Article

Estimation of Biomass Enzymatic Hydrolysis State in Stirred Tank Reactor through Moving Horizon Algorithms with Fixed and Dynamic Fuzzy Weights

Vitor B. Furlong [1], Luciano J. Corrêa [2], Fernando V. Lima [3] , Roberto C. Giordano [1,4] and Marcelo P. A. Ribeiro [1,4,*]

[1] Graduate Program of Chemical Engineering, Federal University of São Carlos, P.O. Box 676, São Carlos 13565-905, SP, Brazil; vitorfurlong@gmail.com (V.B.F.); roberto@ufscar.br (R.C.G.)

[2] Department of Engineering, Federal University of Lavras, P. O. Box 3037, Lavras 37200-000, MG, Brazil; luciano.correa@ufla.br

[3] Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, WV 26506, USA; fernando.lima@mail.wvu.edu

[4] Chemical Engineering Department, Federal University of São Carlos, P.O. Box 676, São Carlos 13565-905, SP, Brazil

* Correspondence: marceloribeiro@ufscar.br

Abstract: Second generation ethanol faces challenges before profitable implementation. Biomass hydrolysis is one of the bottlenecks, especially when this process occurs at high solids loading and with enzymatic catalysts. Under this setting, kinetic modeling and reaction monitoring are hindered due to the conditions of the medium, while increasing the mixing power. An algorithm that addresses these challenges might improve the reactor performance. In this work, a soft sensor that is based on agitation power measurements that uses an Artificial Neural Network (ANN) as an internal model is proposed in order to predict free carbohydrates concentrations. The developed soft sensor is used in a Moving Horizon Estimator (MHE) algorithm to improve the prediction of state variables during biomass hydrolysis. The algorithm is developed and used for batch and fed-batch hydrolysis experimental runs. An alteration of the classical MHE is proposed for improving prediction, using a novel fuzzy rule to alter the filter weights online. This alteration improved the prediction when compared to the original MHE in both training data sets (tracking error decreased 13%) and in test data sets, where the error reduction obtained is 44%.

Keywords: artificial neural network; biomass enzymatic hydrolysis; fuzzy logic; local linear model tree; moving horizon estimation; process monitoring; soft sensing

---

1. Introduction

The optimization of reactors to operate at high solids concentrations is an important step to enable the enzymatic hydrolysis of lignocellulosic biomass as a feasible technology. However, operating under these conditions can cause some issues in the saccharification kinetics, from substrate and product inhibition to heterogeneous agitation, due to several challenges. Some of these challenges can be attributed to reactor homogenization, which is a difficult task at high apparent viscosity. The suspended solids also greatly increase the power consumption from the stirring motor. The energy to agitate the reactor can be as drastic to account for half of the production cost of the final bioethanol [1]. Thus, great attention should be paid to this production state if a financially sound second generation production technology is one's end goal. This challenge motivates the study of the hydrolysis reactor and its architectures.

To overcome these bottlenecks, the torque necessary to stir the reactor, and thus, indirectly, the reactor rheology must be studied. Stirring torque, or agitation power, is a measurement that is

widely used to design reactors and it can be monitored online when a dynamometer is coupled to the stirring apparatus. Thus, the torque can be used to estimate the power consumption of a given reactor architecture during regular operation. This measurement can be used to assess how alterations in the reactor's architecture and its impellers impact the process power consumption [2–5]. Moreover, torque dynamometers that monitor the power consumption of the reactor can be used to control the substrate feeding rate, which diminishes the overall process power consumption while retaining similar productivity and yield [6]. Such control strategies can be used to reduce the process overall cost.

Furthermore, torque measurements can also be used to estimate the kinetics inside the hydrolysis reactor. However, most researches focus on the prediction of viscosity within the reactor for scale up purposes [3,7,8]. The notable exception is the work of Kadhum et al. [9], where a semi-empirical model is proposed to relate torque to insoluble solids and glucose throughout the hydrolysis. However, in this work, the glucose and solids concentration are inputs and not model outputs. The output is the torque itself. This means that using this approach to estimate the products and substrate concentrations would require an inversion of the model, which results in large errors in the direct prediction of these compounds. A model that explicitly estimates the reaction compounds from torque measurements would be easier to use. This model can be generated through several methods. One alternative is the utilization of soft sensors due to the high non-linearity and complex nature of the system. Soft sensors are software capable of predicting non-measured variables from readily available process data. These are used when the direct instrumentation of a process is infeasible or impractical [10].

The models used in data driven soft sensors, those that are trained with experimental data, are usually empirical and several algorithm alternatives are presented in literature [11]. One of the most popular class of models used within soft sensors are Artificial Neural Networks (ANNs). ANNs are empirical mathematical models that are based on how the human brain handles information and they are widely used for pattern recognition [12].

Nonetheless, measurements from soft sensors can be susceptible to data noise. In order to diminish the error in the prediction, the torque measurement can be coupled to a state estimator to further improve the precision of state estimation. State estimators are algorithms that are used to estimate state variables from information that is provided by a previously established system model and instrumentation data [13].

Several estimators can be employed. The most well-established estimator is the Kalman filter, which is only applied to linear systems. The extended Kalman filter is an adaptation of the latter method, based on the local linearization of the dynamic system, allowing for its extension to nonlinear systems. However, despite its popularity and computational efficiency, it might diverge from the true state and not satisfy process constrains [14]. Optimization-based estimators are some alternatives that do not rely on linearization of the nonlinear system. In optimization-based estimators, a nonlinear optimization is run for a time window for each time step to find an estimation of the current state that accounts for both the instrumentation and the model predictions. Furthermore, state constraints can be incorporated into the formulation and the effect of disturbances can be significantly reduced [15].

Optimization-based estimators are divided according to the amount of data points that are optimized and how the optimization occurs. Full information estimation occurs when all of the data points generated by the model and instrumentation are considered in each optimization step. It is the most adequate when stability and optimality are concerned [16]. However, in many applications, especially when the estimation occurs on-line, the mathematical complexity of the system requires a great computational effort, which hampers the use of this method [16].

In contrast, the Moving Horizon Estimator (MHE) uses a dynamic window reducing the computational effort and still retaining favorable characteristics. MHEs are very well established as state estimators for chemical processes, as they can handle nonlinearities in the process dynamics and state constraints that are inherent to most reaction systems in a fairly straightforward manner [17]. This type of algorithm is very versatile and it has been applied to different processes, from solid and

submerged fermentations [18,19], metabolic fluxes estimation [20], and lipids production [21]. Thus, it is a good candidate for application in the enzymatic hydrolysis of biomass.

The aim of this work is first to design and test a soft sensor that is based on agitation power measurements to predict free carbohydrates concentrations while using an ANN as internal model. The soft sensor and a recently developed kinetic model are used in an MHE algorithm. The MHE is then tuned to optimize the monitoring of the enzymatic hydrolysis using sugarcane bagasse as substrate. The Materials and Methods Section was divided into three subsections to explain the performed procedure. Section 2.1 provides information regarding the source of the experimental data and kinetic model, including how experimental data were obtained and how the kinetic model was developed. Section 2.2 details how the soft sensor was developed. Section 2.3 explains how the MHE parameters (weights and membership degree) were optimized. The methods section is followed by the results of the employed methods and the conclusions.

2. Materials and Methods

2.1. Hydrolysis Data and Kinetic Model

The experimental assays that were used in this paper to tune both the ANN and MHE algorithm, as well as the kinetic model used in the state estimator, are presented in [22]. Sections 2.1.1–2.1.3 detail the material and analytical methods.

All of the numerical procedures (model integration, soft sensing and filter tuning) were performed in SCILAB (v.6.0.2, ESO Group, Paris, France, 2019) in a computer with an AMD FXTM-8350 and 15.7 Gb of random access memory, running a 64-bit Linux Mint 18.3 as operating system. For the model integration, a Runge-Kutta 4th order algorithm was used (ode function with "rk" specification).

2.1.1. Enzyme and Substrates

The enzymatic complex employed in all assays was CELIC CTEC 2, donated by Novozymes Latin America (Araucária, Paraná, Brazil). The enzymatic complex protein concentration was 75 $mg_{protein}$ mL^{-1}, with activity of approximately 203 FPU mL^{-1}.

Steam-exploded sugarcane bagasse was used as substrate (1667 kPa and 205 °C for 20 min) provided by Centro de Tecnologia Canavieira (Piracicaba, São Paulo, Brazil). The biomass main components were: 43.1 ± 0.1% cellulose; 12.4 ± 0.1% hemicellulose, 28.8 ± 1.9% lignin and 4.7 ± 0.1% ash ($g_{compound}$ $g_{dry\ biomass}^{-1}$); the remaining are non-analyzed compounds, not considered in the models. The bagasse composition was determined according to Sluiter and collaborators [23]. The moisture content of the pretreated bagasse was 59.9 ± 1.6%. The volume of water at the end of all assays was 3 L.

