Received: 10 March 2019 | Revised: 5 May 2019 | Accepted: 7 May 2019

DOI: 10.1111/jfpe.13103

ORIGINAL ARTICLE

Journal of Food Process Engineering

WILEY

Brewing process optimization by artificial neural network and evolutionary algorithm approach

Maria Beatriz Takahashi[1] | Henrique Coelho de Oliveira[1] | Eutimio Gustavo Fernández Núñez[2,3] | José Celso Rocha[1]

[1]Departamento de Ciências Biológicas, Universidade Estadual Paulista-UNESP/Assis, Assis, São Paulo, Brazil

[2]Centro de Ciências Naturais e Humanas (CCNH), Universidade Federal do ABC, Santo André, São Paulo, Brazil

[3]Escola de Artes, Ciências e Humanidades (EACH), Universidade de São Paulo, São Paulo, São Paulo, Brazil

Correspondence

Eutimio Gustavo Fernández Núñez, Escola de Artes, Ciências e Humanidades (EACH), Universidade de São Paulo, Rua Arlindo Bettio, 1000, Vila Guaraciaba, 03828-000, São Paulo, SP, Brazil.

Email: eutimiocu@yahoo.com, egfnunez@usp.br

Funding information

Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: PIBIC-2015/UNESP/33516; Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: FAPESP (2016/19004-4); Brazilian National Council for Scientific and Technological Development

Abstract

The beer quality can be modulated from changes in their ingredient proportions, as well as in operating parameters. The crossed experimental designs and the multiple optimizations based on desirability functions have demonstrated to be effective methodologies in the unit operation polynomial modeling and optimization of bioprocess, respectively. However, artificial intelligence techniques have been used as an alternative to this modeling in bioprocess. Therefore, this study aimed to implement a software combining artificial neural network (ANN) and differential evolution to optimize the topology of an ANN to model the Ale beer production and to use the optimized ANN in ingredients and operation parameters choice that ensure a beer with high acceptance rate, by the genetic algorithm technique for multiple-objective function. This approach allowed to find ANN models which fitted the process with correlation coefficients higher than 0.85 and high satisfaction level of beer desirable quality attributes (global desirability value = 0.78).

Practical Applications

This manuscript could be useful for bioprocess professionals involved in the development of the brewing process and artificial intelligence applications. The approach applied in this work allows for modeling and optimization of brewing process using a combination of crossed experimental design, artificial neural networks, and evolutionary algorithms with relatively low experimental efforts. At the same time, the quality attributes of the beer are better controlled.

1 | INTRODUCTION

One of the most consumed alcoholic beverages worldwide, beer has an important influence on the economy, culture, and society. Globalization impact in countries as Brazil, China, Russia, and India over the last 50 years increased the beer consumption, overtaking countries that were traditionally known for beer production as Germany and UK (Colen & Swinnen, 2016; Mega, Neves, & Andrade, 2011). Each culture has its preferences when choosing beer, like color, bitterness, and foam stability, which can be achieved by using different compositions of substrate and process variables. Thus, rigorous control of the quality parameters is important for product acceptance by consumers and for beer classification as well (Venturini Filho & Cereda, 2001).

The development of new modeling and optimization strategies in the brewing process is necessary to guarantee product quality due to the demand and consumer requirements are increasing. As a bioprocess, there are few efficient mathematical equations to describe the whole process due to the variables interactions and complex biochemistry reactions (Kumar, Bhalla, & Rathore, 2014; Lee & Gilmore, 2006). The application of empiric mathematical models to correlate process parameters and the desired product variables require the efficiency of the modeling methodologies (Coscione, De Andrade, & May, 2005).

Factorial designs and response surface methodology (RSM) are empiric techniques traditionally applied to screen and optimize factors in the industrial world. The utilization of these statistical tools allows for the setting of unit operations in processes using maps of the

 © 2019 Wiley Periodicals, Inc.

response surface, optimization, and selection of operating conditions by fitting a polynomial equation to the experimental data obtained either at bench or pilot scales (Myers, Montgomery, & Anderson-Cook, 2009). Moreover, this technique performs multivariate analysis of the process instead of the common one-variable-at-a-time, which allows the evaluation of the effects among process and mixture variables (Bezerra, Santelli, Oliveira, Villar, & Escaleira, 2008).

Usually, the factorial design modeling is described by linear functions when represented by first-order models. The nonlinear processes should be described by quadratic response surfaces as those derived from three-level factorial and central composite experimental designs (Bezerra et al., 2008). On the other hand, artificial intelligence techniques such as artificial neural networks (ANN) and genetic algorithms (GA) have been also used to model and optimize process with efficiency (Candioti, De Zan, Cámara, & Goicoechea, 2014; Elsayed & Lacor, 2013; Lin, Su, Wang, Chang, & Juang, 2012). However, these tools have been scarcely applied in brewing processes for both purposes.

In general, ANNs have been used successfully in bioprocess due to its nonlinear feature, which can detect complex nonlinear correlations among the dependent (quality attributes) and independent variables (process parameters) (Agatonovic-Kustrin & Beresford, 1999; Baughman & Liu, 1995). Specifically, ANN is a mature technology in food industry and safety, with applications mainly in classification of food species and food quality, the element content detection and risk management (Lin, Cui, Han, Geng, & Zhong, 2019). However, the selection of ANN topology, which depends on many parameters (i.e., number of hidden layers, number of neurons of hidden layers, neuron transfer function, connections among neurons), is a tricky task (Ding, Li, Su, Yu, & Jin, 2013). On the other hand, evolutionary algorithms, a class of stochastic search and optimization techniques, that simulate the natural evolution of biological systems, have been useful to define the ANN topologies (Basheer & Hajmeer, 2000; Ding et al., 2013; Goethals, Dedecker, Gabriels, Lek, & Pauw, 2007). Among the evolutionary algorithms, differential evolution is highlighted in ANN topology optimizations (Ilonen & Kamarainen, 2003; Magoulas, Plagianakos, & Vrahatis, 2004; Subudhi & Jena, 2011).

