NEURAL NETWORK MODELING AND OPTIMIZATION FOR ENZYMATIC HYDROLYSIS OF XYLOSE FROM RICE STRAW

NUR'ATIQAH BINTI NORHALIM

UNIVERSITI SAINS MALAYSIA

2015

NEURAL NETWORK MODELING AND OPTIMIZATION FOR ENZYMATIC HYDROLYSIS OF XYLOSE FROM RICE STRAW

by

NUR’ATIQAH BINTI NORHALIM

Thesis submitted in fulfilment of the
requirements for the degree of
Master of Science

FEBRUARY 2015

ACKNOWLEDGEMENTS

In the name of The Almighty ALLAH, The Most Gracious and Merciful, with His gracing and blessing has led to success be upon this thesis. All praise to Allah who give me the strength and patience to complete the research successfully. This research project completed with the support of many people. It is therefore a great pleasure to thank those who involved in this thesis.

First and foremost, I would like to offer a sincere thanks and utmost appreciation to my research supervisor, Assoc. Prof. Dr Zainal Ahmad for his guidance, knowledge sharing, encouragement, and undue support to my research project. My gratitude also goes to my co–supervisor, Prof. Dr Mashitah Mat Don for her meaningful advice and encouragement. This research publication would be impossible without their encouragement and advice. I really appreciate all the lessons, guidance and priceless knowledge given to me that I will treasure forever.

I am also very thankful to Universiti Sains Malaysia (USM) for providing good facilities in the campus. To all lecturers, staffs and technicians in the school of Chemical Engineering, a very big thanks to you for all your support and coorperation. A special thanks also goes to Nazira Annisa, Senthil Kumar, Dipesh, and Kiew Peck Loo for willing to share their knowledge in order to assist me during my research study. I am indebted to our group member of Process Control Research Group (PCRG) especially Kak Assyura, Amirah, Hazwani and Ain for their kindness and willingness to spent time for distributing their knowledge, opinion, support and also encouragement to me to complete my research.

My sincere appreciation also extends to all my fellow friends and colleagues who have provided assistance and guidance in preparing this thesis. To my beloved friends, Fatihah, Nadia, Kak Nadzirah, Usman, Farah and not forgetting my sisters for always supporting me through thick and thin. Their unique views and tips are very helpful indeed. Thank you for the time sacrificed to accompany me.

Last but not least, my great appreciation belong to my lovely husband, Zulkarnain bin Mohammed Isa who always supports me all the time especially during my thesis writing. I would like to dedicate this work to our child, Manal Huda binti Zulkarnain and our expecting second child whom I dearly love. I am also grateful to all my family members especially to my parents, Norhalim bin Sirome and Kamariah bt Jaafar, my parent-in-law, Mohammed Isa bin Samsuri and Kurasi binti Sarkan, my siblings and to my relatives for their continuous support for me to be able to successfully accomplish my project.

Nur'atiqah Norhalim,

Campus of Engineering,

School of Chemical Engineering,

Universiti Sains Malaysia,

2014.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AChE	-	Acetylcholinesterase
ANN	-	Artificial neural network
ANOVA	-	Analysis of variance
BP	-	Backpropagation
CCD	-	Central composite design
DE	-	Dextrose equivlent
DNS	-	Dinitrosalicylic acid
DNSA	-	Dinitrosalicylic acid
DOE	-	Design of expert
DRNN	-	Diagonal recurrent neural network
FANN	-	Feed-forward artificial neural network
FEA	-	Finite element analysis
FRNN	-	Fully connected recurrent neural network
gBest	-	global neighborhood
GA	-	Genetic algorithm
IMC	-	Internal model control
lBest	-	local neighborhood
LHDs	-	Latin hypercube designs
LM	-	Levernberg-Marquardt
MLP	-	Multilayer perceptron
MPC	-	Model predictive control
MPX	-	Moldflow Plastic Xpert
MSE	-	Mean square error

OD	-	Optical density
OFAT	-	One factor at time method
PA	-	Polyamide
PID	-	Proportional integral derivative
PSO	-	Particle swarm optimization
RBF	-	Radial basis function
RCNN	-	Recurrent neural network
RNN	-	Recurrent neural network
RSM	-	Response surface methodology
RTNN	-	Recurrent trainable neural network
SSE	-	Sum square error
Std	-	Standard
rpm	-	Round per minute

LIST OF SYMBOLS

A	-	Temperature value during enzymatic hydrolysis
Ai	-	Regression coefficients of variables for linear
Aii	-	Regression coefficients of variables for quadratic
a_{ij}	-	Input
Aij	-	Regression coefficients of variables for interaction
B	-	Agitation speed of incubator shaker
b_j	-	Bias
C	-	Xylanase enzyme concentration.
c_1	-	Acceleration constants
c_2	-	Acceleration constants
c(i)	-	Objective function coefficient corresponding to the ith variable
i	-	Row
j	-	Layer
Ao	-	Regression coefficients of variables for intercept
n	-	Total number of experiments
p_i	-	Prediction value by the network
R	-	Correlation coefficient
R^2	-	Determination coefficient
r_1	-	uniform random real numbers
r_2	-	uniform random real numbers in the range[0,1]
t_i	-	Target output
w_{ij}	-	Weight

$\bar{X}$	-	Corresponds to the values of X_i.
X	-	Xylose concentration
X_i	-	the ith variable decision.
X_i	-	the calculated value of output variables
Xi	-	Independent variables for i
Xj	-	Independent variables for j
Y	-	Response variable
˚C	-	Degree Celsius
%	-	Percentage
±	-	Plus-minus
ω	-	Inertia factor

