



# Hybrid neural network modeling and particle swarm optimization for improved ethanol production from cashew apple juice

Andréa da Silva Pereira<sup>1</sup> · Álvaro Daniel Teles Pinheiro<sup>2</sup> · Maria Valdez Pontes Rocha<sup>1</sup> ·  
Luciana Rocha B. Gonçalves<sup>1</sup> · Samuel Jorge Marques Cartaxo<sup>1</sup>

Received: 17 April 2020 / Accepted: 9 September 2020 / Published online: 29 September 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

A hybrid neural model (HNM) and particle swarm optimization (PSO) was used to optimize ethanol production by a flocculating yeast, grown on cashew apple juice. HNM was obtained by combining artificial neural network (ANN), which predicted reaction specific rates, to mass balance equations for substrate (S), product and biomass (X) concentration, being an alternative method for predicting the behavior of complex systems. ANNs training was conducted using an experimental set of data of X and S, temperature and stirring speed. The HNM was statistically validated against a new dataset, being capable of representing the system behavior. The model was optimized based on a multiobjective function relating efficiency and productivity by applying the PSO. Optimal estimated conditions were:  $S_0 = 127 \text{ g L}^{-1}$ ,  $X_0 = 5.8 \text{ g L}^{-1}$ ,  $35 \text{ °C}$  and  $111 \text{ rpm}$ . In this condition, an efficiency of  $91.5\%$  with a productivity of  $8.0 \text{ g L}^{-1} \text{ h}^{-1}$  was obtained at approximately 7 h of fermentation.

**Keywords** Ethanol production · Cashew apple juice · Artificial neural network (ANN) · Hybrid neural model (HNM) · Particle swarm optimization (PSO)

## Nomenclature

$A_1$	Initial value of horizontal asymptote
$A_2$	Final value of horizontal asymptote
ANN	Artificial neural network
$dx$	Model increment
$F_{\text{tab}}$	Tabulated value for the Fisher Test F
HNM	Hybrid neural model
$n$	Number of samples
$n_v$	Number of variables estimated
$N$	Stirring speed (rpm)
$p$	Number of model parameters
$P$	Product concentration ( $\text{g L}^{-1}$ )
$P_f$	Final product concentration ( $\text{g L}^{-1}$ )
PSO	Particle Swarm Optimization
RSD	Residual standard deviation (%)
$S$	Substrate concentration ( $\text{g L}^{-1}$ )
$S_0$	Initial substrate concentration ( $\text{g L}^{-1}$ )

$S_f$	Final substrate concentration ( $\text{g L}^{-1}$ )
$T$	Time (h)
$t_f$	Final time (h)
$x$	Model variable
$x_0$	Average value between horizontal asymptotes
$X$	Cell concentration ( $\text{g L}^{-1}$ )
$X_0$	Initial cell concentration ( $\text{g L}^{-1}$ )
$\varepsilon$	Error (%)
$\mu_S$	Specific substrate consumption rate ( $\text{g}_{\text{subs}} \cdot \text{g}_{\text{cell}}^{-1} \cdot \text{h}^{-1}$ )
$\mu_P$	Specific product production rate ( $\text{g}_{\text{product}} \cdot \text{g}_{\text{cell}}^{-1} \cdot \text{h}^{-1}$ )
$\mu_X$	Specific growth rate of cells ( $\text{h}^{-1}$ )

## Superscript

cal	Calculated
exp	Experimental
min	Minimum
max	Maximum
n	Normalized

✉ Luciana Rocha B. Gonçalves  
lrg@ufc.br

<sup>1</sup> Chemical Engineering Department, Universidade Federal do Ceará, Campus do Pici, Bloco 709, Fortaleza, CE 60455-76, Brazil

<sup>2</sup> Engineering and Technology Department, Universidade Federal Rural do Semiárido, Mossoró, RN, Brazil

## Introduction

Currently, ethanol is the main substitute for gasoline and can be obtained from alcoholic fermentation [1–3]. The main raw materials used for the industrial production of ethanol are corn and sugarcane. Other sources of biomass have also

been studied and used, such as barley and wheat. However, the increase in world demand encourages the search for other alternative raw materials. In this context, some authors evaluated the potential of cashew apple juice as culture medium [4–8], and ethanol was produced in a laboratory scale at high yields and productivities.

In these processes, yeasts are widely used and *Saccharomyces cerevisiae*, which naturally evolved to efficiently consume sugars such as sucrose, is one of the most important cell used in industrial ethanol production due to its robustness, stress tolerance, genetic accessibility, simple nutrient requirements [9]. Moreover, it is one of the most studied yeast by the scientific community [9–11]. In this work, however, a flocculent *S. cerevisiae* was evaluated aiming to eliminate the cost of separation of cells.

The modified *S. cerevisiae* (FLO5 $\alpha$  gene) tends to form small flocs that decant on the bottom of the fermenter at the end of reaction. This characteristic allows the microorganism to be easily separated from the fermented medium, which reduces process costs once centrifugation becomes unnecessary [12]. Nevertheless, to favor substrate diffusion into the cell or inside the flocs, the stirring speed is essential to avoid that the cells assume a flocculated state during fermentation [13–15]. Therefore, the influence of stirring speed in the reaction medium containing the flocculating *S. cerevisiae* should be considered as a fermentation parameter [7].

Modeling of real fermentation processes presents a high degree of complexity due to genetic characteristics, physico-chemical and biochemical laws involved, besides the non-linearity of its kinetics [16]. Biochemical reactions involve many stages of multiple reactions (in series and parallel) and depend on several transport phenomena that may limit the observed reaction rates [17].

In general, the rigorous approach of the involved processes can be difficult to apply in the kinetic modeling due to inherent non-linearity, lack of information and experimental inaccuracy, as well as deviations from ideal conditions [18]. Thus, finding a faster and simpler way to describe fermentative processes may be more advantageous. Hybrid modeling emerges as an alternative to combine prior knowledge of the process through mass balances with artificial neural network (ANN) that describe the unknown kinetics of the process. Several authors have already proposed the hybrid modeling strategy in their studies and concluded that they are reliable [17, 19–21].

Mathematical model allows to optimize the physical–chemical parameters that influence the general productivity of the process. In this study, the operating conditions such as temperature, stirring speed, initial cell and substrate concentrations play a synergistic role in controlling cell growth and ethanol production. Therefore, the use of advanced modeling and optimization tools, including artificial neural network (ANN) and particle swarm optimization

technique (PSO), was proposed. Those algorithms have been found to be more efficient than other statistical optimization techniques in deriving global optimal solutions for complex and non-linear bioprocesses [22–25].

PSO is a non-deterministic bio-inspired population optimization method and can be applied to optimize non-linear and non-continuous problems with multivariable [26]. It is based on a constructive method to obtain the initial population and a local search technique to improve the solution of the population. With this intention, the individuals (solutions) of this population should evolve according to specific rules that consider the exchange of information among the individuals, leading the population to an optimal solution [27]. Compared with other evolutionary algorithms, PSO has some advantages such as ease implementation, better efficiency, less memory requirement, and constructive cooperation between individuals. Therefore, it is more likely and quicker to “flock” into the better solution areas and discover the optimal results much faster [26, 28].

The originality of this work consists in the use of an HNM-PSO combined strategy to optimize the operational conditions, creating a faster, simpler and more efficient alternative to a mechanistic model. Therefore, a hybrid neural model (HNM) was proposed to represent the alcoholic fermentation of cashew apple juice by a flocculating *S. cerevisiae*, including the influence of cell and substrate concentration, as well as temperature and stirring speed. Then, a combined HNM-PSO algorithm was implemented to maximize ethanol production by changing operational conditions.

## Materials and methods

### Experimental data

Ethanol was produced by *Saccharomyces cerevisiae* CCA008 (with the modified gen FLO5 $\alpha$ ) using cashew apple juice (750 mL of medium at pH 4.5) in a 1 L bench-scale bioreactor (Tec-Bio, Model 1.5, Tecnal, SP, Brazil). All experiments were performed in duplicate as reported by Pinheiro et al. [29].