Another evolutionary algorithm widely known, GA have been applied in several studies of bioprocess optimization from experimental data using ANNs, not mainly to optimize their topologies, if not to deal with optimization of multiple objective functions (Gurunathan & Sahadevan, 2011; Peng et al., 2014; Peng, Meng, & Ai, 2013; Zafar, Kumar, Kumar, & Dhiman, 2012). As a rule, the processes and their derived products have several attributes to describe their performance and quality, respectively. Thus, multiobjective optimization is a usual challenge during the establishment of the commercial processes. The common tool to solve this, from the classical approaches based on RSM, is the use of desirability's functions, which utilized an importance ranking for the attributes under study (He, Zhu, & Park, 2012; Kim, Rhee, & Park, 2002; Ortiz, Simpson, Pignatiello, & Heredia-Langner, 2004; Pasandideh & Niaki, 2006).

Therefore, this study aimed to develop an artificial intelligence software based on differential evolution and ANN techniques for modeling beer's physic-chemical attributes, using the data from an experimental design at the bench scale to ensure better control of the beer quality from brewing critical factors. Moreover, a many-objective optimization approach based on a GA was also developed to define the beer substrates ratio and protease process time for matching a set of beer requirements.

2 | MATERIALS AND METHODS

2.1 | Brewing process at bench scale

Beer production at bench scale was based on Ale-type formula proposed by Palmer (2006), which was previously described (Coelho de Oliveira, Elias da Cunha Filho, Rocha, & Fernández Núñez, 2017). The beer substrate was composed by malted barley (Pilsen malt, Agrágria, Brazil), malted wheat (Light wheat malt, Weyermann, Germany) and high maltose corn powder (Cargill, Brazil) with a total mass of 160 g. Their proportions were defined by an experimental design, which is described in Section 2.2.

The original extract (beer substrate) was mixed with 500 mL of water at 45°C for 5 min to begin the enzymatic activation ramp. Then, the temperature was raised to 52°C for a time specified according to experimental design (see Section 2.2), followed by two additional temperature rise of 68 and 78°C for 60 and 15 min, respectively. At the end of this temperature ramp, the filtration cake was washed with 600 mL of water at 78°C. Then, the filtrate was boiled for 60 min and simultaneously 1.5 g of *Hallertauer Perle* hops (Hopsteiner, Germany) were added. After this step, sterilized water was used to complete 800 mL. The wort was cooled to 25°C for initiating the fermentation stage. The inoculum for brewer's wort was a lyophilized strain of *Saccharomyces cerevisiae* Safbrew WB-06 (Fermentis, Marcq-en-Baroeul, France) rehydrated in 40 mL at 1.6×10^8 cell/mL. The fermentation was performed at 23°C for 120 hr. In the final stage, the beers were maturated at 3°C for 3 days before bottling. The beers were also stored at 3°C for 24 hr until characterization.

2.2 | Experimental design, response surface methodology, and multiobjective optimization

A D-optimal crossed experimental design including mixture and process variables was set and analyzed in Design-Expert software (Trial version, 9.0.4.1, Stat-Ease, Minneapolis). Thirty-one beers were produced to correlate mixture and process variables to beer quality parameters. Five pairs of them correspond to runs performed in duplicate, which were used to estimate the pure or experimental error. The mixture variables were the mass percentages of malted barley (55–100% wt/wt), malted wheat (0–45% wt/wt), and high maltose corn powder (0–45% wt/wt) in the original extract. The process variable under this study was the reaction time (5–25 min) for proteases at 52°C. The number of experimental combinations (31 runs) was sufficient to fit a cubic model for mixture variables and a quadratic model for the process variable (Online Resource 1). The multiobjective optimization was performed in the same software, Design-Expert 9.0.4.1, using the desirability function

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Optimization criterion of the beer quality parameters

Beer quality parameters	Optimization	Importance level
Foam stability	Maximization	Priority
Alcohol content	Maximization	Priority
Bitterness	Minimization	Secondary
Turbidity	Minimization	Secondary
Vicinal diketones	Minimization	Secondary
pH	Minimization	Tertiary
Color	Minimization	Tertiary

implemented on it. The importance level of each beer quality attribute for this optimization approach is detailed in Table 1.

2.3 | Beer characterization

All samples (700 mL) were degassed before analysis by vigorous stirring using a glass rod for 10 min at 20°C. Seven parameters were analyzed under the characterization, all of them are briefly described below. They are included among the most frequent quality attributes used to characterize beers (Briggs, Boulton, Brookes, & Stevens, 2004; Coelho de Oliveira et al., 2017). These parameters are very related to customer perception and palatability.

1. *Foam stability*: it was determined by Rudin's method using nitrogen instead of carbon dioxide (3 L/min) (Rudin, 1957). This parameter was defined as the time in seconds after bubbling stop taken for the boundary between liquid and foam to displace from 5 to 7.5 cm at the bottom of Rudin's cylinder.
2. *Turbidity*: it was measured by the nephelometric method using a portable turbidimeter (Microprocessor Turbidimeter Plus II, Alfakit, Florianópolis, Brazil). Turbidity was expressed in nephelometric turbidity units.
3. *Vicinal diketones*: the official spectrophotometric method defined by the European Brewing Convention was used to quantify vicinal diketones (mg/L) in beer samples (Rodrigues, Barros, Machado Cruz, & Ferreira, 1997).
4. *Color*: The filtered and degassed beer sample was initially diluted fourfold in deionized water. The color in European Brewing Convention (EBC) units was determined taking into account the diluted sample absorbance (A) at 430 nm and dilution factor (f = 4), according to the Equation (1) (Popescu, Soceanu, Dobrinas, & Stanciu, 2013).

$$Color\,(EBC\,units) = A \cdot f \cdot 25 \tag{1}$$

5. *Alcohol content*: One-hundred milliliters of degassed beer were distilled. The condensed vapor was collected in a 100 mL volumetric flask with 10 mL of deionized water. Subsequently, the volumetric flask was completed with deionized water and homogenized. Finally, the density of the hydroalcoholic solution was determined at 20°C by pycnometry and correlated to its alcohol content (% vol/vol) (Zenebon, Pascuet, & Tiglea, 2008).
6. *Bitterness*: The determination of bitterness was carried out according to the accepted methodology by the European Brewery Convention (EBC) (Howard, 1968; Popescu et al., 2013). The values were expressed in (EBC units, EBU).
7. *pH*: The pH value was measured by potentiometry at 28°C (Orion 5-Star Plus, ThermoScientific, Beverly, MA).

