CHEMICAL ENGINEERING TRANSACTIONS
VOL. 43, 2015
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Chief Editors: Sauro Pierucci, Jiří J. Klemeš

ISBN 978-88-95608-34-1; **ISSN** 2283-9216
DOI: 10.3303/CET1543088

Neural Network Modeling to Predict the Total Glucose Yield after Enzymatic Saccharification of H_2SO_4-Catalyzed Hydrothermally Pretreated Sugarcane Bagasse

Elmer Ccopa Rivera*[a], Laura Plazas Tovar[b], Maria Regina Wolf Maciel[b], Antonio Bonomi[a,b], Rubens Maciel Filho[a,b]

[a] Brazilian Bioethanol Science and Technology Laboratory, Brazilian Center for Research in Energy and Materials. P.O. Box 6170, Zipcode 13083–970, Campinas, SP, Brazil
[b] School of Chemical Engineering, University of Campinas, UNICAMP, Zipcode 13083–852, Campinas–SP, Brazil
elmer.rivera@bioetanol.org.br

In this work, sugarcane bagasse was processed by H_2SO_4-catalyzed hydrothermal pretreatment (CHP). CHP was performed under various solid loadings (S^P) 10, 20 and 25 %, acid concentrations (A) 1.0, 2.0 and 3.0 % w/v and reaction times (t^P) 30, 60, 90 and 150 min. The results showed that more than 60 % of the hemicellulose was removed. Cellulignin material was processed by enzymatic hydrolysis using a solid loading of 3.0 % w/v (on a dry basis) and by the action of cellulase from *T. reesei* and cellobiase from Aspergillus niger. Based on this experimental dataset, three Multilayer Perceptron Neural Network-MLPN models were developed (one for each level of acid concentration (A) in CHP) for predicting total glucose yield (G_{Yield}) as a function of solids loading (S^P), pre-treatment reaction time (t^P) and enzymatic hydrolysis time (t). The prediction by the model using the test dataset gave acceptable performance measures (mean absolute error, MAE, and residual standard deviation, RSD), equivalent to those obtained for the validation and training datasets. The dynamic models developed can be used to predict G_{Yield} during enzymatic hydrolysis of H2SO4-catalyzed hydrothermally (CHP) pre-treated sugarcane bagasse, which promote a successful hydrolysis process control.

1. Introduction

Lignocellulosic biomass intended for production of biofuels may be pretreated in an acid thermochemical process stage (i.e. H_2SO_4-catalyzed hydrothermal pretreatment carried out in either concentrated/diluted acid system) in order to improve the susceptibility of the cellulose to enzymatic hydrolysis (Tovar et al., 2014). The process produces a solid material named water-insoluble solid (WIS), which is characterized for presenting amorphous and crystalline regions. WIS is constituted mainly by cellulose plus residual hemicellulosic and lignin fractions, which is susceptible of conversion into fermentable sugars through enzymatic hydrolysis (Canilha et al., 2012).

In recent years, advances have been made in cellulose and/or hemicellulose hydrolysis modeling, which can give a support to provide cost-efficient second-generation processes. Difficulties on modeling the enzymatic hydrolysis of lignocellulosic materials are related to the substrate complexity, mechanism of enzymes action, and enzymes-substrate interaction (Carvalho et al., 2013). Thus, related to the substrate complexity, it is important to include in the model, the effect of pretreatment operating conditions such as: pretreatment time, chemical agent, solid content and temperature used in the pretreatment process. Regarding enzymatic hydrolysis, several models have been reported, which can be grouped into: empirical models, quantifying the effects of various substrate and enzyme properties on hydrolysis, Michaelis - Menten based models, adsorption in cellulose hydrolysis model, using Langmuir adsorption isotherm or kinetic equations and kinetic models for the hydrolysis of soluble cello-oligosaccharides (Bansal et al., 2009). Thus, it is difficult to define a

