Defalque, Cristiane Maria
Silva, Aneirson Francisco da
Defalque, Guilherme Augusto
Marins, Fernando Augusto Silva
Edwards, David J.

A simulation algorithm for optimization of mixture design applied to the assignment of weights in a goal programming problem

http://hdl.handle.net/11067/7393
https://doi.org/10.34628/EYZ9-SK20

Metadados

Data de Publicação 2023

Resumo Experimental design approaches are essential for improving products and processes, and their use is often decisive in achieving a successful target result. Thus, mixture design is a method for designing experiments which considers that the result does not depend on the total amount but on the proportions of the components. Mixture design techniques are often applied to problems in food, beverage, pharmaceutical health, and cement-based materials, among others, and one may also use them to help s...

Tipo bookPart

Editora Universidade Lusíada Editora

A simulation algorithm for optimization of mixture design applied to the assignment of weights in a goal programming problem

DOI: https://doi.org/10.34628/EYZ9-SK20

Cristiane Maria Defalque [1,2*], Aneirson Francisco da Silva [1], Guilherme Augusto Defalque [3], Fernando Augusto Silva Marins [1], David J. Edwards[4]

[1] São Paulo State University "Júlio de Mesquita Filho",
Av. Ariberto Pereira da Cunha, 333, Portal das Colinas, Guaratinguetá, SP, 12516-410, Brazil
cristianedefalque.adm@gmail.com
aneirson.silva@unesp.br
fernando.marins@unesp.br

[2] Escola Preparatória de Cadetes do Exército
Av. Papa Pio XII, 350, Jardim Chapadão, Campinas - SP, 13070-091, Brazil
cristianedefalque.adm@gmail.com

[3] Federal University of Mato Grosso do Sul
Av. Costa e Silva, Pioneiros, Campo Grande, MS, 79070-900, Brazil
guilherme.defalque@ufms.br

[4] Virginia Commonwealth University
4124 Grace E. Harris Hall, 1015 Floyd Avenue, Box 843083, Richmond, VA, 23284, United States
dedwards7@vcu.edu

* Corresponding author

Abstract. Experimental design approaches are essential for improving products and processes, and their use is often decisive in achieving a successful target result. Thus, mixture design is a method for designing experiments which considers that the result does not depend on the total amount but on the proportions of the components. Mixture design techniques are often applied to problems in food, beverage, pharmaceutical health, and cement-based materials, among others, and one may also use them to help solve multi-objective problems when the weights of the objective function components can interfere with the optimization process. Therefore, given the relevance of studies on mixture planning and the increasing use of methods and techniques to consider uncertainty, the objective of this study is to propose an approach to deal with uncertainties in the coefficients of polynomial objective functions for the optimization of mixture design problem considering optimization via Monte Carlo Simulation. Computational tests were made using R software with instances from a literature study on a waste paper recycling logistics problem where the assignment of model weights is part of the process. Comparing the results to those obtained using the General Algebraic Modeling System language and CPLEX solver, they showed that considering uncertainty in the coefficients of objective function assisted in minimizing the difference between the obtained results, allowing for improvement in the representation of several scenarios. The developed approach also provided solution possibilities to help choose the best weights to optimize goal programming problems.

Keywords: Mixture design; Optimization via simulation; Weighted goal programming.

1. Introduction

Many studies in the literature use techniques of Response Surface Methodology (RSM) for modeling and analyzing the relationship between the factors of interest in a given system with applications found in, for example, industrial processes [1, 2], chemical processes [3], among others [4]. In this context, there are studies that use deterministic response surface models [5] and others that use uncertainty as a way to get values for the response variable closer to the actual values [1, 2, 6]

Variables that affect the performance of an experiment may have dependent levels on the proportions of their components. Mixture experiments, for example, consider techniques in which two or more factors are components or ingredients of a mixture, where the level of each factor depends on the levels of the others, constituting the proportions of this mixture [7–10]. Techniques used in mixture design are often applied to problems on food, beverage, pharmaceutical health [7], and when considering that the result does not depend on the total amount but the proportions of the components, they can be used to help solve multi-objective problems associated with the optimization of the weights of the objective function components.

