzing organochloride pesticides in sediments using two-dimensional gas chromatography–time-of-flight mass spectrometry. *Analytical and Bioanalytical Chemistry*, *415*(13), 2629–2640. https://doi.org/10.1007/s00216-022-04464-y

Montgomery, D. C. (2004). *Introdução ao Controle Estatístico da Qualidade* (4aED ed.). LTC.

Provdanov, C. C., & Freitas, E. C. De. (2013). Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico. In *Universidade Feevale* (2º). Universidade Feevale. https://doi.org/10.1017/CBO9781107415324.004

Rehman, M. A., Yusoff, I., Ahmmad, R., & Alias, Y. (2015). Arsenic Adsorption Using Palm Oil Waste Clinker Sand Biotechnology: an Experimental and Optimization Approach. *Water, Air, & Soil Pollution*, *226*(5), 149. https://doi.org/10.1007/s11270-015-2411-9

Rezende, M. D., Rosa, C. S. da, Cardoso, R. P., Reis, J. S. da M., & Sampaio, N. A. de S. (2023). Statistics as a Tool for Decision Making in Agricultural and Environmental Experiments. *Revista de Gestão e Secretariado*, *14*(4), 5204–5217. https://doi.org/10.7769/gesec.v14i4.1978

Stojkovski, S. P. (2009). The Optimization of the Light-Duty Automotive Fleet for Cost Effective Fuel Efficiency. *SAE International Journal of Fuels and Lubricants*, *2*(1), 2009-01–0595. https://doi.org/10.4271/2009-01-0595

Vera Candioti, L., De Zan, M. M., Cámara, M. S., & Goicoechea, H. C. (2014). Experimental design and multiple response optimization. Using the desirability function in analytical methods development. *Talanta*, *124*, 123–138. https://doi.org/10.1016/j.talanta.2014.01.034

Yin, R. K. (2017). Case study research: design and methods. In *SAGE Publications* (6º). SAGE Publications.

Yu, H.-W., Kim, I. S., Niessner, R., & Knopp, D. (2012). Multiplex competitive microbead-based flow cytometric immunoassay using quantum dot fluorescent labels. *Analytica Chimica Acta*, *750*, 191–198. https://doi.org/10.1016/j.aca.2012.05.017

Submetido em: 02.05.2023

Aceito em: 06.06.2023