NEURAL NETWORK MODELING AND OPTIMIZATION FOR ENZYMATIC HYDROLYSIS OF XYLOSE FROM RICE STRAW

ABSTRACT

In this thesis, enzymatic hydrolysis was utilized in the production of xylose from rice straw. The process model was developed by the modeling techniques using feed-forward artificial neural network (FANN) and optimized using both particle swarm optimization (PSO) and genetic algorithm (GA). The parameters studied such as temperature, agitation speed and concentration of enzyme in the process were investigated in order to get an optimum yield of xylose during enzymatic hydrolysis process. Data collected from an experimental design using response surface methodology (RSM) were used to develop the FANN modeling. The data samples has been split into training, testing and validation data set before re-sampling with bootstrap re-sampling method. Then, the FANN model was used to predict the model performance with one hidden layer and the PSO and GA were used to predict the optimum conditions of the process. The number of nodes in the hidden layer obtained is six where the performance on the model is satisfactory with the architecture of FANN, 3-6-1. The correlation coefficient of training and testing set were indicated at 0.9970 and 0.9975 respectively though the correlation coefficient of validation obtained was 0.8501. The optimization of xylose production using the GA method obtained conditions of 50.3˚C, 154 rpm and 1.6944 g/l. The optimum xylose production was predicted as 0.1845 g/l at optimal condition obtained by using GA. Meanwhile with PSO, the optimum temperature observed was at 50 °C, 132

rpm for optimum value of agitation speed and 1.6474 g/l optimum xylanase concentration respectively. The optimal yield of xylose predicted was 0.1845 g/l using PSO for the enzymatic hydrolysis process. The laboratory experiment was carried out to validate the prediction of optimization result. It is shown from the experiment that the concentration of xylose obtained by using prediction optimum parameters for both PSO and GA are 0.2331 g/l and 0.2398 g/l respectively. The average error for the prediction and experimental values for the optimization are 29.97% and 26.34% for GA and PSO respectively. Therefore, the enzymatic hydrolysis on the production of xylose has been enhanced by predicting the optimum conditions utilizing the developed model that fits the experimental data.

PEMODELAN RANGKAIAN NEURAL DAN PENGOPTIMUMAN HIDROLISIS ENZIM DARIPADA XILOSA DARI JERAMI PADI

ABSTRAK

Tesis ini merangkumi kajian terhadap hidrolisis enzim yang telah digunakan dalam penghasilan xilosa daripada jerami padi. Proses hidrolisis enzim terhadap pengeluaran xilosa ini telah dikaji penghasilannya dengan menggunakan teknik permodelan seperti model suap-depan jaringan rangkaian neural (FANN) dan pengoptimuman proses dengan menggunakan kaedah perkumpulan zarah (PSO) dan algoritma genetik (GA). Parameter-parameter yang diuji ketika melakukan proses hidrolisis enzim adalah suhu, kelajuan pengadukan dan kepekatan enzim untuk mendapatkan hasil xilosa yang optima. Data-data yang diperolehi dari rekabentuk eksperimen yang dihasilkan dengan mengunakan kaedah permukaan sambutan (RSM) adalah diperlukan untuk membina model rangkaian neural. Sampel-sampel data telah dibahagikan kepada set latihan, set pengujian dan set pengesahan sebelum data di sampel semula dengan menggunakan kaedah ikat-but sampel semula. Lanjutan itu, model FANN telah digunakan untuk meramalkan prestasi model dengan hanya menggunakan satu lapisan terlindung dalam rangkaiannya manakala kedua-dua PSO dan GA telah digunakan untuk meramalkan keadaan proses yang optimum. Bilangan nod yang didapati dalam lapisan terselindung adalah sebanyak enam nod di mana ramalan keputusan prestasi adalah memuaskan untuk pemodelan, justeru itu, seni bina untuk rangkaian ialah 3-6-1. Pekali sekaitan untuk set latihan dan set pengujian masing-masing adalah 0.997 dan 0.9975, manakala pekali sekaitan

untuk set pengesahan adalah 0.8501. Pengoptimuman penghasilan xilosa dengan menggunakan kaedah GA telah dicatat pada 50.3˚C, 153.5 rpm dan 1.6944 g/l. Penghasilan xilosa optima yang diramalkan adalah sebanyak 0.1845 g/l iaitu pada keadaan yang optima seperti yang diramalkan oleh GA. Sementara itu, suhu optimum dicerap pada 50 °C manakala 132 rpm adalah untuk kelajuan pengadukan yang optimum dan kepekatan enzim yang optimum adalah 1.6474 g/l seperti yang telah dioptimumkan oleh PSO. Penghasilan xilosa yang optimum telah diramalkan sebanyak 0.1845 g/l dengan menggunakan PSO untuk proses hidrolisis enzim. Keputusan ramalan pengoptimuman telah disahkan dengan melakukan ujikaji makmal terhadap proses hidrolisis enzim, di mana kepekatan xilosa telah diperolehi sebanyak 0.2331 g/l dan 0.2398 g/l masing – masing dengan menggunakan PSO dan GA. Purata kesalahan untuk nilai-nilai yang diramalkan dan nilai-nilai yang didapati melalui eksperimen untuk perngoptimuman adalah masing-masing sebanyak 29.97 % dan 26.34% terhadap GA dan PSO. Sehubungan itu, hidrolisis enzim ke atas penghasilan xilosa telah ditingkatkan dengan penemuan keadaan yang optimum dan model yang dibina sesuai dengan data eksperimen.