The operating conditions, efficiency and productivity of the fermentative process are summarized in Table 1. The initial concentration of substrate was varied from 70 to 170 g L<sup>-1</sup>, temperature from 26 to 42 °C, initial cell concentration from 4 to 10 g L<sup>-1</sup> and stirring speed from 80 to 800 rpm with a processing time of 10 h.

### Data processing

Regressions, interpolations, and normalization of experimental data were performed to increase the number of points available per intermediate points prediction. This data

**Table 1** Operational conditions of batch fermentation to produce ethanol by *Saccharomyces cerevisiae* CCA008 using cashew apple juice as substrate

Assay	S <sub>0</sub> (g L <sup>-1</sup> )	T (°C)	X <sub>0</sub> (g.L <sup>-1</sup> )	N (rpm)	Efficiency (%)	Productivity (g L <sup>-1</sup> h <sup>-1</sup> )
1	70	30	4	150	83.6	3.8
2	80	30	4	150	82.6	4.2
3	90	30	4	150	90.0	5.4
4	110	30	4	150	93.0	5.8
5	130	30	4	150	89.5	6.2
6	170	30	4	150	80.2	5.3
7	100	26	5	150	62.3	3.7
8	100	30	5	150	84.5	5.4
9	100	34	5	150	94.4	6.5
10	100	38	5	150	84.3	5.1
11	100	42	5	150	84.0	5.4
12	100	34	3	150	96.8	4.2
13	100	34	8	150	82.7	5.5
14	100	34	10	150	74.6	5.4
15	110	34	7	80	89.7	6.0
16	110	34	4	300	94.1	6.2
17	110	34	5	490	94.2	6.5
18	110	34	5	650	92.8	6.2
19	110	34	5	800	84.2	5.6

<sup>a</sup>Standards uncertainties: Initial Substrate Concentration (S<sub>0</sub>)=5.45 g/L, Temperature (T)=2 °C; Initial Cell Concentration (X<sub>0</sub>)=0.41 g/L, Stirring Speed (N)=5%

processing was necessary, since the ANN algorithms require a large amount of data. Moreover, mathematical model calibration and neural networks training should avoid bias in the data, due for example to experimental measurement error. Therefore, Boltzmann’s regression model [30] (Eq. 1) was selected, because it has the necessary functional features to fit the behavior of curves (inflection points and asymptotes). The least squares method was applied to determine the function parameters for substrate, cell and product concentration profiles at each operational condition. The coefficient of determination (R<sup>2</sup>) was employed aiming to certify the quality of each fitting, varying between 0.97 and 1.00:

$$y = \frac{A_1 - A_2}{1 + e^{\frac{(x-x_0)}{dx}}} + A_2 \tag{1}$$

The functions were interpolated in intervals of 30 min, to quadruple the data intended for ANNs training. Thus, specific rates of glucose consumption, cell growth and ethanol production were estimated, calculating the derivative of the correspondent Boltzmann’s equation for the 19 assays, according to Eqs. 2–4:

$$\mu_X(t) = \frac{\frac{dX(t)}{dt}}{X(t)} \tag{2}$$

$$\mu_S(t) = \frac{\frac{dS(t)}{dt}}{X(t)} \tag{3}$$

$$\mu_P(t) = \frac{\frac{dP(t)}{dt}}{X(t)} \tag{4}$$

In summary, the implemented algorithm (Fig. 1) to treat the data requires as input the experimental dataset, including duplicates, a function for regression (Boltzmann’s regression model) and a time interval in which the data points will be interpolated. As output, the concentration and specific rate profiles are obtained, as well as the interpolated data for both, also called pseudo-experimental data.

Subsequently, the pseudo-experimental data must be normalized before applying the artificial intelligence methods, since there is a significant improvement in the data distribution. It numerically corresponds to adequate the order of magnitude of different variables that can be very diverging in magnitude. In this work the min–max normalization was adopted, Eq. 5, with unitary interval:

$$z_i^n = \frac{z_i - z_i^{min}}{z_i^{max} - z_i^{min}} \tag{5}$$

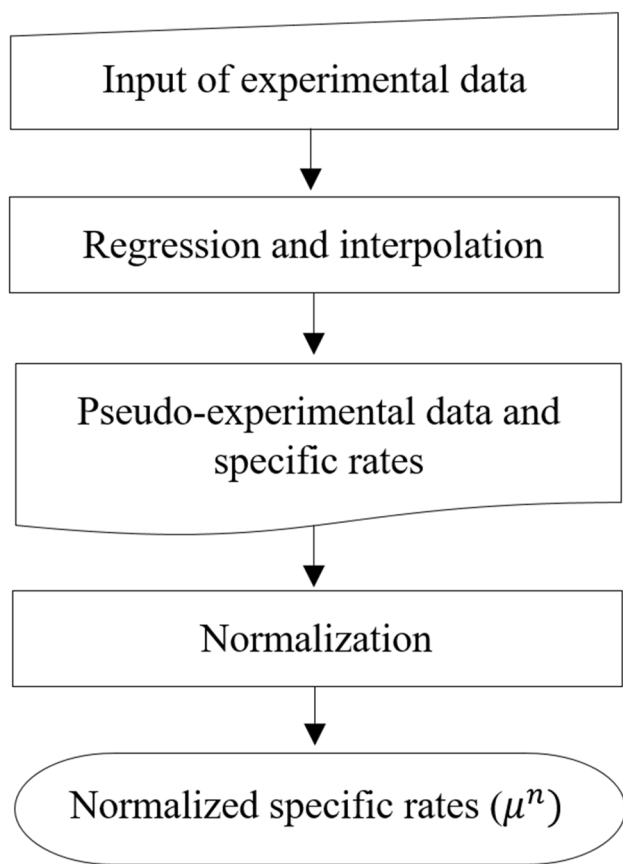


Fig. 1 Algorithm diagram implemented for data processing

### Neural network development

Figure 2 shows the typical structure of the neural networks developed, with input and output data fed for training, consisting of three types of layers: input, hidden and output. The interconnection between the neurons in each layer is defined by weights and biases. ANN learns the cause–effect relationship between input and output variables of the given dataset, updating their weights so that the error between the given data and the simulated output is minimized.

In this work, the described neural networks were developed by an iterative procedure implemented in Python/IPython Notebook version 2.7.8 language associated with PyBrain library for machine learning with backpropagation trainer.

Of the 19 operational conditions, 01 experiment was randomly chosen for validation (assay 13) and 18 experiments were used for train/test the ANN: the experimental data were randomly split into two groups, reserving 75% of data to the training phase and the remaining 25% to test the neural networks.

The architecture of the ANN (number of nodes in each layer and the number of hidden layers) was defined by a trial-and-error procedure. For that, the coefficient of determination ( $R^2$ ), as well as the maximum and average error between experimental and predicted data by ANN was observed.

