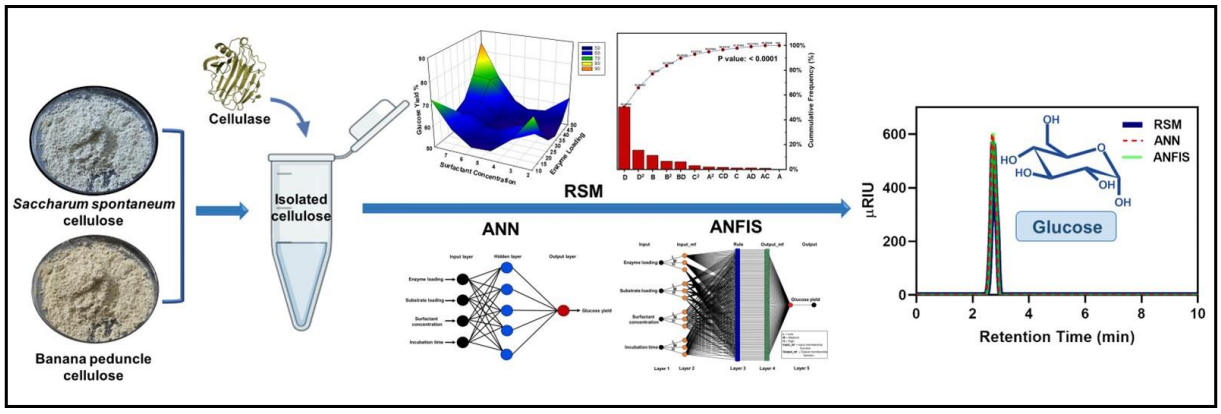
Chapter 4

Optimization of enzymatic hydrolysis of cellulose extracted from Saccharum spontaneum and banana peduncle using RSM, ANN, and ANFIS statistical tools

Abstract

Large-scale production of biomass-derived fuels and chemicals demands the economical and sustainable depolymerization of lignocellulosic biomass or their derived cellulose into sugars. Enzymatic hydrolysis of *Saccharum spontaneum* (*S. spontaneum*) and banana peduncle-derived celluloses for the production of glucose was performed using readily available commercial enzyme Cellulase (from *Aspergillus niger*) in a simple, effective, and ecologically sound way. The major emphasis of this study is the use of response surface methodology (RSM), artificial neural network (ANN), and adaptive neuro-fuzzy inference system (ANFIS) techniques in modeling the experimental parameters of the hydrolysis process to achieve an efficient condition for maximum glucose yield. This was investigated using enzyme loading (10-50 FPU/g), substrate concentration (20-80 mg/mL), surfactant concentration (2-8 mg/mL), and incubation time (24-96 h) as the influencing parameters. The optimum glucose yields predicted were 97.21% and 97.87% for *S. spontaneum* and banana peduncle, respectively. These compared well to ANN validated yields of 95.93% and 93.51% and ANFIS validated yields of 96.1% and 97.42% for *S. spontaneum* and banana peduncle, respectively. Based on the training and validation data sets, all models were statistically compared using the determination coefficient (R^2), adjusted R^2, root mean squares error (RMSE), hybrid fractional error function (HYBRID), average relative error (ARE), absolute average relative error (AARE) and Marquardt's percent standard error deviation (MPSED). According to the statistical indices obtained, RSM performed the least, while ANFIS marginally outperformed ANN. This shows that ANFIS is a powerful tool for modeling and optimizing glucose production in enzymatic hydrolysis.



4.1. Introduction

Glucose is a necessary sugar that serves as a platform component for producing chemicals, biofuels, and materials using chemical and biological methods [1]. Glucose is traditionally derived from starch, however, given the current food shortage globally, the quest for non-food sugar sources is significant and intriguing [2]. Cellulose, an alternate source for glucose production is an abundant polymer of β 1,4-glycosidic bond linked glucose units that can be hydrolyzed either chemically or enzymatically. Over the last few decades, acid catalysts have been demonstrated to be the most accelerating agent for the degradation of cellulose [3, 4, 5]. However, the use of acid catalysts causes significant issues with equipment corrosion, catalysts isolation from products, and waste disposal [6]. On the other hand, enzymatic hydrolysis is advantageous, requires less energy, and operates in mild environmental conditions. Thus, the cost of the process is lower compared to acid hydrolysis [7]. Furthermore, the enzymatic hydrolysis is substrate-specific and does not result in the formation of any byproducts [8]. However, the bioconversion process is currently not economically viable since enzymatic hydrolysis is slow and requires a high enzyme loading to achieve acceptable rates and yields due to the recalcitrant nature of lignocelluloses [9]. Though, the removal of lignin and hemicellulose in the pretreatment process can significantly enhance the hydrolysis of lignocelluloses [10]. In addition, applying surfactants has also shown potential in improving the performance of cellulose hydrolysis due to the higher availability of enzymes for cellulose degradation [11, 12]. Extensive research has demonstrated that the efficiency of enzymatic hydrolysis of cellulose depends on several process parameters including enzyme loading, substrate concentration, reaction time, surfactant addition, and so on [10]. Since these variables often interact, optimizing the enzymatic hydrolysis process is crucial to improve the process efficiency.

Modeling and optimization are perhaps the most essential measures in biological processes as they improve the system and increase process efficiency without increasing costs [13]. The conventional method of optimization, on the other hand, involves changing one independent variable at a time while holding the other variables constant, which is very time-consuming, laborious, and expensive, and often leads to an incomplete understanding of the system behavior, resulting in a lack of prediction ability [14]. In this context, response surface methodology (RSM) is an alternative and strong mathematical technique for quickly and efficiently analyzing the influence of numerous parameters,

alone or in combination, on a particular system with a minimal number of experiments and high statistical significance in the results [15]. This method has already been used to enhance the enzymatic hydrolysis of several substrates, including cellulose [16], wheat straw [17], rice straw [18], sugar beet pulp [19], etc. Artificial neural network (ANN), on the other hand, is a computational approach that replicates the biological processing power of the human brain. Because of its capacity to recognize and replicate cause-effect connections for multiple input-output systems through training, ANN can model and simulate very complex systems. Many researchers in the field of bioprocess engineering have reported on its application [20, 21, 22]. Likewise, Jang's adaptive neuro-fuzzy inference system (ANFIS) combines the characteristics of an adaptive neural network with a fuzzy inference system (FIS) to anticipate correct output from numerical input data [23, 24]. ANFIS was used by Betiku et al. and Onu et al. to simulate the esterification process and estimate dye adsorption ability, respectively [25, 26]. Recently, ANFIS has been utilized to assist different lignocellulosic biomass studies. Rego et al. utilized ANFIS to predict the lignin, glucose, and xylose content in processed sugarcane bagasse [27], while Lerkkasemsan et al. used ANFIS to estimate the composition of *J. curcas* and *Pongamia pinnata* in pyrolysis reactions [28]. Although the approach has been used in modeling a variety of systems, its application to the enzymatic hydrolysis process is scarce.

There is a lot of literature on RSM and ANN performance evaluations of process modeling [29, 30]. The findings are mixed for ANN and ANFIS [27, 31] and RSM and ANFIS [32, 33] in the same studies. ANN has been shown to outperform RSM [29]. While Kim and Park discovered that ANFIS is more accurate than RSM when it comes to prediction abilities, Taheri et al. reported the reverse [34, 35]. Another study found that ANFIS outperformed ANN, whereas Karimi et al. observed that both techniques were equally successful [36, 37]. However, data on comparing the predictive capabilities of RSM, ANN, and ANFIS in the same experiment to clarify these contradicting results are limited.

The current study is a continuation of the previous chapter (Chapter 3), where we have developed an efficient and environmentally friendly conversion method of cellulose to glucose with the readily available commercial enzyme Cellulase (from *Aspergillus niger*). The effects of the experimental parameters (enzyme loading, substrate concentration, surfactant concentration, and incubation time) were studied employing RSM. The findings of RSM were then established using ANN and ANFIS as modeling

tools and their prediction abilities were compared in order to obtain the optimal conditions for maximum production of glucose.

4.2. Materials and methods

4.2.1. Materials and chemicals

Tween-80, sodium citrate ($Na_3C_6H_5O_7$), and citric acid ($C_6H_8O_7$) were purchased from Merck Pvt. Ltd., India. A commercial *Aspergillus niger* cellulase procured from Merck Pvt. Ltd., India was used for enzymatic conversions. Cellulose isolated from *S. spontaneum* and banana peduncle as discussed in Chapter 3 has been used as the substrate material for the reactions.

4.2.2. Enzymatic hydrolysis

Enzymatic hydrolysis of cellulose isolated from *S. spontaneum* and banana peduncle (as discussed in Chapter 3 [29, 38]) was carried out using commercial *Aspergillus niger* cellulase enzyme under different conditions. Cellulase enzyme activity was estimated to be 50 FPU/g (filter paper units per gram of enzymatic solution) using the NREL protocol [39]. The quantity of enzyme required to release 1µmol of glucose/min during the hydrolysis reaction has been specified as one international filter paper unit (FPU).

The enzymatic hydrolysis was conducted in a 2 mL screw-cap tube containing cellulose, cellulase enzyme powder, Tween-80 surfactant, and 1.5 mL of 0.05 M citric acid/ citric sodium buffer (pH 4.8). The tubes were incubated at 50°C in a rotary shaker (ORBITEK range, Scigenics Biotech Pvt. Ltd., India) at 150 rpm. Different enzymatic hydrolysis conditions were tested according to response surface methodology (RSM) of central composite design (CCD). According to the experimental design, samples (1 mL) were taken from the hydrolysis solution at different time intervals. The hydrolysis was terminated by heating immediately for 5 min on a boiling water bath at a temperature of 95°C and the solution was filtered through a 0.22 µm syringe filter before analysis. The glucose concentrations in the supernatants were quantified *via* high-performance liquid chromatography (HPLC). The predicted values of RSM were then validated using ANN and ANFIS models. A schematic representation of the complete experimental procedure is provided in Fig. 4.1.

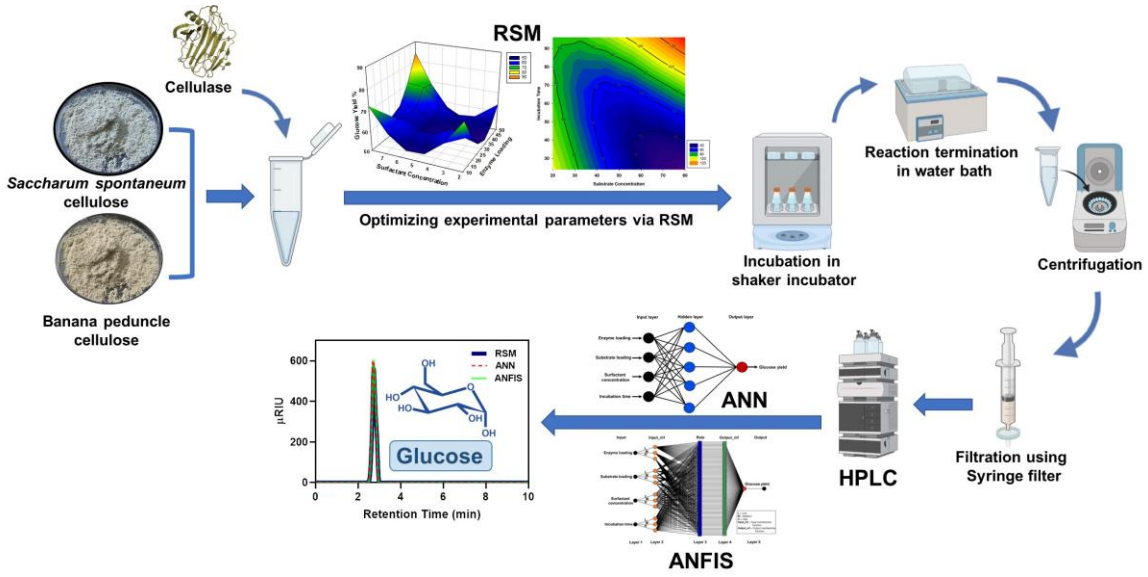


Fig. 4.1. *Schematic representation of the complete enzymatic hydrolysis procedure.*

4.2.3. Optimizing enzymatic hydrolysis variables using RSM

RSM statistical technique was employed to optimize the independent experimental variables influencing cellulase activity and simultaneously achieve the best device output [40, 41]. For this, the effects of univariate and multivariate interactions between the hydrolysis parameters (enzyme loading, substrate concentration, surfactant concentration, and incubation time) on glucose production were studied through RSM. A Central Composite Design (CCD) was used to achieve the fewest possible differences in the four parameters listed above and to analyze the experimental results for each of two different celluloses obtained from *S. spontaneum* and banana peduncle. For the quadratic model with four factors and five levels, a total of 30 experimentations were considered and the values of the response were obtained as a mean of the triplicate runs. The boundary conditions for each parameter were designed based on the preliminary studies and are depicted in Table 4.1. The analysis of variance (ANOVA), regression analysis, design of experiment, and graphical analysis were performed using the statistical software tool, Design-Expert 7.0 (Stat-Ease Inc., USA). To represent the association between prediction response and experimental components the hydrolysis data was modeled utilizing the second-degree polynomial equation shown below [29]:

$$Y = b_0 + \sum_{i=1}^{n} b_i X_i + \sum_{i=1}^{n} b_{ii} X_i^2 + \sum_{i=1}^{n} \sum_{i>j}^{n} b_{ij} X_i X_j \qquad (1)$$

where n is the total number of variables, X_i and X_j are the coded values, b_0 is the model constant, Y is the predictive response, b_{ii} is the interaction coefficients, b_i is the linear coefficient, and b_{ij} is the quadratic coefficient.

