Article

Modelling Acetification with Artificial Neural Networks and Comparison with Alternative Procedures

Jorge E. Jiménez-Hornero [1,*], Inés María Santos-Dueñas [2] and Isidoro García-García [2]

[1] Department of Electrical Engineering and Automatic Control, University of Cordoba, Campus Rabanales, 14071 Cordoba, Spain

[2] Department of Chemical Engineering, University of Cordoba, Campus Rabanales, 14071 Cordoba, Spain; ines.santos@uco.es (I.M.S.-D.); isidoro.garcia@uco.es (I.G.-G.)

* Correspondence: jjimenez@uco.es; Tel.: +34-957-212-079

Abstract: Modelling techniques allow certain processes to be characterized and optimized without the need for experimentation. One of the crucial steps in vinegar production is the biotransformation of ethanol into acetic acid by acetic bacteria. This step has been extensively studied by using two predictive models: first-principles models and black-box models. The fact that first-principles models are less accurate than black-box models under extreme bacterial growth conditions suggests that the kinetic equations used by the former, and hence their goodness of fit, can be further improved. By contrast, black-box models predict acetic acid production accurately enough under virtually any operating conditions. In this work, we trained black-box models based on Artificial Neural Networks (ANNs) of the multilayer perceptron (MLP) type and containing a single hidden layer to model acetification. The small number of data typically available for a bioprocess makes it rather difficult to identify the most suitable type of ANN architecture in terms of indices such as the mean square error (MSE). This places ANN methodology at a disadvantage against alternative techniques and, especially, polynomial modelling.

Keywords: bioreactor systems; acetification; vinegar; modelling; artificial neural networks; Multilayer Perceptron

1. Introduction

1.1. Modelling of Bioprocesses

Biochemical engineering, which aims to develop and optimize bioprocesses, is having an increasingly strong economic impact on developed countries by effect of the wide variety of industrial fields where it is currently used (e.g., agri-food, pharmaceutical and energy production) [1]. As a rule, bioprocesses include complex operations involving intricate biotransformation mechanisms effected by microorganisms that require an appropriate environment for development.

In this scenario, simulation techniques provide powerful tools for the quantitative prediction of state variables (e.g., substrate and product concentrations) and yields under different operating conditions with the need for little or no testing. Among others, this allows substrate-feeding strategies to be precisely designed, dimensioned and controlled [2]. However, simulation requires the use of variably complex mathematical models to compile available knowledge about a product in order to mimic its behavior for a specific purpose [3]. Bioprocesses are typically modelled by using a white-box (mechanistic or first-principles) model, a grey-box (or hybrid) model or a black-box (or empirical) model.

Mechanistic models are based on the physic-chemical and biological factors that govern the target process [4]. Such factors are represented by differential balance equations, kinetic equations and equilibrium relations [5,6]. The greatest difficulty in establishing a mechanistic model for a process is identifying the most suitable mathematical structure for describing its apparent kinetics or stoichiometry. Very often, the mathematical functions used contain a number of parameters to be fitted under specific operating conditions [7], resulting in practice a grey-box model combining one based on available knowledge about the process with a parameter estimation problem from experimental data. These models are often subject to structural [8] and/or practical identifiability problems [9] that can prevent finding unique values for some parameters. Furthermore, however, mechanistic models typically hold over broad experimental ranges and can be easily extrapolated to diverse conditions.

Black-box models describe functional relationships between process inputs and outputs exclusively obtained from experimental data with no provision for internal mechanisms; therefore, they are not based on available knowledge about the processes they are intended to represent [10]. In fact, the mathematical structure of a black-box model is defined beforehand and its parameters, which lack physical significance, are determined by optimization. Overall, black-box models hold narrower ranges than mechanistic models and afford no physical interpretation of the target system. However, they provide highly accurate predictions over the experimental data range used in their development. The black-box models typically applied to bioprocesses are of either of two main types, namely:

- Regression models [11,12], which are usually based on polynomial equations relating the response variables (viz., the dependent variables or outputs) to the factors (viz., the independent variables or inputs) of a process. Regression models are typically used to examine the influence of experimental factors on process outputs, as well as potential interactions between factors.

- Artificial neural networks (ANNs) [13] consist of so-called "neurons". These are elemental computation or processing units that operate in parallel [14] and are mutually connected across a network comprising various layers. The ANNs used to model bioprocesses use mathematical combinations of basis functions that can be fitted to experimental data by determining their connection weights and biases. This makes ANNs highly accurate approximators to both static and dynamic non-linear functions. The process, which is known as "supervised training" or "supervised learning", minimizes differences between ANN outputs and experimental responses to specific inputs (training patterns) by using algorithms such as back-propagation [15]. Topologically, ANNs comprise an input layer that connects process inputs, an output layer that provides the process outputs, and one or more hidden layers between the input and output layer that reflect the black-box behaviour of the ANN. Architecturally, the ANNs typically used with bioprocesses are either feed-forward networks [16], where the only connections between neurons are those from inputs to outputs, or recurrent networks [17], where some outputs are fed back to the input layer (e.g., in models for dynamic systems). Bioprocesses can also be modelled with neuro-fuzzy networks, which use fuzzy logics [18] to increase extrapolability.

1.2. Modelling Acetification

Vinegar is industrially produced in fermentation tanks equipped with a self-aspirating turbine. Generally, the process involves transforming ethanol into acetic acid with the aid of a culture of strictly aerobic acetic acid bacteria (AAB). The process is usually conducted in a semi-continuous mode and each cycle is finished when the substrate (ethanol) is depleted to a preset extent. Then, the reactor is unloaded to an also preset volume and its residual content is used as inoculum in the next cycle, which is started by replenishing the tank with fresh medium. This mode of operation ensures a high productivity and stability. The operational variables that can be changed to control the process are those that influence the average concentrations of ethanol and acetic acid [19,20], namely, the ethanol content of the raw material, that at which the reactor is unloaded; the volume of medium to be unloaded; and the rate at which the reactor is loaded with fresh medium. The activity of AAB depends largely on the substrate and product concentrations [6,19,21–23], so operational variables

must thus be carefully chosen in order to ensure appropriate conditions for the microorganisms to grow. Industrially, the optimum acetification conditions are those that maximize productivity; however, identifying them requires careful modelling of the process.

There are various models for acetification (particularly as regards the biotransformation step). Most are of the first-principles (mechanistic) [6,9] or black-box (polynomial regression) type [24]. Although first-principles models are the more complex, they tend to hold over broader operational ranges because they use available information about the process concerned. Blackbox models are easier to develop because, as noted earlier, they only relate input and output variables through operational variables based on experimental data used to fit them irrespective of the particular mechanisms governing their behavior.

Developing a first-principles model requires examining the influence of each variable on the process, establishing balance and energy equations, defining kinetic equations and estimating their parameters. A number of kinetic equations for the acetification process have been proposed [6], some of which have been established from a small number of experiments or even under conditions markedly departing from those of the industrial process. Experimental plans based on more realistic conditions have led to models considering additional phenomena such as cell lysis and integrating all other variables of the process [5,6].

The accuracy and precision of a first-principles model depend on how accurately its kinetic parameters are estimated. However, the intrinsic complexity of this type of model often poses theoretical and practical identifiability problems that ultimately preclude obtaining a unique value for each parameter. In fact, the structural or theoretical identifiability of a model dictates whether its parameters can be unambiguously determined from the mathematical structure of the model alone [8]. On the other hand, the practical identifiability of a model considers the amount of experimental data used for estimations and their quality [9]. Algorithms for assessing theoretical and practical identifiability are rather complex, and those for the latter purpose can only be as good as the estimates themselves. Therefore, the high computational cost of defining the set of equations to be used adds to that of estimating the parameters concerned—which, as noted earlier, cannot be unambiguously determined. These shortcomings make first-principles models difficult to develop.

Black-box models, which are based on polynomial regression calculations, are usually easier to construct than mechanistic models because they have a preset structure based on polynomial equations and require no prior checking for identifiability. The fitting algorithms used are intended to provide simple relations between the responses (output variables) and factors (operational variables) from experimental data spanning specific operational ranges —and hence those where the model is expected to hold. Bioprocesses have so far been modelled by using various types of linear polynomials and non-linear polynomials of variable order (typically first and second) in addition to diverse experimental designs, such as those of Packett and Burman [25] and Box and Behnken [26]. Estimating the coefficients of a polynomial model requires performing a minimum number of experiments at different values (levels) of the operational variables. The accuracy of the ensuing model will depend on how well polynomial terms are selected and their parameters estimated (i.e., on the amount of data available for as widely different operating conditions as possible).

Because extensive testing is often expensive, the influence of experimental factors, as well as their interactions [27], on the target responses is usually elucidated by using experimental designs that try to establish the minimum number of experiments needed [12] to estimate the corresponding coefficients with accuracy [28]. The most widely used designs for this purpose are of the factorial type and, specifically, face-centered cubic designs [29]. The data obtained by testing are utilized to calculate the coefficients of the polynomials by using statistical methods of the least-squares type, such as best subset regression, backward stepwise regression or forward stepwise regression, in combination with one of various methods for identifying the most significant terms (e.g., Pareto analysis [30] or calculations based on the coefficient of determination, R^2). These algorithms are included in most major statistical software packages and used as supporting tools for the computations needed, all of which can be performed in a highly systematic manner.

Acetic fermentation has been examined with the aid of various polynomial black-box models based on acetic acid production and the mean fermentation rate. Thus, Santos-Duenas et al. [24] used the above-mentioned factors (viz., ethanol concentration at unloading time, volume to be unloaded at the end of a cycle and rate of raw substrate loading) at levels spanning the typical operational ranges for industrial processes in combination with an appropriate experimental design.

First-principles and black-box models for acetification have been used under the operating conditions maximizing acetic acid production [24,31] and provided very similar results.

1.3. Objectives of the Work

As an expansion, the possibility of using ANNs to model the acetification biotransformation is analysed in this work. Additionally, the quality of the ensuing predictions and the complexity of their obtainment, with those of alternative models for the same purpose, have been compared. Although it is well known that ANNs need a lot of experimental data for training and validation, in this work, thanks to the availability of previous models, it might be possible to carry out a quality assessment of this alternative, despite the fact that a large amount of experimental data is not available for this aim. Additionally, the comparison between models could suggest potential improvements over them. On the other hand, though other model types might be considered in future works, only the ANNs will be studied in the present one.

2. Materials and Methods

2.1. Experimental Conditions

The equipment, operational modes, experimental procedure and data used (Tables 1 and 2) to develop the proposed ANN-based model are described in detail elsewhere [6,19,22,24]; see supplemental material (Supplementary Figures S1–S8, Tables S1, S2) for a summary on these aspects.

The experiments in Table 1 were previously employed to develop the first-principles model [9] and those in Table 2 were used for the black-box polynomial regression model [24]. The latter was established under continuous loading conditions and the experiments followed a face-centered cubic design. All experiments were performed in the semi-continuous operational mode.