2.4 | Artificial neural network topology determination by differential evolution

A differential evolution algorithm, proposed by Storn and Price (1995) was implemented to optimize an ANN topology in MATLAB® (MATLAB R2016a, MathWorks Inc., Natick, MA). The algorithm flowchart is described in Figure 1. The vector which defined the population represented the ANN parameters, including the number of neurons and transfer functions from first, second, and third hidden layers, also the transfer function from the output layer, training function to be used, number of hidden layers and training and momentum rates. There are 11 parameters in total, so that represents the vector's dimension is equal to j = [1, 11].

The ANN was a multilayer perceptron neural network with backpropagation training, this approach is one the most utilized ANN learning algorithms based on the gradient descent, it is specially suitable for low-dimensional data or low-complexity properties (Geng, Shang, Han, & Zhong, 2019). The mixture and process variables from the experimental data were used as input set and each beer quality parameter was representing the output variable, totalizing seven ANNs. This dataset was randomly divided into 70:15:15 for training, validation, and test of the ANN topology (Gebler, Kayzer, Szoszkiewicz, & Budka, 2013; Seguritan et al., 2012). The dataset was also normalized between [0,1] due to the variable magnitude differences.

As the differential evolution algorithm (DE) can vary with its parameters, such as scale factor and number of vectors, several values were combined to determine in which conditions population could converge to a better solution. The values to determine an optimal scale factor were 0.5, 0.8, 1, 1.2, and 2, as it varies from [0,2] (Price, Storn, & Lampinen, 2005). The number of vectors in the population defined was 100, 150, 200, and 300. The crossover rate chose was 5% and the number of generations was 400. The algorithm was tested 4 times to each condition described above.

First, all vectors are randomly created to establish the first population of size NP. The ANNs are created based on the parameters of each vector. Then, the second step of the DE is a mutation that follows the DE/best/1 strategy, where the best response in a population (*best*) and the other two vectors $x_{r1i,\,G}$ and $x_{r2i,\,G}$ are randomly chosen to generate a mutate vector $v_{i,\,G}$. Mutation occurs for i = [1, NP], therefore, until NP mutate vectors have been created. This strategy is described in Equation (2) (Mallipeddi, Suganthan, Panb, & Tasgetirenc, 2011).

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

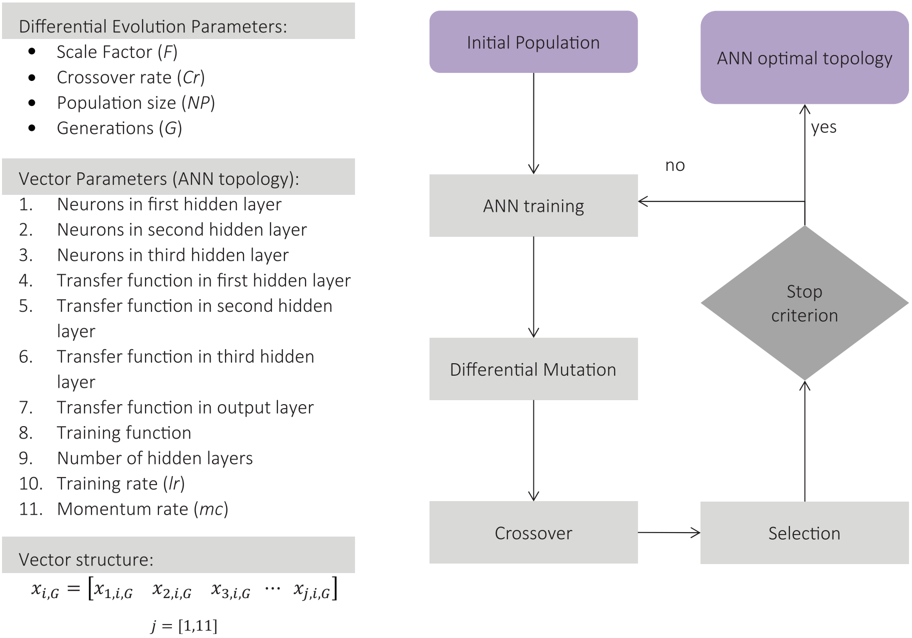

FIGURE 1 Differential evolution algorithm flowchart

$$v_{i,G} = best + F \cdot \left(x_{r1i,G} - x_{r2i,G}\right) \tag{2}$$

in which F is the scale factor.

After mutation, a crossover is applied to population choosing between the parameter from a mutate vector $v_{i,G}$ or from the previous population vector $x_{i,G}$ based on binomia al strategy for a crossover rate (Cr) of 5% (Das & Suganthan, 2011). This step also occurs for i = [1, NP] and is explained in Equation (3). At least one parameter will be changed, accordingly to $j = jrand$.

$$u_{j,i,G} = \begin{cases} v_{j,i,G}, \text{if } (rand\,[0,1] \le Cr) \text{ or } (j = jrand) \\ \qquad\qquad x_{j,i,G}, \text{otherwise.} \end{cases} \tag{3}$$

Selection step is carried out by comparing the performance of the previous population vectors to the current ones [Equation (4)]. The parameter selected to analyze the vectors was the linear regression coefficient R.

$$x_{i,G+1} = \begin{cases} x_{i,G}, \text{if the performance } of\ x_{i,G} > u_{i,G}; \\ \qquad\qquad u_{i,G}, \text{otherwise.} \end{cases} \tag{4}$$

A new population of vectors $x_{i,G+1}$ is then created, following the next generation G, which continues the algorithm. The stop criterion chose was the number of generations. At the end of the algorithm run, it shows the best vector from the population, therefore, the best ANN topology optimized.

2.5 | Beer many-objective optimization by genetic algorithm

Multiobjective optimization is one of the most active research areas in evolutionary computation. It includes problems with at least two objective problems. However, multiobjective problems with four or more objectives are often referred to as many-objective problems. The last ones are very complex to be solved efficiently with classical algorithms, that work for problems with a small number of objectives (Ishibuchi, Tsukamoto, & Nojima, 2008). As there are seven ANN architectures representing the modeling of each beer quality parameter, a many-objective optimization by GA was developed in MATLAB platform. Each ANN represented an objective function. The optimization followed a criterion described in Table 1. The level of importance for each variable was defined according to the authors' criteria about the Brazilian beer customer's perception and preference.