CHAPTER 1

INTRODUCTION

1.1 Research Background

This research studies the bioprocess production of xylose from rice straw, a topic which normally grabs the attention of researchers due to the complexity of the process, thus it is quite challenging in implementing process modeling and optimization of the process itself. In this study, the research has been separated into 2 major studies which are process modeling and optimization for the enzymatic hydrolysis of xylose from rice straw. The first part of the research is basically the model development of the enzymatic hydrolysis by utilizing the capability of feed-forward artificial neural network (FANN) which is known as one of the artificial intelligence tools. Then, the second part of the research is an optimization of the process by using the particle swarm optimization (PSO) and genetic algorithm (GA) to find the optimum condition of the enzymatic hydrolysis process using the model developed by FANN and its comparison with the conventional optimization method using response surface methodology (RSM). The result obtained by the successful performance of the model developed will give the optimum values for the process.

1.2 Enzymatic Hydrolysis Process

The industries nowadays strive toward inventing new techniques to increase the production capacity and efficiency instead of only focusing on producing high capacity or throughput products especially in the biotechnology field. Therefore, new process mechanisms such as enzymatic hydrolysis are proposed to enhance the production of bioprocess as it does not affect the quality of the product itself but increases the process yield. The enzymatic hydrolysis works by adding certain or particular enzymes which be able to increase the efficiency of hydrolysis in terms of yield and selectivity. This idea is aligned with the principle of 'lock and key' which represents the enzyme and substrate during the process (Koshland Jr, 1995). The hydrolysis which involves enzyme beneficially reduces the time of the process and also reduces the energy consumption.

The production of xylose (known as simple sugar or fermentable sugar) from biomass or waste such as rice straw has gained attention amongst researchers as a bio-nature product. In brief, rice straw is a residue after harvesting and removing the rice seed and husk. On the other hand, xylose is the major component of the hemicellulose fractions and the hydrolysates from pretreatment of acid while glucose and galactose are present in smaller amounts which can be considered undesired product at the end of the process (Brodeur *et al.*, 2011).

The pretreatment process is essential for removing the lignin compound which exists in any agriculture crop or residue. This is because by removing the lignin the hydrolysis process can easily react with other compounds hemicelluloses and cellulose of the substrate. Moreover, the hydrolysis may react through the fastest route by adding the specific enzyme which is able to enhance the conversion of substrate to simple sugar.

The enzymatic hydrolysis process is utilized in the production of simple sugar in order to increase the efficiency of the biomass conversion into the desired product. The enzyme used in the process is xylanase which is known to help increase converted hemicelluloses into xylose. The xylanase enzyme can catalyze the random endohydrolysis of β-1,4-xylosidic linkages in xylan to produce xylooligosaccharides and xylose (Zheng *et al.*, 2009). However, the complexity of the enzymatic hydrolysis of lignocellulosic wastes come from the fact that they are heterogeneous insoluble substrates and thus, their enzymatic hydrolysis is always limited (Carrillo *et al.*, 2005). Due to that concern, this study explored the ability and efficiency of the process and utilizing the capability of the process in the industry.

1.3 Feed forward Artificial Neural Network (FANN)

The study has pursued process modeling as well as the optimization process which is known as another step towards the development in the industry widely. In this research, Feed Forward Artificial Neural Network (FANN) modeling is chosen as a tool to model the enzymatic hydrolysis. One of the advantages of the FANN advantages is, it is able to function as an estimator since it performs the correlation without requiring a mathematical description of how the output depends on input (O'Dwyer *et al.*, 2008). The FANN is effective in approximating nonlinear functions, pattern recognitions and classification problems (Ebrahimpour *et al.*, 2008). In addition, FANN acts directly by transferring the data information from input towards output. Then FANN will stimulate the process model and proceed to the model prediction of the process.

However, modeling without optimization is a waste of the model developed. Therefore, other than modeling, FANN was utilized in optimizing the enzymatic hydrolysis process for the production of xylose from rice straw. Two optimization methods which are particle swarm optimization (PSO) and genetic algorithm (GA) were applied in this study. The optimum yield of xylose by using enzymatic hydrolysis was predicted at optimum condition by utilizing these two approaches.

In brief, the particle swarm optimization is widely known as one of the modern heuristic algorithms under the evolutionary algorithms (Eberhart and Kennedy, 1995). The particle swarm algorithm is known to be simple, computationally efficient and easy to implement because it requires only primitive mathematical operators but also has a cognitive component (Skolpap *et al.*, 2008). PSO algorithm can be combined with chaotic theory and employed to model the data of biochemical systems (Shoseyov *et al.*, 2006). In addition, PSO is the imitation and simulation of the natural intelligence according to the survival behavior of living beings.

On the other hand, genetic algorithm (GA) was used for estimating the variables of the enzymatic hydrolysis process. GA is known as one of the heuristic search algorithm by evolutionary natural selection based on genetic itself. GA would attempt to optimize the objective function which was done in an optimization of corn malt drying that estimated the time and temperature parameters (Scardi and Harding Jr, 1999). GA has the biological background (basic genetics) which exploits historical information to direct the search towards the region of better performance within the search space even if it works in randomized pattern (Sivaraj and Ravichandran, 2011).