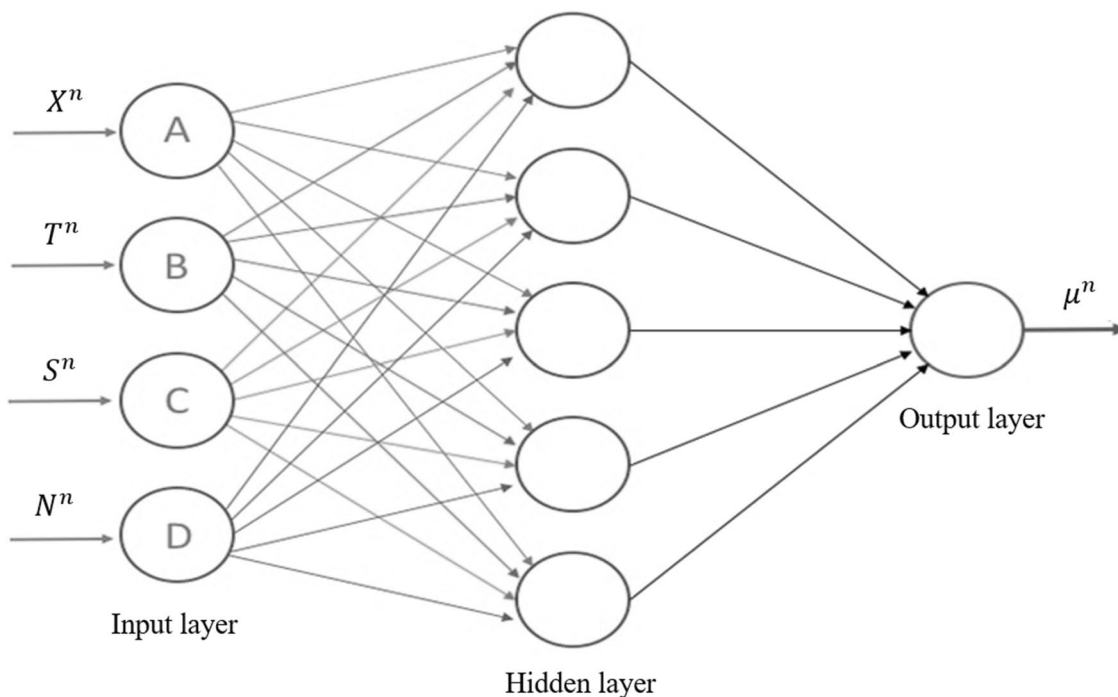


Fig. 2 ANN typical structure

## Mathematical modeling

### Hybrid neural model (HNM)

The fermentative models consist of a set of differential equations obtained by combining batch reactor mass balances by component (cell, substrate and product) to specific rates of reaction ( $\mu$ ). The simplifying hypotheses for the mathematical are:

- i. All cells in the fermentative medium were viable;
- ii. Perfecting mixing system, justified by the presence of a mechanical stirring device;
- iii. Isothermal, since the bioreactor was equipped with a temperature control system;
- iv. Constant reaction volume;
- v. Substrate consumption for the cell maintenance was neglected.

$$\frac{dX}{dt} = \mu_X \cdot X \tag{6}$$

$$\frac{dS}{dt} = -\mu_S \cdot X \tag{7}$$

$$\frac{dP}{dt} = \mu_P \cdot X \tag{8}$$

In this case, the HNM proposed for the alcoholic fermentation of cashew apple juice by *S. cerevisiae* CCA008 combines mass balances with ANNs. The ANNs work as estimators for the specific rates of cells growth, substrate consumption and production formation. Thus, three networks were created: ANN1 ( $\mu_X$ ) ANN2 ( $\mu_S$ ) and ANN3 ( $\mu_P$ ).

For better performance in the ANNs training step, the input and output layers were fed with normalized data (Eq. 9), assigning the same weight for each input variable:

$$\mu^n = f(X^n, T^n, S^n, N^n) \tag{9}$$

$$ANN = f\left(\frac{X - X^{min}}{X^{max} - X^{min}}, \frac{T - T^{min}}{T^{max} - T^{min}}, \frac{S - S^{min}}{S^{max} - S^{min}}, \frac{N - N^{min}}{N^{max} - N^{min}}\right) \tag{10}$$

As the specific rates were normalized (ANNs), an algebraic manipulation between normalization function (Eq. 11) and Eqs. 6 to 8 was necessary, resulting in the hybrid model presented in Eqs. 11–15:

$$\mu^n = \frac{\mu - \mu^{min}}{\mu^{max} - \mu^{min}} \tag{11}$$

$$\mu = \mu^n \cdot (\mu^{max} - \mu^{min}) + \mu^{min} \tag{12}$$

$$\frac{dX}{dt} = [ANN1 \cdot (\mu_X^{max} - \mu_X^{min}) + \mu_X^{min}] \cdot X \tag{13}$$

$$\frac{dS}{dt} = -[ANN2 \cdot (\mu_S^{max} - \mu_S^{min}) + \mu_S^{min}] \cdot X \tag{14}$$

$$\frac{dP}{dt} = [ANN3 \cdot (\mu_P^{max} - \mu_P^{min}) + \mu_P^{min}] \cdot X \tag{15}$$

### HNM implementation

The implemented HNM is schematized as reported in Fig. 3. Initially, the operational conditions are specified to calculate initial specific rates of reaction (ANN1, ANN2 and ANN3). Next, the HNM is resolved by a combination of ANNs previously trained and the developed mathematical model, combined with two mechanistic conditions to guarantee the physical meaning of the model:

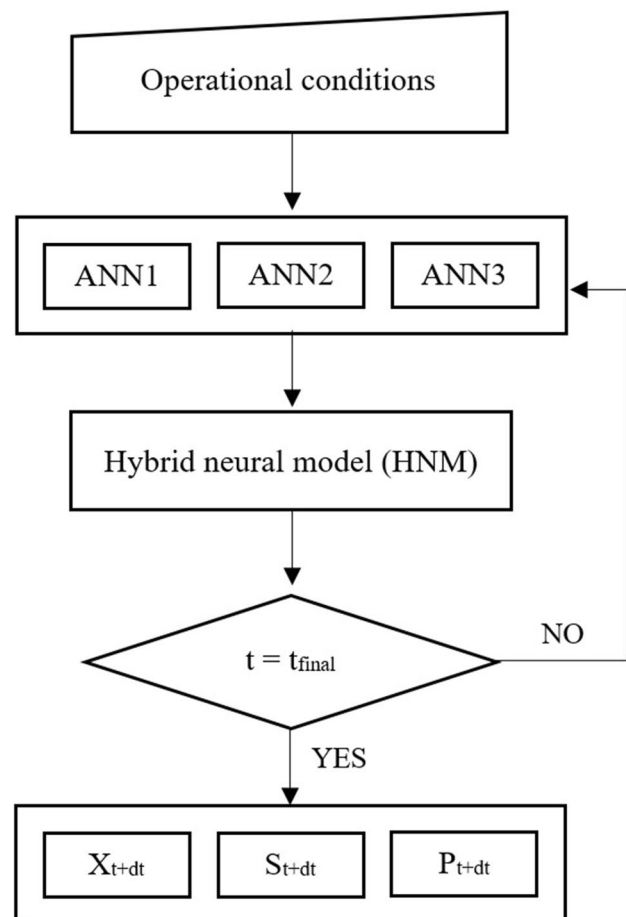


Fig. 3 HNM implementation flowchart

- i. Maximum theoretical yield for ethanol production is  $0.511 \text{ g}_{\text{ethanol}}/\text{g}_{\text{glucose}}$ , due to stoichiometry.
- ii. No reaction takes place in absence of substrate ( $dX/dt = 0$ ,  $dS/dt = 0$  and  $dP/dt = 0$ ).

Thus, the specific rates can be estimated as a function of the fermentative medium conditions (substrate and cells concentrations, temperature and stirring speed) for each instant of the reaction. The new conditions are the responses of the ODE system solution in the instant  $t + \Delta t$  until it reaches the time determined to end the reaction and attainment of the concentration profiles of biomass, substrate and product.

