Vol. 1(1), pp. 1-18, Sept 2022
DOI: 10.0001/AJPC2021.0207
ISSN 2021-xxxx

ASEAN Journal of Process Control

Research Article

Enhanced Production Of Levulinic Acid Using Carbon Source From Oil Palm Biomass And Matured Coconut Water

Sharrmela Raj[1], Norliza Abd. Rahman[1,*] and Jarinah Mohd Ali[1]

[1] Department of Chemical Engineering and Process, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

*Corresponding Author: norlizajkkp@ukm.edu.my
Academic Editor: Marappa Gounder Ramasamy
Received: 6 April 2021; Accepted: 30 April 2021; Published: 1 September 2022

Abstract: Levulinic acid (LA), a powerful platform chemical that can only be synthesised from biomass. LA have the capability to act as a substitute intermediate chemical in the production of hydrocarbon-based products. The initial purpose of this research project was to investigate different types of metal chloride catalyst along with determining the optimum process conditions to synthesis LA from oil palm biomass such as oil palm sap (OPS) and matured coconut water. However, the project was disrupted due to the COVID-19 pandemic. Thus, the completion to this project is via artificial intelligence approach. Artificial neural network (ANN) was employed to create a shallow neural network to predict the LA yield using experimentation data sourced from a relatable journal, whereby the inputs were amount of sugar (g/L), catalyst loading (g), temperature ( ) and reaction time (hr) and the output was LA yield (%). Database for sugar composition in both oil palm biomass and matured coconut water was developed to illustrate that it has the capability to produce LA. Second, a performance comparison and evaluation was carried out to analyse and validate the Levenberg-Marquardt (LM) & Bayesian Regularisation (BR) based-ANN models by altering the number of neurons. Overall, LM generalised better than BR algorithm as all of LMANNs' MSE values were kept below and values were more than 0.88. It was determined that LMANN with 21 hidden neurons is suitable for the experimental database used to build the ANN models.

Keywords: Neural networks, levulinic acid, oil palm biomass.

1. Introduction

In recent years, the depletion of fossil fuel urges the necessity to find a sustainable source to replace it and biomass is deemed as a reliable option. Biomass-based chemicals is a form of renewable energy that has positive attributes to the environment and society such as being a carbon neutral component. In 2004, researchers from the National Renewable Energy Laboratory (NREL) carried out an extensive study to identify valuable biomass-derived chemicals. The list was narrowed down to twelve chemicals which were assessed based on potential markets of the chemical and its derivatives including the synthesis pathway [21]. One of those top twelve value-added chemicals were levulinic acid (LA).

LA consist of a ketone and a carboxylic acid functional group which makes this compound a versatile building block. Figure 1 illustrates the synthesis routes for the various derivatives from LA. The derivatives produced is used a precursor for the production of various applications such as in making of textile dyes, animal feed, coating material, solvent, biodiesel additives, food flavouring agent and as a pharmaceutical compound [15]. Another interesting physical property of LA is that it is also readily soluble in solvents such as water, ethyl acetate and ethanol due to its high dissociation constant (p^{Ka}) which is 4.59 [20].

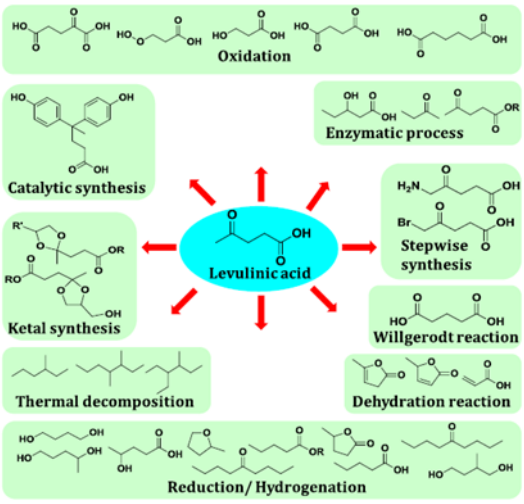

Figure 1: Synthesis routes of various products from LA [7]

The feedstock of LA production varies from starchy edible biomass, lignocellulosic biomass and algae. Initial studies were using edible biomass such as corn stalk and sugar cane molasses as feedstock for the production of LA. However, the usage of edible biomass is not advisable because it imposes a negative impact on food security. Due to the following challenge, the focus was drifted to utilising lignocellulosic biomass for the production of fuels and chemicals. Sources of lignocellulosic biomass includes agricultural wastes, forestry residues and energy crops.

Malaysia has abundance of agricultural wastes and mainly it is contributed from oil palm plantations. Parts of the oil palm such as fresh fruit bunch (FFB), empty fruit bunch (EFB), fronds and trunks have the potential to produce chemical products that were initially produced from petroleum. So, the usage of oil palm residues would help reduce the accumulated agricultural wastage by using it as feedstocks for the production of biochemicals and biofuels. Oil palm sap (OPS) which is extracted from felled trunks consist of 86.9 % glucose and the remaining fraction contains of other C5 and C6 sugars [39]. The OPS can be directly used for the conversion without the need of pre-treatment.

The decomposition of LA from glucose using water as solvent and different types of acid catalysts has been studied extensively. The catalyst studied were either from mineral acids, metal chlorides, zeolites or ionic liquids and the catalyst are classified into three groups as described in Table 1. Based on the studies, conducted by [15,45,25] proved that sulphuric acid can potentially corrode the reactor and suffer difficulty in recovering catalyst as it is a homogeneous catalyst. Another alternative catalyst would be to use ionic liquids (ILs), a dual solvent-catalysts that serve as a reaction media for various fields prominently in production of bio-chemicals such as LA and 5-HMF [74]. However, it is not economically feasible. A better candidate would be to use metal chloride catalyst for the synthesis of lignocellulosic biomass to LA. Metal chlorides can easily be separated from the process stream and be reused several times. Besides, it is also less toxic than a homogenous catalyst [23].

Table 1: Classifications Of Catalyst Used For La Processing

Homogenous Catalyst	**Heterogenous Catalyst**	**Ionic Liquid**
Mineral acids	Metal chlorides Zeolites	Imidazolium Pyridinium

HCl	HY-Zeolite	1-ethyl-3-methyl-imidazolium acetate
H_2SO_4	$InCl_3$	1-ethyl pyridinium chloride
H_3PO_4	$AlCl_3$	

Due to the situation caused by the pandemic, it was impossible to perform the research project as planned, which is via experimentation. So, the alternative to the complete this project was by employing an artificial intelligence (AI) approach. AI techniques such as artificial neural network (ANN) and respond surface methodology (RSM) have been applied extensively to predict the output of biochemical process over the recent years. Analytic tools such as artificial neural network (ANN) has the ability to calculate the relationship between input and output whilst learning and generalising the behaviour of a complex non-linear data [49].

The processing of a neural network mimics the human brain. A human brain is made up of connections of neurons whereby the neuron is the processing element. As for the neural network, the node corresponds to the neurons of human brain and are connected through weights. Using Figure 2 as a reference for explanation, each input element, of the corresponding neuron is associated with a weight, respectively. Bias, is a factor added to each neuron and is associated with the storage of information. To simplify it, weights and biases store the information of a neural network.