Table 1. Experiments used to construct the first-principles model. E_{max} (maximum ethanol concentration allowed during semi-continuous loading mode, % v/v); E_{unload} (ethanol concentration at the end of the cycle, % v/v; $V_{unloaded}$ (% unloaded volume); F_i (loading flow rate (L·min^{-1}); AcH_{cycle} (acetic acid concentration at the end of the cycle, g AcH·L^{-1}); P_{exp} (experimental productivity, g AcH·h^{-1})

Experiment Nº	Loading Mode	E_{max}	E_{unload}	$V_{unloaded}$	F_i	AcH_{cycle}	P_{exp}
1	Continuous	-	2	75	0.035	95 ± 1	15.1 ± 0.5
2	Continuous	-	2	50	0.035	98 ± 1	17.1 ± 0.5
3	Continuous	-	2	25	0.035	97 ± 1	17.3 ± 0.4
4	Continuous	-	3.5	50	0.035	78 ± 5	16.3 ± 0.4
5	Continuous	-	0.5	50	0.035	111 ± 1	14.7 ± 0.3
6	Continuous	-	0.5	75	0.01	110 ± 1	14.3 ± 0.3
7	Continuous	-	3.5	25	0.06	81 ± 1	17.8 ± 0.3
8	Semi-continuous	5	1.5	50	0.02	101 ± 2	14.8 ± 0.4
9	Semi-continuous	5	0.5	50	0.02	110 ± 2	13.8 ± 0.4

Table 2. Experiments used to construct the polynomial model. E_{unload} (ethanol concentration at the end of the cycle, % v/v); $V_{unloaded}$ (% unloaded volume); F_i (loading flow rate (L·min^{-1}); AcH_{cycle} (acetic acid concentration at the end of the cycle, g AcH·L^{-1}); P_{exp} (experimental productivity, g AcH·h^{-1}).

Experiment Nº	E_{unload}	$V_{unloaded}$	F_i	AcH_{cycle}	P_{exp}
1	3.5	75	0.06	81 ± 1	15.4 ± 0.4
2	3.5	75	0.01	81 ± 1	15.1 ± 0.4
3	3.5	25	0.06	81 ± 1	17.8 ± 0.3
4	3.5	25	0.01	81 ± 1	17.4 ± 0.6
5	0.5	75	0.06	108 ± 1	14.4 ± 0.4
6	0.5	75	0.01	110 ± 1	14.3 ± 0.3
7	0.5	25	0.06	112 ± 1	13.6 ± 0.3
8	0.5	25	0.01	111 ± 1	13.8 ± 0.2
9	3.5	50	0.035	78 ± 1	16.3 ± 0.4
10	0.5	50	0.035	111 ± 1	15.5 ± 0.2
11	2	75	0.035	95 ± 1	15.1 ± 0.5
12	2	25	0.035	97 ± 1	17.3 ± 0.4
13	2	50	0.06	95 ± 1	16.7 ± 0.5
14	2	50	0.01	95 ± 1	16.5 ± 0.6
15	2	50	0.035	98 ± 1	17.1 ± 0.5

2.2. Models for Acetic Acid Fermentation

2.2.1. First-Principles Model

The first-principles model used, as stated in [6,9], is defined by Equations (1)–(21). No energy balance has been considered since the process was operated under isothermal conditions.

$$V \frac{dX_v}{dt} + X_v \frac{dV}{dt} = V(r_{X_c} - r_{X_d}) \tag{1}$$

$$V \frac{dX_d}{dt} + X_d \frac{dV}{dt} = V(r_{X_d} - r_{lisis}) \tag{2}$$

$$V \frac{dE}{dt} + E \frac{dV}{dt} = F_i \cdot E_0 - V \cdot r_E \tag{3}$$

$$V \frac{dA}{dt} + A \frac{dV}{dt} = V \cdot r_A \tag{4}$$

$$V \frac{dO}{dt} + O \frac{dV}{dt} = F_i \cdot O^0 + V[\beta(O^0 - O) - r_{OE}] \tag{5}$$

$$\frac{dV}{dt} = F_i \tag{6}$$

$$r_{X_c} = \mu_c \cdot X_v \tag{7}$$

$$\mu_c = \mu_{max} \cdot f_e \cdot f_a \cdot f_o \tag{8}$$

$$f_e = \frac{E}{E + K_{SE} + \frac{E^2}{K_{IE}}} \tag{9}$$

$$f_a = \frac{1}{1 + \left(\frac{A}{K_{IA}}\right)^4} \tag{10}$$

$$f_o = \frac{O}{O + K_{SO}} \tag{11}$$

$$r_{X_d} = \mu_d \cdot X_v \tag{12}$$

$$\mu_d = \mu_d^0 \cdot f_{de} \cdot f_{da} \tag{13}$$

$$f_{dE} = 1 + \left(\frac{E}{K_{mE}}\right)^4 \tag{14}$$

$$f_{dA} = 1 + \left(\frac{A}{K_{mA}}\right)^4 \tag{15}$$

$$r_{lysis} = \mu_{lysis} \cdot X_d \tag{16}$$

$$r_E = a_{E/X} \cdot r_X \tag{17}$$

$$r_A = \frac{r_E}{Y_{E/A}} \tag{18}$$

$$r_{OE} = \frac{r_E}{Y_{E/O}} \tag{19}$$

$$\beta = \frac{K_L a}{1 + \frac{K_L a}{VV_m} \cdot \frac{RT}{H}} \tag{20}$$

$$VV_m = \frac{Q}{V} \tag{21}$$

where
- X_v, X_d, E, A and O are the concentrations of viable cells, dead cells, ethanol, acetic acid and dissolved oxygen (all in g·L⁻¹), respectively.
- t is the time (h).
- V is the volume of the medium (L), F_i the raw material feed rate (L·h⁻¹), E_i the concentration of ethanol in the fed raw material (g·L⁻¹) and O^0 the dissolved oxygen in equilibrium with air (g·L⁻¹).
- r_{X_c} is the cell growth rate (g cell·L⁻¹·h⁻¹), r_{X_d} the cell death rate (g cell·L⁻¹·h⁻¹), r_{lysis} the cell lysis rate (g cell·L⁻¹·h⁻¹), r_E the ethanol uptake rate (g ethanol·L⁻¹·h⁻¹), r_A the acetic acid formation rate (g acetic acid·L⁻¹·h⁻¹) and r_O the dissolved oxygen uptake rate (g oxygen·L⁻¹·h⁻¹).
- μ_c is the specific growth rate (h⁻¹), μ_{max} its maximum value (h⁻¹) and f_e, f_a and f_o are terms representing the influence of ethanol, acetic acid and dissolved oxygen on cell growth, respectively; K_{SE} is the ethanol saturation constant (g ethanol·L⁻¹), K_{IE} is the ethanol inhibition constant (g ethanol·L⁻¹), K_{IA} is the acetic acid inhibition constant (g acetic acid·L⁻¹) and K_{SO} is the dissolved oxygen saturation constant (g oxygen·L⁻¹); μ_d is the specific cell death rate (h⁻¹), μ_d^0 its minimum possible value (h⁻¹) and f_{dE} and f_{dA} are terms representing the influence of ethanol and acetic acid on cell death, respectively; K_{mE} and K_{mA} are the ethanol and acetic acid-induced cell death rate constants (g·L⁻¹), respectively; μ_{lysis} is the specific cell lysis rate (h⁻¹).

- $a_{E/X}$ is the ethanol yield factor required to supply the amount of energy needed for biomass growth (determined experimentally as 116.96 g ethanol·g^{-1} cell), $Y_{E/A}$ is the stoichiometric coefficient of ethanol uptake for acetic acid formation (0.767 g ethanol·g^{-1} acetic acid) and $Y_{E/O}$ is the stoichiometric coefficient of ethanol relative to oxygen (1.44 g ethanol·g^{-1} oxygen).
- β is a constant encompassing the following factors: $K_L a$ is the overall volumetric coefficient of mass transfer for the liquid phase (determined experimentally as 500 h^{-1}), VV_m is the ratio of the air feed rate to the volume of the medium (h^{-1}), R is the universal gas constant (0.082 atm·L·K^{-1}·mol^{-1}), T is the temperature (K), H is the Henry's constant (atm·L·mol^{-1}) and Q is the air feed rate (L·h^{-1}).

The model parameters and their estimated values are shown in Table 3.

Table 3. Model parameters and their estimated values for the first-principles model.

Parameter	Estimated Value
μ_{max}	0.62 h^{-1}
K_{SE}	3.8 g ethanol·L^{-1}
K_{IE}	10.63 g ethanol·L^{-1}
K_{IA}	98.6 g acetic acid·L^{-1}
K_{SO}	3.33 × 10^{-4} g oxygen·L^{-1}
μ_d^0	2.94 × 10^{-5} h^{-1}
K_{mE}	36.81 g ethanol·L^{-1}
K_{mA}	12.51 g acetic acid·L^{-1}
μ_{lysis}	0.52 h^{-1}

Experiments used for developing the model and thus for the parameter estimation are those in Table 1; each experiment corresponds to several repeated production cycles (at least ten), so there are a great number of them in background. Because of the usual identifiability problem (structural and practical ones [8,9,32]) in this type of models, only a couple of these parameters (μ_d^0 and μ_{lysis}) were completely identifiable in practice considering the available experimental data, the influence of such parameters on the model state variables (sensitivity analysis) and the correlations between them.

Regardless of the level of detail considered for the metabolism of this type of bacteria in order to propose the kinetic equations, the parameter identifiability problem will remain and even worsen if the number of parameters and kinetic equations are increased. In any case, the discussion and detailed analysis about metabolic issues of these bacteria is beyond the objectives and needs of this work, and some basic aspects could be consulted elsewhere [33–37].

Acetic acid productivity P_{exp} (g acetic acid·h^{-1}) can be calculated using Equation (22), where AcH_{cycle} is the acetic acid concentration at the end of the production cycle (g acetic acid·L^{-1}), $V_{unloaded}$ is the volume unloaded at the end of the cycle (L) and t_{cycle} is the total duration of the cycle (h).

$$P_{exp} = \frac{AcH_{cycle} \cdot V_{unloaded}}{t_{cycle}} \tag{22}$$

2.2.2. Black-Box Polynomial Model

The black-box polynomial model used for acetic acid productivity, as stated in [24], is based on a second order Box–Behnken model (23), where Y is the dependent or response variable, X are the independent variables or factors and b are the polynomial coefficients. This type of model considers interactions between factors.

$$Y = b_0 + \sum_{i=1}^{n} b_i \cdot X_i + \sum_{\substack{i=1 \\ i<j}}^{n} b_{ij} \cdot X_i \cdot X_j + \sum_{i=1}^{n} b_{ii} \cdot X_i^2 \tag{23}$$

The estimated model was (24), where the factors used were the ethanol concentration in the medium at the time the reactor is unloaded (E_{unload}) and the unloaded volume ($V_{unloaded}$). It was found that the loading flow rate (F_i) was not a significant factor in this case.

$$P_{estimated} = 10.36 + 3.344 \cdot E_{unload} + 0.0118 \cdot V_{unloaded} - 0.413 \cdot E_{unload}{}^2 - 1.01 \cdot 10^{-3}$$
$$\cdot V_{unloaded}{}^2 - 0.02 \cdot E_{unload} \cdot V_{unloaded} \quad (24)$$

Using the above-mentioned experimental design, data used for model estimation are shown in Table 2, where each experiment corresponds to several production cycles (176 cycles in total were carried out to obtain the data shown in the table [24]).

