RESEARCH PAPER

Assessment of artificial neural networks to predict red colorant production by *Talaromyces amestolkiae*

Bianca Dalbem dos Reis[1] · Fernanda de Oliveira[1,3] · Valéria C. Santos-Ebinuma[1] · Érica Regina Filletti[2] · Álvaro de Baptista Neto[1]

Abstract

Consumer choice is typically influenced by color, leading to an increase in the use of artificial colorants by industry. However, several artificial colorants have been banned due to their harmful effects on human health and the environment, leading to increased interest in colorants from natural sources. Natural colorants can be found in plants, insects, and microorganisms. The importance of evaluating the technical and cost feasibility for the production of natural colorants are important factors for the replacement of artificial counterpart. Therefore, it is highly beneficial to predict the productivity of microbial colorants. The use of statistical methods that generate polynomial models through multiple regressions can provide information of interest about a bioprocess. However, modeling and control of biological processes require complex systems models, because they are nonlinear and non-deterministic systems. In this regard, artificial neural networks are suitable for estimating bioprocess variables with systems modeling. In this work, two different strategies were developed to predict the production of red colorants by *Talaromyces amestolkiae*, namely simulation by artificial neural networks (ANN) and response surface methodology (RSM). The results showed that the colorant concentration predicted by ANN is closer to the experimental data than that predicted by polynomial models fitted by multiple regression. Thus, this work suggests that the use of ANN can identify the initial conditions of the culture parameters that have the greatest influence on colorant production and can be a tool to be employed to improve the production of biotechnological products, such as microbial colorants.

Keywords Natural colorants · Artificial neural networks · Filamentous fungi · *Talaromyces amestolkiae*

Introduction

The choice of consumers is influenced by color, especially in food, leading to an increase in the use of artificial colorants by the food industry in the last century [1]. However, several artificial colorants have been banned in industrial products due to their harmful effects on human health [2, 3] and the environment [4]. Thus, there has been an increased interest in colorants from natural sources, which in most cases can be considered safe color additives [5]. According to information from Mordor Intelligence, the market for natural colorants in 2020 was \$1,625.79 million [6].

The sources of natural colorants are the most diverse, including plants, insects, and microorganisms, such as bacteria and filamentous fungi [1, 7]. Among filamentous fungi, the genus *Monascus* is the most reported to produce natural colorants [1, 8]. This genus can produce about 60 secondary metabolites, of which six are the most studied: rubropunctamine and monascorubramine (red); rubropunctamine and monascorubrine (orange); monascine and ankaflavine (yellow) [9]. However, the simultaneous co-production of mycotoxin by *Monascus sp.* is considered a major concern in food, making toxin detection decisive for consumer safety [10, 11]. Thus, one of the strategies proposed to produce

✉ Álvaro de Baptista Neto
alvaro.baptista@unesp.br

1 Department of Engineering Bioprocess and Biotechnology, School of Pharmaceutical Sciences, Universidade Estadual Paulista, Rodovia Araraquara- Jau Km. 01, Araraquara, SP 14801-902, Brazil

2 Department of Engineering, Physics and Mathematics, Institute of Chemistry, Universidade Estadual Paulista, Rodovia Araraquara- Jau Km. 01, SP 14800-903 Araraquara, Brazil

3 Department of Biotechnology, Lorena School of Engineering, University of São Paulo, Lorena, SP 12602-810, Brazil

mycotoxin-free colorants would be the use of fungi of the genus *Talaromyes/Penicillium* [12, 13].

Recently, Oliveira et al. [14] reported that the species *Talaromyces amestolkiae* can secrete five azafilones compounds, all being complexes of the colorants amino-hexanedioic acid and *Monascus* azafilone and being a great alternative for the production of mycotoxin-free colorants. In another work by Oliveira et al. [15], the authors described the importance of evaluating the technical feasibility and cost for the production of natural colorants as important factors for the replacement of artificial colorants. Like other biocompounds, colorant production and microorganism growth are affected by environmental conditions, nutrient sources, and microbiological conditions [1]. Additional supplements such as nitrogen or carbon sources can also enhance growth and colorant production by microorganisms [16]. Therefore, it would be highly beneficial to predict the productivity of microbial colorants.

To predict the growth and bioactive product accumulation of fungi cultivation systems, mathematical modeling using several optimization techniques has been used as a tool [17–20]. The simulation of biochemical processes is a technique that can provide important information about the bioprocess. For example, the use of phenomenological kinetic models is well-reported and described for the most appropriate purposes, mainly aiming to better understand the bioprocesses, optimize the process conditions and evaluate the operation mode. The work by Manan et al. [17] fitted an unstructured model based on logistic and Leudeking–Piret equations to the experimental data from batch cultivation. The authors reported a kinetic model of the process, which was suitable to describe growth, substrate consumption and red colorant production by *M. purpureus*.

In addition to the phenomenological models, the use of empirical models can also provide valuable information, as the influence of conditions process and medium composition about production of Interest metabolites by microorganisms, for the development and optimization of bioprocesses [18]. In this context, the use of statistical methods that generate polynomial models through multiple regression can also provide valuable information about a given bioprocess and it is a tool very explored in the literature to improve the production of biomolecules.