Please cite this article as: Ccopa Rivera E., Plazas Tovar L., Wolf Maciel M.R., Bonomi A., Maciel Filho R., 2015, Neural network modeling to predict the total glucose yield after enzymatic saccharification of h2so4-catalyzed hydrothermally pretreated sugarcane bagasse, Chemical Engineering Transactions, 43, 523-528 DOI: 10.3303/CET1543088

mathematical model that considers the operational conditions of both pretreatment and enzymatic hydrolysis as well as determine its kinetic parameters, mainly when deterministic mathematical model approach is used. Artificial Neural Networks (ANN) has been used successfully for solving biotechnological complex problems for modeling and optimization. Among them, ANN models to describe enzymatic kinetic of cellulose hydrolysis (Rivera et al., 2010). Another study reports the use of ANN models to describe substrate and product inhibition during enzymatic hydrolysis of cellobiose (Corazza et al., 2005). In this work, three Multilayer Perceptron Neural Network-MLPN models (one for each level of acid concentration (*A*) in CHP) for predicting total glucose yield (G_{Yield}) as a function of solids loading (S^P), pre-treatment reaction time (t^P) and enzymatic hydrolysis time (*t*) were developed. This approach aimed to promote a more effective employment of operating conditions during the enzymatic hydrolysis of sugarcane bagasse.

2. Materials and methods

2.1 Raw material

Sugarcane bagasse (SCB) was supplied from the sugar-alcohol-co-generation mill São João, Araras, São Paulo - Brazil. It was air-dried to final moisture content of about 5.0% of dry weight, homogenized and stored in plastic zipper bags. The chemical composition of SCB (on a dry basis, g/100g_{SCB}) was: cellulose 41.81 ± 3.31, hemicellulose 31.00 ± 2.54, lignin 17.08 ± 2.19, extractives 8.93 ± 1.16 and ash 2.53 ± 0.13.

2.2 H_2SO_4-catalyzed Hydrothermal Pre-treatment

Each experiment was submitted to an isothermal hold step at 40 ± 1 °C (during 15 min) in a water bath to guarantee the same temperature distributions in the glass flask. In turn, glass flasks were loaded in a vertical autoclave. The initial reaction time was defined when the programmed temperature was reached (121 ± 1 °C). The residence times for the reaction system were 30, 60, 90 and 150 min. After that, the vertical autoclave was gradually depressurized to atmospheric pressure, and 99 ± 1 °C, and the glass flasks were discharged in sequence. The solid fraction was washed over a 400 mesh wire sieve with warm water at 60 °C to raise the pH to ~6 and prepared to enzymatic digestibility. A portion of WIS fraction was used for subsequent chemical analyses based on standard procedures (Sluiter et al., 2008, Gouveia et al., 2009).

2.3 Enzymatic digestibility

The hydrolysis of WIS fraction was carried out in 100 mL of 50 mM sodium acetate buffer (pH 4.8) at 3.0%, consistency. Cellulase, from *Trichoderma reesei* ATCC 26921 (from Sigma-Aldrich, Brazil) loading was15.0 FPU cellulase/$g_{substrate\ (WIS)}$ and the Cellobiase from Aspergillus niger Novozyme 188 (from Sigma-Aldrich, Brazil) loading was 30.0 IU β-glucosidase/$g_{substrate\ (WIS)}$. Aliquots were taken from the reaction mixture (after 0, 1, 3, 6, 12, 24, 36, 48, 60 and 72 h) and the samples were boiled during 15 min to deactivate the enzymes. The samples were centrifuged to remove the residual insoluble materials. The sugar content was measured by HPLC according to NREL standard method (Sluiter et al., 2006). The hydrolysis yield of the WIS fraction was calculated using Eq(1) and (2).

$$Digestibility = \frac{total\ glucose\ in\ enzymatic\ hydrolysis \times 0.9}{Total\ glucan\ in\ WIS} \times 100 \quad (1)$$

$$G_{Yield} = \frac{total\ glucose\ in\ enzymatic\ hydrolysis \times 0.9}{Total\ glucan\ in\ SCB} \times 100 \quad (2)$$