Despite the differences associated with the domain, the linear dependence of the variables, and the methods used to design experiments, both the RSM and the Mixture design methods employ polynomial models that approximately associate the response with the input variables to describe the studied system and analyze the search space to find the best results [9, 10]. Thus, huge differences between the actual value and the response value obtained when optimizing the components of a polynomial model used to design the response surface in mixture experiments can also make it challenging to improve the performance of processes.

Since techniques for the design of experiments as those used in mixtures consider that the result does not depend on the total amount but the proportions of the components, they can be used to help solve multi-objective problems associated with the optimization of the weights of the objective function components.

Therefore, based on the study developed by [2] the objectives of this paper were to develop an approach to deal with uncertainties in the coefficients of polynomial objective functions for the optimization of mixture design problems considering optimization via Monte Carlo Simulation [11] and to assist with the choice of the best weights to optimize goal programming problems.

To develop the proposed algorithm, a search in the literature was carried out in Scopus [12] and the Web of Science [13] databases to verify the relevance of the publications related to mixture design, prioritization, and Optimization via Simulation, and justify this study. Table 1 summarizes the search method. It should be pointed out that when associating “Optimization via Simulation” and “Monte Carlo” with keywords related to mixture design and prioritization, we found no papers in the consulted databases.

Table 1. Summary of the search carried out in the databases

Steps	Keywords and Connectives	Number of documents found in the databases
1	(“Uncertain” OR “Uncertainty” OR “Risk” OR “Stochastic”) AND (“Mixture Experiment” OR “Mixture Design”)	Scopus: 206 The Web of Science: 107

Steps	Keywords and Connectives	Number of documents found in the databases
2	“Optimization via Monte Carlo Simulation” OR “Optimization by Monte Carlo Simulation”	Scopus: 9 The Web of Science: 5
3	“Mixture Experiment” OR “Mixture Design” AND (“Optimization via Simulation” OR “Optimization by Simulation”)	Scopus: 0 The Web of Science: 0
4	(“Multicriteria” OR “Multiresponse” OR “Multi-objective” OR “Goal Programming”) AND (“Priority” OR “Prioritization” OR “Weight” OR “Preference”)	Scopus: 20,529 The Web of Science: 7,625
5	(“Multicriteria” OR “Multiresponse” OR “Multi-objective” OR “Goal Programming”) AND (“Priority” OR “Prioritization” OR “Weight” OR “Preference”) AND (“Optimization via Simulation” OR “Optimization by Simulation”)	Scopus: 1 The Web of Science: 0
6	(“Multicriteria” OR “Multiresponse” OR “Multi-objective” OR “Goal Programming”) AND (“Priority” OR “Prioritization” OR “Weight” OR “Preference”) AND (“Optimization via Simulation” OR “Optimization by Simulation”) AND “Monte Carlo”	Scopus: 0 The Web of Science: 0

Source: search carried out in Scopus [12] and the Web of Science databases [13].

Therefore, the proposed method is an innovative technique that will help the decision-making process when getting a better fit for the regression model to minimize the difference between the obtained result and the real value.

This paper is organized as follows. Section 2 presents Methodological Procedures, Section 3 comprises Analysis and Discussion, and Section 4 presents the Conclusion, followed by the References.

2. Methodological Procedures

Based on definitions presented by [9], consider p the number of components in a mixture. If x_1, x_2, ..., x_p denote the proportions of these components in the mixture, then $0 \leq x_i \leq_1 1$, $i = 1, ..., p$, and $x_1 + x_2 + ... x_p = 1$. The response surface of a mixture experiment is a $(p - 1)$ - dimensional simplex. For example, for for $p = 2$, there is a line; $p = 3$, the simplex is a triangle; for $p = 4$, a tetrahedron represents the space [9, 10, 14]. Thus, model (1) represents a quadratic model for experiments with mixtures

[9]. The coefficients β_i represent the expected response for the pure mixture, $x_i = 1$ and for $x_j = 0$, and β_{ij} indicate synergism or antagonism of the binary mixture [10, 15].

$$\sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j>i}}^{p} \beta_{ij} x_i x_j \quad (1)$$

Based on that information, we got data and the Weighted Goal Programming Model from [16, 17], which considers eight objectives and, consequently, eight goals, for a waste paper logistics problem. We considered the instances R1 (original) and R3 (with gap = 0%) to carry out tests, and we organized the priorities according to units of measurement of the goals, as well as in some tests presented in [16]. That is x_1, x_2 and x_3 represent the priorities of three blocks of deviation variables associated with goals whose units of measurement are: [R$] (the official currency of Brazil); [km] (kilometers); and [t] (tonne).