Santos-Ebinuma et al. [19] used a statistical approach though factorial design to evaluate the influence of sucrose and yeast extract, carbon and nitrogen source, respectively, in the production of red colorants by *Penicillium purpurogenum*. Oliveira et al. [15] evaluated the influence of glucose (carbon source) and monosodium glutamate (nitrogen source) in the culture medium and the initial pH of the process, using statistical analysis techniques to adjust the model and to study the influence of process variables on red colorant production by *T. amestolkiae*. Zhou et al. [21]

used response surface methodology to optimize the culture medium for yellow colorants production by *M. anka*. Although Response Surface Methodology (RSM) has gained much importance in optimizing colorant production, the biological processes modeling and control require complex system models as they are non-linear and non-deterministic systems [22, 23].

In this sense, Artificial Neural Networks (ANN) are adequate to estimate biological process variables with system modeling, thus estimating different operation modes of the process through its generalization capacity. ANN has the potential to be used in predicting the output variables of a bioprocess and can also provide adjustment points to improve the performance of the process. ANN mimics how a biological brain works with a network of neurons processing information and signals from neighboring neurons [23]. The MATLAB® software is suitable for modeling these systems, which have excellent generalization and data prediction capabilities [24].

ANN has been recognized as a computational tool to model non-linear relationships and it can be used for predictive modeling-related microbial-based colorant products and biomass growth [22, 25, 26]. Information retrieved from the ANN was used to determine the optimal operating conditions for colorant production by *M. purpureus* using bug damaged wheat meal. The developed ANN had R^2 values for training, validation, and testing data sets of 0.993, 0.961, and 0.944, respectively. According to the model, the authors obtained the highest colorant production adjusting the temperature and the agitation speed under light conditions [26]. Singh et al. [25] have investigated the application of ANN in modeling a red colorant production process by *M. purpureus* and reported that ANN model can be used to predict the effects of cultivation parameters on red colorant production with a high correlation. To the best of our knowledge, no study has been reported so far on the production of red colorant from *Penicillium/Talaromyces* species using ANN.

In this sense, this work aims to develop two distinct strategies to predict the red colorants production by *T. amestolkiae*. The response surface methodology was employed to compare the fit obtained using a polynomial model from multiple regression with the fit obtained using the neural network technique. The models were developed by employing experimental data obtained from previous research concerned with studying the effect of cultivation components concentration and pH on the red colorants production by *T. amestolkiae* [15]. The ANN models thus developed would allow the identification of cultivation parameters which have a higher influence on colorant accumulation. Furthermore, the model with a better fit to the experimental data can be used to optimize the cultivation parameters.

Materials and methods

The experimental data used in this study were obtained by Oliveira et al. [15]. The data from the production of red colorant by *T. amestolkiae* were used for the adjustment of the polynomial model obtained by multiple regression and for the adjustment of the model obtained by training ANN. For this purpose, two modeling strategies were adopted to be implemented in the model adjustments: *(i)* a model considering the data achieved using a 2^3 full factorial design and another for the results achieved using a 2^2 central composite design [15], as described in Table 1; and *(ii)* a general model considering all the experimental data described in Table 1. Thus, following this strategy, three ANNs will be trained and their architectures defined to verify which one will present better performance in learning and in simulating the higher colorants production by the microorganism.

Model obtained through neural network training

To adjust the model with ANN simulations, the tool "Neural network fitting tool" *(nftool)* of the MATLAB® software was used, with a graphical interface. The best results for the training of feedforward neural networks were given by the Levemberg–Marquardt algorithm *(trainlm)* with the implementation of the square sum error *(SSE)* performance objective function. The Gradient Descent Backpropagation algorithm *(traingd)* was also tested. In addition, the number of neurons in the intermediate layer, ranging from 2 to 10, and some activation functions were evaluated, using the linear function in the output layer *(purelin)* and testing in the intermediate layers the log-sigmoid functions *(logsig)* and the hyperbolic tangent sigmoid *(tansig)*.

To obtain a good generalization of the ANN, the data set was randomly divided into 70% of the samples available for training, 15% for validation and 15% for the test, as is the standard MATLAB® configuration and as most articles that apply ANN do, such as in the work of Jokić et al. [27]. The performance of the neural network, seeking the closest approximation of the model to the experimental data, was considered by analyzing the values of the correlation coefficient *(R)*, determined by the software, and the *SSE*, described by Eq. 1. Therefore, three neural networks with variations in their architectures were created, two networks considering the strategy *(i)* (analysis for the designs 2^2 and 2^3) and one network considering strategy *(ii)* (all points):

$$SSE = \sum \left(Y_{Calc} - Y_{exp} \right)^2 \tag{1}$$

where Y_{calc} is the colorant concentration calculated by the model and Y_{exp} is the experimental colorant concentration.

Adjustment of the polynomial model

For the adjustment of the polynomial model by multiple regression the Microsoft® Excel 2021 software was used to adjust the model to the experimental data, as shown in the following equation:

$$Y = a + bX_1 + cX_2 + dX_3 + eX_1X_2 + fX_1X_3 + gX_2X_3 \\ + hX_1X_2X_3 + iX_1^2 + jX_2^2 + kX_3^2 \tag{2}$$

where X_1, X_2 and X_3 are the evaluated variables (MSG, pH and glucose, respectively); Y is the response variable,

Table 1 Experimental data were obtained employing a 2^3 full factorial design and 2^2 central composite design for the red colorant production by *T. amestolkiae* [15]