In addition, Pareto analysis was used to determine the sensitivity of each individual parameter to the response variable inside a design model. To determine the relative impact of individual factors, the following equation was utilized for Pareto computation.

$$P_i = \left(\frac{\beta_i^2}{\sum \beta_i^2}\right) \times 100 \ (i \neq 0) \quad (2)$$

where Pi is the percentage effect of individual factors on the response [42].

Table 4.1. *Coded levels of variables for CCD.*

Variables	Units	Coded Levels				
		-α	**-1**	**0**	**1**	**+α**
(A) Enzyme loading	FPU/g	10	20	30	40	50
(B) Substrate loading	mg/mL	20	35	50	65	80
(C) Surfactant Concentration	mg/mL	2	3.5	5	6.5	8
(D) Incubation time	h	24	42	60	78	96

4.2.4. Optimizing enzymatic hydrolysis variables using ANN

For neural network modeling of the process parameters for the enzymatic hydrolysis reactions, a three-layer feed-forward Multi-Layered- Perceptron (MLP) based artificial neural network (ANN) developed using the MATLAB R2013a computing suite (MathWorks USA) was implemented to predict the impact of multivariate associations among the different hydrolysis parameters on the glucose yield. The network contains an input layer with four neurons (enzyme loading, substrate concentration, surfactant concentration, and incubation time), an output layer with one neuron (glucose yield), and a hidden layer (Fig. 4.2 a). The neural system was accomplished by a Levenberg Marquardt (LM) back-propagation learning algorithm with tan sigmoid function (tansig) and a linear transfer function (purelin) in the hidden and output layer, respectively which were used to express the activation function defined by the following equations. The number of hidden neurons was set iteratively by testing a variety of neural networks until the output's mean square error (MSE) was reduced. To avoid overtraining and over parameterization, the

experimental data was divided into three sections training (70%), testing (15%), and validation (15%) [43].

$$X_j = \sum_{i=1}^{P} Y_i w_{ji}^{in} + b_j^{in} \quad (3)$$

$$f(x) = \text{transig}(x) = \frac{1-e^{-x}}{1+e^{-x}} \quad (4)$$

where X_j is the net input to the node j in the hidden layer, Y_i is the input to a neuron, w_{ji}^{in} is the weight associated with each input connection from ith to jth neuron in the hidden layer and b^{in}_j is the bias of the jth neuron in the hidden layer.

4.2.5. Optimizing enzymatic hydrolysis variables using ANFIS

ANFIS is a powerful technique for high-speed modeling of complex nonlinear systems that consists mostly of artificial neural networks (ANN) augmented by fuzzy logic [44]. In our study, the ANFIS model was stimulated using a five-layer Sugeno model that employed the fuzzy IF-THEN principle. Assuming that the fuzzy inference system (FIS) consists of two inputs (x_1, x_2) and one output (y), the IF-THEN rule applies as follows [23]:

Rule 1: IF x_1 is P_1 and x_2 is Q_1, THEN $y_1 = a_1x_1 + b_1x_2 + c_1$

Rule 2: IF x_1 is P_2 and x_2 is Q_2, THEN $y_2 = a_2x_1 + b_2x_2 + c_2$

where P_1, P_2, and Q_1, Q_2 are the fuzzy sets, y_1, y_2 are the system outputs, while a_1, a_2, b_1, b_2, and c_1, c_2 are adjustable parameters.

Fig. 4.2 b depicts the architecture of the proposed ANFIS model, with the first and last layers indicating the input variables (enzyme loading, substrate loading, surfactant concentration, and incubation time) and the response or output variable (glucose yield) respectively. The model matched to first-order Sugeno inference systems in the second layer, which fuzzified input parameters (fuzzification) into membership values using membership functions (MF). Following that, in the fourth layer, output membership functions were used to defuzzify the inference output to actual output values [26]. The fifth layer only had one node, which presented the overall output, which was the percentage of glucose yield, as the sum of all incoming signals.

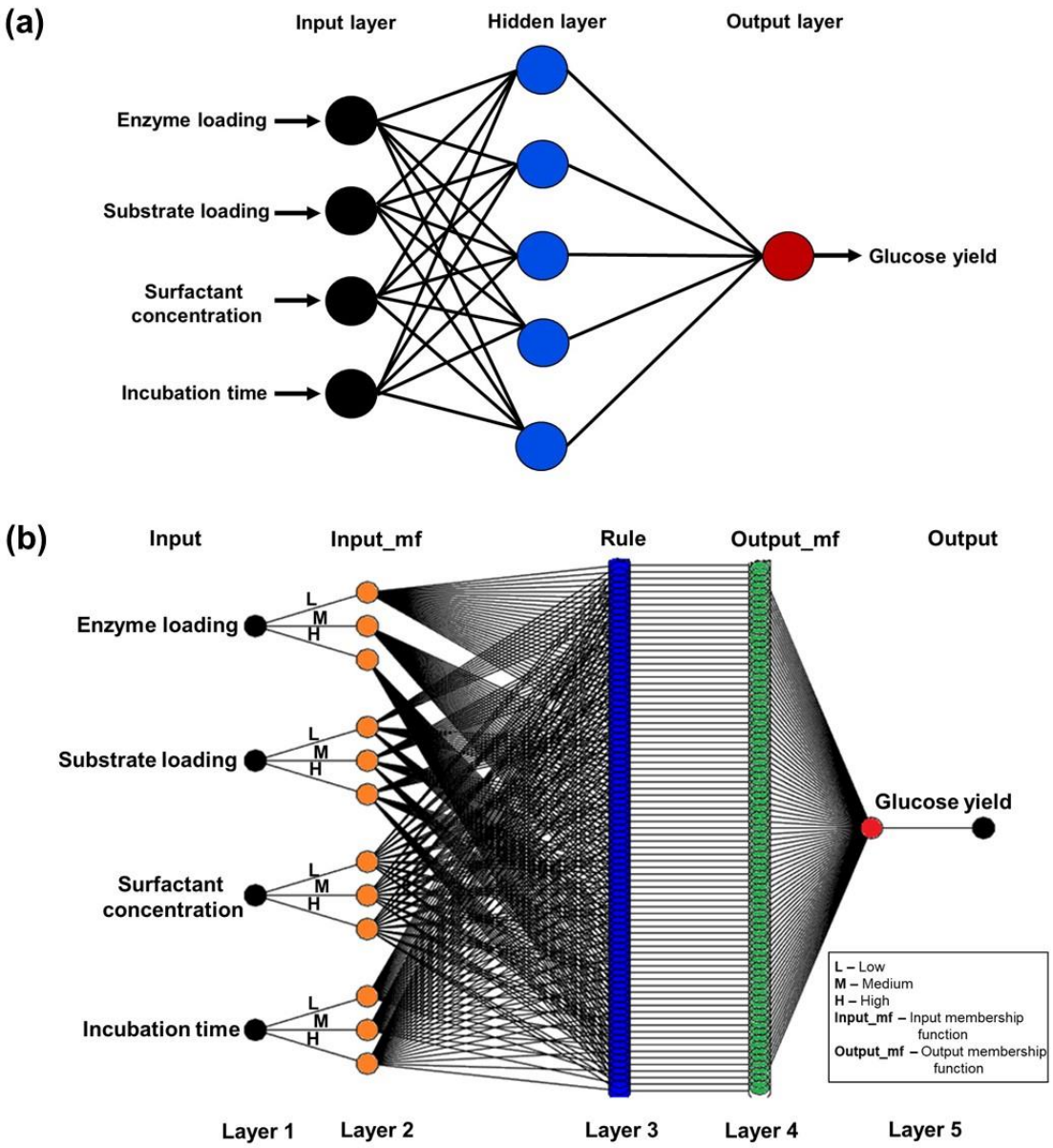


Fig. 4.2. *Architectures for the glucose yield prediction model by ANN (a) and ANFIS (b).*

4.2.6. Model performance indices

The ANFIS, ANN, and RSM modeling predictions were exposed to performance indices to provide a ranking that highlighted the model with the best predictive capability with respect to the experimental data. Seven high-performance statistical error functions are used in the analysis, which are listed in Table 4.2. The appraisal indices that were chosen were based on the characteristics of the data set that was employed [45]. A comparison parity plot was also created, which indicated specific deviation points between RSM, ANN, and ANFIS model predictions based on the experimental data.

Table 4.2. *Statistical error functions.*

Error function	Equation	Reference
Root mean square error	$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{P_{R,\exp(i)}-P_{R,cal(i)}}{P_{R,\exp(i)}}\right)^2}$	[46]
Hybrid fractional error function	$\text{HYBRID (\%)} = \frac{1}{N-P}\sum\left[\frac{\left(P_{R,i,exp}-P_{R,i,cal}\right)^2}{P_{R,i,exp}}\right]100$	[47]
Average relative error	$\text{ARE (\%)} = \frac{100}{N}\sum_{i-1}^{N}\frac{\lvert P_{R,i,exp}-P_{R,i,cal}\rvert}{P_{R,i,exp}}$	[46]
Absolute average relative error	$\text{AARE} = \frac{1}{N}\sum_{i=1}^{N}\left(\left\lvert\frac{P_{R,\exp(i)}-P_{R,cal(i)}}{P_{R,\exp(i)}}\right\rvert\right)$	[47]
Marquardt's percent standard error deviation	$\text{MPSED (\%)} = \sqrt{\frac{\sum\left(\frac{P_{R,exp}-P_{R,cal}}{P_{R,exp}}\right)^2}{N-P}}\times 100$	[46]
Correlation coefficient	$R^2 = \frac{\sum_{i=1}^{N}\left(P_{R,i,cal}-P_{R,exp,ave}\right)^2}{\sum_{i=1}^{N}\left(P_{R,i,cal}-P_{R,exp,ave}\right)^2+\sum_{i=1}^{N}\left(P_{R,i,cal}-P_{R,i,exp}\right)^2}$	[48]
Adjusted R^2	$\text{Adj } R^2 = 1-\left[(1-R^2)\times\frac{N-1}{N-P-2}\right]$	[46]

where N is the number of experimental runs; $P_{R,exp(i)}$, $P_{R,i,exp}$ are the experimental values of the i^{th} experiment; $P_{R,cal(i)}$, $P_{R,i,cal}$ are the model predictions of the i^{th} experiment; $P_{R,exp,ave}$ is the experimentally determined average value, N is the number of experimental runs while P is the number of factors.

4.2.7. Product analysis

Glucose concentrations in the hydrolysis solution were analyzed using high-performance liquid chromatography (HPLC, Thermoscientific) equipped with an Accucore Hilic Amide column (150 X 4.6 mm) and a RI detector. The samples were injected into the column under the conditions: column temperature of 65°C, 0.5 mM of H_2SO_4 as mobile phase at a flow rate of 0.6 mL per minute, and an injection volume of 20 µl. The glucose concentration was calculated using calibration curves obtained from standard glucose solutions run under similar experimental conditions.

4.3. Results and discussion

4.3.1. Optimization of enzymatic hydrolysis using RSM

To optimize the reaction conditions for the enzymatic hydrolysis of cellulose (isolated from *S. spontaneum* and banana peduncle), RSM experiments were conducted for the two substrates separately. After incorporating the experimental value for the response, i.e., the glucose yield, the CCD matrix provided the optimal conditions with four independent variables for each of the substrate, in the form of statistically significant fitted models. P values in the region of 0.0001 indicated that the fitted models had adequate confidence levels. The study was conducted using an α value of 0.05 for both the substrates, which corresponds to a 95% confidence level. The P values for the fitted models were found to be smaller than the α value, rejecting the null hypothesis and confirming the statistical significance of the fitted models for both substrates. In addition, the fitted models had high F values, but the values indicating lack of fit were non-significant, confirming the goodness of fit [49, 50].

Saccharum spontaneum cellulose

For the *S. spontaneum* cellulose, the variation of incubation time, enzyme concentration, substrate concentration, and surfactant concentration, significantly influenced the response parameter, in the fitted single factor interactions (Fig. 4.3 a-c). The optimized conditions for *S. spontaneum* cellulose showing the highest glucose yield were obtained for 30 FPU/g of enzyme loading, 50 mg/ml of substrate loading, 5 mg/ml of surfactant concentration, and 96 h of incubation time (Table 4.3). The quadratic model, which was the best-fitted model was subjected to analysis of variance (ANOVA) resulting in a model determination coefficient R^2=0.97 which suggests that 97% of the total variations of the experimental

parameters could be described by the quadratic model. The statistical significance, fit adequacy, and competency of the predicted model were all evaluated using analysis of variance (ANOVA) (Table 4.4). As a result, the R^2 value was considered to have substantial goodness of fit. The predicted R^2=0.99 was in fair accordance with the adjusted R^2 indicating the accuracy of the model. The accuracy of the model was demonstrated by the variables with a p-value < 0.05 which have a significant effect on the response. However, the p-value for lack of fit should be > 0.05 for higher accuracy. From Table 4.4 it is seen that the lack of fit value for the glucose yield is > 0.05 which confirms the accuracy of the predicted model. The variables, substrate concentration, and incubation time were found to be significant for the hydrolysis reaction. The quadratic terms AB, BC, BD, A^2, B^2, and D^2 were found to have a significant effect on the yield of glucose (Table 4.4). Single factor interaction plots were used to show the effects of interaction among the significant variables (Fig. 4.3 a-c). The significant multi-variate interactions among the variables are depicted in 3D plots and their corresponding contour plots in Fig. 4.4 a-c and a´-c´, respectively.