2.3. Multilayer Perceptron (MLP)

The multilayer perceptron is a feed-forward type of ANN widely used to develop non-linear static models comprising an input layer, an output layer and at least one hidden layer of neurons [1]. Each neuron computes a linear combination of their inputs, the coefficients of which are the weights and biases to be estimated by a training algorithm. Then, a continuous and differentiable nonlinear function (activation function) is applied. Usually, MLP-based models use a sigmoid (25) or a hyperbolic tangent sigmoid function (26). The output ranges for which are 0 to 1 and –1 to +1, respectively. Because inputs can range from –∞ to +∞, the previous ranges span the greatest possible difference.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (25)$$

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (26)$$

While the number of inputs and outputs to be used when applying MLP methodology to a specific problem is preset, those of hidden layers and neurons in each layer are not. There are no general rules for selecting such numbers, which are usually chosen by trial and error. Too large a number of hidden layers or neurons per layer may enhance the predictive ability of an ANN but reduce its extrapolability or generalization capability (i.e., the ability of the ANN to provide accurate predictions under different conditions) and computational cost (especially at the training stage).

2.4. Experimental Data and ANN Training

The experimental data were distributed at random between a training set (80% of data) and a validation set (the remainder 20%), following the k-fold cross-validation strategy (specifically, 5-fold cross-validation) [38]. Data were randomly split into 5 groups (four for training and one for testing) for each of the 50 ANNs estimated for each selected number of neurons (3, 5, 10 or 20). Therefore, 50 repeats have been carried out in each case.

3. Results and Discussion

3.1. Comparing the First-Principles and Black-Box Models

Below are compared the first-principles and black-box models applied here to the acetification process for vinegar production. The models are compared in terms of the experimental data used to fit the predictions of acetic acid production.

Figure 1 shows the mean experimental production values and their standard deviations, in addition to the values estimated by the first-principles model under the conditions of Table 1.
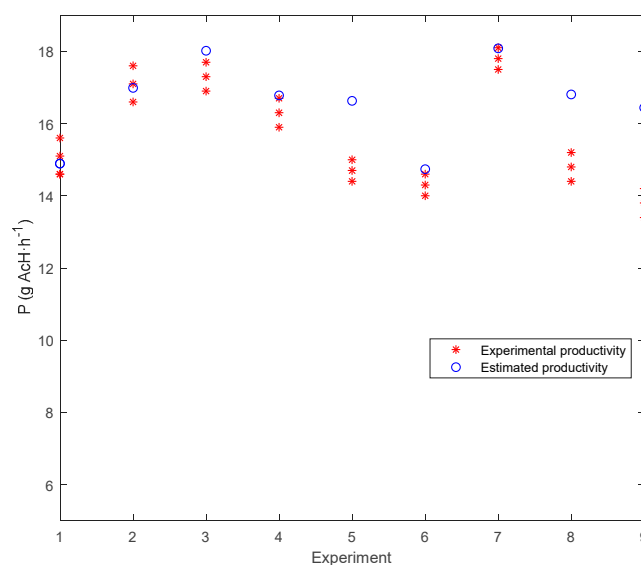
Figure 1. Experimental productivity vs. estimated productivity predicted by the first-principles model.

As can be seen, experiments 5, 8 and 9, were those resulting in the greatest differences between experimental and predicted values. For some reason, the predictions of the first-principles model were not as accurate as those obtained under other conditions. In the previous experiments, the ethanol concentration at the time the reactor was unloaded was low or very low; furthermore, half the reactor content was used as inoculum in the next cycle and the loading rates used fell in the middle of the experimental range. As will be commented further, the scarcity of substrate and the presence of high concentrations of acetic acid under these conditions may have been stressful to acetifying bacteria.

Figure 2 compares the predictions of the polynomial (black-box) model [24] with the experimental data used for estimation (Table 2). As can be seen, the differences between estimated and actual productivity values fell within or very close to the experimental error range. Furthermore, the greatest difference was that of experiment 10, albeit not significant.
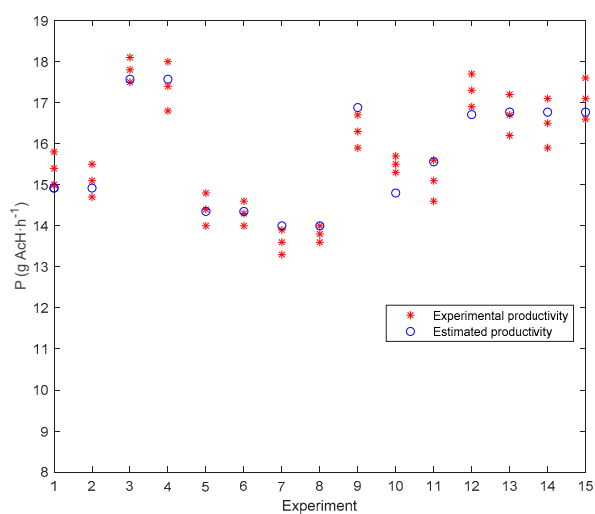
Figure 2. Experimental productivity vs. estimated productivity as predicted by the polynomial model.

Figure 3 compares the productivity values estimated by the first-principles model under the experimental conditions used to fit it. As can be seen, the greatest differences were those of experiments 7 and 8 (Table 2), both of which used extreme conditions (viz., unloading of the reactor

at very low substrate concentrations and low volumes of reaction medium). The resulting low ethanol availability and high acidity must have been highly stressful for bacteria in the medium, so the conditions of these two experiments may have fallen outside the range where the first-principles model would have held and led to inaccurate predictions as a result.

Using the polynomial model to estimate acetic production under the experimental conditions used to fit the first-principles model —under continuous loading conditions only— provided very accurate predictions. As can be seen from Figure 4, the differences between predicted and experimental values fell within the experimental range. Therefore, the polynomial model also afforded accurate predictions under the conditions used to fit the first-principles model.
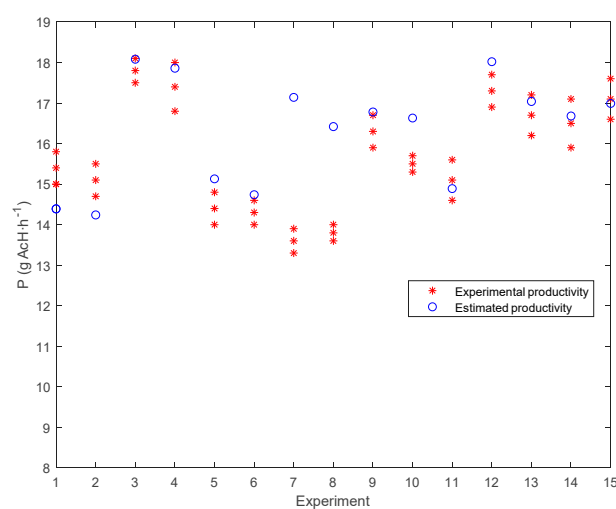
Figure 3. Experimental productivity vs. estimated productivity as predicted by the first-principles models under the conditions used to fit the polynomial model.
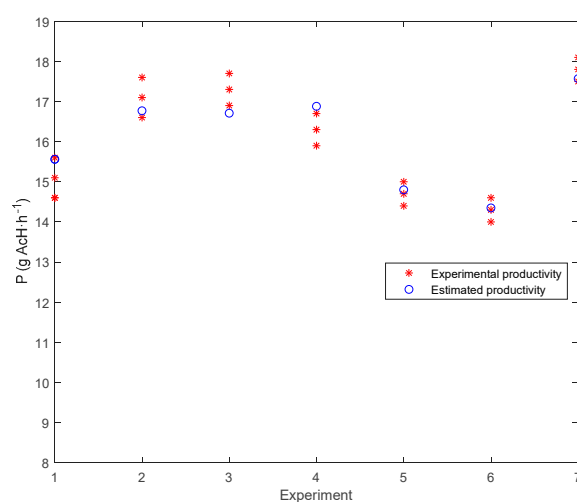
Figure 4. Experimental productivity vs. estimated productivity as predicted by the polynomial model under the conditions used to fit the first-principles model.

The predictive ability of the two models was compared in greater detail in terms of the productivity response surfaces they provided.

The ethanol concentration at unloading time (E_{unload}) and the percent unloaded volume ($V_{unloaded}$) were assumed to span the range from 0.5 to 3.5 % v/v and from 25 to 75 %, respectively. Figure 5a compares the results of the two models at a fixed loading flow rate F_i of 0.01, 0.035 or 0.06 L·min⁻¹. Only one response surface is shown for the polynomial model because acetic acid production is independent of the loading rate [24]. Figure 5b shows the errors or relative residuals between the production estimates obtained with the first-principles and black-box model. Furthermore, as can be

seen from Figure 5a, the response surfaces obtained were virtually identical irrespective of loading flow rate —only at low rates and also low ethanol concentrations at unloading time were differences in productivity appreciable. Therefore, the reactor loading flow rate was virtually uninfluential within the experimental ranges examined.
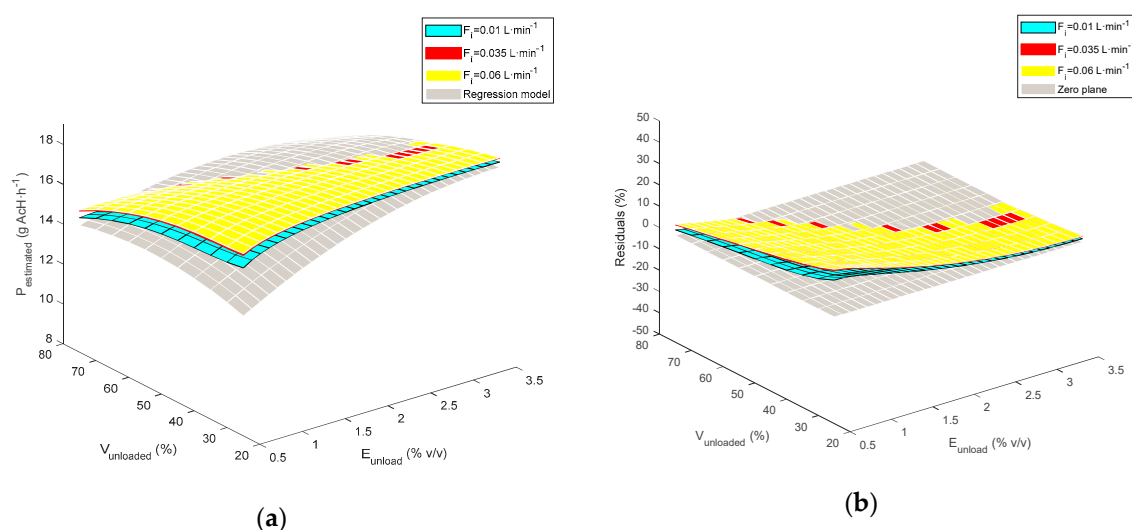

Figure 5. (**a**) Response surfaces obtained with the two models at a variable loading flow rate; (**b**) relative errors from the polynomial model.