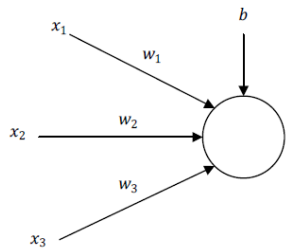

Figure 2: Neuron Structure

The inputs and outputs of the network are related through transfer functions. Transfer functions models the desired output based on the inputs. Types of transfer functions include sigmoid function, hyperbolic tangent function, sine function, linear and saturated linear transfer functions [50] The neurons are normally grouped into subsets which are categorised as layers. The layers are input layer, hidden layer and output layer. An ANN architecture with those three layers are described as multi-layer neural network as illustrated in Figure 3.

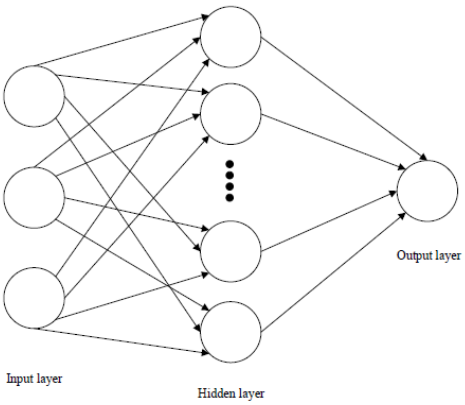

Figure 3: Architecture of a Feed-Forward ANN Model

There are two types of multi-layer neural network; shallow and deep neural network. Shallow neural network is when it consists of only one hidden layer whereas deep neural network consists of two or more hidden layer. Shallow neural network is commonly used as prediction model. Among the

popular architectures recommended for ANN is the multi-layer feed-forward with back propagation learning algorithm (FFBP).

Back-propagation algorithm sets as a systemic method to identify errors in the hidden layers. It essentially means the algorithm propagates the output errors backwards passing through hidden layer and then to input layer. So, during training phase the weights and biases are modified accordingly based on the error. A crucial factor that determines the accuracy of ANN model is the choice of training method. For complex non-linear systems, Levenberg-Marquardt (LM) and Bayesian Regularisation (BR) are recommended as it processes the model fast and reduces overfitting problem.

So, for this study, shallow neural networks were developed to predict the LA yield using experimentation data sourced from a relatable journal, whereby the inputs were amount of sugar (g/L), catalyst loading (g), temperature ($^{\circ}C$) and reaction time (hr) and the output was LA yield (%). Database for sugar composition in both oil palm biomass and matured coconut water was developed to illustrate that it has the capability to produce LA. Second, a performance comparison and evaluation was carried out between Levenberg-Marquardt (LM) and Bayesian Regularisation (BR) based-ANN models by altering the number of neurons. This was done to obtain the most accurate and reliable predictive network.

2. Methodology

2.1. Methodology of Artificial Neural Network

For this study, experimental data sourced from a related journal was fitted using the neural network toolbox in MATLAB R2017a. A multi-layer feed forward back propagation (FFBP) shallow neural network was developed using sigmoid function and linear function as activation functions for the hidden layer and output layer respectively. The data set used in the training phase has 27 experimental points where 70 % is taken up for training, 15 % for testing and another 15 % for validation of ANN model.

The training algorithms chosen to develop the ANN models were altered between training algorithm LM and BR. Using Figure 3 as a guidance, the ANN parameter was optimised by varying the number of neurons in the hidden layer for both training algorithm. The hidden neurons altered for both LMANN and BRANN models are in a range of 3 – 25. A total of 20 ANN models were developed and trained independently.

The performance of ANN is validated based on the calculation of mean square error (MSE) and correlation coefficient (R^2) using Equation 3 and 4 respectively [51]

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=0}^{n}(y_i - y_{di})^2 \quad (1)$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}\left(y_{\mathrm{i}} - y_{\mathrm{di}}\right)^2}{\sum_{\mathrm{i=1}}^{\mathrm{n}}\left(y_{\mathrm{di}} - y_m\right)^2} \quad (2)$$

where n is the number of points, is the predicted LA yield, is the experimental LA yield and is the mean of experimental LA yield. Lower values of MSE indicates that the model is suitable. On the other hand, has to be more than 0.90 for both training algorithms as this shows how well the outcomes are to be predicted by the models developed. So, the training phase for each network is stopped until low MSE and high R2 values were obtained. The properties of the ANN model developed in this study is as described in Table 2. At the plantwide level, several mechanisms and process interactions can determine the SITC dynamics and performance. One critical phenomenon affecting the efficiency of the SITC plant is the formation of different immiscible phases in the Bunsen reactor [8]. Bear in mind that the Bunsen Section is the heart in the SITC process - the reaction yield must be high enough to produce an effluence composition well above the azeotropic (H2O-I2-HI) composition. Meeting this objective in the Bunsen section is vital to enable the smooth operation of those other sections.

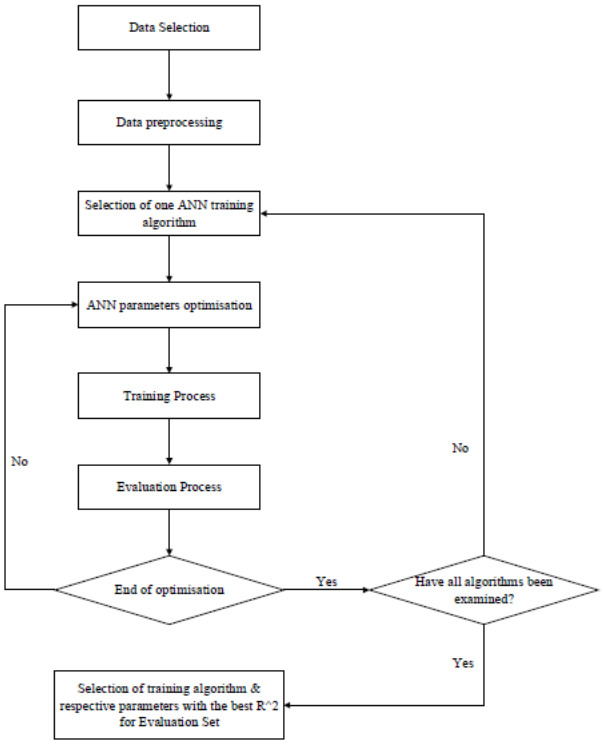

Figure 4: Methodology of training the ANN model

Table 2: Properties Of Ann Model

ANN model	
Network type	Shallow neural network
Activation function (hidden)	hyperbolic tangent
Activation function (output)	linear
Number of layers	2
Number of neurons (input)	4
Number of neurons (hidden layer)	3 – 25
Number of neurons (outputs)	1
Learning algorithm	Back-propagation
Training algorithm	LM and BR

2.2. Development of Artificial Neural Network in Matlab

Input and output tabs were created in Matlab workspace then the respective tabs were filled with the dataset obtained from external journal. 'Transpose' the input and output variables.

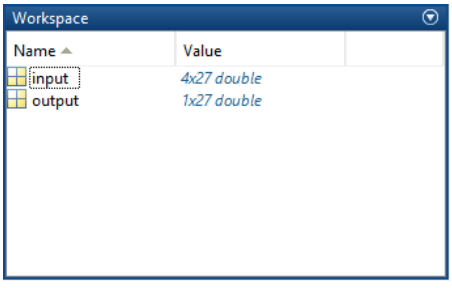

Figure 5: Matlab workspace

Started the neural network fitting toolbox by keying 'nftool' in Matlab command window.