2.1.2. Assay Conditions

Assays were conducted with three feeding profiles. All of the profiles added up to 600 $g_{substrate}$ and 2.22 $g_{protein}$ (approximately 10 FPU.$g_{substrate}^{-1}$) of enzymatic complex at the end of each process, regardless of the feeding policy. The first assay, Data Set 1, is a High Solids Batch (HSB). This assay had all of the compounds added at the beginning of the process. The other two assays were operated in fed-batch mode, where the feeding profiles added up to the same substrate mass as the one of the batch processes. Feedings were performed at discrete times, as presented in Table 1.

Data Set 2 is the Low Solids Fed-Batch (LSF) as the feeding intervals were spaced enough not to generate a visible high solids concentration throughout the hydrolysis. In Data Set 3, additions were closer in time when compared to LSF. In this assay, an initial low solids concentration occurred, but with following substrate additions the medium turned to a high-solids load process. Thus, this assay was named Mixed Profile Fed-Batch (MPF). Duplicates were performed for all assays by conducting the experiment in different days under the same policy.

The reactor used was a 3 L working volume baffled jacketed stirred reactor (reactor's diameter 0.16 m, height 0.37 m and liquid height 0.21 m, baffle's height 0.19 m, and thickness 15 mm) with two

elephant ear impellers (diameter 0.08 m, New Brunswick Scientific®, Edison, NJ, USA) equidistant from one another, the reactor bottom and the liquid surface. The upper impeller generates a downwards pumping flow and the lower impeller pumps upwards. The stirring motor was placed on top of a plate with bearings, so the motor was free to rotate when opposing forces acted upon it. A digital dynamometer (FG 6005 SD, Lutron Electronics, Coopersburg, PA, USA) was coupled to the bearing assembly in order to measure the opposing force. This allows the agitation power to be measured. The reactor media was kept at 50 °C by a thermostatic bath. The agitation speed was 470 RPM.

Table 1. Data Sets Feeding Profiles.

Training Data Sets						Test Data Sets		
Data Set 1—High Solids Batch			Data Set 2—Low Solids Fed-Batch			Data Set 3—Mixed Profile Fed-Batch		
Feeding Time (h)	Solids (g_{solids})	Enzyme ($g_{protein}$)	Feeding Time (h)	Solids (g_{solids})	Enzyme ($g_{protein}$)	Feeding Time (h)	Solids (g_{solids})	Enzyme ($g_{protein}$)
0.0	600.0	2.22	0.0	150.0	2.22	0.0	150.0	2.22
-	-	-	2.0	150.0	-	0.5	150.0	-
-	-	-	12.0	150.0	-	1.0	150.0	-
-	-	-	24.0	150.0	-	2.0	150.0	-

2.1.3. Sampling and Analysis

The sampling was manually done at 0.5, 1, 2, 4, 6, 8, 12, 24, 36, 48, 60, 72, and 96 h. Samples were placed in boiling water bath for 10 min. to stop reaction and then refrigerated. These were analyzed for glucose and xylose by high performance liquid chromatography [24]. The samples were filtered (hydrophilic polyvinylidene fluoride filter, 0.2 μm) into autosampler vials. These were analyzed in a Shimadzu SCL-10A chromatograph (Shimadzu Corp., Kyoto, Japan), with refraction index detector RID10-A (Shimadzu Corp., Kyoto, Japan), column Animex HPX-87H Bio-rad (Bio-rad Laboratories Inc., Rio de Janeiro, Brazil), and mobile phase sulfuric acid 5 mM at 0.6 mL min^{-1}. The sample values were compared to the previously established standards.

2.1.4. Mathematical Modeling

The kinetic model that was used as part of the MHE implementation was developed in a previous work [22]. This kinetic model was the one that best describes glucose and xylose during enzymatic hydrolysis of a steam-exploded sugarcane bagasse under batch and fed-batch operations. A brief description is given in this subsection, and more information can be obtained in the dedicated paper.

To describe the lignocellulosic material hydrolysis, five reactions are considered:

Reaction 1: Cellulose → $\gamma_{Cl\text{-}Cb}$ Cellobiose

Reaction 2: Cellobiose → $\gamma_{Cb\text{-}Gl}$ Glucose

Reaction 3: Hemicellulose → $\gamma_{He\text{-}Xy}$ Xylose

Reaction 4: Lignin → Lignin

Reaction 5: Enzyme → Inactive Enzyme

In the reaction scheme, γ are the pseudo-stoichiometric mass relations between substrates and products for each reaction. The values for these parameters were [5]: $\gamma_{Cl\text{-}Cb}$ = 1.056 $g_{Cellobiose}$ $g_{Cellulose}$$^{-1}$, $\gamma_{Cl\text{-}Gl}$ = 1.111 $g_{Glucose}$ $g_{Cellulose}$$^{-1}$, $\gamma_{Cb\text{-}Gl}$ = 1.056 $g_{Glucose}$ $g_{Cellobiose}$$^{-1}$, $\gamma_{He\text{-}Xy}$ = 0.841 g_{Xylose} $g_{Hemicellulose}$$^{-1}$ [25].

Reactions 1 and 3 are considered to be heterogeneous, as they represent the breakdown of cellulose and hemicellulose. These are modeled via Modified Michaelis–Menten kinetics with competitive product inhibition, presented in Equation (1).

$$\alpha_i = \frac{k_i \cdot [E_i] \cdot [S_i]}{Km_i \cdot \left(1 + \frac{[P_i]}{K_{P,i}}\right) + [E_i]} \tag{1}$$

In this equation, α_i (g L^{-1} min^{-1}) are reaction rates, where the subscripted "i" denotes the reaction used, either 1 or 3. The same nomenclature is used in the subsequent variables, k_i are kinetic constants (min^{-1}), E_i are enzyme concentrations (g L^{-1}), [S_i] are substrate concentrations (g L^{-1}) and Km_i are the Michaelis–Menten modified constant (g$_{Enz}$ L^{-1}), Kp_i are competitive inhibition constants for products (g L^{-1}), and P_i are product concentrations (g L^{-1}).

The hydrolysis of cellobiose into glucose (Reaction 2) is homogeneous. This equation is modeled with a Pseudo-Homogeneous Michaelis–Menten with Competitive Product Inhibition model, as presented in Equation (2).

$$\alpha_2 = \frac{k_2 \cdot [E_2] \cdot [S_2]}{Km_2 \cdot \left(1 + \frac{[P_2]}{K_{P,2}}\right) + [S_2]} \tag{2}$$

The same variables are used in this equation; however, Km_2 is the Michaelis–Menten constant (g L^{-1}) for substrate.

Reaction 5 is included, despite lignin being an inert. Thus, this "reaction" just reflects the accumulation of lignin in the reactor when it operates in fed-batch mode.

Enzymatic inactivation is represented by a generic first order equation, as presented in Equation (3).

$$\alpha_5 = ke \cdot [E] \tag{3}$$

where ke is a first order inactivation parameter (min^{-1}). Table 2 presents the parameters of these equations.

Table 2. High and low models parameters—parameter ± standard error.

Reaction	Solids Model	Parameters			
		k (min^{-1})	Km (g L^{-1})	Kp (g L^{-1})	ke (min^{-1})
1	High	$(3.03 \pm 0.00) \times 10^{-3}$	$(5.31 \pm 0.01) \times 10^{-2}$	$(7.65 \pm 0.01) \times 10^{-4}$	-
	Low	$(2.67 \pm 0.38) \times 10^{-3}$	$(9.75 \pm 0.25) \times 10^{-3}$	$(1.11 \pm 0.03) \times 10^{-3}$	-
3	High	$(9.13 \pm 0.00) \times 10^{-2}$	$(3.82 \pm 0.01) \times 10^{-4}$	$(1.83 \pm 0.00) \times 10^{-1}$	-
	Low	$(6.41 \pm 0.07) \times 10^{-4}$	$(4.49 \pm 0.10) \times 10^{-6}$	$(2.50 \pm 0.00) \times 10^{-1}$	-
4	High	$(1.28 \pm 0.00) \times 10^{-3}$	$(2.02 \pm 0.00) \times 10^{-2}$	$(3.46 \pm 0.00) \times 10^{-1}$	-
	Low	$(1.13 \pm 0.02) \times 10^{-1}$	$(7.80 \pm 0.15) \times 10^{-4}$	$(3.73 \pm 0.35) \times 10^{-3}$	-
5	High	-	-	-	$(1.16 \pm 0.00) \times 10^{-3}$
	Low	-	-	-	$(1.15 \pm 0.35) \times 10^{-3}$

Following the formalism that is proposed in reference [26], Equation (4) represents the mass balance for the components in the reactor. Since substrate feeding was discrete, the fed-batch process was solved as a sequence of batch processes by re-initializing the initial conditions at each feeding with the previous final concentrations.

$$\frac{d}{dt} \cdot \begin{bmatrix} Cl \\ Cb \\ Gl \\ He \\ Xy \\ Lg \\ E \end{bmatrix}_{7 \times 1} = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 & 0 \\ \gamma_{Cl-Cb} & 0 & -1 & 0 & 0 & 0 \\ 0 & \gamma_{Cl-Gl} & \gamma_{Cb-Gl} & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & \gamma_{He-Xy} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}_{7 \times 6} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix}_{6 \times 1} \tag{4}$$

In Equation (4), the column vector in the left-hand side of the equation represents the concentrations of reactive cellulose ([Cl]), cellobiose ([Cb]), glucose ([Gl]), hemicellulose ([He]), xylose ([Xy]), lignin ([Lg]), and enzymatic complex activity ([E]). The resulting 7×5 matrix at the right-hand side of the

equation is the pseudo-stoichiometric matrix and the vector (α_i) are the reaction rates of reaction 1 to 5, as previously described.

