Article

Modeling Based Identifiability and Parametric Estimation of an Enzymatic Hydrolysis Process of Amylaceous Materials

Daniel Padierna-Vanegas,* Juan Camilo Acosta-Pavas,* Laura María Granados-García,*
and Héctor Antonio Botero-Castro*

Read Online

ACCESS | 📊 Metrics & More | 📰 Article Recommendations | 🆂 Supporting Information

ABSTRACT: This work presents the modeling of an enzymatic hydrolysis process of amylaceous materials considering the parameter identification problem as a basis for the construction of the model. For this, a modeling methodology is modified in order to apply the identifiability property and improve the proposed model structure. A brief theoretical explanation of the identifiability is described. This concept is based on the observability property of a nonlinear dynamic system. The used methodology is based on the phenomenological based semiphysical model (PBSM). This methodology visualizes that the structure of a dynamic model can only improve with new mass or energy balances suggested by model suppositions. Additionally, a computer algorithm is included in the methodology to validate if the model is structurally locally identifiable or know if the parameters are unidentifiable. Also, an optimization algorithm is used to obtain the numeric values of the identifiable parameters and, hence, guarantee the validity of the result. The methodology focuses on the liquefaction and saccharification stages of an enzymatic hydrolysis process. The results of the model are compared with experimental data. The comparison shows low errors of 7.96% for liquefaction and 7.35% for saccharification. These errors show a significant improvement in comparison with previous models and validate the proposed modeling methodology.
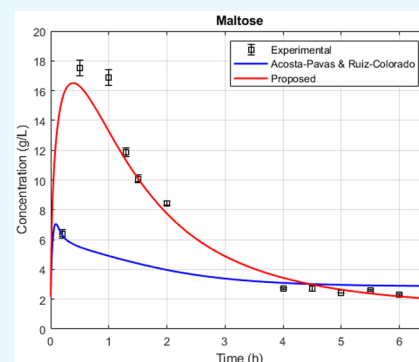
1. INTRODUCTION

Many years ago, various industries introduced mathematical models into their processes to analyze the behavior of the process against variations in initial product concentrations, temperature, or hydrogen potential (pH). Specifically, the introduction of phenomenological models in the biotechnology industry has allowed for better understanding of dynamics of variables interacting in a bioprocess, such as substrates, products, enzymes, or microorganisms, and from there, achievement of improvements in design, optimization, and control. Some examples of these models can be seen in ref 1, where the phenomenology of a submerged membrane bioreactor was developed to couple biomass kinetics with bubble aeration to understand membrane behavior in wastewater treatment. Likewise, in ref 2 an online model was proposed for the growth of Chilean mussel crops in some regions of Chile for predictive purposes. Also in ref 3 a polymerization reactor was modeled to implement a scale-up methodology and to facilitate its implementation at the industrial level.

Although the modeling of processes is becoming increasingly relevant, previous works show difficulties related to a considerable number of parameters involved in the model and the obtainment of numerical values that are able to adjust the dynamics of the nonlinear system. One of the solutions is the use of constitutive equations and bibliographic references. However, sometimes it is necessary to model a bioprocess that has not

been extensively studied, and experimentation must be used to quantify some parameters. This method in dealing with parameter identification can be observed in ref 4, where a bioprocess identification strategy was proposed based on the generalized bioprocess model to improve the parameter adjustment associated with the reaction kinetics used in the modeling stage, or in refs 5 and 6, which use genetic algorithms to perform a global search of the parameters describing bioreactor behavior for the production of ethanol and culture of *E. coli* MC4110 in a semibatch reactor, respectively. Finally, in ref 7 was proposed a methodology for modeling the enzymatic hydrolysis process, considering the estimation of some parameters selected arbitrarily, using the mean square error as a cost function. However, that research does not present an analysis of identifiability, an important aspect to consider in this type of model due to its sensitivity at the moment of changing some of its parameters.

The previous works allow an understanding of the methodologies used for parameter identification, but they do not give an

explanation on whether the implemented algorithms provide a valid solution for the system, that is, if it is possible that, given a set of outputs, $y(x, \psi, t)$, it will be possible to obtain a set of parameters p such that they can solve the dynamic system $\dot{x} = f(x, \theta, t)$.[8] The feasibility of this solution can be evaluated in dynamic systems based on the concept of identifiability. This concept can guarantee reliability in the process of parameter identification. Also, it can be used for iterating the modeling process in such a way that the algorithms used for obtaining these values are more reliable.

Therefore, the main purpose of this work is to present the modeling of an enzymatic hydrolysis process of amylaceous materials, considering the problem of parameter identification, by using the verification of the identifiability as a basis for the construction of the model. These bioprocess are good candidates for applying the proposed methodology because the enzymatic hydrolysis is characterized by its complexity arising from the considerable number of model parameters and nonlinearities present in the bioprocess. Furthermore, the phenomenology of the process will force the designer to focus on the identifiability problem to develop the associated phenomenological based semiphysical model (PBSM). This process is the basis of a major research project in bioprocess estimation and control started by Acosta-Pavas and Ruiz-Colorado.[9] Additionally, a better model of the enzymatic hydrolysis process of amylaceous materials can contribute to understand better how to use agro-industrial wastes for production of value-added products.

The article is organized as follows: First, the concept of identifiability and the method used to check it are explained. Next, the used methodology for process modeling aimed at parameter identification is explained step by step. The next section deals with the enzymatic hydrolysis model and its development, indicating hypotheses, definitions, equations, and identifiability analysis. Finally, the results of the strategy applied to enzymatic hydrolysis are discussed and the conclusions are presented.

2. IDENTIFIABILITY AND THE METHOD USED TO CHECK IT

The concept of identifiability was originally introduced by Bellman and Åström.[10] In this work, an error estimation function was defined as

$$J(p) = \int_0^T (\hat{y}(t) - y(t))^2 \, dt \tag{1}$$

where $y(t)$ is the output vector and $\hat{y}(t)$ is the estimated output vector. In addition, $y(t)$ is related to the nonlinear dynamic system:

$$\frac{dx(t)}{dt} = f(x, \theta, t) \tag{2a}$$

$$y(t) = h(x, \psi, t) \tag{2b}$$

$$x_0 = x(t_0, \theta) \tag{2c}$$

where x is the state vector, θ is the parameter vector associated with the state function $f(x, \theta, t)$, ψ is the parameter vector associated with the output function $h(x, \psi, t)$, and x_0 is the initial condition vector. If a parameter belongs to θ, the parameter is associated with the process phenomenology and is part of the proposed constitutive equations. On the other hand, if a parameter belongs to ψ, the parameter is associated with how

process outputs are measured (installed sensor, signal processing, and output interpretation). Furthermore, the parameter vectors can be summarized in a unique vector $p = [\theta, \psi]^T$.

The system described by eq 2 shows that there are internal couplings between x, p, and t defined by the structure of f. Additionally, available information on the output system y also is defined by the structure of h, i.e., by internal couplings between x, p, and t. The last idea implies that a nonlinear dynamic system will be understood if the structures of related functions (f and h) are well-known, and the structure of related functions will be well-known if the p vector can be known or identifiable.[10]

Considering eq 1, the system defined by eq 2 is structurally locally identifiable if there is a candidate parameter vector $\hat{p}$ such that $J(p)$ has a local minimum. If $\hat{p}$ is also a global minimum, the system is structurally globally identifiable.[10] The last concept is an *a priori* property of the dynamic system; that is, the property only depends on the proposed model structure from eq 2. This paper only focuses on identifiability as an *a priori* property. The decision taken due to this is focused on allowing a phenomenological review of the model in comparison with *a posteriori* concepts based only on the data quality.[8]

2.1. Structural Identifiability as an Observability Redefinition. For obtaining a correct idea of the structural identifiability of some system, many alternative definitions related to the original concept of Bellman and Åström[10] have been proposed. These proposals include the use of Taylor series expansion, differential algebra, or similarity transformation.[8] However, in recent years, the study of structural identifiability for the observability concept has been a novel proposal due to its easy understanding and application in nonlinear systems. That proposal was accurately developed by Villaverde[11] as follows:

With $\mathcal{M}$ the model defined by eq 2 and $L_f^0 h(x)$, $L_f^1 h(x)$, and $L_f^i h(x)$ set to zero, the first and ith Lie derivatives of the output function $h(x)$ (eq 2b) with respect to the state function $f(x)$ (eq 2a) are defined as

$$L_f^0 h(x) = h(x) \tag{3a}$$

$$L_f^1 h(x) = \frac{\partial h(x)}{\partial x} f(x) \tag{3b}$$

$$L_f^i h(x) = \frac{\partial L_f^{i-1} h(x)}{\partial x} f(x) \tag{3c}$$

Then $\mathcal{M}$ is locally observable around the neighborhood $\mathcal{N}$ defined for the initial point x_0 (eq 2c) if $rank(O(x_0) = n)$, where n is the number of states and $O(x)$ is the nonlinear observability matrix calculated as

$$O(x) = \begin{pmatrix} \dfrac{\partial}{\partial x} h(x) \\[2mm] \dfrac{\partial}{\partial x}(L_f h(x)) \\[2mm] \vdots \\[2mm] \dfrac{\partial}{\partial x}(L_f^{n-1} h(x)) \end{pmatrix} \tag{4}$$