2.4 MLPN model development

The main components of a MLPN are its architecture (model structure) and the training procedure. The MLPN used in this work, consisted of three types of layers: an input layer, an output layer and one or more hidden layers, whose numbers of neurons are *N*, *M* and *K*. Each layer may have a different number of neurons, which are interconnected by adjustable parameters associated with them Eq(3).

$$g_k = F\left(\sum_{j=1}^{M} W_{kj} f\left(\sum_{i=1}^{N} w_{ji} x_i + \theta_j\right) + b_k\right), \quad (j = 1,...,M);\ (k = 1,...,K) \tag{3}$$

where w_{ji} is the weight connecting the ith neuron in the input layer and the jth neuron in hidden layer; θ_j is the bias of the jth neuron in the hidden layer; W_{kj} is the weight connecting the jth neuron in the hidden layer and the kth neuron in the output layer. b_k is the bias in the kth neuron in the output layer. $f(\cdot)$ and $F(\cdot)$ are the activation functions of the jth neuron in the hidden layer and of the kth neuron in the output layer, respectively.

Selection of the input variables and training

In this work, after several tests, it was proposed the development of three MLPN models, one for each level of acid concentration (A) in CHP for predicting total glucose yield (G_{Yield}) as a function of solids loading (S^P), pre-treatment reaction time (t^P) and enzymatic hydrolysis time (t). In order to include the kinetic behavior in the modeling, the MLPN training was processed using G_{Yield} profiles. For that, an interpolation procedure was used to fit the experimental data of G_{Yield} with the purpose of increasing the dataset. Thus, each trial used for training has provided 49 interpolated data points (from 10 observed points). Tests have shown that this procedure led to good results when dynamic processes were modeled using MLPN (Rivera et al., 2010). A representative training dataset containing 353 input/output samples was presented in a randomized sequence to each MLPN model for the estimation of its parameters. The predictive capability of neural networks or validation was assessed on a different sequence of experimental observations containing 176 samples. Finally, three datasets, one for each acid concentration in CHP, were used to test the models. The standard back propagation learning algorithm, based on a gradient descendent method implemented in FORTRAN is employed to train each network describing G_{Yield}. Training was stopped after 10,000 epochs. The appropriate number of neurons in the hidden layer was found by the cross-validation technique in order to avoid model over-fitting and to achieve good generalization from the training dataset.

3. Results and discussion

3.1 Effects of pretreatment conditions

Figure 1 presents the cellulose, hemicellulose, lignin and ash mass yields based on solids recovery of WIS after H_2SO_4-catalyzed hydrothermal pretreatment of SCB under different conditions of A, S^P and t^P. The cellulose mass yield in the WIS fractions showed that a portion of the cellulose was dissolved. According to Rocha et al. (2011), it is assumed that it corresponded to the amorphous or low-crystallinity fraction. At the different reaction times, the amount of insoluble hemicellulose in the WIS fractions was shown to decrease as the acid concentration rose from 1.0% w/v to 3.0% w/v, resulting in a lower percent of soluble hemicellulose remaining in the WIS samples (lower than 12 g/100g_{SCB}). This behavior was due to degradation of the xylose during the H_2SO_4-catalyzed hydrothermal pretreatment with the higher acid concentration, forming degradation products such as furfural.

