CHEMICAL ENGINEERING TRANSACTIONS

VOL. 57, 2017

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Sauro Pierucci, Jiří Jaromír Klemeš, Laura Piazza, Serafim Bakalis

ISBN 978-88-95608- 48-8; **ISSN** 2283-9216

DOI: 10.3303/CET1757169

Development of Artificial Intelligence Models to Monitor Biosurfactant Concentration in Real-Time using Waste as Substrate in Bioreactor Through Fermentation by *Bacillus Subtilis*

Brunno F. dos Santos*[a], Alexandre N. Ponezi [b], Ana M. F. Fileti [c]

[a] Department of Chemical and Materials Engineering (DEQM), Pontifical Catholic University of Rio de Janeiro (PUC-Rio). Rua Maequês de São Vicente, 225 – Gávea, Rio de Janeiro - RJ, 22430-060, Brazil.
[b] Research Center for Chemistry, Biology and Agriculture (CPQBA), University of Campinas (UNICAMP). Av. Alexandre Cazellato, 999 – Betel, Paulínia – SP, 13140-000, Brazil.
[c] School of Chemical Engineering (FEQ), Department of Chemical Systems Engineering (DESQ), University of Campinas (UNICAMP). Rua Albert Einstein, 500 – Cidade Universitária, Campinas – SP, 13083-852, Brazil.
bsantos@puc-rio.br

Biosurfactants are biological compounds with active surface interactions. They are produced through metabolism of microorganisms (bacteria, yeast and fungi) and many applications are mentioned in industry (chemical, food, pharmaceutical etc.). However, the production is compromised by the use of methodologies that is unable to compete economically with traditional methods (synthetic substrates, batch without monitoring and no applications). As an alternative indispensable to reduce the costs of the process, artificial intelligence is wide applied currently. Thus, the present work is concerned with the development of intelligent mathematical models to predict the crude biosurfactant concentration using a bench-top bioreactor system. The fermentation substrate was waste material composed of glycerol from biodiesel process and beet peel from restaurants. The microorganism used was *Bacillus subtilis*. In order to improve the final product quality, these models will be used in monitoring through faster decision-making regarding process variables. Two techniques from the artificial intelligence field were used: artificial neural networks (ANN) and neuro-fuzzy (ANFIS). The biosurfactant concentration is the process variable to be predicted using the historical data acquired from the bioreactor plant. Software *MATLAB 2010a* was used to implement ANN and ANFIS models. Both of models were built with six neurons in input layer (microbial concentration - MC, glucose concentration - GC, dissolved oxygen concentration - OD, surface tension – ST, dissolved surface tension in 10× - ST-1 and 100× - ST-2) and the others parameters from networks were developed by factors combination. The results showed that models are appropriate to predict profile of crude biosurfactant concentration successfully and were fast enough to be used in nonlinear predictive strategies with good adjustments.

1. Introduction

Biosurfactant are synthesized by metabolism of microorganism as bacteria, yeast and mold. These molecules are capable of forming stable emulsions and be used in formulations lot of compounds, pharmaceutical products, cosmetic, water and soil remediation and some processes.

The production of many of these biosurfactants at an industrial scale has been limited due to their high cost (Ahmad et al. 2016). An alternative to feasibility of production is to utilize renewable and low cost raw materials such as glycerol (Sousa et al., 2014), corn steep liquor (Gudiña et al., 2015), soybean oil refinery (Tayyebi et al., 2013), groundnut husk (Salihu et al., 2015) and others. Kinetic of fermentation processes are hard to observe, because measurement of some variables during the production, biosurfactant concentration, is currently delayed because of the dependence on other experimental 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.