The above condition is known as the nonlinear observability rank condition (ORC).[11] Now, let $\mathcal{M}$ be the model defined by eq 2, p the static parameter vector ($\dot{p} = 0$), and $\tilde{x}$ the augmented state vector defined by eq 5a; then $\mathcal{M}$ is structurally locally

identifiable around the neighborhood $\mathcal{N}$ defined for the initial point $\widetilde{x}_0$, if $rank(O_I(\widetilde{x}_0)) = n + n_x + n_y$, where n_x is the number of parameters associated with the state function θ, n_y is the number of parameters associated with the output function ψ, and $O_I(\widetilde{x})$ is the augmented nonlinear observability−identifiability matrix given for eq 5b.[11]

$$\widetilde{x} = [x, p]^T \tag{5a}$$

$$O_I(\widetilde{x}) = \begin{pmatrix} \dfrac{\partial}{\partial x} h(\widetilde{x}) \\ \dfrac{\partial}{\partial x}(L_f h(\widetilde{x})) \\ \vdots \\ \dfrac{\partial}{\partial x}(L_f^{n+n_x+n_y-1} h(\widetilde{x})) \end{pmatrix} \tag{5b}$$

The last condition is known as the nonlinear observability−identifiability condition (OIC). If this condition is fulfilled, then all elements into p are structurally locally identifiable and all states into x are locally observable around $\mathcal{N}$. Otherwise, at least one element into p is structurally unidentifiable or one state into x is unobservable.[11]

2.2. Toolbox for the OIC Computation. The computation of $rank(O_I(\widetilde{x}))$ requires the use of symbolic derivation toolboxes. STRIKE-GOLDD (structural identifiability taken as extended-generalized observability with lie derivatives and decomposition) is a MATLAB toolbox that analyses the local structural identifiability, observability, and invertibility of (possibly nonlinear) dynamic models of ordinary differential equations. This toolbox was developed by Villaverde et al.[12] Some advantages of STRIKE-GOLDD are the computation of a minimum number of Lie derivatives for reducing the computational cost for obtaining $O_I(\widetilde{x})$. If the OIC is not met, the toolbox can analyze each one of the model parameters and know which of these are structurally locally identifiable, structurally unidentifiable, and possible identifiable combinations of parameters. This feature allows the user to know problematic parameters in the model that could avoid obtaining a feasible solution in the parametric estimation.

This packet aims to reduce the computational time of the system and obtain new knowledge about each one of the elements. The benefits enable the model to iterate and adjust the system to avoid obtaining the identifiability property. STRIKE-GOLDD is used to determine if the model parameters can be correctly identifiable through an optimization algorithm. If the OIC is not fulfilled, then the optimizer cannot guarantee that $\hat{p}$ might be the solution vector for the identification problem and the model could describe the real process.

2.3. Basis for Parametric Identification. From the point of view of optimization algorithms, the problem of parameter identification in a dynamic system is a topic of constant development. Many approaches about the cost function for its use in parameter identification problems have been proposed in eq 1. Roeva[6] proposed a cost function defined by eq 6, where $y(t)$ is not a continuous measurement but the process can obtain a finite quantity of samples m_s for the output vector with dimension n_y.

$$J(p) = \sum_{i=1}^{m_s} \sum_{j=1}^{n_y} (\hat{y}_j(i) - y_j(i))^2 \tag{6}$$

From eq 6, $J(p)$ can be modified to increase the weight of some samples into the algorithm. For example, Rivera et al.[5] uses the maximum possible value of each output as seen in eq 7. The purpose of that modification is to penalize estimation errors that could be above a defined boundary and affect the optimization trajectory. Meanwhile, Richelle and Bogaerts[4] uses the standard deviation σ of each measurement as seen in eq 8. The objective of that change is to penalize the noisiest measurements.

$$J(p) = \sum_{i=1}^{m_s} \sum_{j=1}^{n_y} \frac{(\hat{y}_j(i) - y_j(i))^2}{(y_{j_{max}})^2} \tag{7}$$

$$J(p) = \sum_{i=1}^{m_s} \sum_{j=1}^{n_y} \frac{(\hat{y}_j(i) - y_j(i))^2}{(\sigma(y_j(i))^2} \tag{8}$$

In addition, $J(p)$ can consider the number of experiments done m_e, as observed in eq 9. This cost function can be solved using a maximum likelihood estimator.[13]

$$J(p) = \sum_{i=1}^{m_s} \sum_{j=1}^{n_y} \sum_{k=1}^{m_e} \frac{(\hat{y}_{j_k}(i) - y_{j_k}(i))^2}{(\sigma(y_{j_k}(i))^2} \tag{9}$$

For executing the parameter identification as an optimization problem, the *fmincon* function available in the MATLAB optimization toolbox was selected for generating a local minimum based on the formulated problem in the last section. In detail, the function requires as input arguments the cost function and matrices including the model constraints (boundaries, equalities, and inequalities). In addition, the toolbox can change the optimization algorithm to execute, but *fmincon* will use the interior point algorithm for both enzymatic hydrolysis stages. The interior point algorithm generates a new cost function by introducing a barrier function as seen in eq 10. This function is compounded for an additional logarithmic term as a barrier value s and a constraint term μ. As a consequence, the problem is changed from an inequality constraint problem to a limited equality problem[14] as follows:

$$\min_{x,s} \gamma_\mu(x, s) = \min_{x,s}\left(\gamma(x) - \mu \sum_i \ln(s_i)\right) \tag{10a}$$

$$subject\ to\ z(x) = 0\ and\ g(x) + s = 0 \tag{10b}$$

In regard to eq 10, the algorithm selects one of the following steps in each iteration:

- **Direct step**: The algorithm tries to solve the Karush−Kuhn−Tucker (KKT) equations for the optimization problem using a linear approximation.

- **Conjugate gradient**: After the algorithm tried to do a direct step but the obtained result was not feasible, the algorithm calculated the Lagrange multipliers for solving approximated KKT equations. In other words, the algorithm tries to minimize the problem in a reliability region, causing a feasible trajectory in each iteration.

3. USED METHODOLOGY FOR PROCESS MODELING

The used methodology is based on the PBSM methodology, developed by Alvarez et al.[15] and applied in ref 9, which can generate gray-box and lumped parameter models. In this work, steps 6−11 were added in order to consider the identifiability of parameters. The methodology steps are as follows, and they are summarized in Figure 1.

Step 1: Describe the process to identify the elements and physical interactions that should be in the model. This action is used as input for the next steps' balances, variables, and constitutive equations.

Step 2: Propose the model suppositions associated with the process. These statements must be considered the main objective of the model, involving variables, physical limits, and available tools for parameter identification and model testing.

Step 3: Define the balances associated with the dynamic process. This step implies an initial PBSM, i.e., a state function f and an output function y with variables, inputs, and outputs.

Step 4: Obtain constitutive equations and the parameter vector p of the system through the suppositions, balances, and variables indicated in steps 2 and 3.

Step 5: Calculate the augmented nonlinear observability−identifiability matrix $O_I(\tilde{x})$ (eq 5). This step is developed by means of the STRIKE-GOLDD toolbox. If the matrix has full rank, then the system is structurally locally identifiable. Otherwise, the modeling process is reset from step 2 considering how to change the model to obtain the identifiability property. This return to step 2 can be possible because the toolbox indicates unidentifiable parameters that must be modified or removed from the PBSM.

Step 6: Choose a cost function $J(p)$ for the development of the optimization problem. If the previous steps were satisfactory, the PBSM will guarantee the existence of a local solution. So, the objective is to obtain a set of optimal values for the system parameters $\hat{p}$ to describe the process through the proposed model.

Step 7: Define the constraints based on the model description and related suppositions. This step is able to define a subspace $\mathcal{P} \subset \mathbb{R}^{n_x + n_y}$ such that $\hat{p}$ is feasible. In the first iteration of the methodology, the constraints should be defined by means of the knowledge of the process and the model suppositions. The reason for the last statement is reflected in the fact that these inputs can give information about the logical scales, units, and values of the system parameters.

Step 8: Get a data set associated with the model for using $J(p)$ and finding $\hat{p}$. For this case of study, the selected data set is based on the measurement of various samples of the process.

Step 9: Filter the obtained data set for removing non-necessary information for the optimization algorithm.