### Accuracy assessment

The precision quality of HNM was evaluated by statistical analysis, as follows: residual standard deviation—RSD (%) and modified  $F$  test. The RSD suggested by Cleran et al. [31], as seen in Eq. 16, was used to assess the quality of the prediction models:

$$RSD = \sqrt{\frac{\sum_{i=1}^n (y_i^{\text{exp}} - y_i^{\text{cal}})^2}{n}} \quad (16)$$

The modified  $F$  test is a way to discern models by calculating the variance of the error between experimental data and theoretical data [32]. This hypothesis test can be used to verify the adequacy of a mathematical model, when the average experimental error of the data is greater than the apparent experimental error calculated by the model ( $\mathcal{E}^{\text{exp}} > \mathcal{E}^{\text{apparent}}$ ). With the Eqs. 17 and 18, it is possible to estimate the apparent experimental error (Eq. 19):

$$S_m^2 = \frac{\sum_j \left[ (X_j^{\text{exp}} - X_j^{\text{cal}})^2 + (S_j^{\text{exp}} - S_j^{\text{cal}})^2 + (P_j^{\text{exp}} - P_j^{\text{cal}})^2 \right]}{3n - p} \quad (17)$$

$$S_e^2 = \frac{\mathcal{E}^2 \sum_j \left[ (X_j^{\text{exp}})^2 + (S_j^{\text{exp}})^2 + (P_j^{\text{exp}})^2 \right]}{3n - n_v} \quad (18)$$

$$\mathcal{E}^{\text{exp}} > \sqrt{\frac{S_m^2 (3n - p)}{F_{\text{tab}} \cdot \sum_j \left[ (X_j^{\text{exp}})^2 + (S_j^{\text{exp}})^2 + (P_j^{\text{exp}})^2 \right]}} \quad (19)$$

## Optimization method

### Particle swarm optimization—PSO

In this work, the optimization aims to determine the optimal operational conditions for the alcoholic fermentation of cashew apple juice, that maximizes efficiency and ethanol productivity. For that, PSO was used combined to the HNM.

PSO uses a swarm population, where each individual within the swarm is denominated particle. According to Jiao, Lian and Gu (2008) [33], a particle  $i$  in an interaction  $k$  moves through the search space with two attributes [33]:

- The current position within the search N-dimensional space  $X_i^k = (x_1^k, \dots, x_n^k, \dots, x_N^k)$  of the problem, with  $x_n^{\text{min}} \leq x_n^k \leq x_n^{\text{max}}$  for each  $n \in [1, N]$ , where  $x_n^{\text{min}}$  and  $x_n^{\text{max}}$  are the limits of coordinate  $n$
- Its speed is vectorially represented by  $V_i^k = (v_1^k, \dots, v_n^k, \dots, v_N^k)$  in the same N-dimensional space of the problem.

After each iteration the speed and position of all particles are updated according to the two best values found during the search. The first one is calculated by the PSO regarding each individual best value found during its lifetime,  $pbest$ . The other one is calculated considering the best value of ensemble of points, swarm, named as  $gbest$ . After finding the two best values, the position and speed of the particles are obtained by Eqs. 20 and 21:

$$V_i^{k+1} = w \cdot V_i^k + c_1 \cdot r_1 \cdot (pbest_i^k - X_i^k) + c_2 \cdot r_2 \cdot (gbest^k - X_i^k) \quad (20)$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \quad (21)$$

### PSO implementation

The first step for PSO implementation is to define the control variables and the objective function to be used for maximization.

- i. Control variables: cells and substrate initial concentrations, temperature and stirring speed.
- ii. Objective function: efficiency and productivity.

**Table 2** Side bounds of the search space

	$X_0$ (g L <sup>-1</sup> )	T (°C)	$S_0$ (g L <sup>-1</sup> )	N (rpm)
Lower bounds of variables	3	26	70	80
Upper bounds of variables	10	42	170	800

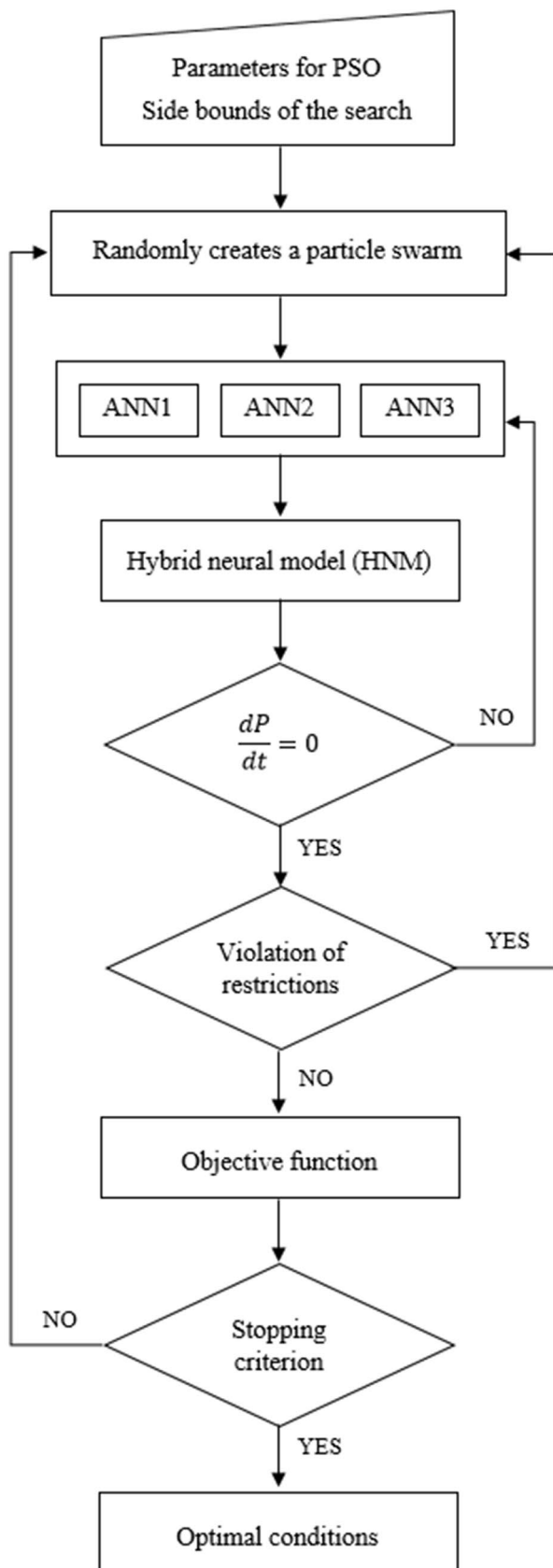


Fig. 4 Process optimization flowchart

Table 3 PSO parameters

Number of particles in the swarm	100
Particle Speed Factor	0.5
Maximum number of iterations	100
Minimum swarm pitch	10 <sup>-3</sup>
Minimum error of objective function value	10 <sup>-3</sup>

Table 4 ANNs architecture

Layer	Neurons number	Activation function
ANN1 ( $\mu_x$ )		
1st hidden layer	20	Hyperbolic tangent
2nd hidden layer	5	Hyperbolic tangent
Output layer	1	Linear
ANN2 ( $\mu_s$ )		
1st hidden layer	25	Hyperbolic tangent
2nd hidden layer	20	Hyperbolic tangent
3rd hidden layer	20	Hyperbolic tangent
Output layer	1	Linear
ANN3 ( $\mu_p$ )		
1st hidden layer	30	Hyperbolic tangent
2nd hidden layer	20	Hyperbolic tangent
3rd hidden layer	5	Hyperbolic tangent
Output layer	1	Linear