Please cite this article as: Santos B., Ponezi A., Fileti A.M.F., 2017, Development of artificial intelligence models to monitor biosurfactant concentration in real-time using waste as substrate in bioreactor through fermentation by bacillus subtilis, Chemical Engineering Transactions, 57, 1009-1014 DOI: 10.3303/CET1757169

Several different research groups dedicated their study to investigating the complex biosurfactant production. Artificial intelligence could be an attainable strategy. Artificial neural network (ANN) is a way to model information from process, which is inspired by the biological nervous system (brain). Dhanarajan et al. (2014) developed an artificial neural network (ANN) modeling coupled with particle swarm optimization (PSO) algorithm to monitor the process variables for enhanced lipopeptide production by marine *Bacillus megaterium*. Temperature, pH, agitation and aeration were used as input variables. Results showed significant enhancement of lipopeptide production around 46% due adopted technique.

Applied in the area of product quality, ANN has been successfully evaluated to predict many variables of processes (Khajeh & Barkhordar, 2013 and Ahmad et al., 2016): temperature (Santos et al.2013), yields (Antwi et al., 2016 and Pappu et al., 2016), reaction kinetics (Yang et al. 2011) and others. Into artificial intelligence, there is the possibility to use also the fuzzy inference system (FIS). It has been applied to identify and model complex non-linear systems (Karaman & Kayacier, 2011) and to predict non-linear evolution of parameters such as physico-chemical or rheological properties. Advanced mathematical model approaches including adaptive neural networks, Adaptive Neuro Fuzzy Inference System (ANFIS) have also been proven to be very powerful tools. The present work is designed to investigate the efficacy of ANN and ANFIS models in estimating biosurfactant production by bacterial *Bacillus subtilis* through fermentation parameters as input in models. With the complexity of the models, it may also be applied as a tool to research and refine knowledge on the correlation of growth of cells. Furthermore, the models provide a first step in developing model-based optimized process control strategies (in real time), which may result in higher biosurfactant production.

2. Materials and methods

2.1 Experiments

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 (625-nm wavelength) to an absorbance range from 0.08 to 0.1, according to McFarland's method.

2.2 Alternative medium

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 peel beet 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.

2.3 Measurement of biomass and crude biosurfactant in bioreactor

The experiments were conducted in a 7.5-L bioreactor (Bioflow 310 New Brunswick Scientific, USA) equipped with standard probes for temperature and OD, as well as auxiliary equipment, allowing a fermentation volume of 4 L. Biomass was determined by the dry weight obtained from the assays. At the end of the assays, a 30-mL sample from the culture broth was centrifuged (10 000 rpm, 10 min, 4 ºC). The biomass obtained was dried at 50 ºC for 24 h and the weight evaluated.To recover the biosurfactant, the cell-free supernatant from the fermentation culture was subjected to acid precipitation. Briefly, the supernatant was acidified (pH 2.0) with 1 M HCl and incubated for 24 h at 7 ºC. Next, it was centrifuged (10 000 rpm, 17 min, 2 ºC). The supernatant was then discarded and the precipitate was washed with acidified water and stored. All assays were performed in duplicates.

2.4 Surface tension

The surface tensions of the cell-free broths were measured by plate method, using a digital tensiometer model k12 (Krüss GmbH, German) at 20ºC.

2.5 Modeling with Artificial Neural Network (ANN)

The ANN employed is the feed forward multilayer perceptron. The model has to be first trained and then tested to use for application. The training was done with *MATLAB R2010a* software. The training algorithms used were the gradient descent backpropagation with adaptive learning rate (*traingdx*), the Levenberg-Marquardt based backpropagation algorithm (*trainlm*) and in conjunction with Bayesian regularization (*trainbr*). The training procedure aims at obtaining an optimal set of the network weights (W) and bias (B) matrices,

which minimize an error function.
In this work, the quantity of neurons in both the input layer (any of these variables: MC, GC, OD, ST, ST-1 and ST-2) and the hidden layer were defined by parameters combination.
Each neuron in the hidden layer has a series of weighted inputs and transforms it by the following activation function, to satisfy feed forward sign. The activation functions used were logsig and tansig.
The ANN was assessed by error criteria, comparing the observed values and the predicted values of the network by means of calculation of the total sum of the square error (SSE) for the data of the training dataset, showed in Equation 1.

