UFSM

CIÊNCIA e NATURA

ISSN 2179-460X
OPEN ACCESS

Ci. e Nat., Santa Maria, v. 46, e71479, 2023 • https://doi.org/10.5902/2179460X71479
Submitted: 29/08/2022 • Approved: 26/03/2024 • Published: 30/07/2024

Statistic

Application of the bootstrap method in the optimization of beet betacyanin extraction

Aplicação do método bootstrap na otimização da extração de betaciania de beterraba

Gabrielli Clímaco[I], Rita de Cássia Bergamasco[II], Diogo Francisco Rossoni[II]

[I] Universidade Estadual de Campinas, Campinas, SP, Brasil
[II] Universidade Estadual de Maringá, Maringá, PR, Brasil

ABSTRACT

The bootstrap method is a statistical tool that allows obtaining a better confidence interval without the need for repetitions, as is the case with the RSM. The objective of this work was to apply this method to the optimization of betacyanin extraction from beets. Time and temperature were used in a second-order CCRD, and the response variable was the betacyanin content. The optimum extraction point was obtained by RSM at 28min and 66°C. From the bootstrap analysis, a confidence ellipse was constructed, and the maximum amounts were located at 20-35min and 60-65ºC, confirming the maximum extraction point. These results suggest that the bootstrap method is useful in evaluating the reliability of the optimization of betacyanin extraction from beet, as predicted by RSM.

Keywords: Optimization; Bioactive compounds; Betacyanin; Resampling; Bootstra

RESUMO

O método bootstrap é uma ferramenta estatística que permite obter um intervalo de confiança melhor sem a necessidade de repetições, como é o caso do RSM. O objetivo deste trabalho foi aplicar este método na otimização da extração de betacianina de beterraba. Tempo e temperatura foram utilizados em um CCRD de segunda ordem, e a variável resposta foi o teor de betacianina. O ponto ótimo de extração foi obtido por RSM, a 28min e 66°C. A partir da análise de bootstrap foi construída uma elipse confiável, e as quantidades máximas foram localizadas a 20-35min e 60-65ºC, confirmando o ponto máximo de extração. Esses resultados sugerem que o método bootstrap é útil na avaliação da confiabilidade da otimização da extração de betacianina da beterraba, prevista pelo RSM.

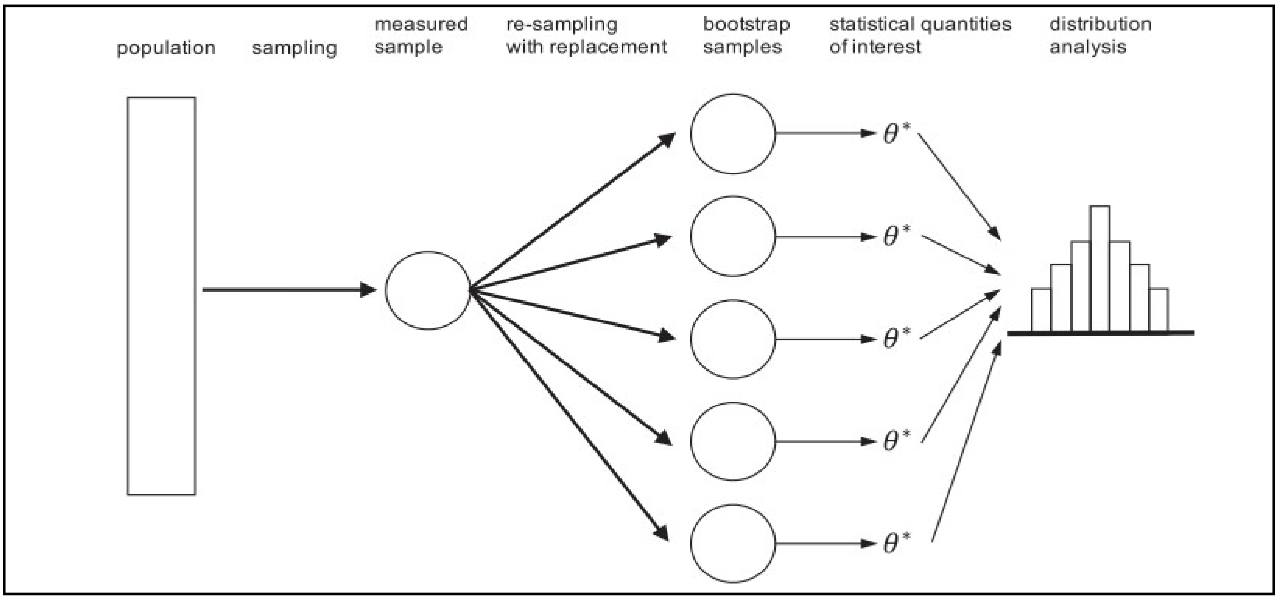
Palavras-chave: Otimização; Compostos bioativos; Betacianina; Reamostragem; Bootstrap