For the single factor interactions, an increase in the time of incubation results in an increase in glucose yield (Fig.4.3 c). On the other hand, an increase in the concentration of substrate initially reduces the glucose yield which on further increase in the concentration results in an increased yield of glucose (Fig.4.3 b). The enzyme loading influences minor variations in glucose yield, when added within the boundary conditions (Fig. 4.3 a). For the multivariate interactions, the 3D plots show the influence of substrate concentration on the glucose yield, when in conjunction with enzyme loading (Fig. 4.4 a) or surfactant concentration (Fig. 4.4 b) or incubation time (Fig. 4.4 c). Interactions between substrate concentration and incubation time were also found to be significant where a maximum glucose yield of ~88% was observed, using a 65 mg/ml of substrate loading for an incubation time of 78 h (Fig. 4.4 c). The corresponding contour plots for the aforementioned 3D plots (Fig. 4.4 a-c) are represented in Fig. 4.4 a´-c´. As shown in Fig.2 a´-c´ the contour plots for the fitted models are elliptical in nature, thereby establishing the significant interactions between the variables. Since the elliptical shape denotes that there is significant interaction between the two variables [51]. The empirical relations between these variables and the glucose yield by the enzymatic hydrolysis can be defined by the following second-order equation:

Glucose yield = + 361.14876 - 0.97624 * A - 6.86090 * B - 5.35827 * C - 4.29331 * D + 4.63171E-003 * A * B + 0.015921 * A * C + 1.31642E-003 * A * D + 0.048652 * B * C

+ 0.046542 * B * D + 0.017219 * C * D + 9.73199E-003 * A^2 + 0.034915 * B^2 + 0.13454 * C^2 + 0.020831 * D^2

where, A = Enzyme loading (FPU/g), B = Substrate concentration (mg/mL), C = Surfactant concentration (mg/mL) and D = incubation time (h).

Table 4.3. *Central composite design matrix, ANN, and ANFIS for the four independent variables on the glucose yield from S. spontaneum cellulose with the predicted and actual response.*

Run	Enzyme loading (FPU/g))	Substrate loading (mg/mL)	Surfactant concentration (mg/mL)	Incubation time	Glucose yield (%)					
					RSM		ANN		ANFIS	
					Predicted	Observed	Predicted	Observed	Predicted	Observed
1	20	35	3.5	42	69.85	69.67	69.67	69.67	69.67	69.67
2	40	65	6.5	42	34.22	37.27	36.6	37.27	37.27	37.27
3	20	35	6.5	42	66.04	65.87	65.9	65.87	65.48	65.87
4	20	65	3.5	42	35.3	35.83	35.66	35.83	35.83	35.83
5	30	50	5	60	44.49	47.66	45.41	47.66	47.18	47.66
6	30	50	5	60	47.49	47.95	47.41	47.95	47.59	47.95
7	50	50	5	60	51.34	50.95	53.9	50.95	50.00	50.95
8	20	65	6.5	78	81.19	86.2	86.17	86.2	86.12	86.2
9	40	65	6.5	78	87.48	88.15	88.46	88.15	87.98	88.15
10	40	65	3.5	78	79.1	84.91	83.92	84.91	84.88	84.91
11	30	50	5	60	47.49	47.7	47.41	47.7	47.68	47.7
12	20	65	3.5	78	82.75	83.21	86.13	83.21	83.57	83.21
13	30	50	8	60	48.49	48.02	48.07	48.02	47.82	48.02
14	30	50	5	60	47.72	47.7	47.41	47.7	47.67	47.7
15	40	35	6.5	78	62.61	64.73	63.73	64.73	63.00	64.73
16	30	50	5	96	97.21	96.07	95.93	96.07	96.1	96.07
17	40	35	3.5	42	67.46	67.09	67.09	67.09	67.13	67.09
18	30	50	5	60	47.49	47.6	47.41	47.6	47.6	47.6
19	30	80	5	60	73.07	72.26	72.15	72.26	72.26	72.26
20	30	50	5	60	48.49	46.31	47.41	46.31	45.45	46.31
21	20	35	6.5	78	65.09	65.91	66.22	65.91	65.19	65.91
22	10	50	5	60	51.42	50.68	47.11	50.68	50.51	50.68
23	30	50	5	24	50.76	51.77	51.78	51.77	52.11	51.77

24	30	20	5	60	80.75	84.43	84.43	84.43	84.58	84.43
25	40	35	6.5	42	62.61	64.64	63.65	64.64	64.32	64.64
26	30	50	2	60	48.91	48.24	48.25	48.24	47.99	48.24
27	40	35	3.5	78	65.6	66.54	66.54	66.54	66.54	66.54
28	20	35	3.5	78	67.04	67.63	67.63	67.63	68.76	67.63
29	20	65	6.5	42	35.87	35.42	35.43	35.42	33.25	35.42
30	40	65	3.5	42	32.69	35.36	33.36	35.36	34.73	35.36

Table 4.4. *ANOVA for the fitted quadratic polynomial model for glucose yield of S. spontaneum cellulose.*

Source	**Sum of Squares**	**Degree of Freedom**	**Mean Square**	**F Value**	**P-value Prob > F**	
Model	8800.756358	14	628.6254541	648.4519512	< 0.0001	significant
A-Enzyme Loading	0.01102102	1	0.01102102	0.011368617	0.9165	
B-Substrate Concentration	204.7066065	1	204.7066065	211.162939	< 0.0001	significant
C-Surfactant Concentration	0.2604375	1	0.2604375	0.268651554	0.6118	
D-Incubation Time	3378.290612	1	3378.290612	3484.840019	< 0.0001	significant
AB	7.722979951	1	7.722979951	7.966558442	0.0129	significant
AC	0.912550326	1	0.912550326	0.941331655	0.3473	
AD	0.898372231	1	0.898372231	0.926706391	0.3510	
BC	19.17279476	1	19.17279476	19.77749404	0.0005	significant
BD	2526.562685	1	2526.562685	2606.249067	< 0.0001	significant
CD	3.458391106	1	3.458391106	3.567466837	0.0784	
A^2	25.97804469	1	25.97804469	26.79737776	0.0001	significant
B^2	1692.742197	1	1692.742197	1746.130344	< 0.0001	significant
C2	2.513603519	1	2.513603519	2.592881175	0.1282	
D^2	1249.482482	1	1249.482482	1288.890465	< 0.0001	significant
Residual	14.54137318	15	0.969424879			
Lack of Fit	12.79963757	10	1.279963757	3.674391644	0.0818	non-significant
Pure Error	1.741735613	5	0.348347123			
Cor Total	8815.297731	29				
Std. Dev. = 0.98		**Mean = 60.19**		**Predicted R^2 = 0.97**		
R^2 = 0.98		**Adj. R^2 = 0.96**		**Adequate Precision = 90.37**		

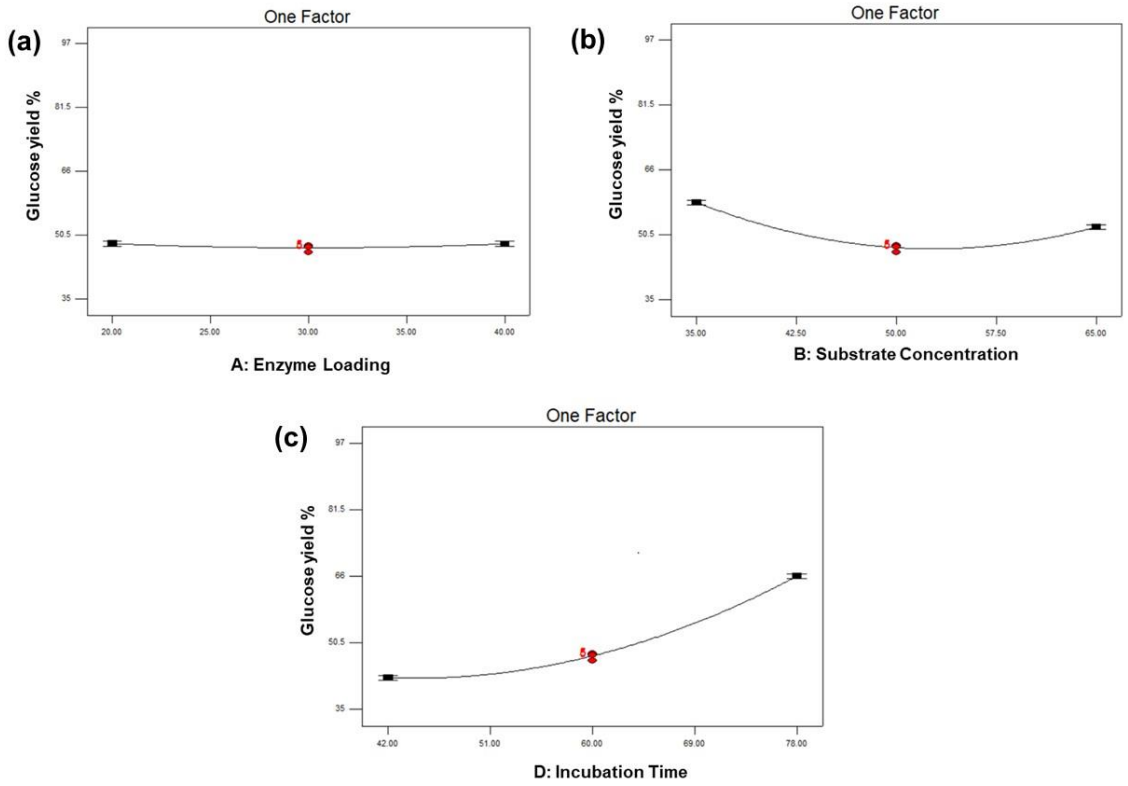


Fig. 4.3. *Effect of univariate interactions described by the model on the glucose yield for S. spontaneum cellulose.*

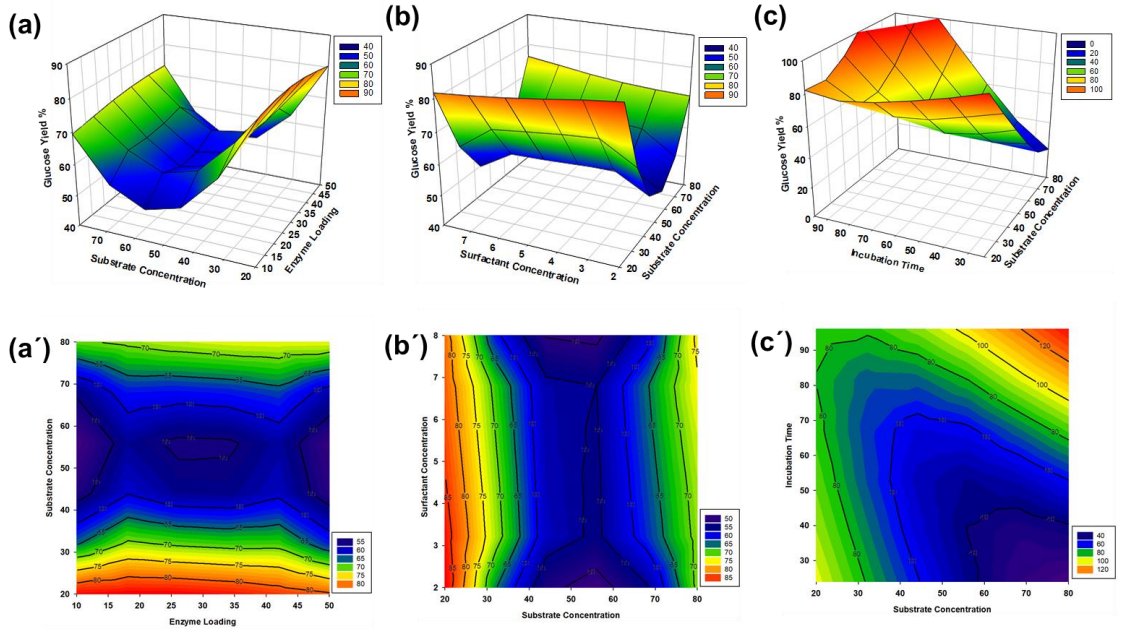


Fig. 4.4. *Effect of multivariate interactions: Response surface 3D plots (a-c) and response surface contour plots (a´-c´) described by the model on the glucose yield for S. spontaneum cellulose.*

Banana peduncle cellulose

The optimized conditions for banana peduncle cellulose showing the highest glucose yield were obtained for 30 FPU/g of enzyme loading, 50 mg/ml of substrate loading, 5 mg/ml of surfactant concentration, and 96 h of incubation time (Table 4.5). The substrate concentration, surfactant concentration, and incubation time were observed to be the most significant parameters influencing glucose yield response for the banana peduncle cellulose (Table 4.6). The expected R^2=0.98 remained in good accordance with the predicted R^2 suggesting that the model was accurate. AC, AD, CD, BD, A^2, C^2, B^2, and D^2 are quadratic relations that have a substantial impact on the glucose yield (Table 4.6). The effect of interaction among the significant variables was visualized using single-factor interaction plots (Fig. 4.5 a-d). Relevant multivariate 3D interaction plots and their corresponding contour plots are represented in Fig. 4.6 a-d and a´-d´, respectively.