1.4 Problem Statement

Enzymatic hydrolysis is widely used as a biochemical process especially in food and beverages industries. The major concern of the enzymatic hydrolysis is to reduce the time consumption as well as the cost of chemical usage for the production process. Moreover, generally the hydrolysis process is demanded as one of the methods that are implemented for producing bio-products such simple sugar or alcohol compound, where the enzymatic hydrolysis progress directly and specifically while work at mild process conditions. The process conditions are considered mild as the process was conducted at lower temperature and agitation speed while the usage of enzyme is less than normal. Regardless of this matter, there are several factors that need to be considered for enzyme usage in hydrolysis as enzyme is literally unstable in certain physical or biological conditions. The physical conditions include temperature and agitation speed of instrument or incubator while the biological conditions include the concentration of enzyme added into the hydrolysis process. Therefore, a study on these conditions has been performed in order to obtain a better value of temperature and agitation speed of incubator in which enzyme concentration is added during the hydrolysis process.

In this case, rice straw used in enzymatic hydrolysis was implemented on the production of xylose. Currently, Malaysia supplies about 65% of the country's rice while another 35% is imported from other countries such Thailand and Vietnam

(Asian Biomass Handbook, 2008). Therefore, the rice residue during harvest and milling process generates a large amount of rice straw and rice husk. The problem arises as the rice straw or biomass residues become a source of pollution as rice straw is usually burnt by open burning. The burnt rice straw may lose its nutrient although it is a cost-effective method for straw disposal. Moreover, the bioconversion of rice waste (biomass waste) is convincing as one of raw material to be processed as renewable resource for complementing from waste to wealth ideology.

The enzymatic hydrolysis may enhance the production of xylose process, however, there is no certain model that could explain every process because it is a nonlinear process. The efficiency of the enzymatic hydrolysis process is necessary for discussion to prove that the process performs well. The parameters correlation brings up several questions about its relevance and significance to enhance the production for enzymatic hydrolysis.

In modeling, it is often necessary to provide some complex description or information. Instead, the feed-forward artificial neural network (FANN) which is proposed in this research has the capability to represent the complex relationship without needing much information. As a nonlinear data analyzer, FANN is a useful tool in advanced technology to investigate the efficiency of predicted data. This is because the FANN is robustness enough to predict the provided process data especially for modeling nonlinear bioprocess such enzymatic hydrolysis. The information is directly transfer from input data to output data, and the training algorithm used in the FANN model to avoid overfit the data. FANN modeling was implemented by using one hidden layer for enzymatic hydrolysis of the production of xylose from rice straw in this study. Furthermore, one hidden layer is often used as the information interconnect with the input and output data which this number of hidden layer may provide a good neural network modeling.

Besides modeling, an optimization method is widely applied among researchers and is also normally compulsory for implementation in their study. Optimization is utilized since this method is able to optimize the yield and the able to reduce the duration time of the experimental work. The study on the bioprocess process modeling requires plenty of time especially on the enzymatic hydrolysis. In addition to that, some of the experimental work need to be done more than one with different parameters and condition.. Therefore, in order to reduce the amount of time and cost for conducting the process, process modeling and the optimization is needed.

The optimization method used in this research is based on the methods of evolution techniques which are genetic algorithm (GA) and particle swarm optimization (PSO) methods. The genetic algorithm is implemented to reduce the searching area of local optima in order to determine the optimum value of parameters in the study. The relationship of parameters of the enzymatic hydrolysis is known to be nonlinear or it is difficult to predict the optimal value. Hence, the GA method is used for the generation of possible solutions through crossover, mutation and evaluation of objective function. On the other hand, the PSO method approach is slightly different than the GA method where the prediction of best solution is done by selecting among the best solutions' search areas to find where the particle swarm is located. The particle swarm and converge toward the next best solution in the search space to find the best particle before sending out the information to others in the optimization problem. This method uses minimum computation time while generating the best solution. This optimization tool is a robust technique for solving nonlinear problems such enzymatic hydrolysis.

1.5 Research objectives

The research basically focuses on modeling and optimizing the enzymatic hydrolysis process. The aims of this research are:

1. To develop a model of enzymatic hydrolysis by xylanase enzyme to produce xylose from rice straw by using feed-forward artificial neural network (FANN).
2. To optimize the enzymatic hydrolysis process using particle swarm optimization (PSO) and genetic algorithm (GA) and compare with the conventional RSM approach.
3. To validate the predicted optimum condition from the model with the experimental work.

1.6 Organization of Thesis

This thesis is organized into five chapters, where Chapter One introduces the current overview of modeling and optimization on enzymatic hydrolysis. The objectives of research and the scope of study are pointed out.

Chapter Two provides reviews on enzymatic hydrolysis process for the production of simple sugar, modeling using neural network and optimization of the parameters studied during enzymatic hydrolysis process.

Chapter Three describes the case study of enzymatic hydrolysis for the production of xylose. This chapter also elaborates on the stages of the research conducted on modeling using neural network and optimization using particle swarm optimization (PSO) and genetic algorithm (GA). It also describes the validation for optimization of enzymatic hydrolysis.

Chapter Four presents the performance evaluation of the model developed for enzymatic hydrolysis process. The optimum values of the conditions for enzymatic hydrolysis obtained were presented and discussed in this chapter.

Last but not least, Chapter 5 concludes the finding of the research and the achievement of the objectives. This chapter includes the recommendations for the future study.

CHAPTER 2

LITERATURE REVIEW

2.1 Production of Xylose

Xylose is known as reducing sugar of monosaccharide group. Xylose can be extracted from polysaccharide with a several methods such as hydrolysis (Zhang *et al.*, 2014*)*. Since few years ago, the production of xylose was recovered from biomass waste which one of reducing sugar that has many beneficial towards mankind. One of the application of xylose, it can converse into xylitol in vitro as in many yeast strain produce high yield of xylitol from xylose (Milessi *et al.*, 2011). Hence, throughout this idea the production of xylose was proposed and recovered from biomass waste as a raw material for this study.