Matrix of 2^3 factorial design					Matrix of 2^2 central composite design		
Run	Initial pH	Glucose (g/L)	MSG** (g/L)	Red colorant production (UA$_{500nm}$)	Initial pH	MSG[b] (g/L)	Red colorant production (UA$_{500nm}$)
1	5	0	8	0.04	3	15	0.3
2	9	0	8	0.05	3	25	0.42
3	5	10	8	0.88	5	15	0.9
4	9	10	8	0.03	5	25	2.88
5	5	0	20	0.77	2.59	20	0.08
6	9	0	20	0.06	5.41	20	1.07
7	5	10	20	1.67	4	12.95	1.61
8	9	10	20	0.05	4	27.05	2.12
9–11[a]	7	5	14	0.37±0.01	4	20	2.11±0.09

[a]Central points

[b]MSG monosodium glutamate

namely, red colorant production; *a, b, c, d, e, f, g, h, i, j* and *k* are the equation parameters. For the first strategy studied and for the results achieved in the 2^2 central composite design, the coefficients i, j and k are equal to zero; and for the results obtained in the 2^3 factorial design the values of *d, f, g, h* and *k* are equal to zero. For the second strategy evaluated all the parameters were considered.

To calculate the *SSE*, it employed Eq. 1. The *R* was determined through the multiple regression for the adjustments of the polynomial models to the experimental data.

Results and discussion

Adjustment of the ANN training data

A study on the architecture of the neural network, the activation function as well as the training algorithm was carried out to evaluate the best strategy to adjust a neural network with a high capacity to predict the red colorant production, considering as input variables the concentration of glucose and MSG; and pH. In the elaboration of artificial neural networks, different architectures were considered to achieve a better fit model, varying the number of neurons in the intermediate layer, testing two transfer functions in the intermediate layer neurons and evaluating the *SSE* and *R* obtained with two learning algorithms: Gradient Descent Backpropagation algorithm (*GD*) and Levenberg–Marquardt algorithm (*LM*). Thus, the simulations with variations listed in Table 2, were carried out, with the results obtained and the comparisons.

Assessing the effects of the Gradient Descent (*traingd*) and Levenberg–Marquardt (*trainlm*) training algorithms, the best optimizer of the weights of the connections between neurons is the *LM* training algorithm, which obtained the highest efficiency in training the neural network, because the *R* between the data was higher and the *SSE* was lower when compared to the values obtained with the Gradient

Descent algorithm (Table 2). Likewise, in the evaluation of the transfer function of the intermediate layer with the best contribution to the minimization of the objective function. The log-sigmoid function (*logsig*) (Eq. 3) was better suited in this model when compared to the hyperbolic tangent sigmoid function (*tansig*) (Eq. 4), since the output points fit in the best way with the experimental data and, consequently, generated fewer errors, with good performance results:

$$\log sig(N) = 1/(1 + \exp(-N)) \tag{3}$$

$$\tan sig(N) = 2/(1 + \exp(-2 * N)) - 1 \tag{4}$$

where N = Input column arrays, specified as an array.

In relation to the number of neurons in the intermediate layer, two neurons showed the best result in all cases, resulting highest *R* and lowest *SSE* in all cases, so it was fixed at two neurons in the intermediate layer for better functioning of the network in relation to learning speed. Considering values of *R* above 0.9, a network with only two neurons was more suitable of predicting experimental data with good quality.

Following the strategies employed in this study, three neural networks were elaborated according to the previously defined architecture. For strategy *(i)* an analysis model for the 2^3 full factorial design and another model for the results achieved using a 2^2 central composite design *(ii)* a general model. The architectures of each neural network referring to the implemented analysis model are represented in Fig. 1. In all networks evaluated in this work, the *bias* was considered for the intermediate and output layers, since they increase or decrease the degree of freedom of the weights adjustments [28].

The neural network follows the analysis model for the design 2^2 and is a good fit with the *R* equal to 0.997 and with a low *SSE*. Likewise, the neural network of the analysis model for the design 2^3 obtained *R* equal to 0.996 and low

Table 2 Artifical neural network simulations showing the architecture in the general model and their respective parameter variations

Adjust parameter

		Learning algorithms				Number of neurons			
		LM		GD		3	6	8	10
		Logsig	Tansig	Logsig	Tansig				
General model	R	0.984	0.968	0.710	0.256	0.702	0.705	0.718	0.562
	sse	0.024	0.049	0.476	0.728	0.462	0.410	0.448	0.768
2^2 model	R	0.997	0.934	0.553	0.548	0.645	0.696	0.858	0.585
	sse	0.006	0.098	0.660	0.695	0.582	0.307	0.273	0.269
2^3 model	R	0.996	0.917	0.836	0.807	0.645	0.696	0.858	0.585
	sse	0.002	0.043	0.182	0.217	0.231	0.208	0.192	0.371