Step 10: Define an initial point p_0 for initializing the optimization algorithm.

Step 11: Execute an optimization algorithm to solve the parameter identification problem. If the obtained vector $\hat{p}$ to the optimization process can fit the PBSM and there is not an alternative to improve the proposed model, then $\hat{p} \approx p$; that is, the identification problem associated with the proposed model is achieved and the methodology generates an acceptable result. If $\hat{p}$ cannot fit the PBSM, the methodology is re-executed from step 7 for improving the algorithm inputs (initial point, constraints, cost function, and data). If there is some feasible improvement to the PBSM, the methodology is reset to step 2.
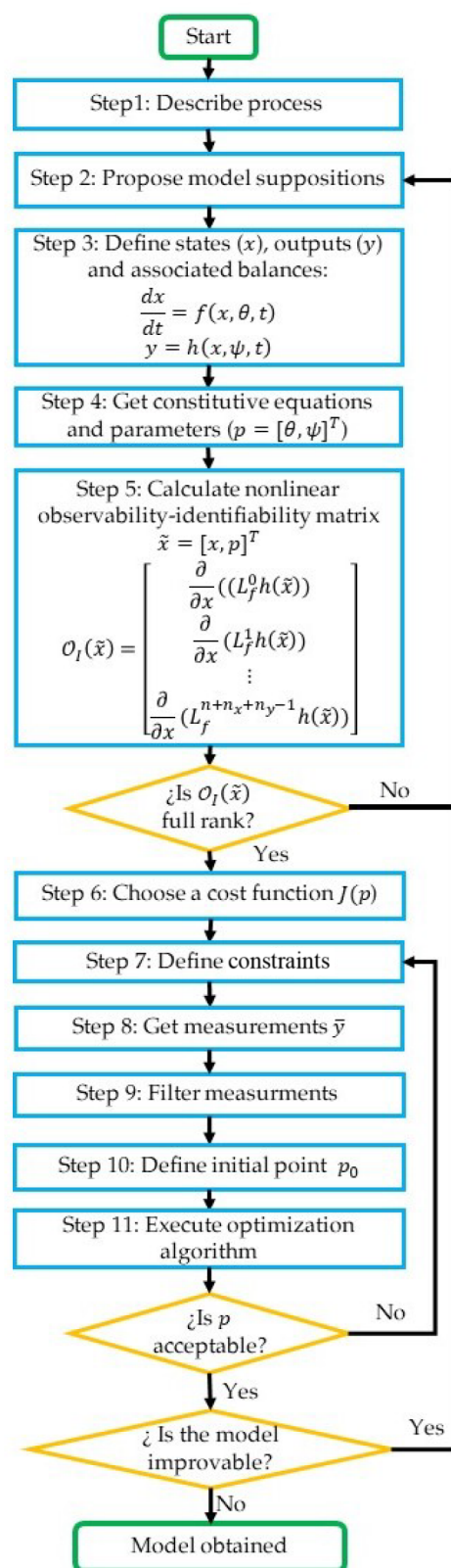

Figure 1. Flow chart for the proposed methodology

The proposed modeling methodology allows design of the PBSM with structurally locally identifiable parameters. Step 5 is very important to aim that property because it functions as a validation test for the parameters used in step 4 and suggested in step 2 and step 3. When the OIC condition ($rank(O_I(\tilde{x}))$)

cannot be guaranteed, the designer must review the results given by STRIKE-GOLDD and identify what parameters are problematic, i.e., unidentifiable. Due to each parameter having at least one associated constitutive equation, the designer must return to step 2 to review what supposition could cause or influence the use of that. When the designer identifies a supposition−parameter pair, it is possible to analyze how to modify this association without contradicting the process description:

- Redefine the supposition for obtaining new constitutive equations and new possible identifiable parameters.
- Change the supposition for modifying balances or constitutive equations and their associated parameters.
- Eliminate the supposition for redefining balances and their associated equations and variables.
- Add a new supposition for obtaining new balances and their associated equations and variables.

Making changes will improve the structure of the dynamic system if modifications are done based on a good knowledge of the real process. Each change in some supposition can modify significantly equations or variables for use in the next steps in the methodology loop.

When the methodology guarantees that the obtained model is structurally locally identifiable, this can add steps for estimating the values of parameters included in the PBSM via the optimization strategy (steps 6−11). The last idea is possible because, as was mentioned in section 2, the obtained identifiable model will have a candidate parameter vector $\hat{p}$ such that an error estimation function $J(p)$ proposed by the designer has a local minimum. Additionally, step 7 allows defining constraints based on the model suppositions such that the p vector helps to limit the feasible neighborhood and simulate correctly the process. For example, a model supposition can suggest rules for setting a logical value or range for the initial condition of the model. For that, if the p vector is obtained in the modeling methodology but the designer can identify an improvement point based on the final results, the methodology will suggest reviewing the model suppositions if a new knowledge of the process or its associated phenomena can set a new parameter neighborhood. However, improvement points also can be proposed based on adding new dynamic behavior. This option suggests proposing news suppositions, and consequently, new variables must be added without losing the identifiability property of the model.

4. RESULTS AND DISCUSSIONS

4.1. Case of Study: Enzymatic Hydrolysis from Wheat Starch.
An interesting case study to apply the methodology is the enzymatic hydrolysis process of amylaceous materials. These materials come from agro-industrial wastes, which, when subjected to hydrolysis processes, allow the production of materials for the production of jams, medicines, fuels, and others.[9,16−21] These agro-industrial wastes are associated with the production of cereal derivatives such as rice, corn, wheat, barley, and oats, legumes such as chickpeas, lentils, soy, and beans, and tubers such as potato, cassava, and yam.[22−24]

4.2. Step 1: Model Description.
This type of agro-industrial residue is composed of two polymers: amylose, with a linear structure of D-glucose linked by α-(1, 4) bonds, and amylopectin, with a branched structure linked by α-(1, 4) and α-(1, 6) bonds.[22,25−27] Sourcing products derived from these polymers can be achieved through the process of enzymatic

Table 1. Initial Values for p and x

Liquefaction model			Saccharification model		
Variable	Value	Units	Variable	Value	Units
G_1	0	$\left[\frac{g}{L}\right]$	G_1	0.0041	$\left[\frac{mol}{L}\right]$
G_2	0	$\left[\frac{g}{L}\right]$	G_2	0.0063	$\left[\frac{mol}{L}\right]$
G_3	0	$\left[\frac{g}{L}\right]$	G_3	0.0259	$\left[\frac{mol}{L}\right]$
G_5	213.49	$\left[\frac{g}{L}\right]$	G_5	0.2419	$\left[\frac{mol}{L}\right]$
T	331.95	$[K]$	T	333.75	$[K]$
$k_{1_{max}}$	1.55	$\left[\frac{1}{h}\right]$	K_I	80000	$\left[\frac{hmL}{mol}\right]$
$k_{2_{max}}$	404	$\left[\frac{1}{h}\right]$	K_S	390	$\left[\frac{hmL}{mol}\right]$
$k_{3_{max}}$	185	$\left[\frac{1}{h}\right]$	K_2	460000	$\left[\frac{hmL}{mol}\right]$
$k_{5_{max}}$	14.25	$\left[\frac{1}{h}\right]$	K_3	2.15	$\left[\frac{hmL}{mol}\right]$
h_1	−0.3407	$[-]$	K_5	3500	$\left[\frac{hmL}{mol}\right]$
h_2	−0.3362	$[-]$	k_{eq_2}	150	$\left[\frac{mol}{L}\right]$
h_3	−0.3228	$[-]$	k_{eq_3}	0.008	$\left[\frac{mol}{L}\right]$
h_5	−0.4567	$[-]$	$v_{2_{max}}$	9	$\left[\frac{mol}{hmL}\right]$
β_1	0.0612	$[-]$	$v_{3_{max}}$	3000	$\left[\frac{mol}{hmL}\right]$
β_2	0.9388	$[-]$	$v_{5_{max}}$	3500	$\left[\frac{mol}{hmL}\right]$
-	-	-	β_1	0.0612	$[-]$
-	-	-	β_2	0.9387	$[-]$

hydrolysis, which consists of swelling an initial substrate (starch) and then breaking the bonds α-(1, 4) and α-(1, 6) from the use of enzymes through a series of synergistic stages: gelatinization, liquefaction, and saccharification.[24,28−31] Each stage aims to modify the type of structure of D-glucose and the quantity in the final product. If there is a polymer composed for s D-glucoses linked, then the product has a polymer with a degree of polymerization (DP) equal to s.