1 INTRODUCTION

Extraction optimization of bioactive compounds is a subject extensively studied by the scientific community because certain compounds have antioxidant, antiviral and antimicrobial properties, sequestering free radicals and preventing oxidation. Bioactive compounds can also be used as dyes, presenting low toxicity when compared to synthetic dyes (Maran, & Priya, 2016). Industries such as cosmetics, pharmaceuticals and food are the most interested in such research, as these ingredients can be used in the manufacture of dietary supplements, nutraceuticals, pharmaceuticals and cosmetics (Gil-Chávez et al., 2013).

The response surface methodology (RSM) is a mathematical and statistical tool that can simultaneously evaluate the influence of different factors and optimize their experimental conditions (Chen, Zhao, & Yu, 2015). In addition to RSM, the bootstrap method can also be used, because it resamples using the original data, obtaining high confidence intervals (Efron, 2003) experiment, without necessarily repeating it (Figure 1) (Silva et al., 2018).

Figure 1 – Schematic illustration of the bootstrap method



Source: MATSUYAMA, 2018

The difference between an experiment with or without bootstrap, in addition to reliability, basically boils down to cost and time reduction, as in the bootstrap the experiment is repeated by programming, with a high confidence factor (which can be above 95%), while to obtain an operational confidence interval using the RSM, that is similar to that obtained by bootstrap, it is necessary to fully repeat the experiment, using raw material, reagents and time. However, in general, the two methodologies manage to reach the same result, just following different paths, which may or may not be more costly on an industrial scale.

Several studies have been carried out in food science using the bootstrap method. Among them is the analysis of growth behavior of pathogenic and nonpathogenic strains of *E. coli*, at different temperatures and media, in which bootstrap models were used to characterize the distributions of deformation variability (Quinto et al., 2018). Other examples are in the obtention of confidence intervals for the parameters analyzed, during contamination quantification by *Campylobacter jejuni* in chicken carcasses (Duqué et al., 2017), and in the quantification of anthocyanins present in sprouts of *Lonicera japonica* and *L. japonica* (Yuan et al., 2014).

The bootstrap method has been successfully applied in the pharmaceutical area to evaluate the reliability of optimal drug formulations predicted by RSM (Arai et al., 2009, 2011; Duangjit et al., 2014).

The extraction of bioactive compounds is a widely studied branch of food engineering, however, aiming for an effective optimization, many factors come into the balance, with the reliability of the results being one of the main ones. For this, finding a confidence interval that allows a variation in factors is essential, because from an industrial point of view, equipment often cannot be programmed exactly at the optimum point of temperature and pressure, for example, so if for this compound there is a confidence interval where this equipment can be adjusted, without the extraction efficiency being reduced, it will optimize the work.

Beet was selected as a case study, because it presents good dye yield and a prominent peak in the visible region of the spectrum, for a better quantitative analysis. Beet contains betalain, which is composed of betacyanin (responsible for red color) and betaxanthin (responsible for yellow color). Betacyanin's account for 75–90% of the total color present in beet. Beet dye is associated with nutritional value and non-toxicity, suitable for application as a dye, where health aspects are priority (Sivakumar et al., 2009). Thus, the application of the bootstrap method in the optimization of betacyanin extraction from beet is helpful to obtain an optimal region of maximum dye yield, with potential application in diverse industries.

The present work aimed to apply the bootstrap method in the optimization of betacyanin extraction from beet, to evaluate the reliability of the optimal extraction point predicted by RSM.

2 MATERIALS AND METHODS

2.1 Material Preparation

Raw beets were obtained in the local commerce in the city of Maringa, Brazil, on the day of the analysis. The beet (approximately 200 g) was washed under running water, peeled (discarding the bark) and grated. The extraction was carried out according to Kushwaha et al., (2017), with some modifications. A previously heavy beet pulp mass was added to an Erlenmeyer containing 50 ml of distilled water. The material was placed in a water bath, without agitation, where the extraction time and temperature were varied. After extraction, the material was filtered on filter paper (0.45 μm) to obtain the extract. The assays were randomly performed, thereby reducing the effects of unexplained variability on responses due to external factors and ensuring independence among the observations.