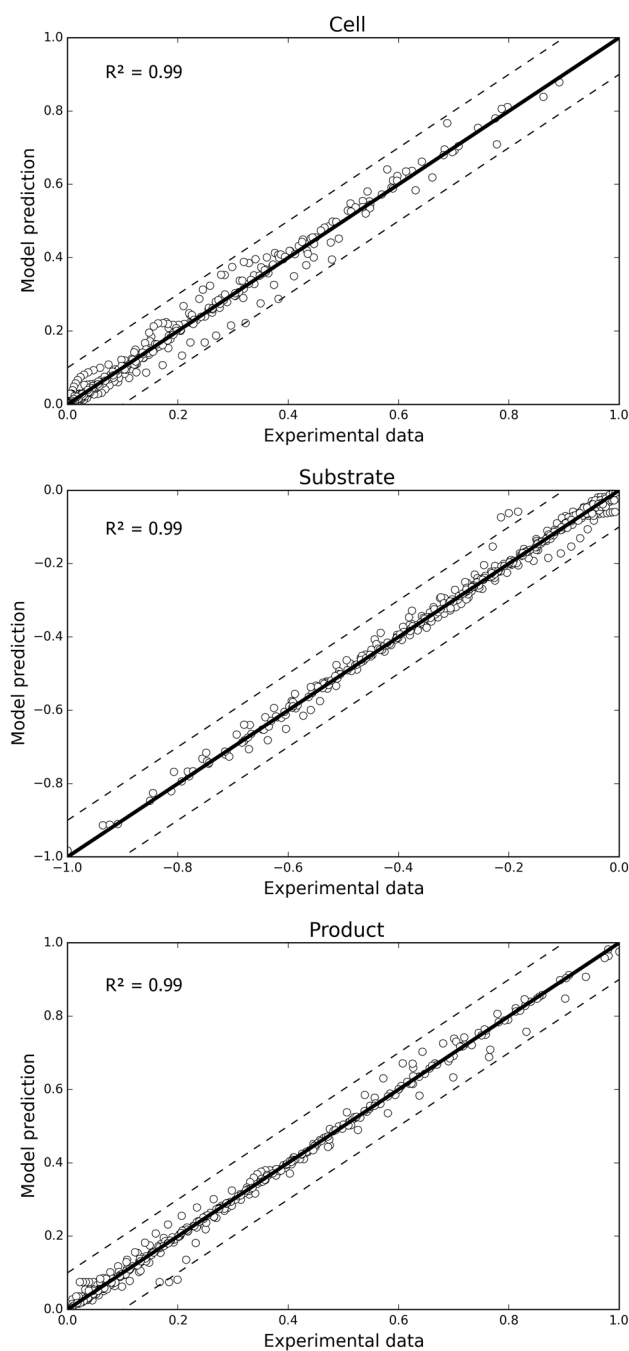
The objective function associates each point of the solutions space to a real number that allows measuring the response quality towards the initial objective. Perceiving that the individual function analysis would not be sufficient to determine the optimal operation point, the problem now becomes multiobjective. To simplify the optimization, the association of objective functions method was chosen, applying a geometric mean between efficiency and productivity (F2). Equations 22, 23 and 24 show the objective functions for efficiency (F1), productivity (F2) and the geometric mean (F3) for both, respectively:

$$F_1 = \max\left(\frac{P_f}{0.511 \cdot (S_0 - S_f)}\right) \tag{22}$$

$$F_2 = \max\left(\frac{P_f}{t_f}\right) \tag{23}$$

$$F_3 = \max\left(\sqrt{\left(\frac{P_f}{t_f}\right) \cdot \left(\frac{P_f}{0.511 \cdot (S_0 - S_f)}\right)}\right) \tag{24}$$





**Fig. 5** ANNs simulations: Comparison between RNA predictions and pseudo-experimental data (circles); (—) Bisector line (---) Specific error of 10%

**Table 5** Maximum and average errors of neural networks

	Maximum error (%)	Average error (%)
ANN1	10.1	1.4
ANN2	14.0	1.1
ANN3	13.5	1.6

Due to the reaction stoichiometry, the theoretical yield ( $Y_{P/S}$ ) for the ethanol production is  $0.511 \frac{\text{g}_{\text{ethanol}}}{\text{g}_{\text{glucose+fructose}}}$ , which limits the efficiency obtained by the HNM, avoiding non-realistic results. Thus, Eqs. 25 and 26 represent the inequality constraints used in the optimization:

$$S_f \geq 0 \quad (25)$$

$$\max \left( \frac{P_f}{S_0 - S_f} \right) < 0.511 \quad (26)$$

Since the objective functions and inequality constraints are defined, lateral restrictions of the search space of interest variables are defined and presented in Table 2.

The optimization algorithm is illustrated in Fig. 4, in which the PSO input parameters are necessary and the limits of the search space. Then, the HNM is solved (ODE system combined with ANN) and the objective functions are calculated for optimal fermentation time ( $dP/dt = 0$ ) testing the inequality constraints when necessary. Finally, when the objective function is maximized, the optimal operational conditions for the process are shown.

PSO algorithm used was obtained from PySwarms library and the optimization algorithm was implemented in Python/IPython Notebook version 2.7.8. The performance parameters used by the PSO algorithm were chosen to reconcile the simulation time and computational cost with the quality of the predictions and the parameters are shown in Table 3.

## Results and discussion

### Mathematical modeling

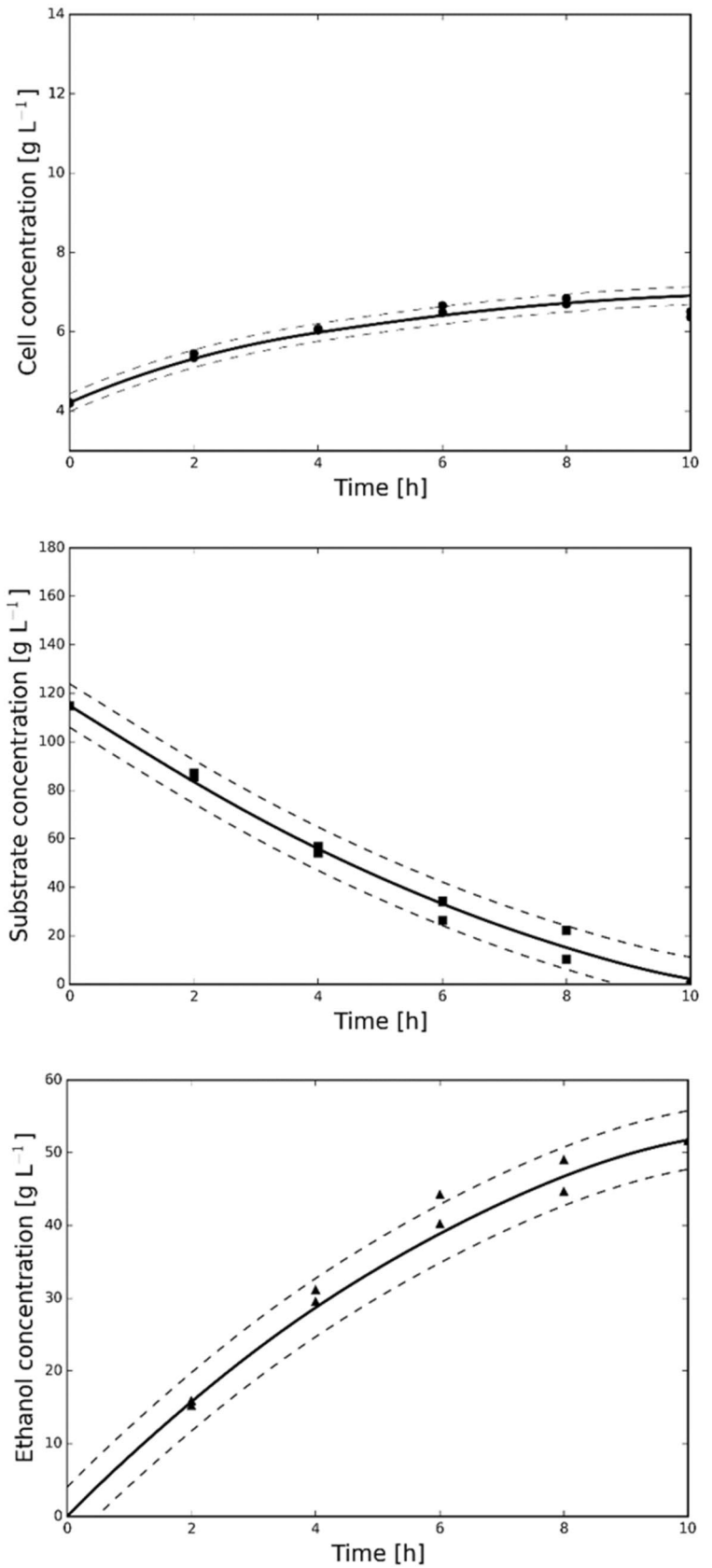
The kinetic model was determined in accordance with the architectures shown in Table 4. The specific rates of cells growth (ANN1), substrate consumption (ANN2) and product formation (ANN3) were estimated as a function of the instant concentration of substrate and cells, temperature and stirring speed. Several training rounds were conducted with different neural network architectures (1 to 4 hidden layers, 5 to 30 neurons per layer) of symmetrical type.