The novelty of this model lies in how it was fitted and how it operates. A fuzzy membership rule interpolates between two different set of parameters for the described kinetic equations, one set was fitted using data from Data Set 1, High Solids Model (HSM), and the other was fitted using Data Set 2, the Low Solids Model (LSM). Thus, when solving for the reaction kinetics, both models generate independent reaction rates, α^{HSM} (High Solids Model reaction rate) and α^{LSM} (Low Solids Model reaction rate) for every equation at a same reaction instant. A Fuzzy Model (FM) then interpolates both standalone reaction rates with an optimized Membership Function (MSF). The MSF calculates the Membership Degree (MD) of each model based on the total amount of Reactive Solids (RS), the addition of the current concentration of cellulose and hemicellulose. The Membership Degree of the HSM (MD^{HSM}) is calculated with the linear fuzzy rule. The upper and lower bounds of the MSF were 75.35 g L^{-1} and 61.52 g L^{-1}, respectively.

With the HSM and LSM membership degrees, the reaction rate for the FM was calculated with Equation (5). It is the output of the Takagi–Sugeno system.

$$\alpha^{Fuzzy} = MD^{HSM} \cdot \alpha^{HSM} + \left(1 - MD^{HSM}\right) \cdot \alpha^{LSM} \qquad (5)$$

where α^{FUZZY} are the reaction rates for each hydrolysis reaction. Thus, a prediction of the entire state vector is obtained. SCILAB's Ordinary Differential Equation Solver (ODE) was used to integrate the state variables, with the generated model parameters.

2.2. Agitation Power Soft Sensing

2.2.1. Soft Sensing Data

ANN training requires a relatively large amount of data to fully recognize the useful patterns. Using only analytical data from the data sets would not be enough in this case, or at least it would generate a network with poor performance. A cubic spline was used in the internal batches to generate new data points where previously there were none (in the periods between feeding times) of an assay to generate more glucose and xylose data points. The interpolated data may not completely match the actual state variable profiles; however, this procedure is necessary when the available data set are not enough for fitting complex empirical models, such as ANNs.

Agitation power measurement was realized during the assays every 10 s, and the average from all of the measurements within 10 min. was used as one data point for that time period. The resulting data was fed through a LOWESS (Locally Weighted Scatterplot Smoothing) filter to reduce noise. This smoothed data was used as an input for the ANN.

Other inputs to the neural network were the first derivative of the agitation power and the biomass added to the reactor until the current data point. The agitation power derivative was calculated within internal batches while using numerical differentiation with forward, central, and backward finite differences for first, internal, and last data points, respectively.

Data from the HSB and LSF were used to train the network, and data from the MSF were used as test assay.

2.2.2. Local Linear Model Trees Algorithm

A Local Linear Model Trees (LOLIMOT) algorithm was used to translate the input variables into the desired carbohydrates concentration. LOLIMOT is a special case of the Local Linear Neuro-Fuzzy Models (LLNFM), which is a subset of ANNs. In local linear models, different models are generated for different regions of the input domain. These models may interact in order to compose the predicted value following different approaches. LLNFMs use different neurons in a hidden layer of an ANN for each Local Linear Model (LLM). Each neuron also possesses a validity function that regulates where each

LLM is used and to what degree. In LLNFMs, a fuzzy membership function is also used alongside the linear model, which allows for multiples models to be "active" for a given data point, generating a more flexible model structure. Several types of membership functions can be used; however, normalized Gaussian functions are usually chosen, since they are highly flexible with a relatively small number of parameters [27].

The "tree" in LOLIMOT regards the manner by which the validity range of each LLM and membership functions are generated. The algorithm that is illustrated here is based on reference [28], and it was implemented while using the Lolimot Scilab ATOM module, based on the same algorithm.

This algorithm uses axis-orthogonal cuts in the input space to generate hyper-rectangles, where the center of the Gaussian function and the linear models are placed. Each additional partitioning of the input variables generates a new LLM and activation function. Figure 1 demonstrates how subsequent cuts are added in a two-dimensional input variables space.
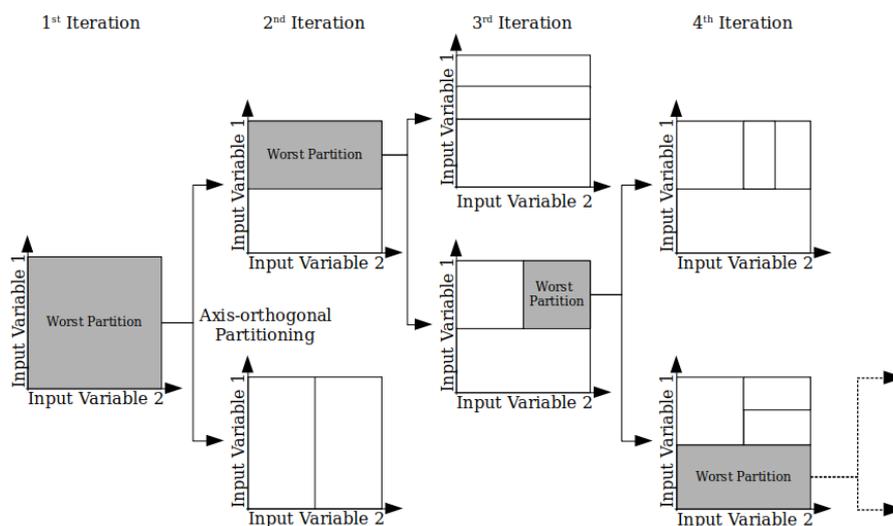

Figure 1. Local Linear Model Tree (LOLIMOT) partitioning procedure for two-dimensional input spaces. Adapted from [28].

Figure 1 illustrates how the LOLIMOT algorithm adds a new partition in a simple model. First, it evaluates the loss function of all existing LLMs, the worst partition is then divided into new hyper rectangles, as many as the number of input variables. New membership functions and LLMs are added and optimized, for all new partitions. The algorithm loops around and looks for the worst partition to be cut if a stopping criterion is not met [28].

Several stopping criteria can be used. In this manuscript, optimum architecture was achieved when the average standard error from the validation data departs from the average standard error from the training data. When this happens, the addition of new neurons starts to model the noise from the samples, which disrupts the network inference and characterizes overfitting [27].

HSB and LSF data were combined and normalized between 0 and 1. Twenty percent of the dataset was randomly separated as a validation set, being used only to evaluate and choose the network complexity. The remainder of the combined dataset was used as a training set to fit the model parameters. The data from the MSF was only used as a test dataset, in order to evaluate the extrapolation capacity of the optimized LOLIMOT.

The LOLIMOT algorithm is highly optimized to be independent of external operators and the Scilab Lolimot Toolbox implementation only leaves one tuning parameter for the operator, a smoothing parameter of the predicted output. As this is an empirical parameter, six different levels were tested for the smoothing parameter, which range from 0.1 to 0.6 and uniformly increasing by 0.1 in each level. Furthermore, separate models were composed for glucose and xylose. Each network was tested with up to 90 rules. The main reason for using different models for glucose and xylose, as noted by Nelles

(2001) [27], is that MISO (multiple-input single-output) models tend to be more flexible, and easier to construct and apply than MIMO (multiple-input multiple-output) models.

2.3. Moving Horizon Estimator

MHE is a modification of the full information estimator to diminish the total number of points used in the optimization problem solved for state estimation. This is achieved by using a moving window of prediction. As the process occurs, new time periods to be used in the optimization are added to the window until the size of the window (N) is filled with T time periods. When the window is full (T = N), the oldest state group is dropped to make room for a new time period set of states. Figure 2 presents a schematic of how the moving window works.
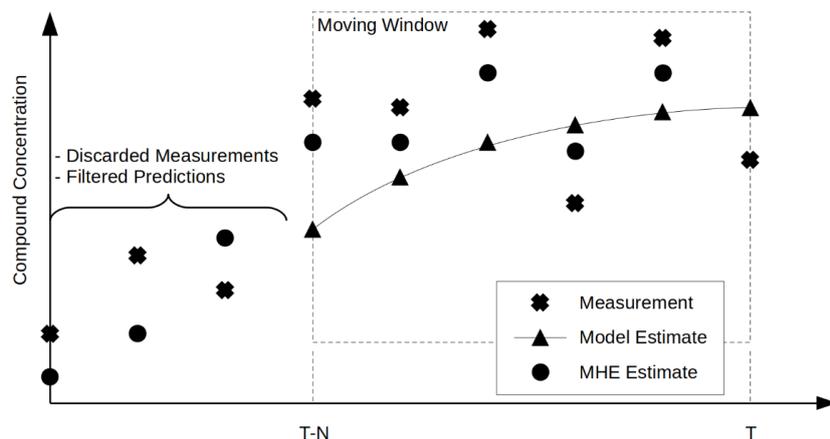

Figure 2. Moving horizon in state estimation schematic. Adapted from [29].