The greatest differences between the predictions of the two models were observed at low ethanol concentrations and unloaded reactor volumes (i.e., under the most extreme conditions for bacterial growth, which included ethanol scarcity, high acidity and scant replenishment of the medium). If one considers the previous finding that the polynomial model was more accurate in predicting the experimental results, then the first-principles model was seemingly unable to accurately predict acetic acid production under such extreme conditions.

From the response surfaces and relative errors obtained over the F_i range from 0.01 to 0.06 L·min⁻¹ and $V_{unloaded}$ range from 25 to 75 % at a fixed E_{unload} value of 0.5, 2.0 or 3.5 % v/v (Figures 6–8), it follows that the residuals found at the latter two E_{unload} values were less than 10% irrespective of the experimental conditions. At E_{unload} = 0.5 % v/v, however, the residuals were considerably greater and, again, increased with increasing unloaded volume.
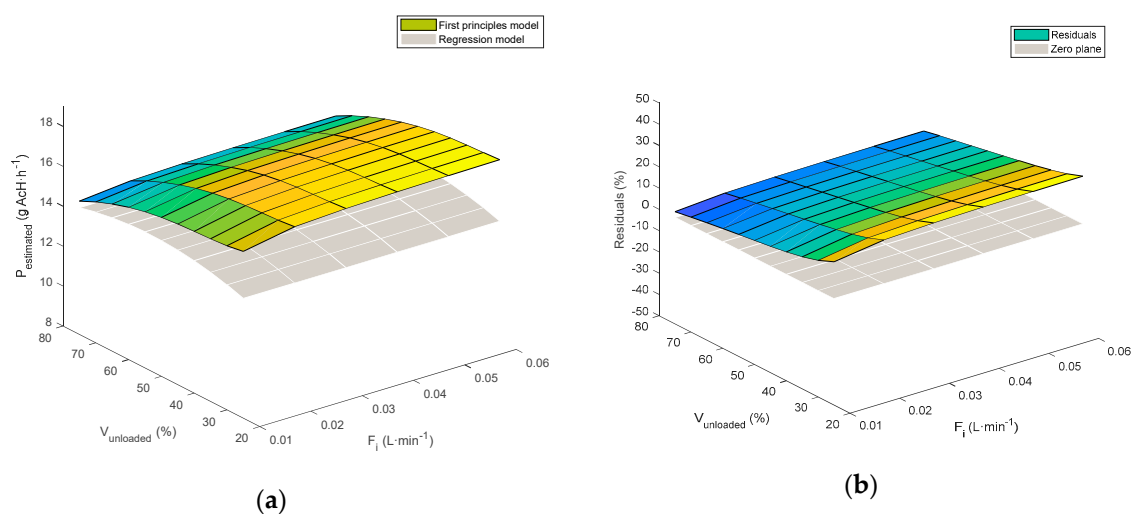
Figure 6. (**a**) Response surfaces obtained with both models at E_{unload}=0.5 % v/v; (**b**) relative errors from the polynomial model.
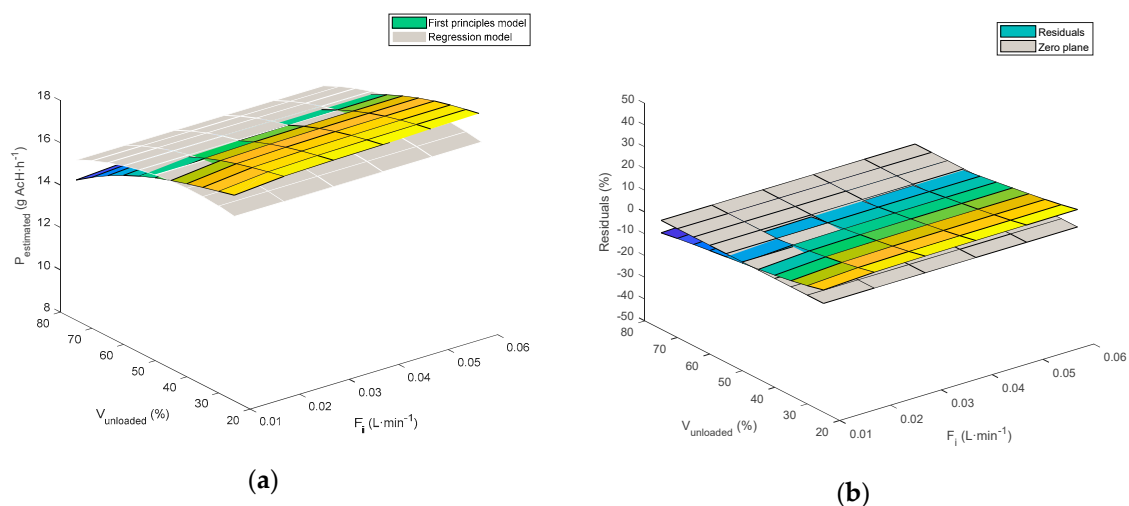
Figure 7. (**a**) Response surfaces obtained with both models at E_{unload}=2 % v/v; (**b**) relative errors from the polynomial model.
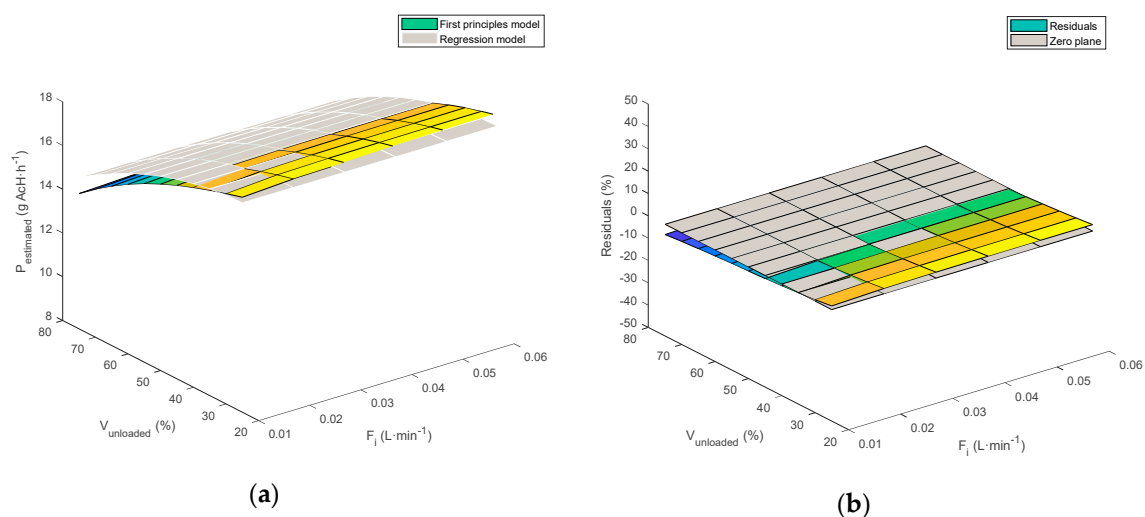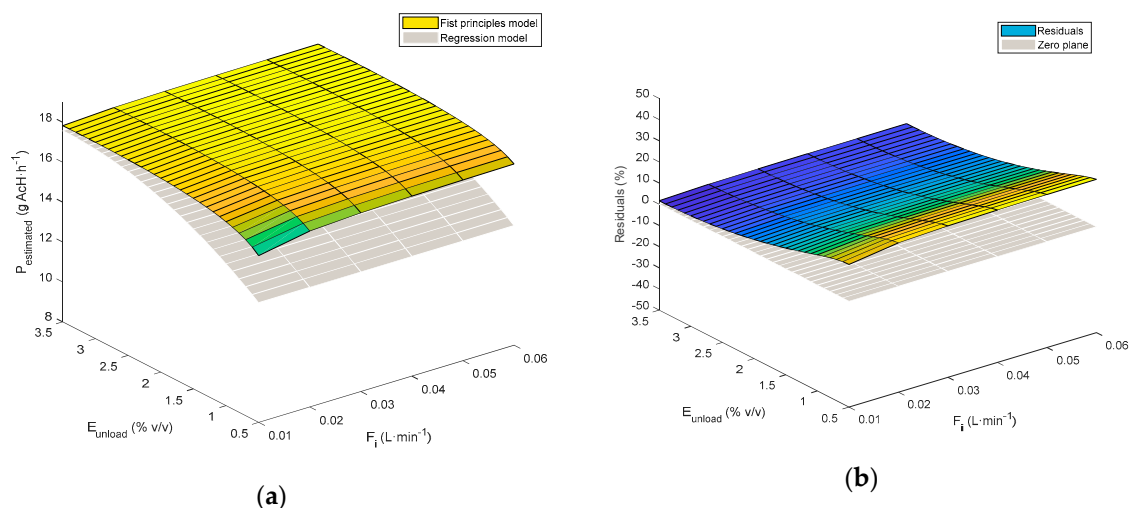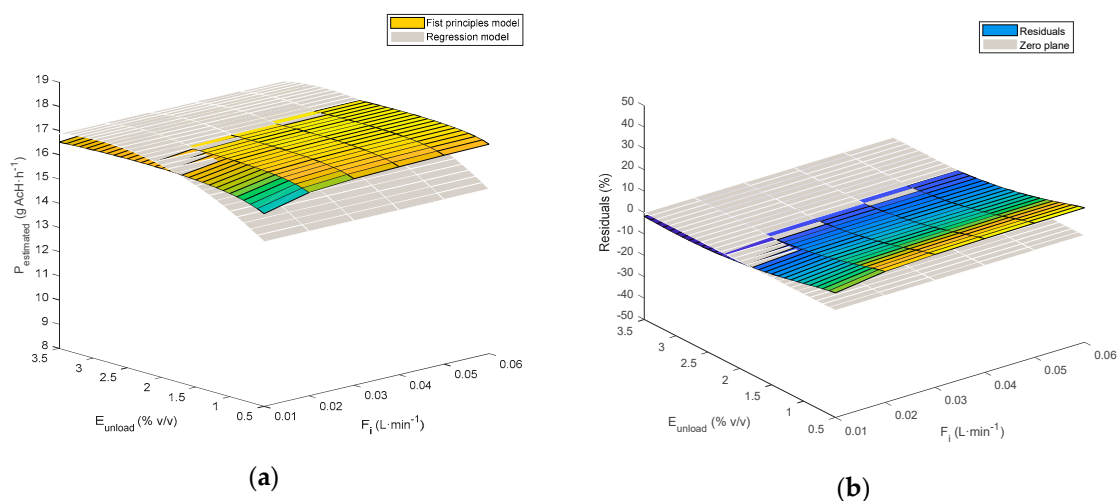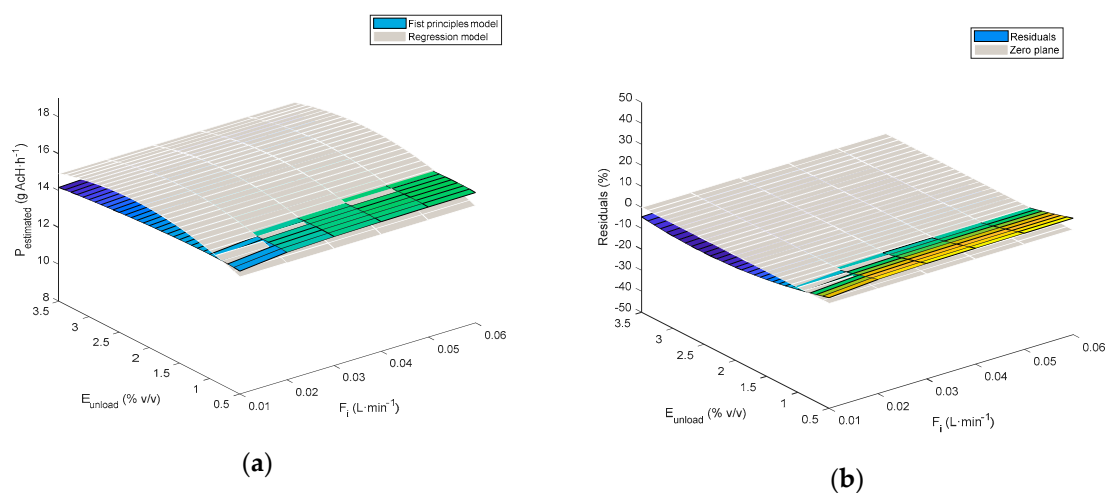