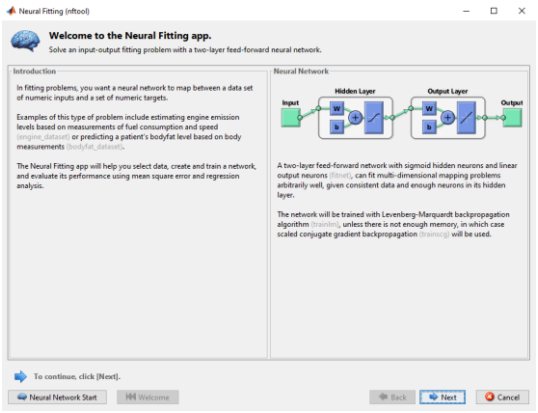

Figure 6: Neural net fitting tool in Matlab

Selected the respective input and output variables.

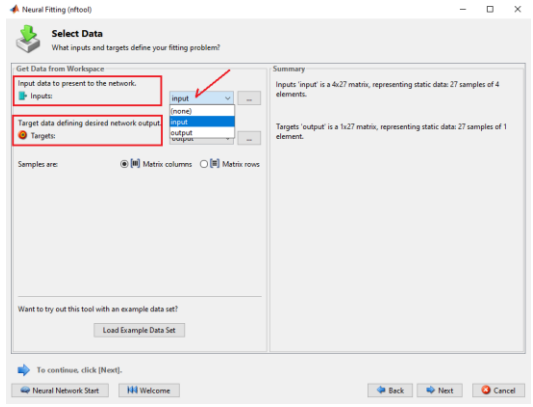

Figure 7: Data selection in neural net fitting tool

Selected a specific number of neurons in the hidden layer within the range of 3 to 25 neurons for network training.

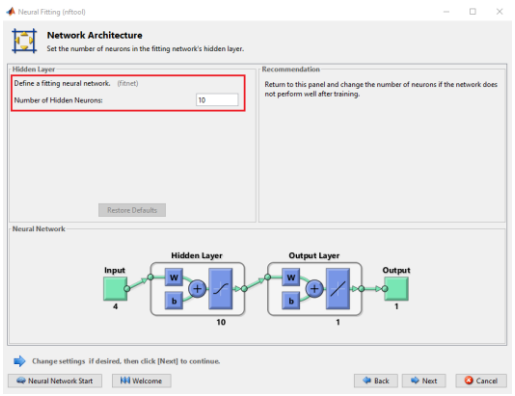

Figure 8: Network architecture in neural net fitting tool

Selected either Levenberg-Marquardt or Bayesian Regularization training algorithm and then network was trained.

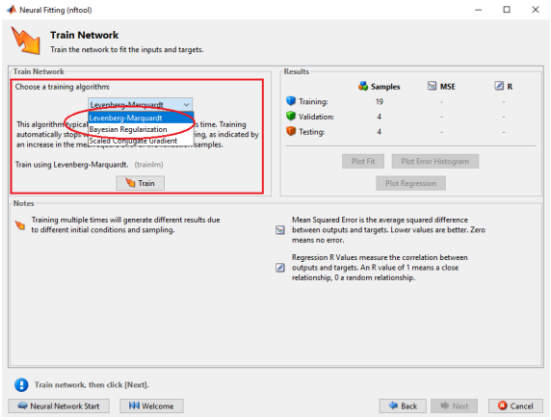

Figure 9: Training network in neural net fitting tool

Extracted the 'Performance' and 'Regression' plots to get the MSE and values.

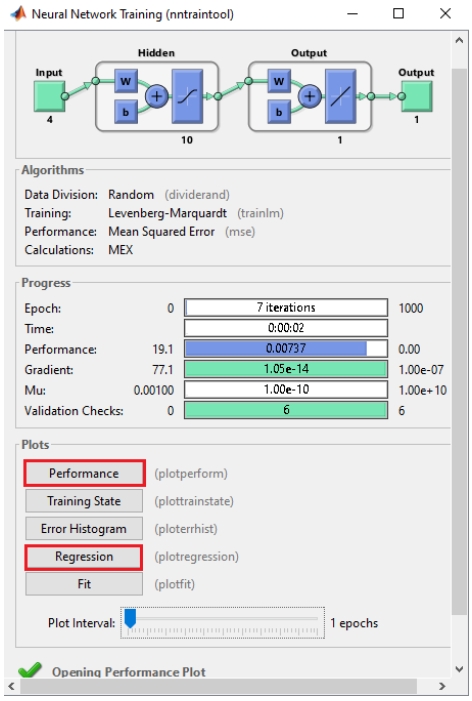

Figure 10: Neural net fitting app

If satisfactory MSE and R^2 values were obtained, then the Matlab script for the neural network architecture developed was generated by clicking onto 'Simple Script'. The results of the predicted outputs as well the errors were saved.

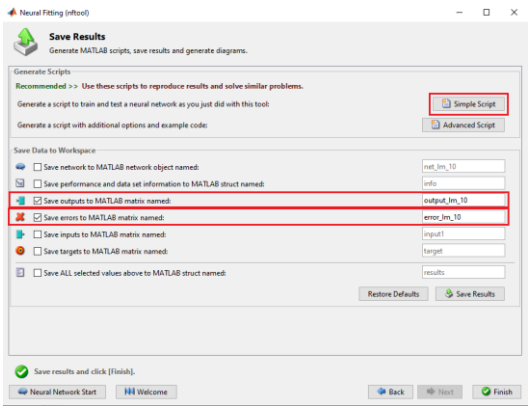

Figure 11: Results in neural net fitting tool

If the results achieved are not desirable, then the network was retrained until satisfactory MSE and values were obtained for that particular architecture. Once that network architecture has produced desirable results, the number of neurons are changed to get the optimum network architecture.

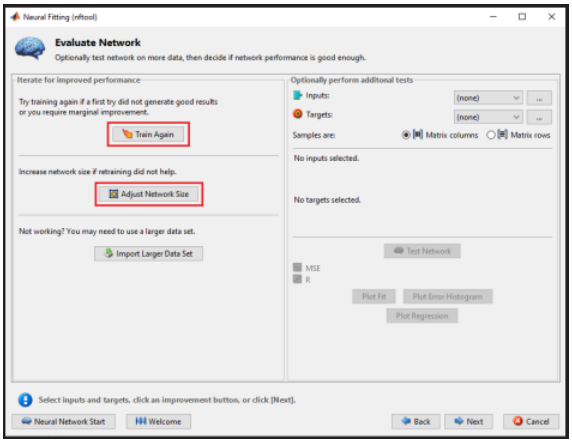

Figure 12: Network evaluation in neural net fitting tool

3. Results and Discussion

3.1. *Database for Experimental Results of Levulinic Acid Yield with Sugar Content in OPS & Matured Coconut Water*

Due to the current exceptional situation caused by the COVID-19 pandemic, the experimentation was impossible to be carried out as planned. Hence, a similar experimentation data sourced from another journal was used in this experiment. The data proposed for the simulation of artificial neural network (ANN) is obtained from a journal from Ramli et al. 2014. In this research journal, the authors produced levulinic acid (LA) from oil palm empty fruit bunch (OPEFB) using ionic liquid as catalyst.

The following journal applied respond surface methodology (RSM) and ANN to find the optimum process parameters for the production of LA using Statsoft Statistica 8.0 software. The parameters studied in this experiment was amount of sugar (g/L), catalyst loading (g), temperature ($^{\circ}C$) and reaction time (hr). The synthesis of LA was carried out in a 100 ml Schott bottle immersed in silicon oil. The data from this journal was selected because the feedstock used was originated from oil palm tree, whereby the juice was extracted from empty fruit bunch (EFB). Table III is the data used as inputs and output for the ANN models developed in this study.in the nomenclature.