2.1.1 Rice Straw as Biomass

Rice straw is considered the largest portion of available biomass feedstock in the world at about 7.31×10^{14} of dry rice straw per year and Asia is responsible for 90% of the annual global production (Kim and Dale, 2004). Rice straw is attractive as a fuel because it is renewable and is considered to be carbon dioxide neutral (Atchison, 1996) but has not yet been commerciallized. The nature of rice straw is limited by the great bulk of material, slow degradation in the field, harboring of rice diseases and high mineral content. However, the straw must be disposed of in order to make way for the next crop (Alexander *et al.*, 2002). Rice straw has been utilized in many bioconversion processes.

Therefore, rice straw have been investigated for its potential as the sole feedstock to produce biogas because it contains a high percentage of polysaccharides and lignin (He *et al.*., 2008). In recent years, a lot of attention have been focused on the biotechnological process for production of several useful feedstock and food products from agro-forest residues and agriculture residues such as rice straw (Roult *et al*., 2008). However, lignocellulose such as rice straw is difficult to hydrolyze using only enzyme due to its recalcitrant and heterogeneous structure, which primarily consists of cellulose, hemicelluloses and lignin (Chandra *et al*., 2007).

2.1.2 Pretreatment Process

Pretreatment of the rice straw has proved to increase both its physical and chemical properties, thereby minimizing the costs of transport, handling and storage. These applications are concerned with improving combustion efficiency and reducing pollution emission. Acid hydrolysis has been investigated as a possible process for treating lignocellulosic materials such as wood chips, rice straw (Almeida, 1991), sugar beet pulp (Chamy *et al*., 1994) and wheat straw (Fanta *et al*., 1984).

Rice straw can be hydrolyzed using dilute acid to obtain a mixture of sugars with xylose as the major component. However, in the hydrolyzate some by-products generated in the hydrolysis, such as acetic acid, furfural, phenolic compounds or lignin degradation products, can be present. These are potential inhibitors of a microbiological utilization of this hydrolyzate (Dominguez *et al.*, 1996). Treatment with dilute sulphuric acid at moderate temperatures in the first stage of acid hydrolysis has proven to be an efficient means of producing xylose from hemicellulose (Roberto *et al.,* 1994; Silva 1996).

On the other hand, the pretreatment used to remove lignin which known as forms one of the part of three the cell wall. They are insoluble in water and is partially soluble in organic solvent because of the hydrogen bonds between polysaccharides and the linkage of lignin to polysaccharides. Thus, the pretreatment have been proven to be one of the most simple and effective methods to improve biodegradability and biogas production of lignocelluloses materials (He *et al*., 2008). The biodegradable part of lignocelluloses materials may be converted into reducing sugar.

There are chemical pretreatments such as alkaline hydrolysis that decreased the lignin content while enhance the enzymatic saccharification. The chemical processes are based on cellulose hydrolysis of acids or cellulose solvents such as alkaline hydrogen peroxide (Mosier *et al.*, 2005). The release of sugars from lignocelluloses biomass was facilitated by pretreatment process.

The lignocelluloses pretreatment uses other chemicals such as ionic liquids to dissolve cellulose in biomass. Ionic liquids can be reused after treatment and are easily applied to enhance the enzymatic hydrolysis and efficiently recover fermentable sugars such as glucose and xylose from lignocellulosic biomass source (Li *et al.*, 2006). Therefore, Nguyen *et al.* (2010) had studied the pretreatment method using ammonia and ionic liquid for the recovery of bio-digestible cellulose from lignocellulosic by-product which effect on the enzymatic glucose conversion.

Lignocellulosic biomass cannot be saccharified by enzyme to obtain high yield without a pretreatment procedure because the lignin in the cell wall is a barrier to enzyme action (Sewalt *et al*., 1997). Rice straw was selected as a substrate for saccharification in a research done by Jeya *et al.* (2009). The lignin components were decreased to 37% in pretreated rice straw with 2% sodium hydroxyide.

2.1.3 Enzymatic Hydrolysis

The hemicellulosic fraction formed after pretreatment process can then be enzymatically hydrolyzed to xylose. Enzymatic hydrolysis of biomass hemicellulose does not produce toxic products. The hydrolysis of hemicellulose is accelerated at elevated temperatures owing to relatively high activation energy in the solid-liquid phase reaction. At high temperatures part of the xylose released from hemicellulose degrades rapidly and cellulose in the amorphous region can yield glucose (Banerjee, 1989). The enzymes that are involved in degradation of hemicelluloses are listed in Table 2.1 (Selinger *et al*., 1996).

Table 2.1: Enzyme involved in the hydrolysis of complex hemicelluloses

Enzyme	**Mode of action**
Endo-xylanase	Hydrolyzes mainly interior β-1,4-xylose linkages of the xylan backbone
Exo-xylanase	Hydrolyzes β-1,4-xylose linkages releasing xylobiose
β-Xylosidase	Releases xylose from xylobiose and short chain xylooligosaccharides
α-Arabinofuranosidase	Hydrolyzes terminal nonreducing α-arabinofuranose from arabinoxylans
α-Glucuronidase	Releases glucuronic acid from glucuronoxylans
Acetylxylan esterase	Hydrolyzes acetylester bonds in acetyl

	xylans
Ferulic acid esterase	Hydrolyzes feruloylester bonds in xylans
γ-Coumaric acid esterase	Hydrolyzes γ-coumaryl ester bonds in xylans

Endo-xylanases are much more common than β-xylosidases, but the latter are necessary in order to produce xylose. Activity was optimum at pH 3.3 and 52 °C. β-Glucuronidase acts in synergism with xylanases and β-xylosidases to hydrolyze glucuronoxylan. The yield of xylose greatly increases in the presence of this enzyme (Puls *et al*., 1980). Thus, Wang and Zhang (2006) produced xylose from corncobs by xylanase through hydrolysis under concentrated ultrafiltration with polyamide (PA) capillary fibres. Most importantly is endoxylanase cleaves β-1,4-xylose backbone in hydrolysis (Sharma *et al*., 2010). In the process, the enzyme (xylanase) is used to increase the surface contact with the substrate (treated rice straw) for incremented productivity of simple sugar (xylose).

