

MONITORING OF BIOSURFACTANT PRODUCTION BY *Bacillus subtilis* USING BEET PEEL AS CULTURE MEDIUM VIA THE DEVELOPMENT OF A NEURAL SOFT-SENSOR IN AN ELECTRONIC SPREADSHEET

B. F. dos Santos^{1*}, A. P. R. Simiqueli², A. N. Ponezi³,
G. M. Pastore⁴ and A. M. F. Fileti¹

¹ School of Chemical Engineering, Department of Chemical Systems Engineering, University of Campinas (UNICAMP) 13083-852, Campinas/SP, Brazil. Corresponding author. Tel: +55 21 3527-2381; E-mail: bsantos@puc-rio.br, frattini@feq.unicamp.br.

² National Agricultural Laboratory, Brazilian Ministry of Agriculture, Livestock and Food Supply (LANAGRO-SP/MAPA), Campinas/SP, Brazil. Tel: +55 21 3527-2381; E-mail: simiqueli@gmail.com.

³ Research Center for Chemistry, Biology and Agriculture (CPQBA), University of Campinas (UNICAMP), 13081-970, Campinas/SP, Brazil. Tel: +55 21 3527-2381; E-mail: ponezi@cpqba.unicamp.br.

⁴ Department of Food Science, Faculty of Food Engineering, University of Campinas (UNICAMP), Campinas/ SP, Brazil. Tel: +55 21 3527-2381; E-mail: glaupast@fea.unicamp.br.

(Submitted: November 28, 2016 ; Revised: June 21, 2017 ; Accepted: July 20, 2017)

Abstract - This work investigated the combination of agitation and aeration conditions in a bench-bioreactor to identify the optimal biosurfactant production from substrate based on beet peel and glycerol from a biodiesel process. Thus, a central composite rotatable design (CCRD) and responses were evaluated by response surface methodology (RSM) modeling. The optimal operation values determined were 200 rpm (agitation) and 0.5 vvm (aeration), reaching values of 1931.2 mg/L of crude biosurfactant concentration and 28.37 mN/m of surface tension. For the development of a mathematical model based on an artificial neural network (ANN), the experimental data from each run (CCRD) of the bioreactor were used. The results indicated a topology of 6-6-1 neurons with an excellent predictive capacity of biosurfactant concentration: dispersion plot with R^2 of 0.995, and error criteria SSE of 0.31, MSE of 7.29×10^{-4} and RSME of 2.7×10^{-2} . A soft sensor was then designed in an electronic spreadsheet, computing the biosurfactant production from secondary measurements. Furthermore, the produced biosurfactant showed the ability to remediate oil spreading, evaluated through the appearance of clear zones on the surface of water covered with oil, and also from high emulsification indexes obtained on most of the solvents tested, such as toluene (~65%).

Keywords: Artificial neural network; Response surface modelling; Soft sensor; Beet peel; Oil spreading.

INTRODUCTION

The development of modern techniques allows the monitoring and control of bioprocesses, particularly when the measurements require lab work involving long processing times and delays. Industrial plants

are expected to show higher efficiencies and product quality through faster decision-making regarding process variables.

Usually in bioprocesses, the cost of downstream steps constitutes the major portion of the total production cost. The cost of biosurfactant production will vary

* Corresponding author: bsantos@puc-rio.br

depending on the biosurfactant application (food/pharmaceutical or pretroleum industry application). Determination of the most important variable during the production process, the biosurfactant concentration, would bring about cost reductions. However, this measurement is currently delayed because of the dependence on other lab methodologies. Its monitoring would thus enable the operator to make decisions in real time and consequently implement the appropriate actions towards a more efficient production rate.

Application of soft sensors for the estimation of hard-to-measure process values is extremely interesting for the process industry, where usually a large number of values are measured continuously (Rogina *et al.*, 2011). A soft sensor is a device that uses a mathematical model of the process, along with experimental measurements of one or more physical variables, to provide an in-line estimation of the product state and other variables of interest (Bosca and Fissore, 2011). Soft sensors have been studied extensively in recent years (Wang *et al.*, 2015; Kaneko and Funatsu, 2013; Ge *et al.*, 2014; Khatibisepehr *et al.*, 2013; Liu *et al.*, 2012). Various techniques have been proposed to design a soft sensor. Ni *et al.* (2014) developed a soft sensor based on localised adaptive techniques, where the preset threshold for updating the local region of the process was re-calculated. The algorithm was used in a moving window to provide the local learning framework for the online prediction with the soft sensor. Jin *et al.* (2015) proposed a multi-model adaptive soft sensor modelling method, based on the local learning framework and online support vector regression, for nonlinear time-variant batch processes.

Artificial intelligence provides mathematical models capable of predicting the dynamic behaviour of processes. Artificial neural networks (ANNs), an artificial intelligence technique, can learn from examples, are fault-tolerant, are able to deal with nonlinear problems, and, when trained, can predict quickly. Caldeira *et al.* (2011) studied the combined effect of incubation time and aspartic acid concentration on the predicted biomass concentrations, as well as the sporulation and anti-fungal activities of compounds, using *Bacillus amyloliquefaciens* as the study organism and ANNs modelling. Sivapathasekaran *et al.* (2010) developed a neural model to predict biosurfactant production, taking into account the four concentrations of the critical medium components glucose, urea, $MgSO_4$, and $SrCl_2$. Other ANN models were also developed by Singh *et al.* (2008), Rahimi *et al.* (2015), Fang *et al.* (2016), Albuquerque *et al.* (2008), Oroian (2015), Fang *et al.* (2016), Fu *et al.* (2013), Oladunjoye *et al.* (2016) and others.

The objectives of the current study include: identifying the factors allowing the optimal biosurfactant production by *Bacillus subtilis* in a batch bioreactor (7-L); in this case, an alternative waste

source substrate (glycerol from biodiesel production and beet peel) without any medium synthetic basal supplements was used; modelling the process using an ANN to predict crude biosurfactant concentrations; applying the intelligent soft sensor and validating it; and employing the produced biosurfactant for oil spreading. The findings allow the implementation of strategies for the maximisation of biosurfactant production and enable future innovations with the use of intelligent sensors that may permit, in the short-term, reduction of losses through the capacity to apply corrective actions on batch processes in real time.

MATERIALS AND METHODS

Inoculum Preparation and Standardisation

The microorganism used in fermentation for biosurfactant production was *Bacillus subtilis*, available from the microorganism bank of the Research Centre for Chemistry, Biology and Agriculture (CPQBA/Unicamp). The medium used for the preparation of the inoculums was nutrient broth. The microorganism was initially added for adaptation to 15 mL of nutrient broth (pre-inoculum), into a 50-mL Erlenmeyer flask, and was incubated in an orbital shaker for 24 h at 37 °C. Then, the inoculum (150 mL of sterile nutrient broth in a 250-mL Erlenmeyer flask) received the pre-inoculum culture and was incubated in an orbital shaker following the same conditions. The standardisation of the inoculum was performed using nutrient broth adjusted in a spectrophotometer Genesys 10S (625-nm wavelength) to an absorbance range from 0.08 to 0.1, according to McFarland's method.