The variables involved in GA were number of individuals (tested for 50, 100, 200, and 300), number of generations (50, 100, and 150), mutation rate (5% of rate), crossover rate (15% of individuals created from crossover), fitness (20% selected as the best individuals), and migration (20% new individuals created and replaced into the population).

The GA began with the creation of individuals from the initial population randomly. In this algorithm, the individuals represented the mixture and process variables of beer production. Then, the selection is carried out by elitism for 20% of the population, in which the fitness function considered was the desirability function, commonly used in RSM (He et al., 2012; Ortiz et al., 2004; Pasandideh & Niaki, 2006). The desirability function for each response is described in Equation (5), which was used to calculate a global desirability value D [Equation (6)] (Myers et al., 2009).

$$d_i(\hat{y}_i) = \begin{cases} \dfrac{\hat{y}_i(x) - L_i}{U_i - L_i}, \text{to maximize;} \\ \dfrac{\hat{y}_i(x) - U_i}{L_i - U_i}, \text{to minimize.} \end{cases} \tag{5}$$

in which L_i is the lower bound of each parameter; U_i is the upper bound and $\hat{y}_i(x)$ is the target response from the objective function.

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

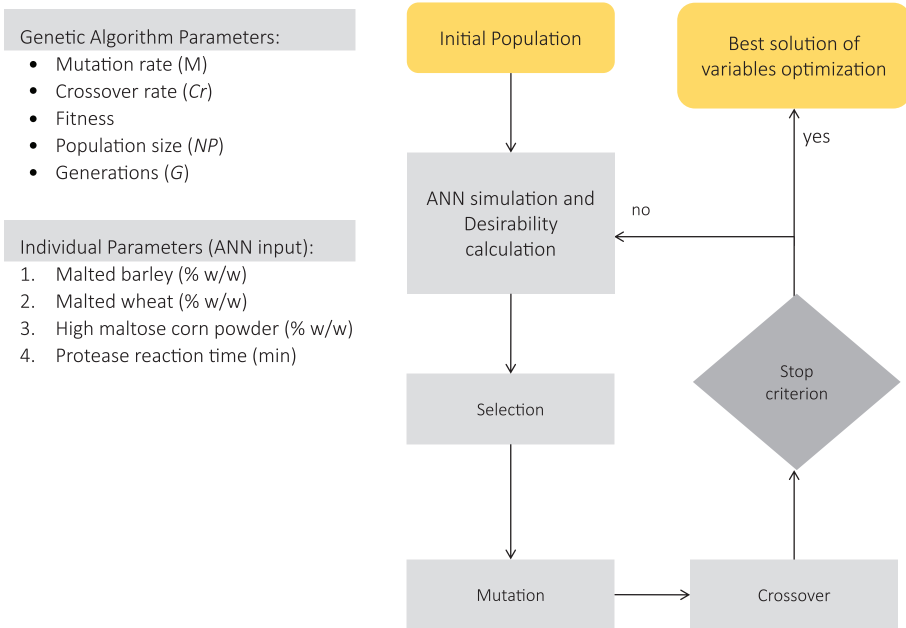

FIGURE 2 Many-objective genetic algorithm flowchart

$$D=\left((d_1)^{w_1}(d_2)^{w_2}\dots(d_p)^{w_p}\right)^{\frac{1}{\sum w_i}} \tag{6}$$

in which w_p is the weight of each parameter according to the Table 1; priority received w_p = 5, secondary w_p = 4 and tertiary w_p = 3.

The algorithm next step is mutation, which followed the binomial method, changing the individual's variables randomly if a value lower than the mutation rate (5%) was obtained. After mutation, 1-point crossover is carried out in 15% of the population, choosing two of the best individuals randomly. A new population is then generated, continuing running the algorithm. The stop criterion was verified based on the maximum number of generations. When the number of generations achieved 50%, there was a migration step, in which 20% of the population was replaced by new random individuals. The overall steps are described in Figure 2.

3 | RESULTS AND DISCUSSION

3.1 | Beer characterization

The primary data utilized in this work, related to beer samples characterization, were previously reported by senior authors in a complementary manuscript dealing with brewery process monitoring (Coelho de Oliveira et al., 2017). The ranges identified in quality parameters of the beer samples were typical for Ale-type beers (Briggs et al., 2004). Besides, the ranges for all beer quality parameters were broad enough to ensure calibration of RSM and ANN models. The variability of quality attributes in the data set was a consequence of changes in beer ingredients and protease reaction time derived from the experimental design (Online Resource 1) (Coelho de Oliveira et al., 2017).

3.2 | Response surface methodology modeling

The models with statistical significance are described in Table 2. The item "Adjusted model" describes the polynomial order fitted to the mixture and process variables. From this modeling, it is possible to identify the influence of variables, as example color, in which the quadratic model was significant to mixture variables; however, there is no influence from the process variable. The numeric values represent the fitted polynomial coefficients, such as in alcohol content modeling [Equation (7)].

$$Y=0.035\cdot A+0.034\cdot B+0.06\cdot C \tag{7}$$

The response surface of each beer quality parameter is shown in Figure 3. Bitterness has not been influenced by any of the variables and, therefore, there is no response surface. This was caused mainly due to hop content, as it was constant in all beer samples. The mean value of bitterness was next to 13.77 IBU. The alcohol content was influenced by maltose in the majority and equal parts of malted barley and wheat (Figure 3a). Process time did not influence this parameter. The alcohol content varied from 3 to 5% vol/vol.

The pH profile was increased by malted barley and wheat, respectively, and decreased by maltose. Process time does not imply in pH changes, however, it was modeled due to the linear relationship between the beer components (Figure 3b).

Beer color has not been influenced by process time as well, and malted barley configured higher color intensity. When high maltose corn powder was absent, malted wheat ratios of 0–20% decreased color intensity (Figure 3c). Above this range, the different proportions of malted cereals did not influence color intensity. When maltose is 20 and 40% (Figure 3c), color intensity is reduced, and malted cereals combination lose their effect. The range of color (5.5–12 EBC) characterizes a light beer type.

Vicinal diketones were dependent only of malted barley even when combined with maximum malted wheat value. The lowest values were defined by 77.5% of malted barley and 22.5% of malted wheat. Maltose reduced the desirable effect of malted cereals interaction, increasing the vicinal diketones concentration, such as when 40% proportion it was 3.5 times higher than the minimum value identified (Figure 3d).