In the gelatinization stage, the starch granules are subjected to high temperatures, which allows swelling by absorption of water and subsequent release of amylose and amylopectin.[28,30,32−36] In the liquefaction stage, the amylose and amylopectin are degraded by the enzyme α-amylase, and the internal bonds α-(1, 4) and α-(1, 6) are hydrolyzed, producing the successive division of both polymers. An intermediate product is obtained, composed of smaller oligosaccharides (DP < 7).[20,29,30,34,36−39] Finally, in the saccharification stage, the products of the last stage are degraded by the enzyme amyloglucosidase until products such as glucose or maltose are obtained.[20,29,30,34,36−39]

4.3. Step 2: Model Suppositions.
4.3.1. For the Liquefaction Stage.

Table 2. Obtained Values for $\hat{p}$

Liquefaction model			Saccharification model		
Variable	Estimated value	Units	Variable	Estimated value	Units
$k_{1_{max}}$	1.548671	$\left[\frac{1}{h}\right]$	K_I	75032.6049	$\left[\frac{hmL}{mol}\right]$
$k_{2_{max}}$	396.445826	$\left[\frac{1}{h}\right]$	K_S	554.9090	$\left[\frac{hmL}{mol}\right]$
$k_{3_{max}}$	182.857411	$\left[\frac{1}{h}\right]$	K_2	460614.7352	$\left[\frac{hmL}{mol}\right]$
$k_{5_{max}}$	7.152166	$\left[\frac{1}{h}\right]$	K_3	0.6699	$\left[\frac{hmL}{mol}\right]$
h_1	-0.343724	$[-]$	K_5	3915.2964	$\left[\frac{hmL}{mol}\right]$
h_2	-0.335595	$[-]$	k_{eq_2}	137.8292	$\left[\frac{hmL}{mol}\right]$
h_3	-0.322011	$[-]$	k_{eq_3}	0.0059	$\left[\frac{hmL}{mol}\right]$
h_5	-0.455272	$[-]$	$v_{2_{max}}$	10.9983	$\left[\frac{mol}{hmL}\right]$
β_1	0.061203	$[-]$	$v_{3_{max}}$	2782.6009	$\left[\frac{mol}{hmL}\right]$
β_2	0.938797	$[-]$	$v_{5_{max}}$	5651.7907	$\left[\frac{mol}{hmL}\right]$
-	-	-	β_1	1.2043	$[-]$
-	-	-	β_2	1.3040	$[-]$

- The bioprocess consists of the formation of oligosaccharides with DP $\leq$ 7. However, the production of maltopentose G_5 includes concentrations of oligosaccharides among 3 < DP $\leq$ 7. The remaining products, glucose G_1, maltose G_2, and maltotriose G_3, are not lumped together.
- On the basis of process data via high-performance liquid chromatography (HPLC), the liquefaction stage shows available measurements of G_1, G_2, and maltodextrins. The last values are equal to the weighted sum of the remaining oligosaccharides in the process.
- As the oligosaccharide DP becomes smaller, oligosaccharides with greater DP compete for the active site of the enzyme. As a consequence, there are reversible reactions among oligosaccharides with DP < 5.
- E_{01} hydrolyzes the internal bonds into the initial substrate S_0, causing oligosaccharide production on a basis to the individual concentrations and the hydrolysis efficiency.
- The initial conditions of the stage are null for all oligosaccharides with DP $\leq$ 3.
- The growth of each oligosaccharide is equal to a first order kinetic model.
- It is an endothermic bioprocess.
- There is heat transfer in the process (temperature T) with the jacketed stirred tank (temperature T_j).
- The reaction velocity is associated with the process temperature by an Arrhenius function.

4.3.2. For the Saccharification Stage.

- The bioprocess consists of the formation of oligosaccharides with DP $\leq$ 7. However, the production of maltopentose G_5 includes concentrations of oligosaccharides among 3 < DP $\leq$ 7. The remaining products, glucose G_1, maltose G_2, and maltotriose G_3, are not lumped together.
- On the basis of process data via high-performance liquid chromatography (HPLC), the liquefaction stage shows available measurements of G_1, G_2, and maltodextrins. The last values are equal to the weighted sum of the remaining oligosaccharides into the process.
- All oligosaccharides compete for the active site of the amyloglucosidase enzyme.
- The inhibition for substrate k_s is only considered for oligosaccharides with DP > 3.
- The initial conditions of the bioprocess are equal to the last values in the liquefaction stage before the introduction of amyloglucosidase enzyme.
- The behavior of each oligosaccharide in this stage is equal to Michaelis−Menten kinetics.
- With respect to maltotriose and maltose, the product of the oligosaccharide with lower DP can decelerate growth.
- In this stage, the reactions among oligosaccharides have a behavior over the following stoichiometry relationships:

$$G_5 + H_2O \xrightarrow{r_{s_5}} G_2 + 3G_1 \tag{11a}$$

$$G_3 + H_2O \underset{r_{-s_3}}{\overset{r_{s_3}}{\rightleftharpoons}} G_2 + G_1 \tag{11b}$$

$$G_2 + H_2O \underset{r_{-s_2}}{\overset{r_{s_2}}{\rightleftharpoons}} 2G_1 \tag{11c}$$

- It is an endothermic bioprocess.
- There is heat transfer between the process (temperature T) with the jacketed stirred tank (temperature T_j).
- The reaction velocity is associated with the process temperature by an Arrhenius function.

4.4. Step 3: Definition of Variables and Dynamics.
4.4.1. For the Liquefaction Stage. Table S1 in the Supporting Information describes the liquefaction process variables, where

$$x = [G_1, G_2, G_3, G_5, T]^T \tag{12}$$

The dynamic system for the state vector (eq 12) is defined as

$$\dot{x}(t) = f(x, \theta, t) = \begin{pmatrix} r_{l_1} \\ r_{l_2} \\ r_{l_3} \\ r_{l_5} \\ Q_c - Q_{g_l} \end{pmatrix} \tag{13}$$

And y is defined as

$$y(t) = h(x, \psi, t) = \begin{pmatrix} G_1 \\ G_2 \\ \beta_1 G_3 + \beta_2 G_5 \\ T \end{pmatrix} \tag{14}$$

The state vector of the system is equal to eq 13, where the dynamic corresponds to the rate concentrations of glucose, maltose, and maltodextrins and the process temperature. The mass balances associated with each oligosaccharide only depend on their concentration rates.[9] Meanwhile, the energy balances
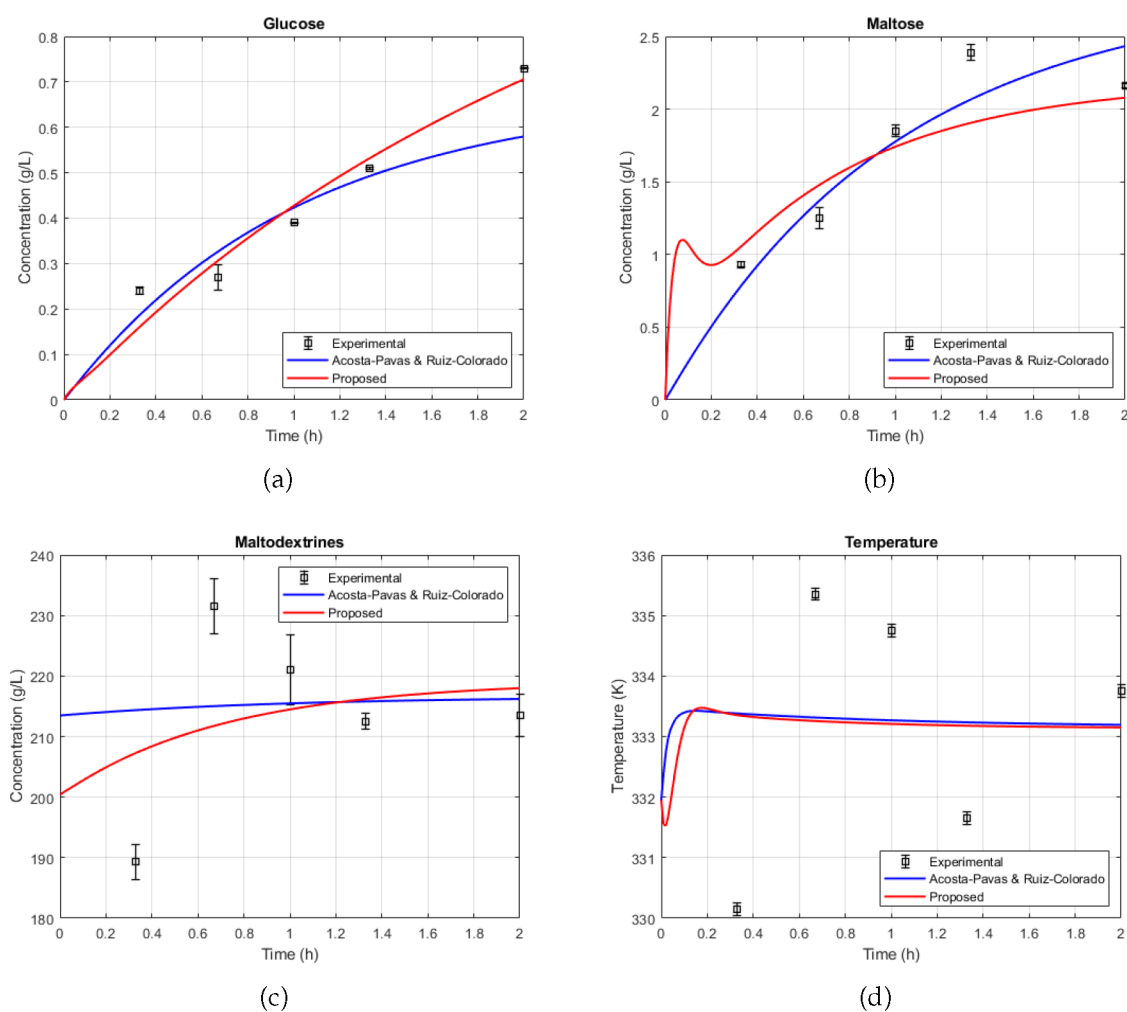
Figure 2. Comparison between the models for the liquefaction stage: (a) glucose; (b) maltose; (c) maltodextrines; (d) temperature.

