COMPUTATIONAL AND EXPERIMENTAL STUDIES ON THE REACTION MECHANISM OF BIO-OIL COMPONENTS WITH ADDITIVES FOR INCREASED STABILITY AND FUEL QUALITY

JIA WEN CHONG, MEng.

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

SEPT 2022

ABSTRACT

As one of the world's largest palm oil producers, Malaysia encountered a major disposal problem as vast amount of oil palm biomass wastes are produced. To overcome this problem, these biomass wastes can be liquefied into biofuel with fast pyrolysis technology. However, further upgradation of fast pyrolysis bio-oil via direct solvent addition was required to overcome it's undesirable attributes. In addition, the high production cost of biofuels often hinders its commercialisation. Thus, the designed solvent-oil blend needs to achieve both fuel functionality and economic targets to be competitive with the conventional diesel fuel.

In this thesis, a multi-stage computer-aided molecular design (CAMD) framework was employed for bio-oil solvent design. In the design problem, molecular signature descriptors were applied to accommodate different classes of property prediction models. However, the complexity of the CAMD problem increases as the height of signature increases due to the combinatorial nature of higher order signature. Thus, a consistency rule was developed reduce the size of the CAMD problem. The CAMD problem was then further extended to address the economic aspects via fuzzy multi-objective optimisation approach.

Next, a rough-set based machine learning (RSML) model has been proposed to correlate the feedstock characterisation and pyrolysis condition with the pyrolysis bio-oil properties by generating decision rules. The generated

decision rules were analysed from a scientific standpoint to identify the underlying patterns, while ensuring the rules were logical. The decision rules generated can be used to select optimal feedstock composition and pyrolysis condition to produce pyrolysis bio-oil of targeted fuel properties.

Next, the results obtained from the computational approaches were verified through experimental study. The generated pyrolysis bio-oils were blended with the identified solvents at various mixing ratio. In addition, emulsification of the solvent-oil blend in diesel was also conducted with the help of surfactants. Lastly, potential extensions and prospective work for this study have been discuss in the later part of this thesis. To conclude, this thesis presented the combination of computational and experimental approaches in upgrading the fuel properties of pyrolysis bio-oil. As a result, high quality biofuel can be generated as a cleaner burning replacement for conventional diesel fuel.

LIST OF PUBLICATIONS

Referred Journals

Radhakrishnapany, K. T., Wong, C. Y., Tan, F. K., **Chong, J. W.**, Tan, R. R., Aviso, K. B., Janairo, J. I. B., & Chemmangattuvalappil, N. G. (2020). Design of fragrant molecules through the incorporation of rough sets into computer-aided molecular design. *Molecular Systems Design & Engineering*, *5*(8), 1391–1416, DOI: 10.1039/D0ME00067A

Chemmangattuvalappil, N. G., Ng, D. K. S., Ng, L. Y., Ooi, J., **Chong, J. W.**, & Eden, M. R. (2020). A Review of Process Systems Engineering (PSE) Tools for the Design of Ionic Liquids and Integrated Biorefineries. *Processes*, 8(12), 1678, DOI: 10.3390/pr8121678

Chong, J. W., Chemmangattuvalappil, N. G., & Thangalazhy-Gopakumar, S. (2021). Aviation Biofuels: Conversion Routes and Challenges. *Journal of Oil Palm, Environment & Health*, 12(5), 69–85, DOI: 10.5366/jope.2021.05

Chong, J. W., Thangalazhy-Gopakumar, S., Muthoosamy, K., & Chemmangattuvalappil, N. G. (2022). Design of bio-oil additives via molecular signature descriptors using a multi-stage computer-aided molecular design framework. *Frontiers of Chemical Science and Engineering*, 16(2), 168–182, DOI: 10.1007/s11705-021-2056-8

Chong, J. W., Ng, L. Y., Aboagwa, O. A., Thangalazhy-Gopakumar, S., Muthoosamy, K., & Chemmangattuvalappil, N. G. (2021). Computer-Aided Framework for the Design of Optimal Bio-Oil/Solvent Blend with Economic Considerations. *Processes*, 9(12), 2159, DOI: 10.3390/pr9122159

Ooi, Y. J., Aung, K. N. G., **Chong, J. W.**, Tan, R. R., Aviso, K. B., & Chemmangattuvalappil, N. G. (2022). Design of fragrance molecules using computer-aided molecular design with machine learning. *Computers & Chemical Engineering*, 157, 107585, DOI: 10.1016/j.compchemeng.2021.107585

Hatamleh, M., **Chong, J. W.**, Tan, R. R., Aviso, K. B., Janairo, J. I. B., & Chemmangattuvalappil, N. G. (2022). Design of Mosquito Repellent Molecules via the Integration of Hyperbox Machine Learning and Computer aided Molecular Design. *Digital Chemical Engineering*, 100018, DOI: 10.1016/j.dche.2022.100018

Chong, J. W., Thangalazhy-Gopakumar, S., Tan, R. R., Aviso, K. B., & Chemmangattuvalappil, N. G. (2022). Estimation of Fast Pyrolysis Bio-Oil Properties from Feedstock Characteristics Using Rough-Set Based Machine Learning. *International Journal of Energy Research*, 1-18, DOI: 10.1002/er.8201

Heng, Y. P., Lee, H. Y., **Chong, J. W.**, Tan, R. R., Aviso, K. B., Chemmangattuvalappil, N. G. (2022). Incorporating Machine Learning in Computer-Aided Molecular Design for Fragrance Molecules. *Processes,* 10 (9), 1767. DOI: 10.3390/pr10091767

Conference Proceedings

Chong, J. W., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N. G. (2022). Design of Bio-Oil Solvents using Multi-Stage Computer-Aided Molecular Design Tools, in: Yoshiyuki Yamashita, Manabu Kano (Eds.), *14th International Symposium on Process Systems Engineering*. Elsevier, Kyoto, Japan, pp. 199–204. DOI: 10.1016/B978-0-323-85159-6.50033-6

Conference Presentations

Chong, J.W., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G. Development of CAMD tools to design solvents

to upgrade bio-oil quality. Presented at: *9th Asian Symposium on Process Systems Engineering*; 4th – 6th November 2020; Taipei, Taiwan, Paper No: 210124.

Chong, J.W., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G. Multi-Stage Computer Aided-Molecular Design (CAMD) Approach in Bio-Oil Solvent Design to Upgrade Bio-Oil Quality. Presented at: *International Conference on Sustainable Biowaste Management 2021*; 12th – 15th April 2021; Hong Kong SAR, P.R. China, Paper No: Oral128.

Chong, J.W., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G. Predicting Pyrolysis Bio-Oil pH and Higher Heating Value using Rough-Set Based Machine Learning. Presented at: *International Conference Series on Alternative Fuels, Energy and Environment: Future and Challenges*; 15th – 18th October 2021; Kayseri, Turkey, Paper No: ICAFEE-2021-F-007.

Chong, J.W., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G. Design of Bio-Oil Solvents using Multi-Stage Computer-Aided Molecular Design Tools. Presented at: *14th International Symposium on Process Systems Engineering*; 19th – 23rd June 2022; Kyoto, Japan, Paper No: 83.

Chong, J.W., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G. Multi-Objective Optimisation Framework for Bio-Oil Solvent Design with Environmental and Economic Considerations. Presented at: *33rd Symposium of Malaysian Chemical Engineers*; 8th – 9th August 2022; Putrajaya, Malaysia, Paper No: 20

Chong, J.W., Ng, L.Y., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G. Design of Bio-oil/Solvent Blend with Economic Considerations Using Computer Aided Molecular Design. Presented at: *19th Asian Pacific Confederation of Chemical Engineering Congress*, 10th – 12th August 2022; Kuala Lumpur, Malaysia, Paper No: 1103

ACKNOWLEDGEMENT

First and most importantly, I would like to express my deep and sincere gratitude to my supervisors, Dr Nishanth G. Chemmangattuvalappil, Dr Suchithra Thangalazhy-Gopakumar and Dr Kasturi Muthoosamy for their invaluable advice, constructive criticism, and continuous support throughout the entire period. I am also grateful to Prof Dominic Foo for his effort and time in reviewing my annual report. I would also like to acknowledge the research ideas and discussion shared by Prof Raymond Tan, Prof Kathleen Aviso, and Dr Lik Yin Ng. My thesis has benefitted greatly from their insightful suggestion.

In addition, I would like to thank the administrative staffs and lab technicians of the Faculty of Science and Engineering, The University of Nottingham Malaysia for their assistance and support. I would also like to express my sincere gratitude to the Ministry of Higher Education Malaysia for the realisation of this research project under the Grant FRGS/1/2019/TK02/UNIM/02/1.

I would like to give special thanks to my Mom and Dad, and to my sister and brother for their continuous love and support. Thank you for your understanding and your never-ending encouragement. Your support has meant more to me than you could ever realise. Last but not least, I am grateful to my friends, Chung Jie Wong, Tan Hwei Hua and Ong Chia Leen, for the emotional support and for always being there for me.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF NOMENCLATURE

Abbreviations

AC	Ash content
AI	Artificial intelligence
ANN	Artificial neural network
ASTM	American Society for Testing and Materials
BGIT	Benson group increment theory
C	Carbon content
CAMD	Computer-aided molecular design
CI	Connectivity index
EFB	Empty fruit bunch
FC	Fixed carbon
FT-IR	Fourier transform infrared
GC	Group contribution
GC-MS	Gas chromatography-mass spectrometry
H	Hydrogen content
HHV	Higher heating value
HLB	Hydrophilic-lipophilic balance
LEM2	Learning from Example Module, Version 2
MC	Moisture content
MILP	Mixed integer linear programming
MINLP	Mixed integer non-linear programming

ML	Machine learning
MLR	Multiple linear regression
MOO	Multi-objective optimisation
N	Nitrogen content
NIST	National Institute of Standards and Technology
O	Oxygen content
O/W	Oil-in-water emulsion
PKS	Palm kernel shell
QSAR	Quantitative structure-activity relationship
QSPR	Quantitative structure-property relationship
RSM	Response surface method
RSML	Rough-set machine learning
RST	Rough set theory
S	Sulphur content
SVM	Support vector machine
TI	Topological index
VM	Volatile matter
W/O	Water-in-oil emulsion

Sets

A	Set of attributes $\{a \mid a = a_1, a_2, \dots, a_m\}$
C	Set of conditional attribute $\{c \mid c = c_1(x), c_2(x), \dots, c_n(x)\}$
D	Set of decision attributes $\{d \mid d = d_1(x), d(x), \dots, d_n(x)\}$
U	Set of objects $\{x \mid x = x_1, x_2, \dots, x_n\}$

Parameters/Variables

${}^{0}c_{m}^{v}$	Connectivity term for zeroth order of valence form for atom m
${}^{0}\chi_{m}^{v}$	Zeroth order connectivity index for atom m
${}^{h}\alpha_{g}$	Number of occurrence of atomic signatures of height h
${}^{h}\sigma_{G}(x)$	Signature of height h of atom x in the subgraph G
A^{C}	Price of competitor's product
$a_{m,n}$	Group interaction parameters of experimental phase equilibrium
A^{P}	Price of new product
C_{blend}	Cost of solvent-oil blend
C_i	Contribution of first-order group i
C_i	Cost of solvent
D_{ij}	Off diagonal elements of distance matrix
D_j	Contribution of second-order group j
E_k	Contribution of third-order group k
F_p	Flash point
g_i	Number of hydrogen atoms bonded to atom i
IGC_{50}	Relative toxicity
K_{oc}	Soil-water partition coefficient
K_{ow}	Octanol/water partition coefficient
M_j	Number of occurrence of second-order group j
n_1	Number of signature with valency of one
n_2	Number of signature with valency of two
n_3	Number of signature with valency of three
n_4	Number of signature with valency of four

N_{Di}	Number of signature with one double bond
N_i	Number of occurrence of first-order group i
N_{Mi}	Number of signature with two double bond n_1
N_{Ti}	Number of signature with one triple bond
O_k	Number of occurrence third-order group k
p_k^L	Lower bound for property k
p_k^U	Upper bound for property k
P_{sat}	Vapour pressure
q_i	Pure component molecular surface areas parameter
Q_k	Group area parameters
r_i	Pure component molecular van der Waals volume parameter
R_k	Group volume parameters
T_b	Normal boiling point
T^C	Demand of competitor's product
T_m	Normal melting point
T^P	Demand of new product
$U_{weighted\ sum}$	Overall objective function for weighted sum method
v_i	Valency of vertex i
$v_k^{(i)}$	Number of groups of type k in molecule i
V_m	Molar volume
w_i	Weighting factor assigned to objective function i
x_d^L	Lower bound for the number of signatures
x_d^U	Upper bound for the number of signatures
x_i	Composition of i in the mole fractions of the trial phase
X_m	Mole fraction of group m in the mixture

Z_i^v	Number of valence electrons for atom i
Z_i	Atomic number of atom i
z_i	Composition of i in the mole fractions of the tested phase
γ_i	Activity coefficient of component i
$\Gamma_k^{(i)}$	Residual activity coefficient of group k in pure component i
Γ_k	Group residual activity coefficient
δ_m^v	Valence delta of atom m
η_L	Dynamic viscosity
θ_i	Area fraction of component i
ϑ_m	Area fraction of group m
λ^C	Consumer's preference function of competitor's product
λ^P	Consumer's preference function of new product
$\sigma_x(C,D)$	Strength factor of the decision rule $C \rightarrow_x D$
ϕ_i	Segment fraction (volume fraction) of component i
$\psi_{m,k}$	Group interaction parameter
BCF	Bioconcentration factor
$cer_x(C,D)$	Certainty factor of the decision rule $C \rightarrow_x D$
$cov_x(C,D)$	Coverage factor of the decision rule $C \rightarrow_x D$
EC_{50}	Aquatic acute toxicity
GWP	Global warming potential
h	Height of signature
$IND_P(A)$	Indiscernibility relation of attribute A
l	Colouring sequence of molecular signature descriptor
LC_{50}	Aquatic acute toxicity
LD_{50}	Oral acute toxicity

PCO	Photochemical oxidation potential
S	Decision system
$TI(root(^{h}\sum))$	Contribution values computed for each atomic signature
v	Viscosity
$W(G)$	Wiener index of molecular graph G
Y	Total market size
α	Consumer's awareness coefficient
β	Consumer's preference coefficient
δ	Elasticity of substitution
λ	Satisfaction degree of fuzzy
ρ	Density

CHAPTER 1

INTRODUCTION

1.1. Oil Palm Biomass's Potential as Renewable Energy Source

Over the century, burning of fossil fuels had generated most of the energy required in our daily life. Even today, the fossil fuels are still serving about 80 % of the world's daily energy needs (EESI, 2021). Report have shown that in July 2022, 98.8 million barrels/day of petroleum and liquid fuels were consumed globally. This amount was forecasted to further increase up to an average of 101.5 million barrels/day by the year of 2023 (EIA, 2022a). Nevertheless, the rising demand and the increased cost of fossil fuels, along with the growing environmental issues from their uses had led to the search for renewable and sustainable fuels resources. In Malaysia, efforts had been made to elevate the renewable energy production by introducing the Malaysia Energy Plan. To date, the installed capacity of renewable energy in Malaysia was reported to be 7995 MW (MIDA, 2021). Under the 12th Malaysia Plan, it is projected for Malaysia to achieve the renewable energy capacity of 31 % by 2025 and 40 % by 2035 (MIDA, 2021).

Biofuel has been recognised as a highly preferable alternative to conventional fossil fuels for their reduced environmental impacts. Generally, the biofuel can be referred to as a renewable and clean burning liquid fuel which can be obtained from biomass feedstock, such as agricultural and industrial

wastes, energy crops and algae sources. The quality and characteristics of biomass varied regionally and are dependent on multidimensional components mechanism and drives, such as water, soil, climate etc. (Kaniapan et al., 2021). In Asia region, specifically in Malaysia and Indonesia, palm oil waste has been perceived as one of the potential biomass sources for its abundant availability.

Malaysia is one of the world's largest palm oil producers and exporters with 25.8 % of world palm oil production and 34.3 % of world exports (MPOC, 2022). In 2021, the total oil palm planted area in Malaysia reached 5.74 million hectares. In addition, production capacity of approximately 39 million tonnes of oil palm products such as crude palm oil, palm kernel, crude palm kernel oil and palm kernel cake were reported by The Malaysia Palm Oil Board (MPOB, 2021). For every tonne of crude palm oil produced, roughly 5.8 tonnes of fresh fruit bunch were used, of which 70% was removed as waste, such as press fibre (30%), empty fruit bunch (28.5%), palm kernel shell (6%), decanter cake (3%) and others (2.5%) (Pleanjai et al., 2007). As of December 2021, there were 451 operating palm oil mills in Malaysia with a total fresh fruit bunch processing capacity of 115 million tonnes per year (MPOB, 2022). In other words, a total of 80.5 million tonnes of palm oil biomass wastes were generated in Malaysia annually.

A variety of thermochemical technologies were available for palm biomass conversion, such as liquefaction, gasification, and pyrolysis where different type of biofuel can be generated. Among these available biomass conversion technologies, pyrolysis process has the advantage of being a relatively simple and inexpensive technology (Fermoso et al., 2017). With

pyrolysis process, solid palm biomass feedstock can be converted into pyrolysis bio-oil along with biochar and gaseous products as by-product. Here, the pyrolysis bio-oil, sometimes also known as the bio-crude, is the liquid product produced from the pyrolysis of biomass through rapid heating at high temperature. However, pyrolysis bio-oil is a complex mixture which mainly composed of water, fine solid particles, and various organic compounds such as hydrocarbons, acids, alcohols, ketones, aldehydes, phenols, as well as large molecular oligomers (Yang et al., 2014). These compounds often lead to poor fuel properties and performance of pyrolysis bio-oil. Problems like corrosiveness, high viscosity, low heating value, thermal and chemical instability, as well as non-miscibility with petroleum fuels were often encountered with the direct application of pyrolysis bio-oil in diesel engines or gas turbines (Khosravanipour Mostafazadeh et al., 2018). To overcome these disadvantages, a vast amount of pyrolysis bio-oil upgrading methods through physical or chemical pathway have been proposed and studied (Ansari et al., 2022; Omar et al., 2019). Direct solvent addition is one of the most popular pyrolysis bio-oil upgrading approaches as it is relatively simple and economically viable. With the addition of solvent, lower viscosity, higher stability, and homogenisation of pyrolysis bio-oil can be achieved (Yang et al., 2014). The greater higher heating value (HHV) of solvent was also proven to improve the HHV of pyrolysis bio-oil (Khosravanipour Mostafazadeh et al., 2018).

1.2. Background Problems

In the past, most of the studies focused on the experimental mixing of pyrolysis bio-oil with one or more solvents selected based on their general properties without knowing the underlying mechanisms between constituent of pyrolysis bio-oil with solvent. The conventional approach in the identification and selection of solvents often involve tedious trial-and-error experimental process within a large set of candidates. This process could be time consuming to test all potential solvents and ineffective in optimising blend performance to meet the property targets (Venkatasubramanian et al., 1994). In response to these challenges, computer-aided molecular design (CAMD) techniques were commonly employed to identify potential candidates that satisfy a set of property targets and constraints in the design and optimisation of solvents (Achenie et al., 2003). CAMD is a reverse engineering approach which predicts, estimates, and construct molecules from a given set of molecular building blocks based on the predefined target properties (Harper and Gani, 2000).

In the past, CAMD approach has been attempted to design solvents for bio-oil applications. To develop a CAMD problem, it is essential to have reliable property prediction models. One of the most widely used methods for the molecular property prediction is the group contribution (GC) methods (Joback and Reid, 1987). However, previous research on the design of bio-oil solvent focused mainly on the solvent functionality that can be predicted using GC prediction models with 1st order groups due to the computational difficulties associated with higher order groups. In addition, the nonavailability of the required GC contributions restricted their applicability in CAMD problem

(Conte et al., 2008). Moreover, the selected GC model may not have all the model parameters required for the estimation of property of a specific chemical (Hukkerikar et al., 2012a). For this reason, topological index (TI) like connectivity index, shape index, or wiener index can be applied as they are a function of the entire molecular graph, which reflect the entire nature of the molecular structure (Austin et al., 2016). Nevertheless, incorporating TI and GC models with higher-order group contributions together is computationally challenging (Chemmangattuvalappil and Eden, 2013). Thus, molecular signature-based algorithms were introduced in this thesis to incorporate higher-order molecular groups from GC models and multiple TIs on a common platform for CAMD.

Molecular signature descriptor is one of the two-dimensional (2-D) fragment-based TI that systematically captures the structural information of a 2-D structural formula. It describes the molecular atoms in terms of extended valencies up to a predefined height (Faulon et al., 2003b). Owing to the fact that molecular signature descriptor is known as the canonical representation of a molecule, all other 2-D classes of descriptors can be represented in terms of molecular signature (Visco and Chen, 2016). Signatures of different height can be used to represent different TI and GC models with higher-order contributions. However, coverage of TIs and higher-order GCs require signatures of higher height due to the requirement of higher structural information. Despite the high accuracy of estimation with the use of signatures of higher height, the complexity of CAMD increases due to the combinatorial nature of higher-order

signatures. Thus, it is crucial to develop an efficient approach to reduce the size of CAMD problem, keeping it at a manageable problem size.

Other than the poor fuel quality and stability of biofuel, its high production cost further hold back on the commercialisation of biofuels (Clemente, 2015). The low heating value of biofuels make them more costly for heat generation (Clemente, 2015). As the biofuel ratio increases in the fuel, the fuel's energy density decreases. Therefore, the addition of solvent is often required to improve the biofuel's properties. These solvents can either be manufactured as commodity or specialty chemicals. However, if the design problem focused solely on optimising the product's performance, it may end up selecting the specialty chemicals as optimal solvents, which are generally associated with high cost. This further increases the production cost of biofuels. In addition, the existing and progressing biofuel legislative framework sparks the urge to assess the cost associated with upgrading bio-oil. To date, most of the research on the design of bio-oil solvent mainly focused on the functionality of the solvent itself. It is also important to incorporate the economic aspects into the development of bio-oil solvent for the designed solvent-oil blend to be competitive with the conventional diesel fuel. The HHV of pyrolysis bio-oil can be increased with the addition of solvent. The higher the mass fraction of solvent in the solvent-oil blend, the higher is the energy content. However, a higher amount of solvent was often associated with higher cost, and thus, lower profitability obtained from the solvent-oil blend. Currently, most of the CAMD techniques focus on optimising a single objective or property of the chemical product (Zimmermann, 1978), but having a multi-objective problem

necessitates the use of more complex optimisation methods. Thus, a multi-objective optimisation (MOO) approach is required to investigate the trade-off between high HHV and high profitability of the solvent-oil blend

It is also to be noticed that the quality and properties of the pyrolysis bio-oil played a major role in determining the final solvent candidates. Meanwhile, the fuel properties of generated pyrolysis bio-oil varied widely depending on the biomass feedstock sources and operating conditions of the pyrolysis process. One of the most significant properties of pyrolysis bio-oil is its HHV, which indicates its quality as a liquid fuel. Many studies had related the HHV of pyrolysis bio-oil to its elemental composition (Vargas-Moreno et al., 2012). However, the studies to estimate HHV of pyrolysis bio-oil from its feedstock properties and operating conditions are not common. In addition, the low pH value of pyrolysis bio-oil often leads to corrosion issues in some materials such as carbon steel, copper, iron, and aluminium which mainly impede its application as engine fuel (Wright et al., 2010). Although characterisations of bio-oil also include the stability, solid content, and water content, these properties are mostly affected by the pyrolysis operating condition such as heating rate, secondary reactions, and separation process. Different from the abovementioned properties, the pyrolysis bio-oil's pH is more directly related to the characterisation of biomass feedstock. Nevertheless, the estimation of bio-oil pH from feedstock and operating conditions have yet to be explored.

In the past, machine learning (ML) tools have been used for predicting the performance of thermochemical biomass processing because of the poor

understanding of the complex underlying reaction mechanisms. The difficulties are further compounded by the highly variable composition and properties of different kinds of biomass feedstocks. As a result, there has been a notable trend in the use of ML tools such as regression, support vector machine (SVM), artificial neural network (ANN) and decision tree to complement mechanistic models for such applications (Kostetskyy and Broadbelt, 2020; Zhu and Yang, 2021). Nevertheless, it is noticed that difficulties were faced in the extraction of practical information from inputs (Tsekos et al., 2021). Popular black-box ML techniques such as ANN and SVM suffer from poor inherent interpretability where the process linking the inputs to outputs is opaque. (Yang et al., 2015). This has led to difficulties in inferring conclusions regarding the underlying mechanisms of a considered process from the network's prediction (Ascher et al., 2022). Other than ML techniques, statistical tools including response surface method (RSM) and multiple linear regression (MLR) approaches will lead to regression models that are also black-box in nature. In the view of this, the rough-set machine learning (RSML) approach can be used to address problems encountered in other ML approaches. With the RSML approach, straightforward and transparent rules can be generated for further information extraction. In addition to being inherently interpretable, ML techniques that use rule-based models allow fusion of expert knowledge with data embedded in information during the training process. Domain knowledge can be introduced via user-defined training parameters to ensure that the final rules are plausible from the perspective of physical mechanisms (Rudin and Ertekin, 2018).

While optimal solvent candidates can be identified from the developed CAMD approach, there are also needs to verify the accuracy of the developed model. On many occasions, the computational solvent design framework ended with the identification of promising chemical compounds. The feasibility and functionality of the identified chemical compounds are usually verified via online database search. However, the results from prediction model often deviates from the experimental results due to the presence of uncertainties in the property relationship. Thus, experimental verification is crucial to ensure the identified solvent candidates are feasible for real-life application while meeting the required targeted properties and performance. In addition, emulsification approach was also required for the application of stable bio-oil with diesel. Pyrolysis oil which was originally immiscible with diesel can be emulsified with the aid of surfactant, resulting in promising stability and ignition characteristic. Nevertheless, previous research on biodiesel blend with surfactant demonstrated varying properties such as low heating value, low cetane number and corrosiveness. Thus, it is crucial to develop an efficient emulsification strategy to generate a stable solvent-oil-diesel blend.

1.3. Thesis Outline

The thesis has been divided into eight chapters. A comprehensive literature review covering the topics of pyrolysis bio-oil, CAMD techniques, MOO approaches, and ML algorithms was presented in Chapter 2. Based on the conducted literature review, the research gaps were determined and described in Chapter 3. This also led to the detailed discussion of research objectives, as well as research methodology in Chapter 3. Following these, four research

scopes were then demonstrated in Chapter 4 to 7, respectively. In Chapter 4, a multi-stage CAMD framework was demonstrated for the design of bio-oil solvent. A novel consistency rule was introduced to reduce the size and complexity of the CAMD problem. Chapter 5 presented an extension of the research scope in Chapter 4, by considering the economic aspect in solvent design. The third research scope is proposed in Chapter 6, where a ML prediction model was developed to predict the fuel properties of pyrolysis bio-oil based on the characterisation of biomass feedstock. Chapter 7 presented the ultrasonic emulsification of pyrolysis bio-oil and diesel fuel with the aid of surfactant and solvents identified in Chapter 4 and 5. The achievements and contributions of this research work were summarised in Chapter 8. The potential future works that can be considered in this area of research work were highlight as well.

CHAPTER 2

Literature Review

The focus of the present work is the combination of computational and experimental approaches in improving pyrolysis bio-oil's fuel properties and combustion quality. This section included the literature review on biomass feedstock, pyrolysis bio-oil, emulsification of pyrolysis bio-oil, computer-aided molecular design (CAMD), machine learning (ML) and rough set theory (RST).

2.1. Biomass Feedstock

Biomass feedstock was referred to as the organic materials, wastes or residues that can be used to produce energy. These biological materials originated from agriculture, including plant and animal wastes as well as aquatic and industrial wastes (OJEU, 2009). Almost all the biomass feedstocks can be converted to energy fuels by either biochemical or thermochemical conversion routes (EIA, 2022b).

2.1.1. Wastes/Low-Valued Co-Products

A large number of agriculture residues are generated annually, especially in agriculture-based countries like China and India (Ross, 2019). However, only 27 % of the agricultural residues generated are further processed as fuel and wood products (DOSM, 2019). On the other hand, biomass wastes generated from wood processing industry came mainly from sawmilling and logging activities such as sawdust, wood chips, damaged or unwanted stem wood and logs.

In addition, the generation of biodiesel from animal fats such as beef tallow, chicken fats and pork lard were reported as well (Banković-Ilić et al., 2014; Toldrá-Reig et al., 2020). These waste animal fats are usually collected as by-products from meat processing facilities and rendering processes. Usage of animal fats as biomass feedstock offered advantages in economic, environmental and food security aspects over the used edible vegetable oils. The high concentration of saturated fatty acid and free fatty acid in animal fats require production techniques with higher complexity, resulting in biodiesel with lower physical and chemical quality. High free fatty acid contents also resulted in soap formation during the base-catalysed transesterification process due to the reaction between free fatty acids and base catalysts. This had led to the loss of catalyst and ester products which increased the production costs (Encinar et al., 2011). On the other hand, low unsaturation of free fatty acids offered several advantages which include high calorific value, high cetane number and high oxidation stability (Adewale et al., 2015).

Waste cooking oil is easily available from restaurants, cafeterias, and household kitchens. It was estimated that approximately 25 million tonnes of waste cooking oil were generated globally, where the United States alone generated 10 million tonnes of waste cooking oil annually (Yaakob et al., 2013). Waste cooking oil collected from kitchen can be categorised as either yellow grease or brown grease. Yellow grease means used cooking oil, spent shortenings or any other vegetable oil, with free fatty acids content of less than 15% (C. E. Goering et al., 1982). It is considered as a “clean” type of grease with little to no contamination and can be recycled. On the other hand, brown

grease included the fat, oil and grease materials collected in grease traps or waste traps, with free fatty acids content of more than 15 % (C. E. Goering et al., 1982). Brown grease is often associated with issues such as food and trash contamination, heavy emulsification, foul odour, and cold flow. Waste cooking oil was reported to be a cheaper alternative of biodiesel feedstock compared to vegetable oils, lowering the biodiesel production cost by 60 % to 90 % (Talebian-Kiakalaieh et al., 2013). Figure 2-1 illustrated the examples of feedstocks categorised under wastes and low-valued co-products. However, in recent years, the drawbacks of using wastes or low-valued co-products for energy have become more apparent. In fact, it was reported that the waste-to-energy approach works against the circular economy (Sara Muznik, 2018). In addition, its negative impacts also include the production of toxic particulate, pollutions, and the possibility to discourage more sustainable waste management solutions (RTS, 2021).

Figure 2-1 Type of wastes/low-valued co-products

2.1.2. Dedicated Energy Crops

Dedicated energy crops are crops grown specifically for energy production which require low cost and maintenance. Energy crops can be generally categorised as grassy, woody, and oil crops as shown in Figure 2-2. Both fast-growing woody and grassy plant species like poplar, willow, miscanthus and switchgrass are cultivated in perennial plantation. Typical rotation periods for woody plants species are between three to seven years and one year for grassy plant species (Rosillo-Calle et al., 2008). Grassy plant species can be classified as high yield and low-energy input crops (Singh and Trivedi, 2017). Switchgrass is a promising biofuel feedstock as it can be cultivated on land with minimal agricultural value while adaptable to a wide range of climates. On the other hand, miscanthus is a perennial, warm-season Asian grass with rapid growth rate and low fertiliser and pesticide inputs (McCalmont et al., 2017).

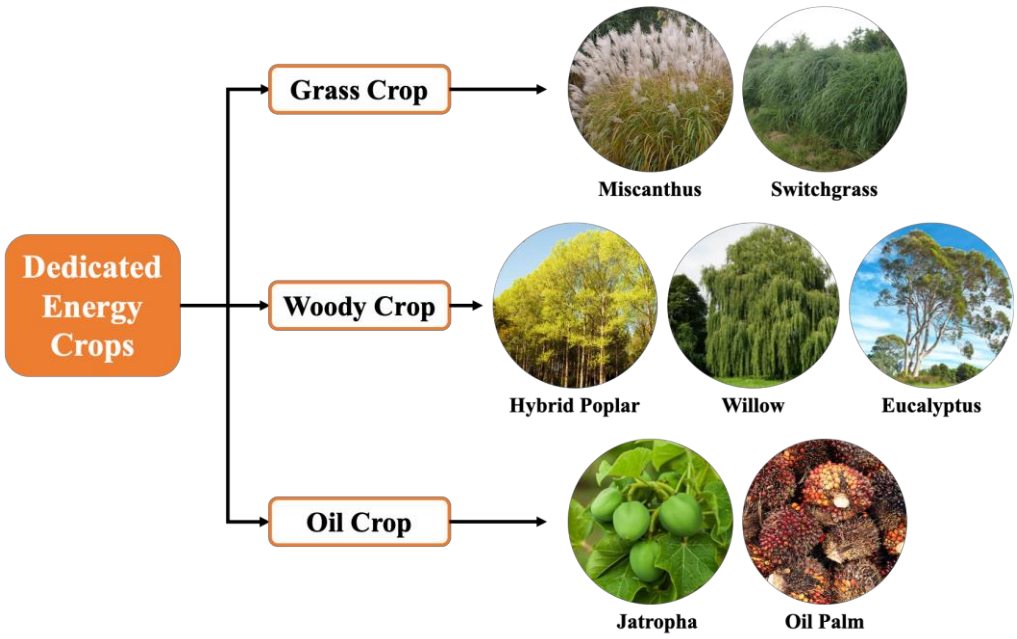


Figure 2-2 Type of dedicated energy crops

Oil energy crops are the base feedstocks for biodiesel production. Grown mainly in Malaysia and Indonesia, oil palm was reported to be the most efficient oil seed crop in the world. Approximately 3.95 tonnes of palm oil and 0.47 tonnes of palm kernel oil can be produced with 1 hectare of land (MPOC, 2020). Based on the total yield of about 4.5 tonnes per hectare, oil palm is 10 times more productive than soybean (MPOB, 2020). Unlike soybean and rapeseed, palm oil is a perennial crop which starts bearing palm fruits for oil about three years after planting. Besides, palm oil has a relatively long productive lifespan of 25 to 30 years (Kurnia et al., 2016). Crude palm oil referred to palm fruit oil extracted from the fruit's flesh. Crude palm oil will be sent to palm oil refineries to be refined, followed by conversion to methyl esters. Methyl esters from palm oil can be directly used as biodiesel or by blending with petroleum diesel. They also exhibit low engine emission and high oxidation stability (Mekhilef et al., 2011).

2.1.3. Algae

Algae are very diverse, ranging from microscopic (microalgae) to large seaweed (macroalgae), as seen in Figure 2-3. They can be found almost everywhere on earth. In recent years, algae have emerged as one of the most promising alternative sources for biodiesel production for their high lipid content, high rate of carbon dioxide absorption, higher growth rate and productivity (Kandaramath Hari et al., 2015).

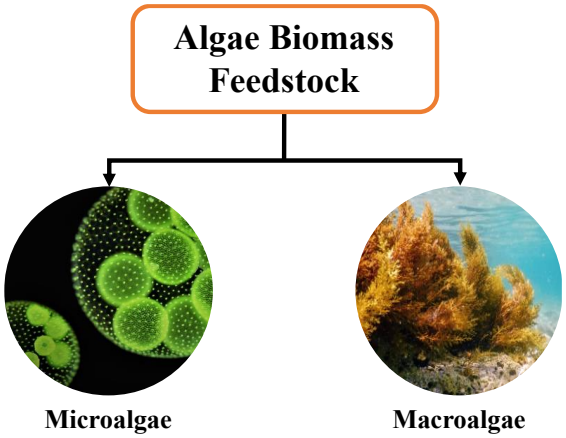


Figure 2-3 Type of algae biomass feedstock

Microalgae are capable to reproduce themselves through photosynthesis, converting sunlight, carbon dioxide and nutrients such as nitrogen and phosphorus into biomass. A complete growth cycle of microalgae takes only a few days, resulting in higher biomass productivity and oil yield compared to other crops (Ahmad et al., 2011). Moreover, they can be grown almost everywhere and can be harvested all year. Algae can be grown using water resources such as seawater, brackish water and wastewater which are unsuitable for human consumption, and thus reduces the food-fuel conflict (Mata et al., 2010). Microalgae with an oil production of at least 70 wt. % of dry biomass only requires 0.1 m^2 year per kg biodiesel of land to produce 121,104 kg of biodiesel per year (Ahmad et al., 2011). In recent years, the focus of research in this area has shifted from direct extraction of lipids to the direct thermal processing and/or fermentation of algae biomass into sustainable alternative fuel (Chuck et al., 2016). Algae biodiesels are reported to be zero-sulphur while having the same performance as petroleum diesel with reduced emissions of particle matters, carbon monoxide, hydrocarbons, and sulphur

oxides. However, emission of nitrogen oxides may be higher in certain engine types (Delucchi, 2003).

Similar to microalgae, seaweed (macroalgae) derived sustainable alternative fuel is also gaining increasing attention as a potential feedstock. Seaweed biomass can be produced either by cultivation and harvesting or by collecting wild drift seaweed, followed by dewatering process(Singh and Trivedi, 2017). In 2017, 30 million fresh weight tonnes of seaweed were produced globally as biomass supply (Buschmann et al., 2017). Chuck et al. (2016) estimated that production from offshore farms could achieve 110 EJ, coastal farms 35 EJ and open sea colonies could even reach 6000 EJ. However, the carbohydrates content of seaweed is low while the ash content is significant, which leads to lower efficiency in conversion through thermal processing and fermentation.

2.2. Pyrolysis Bio-Oil

2.2.1. Pyrolysis Process

Pyrolysis process involves the thermal cracking of biomass feedstock under an inert atmosphere, producing biochar, bio-oil, and syngas. Pyrolysis process can be compared to the charcoal production that was used for centuries (Jenkins et al., 2016). Final product distribution of pyrolysis is dependent on the operating conditions, such as temperature, vapour residence time, heating rate and feedstock size. However, temperature has the most influence on the final pyrolysis product composition. At a lower temperature (< 300 ºC) and heating rate, the conversion of biomass to biochar and gaseous products such as carbon dioxide and water dominates. On the other hand, higher temperature (400 ºC –

500 ºC) and heating rate promote the formation of levoglucosan, a substance that kickstart the production of smaller molecules, thus producing liquid products. At temperature above 650 ºC, the decomposition of the biomass and liquid products dominates, generating gaseous products such as carbon monoxide, hydrogen, and methane (Müller-Langer et al., 2017). Pyrolysis technologies can be categorised into three reaction pathways, which is the slow, fast and flash pyrolysis, depending on the pyrolysis operating conditions. Table 2-1 summarised the operating conditions and approximate product yields for these processes (Nanda et al., 2014).

Table 2-1 Summary of different pyrolysis technologies

Conversion Technology	**Slow Pyrolysis**	**Fast Pyrolysis**	**Flash Pyrolysis**
Operating Conditions:			
Temperature (ºC)	300 – 700	400 – 800	800 – 1000
Vapour Residence Time (min)	10 – 100	0.5 – 5	< 0.5
Heating Rate (ºC/s)	0.1 – 1	10 – 200	>1000
Feedstock Size (mm)	5 – 50	< 3	< 0.2
Product Yield:			
Bio-oil	~ 30 wt. %	~ 50 wt. %	~ 75 wt. %
Biochar	~ 35 wt. %	~ 20 wt. %	~ 12 wt. %
Gases	~ 35 wt. %	~ 30 wt. %	~ 13 wt. %

In fast and flash pyrolysis, the high operating temperature and heating rate enable the pyrolysis process to go beyond the temperature which the first reaction dominates, straight to the second reaction pathway, producing products with mainly liquid fraction. To maximise the bio-oil production, a finely ground feedstock (< 3 mm) is preferred for their low thermal conductivity, which improves the heating rates and heat transfer rates. Besides that, maximum

pyrolysis bio-oil yield can be achieved under the operating temperature of around 500 ºC and short vapour residence time (< 2 seconds) while limiting the secondary reactions. A rapid removal of product char and cooling of pyrolysis vapours results in minimal vapour cracking and higher bio-oil yield (Bridgwater, 2012). Generally, fast pyrolysis method is preferred in industrial bio-oil production due to its scalability and economics of particle grinding (Jenkins et al., 2016).

2.2.1.1. Slow Pyrolysis

In slow pyrolysis, biomass feedstock is pyrolysed at lower heating rates (0.1 – 1 ºC/s). Conversion of biomass to char product dominates at slow pyrolysis, and thus leads to less liquid and gaseous product. A considerable amount of study has been done on this process. Slow pyrolysis of palm kernel cake was carried out in a fixed bed reactor at 700 ºC, heating rate of 20 ºC/min, nitrogen gas flow rate of 200 cm^3/min to obtained bio-oil yield of 54.3 wt. % (Weerachanchai et al., 2011). Ronsse et al. (2013) conducted slow pyrolysis of pine wood, wheat straw, green waste and dried algae and obtained up to 98.4 wt. % yield of bio-char. Hernandez-Mena et al. (2014) performed slow pyrolysis of woody bamboo in a fixed bed reactor at temperature ranging from 300 – 600 ºC, at a 10 ºC/min heating rate. A maximum biochar yield of 80 % was attained at 300 ºC. However, the yield of biochar decreases as the pyrolysis temperature increases, due to the increased in thermal degradation rate. Similar trend was also obtained for slow pyrolysis of coconut fibre and willow (Cai et al., 2020; Dhar et al., 2020). On the other hand, slow pyrolysis of palm empty fruit bunch was conducted to product bio-oil in a bench scale tubular furnace reactor, with

temperature ranging from 400 – 600 ºC at heating rate of 10 ºC/min (Sutrisno and Hidayat, 2018). Maximum bio-oil yield of 44.5 wt. % was obtained at 450 ºC. The most used reactors in this process are fixed bed and tubular reactor (Canabarro et al., 2013).

2.2.1.2. Fast Pyrolysis

In fast pyrolysis, biomass is rapidly heated to a high temperature in the absence of oxygen. However, the heating rate of fast pyrolysis is not as fast as it is in flash pyrolysis. In general, product from fast pyrolysis can be categorised as high-grade bio-oil (Canabarro et al., 2013). Gupta et al. (2019) performed fast pyrolysis of teak sawdust in the temperature range of 400 – 700 ºC. Maximum bio-oil yield of 48.8 % was observed at 600 ºC. In Varma et al. (2019)'s work, maximum yield of wood sawdust pyrolysis bio-oil (44.16 wt.%) was observed at temperature of 500 ºC and heating rate of 50 ºC/min. In another study, date palm tree mixture wastes, and date seed biomass were used to product bio-oil by fast pyrolysis. Maximum yield of 68 wt. % date seed bio-oil was obtained at a temperature of 500 ºC (Bharath et al., 2020). Fast pyrolysis of oil palm empty fruit bunch for bio-oil production was conducted in a bubbling fluidized-bed reactor (J.-W. Park et al., 2019). Improved liquid yield was observed by pre-treating the empty fruit bunch via acid washing with dilute nitric acid. Besides, oil palm frond and palm oil empty fruit bunch were also used to produce pyrolysis bio-oil through fast pyrolysis process (Solikhah et al., 2018). The higher heating value of oil palm frond and empty fruit bunch bio-oils were reported to be 12.19 and 26.49 MJ/kg, respectively. Various reactor configurations such as fluidised bed, entrained flow reactor, wire mesh reactor,

vacuum furnace reactor, vortex reactor, rotating reactor, circulating fluidised bed reactor, etc are suitable to perform fast pyrolysis of biomass for bio-oil (Goyal et al., 2008). Among the abovementioned reactors, fluidised bed is the most promising technology in fast pyrolysis as it allows high heating rate, rapid de-volatilisation, easy control, simple char collection and low cost (Luo et al., 2004).

2.2.1.3. Flash Pyrolysis

Flash pyrolysis is a process in which the reaction time is only a few seconds or even less, which require high heating rate. This involve special reactor configurations such as entrained flow reactor and fluidised bed reactor, where biomass residence times are only a few seconds (Luo et al., 2004). Biomass particle size for flash pyrolysis should be small as rapid heating is required. The effect of temperature and pressure in flash pyrolysis of palm kernel shell as investigated in Matamba et al. (2020)'s work. The study concluded that higher operating temperatures and pressures favoured the generation of polycyclic aromatic hydrocarbons and hydrogen gas. Next, Maliutina et al. (2017) concluded that maximum bio-oil yield of 73.74 wt. % can be obtained with flash pyrolysis of palm kernel shell at 600 ºC. Flash pyrolysis of biomass can achieve up to 75 % of bio-oil yield (Jahirul et al., 2012). However, this process has some technological challenges such as corrosiveness of the bio-oil and solids in the bio-oil. Besides, ash content in the bio-oil has negative impact on the quality and stability of the produced bio-oil. Furthermore, occurrence of catalytic repolymerisation reactions inside the bio-oil due to char fines often lead to a higher viscosity (Canabarro et al., 2013).

2.2.2. Pyrolysis Bio-Oil Upgrading

Bio-oil obtained from the pyrolysis of biomass contains several organic compounds such as hydrocarbons, acids, alcohols, ketones, aldehydes, phenols, and large molecular oligomers, as well as inorganic species. However, these compounds often lead to poor fuel properties and performance of pyrolysis bio-oil. Problems like high oxygen content, high solids content, high viscosity, chemical instability, acidic and corrosive inhibits its application as biofuel. Depending on the operating condition and type of feedstock used, the pyrolysis products are typically distributed among gaseous phase, liquid oil phase, liquid aqueous phase, and solid phase (Müller-Langer et al., 2017). Pyrolysis oil produced will be in stable mixtures of each state, which is impossible to separate them mechanically. Further suspension or thermal drying process is necessary to remove water from the pyrolysis oil. Upgradation of bio-oil can take place in three different pathways:

2.2.2.1. Catalytic Upgrading of Pyrolysis Bio-Oil

In conventional refinery fluid catalytic cracking, petroleum fractions with high molecular weight and high boiling points were converted into lighter products like gasoline or gases. The cracking of these heavy distillates was originally carried out by the thermal cracking process, which requires high temperature with high-pressure hydrogen flow (Zhang et al., 2019). However, this process is then replaced by catalytic cracking as higher production yield of gasoline with a high-octane rating can be achieved. Catalytic vapour cracking, also known as zeolite cracking, has reactions like those of conventional refinery fluid catalytic cracking. Zeolite cracking typically takes place between 300 ºC

and 600 ºC at atmospheric pressure over a zeolite catalyst. HZSM-5 is the common catalyst used in this reaction, which is also applicable in fluid catalytic cracking industry (Chen and Yoshikawa, 2018). An increase in temperature results in a decrease in oil production and an increase in gas production (Mortensen et al., 2011). Higher temperature promotes cracking reactions, resulting in the production of smaller volatile compounds. To remove oxygen elements from bio-oil, a certain amount of cracking at high temperature is required. However, if the rate of cracking is too high, at increased temperature, degradation of the bio-oil to light gases and carbon will occur instead.

2.2.2.2. Chemical Upgrading of Pyrolysis Bio-Oil

Pyrolysis can act as thermal pre-treatment before gasification via synthesis gas (Müller-Langer et al., 2017). A small amount of energy penalty can be observed from the lower pyrolysis energy efficiency, transportation energy and additional bio-oil gasification stand (Bridgwater, 2009). As reported by Dahmen et al. (2012), gasification of solid and liquid suspensions of pyrolysis product is possible to conduct in a large experimental scale. Synthetic hydrocarbons from gasification include diesel, gasoline, kerosene, liquefied petroleum gas and synthetic natural gas, are suitable for conventional fuel application, but much cleaner (Bridgwater, 2012).