Measurement of Glucose Concentration

To determine the glucose concentration, a laboratory biochemical glucose test kit was used. 10 μ L of the supernatant sample was collected and added to 1 μ L of the reagent solution (supplied by the manufacturer) and then immersed in a water bath for 5 minutes at 37 °C. The samples were taken to the spectrophotometer at 505 nm and then absorbance was read immediately. From the calibration curve, glucose concentration was found.

Measurement of Biomass and Analysis of Surface Tension

Cell growth was determined by measuring the optical density of samples, using a UV-VIS spectrophotometer (Genesys 10S) at 600 nm. Biomass concentration was confirmed by using the dry weight (g/L) after removal of cells by centrifugation (10 000 rpm, 10 min, 4 °C) from supernatant.

Surface tension of the cell-free broths was measured by the plate method, using a digital tensiometer model K12 (Krüss GmbH, German) at 20°C. The surface tension determination was replicated at least two times.

Recovery of the Crude Biosurfactant

To recover the biosurfactant, the cell-free supernatant from the fermentation culture was subjected to acid precipitation (pH 2.0) with 1 M HCl and incubated for 24 h at 7 °C. The resultant solution was maintained at rest for 12 h to allow the biosurfactant to settle. The precipitate was collected by centrifugation (10 000 rpm, 17 min, 2 °C). The supernatant was then discarded and the precipitate was washed with acidified water and stored. All assays were performed in duplicates.

Application of Crude Biosurfactant in Oil Spreading

According to the method described by Mouafi et al. (2016), oil spreading was evaluated by adding 20 mL of distilled water to a Petri dish, followed by the addition of 50 µL of oil on its surface. Then, 40 µL of cell-free culture broth was dropped on top of the crude oil surface and the diameter of the clear zone produced on the oil surface was assessed and compared to a negative control (culture medium).

Application of Crude Biosurfactant in Emulsification of Solvents

To analyze the emulsification (E_{24}) in different solvents, the fermented medium was centrifuged (10 000 rpm, 10 min, 4 °C) to obtain cell-free supernatant. 2 mL were collected from the supernatant to mix with 2 mL of solvents (diesel oil, gasoline, soy oil, heptane, hexane, toluene) in test tubes. It was stirred by vortexing for 2 min and the mixture was allowed to stand for 24 h. The E_{24} was calculated by Equation (1):

$$E_{24} = \frac{\text{Height of the emulsion formed in 24h}}{\text{Height of the solution}} \quad (1)$$

Response Surface Methodology (RSM) Modelling

A 2² full factorial design with three assays at the central point and four axial points, totalling 11 runs were performed to optimize bioreactor conditions. Assays were

carried out to verify the combined effect of the agitation (X_1) and aeration (X_2) rates on the crude biosurfactant concentrations and surface tension reduction ratios during the biosurfactant production process. The experimental design was developed and analysed with the help of the *STATISTICA 7* software; both the range and level of each factor are given in Table 1.

The experiments were conducted in a 7-L bioreactor (Bioflow 310 New Brunswick Scientific, USA) equipped with standard probes for temperature and dissolved oxygen, as well as auxiliary equipment, allowing a fermentation volume of 4 L, Figure 1. The composition of the alternative medium was determined as previously described Santos et al. (2014): 6% (v/v) of glycerol from biodiesel production and 7.5 % (v/v) of beet peel from restaurants. The pH was adjusted to around 7, using NaOH or HCl. Each batch was run for 24 h and the sampling performed every 3 h.

A second-order polynomial regression (Equation 2) was used in this study for the estimation of all main and joint effects while central and axial points were for providing replication and curvature terms in the model.

$$y = \beta_0 + \sum_j \beta_j x_j + \sum_{i < j} \beta_{ij} x_i x_j + \sum_j \beta_{jj} x_j^2 + e \quad (2)$$

where x_1 and x_2 are the input variables which are known to affect the response y and $\beta_0, \beta_j, \beta_{ij}, \beta_{jj}$ are the relevant constants of the effects. Analysis of variance (ANOVA) was evaluated to validate the RSM model.

Modelling with ANN and Soft Sensor Design

The adopted ANNs were based on the multilayer perceptron, and the training algorithm used was

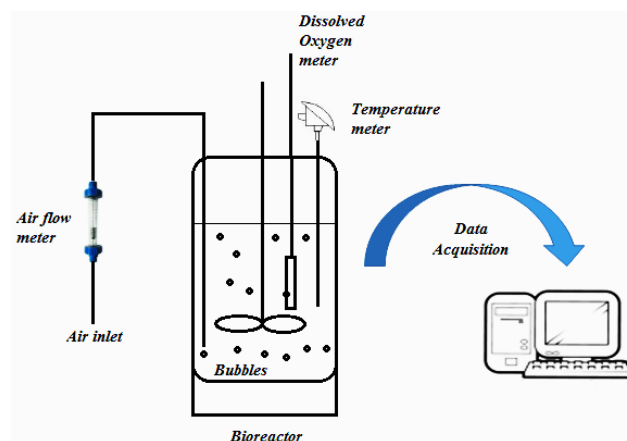


Figure 1. Apparatus scheme used in the experiments representing bioreactor.

Table 1. Values used in the experimental central composite rotatable design (CCRD).

Independent variables	Code	Experimental domain (% v/v)				
		-1.41	-1	0	+1	+1.41
Agitation (rpm)	X_1	100	129	200	270	300
Aeration (vvm)	X_2	0.50	0.86	1.75	2.60	3.00

the Levenberg-Marquardt based backpropagation algorithm, in conjunction with Bayesian regularisation. In the present study, the quantity of neurons in the input layer (any of this variables: microbial concentration, glucose concentration, dissolved oxygen concentration, surface tension, or dissolved surface tension in 10× and 100×) and the hidden layer were defined by parameters combination. This strategy may allow determining the best configurations, using one neuron in the output layer (crude biosurfactant concentration) in all the simulations.

Each neuron in the hidden layer calculates the summation of weighted outputs from previous neurons (n) and transforms it by the following activation function, to satisfy the feed forward sign. The activation functions used were (Equations 3 or 4):

$$\log \operatorname{sig}(n) = \frac{1}{1 + \exp(-n)} \quad (3)$$

$$\tan \operatorname{sig}(n) = \frac{2}{1 + \exp(-n)} - 1 \quad (4)$$

Thus, the simulations were made by blocks, where each scenario was investigated to reach the best values of correlation coefficient (R^2), parameters of error (sum of square error (SSE), mean square error (MSE) and root mean square error (RMSE)), (Equations 5, 6, 7), as well as linear and angular coefficients (A and B), from the dispersion plot.

$$SSE = \sum_{i=1}^n (y_{\text{observed}} - y_{\text{predicted}})^2 \quad (5)$$

$$MSE = \sum_{i=1}^n \frac{(y_{\text{observed}} - y_{\text{predicted}})^2}{n} \quad (6)$$

$$RMSE = \sqrt{MSE} \quad (7)$$

Once the neural network was created, it was trained to accurately model the given dynamic, using the experimental data in MATLAB version 7b (Mathworks Inc., US). All the data sets required to train, validate, and test the neural soft sensors were obtained from biosurfactant production experiments carried out using the bioreactor.