Equation (6) presents the MHE cost function, as optimized in each window.

$$\min_{\hat{X}(T-N+1:T)} \sum_{j=T-N+1}^{T} \|\hat{X}_j - X_j^-\|_{L_w}^2 + \sum_{j=T-N+1}^{T} \|\hat{Y}_j - Y_j\|_{L_v}^2 \tag{6}$$

where the notation $\|z_R^2\|$ represents $z^T R\, z$ and the squared norm of a vector z weighted by a matrix R. $\hat{X}_j$ and $\hat{Y}_j$ are, respectively, the state and output vectors that were predicted by the MHE at the j-th time. X_j^- is the state predicted by the model alone, while not considering the current measurement Y_j. N is the window size and T is the current estimation time. The instrumentation data (Y) are obtained via soft sensing, and the model prediction (X^-) is obtained by solving the kinetic model from T-N until the current time T. The filter estimation ($\hat{X}$ and $\hat{Y}$) is computed at each observation step (each movement of the window). In this work, the instrumentation directly predicts glucose and xylose concentrations, and these are the only two state variables that were subjected to optimization. Other variables are stoichiometrically calculated from the estimated glucose and xylose concentrations.

The MHE cost function is optimized in each step via a Levenberg–Marquardt algorithm, operating up to 1000 iterations, or until the error improvement between two iterations is less than 10^{-6}.

The filter estimation is highly sensitive to the matrices L_w and L_v. In the classical implementation of the MHE, these are the inverse of the covariance matrices of the process and measurement noises, respectively. However, in complex systems, as in the case of biomass hydrolysis, obtaining the noise covariance and the associated weights may not be so trivial. L_w, for instance, could be approximately estimated from the parameter covariance matrix that was obtained during the fitting procedure. It would be a function of state and the parameter sensitivity matrix would have to be evaluated each time step. This approach would be computationally demanding in a system with high output data frequency. Usually, L_w is assumed as a constant matrix and it is estimated by the variances of the errors

between model and analytical data or by the noise covariance. L_v is usually a constant matrix and it might be obtained by output replicates.

In a practical sense for the implementation, L_w and L_v are used as weights for the model and measurement errors. These weights regulate the MHE optimization. Therefore, when L_w and L_v estimates do not provide good inference results due to the complexity of the system, an optimization methodology can be applied to determine these weights. Thus, in this work, a tuning methodology was used to ensure that the value of the matrices L_w and L_v would generate adequate prediction. In this implementation, L_w and L_v are 2×2 diagonal matrices, where the weights are placed in the main diagonal, as only the free carbohydrates (xylose and glucose) are considered in the state estimation.

2.3.1. Fixed Weights Tuning

An optimization loop is used outside of the MHE to better estimate L_w and L_v. A Levenberg–Marquardt algorithm is employed to estimate the values of each weight. The analytical data from the HSB and LSF is used to build a weighted sum of squared errors (F) to generate a performance index for this optimization. Equation (7) demonstrates this cost function.

$$F = \mathrm{e}^T \cdot Q \cdot e \tag{7}$$

where the error vector e is the difference between the final MHE prediction and experimental data. Q is an $n \times n$ diagonal weight matrix and n is the number of experimental data. The elements Q_{ii} are the inverse of the carbohydrate replicate variance ($1/\sigma^2_i$). The variance that is estimated for glucose was 0.392 g^2/L^2, and for xylose was 0.773 g^2/L^2. The analysis of errors that led to these weight values is presented in the Supplementary Material of reference [12]. Data Set 3 (MSF) is used as a test case in order to check the estimator extrapolation capacity.

The tuning algorithm operated for 100 iterations. The only parameter of the filter optimized outside of the tuning algorithm was the size of the window. Tuning was performed with window sizes of 4, 5, and 6 time periods. State variables estimation occurred every 10 min. of real process time, i.e., optimizations occurred six times for every hour of process operation. Table 3 presents a pseudo-code of the tuning algorithm.

Table 3. Moving Horizon Estimator (MHE) Tuning Pseudo-code.

1:	**# Initialization**
2:	Set Initial Parameters: N, L_w and L_v.
3:	**# Main Loop**
4:	**While:** Tuning Algorithm Stopping Criteria = False **do:**
5:	**For**: Every Training Data Set
6:	**For:** Hydrolysis Time
7:	**If:** Window Size < Maximum Size **do:**
8:	Add Measurement from the Neural Network to Window
9:	Integrate Model for Window
10:	**Else:**
11:	Updates Window
12:	Measurement from the Neural Network
13:	Integrate Model for Window
14:	**End If**
15:	**# State Filtering Step**
16:	**While:** Filtering Stopping Criteria = False **do:**
17:	Evaluate MHE Cost Function with Current Weights
18:	Reevaluate Stopping Criteria
19:	**End While**
20:	**If:** Current Time = Feeding Time
21:	Update State With Substrate Addition
22:	Reinitialize Moving Horizon Window
23:	**End If**
24:	**End For:** Hydrolysis Time
25:	**End For:** Number of Feeding Profile
26:	Calculate Current Weights Estimation Squared Errors
27:	Reevaluate Tuning Stopping Criteria
28:	Update weights according to Levenberg-Marquardt algorithm
29:	**End While**
30:	**# Final Procedures**
31:	Save Optimized Errors

2.3.2. Fuzzy Weights Tuning

As mentioned earlier, L_w could be estimated from the parameter covariance, turning it in a time varying matrix. In this work, a different approach is proposed in order to obtain dynamic weights in the state estimation to improve prediction in such complex system. Subsequently, more pertinence is given to the model or to the instrumentation where they can provide more information about the system. Notice that the use of the fuzzy dynamic weights is computationally intensive to build, but it requires less computational effort when the filter is running.

Thus, an alteration of the MHE algorithm is proposed here, in which the diagonal matrices L_w and L_v are multiplied by the linear fuzzy membership function. The membership function is a linear fuzzy rule, ranging from 0 to 1. This function is presented in Figure 3 and it is described by Equation (8).

$$\text{MD}^{\text{Inst}} = \begin{cases} 0 & if \ [S] \leq LL \\ \frac{[S]-LL}{LU-LL} & if \ LL < [S] < LU \\ 1 & if \ LU \leq [S] \geq HL \\ \frac{[S]-HL}{HU-HL} & if \ HL < [S] < HU \\ 0 & if \ [S] \geq LL \end{cases} \qquad (8)$$

where MD^{Inst} is the membership degree of the instrumentation weights, (S) is the solids concentration in the current point, as obtained by the addition of the concentrations of insoluble substrates (cellulose, hemicellulose, and lignin), LL and LU are the Lower and Upper Bounds of the Low Solids side of the fuzzy rule, respectively, and HU and HL are the Lower and Upper Bounds of the High Solids side.

The parameters LL, LU, HU, and HL are optimized together with the weights in the outer optimization loop. The membership degree of the model is obtained by subtracting the MD^{Inst} from 1. The weight vectors then multiply these membership degrees. Thus, as solids are hydrolyzed, the weights would change, altering Equation (6) into (9).

$$\min_{\hat{X}(T-N+1:T)} \sum_{j=T-N+1}^{T} \left(1 - MD_j^{Inst}\right)\|\hat{X}_j - X_j^-\|_{L_w}^2 + \sum_{j=T-N+1}^{T} MD_j^{Inst}\|\hat{Y}_j - Y_j\|_{L_v}^2 \tag{9}$$

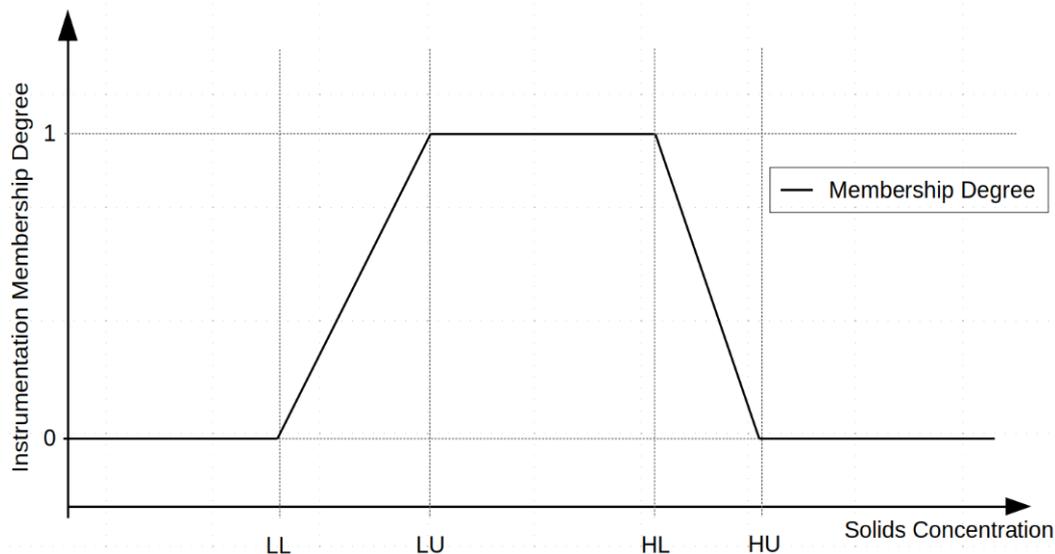The rest of the tuning procedure remains the same as the one that is presented in Table 3.