Table 3: Experimental Data Extracted From External Journal [56]

Run	**Variables**				**Response**
	Temperature (°C)	Reaction time (h)	Amount of sugar (g/L)	Catalyst loading (g)	LA yield (%)
1	130	3	0.2	6	15.7
2	130	5	0.2	6	21.3
3	170	3	0.2	6	23.2
4	170	5	0.2	6	19.6
5	150	4	0.1	2	18.7
6	150	4	0.1	10	23.6
7	150	4	0.3	2	19.2
8	150	4	0.3	10	23.7
9	130	4	0.2	2	17.2
10	130	4	0.2	10	22.3
11	170	4	0.2	2	21.8
12	170	4	0.2	10	20.1
13	150	3	0.1	6	23.7
14	150	3	0.3	6	21.3
15	150	5	0.1	6	21.8

16	150	5	0.2	6	23.7
17	130	4	0.1	6	16.8
18	130	4	0.3	6	21.7
19	170	4	0.1	6	23.2
20	170	4	0.3	6	18.6
21	150	3	0.2	2	20.4
22	150	3	0.2	10	23.3
23	150	5	0.2	2	22.9
24	150	5	0.2	10	22.3
25	150	4	0.2	6	25.7
26	150	4	0.2	6	25.2
27	150	4	0.2	6	25.3

However, the catalyst used for the production of LA in Ramli et al's 2014 journal differs from this study. Ramli et al. 2014 used 1-ethyl-3-methylimidazolium chloride ionic liquid, [EMIM][Cl] to produce LA whereas in this study, the catalyst is supposedly ferric (iii) chloride ( ). Ionic liquid has proven to be a much efficient catalyst compared to heterogenous catalyst such as in terms of catalytic activity and yield just as mentioned in Chapter 2.

That said, the experiment conducted by the authors showed that the highest LA yield obtained was 26 %. It was assumed the LA yield obtained using OPS and matured coconut water with $FeCl_3$ as catalyst will be similar as to the aforementioned journal with the base that the amount of sugar in both combination of sugar composition from inner part of OPS and matured coconut water is much higher than OPEFB alone. Table 4, Table 5 and Table 6 represents the sugar composition in OPEFB, OPS and matured coconut water respectively.

Table 4: Sugar Composition in Opefb [51]

Free Sugars	**Amount (g/L)**
Glucose	48.19
Fructose	8.48
Sucrose	11.91
Total	68.58

Table 5: Sugar Composition in Ops [39]

Free Sugars	**Part (g/L)**		
	Inner	**Middle**	**Outer**
Sucrose	6.5	3.0	1.9
Glucose	85.2	52.2	13.1
Fructose	4.1	3.1	2.1
Xylose	0.7	0.8	1.4
Galactose	0.9	0.8	1.0
Rhamnose	0.4	0.5	0.5
Others	0.3	0.1	0.1
Total	98.1	60.5	20.1

Table 6: Sugar Composition in Matured Coconut Water [53]

Free Sugars	**Age of matured coconut (months) (g/L)**		
	5 to 6	**8 to 9**	**>12**
Sucrose	0.00085	0.00636	0.01437
Glucose	0.03543	0.02996	0.01906
Fructose	0.03904	0.03252	0.02148
Total	0.07532	0.06884	0.05491

To further highlight the importance of amount of free sugars in biomass for LA production, Table 7 summarises the highest LA yield obtained from two different studies that used two different

feedstocks to produce LA. It can be seen that the amount of sugar in OPEFB is 1248 times more than matured coconut water. Therefore, the OPEFB was able to yield LA much more than matured coconut water. To sum this section, higher amount of sugar in feedstock will yield higher amount of LA.

Table 7: Comparison Of La Yield From Two Different Sources

Feedstock	**Amount of sugar (g/L)**	**Experimental LA yield (%)**	**Ref.**
Matured coconut water	0.05491	0.135	[31]
OPEFB	68.58	25.7	[55]

3.2. *Development of Levenberg-Marquardt & Bayesian Regularization Based-ANN Models*

A feed-forward multi-layer ANNs were developed using neural network fitting tool (nftool) in MATLAB R2017a. The input layer has four neurons corresponding to temperature (130 - 170), reaction time (3 - 5 h), amount of sugar (0.1 - 0.3 g/L) and catalyst loading (6 - 10 g) while the output has one neuron that corresponds to LA yield. The training function of the developed ANN updates the weights and biases according to the selected training algorithm. For this study, the training algorithm is back-propagation. Back-propagation algorithm adjusts the weights in the network by following negative direction of the gradient from sum of squared errors (SSEs) with respects to the weight variables [61].

The weights and biases in an ANN is actually a function to the number of neurons in the hidden layer. So, determining the optimum number of hidden neurons in ANN is an important strategy to obtain a network architecture that best fits the data. Therefore, different number of neurons were tested to obtain the optimal number of neurons for both Levenberg-Marquadt artificial neural network (LMANN) and Bayesian regularization artificial neural network (BRANN) algorithm. Each of the network was trained independently until a satisfactory performance results was achieved.

The optimum performance results for each network was selected when minimum MSE and maximum was achieved. MSE is the average squared difference between output and target data. Lower values of MSE indicate higher accuracy of network. Correlation coefficient (R^2) is a measure of the relationship between output and target. A value close to 1 shows that output has a stronger relationship to the target data [65]. The performances for each network developed is tabulated in Table 8.

Table 8: The Performance Data For Lmann & Brann For Determination Optimal Number Of Neuron

LMANN	**MSE**	R^2	**BRANN**	**MSE**	R^2
3	1.8703	0.88419	3	6.55250	0.41047
5	2.5705	0.90825	4	6.36800	0.41622
7	2.5418	0.86242	6	5.32440	0.49608
8	2.7273	0.95291	8	5.26750	0.42245
10	4.4150	0.86906	10	5.97270	0.41959
13	2.3674	0.92001	11	5.95540	0.45940
15	6.2259	0.88550	14	5.69490	0.42399
21	**0.096322**	**0.93762**	18	6.59580	0.41748
23	2.5214	0.90524	22	0.29967	0.84158
25	2.8071	0.90076	**25**	**0.18581**	**0.91968**

Levenberg-Marquadt (LM) algorithm is an efficient algorithm used in deep learning and it is vastly employed in neural network. LM is essentially an iterative method that locates local minimum of multivariate function which is expressed in the SSEs of non-linear, real-valued functions [9]. LM's curve-fitting method is a combination of Gauss-Newton and gradient descent method (GDM). So, when the solution is far from local minimum, the algorithm behaves like gradient descent method; slow convergence. In contrast, when the solution is close to local minimum, algorithm behaves like Gauss-Newton; fast convergence.

Based on Table VIII, all of the LMANNs portrayed descent prediction results on estimating the LA yield. Even though 8 hidden neurons in LMANN gave the highest than the rest, the optimum hidden neurons chosen is 21 because the network has the lowest MSE value, 0.096322 and a reliable coefficient value, 0.94. Figure 13 illustrates the training, testing and validation regression plots for 4-21-1 LMANN structure.