When the ANN model had been validated offline, the structure of the soft sensor was made in the electronic spreadsheet (Excel, Microsoft Office). In this way, the designed soft sensor received the information in the input layer and automatically calculated the output and the values were evaluated.

RESULTS AND DISCUSSION

The procedures for the preparation and standardisation of the inoculum ensured reliable fermentation experiments, making sure that viable cells reached the same final concentrations in the different fermentation cultures. Biosurfactant production was reported earlier (Santos *et al.*, 2014), using glycerol from a biodiesel process and beet peel from restaurants as culture broth, using an orbital shaker. Their results demonstrated that maximum crude biosurfactant concentration was 550 ± 85 mg/L in Erlenmeyer flasks with values of 6% (v/v) glycerol and 7.5 % (v/v) beet peel, established by CCRD, suggesting the feasibility of a batch bioreactor assay.

Central Composite Rotatable Design (CCRD)

The experiments in the batch bioreactor were carried out to provide the best conditions for fermentation by *Bacillus subtilis* and subsequent biosurfactant production, using only waste sources (glycerol from a biodiesel process and beet peel from restaurants). The influence of each factor on the biosurfactant production at determined levels is shown by the responses in the crude biosurfactant concentrations and surface tension reduction ratios (Table 2). Thus, the scenario of possibilities among the variables were performed in the experimental matrix of 2^2 central composite rotatable design (CCRD) in addition to three central points and four axial points, totalling 11 runs.

The CCRD matrix indicated biosurfactant production for all the studied factor combinations (different levels of agitation and aeration). Therefore, these levels yield different concentrations of crude biosurfactant and surface tension reduction ratios. When comparing these results with those previously obtained Santos *et al.* (2014), for which the optimum concentrations in the culture broth (based only on waste sources) were developed in Erlenmeyer flasks, an improvement was observed in crude biosurfactant concentrations upward of 3-fold. Pereira *et al.* (2013) studied biosurfactant production by *Bacillus subtilis* using sucrose as synthetic carbon source (similar to the waste in this study) and reached $2158.5 \pm 240,1$ mg/L of crude biosurfactant concentration; this value is near that observed in this study.

Based on these results, the matrix was evaluated, enabling the calculation of a regression coefficient with a p-value limit of 0.1. Tables 3 and 4 show the scenarios of the regression coefficient for each one of the factors, represented by the linear and quadratic agitation (X_1), the linear and quadratic aeration (X_2), and their interactions, $X_1 \times X_2$. The greatest factor value was the quadratic agitation (X_1^2) in both models, revealing negative influences. Fermentation processes are quite difficult to study, owing to sources of variation,

Table 2. Experimental central composite rotatable design (CCRD) runs, corresponding crude biosurfactant concentrations, and surface tension reduction ratios.

Run number	Factors		Responses	
	Agitation (rpm) X_1	Aeration (vvm) X_2	Crude Biosurfactant concentrations (mg/L)	Surface tension reduction ratios (%)
1	129(-1)	0.86(-1)	855.3±92.0	31.9±5.0
2	270(+1)	0.86(-1)	850.6±6.0	37.7±0.0
3	129(-1)	2.6(+1)	800.2±30.0	37.7±0.0
4	270(+1)	2.6(+1)	792.9±70.0	34.8±3.0
5	100(-1.41)	1.75(0)	550.5±80.0	35.5±2.0
6	300(+1.41)	1.75(0)	203.2±10.0	8.3±0.5
7	200(0)	0.5(-1.41)	1931.2±52.0	46.4±1.5
8	200(0)	3.0(+1.41)	920.6±104.0	40.8±0.0
9	200(0)	1.75(0)	1041.6±8.0	42.6±0.5
10	200(0)	1.75(0)	1028.7±33.0	40.0±0.9
11	200(0)	1.75(0)	1032.8±38.0	41.8±0.0

Table 3. Regression coefficient for the crude biosurfactant concentration model.

Factor	Regression coefficient	Std Err.	t(5)	p-value
Mean/Interaction	1034.41	136.80	7.56	<0.0001
$X_1(L)$	-75.47	83.77	-0.90	0.40
$X_1(Q)$	-347.93	99.71	-3.48	0.01
$X_2(L)$	-180.15	83.77	-2.15	0.08
$X_2(Q)$	176.61	99.77	1.77	0.13
$X_1 \times X_2$	-0.64	118.47	0.00	0.99

Table 4. Regression coefficient for the surface tension reduction ratio model.

Factor	Regression coefficient	Std Err.	t(5)	p-value
Mean/Interc	41.32	4.04	10.22	<0.0001
$X_1(L)$	-4.44	2.47	-1.79	0.13
$X_1(Q)$	-9.01	2.94	-3.06	0.02
$X_2(L)$	-0.63	2.47	-0.25	0.80
$X_2(Q)$	1.86	2.94	0.633	0.55
$X_1 \times X_2$	-2.17	3.49	-0.62	0.56

especially when using waste substrates. By analysing the experimental results through RSM, it was possible to get an empirical model for the best response.

A statistical model was developed (Equations 8 and 9 referring to Table 3 and 4, respectively), which describes the behaviour of crude biosurfactant concentration (mg/L) and surface tension reduction ratio (%), explaining the mathematical relationship between the selected predictors and the response. As seen in CCRD assays, the responses show large deviation due to errors occurred in the data acquisition system (accepted in bioprocesses). This justifies studying the process at a confidence level of 90% (p-value of 0.1). For practical purposes, the model was adjusted through re-parameterisation to make it as simple as possible, with the fewest possible parameters, without losing accuracy.

$$\text{Crude biosurfactant (mg/L)} = 1034.41 - 347.93X_1^2 - 108.15X_2 + 176.61X_2^2 \quad (8)$$

$$\text{Surface tension reduction ratio (\%)} = 41.32 - 4.44X_1 - 9.01X_1^2 \quad (9)$$

Analysis of variance (ANOVA) was performed to ensure confidence of the generated second-order

model to a level of 90 %. Table 5 shows the ANOVA results for those models, explaining 83.05 % of the crude biosurfactant concentration and surface tension reduction ratio behaviour, with F_{cal} larger than F_{tab} in both cases. These results are promising, because these polynomial models can enable the prediction of the responses in valid regions. Figure 2 presents the three-dimensional response surface and contour curve, which are the graphical representations of the equations entered in the optimisation tracks.