2.2.2.3. Physical Upgrading of Pyrolysis Bio-Oil

Addition of polar solvents has been proved to homogenise and reduce the viscosity of bio-oil. Usage of common polar solvents such as ethyl acetate, acetone, methanol and ethanol are extensively studied in the past few decades (Zhang et al., 2019). On the other hand, addition of polar solvents also showed

a significant improvement in oil stability, pH value and high heating value. Blending of bio-oil and solvent under supercritical reactions further enhanced the bio-oil properties by promoting reactions such as esterification and hydrogenation, resulting improvement in physicochemical properties of bio-oil (Omar et al., 2019).

In addition, pyrolysis bio-oils are not miscible with hydrocarbon fuels like diesel. However, the bio-oil can be emulsified with diesel or biodiesel with the aid of surfactants. In work done by (Liu et al., 2021), emulsified oil with fuel properties such as density, viscosity, corrosivity and heating value are close to those of diesel oil can be produced under optimal emulsification conditions, with a mixture of Span 80 and Tween 80 emulsifiers. Emulsification is a comparatively straightforward technique to upgrade bio-oil with diesel or biodiesel, however the downside of this technique includes the high cost of surfactants and high energy requirement for this process.

2.2.3. Emulsification of Pyrolysis Bio-oil and Diesel

Emulsification is a process where two immiscible liquid are blended into a colloidal suspension to form a liquid emulsion. There are many types of emulsion systems, such as the oil-in-water (O/W), water-in-oil (W/O) and multiple emulsions (W/O/W or O/W/O) (Williams, 2001). However, the presence of polar compounds in the pyrolysis bio-oil will lead to stratification of the pyrolysis bio-oil/diesel emulsion (Chong et al., 2017). Thus, a certain amount of surfactants and specific agitation techniques were required in the emulsification in order to generate a stable bio-oil/diesel emulsion. Here, the surfactant can be referred to as a small surface-active materials containing both

lipophilic and hydrophilic groups (Taha et al., 2020). Addition of surfactants with adequate agitation power can form a closely adsorbed monolayer interface on the immiscible liquid droplet to prevent phase separation and coalescence of droplets (Yin et al., 2010). It is reported that the stability of the generated emulsion depends greatly on the surfactant type and the amount of surfactant added to the emulsion (Bertoli et al., 2000). Due to the acidic nature of pyrolysis bio-oil, non-ionic surfactants are favoured for their low sensitivity against electrolytes and compatible nature with cationic and anionic surfactants (Summers et al., 2022). Up to date, non-ionic surfactants such as Tweens, Spans and Brijs has been widely applied in the emulsification of bio-oil and diesel (Liu et al., 2021).

On the other hand, the hydrophilic-lipophilic balance (HLB), depending on the chemical structure and surfactant characteristic, played a major role in choosing a non-ionic surfactant (Lin et al., 2016). The HLB system was initially introduced by Griffin (1954) for the classification of non-ionic surfactants. In the established system, the lipophilic surfactants are featured by low HLB values, while the hydrophilic surfactants are associated with high HLB values. Generally, lipophilic surfactants were ascribed to HLB value ranges of 4 to 8, generating W/O emulsion. Alternatively, O/W emulsions can be obtained with hydrophilic surfactants with HLB value ranging from 9 to 13 (Farooq et al., 2019; Lin et al., 2016).

In recent years, studies involving emulsification of bio-oil and diesel have been reported. Emulsification attributes of four bio-oils from fast pyrolysis of wood wastes and diesel along with the addition of three different commercial

surfactants, namely the Span 80, Tween 80 and Altox 4914 were analysed at different operating conditions (Lin et al., 2016). It was reported that the bio-oil/diesel emulsion with Altox 4914 outperformed the other two surfactants when individually employed. In Farooq et al. (2019)'s work, the resulted ether-extracted bio-oil/diesel emulsion by mixing surfactant with HLB of 7.3 (mixture of Span 80 and Tween 60) showed stable performance with no phase separation up to 40 days. The work was further extended to emulsify the ether-extracted bio-oil/diesel via ultrasonic agitation. Stable emulsion up to 15 days can be obtained at a mixing ratio of 3 wt. % surfactant (mixture of Span 80 and Altox 4916), 15 % pyrolysis oil and 82 % diesel (Farooq et al., 2020). On the other hand, optimal bio-oil/diesel emulsion can be obtained using surfactant with HLB of 7 (by mixing 6 wt. % of Span 80 and Tween 80 as surfactant, 4 wt. % of 2-octanol as co-surfactant), shear velocity of 15,000 rpm, emulsifying time of 5 minutes at emulsifying temperature of 40 ºC. It is expected for the bio-oil/diesel emulsion to be stable up to 384 hours (Liu et al., 2021).

2.2.3.1. Ultrasonic Emulsification of Bio-Oil and Diesel

Emulsification process can be carried out through different agitation approaches such as mechanical agitation (i.e. rotor-stator system), membrane emulsification and ultrasound agitation (i.e. ultrasonic bath). The first application of emulsification via ultrasound agitation was reported in 1927 by Wood and Loomis (1927). From then onwards, extensive research has been conducted on the application ultrasound emulsification, particularly in cosmetic, pharmaceutical and food industries (Mason et al., 1996). In sonication process, ultrasonic sound waves at frequencies greater than 20 kHz were generated (Taha

et al., 2020). The transmission of elastic sound waves will produce alternating positive and negative pressure regions. The generation of these regions were subjected to the rate of frequency at which the ultrasound was applied (BYJU'S, 2022). When the intensity of the ultrasonic sound waves is strong enough, small vacuum bubbles can be generated in the liquid. Then, the formed bubbles will collapse and explode at their saturation level, generating a huge amount of energy (Geetha Bai et al., 2017). This process is also commonly known as the cavitation. The cavitation promotes the disruption of the molecular interaction between the emulsion droplets and thus, minimizing the size of emulsion droplets (Taha et al., 2020).

Nevertheless, the application of ultrasound emulsification to generate pyrolysis bio-oil/diesel emulsion is still limited. An emulsions from crude glycerol and bio-oil was directly prepared by Zhang et al. (2018) via ultrasound and/or mechanical agitation without addition of surfactant. The conducted experiment has proven that the resulted crude glycerol/bio-oil emulsion demonstrated similar fuel properties or even better as compared to the bio-oil alone. The study from Hansen and Mirkouei (2019) demonstrated improved turbidity and stability when ultrasonic emulsification was employed to emulsify bio-oil and diesel with the addition of surfactant and co-surfactant. On the other hand, optimised emulsions of the pyrolysis bio-oil's aqueous fraction and 0# diesel can be obtained under 120 minutes of ultrasonication with the aid of blended surfactants at mass ratio of 72:9:9:10 (Span 80 : Tween 80 : Tween 20 : n-octanol) (Li et al., 2010). Emulsification of hydrothermal liquefaction biocrude from food waste and diesel with the addition of Atlox 4912 surfactant

was performed by Summers et al. (2022) via centrifugation and ultrasonification approach. The results have demonstrated that the ultrasonic approach demonstrated improved energy yield and energy return ratio as compared to the centrifuge method.

2.3. Biodiesel's Standards and Specifications

The quality of biodiesel was significantly affected by various factors such as the composition of biomass feedstock, method of bio-oil extraction, biodiesel synthesis methodology and refining processes. To ensure the biodiesel produced display high performance in engine application, various standards and specifications were established to assess the quality of biodiesel together with proper guidelines for biodiesel fuels testing. Appropriate ranges for various physical and chemical properties of biodiesel fuel were also defined accordingly.

The blend of bio-oil and petroleum diesel was commonly referred to as BX, where X referred to the volume percent of bio-oil in the blend. For example, B5 consisted of 5 % bio-oil, B10 consisted of 10 % bio-oil, and B100 consisted of 100 % bio-oil (Baljet, 2009). To be used in commercial automotive applications, biodiesel must first fulfil the requirements described in the relevant specifications published by various bodies that specialise in fuel certification. These certification bodies include international standards associations, national governments, and corporations. Typically, these specifications were established to control the chemical and physical properties of commercial biodiesel and allow product to be periodically checked for compliance. In 2002, American Society for Testing and Materials (ASTM)

International issued the specification ASTM D 6751 – “Standard Specification for Biodiesel Fuel Blend Stock (B100) for Middle Distillate Fuel”. In October 2013, the European Committee for Standardisation published EN 14214 standard, which described the requirements and testing methods for biodiesel. Table 2-2 listed out the ASTM D6751 and EN 14214 standards of biodiesel together with petroleum diesel fuel standard ASTM D975 for comparison (Sakthivel et al., 2018).

Table 2-2 International standards biodiesel fuels and petroleum diesel fuels.

Property	Units	Petroleum Diesel	Biodiesel	
		ASTM D975	**ASTM D6751**	**EN:14214**
Flash point	ºC	60 – 80	> 130	> 101
Cloud point	ºC	-15 to -5	-3 to -12	–
Pour point	ºC	-35 to -15	-15 to-16	–
Cetane number		46	> 47	> 51
Density @ 15 ºC	kg/m^3	820 – 860	880	860 – 900
Kinematic viscosity at 40 ºC	mm^2/s	2.0 – 4.5	1.9 – 6.0	3.5 – 5.0
Iodine number	g I_2/100 g	–	–	–
Acid number	mg KOH/g	–	< 0.5	< 0.5
Cold filter plugging point	ºC	- 8	Maximum +5	–
Oxidation stability		< 25 mg/L	–	> 3 hrs
Carbon residue	% m/m	< 0.2	< 0.05	< 0.3
Copper corrosion		Maximum Class 1	Maximum No. 3	Maximum Class 1
Distillation temperature	ºC	< 370	360	–
Lubricity (HFRR)	m	< 0.460mm	< 520	–
Sulphated ash content	% mass	–	< 0.002	< 0.02
Ash content	% mass	< 100	–	–
Water and sediment		< 0.05	< 0.005 vol%	500 mg/kg
Monoglycerides	% mass	–	–	< 0.8
Diglycerides	% mass	–	–	< 0.2
Triglycerides	% mass	–	–	< 0.2
Free glycerine	% mass	–	< 0.02	< 0.02
Total glycerine	% mass	–	0.24	0.25
Phosphorus	% mass	–	< 0.001	< 0.001
Sulphur (S 10 grade)	ppm	< 10	–	–
Sulphur (S 15 grade)	ppm	–	< 150	–
Sulphur (S 50 grade)	ppm	< 50	–	–
Sulphur (S 500 grade)	ppm	< 500	< 500	–
Carbon	wt.%	87	77	–
Hydrogen	wt.%	13	12	–
Oxygen	wt.%	–	11	–
BOCLE scuff	G	2000 – 5000	> 7000	–
Total contamination	mg/kg	–	24	24
Boiling point	ºC	–	100 – 615	–
Saponification value	mg KOH/g	–	< 370	–

2.3.1. Flash Point

Flash point of a chemical substance is the lowest temperature at which the combustible vapours of the volatile fuel ignite when exposed to an ignition source. Flash point is an essential criterion for safety purposes. Conventional diesel fuel possesses flash point around 50 – 65 ºC while biodiesel exhibit flash point higher than 93 ºC (National Biodiesel Board, 2020). Compared to conventional diesel fuel, biodiesel fuel poses greater safety aspects when it comes to storage and handling. The minimum flash point for biodiesel is 130 ºC and 101 ºC, according to the ASTM D6751 and EN 14214 standards, respectively. This is to ensure all the alcohol content was removed as the presence of residual alcohol in biodiesel fuel will affect fuel pumps, seals, and elastomers which ultimately leads to poor engine combustion (Alleman and McCormick, 2016).

2.3.2. Heating Value

Heating value or the calorific value of the fuel was defined as the amount of heat energy generated by the combustion of the unit value of fuels. Greater heating value was preferable as it indicated the combustion with higher efficiency. However, the heating value of biodiesel fuel was lower than of conventional diesel fuel due to the higher moisture content in bio-oil. Although heating value was not specified in both ASTM D6751 and EN:14214 standards for biodiesel, it has been prescribed in EN:14214 standard for biodiesel of heating purpose with a minimum value of 35 MJ/kg (Sakthivel et al., 2018).

2.3.3. Density

Fuel density played an important role in engine operation by manipulating the quantity of fuel supplied to the engine for proper combustion. Quantity of fuel injected to the combustion chamber fully relied on the fuel injection system which operated on volume based. As density equals to the ratio of mass per unit volume, increased density resulted in a decreased volume. The density of the biodiesel fuels varied depending on the biomass feedstock used, method of biodiesel conversion and methyl ester profile (Pratas et al., 2011). Biodiesel fuel with a higher percentage of bio-oil resulted in a higher value of density.

2.3.4. Kinematic Viscosity

Viscosity which represented the capability of fuel to flow played a crucial role in injector nozzle design which included the spray atomisation and spray penetration (Sakthivel et al., 2018). Comprising larger chemical structure and molecular mass, biodiesel exhibited viscosity of 10 to 15 times higher than conventional diesel fuel. Highly viscous biodiesel directly impacted the fuel atomisation which led to incomplete combustion and soot deposition in the combustion chamber and thus, lowering the thermal efficiency of the engine. Viscosity of methyl esters in biodiesel fuel was generally reduced by carrying out the transesterification process (Gnanaprakasam et al., 2013).

2.4. Computer-Aided Molecular Design (CAMD)

Traditionally, the design of solvents often involves heuristic rule-based and/or trial-and-error experimental-based approaches, generating safe and reliable product designs (Zhang et al., 2020). However, these traditional solvent

design approaches are usually tedious, time-consuming, and costly. In addition to that, it is not practically feasible to evaluate all the alternatives or to obtain the optimal solution. To overcome these limitations, a more efficient model-based design method, known as the CAMD can be introduced as an alternative to the traditional trial-and-error technique.

The CAMD technique can be defined as a reverse engineering approach to property prediction model-based design. With the CAMD technique, the optimal molecule and/or molecular structure can be identified with a given set of building blocks and specified set of targeted properties (Gani et al., 2003). In other words, the design objective was specified in terms of targeted properties and performance, and property prediction methods were utilised to design the molecule structure with the best performance (Gertig et al., 2020). In the past, CAMD approach has been widely incorporated in the designing of solvent for various applications. A comprehensive review on the solution techniques, applications, and future opportunities of CAMD tools are presented in the review articles of Austin et al. 2016) and Ng et al. (2015). In addition, more detailed discussion on the development of CAMD applications in the design of solvents can be found in the review articles of Zhou et al. (2020) and Chemmangattuvalappil (2020).

Other than the abovementioned application areas, the use of CAMD approach in the design of bio-oil solvents were reported as well. In the past, bio-oil solvents had been applied as additives to the bio-oil to generate a stable solvent-oil blend with promising fuel functionality. Hada et al. (2014) combined property clustering techniques and characterisation-based group contribution

(GC) method in a reverse problem formulation for the design of biodiesel additives. On the other hand, a fuzzy optimisation based CAMD approach was developed in the design of alternative solvents for recovery of palm pressed fibre's residual oil (Khor et al., 2017). Physical properties of the potential solvent along with safety and health attributes were optimised in the study. In addition, Yunus et al. (2018) applied CAMD in the solvent design for palm oil residual extraction from spent bleaching earth. The solvent candidates were screened and evaluated by using a simulation software. Mah et al. (2019) developed a multi-objective optimisation based CAMD framework for bio-oil solvent design. The trade-off between low solvent ratio and high heating value of solvent-oil blend was determined. Due to the rising awareness on environmental issues and stringent environmental regulations, the demand for green solvent has intensified in recent years (Byrne et al., 2016). In the view of this, Neoh et al. (2019) proposed a two-stage multi-objective optimisation problem for the design of bio-oil additives where environmental, health and safety aspect and fuel functionality were optimised simultaneously. However, the environmental, health and safety aspects considered in this work were limited to those properties for which GC property prediction models are available.

Among the different types of property prediction models, some of the prediction models for environmental or non-thermodynamic properties were derived based on semi-empirical quantitative structure-property relationship (QSPR) and quantitative structure-activity relationship (QSAR) models. QSAR/QSPRs are predictive models derived mathematically, correlating

between biological activity or molecular property and one or more physicochemical and/or molecular structural properties, known as descriptors (Dearden, 2017). The ultimate objective in QSAR/QSPR was to convert the chemical structures into molecular descriptors that were relevant to a certain physical property or bioactivity (Dimian et al., 2014). QSAR/QSPR modelling techniques had been widely applied in CAMD problems for their simplicity and efficiency. With the QSAR/QSPR model, the abstract chemical space was associated with a more practical space of quantitative properties, where properties can be estimated on a rather straightforward and efficient approach. In addition to that, QSAR/QSPR in CAMD problems decompose larger molecular structure into smaller sub-molecular collections of atoms and bonds. Each of these sub-structures was assumed to contribute to the property of the final molecule. Owing to that, direct application of combinatorial optimisation in CAMD problem was possible with this type of representation of molecular space (Austin et al., 2016). A simple example of QSAR model was shown by Equation **Error! Reference source not found.** for the acute toxicity of *Pimephales Promelas* by polar and non-polar narcosis (Manuela Pavan et al., 2005):

$$\log LC_{50} = -0.81 \log K_{ow} - 1.744 \qquad (2.1)$$

$$n = 144; r^2 = 0.88; q^2 = 0.87; s = 0.45$$

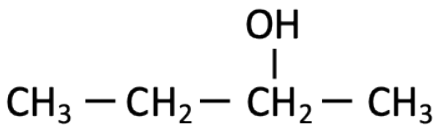
Where LC_{50} is the concentration of narcosis required in feed that was lethal to 50% of the exposed population and $\log K_{ow}$ is the octanol/water partition coefficient. Goodness-of-fit, robustness and predictive ability of the QSAR model can be indicated by the statistical information, such as the total

number of chemical used in QSAR training set (n), the correlation coefficient (r^2), the cross-validated correlation coefficient (q^2) and the predicted standard error(s) (Dearden, 2017).

2.4.1. Group Contribution Method

Group contribution (GC) method is one of the most commonly used property prediction methods for CAMD problem for its simple application and reasonable prediction accuracy. In addition, GC-based model can provide quick estimation without requiring intensive computational resources (Constantinou et al., 1993). In GC method, molecule's properties were estimated by the number of occurrences of each molecular sub-structure known as "groups". For example, 2-butanol can be represented as a combination of the " $–CH_3$ ", " $–CH_2–$ ", and "–OH" groups as shown in Table 2-3.

Table 2-3 GC groups representations for 2-Butanol



Group	No of Occurrences, N_i
$–CH_3$	2
$–CH_2–$	2
–OH	1

Coefficient C_i was assigned to each group to quantify its effect or "contribution" to a specific property X. Properties X can be calculated as Equation 2.1:

$$f(X) = \sum_i N_i C_i \tag{2.1}$$

Where $f(X)$ is the function of the actual properties X, C_i is the contribution of molecular group i that occurred N times. These contributions,

C_i were generated from the regression of property X over a large amount of experimental data for different molecules.

Benson and Buss (1958) developed a simple group additivity rules for the estimation of bond dissociation energies, which was then referred as the group increment theory, or Benson group increment theory (BGIT). In the original paper, BGIT was applied to estimate the entire heat of formation for a molecule by using the experimentally calculated heat of formation for individual groups of atoms. Benson et al. (1969) further expanded the concept, considering a greater diversity of groups for the heat of formation's estimation. Several works have been done by means of Benson's group additivity scheme, estimating the thermodynamic properties of organic compounds in gas, liquid, and solid phases (Domalski and Hearing, 1988; Jalowka and Daubert, 1986; Roganov et al., 2005). However, BGIT does have a few drawbacks such as inaccuracy and group availability issues (Sudlow and Woolf, 1995).

On the other hand, Joback and Reid (1987) developed GC models which estimate heat capacity, liquid viscosity as well as other pure component thermodynamic properties such as the normal boiling and melting point, critical temperature, pressure, and volume. Absence of interaction between groups were assumed, and structurally dependent parameters were thereby calculated by the summation of the number of occurrences of each group multiplied by their group contribution. Constantinou and Gani (1994) then extended the general GC method by introducing two levels of groups, showing significant improvements in accuracy and applicability. This approach was further expanded by Marrero and Gani (2002, 2001) with the introduction of a three-

level group contribution estimation approach. The estimation of basic level, also known as first-order groups, considered the contributions from simple groups in describing a diverse range of organic compounds. However, the first-order group only covered parts of the proximity effects and differences among isomers. The second level was known as the second-order groups, having first-order groups as their building blocks. The second-order groups described the interactions between first-order groups and the effects of certain molecular group combination to the property of a molecule. Although second-order group was capable to provide a more accurate estimation of complex compounds and differentiate among isomers, they were unable to provide a good representation of poly-ring and open-chain polyfunctional compounds with more than four carbon atoms in the main chain (Marrero and Gani, 2001) . This can be further refined by introducing the third-order groups, providing more structural information on fused ring compounds and polycyclic compounds. Analogous to second-order group, third-order group also have first-order group as their building blocks (Marrero and Gani, 2001). The property-estimation model has the form of Equation 2.2:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \qquad (2.2)$$

Where $f(X)$ is the function of the actual properties X, C_i is the contribution of first-order group i that occurred N times, D_j is the contribution of second-order group j that occurred M times and E_k is the contribution of third-order group k that occurred O times.

In GC method, a different combination of functional groups was possible, which resulted in a vast variety of chemical compounds of different structures. This was especially useful from a CAMD perspective (Austin et al., 2016). On the other hand, GC methods described a molecule structure in reference to its functional groups. This can then easily translate into mathematical formulations of CAMD problems. Over the years, the GC method has been widely used to estimate thermodynamic properties of organic compounds. Constantinou et al. (1995) developed GC methods for the estimation of the acentric factor and the liquid molar volume. In addition, GC method was also used to predict the thermodynamic properties such as normal boiling and melting point, critical temperature, pressure and volume, standard enthalpy of formation, vaporisation and fusion, and standard Gibbs energy (Marrero and Gani, 2001). Other than the thermodynamic properties, the GC-based property prediction model also can be used to predict environmental related properties including the acute toxicity, aqueous solubility, bioconcentration factor, global warming potential etc. (Hukkerikar et al., 2012a).

2.4.2. Topological Indexes

Other than the GC method, the topological indexes (TI) is also one of the well-established method in developing property prediction models. TIs are described as the numerical values associated with chemical constitutions for the correlation of chemical structure with physical, chemical or biological activity (Minkin, 1999). TI model considered the connection and interaction between the atoms in a molecule unlike GC models where properties are estimated based on the fixed contributions from the functional groups. In addition to the non-

availability of the required GC contributions, the selected GC model may not have all the model parameters required for the estimation of property of a specific chemical. Moreover, not all properties are not additive. There are properties which are mainly influenced by the interaction between the atoms, types of bonds and degree of branching in the structure. Hence, TI-based property prediction models will be used in this thesis to estimate environmental-related properties including the octanol/water partition coefficient, acute toxicity, global warming potential, etc.

Chemical structures of a molecule contained information of its' molecular geometry which includes the bond lengths, bond angles, torsional angles and other geometrical parameters that determine the position of each atom. In addition, the electronic structure of the molecule can also be determined from its chemical structure, providing insights into the chemical behaviour and the valence shell of the atom. In molecular modelling, relationships between the structure, properties and activity of chemical compounds are studied. Generally, molecules are modelled by a molecular graph. Molecular graph is a simple graph, generally represented as a hydrogen-suppressed graph where molecular skeletons without hydrogen atoms were used, whose vertices represent the atoms and its edges the bonds. For example, the molecular graph for a hydrocarbon molecule consisted of a carbon-atom skeleton. Thus, the vertices of the molecular graph corresponded to the carbon atoms, and its edges corresponded to the carbon-carbon bond.

A molecular graph can be expressed as $G = (V, E)$ for graph G, with vertex set V and edge set E. Vertices were said to be "adjacent" when both

vertices were connected by an edge. The number of vertices of G, adjacent to a given vertex V, is the "degree" of this vertex and will be denoted by d_v. The edge E, however, does not always represent a single covalent bond. In fact, it may be in the form of any type of bond, such as double bond, triple bond, ionic bond, hydrogen bond and more (Basak and Gute, 1997). Figure 2-4 showed the molecular graph of 2,4,4-trimethyl-hexane with its vertex degrees indicated. The 2,4,4-trimethyl-hexane contained five vertices of degree 1, two vertices of degree 2, one vertex of degree 3 and one vertex of degree 4. Vertices of degree 1 were also referred to as the "pendent" vertices, which was described as an end vertex, representing methyl groups. On the other hand, molecular graphs of hydrocarbons cannot possess vertices whose degrees were greater than 4 (Gutman, 2013).

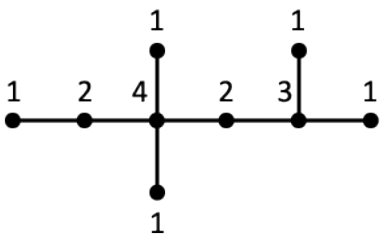


Figure 2-4 Molecular graph of 2,4,4-trimethyl-hexane

Molecular graph can be expressed in different ways, by a drawing, a polynomial, a sequence of numbers, a matrix or by a derived number called TIs, which can take in many forms. One of the main interests of the TI is their ability to segregate very similar structures, such as isomers, providing a more comprehensive picture of the molecule (Austin et al., 2016). Usage of topological descriptors were preferred in QSAR and QSPR in which the

biological activity or other properties of molecules are correlated with their chemical structure studies.

Considering standard molecular graph properties like the degree counts of nodes, connectivity between atoms, branching of molecules, etc, the TIs can be expressed in various forms and expression. The Wiener indexes, or Wiener number, $W(G)$ can be introduced as the "path number", and is the first TI to be used in chemical graph theory (Wiener, 1947). The Wiener indexes can be described as the sum of the distances between any two carbon atoms in the molecule, in terms of the carbon-carbon bond (Wiener, 1947). The Wiener index, $W(G)$ of a molecular graph G is said to be the half sum of the off-diagonal elements of the distance matrix, D_{ij} as shown in Equation 2.3. Wiener index can be applied together with other molecular descriptors, such as adjacency matrix, degree row-matrix and the Laplacian eigenvalues to predict properties like boiling points, heats of formation, heats of vaporisation, molar refractions and molar volume (Nikolić et al., 1995).

$$W = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij} \tag{2.3}$$

On the other hand, the connectivity index (CI) is another form of TI which consider the structural features of a molecule and related to its properties. In the initial development, Randic et al. (1975) proposed a structural descriptor, which is known as the branching index. The branching index of molecular graph G, is defined as the sum of the weights $\left[(v_i)(v_j)\right]^{\alpha}$ of all edges of G, where v is the valency of the considered vertex and α is a constant, -0.5. The larger the branching in a molecular structure, the lower the branching index obtained due

to the inverse relationship. Figure 2-5 illustrates the molecular skeleton of 2,2 dimethylbutane.

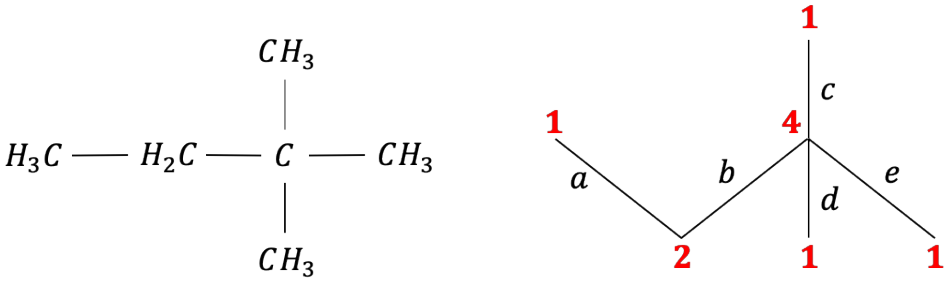


Figure 2-5 Molecular Skeleton of 2,2 Dimethylbutane Molecule

In Figure 2-5, *a, b, c, d,* and *e* indicate different bond, while numbers in red indicates the valency value for each carbon atom. According to the definition above, branching index for 2,2 dimethylbutane can be calculated as :

$$Bond\ a: (1 \times 2)^{-0.5} = 0.707$$

$$Bond\ b: (2 \times 4)^{-0.5} = 0.354$$

$$Bond\ c: (1 \times 4)^{-0.5} = 0.500$$

$$Bond\ d: (1 \times 4)^{-0.5} = 0.500$$

$$Bond\ e: (1 \times 4)^{-0.5} = 0.500$$

$$Branching\ Index_{2,2\ dimethylbutane}$$

$$= 0.707 + 0.354 + 0.500 + 0.500 + 0.500 = 2.561$$

However, the branching index possesses limitations like treatment of cyclic molecules and heteroatoms, and difficulties in differentiating saturated and unsaturated bonds. Thus, further development in connectivity method was proposed, where electronic character of atoms was introduced into the graphic representation of molecules. The new structural descriptor, the valence delta, δ^v was introduced and defined as Equation 2.5:

$$\delta_i^v = Z_i^v - g_i \tag{2.5}$$

where Z_i^v is the number of valence electrons for atom i, and g_i is the number of hydrogen atoms bonded to atom i. However, in Equation 2.5, the number of non-valence core electrons were not considered. Thus Equation 2.6 were redefined by taking the atomic number of an atom, Z into account.

$$\delta^v = \frac{Z^v - h}{Z - Z^v - 1} \tag{2.6}$$

The valence delta, δ^v, encoded the electronic identity of the atom in terms of both valence electron count and core electron count. An individual vertex was said to be in zeroth order when it possesses no edges. Each atom was simply characterised by its valence delta value. The subgraph connectivity term, c was introduced for the zeroth order of valence form for atom m (Equation 2.7):

$${}^0c_m^v = {\delta_m^v}^{-0.5} \tag{2.7}$$

The corresponding molecular CI of zero-order, ${}^0\chi^v$ for a molecule is the sum of ${}^0c^v$ value for all atom (Equation 2.4):

$${}^0\chi^v = \sum {}^0c_m^v \tag{2.4}$$

The zeroth order molecular CI contains a low level of structure information, only with the fact of the presence of the nearest neighbour to each atom was encoded. Subgraph connectivity term for first order, ${}^1c$ can be defined as the product of valence delta values for the pair of atoms involved as shown in Equation 2.5:

$${}^1c_m^v = \left(\delta_i^v \delta_j^v\right)_m^{-0.5} \tag{2.5}$$

The first-order molecular CI, ${}^{v}\chi^{1}$ was defined similarly to ${}^{0}\chi^{v}$ with the summation over all the graph edges in Equation 2.10:

$$ {}^{1}\chi^{v} = \sum {}^{1}c_{m}^{v} \tag{2.60} $$

The first-order molecular CI contained more structure information compared to zeroth order molecular CI. The immediate bonding environment of each atom was encoded by virtue of the edge weight. The ${}^{1}\chi^{v}$ index included both the atom identities as well as the connectedness in the molecular skeleton. For higher order of molecular CI, calculation can be done following the same concept. The n^{th} order of molecular connectivity, ${}^{n}\chi^{v}$ can be calculated by using Equation 2.7 and 2.8:

$$ {}^{n}c_{m}^{v} = \left(\delta_{i}^{v}\delta_{j}^{v} \ldots \delta_{n}^{v}\right)_{m}^{-0.5} \tag{2.7} $$

$$ {}^{n}\chi^{v} = \sum {}^{n}c_{m}^{v} \tag{2.8} $$

As mentioned previously, QSAR/QSPRs were often expressed in terms of more than one TI. Different properties may be expressed in terms of different TI as well. However, different TI exhibit different mathematical expression, which pose challenges in combining and solving it simultaneously on a common platform (Chemmangattuvalappil and Eden, 2013). To overcome this issue, a universal descriptor, namely the molecular signature descriptor was introduced,

where various GC models and TIs can be expressed as a function of the new descriptors on a common platform (Visco et al., 2002).

2.5. Molecular Signature Descriptor

Molecular descriptors are the mathematical representations of molecules' properties that can be generated by an algorithm where these numerical values of molecular descriptors were used to describe the physical and chemical information of the molecules (Chandrasekaran et al., 2018). Molecular descriptors were commonly categorised according to their respective dimensionality which includes zero-dimensional descriptors (0-D), one-dimensional descriptors (1-D), two-dimensional descriptors (2-D), and three-dimensional descriptors (3-D) (Visco and Chen, 2016). 0-D descriptors were established depending solely on the molecular formula of a molecule without any knowledge of structure. 1-D descriptors were a list of substructures in the molecule which contained information that can be calculated from the molecular formula of the molecule. 2-D descriptors had been utilised in most of the molecular descriptor's application. Being more complex than 1-D descriptors, the 2-D descriptors commonly represented molecular topology information such as the size, shape and electronic distribution in the molecule which was handled using graph theory. In 3-D descriptors, the 3-D spatial and features of a molecule were considered.

Molecular signature descriptor is one of the 2-D fragment-based TI that systematically captures the structural information of a 2-D structural formula (Klein, 2002). The molecular signature descriptor was illustrated as a systematic codification system which characterised the molecular atoms in terms of

extended valencies up to a predefined height (Faulon et al., 2003b). Assuming G is a molecular graph and x is an atom of G, the signature of height h of atom x, ${}^{h}\sigma_{G}(x)$ is the canonical representation of subgraph G with all atoms that were at distance h from x. This canonical representation can be established following the procedure as Figure 2-6:

1. Atoms and bonds that were at distance h from atom x were extracted from G and showed up in subgraph ${}^{h}G(x)$.
2. The vertices (atoms) of ${}^{h}G(x)$ were labelled in canonical order, starting from atom x.
3. A tree spanning over all the edges (bonds) of ${}^{h}G(x)$ was constructed one layer at a time up to level h, with the root of the tree being atom x itself. The first layer of the tree was the closest neighbours of atom x, and the second layer consisted of the neighbours of the vertices in the first layer except for atoms x. Assuming the tree has been constructed up to layer $h-n$, layer $h-n+1$ will be constructed considering each vertex y of layer $h-n$. All vertex in the tree will be labelled and coloured with the colouring function accordingly. As vertex can be neighbour of several vertices present in the previous layer, it is normal for vertex appearing more than once in the tree. However, no edge should be repeated in the same tree.
4. The signature was written by reading the tree from atom x, with child-level vertices enclosed in parentheses, level by level, until the required height is reached. The vertex colour must be included in the label even if the vertex appeared several times. All neighbours including the root atom must be written.

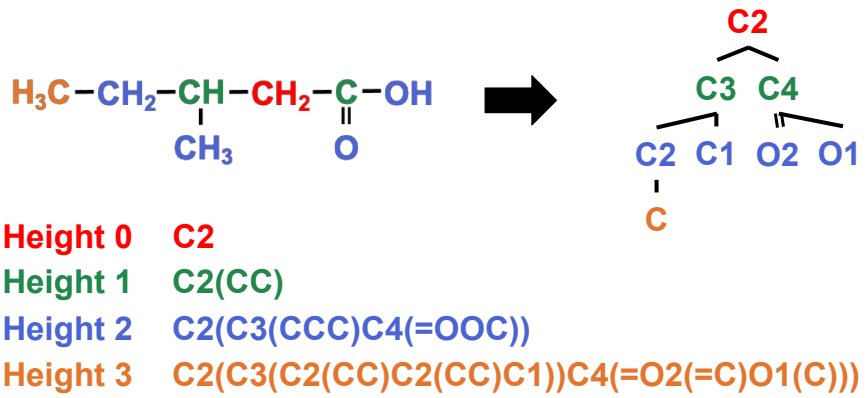


Figure 2-6 Atomic signature of 2-methylbutanoic acid up to height 3

At a given height h, the set of atomic signatures is of finite size. Molecules can be represented by its coordinates in a vectoral space with base vectors as the distinct atomic signatures. Thus, the signature of a molecule can be defined as the linear combination of its atomic signatures as shown in Equation 2.13:

$$ {}^{h}\sigma(G) = \sum_{x \in V_G} {}^{h}\sigma_G(x) = \sum_{i=1}^{{}^{h}K_G} {}^{h}\alpha_i \, {}^{h}\sigma_G({}^{h}X_i) \qquad (2.13) $$

Where ${}^{h}\sigma_G({}^{h}X_i)$ is a base vector, ${}^{h}\alpha_i$ is the number of atoms having the signature of the base vector and ${}^{h}K_G$ is the number of base vectors. Signature descriptors are capable to express various TIs from signature base, representing independent building blocks for a complete molecule while relating the rest of the building blocks (Chemmangattuvalappil and Eden, 2013). The relationship between a TI and its signature can be expressed as a dot product between two vectors, ${}^{h}\alpha_g$, the vector of occurrence number of atomic signatures of height h, and $TI\left(root({}^{h}\sum)\right)$, the vector of predicted values from the model computed for each of the atomic signatures (Equation 2.14).

Meanwhile, the k from Equation 2.14 can be referred as a constant specific to the TI.

$$TI(G) = k^h \alpha_g \cdot TI\left(root\left({}^{h}\sum\right)\right) \tag{2.14}$$

Molecular signature descriptors possess several advantages which make them stand out from other available molecular descriptors. One of its advantages is the complete documentation of atomic topography (Visco and Chen, 2016). Each atom and bond exist in the molecule was identified and tabulated with respective to atomic signature in the molecule. With this complete consideration on the available 2-D structural information, any combination of atomic signature is possible in developing QSARs and QSPRs models (Faulon et al., 2003b; Visco et al., 2002). Owing to the fact that molecular signature descriptor is known as the canonical representation of molecule, all other 2-D classes of descriptors can be represented in terms of molecular signature (Visco and Chen, 2016). This can be beneficial by transforming pre-existing models in other molecular descriptors into models with molecular signature descriptors. An additional key advantage is the degeneracy of molecular signature which is controlled by the height of signatures (Faulon et al., 2005, 2003a). In addition to that, the presence of an efficient algorithm in the molecular signature eases the CAMD process by combining atomic signatures in a structural and comprehensive method (Martin, 2012). This algorithm computes and creates all structures that have a given molecular signature by combining atomic signatures based on their connectivity and valencies. A novel compounds can be created with such an approach as atomic signature fragments will be combined in all possible ways while generating nonintuitive molecules (Visco and Chen, 2016).

Originally formulated as a new TI, the molecular signature was used to create molecules and predict the target properties of these created molecules (Visco and Chen, 2016). In reverse problem application, the non-linearity associated in the original definition of the TIs can be avoided as unknown parameters will be the contribution of individual building blocks instead (Chemmangattuvalappil and Eden, 2013). With the employment of molecular signatures descriptors, mathematical expressions contributing to the non-linearity will be disassembled from other equations, scaling down the complexity of the equation to a certain degree.

In the past, molecular signature descriptors have been applied in various CAMD fields. QSPR based approach with molecular signature descriptor was applied in the design of novel polymers (Brown et al., 2006) and novel glucocorticoid receptor ligands with pulmonary selectivity (Jackson et al., 2008). In Weis and Visco (2010)'s work, ethyl lactate was identified as green industrial solvent by applying the CAMD approach with molecular signature descriptor. Chemmangattuvalappil et al. (2010) redefined the TIs by incorporating molecular signature descriptors in the reverse problem formulation framework. The developed algorithm was then applied in the design of alkyl substituent of fungicide. Ng et al. (2015a) developed a novel two-stage optimisation approach for optimal mixture design in an integrated biorefinery. In addition, molecular signature descriptors were also used to formulate a CAMD model to design fragrance molecules by connecting different prediction models and machine learning algorithm in a common framework (Ooi et al., 2022).

2.6. Economic Consideration in Solvent Design

Aside from the poor fuel quality, the high production cost of pyrolysis bio-oil is another obstacle to the commercialisation of biofuels. Due to their poor heating value and high raw material costs, biofuels are more expensive for heat generation in comparison to the conventional diesel fuel (Clemente, 2015). Additionally, solvents are frequently required to enhance the biofuel's properties. In some cases, the selected solvent might be manufactured as a specialty chemicals which are generally associated with high cost. This further increases the production cost of biofuels. Moreover, the current and evolving biofuel legislation framework further motivates the assessment of cost associated with upgrading bio-oil.

Various recent contributions have included the economic aspects such as product pricing, profitability, market share, and operating cost in the product design. In the design of a traditional Chinese medicinal supplement, customers' preference on product quality and economic considerations was taken into account while developing the chemical product design framework (Cheng et al., 2016). Zhang et al. (2020) provided an overview of chemical product design in the context of a multidisciplinary hierarchical framework including design issues such as project management, market study, economic analysis, product design, and process design. Moreover, an activated carbon production plant from industrial waste nutshells was proposed (León et al., 2020). In the plant design, an economic analysis was developed by considering the cost of the main equipment, the price of the raw materials, basic services, and operations. On the other hand, application of neural network approach is often reported in cost

estimation and investment evaluation. A neural network growth model was proposed to estimate potential of investment in renewables in Japan by (Shahzadah Nayyar Jehan and Mirzosaid Sultonov, 2019). In addition to the existing growth model, the extended growth model considered both environmental risk diversion and risk mitigation. With the developed neural network-based approach, the establishment and operation of a renewable investment opportunity is possible. Further, an artificial neural network method was used to develop a cost estimation model for the tendering of engineering services (Matel et al., 2019).

In this thesis, the pricing model proposed by Bagajewicz (2007) was employed to relate the product quality to demand and price of the product, which is the objectives to be optimised in this Chapter 6 (Fung et al., 2016). In the past, the pricing model has been incorporated in various product design such as wine (Whitnack et al., 2009), carpet deodorizers/ disinfectants (Street et al., 2008), skin moisturising lotion (Bagajewicz et al., 2011), die-attach adhesive (Fung et al., 2016), and dry-cleaning solvent (Lai et al., 2019). The mathematical expression for the pricing model has been shown in Equations 2.15 and 2.16.

$$A^P T^P = A^C (T^P)^\delta \left(\frac{\alpha}{\beta}\right)^\delta \left(\frac{Y - A^P T^P}{A^C}\right)^{1-\delta} \tag{2.15}$$

$$Y \geq A^P T^P + A^c T^c \tag{2.16}$$

Here, A^P and T^P referred to the price and demand of the new product while A^c and T^c referred to the price and demand of the competitor's product. Based on Equation 2.15, Y is the total market size for the new product and δ is

the elasticity of substitution, which is an adjustable parameter that measures the change in the ratio of products demand in response to a change in the ratio of their prices. On the other hand, the parameter α can be expressed as the consumer's awareness on the new product, which can be raised by allocating higher budget in the marketing of new product. The value of parameter α ranges between 0 and 1, where α with the value 0 indicates that the consumers have no knowledge about the new product, and vice versa. Lastly, β is the consumer preference coefficient that relates the consumer's interest in the new product over the competing product, which can be determined using Equation 2.17.

$$\beta = \frac{\lambda^C}{\lambda^P} \tag{2.17}$$

In Equation 2.17, the parameters λ^C and λ^P represented the consumer's preference function of competitor's product and new product, respectively. The new product is said to be preferred by consumers if β is smaller than 1. However, the competitor's product is preferred when the value of β is greater than 1.

Notably, past studies on the design of bio-oil solvents focused primarily on the solvents' functionality. Nevertheless, it is also critical to integrate the economics considerations into the development of bio-oil solvent in order for the solvent-oil blend to compete with conventional diesel fuel. Generally, the heating value of pyrolysis bio-oil increases with the addition of solvent. However, as the amount of solvent increases, the cost of solvent-oil blend increases as well, resulting in lower profitability. Thus, a multi-objective optimisation problem is required to optimise both the fuel property and profitability of the solvent-oil blend, simultaneously.

2.7. Multi-Objective Optimisation Problem

Currently, most of the developed chemical product design frameworks only focused on optimising a single target property. This type of design problem is commonly known as the single-objective optimisation problem. However, in most of the design problems, multiple product properties, such as the thermodynamic properties, safety, health and environmental criteria, and economic aspects need to be considered and optimised simultaneously to design an optimal product. When there is more than one objective function to be optimise, the CAMD design problem have to be solved as a multi-objective optimisation (MOO) problem. For a non-trivial MOO problem, single solution that can simultaneously optimise all the objectives do not exist, but rather is a nondominated set. The nondominated set can also be known as the Pareto set, which was referred to a set of solutions that are non-dominated to each other but are superior to the rest of solutions in the search space (Akbari et al., 2014).

Generally, the methods to solve MOO problem can be categorised into two main categories, which is the preference-based methods and generating methods (Akbari et al., 2014). The goal programming method is one of the oldest and most widely applied methods in the preference-based category (Wang et al., 2017). In this method, the decision-makers are required to set a goal for each objective. The deviations from the set goals will then be minimise using the optimisation approach. Nevertheless, the solution obtained may not be in the Pareto set if the unreasonable goals were set. On the other hand, the ε-constraint method and weighted sum method were reported to be the two common generating approach to MOO problem (Arora, 2012). In the ε-

constraint method, one of the objective function was selected to be optimised, while the other objectives were converted into additional constraints. A number of non-inferior solutions on a non-convex boundary can be identify using the ε-constraint method. However, the selected of ε value does not always ensure a feasible solution. Hence, a large number of iterations were often required to obtain the optimal ε value (Pirouz and Khorram, 2016).

The mathematical expression for the weighted sum method can be found in Equation 2.18:

$$U_{weighted\ sum} = \sum_{i=1}^{k} w_i F_i(\mathbf{x}) \tag{2.18}$$

Here, the w can be referred to as the weighting factor assigned to each of the individual objective function $F_i(\mathbf{x})$, such that $\sum_{i=1}^{k} w_i = 1$ and $w > 0$ (Marler and Arora, 2010). This has allowed multiple objectives to be converted into an aggregated scalar objective function. An optimised overall objective function, $U_{weighted\ sum}$ can be obtained by summing up all the contributions from each individual objective. In the past, the weighted sum method has been vastly applied in various MOO problem for different applications. A weighted sum multi-objective model was proposed by Majidi et al. (2017) to minimise the operation cost and carbon dioxide emission operation of a PV-battery-fuel cell hybrid system in the presence of demand response program. Recently, Madathil et al. (2021) used the weighted sum method for electrical energy management in a residential building by maximising the electrical energy comfort and minimising the drain on electrical energy.

The weighted sum approach is simple and easy to use. Furthermore, the minimum of Equation 2.18 is always Pareto optimal if all the weights are positive (Arora, 2012). Nevertheless, there are several major drawbacks with the weighted sum method. Firstly, the Pareto optimal points could not be located if the MOO problem is non-convex. In addition, an evenly distributed set of weighting coefficient does not guarantee an evenly distributed representation of the Pareto set, even if the problem is convex (Das and Dennis, 1998). Secondly, the decision maker is required to assign a satisfactory weighting factors to each objective function. Hence, there will be a bias in finding the trade-off solution as expert knowledge or personal subjective preferences of the decision maker were involved. As a result, a small change in the weighting coefficient may lead to a big changes in the objective vectors (Diwekar, 2008). However, the weighting factor of each objective function is not always definable in the chemical product design. Furthermore, the relative importance of each objective function is either fuzzy or uncertain, or even might be incomplete, unclear, or contradictory to each other in nature.

2.7.1. Fuzzy Optimisation

Fuzzy optimisation algorithm is another popular alternative in solving the MOO problem. The fuzzy set theory was first proposed by Zadeh (1965) as an improvement over the classical set theory to solve decision-making problem under fuzzy environment. The fuzzy set can be defined as a class of elements with a continuum of grades of membership, ranging from zero to one. Bellman and Zadeh (1970) then introduce fuzzy optimisation approach where preferred alternative can be identified by solving the fuzzy goal in respect to a set of fuzzy

constraints. Here, the fuzzy goals and fuzzy constraints can be described as fuzzy sets in the space of alternatives. Hence, the fuzzy decision can be portrayed as the intersection of the defined goals and constraints. The fuzzy set theory was then employed by Zimmermann (1976) in a fuzzy linear programming problems. Later, the fuzzy set theory was extended to address linear mathematical programming problems involving multiple objectives (Zimmermann, 1978). The fuzzy optimisation approach is capable to solve decision-making problems under the fuzzy environment by defining and quantifying the uncertainties and vagueness. Within the fuzzy optimisation algorithms, the trade-off between the objective functions to be optimised can be identified by introducing the fuzzy membership function. As a result, an optimal compromised solution can be identified by achieving near optimality for all the objectives.