show a heat exchange using the thermal jacket Q_c and the heat taken using concentrations rates Q_{g_l}.

The output vector of the system is equal to eq 14, where the outputs correspond to the concentrations of glucose, maltose, and maltodextrins and the process temperature. In comparison with Acosta-Pavas and Ruiz-Colorado[9] the maltodextrins concentration does not depend on maltotriose and maltopentose and also has a dependency of the values of β_1 and β_2. These parameters indicate a weighted sum relationship between those concentrations during the liquefaction process. For that reason, there are two additional parameters in the model with an equivalent constraint for parameter identification.

4.4.2. For the Saccharification Stage. Table S3 in the Supporting Information describes the saccharification process variables, where

$$x = [G_1, G_2, G_3, G_5, T]^T \qquad (15)$$

The dynamic system for the state vector (eq 15) is defined as

$$\dot{x}(t) = f(x, \theta, t) = \begin{pmatrix} 3r_{s_5} + r_{s_3} + 2r_{s_2} \\ r_{s_5} + r_{s_3} - r_{s_2} \\ -r_{s_3} \\ -r_{s_5} \\ Q_c - Q_{g_s} \end{pmatrix} \qquad (16)$$

And y is defined as equal to eq 14.

The state vector of the system is equal to eq 16, where the dynamic corresponds to the rate concentrations of glucose, maltose, and maltodextrins and the process temperature. The mass balances associated with each oligosaccharide depend on the stoichiometry relationships of the G_5, G_3, and G_2 concentration rates in eq 11. Meanwhile, the energy balances show a heat exchange using thermal jacket Q_c and the heat taken using the concentration rates of the saccharification stage Q_{g_s}.

The output vector of the system is equal to eq 14. In contrast with the liquefaction stage, the numeric values of β_1 and β_2 indicate a linear combination relationship, due to the maltodextrins being compounded by oligosaccharides with $3 \leq DP \leq 5$, causing the system to take different values, although the output equation is the same in both stages. In addition, these parameters have an inequality constraint associated with the change in the relationship. That inequality avoids the
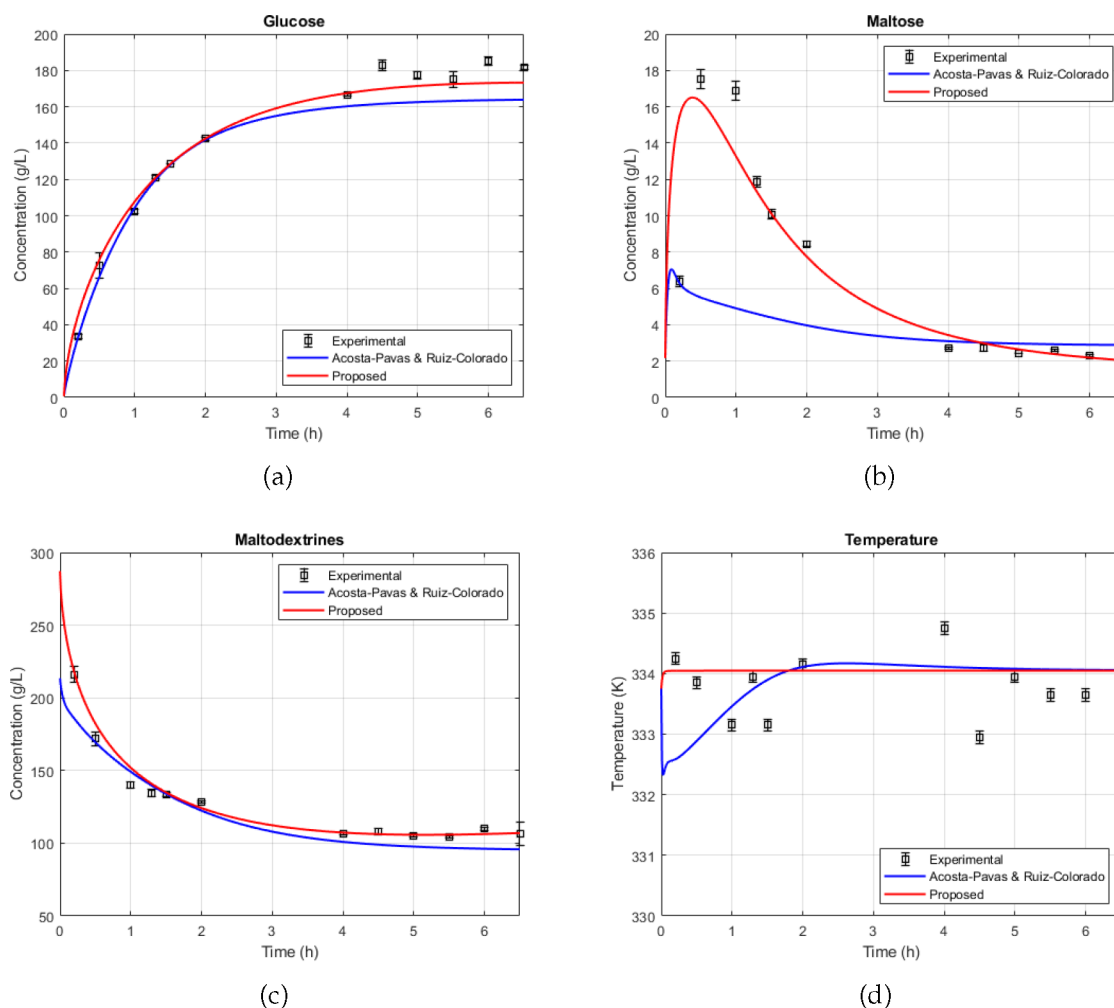
Figure 3. Comparison between the models for the saccharification stage: (a) glucose; (b) maltose; (c) maltodextrines; (d) temperature.

maltodextrin concentration in an extreme case, the sum of oligosaccharides $3 \leq DP \leq 5$ with the same weight or the representation of a unique oligosaccharide in the range.

4.5. Step 4: Constitutive Equations and Parameters.
4.5.1. For the Liquefaction Stage. 4.5.1.1. Reaction Velocity. The reaction velocity k_i associated with the oligosaccharides G_i is defined as

$$k_i = q_{l_i} k_{i_{max}} \exp\left(\frac{E_{a_i}}{RT}\right) \tag{17}$$

where $i \in DP_x = \{1, 2, 3, 5\}$ indicates the DP of the associated oligosaccharide, $k_{i_{max}}$ is the maximum reaction velocity of G_i, E_{a_i} is the activation energy for G_i, R is the Arrhenius constant, and q_{l_i} is a scale factor to compensate the final value of k_i.

4.5.1.2. First Order Reaction. The reaction r_{l_i} for each oligosaccharide with DP = i is defined as

$$r_{l_i} = k_i \left[(1 - h_i)S_0 - \sum_{k \geq i} G_k \right] E_{01} \tag{18}$$

where k_i is defined in relation to eq 17 and h_i is the hydrolysis relationship between S_0 and G_i.

4.5.1.3. Convective Heat Transfer. The heat transferred between the thermal jacket and the process is described as

$$Q_c = \frac{UA}{\rho V c_p}(T_j - T) \tag{19}$$

where U is the heat transfer coefficient, A is the heat transfer area, ρ is the product density, V is the process volume, c_p is the heat capacity, and T_j is the jacket temperature.

4.5.1.4. Heat Produced by the Process. The heat generated by the liquefaction stage is described as

$$Q_{g_l} = \frac{\Delta H}{\rho c_p} \sum_{k \in GP_x} \frac{r_{l_k}}{w_{m_k}} \tag{20}$$

where ΔH is the process enthalpy, r_{l_k} is defined per eq 18, and w_{m_k} is the molecular weight of each oligosaccharide in GP_x.