2.2 Quantification of Betacyanin

The quantification of betacyanin was determined by spectrophotometry Cary 50 (Hayward, CA), according to Equation 1 (Stintzing et al., 2005).

$$(mg/100g\ fresh\ beet) = \frac{AxMMxDFx10^2}{\varepsilon xLxC_a} \quad (1)$$

Where:

is the absorbance value at wavelengths 536 nm (specific for betacyanin);

A is the dilution factor;

DF is the length of the cuvette path (1 cm);

C_a is the concentration of solid/solvent (g/ml).

The molar mass (MM) and the molar extinction coefficient(ε) of betacyanin (MM = 550 g/mol; ε = 60000 M-1 cm-1) was applied to quantify the pigment (Swamy, Sangamithra, & Chandrasekar, 2014).

2.3 Experimental planning

Initially, a 2^3 factorial design was performed, with 6 repetitions at the central point, to find a real interval for the factors (time, temperature, and mass), which presented the maximum point of the compound under analysis. The factors were defined according to literature (Swamy, Sangamithra, & Chandrasekar, 2014) with intervals 42 to 62 minutes, 41 to 61 ° C and 1.5 to 2.5 g of beet. The concentration of betacyanin in the extract was the response variable.

2.4 Extraction optimization

After obtaining the response variable for each point of the experiment, the procedure of steepest ascent was performed, in order to find the maximum extraction peak, according to the (Werkema, & Aguiar, 1996).

With the results of the steepest ascent test, it was observed that the mass factor was limited by 0.5 g. Then, a second-order central composite rotational design (CCRD) (2^2) (Table 1) was developed, with 5 repetitions at the central point and 4 axial points, with time and temperature as factors.

Table 1 – Factors and levels used in the central composite rotational design (CCRD)

Factors	**Coded Levels**				
	-1.414	**-1**	**0**	**1**	**1.414**
Time (min)	16	20	30	40	44
Temperature (ºC)	50	55	65	75	80

Source: Authorship (2022)

The concentration of betacyanin were analyzed using the response surface methodology, based on a second-order polynomial model and the bootstrap method. Statistical analysis was performed in the Software RStudio (RStudio, 2021). Data modeling began with a second-order model, which included linear, quadratic, and interaction terms. The adaptations of the models were analyzed in terms of R^2, R^2 adjusted and the significance of the lack of fit. Using Analysis of variance (ANOVA), the significant terms in the model were calculated for each response and the ANOVA tables were created. From the regression models obtained, the regression coefficients were used for statistical calculations to generate response surface graphs. In addition, Quantil-Quantil plot and Shapiro-Wilk test were produced to evaluate if errors follow a normal distribution.

2.5 Bootstrap Method

The bootstrap analysis starts from an adjusted model, where Y are the experimentally obtained data, and the Bootstrap algorithm for a linear regression model, follows the steps below (Davison, & Hinkley, 1997).

1. Consider
2. Resample the sample with
3. The new sample will be
4. Obtain estimates of interest
5. Return to step 1 and repeat the process at least 1000 times
6. Stop

For greater confidence in the results, the data were resampled 2000 times and, in each surface, a maximum point was established, referring to the highest yield. With the overlap of the surfaces, it was possible to obtain an ellipse with 95% confidence for the maximums, that is, a region within the pre-established time and temperature levels, where it is possible to obtain maximum yield.

This resampling is an effective technique, because it allows uncertainty quantification from the calculations of standard errors and confidence intervals, in addition to performing significance tests, providing more accurate answers than other methods, and using fewer assumptions, which later facilitate an industrial application (Filho, 2010; Mishra, Dolan, & Yang, 2011).

3 RESULTS AND DISCUSSION

The results of betacyanin (mg/100g) obtained in the CCRD for the two variables studied (time and temperature) are shown in Table 2.

The betacyanin content ranged from 48.216 to 88.420 mg/100g, the maximum being found at the central point (tests 5 to 10), under the condition of 30 minutes and 65 ºC.