In recent years, the fuzzy optimisation approaches have been extensively employed in various industries and research disciplines. Among them, the fuzzy max-min aggregation approach introduced by Zimmermann and Zysno (1983) has received great attention. A systematic fuzzy optimisation-based method was developed to design molecules for chemical processes with both property superiority and robustness optimized (Ng et al., 2015b). Khor et al. (2017) adapted the fuzzy optimisation via max-min aggregation in the CAMD of alternative solvents for oil extraction from palm pressed fibre. The developed approach was able to optimise the physical properties of the solvent simultaneously with the safety and health aspects. A similar approach was employed and extended with the introduction of disjunctive programming in the

work by Ten et al. (2017) to design solvent for gas sweetening process. Recently, a max-min aggregation fuzzy optimisation approach was employed along with CAMD method in the design of green solvents for pyrolysis bio-oil upgrading with consideration of environmental, health and safety aspect while ensuring minimal compromise on the fuel functionality (Neoh et al., 2019). Gurav and Regulwar (2020) has utilized fuzzy MOO approach in determining optimal cropping pattern for sustainable irrigation planning by simultaneously maximizing four objectives, including net benefits, crop production, employment generation and manure utilization. In addition, the fuzzy multi-objective model was also adapted in Ghanbarzadeh-Shams et al. (2022)'s work to study the trade-off between profitability, energy consumption and carbon emission in the carpet production industry.

2.8. Machine Learning

Machine learning (ML) is a branch of artificial intelligence (AI) and computer science which provided the machine with capability to self-learn through experience, without being explicitly programmed. The main objective of ML is to train the machine to learn and predict based on the provided labelled or unlabelled data sets to generate results from the specified problem (Dhall et al., 2020). Once the machine's algorithm learned and identified the underlying mechanisms or patterns of the given data, it is capable to carry out its task automatically (Dey, 2016). ML algorithms can be divided into 4 primary categories, depending on the desired outcome of the algorithm.

2.8.1. Supervised Learning

Today, supervised learning method is one of the most popular sub-branches in the ML field. A dataset consisting of training data and validation data is required for supervised learning process (Kotsiantis et al., 2006). In the training data set, input attributes are matched with the desired outputs. During the training process, this training set is used as an example inputs to analyse data and/or predict outputs accurately (Simon et al., 2015). The supervised learning algorithms will identify the underlying patterns in the data that associate with the respective outputs. After the training, the developed model is capable to label and classify the new unseen inputs based on the previous training data. The ML model then achieved the desired outcome by adjusting the error based on the comparison between the computed output and expected output (Das et al., 2015). At the data mining stage, supervised learning can be separated into two main types of problems, classification, and regression.

In classification problem, algorithm was applied to assign training data into pre-defined categories accurately. Common classification algorithms are linear classifiers, support vector machines, naïve bayes, decision trees, k-nearest neighbour, and random forest. In the regression approach, the ML model attempts to identify and understand the relationship between the dependent and independent attributes. Some popular examples from the regression algorithm include the linear regression, logistical regression, and polynomial regression (Wilson, 2019). Supervised learning models were often used in organisation or business applications to solve real-world problems at scale such spam detection, predictive analytic, recognition of handwriting, face and speech, information

retrieval, natural language processing and computer vision (Das et al., 2015; Dhall et al., 2020; IBM Cloud Education, 2020a).

2.8.2. Unsupervised Learning

The unsupervised learning algorithm self-learn through discovering and adopting, based on the pattern of input data (Das et al., 2015). Unlike supervised learning, the unsupervised learning algorithms analyse, discover hidden patterns and cluster unlabelled dataset. The developed algorithms were capable to predict the outcome of the new data introduced based on the past experiences and previously learned features (Dhall et al., 2020). Three common methods used in unsupervised learning are the clustering, association, and dimensionality reduction. Clustering algorithms can be further categorised into a few types, specifically k-means clustering, exclusive clustering, overlapping clustering, hierarchical clustering, and probabilistic clustering. Clustering algorithm is a data mining technique used to process raw and unclassified dataset, by grouping them based on their similarities or patterns (IBM Cloud Education, 2020b). On the other hand, association rule is a rule-based method which discovers the underlying associations between attributes of a large database. Examples of association rules application can be seen in market basket analysis, to develop better cross-selling strategies and recommendation engines (IBM Cloud Education, 2020b). Few algorithms were commonly used for association rule generation, such as Apriori, Eclat and FP-Growth. Although larger dataset can provide more accurate outcome, high number of dimensions or features in a dataset could impact the performance and complexity of the learning process. Thus, the dimensionality reduction was commonly applied in the pre-processing

stage, where large dataset can be reduced into a manageable size with while preserving the integrity of the dataset. Some of the commonly used algorithms in dimensionality reduction methods are principal component analysis, singular value decomposition and autoencoders (IBM Cloud Education, 2020b). Unsupervised learning techniques have become a common approach in organising large dataset for real-world applications such as recommendation engines, DNA classification, market segmentation, social network analysis, computational biology and etc. (Das et al., 2015; Dhall et al., 2020) .

2.8.3. Semi-supervised Learning

Semi-supervised learning algorithm is a combination of both supervised and unsupervised learning algorithm. A smaller size of labelled data set was used to develop and train the algorithm, which was then used to classify and extract information from a larger sized of unlabelled data (IBM Cloud Education, 2020c). This had benefited the learning process by avoiding the challenges of finding a large amount of labelled data, which is a very costly process. In order to fully utilise the unlabelled data, certain assumptions need to be proposed, which formalize the types of expected interaction (Chapelle et al., 2006). The most widely recognized assumptions are the smoothness assumption, the low-density assumption, the cluster assumption, and the manifold assumption (Olivier et al., 2006; van Engelen and Hoos, 2020). Various semi-supervised learning methods were derived based on the abovementioned assumptions such as generative models, low density separation, graph-based methods, heuristic approaches and transductive support-vector machine (Dey, 2016; Olivier et al., 2006). A common example of the practical applications of

semi-supervised learning is the internet content classification. As it is tedious and impractical to manually classify all of the webpage, thus semi-supervised learning algorithm was used to rank the relevance of a webpage for a given query (Gupta, 2019). Other example of semi-supervised learning in real-world application includes the speech analysis, protein sequence classification

2.8.4. Reinforcement Learning

Reinforcement learning algorithm, also termed as behavioural ML model, was trained on a reward and punishment mechanisms, based on trial-and-error search and delayed feedback within an environment (Sutton, 1992). Basic reinforcement learning machine is often modelled as a Markov decision process. Value-based, policy-based, and model-based reinforcement learning method are three of the most used learning approaches in the implementation of reinforcement learning algorithm. Dissimilar to supervised learning algorithm, reinforcement learning algorithm was trained based on the interactions with the environment instead of given sample data. This algorithms take actions based on the more favourable outcomes (Dey, 2016). Correct outcomes will be rewarded while incorrect outcomes will be penalised. A series of positive outcomes will be reinforced to develop the best recommendation or policy for a given problem (IBM Cloud Education, 2020c). The Q-learning, state-action-reward-state-action, deep Q network and deep decision policy gradient are few of the common learning models used in the reinforcement learning. The reinforcement learning algorithm is also applied in other disciplines, such as game theory, control theory, genetic algorithms, multi-agent systems, information theory, simulation-based optimisation etc. Some of the real-world

application of reinforcement learning includes autonomous driving, industry automation, business strategy planning, natural language processing, computer games, traffic light controls, personalised recommendation, and web system configuration (Arel et al., 2010; Bu et al., 2009; Mwiti, 2021; Silver and Hassabis, 2017; Zheng et al., 2018).

2.8.5. Rough Set Theory

The idea of rough set theory (RST) was first proposed by Pawlak (1982) as an extension of set theory, which is a new mathematical approach to overcome vagueness, imprecision, consistencies, and uncertainties in information and knowledge (Zhang et al., 2016). The establishment of RST was based on the assumption that each object of the universe of discourse some information is associated (Pawlak, 2002). Since RST allows for sets with ill-defined boundaries, it is inherently robust and can be used for handling data of poor or inconsistent quality. It is the basis for rough-set machine learning (RSML) which has the advantage of not requiring any preliminary or additional information (such as probability in statistic, or grade of membership in the fuzzy set theory) other than a data set. Furthermore, RST provided an efficient approach and algorithms for identifying the underlying patterns in the data. Other than the abovementioned advantages, RST has many other attractions such as data reduction (discover minimal sets of data), assess the significance of data, automatically generate sets of decision rules from data, apparent interpretation of obtained results, and suited for concurrent (parallel/distributed) processing (Suraj, 2004). In the past, RST has been widely adopted in the field of AI and cognitive sciences, specifically for the applications in ML, knowledge

discovery, data mining, expert systems, approximate reasoning, and pattern recognition.

RSML yields models consisting of "IF-THEN" rules that are particularly useful where direct and transparent interpretations are needed. In addition, meaningful insights and mechanistic plausibility (i.e. consistency with known scientific principles) can be evaluated for the decision rules generated. RSML is especially useful for dealing with sparse datasets that are encountered in many engineering applications. In RSML, the user can thus select the best model based on both statistical performance and subjective judgement or background knowledge. In addition, RSML can differentiate between objects that can be certainly categorised into a certain class and those that may possibly be categorised. These algorithms can be further applied in knowledge reduction, concept approximation, decision rule induction, and object classification (Mahajan et al., 2012).

2.8.5.1. Information Table and Decision Table

In rough set model, data set was represented as a table, which is known as the information table. Each row of the table represents an object, a case, or an event. Meanwhile, each column of the table represents attributes corresponding to each object. This table can be represented as a pair of information system, $S = (U, A)$, where $U = \{x_1, x_2, \dots, x_n\}$ is a non-empty finite set of objects (also called universe) and $A = \{a_1, a_2, \dots, a_m\}$ is a non-empty finite set of attributes, such that $a: U \rightarrow V_a$ and $a \in A$ (Kalaivani et al., 2017). The set V_a is also known as the domain set, which contained the values for a. The set of attributes can be further partitioned into

condition attributes, which describe the object; and the decision attributes which indicate the classes of the object. An information table is often known as a decision table when it consist of the decision attribute/attributes. The decision table can be represented as decision system of the form $S = (U, A \cup \{d\})$ where $d \notin A$ is the decision attribute. Table 2-4 represents a decision table with 2 conditional attributes and a decision attribute.

Table 2-4 Example of a decision table

Object	Conditional Attributes		Decision Attributes
	A1	A2	Dec
P1	1	3	1
P2	1	0	0
P3	2	1	0
P4	2	1	1
P5	3	2	0
P6	1	2	1
P7	3	2	0

2.8.5.2. Reducts and Cores

It is possible for a large dataset to contain two or more objects which perform similarly in the attributes or features, commonly known as the indiscernible objects. In RST, reduct can be referred to as a subset of indispensable attributes which can partition the database with the same level of discrimination as the original set of attributes (Pawlak, 1982). On the other hand, the intersection of all reducts are known as the core. It may also be regarded as the essential attributes set which cannot be excluded from the decision system without losing the equivalence class structure. In other word, if core attributes were removed from the information table, it will result in data inconsistency. To reduce the redundancy of data while retaining its basic features, only one

representative object was stored in the dataset for every set of indiscernible objects (Dutta, 2019)

For the decision system, $S = (U, A)$, with set of attributes $P \subseteq A$, a binary relation denoted by $IND_P(U)$ can be expressed as Equation 2.19:

$$IND_P(U) = \{(x, y) \in U^2 | \forall a \in P, a(x) = a(y)\} \quad (2.19)$$

Equation 2.19 can also be known as P-indiscernibility relation. If $(x, y) \in IND_P$, then objects x and y are indistinguishable from each other over the set of attributes from P (Mahajan et al., 2012). Taking Table 2-4 as example, the indiscernibility relation of attribute $A1$ can be shown as Equation 2.20:

$$IND\ (A1) = \{P1, P2, P6\}, \{P3, P4\}, \{P5, P6\} \quad (2.20)$$

2.8.5.3. Decision Rules and Algorithm

Let the decision table be $S = (U, C, D)$, where C and D is the sets of condition and decision attributes, respectively. Every $x \in U$ determines a sequence $c_1(x), c_2(x), \dots, c_n(x)$ and $d_1(x), d_2(x), \dots, d_m(x)$, where $C = \{c_1, c_2, \dots, c_n\}$ and $D = \{d_1, d_2, \dots, d_m\}$. This sequence is known as a decision rule induced by x in system S. The decision rule can be expressed as $c_1(x), c_2(x), \dots, c_n(x) \rightarrow d_1(x), d_2(x), \dots, d_m(x)$, or in short $C \rightarrow_x D$ (Pawlak, 2002). The strength of the decision rule $C \rightarrow_x D$, can be obtained by Equation 2.21:

$$\sigma_x(C, D) = \frac{supp_x(C, D)}{|U|} \quad (2.21)$$

where $supp_x(C, D) = |A(x)| = |C(x) \cap D(x)|$ is the number of elements that support the decision rule $C \rightarrow_x D$ and $|U|$ is the cardinality of U.

With every decision rule $C \rightarrow_x D$, the conditional probability of an element characterized as $C(x)$ being classified into the decision class $D(x)$ can be known as the certainty factor. The certainty factor can be denoted as $cer_x(C, D)$, and defined as Equation 2.22 and 2.23.

$$cer_x(C, D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{supp_x(C, D)}{|C(x)|} = \frac{\sigma_x(C, D)}{\sigma_x(C)} \quad (2.22)$$

$$cer_x(C, D) = \pi_x(D|C) \quad (2.23)$$

where as $\sigma_x(C) = \frac{|C(x)|}{|U|}$. If a decision rule $C \rightarrow_x D$ have a certainty factor of $cer_x(C, D) = 1$, then the rule will be known as a certain decision rule. However, rule with certainty factor $0 < cer_x(C, D) < 1$ will be referred to as an uncertain decision rule. On the other hand, coverage factor of a decision rule $C \rightarrow_x D$ is defined as the percentage of an element belongs to $D(x)$ being classified under the given rule, and can be denoted as $cov_x(C, D)$ as shown in Equation 2.24 and 2.25:

$$cov_x(C, D) = \frac{|C(x) \cap D(x)|}{|D(x)|} = \frac{supp_x(C, D)}{|D(x)|} = \frac{\sigma_x(C, D)}{\sigma_x(D)} \quad (2.24)$$

$$cov_x(C, D) = \pi_x(C|D) \quad (2.25)$$

where $\sigma_x(D) = \frac{|D(x)|}{|U|}$.

Decision rules in the form of "IF-THEN" are always generated from the decision table by applying the rough set-based algorithms. In other word, the "IF-THEN" decision rules implied that if some of the condition attributes have a given values, then some decision attributes will be having other given values (Anuradha et al., 2011; Qian et al., 2008). The rough set-based decision rule algorithm can be further categorized into three types, algorithm which generates

the minimum set of rules covering all objects from a decision table; algorithm which generates the exhaustive set of rules consisting of all possible rules for a decision table; and algorithms which generates the set of decision rules, even partly discriminant, covering relatively many object but not necessarily all objects from the decision table (Charturvedi et al., 2017). Some of the examples of different algorithms used in generating the decision rules are shown in Table 2-5 (Charturvedi et al., 2017).

Table 2-5 Mechanisms for decision rules generation in RST approach

Algorithm	Approach
LEM2	Decision rules are generated based on the lower and upper approximations. This algorithm is suitable to be used in both increment and non-increment modes.
RLEM2	Decision rules are generated based on the lower and upper approximations with utilization of multi-set decision tables. In this algorithm's approach, the extension of basic SQL operators is used for rules generation.
BLEM2	Decision rules are generated based on the lower and upper approximations with the partition of a boundary set. Generated rules are dependent on Bayes' theorem and rough sets.

2.8.6. Application of Rough Set Theory

RST offers practical approaches in resolving various complex problems in various field including, the medicine, engineering, marketing analysis, bioinformation, banking, financial and others, and has gained achievements in respective field. At present, research on RST mainly focused on three aspects, specifically in the theoretical, application and algorithm research. This section provides a brief overview on the application of the rough set approach in real-world problem.

The rough set-based approach was applied in pattern recognition of partial discharge in noise affected cable data. The proposed rough-set approach was proven to have higher accuracy compared to SVM and back-propagation neural network methods and can be applied for on-line partial discharge monitoring of cable systems after training with valid sample data (Peng et al., 2017). A hybrid of RST and Bayesian reasoning approach was proposed for information retrieval in the ranking of webpages. User queries and web pages are presented in the form of rough sets, while the approximation regions for query and documents are calculated using the Bayesian Rough Set Model. The initial outcome of the presented model shows better performance than some of the existing model (Sharma and Kumar, 2020). Recently, Radhakrishnapany et al. (2020) incorporated RSML approach in CAMD for the design of fragrant molecules. Decision rules were generated based on the relationship between topology of fragrant molecules and their odour characteristic and were applied as constraint in the design problem. Ooi et al. (2022) further extended to work by implementing the hyperbox classifiers to predict the fragrance properties. A quantitative relationship between the structure parameters of molecules to their odour characteristic can be established with the disjunctive decision support rules generated from the developed model.

2.9. Conclusion

Based on the literature review conducted in this chapter, it is clear that the current bio-oil's solvent design mainly focused on the fuel properties that for which GC property prediction models are available. Nevertheless, the incorporation of environmental and economic aspects into the CAMD

framework is crucial in order for the designed solvent-oil blend to compete with conventional diesel fuel. In this chapter, various raw biomass feedstock and different pyrolysis technique were also presented. The final product distribution of pyrolysis bio-oil is highly dependent on the feedstock and pyrolysis operating conditions. Nevertheless, the studies to estimate fuel properties of pyrolysis bio-oil from its feedstock characterisation and pyrolysis operating conditions are not common. Therefore, there is a need to fill the research gap by predicting the optimal feedstock composition and pyrolysis operating condition that can generate pyrolysis bio-oil of desired fuel properties. This can be achieved by employing the RSML approach as a pre-processing and predictive modelling tool. Other than solvent addition, emulsification is also an easy, inexpensive, and effective physical method for pyrolysis bio-oil upgrading. In the past, numerous research focused on the emulsification of pyrolysis bio-oil and diesel via mechanical agitation with the addition of surfactant. However, the study emulsification of pyrolysis bio-oil/diesel via ultrasonification agitation is still scarce. In such cases, there exist an opportunity to explore the potential of generating pyrolysis bio-oil/diesel emulsion at different mixing ratio using sonicator probe with the aid of surfactant. The identified research gaps from this chapter will be further discussed and addressed in Chapter 3.

CHAPTER 3

RESEARCH SCOPES AND METHODOLOGIES

3.1. Research Gaps

Over the decade, the consumption of fossil fuels for energy has taken a toll on the human's health and environment, from air and water pollution to global warming. In addition, the volatilities in crude petroleum oil prices further intensified the need to find a sustainable renewable fuels. Biomass has been identified as one of the potential renewable energy source by converting into biofuels via various conversion techniques. Among the available conversion techniques, fast pyrolysis received great attention for being a relatively simple and inexpensive technology. However, the industrial applications of pyrolysis bio-oil were limited due to its poor fuel properties such as corrosiveness, high viscosity, low heating value, low miscibility, high water content as well as storage instability due to product aging. Attempts had been made in various research to improve the fuel properties of pyrolysis bio-oil via direct solvent addition at different mixing ratio. In most of the studies conducted, alcohol solvents such as methanol, ethanol, 1-butanol and etc were commonly applied as solvent to improve the fuel properties of pyrolysis bio-oil. However, various shortcomings such as short-term stability (24-96 hours), high viscosity, poor miscibility, high water content, high acidity and formation of residual reactive compounds still remained. Thus, designing a solvent which is capable to fulfil

multiple property targets is essential to generate a stable and feasible solvent-oil blend.

The identification and selection of solvents usually involves tedious experimentation through trial-and-error process, and it could be time consuming to test all potential solvents and ineffective in optimising blend performance to meet the property targets. In response to these challenges, computer-aided molecular design (CAMD) techniques were commonly employed to identify potential candidates that satisfy the pre-determined product requirements. Previous research on the design of bio-oil solvent focused only on the property targets that can be predicted using group contribution (GC) method. However, there will be instances where different classes of property prediction models were required to estimate different target properties. For example, some of the property prediction models for environmental or non-thermodynamic properties were expressed in terms of topological indexes (TI). Nevertheless, different classes of properties prediction model may be expressed with different mathematical formulations. Thus, it is difficult to simultaneously accommodate and solve different classes of property prediction models using single calculation approach. To overcome this issue, molecular signature descriptor was introduced to incorporate higher-order molecular groups from GC models and multiple TIs on a common platform for CAMD.

Signatures of different height were required to represent property prediction models of different classes. In order to achieve higher accuracy of estimation, signatures of higher height were essential to portray higher-order contributions from the prediction model due to the requirement of higher

structural information. Despite the high accuracy of estimation with the use of higher height signatures, the complexity of CAMD increases due to the combinatorial nature of higher-order signatures. Having several building blocks will result in a large number of signatures at higher height, which leads to computational difficulties in modelling and solving the CAMD problem. Furthermore, not all the signatures considered in the CAMD problem are consistent with each other to form a feasible molecule. Hence, it is important to develop a framework to reduce the size of CAMD problem by excluding irrelevant molecular signature at a lower height from the building block sets.

A major obstacle in the commercialisation of biofuels is their high production cost. The low energy density of biofuels makes them more costly for heat generation as compared to that of the conventional diesel fuel. In other word, the biofuel's energy density decreases as the bio-oil ratio increases in the biofuel. Furthermore, the addition of solvent is often required to improve the bio-oil's properties. In the past, previous research on the design of bio-oil solvent mainly focused on the functionality of the solvent itself. However, this may result in the selection of specialty chemicals as optimal solvents, which are generally associated with high cost. This further increases the production cost of biofuels. In addition, the existing and progressing biofuel legislative framework sparks the urge to assess the cost associated with upgrading bio-oil. Thus, the incorporation of economic aspects into the development of bio-oil solvent for the designed solvent-oil blend is crucial for the solvent-oil blend to be competitive with the conventional diesel fuel. Generally, the heating value of pyrolysis bio-oil increases with the addition of solvent. However, as the

amount of solvent increases, the cost of solvent-oil blend increases as well, resulting in lower profitability. Thus, it is crucial to investigate the trade-off between high heating values and high profitability of the solvent-oil blend.

On the other hand, it is also important to understand and identify the underlying patterns and relations between biomass with bio-oil to produced bio-oil of targeted fuel properties. Many studies have estimated the properties of pyrolysis bio-oil from its elemental composition. However, the production and the characterisation of pyrolysis bio-oil could be labour intensive, costly, and complicated. In addition, rigorous mechanistic models to allow accurate prediction of product properties from feedstock characteristics are not available. To avoid such difficulties, empirical models have been proposed in various studies to estimate the pyrolysis bio-oil's properties based on its characterisation and pyrolysis condition. It is notable that the empirical models developed in the past mainly focused on the prediction of bio-oil yield, while prediction models for other parameters such as higher heating value (HHV) and pH of bio-oil are still limited. To address this research gap, it is important to develop a predictive model that predicts these pyrolysis bio-oil properties from feedstock characteristics and pyrolysis temperature using rough-set machine learning (RSML).

Normally, most of the work on CAMD for solvent design ended at the identification of promising chemical compounds or solvents through computational approach. However, one shall verify that the solvent identified is feasible for real-life application while meeting the required targeted properties and performance. Therefore, it is crucial to include both computational and

experimental approach in the framework for solvent design. In addition, emulsification was also required to produce a stable biodiesel blend (solvent/bio-oil/diesel blend). Emulsification strategies with surfactant addition were reported to improve the stability of biodiesel. Nevertheless, previous research on biodiesel blend with surfactant showed varied properties. Hence, this lead to another research gap where there is a need to develop a more effective emulsification strategy to generate a stable biodiesel blend with promising fuel properties.

3.2. Scopes of Research

Based on the identified research gaps, this research work can be divided into four main scopes, aiming to generate a stable pyrolysis bio-oil blend which demonstrate promising property targets that fulfils the standards of biodiesel while displaying desirable profit margin. The four scopes presented in this thesis are as below:

3.2.1. Multi-stage computational framework for solvent design to upgrade pyrolysis bio-oil

Solvent with promising physical properties and low environmental impact can be designed using CAMD tools. Generally, GC methods have been widely employed in the estimation of thermodynamic related properties. However, most of the available prediction models for environmental related properties are expressed in TI models. As different classes of prediction models use different mathematical formulations, hence it is mathematical challenging to combine and solve prediction models of different classes simultaneously. For this reason, molecular signature descriptors can be introduced as a common

platform to formulate the design problem consisting different property prediction models. However, the complexity of CAMD problem increases with the height of signatures, due to the combinatorial nature of higher-order signatures. Thus, there is a need to develop multi-stage framework to systematically eliminate infeasible signatures at different level to keep a manageable CAMD problem size.

3.2.2. Multi-objective optimisation of solvent-oil blend's fuel quality and economic targets

In research scope 1, CAMD approach has been utilised to identified solvent which satisfy the physical and environmental related properties. Nevertheless, it is also important to incorporate the economic aspects into the design of solvent-oil blend in order to be competitive with the conventional diesel fuel. Generally, the heating value of pyrolysis bio-oil increases with the addition of solvent. However, the cost of solvent-oil blend increases as the amount of solvent increases. To address this issue, multi-objective optimisation (MOO) approach has to be applied to investigate the trade-off between high heating value and high profitability of the solvent-oil blend.

3.2.3. Estimation of fast pyrolysis bio-oil properties from the biomass feedstock characterisation and pyrolysis condition

Other than the solvent addition, the fuel quality of pyrolysis bio-oil can also be improved by selecting the suitable biomass feedstock and pyrolysis condition. In the past, various machine learning (ML) tools have been used for predicting the performance of thermochemical biomass processing. However, popular black-box ML techniques such as artificial neural network (ANN) and

support vector machine (SVM) suffer from poor interpretability. Other than ML techniques, statistical tools including response surface model (RSM) and multiple linear regression (MLR) approaches also lead to regression models that are also black-box in nature. Nevertheless, interpretability is important for ML tools to provide explanation on their models and better understanding on the value and accuracy of their findings. To overcome this limitation, application of an interpretable ML tool that is capable to generate straightforward and transparent rules is crucial for further information extraction, other than unveiling the relationship between the input and output attributes.

3.2.4. Generate stable solvent-bio-oil-diesel blend with the aid of emulsifiers through experiment.

In most of the studies, the solvents generated from CAMD models are validated with the existing online database. Nevertheless, it is crucial to verify the feasibility of identified solvent in real-life application via experimental verification. Thus, there is a need to develop a solvent design framework which combines both computational and experimental approach. Instead of verifying the performance of all available solvents via trial-and-error approach, only the shortlisted potential solvents identified from the CAMD model need to be verify now. Furthermore, emulsification of solvent-oil blend and diesel fuel with the aid of emulsifier is essential to produce a stable biodiesel blend. However, previous research on emulsification of biodiesel demonstrated varied outcomes. Hence, it is crucial to develop an efficient emulsification strategies and determine the optimal mixing ratio to generate a stable solvent-bio-oil-diesel blend.

3.3. Research Objectives

The research scopes identified in Section 3.2 leads to the specific objectives which were stated as followings:

1. To develop a multi-stage CAMD framework for the design of pyrolysis bio-oil's solvent via molecular signature descriptor to upgrade pyrolysis bio-oil's fuel properties.
2. To optimise the fuel functionality and economic aspect of solvent-oil blend via fuzzy multi-objective optimisation approach.
3. To estimate the fast pyrolysis bio-oil's HHV and pH value from biomass feedstock characterisation and pyrolysis condition via RSML model.
4. To generate a stable biodiesel blend with promising fuel performance via ultrasonification approach.

3.4. Overall Research Methodology

The overall methodology for this research work can be generalised into four main stages: research scope 1, research scope 2, research scope 3 and research scope 4, which is briefly discussed in Section 3.4.1 to 3.4.4. The detailed methodologies of all the four scopes are presented in the following Chapters 4 to 7. Figure 3-1 demonstrated the graphical representation of the overall research methodology.

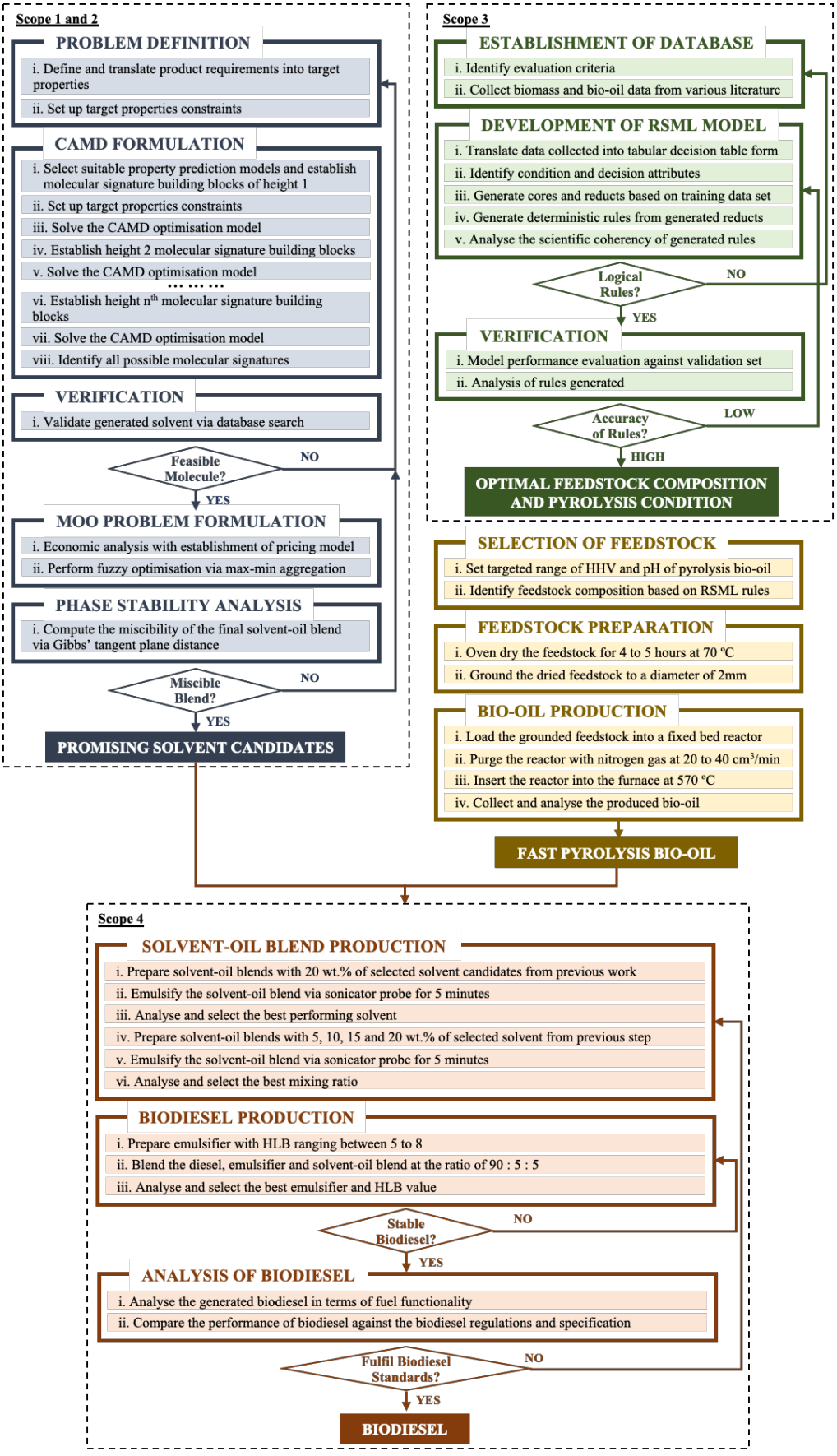


Figure 3-1 Overall methodology developed for this research work

3.4.1. Multi-stage computational framework for solvent design to upgrade pyrolysis bio-oil

Solvent addition is a simple method to overcome the poor fuel properties of pyrolysis bio-oil to allow further processing and storage. To design a solvent that is capable to improve the stability and fuel properties of the pyrolysis bio-oil, CAMD tools can be employed to identified molecules possessing desired properties based on the pre-determined product requirements. Initially, target solvent requirements can be translated into measurable physical properties. The identified target properties will either be used as constraints or optimisation objective. As different property prediction models consist different levels of structural information, molecular signature descriptor can be used as a common platform to formulate the design problem. Due to the combinatorial nature of higher order signatures, the complexity of a CAMD problem increases with the height of signatures. Thus, a multi-stage framework can be developed by introducing the consistency rules that restrict the number of higher-order signatures. Next, database search is conducted on the all the generated solvent to ensure its feasibility and is practical to be applied in real-life applications. In addition, phase stability analysis is required to evaluate the stability of the solvent-oil blend. Sensitivity analysis was also conducted on bio-oil's water content to investigate its effect on the solvent ratio and miscibility of the final solvent-oil blend. The detailed methodology of this scope can be found in Chapter 4.

3.4.2. Multi objective optimisation of solvent-oil blend's fuel quality and economic targets

In Section 3.4.1, solvent design was performed to generate a solvent which is capable to upgrade bio-oil's properties in terms of bio-oil's fuel performance and environmental aspect. In the second scope, the previous work is further extended to incorporate the consideration of economic aspect into the solvent design of bio-oil. In this study, the solvent-oil blend's quality can be related to the demand and price of the blend by employing the pricing model. In addition, a pyrolysis plant is proposed to aid the estimation of pyrolysis bio-oil's production cost. Generally, the higher the mass fraction of solvent in the solvent-oil blend, the higher is the energy content. However, a higher amount of solvent was often associated with higher cost, and thus, lower profitability obtained from the solvent-oil blend. Thus, fuzzy optimisation approach can be employed to investigate the trade-off between the functionality and profitability of the solvent-oil blend. With the identification of optimal solvent-oil blend ratio, phase stability analysis will be carried out to ensure the miscibility of the designed solvent-oil blend at the targeted mixing ratio. In addition to that, an economic study on the bio-oil-diesel blend will also be conducted to investigate the relationship between the ratio of bio-oil-diesel blend, the price and the HHV of the bio-oil/diesel blend. The proposed methodology for this scope is further discussed in Chapter 5.

3.4.3. Estimation of fast pyrolysis bio-oil properties from the biomass feedstock characterisation and pyrolysis condition

For the third scope, a data-driven rough-set based ML model was proposed as a pre-processing and predictive modelling to predict the pyrolysis bio-oil properties based on pyrolysis temperature and feedstock characteristics. At the first stage, a database consisting of biomass feedstock proximate and ultimate analyses, pyrolysis temperature, bio-oil's pH value, and bio-oil's HHV are compiled based on published literature and experimental studies. The generated decision rules will be analysed from a scientific standpoint to identify the underlying trends or patterns, while ensuring the logical and feasibility for the rules to be applied in the later stage. The most suitable prediction models and decision rules will be select based on the novel insights and mechanistic plausibility of the decision rules. In addition, validation dataset will be used to evaluate the statistical performance of rule-based models to ensure good predictive performance and low overfitting. Four case studies were proposed to investigate the effect of different combination of biomass characteristics and pyrolysis temperature on the HHV and pH values of the pyrolysis bio-oil.

3.4.4. Generate stable solvent-bio-oil-diesel blend with the aid of emulsifiers.

In order to validate feasibility and applicability of the solvents identified from scope 1 and 2 (as mentioned in Section 3.2.1 and 3.2.2, respectively), further experimental verifications are required. In the fourth scope, an efficient emulsification strategy via sonification is proposed to generate stable solvent-oil-diesel blend. At the first stage, the best performing solvent are selected based

on its fuel performance at 20 wt. % of solvent mixing ratio. Then, the selected solvent will be mixed with pyrolysis bio-oil at different mixing ratio to find out the optimal solvent-oil blend ratio. Once the optimal mixing ratio are determined, the solvent-oil blend will be mixed with diesel with the aid of surfactants. The generated biodiesel blend will be analysed in terms of its fuel functionality and stability.

3.5. Summary

This chapter presents the research scopes and the overall research methodology. Based on the research gaps identified in Section 3.1, there is a need to design a solvent which could overcome all the shortcomings of pyrolysis bio-oil via CAMD framework. Owing to the combinatorial nature of the molecular signature descriptor, the complexity of the CAMD problem increases as the height of signature increases. Thus, a multi-stage framework needs to be developed to restrict the number of higher-order signature. In addition, it is also crucial to incorporate economic aspects into the solvent design. Hence, a fuzzy MOO approach should be developed to simultaneously optimise both fuel functionality and economic targets of solvent-oil blend. Besides, it is found that the properties of pyrolysis bio-oil can be related with the characterisation of biomass feedstock and pyrolysis condition. Therefore, straightforward rules can be generated using ML approach to estimate the HHV and pH of pyrolysis bio-oil based on the characterisation of feedstock and pyrolysis condition. Nevertheless, it is also important verify the feasibility of the designed solvent in real-life application via experimental verification. The methodology should

be further extended to include the experimental procedures where pyrolysis bio-oil are emulsified with diesel with the aid of solvent and emulsifiers.

CHAPTER 4

DESIGN OF BIO-OIL SOLVENTS VIA MOLECULAR SIGNATURE DESCRIPTORS USING A MULTI-STAGE COMPUTER-AIDED MOLECULAR DESIGN FRAMEWORK

4.1. Introduction

In this chapter, computer-aided molecular design (CAMD) tools were developed to design optimal solvents to upgrade bio-oil whilst having low environmental impact. Firstly, target solvent requirements were identified and translated into measurable physical properties. As different property prediction models consist different levels of structural information, molecular signature descriptor was used as a common platform to formulate the design problem. Because of the differences in the required structural information of different property prediction models, signatures of different heights were needed in formulating the design problem. Due to the combinatorial nature of higher-order signatures, the complexity of a CAMD problem increases with the height of signatures. Thus, a multi-stage framework was developed by introducing a novel consistency rule that restricts the number of higher-order signatures. Finally, phase stability analysis was conducted to evaluate the stability of the solvent-oil blend. As a result, optimal solvents that improve the solvent-oil blend properties while displaying low environmental impact were identified. A

case study on solvent design for pyrolysis bio-oil was conducted to illustrate the application of this proposed methodology.

4.2. Problem Statement

In recent years, bio-oil has risen as one of the most promising alternatives for conventional diesel fuels. However, poor fuel quality and stability of bio-oil often limit their applications in diesel engines. Various research has been done on enhancing the bio-oil fuel properties by blending it with solvent at different mixing ratio. Having a solvent that can fulfil multiple property targets is essential to produce a stable solvent-oil blend with improved physicochemical properties. By utilising the CAMD tools, an optimal solvent molecule can be predicted together with identifying the underlying reaction mechanisms. At present, most of the solvent design via CAMD approach had been done on group contribution (GC)-based methods. However, some property prediction models were expressed in indexes other than the GC method. To consider all related property prediction models, CAMD approach with molecular signature descriptors can be applied. Despite the higher accuracy of estimation at a higher height of signature, the complexity of CAMD increases as well due to the vast amount of possible molecular signatures to be considered. Thus, it is important to develop a framework in reducing the size of problem by removing irrelevant molecular signature at a lower height from the building blocks set.

4.3. Methodology

The main objective of this work is to develop a systematic multi-stage framework in reducing the size of CAMD problem due to the combinatorial

nature of molecular signature descriptor. Solvents that form stable blends with pyrolysis bio-oil and possess optimal properties can be generated with this framework by considering physical, environmental, and thermodynamic properties. An algorithm of GC method coupled with topological index (TI) approach was used to solve the multiple property indexes involved in the CAMD problem. This framework can be divided into 4 main stages and their correlated sub-steps are shown in Figure 4-1.

Error! Reference source not found.

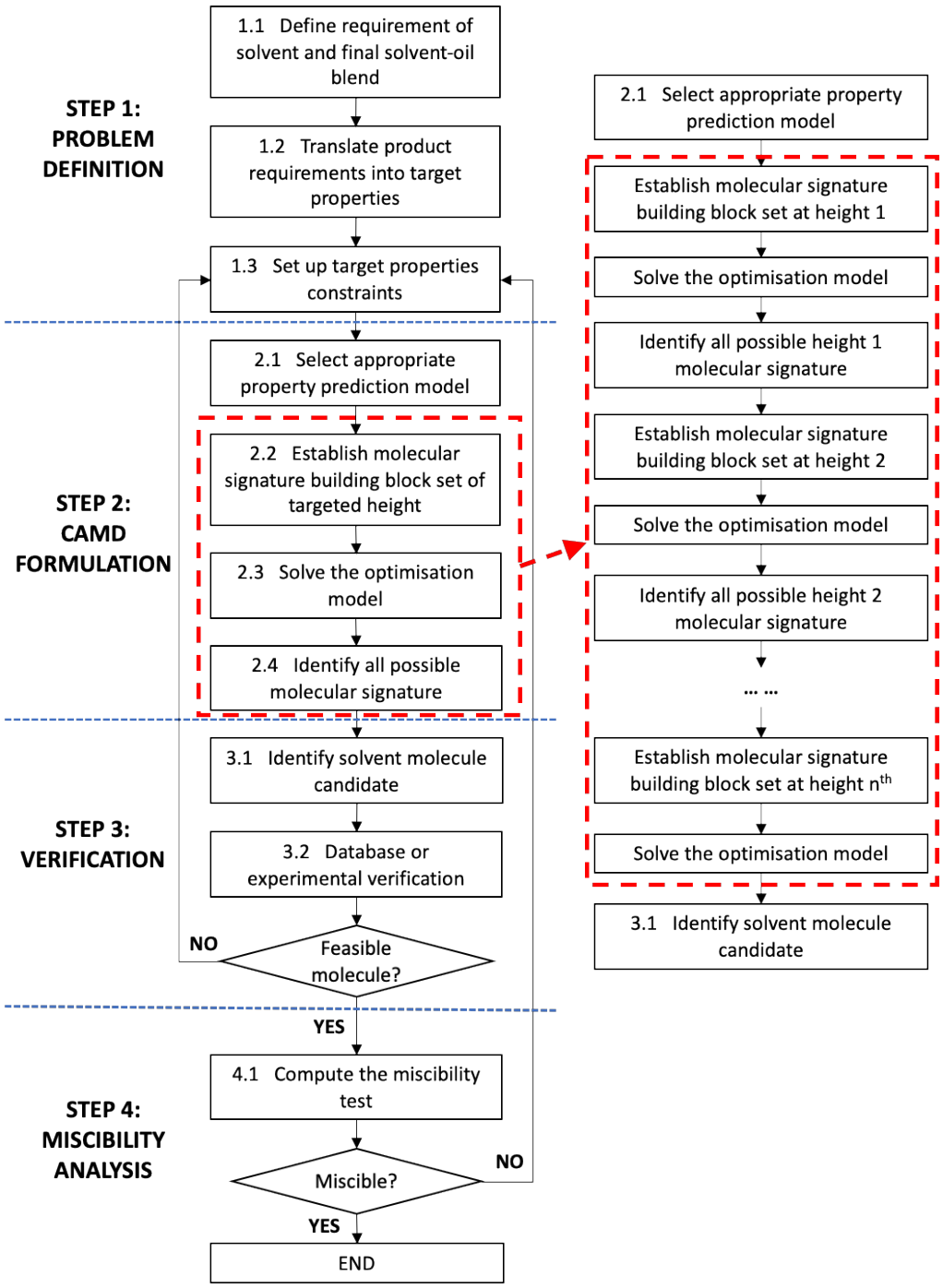


Figure 4-1 Framework for the development of CAMD model for the design of solvent additives

4.3.1. Problem Definition

Firstly, the problem definition was formulated, where the product needs were determined based on the requirements from regulations and specifications. This usually requires data on physical and thermodynamic properties as they contribute to the functionality of the designed product. In addition, environmental properties were considered to ensure that the generated solvent molecules have low environmental impact. The selected desired properties will be serving as the design objective to generate promising molecules.

The identified product requirements were then translated into measurable quantitative target properties. For example, the flow consistency of solvent can be expressed in terms of its density and viscosity. These identified target properties will either be used as constraints or optimisation objective in the CAMD formulation stage. Upper and lower limits were defined for these target properties to ensure the designed solvents display similar physical characteristics as a conventional solvent.

4.3.2. Property Prediction Models

Next, suitable property prediction models were selected to estimate the target properties of the solvent. In this work, property prediction models in terms of GC and TI approaches were considered and expressed as a function of the molecular signature descriptor. For GC based property prediction models, higher-order molecular groups demonstrated higher prediction quality compared to the 1st order approach. In this work, higher-order molecular groups (2nd and 3rd order groups) along with the basic molecular building blocks (1st order) have been considered. Molecular signatures of desirable heights were

generated based on the root atoms and chemical families selected for the solvent design. The height and number of signatures required to describe the molecular group in GC models depend on the number of atoms present for the molecular structure and the nature of the final molecule. Thus, maximum signature height for the CAMD problem can be determined from the available property prediction models.

4.3.3. CAMD Formulation

The CAMD optimisation model were represented using the following set of generalised mathematical expression (Austin et al., 2016):

$$F_{obj} = \max F(x, p) \tag{4.1}$$

$$p = f(x) \tag{4.2}$$

$$h_1(p, x) \leq 0 \tag{4.3}$$

$$h_2(p, x) = 0 \tag{4.4}$$

$$s_1(x) \leq 0 \tag{4.5}$$

$$s_2(x) = 0 \tag{4.6}$$

$$p_k^L \leq p_k \leq p_k^U \quad \forall k \tag{4.7}$$

$$x_d^L \leq x_d \leq x_d^U \quad \forall d \tag{4.8}$$

For the above expressions, p is the vector of properties and p_k is the property values for each property k. Meanwhile, n is the vector representing the structural information of designed molecules. The x_d vector indicates the number of occurrences of each molecular signature d. The function f then transforms this structural information into a property estimate using the appropriate QSPR relationship. Equation 4.1 is a general objective function for the CAMD problem. The $F(x, p)$ is known as the vector of objective function

which quantify the performance of the designed molecule based on its properties p. The F function can either be maximised or minimised depending on the design problem. Equation 4.2 is the function f which estimates vector of properties p from attributes such as number of molecular signatures. Equations 4.3 and 4.4 are the general function representing the inequality and equality constraints, respectively. These equations corresponded to product design specifications such as property's value for thermodynamic and environmental properties. As the property depends on the presence of signatures, these constraints can control the number of appearances of specific signatures in the designed molecule. As for Equations 4.5 and 4.6, they are the general function representing the inequality and equality constraints, respectively, related to the molecular structure generation. These structural constraints ensure the generated molecule is structurally feasible. Equations 4.7 and 4.8 are the boundaries set on property values and the number of signatures. The parameters p_k^L and x_d^L are the lower bounds for property k and x_d, respectively. Similarly, parameter p_k^U and x_d^U are the upper bounds for property k and x_d, respectively.