As it can be seen in Fig. 5, the simulated data are randomly spread in relation to the bisector line. These results indicate that the values predicted by the neural networks are satisfactory to represent the specific rates, as shown in Table 5.

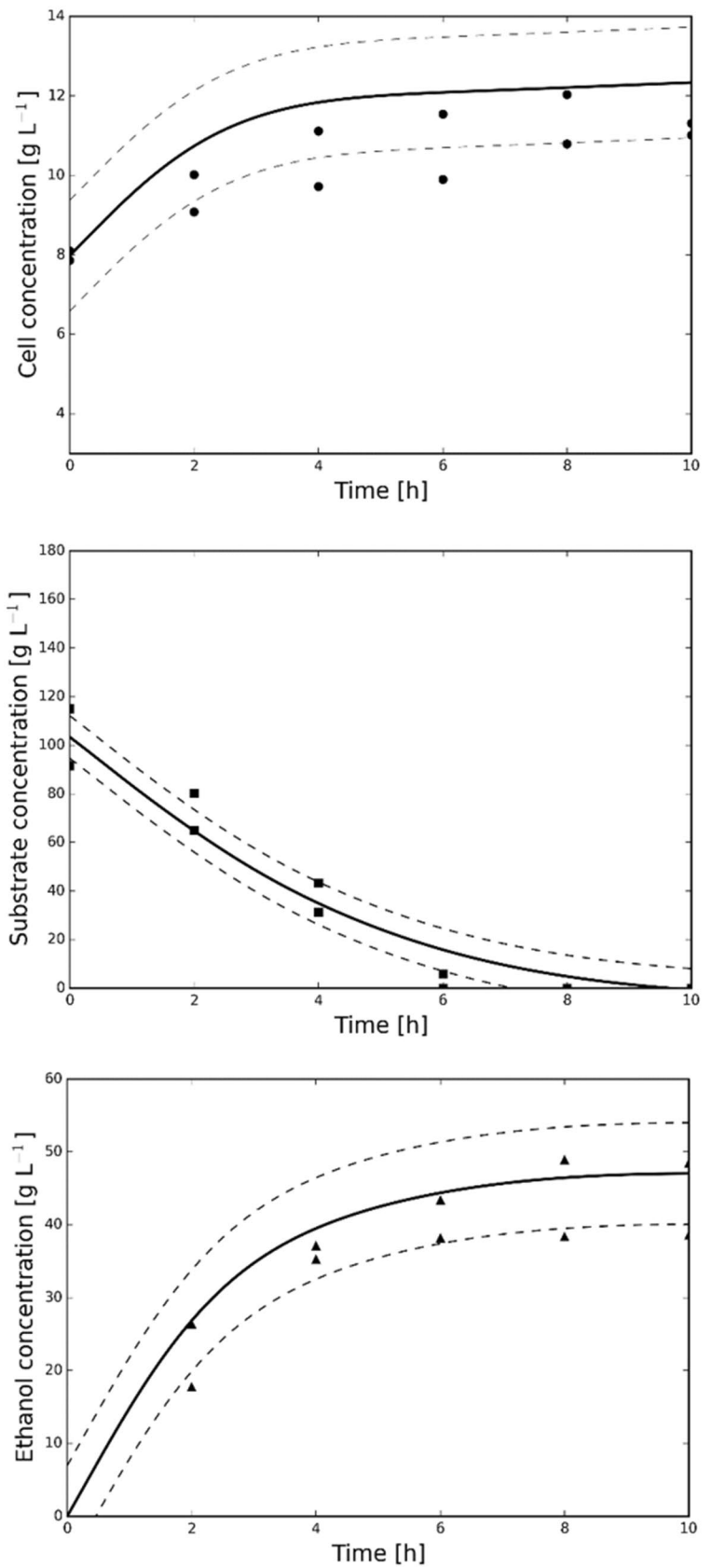
The maximum errors of the ANN are higher than the 10%, which is the standard error value accepted in bio-process [34]. However, the amount of data above this limit is irrelevant against the total set. Moreover, all the mean deviations calculated were below 2% (Table 5), indicating the quality of the ANN's prediction.



**Fig. 6** Experimental and simulated data for assay 4: (closed circles) cell concentration ( $\text{g L}^{-1}$ ); (closed square) substrate concentration ( $\text{g L}^{-1}$ ); (closed triangles) ethanol concentration ( $\text{g L}^{-1}$ ); (—) HNM, (---) Confidence interval with 90% significance level for the experimental data



**Fig. 7** HNM validation (Assay 13): (closed circles) cell concentration ( $\text{g L}^{-1}$ ); (closed square) substrate concentration ( $\text{g L}^{-1}$ ); (closed triangles) ethanol concentration ( $\text{g L}^{-1}$ ); (—) HNM, (---) Confidence interval with 90% significance level for the experimental data



**Table 6** HNM statistical analysis. The RSDs values represent the standard deviation of the data in comparison to the model.  $\epsilon_{exp}$  is the experimental error

		Biomass	Substrate	Product
Training	$\epsilon_{exp}(\%) >$	7	10.9	9.4
	RSD	0.6	8.0	3.3
Validation	$\epsilon_{exp}(\%) >$	8	10.8	8.2
	RSD	1.2	9.0	4.9

Table 6 shows the statistical analysis for the ANNs learning process, where the RSDs values represent the standard deviation of the data in comparison to the model. The apparent error calculated for the process was  $\epsilon_{exp} > 8.37\%$ , proving that the HNM is a more direct and efficient alternative to represent the process of ethanol production by *S. cerevisiae* CCA008 using cashew apple juice as substrate when compared to the mechanistic model previously reported by our group [7].

To verify the adequacy of the model to the experimental data used in the training of the ANN, experimental and simulated data for assay 4 (see operational conditions in Table 1) were compared (Fig. 6). The model is capable to describe the behavior of the fermentation, since the model not only represent the experimental points but also is contained in the confidence interval of the duplicates, with 90% significance level.

## Model validation

After confirming the HNM statistical adequacy, the next step was the model validation by checking if it fits a new dataset (Assay 13), not included in learning. Figure 7 shows the comparison between the simulated and experimental data of assay 13, where it is possible to see that the model was able to predict the experimental data. Therefore, this simulation represents the general validation of the proposed model, since it tests the capability of the HNM to properly predict the system behavior, as verified on the statistical analysis presented on Table 6.

The hybrid model validation returns higher RSD values and a lower apparent mean error when compared to the training step, reproducing satisfactorily the assays.

The performance of the HNM is comparable to the one obtained by the experimental data used to execute the phases of the training and test. Reasonable values of RSDs and apparent errors (Test F) were obtained, lower than experimental mean errors of the assays for biomass (11.8%), substrate (16.2%) and product (11.9%) concentrations.

Thus, the results obtained show that the ANNs are capable to adequately predict the system behavior, even when

**Table 7** Maximum values and optimal conditions of objective functions for efficiency ( $F_1$ ), productivity ( $F_2$ ) and the geometric mean between efficiency and productivity ( $F_3$ )

	$F_1=94.5\%$	$F_2=8.5$ ( $\text{g L}^{-1} \text{h}^{-1}$ )	$F_3=2.7$
$X_0$ ( $\text{g L}^{-1}$ )	6.0	6.0	5.8
$S_0$ ( $\text{g L}^{-1}$ )	101.5	148.2	127.0
T ( $^{\circ}\text{C}$ )	34.0	36.0	35.0
N (rpm)	110.0	107.0	111.0

operating in non-explored conditions. The HNM developed is particularly useful regarding the control and optimization of processes, providing trustworthy predictions for biotechnological applications.