After that, tests were performed using the General Algebraic Modeling System language, GAMS 23.5.2, [18–20] and CPLEX solver 12.2 [21], as well as the software used by [16, 17], considering thirteen different priorities, that is, thirteen combinations of values for three mixture components: three constituted of pure components, three of binary mixtures, combining 50% of each two components, one ternary mixture (1/3 of each component), and six formulations combining 2/3 and 1/3 of each two components. The tests were performed on an Intel Core i7-12650H computer with 2.30GHz and 16GB of RAM. The criterion for interrupting the considered programs was the time limit of 10.800 seconds.

Then, based on the information about the company's need, available in [16, 17], we selected the results of the following deviation variables to consider as response variables: sum of the "negative deviation from the goal of sales of bales of material" per material *i*, per customer *c*, in each period *t* (df_{ict}^{-}), sum of the "negative deviation from the goal of sales of bales of material" per material *i*, over the planning horizon (dq_i^{-}) [16, 17], and the sum of the total value of deviation variables whose unit measure is tonne (S_{total}).

Thus, the empirical function and the confidence intervals (CI) of 95% for all the coefficients of the independent variables were generated. Ordinary Least Square Algorithm (OLS) [10, 15] in the "mixexp" package from software R was used to draw objective functions [22, 23], which was optimized by the Augmented Lagrange Method available in the "Rsolnp" package [24, 25]. Based on [10], for choosing the model, we analyzed the value of R^2, normality, homoscedasticity of variances, and autocorrelation of the residuals. Considering the hierarchy, we also disregarded the terms that were not significant with $\alpha = 5\%$ in equation (1). In all tests, the objective was to minimize the objective function value.

In addition, we developed a stochastic simulation algorithm that uses the CI generated during the previous step for writing the objective function coefficients (β_i) as uniform random values, and Monte Carlo Simulation was inserted into the process. To solve problems with uncertainty, the value of the best feasible solution was obtained considering a minimum of 50,000 replications.

After that, analyses were performed comparing the optimized responses obtained from empirical functions with deterministic and stochastic coefficients with those obtained using the General Algebraic Modeling System language, GAMS 23.5.2, [18–20] and CPLEX solver 12.2 [21]. Figure 1 summarizes the steps of the process and Figure 2 presents the pseudo-code of the developed algorithm.

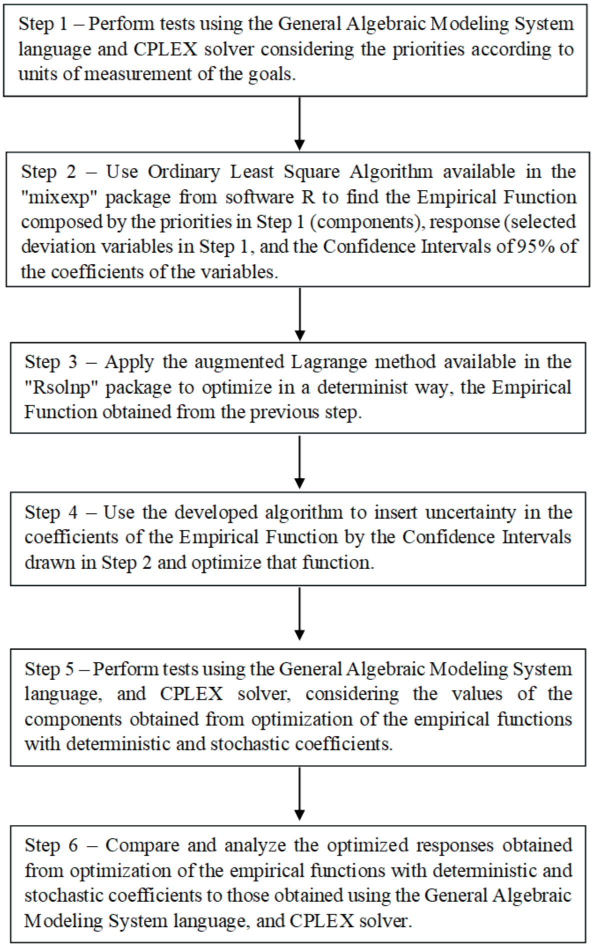

Figure 1. Steps of the Process.