For the CAMD problem, the function $f(x)$ may be formulated as a mixed integer non-linear programming (MINLP) model. However, due to the increasing size of the mathematical problem, it is usually challenging to solve such a MINLP models with the structural information included (Zhang et al., 2015). In this work, molecular signature descriptors were used to present the CAMD problem as an equivalent mixed integer linear programming (MILP). The two-dimensional descriptor (both TI and GC models) of molecule G, $TI(G)$ can be expressed as a dot product between two vectors, ${}^{h}\alpha_g$, the vector of

occurrence number of atomic signatures of height h, and $TI\left(root(\,{}^{h}\sum\,)\right)$, the vector of predicted values from the model computed for each of the atomic signatures as shown in (Equation 4.9) (Chemmangattuvalappil and Eden, 2013):

$$TI(G) = k^{h}\alpha_{g} \cdot TI\left(root\left({}^{h}\sum\right)\right) \tag{4.9}$$

By using the signatures, the non-linear part of the mathematical formulation can be hidden inside the molecular signature building blocks. As shown in Equation 4.9, the prediction model was expressed as linear equation, where the only variable is the number of respective building blocks, which are atomic signatures. In each of the building blocks, contribution to the target property can be estimated using the original property prediction model. When GC model was used for property prediction, the term in the bracket referred to the property contribution from the signature that represented the set of molecular groups present in the higher-order groups. In CAMD, since the structure of molecule is not known prior to the design, the presence of higher-order groups is not known during the design. Molecular signatures were used to track the presence of all possible 2nd and 3rd order contributions by considering a signature height of 2 or 3. These models can be linear or non-linear. However, since these non-linear expressions were only used in the estimation of contribution of the building blocks, hence it will not be part of the variables. The only variable involved will be the number of appearances of each signature. Considering the prediction model for normal melting point, T_m as an example (Equation 4.10)(Marrero and Gani, 2001):

$$\exp\left(\frac{T_m}{T_{m0}}\right) = \sum_i N_i T_{mi} + \sum_i M_j T_{mj} + \sum_i O_k T_{mk} \qquad (4.10)$$

$$T_{m0} = 147.45K$$

Where N_i, M_j and O_k are the number of individual building blocks from 1st order, 2nd order and 3rd order groups, respectively. The parameters T_{mi}, T_{mj} and T_{mk} are the contributions for 1st order, 2nd order and 3rd order groups, respectively. T_m is the normal melting point and was set to be lower than 298.15 K. The exponential term, $\exp\left(\frac{T_m}{T_{m0}}\right)$ at the left-hand side of Equation 4.10 contributes to the non-linearity of the expressions. By substituting and solving the left-hand side of the equation, the prediction model is now a linear equation, as shown in Equation 4.11. The only variable in the equation is the number of building blocks.

$$7.554 > \sum_i N_i T_{mi} + \sum_i M_j T_{mj} + \sum_i O_k T_{mk} \qquad (4.11)$$

However, this approach will lead to the generation of a substantial amount of molecular signature building blocks to be considered for CAMD. To address this issue, a multi-level approach has been developed where the amount of generated molecular signature building blocks can be controlled.

4.3.3.1. Feasibility Rules

To ensure the feasibility of the final molecule, the selected signature building blocks should fulfil the requirements to form effective solvents. An efficient, structured algorithm for joining groups to form feasible chemical compounds was integrated into the signature based CAMD (Gani et al., 1991; van Dyk and Nieuwoudt, 2002). Generally, molecules are reported to be

unstable if two heteroatoms are bonded to the same carbon atom, and at least one of the atoms is also bonded to a hydrogen atom. Combination like heteroatom bonding with another should be avoided as these compounds are usually highly reactive and not suitable to be considered as solvents. Constraints from the work of van Dyk and Nieuwoudt (2002) classified the groups of molecules according to the type of free bonds as shown in Table 4-1. In general, Table 4-2 can be summarised based on Equation 4.12; where n_i is the total number of free bond group i in the molecule (van Dyk and Nieuwoudt, 2002).

$$n_4 + n_5 \geq n_1 + n_2 \tag{4.12}$$

Table 4-1 Free bond groups in terms of signature of height 2.

Group	Description	Example
I	Bonding atom is a heteroatom bonded to a hydrogen atom	O1(C2(CO))
II	Bonding atom is a heteroatom bonded to a carbon atom	O2(C2(CO)C3(CCO))
III	Bonding atom is a carbon atom bonded to a heteroatom, which is bonded to a hydrogen atom	C2(O1(C)C2(CC))
IV	Bonding atom is a carbon atom bonded to a heteroatom, which is bonded to a carbon atom	C2(O2(CC)C2(CC))
V	Bonding atom is a carbon atom bonded to another carbon atom	C2(C2(CC)C3(CCC))

Table 4-2 Allowed combination of group

Group	I	II	III	IV	V
I	✗	✗	✗	✗	✓
II	✗	✗	✗	✓	✓
III	✗	✗	✓	✓	✓
IV	✗	✓	✓	✓	✓
V	✓	✓	✓	✓	✓

As some of the property prediction models used GC method, signature descriptors were translated and assigned to their corresponding groups from GC approach. Only the root atom of each atomic signature was considered to prevent the overlapping issues during the property estimation. Taking the molecular signature C2(CC) as reference, the root atom C was connected to 2 carbon atoms by single bonds and the rest was bonded to hydrogen atoms. Thus 'CH_2' is the corresponding GC group for this signature. In another example, the signature C4(=CCO) has the root atom C connected to an oxygen and two carbons by single and double bonds, respectively. To ensure no overlapping of groups, the simplest equivalent group was chosen. In this case, the group 'C = C' was chosen. The other examples of height 1 signatures and their corresponding groups are listed in Table 4-3.

Table 4-3 Height 1 signature and their corresponding group

Signatures	Corresponding Group
C4(=C=C)	C=C
C4(=CCC)	C=C
C4(CCCC)	C
C3(=CC)	CH=C
C3(CCC)	CH
C2(=C)	CH_2=C
C2(CC)	CH_2
C1(C)	CH_3

4.3.3.2. Structural Constraints

Structural constraints are essential in a CAMD problem to ensure the formation of a complete molecular graph with all signatures connected to generate a feasible solution. The structural constraints used in molecular

signature-based algorithms must follow a few rules in order to generate a complete structure (Chemmangattuvalappil et al., 2010):

1. Signatures must be connected without any free bonds in the structure. Thus, the total number of available degrees (valencies) should be matching with the total number of vertices (atoms) in the graph (molecules).
2. The number of bonds in each signature should be consistent with the bonds of the rest of signatures.

To obey rule (I), Equation 4.13 was developed to express the relation between the number of signatures and the bonds where n_1, n_2, n_3, and n_4 are the number of signatures x_i with valency of one, two, three and four respectively. Here N_{Di}, N_{Mi} and N_{Ti} are the signatures with one double bond, two double bond and one triple bond respectively (Chemmangattuvalappil et al., 2010).

$$\sum_{i=1}^{n_1} x_i + 2\sum_{n_1}^{n_2} x_i + 3\sum_{n_2}^{n_3} x_i + 4\sum_{n_3}^{n_4} x_i = 2\left[\left(\sum_{i=1}^{N} x_i + \frac{1}{2}\sum_{i=0}^{N_{Di}} x_i + \sum_{i=0}^{N_{Mi}} x_i + \sum_{i=1}^{N_{Ti}} x_i\right) - 1\right] \tag{4.13}$$

On the other hand, rule (II) can be mathematically represented as Equation 4.14, which must be fulfilled by all colour sequences, including colour sequences in which $i = j$ at each height. The expression $\left(l_i \rightarrow l_j\right)_h$ is for colouring sequence $l_i \rightarrow l_j$ at level h (Chemmangattuvalappil et al., 2010):

$$\sum \left(l_i \rightarrow l_j\right)_h = \sum \left(l_j \rightarrow l_i\right)_h \tag{4.14}$$

From Equation 4.14 we can know that if there is a signature with a colour sequence of $l_i \rightarrow l_j$ with $i = j$, then there must be another signature with the

same colour sequence present from the same signature set to complement the previous signature. Considering a signature of height 2, $C4(C4(CCC)C1C4(CCC)C3(CC))$ as example. This signature contains two $C4 \rightarrow C4$ colourings, one $C4 \rightarrow C1$ colouring and one $C4 \rightarrow C3$ colouring. Thus, there must be two more $C4 \rightarrow C4$ colourings, one $C1 \rightarrow C4$ colouring and one $C3 \rightarrow C4$ colouring present in other signatures of the same molecule. This can be expressed mathematically as Equation 4.15 to confirm the consistency of the signatures with the same colour sequence on the same edges where η is the number of colour sequence of $l_i \rightarrow l_j$ on a signature with $i = j$, x is the number of such colour sequence and K is a constant integer (Chemmangattuvalappil et al., 2010).

$$\sum_{i=j} \eta_i x_i = 2K \tag{4.15}$$

It is also important to make sure that the number of a specific colour in the child level would not exceed the total number of the same colour when it is in parent level. In other words, the total number of signatures with more degree of the vertex at a higher level should be lesser than the total number of vertices with the higher degree which can be represented by Equation 4.16 (Chemmangattuvalappil et al., 2010):

$$\sum x_i n_i \leq \sum x_j \tag{4.16}$$

Where n_i is the number of child vertex with higher degree than the parent vertex, while both i and j correspond to the colour of child and parent, respectively.

4.3.3.3. Consistency rules

The CAMD problem was initially solved at height 1 level to identify promising signatures generated from the previous stage. Subsequently, height 2 signatures were generated based on the identified height 1 signatures. However, to ensure the final generated molecule is structurally feasible, only signatures that fulfil the structural constraints were considered.

To generate a feasible molecular structure from the signature building blocks, each signature must be connected to another signature that with the same structure at a level $h-1$. An example on the enumeration of molecular structures from signatures are shown in Table 4-4. The collection of signatures presented in this example is one of the solutions obtained for the bio-oil solvent case study in Section 4.4.

Table 4-4 Set of signatures for 2-Octanol with its corresponding height 2 signatures

No	Height 3 Signatures	Corresponding Height 2 Signature
1	C1(C3(C1(C)C2(CC)O1(C)))	C1(C3(CCO)
2	C1(C2(C1(C)C2(CC))	C1(C2(CC))
3	C2(C1(C2(CC))C2(C2(CC)C2(CC)))	C2(C1(C)C2(CC))
4	C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))	C2(C2(CC)C2(CC))
5	C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))	C2(C2(CC)C2(CC))
6	C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO)))	C2(C2(CC)C2(CC))
7	C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C)))	C2(C2(CC)C3(CCO))
8	C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO)))	C3(C1(C)C2(CC)O1(C))
9	O1(C3(C1(C)C2(CC)O1(C)))	O1(C3(CCO))

First, any signature of height 3 was selected. In this case, Signature (1), C1(C3(C1(C)C2(CC)O1(C))) was selected. Next, it was inferred that there is only one signature possible from the first layer, which is

C3(C1(C)C2(CC)O1(C)). Referring to Table 4-4, the height 2 signature of Signature (8), C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))) was exactly same as the signature from the first layer. Thus, Signature (1) was connected with Signature (8). The same procedure was then repeated on Signature (8) to get the next bond. In this study, an algorithm was developed based on the graph signature enumeration method by Faulon et al. (2003).

In the developed approach, signatures of height h were generated based on the collection of height $h-1$ signatures identified from the CAMD problem. The first layer of signature generated must contain one of the height $h-1$ signatures from the previous result. For example, assuming the signatures C1(C), C2(CC), C2(CO) and C3(CCO) were identified as the promising height 1 signature from the CAMD problem, the generated height 2 signatures based on C1(C) are shown as below:

1. C1(C2(CC))
2. C1(C2(CO))
3. C1(C3(CCO))

With this approach, the total number of generated height 2 signatures was reduced from 13 signatures to 3 signatures. In another example, taking the collection of height 2 signatures, the following set is obtained:

1. C1(C3(CCO)
2. C1(C2(CC))
3. C2(C1(C)C2(CC))
4. C2(C2(CC)C2(CC))
5. C2(C2(CC)C3(CCO))
6. C3(C1(C)C2(CC)O1(C))

7. O1(C3(CCO))

In this case, height 3 signatures generated based on the Signature (3), C2(C1(C)C2(CC)) are listed as:

1. C2(C1(C2(CC))C2(C1(C)C2(CC)))
2. C2(C1(C2(CC))C2(C2(CC)C2(CC)))
3. C2(C1(C2(CC))C2(C2(CC)C3(CCO)))

Similar approach was applied to the rest of signatures to generate the remaining height 3 signatures.

4.3.4. Verification

Verification step is crucial to ensure that the molecules generated from previous steps are feasible and practical. In this step, generated molecules were verified through database search from various platforms like ChemSpider, PubChem, etc. For compounds that exist in the database, comparison was made to verify the property values obtained from the CAMD result. As for compounds that do not exist in the database or proved to be infeasible, the previous step was repeated by modifying the property attributes and constraints.

4.3.5. Miscibility Analysis

It is crucial to ensure the designed solvent is miscible with the blend of bio-oil and diesel to avoid phase separation in the final solvent-oil blend. Phase stability test was conducted by computing the tangent plane distance. For an n-component mixture at constant temperature and pressure, the phase stability analysis employed the Gibbs tangent plane distance function as shown in Equation 4.17 (Prausnitz, 1969):

$$d(x) = \sum_{i=1}^{n} x_i[\ln x_i \gamma_i(x) - \ln z_i \gamma_i(z)] \quad (4.17)$$

Where z is the compositions of component i in mole fractions of the tested phase, x is the composition component i of a trial phase and γ indicates the activity coefficient of component i in respective phase. For mixture that is stable and exhibits homogenous single-phase, the following equation (Equation 4.18) can be followed (Baker et al., 1982):

$$d(x) \geq 0 \quad (4.18)$$

To estimate the activity coefficients in non-ideal liquid mixture, group contribution estimation approach developed by Fredenslund et al. (1975) was applied. In this work, the GC prediction model combines the solution-of-functional-groups concept with a model for activity coefficient based on UNIQUAC. In a multi-component mixture, the UNIQUAC equation for the activity coefficient of component i is given by Equation 4.19 (Fredenslund et al., 1975):

$$\ln \gamma_i = \ln \gamma_i^C + \ln \gamma_i^R \quad (4.19)$$

In Equation 4.19, C represented the combinatorial part while the residual part is denoted as R. Here, Equation 4.20 and 4.21 calculates the value $\ln \gamma_i^C$ and $\ln \gamma_i^R$ (Fredenslund et al., 1975):

$$\ln \gamma_i^C = \ln \frac{\phi_i}{x_i} + 5q_i \ln \frac{\theta_i}{\phi_i} + l_i - \frac{\phi_i}{x_i} \sum_j x_j l_j \quad (4.20)$$

$$\ln \gamma_i^R = \sum_k v_k^{(i)} (\ln \Gamma_k - \ln \Gamma_k^{(i)}) \quad (4.21)$$

Equation 4.22 to Equation 4.29 represents the calculation for terms in Equation 4.20 and 4.21 (Fredenslund et al., 1975):

$$l_j = 5(r_i - q_i) - (r_i - 1) \tag{4.22}$$

$$\phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \tag{4.23}$$

$$\theta_i = \frac{q_i x_i}{\sum_j q_j x_j} \tag{4.24}$$

$$r_i = \sum_k v_k^{(i)} R_k \tag{4.25}$$

$$q_i = \sum_k v_k^{(i)} Q_k \tag{4.26}$$

$$\ln \Gamma_k = Q_k \left[1 - \ln \sum_m \vartheta_m \psi_{m,k} - \sum_m \frac{\vartheta_m \psi_{m,k}}{\sum_n \vartheta_n \psi_{n,m}} \right] \tag{4.27}$$

$$\vartheta_m = \frac{Q_m X_m}{\sum_n Q_n X_n} \tag{4.28}$$

$$\psi_{m,n} = -\exp(\frac{a_{mn}}{T}) \tag{4.29}$$

Where γ_i = activity coefficient of component i

ϕ_i = segment fraction (volume fraction) of component i

θ_i = area fraction of component i

x_i = mole fraction of component i

r_i = pure component molecular van der Waals volume parameter

q_i = pure component molecular surface areas parameter

$v_k^{(i)}$= number of groups of type k in molecule i

R_k = group volume parameters

Q_k = group area parameters

Γ_k = group residual activity coefficient

$\Gamma_k^{(i)}$ = residual activity coefficient of group k in pure component i

ϑ_m = area fraction of group m

$\psi_{m,k}$ = group interaction parameter

X_m = mole fraction of group m in the mixture

$a_{m,n}$ = group interaction parameters of experimental phase equilibrium data

After obtaining feasible solvent candidates from the previous expressions, solvents and bio-oil are tested at different mixing ratio with the application of tangent plane criterion, to obtain the optimal mixing ratio. The Gibbs energy curve is constructed by reviewing the expression $x_i \ln x_i \gamma_i(x)$ for different mixture composition of x, followed by calculation of expression $x_i \ln z_i \gamma_i(z)$ to obtain the tangent to the curve where z denotes the designed mixing ratio generated from the optimisation results. The solvent-oil blend is said to be in stable if the tangent plane distance is non-negative. If otherwise, the previous steps will be revisited by modifying the property attributes and constraints.

4.4. Case Study

4.4.1. Defining Target Properties and Constraints

The main objective of the designed solvent is to improve the fuel quality of the pyrolysis bio-oil. The designed solvent should always be in a liquid state at room temperature for ease of handling and storage. Thus, the constraints for the normal melting and boiling points of solvent were set at 298.15 K and 400.15 K, respectively. On the other hand, a greater higher heating value (HHV) is preferable for better fuel combustion. In present work, HHV for the designed solvent was maximized, which serves as the objective function. Besides, the final solvent-oil blends are expected to display good continuous flow. The

designed solvent should also exhibit high miscibility in pyrolysis bio-oil to ensure the homogeneity of the final blend. Finally, the designed solvent should also comply with the environmental regulations set by authorities for low environmental impact. The generated solvents should possess low toxicity with minimal accumulation in both land and aquatic ecosystem. The final solvent-oil blends are also required to be environmentally sustainable, which is usually measured by the global warming potential (Pacheco and Silva, 2019). In order to reduce the formation of photochemical smog, low photochemical oxidation potential is expected for the final biodiesel blend (Ooi et al., 2018). Constraints for the properties mentioned above were defined according to the ASTM D6751 and EN:14214 standards. Table 4-5 shows the respective targeted properties and identified constraints for each product requirements.

Table 4-5 Translation of product requirements into target properties and constraints

Product Requirements	Targeted Properties	Constraints
Liquid state at room temperature	Normal boiling point (K)	> 400.15
	Normal melting point (K)	< 298.15
Fuel combustion quality	Higher heating value	To be maximised
Fuel flow consistency	Viscosity (mPa s)	$1 > v > 6$
	Density (kg m^{-3})	$800 > \rho > 1000$
Homogenous form	Tangent plane distance	To be determined
Environmental related properties and toxicology	Aquatic acute toxicity, LC_{50}	> 100
	Aquatic acute toxicity, EC_{50}	> 100
	Oral acute toxicity, LD_{50}	> 100
	Bioconcentration factor	< 1000
	Soil-water partition coefficient (L kg^{-1})	< 31622
	Global Warming Potential	< 10
	Photochemical Oxidation Potential	< 10

4.4.2. Selecting Appropriate Property Prediction Model

Based on the target properties identified in Section 4.4.1, the respective property prediction models were selected to estimate the properties of the designed solvents as shown in Table 4-6.

Table 4-6 Property prediction models

Property	Property Model
Normal Melting Point (T_m) (Marrero and Gani, 2001)	$\exp\left(\frac{T_m}{T_{m0}}\right) = \sum_i N_i T_{mi} + \sum_i M_j T_{mj} + \sum_i O_k T_{mk}$ $T_{m0} = 147.450K$
Normal Boiling Point (T_b) (Marrero and Gani, 2001)	$\exp\left(\frac{T_b}{T_{b0}}\right) = \sum_i N_i T_{bi} + \sum_i M_j T_{bj} + \sum_i O_k T_{bk}$ $T_{b0} = 222.543K$
Flash Point (F_p) (Hukkerikar et al., 2012b)	$F_p - F_{p0} = \sum_i N_i T_{pi} + \sum_i M_j T_{pj} + \sum_i E_k T_{pk}$ $F_{p0} = 150.0218K$
Molar Volume (V_m) (Constantinou et al., 1995)	$V_{m(25^oC)} - d = \sum_i N_i v_{m1i} + \sum_j M_j v_{m2j}$ $d = 0.01211$
Vapour Pressure (P_{sat}) (Sinha and Achenie, 2001)	$P_{sat} = 5.58 - 2.7\left(\frac{T_b}{298.15}\right)^{1.7}$ $T_b = Normal\ Boiling\ Point$
Higher Heating Value (HHV) (Walters, 2002)	$HHV = \frac{\sum_i N_i H_i}{\sum_i N_i M_i}\left({}^{kJ}/_{g}\right)$ $M = molecular\ weight\ (g/mol)$
Density (ρ)	$\rho = \frac{M_i}{V_m}$
Dynamic Viscosity (η_L) (Conte et al., 2008)	$\ln \eta_L = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k$
Octanol/water partition coefficient (K_{ow}) (Boyd et al., 1982)	$\log K_{ow} = 1.267\ ({}^v\chi^1) + 0.612({}^v\chi^3) - 0.976({}^v\chi^3) - 2.130$
Acute Toxicity (Aquatic, LC_{50}) (Manuela Pavan et al., 2005)	$\log LC_{50}^{-1} = 0.81 \log K_{ow} + 1.744$
Acute Toxicity (Aquatic, EC_{50}) (FM Carpanini, 1998)	$\log EC_{50} = -0.95 \log K_{ow} - 1.32$
Acute Toxicity (Oral, LD_{50}) (Lipnick, 1989)	$\log LD_{50}^{-1} = 0.805 \log K_{ow} - 0.971 \log(0.0807K_{ow} + 1) + 0.984$
Relative Toxicity (IGC_{50}) (Schultz et al., 2002)	$\log IGC_{50}^{-1} = 0.723(0.14) \log K_{ow} - 1.79(0.031)$
Bioconcentration Factor (Lu et al., 2000)	$\log BCF = 0.032 + 0.636 \log K_{ow}$

Table 4-6 Property prediction models (continued)

Soil-water partition coefficient (K_{oc}) (Liao et al., 1996)	$\log K_{oc} = 0.59\,({}^{v}\chi^{1}) - 0.97$
Global Warming Potential (Hukkerikar et al., 2012a)	$\log GWP = \sum_i a_i A_i + b({}^{v}\chi^{0}) + 2c({}^{v}\chi^{1}) + d$ $b = -0.01877; c = -1.52848; d = -0.52073$
Photochemical Oxidation Potential (Hukkerikar et al., 2012a)	$-\log PCO = \sum_i a_i A_i + b({}^{v}\chi^{0}) + 2c({}^{v}\chi^{1}) + d$ $b = -0.10486; c = 0.005087; d = -0.25708$

In this case study, the chosen property prediction models were expressed in terms of GC and TI method. These property prediction models require different degree of details on the structural knowledge to estimate the properties of the designed molecules. Different TI and GC models require different levels of structural information. Thus, the targeted signature height depends on the required structural information of the TI or GC models. Signatures with higher height contain more structural information of the molecules. It is possible to enumerate the lower order signature from a higher-order signature. Thus, signatures of lower height can be estimated as the sum of higher-order signatures. For GC models, higher-order (2nd and 3rd order) groups were considered as they can provide a better description on the interaction between 1st order groups and the effects of certain molecular group combinations to the property of a molecule. Despite the higher accuracy of estimation for complex compounds, higher-order GC groups require more details on structural knowledge. Generally, a 2nd order group from GC method can be represented in a molecular signature of height 2 or 3, with examples shown in Table 4-7. Other than GC models, the height of signature also dependent on the TI models. For instance, 1st order connectivity index requires signature of height 2; 2nd order

connectivity index requires signature of height 3, etc. From Table 4-6, the prediction model for octanol/water partition coefficient requires connectivity index of 3rd order. Therefore, the maximum signature height required in this problem was set at 4. However, all possible height 4 signatures need to be generated to solve the CAMD problem using molecular signature descriptor. As the height of signature increases, the possible combination of molecular signature increases as well. In this case study, the number of generated height 4 signatures were expected to exceed 100,000 signatures. Pre-screening was conducted by applying feasibility rules on the generated signatures. As a result, the total number of height 4 signatures were reduced to around 10000 signatures.

Table 4-7 Example of 2nd order group expressed in terms of signature of height 2 or 3.

2nd Order Group	**Molecular Signature**
$(CH_3)_2CH$	C3(C1(C)C1(C)C2(CC))
$CH(CH_3)CH(CH_3)$	C3(C1(C3(CCC)) C1(C3(CCC)) C3(C3(CCC)C1(C)C1(C))
CH_3COOCH	C4(C1(C4(=OOC) =O2(=C4(=OOC) O2(C4(=OOC)C2(CO)))

4.4.3. CAMD Formulations

The atoms that are commonly present in solvents, which includes: hydrogen (H), carbon (C), nitrogen (N) and oxygen (O) was chosen for the design of bio-oil solvent. The hydrocarbon groups considered in this study were limited to alkanes, alkenes, alcohol, carboxylic acid, ketones, aldehyde, esters, ethers, and nitriles which can be predominately found in solvents. The chemical groups are listed in Table 4-8 (Conte et al., 2008).

Table 4-8 Chemical classes considered and their respective chemical groups

Chemical Class	Chemical Group
Alkanes	CH_3, CH_2, CH, C
Alkenes	CH_2=CH, CH=CH, CH_2=C, CH=C, C=C
Alcohol	OH
Carboxylic Acid	COOH
Ketones	CH_3CO, CH_2CO, CHCO, CCO
Aldehyde	CHO
Esters	CH_3COO, CH_2COO, CHCOO, CCOO, HCOO
Ethers	CH_3O, CH_2O, CH-O, C-O
Nitriles	CH_2CN, CHCN, CCN

Initially, the signatures of height 1 were generated based on the selected atoms' type and chemical families, resulting in a total of 65 different molecular signature combinations. By applying feasibility rules mentioned in Section 4.3.3.1, the set of height 1 signatures was then reduced to a total of 24 signatures. As some of the property prediction models were expressed in GC method, signature descriptors were translated and assigned to their corresponding groups from GC method as shown in Table 4-9.

Table 4-9 Height 1 molecular signature set and their corresponding groups.

No.	Signature	Corresponding Group	No.	Signature	Corresponding Group
S1	C1(C)	CH_3	S13	C4(=C=O)	C = C
S2	C1(O)	CH_3O	S14	C4(=CCC)	C = C
S3	C2(=C)	CH_2 = C	S15	C4(=CCO)	C = C
S4	C2(CC)	CH_2	S16	C4(=OCC)	CHCO
S5	C2(CO)	CH_2	S17	C4(=OCO)	CH2COO
S6	C3(=CC)	CH = C	S18	C4(≡NC)	CCN
S7	C3(=CO)	CH = C	S19	C4(CCCC)	C
S8	C3(=OC)	CHO	S20	C4(CCCO)	C
S9	C3(=OO)	HCOO	S21	N3(≡C)	CCN
S10	C3(CCC)	CH	S22	O1(C)	OH
S11	C4(CCO)	CH	S23	O2(=C)	COOH
S12	C4(=C=C)	C = C	S24	O2(CC)	CH-O

CAMD problem was then solved using global solver by LINGO extended version 18.0.56. By solving the CAMD problem, 5 height 1 signatures were identified as promising signature candidates as shown in Table 4-13. Next, height 2 signatures were generated based on these 5 identified signatures of height 1. A total of 147 height 2 signatures were generated if only pre-screening step was conducted. Taking C1(C) from the resulting signature of height 1 candidates as an example, a total of 23 signatures were generated by considering only the feasibility rules as shown in Figure 4-2,. However, not all these 23 signatures were consistent with each other to form a feasible molecule. By applying the consistency rule, only 3 signatures out of the 23 signatures can fulfil the requirement, which include:

1. C1(C3(CCO))
2. C1(C2(CC))
3. C1(C2(CO))

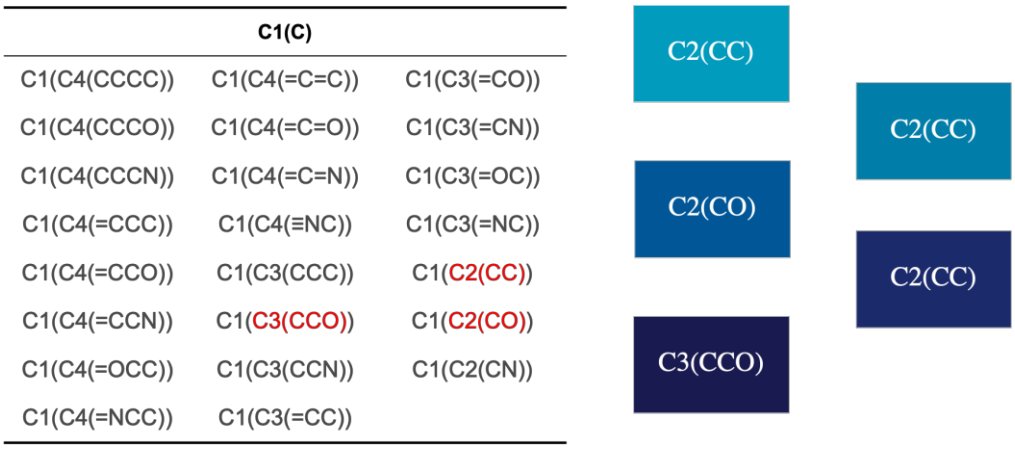


Figure 4-2 Generation of height 2 signature based on the height 1 signature, CI(C)

Same approach was applied to the remaining 4 height 1 signatures as shown in Figure 4-2. As a result, a total of 17 height 2 signatures were generated by applying both feasibility and consistency rules. The generated height 2

signatures together with their corresponding GC group are shown in Table 4-10. The CAMD problem was then solved again for the 17 height 2 signatures set. As a result, 7 signatures from the height 2 set were identified as promising signature candidates. Similar methodology was then applied to generate height 3 and height 4 signatures. List of generated height 3 and height 4 signatures are shown in Table 4-11 and Table 4-12.

Table 4-10 Height 2 signature and their corresponding GC group

No	Signature	Corresponding Group
D1	C1(C3(CCO))	CH_3
D2	C1(C2(CC))	CH_3
D3	C1(C2(CO))	CH_3
D4	C2(C1(C)C2(CC))	CH_2
D5	C2(C1(C)C2(CO))	CH_2
D6	C2(C1(C)C3(CCO))	CH_2
D7	C2(C2(CC)C2(CC))	CH_2
D8	C2(C2(CC)C2(CO))	CH_2
D9	C2(C2(CC)C3(CCO))	CH_2
D10	C2(C2(CO)C3(CCO))	CH_2
D11	C2(C1(CO)O1(C))	CH_2
D12	C2(C2(CC)O1(C))	CH_2
D13	C3(C1(C)C1(C)O1(C))	CH
D14	C3(C1(C)C2(CC)O1(C))	CH
D15	C3(C2(CC)C2(CC)O1(C))	CH
D16	O1(C2(CO))	OH
D17	O1(C3(CCO))	OH

Table 4-11 Height 3 signature and their corresponding GC group

No	Signature	Corresponding Group
T1	C1(C3(C1(C)C2(CC)O1(C)))	CH_3
T2	C1(C2(C1(C)C2(CC)))	CH_3
T3	C2(C1(C2(CC))C2(C1(C)C2(CC)))	CH_2
T4	C2(C1(C2(CC))C2(C2(CC)C2(CC)))	CH_2
T5	C2(C1(C2(CC))C2(C2(CC)C3(CCO)))	CH_2
T6	C2(C2(C1(C)C2(CC))C2(C1(C)C2(CC)))	CH_2
T7	C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))	CH_2
T8	C2(C2(C1(C)C2(CC))C2(C2(CC)C3(CCO)))	CH_2
T9	C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))	CH_2
T10	C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO)))	CH_2
T11	C2(C2(C1(C)C2(CC))C3(C1(C)C2(CC)O1(C)))	CH_2
T12	C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C)))	CH_2
T13	C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO)))	CH
T14	O1(C3(C1(C)C2(CC)O1(C)))	OH

Table 4-12 Height 4 signature and their corresponding GC group

No	Signature
Q1	C1(C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))
Q2	C1(C2(C1(C2(CC))C2(C2(CC)C2(CC))))
Q3	C2(C1(C2(C1(C)C2(CC))C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC))))
Q4	C2(C1(C2(C1(C)C2(CC))C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC))))
Q5	C2(C1(C2(C1(C)C2(CC))C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO))))
Q6	C2(C2(C1(C2(CC))C2(C2(CC)C2(CC)))C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC))))
Q7	C2(C2(C1(C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC))))
Q8	C2(C2(C1(C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO))))
Q9	C2(C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC))))
Q10	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC))))
Q11	C2(C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC))))
Q12	C2(C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO))))
Q13	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO))))
Q14	C2(C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C))))
Q15	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C))))
Q16	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO)))C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C))))
Q17	C2(C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))
Q18	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO)))C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))
Q19	C2(C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C)))C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))
Q20	C3(C1(C3(C1(C)C2(CC)O1(C)))C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C)))O1(C3(C1(C)C2(CC)O1(C))))
Q21	O1(C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))

With this approach, the signature set size was reduced from a set of more than 10,000 height 4 signatures to the final 21 height 4 signatures. Finally, the CAMD problem was solved and promising molecular signatures of height 4 identified are tabulated in Table 4-13.

Table 4-13 Potential height 1, 2, 3 and 4 signatures generated

No	Signature
	Height 1
S1	C1(C)
S4	C2(CC)
S5	C2(CO)
S11	C3(CCO)
S22	O1(C)
	Height 2
D1	C1(C3(CCO))
D2	C1(C2(CC))
D4	C2(C1(C)C2(CC))
D7	C2(C2(CC)C2(CC))
D9	C2(C2(CC)C3(CCO))
D14	C3(C1(C)C2(CC)O1(C))
D17	O1(C3(CCO))
	Height 3
T1	C1(C3(C1(C)C2(CC)O1(C)))
T2	C1(C2(C1(C)C2(CC)))
T4	C2(C1(C2(CC))C2(C2(CC)C2(CC)))
T7	C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))
T9	C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))
T10	C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO)))
T12	C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C)))
T13	C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO)))
T14	O1(C3(C1(C)C2(CC)O1(C)))
	Height 4
Q1	C1(C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))
Q2	C1(C2(C1(C2(CC))C2(C2(CC)C2(CC))))
Q3	C2(C1(C2(C1(C)C2(CC))C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC))))
Q7	C2(C2(C1(C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC))))
Q12	C2(C2(C2(C1(C)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO))))
Q15	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C2(CC)))C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C))))
Q18	C2(C2(C2(C2(CC)C2(CC))C2(C2(CC)C3(CCO)))C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))
Q20	C3(C1(C3(C1(C)C2(CC)O1(C)))C2(C2(C2(CC)C2(CC))C3(C1(C)C2(CC)O1(C)))O1(C3(C1(C)C2(CC)O1(C))))
Q21	O1(C3(C1(C3(CCO))C2(C2(CC)C3(CCO))O1(C3(CCO))))

The molecular structures of promising solvents were generated from the identified height 4 signature building blocks. Database search for the feasible molecules was then carried out. The feasible solvent molecules were identified as 2-octanol, 2-heptanol, 2-hexanol and 2-pentanol, respectively. Higher heating value of the identified solvent candidates were verified through NIST's database as shown in Table 4-14 (NIST, 2018). The higher heating value estimated in present work for the abovementioned solvent candidates were close to the actual HHV obtained from NIST database, with less than 1% differences. According to Equation 4.30, the higher heating value for the final solvent-oil blend was expected to increase as the amount of solvent fraction increases. However, the solvent-oil blend will be mixed with a large portion of diesel, forming a solvent-oil-diesel blend. Thus, effect of the amount of solvent added on the higher heating value of solvent-oil blend will be negligible as compared to the amount of diesel present in the blend.

$$HHV_{mix} = \sum x_i HHV_i \tag{4.30}$$

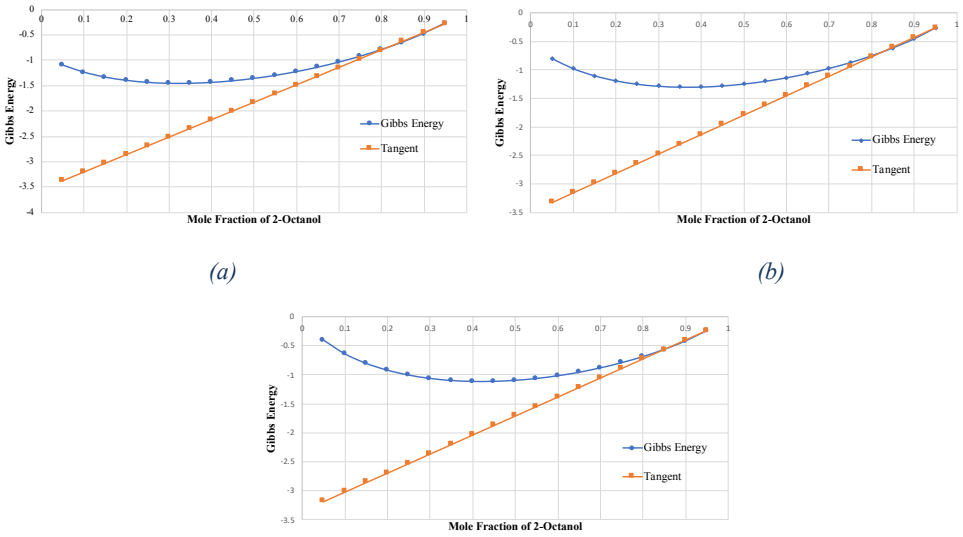
Table 4-14 HHV of solvent candidates from NIST's database and present work

Molecular Name	**Higher heating value from NIST / MJ kg^{-1}** (NIST, 2018)	**Higher heating value from Present work / MJ kg^{-1}**
2-Octanol	40.66	40.89
2-Heptanol	39.72	40.00
2-Hexanol	38.98	38.92
2-Pentanol	37.72	37.50

Other than contributing to the HHV of bio-oil, the solvent candidates also play a major role in improving the miscibility of the final blend. With the absence of solvent, strong intermolecular forces of the bio-oil will attract the

molecules instead of dispersing in aqueous phase and petroleum fraction (Manara et al., 2018). However, the amphiphilic properties of the identified solvent candidates are capable to help in dispersion of the bio-oil. Phase stability was conducted by computing the tangent plane distance for the 4 identified solvent molecules. Sensitivity analysis on the phase behaviour of solvent-oil blend was also conducted for different water content. The average water mass fraction for crude pyrolysis bio-oil was reported to be around 38 – 42% (Asadullah et al., 2013). Thus, crude pyrolysis bio-oil containing 40 wt.% of water was considered as the maximum water content. To investigate the effect of water content on the final blend's miscibility, bio-oil with reduced water content will also be considered in the analysis. The water content in bio-oil can be reduced to 16 wt.% by eliminating the aqueous phase. In addition, water content of 25 wt.% was taken as the median value and considered in the sensitivity analysis. Figure 4-3 shows the Gibbs energy and tangent plot for 2-octanol-oil blend at 16% (Figure 4-3a), 25% (Figure 4-3b) and 40% water content (Figure 4-3c), respectively. The optimal mole fraction obtained for 2-octanol was 0.805, 0.83 and 0.85 at bio-oil's water content of 16%, 25% and 40%, respectively. The amount of solvent required in the solvent-oil blend increases slightly as the water content in bio-oil increases. Similar trends were obtained for 2-heptanol-, 2-hexanol- and 2-pentanol-oil blends where the Gibbs energy and tangent plots for these solvent-oil blends are shown in Figure 4-4, Figure 4-5, and Figure 4-6, respectively. From Figure 4-3(a), (b) and (c), the blend of 2-octanol and bio-oil is stable and exhibit homogenous single-phase as the tangent line was plotted below the Gibb's curve. This could be explained by

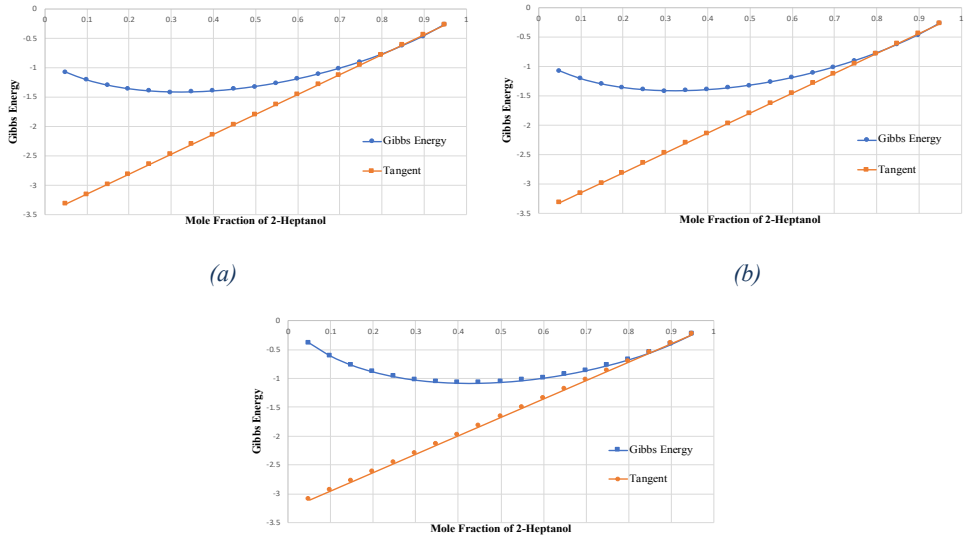
the presence of alcohol (OH) group in the solvent's molecular structure which aids in promoting miscibility of the blend.



(a) *(b)*

(c)

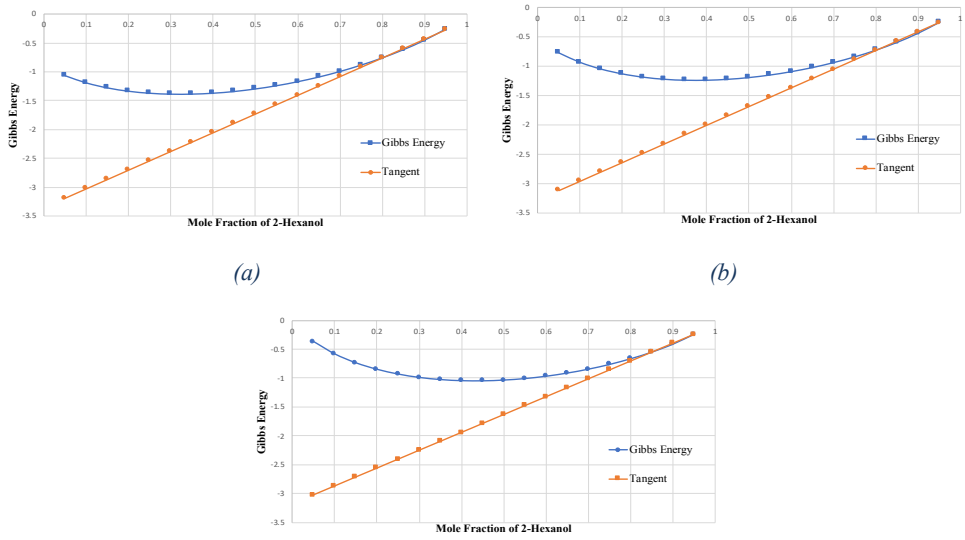
Figure 4-3 Gibb's energy and tangent plot for 2-octanol and bio-oil at (a) 16% water content (b) 25% water content and (c) 40% water content



(a) *(b)*

(c)

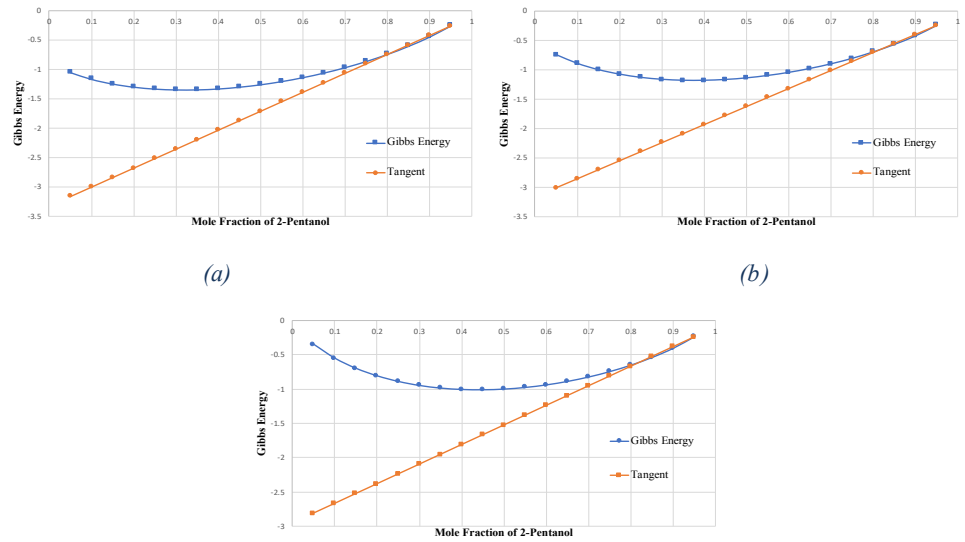
Figure 4-4 Gibb's energy and tangent plot for 2-heptanol and bio-oil at (a) 16% water content (b) 25% water content and (c) 40% water content



(a) *(b)*

(c)

Figure 4-5 Gibb's energy and tangent plot for 2-hexanol and bio-oil at (a) 16% water content (b) 25% water content and (c) 40% water content



(a) *(b)*

(c)

Figure 4-6 Gibb's energy and tangent plot for 2-pentanol and bio-oil at (a) 16% water content (b) 25% water content and (c) 40% water content

Besides, a ternary phase diagram was plotted for the mixtures of bio-oil, water, and 2-octanol (solvent) to evaluate the miscibility of final blend at various mixing compositions as shown in Figure 4-7(a). In the phase diagram, the red dots represent the immiscible blend while the green dots represent the miscible blend. The solvent-oil-water blend was miscible over most of the composition range. However, the blend was immiscible at the mixing ratio, 2-octanol: bio-oil: water of 0 : 10 : 90 and 10: 0 : 90. Similar results were obtained for 2-heptanol-, 2-hexanol- and 2-pentanol-oil-water blends, as shown in Figure 4-7(b), (c) and (d), respectively.