Analyzing the influence of temperature, the lowest yield was observed in test 3, with 20 minutes and 75ºC, being 48.216 mg/100g. A similar result was observed (Swamy, Sangamithra, & Chandrasekar, 2014) in the aqueous extraction of beet pigments, where the authors observed a decrease in the amount of betacyanin at 70ºC, thus evidencing its instability at higher temperatures. From the Analysis of variance

(ANOVA) (Table 3), a non-significant adjustment ($p > 0.05$) for the lack of fit and a R^2 value of 0.942 are observed, implying a good fit of the model. It is important to note that only the analysis of the coefficient of determination does not justify the lack of fit of the model, because it measures the reduction of response variability using the regressor variables in the model.

Table 2 – Response Variable obtained in the experimental design

Test	Time (minutes)	Temperature (ºC)	Betacyanin (mg/100g)
1	20	55	79.565
2	40	55	56.452
3	20	75	48.216
4	40	75	53.441
5	30	65	88.422
6	30	65	85.063
7	30	65	83.077
8	30	65	84.299
9	30	65	76.816
10	30	65	82.772
11	16	65	57.475
12	44	65	62.516
13	30	50	76.358
14	30	80	50.691

Analyzing the influence of temperature, the lowest yield was observed in test 3, with 20 minutes and 75ºC, being 48.216 mg/100g. A similar result was observed (Swamy, Sangamithra, & Chandrasekar, 2014) in the aqueous extraction of beet pigments, where the authors observed a decrease in the amount of betacyanin at 70ºC, thus evidencing its instability at higher temperatures. From the Analysis of variance (ANOVA) (Table 3), a non-significant adjustment ($p > 0.05$) for the lack of fit and a R^2 value of 0.942 are observed, implying a good fit of the model. It is important to note that only the analysis of the coefficient of determination does not justify the lack of fit of the model, because it measures the reduction of response variability using the regressor variables in the model.

Thus, a high value does not necessarily imply a good adjustment (Bas, & Boyaci, 2007). In addition, it is interesting that both the linear model (FO) and the quadratic model (PQ) are significant (p 0.05) for betacyanin extraction.

Table 3 – Analysis of variance for the optimization of betacyanin extraction

	Gl	Sq	Qm	F-value	Pr(F)
FO (x1, x2)	2	638.53	319.27	15.7421	0.0016
TWI (x1, x2)	1	200.76	200.76	9.8988	0.0136
PQ (x1, x2)	2	1797.80	898.90	44.3222	4695×10^{-5}
Residue	8	162.25	20.28		
"Lack of fit"	3	89.60	29.87	2.0557	0.2248
Pure Error	5	72.65	14.53		
R^2	0.942				
R^2 adjusted	0.9058				

FO: First order polynomial; TWI: 2-way interaction; PQ: Quadratic model; Significant factors

In the coefficient estimates (Table 4), it is noticeable that temperature had a significant effect on extraction, as well as the interaction between time and temperature and its quadratic terms. This fact is also verified in Table 2, in which it is possible to observe an increase in the amount of compound in the central region (central points), but also its degradation when exposed to high temperatures for a long period of time.

Table 4 – Coefficient estimation for the optimization of extraction of betacyanin

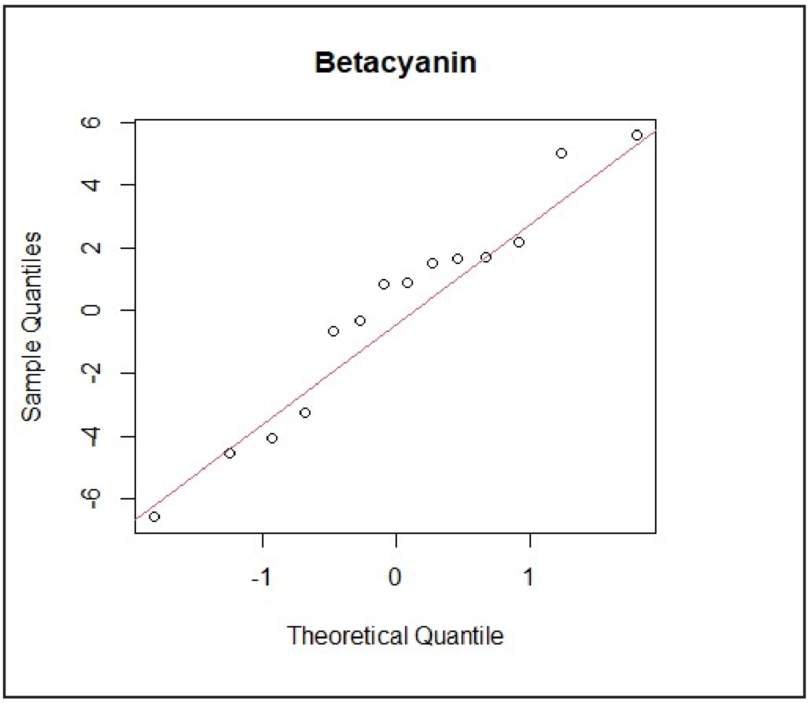
	Estimate	Standard error	t-value	Pr (|t|)
(Intercept)	83.409	1.838	45.36	6.155×10^{-11} [a]
x1 (Time)	-1.345	1.592	-0.844	0.422
x2 (Temperature)	-8.832	1.592	-5.547	5.428×10^{-4} [a]
x1:x2	7.084	2.251	3.146	0.013[a]
x1^2	-12.29	1.657	-7.416	7.49×10^{-5} [a]
x2^2	-10.52	1.657	-6.351	2.20×10^{-4} [a]