## Optimization

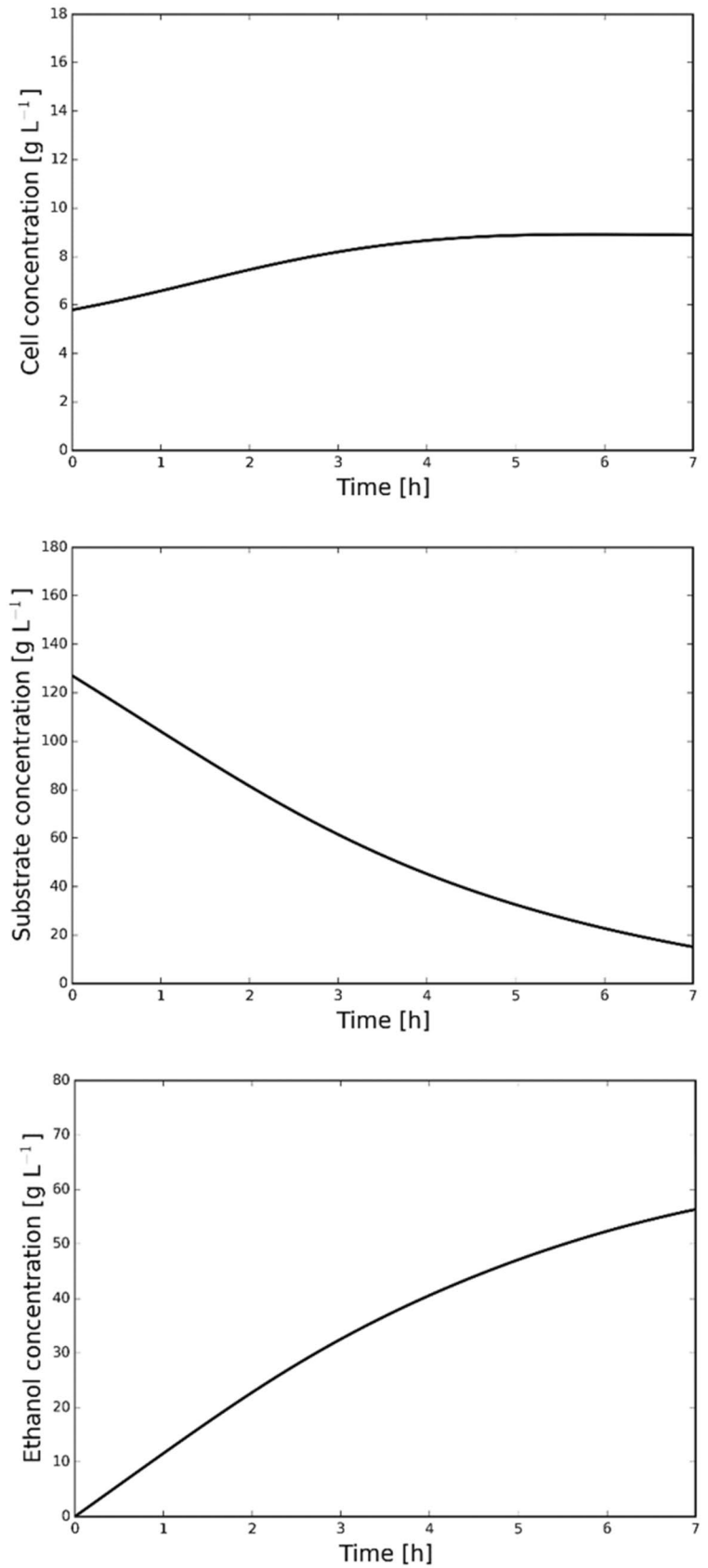
One of the biggest challenges of this work was the optimization of objective multimodal functions deriving from the inherent complexity of the biochemical processes and the number of optimization variables. Due to this, the analysis was restricted to a maximal time of 10 h of fermentation with low stirring speed, but enough to keep the process well mixed. This selection was made to minimize the energy consumption costs of the process. Table 7 presents the maximal values and operation conditions obtained by PSO to each objective function chosen.

The optimization of the objective function Efficiency ( $F_1$ ) is in accordance to the statistical model proposed and experimentally validated in a previous work [29], using the Monod model [6, 29]. This fact indirectly validates the optimization model HNM-PSO. In these conditions, the productivity was  $6.3 \text{ g L}^{-1}\text{h}^{-1}$  after 8 h of reaction.

High values of productivity ( $F_2$ ) were obtained ( $8.5 \text{ g L}^{-1}$ ) after 7 h of fermentation with an efficiency of 83%, which is also in accordance with the optimization of the mechanistic model proposed before [7]. Comparing  $F_1$  and  $F_2$ , the initial concentration of substrate and temperature are different. To maximize the productivity, high levels of substrate should be available in the reaction medium, which promotes a higher ethanol concentration. Temperature also improves rates, which has an impact on productivity.

Therefore, to evaluate a combination of productivity and efficiency, a multiobjective function ( $F_3$ ) was proposed.  $F_3$  was defined to equilibrate the relation between substrate consumption to produced ethanol and time of reaction. Figure 8 shows a simulation for the ideal condition of fermentation (initial substrate concentration  $127 \text{ g L}^{-1}$ , temperature  $35 \text{ }^{\circ}\text{C}$ , initial cells concentration  $5.8 \text{ g L}^{-1}$ , and stirring speed 111 rpm), achieving an efficiency and productivity of

**Fig. 8** HNM simulation for objective function optimization  $F_3$



91.5% and  $8.0 \text{ g L}^{-1} \text{ h}^{-1}$ , respectively, at approximately 7 h of bioprocess.

The optimal conditions of initial substrate concentration and temperature for function  $F_3$  are comprised in the interval of  $F_1$  and  $F_2$ , while the initial cell concentration and stirring remained almost constant. This fact was expected, because the objective function  $F_3$  captures the combined effect of the efficiency and productivity parameters.

## Conclusion

The HNM approach proved to be a very efficient tool to analyze and simulate the alcoholic fermentation of cashew apple juice by a flocculant yeast (*S. cerevisiae* CCA008). The algorithm appears as an alternative for biotechnological processes that presents a high level of complexity due to physicochemical and biochemical laws and evolved genetics. In this work, the use of ANN allowed to disregard a complex reaction mechanism. In addition, the HNM presented a high level of generality, allowing this model to be applied to other fermentation processes. Last but not least, the combination of advanced modeling techniques and optimization was successfully applied to maximize efficiency and reaction productivity. Generally speaking, the HMN-PSO optimization technique can be very useful in the optimization of bioprocesses, traditionally non-linear and involving multiple variables.