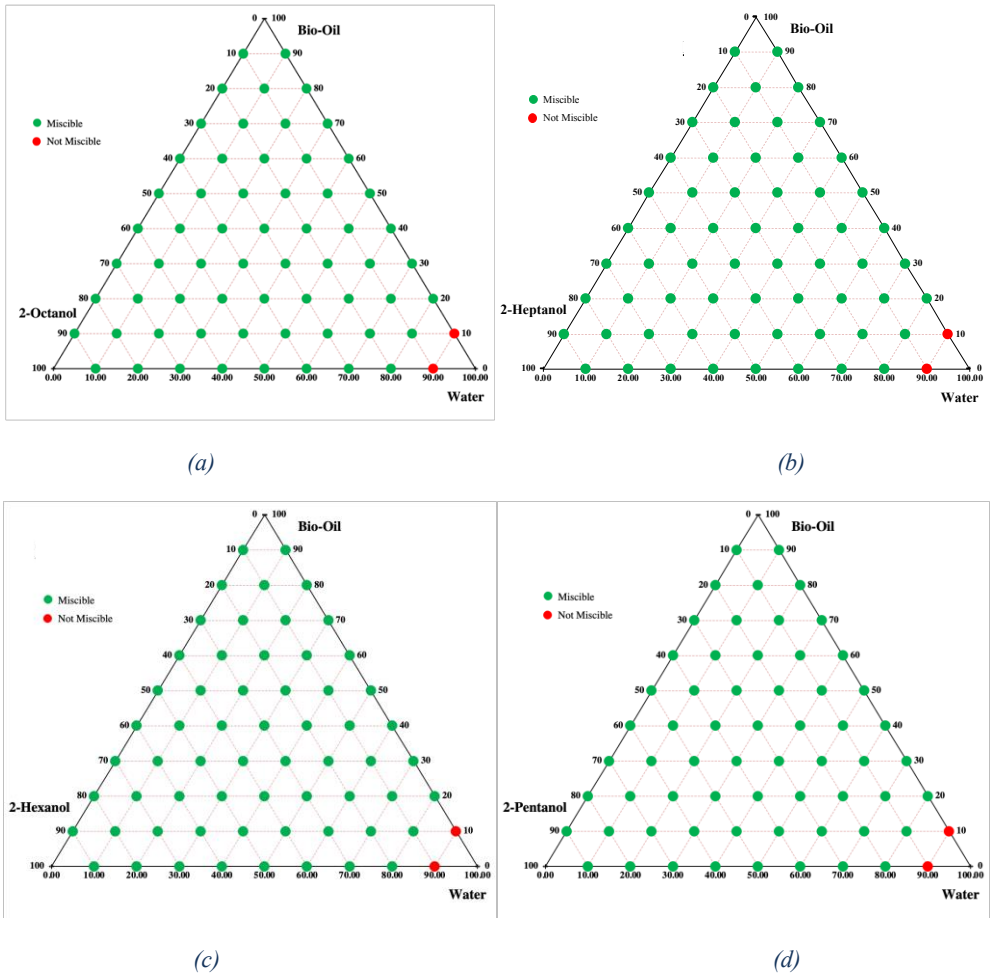
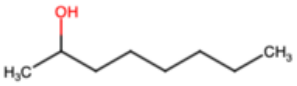
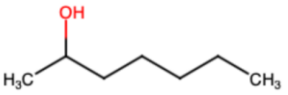
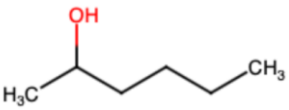
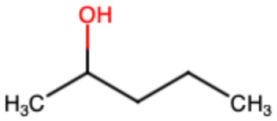


Figure 4-7 Gibb's phase ternary graph of bio-oil, water, and (a)2-octanol, (b) 2-heptanol, (c) 2-hexanol and (d) 2-pentanol

Table 4-15 summarises the key properties and information of the identified candidate solvents. All the resulting molecules possess a higher heating value of at least 37.5 MJ/kg. The solvent-oil blends were expected to be homogenous as the tangent plane distance calculated is non-negative for all solvent candidates. It can be concluded that 2-octanol is the most suitable solvent candidate with the highest higher heating value at 40.89 MJ/kg.

Table 4-15 The identified feasible solvent candidates

Molecular Name	Formula	Molecular Structure	Higher heating value / MJ kg^{-1}	Miscibility
2-Octanol	$CH_3(CH_2)_5CH(OH)CH_3$		40.89	Miscible
2-Heptanol	$CH_3(CH_2)_4CH(OH)CH_3$		40.00	Miscible
2-Hexanol	$CH_3(CH_2)_3CH(OH)CH_3$		38.92	Miscible
2-Pentanol	$CH_3(CH_2)_2CH(OH)CH_3$		37.50	Miscible

4.5. Summary

This chapter introduces a multi-stage CAMD approach to design a solvent for pyrolysis bio-oil upgradation. At the initial stage, solvent requirements were determined and translated into target properties. Suitable GC and TI based property prediction models were selected to estimate the identified targeted physicochemical and environmental properties. Different property

prediction models possess different structures, thus, requiring different topological information. In this chapter, molecular signature descriptors were applied in the CAMD problem to accommodate different indexes on a common platform. Signature building blocks were formulated up to height 4 to cover higher order GC groups. Next, relevant structural constraints were incorporated in the CAMD model to ensure the feasibility of the designed molecules. In addition, the consistency rule developed was applied to reduce the size of CAMD problem by ensuring only relevant and consistent signatures are generated. In the next stage, stability analysis was conducted on the identified solvent candidates to ensure the miscibility and stability of final solvent-oil blend. Based on the case study conducted, 2-pentanol, 2-hexanol, 2-heptanol and 2-octanol were identified as the promising solvents candidates. Among the identified solvents, 2-octanol was selected as the most promising solvent candidate with a HHV of 40.89 MJ/kg along with other desirable attributes. To conclude, the developed methodology in this chapter can be applied in the design of solvents for any application. Further improvements can be made by incorporating the economic consideration into solvent design to ensure the competitiveness of the designed solvent. This design consideration will be further discussed in Chapter 5.

CHAPTER 5

COMPUTER-AIDED FRAMEWORK FOR THE DESIGN OF OPTIMAL BIO-OIL/SOLVENT BLEND WITH ECONOMIC CONSIDERATION

5.1. Introduction

In this chapter, the second research scope is presented as an extension of the first research scope. Other than the fuel functionality, it is also crucial for the designed solvent-oil blend to achieve economic targets in order to be competitive with the conventional diesel fuel. Hence, the economic aspect such as the product pricing, profitability, market share, and operating cost was incorporated into the design problem. The main objective of this chapter is to generate feasible solvent candidates by solvent the multi-objective optimisation (MOO) problem via computer-aided molecular design (CAMD) approach. Initially, an optimisation model was developed to identify potential solvents that satisfied the predefined targeted properties. Next, a MOO model was developed via fuzzy optimisation approach to identify the trade-off between the profitability and heating value of the solvent-oil blend. In this work, a pricing model was employed to estimate the profitability of the solvent-oil blend. The production of bio-oil in a pyrolysis plant was used to illustrate the applicability of the pricing model. Lastly, phase stability analysis was conducted to ensure the stability and miscibility of the solvent-oil blend at their optimal mixing ratio.

With the developed framework, a promising and cost-effective solvent-oil blend can be generated while displaying promising biofuel properties.

5.2. Problem Statement

A major obstacle in the commercialisation of biofuels is their high production cost as compared to that of the conventional diesel fuel. Their low heating value and the high cost of raw materials make biofuels more costly for heat generation. As the biofuel ratio increases in the fuel, the fuel's energy density decrease. Furthermore, the addition of solvents is often required to improve the biofuel's properties. However, the solvents are generally associated with high cost as most of them are manufactured as specialty chemicals. This further increases the production costs of biofuels. In addition, the existing and progressing biofuel legislative framework sparks the urge to assess the cost associated with upgrading bio-oil. Notably, previous research on the design of bio-oil solvent mainly focused on the functionality of the solvent itself. It is also important to incorporate the economic aspects into the development of bio-oil solvent for the designed solvent-oil blend to be competitive with the conventional diesel fuel. Generally, the higher heating value (HHV) of pyrolysis bio-oil increases with the addition of solvent. However, as the amount of solvent increases, the cost of the solvent-oil blend increases as well, resulting in lower profitability. Thus, a MOO approach was adapted in this chapter to optimise both objectives, which are the heating value and profitability of the solvent-oil blend, simultaneously. The weighted sum method is a more common approach for handling MOO problem, where each objective function was allocated with a weighting factor to convert different objectives into an aggregated scalar

objective function (Fishburn, 1967). Nevertheless, the weighting factor of each objective is not always definable. In most of the cases, the relative importance of each objective is either fuzzy or uncertain. In addition, these objectives might be contradictory to each other in nature. Thus, fuzzy optimisation via max-min aggregation approach was employed in this chapter to solve the MOO problem. The fuzzy optimisation approach is capable to solve decision-making problems under the fuzzy environment by defining and quantifying the uncertainties and vagueness. Within the fuzzy optimisation algorithms, the trade-off between the objective functions to be optimised can be identified by introducing the fuzzy membership function. As a result, an optimal solution can be identified by achieving near optimality for all the objectives without compromising on each other.

5.3. Methodology

In this work, a MOO framework was established to solve the CAMD problem using fuzzy optimisation approach. The developed framework can be divided into four main stages and their correlated sub-steps. The detailed methodology for each stage will be further discussed in this section. Figure 4-1 demonstrates the overview of methodology for this solvent design problem.

Error! Reference source not found.

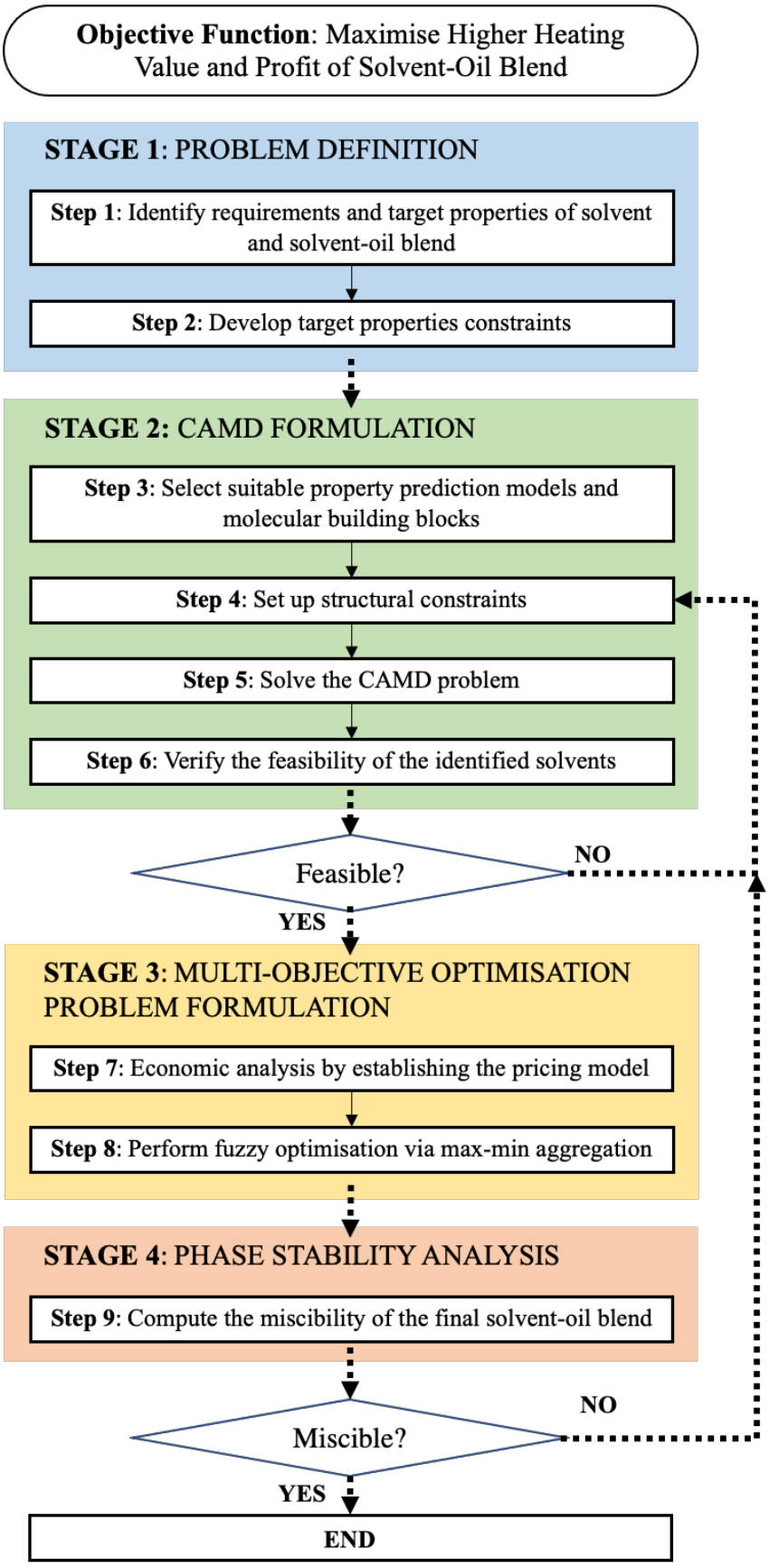


Figure 5-1 Graphical summary of work carried out in this chapter

5.3.1. Problem Definition

Firstly, the problem definition was formulated, where the product needs of solvent and final solvent-oil blend were defined based on the requirements form the currently available regulations, specifications of biodiesel. In addition, the general safety and environmental regulatory requirements were also considered. The identified product requirements were then translated into measurable quantitative target properties. In this study, the crude bio-oil derived via fast pyrolysis of palm kernel shell (PKS) was used as the basis. However, only the organic phase of pyrolysis bio-oil was considered. The properties and components of the pyrolysis bio-oil applied in this study was summarised and listed in Table 5-1.

Table 5-1 Pyrolysis bio-oil's properties

Property	**Values**
Moisture content	16 wt.%
Kinematic viscosity	17.4 mm^2/s
HHV	19.0 MJ/kg
Density	1150 kg/m^3
Components	**Mole Fraction**
Phenol	0.62
2,6 – dimethoxyphenol	0.11
2 – methoxyphenol	0.11
Furfural	0.08
1.2 – benzenediol	0.08
Acetic acid	0.07

5.3.2. Property Prediction Models

In the following steps, suitable property prediction models were selected to estimate the solvent's target properties. Here, property prediction models in

terms of GC and TI approaches were employed and expressed as a function of the molecular signature descriptor. The upper and lower limits of property constraints were then defined based on the process requirements or customers' need.

5.3.3. CAMD Formulation

The generalised mathematical expression (Equation 4.1 to 4.8) for CAMD optimisation model presented in Section 4.3.3 was implemented in this chapter. To generate structurally feasible solvent compounds, structural constrains presented in Section 4.3.3.1 to Section 4.3.3.3 were applied. Next, the generated molecules were verified through database search to ensure its feasibility and practicability.

5.3.4. Multi-Objective Optimisation Problem Formulation

5.3.4.1. Formulation of Pricing Model

After the potential solvent candidates were identified in the previous stage, a detailed economic analysis was conducted to determine the selling price of the new solvent-oil blend based on the current market demand, availability and existing competitors identified via a comprehensive background study. The pricing model proposed by Bagajewicz (2007) was employed to relate the product quality to demand and price of the product (Fung et al., 2016). The mathematical expression for the pricing model has been shown in Equations 5.1 and 5.2 (Bagajewicz, 2007).

$$A^P T^P = A^C (T^P)^\delta \left(\frac{\alpha}{\beta}\right)^\delta \left(\frac{Y - A^P T^P}{A^C}\right)^{1-\delta} \tag{5.1}$$

$$Y \geq A^P T^P + A^c T^c \tag{5.2}$$

Here, A^P and T^P referred to the price and demand of the new solvent-oil blend while A^c and T^c referred to the price and demand of the competitor's product. In this study, the value for A^P can be obtained by summing up the cost of pyrolysis bio-oil production, cost of solvent and the profit obtained by selling the solvent-oil blend (Equation 5.3).

$$A^p = Cost_{Bio-oil} + Cost_{Solvent} + Profit \tag{5.3}$$

Based on Equation (5.1), Y is the total market size for the solvent-oil blend and δ is the elasticity of substitution, which is an adjustable parameter that measures the change in the ratio of products demand in response to a change in the ratio of their prices. On the other hand, the parameter α can be expressed as the consumer's awareness on the new product, which can be raised by allocating higher budget in the marketing of new product. The value of parameter α ranges between 0 and 1, where α with the value 0 indicates that the consumers have no knowledge about the new product, and vice versa. Lastly, β is the consumer preference coefficient that relates the consumer's interest in the new product over the competing product, which can be determined using Equation 5.4 (Bagajewicz, 2007).

$$\beta = \frac{\lambda^C}{\lambda^P} \tag{5.4}$$

In Equation 5.4, the parameters λ^C and λ^P represented the consumer's preference function of competitor's product and new product, respectively. In this study, the consumer's preference was related to the HHV of the solvent-oil blend, which possess a significant influence on the functionality of the solvent-

oil blend. The new solvent-oil blend is said to be preferred by consumers if β is smaller than 1. However, the competitor's product is preferred when the value of β is greater than 1.

Based on the market analysis conducted, the total market size, Y of solvent-oil blend was identified to be US$ 500 million annually (Energy, 2020). Based on the previous studies for diesel fuel, the price elasticity was defined between the range of 0.11 to 0.33. The demand of solvent-oil blend was said to be price inelastic when the parameter δ lies between 0.1 to 1. Thus, in this work, the parameter δ was assumed to be 0.1 (Dahl, 2012). On the other hand, the parameter α was estimated to be at the value 0.85. These values should be revised and updated based on the response received once the solvent-oil blend was introduced into the market. The benchmark for this work is the reported solvent-oil blend consisting of 50 wt.% pyrolysis bio-oil and 50 wt.% of iso-propanol, with a HHV of 27.55 MJ/kg (Omar et al., 2019). The cost of iso-propanol and pyrolysis bio-oil was assumed to be US$ 1336.57 per ton of iso-propanol and US$ 354 per ton of pyrolysis bio-oil (China Petroleum and Chemical Industry Federation, 2021; EUBIA, 2021). The cost of the competitor's solvent-oil blend can be calculated using Equation (5.5), where x_i and C_i is the ratio and costs of the solvent and bio-oil, respectively

$$C_{blend} = \Sigma_i x_i C_i \tag{5.5}$$

Next, the selling price of competitor's blend, A^c can be calculated by summing up the cost of solvent-oil blend and the profit obtained, which was assumed to be US$ 50 per ton solvent-oil blend, in this case. Table 5-2

summarises the parameters and its respective values obtained from this market analysis.

Table 5-2 Parameters and values from market analysis.

Parameters	Values
Total market size, Y	US$ 500,000,000 per year
Consumer's awareness coefficient, α	0.85
Elasticity of substitution, δ	0.10
Price of competitor's product, A^C	US$ 895.29 per ton solvent-oil blend
HHV of competitor's product, λ^C	27.55 MJ/kg

5.3.4.2. Formulation of Fuzzy Multi-Objective Optimisation

The HHV of pyrolysis bio-oil can be increased with the addition of solvent. The higher the mass fraction of solvent in the solvent-oil blend, the higher is the energy content. However, a higher amount of solvent was often associated with higher cost, and thus, lower profitability obtained from the solvent-oil blend. In this study, a MOO problem was developed to investigate the trade-off between high HHV and high profitability. Most of the current CAMD techniques focus on optimising a single objective or property of the chemical product, but having a multi-objective problem necessitates the use of more complex optimisation methods (Khor et al., 2017).

Thus, fuzzy mathematical programming was applied to solve the MOO design problem. The satisfaction degree of fuzzy, λ was introduced to both property functions that were to be optimised. The degree λ is a continuous variable that lies between the value 0 and 1, where 0 indicates unsatisfactory and 1 is complete satisfactory (Equation 5.6).

$$0 \leq \lambda \leq 1 \tag{5.6}$$

The objective function of the fuzzy optimisation model was to maximise the overall satisfaction degree of fuzzy constraint λ as shown in Equation 5.7 . The max-min aggregation was applied to the fuzzy programming, where every fuzzy constraint should be satisfied partially at least to the degree λ .

$$f_{obj} = \max \lambda \tag{5.7}$$

Fuzzy goals for the HHV and profitability of the solvent-oil blend were expressed using a linear membership function, as shown in Equations 5.8 and 5.9 (Khor et al., 2017).

$$\lambda_{p(\max)} = \begin{cases} 0, & V_p \leq v_p^L \\ \dfrac{V_p - v_p^L}{v_p^L - v_p^L}, & v_p^L \leq V_p \leq v_p^U \quad \forall p \in P \\ 1, & V_p \geq v_p^U \end{cases} \tag{5.8}$$

$$\lambda_p \geq \lambda \tag{5.9}$$

Where V_p is the target property values bounded by the lower and upper bounds, v_p^L and v_p^U, respectively. The values for the lower and upper bound can be obtained by performing single objective optimisation for both objective functions.

5.3.5. Phase Stability Analysis

With the identification of optimal solvent-oil blend ratio from the previous step, phase stability analysis was carried out to ensure the miscibility of the designed solvent-oil blend at the targeted mixing ratio. In this chapter, the phase stability analysis via computation of the Gibbs tangent plane distance, as presented in Section 4.3.5 was conducted. The solvent-oil blend can be

concluded as stable if the tangent plane distance is non-negative. If otherwise, the previous steps are repeated by revising the property attributes and constraints.

5.4. Case Study

A case study on the solvent design for pyrolysis bio-oil application was conducted to illustrate the application of this proposed methodology. The fast pyrolysis process considered in this chapter is related to an application in Malaysia. All pricing in this study was converted to U.S. Dollar at the exchange rate of RM 1 = US$. 0.24, and was adjusted to 2021 values using appropriate indices.

5.4.1. Identification of Feasible Solvent Candidates

The promising solvent candidates identified from Chapter 4 were used to demonstrate the proposed methodology. In this case study, 2-octanol, 2-heptanol, 2-hexanol and 2-pentanol were considered. The HHV value estimated in Chapter 4 for the abovementioned candidates were close to the actual HHV obtained from the NIST database, with less than 1% differences (NIST, 2018). The solvents were of analytical grade and the costs were obtained from chemical vendors as shown in Table 5-3.

Table 5-3 Cost and HHV for solvent candidates generated from CAMD optimization model

Compound Name	Chemical Formula	HHV(MJ/kg)	Cost (US$/g solvent)
2-Octanol	$C_8H_{18}O$	40.89	0.0026
2-Heptanol	$C_7H_{16}O$	40.00	0.0030
2-Hexanol	$C_6H_{14}O$	38.92	0.3958
2-Pentanol	$C_5H_{12}O$	37.50	0.1010

5.4.2. Multi-Objective Optimisation Model

Here, a MOO model was developed via fuzzy max-min aggregation approach to optimise the HHV and profitability of the solvent-oil blend, simultaneously. Two different case studies were presented to investigate the effect of different constrains on the outcome while optimising both objective functions.

5.4.2.1. Estimation of Pyrolysis Bio-Oil Production Cost

In this chapter, a pyrolysis plant was proposed to aid the estimation of pyrolysis bio-oil production cost. The pyrolysis plant was designed to produce 120 tonne of pyrolysis bio-oil from 200 dry tonne of palm kernel shell biomass feedstock daily via fast pyrolysis. The typical value for pyrolysis bio-oil yield, which is 60 % was considered in this study (Barik, 2019). It is expected for the pyrolysis plant to operate on a continuous operation daily for 24 hours and 300 days, with a plant lifetime of 30 years. The production costs of the pyrolysis plant include the raw biomass cost, capital cost, labour cost, electrical cost, and other operational cost. Assumption was made that the palm kernel shell biomass used in the pyrolysis plant were supplied by a palm oil mill at no cost.

On the other hand, the capital cost of the pyrolysis plant was estimated based on the sizing curve developed by Rogers and Brammer (2012), which relates both the total plant cost and the plant capacity. In this case, the total plant cost allocated for the designed pyrolysis plant was estimated to be US$ 16 million. In addition, the capital cost for the biomass pre-processing plant was included, with an estimated cost of US$ 2.98 million. As for the labour cost

estimation, the following scenario was assumed where the designed pyrolysis plant operates on a shift-work basis, with 5 operators and 1 supervisor per shift. Three 8-hour shifts pattern were implemented with 4 teams to provide 24/7 coverage. An average annual salary of US$ 13 K was allocated for each employee, which covers the employers' insurance cost, pension contribution, anti-social hours payments, training, and administration charges (Salary Explorer, 2021).

A total electrical consumption of 240 kWh per dry tonne of biomass feedstock was estimated for both the biomass pre-processing plant and the pyrolysis plant (Rogers and Brammer, 2012). The electric tariff of E1 for general industry with medium voltage as defined by Malaysia's energy provider (Tenaga Nasional Berhad) was considered in this study. The price of tariff E1, US$ 0.08 kWh was used to calculate the total cost of electricity (TNB, 2021). Lastly, an allowance of 4 % of the total plant cost (US$ 771.48 K per year) has been made to cover other miscellaneous cost such as repair, maintenance, insurance, and business costs (Rogers and Brammer, 2012). Thus, the total cost to produce 1 tonne of pyrolysis bio-oil were calculated to be US$ 80.37, as shown in Table 5-4.

Table 5-4 Summarised pyrolysis bio-oil production cost.

Production Cost	Cost (US$/ton of bio-oil)
Biomass Cost	N/A
Capital Cost	17.86
Labour Cost	8.67
Electrical Cost	32.35
Other Operating Cost	21.43

Total Production Cost	**80.37**

5.4.2.2. Fuzzy optimisation

In case study 1, the constraint on solvent fraction added to the blend was relaxed to allow higher HHV value of the generated solvent-oil blend. Here, the consumer preference coefficient, β was set to be lesser than 0.75 to achieve HHV of at least 35 MJ/kg. In case study 2, the constraint on parameter β was relaxed, thus lowering the HHV requirement of solvent-oil blend to allow higher profitability. As mentioned above, the competitor's product consisted of 50 wt.% solvent. Hence, solvent fraction was set to be lesser than 0.5. Table 5-5 summarised the constraints defined in case study 1 and 2, respectively.

Table 5-5 Comparison of constraints for case study 1 and 2.

Case	Consumer preference coefficient, β	Solvent Fraction
Case 1	< 0.75	< 0.99
Case 2	< 0.99	< 0.50

Single objective optimisation was conducted to identify the upper and lower bounds for both the objective functions. The results from single objective optimisation can be found in Table 5-6.

Table 5-6 Results from single objective optimisation of solvent-oil blend.

Case	Objective Function	Max HHV_{blend}	Max $Profitability_{blend}$
Case 1	HHV_{blend} (MJ/kg)	40.67	36.73
	$Profitability_{blend}$ (US$/ton of blend)	3303.66	4091.74
Case 2	HHV_{blend} (MJ/kg)	29.95	27.83

$Profitability_{blend}$ (US$/ton of blend)	4626.73	5033.32

Among the 4 solvents identified in Chapter 4, only 2 solvents, which are the 2-octanol and 2-heptanol demonstrated promising performance in terms of functionality, environment, and economics. Table 5-7 shows the results obtained from case study 1 and 2. As the constraints on the HHV of the solvent-oil blend were relaxed in case study 1, higher HHV can be observed, ranging from 37.94 to 39.30 MJ/kg. However, a large amount of solvent was required to be blended with pyrolysis bio-oil, thus leading to the increased cost and low profitability. From Table 5-7, higher profit was obtained for case 2. Nonetheless, this was compensated with the lower HHV of solvent-oil blend ranging from 28.49 to 29.20 MJ/kg.

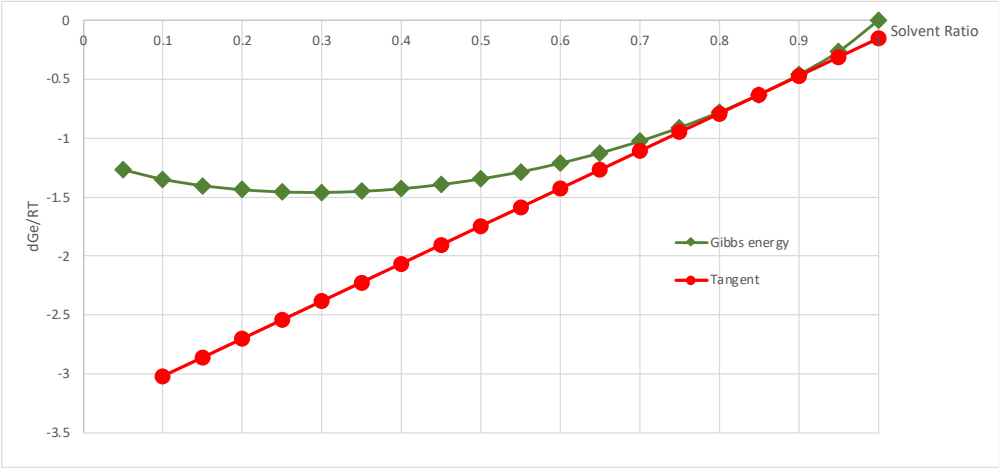
Table 5-7 Results for solvent blend candidates.

Case Study	Solvent	HHV_{blend} (MJ/kg)	Profit (US$/tonne of blend)	Solvent Ratio	Miscible
Case 1	2-Octanol	39.30	3816.42	0.93	Yes
	2-Heptanol	37.94	3544.29	0.90	Yes
Case 2	2-Octanol	29.20	4890.19	0.47	Yes
	2-Heptanol	28.49	4754.33	0.45	Yes

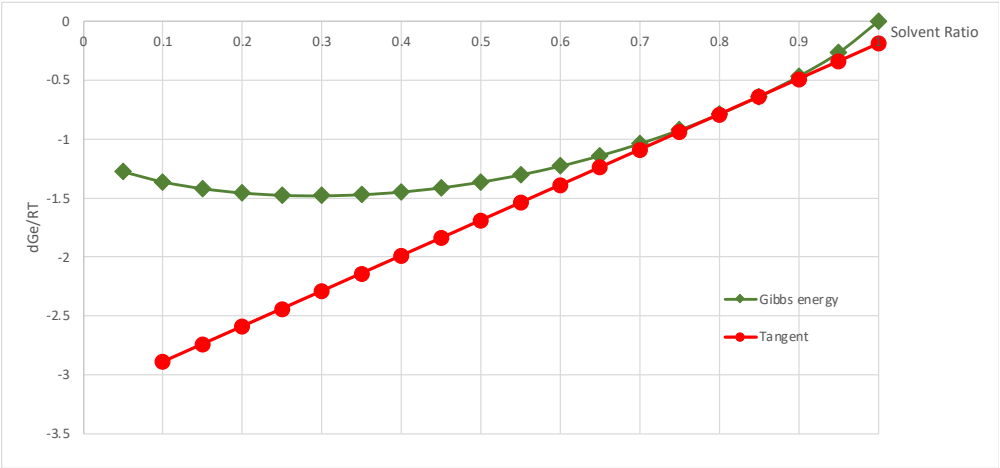
5.4.3. Phase Stability Analysis

Apart from contributing to the HHV of the solvent-oil blend, the solvent candidates also play a major role in improving the miscibility of the final blend. Instead of dispersing in aqueous and organic phase, the strong intermolecular forces between the molecules in the crude pyrolysis bio-oil will attract each

other (Manara et al., 2018). However, the dispersion of pyrolysis bio-oil can be improved with addition of solvent candidates due to its amphiphilic properties, and thus improving the phase separation of pyrolysis bio-oil. In this work, the phase stability analysis was carried out by computing the tangent plant distance against the identified solvent candidates. All the solvent candidates identified in both case 1 and 2 were miscible with pyrolysis bio-oil at their respective mixing ratio. Figure 5-2 (a) and (b) demonstrated the Gibbs energy and tangent plot for 2-octanol-oil blend and 2-heptanol-oil blend obtained from case 1, respectively. On the other hand, Figure 5-3 (a) and (b) demonstrated the Gibbs energy and tangent plot for 2-octanol-oil blend and 2-heptanol-oil blend obtained from case 2, respectively. From Figure 5-2 (a), the blend is said to be stable and demonstrated homogenous single-phase as the tangent line was plotted below the Gibbs energy curve. This could be explained by the presence of alcohol (OH) group in the solvent's molecular structure which aids in promoting miscibility of the blend. As the final solvent0oil blend was expected to be homogenous while demonstrating promising fuel properties, it can be concluded that the solvent-oil blend with 93 wt.% of 2-octanol solvent was chosen as the most promising blend with HHV of 39.30 MJ/kg and profit of US$ 3816.42 per tonne of solvent-oil blend.

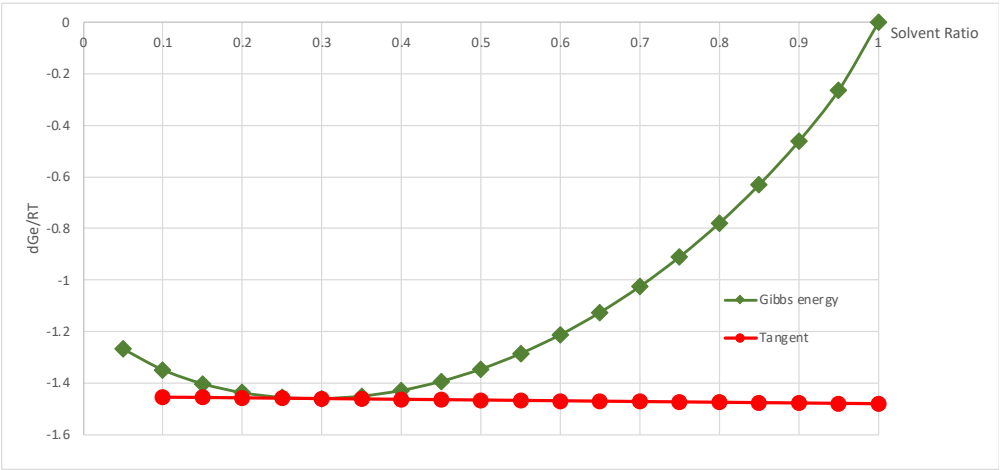


(a)

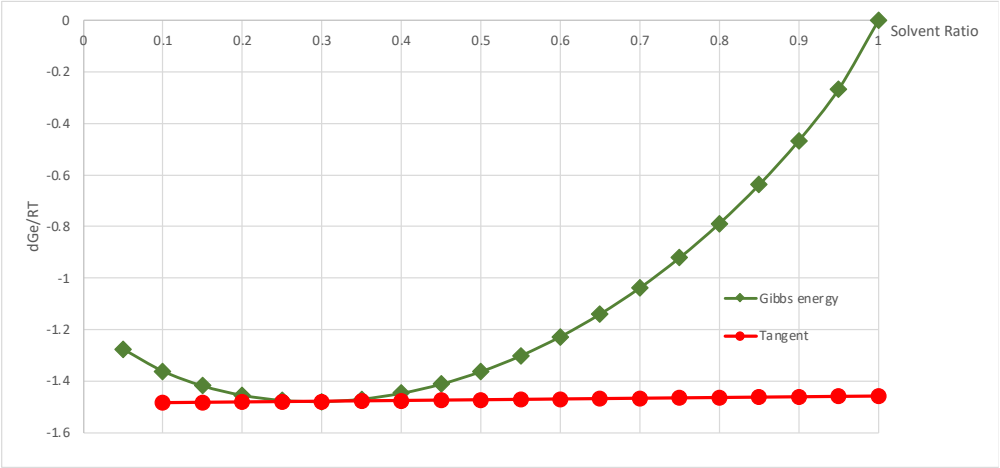


(b)

Figure 5-2 Gibbs energy and tangent plot for (a) 2-Octanol and (b) 2-Heptanol of case 1



(a)



(b)

Figure 5-3 Gibbs energy and tangent plot for (a) 2-Octanol and (b) 2-Heptanol of case 2

5.4.4. Economic Study on the Solvent-Bio-oil-Diesel Blend

By referring to the optimal solvent-oil blend identified from the previous section, an economic analysis was conducted to investigate the relationship between the ratio of solvent-bio-oil-diesel blend, the price and the HHV of the blend. The blend of pyrolysis bio-oil and petroleum diesel was commonly referred to as BX, where X refers to the volume percent of bio-oil in the blend. For example, B5, B10 and B100 consist of 5 %, 10 % and 100 % bio-oil, respectively. Generally, the bio-oil-diesel blend has a lower energy content and higher fuel consumption as compared to that of the conventional diesel. Despite the environmental benefits provided via the application of bio-oil-diesel blend, the price of this blend is costlier than the conventional diesel fuel. As of September 2021, the average price of diesel around the world is US$ 1.07 per litre (Global Petrol Prices, 2021). Meanwhile, the energy content of the conventional diesel fuel generally ranged between 44 to 48 MJ/kg (World Nuclear Association, 2021). From the results obtained, the price of 2-octanol-oil blend was observed to cost US$ 6269.56 per tonne of solvent-oil blend, with

HHV of 39.30 MJ/kg. Table 5-8 summarised the price and HHV of the solvent-oil blend and diesel fuel used in this study.

Table 5-8 Price and HHV for both solvent-oil blend and diesel fuel

Properties	Price (US$/tonne)	HHV (MJ/kg)
Solvent-oil blend	6269.56	39.30
Diesel fuel	1258.82	46.00

Figure 5-4 illustrated the effects of solvent-oil blend ratio on the price and HHV of the bio-oil-diesel blend. It was observed that as the ratio of solvent-oil blend increases, the price of the bio-oil-diesel blend increases proportionately. However, the HHV of the bio-oil-diesel blend decreases as the amount of solvent-oil blend increases. In this study, biodiesel with HHV of 40 MJ/kg was used as benchmark to determine the desired ratio of solvent-oil-diesel blend. As shown in Figure 5-4, blending with at least 20 wt.% of diesel fuels, or 80 wt.% of solvent-oil blend was required to generate biodiesel with HHV of at least 40 MJ/kg. However, blending with 80 wt.% solvent-oil blend will cost approximately US$ 5.2 K per tonne bio-oil-diesel, which is equivalent to 4.2-fold increase as compared to pure diesel fuel. To be competitive with the conventional diesel fuel, substantial subsidies and tax incentives from government are crucial. In addition, the demand for bio-oil-diesel could be stimulated with the introduction of legislation mandating the blending of biofuel in conventional diesel fuel, thus making the bio-oil-diesel demand to be independent of the diesel fuel price (Chin, 2011).

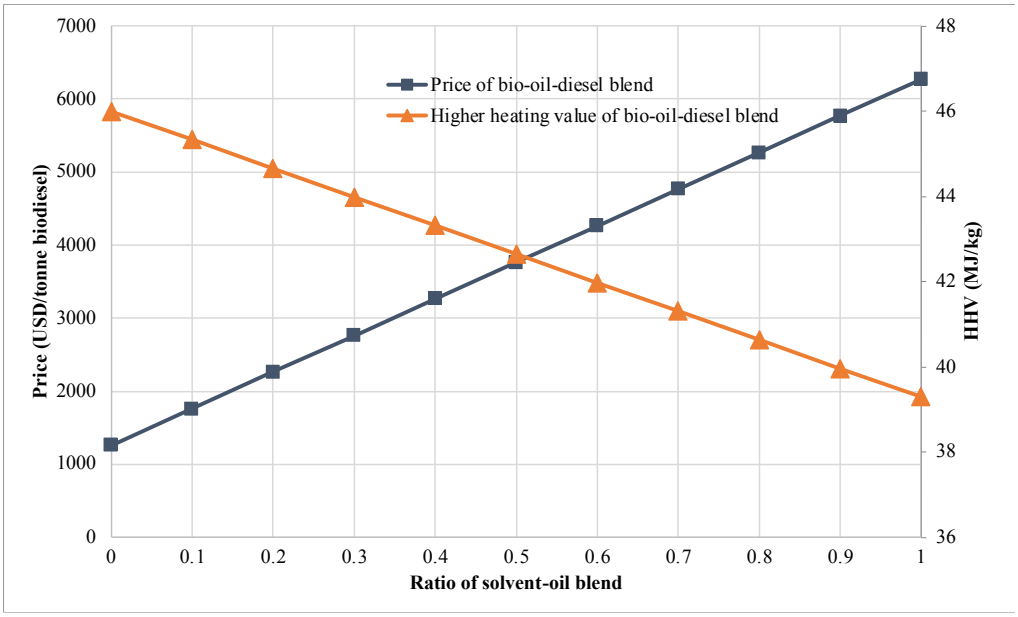


Figure 5-4 Graph of the ratio of solvent-oil blend against the blend's price and HHV.

5.5. Summary

In this chapter, a CAMD framework was developed to design solvent molecules that can upgrade the properties of pyrolysis bio-oil upon blending, while achieving low mixing ratio and maintaining a promising profitability. At the initial stage, the requirements of the solvent and solvent-oil blend were identified and translated into target properties. Suitable property prediction models were selected to estimate the targeted fuel and environmental properties. Here, molecular signature descriptors were applied as a common platform to accommodate all the indexes used in the CAMD problem. Consistency rules were applied to ensure only relevant and consistent signatures are generated and thus, reducing the size and complexity of the CAMD problem. In the second stage, a MOO model was developed via fuzzy max-min aggregation approach to study the trade-off between high HHV and high profitability of the solvent-oil blend. In addition, a pricing model was introduced to evaluate the

profitability of the solvent-oil blend. Also, a pyrolysis plant was proposed in this chapter to aid the estimation of pyrolysis bio-oil production cost. Solvent-oil blend with 2-octanol and 2-heptanol demonstrated positive performance in terms of functionality and economical. Among the identified solvent-oil blend, blends with 93 wt.% of 2-octanol was selected as the best performing solvent-oil blend with HHV of 39.30 MJ/kg and profit of US$ 3816.42 per tonne of blend. To conclude, the proposed framework can be applied in the design of bio-oil solvents with different bio-oil type and composition. However, it is also crucial to obtain financial and legislative support from the government for the commercialisation of bio-oil-diesel blend. Other than the solvent addition, the source and characteristic of biomass feedstock also contributed to the HHV of the pyrolysis bio-oil produced. Thus, the consideration of biomass characteristics and the pyrolysis operating condition in upgrading pyrolysis bio-oil's HHV will be presented in Chapter 6.

CHAPTER 6

ESTIMATION OF FAST PYROLYSIS BIO-OIL PROPERTIES FROM FEEDSTOCK CHARACTERISTICS USING ROUGH-SET BASED MACHINE LEARNING

6.1. Introduction

The production and characterisation of pyrolysis bio-oil are generally labour intensive, costly, time consuming and complicated. To produce pyrolysis bio-oil with desired fuel properties, trial-and-error process was required to select the most suitable feedstock composition and pyrolysis operating condition. In addition, rigorous mechanistic models to allow accurate prediction of product properties from feedstock characteristics are not available. To avoid such difficulties, empirical models have been proposed to estimate the pyrolysis bio-oil's properties based on its characterisation and pyrolysis condition. Nevertheless, most of the studies focused on estimating the pyrolysis bio-oil properties from its elemental composition. The studies to estimate pyrolysis bio-oil from its feedstock properties and operating condition is still rare. Hence, a data-driven rough-set machine learning (RSML) model has been proposed in this chapter as a pre-processing and predictive modelling tool to predict the fast pyrolysis bio-oil's properties based on pyrolysis temperature and the characterisation of biomass feedstock. A database consisting of biomass feedstocks' proximate and ultimate analyses, pyrolysis temperature, fast

pyrolysis bio-oil's pH value, and fast pyrolysis bio-oil's higher heating value (HHV) was compiled and used to train the RSML model. To get a better range of parameters and to reduce sample bias, different types of feedstock such as rice straw, oil palm shell, sewage sludge, municipal solid waste, waste tyre, algae, etc. were included in the developed database. Four case studies with varied combination of attributes were solved in this study. The resulting rule-based RSML model demonstrated promising strength, certainty, and coverage factor. Furthermore, the generated decision rules were analysed from a scientific standpoint to identify underlying trends or patterns, while ensuring the rules were logical and feasible to be applied in later stage. The generated rules illustrated reasonable predictive capability in estimating the HHV and pH value of pyrolysis bio-oil based on the feedstock characterisation and pyrolysis temperature. RSML model is thus demonstrated to be a simple and straightforward approach for feedstock composition and pyrolysis temperature selection in pyrolysis/co-pyrolysis bio-oil production.

6.2. Problem Statement

One of the most significant properties of fast pyrolysis bio-oil is the HHV, where it indicates its quality as a liquid fuel and is directly related to the elemental composition of pyrolysis bio-oil. However, the studies to estimate HHV of pyrolysis bio-oil from its biomass feedstock properties and operating conditions are not common. In addition, the low pH value of pyrolysis bio-oil often leads to corrosion issues impede its application as engine fuel. The bio-oil's pH is often related to the characterisation of biomass feedstock. Nevertheless, the estimations of bio-oil's pH from feedstock and operating

conditions have yet to be explored. In the past, pyrolysis bio-oil's quality was often linked with the feedstock characterisation such as the ultimate analysis (i.e., carbon C, hydrogen H, oxygen O, nitrogen N, and sulphur S content), proximate analysis (i.e., volatile matter VM, ash content AC, moisture content MC, and fixed carbon FC), and the pyrolysis temperature (Tang et al., 2021). However, the production and the characterisation of pyrolysis bio-oil could be labour intensive, costly, and complicated. In addition, rigorous mechanistic models to allow accurate prediction of product properties from feedstock characteristics are not available. To avoid such difficulties, empirical models have been proposed to estimate the pyrolysis bio-oil's properties based on its characterisation and pyrolysis condition. It is notable that the empirical models developed in the past mainly focused on the prediction of bio-oil yield, while prediction models for other parameters such as HHV and pH of bio-oil are still limited. However, popular black-box machine learning (ML) techniques such as artificial neural network (ANN) and support vector machine (SVM) suffer from poor inherent interpretability (Yang et al., 2015). Other than ML techniques, statistical tools including response surface method (RSM) and multiple linear regression (MLR) approaches will lead to regression models that are also black-box in nature. To address this research gap, it is important to develop a predictive model that predicts these bio-oil properties from feedstock characteristics and pyrolysis temperature using RSML tools. With RSML approach, straightforward and transparent rules can be generated for further information extraction, thus addressing problems encountered in previous work. In addition to being inherently interpretable, ML techniques that use rule-based models allow fusion of expert knowledge with data embedded in information

during the training process. Domain knowledge can be introduced via user-defined training parameters to ensure that the final rules are plausible from the perspective of physical mechanisms (Rudin and Ertekin, 2018).

6.3. Methodology

The main objective of this chapter is to predict the properties of fast pyrolysis bio-oil by deducing the underlying patterns and relationships between the characterisation of biomass feedstock and pyrolysis temperature, with the pH and HHV of fast pyrolysis bio-oil. An RSML model was used as a predictive modelling tool to generate decision rules for the prediction of fast pyrolysis bio-oil properties. Figure 6-1 demonstrates the schematic diagram of the developed framework.