Therefore, it is necessary to develop enzymes and microorganisms that are resistant to such inhibitory substances or to employ additional steps to remove the inhibitor. The pretreatment and enzymatic hydrolysis steps to achieve fermentable sugar are currently known to provide much more room for reducing processing cost than other processes (Lynd *et al.*, 2008). The enzymatic hydrolysis is utilized on recovery of reducing sugar. The crystalline structure of cellulose has been hydrolyzed by cellulase during enzymatic hydrolysis (Jeoh *et al*., 2007). The loading of cellulase increases the production of glucose (Bak *et al*., 2009).

Furthermore, the enzyme enhances conversions of pretreated substrate. Thus, Kovacs *et al.*. (2009) have studied enzymatic hydrolysis of steam-pretreated lignocellulosic materials with Trichoderma atroviride enzymes produced in-house. In the research claimed that supplementation of Trichoderma atroviride with xylanase enzyme resulted in an increase of 40% in the xylose level and an improvement of 21% in the glucose concentration.

2.1.4 Independent Variables Affecting the Enzymatic Hydrolysis

The enzymatic hydrolysis has been studied for many years especially in bioprocess, and the researcher is interested to enhance and to recover the production of the process. In addition, the enzymatic hydrolysis is one of way to produce reducing sugar in mild operating condition such pH and temperature (Carrillo *et. al*., 2005). There are several factors that affecting the production process such as mechanical or physical effect and chemical effect such as substrate concentration, enzyme concentration, temperature, agitation speed, pH and time for the enzymatic hydrolysis process. However, the main focus in this study are on the effect of temperature, agitation speed and enzyme concentration for the enzymatic hydrolysis process in the production of xylose. These three parameters chosen as it is significant toward the enzymatic hydrolysis.

2.1.4.1 Temperature

Temperature is a vital condition for every process, so is the case for the enzymatic hydrolysis. Thus the effects of temperature towards the bioprocess production have been studied in various ways, both either in large scale and flask scale. The study on this effect have been reported by Srivastava and Tyagi (2003) for investigating the temperature effects on the juice from apple fruit during the enzymatic hydrolysis. They studied the effects of temperature within the range of 35 – 55 °C of the enzymatic hydrolysis.

The temperature of the enzymatic hydrolysis was investigated for obtaining the suitable value for perfectly operates on the process. Mild conditions are known to give the best production results whenever an enzyme is involved during the process. The study on the temperature is of vital significance for example in the enzymatic hydrolysis of breadfruit starch which has been studied in the case study with its utilization for gluconic acid production by Betiku and Ajala (2010). The temperature was studied on both liquefaction and saccharification in order to determine the biomass concentration. The measurement of the dextrose equivalent (DE) at certain readings indicated that the glucoamylase operates effectively for the hydrolysis of the breadfruit starch.

2.1.4.2 Agitation Speed

Study on the effect of agitation speed has the potential to observe the mixture effect on the homogenous solution for the production of bioprocess. The agitation speed has been studied for the fermentation process on the biotransformation of fenofibrate which looks for maximum metabolite production and cell dry weight (Vidyavathi *et al.*, 2013). Therefore, the study on the agitation speed for bioprocess have been studied widely in different areas and also for different productions.

The study on this effect is believed to have contributed to the better performance of bioprocess production, specifically in the enzymatic hydrolysis. This is due to the fact that the agitation speed enhances the interaction within the substrate and enzyme literally without interfering with the enzyme activity during the enzymatic hydrolysis (Inggeson *et al.*, 2001). This effect makes its study for enzymatic hydrolysis significant as it is able to increase the production with a certain time.

2.1.4.3 Enzyme Concentration

Independent variables, especially such as the enzyme concentration is an interesting factor for study since it is widely utilized in industrial chemical process. Studies have been done to observe the enzyme treatment effect on different situations such as in pressing operation of borage seed oil extraction (Soto *et al.*, 2006). The enzymatic hydrolysis was evaluated on the oil extraction from different kinds of oilseed and the production qualities for this treatment has also been studied. The moisture of enzyme was analyzed to find better enzyme activity whenever there is water bonding to compare to the total water of the systems.

On the other hand, the enzyme concentration such xylanase has been studied by Normah *et al.* (2012) in order to determine the enzyme activity during enzymatic hydrolysis. The enzyme concentration was varied at 2, 4 and 6 units/ml to optimize the process of hydrolysis in order to enhance the production of xylooligosaccharides. It was determined in the study that the lowest concentration contributed to higher yield of the xylooligosaccharides production at reaction time periods.

2.1.5 Experimental Design on the Production of xylose

The production of xylose from biomass or rice straw has reviewed on used of the experimental design during the enzymatic hydrolysis process. The experimental design by using different kinds of tool which used for analysis the experimental result or even arrange the experimental work that includes the independent variables study. There several tools applicable for design the experimental of the process such response surface methodology (RSM), design of experiment (DOE) and other statistical tools as Minitab.