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Fitted polynomial models for beer parameters based on mixture and process variables

	Beer's quality attributes						
Factor types	**Bitterness**	**Alcohol content**	**pH**	**Color**	**Vicinal diketones[a]**	**Foam stability**	**Turbidity**
Mixture	M	L	L	Q	SC	M	L
Process	M	M	M	M	M	C	C
Model terms	Values of the significant terms of the adjusted models						
Intercept	13.772	-	-	-	-	−51.537	-
A	-	0.035	0.041	0.097	−0.011	-	0.134
B	-	0.034	0.038	0.223	0.056	-	−0.042
C	-	0.06	0.03	0.025	−0.015	-	−1.399
D	-	-	-	-	-	44.822	-
D^2	-	-	-	-	-	−3.262	-
D^3	-	-	-	-	-	0.068	-
AB		-	-	−0.003	−0.001	-	-
CD	-	-	-	-	-	-	0.472
CD^2	-	-	-	-	-	-	−0.035
CD^3	-	-	-	-	-	-	0.001
Statistical parameter	Values of the statistical parameters						
p-value	-	.007	<.0001	.002	.039	.027	.004
R^2	-	.299	.757	.425	.252	.283	.478
R	-	.547	.87	.652	.502	.532	.691
MSE	4.186	0.288	0.006	0.95	0.002	443.943	23.957
Fit loss	-	0.634	0.918	0.128	0.546	0.762	0.573
R^2 fitted	-	.249	.74	.361	.193	.204	.373
R^2 predicted	-	.172	.713	.138	.08	.076	.253

Note: A: malted barley; B: malted wheat; C: high maltose corn powder; D: protease time process; M: mean; L: linear; Q: quadratic; SC: special cubic and C: cubic.
Abbreviation: MSE, mean squared error.
[a]For a better modeling fit, the data were transformed with Log_{10} function.

A cubic model for the process variable was fitted, which showed that the longer reaction time, the lower foam stability was found (Figure 3e). Reaction time above 10 min decreased foam stability considerably.

Turbidity was the most complex parameter, due to the linear correlation of mixture variables and the cubic relationship of process time. Maltose increase influenced turbidity for a processing time of 10 min. Above 20 min, turbidity was less impact (Figure 3f).

3.3 | Artificial neural network topology optimized by differential evolution

The optimal ANN topologies for each beer quality parameter and the DE initial conditions that provided the best solutions are described in Table 3. The scale factor ranging 0.8 to 1.2 and around 300 vectors in the population provided better convergence of the population for an optimal solution. The classical method of the DE suggests that the *NP* value should be 5–10 times higher than the vector dimension (Das & Suganthan, 2011) and trial value of $F = 0.5$ (Price et al., 2005). In this study, the vector dimension was 11 and the number of vectors of the population was higher than 10-fold vector dimension. However, the software analyzed a broad range of each parameter, such as the number of neurons (1–100). Therefore, higher values of these parameters were required to converge an optimal response.

All the optimal neural architectures presented coefficients of correlation (*R*) > .85 when comparing all predicted and observed values, so then the method performed a good calibration of the process (Table 3). Besides, the values of mean squared error were lower than 0.1, a low value, which also validates the method. The algorithm also validates the ANN architecture in terms of data, when randomizing the dataset division in training, validation, and test, as in cross-validation methodologies (Xu, Liang, & Du, 2004).

Previous works using similar approaches (ANN-DE, ANN-GA) to optimize ANN topologies for modeling quality attributes, applied to other bioprocesses (Gurunathan & Sahadevan, 2011; Satya Eswari & Venkateswarlu, 2016), reported values of *R* (.89–.99) in the same range obtained in this application to the brewery (Table 3). The results derived from the integration of both artificial intelligence tools confirmed the power of this combination to make ANNs have more excellent performance (Ding et al., 2013). Likely, this current work is the first application of DE to optimize the ANN topology for modeling