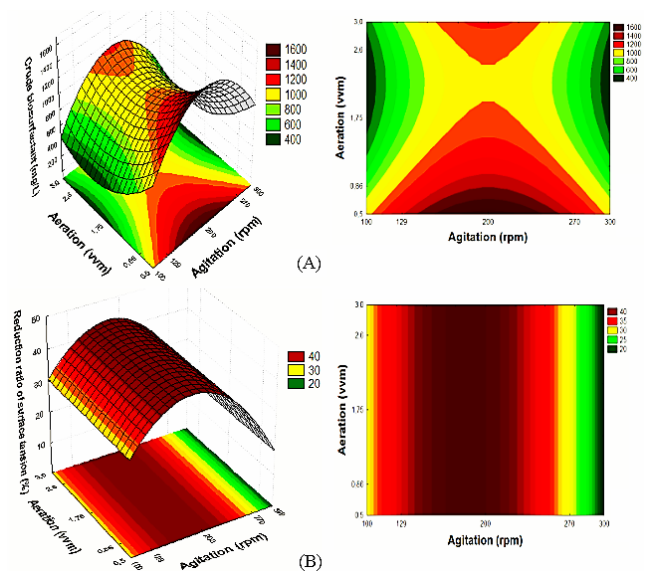
According to each plot, the effect of two independent variables was diversified within the experimental range for optimal values. Thus, it was possible to suggest the type of interactions between variables and establish the best operation conditions. Agitation was set at the middle level, 200 rpm, and aeration was set at the lower level, 0.5 vvm. For validation, experiments were performed with the optimised conditions and values were obtained of 1780.60 ± 80.50 mg/L for crude biosurfactant concentration and 45.03 ± 4.80 % for surface tension reduction ratio (~ 28 mN/m), showing good agreement with the models. When compared with the previous study, the crude biosurfactant concentration was enhanced three-fold, confirming

Table 5. Analysis of variance (ANOVA) for the crude biosurfactant concentration and surface tension reduction ratio.

Source of variation	d.f.		SS		MS		Fcal	
	CB	ST	CB	ST	CB	ST	CB	ST
Regression	3	2	1119427	723.6	373142.33	361.8	8.0	10.10
Residual	7	8	326327	286.7	46618.14	35.8		
Total	10	10	1445754					

Crude biosurfactant (CB): $F_{3;7;0,10} = 3.07$; correlation coefficient: $R^2 = 83.96$.

Surface tension reduction ratio (ST): $F_{2;8;0,10} = 3.11$; correlation coefficient: $R^2 = 75.75$.

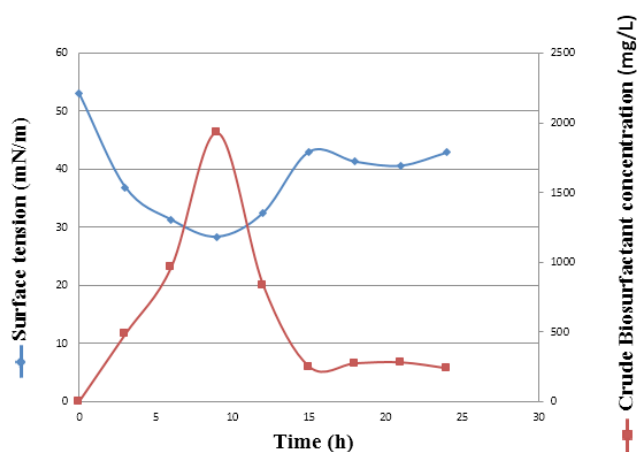
**Figure 2.** Response surface and curve contour: (A) crude biosurfactant concentration; and (B) surface tension reduction ratio.

the importance of the optimisation of operating parameters, such as agitation and aeration.

It is important to understand the experimental characteristics of this process. All responses used in the CCRD corresponded to high values reached in batch (around 9h of 24h). The 24 hours for batch duration were chosen by observing the behaviour of the process using only waste in fermentation.

The biosurfactant production was made evident in the first 9h by the accentuated fall in surface tension from 53 to 28.37 mN/m and increase of crude biosurfactant concentration to 1931.2 mg/L. After this period, the production of biosurfactant was reduced in the medium, resulting in an increase in surface tension. Barros *et al.* (2008) showed similar behaviour that could be associated with the exponential growth and, consequently, with the synthesis of compounds of interest. The substrate was consumed over time and the biosurfactant produced may have been used to facilitate the maintenance of the microorganism for survival in the fermentation. This theory is reinforced when the graph of kinetics of microbial concentration (g/L), glucose concentration (g/L) and dissolved oxygen was analyzed (Figure 4).

It is possible to observe in Figure 3, that as, biosurfactant concentration increases in the culture

**Figure 3.** Kinetics of biosurfactant production: Surface tension (mN/m) and crude biosurfactant concentration (mg/L).

medium, lower values of surface tension are achieved up to a minimum value.

The microbial concentration increases in fermentation as the dissolved oxygen and glucose concentration fall. The consumption of micronutrients results in a critical medium (approximately 11h) for the microorganism and the microbial growth was not harmed because of the presence of biosurfactant. Microorganisms have adapted to living in these challenging environments, surviving as free-living as is mentioned by Kennedy *et al.* (2011). Oliveira *et al.* (2013) working with *Bacillus subtilis*, grown in clarified cashew apple juice, report that the biosurfactant tested exhibited excellent surface activity, reducing the tension of water from 58 to 30 mN/m in 24h and then increasing from 30 to 45 mN/m in 72h. As biosurfactant production increases the surface tension of the medium decreases up to a certain value and then becomes almost constant due to the interface saturation with the biosurfactant molecules. The authors attributed the contradiction to the low purity of the preparation, since it was extracted from the culture medium.

It was noted that the glucose concentration increased from time 0 to 3 hours. The reason for this phenomenon is the primary carbon source, sucrose. Increasing microbial concentration also increases enzyme production to convert sucrose into glucose and fructose. Thus, there is an increase of glucose concentration to 5.5 g/L at 3

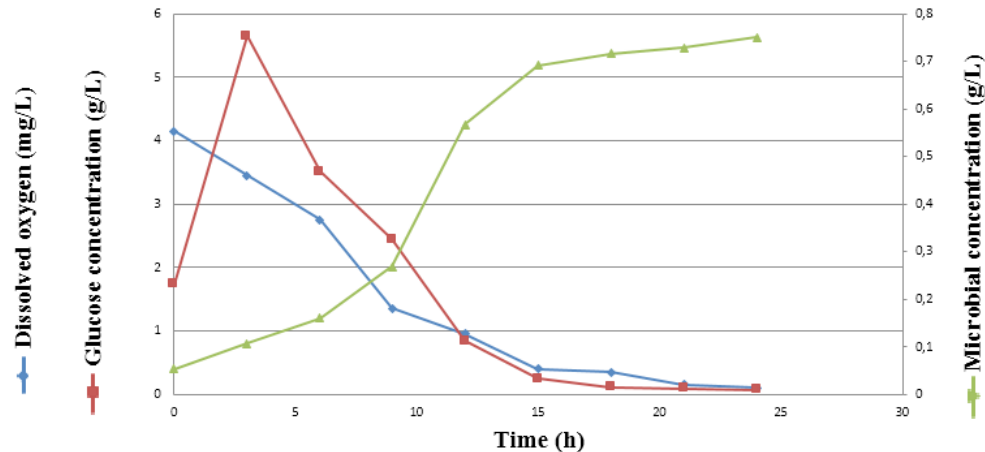


Figure 4. Kinetics of dissolved oxygen (mg/L), glucose concentration (g/L) and microbial concentration (g/L).

hours-time. From this moment, the concentration starts decreasing until vanishing.