$$SSE = \sum_{i=1}^{n} (Y_{observed} - Y_{predicted}) \qquad (1)$$

2.6 Adaptive Neuro-Fuzzy Inference System (ANFIS)

The ANFIS is a hybrid model in which the nodes in the different layers of the network handle fuzzy parameters, representing a useful neural network approach for the solution of function approximation problems. Data driven procedures for the synthesis of ANFIS networks are typically based on clustering a training set of numerical samples of the unknown function to be approximated. Each layer in the network corresponds to a part of the fuzzy inference system (FIS) called: input fuzzification, rule inference and fire strength computation, and output defuzzification. The main advantage of this kind of representation is that the FIS parameters are encoded as weights in the neural network and, thus, can be optimized via powerful well known neural net learning methods. This model is mostly suited to the modeling of nonlinear systems. In this work, the training of the neuro-fuzzy estimator was performed using the ANFIS toolbox of MATLAB R2010a.

3. Results

The procedures for the preparation and standardization 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 (19), using glycerol from biodiesel process and beet peel from restaurants as culture broth. The maximum crude biosurfactant concentration in bioreactor was 550 ± 85 mg/L with an established 6% of glycerol from biodiesel and 7.5 % (v/v) of beet peel. The operating conditions of bioreactor were 200 rpm for agitation and 0,5 vvm for aeration.
It is very important the development of mathematical model to allow the monitoring and control of process, particularly because the measurements of crude biosurfactant require lab work involving long processing times.To demonstrate the effectiveness of the proposed approaches, the concentration of crude biosurfactant prediction from ANN model were compared to that obtained from ANFIS model, under offline.

3.1 Development of ANN model

The experimental data, obtained from eight runs were normalized in the range [-1,1]. The following experimental data were collected during the study in duplicate and the average between them was used. Each run set, containing about 25 arrays (due cubic spline data interpolation), that was randomly split into sets of training data (75%) and testing data (25%). Initially, for the prediction of the crude biosurfactant, it was proposed a network structure with sex input neurons - microbial concentration - MC, glucose concentration - GC, dissolved oxygen concentration - OD, surface tension – ST, dissolved surface tension in 10× - ST^{-1} and 100× - ST^{-2}. Thus, the combinations of several scenarios allowed reaching a topology of neural network, seen in Table 1. A hidden layer containing 6 and 8 neurons were evaluated by observing the SSE value. The minimum SSE obtained for the training set was 0.283 (Figure 2).

Table 1: Scenarios of results of the artificial neural network (ANN) model with varying number of neurons in the hidden layer

Scenarios	Hidden Layer	Training	SSE	R^2	Activation Function
1	6	Traingdx	5.89	0.987	Tansig
2	6	Traingdx	7.85	0.919	Logsig
3	8	Traingdx	6.45	0.925	Tansig
4	8	Traingdx	7.77	0.944	Logsig
5	6	Trainlm	0.305	0.996	Tansig
6	6	Trainlm	0.293	0.995	Logsig
7	6	Trainbr	0.292	0.994	Tansig
8	6	Trainbr	0.283	0.995	Logsig

The best scenario simulated was 8. Extracting the square root of MSE (SSE/number of vectors) value (6.65E-04) and denormalizing it, the deviation of biosurfactant production was found to be 1.5 g/L, lower than precision of crude biossurfactant measurement. Hyperbolic tangent, logistic and linear functions were used as activation fuction and *Trainbr* was chose as training algorithm (Figure 1). Using the test set to evaluate the generalization capacity of the model, a good agreement (slope=1 and linear coefficient = 0.32) was observed between predictions and unseen points in the dispersion plot (Figure 2A). The developed neural model provides crude biosurfactant predictions with 99.5% certainty.

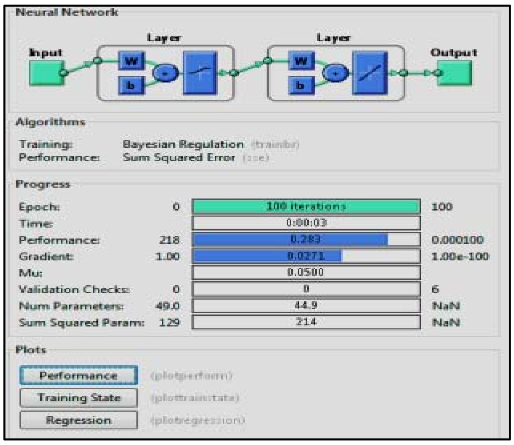

Figure 1: Evaluation of ANN showing the training patterns and epoch iteration.