Figure 3. Instrumentation membership degree fuzzy rule.

3. Results and Discussion

3.1. Soft Sensing Optimization

The best results for the network number of rules optimization are obtained when using the smoothing parameter of 0.4 for both networks based on the MSE of training and validation data sets. Figure 4 presents the error plots for these optimizations.

Figure 4 demonstrates that, for both networks, the validation MSE initially presents a somewhat erratic pattern. This situation is addressed when more rules are added.

As more rules are added, the validation MSE becomes greater than the training error, which is the expected behavior, and it follows the decrease of the training MSE. It is noticeable that, at some points, the validation error increases from one quantity of rules to another. At such points, one could be tempted to consider the previous smaller error condition for the optimum architecture. However, if the validation MSE continues to decrease at following rule additions, it is possible to attribute the increase in the error to small variations in the current fitting stage. The optimum number of rules is the one where the validation MSE deviates clearly from the training MSE decreasing tendency. In the presented algorithm, the number of rules where this was considered to occur was at 38 and 67 rules for the glucose and xylose networks, respectively. These points are marked in Figure 4 with vertical solid bars.

Figure 5 presents the prediction of the optimum networks for the training sets HSB and LSF input data and analytical data.
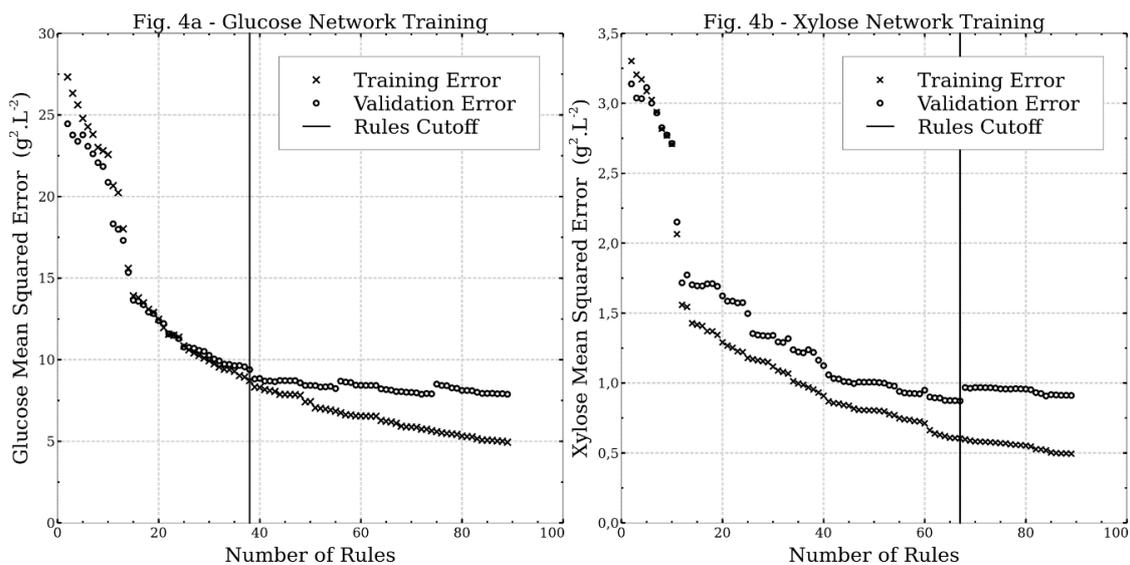
Figure 4. Error dispersion for Artificial Neural Network (ANN) number of rule optimization with smoothing parameter 0.4. (**a**) Glucose network error dispersion; and, (**b**) Xylose network error dispersion; crosses represent training mean squared error (20% of data) and circles training mean squared errors (80% of data); solid vertical line is the optimum number of rules.
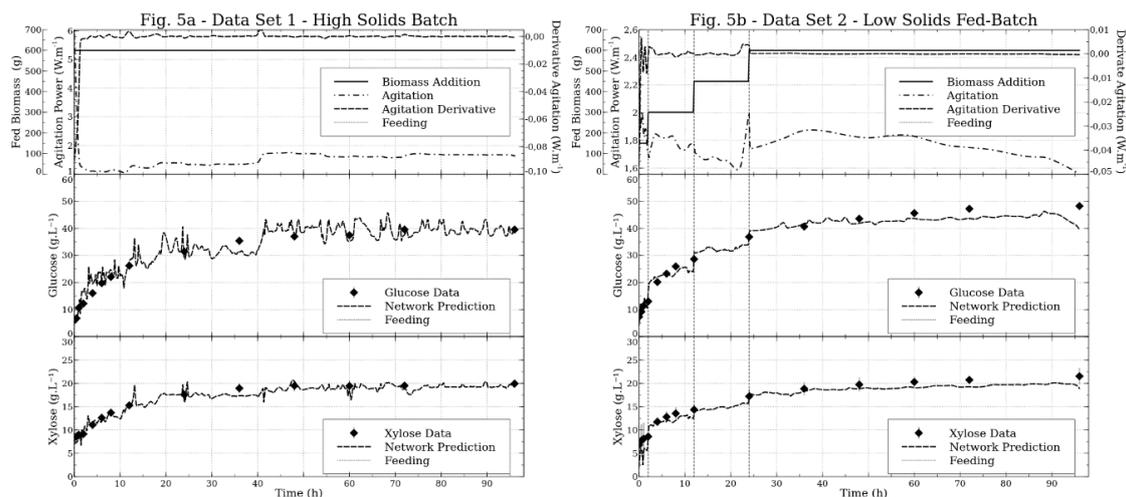

Figure 5. Carbohydrates concentration prediction for optimum Artificial Neural Network (ANN) using the training data sets. (**a**) High Solid Batch (HSB) predictions; and, (**b**) Low Solids Fed-batch (LSB) predictions; top graphs are input variables to the network for each data set, middle and bottom graphs are glucose and xylose predictions; error bars in data points are standard errors between experiment duplicates for compound concentration; vertical dotted lines are the biomass feeding times.

Figure 5 shows that the estimates in the training data have high inherent noise; however, for a prediction with no phenomenological or kinetic data, using only instantaneous measurements, the prediction follows the analytical values fairly well, being capable of predicting the steps in concentrations, due to biomass feedings in the LSF. Figure 6 presents the network prediction when dealing with data that are not present in the training/validation data set.

The prediction in Figure 6 is clearly less accurate than the one that is displayed in Figure 5. This is expected for the LOLIMOT, as it is an empirical model and, thus, might have poor extrapolation capability. The subsequent feedings in the MSF profile (from 0 h to 2 h), much closer than the times in the LSF (from 0 h to 24 h), generates an overestimation of the carbohydrates' concentrations. However, towards the end of the process, the prediction becomes more accurate. From Tables 4 and 5, it is also

clear that the soft sensor is less accurate than the model prediction. Nevertheless, the behavior of such predictions can be improved using the process model and the state estimation implementation.
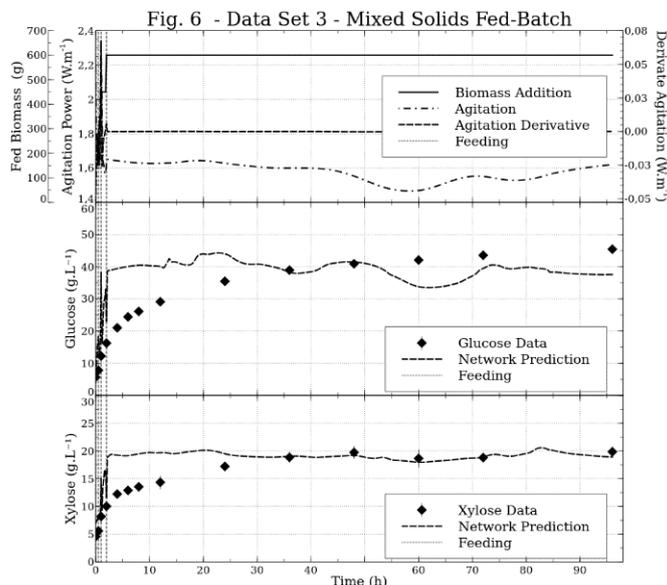

Figure 6. Carbohydrates concentration prediction for optimum Artificial Neural Network (ANN) for Mixed Solids Fed-batch (MSF) predictions; top graph is input variables to the network, middle and bottom graphs are glucose and xylose predictions; error bars in data points are standard errors between experiment duplicates for compound concentration; vertical dotted lines are the biomass feeding times.