[a]Fixed with 2 neurons in the intermediate layer

[b]Fixed with the learning algorithm LM and the log-sigmoid functions (*logsig*)
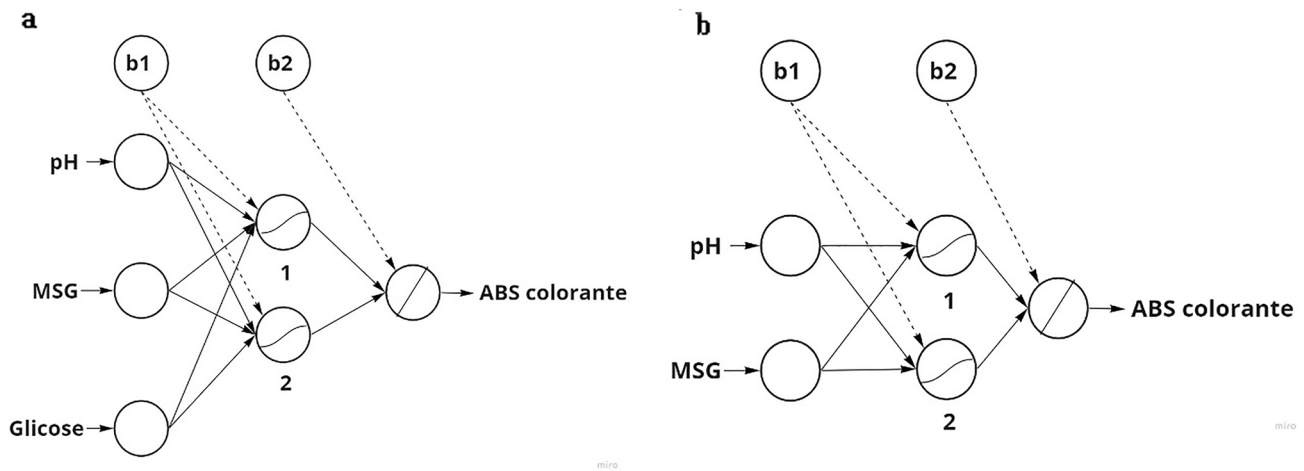
Fig. 1 Schematic representation of ANN models: an analysis model for the 2^3 full factorial design and mixed (**a**) and the 2^2 central composite design (**b**). Neurons identified as b_1 and b_2 are the bias of the intermediate and output layers, respectively

SSE. Regarding the artificial neural network of the general model, considering all experimental data, it had a great performance, with *R* equal to 0.984 and *SSE* equal to 0.024. The adjustments of the artificial neural networks of the designs 2^2, 2^3 and general analysis model are shown in Fig. 2.

The results presented in Fig. 2 show that with the artificial neural network it was possible to obtain a model capable of simulating values close to the experimental data, considering the two adjustment strategies. Therefore, the analysis suggests a robust model capable of fitting well to experimental models.

In this study, the architecture of the artificial neural network, the activation function as well as the training algorithm was carried out to evaluate the best strategy to adjust an artificial neural network with a high capacity to predict

the red colorant production, considering as input variables the concentration of glucose and MSG; and pH. Thus, it is concluded that an artificial neural network containing two neurons in the hidden layer was good to describe the process. In addition, using the best activation function was *logsig* to determine the input values for the hidden layer and *purelin* for the output layer. For neural network training, the best training algorithm was the Levenberg–Marquardt.

Adjustment of the polynomial model

To elaborate the polynomial model following the strategy *(i)*, adjustments were made to two empirical models by multiple regression. The models obtained are described by Eq. 5
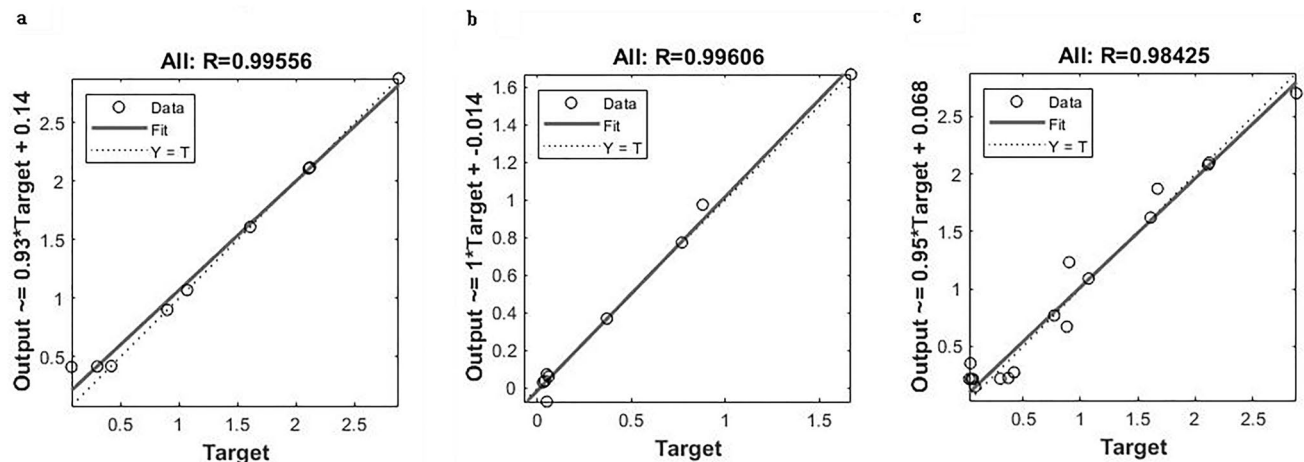