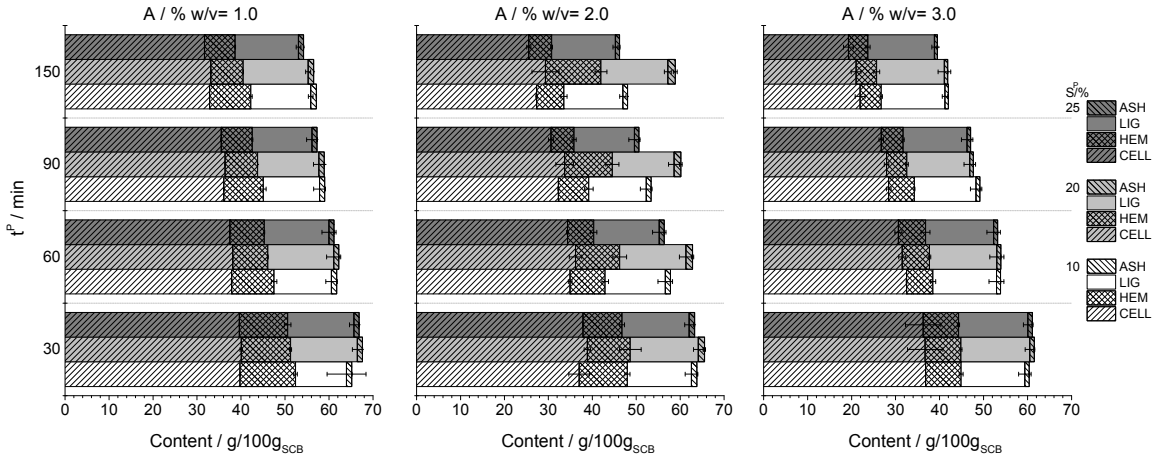

Figure 1: Cellulose (CELL), hemicellulose (HEM), lignin (LIG) and ash mass yields based on solids recovery of WIS after H_2SO_4-catalyzed hydrothermal pretreatment of SCB under different conditions of A, S^P and t^P. The data represent the means and standard deviations of experiments carried out in triplicate.

3.2 Effect of high-solid loading used in the pre-treatment process on enzymatic digestibility

Previous studies have shown that high-solids loadings (>10 %) improved the efficiency of the enzymatic hydrolysis process (Roche et al., 2009). However, in the present study it was not possible to confirm this assumption since the efficiency of the pretreatment coupled to enzymatic hydrolysis was dependent on the combined effect of the pretreatment conditions such as: reaction time, solids loading and acid concentration, as well as the effect of the heterogeneous reaction catalyzed by cellulase.
In order to take these effects into account, Chen et al. (2012) studied the biomass acid loading (grams of acid/grams of biomass) and found that this was the more dominant factor in determining biomass digestibility. In this work, WIS fraction digestibility obtained with a solids loading of 20%, acid concentration of 1.0% w/v and H_2SO_4 catalyzed hydrothermal pretreatment time of 150 min at 121 ± 1°C was investigated. It was found that the increases in grams of acid per gram of sugarcane bagasse (0.05; 0.10 and 0.15) associated with the acid concentration used in the pretreatment (1.0, 2.0 and 3.0% w/v) significantly affected the cellulose in WIS, reporting losses of 20.68 ± 0.06 %, 33.94 ± 2.48 % and 49.74 ± 2.68 %.

Table 1: Effect of acid concentration (A / % w/v), solids loading (S^P / %) and reaction time (t^P / min) on glucan digestibility, total glucose yield (G_{Yield}) after 72 h of enzymatic hydrolysis. The data represent the means and standard deviations of experiments (enzymatic hydrolysis, EH) carried out in duplicate