For single-factor interactions, increasing incubation time results in an improvement of the glucose yield (Fig. 4.5 d). While increasing the concentration of substrate initially lowers the glucose yield but with an increase in the concentration, the glucose yield increases (Fig. 4.5 b). The enzyme loading and surfactant concentrations negligibly influence glucose yield, after introduction within the boundary levels (Fig. 4.5 a, c). The 3D graphs indicate the effect of enzyme loading on glucose yield for multivariate interactions, once combined with either surfactant concentration (Fig. 4.6 a) or incubation time (Fig. 4.6 b), for multivariate interactions. In contrast, incubation time affected glucose yield while conjunction with substrate concentration (Fig. 4.6 c) and surfactant concentration (Fig. 4.6 d). As apparent from the single factor interaction plot (Fig. 4.5 d) and response surface plot (Fig. 4.6 b, c, d), the incubation time was found to have a substantial effect on the hydrolysis and a maximum glucose yield of 97% was observed with 30 FPU/g of enzyme loading, 50 mg/ml of substrate loading, 5 mg/ml of surfactant concentration and 96 h of incubation time. Fig. 4.6 a´-d´ shows the contour plots for the aforesaid 3D plots are elliptical representing significant interactions between the variables. The following second-order equation can be used to describe the empirical relationships between these variables and the glucose yield by the enzymatic hydrolysis:

Glucose yield (%) = + 44.91 + 0.24 * A - 6.16 * B + 2.04 * C + 13.01 * D + 1.27 * A * B + 2.09 * A * C - 2.50 * A * D + 0.72 * B * C + 5.55 * B * D - 2.89 * C * D + 2.45 + 4.39 + 3.02 * C^2 + 6.73 * D^2

where, A = Enzyme loading (FPU/g), B = Substrate concentration (mg/mL), C = Surfactant concentration (mg/mL) and D = incubation time (h).

Table 4.5. *Central composite design matrix, ANN, and, ANFIS for the four independent variables on the glucose yield from banana peduncle with the predicted and actual responses.*

Experimental runs	Enzyme loading (FPU/g)	Substrate loading (mg/mL)	Surfactant concentration (mg/mL)	Incubation time (h)	Glucose Yield%					
					RSM		ANN		ANFIS	
					Predicted	Observed	Predicted	Observed	Predicted	Observed
1	20	35	3.5	42	56.61	56.84	56.19	56.84	56.84	56.84
2	40	65	6.5	42	45.54	49.79	49.75	49.79	49.79	49.79
3	20	35	6.5	42	58.85	60.91	60.52	60.91	60.91	60.91
4	20	65	3.5	42	29.21	30.76	32.36	30.76	30.76	30.74
5	30	50	5	60	43.91	45.4	45.4	45.4	45.85	44.58
6	30	50	5	60	44.91	45.14	44.4	45.14	44.85	44.54
7	50	50	5	60	55.18	54.27	53.17	54.27	54.27	54.02
8	20	65	6.5	78	72.67	75	75.03	75	75	75.00
9	40	65	6.5	78	69.89	74.96	74.92	74.96	74.96	74.97
10	40	65	3.5	78	69.96	72.42	72.89	72.42	72.42	72.00
11	30	50	5	60	44.91	43.26	43.4	43.26	43.85	44.50
12	20	65	3.5	78	77.11	76.09	77	76.09	76.09	76.09
13	30	50	8	60	61.09	59.38	58.32	59.38	59.38	58.32
14	30	50	5	60	48.91	45.58	44.4	45.58	45.85	44.24
15	40	35	6.5	78	72.12	72.09	72.03	72.09	72.09	71.91
16	30	50	5	96	97.87	97.42	93.51	97.42	97.42	96.79
17	40	35	3.5	42	55.35	54.55	47.73	54.55	54.55	52.65
18	30	50	5	60	44.91	45.4	44.4	45.4	45.85	43.23
19	30	80	5	60	50.15	47.76	46.75	47.76	48.09	54.50
20	30	50	5	60	40.91	44.69	44.4	44.69	44.85	44.00
21	20	35	6.5	78	74.98	74.77	74.57	74.77	72.64	71.79
22	10	50	5	60	54.22	53.3	51.62	53.3	54.05	52.54
23	30	50	5	24	45.82	44.45	44.03	44.45	45.61	46.10
24	30	20	5	60	74.78	75.34	75.56	75.34	74.9	75.00
25	40	35	6.5	42	74.97	69.29	70.44	69.29	69.33	69.22
26	30	50	2	60	52.91	52.79	51.79	52.79	52.43	52.00
27	40	35	3.5	78	71.07	70.92	70.47	70.92	70.87	68.98
28	20	35	3.5	78	82.31	82.58	81.85	82.58	82.46	81.67
29	20	65	6.5	42	33.34	36.8	36.74	36.8	36.74	35.12
30	40	65	3.5	42	30.04	33.55	32.14	33.55	33.65	33.33

Table 4.6. *ANOVA for the fitted quadratic polynomial model for glucose yield of banana peduncle cellulose.*

Source	Sum of Squares	Degree of Freedom	Mean Square	F Value	P-value Prob > F	
Model	7575.65	14	541.1179	252.9088	< 0.0001	significant
A-Enzyme Loading	1.382832	1	1.382832	0.646311	0.434	non-significant
B-Substrate Concentration	909.5059	1	909.5059	425.0868	< 0.0001	significant
C-Surfactant Concentration	100.3379	1	100.3379	46.89615	< 0.0001	significant
D-Incubation Time	4063.489	1	4063.489	1899.202	< 0.0001	significant
AB	25.81936	1	25.81936	12.0675	0.0034	non-significant
AC	70.04979	1	70.04979	32.74001	< 0.0001	significant
AD	99.66877	1	99.66877	46.5834	< 0.0001	significant
BC	8.312121	1	8.312121	3.884937	0.0675	non-significant
BD	492.8389	1	492.8389	230.3441	< 0.0001	significant
CD	133.8724	1	133.8724	62.56957	< 0.0001	significant
A^2	164.2707	1	164.2707	76.7772	< 0.0001	significant
B^2	528.1971	1	528.1971	246.8699	< 0.0001	significant
C^2	250.4912	1	250.4912	117.0751	< 0.0001	significant
D^2	1243.857	1	1243.857	581.3565	< 0.0001	significant
Residual	32.09366	15	2.139577			
Lack of Fit	28.35623	10	2.835623	3.793544	0.0771	non-significant
Pure Error	3.737432	5	0.747486			
Cor Total	7607.744	29				
Std. Dev. = 1.46		**Mean = 58.18**		**Predicted R^2 = 0.97**		
R^2 = 0.98		**Adj. R^2 = 0.98**		**Adequate Precision = 66.38**		

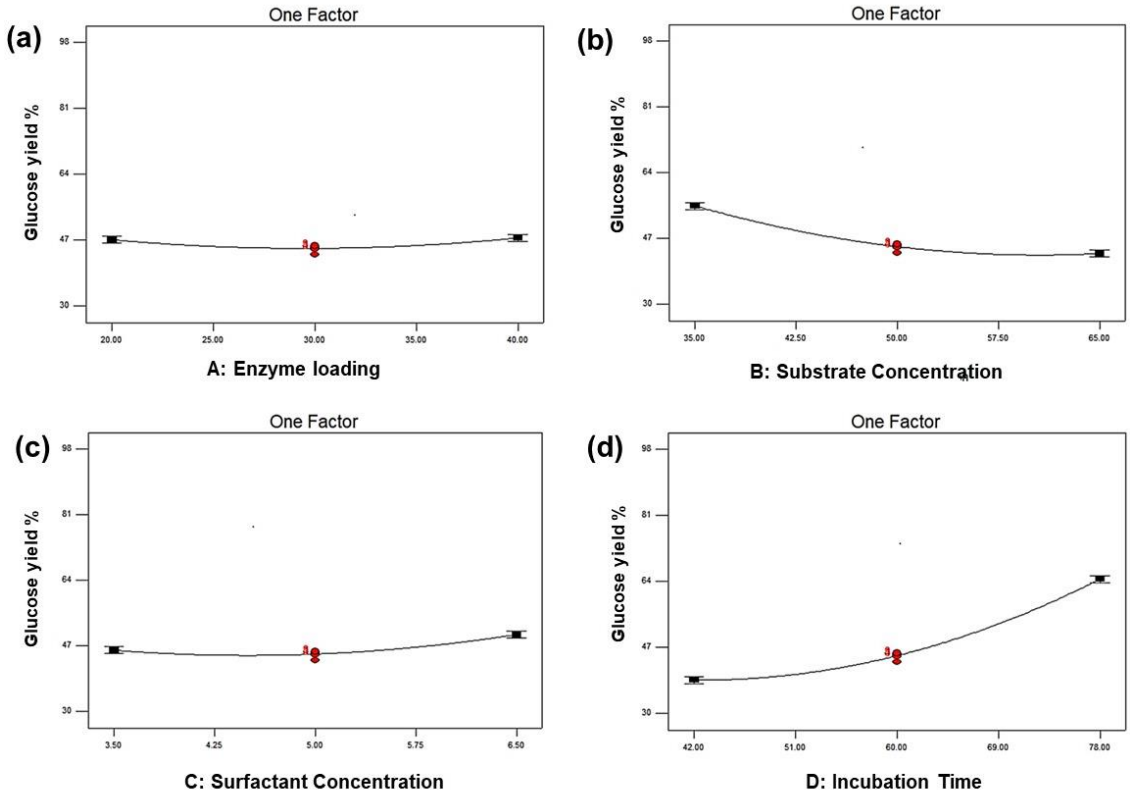


Fig. 4.5. *Effect of univariate interactions described by the model on the glucose yield for banana peduncle cellulose.*

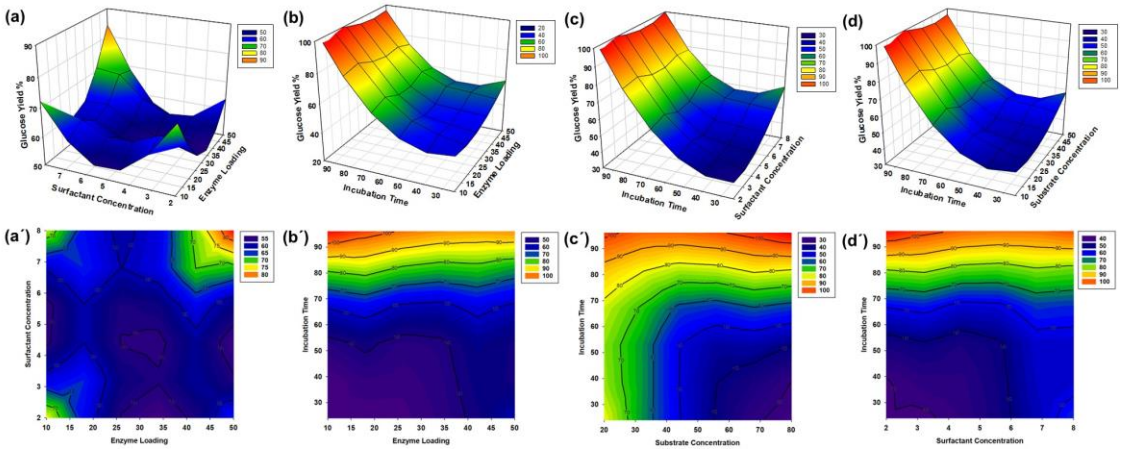


Fig. 4.6. *Effect of multivariate interactions: Response surface 3D plots (a-d) and response surface contour plots (a´-d´) described by the model on the glucose yield for banana peduncle cellulose.*

The Pareto graphic analysis was carried out to determine the percentage effect of the most influential parameters as described by the model equation. The effect of incubation time was found to be the most significant parameter for both *S. spontaneum*

and banana peduncle celluloses as illustrated in Fig. 4.7 a and b and is consistent with the results obtained (Table 4.4 and Table 4.6). Furthermore, the normal probability plots (Fig. 4.7 c and d), show the data points are distributed roughly linearly in both cases, forming a straight line that is in proper compliance with normally distributed response residuals confirming the adequacy of the models.