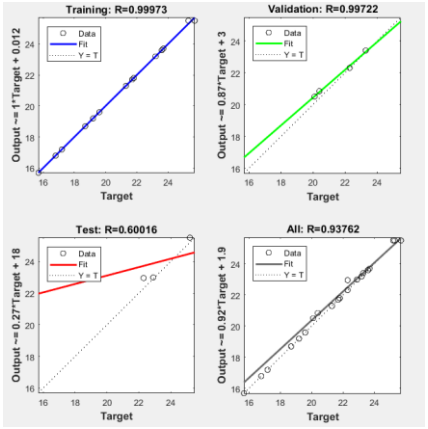

Figure 13: LMANN regression plots for 4-21-1

Bayesian regularization (BR) algorithm is a method that performs shrinkage to estimate the model with the least number of weights. A BRANN penalises large weights to prevent overfitting and to get a smoother mapping of the datasets. Thus, BRANN takes a much longer prediction time and it also performs more iterations (epochs) than other training algorithms [51]. The datasets plugged in as inputs and target variables in BRANN is divided into three parts; training, testing and validation. However, this may not be the case if the datasets are small. Instead, the data will be segregated for training and testing only.

After examining LMANN architectures, BR algorithm was applied to train the dataset with a different number of hidden neurons. For BRANNs, the performance of networks was moderate as most coefficient values were in the range of 0.40 - 0.45 while most of the MSE values were between 5.20 and 6.60. The BRANN architecture selected is 4-25-1 as it gave the lowest MSE and highest R values compared to the rest. Figure 14 illustrates the training and testing regression plots for 4-25-1 BRANN structure.

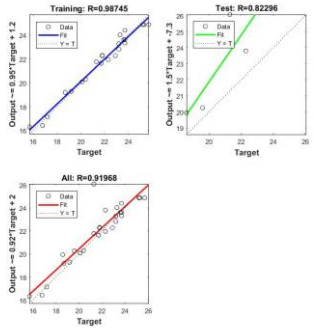

Figure 14: BRANN regression plots for 4-25-1

3.3. *Performance Evaluation of Levenberg-Marquardt & Bayesian Regularization Based-ANN Models*

Random nineteen experimental data were used for network training while the remaining eight data were shared equally for network testing and validation. Since the number of hidden neurons plays a major part in producing a reliable predictive network, different LMANN and BRANN architectures were developed to predict the LA yield. A performance comparison was carried out to compare experimental data with the predicted data to evaluate the predictive performances of the chosen LMANN and BRANN architectures respectively. Table 9 is the results for 4-21-1 LMANN and 4-25-1 BRANN.

Table 9: The Prediction Data for Lmann And Brann

Run	Response LA yield (%)		
	Experimental	4-21-1 LMANN	4-25-1 BRANN
1	15.700	15.700	16.312
2	21.300	21.300	21.772
3	23.200	23.200	23.269
4	19.600	19.600	20.261
5	18.700	18.700	19.195
6	23.600	23.600	23.606
7	19.200	19.200	19.284
8	23.700	23.700	23.622
9	17.200	17.200	17.135
10	22.300	22.314	21.953
11	21.800	21.800	22.298
12	20.100	20.511	20.080
13	23.700	23.700	23.324
14	21.300	21.300	26.064
15	21.800	21.800	22.267
16	23.700	23.700	24.399
17	16.800	16.800	16.426
18	21.700	21.700	21.633
19	23.200	23.200	22.816
20	18.600	23.383	19.944
21	20.400	20.852	20.294
22	23.300	23.410	24.021
23	22.900	22.988	22.269
24	22.300	22.953	23.801
25	25.700	25.500	24.873
26	25.200	25.500	24.873
27	25.300	25.500	24.873

Based on Table IX, the LMANN has portrayed more accurate result than BRANN. About two-thirds of LMANN data were correctly predicted by the 4-21-1 network. The remaining one-third predicted data has a deviation of less than 0.02 % from the actual experimental data. As for 4-25-1 BRANN, the accuracy of the predicted LA yield data was excellent too. The deviation of the predicted data is less than 1 % from the actual data. Figure 15 exhibits the comparative parity plots for LMANN and BRANN predictions against the actual experimental values.

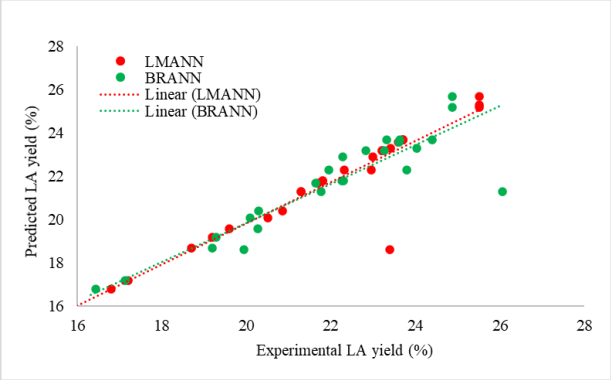

Figure 15: The parity plots for LMANN and BRANN

It is visible that the predicted values from LMANN is scattered densely to its linear regression line. As for BRANN, the predicted values are rather scattered a bit further from its regression line. The plot provides a better visual that LMANN has a better predictive capacity in comparison to BRANN. Next, the abilities of both networks were statistically evaluated in terms of correlation coefficient (R^2). The

LMANN has a R^2 value of 0.93762 while BRANN was 0.91968 which shows that LMANN has a stronger relationship with the experimental data.

To further statistically evaluate the 4-21-1 LMANN architecture developed, the model is compared with the LMANN developed by Ramli et al. 2014, which used the same dataset as inputs and outputs for the ANN training. Ramli et al. created multiple LMANN models to obtain the best network by varying the number of hidden neurons from 3 - 10. However, the author did not specify the optimum LMANN architecture but it is mentioned that the best network gave a of 0.79.

Table 10: Comparison of LMANN Architectures

LMANN		
Source	Present study	Ramli et al. 2014
R^2	0.94	0.79
MSE	0.0963	1.98

Between the both LMANN models in Table 10, LMANN model created in the present study has a higher value than the LMANN created by Ramli et al. 2014. Therefore, LMANN in the present study shares a stronger relationship with the experiment than the LMANN created by Ramli et al. 2014. Possible reason of lower value is due to the insufficient number of hidden neurons used by Ramli et al. 2014. While it is not specified by Ramli et al, the hidden neurons used by the author is definitely within the range of 3 -10 whereas in this present study, 21 hidden neurons gave a better . To sum this section, 21 hidden neurons is ideal for the experimentation dataset listed in Table 3.

3.4. *Conclusion of study*

In this study, LMANN model has proven to have the most accurate prediction and stronger relationship with the experimental data. Since BR method has the ability to remove large weights using the penalty term, it is supposed to deliver better prediction results than LM algorithm. However, it has proven to be otherwise in this study. All of the BRANNs architecture were giving high MSEs and low R^2 values except for the last two architectures. Overall, the relationship between BRANNs predicted data and experimental data were rather random and has large deviation from each other.

LM's method focusses on minimising the sum of squared errors (SSEs) and produce a network with fast convergence [38] All of LMANNs gave satisfactory performances whereby the MSE values were kept below and coefficient value of more than 0.88. It can be concluded that LM algorithm is suitable for the datasets in Table 4. Despite LMANN demonstrated high prediction accuracy, the deviation rates were possibly caused by insufficient data. To reduce the deviation rates and to build a much more stable network, more datasets should be employed to build a reliable predictive network.