2.1.5.1 Response Surface Methodology

The design of experiment (DOE) tool using response surface methodology (RSM) is implemented in order to obtain the optimum conditions of production. RSM is a collection of statistical techniques for designing experiments, building models, evaluating the effects of parameters and searching for the optimum conditions, and has successfully been used in the optimization of bioprocesses (Hao *et al.*, 2006). In order to optimize the xylose production by hydrolyzing sugarcane bagasse, RSM was used to maximize the temperature and sulphuric acid concentration of the samples (Paiva *et al.*, 2002). Based on the principle of DOE, the methodology encompasses the use of various types of design such as experimental designs, generation of polynomial equations and mapping of the response over the experimental domain to determine the optimum product (Box and Draper, 1987).

Statistical experimental design was used in order to optimize hydrolysis parameters such as pH, temperature, and concentrations of substrates and enzymes to achieve the highest saccharification yield. Enzyme concentration was identified as the limiting factor for saccharification of rice straw (Jeya *et al.*, 2009). On the other hand, Silva and Roberto (2001) have used RSM in order to optimize their production of xylitol by *Candida guilliermondii* FTI 20037 based on the effect of initial xylose concentration and inoculums level.

Response Surface Methodology is able to evaluate multiple parameters and their interactions while at the same time reducing the number of experimental trials. Thus, the process conditions such as pressure, temperature, camel hump fat ratio, water content and incubation time can be optimized at five different levels using RSM as done by Shekarchizadeh *et al.* (2009). A second order polynomial response surface equation was developed indicating the effect of the mentioned variables on cocoa butter analog yield.

2.1.5.2 Regression Analysis

There are several regression analysis tools that are widely used in research studies such as minitab software, design of experts, polymath, statisca software and other analysis tools. Bioprocess or specifically the biotechnology field often uses RSM designed by experts to demonstrate the mathematical and statistical analysis equation. This is also known as regression analysis where the mathematical equation used to express the relationship of process variables such as pH, temperature, substrate concentration, inoculum size, and agitation speed on xylitol yield (Ramesh *et al.*, 2013).

Regression analysis may also help to study modeling and optimization process. The regression analysis was performed based on the experimental data and expressed as an empirical model second order polynomial equation (Long et.al, 2010) such as in Equation 2.1.

$$Y = \sum_{i=1}^{3} AiXi + \sum_{i=1}^{3} AiiXi^2 + \sum_{i=1}^{2}\sum_{j=i+1}^{3} AijXj \qquad (2.1)$$

Whereby Y is represents the response variable, while Ao, Ai, Aii, Aij refer to regression coefficients of variables for intercept, linear quadratic and interaction terms respectively. Xi and Xj noted independent variables that are studied.

In normal cases, the statistical analysis is performed in the form of variance known as analysis of variance (ANOVA). The analysis comes out with the result of correlation coefficient (R), Fisher's F-Test, determination coefficient (R^2) for measuring the proportion of variance obtained (Paiva *et al.*, 2008). The fit model of second order polynomial model equation is verified by the determination coefficient and the regression coefficient significance needs to be checked by Fisher's F-Test (Long *et al.*, 2010).

2.2 Empirical Model for Enzymatic Hydrolysis Process

There are two types of modelling approaches which are empirical and mechanistic modelling (Peri *et al.*, 2007). According to Peri *et al.* (2007) the empirical models relate the factors using mathematical correlations without underlying mechanisms. These define models that are easy to develop and are useful in enzyme characterizations and substrate preparation. Meanwhile, the mechanistic model is developed from the reaction mechanisms, mass transfer considerations and other physical parameters that affect the extent of hydrolysis.

An empirical model develops for certain process without mechanistic consideration within the process. The model had been studied for prediction of drugs skin permeability where the approach was based on the selection of molecular or structural descriptors (Yamashita and Hashida, 2003). The study mentions that an artificial neural network have been used as tool for non-linear modeling of complex causal-effect relationship.

2.2.1 Artificial Neural Network (ANN)

Modeling using neural network has been practiced in many areas such as in chemical and biotechnology processes and even in the pharmacological field. The neural network is robust in solving some complex problems by the linking of inputs with the predicted dependent variables with either linear or nonlinear model (Szaleniec, 2012). The report discussed the input variables selection which a must in the neural network modeling, data set division and model validation while also reviewing the optimization of network internal structure.

The neural network was literally inspired by the human brain which works as biological model and building blocks, where the neurons are combined and connected into layers for transferring the information data (Li *et al.*, 2006). The development of the artificial neural network (ANN) model or known as one of types of artificial intelligence was initially made to understand how the brain works and to construct a mechanism that functions or mimics the same way (Cheng and Titterington, 1994). Figure 2.1 illustrates the neuron system in human brain.