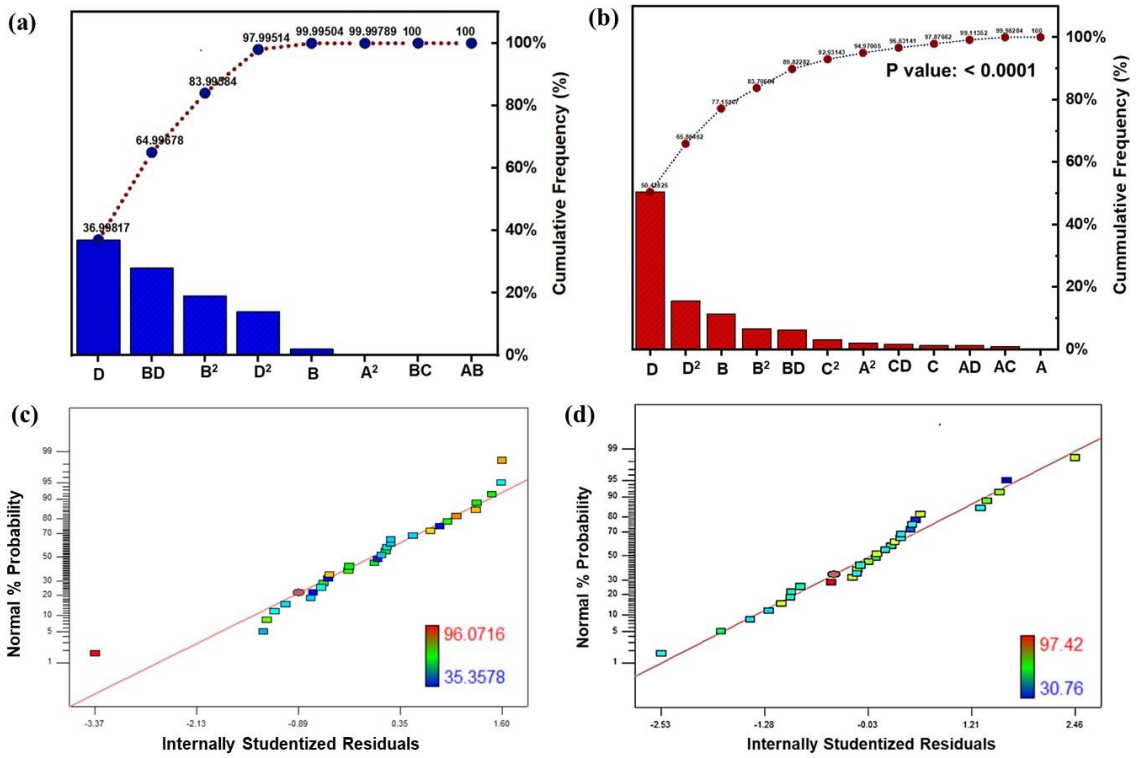


Fig. 4.7. *Pareto graphic analysis and normal probability plot described by the models for S. spontaneum cellulose (a, c) and banana peduncle (b, d), respectively.*

4.3.2. Optimization of enzymatic hydrolysis using ANN

The ANN model for predicting glucose yield was built using the dataset derived from the RSM model. Table 4.3 and Table 4.5 shows the expected glucose yield from *S. spontaneum* cellulose and banana peduncle cellulose, respectively for all 30 experiments using ANN. The experimental data was trained using the LM algorithm applying second-degree derivatives of mean squared error (MSE). One of the most important exercises in developing an ANN is the standardization of the number of neurons and hidden layers of the neural network [52]. In this study, the simulation of various topologies with different feed-forward networks was used to evaluate the determination of suitable ANN

architecture based on minimizing the mean square error (MSE), different algorithms for different hidden layers, and the number of neurons. Based on these criteria, the network was optimized by varying the number of neurons in the hidden neuronal layer from 5–100 with a minimum of 10 training runs for each ANN topology [53]. An optimal ANN topology was recognized for both *S. spontaneum* and banana peduncle cellulose using four neurons in the input layer, twenty neurons in the hidden layer, and one neuron in the output layer (4-10-1) based on minimizing average MSE values in a total of 2850 (95 × 30) experimental set-ups. The notably high R^2 value of 0.99 indicates that 99% of the total variations in the experimental parameters could be explained by these ANN predictive models in both cases (Fig. 4.8 a and b, Table 4.4 and 4.6). Furthermore, the ANN topologies of 4-10-1 showed the lowest MSE values of 0.007 for *S. spontaneum* cellulose (Fig. 4.8 a) and 7.8023 for banana peduncle cellulose (Fig. 4.8 b) and were used for the predictive modeling of enzymatic hydrolysis. The regression plots of the developed models are also shown in Fig. 4.8 a and Fig. 4.8 b. The association attained were 0.9976, 0.9999, 0.9979, and 0.9975 for training, testing, validation, and overall data respectively for *S. spontaneum* cellulose. While for banana peduncle glucose these values are 0.9999, 0.9937, 0.9969, and 0.9962 for training, testing, validation, and overall data correspondingly. In all data sets, the correlation coefficients were near one, representing that the fit was suitable for all data sets. Additionally, the fit line for the training and overall data sets, where the objectives were nearly equivalent to the network outputs, lay on the 45-degree line. As a result, the glucose yield prediction accuracy of the ANN output network response was satisfactory [26]. The complete ANN predicted glucose yield based on each experimental run and the assessment with the investigational data set was depicted in Table 4.3 and Table 4.5.

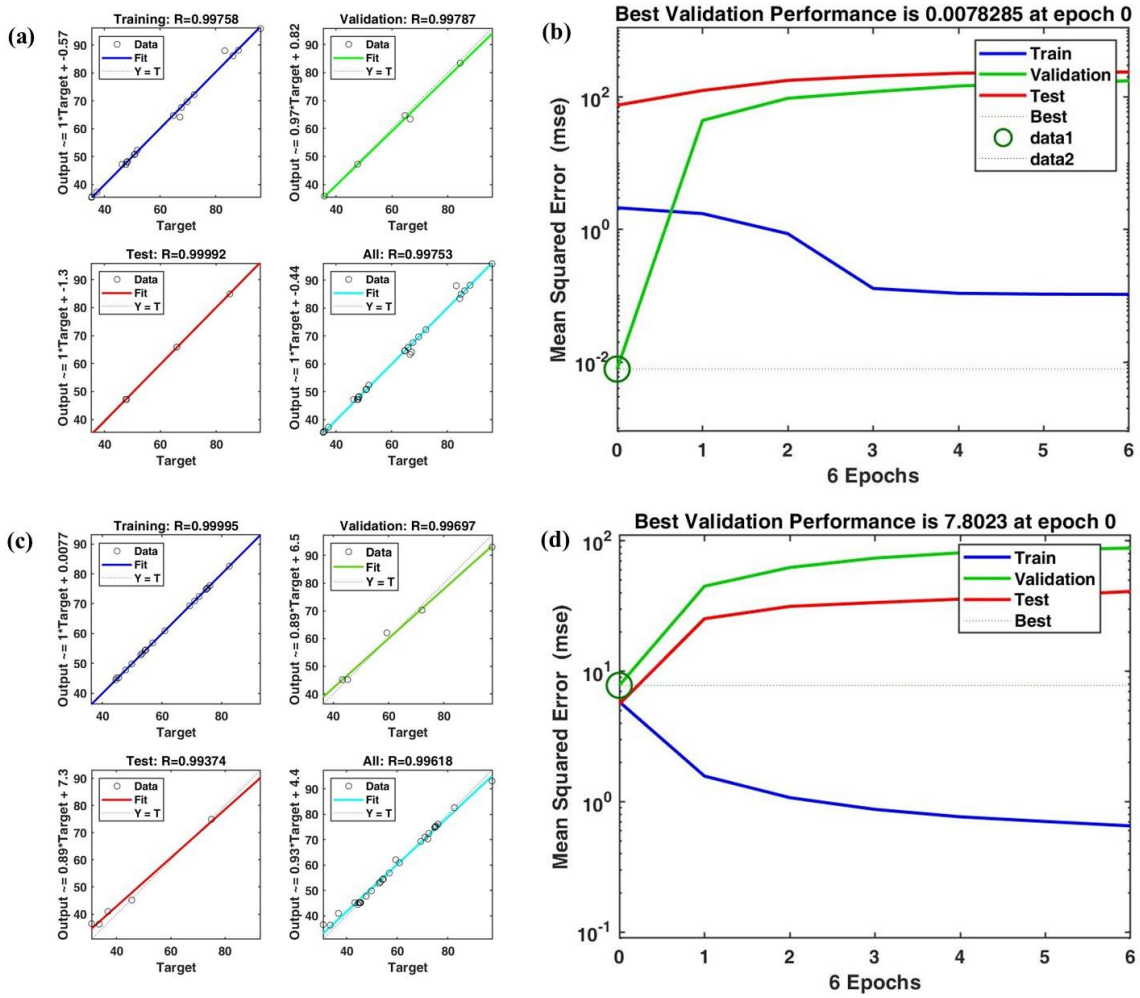


Fig. 4.8. *Selection of the best ANN architecture based on the determination coefficients (R^2) for S. spontaneum cellulose (a), banana peduncle cellulose (c) and mean squared error (MSE) values for S. spontaneum cellulose (b), banana peduncle cellulose (d) for the training and validation of the optimum ANN topology.*

4.3.3. Optimization of enzymatic hydrolysis using ANFIS

The datasets derived from the RSM models have been used to construct the ANFIS models. Fig. 4.9 a and Fig. 4.9 b show the plots of Gaussian membership functions (MF) for the four input variables (enzyme loading, substrate concentration, surfactant concentration, and, incubation time) for the *S. spontaneum* cellulose model and banana peduncle cellulose model, respectively. The predicted results of glucose yield given by both the ANFIS models are presented in Table 4.3 and Table 4.5. The calculated R, $R^{2,}$ and adjusted R^2 of the ANFIS model for *S. spontaneum* were 0.9976, 0.9989, and 0.9928, respectively and for banana peduncle these were 0.9969, 0.9991, and 0.9925, respectively (Table 4.7). The closeness of R to unity indicates that the experimental and predicted values were in good

agreement. Furthermore, the R^2 result implies that the model can explain 99.34% (in the case of *S. spontaneum*) and 99.31% (in the case of banana peduncle) of the difference in experimental and predicted values. The high R^2 value also indicates that the model is fitted well [54].

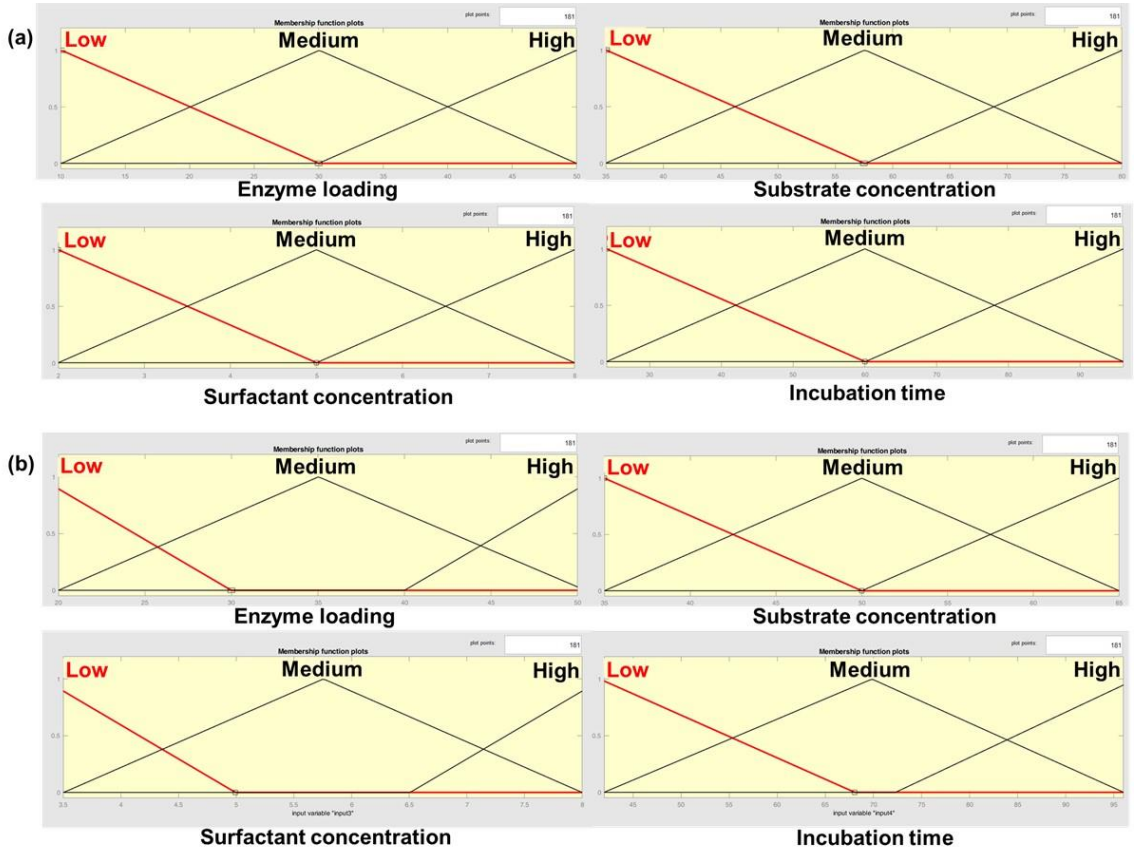


Fig. 4.9. *Membership function plots of input variables for S. spontaneum cellulose model (a) and banana peduncle cellulose model (b).*

Saccharum spontaneum cellulose

The interactions among the four process parameters investigated for the glucose yield from *S. spontaneum* cellulose were examined using 3D surface viewer plots. The relationship between substrate concentration and enzyme loading on the glucose yield is depicted in Fig. 4.10 a. The figure shows that the interaction between the two parameters significantly affected the glucose yield. As the substrate concentration increases, the glucose yield decreases attaining minimum and then further increases with further increase in substrate concentration. On the other hand, increasing the enzyme loading considerably increases the glucose yield to attain a maximum and then decreases with further increase in enzyme loading. The lowest value is observed at an enzyme loading of 10 FPU/g and 40 mg/ml of

substrate concentration. Both parameters are known to have a strong influence on the glucose yield. The 3D surface plot of substrate concentration and surfactant concentration together with the glucose yield is presented in Fig. 4.10 b. The highest glucose yield was also observed at the moderate surfactant concentration and highest substrate concentration. Likewise, Fig. 4.10 c shows the 3D surface plot of incubation time and substrate concentration with the glucose yield. The surface plot suggests that the glucose value was highest at the highest incubation time and moderate substrate concentration. The highest value recorded at 96 h incubation time and 50 mg/ml substrate concentration was 96.07% and the lowest value was 33.36% at 42 h incubation time and 65 mg/ml substrate concentration. The rule viewer plot (Fig. 4.11 a) provides a set of 31 values of responses out of 81 total rules by varying the process parameters.