Considering the relationship of the main variables, the process displays a strong nonlinear behavior, difficult to model from first principles phenomena (Figure 5). This fact justifies the use of artificial neural network modeling.

Design of a Soft Sensor with Artificial Neural Network (ANN) Modelling

The first approach to design a soft sensor is to develop a mathematic model. This study uses an

artificial neural network (ANN) as the inferential model. This artificial intelligence technique needs process examples to learn its behaviour. Therefore, the database from the eight runs (dynamic behaviour of available points) in the batch bioreactor was used to train and test the ANN (corresponding to 9 experimental points).

The data sets required to train and test the neural soft sensors were obtained from biosurfactant production experiments carried out using substrate based on combinations of glycerol from a biodiesel process and peel beet in a 7-L bioreactor, under different agitation and aeration conditions specified in the CCRD. The

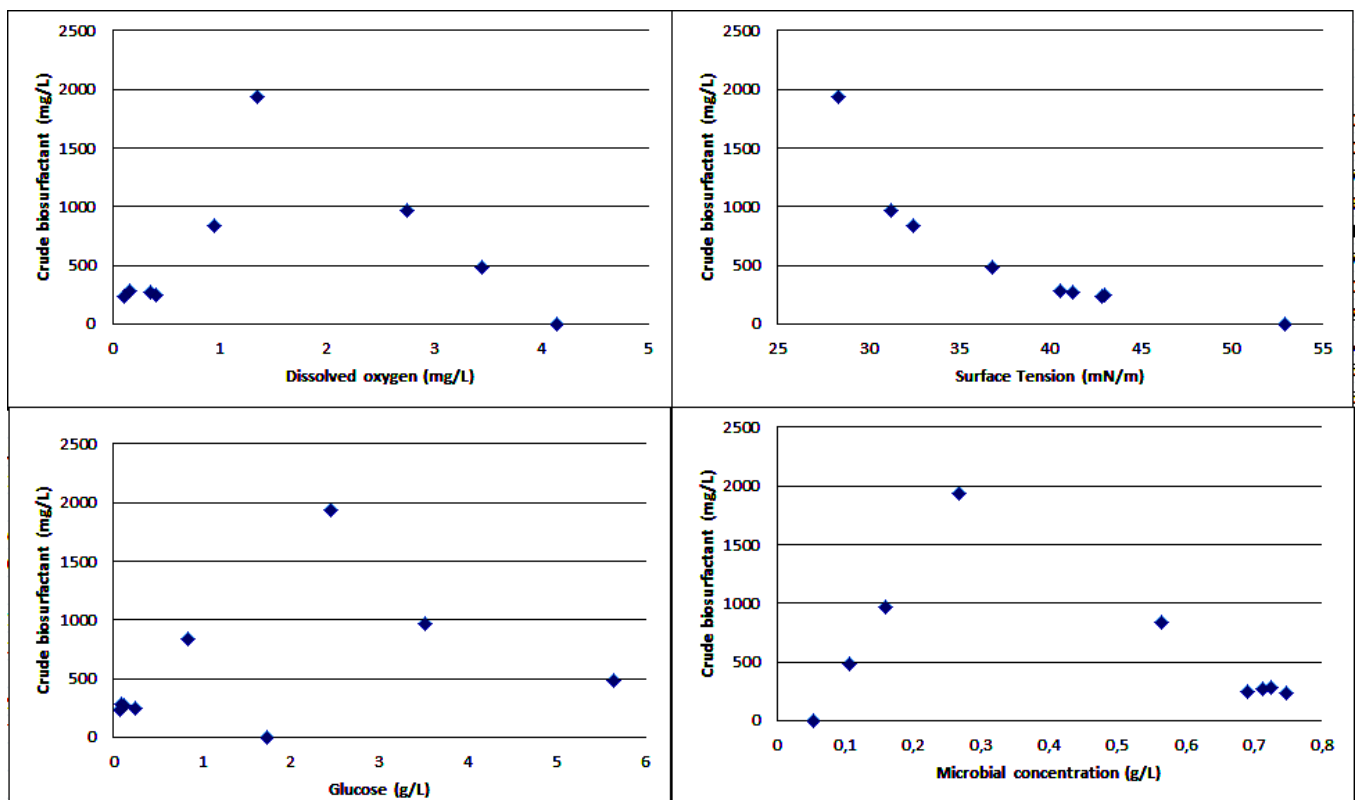


Figure 5. Relationship between biosurfactant production (crude biosurfactant) and dissolved oxygen, surface tension, glucose and microbial concentration.

experimental data sets were smoothed and expanded by interpolation using piecewise cubic spline smoothing. The smoothing expanded each data set (from 8 runs) from 9 to 25 data. As each data set was performed in duplicate, then the database transitioned to having 600 data from the original 216 data points.

The variables assessed from the bioreactor through sampling were microbial concentration (MC), glucose concentration (GC), dissolved oxygen concentration (OD), surface tension (ST), dissolved surface tension ($10 \times (ST-1)$ and $100 \times (ST-2)$), and crude biosurfactant concentration; sampling was carried out every 3 h. The choice of the variables was according to availability and time for measurement, with some of them being measured by analytical methods (glucose concentration). The microbial concentration was measured using optical density and a gravimetric method as well. The optical density measurement was chosen to provide ANN input data because it is faster than the gravimetric methodology. It is important to highlight that the crude biosurfactant concentration

(output neuron in the ANN model) variable takes about 48 h to measure. The strategy was to build a soft sensor that could analyse this variable in real time. The database was separated into two parts: 75% for training and 25% for testing.

The simulation was performed to seek the best scenario of neural topologies. Table 6 shows the simulations in blocks with combinations of neurons in the input layer, activation functions and neurons in the output layer.

For each simulation by blocks, neurons in the hidden layer were diversified as for the neuron quantity in the input layer. The activation functions used for the hidden layer were logistic (logsig) and hyperbolic tangent (tansig), while linear for the output layer. For all cases, the learning algorithm was Levenberg-Marquardt with Bayesian regularisation (trainbr).

The analysis of the simulation is presented in Table 7. The values of correlation coefficients (R^2), error parameters (SSE, MSE, and RMSE), and linear and angular coefficients (from the equation of the line),

Table 6. Scenarios of ANN topologies.

Blocks	Input layer						Activation function	Output layer
	MC*	GC*	OD*	ST*	ST ¹ *	ST ² *		
I	X	X	X	-	-	-	Logsig	CB*
II	X	X	X	-	-	-	Tansig	CB
III	X	X	X	X	-	-	Logsig	CB
IV	X	X	X	X	-	-	Tansig	CB
V	X	X	X	X	X	X	Logsig	CB
VI	X	X	X	X	X	X	Tansig	CB

*MC - Microbial concentration; GC - Glucose concentration; OD - Dissolved oxygen concentration; ST - Surface tension; ST¹ - Dissolved surface tension $10 \times$; ST² - Dissolved surface tension $100 \times$; CB - Crude biosurfactant concentration.