4.6. Parameters. Based on constitutive equations, Table S1 in the Supporting Information describes the liquefaction parameters, where

$$\theta = [k_{1_{max}}, k_{2_{max}}, k_{3_{max}}, k_{5_{max}}, h_1, h_2, h_3, h_5]^T \tag{21a}$$

$$\psi = [\beta_1, \beta_2]^T \tag{21b}$$

In addition, Table S2 in the Supporting Information provides the model constants used in the liquefaction stage.

4.6.1. For the Saccharification Stage. 4.6.1.1. Reaction Velocity. The reaction velocity v_i associated with the reaction rate r_{s_i} is defined as

$$v_i = q_{s_i} v_{i_{max}} \exp\left(\frac{E_a}{RT}\right) \tag{22}$$

where $i \in r_x = \{2, 3, 5\}$ indicates the reaction rate associated with the oligosaccharide with DP = i, $v_{i_{max}}$ is the maximum reaction velocity of G_i, E_a is the activation energy for the final product (in this case, G_1), R is the Arrhenius constant, and q_{s_i} is a scale factor used to compensate for the magnitude of the exponential expression. In a similar way to q_{l_i} in eq 17, q_{s_i} is a scale factor including an offset to the magnitude of the exponential expression. The use of q_{s_i} has a similar objective to q_{l_i} in the liquefaction stage.

4.6.1.2. Reaction Rate of Maltopentose. The reaction rate r_{s_5} associated with G_5 is defined as

$$r_{s_5} = \frac{v_5 K_5 G_5 E_{02}}{1 + K_I G_1 + K_S + K_2 G_2 + K_3 G_3 + K_5 G_5} \tag{23}$$

where K_5 is the inverse of the inhibition constant for $G_5 \left(K_5 = \frac{1}{k_{m_5}}\right)$, K_3 is the inverse of the inhibition constant for $G_3 \left(K_3 = \frac{1}{k_{m_3}}\right)$, K_2 is the inverse of the inhibition constant for $G_2 \left(K_2 = \frac{1}{k_{m_2}}\right)$, K_I is the inverse of the inhibition constant for $G_1 \left(K_I = \frac{1}{k_I}\right)$, and K_S is the inverse of the inhibition constant for the substrate $\left(K_S = \frac{1}{k_S}\right)$.

4.6.1.3. Reaction Rate of Maltotriose. The reaction rate r_{s_3} associated with G_3 is defined as

$$r_{s_3} = \frac{v_3 K_3 \left(G_3 - \frac{G_1 G_2}{k_{eq_3}}\right) E_{02}}{1 + K_I G_1 + K_2 G_2 + K_3 G_3 + K_5 G_5} \tag{24}$$

where k_{eq_3} is the equilibrium constant associated with G_3.

4.6.1.4. Reaction Rate of Maltose. The reaction rate r_{s_2} associated with G_2 is defined as

$$r_{s_2} = \frac{v_2 K_2 \left(G_2 - \frac{G_1^2}{k_{eq_2}}\right) E_{02}}{1 + K_I G_1 + K_2 G_2 + K_3 G_3 + K_5 G_5} \tag{25}$$

where k_{eq_2} is the equilibrium constant associated with G_2.

In comparison with the traditional structure of Michaelis−Menten kinetics, in eqs 23, 24, and 25 the inverses of the inhibitions present in the stage are used. This change generated in each r_{s_i} could be simpler to operate mathematically due to not adding the derivative of the quotient. Therefore, the model is easier for OIC computing in STRIKE-GOLDD concerning other models and permits validation of parameter identification.

4.6.1.5. Convective Heat Transfer. The heat transferred between the thermal jacket and the process is described in eq 19.

4.6.1.6. Heat Produced by the Process. The heat generated by the saccharification stage is described as

$$Q_{g_s} = \frac{\Delta H}{\rho c_p} \sum_{i \in r_x} r_{s_i} \tag{26}$$

where r_{s_i} is defined in eqs 23, 24, and 25.

4.6.1.7. Parameters. Based on constitutive equations, Table S3 in the Supporting Information describes the saccharification parameters, where

$$\theta = [K_I, K_S, K_2, K_3, K_5, k_{eq_2}, k_{eq_3}, v_{2_{max}}, v_{3_{max}}, v_{5_{max}}]^T \tag{27a}$$

$$\psi = [\beta_1, \beta_2]^T \tag{27b}$$

In addition, Table S4 in the Supporting Information provides the model constants used in the saccharification stage.

4.7. Step 5: Identifiability Analysis. *4.7.1. For the Liquefaction Stage.* The STRIKE-GOLDD packet in MATLAB was used for computing the OIC. The initial model conditions were known. The software obtained $O_I(\tilde{x})$ rank equal to 15 in 2.8 s of executing time, and it calculated up to the third Lie derivative. This result indicates that $O_I(\tilde{x})$ is full rank, i.e., the system is structurally identifiable for the whole parameters. In previous proposals were not possible to guarantee this condition.[17,24,40] However, it was possible with the current model due to the changes in constitutive equations.

4.7.2. For the Saccharification Stage. The STRIKE-GOLDD packet in MATLAB was used for computing the OIC. The initial model conditions were known. The software gave $O_I(\tilde{x})$ rank equal to 17 in 398.4 s executing time, and it calculated up to the fourth Lie derivative. This result indicates that $O_I(\tilde{x})$ is full rank; that is, the system is structurally identifiable for all of the parameters. In previous proposals, it was not possible to guarantee this condition.[17,24,40] However, it was possible with the current model due to the changes in the constitutive equations.

4.8. Step 6: Cost Function in the Liquefaction and Saccharification Stages. As intended in the proposed methodology, the enzymatic hydrolysis model in both stages has numerous parameters, 10 in the liquefaction stage and 12 in the saccharification stage.

Let $g_g(\psi)$ be an equality constraint function between ψ parameters and a constant c_g, and let $g_h(\psi)$ be an inequality constraint function between ψ and another constant c_h, where the function must be less than or equal to the constant. Let p_l and p_u be the lower and upper boundaries of the vector p, respectively. In developing parameter identification in both the liquefaction stage and saccharification stage, $J(p)$ is proposed as

$$\min_p J(p) = \sum_{i=1}^{m_s} \sum_{j=1}^{m_y} \frac{(\hat{y}_j(i) - \bar{y}_j(i))^2}{(\sigma(\bar{y}_j(i))^2} \tag{28a}$$

subject to

$$\dot{x} = f(x, \theta, t) \tag{28b}$$

$$p_l \leq p \leq p_u \tag{28c}$$

$$g_g(\psi) = c_g \tag{28d}$$

$$g_h(\psi) \leq c_h \tag{28e}$$

where $\bar{y}_j$ and $\sigma(\bar{y}_j)$ are the mean value of the output j and the standard deviation associated with the sample i, respectively, and are defined as

$$\overline{y}_j(i) = \frac{1}{m_e} \sum_{k=1}^{m_e} y_{m_{j_k}}(i) \tag{29a}$$

$$\sigma(\overline{y}_j(i)) = \sqrt{\frac{1}{m_e} \sum_{k=1}^{m_e} (\overline{y}_j(i) - y_{m_{j_k}}(i))^2} \tag{29b}$$

$J(p)$ is subject to the mean values describe in eq 29. This structure allows use of the experimental data shown in ref 9 and obtains a point of comparison between the proposed model and the model included in the previously mentioned reference. Therefore, $J(p)$ is a measurement of the integral quadratic error of the parameter identification adjusted to the standard deviation and mean of each sample for including all information on the measured data in comparison with previous works.[7,9] While $J(p)$ takes a lower value, $\hat{p}$ will be a local minimum nearer to the real value of p.

4.9. Step 7: Constraints' Definition in the Liquefaction and Saccharification Stages. In the liquefaction stage, the algorithm does not consider eq 28e and uses eq 28d as the basis of eq 30 for including the weighted sum constraint in the maltodextrins concentration. Meanwhile, the saccharification stage only considers eqs 28e−31 with the objective that ψ can adjust the maltodextrin concentration between the limits.

$$g_g(\psi) = \beta_1 + \beta_2 = 1 \tag{30}$$

$$g_h(\psi) = \beta_1 + \beta_2 \leq 3 \tag{31}$$

Table S5 in the Supporting Information shows the numeric values of the constraints in the lower boundary p_l and the upper boundary p_u. In the liquefaction stage, the boundaries are based on previous results shown in the literature,[9,17] except the h_i parameters that must have a range between −1 and 0 in the proposed model. The last decision is due to the fact that each h_i must indicate the growing possibility of the respective oligosaccharide concerning S_0. In the saccharification stage, the first numerical values of the boundaries are based on previous results shown in the literature.[9] Moreover, those values were iterated following expected values selected using the phenomenological knowledge of the system. The objective of that is to reduce the neighborhood $\mathcal{N}$ and improve the solution obtained in $\hat{p}$.