Banana peduncle cellulose

The interactions among the four process parameters investigated for the glucose yield from banana peduncle cellulose were also examined using 3D surface plots. The relationship between enzyme loading and surfactant concentration on the glucose yield is depicted in Fig. 4.10 d. The graph demonstrates that the two parameters interacted and had a considerable impact on the glucose value. The glucose value increases with an increase in enzyme loading and surfactant concentration and attains a maximum. As both the enzyme and surfactant loading increase, the glucose yield decreases. The glucose yield decreased from over 70% to less than 10%. The lowest glucose yield was observed at 20 FPU/g enzyme loading and 2 mg/mL surfactant concentration. Fig. 4.10 e shows the interaction plot between enzyme loading and incubation time on the hydrolysis reaction. The glucose yield is highest at the highest incubation time and a moderate enzyme loading. The plot suggests that a further increase in enzyme loading with incubation time gradually decreases the glucose yield, which also corroborates with the RSM results. The highest glucose yield (97%) was recorded at an enzyme loading of 30 FPU/g and 96 h incubation time. The 3D surface plot of substrate concentration and incubation time together with glucose yield is presented in Fig. 4.10 f. The surface plot is comparable to Fig. 4.10 c and suggests that glucose yield increases with an increase in incubation time and substrate concentration, accomplished the highest glucose value, and then decreases with a further increase in substrate concentration (65 mg/ml). Lastly, the 3D interaction plot between the surfactant concentration and incubation time (Fig. 4.10 g) is quite similar to that of Fig.

4.10 e. The glucose yield is highest at the highest incubation time (96 h) and has a moderate surfactant concentration. The plot suggests that a further increase in surfactant concentration with incubation time gradually decreases the glucose yield. Overall, both the ANFIS models for *S. spontaneum* cellulose and banana peduncle cellulose suggest incubation time to be the most influential parameter and is in accordance with the findings (Table 4.3 and Table 4.5). The rule viewer plot (Fig. 4.11 b) provides a set of 31 values of responses out of 81 total rules by varying the process parameters.

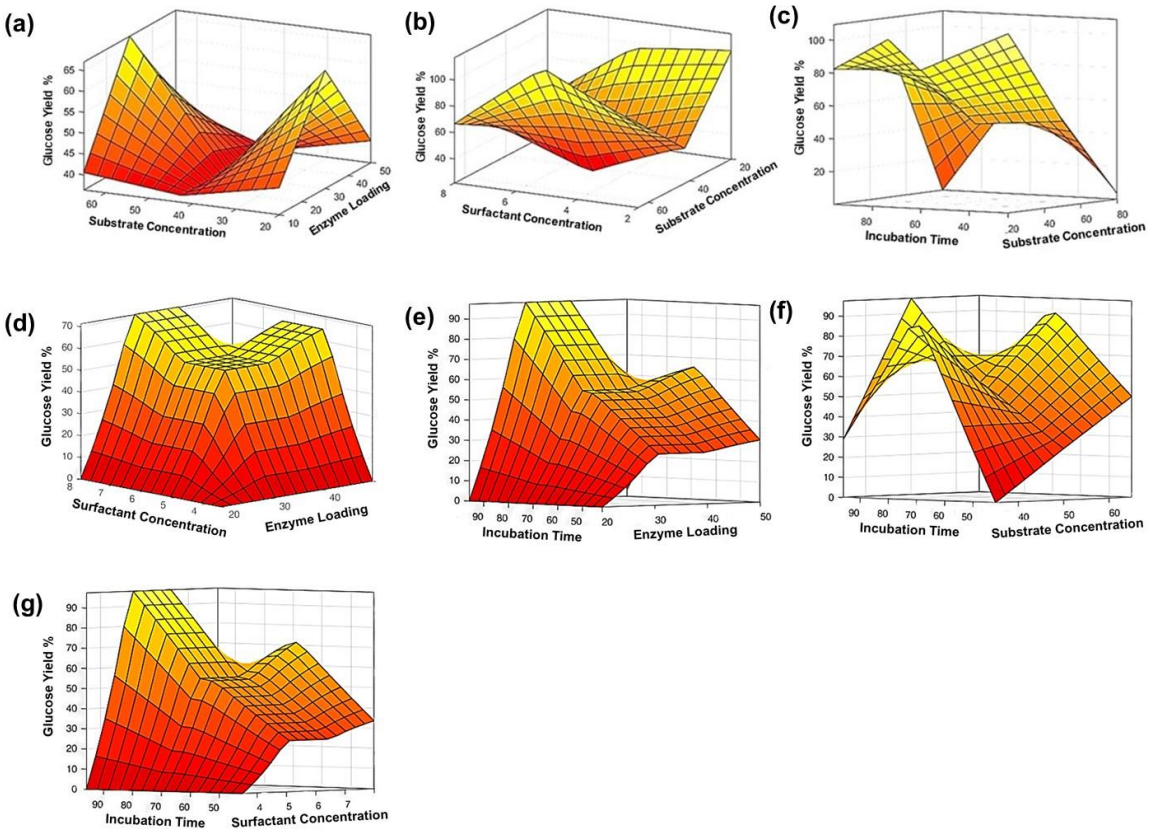


Fig. 4.10. *3D surface plots described by the ANFIS models for S. spontaneum cellulose (a-c) and banana peduncle cellulose (d-g) on the glucose yield.*

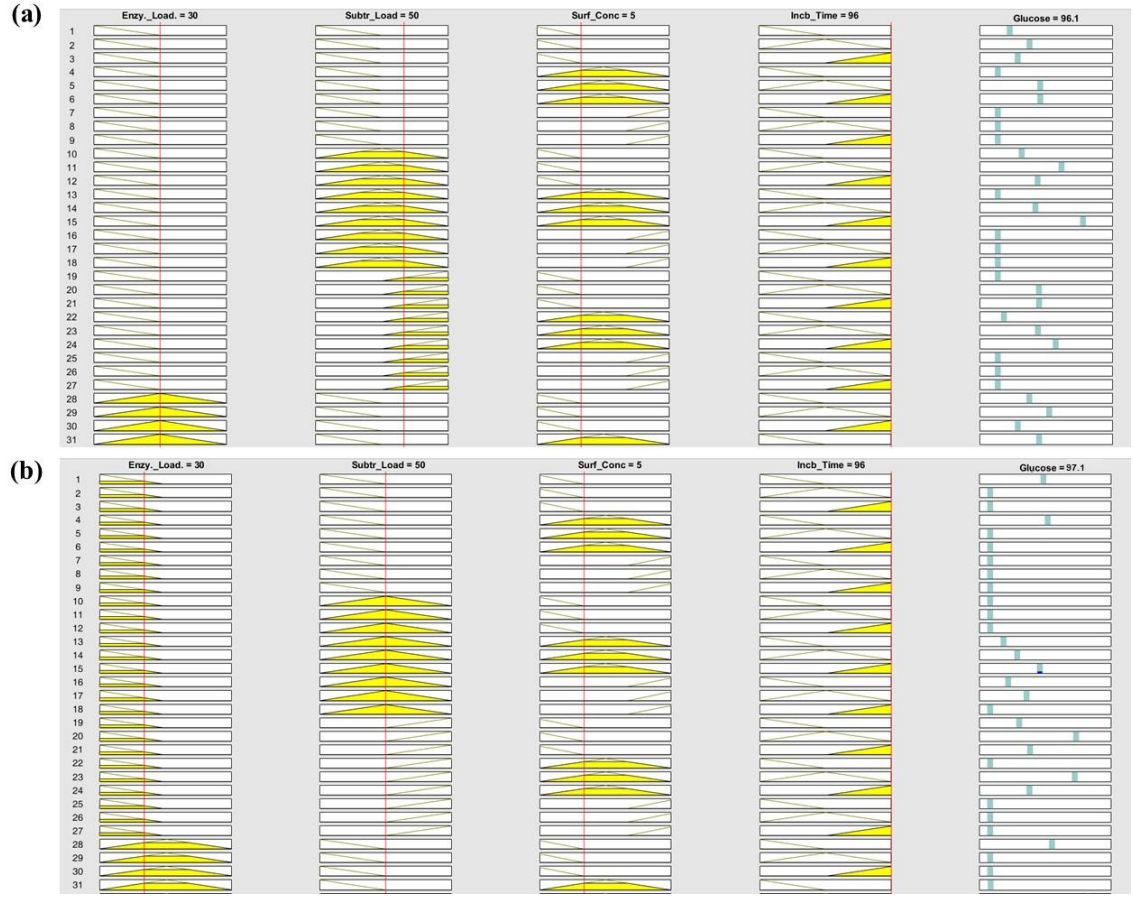


Fig. 4.11. *ANFIS rule viewer for the effect of process variables on responses for enzymatic hydrolysis of S. spontaneum (a) and banana peduncle (b).*

4.3.4. Performance assessment of the predictive ability of RSM, ANN, and, ANFIS models

Table 4.3 and Table 4.5 shows the predictive outputs of the three models, as well as the experimental findings for *S. spontaneum* cellulose and banana peduncle cellulose, respectively. The outputs of the three models were evaluated using linear correlation plots (Fig. 4.12). Usually, the three models were effective and near precise in predicting the glucose yield outputs. The results exhibited that the experimental value and the RSM, ANN, and ANFIS model estimates were quite close, resulting in low residual values. However, based on many insignificant standards in their residues, the ANFIS model appears to be more appropriate for estimating the glucose yield. The comparative plots in Fig. 4.12 showed a graphical association between investigational value and model calculations by RSM, ANN, and ANFIS models.

In addition, seven statistical error functions were applied to the model predictions to investigate further the model precision abilities, as shown in Table 4.7. RMSE, HYBRID, ARE, AARE, and MPSED error functions were assessed for each model. The model's ability to anticipate was demonstrated by the error functions' low values. The result showed that all three models have insignificant error values. Besides, R^2 and adjusted R^2 were also assessed. Adjusted-R^2 was used to investigate the overestimation of the R^2, and the values attained for both models were adequately high, demonstrating their significance. The more precise the model predictions, the higher the R^2 and adjusted R^2 values.

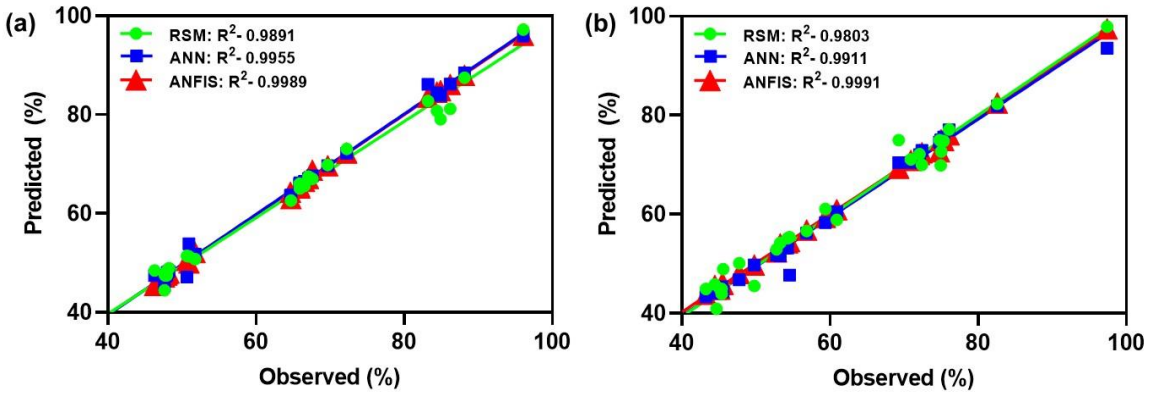


Fig. 4.12. *Comparison of predictive outputs of RSM, ANN, and ANFIS design matrices based on determination coefficients for S. spontaneum model (a) and banana peduncle model (b).*

Although it has been demonstrated that ANN outperforms RSM in terms of prediction capability [25], studies on the performance of ANFIS and ANN have been conflicting. This is because when one study found ANN to be more accurate than ANFIS [55] in terms of prediction capability, another study reported ANFIS to be more accurate than ANN [56]. On the other hand, Karimi et al. found that both ANN and ANFIS produced comparable results in their study [37]. In the current study, both ANN and ANFIS were found to be superior to RSM in terms of prediction accuracy and precision, however, though the results from the two models were similar, ANFIS was only marginally better than ANN.