Check runs using the same conditions in bioreactor were carried out to validate the developed model. Figure 2B shows the behaviour of observed and predicted values by ANN in validation.

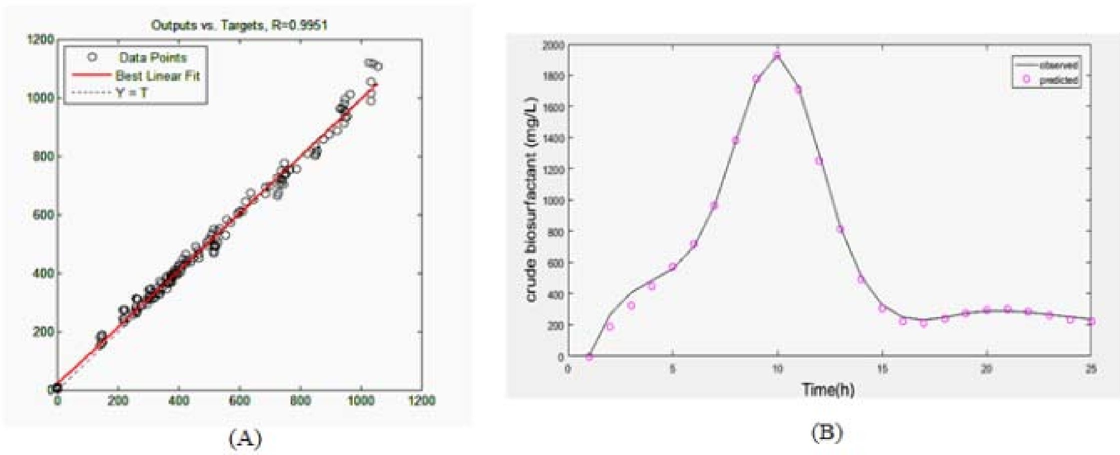

Figure 2: Representation of regression curve (A) and curve of predict and observed data (B).

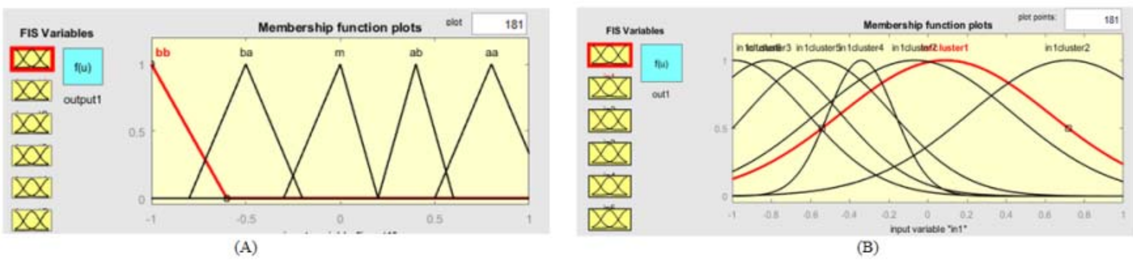

Figure 3: Initial membership function (A) and final membership function for inputs for the prediction model using ANFIS clustering.