[a] significant factor p 0.05

By estimating the coefficients, it is observed that time and temperature have negative effects on the response variable, higher betacyanin content is extracted at lower temperature and time, as it can suffer degradation when exposed to high temperatures for a long period. However, the time versus temperature interaction causes a positive effect on the content extracted.

To analyze the normality of the residues, Quantil-Quantil plot (Figure 2) and Shapiro-Wilk test were performed.

Figure 2 – Quantil-Quantil plot of residuals from second-order model for betacyanin response



Source: Authorship, 2022

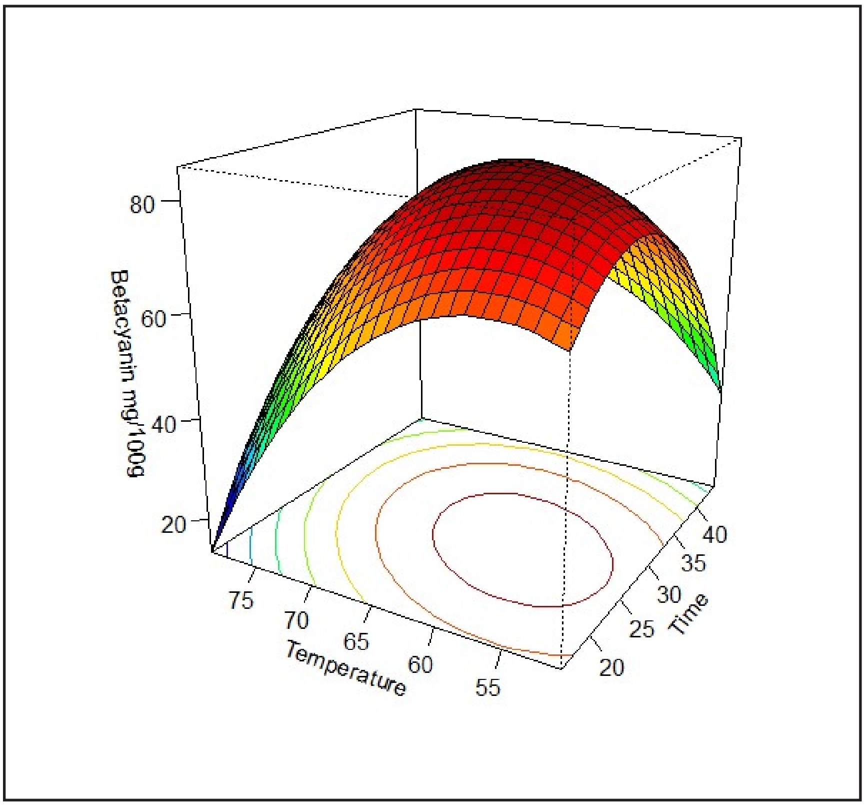
In the Quantil-Quantil plot, it is observed that the points are close to a line, indicating normality. In the Shapiro-Wilk normality test, the p-value was 0.95071. Thus, it can be affirmed that the residues can be considered normal with 5% significance.

3.1 Stationary point

The stationary point can be located in three ways on a response surface, being a maximum, minimum or saddle point (Montgomery, 1976). For betacyanin, the stationary point is present at – 0.1944 and – 0.4484, generating an optimal point for data close to the estimated ideal time of 28minutes and temperature of 66 °C, obtaining 83.40mg/100g of betacyanin.