Table 7. Performances of simulation blocks.

Block	Number of neurons in hidden layer	R^2	SSE*	MSE*	RMSE*	A	B
I	3	0.627	13.69	3.22E-02	1.79E-01	1.3E02	0.59
	6	0.555	9.13	2.14E-02	1.46E-01	1.7E02	0.53
	8	0.578	7.17	1.64E-02	1.30E-01	1.4E02	0.62
	10	0.651	7.30	1.71E-02	1.31E-01	1.6E02	0.60
II	3	0.737	13.60	3.2E-02	1.78E-01	94.0	0.66
	6	0.545	9.10	2.14E-02	1.46E-01	1.7E02	0.52
	8	0.645	8.05	1.89E-02	1.37E-01	1.3E02	0.64
III	10	0.605	7.80	1.83E-02	1.35E-01	1.6E02	0.57
	4	0.984	0.376	8.84E-04	2.97E-02	8.7	1.00
	6	0.978	0.286	6.73E-04	2.60E-02	61.0	0.90
IV	10	0.933	0.166	3.90E-04	1.97E-02	52.0	0.88
	12	0.945	0.152	3.57E-04	1.89E-02	51.0	0.92
	4	0.984	0.376	8.84E-04	2.97E-02	8.7	1.00
	6	0.978	0.286	6.72E-04	2.59E-02	61.0	0.90
V	10	0.933	0.166	3.90E-04	1.97E-02	52.0	0.88
	12	0.978	0.247	5.81E-04	2.41E-02	34.0	0.92
	6	0.967	0.185	4.11E-03	6.41E-02	21.0	1.00
	10	0.956	0.120	2.82E-04	1.68E-02	56.0	0.89
VI	14	0.946	0.124	2.91E-04	1.70E-02	0.56	1.10
	16	0.975	0.078	1.83E-04	1.35E-02	27.0	0.96
	6	0.995	0.310	7.29E-04	2.70E-02	-0.38	0.99
	10	0.994	0.202	4.75E-04	2.18E-02	4.2	1.0
VI	14	0.993	0.228	5.36E-04	2.31E-02	14.0	1.0
	16	0.996	0.158	3.71E-04	1.92E-02	-1.6	1.0

refer to testing data. Ideally, the values of R^2 should be close to 1, the values of error indices (SSE, MSE, RMSE) should be close to 0, and linear and angular coefficients (A and B) should be 0 and 1, respectively. These results indicate that the model describes well the relationship between the input and output parameters.

The simulations with the neural model were quite different depending on each block. Block I was the worst simulation scenario, even with high amounts of effective parameters. The R^2 did not reach values higher than 0.6, the sum of square error (SSE) decreased from 13.69 to 7.30 (from low- to high-parameter values), and so did the mean square error (MSE; from $3.22E-02$ to $1.71E-02$) and the root mean square error (RMSE; from $1.79E-01$ to $1.31E-01$). However, the linear and angular coefficients (A and B) were balanced.

Block II was not different from block I, because the only difference between both scenarios was the change in the activation function, from *logsig* to *tansig*.

For blocks III and IV, we observed improved values. The R^2 reached values around 0.9 and the error parameters (SSE, MSE, and RMSE) decreased markedly (around 0.1, $3.9E-04$, and $2E-02$, respectively). The linear coefficient reached values lower than those observed before, but the angular coefficient showed values close to one (better than seen before).

The last blocks of simulations (V and VI) gave the best results, with nearly all the values showing improvement. More particularly, block VI showed the best simulation outcomes. The condition with six neurons in the hidden layer was the best scenario of all, showing a R^2 of 0.995 (hence very close to one), the lowest error values (SSE of 0.31, MSE of $7.29E-04$, and RMSE of $2.70E-02$), and linear and angular coefficients of -0.38 and 0.99, respectively.

The performance of the ANN model for the prediction of crude biosurfactant concentration in block VI (with six neurons in the input layer) was also evaluated through a graph of dispersion (Figure 6). The comparison between the ANN model prediction and the observed values is illustrated in Figure 7. This process is important to understand the precision of the developed ANN model.

The results show that the ANN model accurately estimated the variation in real crude biosurfactant concentration values. This leads to the conclusion that the topology 6-6-1 (Figure 8) could be used, because the model achieved a good agreement with few parameters (49). A neural model with a large set of parameters can cause overfitting, leading to large errors in performance for prediction. Small-size neural networks are important for real-time applications, due to their better generalization capability and less computational effort (Albuquerque et al., 2008). Thus, the developed neural model can be applied to predict dynamic behaviour in batch bioreactors for

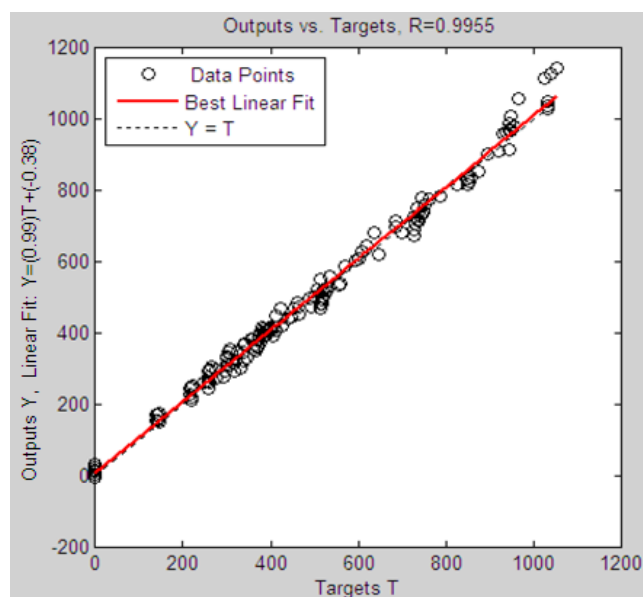


Figure 6. Dispersion plot of crude biosurfactant production (mg/L).

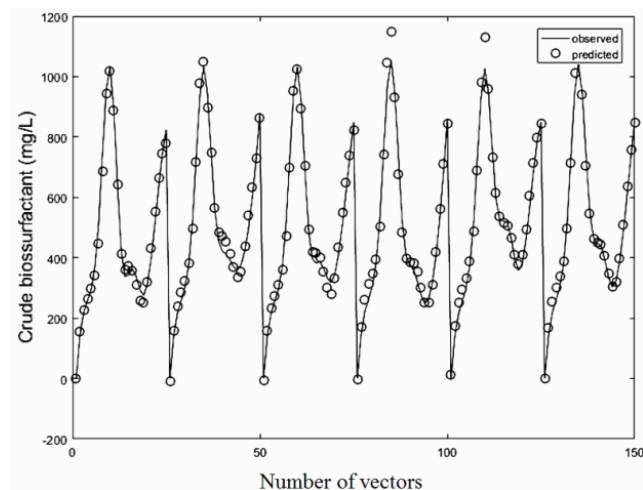


Figure 7. Comparison between observed and predicted values of crude biosurfactant concentrations in the test data set performed by the ANN model.

biosurfactant production, enabling faster actions for adjustment.