4.10. Steps 8 and 9: Obtain and Filter Measurements of the Liquefaction and Saccharification Stages. For each stage, Acosta-Pavas and Ruiz-Colorado[9] ran three experiments (m_e). For each experiment, $\overline{y}_j$ and $\sigma(\overline{y}_j)$ were calculated. The details of the experiment can be found in refs 9 and 41. Based on this data set, some considerations were done for each model. In the liquefaction stage, the data set has six available measurements (m_s) taken in a process time of 2 h for each of the outputs defined in eq 14. In this data set, the first measurements (initial condition) were ignored because the data causes problems with the cost function. In detail, these data points might set singularities into the cost function due to standard deviations being equal to zero. In the saccharification stage, the data set has 15 available measurements (m_s) taken in a process time of 6.5 h for each of the outputs defined in eq 14. In this data set, the eighth and ninth measurements (initial condition) were ignored, justified on the same basis as in the liquefaction stage.

4.11. Step 10: Initial Point in the Liquefaction and Saccharification Stages. For the initial conditions x_0 in each stage, the algorithm took the published values in ref 9. With respect to the initial values for parameter vector p_0 for executing

fmincon, these were selected such that the algorithm can obtain a feasible $\hat{p}$ over the parameter boundaries p_l and p_u. These values are shown in Table 1.

4.12. Step 11. Optimization Results in the Liquefaction and Saccharification Stages. Table 2 shows the $\hat{p}$ values obtained through the optimization algorithm. In the liquefaction stage, the proposed methodology obtained a lower $J(p)$ value (3.8665×10^3) than the evaluation of the cost function over the developed model in ref 9 (2.0461×10^4). In terms of the relative error between the model and data, the proposed model obtains 7.64% concerning 7.96% of the previous model. This result can be interpreted as an improvement of this model. For the other side, in the saccharification stage, the $J(p)$ value obtained by the proposed methodology is also lower (2.2062×10^3) than the evaluation of the cost function over the developed model in ref 9 (7.4025×10^3), but the difference is less than the result obtained in the liquefaction stage. Moreover, the improvement is more significant, indicating a better adjustment using the proposed model. Model improvement was also observed, with the relative error between each model and the respective experimental data being calculated. The proposed model obtains 7.35% concerning 11.99% of the previous model.

Figure 2 shows the simulation of the proposed model with the identified parameters. In Figure 2a it is possible to see the match between the data and the model. In Figure 2b there is a difference between the data and the model in the initial phase of growth because of the supposition about oligosaccharides with more DP competing for the active site of the enzyme. As a consequence, there are reversible reactions among oligosaccharides with DP < 5 in the model.

In Figure 2c it was impossible to obtain a better result because this output variable is a combination of two state variables; however, this result is better than the previous result obtained in ref 9. Finally, Figure 2d shows no significant change in temperature because the process has a feedback control loop of control not incorporated in this study.

Despite this, the developed methodology aims at adjusting the bioprocess dynamics correctly for all outputs. For G_1, G_2, $\beta_1 G_3 + \beta_2 G_5$, and T, the relative errors in the liquefaction stage were 12.72%, 12.33%, 4.96%, and 0.54%, respectively. In other words, the proposed model obtained relative errors in the liquefaction stage <13%. It is important that the best adjustment in the concentration dynamic in the liquefaction stage was the maltodextrin concentration. The last idea indicates that the changes introduced into the model for that output were positive for bioprocess modeling.

Figure 3 shows the simulation of the proposed model with the identified parameters. In Figure 3a−c it is possible to see the match between the data and the model. This result was possible because the amount of data was significantly greater than in the liquefaction stage. Finally, Figure 2d shows no significant change in temperature because this variable was controlled. Therefore, the developed methodology aimed to adjust the bioprocess dynamics correctly for all outputs. For G_1, G_2, $\beta_1 G_3 + \beta_2 G_5$, and T, the relative errors in the saccharification stage were 5.66%, 20.87%, 2.74%, and 0.13%, respectively. In other words, the proposed model obtained relative errors in the saccharification stage <6%, except for maltose, where the error was 20.87% (Figure 3). Similar to the liquefaction stage, the best adjustment in a concentration dynamic in the liquefaction stage was the maltodextrin concentration. The last idea indicates that the change in the use of the constraint for ψ was positive for bioprocess modeling.

5. CONCLUSIONS

A methodology to get a PBSM of the enzymatic hydrolysis process of amylaceous materials was applied in this work. The methodology considers the identifiability concept which is evaluated by using the STRIKE-GOLDD tool. This new perspective takes into account the viability of the parameter identification on the mathematical structure of the dynamic system to be iterated. Therefore, it is possible to reach a simplification of the modeling process and to obtain better results when applying optimization algorithms.

To show the effectiveness of the methodology, the enzymatic hydrolysis of amylaceous materials was carried out in the liquefaction and saccharification stages. The obtained results have a less complex mathematical structure, in comparison with previous models found in the literature. The errors with respect to the experimental data were 7.64% and 7.35% for the liquefaction and saccharification stages, respectively. A relevant modification to the model was the change in the definition of the concentration of maltodextrins based on the parameters β_1 and β_2 depending on the stage. This fact was associated with additional restrictions for ψ that allowed a better fit of the model.

Finally, the aim is to present greater robustness in the model and simulation of biological processes based on their phenomenology and parametric identification, achieving user-friendly methodologies.

ASSOCIATED CONTENT

ⓈI Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.1c06193.

Table S1: Description of used model variables for the liquefaction model. Table S2: Description of liquefaction model constants and their values. Table S3: Description of used model variables for the saccharification model. Table S4: Description of saccharification model constants and their values. Table S5: Boundary values for the parameter identification process in both models. (PDF)

AUTHOR INFORMATION

Corresponding Authors

Daniel Padierna-Vanegas − *Departamento de Energía Eléctrica y Automática, Facultad de Minas, Universidad Nacional de Colombia, Medellín 050034, Colombia; KALMAN, Grupo de investigación en Procesos Dinámicos, Universidad Nacional de Colombia, Medellín 050034, Colombia;* Email: dpadiernav@unal.edu.co

Juan Camilo Acosta-Pavas − *Departamento de Procesos y Energía, Facultad de Minas and BIOFRUN, Grupo de investigación Bioprocesos y Flujos Reactivos, Universidad Nacional de Colombia, Medellín 050034, Colombia;* Ⓞorcid.org/0000-0003-0398-7712; Email: jucacostapa@unal.edu.co

Laura María Granados-García − *Departamento de Energía Eléctrica y Automática, Facultad de Minas, Universidad Nacional de Colombia, Medellín 050034, Colombia;* Email: lmgranadosg@unal.edu.co

Héctor Antonio Botero-Castro − *KALMAN, Grupo de investigación en Procesos Dinámicos, Universidad Nacional de Colombia, Medellín 050034, Colombia;* Email: habotero@unal.edu.co

Complete contact information is available at:

https://pubs.acs.org/10.1021/acsomega.1c06193

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Zarragoitia-González, A.; Schetrite, S.; Alliet, M.; Jáuregui-Haza, U.; Albasi, C. Modelling of submerged membrane bioreactor: Conceptual study about link between activated slugde biokinetics, aeration and fouling process. *J. Membr. Sci.* **2008**, *325*, 612−624.

(2) Marambio, J. A.; Maturana, S.; Campos, B. Modelo dinámico de crecimiento de la biomasa para *Mytilus chilensis* en sistemas de cultivo en líneas. *Rev. biol. mar. oceanogr.* **2012**, *47*, 51−64.

(3) Monsalve-Bravo, G. M.; Garelli, F.; Mozumder, M. S. I.; Alvarez, H.; De Battista, H. Model-based scale-up methodology for aerobic fed-batch bioprocesses: application to polyhydroxybutyrate (PHB) production. *Bioprocess Biosyst. Eng.* **2015**, *38*, 1179−1190.

(4) Richelle, A.; Bogaerts, P. Systematic methodology for bioprocess model identification based on generalized kinetic functions. *Bioprocess Biosyst. Eng.* **2015**, *100*, 41−49.

(5) Rivera, E. C.; Costa, A. C.; Lunelli, B.; Maciel, M.; Filho, R. Kinetic Modeling and Parameter Estimation in a Tower Bioreactor for Bioethanol Production. *Appl. Biochem. Biotechnol.* **2008**, *148*, 163−73.

(6) Roeva, O. A Hybrid Genetic Algorithm for Parameter Identification of Bioprocess Models. *Large-Scale Scientific Computing*; Berlin, Heidelberg, 2012; pp 247−255.