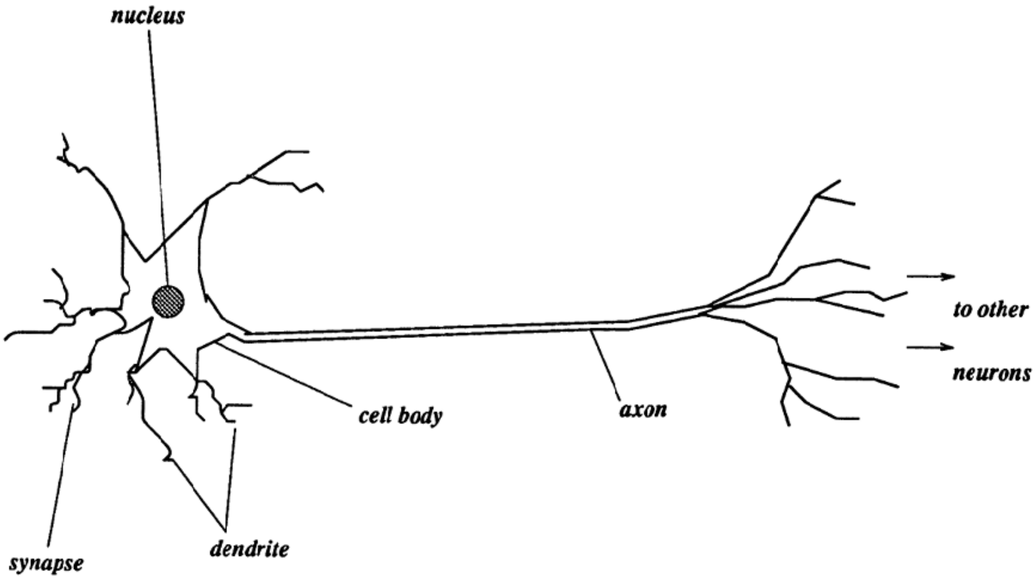

Figure 2.1: Schematic diagram of real neuron systems in the human brain by Cheng and Titterington (1994).

The neural network model selection is based on weight and bias that is employed in the neurons of network. The weight and bias is calculated as in Equation 2.2 as an example for a single neuron (Murad *et al.*, 2010).

$$y_{pi}^{l} = \sum_{i=1}^{n} \left(w_{ij} a_{ij} \right) + b_j \qquad (2.2)$$

where w_{ij}, a_{ij} and b_j represent as weight, input and bias respectively. i and j denotes the row and layer of the neuron in the network. The number of nodes is calculated and then passed to the transfer function which is then employed in the neural network for the simulation. The transfer function using log- sigmoidal transfer function and purelin transfer function are expressed as in Equation 2.3 and Equation 2.4.

$$f(y) = \frac{1}{1+e^{-y}} \tag{2.3}$$

$$f(y) = y \tag{2.4}$$

2.2.2 Type of Neural Network Model

The neural network model has numerous types of model which can be used in the process modeling. Several of model are known as feed-forward neural network (FANN), radial basis function (RBF) and recurrent neural network (RCNN) which normally implemented in the process for modeling. One of the popular artificial neural network architecture is feed forward neural network (FANN) where the neurons of the network are grouped in layers and connect together with forward connection thus the connection can learn any kind of continuous nonlinear mapping applications (Mendes *et al.,* 2002). The feed-forward networks training using Levernberg-Marquardt algorithm was used in the optimization the network parameters due to its quadratic approximation accuracy and converges faster than gradient descent method (Ismail *et al*., 2013).

The neural network model is usually known as multilayer perceptron (MLP) which consists of several layers that are connected to each other. MLP interconnects the neurons that are arranged in layers corresponding to the input layer hidden layer and output layer (Nodeh, 2012). The study was investigated on two layers with different neurons in which the neural network was applied for modelling the saccharification of biomass by concentrated acid hydrolysis.

The MLP is the most common architecture of artificial neural network which is normally applied for model development whereby complex or nonlinear relationship is involved in certain processes. The processing units in the input layer and output layer of the MLP is referred to as number nodes which is determined by the independent variables and dependent variables (Masoumi, *et al.*, 2011). The MLP diagram which consists of an input layer, one hidden layer and one output layer (Kana *et al.*, 2012) is shown as in Figure 2.2.

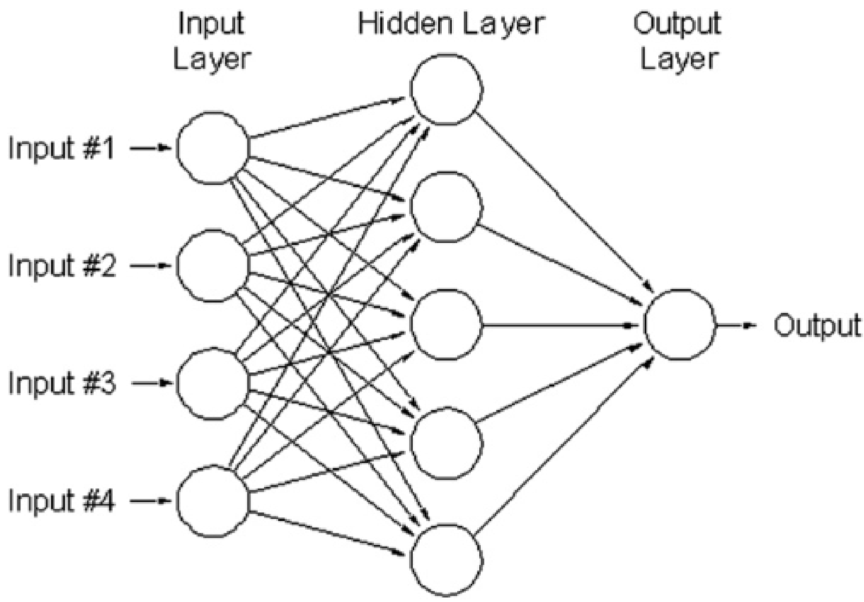

Figure 2.2: The diagram of MLP that consists of input layer, hidden layer and output layer.

The neural network that learns directly from input data set and output data set is called a feed-forward neural network. The principle of learning the input and output data without the requirement of the phenomenological description of how the