Kashkouli et al. (2011) used ANN modelling in fermentation parameters for biosurfactant production by *Bacillus subtilis* using sugar cane molasses. The topology reached was 3 neurons in the input layer, 6 in the hidden layer and 1 in the output layer and a network was trained by Levenberg-Marquardt algorithm. The authors observed MSE of $1.33E-06$ and R^2 of 0.997.

Application of the Soft Sensor Using an Electronic Spreadsheet

The soft sensor was based on the neural model (with optimised weight and bias) and its structure was entered into an electronic spreadsheet. Consequently,

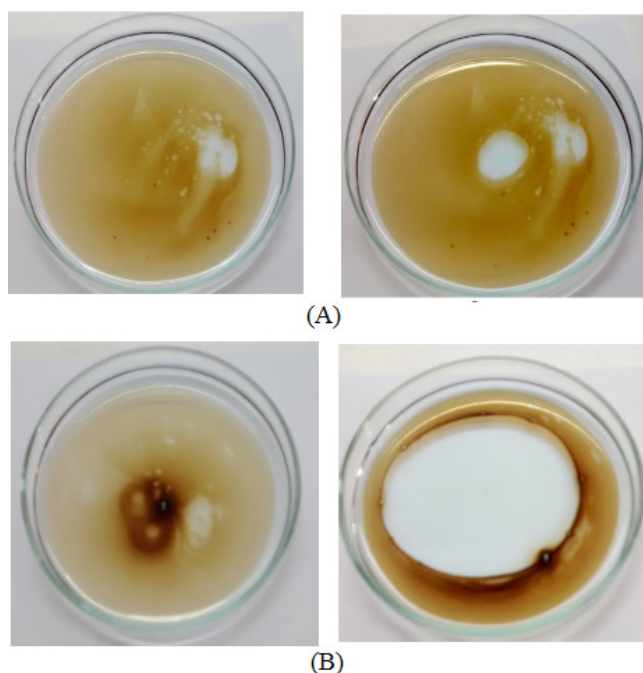


Figure 11. Crude biosurfactant application: (A) Petri dish before (left) and after (right) addition of negative control. (B) Petri dish before (left) and after (right) addition of crude biosurfactant.

Emulsifying Index (%)

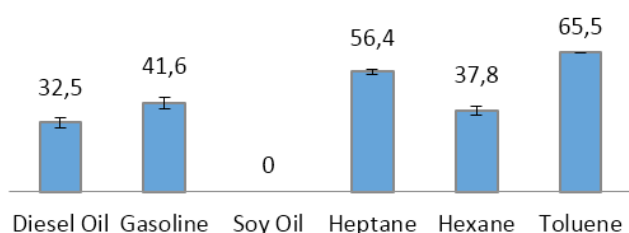


Figure 12. Emulsifying index (%) in different solvents such as diesel oil, gasoline, soy oil, heptane, hexane, toluene.

%) and maximum with toluene (65.5 %). It is known the biosurfactants interact with hydrocarbons and form emulsions with the aqueous medium.

Al-Wahaibi et al. (2014) reported biosurfactant production by *Bacillus subtilis* that emulsified various hydrocarbons, among them hexane 48% and heptane

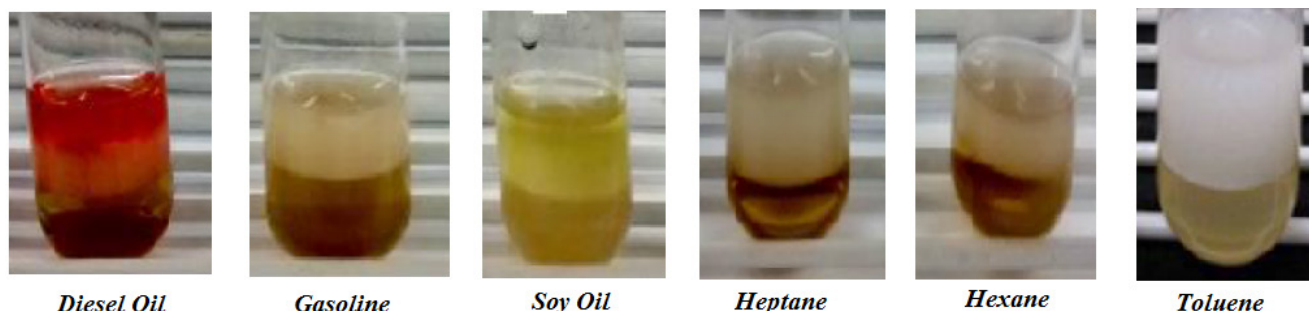


Figure 13. Emulsifying index using diesel oil, gasoline, soy oil, heptane and hexane.

52%. The emulsification index values of *Bacillus subtilis* K1 culture supernatant (Pathak et al., 2013) for hexane was 33.3% and heptane was 40%. In this study, the emulsification index obtained is in accord with that observed in other papers, indicating its application, for instance, in environmental contamination.

CONCLUSION

The biosurfactant produced by *Bacillus subtilis* fermentation could be followed in a batch bioreactor with an alternative substrate from waste, without the addition of any synthetic supplements. Through the statistical model, the optimum points were reached by analysis of response surface methodology to define the operation parameters, yielding a good agreement between the values for crude biosurfactant concentration (~1931.2) and surface tension reduction ratio (~46.4). To our knowledge this is first report on the use of beet peel for biosurfactant production in a batch bioreactor. The use of alternative substrate from waste (beet peel from restaurants and glycerol from biodiesel production) as culture medium would have a huge impact on biosurfactant production costs. Thus, the application of an ANN-based approach with hyperbolic tangent activation function and trainbr training algorithm was able to accurately model ($R^2 = 0.9955$) the desired dynamic behaviour of crude biosurfactant concentrations. These results allowed the development of a soft sensor aided by Excel to monitor and infer actions. The biosurfactant produced was applied in Petri dishes containing an oil layer on water and showed promising potential (observation of clear zones) in remediation applications of oil spill. The values of the emulsification index observed in organic solvents, such as gasoline (41.6%), heptane (56.4%) and toluene (65.5%), suggest its application in industrial processes.

ACKNOWLEDGEMENTS

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq- National Council of Technological and Scientific Development).