Figure 8. (**a**) Response surfaces obtained with both models at E_{unload}=3.5 % v/v; (**b**) relative errors from the polynomial model.

Figures 9–11 show the response surfaces and relative errors obtained at a constant volume $V_{unloaded}$ of 25, 50 or 75 %, a F_i value from 0.01 to 0.06 L min⁻¹ and an E_{unload} value from 0.5 to 3.5 % v/v. Again, the greatest differences were observed with low unloaded volumes and low ethanol concentrations at unloading time, with residuals exceeding 20 % in some cases (Figure 9b).

Figure 9. (**a**) Response surfaces obtained with both models at $V_{unloaded}$=25%; (**b**) relative errors from the polynomial model.

Figure 10. (**a**) Response surfaces obtained with both models at $V_{unloaded}$=50%; (**b**) relative errors from the polynomial model.

Figure 11. (**a**) Response surfaces obtained with both models at $V_{unloaded}$=75%; (**b**) relative errors from the polynomial model.

Based on the previous results, both models predicted roughly the same productivity values under conditions favoring bacterial growth (viz., no alcohol depletion, low acidity); additionally, such values were very similar to their experimental counterparts.

Conversely, the first-principles model was much less accurate in predicting acetic acid production under the most unfavorable conditions for bacterial growth. This may have been the result of its kinetic equations disregarding the effect of some factor under extreme conditions and/or the difficulty of estimating the parameters of such equations (i.e., of solving the structural and practical identifiability problems they pose).

Although first-principles models are theoretically valid over a wide range of conditions, their results are strongly dependent on the kinetic equations used, which can rarely consider all phenomena potentially affecting acetifying bacteria under conditions other than those affording unrestricted growth (exponential growth phase). For example, based on eq. 9, which is the kinetic equation reflecting ethanol-based cell growth [6,8], there will be bacterial growth regardless of how low the ethanol concentration in the medium is. This, however, may not be the case since a scarcity of substrate can lead many microbes to aim their metabolic activity at maintenance functions. Based on eq. 10, which describes acetic acid-based cell growth [6,8], acetic acid only acts as a bacterial growth inhibitor when, in fact, it may also act as a booster at very low concentrations [39–41].

As can be inferred from the above-described problems, accurate first-principles models are more difficult to construct than are other types of models such as those based on polynomial regression equations. In contrast, polynomial models usually require greater numbers of experimental data and are less readily extrapolated to other scenarios. However, the alternative models can be more easily and systematically developed; furthermore, they provide accurate predictions—at least under the range of experimental conditions used in their development. The increased predictive ability and prediction quality of polynomial models allows them to be used as references for improving first-principles models or even to construct alternative black-box approaches such as those based on artificial neural networks (ANNs).

3.2. Artificial Neural Network Model for Productivity in the Acetic Fermentation Process

Artificial neural networks allow effective black-box models to be developed; although a lot of experimental data are normally needed to carry out the training and validation of ANNs, in this case, thanks to the availability of two previous models, a feasibility analysis about their use has been completed, even though the available data can be scant. Like polynomial regression models, ANN-based models can be constructed from experimental data obtained under conditions spanning the operational ranges of the target bioprocess. Furthermore, ANN-based models established from appropriate datasets to avoid overfitting are usually easy to extrapolate to alternative conditions.

In this work, we used a neural network in the form of a multilayer perceptron (MLP) comprising a single hidden layer containing a variable number of neurons and an output layer. The sum of the weighted inputs and bias for each neuron in the hidden layer was used as input of a hyperbolic tangent sigmoid transformation function to obtain its output. The output layer differed from the hidden layer in that the former used a linear transfer function, which is better suited to non-linear regression problems such as that addressed here because it imposes no restriction on output values. This type of network represents a universal approximator to any non-linear function [42] provided an adequate number of neurons is used in the hidden layer. Therefore, it allows non-linear models of arbitrary accuracy to be constructed.

In the modelling process, the experimental data previously used to fit the polynomial and first-principles models (Tables 1 and 2, data for continuous loading operation only) were used to train multilayer perceptrons with 3, 5, 10 and 20 neurons in the hidden layer (50 networks in all cases) by supervised learning. The variables used as ANN inputs were the ethanol concentration in the medium at the time the reactor is unloaded (E_{unload}), the unloaded volume ($V_{unloaded}$) and the loading flow rate (F_i). The experimental data distribution and ANN training strategy were described in section 2.4. Data were fitted by back-propagation in combination with the Levenberg–Marquardt method to solve the least-squares problem arising in estimating the parameter values for each

network with weights and biases as decision variables. The cost function for each network was taken to be the Mean Square Error (MSE) between predicted and experimental acetic acid production values. The starting parameter values for each network were chosen at random in order to allow the optimization algorithm to obtain different solutions.

The 50 networks used in each case were used to select those providing the best MSE compromise for the estimation and validation sets (viz., one where neither error was high relative to the other networks in order to avoid overfitting and poorer fitting to the estimation data). The training process was stopped when no improvement in estimation error or increase in validation error was observed after 3 epochs (i.e., three iterations of the training algorithm).

By way of example, Figures 12 and 13 show the results of the fitting of the network most accurately estimated with 3 neurons (viz., no. 1 in Table 3). Figure 12 shows the variation of the training and validation MSE as a function of the number of epochs. As can be seen, both MSE values stopped decreasing after epoch 23, so the criterion established to halt training was fulfilled by stopping the process at epoch 26. Figure 13 compares the bisector of the first quadrant (i.e., perfect fitting) to the linear regression between the experimental productivity and that predicted by the ANN model under identical operating conditions. The coefficient of determination of the regression was $R^2 = 0.97452$, so the fitting was quite good.
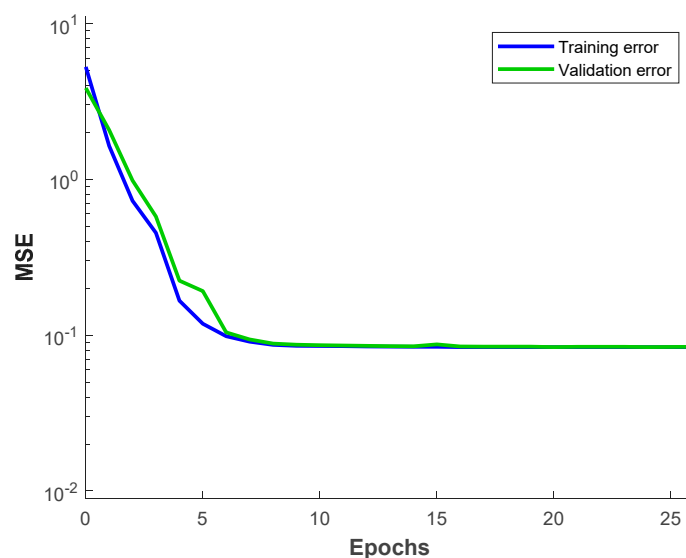

Figure 12. Variation of the training and validation MSE (Mean Square Error) with the number of epochs.
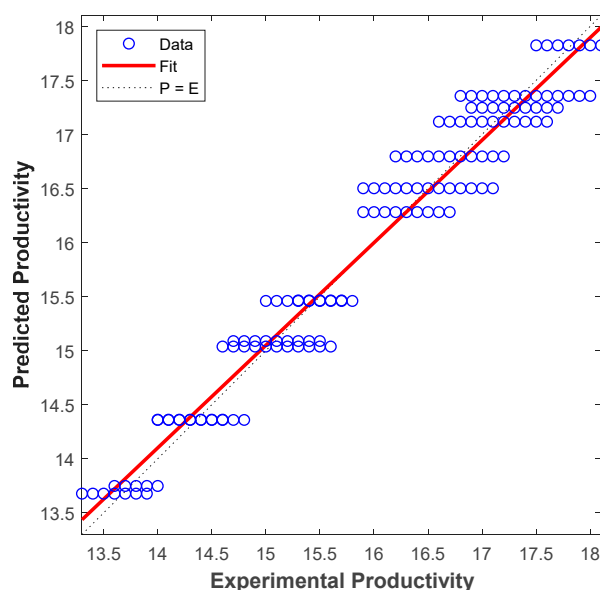
Figure 13. Linear regression between experimental productivity and predicted values by ANN.

Table 4 shows the results provided by four selected networks with a different number of neurons in the hidden layer. The coefficient of determination shown is that for the linear regression between the predicted production values of each network and the experimental values obtained under each set of operating conditions. As can be seen, the estimates were all similarly good, the only appreciable difference being that the number of epochs needed to estimate the networks increased with increasing number of neurons in the hidden layer.

Table 4. Fitting of ANNs with a variable number of neurons in the hidden layer.