Table 4.7. *Statistical error indices of RSM, ANN, and ANFIS.*

Error function	Results					
	S. spontaneum			**Banana peduncle**		
	RSM	**ANN**	**ANFIS**	**RSM**	**ANN**	**ANFIS**
RMSE	0.018	0.0089	0.0082	0.0146	0.0085	0.0076
HYBRID (%)	0.654	2.691	2.75	0.456	1.986	2.034
ARE (%)	0.401	1.23	1.04	0.5	1.026	1.001
AARE	0.0044	0.0189	0.021	0.0064	0.0184	0.0181
MPSED (%)	0.8886	2.0114	2.1503	0.9857	2.009	2.1212
R	0.9993	0.9969	0.9976	0.9939	0.9974	0.9969
R^2	0.9891	0.9955	0.9989	0.9803	0.9911	0.9991
Adj R^2	0.9816	0.9918	0.9928	0.9789	0.9894	0.9925

4.3.5. Experimental validation of RSM, ANN, and, ANFIS optimal predictive outputs

The optimal predictive outputs of RSM, ANN, and ANFIS were further validated experimentally by performing the enzymatic hydrolysis reactions at optimized conditions (Experimental run 16 of Table 4.3 and Table 4.5). The experimental results showed 96.66%, 96.43%, and 96.96% in the case of *S. spontaneum* cellulose and 97.32%, 96.94%, and 97.10% of glucose yield in the case of banana peduncle cellulose. The findings are close to our predictive results. The glucose peaks in each condition are shown in the HPLC chromatogram (Fig. 4.13). A comparative table of validation of optimal conditions predicted by RSM and ANN is shown in Table 4.8.

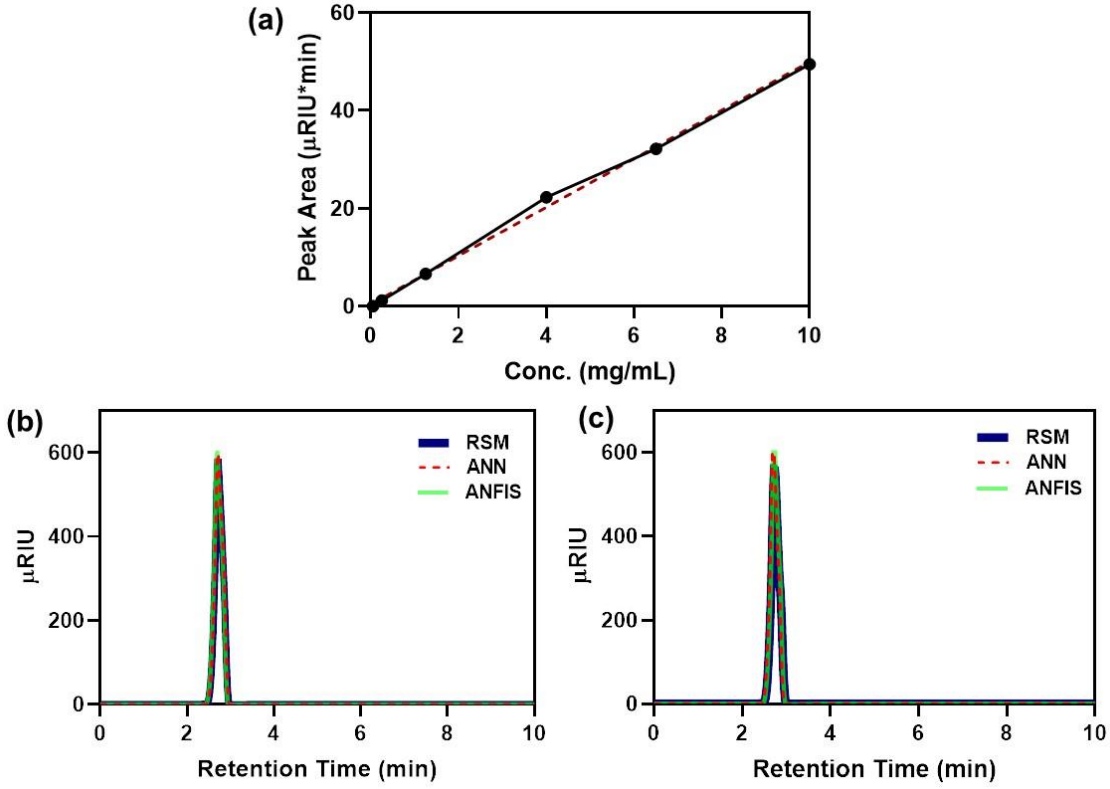


Fig. 4.13. *Glucose standard curve of HPLC (a); HPLC chromatograms showing the glucose peaks at optimal conditions of RSM, ANN, and ANFIS obtained from S. spontaneum cellulose (b) and banana peduncle cellulose (c).*

Table 4.8. *Table of validation of optimal conditions predicted by RSM, ANN, and ANFIS.*

Source	Optimized Condition	Enzyme Loading (FPU/g)	Substrate Loading (mg/mL)	Surfactant Concentration (mg/mL)	Incubation Time (h)	Glucose Yield (Predicted)	Glucose Yield* (Observed)
***S. spontaneum* cellulose**	**RSM**	30	50	5	96	97.21	96.66 ± 1.12
	ANN	30	50	5	96	95.93	96.43 ± 0.95
	ANFIS	30	50	5	96	96.1	96.96 ± 1.04
Banana peduncle cellulose	**RSM**	30	50	5	96	97.87	97.32 ± 0.56
	ANN	30	50	5	96	93.51	96.94 ± 1.89
	ANFIS	30	50	5	96	97.42	97.10 ± 1.19

* Values are Mean ± SD of triplicate analysis.

4.3.6. Comparison between the glucose yield obtained from this work to that of previous work

A comparative table (Table 4.9), exhibiting the glucose yield obtained from different lignocellulosic sources using different pretreatment approaches is shown below. It is pertinent to note from these studies that the pretreatment procedures of the biomass play a crucial role in the enhancement of glucose production. In this context, a high yield of glucose by the hydrolysis of cellulose isolated from banana peduncle using an integrated methodology of pretreatment has been encouraging. It is apparent from Table 4.9, that feedstocks like banana peduncle and corncob having a higher amorphous cellulose content but a low overall CrI, post pretreatments would promote higher enzymatic hydrolysis mediated glucose production, compared to others [57]. In this context, the removal of the non-cellulosic entities and pre-treatment mediated size reduction promote the higher enzymatic hydrolysis yields.

Table 4.9. *An assessment of diverse pretreatment and enzymatic hydrolysis procedures to obtain glucose with different yields from reported studies.*

Lignocellulosic sources	Pre-treatment techniques	Enzymes used	Glucose yield (%)	References
Saccharum spontaneum	Integrated approach	Cellulase (*Aspergillus niger*)	96	This study
Banana peduncle	Integrated approach	Cellulase (*Aspergillus niger*)	97	This study
Corncob	Fenton	Cellulase (*Trichoderma reesei*)	92	[58]
Alfalfa	CO_2 explosion	Not mentioned	75	[59]
Biomass	H_2SO_4 (above 30%) hydrolysis	Not mentioned	70	[60]
Sugarcane bagasse	CO_2 explosion	Cellulase	72.6	[61]
Cornstalk	Electron beam irradiation	Cellulase	43	[62]
Wheat straw	Combined wet oxidation and alkaline hydrolysis	Cellulase (*Aspergillus niger*)	85	[63]
Eucalyptus urophylla	Hydrothermal and alkali fractionation	Cellulase and β-Glucosidase	66.3	[64]

4.4. Conclusion

The current chapter describes the enzymatic hydrolysis of *Saccharum spontaneum* and banana peduncle-derived celluloses for the production of glucose. To maximize the glucose yield values, the hydrolysis process was optimized utilizing three optimization techniques, namely, RSM, ANN, and ANFIS, and their prediction efficacies were compared. The best conditions were anticipated by all three models with a combination of process parameters such as enzyme loading 30 FPU/g, substrate concentration 50 mg/mL, surfactant concentration 5 mg/mL, and incubation time 96 h. The optimal condition led to a glucose turnover of > 96% from both *Saccharum spontaneum* and banana peduncle-derived celluloses. According to the seven statistical error indices, the ANN and ANFIS models for glucose yield prediction had a better performance when compared with the RSM model, however, ANFIS was marginally better than ANN. This is the first sequential study of an economical and environmentally sustainable method for enhanced enzymatic hydrolysis for glucose production from the lignocellulosic extracted celluloses (as discussed in Chapter 3). The enzymatic approach is efficient, practical and economical, and environmentally benign, streamlining the transformation of lignocellulosic biomass to glucose.

References

[1] Li, X., Xu, Q., Shen, H., Guo, Y., Wu, M., Peng, Y., Zhang, L., Zhao, Z. K., Liu, Y., Xie, H. Capturing CO_2 to reversible ionic liquids for dissolution pretreatment of cellulose towards enhanced enzymatic hydrolysis. *Carbohydrate Polymers*, 204:50–58, 2019.

[2] He, F., Chen, J., Gong, Z., Xu, Q., Yue, W., and Xie, H. Dissolution pretreatment of cellulose by using levulinic acid-based protic ionic liquids towards enhanced enzymatic hydrolysis. *Carbohydrate Polymers*, 269:118271, 2021.

[3] Li, W., Zhu, Y., Lu, Y., Liu, Q., Guan, S., Chang, H. M., Jameel, H., and Ma, L. Enhanced furfural production from raw corn stover employing a novel heterogeneous acid catalyst. *Bioresource Technology*, 245:258-265, 2017.

[4] Yu, P. R., Hung, W. C., and Wan, H. P. LiCl/HCl ionic solution for efficient conversion of lignocellulose into glucose under mild conditions. *Journal of the Taiwan Institute of Chemical Engineers*, 93:193-200, 2018.

[5] Wang, H., Zhu, C., Li, D., Liu, Q., Tan, J., Wang, C., Cai, C., and Ma, L. Recent advances in catalytic conversion of biomass to 5-hydroxymethylfurfural and 2, 5-dimethylfuran. *Renewable and Sustainable Energy Reviews*, 103:227-247, 2019.

[6] Suganuma, S., Nakajima, K., Kitano, M., Yamaguchi, D., Kato, H., Hayashi, S., and Hara, M. Hydrolysis of cellulose by amorphous carbon bearing SO_3H, COOH, and OH groups. *Journal of the American Chemical Society*, 130(38):12787-12793, 2008.

[7] Tursi, A. A review on biomass: importance, chemistry, classification, and conversion. *Biofuel Research Journal*, 6(2):962, 2019.

[8] Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M., and Ladisch, M. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*, 96(6):673-686, 2005.

[9] Wyman, C. E. Biomass ethanol: technical progress, opportunities, and commercial challenges. *Annual Review of Energy and the Environment*, 24(1):189-226, 1999.

[10] Sun, Y., and Cheng, J. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology*, 83(1):1-11, 2002.

[11] Alkasrawi, M., Eriksson, T., Börjesson, J., Wingren, A., Galbe, M., Tjerneld, F., and Zacchi, G. The effect of Tween-20 on simultaneous saccharification and fermentation of softwood to ethanol. *Enzyme and Microbial Technology*, 33(1):71-78, 2003.

[12] Eriksson, T., Börjesson, J., and Tjerneld, F. Mechanism of surfactant effect in enzymatic hydrolysis of lignocellulose. *Enzyme and Microbial Technology*, 31(3):353-364, 2002.

[13] Olusola, A., and Okewale Akindele, O. A. Comparative studies of response surface methodology (RSM) and artificial neural network (ANN) predictive capabilities on enzymatic hydrolysis optimization of sweet potato starch. *International Journal*, 2(10):849-860, 2014.

[14] Mäkelä, M. Experimental design and response surface methodology in energy applications: A tutorial review. *Energy Conversion and Management*, 151:630-640, 2017.

[15] Guan, X., and Yao, H. Optimization of Viscozyme L-assisted extraction of oat bran protein using response surface methodology. *Food Chemistry*, 106(1):345-351, 2008.

[16] Guo, R., Zheng, X., Wang, Y., Yang, Y., Ma, Y., Zou, D., and Liu, Y. Optimization of cellulase immobilization with sodium alginate-polyethylene for enhancement of enzymatic hydrolysis of microcrystalline cellulose using response surface methodology. *Applied Biochemistry and Biotechnology*, 193(7):2043-2060, 2021.