5. Conclusion and Recommendation

4.1 Conclusions

Levulinic acid (LA) is a beneficial compound that can be used as derivatives for various applications. It is synthesised from biomass with the help of acid catalyst or ionic liquids. Popular catalysts includes sulphuric acid (H_2SO_4) and chromium chloride ($CrCl_3$), however in-depth study is being carried out to utilise ionic liquid for the production of LA. In the present study, juice from oil palm biomass (trunk and empty fruit bunch) and coconut water using $FeCl_3$ catalyst was used to produce LA. Given the current exceptional situation caused by COVID-19, the experimentation procedure did not take place. The project was completed via AI approach using artificial neural network (ANN). Experimental data used to develop the ANN models were obtained from a journal that produced LA using oil palm biomass.

The combination of oil palm biomass and matured coconut water as feedstocks has the potential to produce (LA). The amount of sugar in both feedstocks are sufficient to be synthesised to levulinic acid with the aid of $FeCl_3$. A conducive study was carried out to compare different models of artificial neural network (ANN) to predict the yield of LA. The ANN models were developed based on two different types of algorithms, each tested with different number of neurons.

Number of neurons essentially controls the weights and biases in the network. It is a very important strategy to determine the optimal number of neurons in the hidden layer of ANN. An ANN model with insufficient hidden neurons may fail to grasp the relationship of input and target variables. On the contrary, ANN with excess hidden neurons may lead to over-regularization. Hence, getting the optimal number of neurons in the hidden layer will give a good generalization network.

Besides, the type of training algorithm selected also determines the accuracy and relationship of network with the experimental data. The suitable training algorithm along with the optimum number of hidden neurons can produce an efficient and accurate ANN model. For this study, Levenberg-Marquardt (LM) and Bayesian regularisation (BR) were assessed with a range of 3 to 25 neurons in the hidden layer. The LMANN and BRANN models were analysed statistically based on MSE and R^2.

For a non-linear and small dataset, LMANN has a better predictive ability than BRANN. It is revealed that LMANN with 21 neurons in the hidden layer produced MSE that is 9.6×10^{-2} and an overall R^2 of 0.94, whereas BRANN with 25 neurons only gave MSE of 1.8×10^{-1} and R^2 of 0.92. Therefore, it can be concluded that the predicted LA yield from LMANN is accurate and have a stronger relationship to the experimental data. To conclude the finding in this study, LM algorithm is suitable for small datasets and ANN is an excellent prediction tool for non-linear and simple datasets.

4.2 Recommendations

The neural network structure could be improved by employing large datasets. A network with inadequate input often makes a simple model that leads to poor generalization. This is probably what caused the BRANN architectures to have poor MSE and values. So a recommendation would be to perform more experiment runs to collect adequate datasets to get more points in regards to using oil palm biomass as feedstocks.

Another suggestion to improve ANN would be to tune the hyperparameters. Hyperparameters in ANN essentially controls how the network fits the model to the data. Changing the values of the hyperparameters does not only generalize the data well but it also produces a much reliable network. Hyperparameters in ANN are number of neurons in hidden layer, learning rate and stochastic gradient descent and can be tuned using Bayesian Optimisation.

Acknowledgement

My deepest gratitude goes firstly to my supervisor, Prof. Madya Dr. Norliza Abd. Rahman, who expertly guided me through this journey and introduced me to exploring different aspects of Chemical Engineering. Her enthusiasm towards this project has been truly overwhelming and keeps me constantly engaging to my work. Her personality is so warm which makes it easier for our communication. I would also like to appreciate Dr Rafidah from FRIM for her advices and inputs regarding this research project.

References

1. Alipour, S. & Omidvarborna, H. 2017. Enzymatic and catalytic hybrid method for levulinic acid synthesis from biomass sugars. J. Clean. Prod. 143: 490–496.
2. Anon. 2015. Levulinic acid market to grow at 5.7% CAGR from 2014 to 2020. Grand View Research.
3. Anon. 2016. The Main Types of Biomass Energy. The Earth Project. 16 February.
4. Anon. 2017. Global Levulinic Acid Market Analysis & Trends - Industry Forecast to 2025. Research and Markets.
5. Antonetti, C., Bonari, E., Licursi, D., Di Nasso, N.N., Galletti, A.M.R., Cravotto, G. & Chemat, F. 2015. Hydrothermal conversion of giant reed to furfural and levulinic acid: Optimization of the process under microwave irradiation and investigation of distinctive agronomic parameters. Molecules 20: 21232–21353.
6. Badgujar, K.C., Wilson, L.D. & Bhanage, B.M. 2019. Recent advances for sustainable production of levulinic acid in ionic liquids from biomass: Current scenario, opportunities and challenges. Renewable and Sustainable Energy Reviews 102: 266 – 284.
7. Badgujar, K.C., Wilson, L.D. & Bhanage, M. 2019. Recent advances for sustainable production of levulinic acid in ionic liquids from biomass: Current scenario, opportunities and challenges. Renewable and Sustainable Energy Reviews 102 : 266 – 284.

8. Balan V. 2014. Current challenges in commercially producing biofuels from lignocellulosic biomass. ISRN Biotechnol.
9. Baptista, D. & Morgado-Dias, F. 2012. Comparing Different Levenberg-Marquardt algorithm implementations.
10. Basu, P. 2013. Biomass Conversion. Biomass Gasification, Pyrolysis and Torrefaction (2): 7-15, 2013
11. Bio-based Chemicals: Value Added Products from Biorefineries. 2004. IEA Bioenergy.
12. Bozell, J.J. & Petersen G.R. 2010. Technology development for the production of biobased products from biorefinery carbohydrates - the US Department of Energy's "Top 10" revisited. Green Chem 12(4): 539–554.
13. Bozell, J.J., Moens, L., Elliot, D.C., Wang, Y., Neuenscwander, G.G., Fitzpatrick, S.W., Bilski, R.J. & Jarnefeld, J.L. 2000. Production of levulinic acid and use as a platform chemical for derived products. Resources, Conservation and Recycling 28 (4): 227- 239.
14. BP. 2012. Statistical review of world energy. In: British Petroleum.
15. Cha, J.Y. & . Hanna, M.A. 2002. Levulinic acid production based on extrusion and pressurized batch reaction. Industrial Crops and Products 16: 109 - 118.
16. Chen, H., Yu, B. & Jin, S. 2011. Production of levulinic acid from steam exploded rice straw via solid superacid, S2O8 2-/ZrO2-SiO2-Sm2O3. Bioresour. Technol. 102: 3568–3570.
17. Chen, S.S., Maneerung, T., Tsang, D.C.W., Ok, Y.S. & Wang, C.H. 2017.Valorization of biomass to hydroxymethylfurfural, levulinic acid, and fatty acid methyl ester by heterogeneous catalysts. Chem. Eng. J. 328: 246–273.
18. Chong, C., Ni, W., Ma, L., Liu, P. & Li, Z. 2015. The Use of Energy in Malaysia: Tracing Energy Flows from Primary Source to End Use. Energies 8: 2828-2866.
19. Dalibard, C. 1999. Overall view on the tradition of tapping palm trees and prospects for animal production. Livest. Res. Rural Dev. 11: 1-39.
20. Daniel, K. & Markus, A. 2010. Transfer hydrogenation of levulinic acid under hydrothermal conditions catalyzed by sulfate as a temperature-switchable base. Green Chemistry 4 (12): 656-660
21. DOE. Top Value Added Chemicals from Biomass - Results of Screening for Potential Candidates from Sugars and Synthesis Gas. 2004. United States: Department of Energy: Energy Efficiency and Renewable Energy (1)1 -76.
22. Enerdata. 2018. World Energy Statistical Yearbook 2019.
23. Fernandes, D.R., Rocha, A.S., Mai, E.F., Mota, C.J.A. & Teixeira Da Silva, V. 2012. Levulinic acid esterification with ethanol to ethyl levulinate production over solid acid catalysts. Appl. Catal. A General 425-426, 199-204.
24. Galletti, AMR., Antonetti, C., Luies, V., Licursi, D. & Nasso, N.N. 2012. Levulinic acid from waste. Bioresources 7 (2):1824–1835.
25. Girisuta, B., 2007. Levulinic Acid from Lignocellulosic Biomass. University of Groningen, 64 – 83.
26. Girisuta, B., Danon, B., Manurung, R., Janssen, L.P. & Heeres, H.J. 2008. Experimental and kinetic modeling studies on the acid-catalysed hydrolysis of the water hyacinth plant to levulinic acid. Bioresor. Techno. l99(17):8367–8375.
27. Guzman, D. 2013. Segetis starts levulinic acid pilot production. Green Chemicals Blog.
28. Guzman, D. 2015. Levulinic acid commercialization expands. Green Chemicals Blog.
29. Hartono, C.D., Marlie, K.J., Putro, J.N., Soetardjo, F.E., Ju, Y.H., Sirodj, D.A.N. & Ismadji, S. 2016. Levulinic acid from corncob by subcritical water process. Int. J. Ind. Chem. 7: 401–409.
30. Husin, M. 2000. Utilization of oil palm biomass for various wood-based and other products. Advances in oil palm research. Malaysian Palm Oil Board. 2000: 1346-1412.
31. Ismail, F., Rahman, N.A. & Esa, F. 2019. Catalytic production of levulinic acid from matured coconut water. IOP Conference Series: Materials Science and Engineering 709: 1 – 10.
32. Isoni, V., Kumbang, D., Sharratt, P. & Khoo, H. 2004. Biomass to Levulinic Acid: A techno-economic analysis and sustainability of biorefinery processes in Southeast Asia. Journal of Environmental Management 214(1): 267 - 275.
33. Jeong, H., Jang, S.K., Hong, C.Y., Kim, S.H., Lee, S.Y., Lee, S.M., Choi, J.W. & Choi, I.G. 2017. Levulinic acid production by two-step acid-catalyzed treatment of Quercus mongolica using dilute sulfuric acid. Bioresour. Technol. 225: 183–190.