Fig. 2 Representation of the linear regression of the fits of the analysis models: 2^2 central composite design (**a**); 2^3 full factorial design (**b**); general (**c**) obtained by simulation method of artificial neural networks in MATLAB software
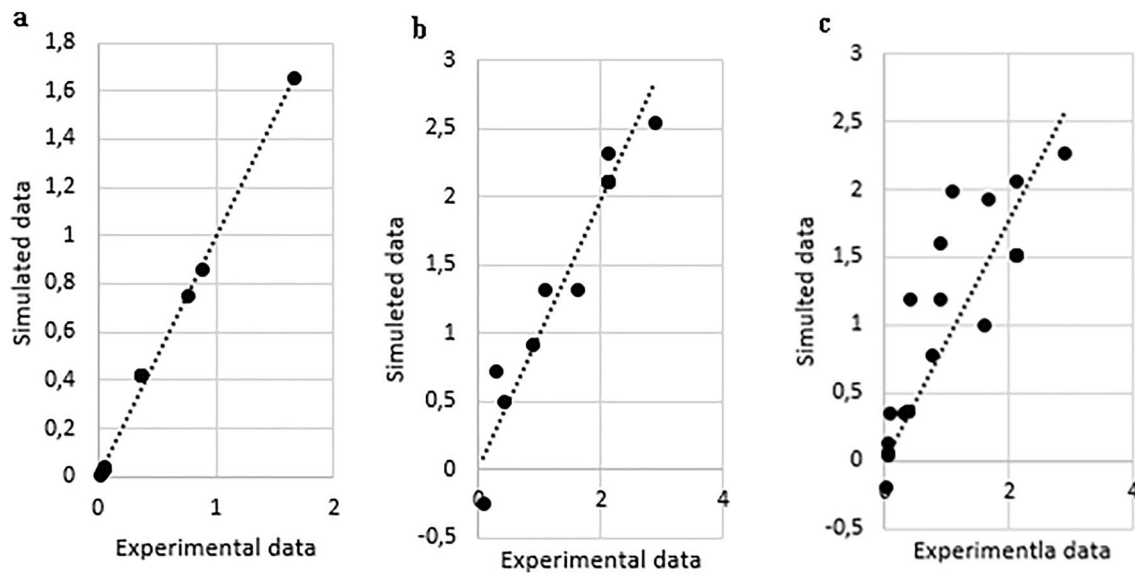
Fig. 3 Representation comparing the experimental data and the ones achieved by the simulation method of artificial neural networks in MATLAB for the 2^3 full factorial design (**a**); $2.^2$ central composite design (**b**); all the results (**c**)

(analysis model for the 2^3 full factorial design) and 6 (analysis model for the 2^2 central composite design):

$$Y = -1.079 + 0.122 * pH + 0.136 * MSG + 0.183 * Gly - 0.015$$
$$* pH * MSG \, 0.0207 * pH * Gly + 0.0001 * pH * MSG * Gly \tag{5}$$

$$Y = -9.151 + 5.057 * pH - 0.068 * MSG + 0.093 * pH * MSG$$
$$- 0.795 * pH^2 - 0.006 * MSG^2 \tag{6}$$

With this, it was possible to obtain results similar to those achieved by Oliveira et al. [15], achieving a *R* equal to 0.998 and an *SSE* equal to 0.012 for the analysis model of the design 2^3 full factorial design. The fit of the 2^2 central composite design data model obtained a *R* equal to 0.965 and *SSE* of 0.594.

Adopting the strategy *(ii)*, an empirical polynomial model was fitted to the data from Oliveira et al. [12]. Equation 7 shows the equation adjusted to the experimental data:

$$Y = -8.722 + 2.530 * pH + 0.122 * MSG - 0.003 * Gly$$
$$- 0.015 * pH * MSG - 0.04017 * pH * Gly$$
$$+ 0.0007 * pH * MSG * Gly - 0.172 * pH^2$$
$$- 0.0005 * MSG^2 + 0.032 * Gly^2 \tag{7}$$

The polynomial model gave a high *SSE* equal to 4.039 and a *R* equal to 0.871. Such values show the lowest correlation and the greatest difference between experimental and simulated data in all evaluated tests.

Figure 3 shows the comparison between the experimental data and the results obtained by the polynomial models.

Thus, the polynomial model obtained through multiple regression presented a good fit when considering the strategy (i). However, when considering the strategy (ii), the model presented a fit with the worst fit, reflecting the lowest correlation coefficient as well as the largest SSE obtained. Thus, the polynomial model presented a good fit when the strategy (i) was used, but it was not possible to obtain a good fit of the model to the experimental data using the strategy (ii). For the statistical analysis of the results, Oliveira et al. [15] analyzed the 2^3 full factorial design, which considered the concentration of MSG and glucose; and pH; and then the 2^2 central composite design, which considered the concentrations of MSG and the pH. Therefore, the fact that glucose was not considered in the second set of experiments must have contributed to the worst fit using polynomial models through strategy (ii).

Models analysis

Table 3 shows the *R* and *SSE* determined by adjust of the models evaluated to the experimental data described by Oliveira et al. [15]. For strategy *(i)* for the designs 2^3 and

Table 3 Results of polynomial adjustments (PM) and training of artificial neural networks. (ANN)

	Data from	*R*	PM	ANN
			0.998	0.996
Strategy (i)	Design 2^3	*SSE*	0.012	0.002
	Data from	*R*	0.965	0.997
	Design 2^2	*SSE*	0.594	0.006
Strategy (ii)	All data	*R*	0.871	0.984
		SEE	4.039	0.024