Algorithm 1 Optimization via Simulation

```
data ← DATAFRAME()
model ← OLS_COEFFICIENTS(data)                          ▷ model is empirical function
β ← model_coefficients                                  ▷ β is coefficients of model
x ← model_variables                                     ▷ x is variables of model
β_confidence_interval ← CI(β)                           ▷ get confidence interval of β
β_u ← (0, 0, ..., n)                                    ▷ n is number of βs
x_u ← (0, 0, ..., m)                                    ▷ m is number of xs
y_u ← −∞
b ← (0, 0, ..., n)
for i = 0 to k do
    for j = 0 to n do
        b[j] ← RANDOM(β_confidence_interval[j])         ▷ uniform values
    end for
    obj ← OBJ(b, x)                                     ▷ objective function
    constraints ← CONSTRAINTS(x)                        ▷ constraints
    opt ← OPTIMIZE(obj, constraints)
    xs_optim ← opt_x                                    ▷ optimum xs
    y_optim ← opt_y                                     ▷ optimum y
    if y_optim > y_u then
        y_u ← y_optim
        for j = 0 to m do
            x_u ← xs[j]
        end for
        for j = 0 to n do
            β_u ← b[j]
        end for
    end if
end for
return y_u, β_u, x_u
```

Figure 2. Pseudo-code of the developed algorithm.

3. Analysis and Discussion

Table 2 shows statistically significant results obtained by R software [22, 23] of the p-values of Shapiro-Wilk, Breusch-Pagan, Durbin-Watson tests and Adjusted-R^2 of the quadratic model without the iteration x_1x_2 defined to design the experiment. Tables 3 and 4 show the deterministic coefficients of the empirical function, the new coefficients generated with the insertion of uncertainty, the optimized values of the independent variables x_1, x_2, and x_3, and the optimized values of the response variable obtained by R software [22-25] and GAMS/CPLEX [18-21] in the instances R1 and R3.

All the optimized values of the independent variables, calculated by the proposed algorithm, differed from those obtained with the deterministic algorithm. Although we could verify that the time limit considered to find a solution by GAMS/CPLEX was not enough to reach out GAP equal to zero, we stressed that when considering uncertainty, the values of the response variables were smaller. The checking tests by GAMS/CPLEX resulted in a minimal difference. Observe, for example, the values of dq_i^- in R1 and R3, S_{total} in R1 and df_{ict}^- in R3. While the differences between the values obtained with the stochastic algorithm and then with the GAMS/CPLEX were at most 5.355, the differences between the values obtained with the deterministic algorithm and with the GAMS/CPLEX reached 209.46 units.

The developed strategy to identify and consider weights in the model allows the analysis of different scenarios associated with, for example, variations in the availability of waste. That strategy can help the manager plan and optimize the development of processes, such as acquiring material and meeting demand [16,17].

Table 2. P-values of Empirical Functions associated with each response

Instances	Responses	Shapiro-Wilk test (p-value)	Breusch-Pagan test (p-value)	Durbin-Watson teste (p-value)	Adjusted-R^2
R1		0.33	0.19	0.14	0.91
		0.42	0.22	0.20	0.91
		0.37	0.19	0.16	0.91
R3		0.06	0.17	0.15	0.92
		0.04	0.12	0.10	0.94
		0.07	0.20	0.10	0.92

Source: Tests performed using R software [22, 23].