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

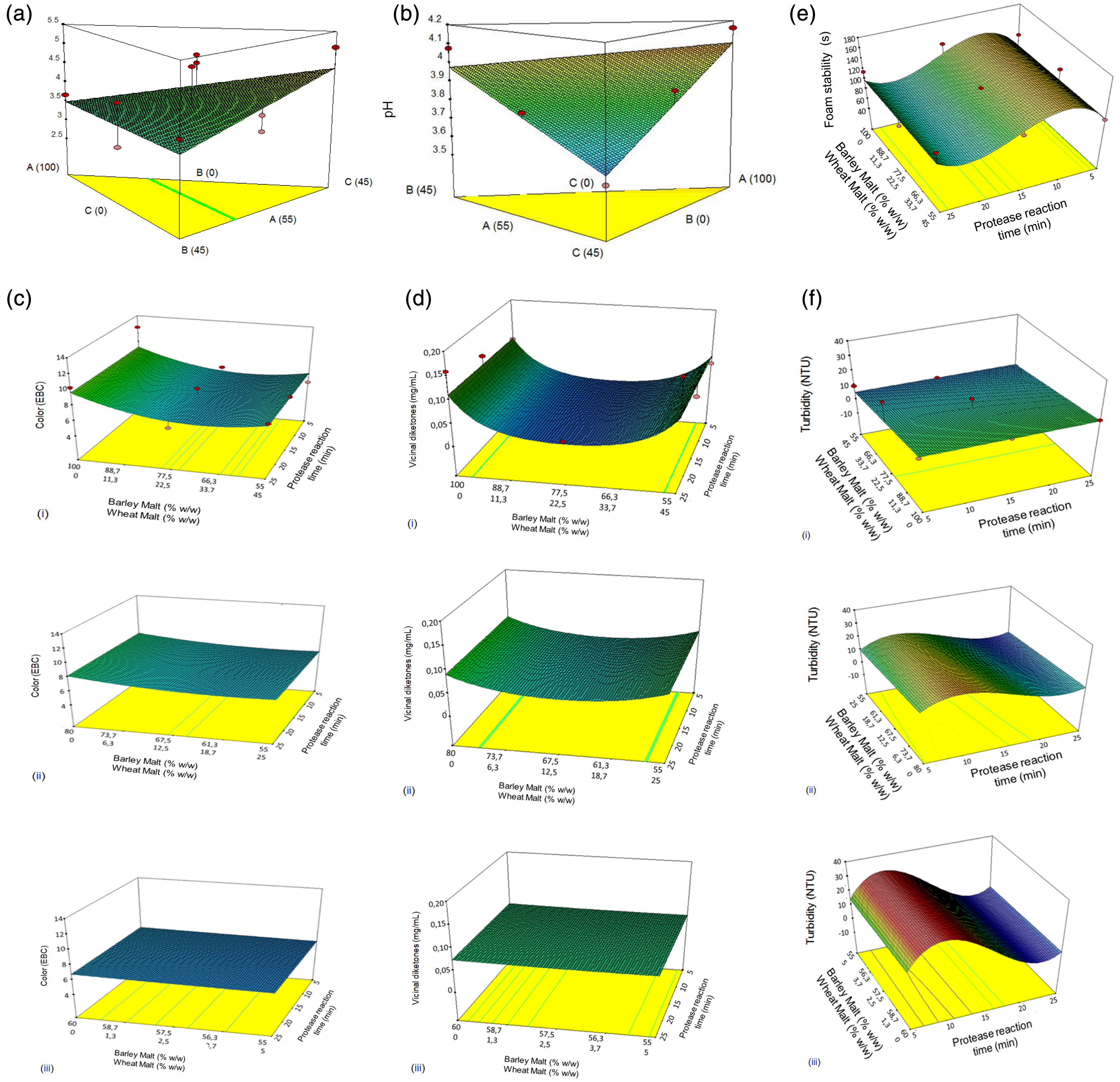

FIGURE 3 Response surface of the fitted models for beer quality parameters. (a) Alcohol content, (b) pH, (c) color, (d) vicinal diketones, (e) foam stability, and (f) turbidity. The figures represented by three graphs (i, ii, iii) are correspondent to (a) 0%, (b) 20%, and (c) 40% of high maltose corn powder respectively

brewery process. The approach allows defining robustly the ANN topology, which is able to make an accurate prediction of beer attributes as a function of process parameters.

3.4 | Beer parameter optimization

Both RSM and ANNs—GA approach—under this study used desirability function as fitness for determining the best solution based on the optimization criterion (Table 1). The GA performance was effective with 200 individuals and 50 generations.

According to the desirability values, the ANN-GA method (0.78) was more effective than classical polynomial approach (0.55) to optimize the beer process due to the ANN nonlinear modeling and the GA optimization strategy (Table 4). Desirability values closer to 1 define a more desirable response according to how well were matched the set of the optimization criteria for each variable (Table 1) (Salmasnia, Baradaran Kazemzadeh, & Niaki, 2012). Similar results were found in literature when comparing these methodologies (Elsayed & Lacor, 2013; Lin et al., 2012) and also several studies applied GA and desirability approach with success (He et al., 2012; Kim et al., 2002; Pasandideh & Niaki, 2006). The

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Description of the optimal artificial neural network topologies and differential evolution parameters

Beer quality parameters	ANN topology: Number of neurons in each hidden layer	Transfer function in each layer	Training function	Learning rate (lr)	Momentum rate (me)	ANN evaluation: *R* (total)	*R* (test)	MSE (test)	Differential evolution parameters: *NP*	*F*	*G** convergence
Alcohol content	44 60 9	*hardlim-hardlim-tansig-tansig*	*trainscg*	-	-	.91	.93	0.027	300	1.2	372
Bitterness	97	*radbas-tansig*	*trainscg*	-	-	.86	.97	0.013	300	0.8	161
Color	76 33	*purelin-radbas-tansig*	*traingdx*	0.01	0.45	.97	.92	0.003	300	1.2	93
Vicinal diketones	55 53 55	*tansig-radbas-tansig-tansig*	*traingda*	0.001	-	.92	.92	0.012	150	1	225
Foam stability	20 15 69	*radbas-radbas-tansig-tansig*	*trainscg*	-	-	.86	.98	0.016	200	1	62
pH	100 93 38	*radbas-tansig-tansig-tansig*	*trainrp*	0.001	-	.95	.94	0.005	300	1.2	120
Turbidity	100 91 100	*tansig-tansig-radbas-tansig*	*traingdm*	0.01	0.45	.92	.95	0.006	300	1.2	315

Notes: *trainscg*: Scaled conjugate gradient backpropagation. *traingdx*: Gradient descent with momentum and adaptive learning rate backpropagation, *traingda*: Gradient descent with adaptive learning rate backpropagation, *trainrp*: Resilient backpropagation, *traingdm*: Gradient descent with momentum backpropagation. The transfer functions in each layer are defined detailed in https://la.mathworks.com/help/#.
*G convergence means the generation number in which convergence was achieved, that is, the best solution was found by the algorithm and this solution remains the best among the population until the algorithm is stopped.
Abbreviations: ANN, artificial neural network; MSE, mean squared error.

TABLE 4 Comparison of the methods to optimize quality parameters of beer

		Polynomial model	Neural network-genetic algorithm
Desirability		0.55	0.78
Solution	Malted barley (% wt/wt)	55.0	64.04
	Malted wheat (% wt/wt)	-	19.45
	High maltose corn powder (% wt/wt)	45.0	16.58
	Protease reaction time (min)	16.0	18.15
Optimized parameters	Alcohol content (% vol/vol)	4.60	4.56
	Foam stability (s)	107.00	142.70
	Bitterness (IBU)	13.77	2.90
	Turbidity (NTU)	14.83	8.68
	Vicinal diketones (mg/L)	0.07	0.03
	pH	3.58	3.71
	Color (EBC)	6.45	7.56

Abbreviations: EBC, European Brewing Convention; NTU, nephelometric turbidity unit.

best solution of ANN-GA, with the addition of 19.45% of malted wheat and considerably less content of high maltose corn powder, satisfies better the desired beer characteristics comparing to the polynomial approach. ANN-GA method was able to hunt up a solution which has similar content of alcohol, and increasing significantly foam stability and decreasing bitterness, turbidity, and vicinal diketones, as well.

Therefore, this work could find application in the current brewery industry momentum. Beer is the most consumed alcoholic drink globally (Colen & Swinnen, 2016). Specifically, in Brazil, beer consumption has been increasing and new formulations for beer production have been included in the market. Just in 2018, there was an increase of 30% in the registration of new brewery factories in Brazil, totalizing 889 companies (Marcusso & Müller, 2018). Within this favorable and competitive framework, the method from this study allowed with a relatively small amount of experiments, the development of a tool that can optimize the parameters of a beer based on desired criteria. Thus, this approach can be used for companies that want to create new types of beer based on customers' profiles.