Fig. 4 Response surface generated by the empirical model for glucose concentration of 10 g/L of Eq. 1 (data from 2^3 design) (**a**); training of artificial neural networks for concentration and glucose of 10 g/L (data from 2^3 design) (**b**); empirical model of Eq. 2 (data from 2^2 design) (**c**); training of artificial neural networks for the data from 2^2 design (**d**); empirical model fitted to all experimental data (**e**); training of artificial neural networks for the simulation of the model fitted to all experimental data (**f**)
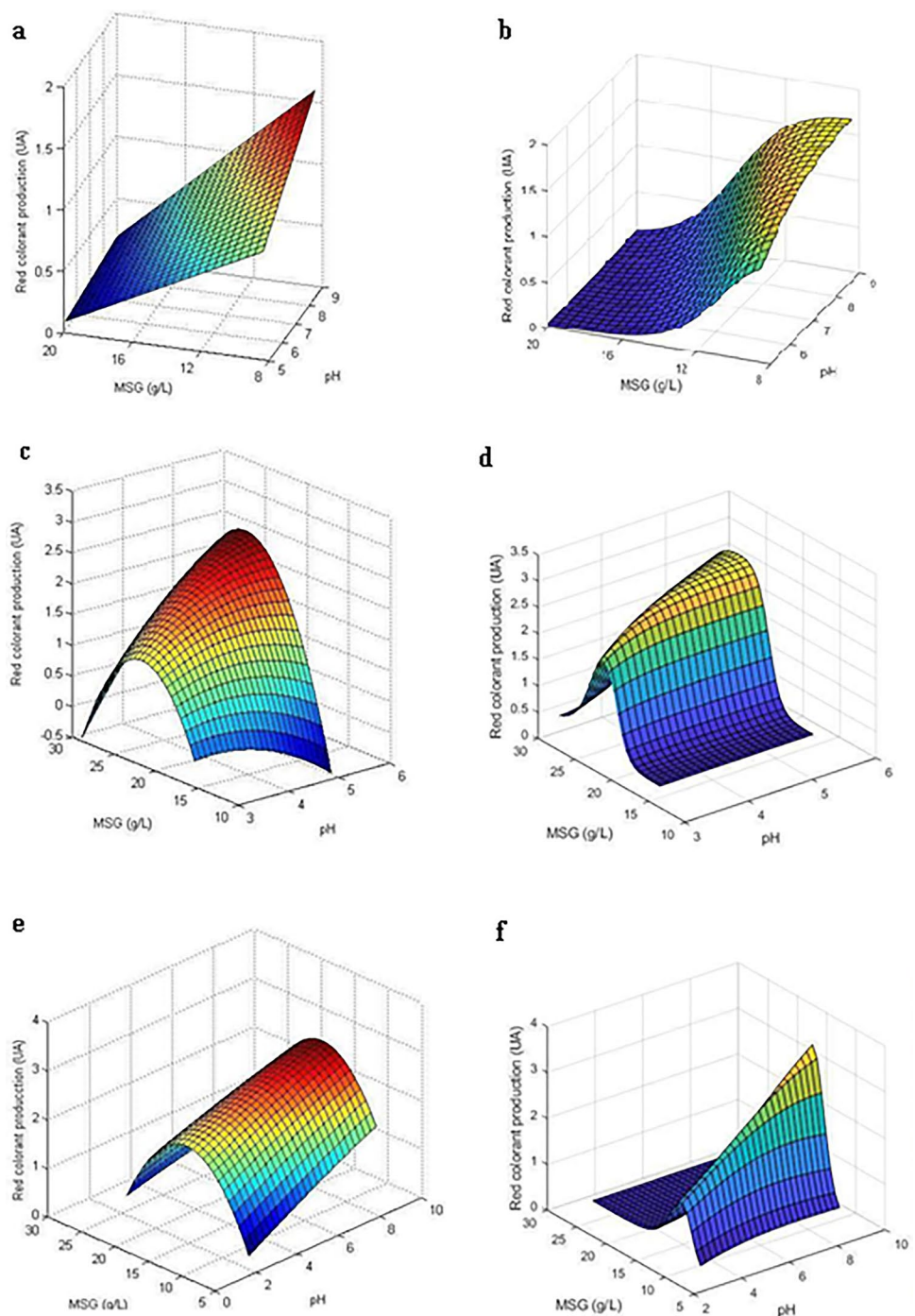
2^2 the R are low, above 0.9, low values were also obtained for *SSE* in both cases: polynomial model (PM) and ANN models.

However, when analyzing the results of the strategy *(ii)*, it is noted that only the correlation coefficient obtained through ANN model was above 0.9 and a low *SSE*. Considering the

use of the adjustment of the PM to the experimental data obtained through multiple regression, it presented the worst R, below 0.9, and the highest *SSE* value.

Comparing the fit obtained through PM´s and the ANN model, it was found that the output values of the ANN are closer to the experimental data in all sets of results evaluated

(Figs. 2 and 3), presenting higher value of R and lower value of *SSE* in all selected artificial neural networks. Such comparison highlights the potential of using artificial neural networks and the architecture used to predict the bioprocess used in this study.

The advantages of using the ANN against the PM is the higher level of prediction accuracy of the results. ANN can be more sensitive in predicting the initial conditions that improve the production of the final product. Other advantage is that the model is not linear and can better understand the variation of the microorganism's metabolism during the production process and, thus, obtain a better prediction of the final production. In addition, the ANN can assist in the prediction and/or optimization of the colorants production or any other bioprocesses [29].

One of the main problems in the bioprocess industry is being able to follow the production process of a microorganism and ANN has shown itself to be able to predict this production with better precision than the most used model, like the PM. In this way, the trained networks can be used to simulate the biological process to further adequate the initial conditions to obtain a better output result.

Comparing the surface graphs described by the polynomial models (Fig. 4a, c and e) and the graphs obtained using artificial neural networks (Fig. 4b, d and f), it was observed similar trends, but a different profile. In graphs using results simulated by ANN and when compared to the polynomial model, it was observed a lower concentration range of MSG in which colorant production occurs. In relation to the pH, it is observed that the prediction of colorant production by the neural network model was higher than that of the polynomial model. This can be explained by the fact that artificial neural networks are composed of activation functions, which vary from zero to one abruptly, which can help to predict regions in which a given input variable has a greater influence. A polynomial model presents smoother curves, which makes it difficult to predict a region of greater influence of a given variable.