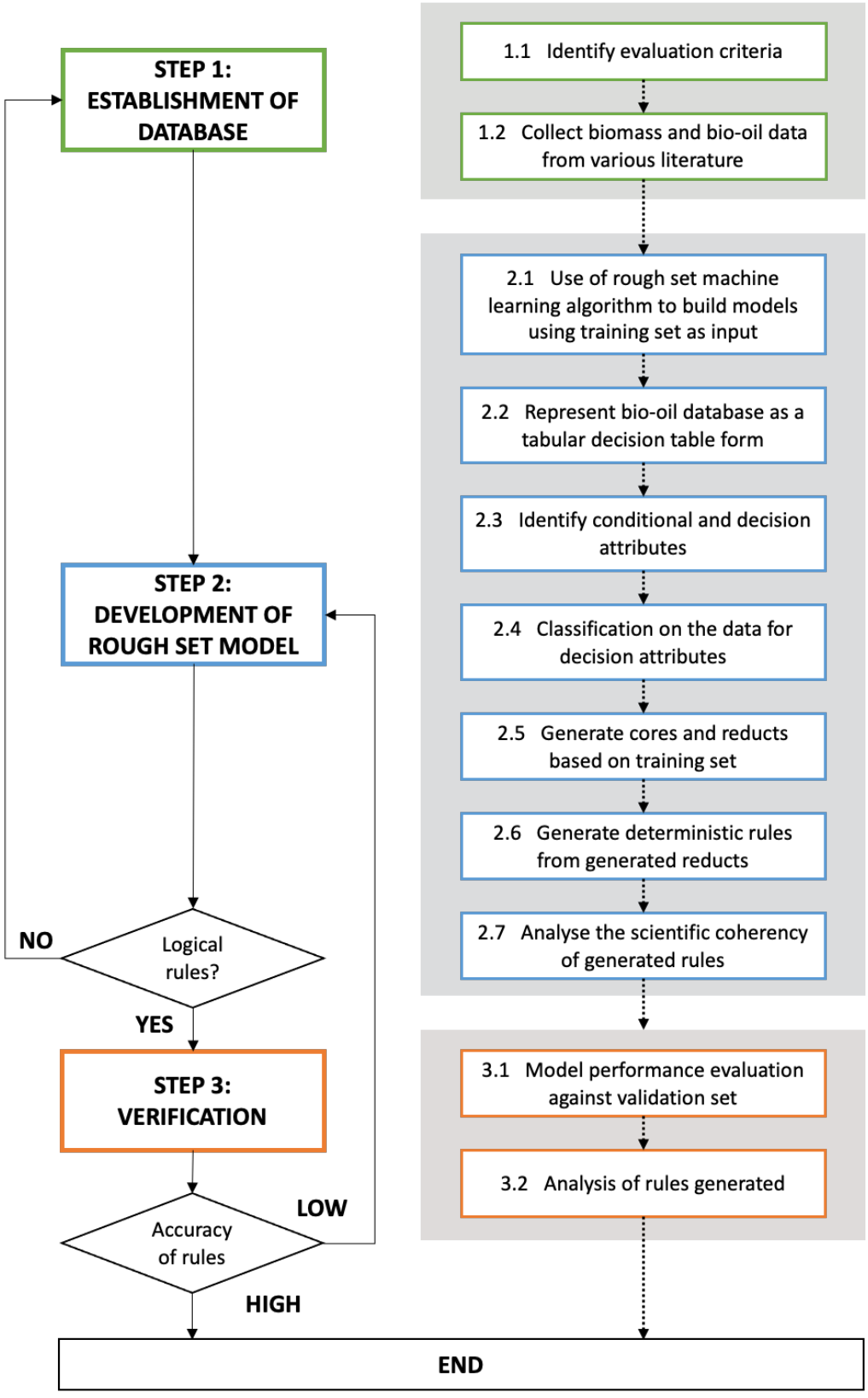


Figure 6-1 Framework for the development of rough set model for the optimization of pyrolysis bio-oil's properties

6.3.1. Compilation of Database

In this chapter, a fast pyrolysis bio-oil database consisting of 207 data points for bio-oil HHV and 128 data points for bio-oil pH value was developed based on the compiled data from various published literature. The complete database can be found at Table C-1 in Appendix C. In this study, 70 % of samples was randomly chosen from the established database as training subset, which was then used to train the RSML model. Meanwhile, the remaining (30 % of entire database) was set aside for validation. To get a better range of parameters and to reduce sample bias, different types of feedstock such as rice straw, oil palm shell, sewage sludge, municipal solid waste, waste tyre, algae, etc. were included in the developed database. Both the HHV and pH values of pyrolysis bio-oil were evaluated in the state of whole fast pyrolysis bio-oil to consider the influence from both aqueous and organic phases. Meanwhile, the ultimate analysis information (i.e., C-H-O-N-S content), proximate analysis information (i.e., VM-AC-MC-FC), HHV, and pyrolysis temperature served as the feedstock attributes for establishing the prediction models. The ultimate analysis (i.e., C-H-O-N-S content) of feedstock were in the range of 20.34 % to 86.40 %, 3.39 % to 23.68 %, 0.80 % to 54.90 %, 0.01 % to 8.23 % and 0.01 % to 3.30 %, respectively. Meanwhile, the proximate analysis (i.e., VM-AC-MC-FC) of feedstock ranged from 18.33 % to 96.88 %, 0.10 % to 57.78 %, 0.41 % to 77.52 %, and 0.28 % to 32.31 %, respectively. The pyrolysis temperature ranged from 300 ºC to 800 ºC. On the other hand, the HHV of feedstock and pyrolysis bio-oil ranged from 3.70 MJ/kg to 44.54 MJ/kg and 3.72 MJ/kg to 45.71 MJ/kg, respectively, while the pH values of pyrolysis bio-oil ranged from 1.50 to 8.50. It is worth noting that there were some uncertainties in the dataset,

for example, the measurement errors found in samples and the uncertainties of experimental environments. To minimise these uncertainties, the experimental environmental was strictly limited to only fast pyrolysis without any catalysts. In other words, the data collected in this work only include the bio-oil produced from fast pyrolysis without any catalysts. Here, the fast pyrolysis is defined as a process where the biomass is heated to the temperature ranging between 500 ºC to 800 ºC, at a heating rate of 10 ºC/s to 200 ºC/s, and solid residence time of 0.5 s to 10 s.

6.3.2. Development of Rough Set Model

To develop the RSML model, the fast pyrolysis bio-oil database was represented as a tabular decision table. In the decision table, each row represents an object, while the columns represent the attributes corresponding to each object. Here, the decision table was represented as the decision system, $S = (U, A)$, where $U = \{x_1, x_2, \dots, x_n\}$ is a non-empty finite set of objects, also known as the universe and $A = \{a_1, a_2, \dots, a_m\}$ is a non-empty finite set of attributes, such that $a: U \rightarrow V_a$ and $a \in A$ (Kalaivani et al., 2017). This set of attributes are further categorised as condition attributes which describe the object (i.e., pyrolysis temperature, ultimate and proximate analyses of feedstock samples) and decision attributes which indicate the classes of the object (i.e., HHV and pH value of pyrolysis bio-oil samples). In this study, classification of decision attributes was conducted based on the HHV and pH of pyrolysis bio-oil. An example of a simplified decision table for pyrolysis bio-oil's HHV is shown in Table 6-1.

Table 6-1 Simplified pyrolysis bio-oil's HHV decision table

Feedstock	Condition Attributes			Decision Attributes
	A1	A2	A3	HHV (MJ/kg)
B1	19.73	2.41	19.73	1
B2	38.12	1.94	38.12	3
B3	24.55	1.94	24.55	2
B4	19.73	2.41	19.73	1
B5	19.73	2.41	19.73	1
B6	38.12	1.94	38.12	3
B7	19.73	3.09	19.73	1

From Table 6-1, the objects in the model (biomass samples) were represented as B1 to B7 under the "Feedstock" column. Meanwhile, the column "A1", "A2", and "A3" represent the conditional attributes (i.e., pyrolysis temperature, ultimate and proximate analyses of feedstock samples) corresponding to each biomass sample. Lastly, the column "Decision Attributes" listed out the classification of HHV values of the pyrolysis bio-oil. Each bio-oil sample demonstrated different HHV values, ranging from 3.72 MJ/kg to 45.71 MJ/kg. For ease of data handling, the HHV of bio-oil were classified into 3 class intervals as shown in Table 6-1. For example, Class 1 represents pyrolysis bio-oil with HHV values between 10 MJ/kg to 19 MJ/kg; Class 2 represents pyrolysis bio-oil with HHV values between 20 MJ/kg to 29 MJ/kg; and Class 3 represents pyrolysis bio-oil with HHV values between 30 MJ/kg to 39 MJ/kg.

Once the decision table was established, reduction of condition attributes was performed. It is possible for a large dataset to contain two or more objects which perform similarly in the attributes or features, commonly known as the indiscernible objects. In rough set theory, reduct refers to a subset of indispensable attributes which can partition the database with the same level of discrimination as the original set of attributes (Pawlak, 1982). On the other hand,

the intersection of all reducts are known as the core. It may also be regarded as the essential attributes set which cannot be excluded from the decision system without losing the equivalence class structure. To reduce the redundancy of data while retaining its basic features, only one representative object was stored in the dataset for every set of indiscernible objects (Dutta, 2019).

For the decision system, S with set of attributes $B \subseteq A$, the indiscernibility relation can be denoted as Equation 6.1.

$$IND_B(U) = \{(x, y) \in U^2 | \forall a \in B, a(x) = a(y)\} \quad (6.1)$$

Taking Table 6-1 as an example and letting $B \in U$, the indiscernibility relation of U can be shown as Equation 6.2. Meanwhile, the indiscernibility of attributes set $\{A2, A3\}$, attributes set $\{A1, A3\}$ and attributes set $\{A1, A2\}$ are shown in Equation 6.3, 6.4 and 6.5, respectively.

$$IND\ (U) = \{B1, B4, B5\}, \{B2, B6\}, \{B3\}, \{B7\} \quad (6.2)$$

$$IND\ (U - A1) = \{B1, B4, B5\}, \{B2, B6\}, \{B3\}, \{B7\} = \boldsymbol{IND}\ (\boldsymbol{U}) \quad (6.3)$$

$$IND\ (U - A2) = \{B1, B4, B5, B7\}, \{B2, B6\}, \{B3\} \quad (6.4)$$

$$IND\ (U - A3) = \{B1, B4, B5\}, \{B2, B6\}, \{B3\}, \{B7\} = \boldsymbol{IND}\ (\boldsymbol{U}) \quad (6.5)$$

According to Equation 6.3 and 6.5, both indiscernibility of attributes set $\{A2, A3\}$ and attributes set $\{A1, A2\}$ were equivalent to the indiscernibility relation of U, $\boldsymbol{IND}\ (\boldsymbol{U})$. This indicated that attributes $A1$ and $A3$ were dispensable as the removal of these attributes from relation U does not affect the results. However, attribute $A2$ was indispensable as the removal of this attribute affected the result as shown in Equation 6.4. The classification defined

by U was equivalent to the classification defined by attributes set of $\{A1, A2\}$ or attributes set of $\{A2, A3\}$. Thus, the attribute sets $\{A1, A2\}$ and $\{A2, A3\}$ were also known as the reducts for the dataset R. In this case, attribute $A2$ was identified as the core for the decision system. The generated reduct sets were used to derive a set of decision rules, which were used to predict the HHV and pH value of pyrolysis bio-oil in a later stage. For the reduct set $\{A1, A2\}$, the example of decision rules generated are shown in Table 6-2.

Table 6-2 Example of rules generated for reduct set {A1, A2}

No	Rule	Decision
1	$A1 = 19.73; A2 \geq 2.41$	1
2	$A1 = 24.55; A2 = 1.94$	2
3	$A1 = 38.12; A2 = 1.94$	3

In the current work, Rough Sets Data Explorer (ROSE2) software was used to generate the cores, reducts and decision rules, based on Learning from Examples Module, version 2 (LEM2) algorithms (Grzymala-Busse and Grzymala-Busse, 1995). ROSE2 is an interactive system running on 32-bit Microsoft Windows operating systems, developed at the Laboratory of Intelligent Decision Support systems of the Institute of Computing Science in Poznań, Poland (IDSS, 1999; Predki et al., 1998; Prędki and Wilk, 1999). The software consists of both basic and advanced data analysis methods based on the classical rough set theory and the variable precision rough set theory (Pięta et al., 2019). Meanwhile, LEM2 algorithm is a rule induction algorithm based on rough set theory (Grzymala-Busse and Grzymala-Busse, 1995). In LEM2 algorithm's approach, decision rules in the form of "IF-THEN" are generated by considering the lower and upper approximations. In other words, the "IF-

THEN" decision rules implied that *"if* the condition attribute is ***A***, *then* the decision attribute is ***B**",* where ***A*** and ***B*** are the hypotheses and conclusion, respectively (Anuradha et al., 2011; Qian et al., 2008). This algorithm is suitable for use in both increment and non-increment modes. In addition, the underlying trends in the generated rules were analysed and interpreted from a scientific standpoint. Generated rules are valid if it can be explained scientifically. If otherwise, the previous step was revisited and revised to modify the condition and decision attributes.

6.3.3. Validation

In the validation step, performance, and prediction accuracy of the reduct sets and generated decision rules were evaluated using the validation data set. Data in the validation set consisted 30 % of the overall data, with no overlaps with the training data set. The purpose of this step is to gauge the model's performance on data it has not previously encountered. Promising validation set performance indicates that the model has learned the useful generalisable principles. Underlying patterns in the data should be converted into rules with a good balance of prediction accuracy and generalization power as indicated by the number of examples covered. However, if the generated rules exhibit low accuracy or coverage, it is due to the occurrences of coverage clusters in the dataset where each of the cluster demonstrates unique behaviour. Thus, the rough set model will be re-examined and revised to refine the prediction accuracy and coverage.

In RSML, the strength, certainty, and coverage factors are normally employed to assess the performance of the rule-based models vis-à-vis the

training data. Strength and coverage measure the generalization power of a rule, while certainty measures its predictive reliability. Higher strength, certainty, and coverage factors are preferable.

The strength of a decision rule can be referred to as the fraction of data points in a dataset that supports it and can be obtained by Equation 6.6. From Equation 6.6, C and D refer to the set of condition and decision attributes, respectively. Meanwhile, $supp_x(C, D) = |A(x)| = |C(x) \cap D(x)|$ refers to the number of elements that supported the decision rule $C \rightarrow_x D$ and $|U|$ is the cardinality of U.

$$\sigma_x(C, D) = \frac{supp_x(C, D)}{|U|} \tag{6.6}$$

Since variations in training data quality generally prevents perfect classification, it is necessary to have a metric to quantify inconsistencies. For decision rule $C \rightarrow_x D$, the conditional probability of an element characterised as $C(x)$ being classified into the decision class $D(x)$ is known as the certainty factor. The certainty factor can be denoted as $cer_x(C, D)$, and is defined by Equation 6.7 and 6.8, where as $\sigma_x(C) = |C(x)|/|U|$. If a decision rule $C \rightarrow_x D$ has a certainty factor of $cer_x(C, D) = 1$, then the rule will be known as a certain or deterministic decision rule. However, rules with certainty factor $0 < cer_x(C, D) < 1$ is known as an uncertain or rough decision rule.

$$cer_x(C, D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{supp_x(C, D)}{|C(x)|} = \frac{\sigma_x(C, D)}{\sigma_x(C)} \tag{6.7}$$

$$cer_x(C, D) = \pi_x(D|C) \tag{6.8}$$

On the other hand, coverage factor of a decision rule $C \rightarrow_x D$ is expressed as the percentage of elements belonging to $D(x)$ being classified under the given rule and can be denoted as $cov_x(C, D)$ as shown in Equation 6.9 and 6.10, where $\sigma_x(D) = |D(x)|/|U|$. It quantifies the extent to which a given rule accounts for a particular decision outcome.

$$cov_x(C, D) = \frac{|C(x) \cap D(x)|}{|D(x)|} = \frac{supp_x(C, D)}{|D(x)|} = \frac{\sigma_x(C, D)}{\sigma_x(D)} \tag{6.9}$$

$$cov_x(C, D) = \pi_x(C|D) \tag{6.10}$$

Other than the standard RSML metrics relative to the training data, the performance of the model can be empirically assessed using the validation data set. The Type I and Type II error rates were also included to further analyse the performance of the prediction model. The Type I error, also known as "false positive", occurs when a true null hypothesis was mistakenly rejected. For example, Type I error occurred when a bio-oil with HHV of Class 1 was mistakenly classified as Class 3. On the other hand, the Type II error rate, also known as the "false negative" occurs when a false null hypothesis was mistakenly accepted. For example, Type II error occurred when a bio-oil with HHV of Class 1 was not classified under any rules for Class 1. However, in this study, only Type I error rate was reported as it carries much more consequences than the Type II error rate.

6.4. Case Study

6.4.1. Data Classification

In this study, the pyrolysis temperature, feedstock's ultimate and proximate analyses were identified as the condition attributes. Meanwhile, the HHV and pH values of pyrolysis bio-oil were selected as the decision attributes. Four rough set models with different combination of attributes were developed. Figure 6-2 illustrates the combinations of condition and decision attributes for Case Study 1, 2, 3 and 4, respectively. The data points considered in each Case Study differs as some information for certain data points were missing from the reported literature.

The HHV and pH values of pyrolysis bio-oil were predicted separately to investigate the effect of different combination of biomass characteristics and pyrolysis temperature on these properties. These properties are treated as outputs that are predicted in parallel via the rule-based models with a common set of input variables and are hence already indirectly linked to each other.

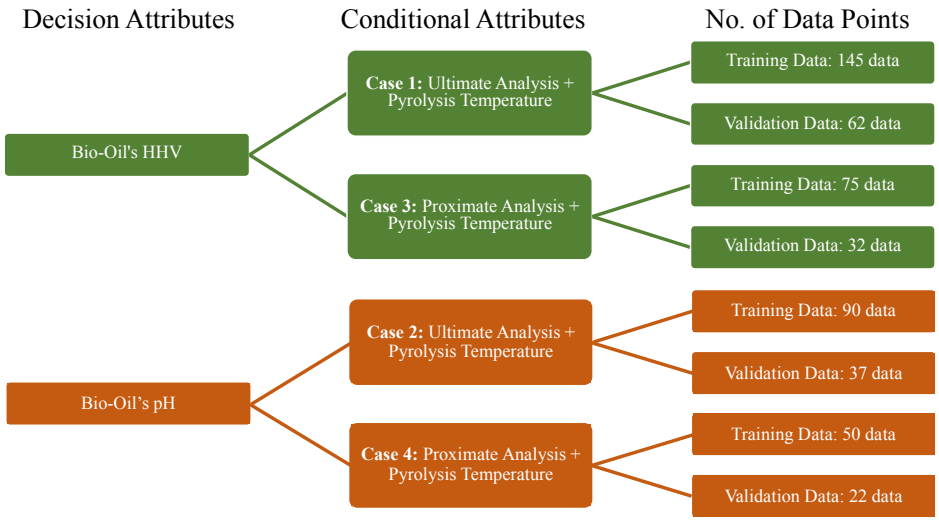


Figure 6-2 Attribute sets and decision sets along with the number of data points for each case study

The rough set model was formulated by classifying the data for decision attributes (i.e., HHV and pH values of pyrolysis bio-oil) into 5 and 6 class intervals for bio-oil's HHV and pH value, respectively, using continuous frequency distribution. Class categories were defined in increments of 10.00 MJ/kg for the bio-oil's HHV while increments of 1 was applied for bio-oil's pH value, as shown in Table 6-3. For each case study, 70 % of the entire database were categorised as the training data set, which was used to train the rough set model. Tabular decision tables were established with these data as explained in Section 6.3.2.

Table 6-3 Categorization of bio-oil's HHV and pH value for each case study

Decision Attribute	Value	Category
Bio-oil's HHV	0.00 – 9.00 MJ/kg	Class 0
	10.00 – 19.00 MJ/kg	Class 1
	20.00 – 29.00 MJ/kg	Class 2
	30.00 – 39.00 MJ/kg	Class 3
	40.00 MJ/kg and above	Class 4
Bio-oil's pH	1.00 – 1.90	Class 1
	2.00 – 2.90	Class 2
	3.00 – 3.90	Class 3
	4.00 – 4.90	Class 4
	5.00 – 5.90	Class 5
	6.00 and above	Class 6

6.4.2. Cores and Reducts

With the established decision system in Case Study 1, 2, 3 and 4, the cores and reducts were identified and generated by the rough set model using ROSE2 software. The complete core and reduct sets for each case study are listed in the Table 6-4. For Case Study 1, 2 and 3, the pyrolysis temperature was identified as the core. Being the core, pyrolysis temperature was also known as the most decisive subset of attributes in the decision table. In other words, the

pyrolysis temperature attribute cannot be excluded from the decision system without influencing the classification power of the attributes. Meanwhile, no core was identified for Case Study 4. In other words, any single conditional attributes can be removed without altering the equivalence-class structure. In such cases, there is no essential or necessary attribute which is required for the class structure to be represented. A total of 30 reduct sets were generated for all the 4 case studies. These reduct sets were employed to develop decision rules by running the rough set models.

Table 6-4 Cores and reducts generated for Case 1, 2, 3 and 4

Case	Cores	Reducts	Total No. of Rules Generated
1	Temperature	Reduct 1: Carbon, Hydrogen, Temperature	48
		Reduct 2: Carbon, Oxygen, Temperature	41
		Reduct 3: Hydrogen, Oxygen, Temperature	47
		Reduct 4: Carbon, Nitrogen, Temperature	43
		Reduct 5: Oxygen, Nitrogen, Temperature	45
2	Temperature	Reduct 1: Carbon, Oxygen, Temperature	37
		Reduct 2: Hydrogen, Oxygen, Temperature	35
		Reduct 3: Carbon, Hydrogen, Sulphur, Temperature	36
		Reduct 4: Oxygen, Nitrogen, Temperature	33
		Reduct 5: Carbon, Nitrogen, Temperature	34
		Reduct 6: Oxygen, Sulphur, Temperature	40
3	Temperature	Reduct 1: Volatile Matter, Ash, Temperature	23
		Reduct 2: Volatile Matter, HHV, Temperature	29
		Reduct 3: Ash, HHV, Temperature	26
		Reduct 4: Volatile Matter, Fixed Carbon, Temperature	26
		Reduct 5: Ash, Fixed Carbon, Temperature	24
		Reduct 6: Moisture Content, Temperature	31
		Reduct 7: Fixed Carbon, HHV, Temperature	24
4	-	Reduct 1: Volatile Matter, Moisture Content	15
		Reduct 2: Ash, Moisture Content	15
		Reduct 3: Volatile Matter, Fixed Carbon	18
		Reduct 4: Ash, HHV	18
		Reduct 5: Moisture Content, HHV	18
		Reduct 6: Moisture Content, Fixed Carbon	15
		Reduct 7: Fixed Carbon, HHV	22
		Reduct 8: Ash, Fixed Carbon	15
		Reduct 9: Volatile Matter, Temperature	26
		Reduct 10: Volatile Matter, HHV	21
		Reduct 11: Moisture Content, Temperature	21
		Reduct 12: HHV, Temperature	27

6.4.3. Rule-Based Model

In this study, a total of 833 decision rules were generated from these 30 reduct sets. Out of these 833 decision rules, 225 rules were generated in Case Study 1, 215 rules in Case Study 2, 183 rules in Case Study 3 and 232 rules in Case Study 4. All the generated rules along with the respective strength, certainty factor, and coverage factor are listed out in Table C-2 to Table C-31 in Appendix. Decision rules of different decision classes generated from Case Study 1: Reduct 1 was listed in Table 6-5.

Table 6-5 Generated rules from Case Study 1: Reduct 1

No.	Attributes	Rule	Decision	Strength	Coverage	Certainty
10	Carbon Hydrogen Temperature	$47.35 - 49.35$ < 6.15 ≥ 452.50	1	7.59 %	15.07 %	100 %
25	Carbon Hydrogen Temperature	< 48.05 $5.95 - 6.05$ -	2	6.90 %	22.22 %	100 %
47	Carbon Hydrogen Temperature	≥ 80.25 - -	4	5.52 %	88.89 %	100 %

Based on the details from Table 6-5, Rule 10 can be interpreted as the statement " *If the feedstock sample possessed carbon content between the range of 47.35 to 49.35 %, hydrogen content of below 6.15 % and pyrolysis temperature of greater or equal to 452.50 ℃, then the pyrolysis bio-oil produced will have a HHV within the range of 10.00 to 19.00 MJ/kg.* " The strength of Rule 10 was 7.59 % as it was supported by 11 out of the 145 feedstock samples from the training data set. On the other hand, the coverage factor for this rule was observed to be 15.07 % where 11 out of the 73 feedstock samples of Class 1 were classified by the rule into decision Class 1. Finally, Rule 10 can be known as a certain decision rule as the certainty factor obtained

for this rule was 100 %. In other words, all 11 feedstock samples that supported this rule were correctly classified as Class 1.

On the other hand, Rule 25 and 47 can be interpreted as " *If the feedstock sample has a carbon content of below 48.05% and hydrogen content between the range of 5.95 to 6.05 %, then the produced bio-oil will be having a HHV values within the range of 20.00 to 29.00 MJ/kg."* and " *If the feedstock sample has a carbon content of greater or equal to 80.25 %, then the produced bio-oil will be having a HHV values of greater or equal to 40.00 MJ/kg. "* respectively.

By observing the decision rules generated (Table C-2 to Table C-31 in Appendix C), all the rules generated have a certainty factor of 100 %. This indicates that all the feedstock samples were correctly categorised under respective classes with the generated decision rules. In addition, further evaluation on the performance of the rules were conducted and will be discussed in Section 6.4.5. Thus, we can conclude that the rules generated by the RSML model still possess high potential in predicting the HHV and pH of the pyrolysis bio-oil.

6.4.4. Mechanistic Plausibility of Rules

In Case Study 1, five reduct sets were identified from the decision table for Case Study 1 and pyrolysis temperature was obtained as the core attribute. This indicated that the HHV of pyrolysis bio-oil is highly dependent on the pyrolysis temperature. From the decision rules generated, it was observed that pyrolysis bio-oil with HHV values of at least 10.00 MJ/kg can be obtained by conducting the pyrolysis process with temperature ranging between 400 ºC to 600 ºC.

Among the reduct sets in Case Study 1, decision rules in Case Study 1: Reduct 2 showed the highest overall coverage and certainty factor. The condition attributes in this reduct included the pyrolysis temperature, carbon, and oxygen content. Furthermore, all the reduct sets in Case Study 1 consisted of either carbon content and/or oxygen content as the condition attributes. This may be explained by the domination of the carbon and oxygen content in the ultimate analysis of the feedstock samples, which make up to 80 to 90 % of the entire feedstock samples. In the past, various literature often correlates the carbon and oxygen content of feedstock samples with the energy content of the fuel. Higher carbon content was reported to improve the combustion properties of fuel and thus resulting in higher HHV of bio-oil (Ben et al., 2019). However, lower oxygen content was preferable in any fuel as it reduced the HHV of fuel (Kasar and Ahmaruzzaman, 2021). From the decision rules generated as shown in Table 6-6, we can conclude that to achieve HHV of at least Class 2, the carbon and oxygen content of the feedstock sample needs to be greater than 44 % and below 45 %, respectively. For example, Rule 27 from Case Study 1: Reduct 2 stated that " *If the feedstock sample has a carbon content of greater or equal to 48.5% and oxygen content of greater or equal to 43.45 %, then the produced bio-oil will be having a HHV values within the range of 20.00 to 29.00 MJ/kg."*

Table 6-6 Rules involving the attributes: Carbon and Oxygen to generate bio-oil with HHV greater or equal to 20 MJ/kg.

Reduct 2	**Rule No.**	**Carbon (%)**	**Oxygen (%)**	**Temperature (ºC)**
	27	≥ 48.5	≥ 43.45	445.00 - 477.50
	32	46.15 - 47.70	-	-
Reduct 4	**Rule No.**	**Carbon (%)**	**Nitrogen (%)**	**Temperature (ºC)**
	25	45.85 - 46.65	< 2.15	-
Reduct 5	**Rule No.**	**Oxygen (%)**	**Nitrogen (%)**	**Temperature (ºC)**
	26	21.60 - 38.55	≥ 0.85	-
	30	≥ 40.35	0.85 - 2.75	≥ 502.50
	36	44.00 - 44.45	-	-

On the other hand, the sulphur content was not included in any of the reduct sets as a condition attribute. This can be explained by the small amount of sulphur content where the sulphur content was the least elemental constituent of the feedstock samples, followed by nitrogen content. Thus, we can conclude that the effect of sulphur content in feedstock was negligible in determining the HHV of pyrolysis bio-oil. Lastly, no obvious trend can be observed between the hydrogen content of feedstock sample and the HHV of pyrolysis bio-oil. This can be supported by the literature which reported that no distinct relation can be observed between HHV and hydrogen content (Hasan et al., 2018) .

In Case Study 2, the ultimate analysis of feedstock and pyrolysis temperature were employed to predict the pH value of pyrolysis bio-oil. Up to date, no study was done on correlating the ultimate analysis of feedstock with the pH value of pyrolysis bio-oil. From the results obtained from the rough set model, the rules generated in Case Study 2: Reduct 4, 5, and 6 showed higher certainty and coverage factor as compared to other reduct sets. The common attributes found in these three reducts were the nitrogen and sulphur content. Thus, we can deduce that the nitrogen and sulphur content affect the final pH value of pyrolysis bio-oil the most as compared to the remaining content such as carbon, oxygen, and hydrogen content.

From the decision rules generated in Case Study 2: Reduct 4 and 5, it is observed that the pH values of pyrolysis bio-oil increases as the nitrogen content in the feedstock sample increases. This can be explained by the nitrogenous compound formed in the pyrolysis bio-oil such as amines, amides, etc. Commonly, more than 50.00 % of the nitrogen content in feedstock sample was

converted into the bio-oil in the form of organic or ammonium as nitrogen containing product (Leng and Zhou, 2018). Amines are the most found nitrogenous compounds in bio-oil, which contributes to the alkalinity of bio-oil (Leng et al., 2020). This can be further confirmed by the results of the rough set model where both Case Study 2: Reduct 4 and 5 have the highest overall performance. Based on Table 6-7, the decision rules generated from both reducts, nitrogen content of greater than 0.35 % was required to achieve pH value of greater than 3. In addition, to achieve pH values of greater than 8, the nitrogen content of feedstock sample needs to be greater than 6.45 %. On the other hand, the sulphur content of feedstock samples was observed to increase the pH value of bio-oil as well. However, the trend was not obvious as the amount of sulphur content in the feedstock sample was too little.

Table 6-7 Rules involving the attribute: Nitrogen to generate bio-oil with pH greater or equal to 3.

	Rule No.	**Nitrogen (%)**	**Oxygen (%)**	**Temperature (ºC)**
Reduct 4	15	0.35 - 0.65	-	< 505.00
	16	≥ 0.55	≥ 44.25	-
	17	≥ 0.35	43.65 - 43.95	-
Reduct	**Rule No.**	**Nitrogen (%)**	**Carbon (%)**	**Temperature (ºC)**
5	13	0.35 - 0.95	< 49.20	-

In Case Study 3, the relationship between the proximate analysis of feedstock and HHV of bio-oil were studied. In the past decades, many correlations were established in the literature to estimate the HHV of fuel using proximate analysis of feedstock. The proximate analysis is one of the main attributes which measures the fuel or feedstock's energy content via ratio of combustibles (VM and FC) to non-combustibles (AC and MC) (Qureshi et al., 2019).

The VM of feedstock samples can be referred to as the fuel fraction released when heated at high temperature (950 ºC). From the rules generated in Case study 3, higher VM in feedstock sample was correlated to higher HHV of pyrolysis bio-oil. To achieve HHV of at least Class 2, the VM of feedstock sample needs to be greater than 78.00 %, as shown in Table 6-8. Higher VM of fuel or feedstock implied higher reactivity and enriched combustion process which supported the liquid fuel production (Omar et al., 2011). On the other hand, feedstock with higher FC content will usually produce bio-oil with higher carbon content, and thus higher HHV (Abnisa et al., 2013a). However, no obvious trend was observed from the decision rules generated from the reduct sets with FC as attribute. FC was defined as the fraction of feedstock sample that remained after the VM was completely released, excluding AC and MC, forming the char (Obernberger and Thek, 2004). Thus, the FC content contributed more to the HHV of biochar rather than the bio-oil.

Table 6-8 Rules involving the attribute: Volatile Matter to generate bio-oil with HHV greater or equal to 20 MJ/kg

	Rule No.	**Volatile Matter (%)**	**Ash Content (%)**	**Temperature (ºC)**
Reduct 1	12	78.40 - 81.75	1.45 - 5.85	-
	13	78.40 - 84.60	1.45 - 2.85	-
	15	≥ 68.00	0.85 – 2.00	-
	Rule No.	**Volatile Matter (%)**	**HHV (MJ/kg)**	**Temperature (ºC)**
Reduct 2	16	68.00 - 81.75	16.70 - 20.10	537.50 – 675.00
	20	80.60 - 84.60	-	< 555.00

Next, high amounts of MC in feedstock sample was undesirable as it reduced the HHV of the fuel. During the bio-oil production process, the water content in feedstock goes to the bio-oil, leading to an increment of water content in bio-oil and thus lowering the HHV of the bio-oil (Asadullah et al., 2013).

From the results generated in Case Study 3, the MC of feedstock was observed to have no effect to the final HHV of pyrolysis oil. This may be due to the low MC in the feedstock samples, which were typically below 10.00 %. Lastly, the AC is the inorganic waste that remained after FC combustion. According to the rules generated, the AC was found to be unrelated to the HHV of pyrolysis bio-oil. AC was usually used to estimate and determine the yield percentage distribution of solid, liquid, and gaseous products (Abnisa et al., 2013b) .

In Case Study 4, the relationship between the proximate analysis of feedstock and the pH value of pyrolysis bio-oil were investigated. However, no obvious trend can be concluded from the rules generated in this case study. As the proximate analyses contributed more to the calorific values and the yield percentage of product, it can be concluded that the proximate analyses of feedstock samples do not affect the pH value of pyrolysis bio-oil (Venderbosch and Prins, 2010). Nevertheless, high coverage and certainty factor can be obtained from the rules generated by the reduct sets in Case Study 4. This can be explained by the smaller size of training and validation data size where most of the data might be coming from the same source.

6.4.5. Evaluation of Model Performance

The resulting decision rules from the rough set models were verified against the validation set, which consisted 30 % of the entire data set. Considering the generated decision rules from Case Study 1: Reduct 2 (Conditional Attributes: C, O, Pyrolysis Temperature; Decision Attributes: HHV), the decision rules were verified on 62 feedstock samples from the validation data set. By applying the decision rules generated, the pyrolysis bio-

oil from validation set can be classified under one of the five classes. For instance, the pyrolysis bio-oil was predicted to have a HHV ranging from 10.00 – 19.00 MJ/kg if the attributes from feedstock sample fit in one or more rules under Class 1. **Error! Reference source not found.** illustrates the validation result for the selected promising rules from Case Study 1: Reduct 2, where the certainty and coverage factor along with the respective number of molecules that supported a particular rule are shown. Based on **Error! Reference source not found.**, decision rules under Class 4 showed higher coverage factor compared to other classes (i.e., Class 1 and 2). This may be due to the smaller size of verification data set for Class 4 (Reduct 2: Case Study 1), which only consisted of 4 feedstock samples.

From Figure 6-3, 11 decision rules (Rule 6, 7, 8, 12, 13, 14, 16, 17, 18, 19 and 22) were classified under Class 1. However, not all the decision rules exhibited high coverage and certainty factor. Thus, the validation results were further analysed to select rules with highest confidence level for the application in further stages. Rule 17, 18, 19 and 21 displayed a prediction accuracy of 100 %. However, Rules 18 and 21 may not hold a high power of generalisation as they were only supported by one feedstock sample each. In this case, Rule 7, 8, 16 and 19 showed satisfactory prediction accuracy where each of the rules were supported by 3 or more feedstock samples with minimum accuracy of 60 %. In other words, the HHV of pyrolysis bio-oil was predicted to fall within the range of 10.00 to 19.00 MJ/kg if the ultimate analysis of feedstock samples fulfils either one of these decision rules (Rule 7, 8, 16 and 19). The general coverage and certainty factor for Rule 7, 8, 16, and 19 was listed in Table 6-9.

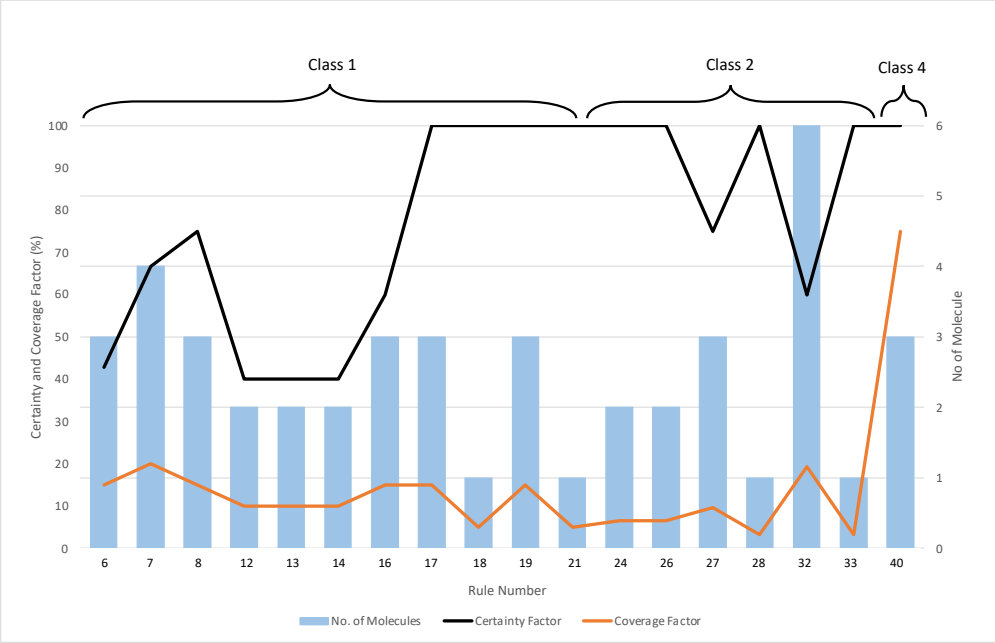


Figure 6-3 Validation results on Case 1: Reduct 2

On the other hand, Figure 6-4 demonstrated the validation result for Case Study 2: Reduct 4 (Conditional Attributes: O, N, Pyrolysis Temperature; Decision Attributes: pH). The generated decision rules were verified on 38 feedstock samples from the validation data set. With the application of the generated decision rules, the pyrolysis bio-oil from validation set can be classified under one of the six classes. The pyrolysis bio-oil are expected to demonstrate pH value ranging from 1.00 – 1.90 if the attributes from the feedstock sample fit in one or more rules under Class 1. Based on Figure 6-4, the resulting decision rules possesses high certainty factor where decision rules generated under Class 1, 2, 5, and 6 demonstrated certainty factor of 100%. Thus, it can be concluded that the model has successfully learned the useful generalisable principles. Furthermore, the decision rules under Class 5 and 6 also showed higher coverage factor compared to other classes (i.e., Class 1, 2 and 3). This can be explained by the smaller size of verification data set for Class 5 and 6 which only consisted of 1 and 2 feedstock sample, respectively.

From Figure 6-4, 6 decision rules (Rule 3, 4, 6, 8, 9 and 10) were classified under Class 2 with certainty factor of 100%. Despite the high prediction accuracy, Rule 6, 8, 9 and 10 were only supported by one feedstock sample, respectively. Meanwhile, Rule 3 and 4 were supported by 4 feedstock samples each, thus demonstrating promising validation performance. In other word, the pH of pyrolysis bio-oil was estimated to be within the range of 2.00 – 2.90 if the ultimate analysis of feedstock samples fulfils either one of these decision rules (Rule 3 and 4).

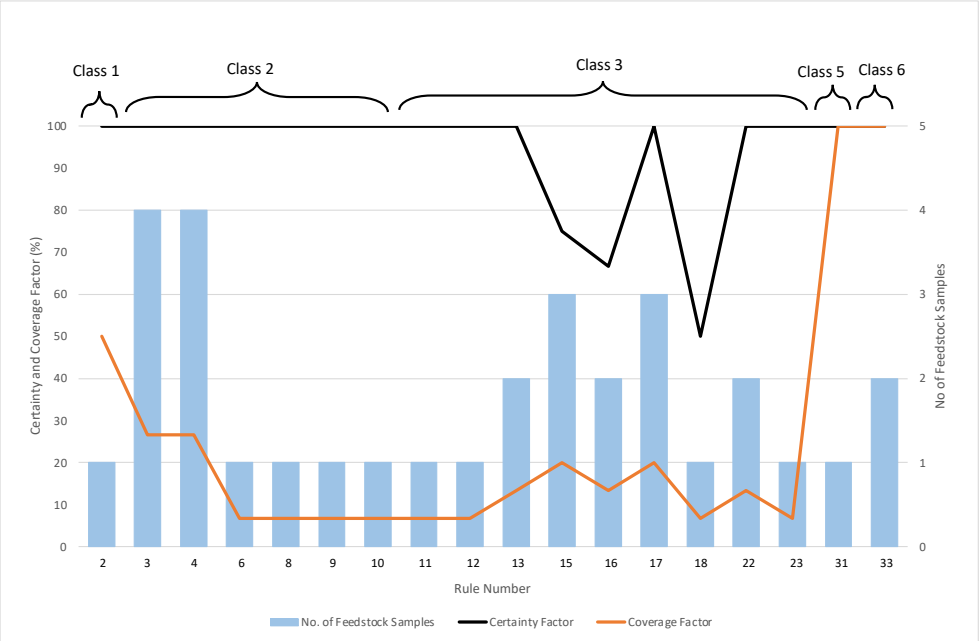


Figure 6-4 Validation results on Case 2: Reduct 4

A similar approach was applied for the remaining decision rules generated. The validation results for the best performing generated rules from Case Study 3: Reduct 2 and Case Study 4: Reduct 7 can be seen in Figure 6-5 and Figure 6-6, respectively.

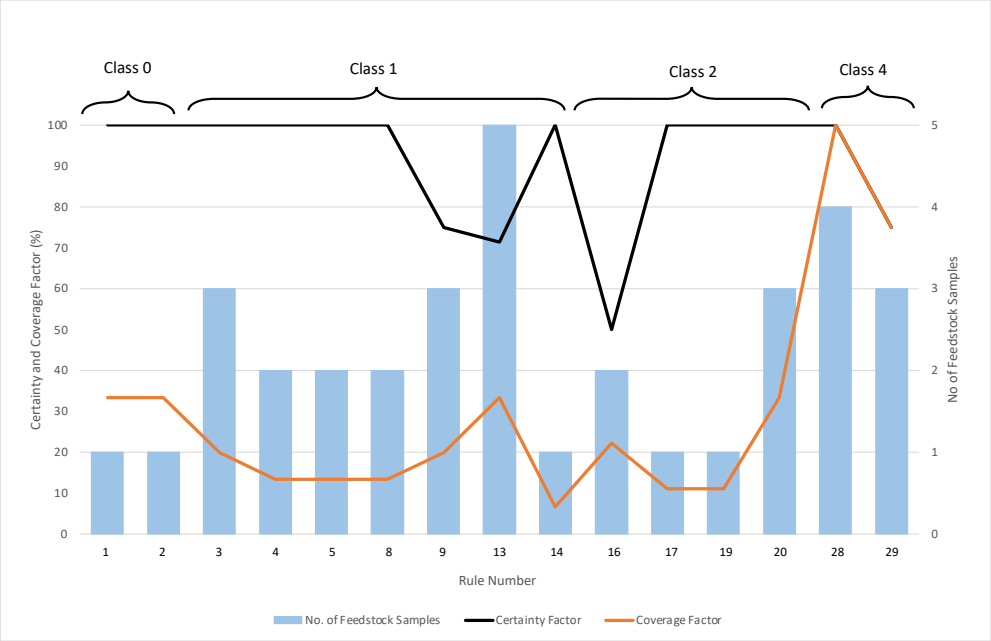


Figure 6-5 Validation results on Case 3: Reduct 2

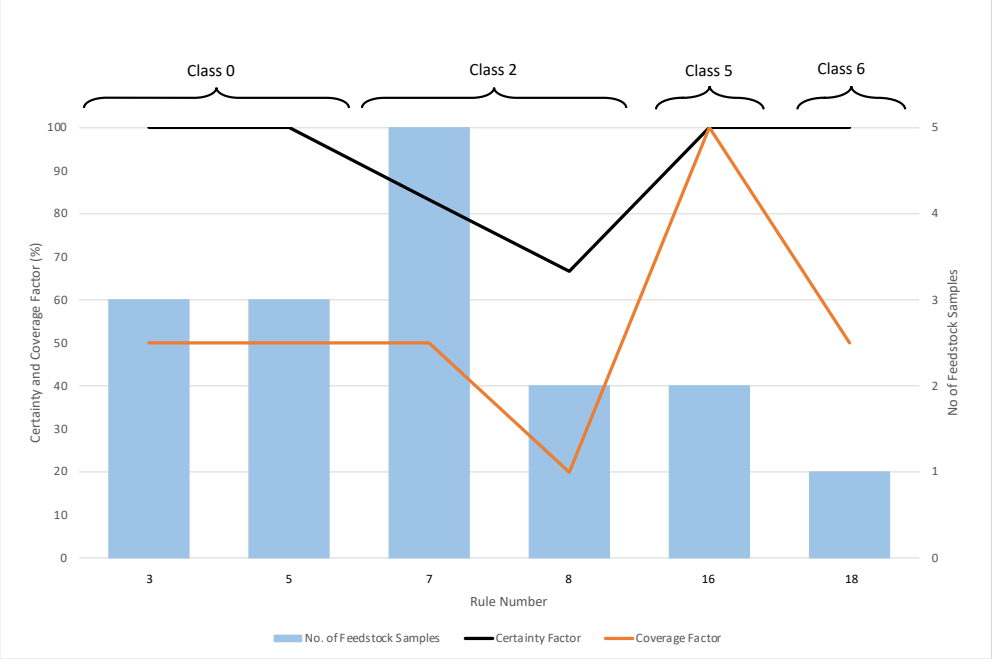


Figure 6-6 Validation results on Case 4: Reduct 7

The general validation results for decision rules for each class, from the best performing reduct sets of Case Study 1, 2, 3, and 4 were shown in Table 6-9. The summary of analysed validation results for the remaining rules are shown in Table C-32 to Table C-61 in the Appendix C.

Table 6-9 Summarized validation results for the best performing reduct set from Case Study 1, 2, 3 and 4

	Rule No	Rules			Bio-oil' s HHV (MJ/kg)	Strength (%)	Coverage (%)	Certainty (%)	Type I Error (%)
		Carbon (%)	Oxygen (%)	Temperature (°C)					
Case 1: Reduct 2	7	42.90 – 46.30	41.90 – 49.95	452.50 – 522.50					
	8	≥ 49.75	≥ 41.90	-					
	16	33.20 – 41.10	-	-	10 – 19	20.97	65.00	72.22	27.77
	19	48.05 – 48.85	41.75 – 44.00	-					
	27	≥ 48.50	≥ 43.45	445.00 – 477.50					
	32	46.15 - 47.70	-	-	20 – 29	14.52	29.03	64.29	35.71
	40	≥ 80.25	-	-	40 and above	4.84	75.00	100.00	0.00
	Rule No	Rules			Bio-oil' s pH	Strength (%)	Coverage (%)	Certainty (%)	Type I Error (%)
		Oxygen (%)	Nitrogen (%)	Temperature (°C)					
Case 2: Reduct 4	2	52.20 - 54.05	-	-	1.0 – 1.9	2.70	50.00	100.00	0.00
	3	≥ 48.95	0.23 – 0.35	-					
	4	44.55 - 47.70	< 0.35	≥ 495.00	2.0 – 2.9	21.62	53.33	100.00	0.00
	15	-	0.35 – 0.65	< 505.00					
	16	≥ 44.25	≥ 0.55	-	3.0 – 3.9	18.92	46.67	77.78	22.22
	17	43.65 - 43.95	≥ 0.35	-					
	31	42.35 - 42.80	-	-	5.0 – 5.9	2.70	100.00	100.00	0.00
	33	-	≥ 6.45	-	6.0 and above	5.41	100.00	100.00	0.00

Table 6-9 Summarized validation results for the best performing reduct set from Case Study 1, 2, 3 and 4 (continued)

	Rule No	Rules			Bio-oil' s HHV (MJ/kg)	Strength (%)	Coverage (%)	Certainty (%)	Type I Error (%)
		Volatile Matter (%)	HHV (MJ/kg)	Temperature (°C)					
Case 3: Reduct 2	1	-	15.75 - 15.90	≥ 537.50	0 – 9	6.25	66.67	100.00	0.00
	2	74.45 - 74.70	-	< 412.50					
	3	75.20 - 78.00	-	< 537.50	10 – 19	34.38	73.33	78.57	21.43
	4	75.20 - 80.00	< 18.15	-					
	9	≥ 75.20	< 18.25	< 502.50					
	13	< 61.00	-	-					
	16	68.00 - 81.75	16.70 - 20.10	537.50 - 675.00	20 – 29	12.50	44.44	66.67	33.33
	20	80.60 - 84.60	-	< 555.00					
	28	-	≥ 28.05	-	40 and above	12.50	100.00	100.00	0.00

	Rule No	Rules			Bio-oil' s pH	Strength (%)	Coverage (%)	Certainty (%)	Type I Error (%)
		Moisture Content (%)		HHV (MJ/kg)					
Case 4: Reduct 7	3	-		15.65 - 15.90	2.0 – 2.9	18.18	66.67	100.00	0.00
	5	9.60 - 11.40		-					
	7	4.35 – 8.00		≥ 14.25	3.0 – 3.9	31.82	70.00	77.78	22.22
	8	≥ 11.40		< 17.90					
	16	3.15 - 3.65		-	5.0 – 5.9	9.09	100.00	100.00	0.00
	18	< 5.80		< 12.00	6.0 and above	4.55	50.00	100.00	0.00

6.5. Conclusion

In this chapter, rule-based models were developed using RSML algorithms for use as a predictive modelling tool to estimate pyrolysis bio-oil's HHV and pH values based on the pyrolysis temperature and feedstock characteristics. Pyrolysis bio-oil's HHV and pH were classified into different decision classes based on the corresponding property's value. Decision rules from Case Study 1, 2 and 3 exhibit promising certainty and coverage factor while having logical sense from scientific standpoint. Based on the results, the carbon and oxygen content of the feedstock sample need to be greater than 44 % and lesser than 45 %, respectively, to produce pyrolysis bio-oil with HHV of at least 20 MJ/kg, Besides, nitrogen content of greater than 0.35 % was required to achieve bio-oil with pH value of greater than 3. Decision rules generated using this developed approached are capable to identify the required feedstock composition and pyrolysis temperature by applying the reverse engineering approach, which can then use to generate pyrolysis bio-oil with targeted range of HHV and pH via pyrolysis/co-pyrolysis process. A similar RSML approach can also be used to predict the properties of the other pyrolysis fractions. This current computational framework can be further extended to include the experimental studies to validate the results obtained from the CAMD and prediction model. Hence, the experimental methodology will be presented in the next chapter.