Table 4. Root Mean Squared Error (RMSE) of product predictions for all Moving Horizon Estimator (MHE) tuning conditions as well as pure model and soft sensing.

Prediction Source	Training Data Sets		Test Data Set
	HSB (g L^{-1})	LSF (g L^{-1})	MSF (g L^{-1})
Model Prediction	0.673	1.696	3.066
Soft Sensing	5.340	3.680	7.589
Window Size	Fixed Weights MHE		
4	0.748	1.399	3.476
5	0.758	1.400	3.513
6	0.824	1.461	4.371
Window Size	Fuzzy Weights MHE		
4	0.662	1.172	1.937
5	0.664	1.178	15.581
6	0.664	2.417	16.365

Table 5. Root Mean Squared Error (RMSE) for xylose and glucose predictions for Moving Horizon Estimator (MHE, window size of four elements), as well as pure model and soft sensing.

Prediction Source	Training Data Sets		Test Data Set	
	Glu (g L^{-1})	Xyl (g L^{-1})	Glu (g L^{-1})	Xyl (g L^{-1})
Model Prediction	1.67	0.90	4.16	1.73
Soft Sensing	2.67	1.69	11.84	3.92
Fixed Weights MHE	1.52	0.64	4.90	1.41
Fuzzy Weights MHE	1.28	0.60	1.85	1.49

3.2. Moving Horizon Tuning

The soft sensing predictions are then used in the MHE algorithm. In a first attempt to use MHE, L_w was a diagonal matrix that was estimated by the model prediction RMSE obtained for the Training Data Sets (Table 5), which results in the elements of 0.359 for glucose and 1.23 for xylose. On the other hand, L_v was a diagonal matrix that was estimated by the Soft Sensing RMSE (Table 5), resulting in elements of 0.140 for glucose and 0.350 for xylose. Although the variance of the soft sensor error is greater than the variance of the model, the MHE still gives too much weight to the soft sensor. The MHE results practically follow soft sensor prediction (data not shown). Hence, no significant improvement was obtained in the RMSE when compared to the use of soft sensor alone.

The next approach was to find the optimal fixed and dynamic weights, L_w and L_v (Sections 2.3.1 and 2.3.2), to predict the training data set. A summary of the tuning algorithms' Root Mean Squared Errors (RMSE) is presented in Table 4, for both the fixed weights and fuzzy weights and for all of the window sizes. Table 5 shows the results separated by component.

From Table 4, it is clear that using the smallest window of prediction (four elements) yields the best results for the MHE.

The fixed weights MHE with four elements in the window employed optimized weights of 32.06 for glucose and 78.72 for xylose in L_w, and 3.42 for glucose and 2.65 for xylose in the L_v matrix. Effectively, this means that, in this implementation, the knowledge from the model was weighted more than the one from the instrumentation. This can be observed in the training data set predictions, as presented in Figure 7.
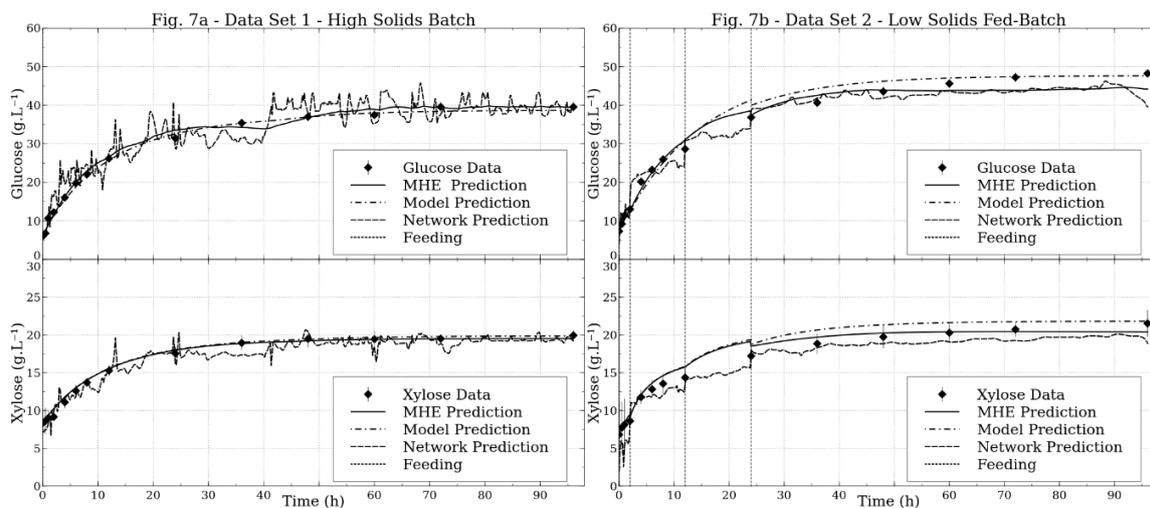

Figure 7. Fixed weights Moving Horizon Estimator (MHE) with four elements in moving window carbohydrates concentration prediction. (**a**) High Solid Batch (HSB) predictions; and, (**b**) Low Solids Fed-batch (LSB) predictions; top and bottom graphs are glucose and xylose predictions; error bars in data points are standard errors between experiment duplicates for compound concentration; vertical dotted lines are the biomass feeding times.

Figure 7 illustrates the results that are presented in Table 4. A slight decrease in prediction performance (increase of approximately 11% in the RMSE in relation to the model) occurs in the HSB data set to improve prediction in the LSF data set. However, as depicted in Table 5, when compared to the pure model prediction, MHE reduced the errors for both products when HSB and LSF are treated together. Thus, the overall decrease in the RMSE when using the MHE algorithm was of approximately 17% across assays, showing that the use of the agitation measurements may improve the prediction. However, an issue occurs when the prediction of the test data set (MSF) is considered, as presented in Figure 8.
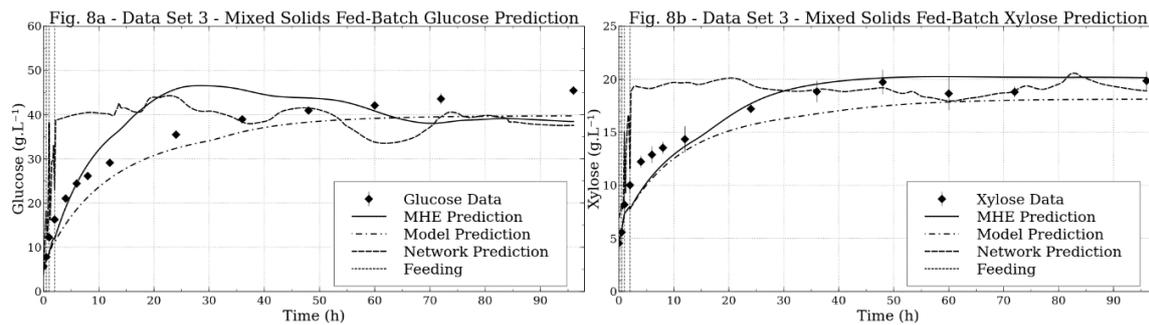
Figure 8. Fixed weights Moving Horizon Estimator (MHE) with four elements in moving window carbohydrates concentration predictions for Mixed Solids Fed-batch (MSF); top and bottom graphs are glucose and xylose predictions; error bars in data points are standard errors between experiment duplicates for compound concentration; vertical dotted lines are the biomass feeding times.

In the test data set, the prediction performance is less than ideal. The fixed weights that are tuned in other data sets are not capable of sustaining the MHE performance in this situation. Although an initial prediction for glucose concentration is improved, it reaches concentrations that are far above the actual value. Additionally, glucose concentration is underestimated due to a decrease in the network prediction towards the end of the reaction. The algorithm is more accurate for xylose, as the final concentration is predicted to be better than when using the pure model (see Table 5). However, the MHE with fixed weights does little to improve the initial concentration predictions. Nevertheless, this behavior does not occur when using the fuzzy weights implementation.

When analyzing the MHE with fuzzy weight results in Tables 4 and 5, it is clear that the dynamic weight approach greatly improves the state prediction within the training data, even when considering the HSB model prediction, especially for the smallest windows, for which the best results are obtained. In this implementation, the value for the weights that are optimized by the tuning algorithm are L_w of 69.01 for glucose and 126.90 for xylose, as well as L_v of 1.08 for glucose and 0.54 for xylose. This followed the same trend of giving higher values to the weights that multiply the model prediction, as this clearly provides a more accurate model prediction.