REFERENCES

- Albuquerque, C. D. C., Takaki, G. M. C., and Fileti, A. M. F., On-line biomass estimation in biosurfactant production process by *Candida lipolytica* UCP 988. *Journal of Industrial Microbiology & Biotechnology*, 35(11), 1425-1433 (2008).
- Al-Wahaibi, Y., Joshi, S., Al-Bahry, S., Elshafie, A., Al-Bemani, A., Shibulal, B., Biosurfactant production by *Bacillus subtilis* B30 and its application in enhancing oil recovery. *Colloids and Surfaces B: Biointerfaces*, 114, 324-333 (2014).
- Barros, F. F. C., Quadros, C. P.; Pastore, G. M., Studies of emulsifying properties and stability of the biosurfactant produced by *Bacillus subtilis* in cassava wastewater. *Food Science and Technology*, 28, 979-985 (2008).
- Bosca, S., Fissore, D., Design and validation of an innovative soft-sensor for pharmaceuticals freeze-drying monitoring. *Chemical Engineering Science*, 66(21), 5127-5136 (2011).
- Caldeira, A. T., Arteiro, J. M., Roseiro, J. C., Neves, J., Vicente, H., An artificial intelligence approach to *Bacillus amyloliquefaciens* CCM1 1051 cultures: application to the production of anti-fungal compounds. *Bioresource Technology*, 102(2), 1496-1502 (2011).
- Fang, N., Pai, P. S., Edwards, N., Neural network modeling and prediction of surface roughness in machining aluminum alloys. *Journal of Computer and Communications*, 4, 1-9 (2016).
- Fang, Z., Liu, X., Zhang, M., Sun, J., Mao, S., Lu, J., Rohani, S., A neural network approach to simulating the dynamic extraction process of L-phenylalanine from sodium chloride aqueous by emulsion liquid membrane. *Chemical Engineering Research and Design*, 105, 188-199 (2016).
- Fu, K., Chen, G., Sema, T., Zhang, X., Liang, Z., Idem, R., Tontiwachwuthikul, P. Experimental study on mass transfer and prediction using artificial neural network for CO₂ absorption into aqueous DETA. *Chemical Engineering Science*, 100, 195-202 (2013).
- Ge, Z. Huang, B. Song, Z., Mixture semisupervised principal component regression model and soft sensor application. *Process Systems Engineering*. 60, 533-545 (2014).
- Gudiña, E. J., Fernandes, E. C. Rodrigues, A. J., Teixeira, J. A., Rodrigues, L. R., Biosurfactant production by *Bacillus subtilis* using corn steep liquor as culture medium. *Frontiers in Microbiology*, 6, 1-7 (2015). <https://doi.org/103389/fmicb.2015.00059>
- Jin, H., Chen, X., Yang, J.; Zhang, H., Wang, L., Wu, L., Multi-model adaptive soft sensor modeling method using local learning and online support vector regression for nonlinear time-variant batch process. *Chemical Engineering Science*, 131, 282-303 (2015).
- Kaneko, H., Funatsu, K. Classification of the degradation of soft sensor models and discussion on adaptive models process systems engineering, *AIChE Journal*, 59, 2339-2347 (2013).
- Khatibisepehr, S., Huang, B., Domlan, E. Naghooshi, E., Zhao, Y., Miao, Y., Shao, X., Khare, S., Keshavarz, M. Feng, E., Xu, F., Espejo, A., Kadil, R., Soft Sensor Solutions for Control of Oil Sands Processes, *Canadian Journal of Chemical Engineering*, 91, 1416-1426 (2013).
- Liu, Y., Gao, Z., Li, P., Wang, H., Just-in-time Kernel learning with adaptive parameter selection for soft sensor modeling of batch processes. *Industrial & Engineering Chemistry Research*. 51, 4313-4327 (2012).
- Mouafi, F., Abo, E. M. M. and Moharam, M. E., Optimization of biosurfactant production by *Bacillus brevis* using response surface methodology. *Biotechnology Reports*. 9, 31-37 (2016).
- Ni, W., Brown, S., Man, R., A localized adaptive soft sensor for dynamic system modeling. *Chemical Engineering Science*, 111, 350-363 (2014).
- Oladunjoye, A. O., Oyewole, S. A., Singh, S., Ijabadeniyi, O. A., Prediction of *Listeria monocytogenes* ATCC 7644 growth on fresh-cut produce treated with bacteriophage and sucrose monolaurate by using artificial neural network. *LWT-Food Science and Technology*, 76, 9-17 (2016).
- Oliveira, D. W. F., França, I. W. L., Felix, A. K. N., Martins, J. J. L., Giro, M. E. A., Melo, V. M. M., Gonçalves, L. R. B., Kinetic study of biosurfactant production by *Bacillus subtilis* LAMI005 grown in clarified cashew apple juice. *Colloids and Surface B: Biointerfaces*, 101, 34-43 (2012).
- Oroian, M., Influence of temperature, frequency and moisture content on honey viscoelastic parameters -Neural networks and adaptive neuro-fuzzy inference system prediction. *LWT-Food Science and Technology*, 63, 1309-1316 (2015).
- Pathak, K. V., Keharia, H., Application of extracellular lipopeptide biosurfactant produced by endophytic *Bacillus subtilis* K1 isolated from aerial roots of banyan (*Ficus benghalensis*) in microbially enhanced oil recovery (MEOR). *Biotech*, 4(1), 41-48 (2014). <https://doi.org/10.1007/s13205-013-0119-3>
- Pereira, J. F. B., Gudiña, E. J., Costa, R., Vitorino, R., Teixeira, J. A., Coutinho, J. A. P., Rodrigues, L. R., Optimization and characterization of biosurfactant production by *Bacillus subtilis* isolates towards microbial enhanced oil recovery applications. *Fuel*, 111, 259-268 (2013).
- Rahimi, M., Hajialyani, M., Beigzadeh, R. Alsairafi, A. A., Application of artificial neural network and genetic algorithm approaches for prediction of flow characteristic in serpentine microchannels. *Chemical Engineering Research and Design*, 98, 147-156 (2015).

- Rogina, A. A., Sisko, I., Mohler, I., Ujevic, Z., Bolf, N., Soft sensor for continuous product quality estimation (in crude distillation unit). *Chemical Engineering Research and Design*, 89, 2070-2077 (2011).
- Santos, B. F., Ponezi, A. N., and Fileti, A. M. F., Strategy of using waste for biosurfactant production through fermentation by *Bacillus subtilis*. *Chemical Engineering Transactions*. 37, 727-732 (2014).
- Singh, K. P., Basant, A., Malik, A. and Jain, G., Artificial neural network modelling of the river water quality-A case study. *Ecological Modelling*. 220, 888-895 (2008).
- Sivapathasekaran, C., Mukherjee, S., Ray, A. Gupta, A., Sen, R., Artificial neural network modeling and genetic algorithm based medium optimization for the improved production of marine biosurfactant. *Bioresource Technology*, 101, 2884-2887 (2010).
- Sousa, M., Dantas, I. T., Feitosa, F. X., Alencar, A. E. V., Soares, S. A., Melo, V. M. M., Gonçalves, L. R. B., Sant'ana, H. B., Performance of biosurfactant produced by *Bacillus subtilis* LAMI005 on the formation of oil/biosurfactant/water emulsion: Study of the phase behaviour of emulsified systems. *Brazilian Journal of Chemical Engineering*, 31, 613-623 (2014).
- Wang, Z. X., He, Q. P., Wang, J., Comparison of variable selection methods for PLS-based soft sensor modeling. *Journal of Process Control*, 26, 56-72 (2015).