The response surface graph was constructed for the compound (Figure 3), to visually verify the interaction between the two factors analyzed and allow visualization of the ideal extraction point.

Figure 3 – Betacyanin response surface, in the operating conditions Time x Temperature



Source: Authorship, 2019

Again, the effect of temperature on betacyanin extraction is verified, in which the best results were observed at temperatures close to the central point (approximately

65 °C), with time ranging from 20 to 30 minutes. In general, temperature acts to soften the plant tissue, disturbing the interaction between its constituents, such as phenolic compounds and carbohydrates, increasing its solubility and, consequently, increasing extraction (Maran and Priya, 2016; Swamy et al., 2014)time and mass of beetroot on the aqueous extraction of betalain from beetroot. The optimum conditions for the aqueous extraction of betalamic acid, betaxanthin and betacyanin from beetroot were performed using a three-factor and three-level Box-Behnken design under response surface methodology. The pigments were extracted from beetroot at temperature (40-70 °C).

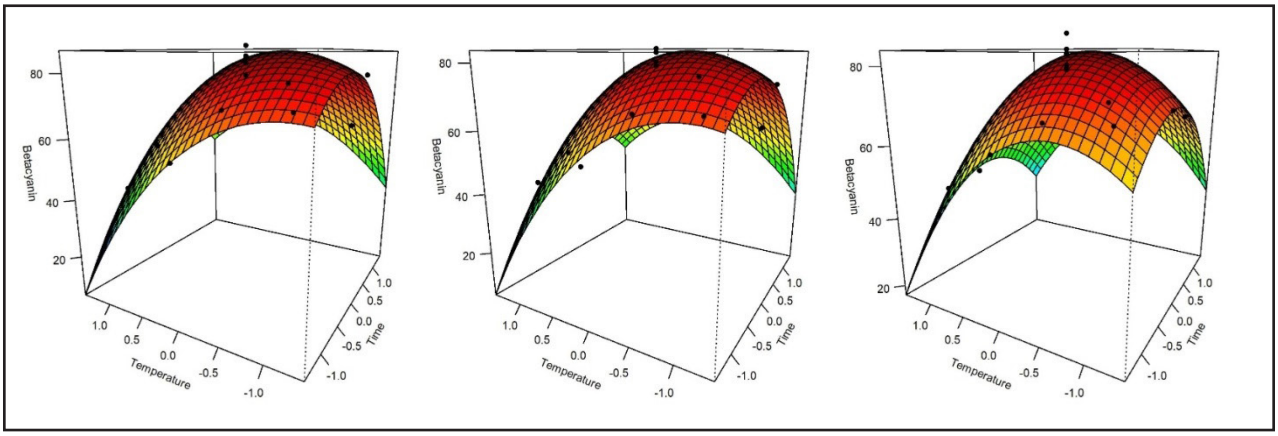
3.2 Bootstrap method analysis

Correlation coefficients are values that indicate the stability of the response surface; however, the reliability of the optimal point cannot be evaluated using these values. In this case, the bootstrap method was used to evaluate the reliability of the optimal point predicted by RSM (Duangjit et al., 2014)although the direct reliability of the optimal formulation must be evaluated. In this study, we demonstrated the feasibility of using the bootstrap (BS).

Initially, data resampling was performed (Figure 4), to demonstrate what the method can do. The new resampled surfaces are very similar to those generated by the experimental data (Figure 3), because these values are very close to the predicted values, this affirms the feasibility of this approach for the extraction of bioactive compounds. It is noteworthy thatthe data were resampled 2000 times.

It was also possible to observe that the maximum region was located in the central region, that is, around zero, both for time and temperature. Remembering that in Figure 4, the factors are presented at coded levels of the design. This only proves what was previously analyzed, that the highest yield is obtained in the region where the central points of the design were delimited, being 30 minutes and 65ºC.