H_2SO_4 acid-catalyzed hydrothermal pretreatment conditions				Glucan digestibility after 72 h of EH	
A / % w/v	S^P / %	$g_{acid}/g_{dry\ matter}$	t^P / min	Digestibility / %	G_{Yield} / %
Control (SCB)				12.97 ± 1.14	12.97 ± 1.13
1.0	10	0.10	30	46.49 ± 0.37	44.24 ± 0.35
1.0	20	0.05	30	50.08 ± 0.18	47.90 ± 0.17
1.0	25	0.04	30	47.20 ± 4.70	44.68 ± 4.44
1.0	10	0.10	60	60.40 ± 0.39	54.75 ± 0.35
1.0	20	0.05	60	51.55 ± 4.62	47.07 ± 4.22
1.0	25	0.04	60	42.13 ± 0.58	37.81 ± 0.52
1.0	10	0.10	90	55.99 ± 0.66	48.39 ± 0.57
1.0	20	0.05	90	53.38 ± 3.41	46.54 ± 2.97
1.0	25	0.04	90	44.90 ± 3.03	38.17 ± 2.57
1.0	10	0.10	150	72.87 ± 0.19	57.38 ± 0.15
1.0	20	0.05	150	58.75 ± 0.98	46.71 ± 0.78
1.0	25	0.04	150	63.32 ± 1.88	48.12 ± 1.43
2.0	10	0.20	30	72.04 ± 0.55	63.75 ± 0.49
2.0	20	0.10	30	58.56 ± 0.21	54.63 ± 0.20
2.0	25	0.08	30	48.57 ± 4.81	44.08 ± 4.37
2.0	10	0.20	60	57.02 ± 0.37	47.71 ± 0.31
2.0	20	0.10	60	63.84 ± 5.67	55.49 ± 4.93
2.0	25	0.08	60	61.24 ± 0.81	50.41 ± 0.67
2.0	10	0.20	90	50.42 ± 0.60	38.98 ± 0.46
2.0	20	0.10	90	70.56 ± 4.45	57.13 ± 3.61
2.0	25	0.08	90	47.91 ± 3.22	35.18 ± 2.36
2.0	10	0.20	150	44.63 ± 0.12	29.23 ± 0.08
2.0	20	0.10	150	56.15 ± 0.96	39.56 ± 0.68
2.0	25	0.08	150	58.23 ± 1.75	35.63 ± 1.07
3.0	10	0.30	30	50.62 ± 0.40	44.60 ± 0.35
3.0	20	0.15	30	57.18 ± 0.21	50.19 ± 0.18
3.0	25	0.12	30	44.51 ± 4.45	38.58 ± 3.86
3.0	10	0.30	60	56.88 ± 0.37	44.26 ± 0.29
3.0	20	0.15	60	51.35 ± 4.63	38.77 ± 3.49
3.0	25	0.12	60	57.37 ± 0.77	42.15 ± 0.57
3.0	10	0.30	90	66.09 ± 0.77	45.05 ± 0.53
3.0	20	0.15	90	59.04 ± 3.76	39.57 ± 2.52
3.0	25	0.12	90	72.04 ± 4.72	46.23 ± 3.03
3.0	10	0.30	150	74.00 ± 0.19	38.87 ± 0.10
3.0	20	0.15	150	69.55 ± 1.16	35.02 ± 0.59
3.0	25	0.12	150	79.82 ± 2.37	36.81 ± 1.09

The WIS digestibility from sugarcane bagasse pretreated at 121 ± 1°C with a solids loading of 20 % and 0.15 g of acid/g of sugarcane bagasse (69.55 ± 1.16 %) was higher than those pretreated with 0.10 and 0.05 g of acid/g of sugarcane bagasse (56.15 ± 0.96 % and 58.75 ± 0.98 %) as can be seen from Table 1. However, at higher acid loading (0.15 g of acid per gram of sugarcane bagasse) the G_{Yield} from enzymatic hydrolysis was lower (35.02 ± 0.59 %) than from those pretreated with 0.10 and 0.05 g of acid per g of sugarcane bagasse, which gave values of 39.56 ± 0.68% and 46.71 ± 0.78 %. The increase in digestibility from 0.05 to 0.15 g of acid per g of sugarcane bagasse was over 60 %. Nevertheless, this rise was governed by the decrease in G_{Yield} to below 35 %, clearly indicating that pretreatment using relatively high acid concentrations had a significant effect on the losses of cellulose from the WIS, which is the principal supplier of the release of C6 fermentable sugars for ethanol production. The G_{Yield} of the WIS fraction was highly dependent on the biomass acid loading (grams of biomass per gram of sulfuric acid). Thus, the combination of acid concentration 1.0 % w/v and solids loadings 20 – 25 % governed the effectiveness of the H_2SO_4-catalyzed hydrothermal pretreatment coupled to enzymatic hydrolysis to attain a G_{Yield} of nearly 50 %.

3.3 MLPN model prediction

The potential of different MLPN structures with a single hidden layer for estimation of G_{Yield} was assessed for all three models, one for each level of acid concentration (*A*) in CHP. The validation dataset was assembled to evaluate the performance of the trained MLPN structures. During the validation process, the number of hidden neurons was varied from two to twenty four, and the optimal number chosen by the cross-validation criterion with the number of epochs fixed at 10,000 for all the structures studied.