3.2 ANFIS model

ANFIS uses a hybrid learning algorithm to tune the parameters of a Sugeno-type fuzzy inference system (FIS). The same training and test sets containing six inputs were used for the ANFIS model. In this work, the fuzzy clustering algorithm was implemented to become a neural network system more useful. The fuzzy stage is responsible for the analysis of the distribution of data and grouping them into clusters with different membership values (Figure 3). The training set is reduced using ANFIS clustering, therefore, training period of the neural network decreased because of the low computational effort. ANFIS also validates models using a checking data set to test for overfitting of the training data.
The algorithm uses a combination of the least-squares and back-propagation gradient descent methods to model a training data set. ANFIS also validates models using a checking data set to test for overfitting of the training data.
The maximum error tolerance and epochs were set to 0.001 and 100, respectively. Test set data were fitted to the ANFIS model output, with a satisfactory error of 0.025 mg/L.
The total number fuzzy rules were seven and the membership function was of Gaussian shape, defined in Equation 2, where σ and are parameters of the Gaussian membership function.

$$f_g(x,\sigma,c) = \exp\left[-\frac{1}{2}\left[\frac{x-c}{\sigma}\right]^2\right] \tag{2}$$

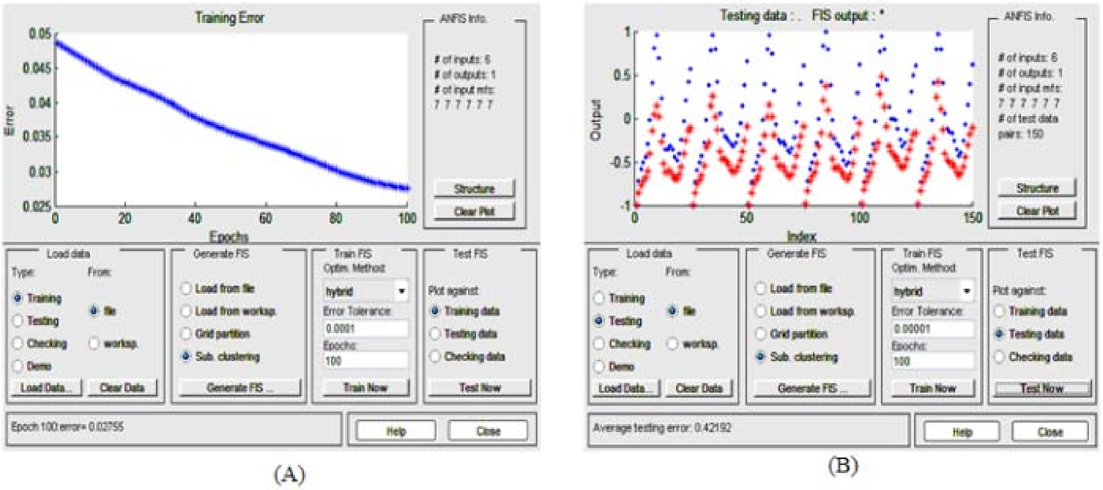

Figure 4: Training data set: SSE profile (A); Testing data set (B)

The value of SSE during ANFIS training error was (Figure 4) better than reached SSE in ANN training error, but the prediction by ANN models were better than ANFIS mode (Figure 2B), nevertheless, both of adjusts were good. Leite et al. (2011) also used two techniques from the artificial neural network (ANN and ANFIS) applied to pilot plant of styrene production using MATLAB. Their results demonstrated that both models were able to predict the reactor temperature profile successfully.
The ANFIS model was considered acceptable because the dispersion plot in Figure 5 presented a slope equal to the unity and the linear coefficient vanished.

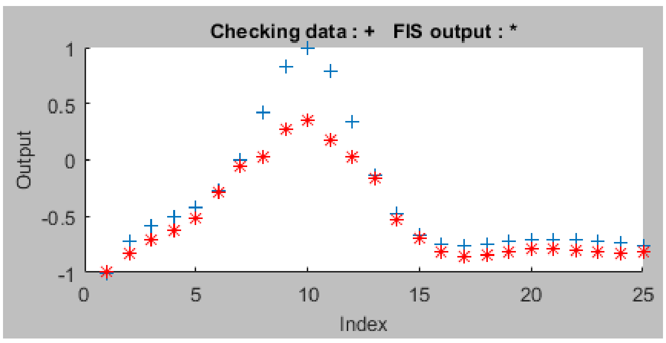

Figure 5: Prediction curve versus observed data using ANFIS.