Table 3. Instance R1 - Results of the Empirical Functions with and without uncertainty

Parameters/Variables	Type of coefficients	$df_{\bar{ict}}$	$dq_{\bar{i}}$	S_{total}
β_1	Deterministic	766	692	1981
	Stochastic	839	663	2162
β_2	Deterministic	817	770	2100
	Stochastic	873	847	2216
β_3	Deterministic	426	327	829
	Stochastic	403	184	368
β_{13}	Deterministic	-906	-1018	-3147
	Stochastic	-1045	-1379	-2099
β_{23}	Deterministic	-1088	-1212	-3524
	Stochastic	-879	-711	-4068
x_1	Deterministic	0	0.09	0.11
	Stochastic	0.19	0.41	0.35
x_2	Deterministic	0.55	0.01	0
	Stochastic	0.26	0.23	0.21
x_3	Deterministic	0.45	0.9	0.89
	Stochastic	0.55	0.36	0.44

Parameters/Variables	Type of coefficients	df_{ict}^{-}	dq_i^{-}	S_{total}
y (proposed)	Deterministic	373	270	644
	Stochastic	373	270	644
y (GAMS)	Deterministic	373	277	853
	Stochastic	373	270	649
GAP (GAMS)	Deterministic	0	0.04	0.14
	Stochastic	0	0.05	0.004

Source: Tests performed using R software [22-25], GAMS 23.5.2 [18–20] and CPLEX solver 12.2 [21]

Table 4. Instance R3 - Results of the Empirical Functions with and without uncertainty.

Parameters/Variables	Type of coefficients	df_{ict}^{-}	dq_i^{-}	S_{total}
β_1	Deterministic	834	751	2110
	Stochastic	639	713	1854
β_2	Deterministic	823	742	2079
	Stochastic	674	777	2225
β_3	Deterministic	111	86	272
	Stochastic	-149	80	-258
β_{13}	Deterministic	-2067	-1877	-5125
	Stochastic	-2907	-1626	-7298
β_{23}	Deterministic	-1987	-1857	-5026
	Stochastic	-2906	-1288	-6870
x_1	Deterministic	0.55	0	0.54
	Stochastic	0	0.52	0.76
x_2	Deterministic	0	0.57	0
	Stochastic	0.78	0	0
x_3	Deterministic	0.45	0.43	0.46
	Stochastic	0.22	0.49	0.24
y (proposed)	Deterministic	0	0.0013	0
	Stochastic	0	0	0
y (GAMS)	Deterministic	67.08	0	14.846
	Stochastic	2.16	0	4.247
GAP (GAMS)	Deterministic	1	1	1
	Stochastic	1	1	1

Source: Tests performed using R software [22-25], GAMS 23.5.2 [18–20] and CPLEX solver 12.2 [21]

4. Conclusions and Future Research

The objective of this study was to develop an approach to deal with uncertainties in the coefficients of polynomial objective functions for the optimization of mixture design problems considering optimization via Monte Carlo Simulation. The proposal also aimed to assist with the choice of the best weights to optimize goal programming problems.

We verified that the proposed algorithm has shown competitive results concerning the deterministic model. When considering uncertainty in the coefficients of the objective function, the results obtained with the proposed method allowed for improvement in the representation of several scenarios. The proposal also provided solution possibilities to help choose the best weights to the optimize goal programming problem.

The algorithm can be adapted for considering and optimizing multiple responses in future research. In addition, the proposal can be applied to assist in solving other actual problems related to mixture design, such as detergents, soaps, food, and polymer concrete.

Acknowledgements

This research was partially supported by the National Council for Scientific and Technological Development (CNPq-303090/2021-9; CNPq-306868/2020-2; CNPq-304197/2021-1).

References

1. da Silva AF, Marins FAS, da Silva Oliveira JB, Dias EX (2021) Multi-objective optimization and finite element method combined with optimization via Monte Carlo simulation in a stamping process under uncertainty. The International Journal of Advanced Manufacturing Technology 117:305–327. https://doi.org/10.1007/s00170-021-07644-9

2. da Silva AF, Marins FAS, Dias EX, da Silva Oliveira JB (2019) Modeling the uncertainty in response surface methodology through optimization and Monte Carlo simulation: An application in stamping process. Mater Des 173:107776. https://doi.org/10.1016/j.matdes.2019.107776