CHAPTER 7

ENHANCED FUEL PROPERTIES OF PYROLYSIS BIO-OIL AND DIESEL FUEL EMULSION VIA SOLVENT ADDITION AND ULTRASONIC EMULSIFICATION

7.1. Introduction

In the previous chapters, computational solvent design framework was developed to identify optimal bio-oil solvent for pyrolysis bio-oil upgradation. Nevertheless, further experimental verifications were still required to verify the applicability of the identified solvents in real-life application. Hence, in this chapter, an efficient emulsification strategy was developed to generate stable bio-oil/diesel emulsions through solvent addition and ultrasonic mixing. Fast pyrolysis of palm kernel shell (PKS) was conducted to produce PKS bio-oil at 550 ºC in a fixed bed tubular reactor. The generated PKS bio-oil was then blended with solvents identified from the developed computational solvent design framework in Chapter 4 and 5. Three solvents, namely 2-Octanol, 2-Heptanol and 2-Octanone were selected for their promising performance in terms of fuel functionality and economic. In addition, the PKS bio-oil was also blended with 5 wt. %, 10 wt. %, 15 wt. % and 20 wt. % of 2-Octanol, respectively to study on how different solvent mixing ratio affect the yield of solvent-extracted bio-oil. Next, the solvent-extracted bio-oil with 20 wt. % of 2-Octanol was processed further for emulsification with diesel fuel. Two composite surfactants (Tween 80 and Span 80) with a hydrophilic-lipophilic

balance (HLB) value of 7 was used to aid the ultrasonic emulsification process. The stability of solvent-extracted bio-oil/diesel emulsion was investigated as well. Lastly, the 2-Octanol-, 2-Heptanol- and 2-Octanone-extracted bio-oil/diesel emulsions were prepared, and their fuel properties were analysed, respectively.

7.2. Problem Statement

In recent years, bio-oil generated from fast pyrolysis of biomass feedstock has attracted increasing attention as a potential drop-in fuel. Nevertheless, the direct application of bio-oil is still limited due its poor fuel properties and instability. Therefore, downstream upgradation is still required for the pyrolysis bio-oil in order to obtain desired specifications for final applications. In the view of this, the emulsification of pyrolysis bio-oil with the conventional diesel fuel is one of the popular approaches to utilise bio-oil. However, the blending of bio-oil and diesel could be challenging as bio-oil is a complex mixture consisting of both polar and non-polar compounds. Hence, it is important to develop an efficient emulsification strategies and determine the desired mixing ratio to generate a stable bio-oil/diesel emulsion.

7.3. Materials and Methods

7.3.1. Feedstock Preparation

Euro 5 Diesel fuel was purchased from a Shell petrol station in Semenyih, Selangor, Malaysia. The PKS biomass feedstocks were supplied from Seri Ulu Langat Palm Oil Mill Sdn. Bhd., Dengkil, Selangor, Malaysia. The proximate analysis and higher heating value (HHV) of same PKS feedstock has been presented in previous work by Vasu et al. (2020). The collected PKS

was dried in an oven (Memmert, Germany) at 75 ºC for 16 hours. Then, the dried PKS was grounded into particle size of smaller than 0.2mm using an electric grinder (ormiSmart, China). The moisture content of PKS was determine by calculating the weight loss after oven drying for 16 hours at 103ºC.

On the other hand, solvents with promising performance from previous work were also selected to upgrade the PKS bio-oil quality. In Chapter 5, the solvent-oil blend with 2-Octanol and 2-Heptanol demonstrated positive performance in terms of fuel functionality, environmental and economical aspect. In addition, 2-Octanone identified from Chong et al. (2021)'s work was considered in this chapter as well. The organic phase of the crude pyrolysis PKS bio-oil derived in Asadullah et al. (2013)'s work was used as the basis in both computational studies. Their PKS bio-oil was reported to have a HHV of 19.0 MJ/kg with a moisture content of 16 wt.%.

7.3.2. Fast Pyrolysis in Fixed Bed Reactor

The fixed-bed fast pyrolysis of PKS was performed in a horizontally tubular stainless-steel reactor (Outer Diameter: 28.4mm; Length: 600mm), which was heated using a splitable mini tube furnace (BSO-1200 G, Malaysia). Figure 7.1 demonstrated the schematic diagram of the reactor configuration.

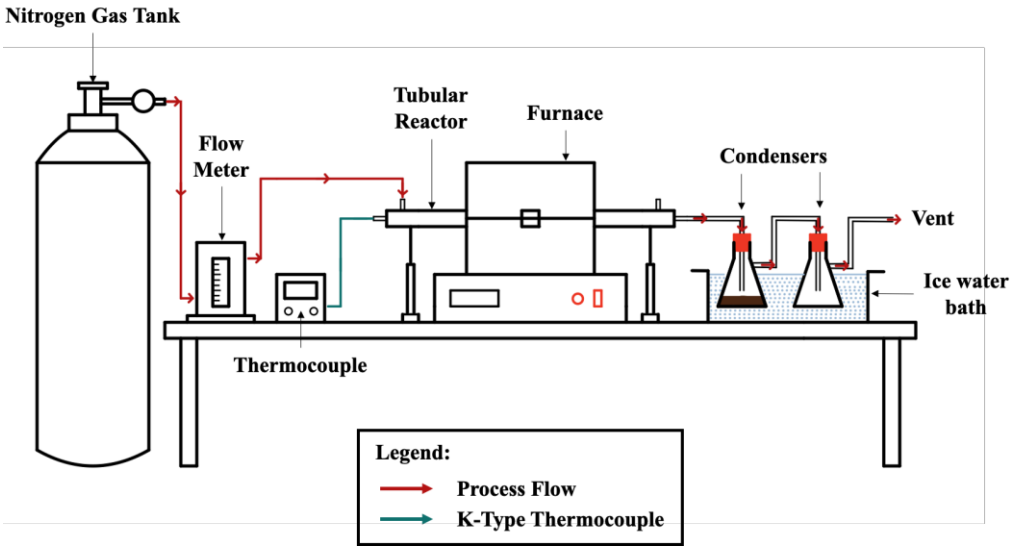


Figure 7-1 Schematic representation of pyrolysis setup

Initially, a known amount of grounded PKS biomass (around 25 g) was loaded into the tubular reactor. Glass wool was inserted at the effluent end of the reactor to prevent the solid char particles from entering the vapours. In order to create an inert atmosphere in the reactor, nitrogen gas was purged through the reactor at a flowrate of 20 – 40 cm^3/min. K-type thermocouple was used to measure the actual temperature in the reactor during the pyrolysis experiment. In order to achieve pyrolysis temperature of 550 ± 20 ºC, the tubular reactor was inserted into the furnace once it achieves the pre-set temperature at 575 ºC for sufficient heat transfer. The heating rate inside the reactor was obtained as 35.25 ± 1.7 ºC/min, indicating the occurrence of intermediate/fast pyrolysis reaction for PKS bio-oil production.

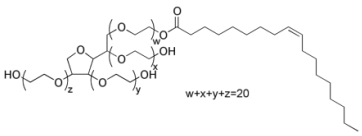
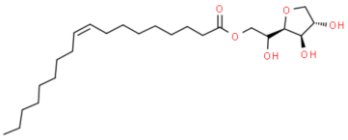
White fumes were observed within 3 to 4 minutes of reactor insertion, with reactor temperature within the range of 200 ºC to 300 ºC. The temperature of the furnace and nitrogen gas flow were maintained until no white fumes were

observed any longer. Two condensers were placed in series and submerged in ice bath for the collection of condensed liquid product, which is also known as the PKS bio-oil. The non-condensable vapour product was then vented out. The resulting bio-oil and biochar yield were calculated by measuring the weight at the end of each pyrolysis experiment. On the other hand, the gas yield was determined by subtracting the total bio-oil and biochar yield from a mass basis of 100 wt. %.

7.3.3. Preparation of Composite Surfactant

In this chapter, the non-ionic surfactants, Tween 80 and Span 80 were used to aid the ultrasonic emulsification process. The physicochemical properties of the surfactants are shown in Table 7-1 (Farooq et al., 2019; Lin et al., 2016).

Table 7-1 Properties of surfactant Tween 80 and Span 80.

Surfactant	Tween 80	Span 80
Physical Appearance	Amber liquid	Brownish yellow liquid
Chemical Structure		
Hydrophilic group	Polyethylene glycol	Sorbitan
Lipophilic group	Oleic acid	Oleic acid
HLB value	15	4.3
Molecular weight (g/mol)	1310	429
Density (g/cm^3)	1.08	0.99

The HLB value was first established by Griffin (1954) to classify non-ionic surfactants. In most cases, the HLB value is an important criteria in choosing the ideal surfactant, depending on the chemical structure and

surfactant properties (Lin et al., 2016). Reports have shown that the optimal HLB value of the surfactants for the emulsification of solvent-extracted bio-oil and diesel ranged from 4 – 8, resulting in water-in-oil (W/O) emulsions (Farooq et al., 2019; Lin et al., 2016). Hence, in this chapter, composite surfactants with HLB value of 7 were used to emulsify the solvent-extracted bio-oil/diesel emulsion, which can be obtained by mixing Tween 80 and Span 80. The specific weight percentages of Tween 80 and Span 80 can be calculated using Equation 7.1 (Griffin, 1954):

$$HLB_{Surfactant} = \frac{HLB_1 \times W_1 + HLB_2 \times W_2}{W_1 + W_2} \quad (7.1)$$

where HLB_1 and HLB_2 represent the HLB of surfactant 1 and 2 while W_1 and W_2 represent the weight percentage of surfactant 1 and 2, respectively.

7.3.4. Ultrasonic Emulsification

Initially, the PKS bio-oil was emulsified with 5, 10, 15 and 20 wt. % of 2-Octanol, respectively using ultrasonic emulsification setup (Bandelin Sonopuls, Germany). The solvent-oil mixture was then sonicated for 5 minutes (30% amplitude, 200 W, pulse mode). As a result, a two-phase immiscible solvent-oil emulsion was obtained as shown in Figure 7-2. The upper layer of the blend is the oil phase while the bottom part is the aqueous phase. Then, the upper oil layer was extracted out using the separating funnel. In this chapter, only the upper oil layer was considered and will be referred to as the solvent-extracted bio-oil in the rest of the chapter.

The yield and HHV of the solvent-extracted bio-oils with 5, 10, 15, and 20 wt. % of 2-Octanol were determined, respectively. In another experiment,

the solvent-oil emulsion of different solvents type (2-Octanol, 2-Heptanol and 2-Octanone) at 20 wt. % were also prepared to study the effect of different solvent type on the yield and properties of solvent-extracted bio-oil.

Figure 7-2 Emulsified solvent-oil blend with 20 wt. % 2-Octanol

Then, the solvent-extracted bio-oil was emulsified with the composite surfactant and diesel fuel at a predetermined ratio using the sonotrode. Three diesel-to-surfactant-to-solvent-extracted bio-oil weight ratios of 90:5:5, 85:10:5 and 80:15:5 were taken into account to test the emulsification behaviour. Lastly, the emulsions of diesel/surfactant/2-Heptanol-extracted bio-oil and diesel/surfactant/2-Octanone-extracted bio-oil were also prepared at weight ratios of 80:15:5, respectively. All the experiments were repeated twice to ascertain the reproducibility and consistency .

7.3.5. Extracted Solvent-Oil/Diesel Emulsions Stability Analysis

Accelerated aging test was conducted on all the generated solvent-extracted bio-oil/diesel emulsions. The sample emulsions were immersed in a hot water bath at 80 ºC for 24 hours, which is equivalent to approximate 6 – 12 months of storage at room temperature (Elliott et al., 2012). After the aging, the samples were cooled down in ice-cooled water to stop any further aging reaction. In addition, the solvent-extracted bio-oil/diesel emulsions were also place still

under room temperature for normal aging process. The phase separation of the emulsions was checked every 4 hours for the first 24 hours and every 24 hours for the remaining of the stability test. Phase separation was considered as an indication of instability of the solvent-extracted bio-oil/diesel emulsion. The functional groups in the aged emulsions were detected using FT-IR spectroscopy.

7.3.6. Product Characterisation

7.3.6.1. pH value

The pH values of all samples were measured using a digital benchtop pH meter (Sartorius PB-10, Germany). The pH meter was pre-calibrated using buffer solution of pH 4.0, 7.0 and 10.0.

7.3.6.2. Higher Heating Value

The heating values of all the samples were determined using bomb calorimeter (IKA C 200, Germany).

7.3.6.3. Fourier Transform Infrared (FT-IR) Spectroscopy

All the liquid and solid samples were analysed using the FT-IR spectroscopy (PerkinElmer Frontier, USA). The FT-IR spectra were recorded at wavenumber ranging from 400 – 4000 cm^{-1}, with 8 scans at a scanning resolution of 8 cm^{-1}.

7.3.6.4. Gas Chromatography-Mass Spectrometry (GC-MS)

The volatile compounds present in liquid samples were analysed using the GC-MS (PerkinElmer 680 GC/SQ8S, USA). 0.75 g of each bio-oil and solvent-extracted bio-oil samples were diluted in 50 mL of methanol.

Meanwhile, 0.75 g of diesel and solvent-extracted bio-oil/diesel emulsion samples were diluted in 50 mL of dichloromethane (DCM). The diluted solutions were then gently shaken to ensure complete mixing. Then, approximately 2.5 mL of each diluted samples was withdrawn and filtered (syringe filter, Agilent PTFE 0.45 μm) into Agilent vials. The diluted bio-oil samples were injected into a DB 1701 column (Length: 30 m; Inner diameter: 0.32 mm; Film thickness: 0.25 μm) while the diluted solvent-extracted bio-oil, diesel and solvent-extracted bio-oil/diesel emulsion samples were injected into an Elite-5MS capillary column (Length: 30 m; Inner diameter: 0.25 mm; Film thickness: 0.25 μm). The 1071 GC-column is reported to be more a more sensitive column for the analysis of alcohols, oxygenates and anhydrous sugar, which are commonly found in the crude bio-oil (Agilent, 2022). Therefore, DB-1701 column was used for the chemical characterization of PKS bio-oil. However, later studies include diesel and diesel emulsion with bio-oil, Elite-5MS capillary column is used for further GC-MS analysis (Perkin Elmer, 2022). The injector temperature was kept constant at 280 ºC with a split ratio of 50:1. The column's initial temperature was set at 40 ºC and was held for 3 minutes. The column's temperature was then increased at a constant ramping rate of 10 ºC/min to reach 250 ºC, which was then held for another 3 minutes at the final temperature. The carrier gas (helium) was maintained at a flow rate of 1.35 mL/min. In addition, the GC-MS interface was operated at 280 ºC. The compounds were identified by comparing the mass spectra with the National Institute of Standards and Technology (NIST) mass spectra library.

7.4. Results and discussion

7.4.1. Biomass and Diesel Characterisation

The physicochemical properties of the PKS biomass have great influences on the production yield and characteristic of the produced pyrolysis bio-oil. The moisture content of the biomass plays an important role in the production of pyrolysis bio-oil. High amount of moisture content (> 10 wt. %) was undesirable as it reduced the yield and HHV of the pyrolysis bio-oil generated. During the pyrolysis of PKS, the water content in the feedstock goes to the generated bio-oil, leading to an increment of water content in bio-oil and thus lowering bio-oil's HHV. In this chapter, the moisture content of the PKS biomass was identified to be 3.3 ± 0.1 wt. %. Hence, no additional actions were required to further reduce the moisture content of PKS of biomass to improve the heating rate during pyrolysis reaction. In addition, the HHV of PKS biomass was obtained as 19.4 MJ/kg, indicating its potential to be used as a feedstock for energy generation.

On the other hand, the pH value of the diesel was measured to be 8.1 ± 0.1, which is consistence with the typical values (5.5 – 8) reported by Craig Hartman (2020). Meanwhile, HHV value of 45.1 MJ/kg was obtained for the diesel. In addition, FT-IR and GC-MS analyses were also conducted to identify the functional groups and compounds present in the diesel. Forty-four compounds were identified via GC-MS, which mainly can be grouped as paraffinic hydrocarbons, naphthenic hydrocarbons, and fatty acid methyl esters. Table 7-2 shows the major compound groups along with their respective total

peak area percentage. The detailed lists of compounds and FT-IR spectra of diesel can be found in the Appendix D.

Table 7-2 GC-MS results of main components from diesel

Compound groups	Total peak area (%)
Paraffinic hydrocarbons	70.8
Naphthenic hydrocarbons	5.6
Fatty acid methyl esters	23.6

7.4.2. Pyrolysis Product Yield and Properties

The PKS bio-oil yield of 41.7 ± 3.4 wt. % on PKS weight basis was achieved in this chapter. Meanwhile, the gas and biochar yield were obtained as 33.0 ± 2.4 wt. % and 25.3 ± 1.5 wt. % on PKS weight basis, respectively. This results were in line with the yields reported in Chong et al. (2017) and Vasu et al. (2020)'s work. On the other hand, the HHV value of PKS bio-oil and biochar was determined to be 15.6 MJ/kg and 27.89 MJ/kg, respectively. Nevertheless, lower HHV of PKS bio-oil in current work as compared to Asadullah et al. (2013)'s study was due to the difference in water content of bio-oil. In Asadullah et al. (2013)'s work, only the HHV of the crude bio-oil's organic phase was considered. Meanwhile, whole bio-oil consisting of both aqueous and organic phase was considered in the current study, and thus lower HHV of the PKS bio-oil. In addition, The pH value of PKS bio-oil was found to be 2.35, which is rather acidic in nature due to the presence of organic acids such as formic and acetic acid.

A total of twenty-two compounds from PKS bio-oil were identified by GC-MS and those were grouped as carbonyl compounds, furan, nitrogen containing compounds as well as phenol and phenolic derivatives. Table 7-3

summarised the major compounds groups of PKS bio-oil along with their respective total peak area percentage. Based on Table 7-3, the total peak area of phenolic compounds was the highest where phenol alone contributed 38.2 % to the total peak area percentage of PKS bio-oil. The complete list of compounds present in the PKS bio-oil and their respective peak area percentage were included in the Appendix D.

Table 7-3 Major compound groups of PKS bio-oil

Compound group	Total peak area (%)
Carbonyl group	16.2
N-containing group	22.1
Furan group	0.5
Phenol group	59.9

7.4.3. Effect of solvent ratio on solvent-extracted bio-oil properties

In order to study the effect of different solvent ratio on the properties of solvent-extracted bio-oil, four 2-Octanol to PKS bio-oil ratios were proposed in this work. The resulted yield, HHV and pH were analysed to characterise the solvent-extracted bio-oil obtained from different mixing ratio, as shown in Table 7-4.

Table 7-4 Yields and properties of 2-Octanol-extracted bio-oil at different 2-Octanol : PKS bio-oil ratio

Solvent : PKS bio-oil	5 : 95	10 : 90	15 : 85	20 : 80
Yield (%)	21.0 ± 0.4	27.2 ± 0.2	32.4 ± 0.7	37.9 ± 1.3
HHV (MJ/kg)	26.8	28.3	29.9	31.9
pH	2.3 ± 0.1	2.6 ± 0.1	2.8 ± 0.1	2.9 ± 0.1

The 2-Octanol-extracted bio-oil's yields obtained in Table 7-4 were calculated on 2-Octanol/PKS bio-oil blend's weight basis . From Table 7-4 it

was observed that the yield of 2-Octanol-extracted bio-oil increases as the ratio of solvent increases in the 2-Octanol/PKS bio-oil blend. At 10 wt. % of 2-Octanol addition, 2-Octanol-extracted bio-oil's yield of 27.2 ± 0.2 % was observed, which is about 6.2 % increment as compared to the yield obtained from 5 wt.% of 2-Octanol addition. However, at 15 wt. % and 20 wt. % of 2-Octanol addition, only 5.2 % and 5.4 % increment in the 2-Octanol-extracted bio-oil's yields were observed, respectively. This value may be due to the increment of 2-Octanol in the solvent-oil blending, which was also 5 wt. % on 2-Octanol/PKS bio-oil mixture weight basis. In other word, further increment on the solvent ratio in the blend would not increase the extraction of bio-oil content in the 2-Octanol-extracted bio-oil.

As expected, as the ratio of 2-Octanol increases in the 2-Octanol/PKS bio-oil mixture, higher HHV of the resulted 2-Octanol-extracted bio-oil can be achieved. The generated 2-Octanol-extracted bio-oil at 20 wt. % of 2-Octanol had achieved a HHV value of 31.9 MJ/kg, which is approximately 100 % increase as compared to the HHV value of initial PKS bio-oil, which was only 15.6 MJ/kg. This can be explained with the removal of water content from the PKS bio-oil as aqueous bio-oil, and thus increasing the HHV value of the organic-rich 2-Octanol-extracted bio-oil. Other than that, the addition of 2-Octanol (40.7 MJ/kg) as solvent also contributed to the high HHV value of the 2-Octanol-extracted bio-oil.

Similar trends and FT-IR peaks were observed from the FT-IR spectra of 2-Octanol-extracted bio-oil with 5, 10, 15 and 20 wt. % of 2-Octanol addition. Hence, only the FT-IR spectra of 2-Octanol-extracted bio-oil with 5 and 20 wt. %

of 2-Octanol addition, along with the FT-IR spectra of PKS bio-oil were presented in Figure 7-3. The FT-IR results of PKS bio-oil were similar to those reported by Chong et al. (2019). The presence of a transmittance broad band in the range of 3200 – 3550 cm^{-1} was assigned to the stretching vibration of O-H bonds, indicating the presence of phenol, alcohol, and water in the samples (Hassan et al., 2009). Generally, increase in the peak intensity was often associated with increased amount (per unit volume) of the functional group correlated with the particular molecular bond. Here, the broad band from PKS bio-oil demonstrated higher intensity as compared to the other two samples. This may be due to the high water content present in PKS bio-oil, which contribute to the high concentration of O-H bond. On the other hand, the peaks at 2960 cm^{-1}, 2930 cm^{-1} and 2858 cm^{-1} can be accredited to the stretching vibration of the C-H bonds, signifying the presence of alkane groups in the samples. Peaks with higher intensity within this range was observed for the 20 wt.% 2-Octanol-extracted bio-oil sample, followed by 5 wt.% 2-Octanol extracted bio-oil and lastly PKS bio-oil with the lowest intensity. This may be due to the presence of long carbon chain in 2-Octanol, which contributed to the high concentration of C-H bond in the sample. This can be further confirmed with the C-H bonds deformation vibrations at peaks 1366 cm^{-1} and 1464 cm^{-1}.

Furthermore, the peak at the 1704 cm^{-1} can be accounted for the presence of carbonyl group (C=O) in ketones, aldehydes, and carboxylic acids. Here, the PKS bio-oil was noticed to demonstrate peak with higher intensity as it contained higher amount of carboxylic acid such as acetic and formic acids. The presence of aromatics in samples was supported by the C=C stretching vibration

(in ring) at the peak 1594 cm^{-1}. Various peaks between 970 cm^{-1} and 1250 cm^{-1} can be ascribed to the stretching vibrations of C-O bond due to the presence of carboxylic acid, phenol and alcohol groups (Coates, 2006). The high concentration of acetic and formic acid in the PKS bio-oil contributed to the higher intensity peak at 1232 cm^{-1}. Aside from that, the transmittance peaks in the 900 cm^{-1} to 700 cm^{-1} range indicate possibility of the presence of mono, polycyclic, and substituted aromatic groups (Chong et al., 2019). Similar peak intensity was observed for all three samples. This phenomenon indicated that most of the aromatic groups present in PKS bio-oil was extracted into the solvent-extracted bio-oil.

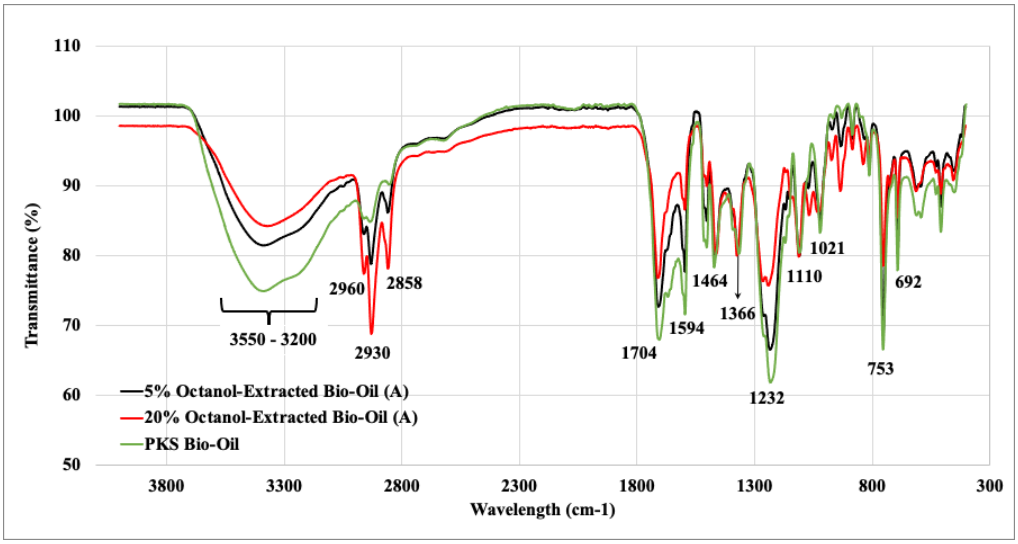


Figure 7-3 FT-IR spectra of PKS bio-oil, 5wt.% and 20wt. % 2-Octanol-extracted bio-oil

The compounds present in the solvent-extracted and aqueous bio-oils from different 2-Octanol mixing ratios were identified via GC-MS and grouped as phenols and phenolic derivatives, carbonyl compounds and furans. Table 7-5 shows the total peak area percentages of these major compound groups and 2-Octanol, respectively. The detailed lists of the compounds for these samples were provided in Appendix D.

Table 7-5 Major compound groups of solvent-extracted and aqueous bio-oil from different solvent addition ratios

Samples	Phenols (area %)	Carbonyls (area %)	Furans (area %)	2-Octanol (area %)
Aqueous bio-oil:				
5 wt. % 2-Octanol	19.2	64.3	11.5	5.0
10 wt. % 2-Octanol	17.4	67.5	10.0	4.9
15 wt. % 2-Octanol	13.7	71.4	10.7	4.2
20 wt. % 2-Octanol	10.6	73.4	11.8	4.2
2-Octanol-extracted bio-oil:				
5 wt. % 2-Octanol	32.9	7.7	5.5	50.4
10 wt. % 2-Octanol	22.2	6.3	4.4	65.0
15 wt. % 2-Octanol	15.8	4.9	3.8	74.5
20 wt. % 2-Octanol	14.4	4.4	3.0	77.1

Phenols and its derivatives such as creosol, catechol and isoeugenol were identified as the second largest compound groups in 2-Octanol-extracted bio-oils, where 2-Octanol is the dominant component. On the other hand, it is noticed that the carbonyl compounds like acetic acid, formic acid, butanone, and propanoic acid constituted the major proportions in aqueous bio-oil. The acetic acid was regarded as the main component in the carbonyl group as it possessed the highest peak area percentage as compared to other compounds. According to Table 7-5, the peak area percentage of carbonyl groups in aqueous bio-oil was the highest at 20 wt. % of 2-Octanol, followed by 15, 10 and 5 wt.% of 2-Octanol. Nevertheless, the carbonyl groups demonstrated smaller peak area percentage in 2-Octanol-extracted bio-oils, as compared to the aqueous bio-oil. This has proven that most of the carbonyl groups including the acetic and formic acid were left behind in the aqueous bio-oil during the solvent extraction, thus resulting in a less-acidic solvent-extracted bio-oil with higher HHV value.

7.4.4. Effect of Solvent Type on Properties of Bio-Oil

Other than the solvent ratio, the effect of different solvent type on the properties of PKS bio-oil were also studied in this chapter. The resulting yield, HHV and pH of the solvent-extracted bio-oil were analysed to evaluate the performance of 2-Octanol, 2-Heptanol and 2-Octanone as solvent, respectively. Table 7-6 shows the yields and properties of solvent-extracted bio-oil with different type of solvent. All three 2-Octanol-, 2-Heptanol-, and 2-Octanone-extracted bio-oils demonstrated similar performances in terms of pH values. However, it was noticed that the 2-Octanone-extracted bio-oil has the lowest solvent-extracted bio-oil yield, followed by 2-Heptanol-extracted bio-oil and 2-Octanol-extracted bio-oil. The variation of extraction results was mainly caused by different solvent polarities and solubilities. The water solubility of 2-Heptanol was the highest at 3.3 g/L, followed by 2-Octanol at 1.1 g/L. On the other hand, the 2-Octanone has the lowest water solubility which is at 0.09 g/L. The higher the solvent's water solubility, the larger amount of solvent can be dissolved in the aqueous bio-oil, and thus resulting in lower yield of the solvent-extracted bio-oil. Theoretically, 2-Octanone should have the highest yield as it is the least water-soluble solvent compared to 2-Octanol and 2-Heptanol. However, from Table 7-6, it is observed that 2-Octanone has the lowest solvent-extracted bio-oil yield at 30.7 ± 0.3 %. This may be due to the absence of polar compounds such as acetic and formic acid in the 2-Octanone-extracted bio-oil, and thus resulting in a lower yield. Nevertheless, the HHV of the 2-Octanone-extracted bio-oil was comparable to the 2-Octanol-extracted bio-oil. On the other hand, a small fraction of 2-Heptanol was dissolved in the aqueous phase

due to its higher water solubility, hence resulting in a lower solvent-extracted bio-oil yield and HHV.

In addition, the performance of the solvent-extracted bio-oil was also compared with the results obtained from computational models reported in Chapter 5. In Chapter 5, the Gibbs energy and tangent plot for all the identified solvent-oil blend showed that the final blend should demonstrate single homogenous phase (Figure 5-2 and Figure 5-3). However, in this chapter, two distinct separated phase, which is the upper oil layer and bottom aqueous layer was obtained from the solvent-oil blend. This deviation from the computational modelling may be explained with the difference in solvent-oil blend mixing mechanism. In Chapter 5, the solvent-oil blend design problem was modelled under the assumption that the blend was mixed using mechanical agitation. Meanwhile, in this chapter, the solvent-oil blend was emulsified using the ultrasonic emulsification approach. Although this observation was unexpected, however the outcome is favourable. With the removal of water content as aqueous phase, the solvent-extracted bio-oil can now achieve higher HHV with lower solvent mixing ratio. Referring to Table 5-7 in Chapter 5, results from case study 2 shows that solvent-oil blend with HHV of 29.2 MJ/kg and 28.49 MJ/kg can be obtained with the addition of 47 wt. % 2-Octanol and 45 wt. % of 2-Heptanol, respectively. However, in this chapter, solvent-extracted bio-oil with HHV of 31.9 MJ/kg and 29.8 MJ/kg can be achieved with the addition of only 20 wt. % 2-Octanol and 2-Heptanol, which is around 55 % decrease in the amount of solvent required.

Table 7-6 Solvent-extracted bio-oil yields and their properties with different type of solvent

Solvent	2-Octanol	2-Heptanol	2-Octanone
Yield (%)	37.9 ± 1.3	35.4 ± 0.3	30.7 ± 0.3
HHV (MJ/kg)	31.9	29.8	31.5
pH	2.9 ± 0.1	2.8 ± 0.1	2.8 ± 0.1

The FT-IR spectra of the 2-Octanol-, 2-Heptanol-, and 2-Octanone-extracted bio-oils were shown in Figure 7-4. The transmittance peaks for 2-Octanol- and 2-Heptanol-extracted bio-oil were alike to that of the peaks obtained from FT-IR of PKS bio-oil (Figure 7-3). However, the peaks for 2-Octanone-extracted bio-oil demonstrated slight differences as compared to the other two solvent-extracted bio-oils. From Figure 7-4, less intense broad band within the range of 3200 cm^{-1} to 3550 cm^{-1} was observed for the 2-Octanone-extracted bio-oil. The alcohol groups from 2-Octanol and 2-Heptanol contributed to the O-H bonds in the solvent-extracted bio-oil, and thus resulting in a higher intensity broad band within this range. On the other hand, peak with higher intensity at 1709 cm^{-1} was noticed for 2-Octanone-extracted bio-oil. This may be due to the presence of C=O bond in ketone groups, which can be dominantly found in the 2-Octanone solvent.

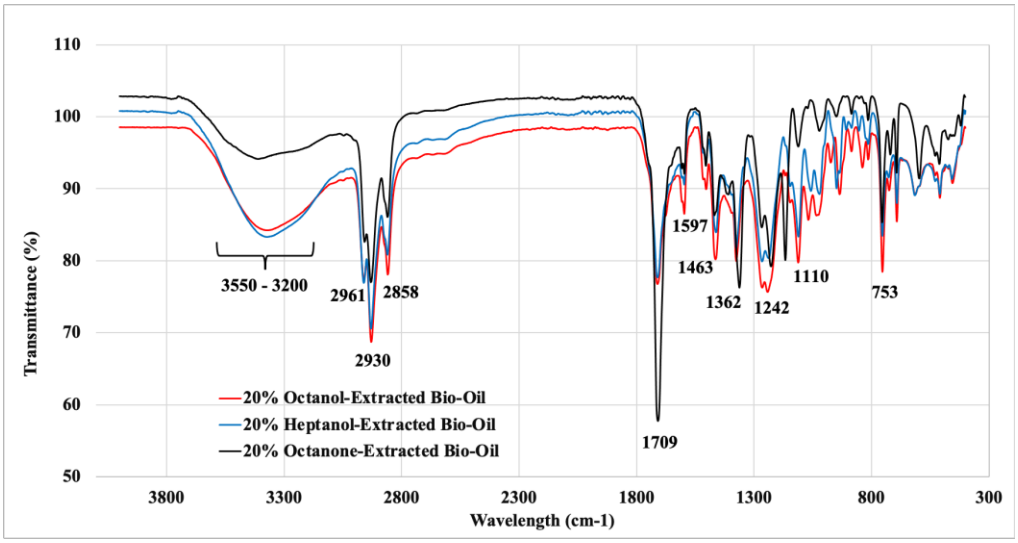


Figure 7-4 FT-IR spectra of solvent-extracted bio-oil with 20 wt. % of 2-Octanol, 2-Heptanol and 2-Octanone, respectively.

7.4.5. Effect of Surfactant Ratio on the Stability of Solvent-Extracted Bio-Oil/Diesel Emulsion

Based on the results from Section 7.4.3 and 7.4.4, solvent-extracted bio-oil with 20 wt. % of 2-Octanol demonstrated the most promising performance in terms of fuel functionality and extraction yield. Hence, the solvent-extracted bio-oil with 20 wt. % of 2-Octanol will be used for the emulsification with diesel fuel, with the aid of composite surfactant. Here, the diesel to surfactant ratio was varied to study how it affects the stability of the solvent-extracted bio-oil/diesel emulsion. Table 7-7 shows the HHV and pH values of the generated solvent-extracted bio-oil/diesel emulsions at diesel to surfactant ratio of 90 : 5, 85 : 10 and 80 : 15, respectively. The results of the emulsions' HHV were in accordance with those reported by Chong et al. (2017) and Farooq et al. (2019). According to Table 7-7, slight increment in the HHV can be noticed with the increasing of diesel content. On the other hand, all three solvent-extracted bio-

oil/diesel emulsions at different diesel to surfactant ratio demonstrated similar pH values.

Table 7-7 The properties of solvent-extracted bio-oil/diesel emulsion at different diesel : surfactant ratio

Diesel : Surfactant	80 : 15	85 : 10	90 : 5
HHV (MJ/kg)	42.7	42.6	44.0
pH	8.0 ± 0.1	8.0 ± 0.1	8.0 ± 0.1

The compounds in the solvent-extracted bio-oil/diesel emulsions were analysed using the GC-MS. The identified compounds were grouped into paraffinic hydrocarbons, naphthenic hydrocarbons and fatty acid methyl esters. In addition, the peak area percentages of 2-Octanol and surfactant's compounds were shown in Table 7-8 as well. The compounds identify under paraffinic hydrocarbons, naphthenic hydrocarbons and fatty acid methyl esters groups for all three samples were comparable to the diesel's GC-MS result. Compounds related to the surfactant such as the methylene chloride and 2-chloro-2-methyl-butane can be found in the solvent-extracted bio-oil/diesel emulsions with diesel to surfactant ratio of 80 : 15 and 85 : 10. However, these compounds were not detected in the emulsion with diesel to surfactant ratio of 90 : 5. This may be due to the lower amount of surfactant being added to the emulsions.

Table 7-8 Major compound groups of solvent-extracted bio-oil/diesel emulsion with different diesel to surfactant ratios

Compounds	Diesel:Surfactant		
	80:15	85:15	90:5
Paraffinic Hydrocarbons (area %)	54.9	56.0	64.9
Naphthenic Hydrocarbons (area %)	9.6	9.9	7.3
Fatty acid methyl esters (area %)	20.6	19.2	22.1
2-Octanol (area %)	5.0	5.1	4.8
Surfactant (area %)	9.4	8.9	0.00

Accelerated aging was conducted on the emulsified fuel to assess their thermal stability at different diesel to surfactant ratio. Prior to the accelerated aging process, all the solvent-extracted bio-oil/diesel emulsions demonstrated homogenous single phase. Nevertheless, phase separation was observed for the fuel emulsions with diesel to surfactant ratio of 90 : 5 and 85 : 10 after heating in water bath at 80 ºC for 24 hours. In addition, the changes in the appearance of the fuel emulsions were observed as well. Solvent-extracted bio-oil/diesel emulsions with diesel to surfactant of 90 : 5 has the lighter yellowish appearance, indicating higher degree of solvent-extracted bio-oil's settlement at the bottom of the bottle. At lower surfactant concentration, the emulsion is not stable due to the agglomeration of the oil droplet, thus resulting in phase separation (Jiang and Ellis, 2010). In contrast, the solvent-extracted bio-oil/diesel emulsions with diesel to surfactant of 85 : 15 were appeared to be in a darker brown, with no phase separation observed.

Other than physical observations, the FT-IR spectroscopy was also used to examine the homogeneity and stability of solvent-extracted bio-oil diesel emulsions of different surfactant ratio after accelerated aging via hot water bath. Samples were collected from the middle fraction of aged-solvent-extracted bio-oil diesel emulsions for FT-IR examination. Figure 7-5 illustrated the FT-IR spectra of the 2-Octanol-extracted bio-oil/diesel emulsions before and after accelerated aging, with diesel/surfactant/solvent-extracted bio-oil mixing ratio of 90:5:5, 85:10:5 and 80:15:5, respectively. As shown clearly in Figure 7-5, broad band was observed within the 3550 cm^{-1} to 3200 cm^{-1} range for all three of the emulsified fuels before aging. However, no broad band within the

abovementioned range was observed for the FT-IR spectra of the emulsified fuels with diesel to surfactant ratio of 90 : 5 and 85 : 10 after aging. This indicated the absence of alcohol groups (O-H bonds) in the middle fraction of the emulsified fuels, which was one of the major components of the 2-Octanol-extracted bio-oil. The result from FT-IR further confirmed on the phase separation of the 2-Octanol-extracted bio-oil from the emulsified fuel. On the other hand, the spectra of the emulsified fuels at diesel to surfactant ratio of 80 : 15 before and after accelerated aging overlapped with each other, indicating stable and homogenous 2-Octanol-extracted bio-oil/diesel emulsions with diesel to surfactant ratio of 80 : 15. Consequently, the diesel to surfactant ratio of 80 : 15 was selected as the optimum diesel to surfactant ratio in the emulsification process.

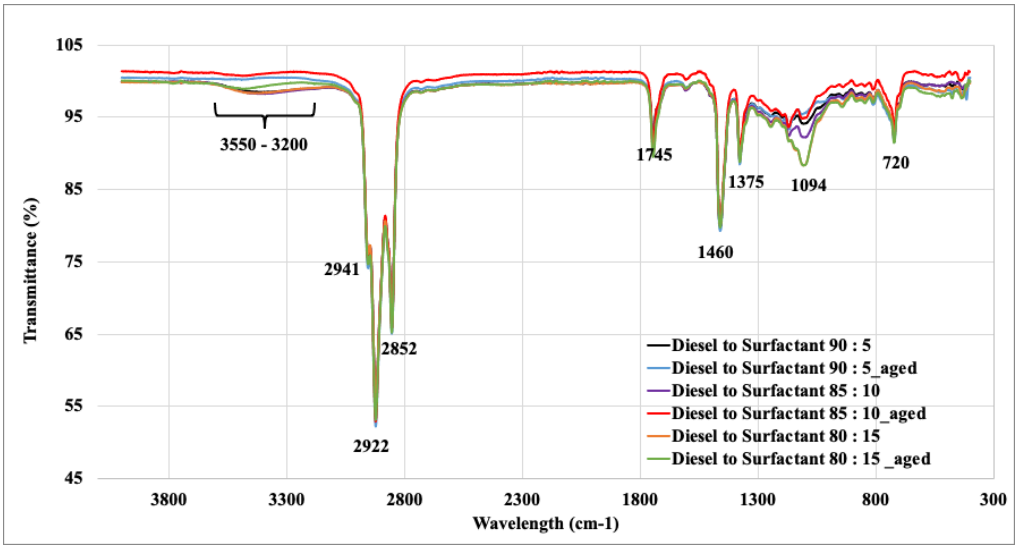


Figure 7-5 FT-IR spectra of solvent-extracted bio-oil/diesel emulsion before and after accelerated aging at different diesel to surfactant ratio

However, in this work, a higher amount of surfactants at 15 wt. % were required to generate a stable emulsion as compared to the other reported work.

For instance, a stable emulsion up to 40 days could be obtained with diesel, ether-extracted bio-oil and surfactant content of 90, 5, 5 wt. % respectively (Farooq et al., 2019). In addition, the diesel-bio-oil emulsion with 6 wt. % of surfactant (HLB of 7) and 4 % of co-surfactant are reported to stable up to 384 hours with emulsifying time of 5 minutes and shear velocity of 15,000 rpm at 40 ºC (Liu et al., 2021).

A study by Jiang and Ellis, (2010) reported that the stability of bio-oil/diesel emulsion depends greatly on the mixing temperature. The emulsion's stability increases as the mixing temperature increases due to the decreases in viscosity and interfacial tension (Ostberg et al., 1995). Nevertheless, the stability decreases dramatically once the mixing temperature exceed 30 ºC. At temperature after 30 ºC, the droplets in the emulsion tend to coagulate and thus leads to the destabilisation of emulsion. In this study, the cavitation from the ultrasonication emulsification created energy which in turn created heat in the emulsion. Temperature up to 80 ºC was observed for the solvent-extracted bio-oil/diesel emulsion during the ultrasonic emulsification. The high emulsion temperature degrades the interfacial adsorption of the surfactant, thereby increasing the tendency of collision and coalescences and resulted in deterioration in the stability of the emulsion (Chen and Tao, 2005). Thus, extra measurements are required to keep the emulsion at the optimal temperature during the ultrasonic emulsification process.

7.4.6. Emulsification of different solvent-extracted bio-oil with diesel fuel

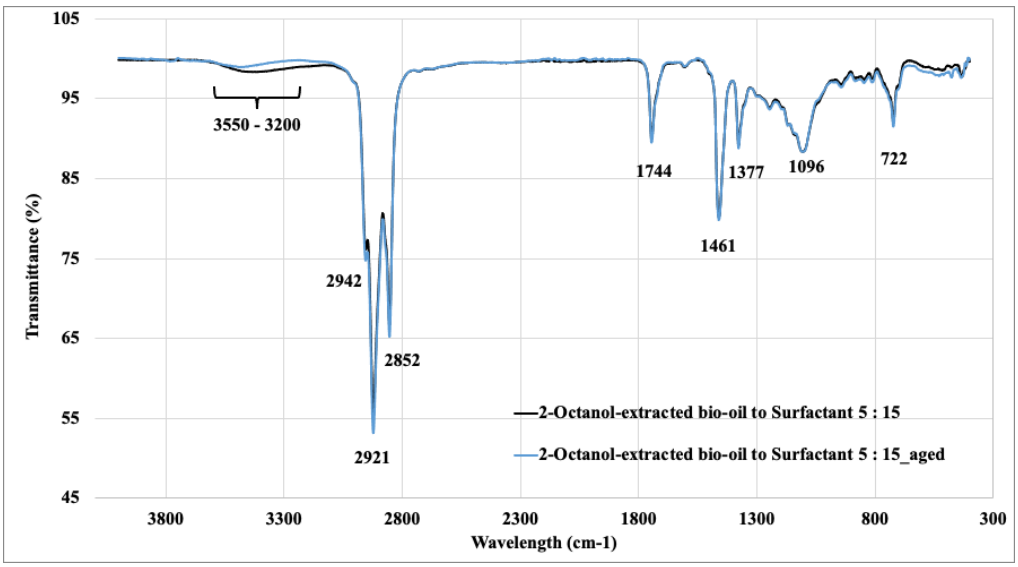
The fuel properties of 2-Octanol-, 2-Heptanol- and 2-Octanone-extracted bio-oils were reported as well to verify the feasibility and applicability of these three solvents in upgrading the PKS bio-oil. Solvent-extracted bio-oils with 20 wt. % of 2-Octanol, 2-Heptanol and 2-Octanone were emulsified with diesel fuel at a diesel : surfactant : solvent-extracted bio-oil ratio of 80 : 15 : 5. The resulted HHV and pH of these emulsified fuel were shown in Table 7-9. Here, the compounds and its respective GC peak area percentage identified for different solvent-extracted bio-oil were comparable to the GC-MS results reported in Section 7.4.5. The complete lists of components identified can be found in Appendix D.

Table 7-9 The properties of solvent-extracted bio-oil/diesel emulsion at with different solvents.