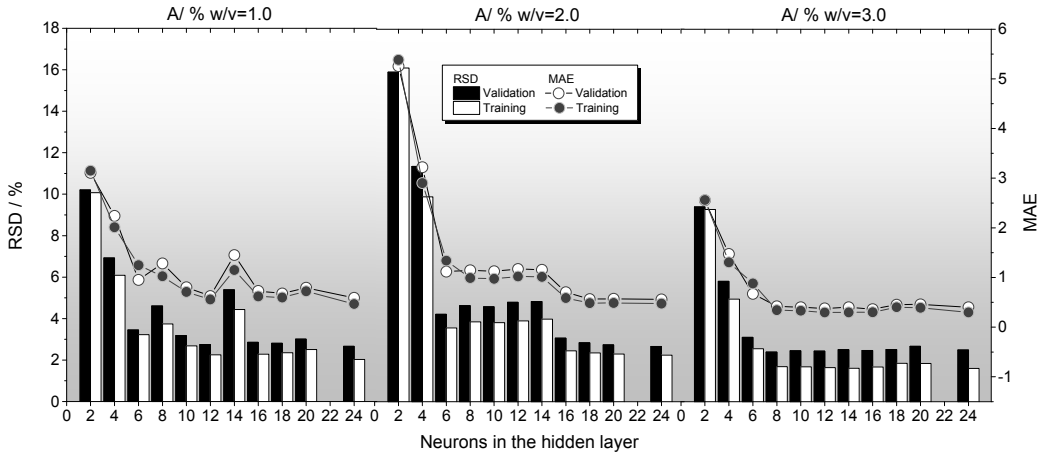

Figure 2: Cross-validation technique to achieve good generalization for the MLPN model, one for each level of acid concentration.

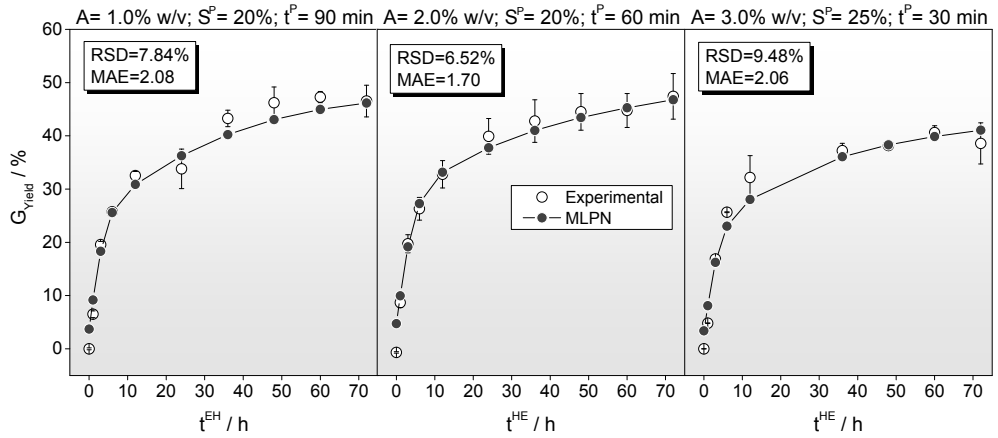

Figure 3: Experimental data (solid symbols) and performance of the MLPN models (solid symbols and lines) using the test datasets.

From Figure 2 it can be seen that the neural network with six hidden neurons for G_{Yield} prediction was found to give the lowest RSD and MAE for the validation dataset, before both criteria start to oscillate. The G_{Yield} prediction as a function of S^P, t^P and t was then modeled using three MLPN with 31 scalar parameters each. The performance of the optimal MLPN models in describing the G_{Yield} for the test datasets are depicted in Figure 3. The models described the experimental observations accurately, as judged by the RSD and MAE.