34. Jeong, H., Park, S.Y., Ryu, G.H., Choi, J.H., Kim, J.H., Choi,W.S., Lee, S.M., Choi, J.W. & Choi, I.G. 2018. Catalytic conversion of hemicellulosic sugars derived from biomass to levulinic acid. Catal. Commun. 117: 19–25.
35. Kamm, B. & Kamm, M. 2004. Principles of biorefineries. Applied Microbiology and Biotechnology 64: 137-145
36. Kang, S. & Yu, J. 2016. An intensified reaction technology for high levulinic acid concentration from lignocellulosic biomass. Biomass Bioenergy 95: 214–220.
37. Karadurmuş, E., Taşkın, N., Göz, E., Yüceer, M. 2019. Prediction of Bromate Removal in Drinking Water Using Artificial Neural Networks. Ozone: Science and Engineering 41: 118–127.
38. Kayri, M. 2016. Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. Mathematical and Computational Applications 21:1-11.
39. Kosugi, A. & Tanaka, R. & Magara, K. & Murata, Y. & Arai, T. & Sulaiman, O. & Hashim, R. & Abdul Hamid, Z. & Yahya, M. & Yusof, M. & Ibrahim, W. & Mori, Y. 2010. Ethanol and Lactic Acid Production Using Sap Squeezed from Old Oil Palm Trunks Felled for Replanting. Journal of bioscience and bioengineering 110: 322-355.
40. Kupiainen, L. Ahola, J. & Tanskanen, J. 2011. Kinetics of glucose decomposition in formic acid. Chem Eng Res Des 89 (12): 2706 - 2713.
41. Lange, JP. 2007. Lignocellulose Conversion: An Introduction to Chemistry, Process and Economics. Catalysis for Renewables: from Feedstock to Energy Production: 21–51.
42. Leonard, R.H. 1956. Levulinic Acid as a Basic Raw Material. Industrial and Engineering Chemistry 48: 1331 – 1341.
43. Li, H., Zhang, Q., Liu, X., Chang, F., Hu, D., Zhang, Y., Xue, W. & Yang, S. 2013. InCl3-ionic liquid catalytic system for efficient and selective conversion of cellulose into 5-hydroxymethylfurfural. RSC Advance 3: 3648-3654
44. Loow, Y.L., Wu, T.Y., Tan, K.A., Lim, Y.S., Siow, L.F., Md Jahim, J., Mohammad, A.W. & Teoh, W.H. 2015. Recent advances in the application of inorganic salt pretreatment for transforming lignocellulosic biomass into reducing sugars. J. Agricultural Food Chemistry (63): 8349-8363.
45. M. Hoogwijk , A. Faaij , R. van den Broek , G. Berndes , D. Gielen & W. Turkenburg , Biomass and Bioenergy , 2003 , 25 , 119 – 133.
46. Morone, A., Apte, M. & Pandey, R.A. 2015. Levulinic acid production from renewable waste resources: Bottlenecks, potential remedies, advancements and applications. Renewable and Sustainable Energy 51: 548 – 565.
47. N. Mosier , C. Wyman , B. Dale , R. Elander , Y. Y. Lee , M. Holtzapple & M. Ladisch , Bioresource Technology , 2005 , 96 , 673 – 686.
48. Nagi, S. & Pandey, A.K. 2015. Chapter 8 - Ionic Liquid Pretreatment. Pretreatment of Biomass: Processes and Technologies: 137-155.
49. Nagy, Z.K. 2007. Model based control of a yeast fermentation bioreactor using optimally designed artificial neural networks. Chemical Engineering Journal 127: 95–109
50. Nasrah, N.S.M. 2017. Statistical Optimization for Biobutanol Production by Clostridium acetobutylicum ATCC 824 from Oil Palm Frond (OPF) Juice Using Response Surface Methodology. EDP Sciences.
51. Okut, H. 2016. Bayesian Regularized Neural Networks for Small n Big p Data, Artificial Neural Networks - Models and Applications.
52. Peng, L., Lin, L., Zhang, J., Zhuang, J., Zhang, B. & Gong, Y. 2010. Catalytic conversion of cellulose to levulinic acid by metal chlorides. Molecules (15): 5258-5272.
53. Prades, A., Dornier, M., Diop, N. & Pain, J. P. .2012. Coconut water uses, composition and properties: a review. Fruits 67(2): 87–107.
54. Rackemann, D.W. 2014. Production of Levulinic Acid and Other Chemicals from Sugarcane Fibre. Faculty of Science and Technology: Queensland University of Technology (QUT).
55. Ramli, N.A.S. & Amin, N.A.S. 2014. Catalytic hydrolysis of cellulose and oil palm biomass in ionic liquid to reducing sugar for levulinic acid production. Fuel Process. Technology (128): 490-498.
56. Ramli, N.A.S. & Amin, N.A.S. 2015. A new functionalized ionic liquid for efficient glucose conversion to 5-hydroxymethyl furfural and levulinic acid. Journal of Molecular Catalyst A: Chemical 407: 113 – 121.