**Acknowledgements** The authors would like to thank CAPES, CNPq and FUNCAP (from Brazil) for the financial support that made this work possible. The authors also thank Professor Sandra Regina Ceccato Antonini for the yeast *Saccharomyces cerevisiae* CCA008 and Embrapa Agroindustria Tropical (Ceará, Brazil) for the Cashew Apple Juice.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Anderson ST (2012) The demand for ethanol as a gasoline substitute. *J Environ Econ Manage* 63(2):151–168. <https://doi.org/10.1016/j.jeem.2011.08.002>
- (IEA), I.E. A. (2015) Medium-term renewable energy market report. OECD/IEA, Paris
- International Energy Agency (2017) Renewables 2017: analysis and forecasts to 2022—executive summary. *J Qual Participat*. <https://doi.org/10.1073/pnas.0603395103>
- Neelakandan T, Usharani G (2009) Optimization and production of bioethanol from cashew Apple juice using immobilized yeast cells by *Saccharomyces cerevisiae*. *Am Eur J Sci Res* 4:85–88
- Karuppaya M, Sasikumar E, Viruthagiri T, Vijayagopal V (2010) Optimization of process variables using response surface methodology (RMS) for ethanol production from cashew apple juice by *Saccharomyces cerevisiae*. *Asian J Food Agro Ind* 3:462–473
- Pinheiro ÁDT, da Silva Pereira A, Barros EM, Antonini SRC, Cartaxo SJM, Rocha MVP, Gonçalves LRB (2017) Mathematical modeling of the ethanol fermentation of cashew apple juice by a flocculent yeast: the effect of initial substrate concentration and temperature. *Bioprocess Biosyst Eng* 40(8):1221–1235. <https://doi.org/10.1007/s00449-017-1782-2>
- Pereira AD, Pinheiro ÁD, Rocha MV, Gonçalves LR, Cartaxo SJ (2019) A new approach to model the influence of stirring intensity on ethanol production by a flocculant yeast grown on cashew apple juice. *Canad J Chem Eng* 97:1253–1262. <https://doi.org/10.1002/cjce.23419>
- Pinheiro ADT, Rocha MVP, Macedo GR, Gonçalves LRB (2008) Evaluation of cashew apple juice for the production of fuel ethanol. *Appl Biochem Biotechnol* 148(1–3):227–234. <https://doi.org/10.1007/s12010-007-8118-7>
- Marques WL, Raghavendran V, Ugarte B (2015) Sucrose and *Saccharomyces cerevisiae*: a relationship most sweet. *FEMS Yeast Res*. <https://doi.org/10.1093/femsyr/fov107>
- Rich JO, Anderson AM, Leathers TD, Bischoff KM, Liu S, Skory CD (2020) Microbial contamination of commercial corn-based fuel ethanol fermentations. *Bioresource Technol Rep* 11(April):100433. <https://doi.org/10.1016/j.biteb.2020.100433>
- Ruchala J, Kurylenko OO, Dmytruk KV, Sibirny AA (2020) Construction of advanced producers of first- and second-generation ethanol in *Saccharomyces cerevisiae* and selected species of non-conventional yeasts (*Scheffersomyces stipitidis*, *Ogataea polymorpha*). *J Ind Microbiol Biotechnol* 47(1):109–132. <https://doi.org/10.1007/s10295-019-02242-x>
- Gong X, Cao N, Du J, Tsao G (1999) Ethanol production from renewable resources. *Adv Biochem Eng Biotechnol* 65:207–242
- Lei J, Zhao X, Ge X, Bai F (1998) Ethanol tolerance and the variation of plasma membrane composition of yeast floc populations with different size distribution. *J Biotechnol* 2:35–43
- Ortiz-Muñiz B, Carvajal-Zarrabal O, Torrestiana-Sanchez B, Aguilar-Uscanga MG (2010) Kinetic study on ethanol production using *Saccharomyces cerevisiae* ITV-01 yeast isolated from sugar canemolasses. *J Chem Technol Biotechnol* 85(10):1361–1367. <https://doi.org/10.1002/jctb.2441>
- Domingues L, Vicente AA, Lima N, Teixeira JA (2000) Applications of yeast flocculation in biotechnological processes. *Biotechnol Bioprocess Eng* 5(4):288–305. <https://doi.org/10.1007/BF02942185>
- Volesky B, Votruba J (1992) Modeling and optimization of fermentation processes. Elsevier, Boca Raton
- Saraceno A, Curcio S, Calabrò V, Iorio G (2010) A hybrid neural approach to model batch fermentation of “ricotta cheese whey” to ethanol. *Comput Chem Eng* 34:1590–1596
- Feyo de Azevedo S, Dahm B, Oliveira FR (1997) Hybrid modeling of biochemical processes: a comparison with the conventional approach. *Comput Chem Eng* 21:751–756
- Thompson ML, Kramer MA (1994) Modeling chemical processes using prior knowledge and neural networks. *AIChE J* 40(8):1328–1340. <https://doi.org/10.1002/aic.690400806>
- Costa A, Henrique ASW, Alves T, Filho R, Lima EL (1999) A hybrid neural model for the optimization of fed-batch fermentations. *Braz J Chem Eng*. <https://doi.org/10.1590/S0104-66321999000100006>
- Silva RG, Cruz AJG, Hokka CO, Giordano RLC, Giordano RC (2000) A hybrid feedforward neural network model for the cephalosporin C production process. *Braz J Chem Eng* 17:587–598

22. Sivapathasekaran C, Mukherjee S, Ray A, Gupta A, Sen R (2010) Artificial neural network modeling and genetic algorithm based medium optimization for the improved production of marine biosurfactant. *Bioresour Technol* 101(8):2884–2887. <https://doi.org/10.1016/j.biortech.2009.09.093>
23. Dhanarajan G, Mandal M, Sen R (2014) A combined artificial neural network modeling-particle swarm optimization strategy for improved production of marine bacterial lipopeptide from food waste. *Biochem Eng J* 84:59–65. <https://doi.org/10.1016/j.bej.2014.01.002>
24. Bhattacharya S, Dineshkumar R, Dhanarajan G, Sen R, Mishra S (2017) Improvement of  $\epsilon$ -polylysine production by marine bacterium *Bacillus licheniformis* using artificial neural network modeling and particle swarm optimization technique. *Biochem Eng J* 126:8–15. <https://doi.org/10.1016/j.bej.2017.06.020>
25. Huang J, Mei L-H, Xia J (2006) Application of artificial neural network coupling particle swarm optimization algorithm to biocatalytic production of GABA. *Biotechnol Bioeng* 96(5):924–931. <https://doi.org/10.1002/bit>
26. Cockshott AR, Sullivan GR (2001) Improving the fermentation medium for Echinocandin B production part II: particle swarm optimization. *Process Biochem* 36(7):661–669. [https://doi.org/10.1016/S0032-9592\(00\)00261-2](https://doi.org/10.1016/S0032-9592(00)00261-2)
27. Serapião ABS (2009) PID Tuning By Swarm Optimization Strategies. In *Proceedings of the 8th Brazilian Conference on Dynamics*. Bauru-SP
28. Dineshkumar R, Dhanarajan G, Dash SK, Sen R (2015) An advanced hybrid medium optimization strategy for the enhanced productivity of lutein in *Chlorella minutissima*. *Algal Res* 7:24–32. <https://doi.org/10.1016/j.algal.2014.11.010>
29. Pinheiro AD, Barros EM, Rocha LA, da Rocha Ponte VM, de Macedo AC, Rocha MV, Gonçalves LR (2020) Optimization and scale-up of ethanol production by a flocculent yeast using cashew apple juice as feedstock. *Braz J Chem Eng*. <https://doi.org/10.1007/s43153-020-00068-0>
30. Wisselink HW, Toirkens MJ, Berriel MDRF, Winkler AA, Van Dijken JP, Pronk JT, Van Maris AJA (2007) Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl Environ Microbiol* 73(15):4881–4891. <https://doi.org/10.1128/AEM.00177-07>
31. Cleran Y, Thibault J, Cheruy A, Corrieu G (1991) Comparison of prediction performances between models obtained by the group method of data handling and neural networks for the alcoholic fermentation rate in enology. *J Ferment Bioeng* 71(5):356–362. [https://doi.org/10.1016/0922-338X\(91\)90350-P](https://doi.org/10.1016/0922-338X(91)90350-P)
32. Salehi M, Mohammadpour A, Mohammadi M, Aminghafari M (2018) A modified F-test for hypothesis testing in large-scale data. *J Biopharm Stat* 28(6):1078–1089. <https://doi.org/10.1080/10543406.2018.1436557>
33. Jiao B, Lian Z, Gu X (2008) A novel particle swarm optimization algorithm for permutation flow-shop scheduling to minimize makespan. *Chaos Sol Fract*. <https://doi.org/10.1016/j.chaos.2006.05.082>
34. Balat M, Balat H (2009) Recent trends in global production and utilization of bio-ethanol fuel. *Appl Energy* 86(11):2273–2282. <https://doi.org/10.1016/j.apenergy.2009.03.015>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.