4 | CONCLUSIONS

Brewing process modeling and optimization using the ANN and evolutionary algorithm approach showed to be an efficient and accurate tool to reduce experimental costs over the product development process and guarantee an Ale-type beer with suitable quality standards. Comparing to the traditional method used in industry (RSM), the

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

fitness quality of process models based on ANN and evolutionary algorithms were higher (correlation coefficients > .85) and superior satisfaction level of desirable quality attributes (global desirability value = 0.78) was also achieved. The current methodology could be applied either to speed up new beer commercialization or to develop new ones in a low-level investment environment.

ACKNOWLEDGMENTS

This work was funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP: 2016/19004-4) and the Brazilian National Council for Scientific and Technological Development (CNPq/Brazil: PIBIC-2015/UNESP/33516).

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Eutimio Gustavo Fernández Núñez https://orcid.org/0000-0002-2800-392X

REFERENCES

Agatonovic-Kustrin, S., & Beresford, R. (1999). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, *22*(2000), 717–727.

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods*, *43*(1), 3–31. https://doi.org/10.1016/S0167-7012(00)00201-3

Baughman, D. R., & Liu, Y. A. (1995). *Neural networks in bioprocessing and chemical engineering*. San Diego, CA: Elsevier.

Bezerra, M. A., Santelli, R. E., Oliveira, E. P., Villar, L. S., & Escaleira, L. A. (2008). Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta*, *76*(5), 965–977. https://doi.org/10.1016/j.talanta.2008.05.019

Briggs, D. E., Boulton, C. A., Brookes, P. A., & Stevens, R. (2004). *Brewing: Science and practice*. Boca Raton, FL: CRC Press.

Candioti, L. V., 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.

Coelho de Oliveira, H., Elias da Cunha Filho, J. C., Rocha, J. C., & Fernández Núñez, E. G. (2017). Rapid monitoring of beer quality attributes based on UV-Vis spectral data. *International Journal of Food Properties*, *20*(S2), S1686–S1699. https://doi.org/10.1080/10942912.2017.1352602

Colen, L., & Swinnen, J. (2016). Economic growth, globalisation and beer consumption. *Journal of Agricultural Economics*, *67*(1), 186–207. https://doi.org/10.1111/1477-9552.12128

Coscione, A. R., De Andrade, J. C., & May, G. M. (2005). O modelamento estatístico de misturas: Experimento tutorial usando voltametria de redissolução anódica. *Química Nova*, *28*(6), 1116–1122.

Das, S., & Suganthan, P. N. (2011). Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation*, *15*(1), 4–31. https://doi.org/10.1109/TEVC.2010.2059031

Ding, S., Li, H., Su, C., Yu, J., & Jin, F. (2013). Evolutionary artificial neural networks: A review. *Artificial Intelligence Review*, *39*(3), 251–260. https://doi.org/10.1007/s10462-011-9270-6

Elsayed, K., & Lacor, C. (2013). CFD modeling and multi-objective optimization of cyclone geometry using desirability function, artificial neural networks and genetic algorithms. *Applied Mathematical Modelling*, *37*(8), 5680–5704. https://doi.org/10.1016/j.apm.2012.11.010

Gebler, D., Kayzer, D., Szoszkiewicz, K., & Budka, A. (2013). Artificial neural network modelling of macrophyte indices based on physico-chemical characteristics of water. *Hydrobiologia*, *737*, 1–10. https://doi.org/10.1007/s10750-013-1585-7

Geng, Z., Shang, D., Han, Y., & Zhong, Y. (2019). Early warning modeling and analysis based on a deep radial basis function neural network integrating an analytic hierarchy process: A case study for food safety. *Food Control*, *96*, 329–342. https://doi.org/10.1016/j.foodcont.2018.09.027

Goethals, P. L. M., Dedecker, A. P., Gabriels, W., Lek, S., & Pauw, N. D. (2007). Applications of artificial neural networks predicting macroinvertebrates in freshwaters. *Aquatic Ecology*, *41*(3), 491–508.

Gurunathan, B., & Sahadevan, R. (2011). Design of experiments and artificial neural network linked genetic algorithm for modeling and optimization of L-asparaginase production by Aspergillus terreus MTCC 1782. *Biotechnology and Bioprocess Engineering*, *16*(1), 50–58. https://doi.org/10.1007/s12257-010-0119-7

He, Z., Zhu, P., & Park, S. (2012). A robust desirability function method for multi-response surface optimization considering model uncertainty. *European Journal of Operational Research*, *221*(1), 241–247. https://doi.org/10.1016/j.ejor.2012.03.009

Howard, G. A. (1968). Estimation of the bitterness of beer. *Journal of the Institute of Brewing*, *74*, 249–251. https://doi.org/10.1002/j.2050-0416.1968.tb03121.x

Ilonen, J., & Kamarainen, J. (2003). Differential evolution training algorithm for feed-forward neural networks. *Neural Processing Letters*, *17*, 93–105.

Ishibuchi, H., Tsukamoto, N., & Nojima, Y. (2008). *Evolutionary many-objective optimization: A short review*. Hong Kong: IEE Congress on Evolutionary Computation, 2419–2426.

Kim, D., Rhee, S., & Park, H. (2002). Modelling and optimization of a GMA welding process by genetic algorithm and response surface methodology. *International Journal of Production Research*, *40*(7), 1699–1711. https://doi.org/10.1080/00207540110119964

Kumar, V., Bhalla, A., & Rathore, A. S. (2014). Design of experiments applications in bioprocessing: Concepts and approach. *Biotechnology Progress*, *30*(1), 86–99. https://doi.org/10.1002/btpr.1821

Lee, K.-M., & Gilmore, D. F. (2006). Statistical experimental design for bioprocess modeling and optimization analysis. *Applied Biochemistry and Biotechnology*, *135*, 101–115.