4. Conclusions

An accurate estimation and prediction of biosurfactant production should assist in reducing the total cost in bioreactor and may facilitate development of processes that utilize renewable medium. Artificial intelligence systems provide powerful tools that are able to model complex relationship between inputs variables. Thus, two modeling approaches, ANN and ANFIS, were evaluated for their efficacy to predict accurately estimation of biosurfactant production based on prior obtained input data. Based on several experimental runs, under different operating conditions, the models were built in order to predict the biosurfactant production. Implementing of such modeling strategies in this system is very promising because optimal and/or predictive control schemes can be developed so that the end-of-batch product quality is optimized. This approach can be effectively used for identifying various important parameters for developing real-time software-based sensor estimation for biosurfactant production. The ANN and the ANFIS models showed to be very effective in the dynamic modeling of this nonlinear fermentation process, offering accurate long-range predictions as well.

Acknowledgments

National Council of Technological and Scientific Development —CNPq.

Reference

Ahmad Z., Crowley D., Marina N., Jha S. Kr., 2016a, Estimation of biosurfactant yield produced by Klebseilla SP FKOD36 bacteria using artificial network approach. Measurement, 81,163-173.

Ahmad Z., Arshad M., Crowley D., Hhoshnevisan B., Yousefi M., Imram M., Hussain S., 2016b, Comparative efficacy of ANN and ANFIS models in estimating biosurfactant production produced by Klebseilla SP FKOD36. Stoch Environ Res Risk Assess, 30, 353-363.

Antwi P., Li J., Boadi P. O., Meng J., Shi E., Deng K., Bondinuba F. K., 2017, Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network, Bioresource Technology, 228, 106-115.

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, Biochemical Engineering Journal, 84, 59-65.

Gudiña E. J., Fernandes E. C., Rodrigues A. I., Teixeira J. A., Rodrigues L. R., 2015, Biosurfactant production by Bacillus sibtilis using corn steep liquor as culture medium, Front Microbiol, 6, 1-7. DOI: 10.3389/fmicb.2015.00059.

Karaman S., Kayacier A., 2011, Effect of temperature on rheological characteristics of molasses: modeling of apparent viscosity using adaptive neuro-fuzzy inference system (ANFIS), LWT-Food Science and Technology, 44, 1717-1725.

Khajeh M., Barkhordar A., 2013, Modelling of solid -phase tea waste extraction for the removal of manganese from food samples by using artificial neural network approach, Food Chemistry, 141, 712-717.

Leite M. S., Santos B. F., Lona L. M. F., Silva F. V., Fileti A. M. F., 2011, Application of artificial intelligence techniques for temperature prediction in a polymerization process, Chemical Engineering Transactions, 24, 385-390.

Pappu J. S. M., Gummadi S. N., 2016, Modeling and simulation of xylitol production in bioreaction by Debaryomyces nepalensis NCYC 3413 using unstructured and artificial neural network models, Bioresource Technology, 220, 490-499.

Salihu A., Abbas O., Sallau A. B., Alam M. Z., 2015, Agricultural residues for cellulolytic enzyme production by *Aspergillus niger*: effects of pretreatment, 3 Biotech, 5, 1101-1106.

Santos B. F., Ponezi A. N., Fileti A. M. F., 2014, Strategy of using waste for biosurfactant production through fermentation by Bacillus subtilis, Chemical Engineering Transactions, 37, 727-732.

Santos R. R. C., Santos B. F., Fileti A. M. F., Silva F. V., Zemp R. J., 2013, Application of artificial neural networks in na experimental batch reactor of styrene polymerization for predictive model development, Chemical Engineering Transactions, 32,1399-1404.

Tayyebi S., Lotfabad T. B., Roostaazad R., 2013, Appling neural network to dynamic modeling of biosurfactant production using soybean oil refinery waste, Iranica Journal of Energy & Environment, 4, 161-170. ISSN: 2079-2115.

Yang Y., Wang G., Wang B., Li Z., Jia X., Zhou Q., Zhao Y., 2011, Biosorption of acid black and Congo red from aqueous solution by nonviable Penicillium YW 01: Kinetic study, equilibrium isotherm and artificial neural network modeling, Bioresource Technology, 102, 828-834.