In relation to the pH, it was also observed the same profile and trend by the results simulated by both models, indicating the use of more alkaline media to produce red colorants. Regarding the simulation of the models adjusted to the data from the design 2^2, the results simulated by the artificial neural network show that there is a greater pH range in which the greatest colorant production occurs. In this case, the simulated values by the artificial neural network of colorant concentration were from 1.5 to 3.5 uABS and by the polynomial model were 1 to 2.25, considering the best condition of MSG and pH range between 2.5 and 5.4.

Regarding glucose concentration, graphs were not performed, as the effect of glucose was not considered significant in the range evaluated. This result was described by Oliveira et al. [15]. In all strategies used in this work, the ANN models fitted the experimental results better than the polynomial models. Therefore, ANN models can be used for predicting results as well as optimizing process conditions, making it possible to obtain better results than those determined through polynomial models.

Conclusions

In this work, it was possible to conclude that the application of artificial neural network techniques to predict the production of secondary metabolites by filamentous fungi may be a promising technique. The artificial neural network proved to be robust to be adjusted to the experimental data, considering any strategy addressed in this study. The results also show that the red colorant concentration predicted by artificial neural networks are closer to the experimental data for red colorant production than the red colorant concentration predicted by the polynomial models adjusted by the multiple regression method, since they did not present a good fit in strategy *(ii)*, in which all results were considered.

It is also noteworthy that the number of experimental results used to adjust the polynomial models and the ANN were the same, so there was no need for a greater number of experimental results than those already predicted by the experimental designs. Considering the two strategies evaluated, a simple ANN, containing from two to three input neurons (MSG and glucose concentration; and pH) and two neurons in the intermediate layer was able to adjust well to the experimental results of the red colorant production bioprocesses. Thus, this work suggests that the use of ANN for the optimization of experimental conditions can be a tool to be used, since the predicted values are closer to the experimental ones.

(Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil), finance code 001.

Data availability The authors declare that the data supporting the findings of this study are available within the article.

Declarations

Conflict of interest The authors declare that they have no interest in conflict with the work described in this paper.

References

1. Morales-Oyervides L, Ruiz-Sánchez JP, Oliveira JC, Sousa-Gallagher MJ, Méndez-Zavala A, Giuffrida D, Dufossé L, Montañez J (2020) Biotechnological approaches for the production of natural colorants by Talaromyces/Penicillium: a review. Biotechnol Adv. https://doi.org/10.1016/j.biotechadv.2020.107601

2. Lehto S, Buchweitz M, Klimm A, Straßburger R, Bechtold C, Ulberth F (2017) Comparison of food colour regulations in the EU and the US: a review of current provisions. Food Addit Contam Part A Chem Anal Control Expo Risk Assess 34:335–355. https://doi.org/10.1080/19440049.2016.1274431

3. Sen T, Barrow CJ, Deshmukh SK (2019) Microbial pigments in the food industry—challenges and the way forward. Front Nutr 6:1–14. https://doi.org/10.3389/fnut.2019.00007

4. Sharma J, Sharma S, Soni V (2021) Classification and impact of synthetic textile dyes on Aquatic Flora: a review. Reg Stud Mar Sci. https://doi.org/10.1016/j.rsma.2021.101802

5. Lagashetti AC, Dufossé L, Singh SK, Singh PN (2019) Fungal pigments and their prospects in different industries. Microorganisms 7:1–36. https://doi.org/10.3390/microorganisms7120604

6. Mordor Intelligence, Natural food colorants market: growth, trend, COVID-19 impact, and forecast (2022–2027). Available online:, mordorintelligence.com/industry-reports/global-natural-food-colorants-market. Accessed 7 Febr. 2022. (2021)

7. Torres FAE, Zaccarim BR, de Lencastre Novaes LC, Jozala AF, dos Santos CA, Teixeira MFS, Santos-Ebinuma VC (2016) Natural colorants from filamentous fungi. Appl Microbiol Biotechnol 100:2511–2521. https://doi.org/10.1007/s00253-015-7274-x

8. Chen W, Chen R, Liu Q, He Y, He K, Ding X, Kang L, Guo X, Xie N, Zhou Y, Lu Y, Cox RJ, Molnár I, Li M, Shao Y, Chen F (2017) Orange, red, yellow: biosynthesis of azaphilone pigments in Monascus fungi. Chem Sci 8:4917–4925. https://doi.org/10.1039/C7SC00475C

9. Mapari SAS, Thrane U, Meyer AS (2010) Fungal polyketide azaphilone pigments as future natural food colorants? Trends Biotechnol 28:300–307. https://doi.org/10.1016/j.tibtech.2010.03.004

10. Caro Y, Anamale L, Fouillaud M, Laurent P, Petit T, Dufosse L (2012) Natural hydroxyanthraquinoid pigments as potent food grade colorants: an overview. Nat Products Bioprospect 2:174–193. https://doi.org/10.1007/s13659-012-0086-0

11. Yang J, Chen Q, Wang W, Hu J, Hu C (2015) Effect of oxygen supply on Monascus pigments and citrinin production in submerged fermentation. J Biosci Bioeng 119:564–569. https://doi.org/10.1016/j.jbiosc.2014.10.014