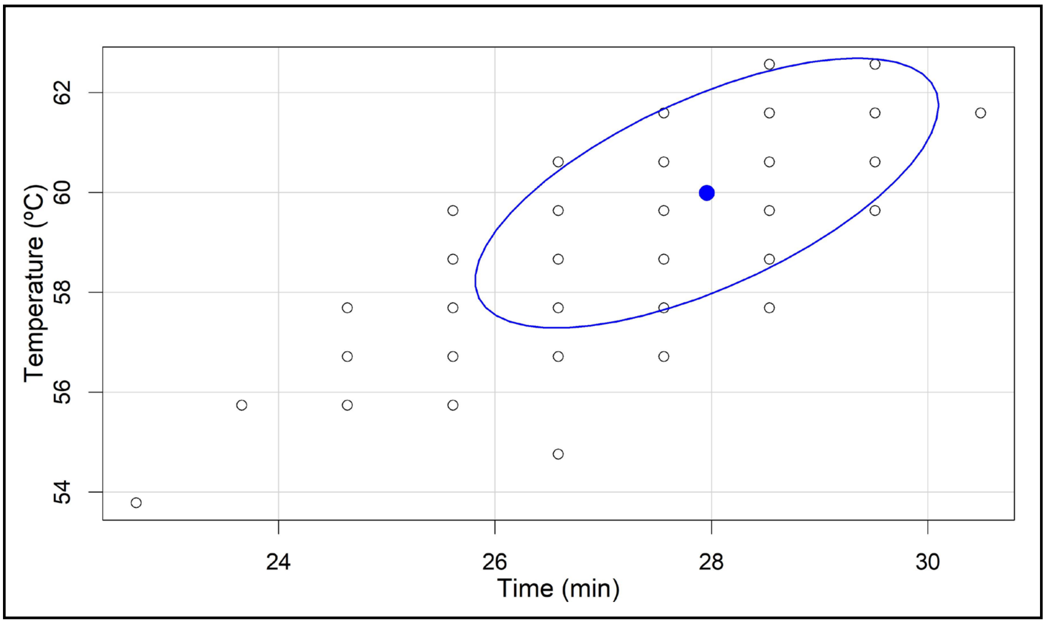
Figure 4 – Surfaces with resampled data for betacyanin, extraction conditions Time x Temperature



Source: Authorship, 2019

From the experimental data (Table 3), it was possible to obtain an optimal point, using the response surface methodology, so that, for each new surface generated, new optimal points were also obtained, thus creating an optimal two-dimensional distribution (Figure 5), from which a reliable ellipse (95%) was obtained, where it was possible to obtain maximum compound amounts.

Figure 5 – Confiance ellipse (95%) of optimal points for betacyanin



Source: Authorship, 2019

From the experimental data (Table 3), it was possible to obtain an optimal point using the response surface methodology, so that for each new surface generated (through bootstrap resampling), new optimal points were also obtained, thus creating an optimal two – dimensional distribution (Figure 5), from which a reliable ellipse (95%) was obtained, where it was possible to obtain maximum amounts of the compound. Thus, it can be stated that within this confidence interval it is possible to obtain the same amount of compounds from the optimal (central) point, thus allowing for variations in time and temperature, according to the specifications of each equipment.

Similar coding was used in Figure 4, where the X and Y axes represent time, ranging from 20 to 40 minutes and temperature, from 65 to 76°C. Thus, the reliable ellipseis present near the central point, between 20-30 minutes and 60-65ºC.

4 CONCLUSIONS

In this work, the bootstrap method with resampling was applied in the optimization of betacyanin extraction from beet. A second-order central composite rotational design was carried out in the experiment, with extraction time and temperature as factors. The betacyanin content in beet aqueous extract was defined as response variable. The optimal point of extraction was determined by RSM. The reliability of the results obtained was evaluated by the bootstrap method, with the construction of a reliable ellipse (95% of confidence) confirming the location of the maximum extraction point. Thus, the bootstrap method is a useful tool that can be successfully applied to optimize the extraction of bioactive compounds from plants, as in addition to evaluating the reliability of the results predicted by the RSM, it allows the creation of an operational confidence interval, thus facilitating extractions on an industrial scale, with cost reduction involving reagents and raw materials and time optimization.

REFERENCES

Arai, H. et al. (2009). Effect of an experimental design for evaluating the nonlinear optimal formulation of theophylline tablets using a bootstrap resampling technique. Chemical and Pharmaceutical Bulletin, 57(6), 572–579.

Arai, H. et al. (2011). Reliability Evaluation of the Design Space of the Granulation Process of Mefenamic Acid Tablets Using a Bootstrap Resampling Technique. Chemical and Pharmaceutical Bulletin, 59(5), 608–617, 2011.

Bas, D., & Boyaci, I. H. (2007). Modeling and optimization i: Usability of response surface methodology. Journal of Food Engineering, 78(3), 836–845.

Chen, M., Zhao, Y., & Yu, S. (2015). Optimization of ultrasonic-assisted extraction of phenolic compounds, antioxidants, and anthocyanins from sugar beet molasses. Food Chemistry, 172, 543–550.