Network no.	Number of Neurons	Estimation MSE	Validation MSE	Estimation R^2	Number of Epochs
1	3	0.0839	0.0838	0.97452	20
2	5	0.0836	0.0795	0.975	7
3	10	0.0838	0.0852	0.975	6
4	20	0.0831	0.0822	0.97485	3

The validity of the predictions over the variation range of each variable is illustrated in Figures 14–17, which compare the response surfaces for the networks in Table 4 with that constructed from the polynomial model —which was used as reference for the above-described reasons. By way of example, the graphs in the figures were obtained at a medium loading flow rate (F_i = 0.04 L·min^{-1}), and $V_{unloaded}$ and E_{unload} values spanning the ranges from 25 to 75 % and from 0.5 to 3.5 % v/v, respectively.

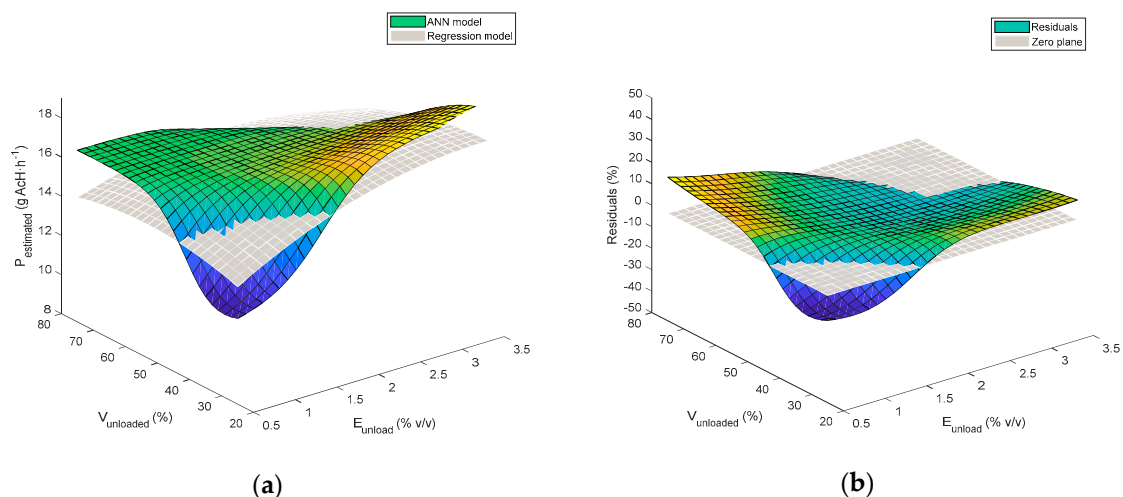(**a**)　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 14. (**a**) Response surfaces obtained with network no. 1 and the polynomial model; (**b**) relative errors from the polynomial model.
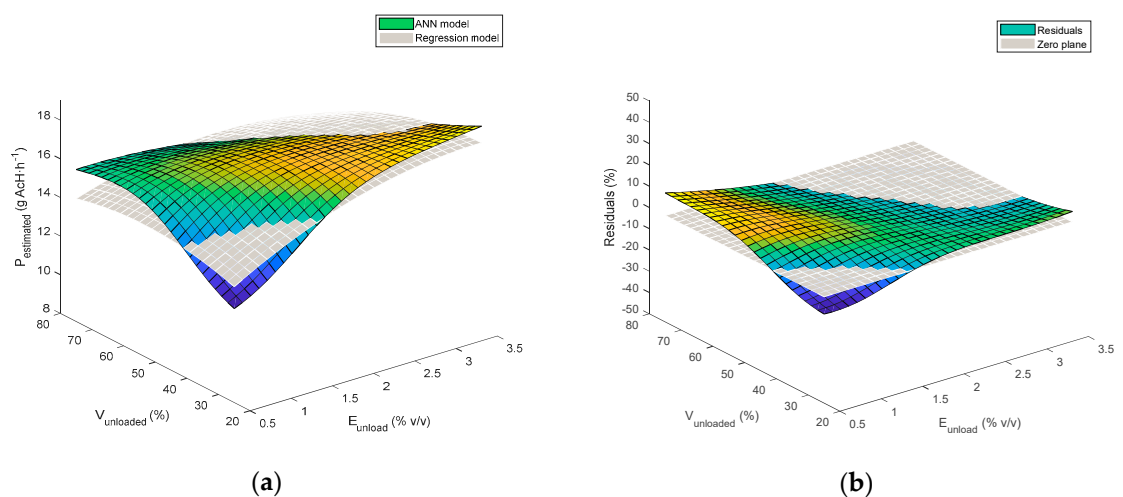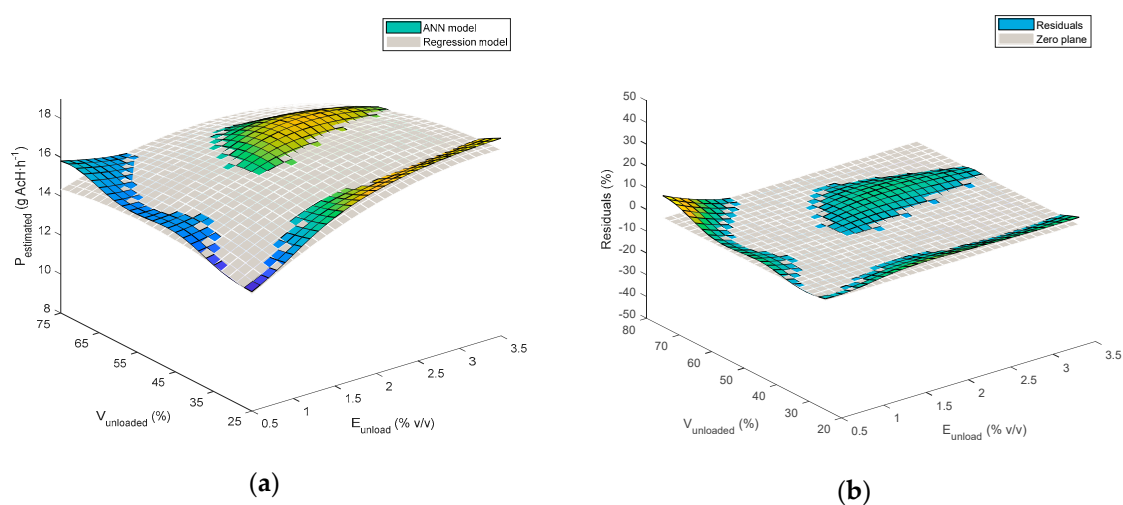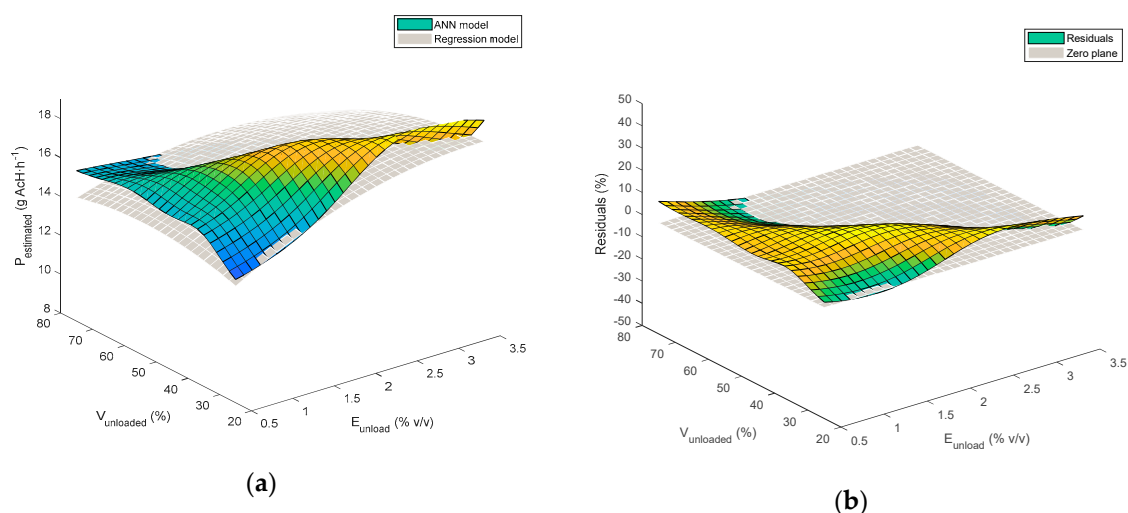

(**a**)　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 15. (**a**) Response surfaces obtained with network no. 2 and the polynomial model; (**b**) relative errors from the polynomial model.

(**a**)　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 16. (**a**) Response surfaces obtained with network no. 3 and the polynomial model; (**b**) relative errors from the polynomial model.

Figure 17. (**a**) Response surfaces obtained with network no. 4 and the polynomial model; (**b**) relative errors from the polynomial model.

Despite the overall goodness of the estimates (Table 4), the results for conditions outside the experimental data were not so good. This led us to examine the quality of the predictions obtained under conditions other than those of Table 4. It should be noted that the differences between the estimation and validation MSE values were relatively small. Furthermore, although the loading flow rate was scarcely influential, its actual effect was checked by comparing networks constructed at four different flow rates, namely: 0.01, 0.02, 0.04 and 0.06 L·min^{-1}.