Lin, H. C., Su, C. T., Wang, C. C., Chang, B. H., & Juang, R. C. (2012). Parameter optimization of continuous sputtering process based on Taguchi methods, neural networks, desirability function, and genetic algorithms. *Expert Systems with Applications*, *39*(17), 12918–12925. https://doi.org/10.1016/j.eswa.2012.05.032

Lin, X., Cui, S., Han, Y., Geng, Z., & Zhong, Y. (2019). An improved ISM method based on GRA for hierarchical analyzing the influencing factors of food safety. *Food Control*, *99*, 48–56. https://doi.org/10.1016/j.foodcont.2018.12.020

Magoulas, G. D., Plagianakos, V. P., & Vrahatis, M. N. (2004). Neural network-based colonoscopic diagnosis using on-line learning and differential evolution. *Applied Soft Computing*, *4*, 369–379. https://doi.org/10.1016/j.asoc.2004.01.005

Mallipeddi, R., Suganthan, P. N., Panb, Q. K., & Tasgetirenc, M. F. (2011). Differential evolution algorithm with ensemble of parameters and mutation and crossover strategies. *Applied Soft Computing*, *11*, 1679–1696. https://doi.org/10.1007/978-3-642-17563-3_9

Marcusso, E. F., & Müller, C. V. (2018). *Ministério da Agricultura, Pecuária e Abastecimento*. Retrieved from http://www.agricultura.gov.br/

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

assuntos/inspecao/produtos-vegetal/pasta-publicacoes-DIPOV/anuario-da-cerveja-no-brasil-2018/view

Mega, J. F., Neves, E., & Andrade, C. J. (2011). A produção da cerveja no brasil. *Revista Citino*, *1*(1), 34–42.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2009). *Response surface methodology: Process and product optimization using designed experiments* (3rd ed.). New York, NY: John Wiley & Sons.

Ortiz, F. J., Simpson, J. R., Pignatiello, J. J. J., & Heredia-Langner, A. (2004). A genetic algorithm approach to multiple-response optimization. *Journal of Quality Technology*, *36*(4), 432–450.

Palmer, J. J. (2006). *How to brew: Everything you need to know to brew beer right the first time*. Boulder, CO: Brewers Publications.

Pasandideh, S. H. R., & Niaki, S. T. A. (2006). Multi-response simulation optimization using genetic algorithm within desirability function framework. *Applied Mathematics and Computation*, *175*, 366–382. https://doi.org/10.1016/j.amc.2005.07.023

Peng, J., Meng, F., & Ai, Y. (2013). Time-dependent fermentation control strategies for enhancing synthesis of marine bacteriocin 1701 using artificial neural network and genetic algorithm. *Bioresource Technology*, *138*, 345–352. https://doi.org/10.1016/j.biortech.2013.03.194

Peng, W., Zhong, J., Yang, J., Ren, Y., Xu, T., Xiao, S., … Tan, H. (2014). The artificial neural network approach based on uniform design to optimize the fed-batch fermentation condition: Application to the production of iturin A. *Microbial Cell Factories*, *13*(1), 54. https://doi.org/10.1186/1475-2859-13-54

Popescu, V., Soceanu, A., Dobrinas, S., & Stanciu, G. (2013). A study of beer bitterness loss during the various stages of the Romanian beer production process. *Journal of Institute Brewing*, *119*, 111–115.

Price, K. V., Storn, R. M., & Lampinen, J. A. (2005). *Differential evolution: A practical approach to global optimization*. Berlin, Germany: Springer Berlin Heidelberg.

Rodrigues, J. A., Barros, A. A., Machado Cruz, J. M., & Ferreira, A. A. (1997). Determination of diacetyl in beer using differential-pulse polarography. *Journal of the Institute of Brewing*, *103*, 311–314.

Rudin, A. O. (1957). Measurement of the foam stability of beers. *Journal of the Institute of Brewing*, *63*, 506–509.

Salmasnia, A., Baradaran Kazemzadeh, R., & Niaki, S. T. A. (2012). An approach to optimize correlated multiple responses using principal component analysis and desirability function. *The International Journal of Advanced Manufacturing Technology*, *62*(5–8), 835–846. https://doi.org/10.1007/s00170-011-3824-2

Satya Eswari, J., & Venkateswarlu, C. (2016). Multiobjective simultaneous optimization of biosurfactant process medium by integrating differential evolution with artificial neural networks. *Indian Journal of Chemical Technology*, *23*(5), 335–344.

Seguritan, V., Alves, N., Jr., Arnoult, M., Raymond, A., Lorimer, D., Burgin, A. B., Jr., … Segall, A. M. (2012). Artificial neural networks trained to detect viral and phage structural proteins. *PLoS Computational Biology*, *8*(8), e1002657. https://doi.org/10.1371/journal.pcbi.1002657

Storn, R., & Price, K. (1995). *Differential evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces* (Vol. 3). Berkeley, CA: ICSI.

Subudhi, B., & Jena, D. (2011). A differential evolution based neural network approach to nonlinear system identification. *Applied Soft Computing*, *11*, 861–871. https://doi.org/10.1016/j.asoc.2010.01.006

Venturini Filho, W. G., & Cereda, M. P. (2001). Cerveja. In E. Aquarone (Ed.), *Biotecnologia industrial: Biotecnologia na produção de alimentos* (pp. 91–139). São Paulo, SP: Ed. Edgard Blucher.

Xu, Q. S., Liang, Y. Z., & Du, Y. P. (2004). Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *Journal of Chemometrics*, *18*(2), 112–120. https://doi.org/10.1002/cem.858

Zafar, M., Kumar, S., Kumar, S., & Dhiman, A. K. (2012). Bioresource technology artificial intelligence based modeling and optimization of poly (3-hydroxybutyrate- co-3-hydroxyvalerate) production process by using Azohydromonas lata MTCC 2311 from cane molasses supplemented with volatile fatty acids : A genet. *Bioresource Technology*, *104*, 631–641. https://doi.org/10.1016/j.biortech.2011.10.024

Zenebon, O., Pascuet, N. S., & Tiglea, P. (2008). Métodos físico-químicos Para análise de alimentos.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Takahashi MB, Coelho de Oliveira H, Fernández Núñez EG, Rocha JC. Brewing process optimization by artificial neural network and evolutionary algorithm approach. *J Food Process Eng*. 2019;42:e13103. https://doi.org/10.1111/jfpe.13103

17454530, 2019, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jfpe.13103 by Fraunhofer IPK, Wiley Online Library on [28/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License