57. Ramli, N.A.S. & Amin, N.A.S. 2016. Optimization of biomass conversion to levulinic acid in acidic ionic liquid and upgrading of levulinic acid to ethyl levulinate. BioEnergy Resource :1-14.
58. Ramli, N.A.S. & Amin, N.A.S. 2017. Optimization of Biomass Conversion to Levulinic Acid in Acidic Ionic Liquid and Upgrading of Levulinic Acid to Ethyl Levulinate. Bioenergy Res. 10: 50–63.
59. Ren, H., Girisuta, B., Zhou, Y. & Liu, L. 2015. Selective and recyclable depolymerization of cellulose to levulinic acid catalysed by acidic ionic liquid. Carbohydrate Polymer 117: 569 – 576.
60. Ren. H., Zhou, Y. & Liu, L. 2013. Selective conversion of cellulose to levulinic acid via microwave-assisted synthesis in ionic liquids. Bioresource Technol. 129: 616 - 619.
61. A. Reynaldi, S. Lukas, H. Margaretha. 2012.Backpropagation and Levenberg-Marquardt algorithm for training finite element neural network. 2012 Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation, IEEE: 89-94
62. Ritter, S. 2006. Biorefinery gets ready to deliver the goods. Science/Technology 84 (34): 47.
63. Rivas, S., González-Muñoz, M.J., Vila, C., Santos, V. & Parajó, J.C. 2013. Manufacture of levulinic acid from pine wood hemicelluloses: A kinetic assessment. Ind. Eng. Chem. Res. 52: 3951–3957.
64. Saidina Amin, Nor Aishah & Ramli, N.A.S., & Ya'aini, N. 2013. Comparison of Response Surface Methodology and Artificial Neural Network for Optimum Levulinic Acid Production from Glucose, Empty Fruit Bunch and Kenaf. Special Issue of International Journal of Nano and Biomaterials: Green Materials and Technologies 5 (1): 59 - 74.
65. Sapna, S., Tamilarasi, A., & Kumar, M. P. 2012. Backpropagation learning algorithm based on Levenberg Marquardt Algorithm. Comp Sci Inform Technol (CS and IT) 2: 393-398.
66. Schwartz D. 2015. Incitor's high performance fuel from algae biomass. Algae Industry.
67. Shen, Y., Kui-Sun, J., Wang, Bo., Xu, F. & Sun, R.C. 2015. One-pot synthesis of levulinic acid from cellulose in ionic liquids. Bioresource Technology 2015 192: 812 – 816.
68. Shen, Y., Xu, Y., Sun, J., Wang, B., Xu, F. & Sun, R. 2014. Efficient conversion of monosaccharides into 5-hydroxymethylfurfural and levulinic acid in InCl3-H2O medium. Catalyst Community 50: 17-20.
69. Signoretto, M., Taghavi, S., Ghedini, E. & Menegazzo, F. 2019. Catalytic Production of Levulinic Acid (LA) from Actual Biomass. Molecules 24(15): 2415 – 2760.
70. Sikarwar, V., Zhao, M., Fennell, P.S., Shah, N. & Anthony, E. 2017. Progress in biofuel production from gasification. Progress in Energy and Combustion Science 61: 189-248.
71. Tan-Soetedjo, J.N.M., Van de Bovenkamp, H.H., Abdila, R.M., Rasrendra, C.B., Ginkel, J.V. & Heeres, H.J. 2017. Ind. Eng. Chem. Res. 56: 13228 – 13239
72. Tao, D.J., Lu, X.M., Lu, J.F., Huang, K., Zhou, Z., Wu Y.T. 2011. Noncorrosive ionic liquids composed of [HSO4] as esterification catalysts. Chem. Eng. J. (171): 1333-1339.
73. Tao, F., Song, H. & Chou, L.J. 2011. Catalytic conversion of cellulose to chemicals in ionic liquid. Carbohydrate research 346 (1): 58-63
74. Tao, F., Song, H., Chou, L. 2011. Hydrolysis of cellulose in SO3H-functionalized ionic liquids. Bioresource Technology. 102: 9000 - 9006.
75. Tiong, Y.W., Yap, C.L., Gan, S. & Yap, W.S.P. 2017. One-pot conversion of oil palm empty fruit bunch and mesocarp fiber biomass to levulinic acid and upgrading to ethyl levulinate via indium trichloride-ionic liquids. Journal of Cleaner Production 168: 1251 – 1261.
76. Vijander Singh, Indra Gupta, H.O. Gupta. 2007. ANN-based estimator for distillation using Levenberg–Marquardt approach. Engineering Applications of Artificial Intelligence 20 (2):249-259,
77. Wilson, K. & Lee, A.F. 2014. Bio-Based Chemicals From Biorefining: Carbohydrate Conversion And Utilisation. Advances in Biorefineries: Biomass and Waste Supply Chain Exploitation: 624 – 658.
78. Ya'aini, N., & Amin, N.A.S. 2013. Catalytic conversion of lignocellulosic biomass to levulinic acid in ionic liquid. Bioresources (8): 5761-5772.
79. Ya'aini, N., Amin, N.A.S. & Asmadi, M. 2012. Optimization of levulinic acid from lignocellulosic biomass using a new hybrid catalyst. Bioresour. Technol. 116: 58–65.
80. Yamada, H., Tanaka, R., Sulaiman, O., Hashim, R., Hamid, Z., Yahya, M., Kosugi, A., Arai, T., Murata, Y. & Nirasawa, S. 2010. Old oil palm trunk: A promising source of sugars for bioethanol production. Biomass and Bioenergy 34(11): 1608-1613.
81. Yan, K., Jarvis, C., Gu, Jing. & Yan, Y. 2015. Production and catalytic transformation of levulinic acid: A platform for speciality chemicals and fuels. Renewable and Sustainable Energy Reviews 51: 986 – 997.

82. Yan, L., Yang, N., Pang, H. & Liao, B. 2008. Production of levulinic acid from bagasse and paddy straw by liquefaction in the presence of hydrochloride acid. Clean: Soil Air Water 36(2):158–163.
83. Yong., W.T., Chiew., L.Y., Suyin., G. & Winnie., S.P.Y. 2017. One-pot conversion of oil palm empty fruit bunch and mesocarp fibre biomass to levulinic acid and upgrading to ethyl levulinate via indium trichloride-ionic liquids. Journal of Cleaner Production 168(1): 1251 – 1261.
84. Yosuke, M., Tatsuya, S., Hiroyuki, S., Isao, S. & Kazuhiro, M. 2014. Effective Production of Levulinic Acid from Biomass through Pretreatment Using Phosphoric Acid, Hydrochloric Acid, or Ionic Liquid. Ind. Eng. Che,. Res 53 (1): 11611-11621.
85. Zanghellini, A.L. 2013. Fermentation routes for the production of levulinic acid, levulinate esters valerolactone, and derivatives thereof. European patent 2611764.
86. Zhou, C., Zhao, J., Yagoub, A.E.A., Ma, H., Yu, X., Hu, J., Bao, X. & Liu, S. 2017. Conversion of glucose into 5-hydroxymethylfurfural in different solvents and catalysts: Reaction kinetics and mechanism. Egyptian Journal of Petroleum 26: 477 – 487.
87. Zhou, X., Zhang, Z., Liu, B., Xu, Z. & Deng, K. 2013. Microwave-assisted rapid conversion of carbohydrates into 5-hydroxymethylfurfural by $ScCl_3$ in ionic liquids. Carbohydrate Research 375: 68-72.