3. Choi H-J, Naznin M, Alam MB, et al (2022) Optimization of the extraction conditions of Nypa fruticans Wurmb. using response surface methodology and artificial neural network. Food Chem 381:132086. https://doi.org/10.1016/j.foodchem.2022.132086

4. Smucker BJ, Edwards DJ, Weese ML (2021) Response surface models: To reduce or not to reduce? Journal of Quality Technology 53:197–216. https://doi.or

g/10.1080/00224065.2019.1705208

5. Babaki M, Yousefi M, Habibi Z, Mohammadi M (2017) Process optimization for biodiesel production from waste cooking oil using multi-enzyme systems through response surface methodology. Renew Energy 105:465–472. https://doi.org/10.1016/j.renene.2016.12.086

6. Silva AF, Silva Marins FA, Dias EX, Carvalho Miranda R (2022) Goal programming and multiple criteria data envelopment analysis combined with optimization and Monte Carlo simulation: An application in railway components. Expert Syst 39:e12840. https://doi.org/10.1111/exsy.12840

7. Galvan D, Effting L, Cremasco H, Conte-Junior CA (2021) Recent Applications of Mixture Designs in Beverages, Foods, and Pharmaceutical Health: A Systematic Review and Meta-Analysis. Foods 10:1941. https://doi.org/10.3390/foods10081941

8. Goupy Jacques, Creighton Lee (2007) Introduction to design of experiments with JMP examples, 3rd ed. SAS Institute, Cary, NC, USA

9. Scheffe H (1958) Experiments With Mixtures. Journal of the Royal Statistical Society: Series B (Methodological) 20:344–360. https://doi.org/10.1111/j.2517-6161.1958.tb00299.x

10. Montgomery DC (2013) Design and analysis of experiments, 8th ed. Wiley, New York

11. Rubinstein RY (1981) Simulation and The Monte Carlo Method. Wiley, New York

12. Scopus (2023) Scopus database. In: Elsevier B. V. www.scopus.com

13. Web of Science (2023) Web of Science database. In: Clarivate Analytics. www.webofknowledge.com

14. Claringbold PJ (1955) Use of the Simplex Design in the Study of Joint Action of Related Hormones. International Biometric Society 11:174–185. https://doi.org/10.2307/3001794

15. Cornell JA (2002) Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data, 3rd ed. Wiley, New York

16. Defalque CM (2021) Otimização Multiobjetivo Aplicada a Processos Logísticos de Resíduos de Papel: Abordagens de Programação por Metas Determinísticas e Sob Incerteza (in Portuguese). Faculdade de Engenharia de Guaratinguetá, Universidade Estadual Paulista “Júlio de Mesquita Filho”

17. Defalque CM, Silva AF, Marins FAS (2021) Goal Programming Model Applied to Waste Paper Logistics Processes. Appl Math Model 98:185–206. https://doi.org/10.1016/j.apm.2021.05.002
18. Brooke A, Kendrick D, Meeraus A (1997) GAMS: Sistema Geral de Modelagem Algébrica (in Portuguese), 1st ed. Edgard Blücher LTDA, São Paulo - SP - Brazil
19. GAMS Development Corporation (2023) User's Guide. https://www.gams.com/latest/docs/UG_MAIN.html#UG_Language_Environment
20. GAMS Development Corporation (2010) General Algebraic Modeling System (GAMS) Release 23.5.2. https://www.gams.com/latest/docs/RN_235.html#RN_2352
21. IBM ILOG (2010) IBM ILOG CPLEX 12.2. . https://www.ibm.com/support/pages/apar/RS00416
22. Lawson J, Willden C (2016) Mixture experiments in R using mixexp. J Stat Softw 72:. https://doi.org/10.18637/jss.v072.c02
23. R Core Team (2023) R: A Language and Environment for Statistical Computing. In: R Foundation for Statistical Computing. https://www.R-project.org/
24. Ghalanos A, Theussl S (2015) Rsolnp: General Non-linear Optimization Using Augmented Lagrange Multiplier Method
25. Ye Y (1987) Interior Algorithms for Linear, Quadratic, and Linearly Constrained Non-Linear Programming. Department of ESS, Stanford University