By way of example, Figure 18 compares the results obtained at each flow rate with the best network containing three neurons in the hidden layer and the response surface for the polynomial model. As can be seen, these networks do not coincide with no. 1 in Table 4. This suggested that alternative networks among those trained here could perform better than those initially selected. In fact, as can be seen from Figure 18, the loading flow rate was scarcely influential —the response surfaces obtained at the four different values were very similar. Therefore, the discussion that follows applies to a single, medium flow rate value (F_i = 0.04 L·min^{-1}).
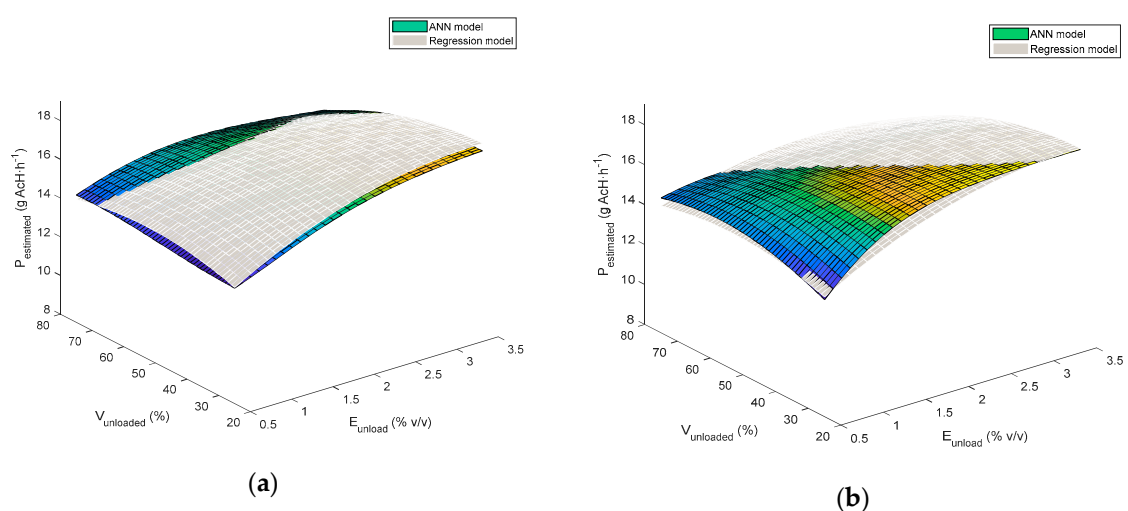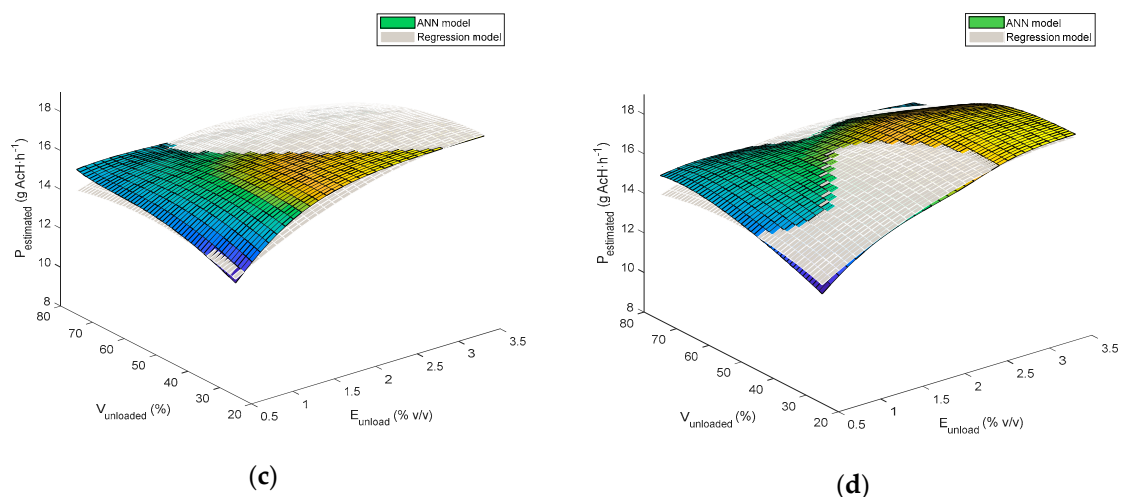
Figure 18. Response surfaces obtained with the best network containing three neurons in the hidden layer as compared with the polynomial model. Simulations at a loading flow rate of (**a**) 0.01; (**b**) 0.02; (**c**) 0.04; (**d**) 0.06 L·min⁻¹.

None of the networks constructed with other numbers of neurons in the hidden layer that improved on the results of the polynomial model coincided with those in Table 4. Therefore, the network selection criterion used with the relatively narrow range of experimental conditions available, which was based on the goodness of fitting of the networks, was probably not the most suitable. Table 5 shows the MSE relative to the polynomial model of the networks of Table 4 and those providing the best results including intermediate experimental conditions, the response surfaces of which are shown in Figure 19.

Table 5. MSE for the networks of Table 4 as compared with the polynomial model ($F_i = 0.04$ L·min⁻¹).

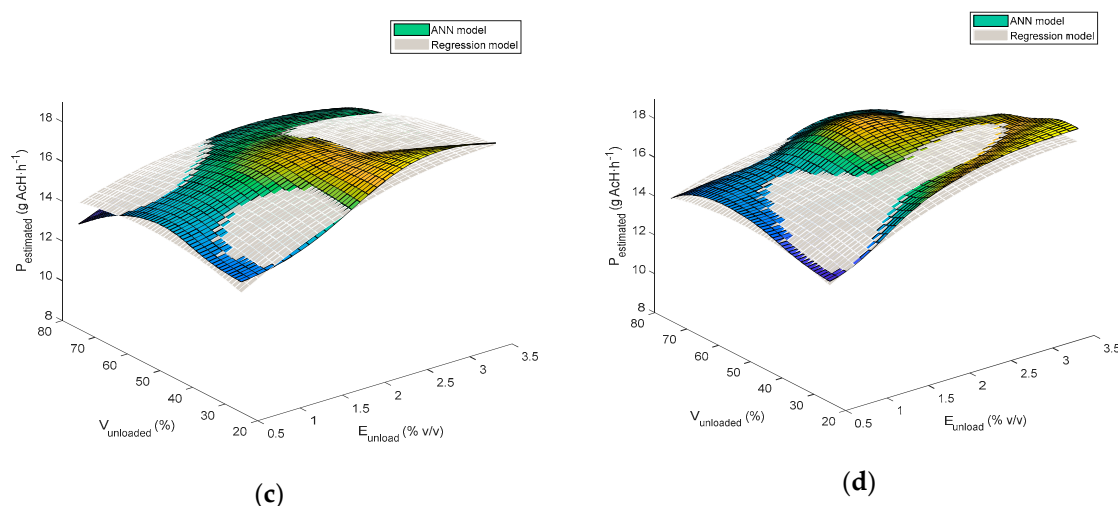| Number of Neurons	MSE (Networks of Table 4)	MSE (Best Prediction Networks)
3	0.9446	0.1304
5	0.3857	0.0633
10	0.1704	0.0879
20	1.1636	0.1336
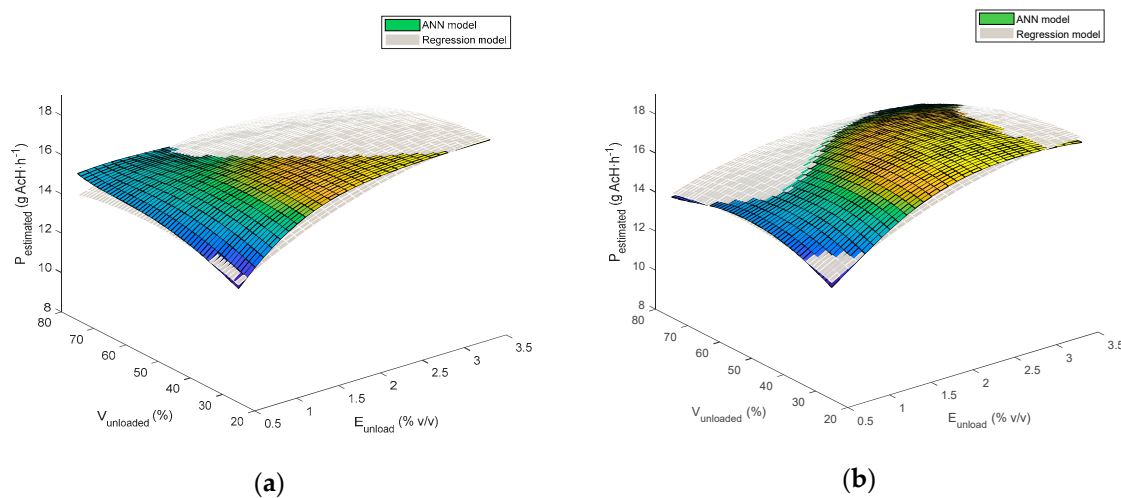
Figure 19. Response surfaces for the best predictive networks with (**a**) 3; (**b**) 5; (**c**) 10; (**d**) 20 neurons in the hidden layer in comparison with the polynomial model (F_i = 0.04 L·min^{-1}).

Based on this figure and on the MSE values of Table 5, the best predictions were obtained with 5 neurons in the hidden layer, albeit with only slight differences from the network with 10 neurons in that layer. This suggests that the most suitable number of neurons in the hidden layer of our predictive ANNs was 5 or a slightly greater number.

Based on the foregoing, selecting an effective ANN for modelling a bioprocess over a broad enough range of operating conditions when experimental data are scant is rather difficult. Nevertheless, there is a possibility to find ANNs with a better fit; the main problem is to choose suitable assessment criteria when reference models do not exist, an issue that might be analysed in future works. In fact, this would be of great interest, considering that obtaining large amounts of experimental data from a bioprocess is a difficult, time-consuming task. As shown here, in this case, one or more artificial networks were able to be constructed to predict its results under a broad range of operating conditions only because the surface response of a polynomial model for the given bioprocess was known.

Then, polynomial modelling approaches are subject to fewer problems than ANN-based predictive models when experimental data are scant. Additionally, the polynomial regression can be easily obtained through a systematic process and its predictions are usually of good quality.

4. Conclusions

Existing mathematical models provide a powerful tool for examining, analyzing and optimizing bioprocesses, each type of model having specific advantages and disadvantages.

The acetification process used in the industrial production of vinegar has largely been modelled with mechanistic (first-principles) or polynomial (black-box) models. The former models have the disadvantages that their kinetic equations are difficult to establish and that estimation of their parameters is usually subject to structural and/or practical identifiability problems. However, mechanistic models afford better understanding of the internal aspects of the target processes and usually hold broader ranges of operating conditions. On the other hand, polynomial regression black-box models are easier to develop but use to have more limited validity ranges than first-principles models. A comparison of the predictions of acetic acid production with the two models revealed that the mechanistic model performed worse than the polynomial model under extreme conditions for bacterial growth, namely, a low substrate (ethanol) concentration and a high product (acetic acid) concentration. This suggests that the kinetic equations of the mechanistic model failed to consider factors such as cell growth below a given substrate concentration or a boosting —not purely inhibitory— effect of the product (acetic acid) at low concentrations. One other reason for the differences may be inaccuracy in estimating the parameter values of the kinetic equations by effect of

identifiability problems. The polynomial model is seemingly accurate irrespective of the operating conditions —even those under which the first-principles model was established. This led us to use such a model as reference for comparison of the predictions of the other models.

Black-box models using artificial neural networks (ANNs) of the multilayer perceptron (MLP) type have been analyzed. The networks contained a single hidden layer and were used in combination with all experimental data available for the acetification process, some for training and other for validation, and the mean square error as training target function. A comparison of the results obtained with ANNs containing a variable number of neurons in the hidden layer and the predictions of the polynomial model revealed the optimum number of neurons to be 5–10. However, the predictions of the networks with the smallest MSE values under operating conditions in the middle of the range used for training were not good, as expected, which made identifying the most suitable network rather difficult or impossible without a reference model. Because of the large number of degrees of freedom of this type of model, the problem largely arises from the usually small number of experiments available for a bioprocess, but if it were possible to find the suitable selection criteria, as has been shown in this work, ANNs with a better fit can be found. Due to the lack of a very high amount of experimental data, future research with the aim of developing these selection criteria could be of great interest.

Based on the results, the best choice for modelling acetic fermentation in terms of ease of development and accuracy of predictions irrespective of the particular operating conditions is the polynomial regression black-box model.