The fuzzy rule thresholds, for the four-element moving window, are 10.77, 63.63, 80.06, and 190.23 g L^{-1} for LL, LU, HL, and HU, respectively. These values mean that above 190 g L^{-1} of solids within the reactor, only the model prediction is used and below 10 g L^{-1} only the soft sensing is used. However, these situations do not occur within the data set, meaning that the algorithm is always using both predictions to compose the new state estimation, while using more the model when solid concentration is high, and more the network when solids are low. The results for the tuned fuzzy weights MHE with four elements in the prediction windows for the training data sets are presented in Figure 9.

The behavior that is expected from the fuzzy rule is clearly seen in Figure 9. In the HSB, the prediction with the MHE closely follows the model prediction. The use of this methodology also represents a small improvement when compared to using only the model in this data set (1.63%). A more interesting trend is observed in the LSF. During low solids concentrations (before the second addition of biomass at 24 h), the MHE prediction closely follows the network prediction. After this point, an interpolation between both predictions prevails. This greatly improves prediction, decreasing the RMSE in approximately 30.8% when compared with the model results alone.

Yet, the great strength of this version of the MHE is observed when analyzing the test data set prediction, as presented in Figure 10.
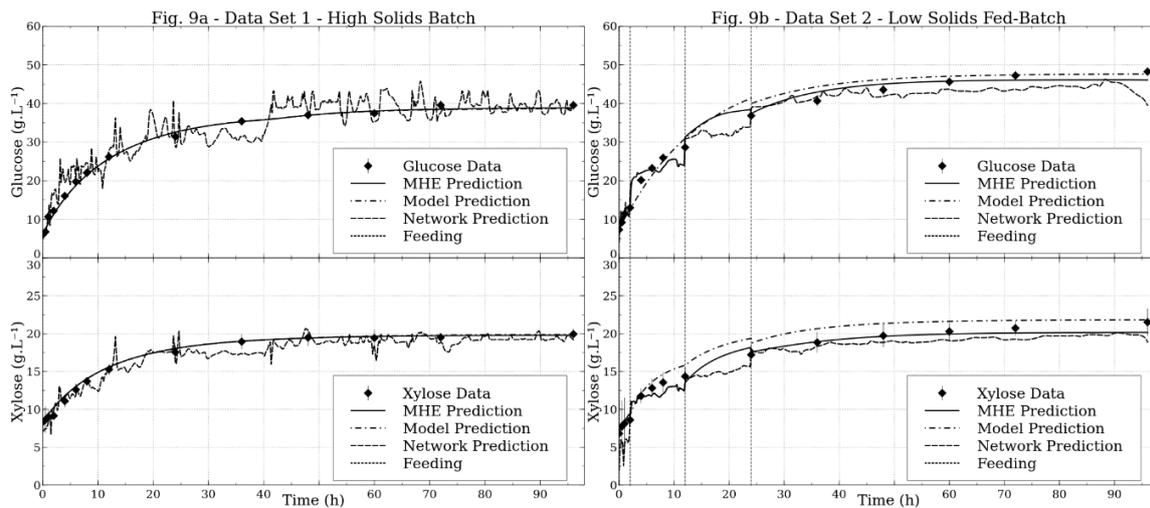
Figure 9. Fuzzy weights Moving Horizon Estimator (MHE) with 4 elements in moving window carbohydrates concentration predictions. (**a**) High Solid Batch (HSB) predictions; and, (**b**) Low Solids Fed-batch (LSB) predictions; top and bottom graphs are glucose and xylose predictions; error bars in data points are standard errors between experiment duplicates for compound concentration; vertical dotted lines are the biomass feeding times.

Figure 10. Fuzzy weights Moving Horizon Estimator (MHE) with four elements in moving window: carbohydrate concentration predictions for Mixed Solids Fed-batch (MSF); left and right graphs are glucose and xylose predictions; error bars in data points are standard errors between experiment duplicates for compound concentrations; vertical dotted lines are the biomass feeding times.
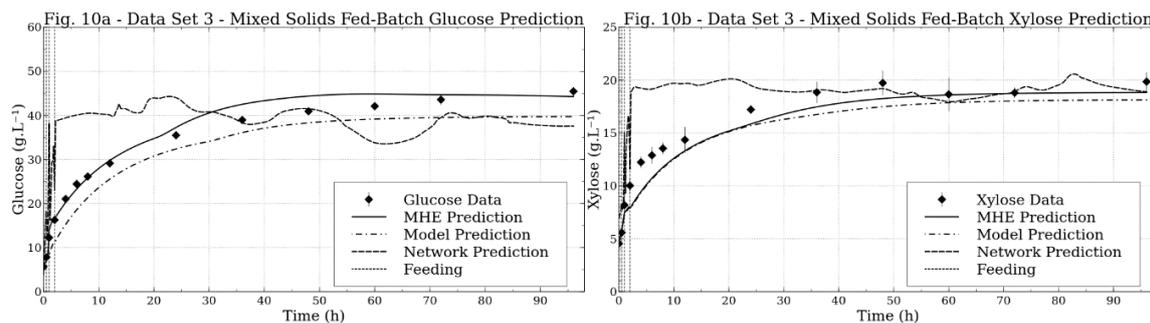
In the test data set, a decrease of 36.8% in the overall RMSE was obtained while using Fuzzy MHE, when comparing to the model prediction alone. The error was reduced for both Xylose and Glucose when compared to the pure model prediction (Table 5). Xylose prediction improved more towards the end of the process. Glucose prediction is improved throughout the entire reaction, especially at the beginning, where the interpolation between model and soft sensing predictions generated a more accurate estimation.

Therefore, it is clear that using the proposed approach not only improves the prediction within training data, but it also improves prediction in previously unforeseen circumstances. This helps to enhance the robustness of the state estimation and it can greatly improve the control of the process if a controller is used alongside the proposed MHE implementation.

4. Conclusions

In this study, a LOLIMOT-based soft sensor using torque measurements was successfully built to monitor the glucose and xylose profiles during batch and fed-batch enzymatic hydrolysis of sugarcane bagasse. The developed soft sensor did not present good performance when considering the extrapolation capabilities. The use of MHE together with the soft sensing was successful in improving

the prediction of desired state variables, regardless of the weights used within the implementation. In particular, the novel fuzzy dynamic weights for MHE greatly improved the prediction in both training data sets (16.2% RMSE reduction when compared to the model prediction alone) and test data sets (36.8% RMSE reduction when compared to the model prediction alone). The approach proposed here might be a valuable tool for monitoring the hydrolysis of lignocellulosic material, or any other type of process that has online agitation measurements, especially when the state estimation is needed in a feedback control system.

Author Contributions: Conceptualization, Methodology, Writing, Review & Editing, V.B.F., L.J.C., F.V.L., R.C.G., M.P.A.R.; Investigation, V.B.F., L.J.C.; Formal analysis, Visualization, V.B.F., F.V.L., R.C.G., M.P.A.R.; Resources, Supervision, F.V.L., R.C.G., M.P.A.R.; Data curation, Project administration, Software, Writing—Original draft, V.B.F. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Variables

e	Residuals Vector
F	Objective Function Optimum Value
HL	High Solids Fuzzy Rule Lower Bound
HL	High Solids Fuzzy Rule Upper Bound
k_i	Kinetic constants (min^{-1})
ke	First order inactivation constant (min^{-1})
Km_i	Michaelis-Menten constant (g L^{-1})
Kp_i	Products competitive inhibition constant (g L^{-1})
LL	Low Solids Fuzzy Rule Lower Bound
L_v	Instrumentation Prediction Weight Vector
L_w	Model Prediction Weight Vector
LH	Low Solids Fuzzy Rule Upper Bound
MD$^{\text{Inst}}$	Membership Degree of the Instrumentation Weights
MD$^{\text{HSM}}$	Membership Degree of the High Solids Model
N	Window Size
Q	Cost Function Weights Matrix
T	Current Estimation Time
$\hat{X}$	Moving Horizon Estimator Prediction
X^-	Model Prediction
Y	Instrumentation Data
$\hat{Y}$	Moving Horizon Estimator Prediction in Instrumentation Unit
Cb	Cellobiose Concentration (g L^{-1})
Cl	Cellulose Concentration (g L^{-1})
E_i	Enzyme Concentration (g L^{-1})
Gl	Glucose Concentration (g L^{-1})
He	Hemicellulose Concentration (g L^{-1})
Lg	Lignin Concentration (g L^{-1})
P_i	Product concentration (g L^{-1})
S_i	Substrate Concentration (g L^{-1})
Xy	Xylose Concentration (g L^{-1})

Greek Letters

α^{FUZZY}	Fuzzy Model reaction rate (g L^{-1} min^{-1})
α^{HSM}	High Solids Model reaction rate (g L^{-1} min^{-1})
α_i	Reaction rate for "i" reaction, where "i" are reactions 1 through 6 (g L^{-1} min^{-1})
α^{LSM}	Low Solids Model reaction rate (g L^{-1} min^{-1})