Davison, A. C., & Hinkley, D. V. (1997). Bootstrap Methods and their Application. United States of America: Cambridge University Press 1997.

Duangjit, S. et al. (2014). Bootstrap resampling technique to evaluate the reliability of the optimal liposome formulation: Skin permeability and stability response variables. Biological and Pharmaceutical Bulletin, 37(9), 1543–1549.

Duqué, B. et al. (2017). Quantification of Campylobacter jejuni contamination on chicken carcasses in France. Food Research International, n. August, 1–9.

Efron, B. (2003). Second Thoughts on the Bootstrap. Statistical Science, 18(2), 135–140.

Filho, A. S. Da S.(2010). Inferência em amostras pequenas: métodos bootstrap. Revista de Ciência Exata e Tecnologia, 5(5), 115–126.

Gil-Chávez, G. J. (2013). et al. Technologies for Extraction and Production of Bioactive Compounds to be Used as Nutraceuticals and Food Ingredients: An Overview. Comprehensive Reviews in Food Science and Food Safety, 12(2008), 5– 23.

Maran, J. P.,& Priya, B. (2016). Multivariate statistical analysis and optimization of ultrasound – assisted extraction of natural pigments from waste red beet stalks. Journal of Food Science and Technology, 53(1), 792–799.

Matsuyama, T. (2018). An application of bootstrap method for analysis of particle size distribution. Advanced Powder Technology, 29(6), 1404–1408.

Mishra, D. K., Dolan, K. D., & Yang, L. (2011). Bootstrap confidence intervals for the kinetic parameters of degradation of anthocyanins in grape pomace. Journal of Food Process Engineering, 34(517), 1220–1233.

Montgomery, D. C. (1976). Design and Analysis of Experiments. 5. ed. [s.l.] John Wiley & Sons.

Quinto, E. J. et al. (2018). Bootstrap parametric GB2 and bootstrap nonparametric distributions for studying shiga toxin-producing Escherichia coli strains growth rate variability. Food Research International.

RStudio. (2021). RStudio: Integrated Development for R.Bostom, MA. Disponível em: http://www.rstudio.com/

Silva, C. M. et al. (2018). Controle estatístico de qualidade usando o método bootstrap. Scientia Pllena, 14(3), 1–9.

Sivakumar, V. et al. (2009). Ultrasound assisted enhancement in natural dye extraction from beetroot for industrial applications and natural dyeing of leather. Ultrasonics – Sonochemistry, 16(6), 782–789.

Stintzing, F. C. et al. (2005). Color, betalain pattern, and antioxidant properties of cactus pear (Opuntia spp.) clones. Journal of Agricultural and Food Chemistry, 53(2), 442– 451.

Swamy, G. J., Sangamithra, A., & Chandrasekar, V.(2014). Response surface modeling and process optimization of aqueous extraction of natural pigments from Beta vulgaris using Box-Behnken design of experiments. Dyes and Pigments, 111, 64–74.

Werkema, M. C. C., & Aguiar, S. (1996). Otimização estatística de processos: como determinar a condição de operação de um processo que leva ao alcance de uma meta de melhoria. [s.l.] Editora QFCO.

Yuan, Y. et al. (2014). Anthocyanins from buds of Lonicera japonica Thunb. var. chinensis (Wats.) Bak. FRIN, 62, 812–818.

Authorship contributions

1 – Gabrielli Clímaco

Universidade Estadual de Campinas - UNICAMP
https://orcid.org/0000-0002-6113-1616 • gabiclimaco7@gmail.com
Contribution: conceptualization, data curation, formal analysis, investigation, validation, visualization, writing – original draft.

2 – Diogo Francisco Rossoni

Universidade Estadual de Maringá - UEM
https://orcid.org/0000-0001-6337-6628 • dfrossoni@uem.br
Contribution: methodology, project administration, resources, software, supervision, writing, review & editing.

3 –Rita de Cássia Bergamasco

Universidade Estadual de Maringá - UEM
https://orcid.org/0000-0003-4318-2353 • rcbergamasco@uem.br
Contribution: methodology, resources, supervision, writing, review & editing.

How to quote this article

Clímaco, G., Rossoni, D. F., & Bergamasco, R. de C. (2024). Application of the bootstrap method in the optimization of beet betacyanin extraction. Ciência e Natura, 46. e71479. https://doi.org/10.5902/2179460X71479.