4. Conclusions

The complex relationship between the G_{Yield} and operational parameters of enzymatic hydrolysis of H_2SO_4-catalyzed hydrothermally (CHP) pretreated sugarcane bagasse is not easy to identify and difficult to model accurately by deterministic approach. However, the methodology developed in this work, based on MLPN, can convert the correlation of these variables into a predictive mathematical model, while maximizing the amount of experimental data that can be obtained from the process. Based on this methodology, the models performance in describing the dynamic behavior of G_{Yield} during hydrolysis was assessed. The prediction by the model using the test datasets gave acceptable performance, equivalent to those obtained for the training and validation datasets. These models can be used to facilitate the implementation of suitable operating strategies to achieve high operational performance of enzymatic hydrolysis.

Acknowledgements

The authors thank FAPESP (Grant N°. 2012/10857-3 and 2008/57873-8) for financial support.

References

Bansal, P., Hall, M., Realff, M. J., Lee, J. H., Bommarius, A. S. 2009. Modeling cellulase kinetics on lignocellulosic substrates. Biotechnology Advances, 27, 833-848.

Canilha, L., Chandel, A. K., Milessi, T. S. D. S., Antunes, F. A. F., Freitas, W. L. D. C., Felipe, M. D. G. A., Da Silva, S. S. 2012. Bioconversion of Sugarcane Biomass into Ethanol: An Overview about Composition, Pretreatment Methods, Detoxification of Hydrolysates, Enzymatic Saccharification, and Ethanol Fermentation. Journal of Biomedicine and Biotechnology, 2012, 15.

Carvalho, M. L., Sousa Jr., R., Rodríguez-Zúñiga, U. F., Suarez, C. A. G., Rodrigues, D. S., Giordano, R. C., Giordano, R. L. C. 2013. Kinetic study of the enzymatic hydrolysis of sugarcane bagasse. Brazilian Journal of Chemical Engineering, 30, 437-447.

Chen, Y., Stevens, M., Zhu, Y., Holmes, J., Moxley, G., Xu, H. 2012. Reducing acid in dilute acid pretreatment and the impact on enzymatic saccharification. Journal of Industrial Microbiology & Biotechnology, 39, 691-700.

Corazza, F. C., Calsavara, L. P. V., Moraes, F. F., Zanin, G. M., Neitzel, I. 2005. Determination of inhibition in the enzymatic hydrolysis of cellobiose using hybrid neural modeling. Brazilian Journal of Chemical Engineering, 22, 19-29.

Gouveia, E. R., Nascimento, R. T. D., Souto-Maior, A. M., Rocha, G. J. D. M. 2009. Validação de metodologia para a caracterização química de bagaço de cana-de-açúcar. Química Nova, 32, 1500-1503.

Rivera, E. C., Rabelo, S. C., Dos Reis Garcia, D., Filho, R. M., Da Costa, A. C. 2010. Enzymatic hydrolysis of sugarcane bagasse for bioethanol production: determining optimal enzyme loading using neural networks. Journal of Chemical Technology and Biotechnology, 85, 983-992.

Rocha, G. J. M., Martin, C., Soares, I. B., Souto Maior, A. M., Baudel, H. M., Moraes De Abreu, C. A. 2011. Dilute mixed-acid pretreatment of sugarcane bagasse for ethanol production. Biomass and Bioenergy, 35, 663-670.

Roche, C. M., Dibble, C. J., Knutsen, J. S., Stickel, J. J., Liberatore, M. W. 2009. Particle concentration and yield stress of biomass slurries during enzymatic hydrolysis at high-solids loadings. Biotechnology and Bioengineering, 104, 290-300.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. 2006. Determination of sugars, by products, and degradation products in liquid fraction samples - NREL/TP-510-42623. Laboratory Analytical Procedure (LAP). Revised January 2012 ed.: National Renewable Energy Laboratory.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., Crocker, D. 2008. Determination of structural carbohydrates and lignin in biomass. Laboratory Analytical Procedure (LAP). Revised August 2012 ed.: National Renewable Energy Laboratory.

Tovar, L. P., Maciel, M. R. W., Filho, R. M. 2014. Plug flow reactor model to analyse operating conditions on the dilute H2SO4-acid hydrolysis of sugarcane bagasse at high-solids loading. Chemical Engineering Transactions, 37, 325-300.