$\gamma_{\text{Cl-Cb}}$	Pseudo-stoichiometric mass relation between cellulose and cellobiose ($g_{\text{Cellobiose}} \cdot g_{\text{Cellulose}}^{-1}$)
$\gamma_{\text{Cb-Gl}}$	Pseudo-stoichiometric mass relation between cellobiose and glucose ($g_{\text{Glucose}} \cdot g_{\text{Cellobiose}}^{-1}$)
$\gamma_{\text{He-Xy}}$	Pseudo-stoichiometric mass relation between Hemicellulose and Xylose $g_{\text{Xylose}} \cdot g_{\text{Hemicellulose}}^{-1}$

Abbreviations

ANN	Artificial Neural Network
FM	Fuzzy Model
HSB	High Solids Batch
HSM	High Solids Model
LLM	Local Linear Model
LLNFM	Local Linear Neuro-Fuzzy Models
LOLIMOT	Local Linear Model Trees
LOWESS	Locally Weighted Scatterplot Smoothing
LSF	Low Solids Fed-batch
LSM	Low Solids Model
MHE	Moving Horizon Estimator
MSE	Mean Squared Error
MPF	Mixed Profile Fed-batch
ODE	Ordinary Differential Equation Solver
RMSE	Root Mean Squared Error
RS	Reactive Solids

References

1. Zhang, J.; Chu, D.; Huang, J.; Yu, Z.; Dai, G.; Bao, J. Simultaneous saccharification and ethanol fermentation at high corn stover solids loading in a helical stirring bioreactor. *Biotechnol. Bioeng.* **2010**, *105*, 718–728. [CrossRef] [PubMed]
2. Correâ, L.J.; Badino, A.C.; Cruz, A.J.G. Mixing design for enzymatic hydrolysis of sugarcane bagasse: Methodology for selection of impeller configuration. *Bioprocess Biosyst. Eng.* **2016**, *39*, 285–294. [CrossRef] [PubMed]
3. Bondancia, T.J.; Corrêa, L.J.; Cruz, A.J.G.; Badino, A.C.; Mattoso, L.H.C.; Marconcini, J.M.; Farinas, C.S. Enzymatic production of cellulose nanofibers and sugars in a stirred-tank reactor: Determination of impeller speed, power consumption, and rheological behavior. *Cellulose* **2018**, *25*, 4499–4511. [CrossRef]
4. Palmqvist, B.; Lidén, G. Torque measurements reveal large process differences between materials during high solid enzymatic hydrolysis of pretreated lignocellulose. *Biotechnol. Biofuels* **2012**, *5*, 1–9. [CrossRef]
5. Nguyen, T.C.; Anne-Archard, D.; Coma, V.; Cameleyre, X.; Lombard, E.; Binet, C.; Nouhen, A.; To, K.A.; Fillaudeau, L. In situ rheometry of concentrated cellulose fibre suspensions and relationships with enzymatic hydrolysis. *Bioresour. Technol.* **2013**, *133*, 563–572. [CrossRef]
6. Sotaniemi, V.H.; Taskila, S.; Ojamo, H.; Tanskanen, J. Controlled feeding of lignocellulosic substrate enhances the performance of fed-batch enzymatic hydrolysis in a stirred tank reactor. *Biomass Bioenergy* **2016**, *91*, 271–277. [CrossRef]
7. Samaniuk, J.R.; Scott, C.T.; Root, T.W.; Klingenberg, D.J. Rheological modification of corn stover biomass at high solids concentrations solids concentrations. *J. Rheol.* **2012**, *56*, 649–665. [CrossRef]
8. Du, J.; Zhang, F.; Li, Y.; Zhang, H.; Liang, J.; Zheng, H.; Huang, H. Enzymatic liquefaction and saccharification of pretreated corn stover at high-solids concentrations in a horizontal rotating bioreactor. *Bioprocess Biosyst. Eng.* **2014**, *37*, 173–181. [CrossRef]
9. Kadhum, H.J.; Mahapatra, D.M.; Murthy, G.S. A novel method for real-time estimation of insoluble solids and glucose concentrations during enzymatic hydrolysis of biomass. *Bioresour. Technol.* **2019**, *275*, 328–337. [CrossRef]
10. Sagmeister, P.; Wechselberger, P.; Jazini, M.; Meitz, A.; Langemann, T.; Herwig, C. Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development. *Chem. Eng. Sci.* **2013**, *96*, 190–198. [CrossRef]

11. Kadam, K.L.; Rydholm, E.C.; McMillan, J.D. Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnol. Prog.* **2004**, *20*, 698–705. [CrossRef] [PubMed]

12. Furlong, V.B.; Pereira Filho, R.D.; Margarites, A.C.; Goularte, P.G.; Costa, J.A.V. Estimating microalgae Synechococcus nidulans daily biomass concentration using neuro-fuzzy network. *Food Sci. Technol.* **2013**, *33*, 142–147. [CrossRef]

13. Mohd Ali, J.; Hussain, M.A.; Tade, M.O.; Zhang, J. Artificial Intelligence techniques applied as estimator in chemical process systems—A literature survey. *Expert Syst. Appl.* **2015**, *42*, 5915–5931. [CrossRef]

14. Haseltine, E.L.; Rawlings, J.B. Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Ind. Eng. Chem. Res.* **2005**, *44*, 2451–2460. [CrossRef]

15. Rincón, F.D.; Le Roux, G.A.C.; Lima, F.V. The autocovariance least-squares method for batch processes: Application to experimental chemical systems. *Ind. Eng. Chem. Res.* **2014**, *53*, 18005–18015. [CrossRef]

16. Rawlings, J.B.; Ji, L. Optimization-based state estimation: Current status and some new results. *J. Process Control* **2012**, *22*, 1439–1444. [CrossRef]

17. Lima, F.V.; Rawlings, J.B. Nonlinear stochastic modeling to improve state estimation in process monitoring and control. *AIChE J.* **2011**, *57*, 996–1007. [CrossRef]

18. Valdés-González, H.; Flaus, J.M.; Acuña, G. Moving horizon state estimation with global convergence using interval techniques: Application to biotechnological processes. *J. Process Control* **2003**, *13*, 325–336. [CrossRef]

19. Campani, G.; Ribeiro, M.P.A.; Zangirolami, T.C.; Lima, F.V. A hierarchical state estimation and control framework for monitoring and dissolved oxygen regulation in bioprocesses. *Bioprocess Biosyst. Eng.* **2019**, *42*, 1467–1481. [CrossRef]

20. Vercammen, D.; Logist, F.; Impe, J. Van Online moving horizon estimation of fluxes in metabolic reaction networks. *J. Process Control* **2016**, *37*, 1–20. [CrossRef]

21. Abdollahi, J.; Dubljevic, S. Lipid production optimization and optimal control of heterotrophic microalgae fed-batch bioreactor. *Chem. Eng. Sci.* **2012**, *84*, 619–627. [CrossRef]

22. Furlong, V.B.; Corrêa, L.J.; Giordano, R.C.; Ribeiro, M.P.A. Fuzzy-enhanced modeling of lignocellulosic biomass enzymatic saccharification. *Energies* **2019**, *12*, 2110. [CrossRef]

23. Sluiter, A.; Hames, B.; Ruiz, R.O.; Scarlata, C.; Sluiter, J.; Templeton, D.; Energy, D. of Determination of Structural Carbohydrates and Lignin in Biomass. Laboratory ANalytical Procedure (LAP). *Biomass Anal. Technol. Team Lab. Anal. Proced.* **2004**, *2011*, 1–14.

24. Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D. *Determination of Sugars, Byproducts, and Degradation Products in Liquid Fraction Process Samples*; Technical Report NREL/TP-510-42623; National Renewable Energy Laboratory: Golden State, CO, USA, 2008; pp. 1–14.

25. Yao, M.; Wang, Z.; Wu, Z.; Qi, H. Evaluating kinetics of enzymatic saccharification of lignocellulose by fractal kinetic analysis. *Biotechnol. Bioprocess Eng.* **2011**, *16*, 1240–1247. [CrossRef]

26. Bastin, G.; Dochain, D. *On-Line Estimation and Adaptive Control of Bioreactors*; Elsevier: Amsterdam, The Netherlands, 1991; Volume 243, ISBN 0444884300.

27. Nelles, O. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*; Springer: Berlin/Heidelberg, Germany, 2001; ISBN 3540673695.

28. Nelles, O.; Fink, A.; Isermann, R. Local Linear Model Trees (LOLIMOT) Toolbox for Nonlinear System Identification. *IFAC Proc. Vol.* **2000**, *33*, 845–850. [CrossRef]

29. Rawlings, J.B.; Lima, F.V. *State Estimation of Linear and Nonlinear Dynamic Systems. Part IV: Nonlinear Systems: Moving Horizon Estimation (MHE) and Particle Filtering (PF)*; AICES Regional School, RWTH Aachen: Aachen, Germany, 2008; Available online: https://fernandolima.faculty.wvu.edu/teaching-outreach (accessed on 28 January 2020).