[17] Qi, B., Chen, X., Shen, F., Su, Y., and Wan, Y. Optimization of enzymatic hydrolysis of wheat straw pretreated by alkaline peroxide using response surface methodology. *Industrial & Engineering Chemistry Research*, 48(15):7346-7353, 2009.

[18] Kshirsagar, S. D., Waghmare, P. R., Loni, P. C., Patil, S. A., and Govindwar, S. P. Dilute acid pretreatment of rice straw, structural characterization and optimization of enzymatic hydrolysis conditions by response surface methodology. *RSC Advances*, 5(58):46525-46533, 2015.

[19] Donkoh, E., Degenstein, J., Tucker, M., and Ji, Y. Optimization of enzymatic hydrolysis of dilute acid pretreated sugar beet pulp using response surface design. *Journal of Sugar Beet Research*, 49(1):26, 2012.

[20] Jaya, E. M. J., Norhalim, N. A., and Ahmad, Z. Artificial neural network model prediction of glucose by enzymatic hydrolysis of rice straw. *Journal of Engineering Science*, 10:85, 2014.

[21] Suresh, T., Sivarajasekar, N., Balasubramani, K., Ahamad, T., Alam, M., and Naushad, M. Process intensification and comparison of bioethanol production from food industry waste (potatoes) by ultrasonic assisted acid hydrolysis and enzymatic

hydrolysis: Statistical modelling and optimization. *Biomass and Bioenergy*, 142:105752, 2020.

[22] Khamparia, A., Pandey, B., Pandey, D. K., Gupta, D., Khanna, A., and de Albuquerque, V. H. C. Comparison of RSM, ANN and Fuzzy Logic for extraction of *Oleonolic Acid* from *Ocimum sanctum*. *Computers in Industry*, 117:103200, 2020.

[23] Jang, J. S. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665-685, 1993.

[24] Sonmez, A. Y., Kale, S., Ozdemir, R. C., and Kadak, A. E. An adaptive neuro-fuzzy inference system (ANFIS) to predict of cadmium (Cd) concentrations in the Filyos River, Turkey. *Turkish Journal of Fisheries and Aquatic Sciences*, 18(12):1333-1343, 2018.

[25] Betiku, E., Odude, V. O., Ishola, N. B., Bamimore, A., Osunleke, A. S., and Okeleye, A. A. Predictive capability evaluation of RSM, ANFIS and ANN: a case of reduction of high free fatty acid of palm kernel oil via esterification process. *Energy Conversion and Management*, 124:219-230, 2016.

[26] Onu, C. E., Nwabanne, J. T., Ohale, P. E., and Asadu, C. O. Comparative analysis of RSM, ANN and ANFIS and the mechanistic modeling in eriochrome black-T dye adsorption using modified clay. *South African Journal of Chemical Engineering*, 36:24-42, 2021.

[27] Rego, A. S., Valim, I. C., Vieira, A. A., Vilani, C., and Santos, B. F. Optimization of sugarcane bagasse pretreatment using alkaline hydrogen peroxide through ANN and ANFIS modelling. *Bioresource Technology*, 267:634-641, 2018.

[28] Lerkkasemsan, N. Fuzzy logic-based predictive model for biomass pyrolysis. *Applied Energy*, 185:1019-1030, 2017.

[29] Baruah, J., Bardhan, P., Mukherjee, A. K., Deka, R. C., Mandal, M., and Kalita, E. Integrated pretreatment of banana agrowastes: Structural characterization and enhancement of enzymatic hydrolysis of cellulose obtained from banana peduncle. *International Journal of Biological Macromolecules*, 201:298-307, 2022.

[30] Astray, G., Gullón, B., Labidi, J., and Gullón, P. Comparison between developed models using response surface methodology (RSM) and artificial neural networks (ANNs) with the purpose to optimize oligosaccharide mixtures production from sugar beet pulp. *Industrial Crops and Products*, 92:290-299, 2016.

[31] Kumar, V., Kumar, A., Chhabra, D., and Shukla, P. Improved biobleaching of mixed hardwood pulp and process optimization using novel GA-ANN and GA-ANFIS hybrid statistical tools. *Bioresource Technology*, 271:274-282, 2019.

[32] Mostafaei, M., Javadikia, H., and Naderloo, L. Modeling the effects of ultrasound power and reactor dimension on the biodiesel production yield: Comparison of prediction abilities between response surface methodology (RSM) and adaptive neuro-fuzzy inference system (ANFIS). *Energy*, 115:626-636, 2016.

[33] Ganesan, V., Gurumani, V., Kunjiappan, S., Panneerselvam, T., Somasundaram, B., Kannan, S., Chowdhury, A., Saravanan, G., and Bhattacharjee, C. Optimization and analysis of microwave-assisted extraction of bioactive compounds from *Mimosa pudica* L. using RSM & ANFIS modeling. *Journal of Food Measurement and Characterization*, 12(1):228-242, 2018.

[34] Kim, B., and Park, J. H. Qualitative fuzzy logic model of plasma etching process. *IEEE Transactions on Plasma Science*, 30(2):673-678, 2002.

[35] Taheri, M., Moghaddam, M. A., and Arami, M. Techno-economical optimization of Reactive Blue 19 removal by combined electrocoagulation/coagulation process through MOPSO using RSM and ANFIS models. *Journal of Environmental Management*, 128:798-806, 2013.

[36] Das, A., Maiti, J., and Banerjee, R. N. Process control strategies for a steel making furnace using ANN with bayesian regularization and ANFIS. *Expert Systems with Applications*, 37(2):1075-1085, 2010.

[37] Karimi, S., Kisi, O., Shiri, J., and Makarynskyy, O. Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin Harbor, Australia. *Computers & Geosciences*, 52:50-59, 2013.

[38] Baruah, J., Deka, R. C., and Kalita, E. Greener production of microcrystalline cellulose (MCC) from *Saccharum spontaneum* (Kans grass): Statistical optimization. *International Journal of Biological Macromolecules*, 154:672-682, 2020.

[39] Adney, B., and Baker, J. Measurement of Cellulase Activities. Technical Report NREL/TP-510-42628, National Renewable Energy Laboratory, Golden, CO, 2008.

[40] Rai, S. K., and Mukherjee, A. K. Statistical optimization of production, purification and industrial application of a laundry detergent and organic solvent-stable subtilisin-

like serine protease (Alzwiprase) from *Bacillus subtilis* DM-04. *Biochemical Engineering Journal*, 48:173-180, 2010.

[41] Roy, J. K., Rai, S. K., and Mukherjee, A. K. Characterization and application of a detergent-stable alkaline α-amylase from *Bacillus subtilis* strain AS-S01a. *International Journal of Biological Macromolecules*, 50:219-229, 2012.

[42] Okoye, C. C., Onukwuli, O. D., and Okey-Onyesolu, C. F. Predictive capability evaluation of RSM and ANN models in adsorptive treatment of crystal violet dye simulated wastewater using activated carbon prepared from *Raphia hookeri* seeds. *Journal of the Chinese Advanced Materials Society*, 6(4):478-496, 2018.

[43] Verma, M. Medical diagnosis using backpropagation algorithm in ANN. *International Journal of Science, Engineering and Technology Research (IJSETR)*, 3(1):94-99, 2014.

[44] del Cerro, R. T. G., Subathra, M. S. P., Kumar, N. M., Verrastro, S., and George, S. T. Modelling the daily reference evapotranspiration in semi-arid region of South India: a case study comparing ANFIS and empirical models. *Information Processing in Agriculture*, 8(1):173-184, 2021.

[45] Zaghloul, M. S., Hamza, R. A., Iorhemen, O. T., and Tay, J. H. Comparison of adaptive neuro-fuzzy inference systems (ANFIS) and support vector regression (SVR) for data-driven modelling of aerobic granular sludge reactors. *Journal of Environmental Chemical Engineering*, 8(3):103742, 2020.

[46] Abdel-Halim, E. Chemical modification of cellulose extracted from sugarcane bagasse: Preparation of hydroxyethyl cellulose. *Arabian Journal of Chemistry*, 7:362-371, 2014.

[47] Wakkel, M., Khiari, B., and Zagrouba, F. Textile wastewater treatment by agro-industrial waste: equilibrium modelling, thermodynamics and mass transfer mechanisms of cationic dyes adsorption onto low-cost lignocellulosic adsorbent. *Journal of the Taiwan Institute of Chemical Engineers*, 96:439-452, 2019.

[48] Morais, J. P. S., de Freitas Rosa, M., Nascimento, L. D., do Nascimento, D. M., and Cassales, A. R. Extraction and characterization of nanocellulose structures from raw cotton linter. *Carbohydrate Polymers*, 91(1):229-235, 2013.

[49] Fang, H., Dong, H., Cai, T., Zheng, P., Li, H., Zhang, D., and Sun, J. In vitro optimization of enzymes involved in precorrin-2 synthesis using response surface methodology. *PloS one*, 11(3):e0151149, 2016.

[50] Zhang, Q. Y., Zhou, W. W., Zhou, Y., Wang, X. F., and Xu, J. F. Response surface methodology to design a selective co-enrichment broth of *Escherichia coli*, Salmonella spp. and Staphylococcus aureus for simultaneous detection by multiplex PCR. *Microbiological Research*, 167(7):405-412, 2012.

[51] Shang, H., Zhou, H., Duan, M., Li, R., Wu, H., and Lou, Y. Extraction condition optimization and effects of drying methods on physicochemical properties and antioxidant activities of polysaccharides from comfrey (*Symphytum officinale* L.) root. *International Journal of Biological Macromolecules*, 112:889-899, 2018.

[52] Badkar, D. S., Pandey, K. S., and Buvanashekaran, G. Development of RSM-and ANN-based models to predict and analyze the effects of process parameters of laser-hardened commercially pure titanium on heat input and tensile strength. *The International Journal of Advanced Manufacturing Technology*, 65(9):1319-1338, 2013.

[53] Nath, B. K., Chaliha, C., and Kalita, E. Iron oxide Permeated Mesoporous rice-husk nanobiochar (IPMN) mediated removal of dissolved arsenic (As): Chemometric modelling and adsorption dynamics. *Journal of Environmental Management*, 246:397-409, 2019.

[54] Pongsumpun, P., Iwamoto, S., and Siripatrawan, U. Response surface methodology for optimization of cinnamon essential oil nanoemulsion with improved stability and antifungal activity. *Ultrasonics Sonochemistry*, 60:104604, 2020.

[55] Kiran, T. R., and Rajput, S. P. S. An effectiveness model for an indirect evaporative cooling (IEC) system: Comparison of artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS) and fuzzy inference system (FIS) approach. *Applied Soft Computing*, 11(4):3525-3533, 2011.

[56] Dastorani, M. T., Moghadamnia, A., Piri, J., and Rico-Ramirez, M. Application of ANN and ANFIS models for reconstructing missing flow data. *Environmental Monitoring and Assessment*, 166(1):421-434, 2010.

[57] Chávez-Guerrero, L., Silva-Mendoza, J., Toxqui-Terán, A., Vega-Becerra, O. E., Salinas-Montelongo, J. A., and Pérez-Camacho, O. Direct observation of endoglucanase fibrillation and rapid thickness identification of cellulose nanoplatelets using constructive interference. *Carbohydrate Polymers*, 254:117463, 2021.

[58] Yu, H. T., Chen, B. Y., Li, B. Y., Tseng, M. C., Han, C. C., and Shyu, S. G. Efficient pretreatment of lignocellulosic biomass with high recovery of solid lignin and

fermentable sugars using Fenton reaction in a mixed solvent. *Biotechnology for biofuels*, 11(1):1-11, 2018.

[59] Sankaran, R., Cruz, R. A. P., Pakalapati, H., Show, P. L., Ling, T. C., Chen, W. H., and Tao, Y. Recent advances in the pretreatment of microalgal and lignocellulosic biomass: A comprehensive review. *Bioresource Technology*, 298:122476, 2020.

[60] Kucharska, K., Rybarczyk, P., Hołowacz, I., Łukajtis, R., Glinka, M., and Kamiński, M. Pretreatment of lignocellulosic materials as substrates for fermentation processes. *Molecules*, 23(11):2937, 2018.

[61] Zheng, Y., Lin, H. M., and Tsao, G. T. Pretreatment for cellulose hydrolysis by carbon dioxide explosion. *Biotechnology Progress*, 14(6):890-896, 1998.

[62] Abou Elmaaty, T., Okubayashi, S., Elsisi, H., and Abouelenin, S. Electron beam irradiation treatment of textiles materials: a review. *Journal of Polymer Research*, 29(4):117, 2022.

[63] Bjerre, A. B., Olesen, A. B., Fernqvist, T., Plöger, A., and Schmidt, A. S. Pretreatment of wheat straw using combined wet oxidation and alkaline hydrolysis resulting in convertible cellulose and hemicellulose. *Biotechnology and Bioengineering*, 49(5):568-577, 1996.

[64] Sun, S., Cao, X., Sun, S., Xu, F., Song, X., Sun, R. C., and Jones, G. L. Improving the enzymatic hydrolysis of thermo-mechanical fiber from Eucalyptus urophylla by a combination of hydrothermal pretreatment and alkali fractionation. *Biotechnology for Biofuels*, 7:1-12, 2014.