(7) Jiménez-Villota, D. S.; Acosta-Pavas, J. C.; Betancur-Ramírez, K. J.; Ruiz-Colorado, A. A. Modeling and Kinetic Parameter Estimation of the Enzymatic Hydrolysis Process of Lignocellulosic Materials for Glucose Production. *Ind. Eng. Chem. Res.* **2020**, *59*, 16851−16867.

(8) Ashyraliyev, M.; Fomekong-Nanfack, Y.; Kaandorp, J.; Blom, J. Systems biology: Parameter estimation for biochemical models. *FEBS J.* **2009**, *276*, 886−902.

(9) Acosta-Pavas, J. C.; Ruiz-Colorado, Á. A. Approximation of Scale-Up of Enzymatic Hydrolysis Process from Phenomenological-Based Semiphysical Model and Control Theory Tools. *Ind. Eng. Chem. Res.* **2020**, *59*, 8046−8065.

(10) Bellman, R.; Åström, K. On structural identifiability. *Math. Biosci.* **1970**, *7*, 329−339.

(11) Villaverde, A. Observability and Structural Identifiability of Nonlinear Biological Systems. *Complexity* **2019**, *2019*, 1−12.

(12) Villaverde, A.; Barreiro, A.; Papachristodoulou, A. Structural Identifiability of Dynamic Systems Biology Models. *PLOS Comput. Biol.* **2016**, *12*, e1005153.

(13) Rodríguez-Fernández, M. Modelado e Identicación de Bioprocesos. Ph.D. Thesis, Universidad de Vigo, 2006.

(14) Nesterov, Y.; Nemirovskii, A. *Interior-Point Polynomial Algorithms in Convex Programming*; Society for Industrial and Applied Mathematics, 1994.

(15) Alvarez, H.; Lamanna, R.; Vega, P.; Revollar, S. Metodología para la Obtención de Modelos Semifísicos de Base Fenomenológica Aplicada a una Sulfitadora de Jugo de Caña de Azúcar. *Rev. Iberoam. Autom. Inf. Ind. RIAI* **2009**, *6*, 10−20.

(16) Pereira, C.; Resende, J.; Guerra, E.; Lima, V.; Martins, M.; Knob, A. Enzymatic conversion of sweet potato granular starch into fermentable sugars: Feasibility of sweet potato peel as alternative substrate for -amylase production. *Biocatal. Agric. Biotechnol.* **2017**, *11*, 231−238.

(17) Morales, S.; Álvarez, H.; Sánchez, C. Dynamic models for the production of glucose syrups from cassava starch. *Food Bioprod. Process.* **2008**, *86*, 25−30.

(18) Collar, C.; Jiménez, T.; Conte, P.; Fadda, C. Impact of ancient cereals, pseudocereals and legumes on starch hydrolysis and antiradical activity of technologically viable blended breads. *Carbohydr. Polym.* **2014**, *113*, 149−158.

(19) Rosicka-Kaczmarek, J.; Kwaśniewska-Karolak, I.; Nebesny, E.; Komisarczyk, A. In *Starch in Food*, 2nd ed.; Sjöö, M., Nilsson, L., Eds.; Woodhead Publishing Series in Food Science, Technology and Nutrition; Woodhead Publishing, 2018; pp 325−352.

(20) Collares, R. M.; Miklasevicius, L. V. S.; Bassaco, M. M.; Salau, N. P. G.; Mazutti, M. A.; Bisognin, D. A.; Terra, L. M. Optimization of enzymatic hydrolysis of cassava to obtain fermentable sugars. *J. Zhejiang Univ., Sci., B* **2012**, *13*, 579−586.

(21) Virunanon, C.; Ouephanit, C.; Burapatana, V.; Chulalaksananukul, W. Cassava pulp enzymatic hydrolysis process as a preliminary step in bio-alcohols production from waste starchy resources. *J. Cleaner Prod.* **2013**, *39*, 273−279.

(22) Tovar Benítez, T. Caracterización Morfológica y térmica del almidón de maíz (Zea mays L) obtenido por diferentes métodos de aislamiento. Thesis, 2008; http://dgsa.uaeh.edu.mx:8080/bibliotecadigital/handle/231104/508.

(23) Prada Ospina, R. Alternativa de aprovechamiento eficiente de residuos biodegradables: el caso del almidón residual derivado de la industrialización de la papa Bogotá. *Rev. esc.adm.neg* **2012**, 182−192.

(24) Acosta Pavas, C. J. *Escalado de hidrólisis enzimática de materiales amiláceos a partir de la simulación del proceso aplicando teoría de control.* 2019.

(25) Curvelo-Santana, J. A. C.; Ehrhardt, D. D.; Tambourgi, E. B. Otimização da produção de álcool de mandioca. *Food Sci. Technol.* **2010**, *30*, 613−617.

(26) Chung, H.-J.; Liu, Q. Impact of molecular structure of amylopectin and amylose on amylose chain association during cooling. *Carbohydr. Polym.* **2009**, *77*, 807−815.

(27) Pandey, A.; Negi, S.; Soccol, C. R. In *Current Developments in Biotechnology and Bioengineering*; Pandey, A., Negi, S., Soccol, C. R., Eds.; Elsevier, 2017.

(28) Ji, Z.; Yu, L.; Liu, H.; Bao, X.; Wang, Y.; Chen, L. Effect of pressure with shear stress on gelatinization of starches with different amylose/amylopectin ratios. *Food Hydrocolloids* **2017**, *72*, 331−337.

(29) Kong, H.; Yang, X.; Gu, Z.; Li, Z.; Cheng, L.; Hong, Y.; Li, C. Heat pretreatment improves the enzymatic hydrolysis of granular corn starch at high concentration. *Process Biochem. (Oxford, U. K.)* **2018**, *64*, 193−199.

(30) A. Gonzalez Tello, F. C. R.; Medina, A. R. Hidrolizados enzimáticos de interés en la iaa. *Aliment., Equipos Tecnol.* **1989**, 201−207.

(31) Avalos Zavaleta, L.; Llenque-Díaz, R.; Segura-Vega, R. Aislamiento y selección de cultivos nativos de Bacillus sp. productor de amilasas a partir de residuos amiláceos del mercado La Hermelinda, Trujillo, Perú. *Rev. Invest. REBIOL.* **2016**, *36*, 16−26.

(32) Cruz, K. A. R. Modelado del proceso de hidrólisis enzimática de almidónes gelatinizados del fruto de la planta de banano. M.Sc. Thesis, Universidad Nacional de Colombia, 2012.

(33) Li, Z.; Liu, W.; Gu, Z.; Li, C.; Hong, Y.; Cheng, L. The effect of starch concentration on the gelatinization and liquefaction of corn starch. *Food Hydrocolloids* **2015**, *48*, 189−196.

(34) Dura, A.; Błaszczak, W.; Rosell, C. Functionality of porous starch obtained by amylase or amyloglucosidase treatments. *Carbohydr. Polym.* **2014**, *101*, 837−845.

(35) Betiku, E.; Akindolani, O.; Ismaila, A. Enzymatic hydrolysis optimization of sweet potato (Ipomoea batatas) peel using a statistical approach. *Braz. J. Chem. Eng.* **2013**, *30*, 467−476.

(36) Baks, T.; Kappen, F. H.; Janssen, A. E.; Boom, R. M. Towards an optimal process for gelatinisation and hydrolysis of highly concentrated starch−water mixtures with alpha-amylase from B. licheniformis. *J. Cereal Sci.* **2008**, *47*, 214−225.

(37) Li, L.; Blanco, M.; lin Jane, J. Physicochemical properties of endosperm and pericarp starches during maize development. *Carbohydr. Polym.* **2007**, *67*, 630−639.

(38) Smrčková, P.; Horský, J.; Šárka, E.; Koláček, J.; Netopilík, M.; Walterová, Z.; Kruliš, Z.; Synytsya, A.; Hrušková, K. Hydrolysis of wheat B-starch and characterisation of acetylated maltodextrin. *Carbohydr. Polym.* **2013**, *98*, 43−49.

(39) Roman, L.; Gomez, M.; Li, C.; Hamaker, B. R.; Martinez, M. M. Biophysical features of cereal endosperm that decrease starch digestibility. *Carbohydr. Polym.* **2017**, *165*, 180−188.

(40) Zanin, G. M.; De Moraes, F. F. Modeling fixed and fluidized reactors for cassava starch saccharification with immobilized enzyme. *Appl. Biochem. Biotechnol.* **1997**, *63*, 527.

(41) Acosta-Pavas, J. C.; Alzate-Blandon, L.; Ruiz-Colorado, A. A. Enzymatic hydrolysis of wheat starch for glucose syrup production. *Dyna* **2020**, *87*, 173−182.