References

1. García-García, I.; Santos-Dueñas, I.M.; Jiménez-Ot, C.; Jiménez-Hornero, J.E.; Bonilla-Venceslada, J.L. Vinegar engineering. In *Vinegars of the World*; Solieri, L., Giudici, P., Eds.; Springer: Milano, Italy, 2009; Chapter 9, pp. 97–120. doi:10.1007/978-88-470-0866-3_6.
2. Julien, C.; Whitford, W. Bioreactor Monitoring, Modeling, and Simulation. *BioProcess Int.* Suppl. **2007**, *5(1)*.
3. Agger, T.; Nielsen, J. Mathematical Modelling of Microbial Processes-Motivation and Means. In *Engineering and Manufacturing for Biotechnology. Focus on Biotechnology*; Hofman, M., Thonart, P., Eds.; Springer: Dordrecht, The Netherlands, 2001; Volume 4, pp. 61–75. doi:10.1007/0-306-46889-1_5.
4. Gernaey, K.V.; Lantz, A.E.; Tufvesson, P.; Woodley, J.M.; Sin, G. Application of mechanistic models to fermentation and biocatalysis for next-generation processes. *Trends Biotechnol.* **2010**, *28*, 346–354. doi:10.1016/j.tibtech.2010.03.006.
5. Jiménez-Hornero, J.E. Contribuciones al Modelado y Optimización del Proceso de Fermentación Acética. Ph.D. Thesis, Universidad Nacional de Educación a Distancia, Madrid, Spain, 2007.

6. Jiménez-Hornero, J.E.; Santos-Dueñas, I.M.; García-García, I. Optimization of biotechnological processes. The acetic acid fermentation. Part I: The proposed model. *Biochem. Eng. J.* **2009**, *45*, 1–6. doi:10.1016/j.bej.2009.01.009.

7. Banga, J.R.; Balsa-Canto, E.; Moles, C.G.; Alonso, A.A. Improving food processing using modern optimization methods. *Trends Food Sci. Technol.* **2003**, *14*, 131–144. doi:10.1016/S0924-2244(03)00048-7.

8. Jiménez-Hornero, J.E.; Santos-Dueñas, I.M.; Garcia-Garcia, I. Structural identifiability of a model for the acetic acid fermentation process. *Math. Biosci.* **2008**, *216*, 154–162. doi:10.1016/j.mbs.2008.09.004.

9. Jiménez-Hornero, J.E.; Santos-Dueñas, I.M.; Garcia-Garcia, I. Optimization of biotechnological processes. The acetic acid fermentation. Part II: Practical identifiability analysis and parameter estimation. *Biochem. Eng. J.* **2009**, *45*, 7–21. doi:10.1016/j.bej.2009.01.010.

10. Ljung, L. *System Identification: Theory for the User*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1987; p. 519.

11. Santos-Dueñas, I.M. Modelización Polinominal y Optimización de la Acetificación de Vino. Ph.D. Thesis, Universidad de Córdoba, Córdoba, Spain, 2009.

12. Miller, N.; Miller, C. *Estadística y Quimiometria para Quimica Analitica*, 4th ed.; Pearson educación SA: Madrid, España, 2002; p. 296.

13. Anderson, J.A. *An Introduction to Neural Networks*; MIT Press: Cambridge, MA, USA, 1995; p. 672.

14. Bogaerts, P.; Hanus, R. Macroscopic Modelling of Bioprocesses with a View to Engineering Applications. In *Engineering and Manufacturing for Biotechnology. Focus on Biotechnology*; Hofman, M., Thonart, P., Eds.; Springer: Dordrecht, The Netherlands, 2001; Volume 4, pp. 77–109. doi:10.1007/0-306-46889-1_6.

15. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. doi:10.1038/323533a0.

16. Jeong, D.H.; Lee, J.M. Enhancement of modifier adaptation scheme via feedforward decision maker using historical disturbance data and deep machine learning. *Comput. Chem. Eng.* **2018**, *108*, 31–46. doi:10.1016/j.compchemeng.2017.08.005.

17. Petsagkourakis, P.; Orson Sandoval, I.; Bradford, E.; Zhang, D.; del Rio-Chanona, E.A. Reinforcement Learning for Batch-to-Batch Bioprocess Optimisation. *Comput. Aided Chem. Eng.* **2019**, *46*, 919–924. doi:10.1016/B978-0-12-818634-3.50154-5.

18. Nelles, O. *Nonlinear System Identification*; Springer-Verlag: Berlin/Heidelberg, Germany, 2001; p. 785. doi:10.1007/978-3-662-04323-3.

19. Baena-Ruano, S.; Jiménez-Ot, C.; Santos-Dueñas, I.M.; Jiménez-Hornero, J.E.; Bonilla-Venceslada, J.L.; Álvarez-Cáliz, C.; García-García, I. Influence of the final ethanol concentration on the acetification and production rate in the wine vinegar process. *J. Chem. Technol. Biotechnol.* **2010**, *85*, 908–912. doi:10.1002/jctb.2368.

20. Álvarez-Cáliz, C.; Santos-Dueñas, I.M.; García-Martínez, T.; Cañete-Rodríguez, A.M.; Millán-Pérez, M.C.; Maurico, J.C.; García-García, I. Effect of biological ageing of wine on its nitrogen composition for producing high quality vinegar. *Food Bioprod. Process* **2014**, *92*, 291–297. doi:10.1016/j.fbp.2013.07.005.

21. Baena-Ruano, S.; Jiménez-Ot, C.; Santos-Dueñas, I.M.; Cantero-Moreno, D.; Barja, F.; García-García, I. Rapid method for total, viable and non-viable acetic acid bacteria determination during acetification process. *Process Biochem.* **2006**, *41*, 1160–1164. doi:10.1016/j.procbio.2005.12.016.

22. García-García, I.; Cantero-Moreno, D.; Jiménez-Ot, C.; Baena-Ruano, S.; Jiménez-Hornero, J.; Santos-Duenas, I.; Bonilla-Venceslada, J.L.; Barja, F. Estimating the mean acetification rate via on-line monitored changes in ethanol during a semi-continuous vinegar production cycle. *J. Food Eng.* **2007**, *80*, 460–464. doi:10.1016/j.jfoodeng.2006.05.028.

23. Álvarez-Cáliz, C.; Santos-Dueñas, I.M.; Cañete-Rodríguez, A.M.; García-Martínez, T.; Maurico, J.C.; García-García, I. Free amino acids, urea and ammonium ion contents for submerged wine vinegar production: Influence of loading rate and air-flow rate. *Acetic Acid Bact.* **2012**, *1*, 1–6. doi:10.4081/aab.2012.e1.

24. Santos-Dueñas, I.M.; Jimenez-Hornero, J.E.; Cañete-Rodríguez, A.M.; García-García, I. Modeling and optimization of acetic acid fermentation: A polynomial-based approach. *Biochem. Eng. J.* **2015**, *99*, 35–43. doi:10.1016/j.bej.2015.03.002.

25. Packett, R.L.; Burman, J.P. The desing of optimum multi-factorial experiments. *Biometrika* **1946**, *33*, 305–325. doi:10.1093/biomet/33.4.305.

26. Box, G.E.P.; Behnken, D.W. Some new three-level designs for the study of quantitative variables, *Technometrics* **1960**, *2*, 455–475. doi:10.2307/1266454.

27. Morgan, E. *Chemometrics: Experimental Design*; Wiley: Chichester, UK, 1991; p. 275.

28.　Castro Mejías, R.; Natera Marín, R.; García Moreno, M.V.; García Barroso, C. Optimisation of headspace solid-phase microextraction for analysis of aromatic compounds in vinegar. *J. Chromatogr. A* **2002**, *953*, 7–15. doi:10.1016/S0021-9673(02)00122-X.

29.　Ramis Ramos, G.; García Álvarez-Coque, M.C. *Quimiometría*; Sintesis: Madrid, Spain, 2001; p. 240.

30.　Grierson, D.E. Pareto multi-criteria decision making. *Adv. Eng. Inform.* **2008**, *22*, 371–384. doi:10.1016/j.aei.2008.03.001.

31.　Jiménez-Hornero, J.E.; Santos-Dueñas, I.M.; Garcia-Garcia, I. Optimization of biotechnological processes. The acetic acid fermentation. Part III: Dynamic optimization. *Biochem. Eng. J.* **2009**, *45*, 22–29. doi:10.1016/j.bej.2009.01.011.

32.　García-García, I.; Jiménez-Hornero, J.E.; Santos-Dueñas, I.M.; González-Granados, Z.; Cañete-Rodríguez, A.M. Modelling and optimization of acetic acid fermentation. In *Advances in Vinegar Production*; Bekatorou, A., Ed.; CRC Press: Boca Raton, FL, USA, 2019; Chapter 15, pp. 97–120. ISBN 978-0-815-36599-0.

33.　Saichana, N.; Matsushita, K.; Adachi, O.; Frébort, I.; Frebortova, J. Acetic acid bacteria: A group of bacteria with versatile biotechnological applications. *Biotechnol. Adv.* **2015**, *33*, 1260–1271, doi:10.1016/j.biotechadv.2014.12.001.

34.　Matsushita, K.; Toyama, H.; Tonouchi, N.; Okamoto-Kainuma, A. (Eds.) *Acetic Acid Bacteria: Ecology and Physiology*; Springer Japan: Osaka, Japan, 2016; ISBN 978-4-431-55931-3.

35.　Mamlouk, D.; Gullo, M. Acetic Acid Bacteria: Physiology and Carbon Sources Oxidation. *Indian J. Microbiol.* **2013**, *53*, 377–384, doi:10.1007/s12088-013-0414-z.

36.　Deppenmeier, U.; Ehrenreich, A. Physiology of Acetic Acid Bacteria in Light of the Genome Sequence of Gluconobacter oxydans. *MMB* **2009**, *16*, 69–80, doi:10.1159/000142895.

37.　Adler, P.; Frey, L.J.; Berger, A.; Bolten, C.J.; Hansen, C.E.; Wittmann, C. The key to acetate: Metabolic fluxes of acetic acid bacteria under cocoa pulp fermentation-simulating conditions. *Appl. Environ. Microbiol.* **2014**, *80*, 4702–4716, doi:10.1128/AEM.01048-14.

38.　James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning with Application in R.*; Springer: Berlin, Germany, 2017; ISBN 978-1-461-47137-0.

39.　Park, Y.S.; Fukaya, M.; Okumura, H.; Kawamura, Y.; Toda, K. Production of acetic acid by a repeated batch culture with cell recycle of *Acetobacter aceti*. *Biotechnol. Lett.* **1991**, *13*, 271–276. doi:10.1007/BF01041483.

40.　Romero, L.E.; Gómez, J.M.; Caro, I.; Cantero. D. A kinetic model for growth of Acetobacter aceti in submerged culture. *Chem. Eng. J. Biochem. Eng. J.* **1994**, *54*, B15–B24. doi:10.1016/0923-0467(93)06046-S.

41.　González-Sáiz, J.M.; Pizarro, C.; Garrido-Vidal, D. Evaluation of kinetic models for industrial acetic fermentation: Proposal of a new model optimized by genetic algorithms. *Biotechnol. Prog.* **2003**, *19*, 599–611. doi:10.1021/bp0256871.

42.　Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257. doi:10.1016/0893-6080(91)90009-T.