12. Teixeira MFS, Martins MS, da Silva JC, Kirsch LS, Fernandes OCC, Carneiro ALB, de Conti R, Durán N (2012) Amazonian biodiversity: Pigments from Aspergillus and Penicillium-characterizations, antibacterial activities and their Toxicities. Curr Trends Biotechnol Pharm 6:300–311

13. Mapari SAS, Hansen ME, Meyer AS, Thrane U (2008) Computerized screening for novel producers of Monascus-like food pigments in Penicillium species. J Agric Food Chem 56:9981–9989. https://doi.org/10.1021/jf801817q

14. de Oliveira F, Rocha ILD, Claudia Gouveia Alves Pinto D, Ventura SPM, Gonzaga A, dos Santos E, De José Crevelin V, Ebinuma CS (2022) Identification of azaphilone derivatives of Monascus colorants from Talaromyces amestolkiae and their halochromic properties. Food Chem. https://doi.org/10.1016/j.foodchem.2021.131214

15. de Oliveira F, Pedrolli DB, Teixeira MFS, de Carvalho Santos-Ebinuma V (2019) Water-soluble fluorescent red colorant production by Talaromyces amestolkiae. Appl Microbiol Biotechnol 103:6529–6541. https://doi.org/10.1007/s00253-019-09972-z

16. de Oliveira F, Ferreira LC, Neto ÁB, Simas Teixeira MF, de Carvalho Santos V, Ebinuma, (2020) Biosynthesis of natural colorant by Talaromyces amestolkiae: Mycelium accumulation and colorant formation in incubator shaker and in bioreactor. Biochem Eng J. https://doi.org/10.1016/j.bej.2020.107694

17. Abdul Manan M, Ariff A, Mohamad R, Abdul Karim M (2005) Kinetics and modeling of red pigment fermentation by Monascus purpureus FTC 5391 in 2-litre stirred tank fermenter using glucose as a carbon source. J Trop Agric Food Sci 33:277–284

18. Panagou EZ, Skandamis PN, Nychas GJE (2003) Modelling the combined effect of temperature, pH and aw on the growth rate of Monascus ruber, a heat-resistant fungus isolated from green table olives. J Appl Microbiol 94:146–156. https://doi.org/10.1046/j.1365-2672.2003.01818.x

19. Santos-Ebinuma VC, Roberto IC, Simas Teixeira MF, Pessoa A (2013) Improving of red colorants production by a new Penicillium purpurogenum strain in submerged culture and the effect of different parameters in their stability. Biotechnol Prog 29:778–785. https://doi.org/10.1002/btpr.1720

20. Tolborg G, Ødum ASR, Isbrandt T, Larsen TO, Workman M (2020) Unique processes yielding pure azaphilones in Talaromyces atroroseus. Appl Microbiol Biotechnol 104:603–613. https://doi.org/10.1007/s00253-019-10112-w

21. Zhou B, Wang J, Pu Y, Zhu M, Liu S, Liang S (2009) Optimization of culture medium for yellow pigments production with Monascus anka mutant using response surface methodology. Eur Food Res Technol 228:895–901. https://doi.org/10.1007/s00217-008-1002-z

22. V.C. Liyanaarachchi, M. Premaratne, P.H. Viraj Nimarshana, T. Udayangani Ariyadasa, Investigation of the Effect of Organic and Inorganic Carbon on Biomass Production and Astaxanthin Accumulation of the Microalga Haematococcus pluvialis Using Artificial Neural Network, 2020 IEEE 17th India Counc Int Conf INDICON 2020. (2020). https://doi.org/10.1109/INDICON49873.2020.9342373

23. Sommer R, Paxson V (2010) Outside the closed world: On using machine learning for network intrusion detection. Proc IEEE Symp Secur Priv. https://doi.org/10.1109/SP.2010.25

24. Karim MN, Yoshida T, Rivera SL, Saucedo VM, Eikens B, Gyu-Seop OH (1997) Global and local neural network models in biotechnology: Application to different cultivation processes. J Ferment Bioeng 83:1–11. https://doi.org/10.1016/S0922-338X(97)87318-7

25. Singh N, Goel G, Singh N, Pathak BK, Kaushik D (2015) Modeling the red pigment production by Monascus purpureus

MTCC 369 by Artificial Neural Network using rice water based medium. Food Biosci 11:17–22. https://doi.org/10.1016/j.fbio.2015.04.001

26. Durakli-Velioğlu S, Boyaci IH, Şimşek O, Gümüş T (2013) Optimizing a submerged Monascus cultivation for production of red pigment with bug damaged wheat using artificial neural networks. Food Sci Biotechnol 22:1639–1648. https://doi.org/10.1007/s10068-013-0261-z

27. Jokić A, Pajčin I, Grahovac J, Lukić N, Ikonić B, Nikolić N, Vlajkov V (2020) Dynamic modeling using artificial neural network of bacillus velezensis broth cross-flow microfiltration enhanced by air-sparging and turbulence promoter. Membranes (Basel) 10:1–14. https://doi.org/10.3390/membranes10120372

28. Haykin S (2001) Neural Networks: a comprehensive foundation. 2. ed. Delhi, Índia: Pearson Education 1:33.

29. Sewsynker-Sukai Y, Faloye F, Kana EBG (2016) Artificial neural networks: an efficient tool for modelling and optimization of biofuel production (a mini review). Biotechnol Biotechnolog Equip 31:221–235. https://doi.org/10.1080/13102818.2016.1269616