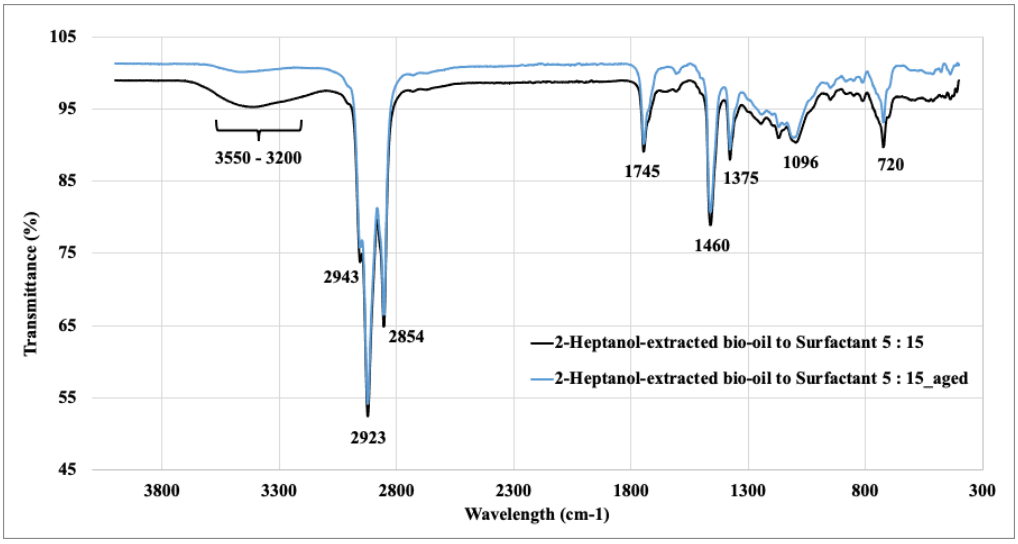
Solvent-extracted bio-oil	2-Octanol	2-Heptanol	2-Octanone
HHV (MJ/kg)	42.7	42.8	42.9
pH	8.0 ± 0.1	8.0 ± 0.1	8.0 ± 0.1

From Table 7-9, it is observed that the all three 2-Octanol-, 2-Heptanol- and 2-Octanone-extracted bio-oil/diesel emulsions demonstrated comparable HHV and pH values. In addition, the stability and homogeneity of the fuel emulsions were studied as well by undergoing accelerated aging. A small amount of bio-oil sediments was noticed for 2-Heptanol-extracted bio-oil/diesel emulsions after the accelerated aging process. However, no phase separation or bio-oil sediment was noticed in both 2-Octanol- and 2-Octanone-extracted bio-oil/diesel emulsion after accelerated aging process.

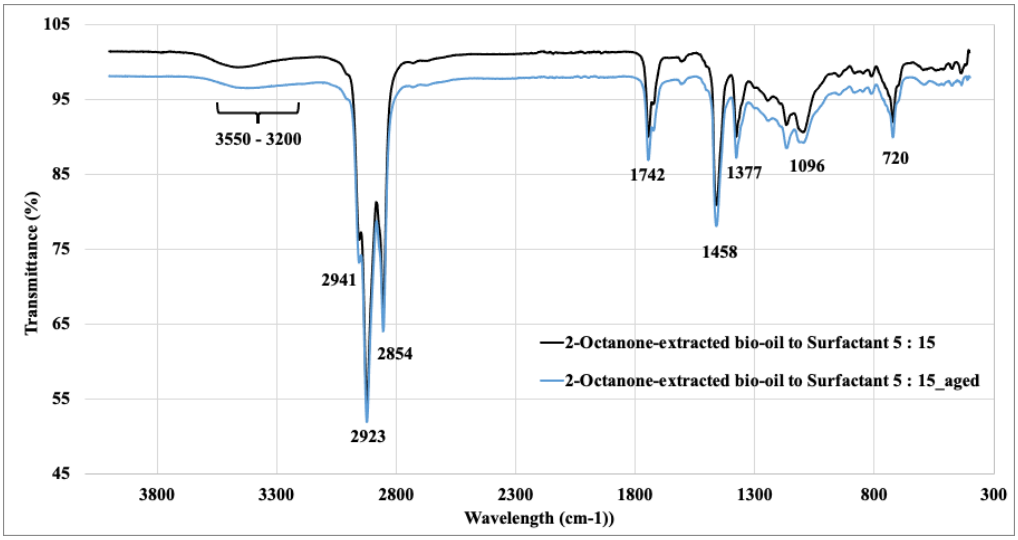
Figure 7-6 (a), (b) and (c) demonstrated the FT-IR spectra obtained for 2-Octanol-, 2-Heptanol-, and 2-Octanone-extracted bio-oil/diesel emulsion before and after the accelerated aging process. The FT-IR spectra of 2-Octanol-extracted bio-oil/diesel emulsion before and after the accelerated aging from Figure 7-6 (a) overlap with each other, and hence indicating the emulsion is stable and demonstrated homogeneous single phase. Similar trend was also observed for the FT-IR spectra of 2-Octanone-extracted bio-oil/diesel emulsion as shown in Figure 7-6 (c). On the other hand, the FT-IR spectra of 2-Heptanol-extracted bio-oil/diesel emulsions demonstrated slight differences before and after the accelerated aging process. From Figure 7-6 (b), a lower intensity of broad band correlated to the O-H bonds was observed for the emulsions after the accelerated aging process, indicating the absence of alcohol groups in the emulsions. This further proved that the separation and sedimentation of 2-Heptanol-extracted bio-oil in the emulsions after accelerated aging.



(a)



(b)



(c)

Figure 7-6 FT-IR spectra of (a) 2-Octanol- (b) 2-Heptanol- (c)2-Octanone-extracted bio-oil/diesel emulsion before and after accelerated aging

7.5. Conclusion

In this chapter, ultrasonic emulsification of bio-oil and diesel, in the presence of solvents and composite surfactants (Tween 80 and Span 80) was conducted. Pyrolysis bio-oil was generated from the fast pyrolysis of palm kernel shell in a fixed bed tubular reactor at a temperature of 550 ºC. PKS bio-

oil production yield of 41.7 ± 3.4 wt. % was achieved in this chapter. The HHV of the PKS bio-oil was obtained as 15.6 MJ/kg. Thus, further upgradation on the PKS bio-oil were required for downstream applications. In the view of this, promising solvents identified from the Chapter 4 and 5, which include the 2-Octanol, 2-Heptanol and 2-Octanone were selected for bio-oil upgradation purposes. Initially, emulsification of PKS bio-oil and solvent were conducted with PKS bio-oil to 2-Octanol ratios at 5, 10, 15 and 20 wt. % solvent. In addition, the effect of different solvent type on the properties of solvent-extracted bio-oil were also studied by emulsifying PKS bio-oil with 20 wt. % of 2-Octanol, 2-Heptanol and 2-Octanone. At 20 wt. % of 2-Octanol addition, the of solvent-extracted bio-oil's yield of 37.9 ± 1.3 % and HHV of 31.9 MJ/kg was reported. Then, solvent-extracted bio-oil with 20 wt. % of 2-Octanol was emulsified with diesel. Two composite surfactants (Tween 80 and Span 80) with HLB value of 7 was used to aid the ultrasonic emulsification process. The emulsification of solvent-extracted bio-oil with diesel were conducted by varying the diesel to surfactant ratio. The optimum ratio was found to be 80 : 15 : 5 (diesel : surfactant : solvent-extracted bio-oil). The optimised emulsion samples demonstrated promising quality with HHV of 44.0 and pH value of 8.0 ± 0.1. In addition, no phase separation was observed even after undergoing the accelerated aging process at 80 ºC for 24 hours.

CHAPTER 8

CONCLUSION

This chapter summarises all the research works conducted for this thesis and provides an overview of some of the possible research directions for the future. Based on the identified research gaps, four research scopes were established in this work. The detailed discussion of the research scopes, research objectives and research methodology were presented in Chapter 3. The proposed research scopes had provided several significant contributions which aided in the upgradation of the pyrolysis bio-oil's properties. Nonetheless, there are still a number of opportunities to extend and improve the developed methodology in this area of interest.

8.1. Achievements

The overall achievements of the research described in this thesis include the development of a combined computational and experimental approach in upgrading the properties of pyrolysis bio-oil. In this research work, pyrolysis bio-oil solvent which fulfils the pre-defined target properties can be designed using the computer-aided molecular design (CAMD) techniques. Later, the research work was further extended to simultaneously optimise the performance of solvent-oil blend in terms of economic and fuel properties using the multi-objective optimisation (MOO) approach. Moreover, understanding that the pyrolysis operating condition and the characterisation of biomass feedstock are often related to the final properties of the pyrolysis bio-oil, hence, a rough-set

machine learning (RSML) approach was employed to study the correlations between these attributes. Lastly, an experimental work on the emulsification of pyrolysis bio-oil and diesel with the aid of surfactant and solvents identified from the previous research work was conducted. The major contributions of each research scope are summarised in the following paragraphs.

In Chapter 4, a systematic multi-stage CAMD framework was developed to design bio-oil solvents that can fulfil multiple desired target properties while possessing low environmental impact. Signature-based molecular design techniques have been utilised to include property prediction models of various indexes in designing the bio-oil solvent. However, in most of the existing CAMD formulations, signatures with higher height had to be excluded from the formulation due to the combinatorial nature of molecular signature descriptors. Thus, the consistency rules developed in this research scope addressed this limitation by removing irrelevant molecular signature at a lower height from the building block sets. This was able to keep the CAMD problem in a manageable size by eliminating signatures that do not fulfil the consistency rule at different height. With the consistency rules, it is possible to consider all promising building blocks with higher-order contribution from group contribution (GC) and topological indexes (TI) based property prediction models. As a result, solvent candidate with optimum product functionality and environmental properties could be identified.

Nevertheless, the research framework developed in first scope only guarantees that the identified solvent candidates possess optimal physical functionality. Other than the solvents' physical functionality, there is also a need

for the designed solvent-oil blend to demonstrate promising economic performance to be competitive with the conventional diesel fuel. Therefore, a MOO approach for bio-oil solvent design has been presented in Chapter 5 to simultaneously consider and optimise the fuel functionality and profitability of the solvent-oil blend. However, in this case, the relative importance of each objective is fuzzy and contradict with each other in nature. Thus, an optimisation approach was developed by adapting the fuzzy optimisation algorithm into the signature based CAMD problem. Through this approach, optimal solvent candidates were identified by achieving near optimality for both objective functions without compromising on each other.

The production and characterisation of pyrolysis bio-oil are usually complicated, time consuming and costly. In addition, trial-and-error processes were often required to select the most suitable biomass feedstock composition and operating condition to generate pyrolysis bio-oil of desired properties. Thus, a rule-based predictive model via RSML algorithms was developed in this Chapter 6 to estimate the pyrolysis bio-oil's properties based on the pyrolysis operating condition and biomass feedstock characteristics. Being inherently interpretable, transparent and straightforward decision rules were generated from the developed RSML model. Based on the conducted case studies, the generated decision rules demonstrated logical sense from a scientific standpoint. With these decision rules, the ideal biomass feedstock compositions and pyrolysis operating conditions were identified via reverse engineering approach, to generate pyrolysis bio-oil within the targeted property range.

In the first three scopes, the upgradations of pyrolysis bio-oil's properties were conducted via different computational approaches. However, the results obtained from computational studies might deviates from the experimental observations. Thus, an experimental methodology on pyrolysis bio-oil upgradation via solvent addition was presented in Chapter 7. In the proposed approach, the bio-oil solvents identified from Chapter 4 and 5 were blended into the pyrolysis bio-oil/diesel emulsion via ultrasonification technology. Thus, by utilising the proposed approach, stable pyrolysis bio-oil/diesel emulsions were generated with stability more than 15 days. To conclude, this thesis presented various novel computational and experimental approaches to upgrade the fuel and environmental characteristics of pyrolysis bio-oil while maximising its economic performance.

8.2. Future works

Other than the abovementioned achievements, some challenges and limitations were also encountered in the presented methodologies. The potential future works for further improvement and enhancement of the presented thesis can be summarised as below:

8.2.1. Incorporation of sustainability aspects into CAMD solvent design framework

Driven by the legislation and growing awareness towards sustainability issues, designing green solvents has become an increasingly important area of research. The incorporation of the sustainability idea into the solvent design, manufacturing, and value chain management to minimise the resource utilisation and unfavourable environmental impact is pivotal. Other than the

performance and economics evaluation, there is also a need to incorporate the environmental, health and safety hazard assessment and life cycle values into the CAMD design. A broadly accepted approach for the sustainability assessment is the life cycle assessment, where the emissions associated with all life cycle stages of the products was considered. The complete life cycle impact can be included into the CAMD framework as an index for sustainability of the designed solvent. Nevertheless, having a trade-off solution with high sustainability characteristics may come at a significant cost to the solvent's profit and quality. In the view of this, there is the need to develop a systematic solvent design strategy to optimise the minimum deviation permitted from the best possible solvent's quality and profit, while having a huge improvement in its' sustainability characteristics.

8.2.2. Development of CAMD framework that considers the reaction mechanisms between solvents and pyrolysis bio-oil

Apart from the fuel functionality and environmental aspects, the reaction mechanisms between the solvents and bio-oil components should also be considered to avoid undesired side-reaction. When blending with alcohol solvents, several complex and simultaneous reaction equilibria occurred in the solvent-oil blend (Zhang et al., 2013). These potential reactions could be phenol alkylation, olefins hydration, esterification, isomerisation, cracking, oligomerisation, etc (Chin et al., 2021). For example, most of the reactions involving alcohol solvents produce water as by-product, which will reduce the heating value of the solvent-oil blend. However, blending with alkene solvents would result in higher viscosity due to the absence of reactions that generate

water. Hence, the use of alkene-alcohol solvent mixture may be beneficial by compensating each other to meet the constraints on heating value and viscosity simultaneously. In addition, a detailed molecular dynamic simulation should be conducted to account for the changes of bond energies during the interaction of solvent candidates with constituents inside pyrolysis bio-oil.

8.2.3. Consideration of pyrolysis process parameters for prediction of bio-oil properties via RSML approach

Other than the biomass feedstock characterisation and pyrolysis operating temperature, the process parameters also strongly affect the yield and quality of the pyrolysis bio-oil due to the various chemical reactions that are favoured in each condition. These main parameters of pyrolysis process include the biomass particle size, pyrolysis heating rate, gas flow rate, holding time, type of catalyst used, operation mode etc. The performance and accuracy of the RSML prediction model can be further improved by including the above-mentioned process parameters as attributes. In addition, the type of reactor used in the pyrolysis processes also affect the bio-oil production yield (Guedes et al., 2018). Further research should also be carried out to predict other pyrolysis bio-oil characteristics such as viscosity, oxygen content and bio-oil yield to optimise the pyrolysis bio-oil's quality and product yield.

8.2.4. Production of pyrolysis bio-oil via catalytic pyrolysis

Aside from the non-catalytic fast pyrolysis bio-oil production route, catalytic fast pyrolysis is another option to generate high quality pyrolysis bio-oil. The limitations faced in the thermal pyrolysis could be addressed via catalytic pyrolysis. With the addition of catalyst, the activation energy can be

lowered in order for the pyrolysis reaction to take place. This will also lower the requirement on pyrolysis operating temperature, therefore reducing energy consumption and process cost (Hafeez et al., 2019). Furthermore, the catalysed pyrolysis bio-oil often demonstrates lower oxygen content and increased hydrogen to carbon ratio. As a result, higher energy content can be observed in the pyrolysis bio-oil. The low oxygen content also reduced the acidic component present in pyrolysis bio-oil and thus a less corrosive nature as compared to non-catalytic pyrolysis oil (Dolah et al., 2021). From the techno-economic perspective, however, there hasn't been much research on low-cost catalyst with high efficiency and lifetime. The design of an optimal catalyst for the pyrolysis of bio-oil could be challenging, where both the lifetime and the selectivity of the designed catalyst towards a particular product become important in the design of the reaction set up (Fadillah et al., 2021).

To conclude, these suggested future works are important in bringing a significant advancement in this area of research. The design of green solvent for pyrolysis bio-oil upgradation is possible by including the sustainability aspects into the solvent design framework. Furthermore, the incorporation of the potential reaction mechanisms between the designed solvent and bio-oil compounds in CAMD could avoid the occurrence of undesired side-reaction upon blending. In addition, the main pyrolysis process parameters should be taken into account in the development of the RSML prediction model to improve its performance and prediction accuracy. Last but not least, catalytic pyrolysis process could be performed to generate high quality pyrolysis bio-oil with lesser energy consumption and process cost.

REFERENCES

Abnisa, F., Arami-Niya, A., Daud, W.M.A.W., Sahu, J.N., 2013a. Characterization of Bio-oil and Bio-char from Pyrolysis of Palm Oil Wastes. Bioenergy Res 6, 830–840. https://doi.org/10.1007/s12155-013-9313-8

Abnisa, F., Arami-Niya, A., Wan Daud, W.M.A., Sahu, J.N., Noor, I.M., 2013b. Utilization of oil palm tree residues to produce bio-oil and bio-char via pyrolysis. Energy Convers Manag 76, 1073–1082. https://doi.org/10.1016/j.enconman.2013.08.038

Abnisa, F., Daud, W.M.A.W., Husin, W.N.W., Sahu, J.N., 2011a. Utilization possibilities of palm shell as a source of biomass energy in Malaysia by producing bio-oil in pyrolysis process. Biomass Bioenergy 35, 1863–1872. https://doi.org/10.1016/j.biombioe.2011.01.033

Abnisa, F., Wan Daud, W.M.A., Sahu, J.N., 2011b. Optimization and characterization studies on bio-oil production from palm shell by pyrolysis using response surface methodology. Biomass Bioenergy 35, 3604–3616. https://doi.org/10.1016/j.biombioe.2011.05.011

Achenie, L.E.K., Gani, R., Venkatasubramanian, V., 2003. Computer Aided Molecular Design: Theory and Practice, 1st ed. Elsevier B.V., Amsterdam.

Adewale, P., Dumont, M.-J., Ngadi, M., 2015. Recent trends of biodiesel production from animal fat wastes and associated production techniques.

Renewable and Sustainable Energy Reviews 45, 574–588. https://doi.org/10.1016/j.rser.2015.02.039

Agblevor, F.A., Besler, S., Wiselogel, A.E., 1995. Fast Pyrolysis of Stored Biomass Feedstocks. Energy & Fuels 9, 635–640. https://doi.org/10.1021/ef00052a010

Agilent, 2022. Standard Polysiloxane GC Columns: DB-1701 Columns [WWW Document]. Agilent. URL https://www.agilent.com/en/product/gc-columns/standard-polysiloxane-gc-columns/db-1701-columns (accessed 9.22.22).

Ahmad, A.L., Yasin, N.H.M., Derek, C.J.C., Lim, J.K., 2011. Microalgae as a sustainable energy source for biodiesel production: A review. Renewable and Sustainable Energy Reviews 15, 584–593. https://doi.org/10.1016/j.rser.2010.09.018

Ahmed, A., Abu Bakar, M.S., Azad, A.K., Sukri, R.S., Phusunti, N., 2018. Intermediate pyrolysis of Acacia cincinnata and Acacia holosericea species for bio-oil and biochar production. Energy Convers Manag 176, 393–408. https://doi.org/10.1016/j.enconman.2018.09.041

Akbari, M., Asadi, P., Besharati Givi, M.K., Khodabandehlouie, G., 2014. Artificial neural network and optimization, in: Advances in Friction-Stir Welding and Processing. Elsevier, pp. 543–599. https://doi.org/10.1533/9780857094551.543

Ali, N., Saleem, M., Shahzad, K., Hussain, S., Chughtai, A., 2016. Effect of operating parameters on production of bio-oil from fast pyrolysis of maize

stalk in bubbling fluidized bed reactor. Polish Journal of Chemical Technology 18, 88–96. https://doi.org/10.1515/pjct-2016-0053

Alleman, T.L., McCormick, R.L., 2016. Biodiesel Handling and Use Guide, Biodiesel Handling and Use Guide. US.

Alvarez, J., Lopez, G., Amutio, M., Artetxe, M., Barbarias, I., Arregi, A., Bilbao, J., Olazar, M., 2016. Characterization of the bio-oil obtained by fast pyrolysis of sewage sludge in a conical spouted bed reactor. Fuel Processing Technology 149, 169–175. https://doi.org/10.1016/j.fuproc.2016.04.015

Amutio, M., Lopez, G., Artetxe, M., Elordi, G., Olazar, M., Bilbao, J., 2012. Influence of temperature on biomass pyrolysis in a conical spouted bed reactor. Resour Conserv Recycl 59, 23–31. https://doi.org/10.1016/j.resconrec.2011.04.002

Ansari, K.B., Banerjee, A., Danish, Mohd., Hassan, S.Z., Sahayaraj, D. v., Khan, M.S., Phan, T.T.N., Trinh, Q.T., 2022. State-of-the-art practices to upgrade biomass fast pyrolysis derived bio-oil, in: Innovations in Thermochemical Technologies for Biofuel Processing. Elsevier, pp. 115–147. https://doi.org/10.1016/B978-0-323-85586-0.00003-2

Anuradha, J., Karthik, S., Tripathy, B.K., 2011. Classifcation and rule extraction using LEM1 algorithm for diagnosis of liver disease and its type. Computer Science and Engneering 38, 4370–4374.

Arel, I., Liu, C., Urbanik, T., Kohls, A.G., 2010. Reinforcement learning-based multi-agent system for network traffic signal control. IET Intelligent Transport Systems 4, 128. https://doi.org/10.1049/iet-its.2009.0070

Arora, J.S., 2012. Multi-objective Optimum Design Concepts and Methods, in: Introduction to Optimum Design. Elsevier, pp. 657–679. https://doi.org/10.1016/B978-0-12-381375-6.00017-6

Asadullah, M., Ab Rasid, N.S., Kadir, S.A.S.A., Azdarpour, A., 2013. Production and detailed characterization of bio-oil from fast pyrolysis of palm kernel shell. Biomass Bioenergy 59, 316–324. https://doi.org/10.1016/j.biombioe.2013.08.037

Asadullah, M., Anisur Rahman, M., Mohsin Ali, M., Abdul Motin, M., Borhanus Sultan, M., Robiul Alam, M., Sahedur Rahman, M., 2008. Jute stick pyrolysis for bio-oil production in fluidized bed reactor. Bioresource Technol 99, 44–50. https://doi.org/10.1016/j.biortech.2006.12.002

Asadullah, M., Rahman, M.A., Ali, M.M., Rahman, M.S., Motin, M.A., Sultan, M.B., Alam, M.R., 2007. Production of bio-oil from fixed bed pyrolysis of bagasse. Fuel 86, 2514–2520. https://doi.org/10.1016/j.fuel.2007.02.007

Ascher, S., Watson, I., You, S., 2022. Machine learning methods for modelling the gasification and pyrolysis of biomass and waste. Renewable and Sustainable Energy Reviews 155, 111902. https://doi.org/10.1016/j.rser.2021.111902

Austin, N.D., Sahinidis, N. v., Trahan, D.W., 2016. Computer-aided molecular design: An introduction and review of tools, applications, and solution

techniques. Chemical Engineering Research and Design 116, 2–26. https://doi.org/10.1016/j.cherd.2016.10.014

Bagajewicz, M., Hill, S., Robben, A., Lopez, H., Sanders, M., Sposato, E., Baade, C., Manora, S., Hey Coradin, J., 2011. Product design in price-competitive markets: A case study of a skin moisturizing lotion. AIChE Journal 57, 160–177. https://doi.org/10.1002/aic.12242

Bagajewicz, M.J., 2007. On the role of microeconomics, planning, and finances in product design. AIChE Journal 53, 3155–3170. https://doi.org/10.1002/aic.11332

Baker, L.E., Pierce, A.C., Luks, K.D., 1982. Gibbs Energy Analysis of Phase Equilibria. Society of Petroleum Engineers Journal 22, 731–742. https://doi.org/10.2118/9806-PA

Baljet, M., 2009. FAME/Airport Aviation Fueling. International Airport Review.

Banković-Ilić, I.B., Stojković, I.J., Stamenković, O.S., Veljkovic, V.B., Hung, Y.-T., 2014. Waste animal fats as feedstocks for biodiesel production. Renewable and Sustainable Energy Reviews 32, 238–254. https://doi.org/10.1016/j.rser.2014.01.038

Barik, D., 2019. Energy Extraction From Toxic Waste Originating From Food Processing Industries, in: Energy from Toxic Organic Waste for Heat and Power Generation. Elsevier, pp. 17–42. https://doi.org/10.1016/B978-0-08-102528-4.00003-1

Basak, S.C., Gute, B.D., 1997. Characterization of Molecular Structures Using Topological Indices. SAR QSAR Environ Res 7, 1–21. https://doi.org/10.1080/10629369708039122

Bellman, R.E., Zadeh, L.A., 1970. Decision-Making in a Fuzzy Environment. Manage Sci 17, B-141-B-164. https://doi.org/10.1287/mnsc.17.4.B141

Ben, H., Wu, F., Wu, Z., Han, G., Jiang, W., Ragauskas, A.J., 2019. A Comprehensive Characterization of Pyrolysis Oil from Softwood Barks. Polymers (Basel) 11, 1387. https://doi.org/10.3390/polym11091387

Benson, S.W., Buss, J.H., 1958. Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties. J Chem Phys 29, 546–572. https://doi.org/10.1063/1.1744539

Benson, S.W., Cruickshank, F.R., Golden, D.M., Haugen, G.R., O'Neal, H.E., Rodgers, A.S., Shaw, R., Walsh, R., 1969. Additivity rules for the estimation of thermochemical properties. Chem Rev 69, 279–324. https://doi.org/10.1021/cr60259a002

Bertoli, C., D'Alessio, J., del Giacomo, N., Lazzaro, M., Massoli, P., Moccia, V., 2000. Running Light-Duty DI Diesel Engines with Wood Pyrolysis Oil. https://doi.org/10.4271/2000-01-2975

Bharath, G., Hai, A., Rambabu, K., Banat, F., Jayaraman, R., Taher, H., Bastidas-Oyanedel, J.-R., Ashraf, M.T., Schmidt, J.E., 2020. Systematic production and characterization of pyrolysis-oil from date tree wastes for bio-fuel applications. Biomass Bioenergy 135, 105523. https://doi.org/10.1016/j.biombioe.2020.105523

Boubacar Laougé, Z., Çığgın, A.S., Merdun, H., 2020. Optimization and characterization of bio-oil from fast pyrolysis of Pearl Millet and Sida cordifolia L. by using response surface methodology. Fuel 274, 117842. https://doi.org/10.1016/j.fuel.2020.117842

Boyd, J.C., Millership, J.S., Woolfson, A.D., 1982. The Relationship between Molecular Connectivity and Partition Coefficients. Journal of Pharmacy and Pharmacology 34, 364–366. https://doi.org/10.1111/j.2042-7158.1982.tb04730.x

Bridgwater, A.V., 2012. Review of fast pyrolysis of biomass and product upgrading. Biomass Bioenergy 38, 68–94. https://doi.org/10.1016/j.biombioe.2011.01.048

Bridgwater, A.V., 2009. Technical and Economic Assessment of Thermal Processes for Biofuels.

Brown, W.M., Martin, S., Rintoul, M.D., Faulon, J.L., 2006. Designing novel polymers with targeted properties using the signature molecular descriptor. J Chem Inf Model 46, 826–835. https://doi.org/10.1021/ci0504521

Bu, X., Rao, J., Xu, C.-Z., 2009. A Reinforcement Learning Approach to Online Web Systems Auto-configuration, in: 2009 29th IEEE International Conference on Distributed Computing Systems. IEEE, pp. 2–11. https://doi.org/10.1109/ICDCS.2009.76

Buschmann, A.H., Camus, C., Infante, J., Neori, A., Israel, Á., Hernández-González, M.C., Pereda, S. v., Gomez-Pinchetti, J.L., Golberg, A., Tadmor-Shalev, N., Critchley, A.T., 2017. Seaweed production: overview

of the global state of exploitation, farming and emerging research activity. Eur J Phycol 52, 391–406. https://doi.org/10.1080/09670262.2017.1365175

BYJU'S, 2022. Sonication: Definition, Working Principle, Applications and Methods [WWW Document]. Sonication. URL https://byjus.com/physics/sonication/ (accessed 8.28.22).

Byrne, F.P., Jin, S., Paggiola, G., Petchey, T.H.M., Clark, J.H., Farmer, T.J., Hunt, A.J., Robert McElroy, C., Sherwood, J., 2016. Tools and techniques for solvent selection: green solvent selection guides. Sustainable Chemical Processes 4, 7. https://doi.org/10.1186/s40508-016-0051-z

C. E. Goering, A. W. Schwab, M. J. Daugherty, E. H. Pryde, A. J. Heakin, 1982. Fuel Properties of Eleven Vegetable Oils. Transactions of the ASAE 25, 1472–1477. https://doi.org/10.13031/2013.33748

Cai, N., Zhang, H., Nie, J., Deng, Y., Baeyens, J., 2020. Biochar from Biomass Slow Pyrolysis. IOP Conf Ser Earth Environ Sci 586, 012001. https://doi.org/10.1088/1755-1315/586/1/012001

Canabarro, N., Soares, J.F., Anchieta, C.G., Kelling, C.S., Mazutti, M.A., 2013. Thermochemical processes for biofuels production from biomass. Sustainable Chemical Processes 1, 22. https://doi.org/10.1186/2043-7129-1-22

Chandrasekaran, B., Abed, S.N., Al-Attraqchi, O., Kuche, K., Tekade, R.K., 2018. Computer-Aided Prediction of Pharmacokinetic (ADMET)

Properties, in: Dosage Form Design Parameters. Elsevier, pp. 731–755. https://doi.org/10.1016/B978-0-12-814421-3.00021-X

Chapelle, O., Scholkopf, B., Zien, A., 2006. Semi-Supervised Learning. MIT Press, United Kingdom.

Charturvedi, P., Daniel, A.K., Khusboo, K., 2017. Concepts of Rough Set Theory and its Applications in Decision Making Process. International Journal of Advanced Research in Computer and Communication Engineering 6, 43–46. https://doi.org/10.17148/IJARCCE

Chemmangattuvalappil, N.G., 2020. Development of solvent design methodologies using computer-aided molecular design tools. Curr Opin Chem Eng 27, 51–59. https://doi.org/10.1016/j.coche.2019.11.005

Chemmangattuvalappil, N.G., Eden, M.R., 2013. A Novel Methodology for Property-Based Molecular Design Using Multiple Topological Indices. Ind Eng Chem Res 52, 7090–7103. https://doi.org/10.1021/ie302516v

Chemmangattuvalappil, N.G., Solvason, C.C., Bommareddy, S., Eden, M.R., 2010. Reverse problem formulation approach to molecular design using property operators based on signature descriptors. Comput Chem Eng 34, 2062–2071. https://doi.org/10.1016/j.compchemeng.2010.07.009

Chen, D., Li, Y., Cen, K., Luo, M., Li, H., Lu, B., 2016. Pyrolysis polygeneration of poplar wood: Effect of heating rate and pyrolysis temperature. Bioresour Technol 218, 780–788. https://doi.org/10.1016/j.biortech.2016.07.049

Chen, G., Tao, D., 2005. An experimental study of stability of oil–water emulsion. Fuel Processing Technology 86, 499–508. https://doi.org/10.1016/j.fuproc.2004.03.010

Chen, L., Yoshikawa, K., 2018. Bio-oil upgrading by cracking in two-stage heated reactors. AIMS Energy 6, 203–315. https://doi.org/10.3934/energy.2018.1.203

Cheng, Y.S., Fung, K.Y., Ng, K.M., Wibowo, C., 2016. Economic analysis in product design — A case study of a TCM dietary supplement. Chin J Chem Eng 24, 202–214. https://doi.org/10.1016/j.cjche.2015.06.014

Chin, H.H., Mah, A.X.Y., Neoh, J.Q., Aboagwa, O.A., Thangalazhy-Gopakumar, S., Chemmangattuvalappil, N.G., 2021. Reactive and non-reactive solvents as bio-oil blends: a computer-aided molecular design approach. Biomass Convers Biorefin 11, 1633–1649. https://doi.org/10.1007/s13399-019-00565-4

Chin, M., 2011. Biofuels in Malaysia: An analysis of the legal and institutional framework. Bogor, Indonesia.

China Petroleum and Chemical Industry Federation, 2021. China Market Price: Monthly Avg: Organic Chemical Material: Propan 2 ol, Isopropyl Alcohol 99% [WWW Document]. CEIC.

Chong, J.W., Ng, L.Y., Aboagwa, O.A., Thangalazhy-Gopakumar, S., Muthoosamy, K., Chemmangattuvalappil, N.G., 2021. Computer-Aided Framework for the Design of Optimal Bio-Oil/Solvent Blend with

Economic Considerations. Processes 9, 2159. https://doi.org/10.3390/pr9122159

Chong, Y.Y., Thangalazhy-Gopakumar, S., Ng, H.K., Ganesan, P.B., Gan, S., Lee, L.Y., Manickavel, V.S.A.R., Ong, C.M., al HinaI, H.S.R., 2017. Emulsification of Bio-Oil and Diesel. Chem Eng Trans 56, 1801–1806.

Chong, Y.Y., Thangalazhy-Gopakumar, S., Ng, H.K., Lee, L.Y., Gan, S., 2019. Effect of oxide catalysts on the properties of bio-oil from in-situ catalytic pyrolysis of palm empty fruit bunch fiber. J Environ Manage 247, 38–45. https://doi.org/10.1016/j.jenvman.2019.06.049

Chuck, C.J., McManus, M., Allen, M.J., Singh, S., 2016. Feedstocks for Aviation Biofuels, in: Biofuels for Aviation. Elsevier, pp. 17–34. https://doi.org/10.1016/B978-0-12-804568-8.00002-0

Clemente, J., 2015. Why Biofuels Can't Replace Oil. Forbes 1.

Coates, J., 2006. Interpretation of Infrared Spectra, A Practical Approach, in: Encyclopedia of Analytical Chemistry. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/9780470027318.a5606

Constantinou, L., Gani, R., 1994. New group contribution method for estimating properties of pure compounds. AIChE Journal 40, 1697–1710. https://doi.org/10.1002/aic.690401011

Constantinou, L., Gani, R., O'Connell, J.P., 1995. Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution

method. Fluid Phase Equilib 103, 11–22. https://doi.org/10.1016/0378-3812(94)02593-P

Constantinou, L., Prickett, S.E., Mavrovouniotis, M.L., 1993. Estimation of thermodynamic and physical properties of acyclic hydrocarbons using the ABC approach and conjugation operators. Ind Eng Chem Res 32, 1734–1746. https://doi.org/10.1021/ie00020a030

Conte, E., Martinho, A., Matos, H.A., Gani, R., 2008. Combined Group-Contribution and Atom Connectivity Index-Based Methods for Estimation of Surface Tension and Viscosity. Ind Eng Chem Res 47, 7940–7954. https://doi.org/10.1021/ie071572w

Craig Hartman, 2020. Diesel Fuel Maintenance [WWW Document]. Energy Management Corporation. URL https://emcsolutions.com/2020/09/23/diesel-fuel-maintenance/ (accessed 9.23.22).

Cunliffe, A.M., Williams, P.T., 1998. Composition of oils derived from the batch pyrolysis of tyres. J Anal Appl Pyrolysis 44, 131–152. https://doi.org/10.1016/S0165-2370(97)00085-5

Dahl, C.A., 2012. Measuring global gasoline and diesel price and income elasticities. Energy Policy 41, 2–13. https://doi.org/10.1016/j.enpol.2010.11.055

Dahmen, N., Dinjus, E., Kolb, T., Arnold, U., Leibold, H., Stahl, R., 2012. State of the art of the bioliq® process for synthetic biofuels production. Environ Prog Sustain Energy 31, 176–181. https://doi.org/10.1002/ep.10624

Dai, L., Wang, Y., Liu, Y., Ruan, R., Yu, Z., Jiang, L., 2019. Comparative study on characteristics of the bio-oil from microwave-assisted pyrolysis of lignocellulose and triacylglycerol. Science of The Total Environment 659, 95–100. https://doi.org/10.1016/j.scitotenv.2018.12.241

Das, I., Dennis, J.E., 1998. Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems. SIAM Journal on Optimization 8, 631–657. https://doi.org/10.1137/S1052623496307510

Das, S., Dey, A., Pal, A., Roy, N., 2015. Applications of Artificial Intelligence in Machine Learning: Review and Prospect. Int J Comput Appl 115, 31–41. https://doi.org/10.5120/20182-2402

Dearden, J.C., 2017. The Use of Topological Indices in QSAR and QSPR Modeling. pp. 57–88. https://doi.org/10.1007/978-3-319-56850-8_2

Delucchi, M.A., 2003. A Lifecycle Emissions Model (LEM): Lifecycle Emissions from Transportation Fuels, Motor Vehicles, Transportation Modes, Electricity Use, Heating and Cooking Fuels, and Materials. California.

Dey, A., 2016. Machine Learning Algorithms: A Review. International Journal of Computer Science and Information Technologies 7, 1174–1179.

Dhall, D., Kaur, R., Juneja, M., 2020. Machine Learning: A Review of the Algorithms and Its Applications. pp. 47–63. https://doi.org/10.1007/978-3-030-29407-6_5

Dhar, S.A., Sakib, T.U., Hilary, L.N., 2020. Effects of pyrolysis temperature on production and physicochemical characterization of biochar derived from coconut fiber biomass through slow pyrolysis process. Biomass Convers Biorefin. https://doi.org/10.1007/s13399-020-01116-y

Dimian, A.C., Bildea, C.S., Kiss, A.A., 2014. Chemical Product Design. pp. 489–523. https://doi.org/10.1016/B978-0-444-62700-1.00012-7

Ding, T., Li, S., Xie, J., Song, W., Yao, J., Lin, W., 2012. Rapid Pyrolysis of Wheat Straw in a Bench-Scale Circulating Fluidized-Bed Downer Reactor. Chem Eng Technol 35, 2170–2176. https://doi.org/10.1002/ceat.201200140

Diwekar, U., 2008. Introduction to Applied Optimization. Springer US, Boston, MA. https://doi.org/10.1007/978-0-387-76635-5

Dolah, R., Karnik, R., Hamdan, H., 2021. A Comprehensive Review on Biofuels from Oil Palm Empty Bunch (EFB): Current Status, Potential, Barriers and Way Forward. Sustainability 13, 10210. https://doi.org/10.3390/su131810210

Domalski, E.S., Hearing, E.D., 1988. Estimation of the Thermodynamic Properties of Hydrocarbons at 298.15 K. J Phys Chem Ref Data 17, 1637–1678. https://doi.org/10.1063/1.555814

DOSM, 2019. Selected Agricultural Indicators, Malaysia, 2019, Departmen of Statistics Malaysia Offical Portal. Kuala Lumpur.

Dutta, A., 2019. Rough Set Theory - An Introduction [WWW Document]. GeeksforGeeks. URL https://www.geeksforgeeks.org/rough-set-theory-an-introduction/

Edwin Raj, R., Robert Kennedy, Z., Pillai, B.C., 2013. Optimization of process parameters in flash pyrolysis of waste tyres to liquid and gaseous fuel in a fluidized bed reactor. Energy Convers Manag 67, 145–151. https://doi.org/10.1016/j.enconman.2012.11.012

EESI, 2021. Fossil Fuels [WWW Document]. Environmental and Energy Study Institute. URL https://www.eesi.org/topics/fossil-fuels/description (accessed 9.3.22).

EIA, 2022a. Short-Term Energy Outlook. Washington, D.C.

EIA, 2022b. Biomass explained [WWW Document]. U.S. Energy Information Administration. URL https://www.eia.gov/energyexplained/biomass/ (accessed 9.4.22).

Elliott, D.C., Oasmaa, A., Meier, D., Preto, F., Bridgwater, A. v., 2012. Results of the IEA Round Robin on Viscosity and Aging of Fast Pyrolysis Bio-oils: Long-Term Tests and Repeatability. Energy & Fuels 26, 7362–7366. https://doi.org/10.1021/ef301607v

Encinar, J.M., Sánchez, N., Martínez, G., García, L., 2011. Study of biodiesel production from animal fats with high free fatty acid content. Bioresour Technol 102, 10907–10914. https://doi.org/10.1016/j.biortech.2011.09.068

Energy, 2020. Waste-derived Pyrolysis Oil Market to Close in on USD 500 Mn Valuation by 2029. Energy Industry Review.

EUBIA, 2021. Pyrolysis [WWW Document]. European Biomass Industry Association.

Fadillah, G., Fatimah, I., Sahroni, I., Musawwa, M.M., Mahlia, T.M.I., Muraza, O., 2021. Recent Progress in Low-Cost Catalysts for Pyrolysis of Plastic Waste to Fuels. Catalysts 11, 837. https://doi.org/10.3390/catal11070837

Fahmi, R., Bridgwater, A.V., Donnison, I., Yates, N., Jones, J.M., 2008. The effect of lignin and inorganic species in biomass on pyrolysis oil yields, quality and stability. Fuel 87, 1230–1240. https://doi.org/10.1016/j.fuel.2007.07.026

Farooq, A., Lee, H.W., Jae, J., Kwon, E.E., Park, Y.-K., 2020. Emulsification characteristics of ether extracted pyrolysis-oil in diesel using various combinations of emulsifiers (Span 80, Atlox 4916 and Zephrym PD3315) in double reactor system. Environ Res 184, 109267. https://doi.org/10.1016/j.envres.2020.109267

Farooq, A., Shafaghat, H., Jae, J., Jung, S.-C., Park, Y.-K., 2019. Enhanced stability of bio-oil and diesel fuel emulsion using Span 80 and Tween 60 emulsifiers. J Environ Manage 231, 694–700. https://doi.org/10.1016/j.jenvman.2018.10.098

Faulon, J.-L., Brown, W.M., Martin, S., 2005. Reverse engineering chemical structures from molecular descriptors: how many solutions? J Comput Aided Mol Des 19, 637–650. https://doi.org/10.1007/s10822-005-9007-1

Faulon, J.-L., Churchwell, C.J., Visco, D.P., 2003a. The Signature Molecular Descriptor. 2. Enumerating Molecules from Their Extended Valence Sequences. J Chem Inf Comput Sci 43, 721–734. https://doi.org/10.1021/ci020346o

Faulon, J.-L., Visco, D.P., Pophale, R.S., 2003b. The Signature Molecular Descriptor. 1. Using Extended Valence Sequences in QSAR and QSPR Studies. J Chem Inf Comput Sci 43, 707–720. https://doi.org/10.1021/ci020345w

Fermoso, J., Pizarro, P., Coronado, J.M., Serrano, D.P., 2017. Advanced biofuels production by upgrading of pyrolysis bio-oil. Wiley Interdiscip Rev Energy Environ 6, e245. https://doi.org/10.1002/wene.245

Fishburn, P.C., 1967. Method of Estimating Additives Utilities. Manage Sci 13, 435–453.

FM Carpanini, 1998. TECHNICAL REPORT No. 74: QSARs in the Assessment of the Environmental Fate and Effects of Chemicals. Brussels, Belgium.

Fonts, I., Kuoppala, E., Oasmaa, A., 2009. Physicochemical Properties of Product Liquid from Pyrolysis of Sewage Sludge. Energy & Fuels 23, 4121–4128. https://doi.org/10.1021/ef900300n

Fredenslund, A., Jones, R.L., Prausnitz, J.M., 1975. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. AIChE Journal 21, 1086–1099. https://doi.org/10.1002/aic.690210607

Fung, K.Y., Ng, K.M., Zhang, L., Gani, R., 2016. A grand model for chemical product design. Comput Chem Eng 91, 15–27. https://doi.org/10.1016/j.compchemeng.2016.03.009

Gani, R., Achenie, L.E.K., Venkatasubramanian, V., 2003. Introduction to CAMD. pp. 3–21. https://doi.org/10.1016/S1570-7946(03)80003-2

Gani, R., Nielsen, B., Fredenslund, A., 1991. A group contribution approach to computer-aided molecular design. AIChE Journal 37, 1318–1332. https://doi.org/10.1002/aic.690370905

Geetha Bai, R., Muthoosamy, K., Shipton, F.N., Manickam, S., 2017. Acoustic cavitation induced generation of stabilizer-free, extremely stable reduced graphene oxide nanodispersion for efficient delivery of paclitaxel in cancer cells. Ultrason Sonochem 36, 129–138. https://doi.org/10.1016/j.ultsonch.2016.11.021

Gertig, C., Leonhard, K., Bardow, A., 2020. Computer-aided molecular and processes design based on quantum chemistry: current status and future prospects. Curr Opin Chem Eng 27, 89–97. https://doi.org/10.1016/j.coche.2019.11.007

Ghanbarzadeh-Shams, M., Ghasemy Yaghin, R., Sadeghi, A.H., 2022. A hybrid fuzzy multi-objective model for carpet production planning with reverse logistics under uncertainty. Socioecon Plann Sci 83, 101344. https://doi.org/10.1016/j.seps.2022.101344

Global Petrol Prices, 2021. Diesel prices, liter, 27-Sep-2021 [WWW Document]. GlobalPetrolPrices.com.

Gnanaprakasam, A., Sivakumar, V.M., Surendhar, A., Thirumarimurugan, M., Kannadasan, T., 2013. Recent Strategy of Biodiesel Production from Waste Cooking Oil and Process Influencing Parameters: A Review. Journal of Energy 2013, 1–10. https://doi.org/10.1155/2013/926392

González, J.F., Encinar, J.M., Canito, J.L., Sabio, E., Chacón, M., 2003. Pyrolysis of cherry stones: energy uses of the different fractions and kinetic study. J Anal Appl Pyrolysis 67, 165–190. https://doi.org/10.1016/S0165-2370(02)00060-8

González, J.F., Ramiro, A., González-García, C.M., Gañán, J., Encinar, J.M., Sabio, E., Rubiales, J., 2005. Pyrolysis of Almond Shells. Energy Applications of Fractions. Ind Eng Chem Res 44, 3003–3012. https://doi.org/10.1021/ie0490942

Goyal, H.B., Seal, D., Saxena, R.C., 2008. Bio-fuels from thermochemical conversion of renewable resources: A review. Renewable and Sustainable Energy Reviews 12, 504–517. https://doi.org/10.1016/j.rser.2006.07.014

Griffin, W.C., 1954. Calculation of HLB value of Nonionic Surfactants. J Soc Cosmet Chem 5, 249–256.

Grzymala-Busse, D.M., Grzymala-Busse, J.W., 1995. The Usefulness of a Machine Learning Approach to Knowledge Acquisition. Comput Intell 11, 268–279. https://doi.org/10.1111/j.1467-8640.1995.tb00032.x

Guedes, R.E., Luna, A.S., Torres, A.R., 2018. Operating parameters for bio-oil production in biomass pyrolysis: A review. J Anal Appl Pyrolysis 129, 134–149. https://doi.org/10.1016/j.jaap.2017.11.019

Guo, X., Wang, S., Wang, Q., Guo, Z., Luo, Z., 2011. Properties of Bio-oil from Fast Pyrolysis of Rice Husk. Chin J Chem Eng 19, 116–121. https://doi.org/10.1016/S1004-9541(09)60186-5

Gupta, A., 2019. Machine Learning - Semi-Supervised Learning [WWW Document]. GeeksforGeeks. URL https://www.geeksforgeeks.org/ml-semi-supervised-learning/

Gupta, G.K., Gupta, P.K., Mondal, M.K., 2019. Experimental process parameters optimization and in-depth product characterizations for teak sawdust pyrolysis. Waste Management 87, 499–511. https://doi.org/10.1016/j.wasman.2019.02.035

Gurav, J.B., Regulwar, D.G., 2020. Multi-objective fuzzy optimization for sustainable irrigation planning. H2Open Journal 3, 373–389. https://doi.org/10.2166/h2oj.2020.032

Gutman, I., 2013. Degree-based topological indices. Croatica Chemica Acta 86, 351–361. https://doi.org/10.5562/cca2294

Hada, S., Solvason, C.C., Eden, M.R., 2014. Characterization-Based Molecular Design of Bio-Fuel Additives Using Chemometric and Property Clustering Techniques. Front Energy Res 2. https://doi.org/10.3389/fenrg.2014.00020

Hafeez, S., Pallari, E., Manos, G., Constantinou, A., 2019. Catalytic Conversion and Chemical Recovery, in: Plastics to Energy. Elsevier, pp. 147–172. https://doi.org/10.1016/B978-0-12-813140-4.00006-6

Hansen, S., Mirkouei, A., 2019. Bio-Oil Upgrading via Micro-Emulsification and Ultrasound Treatment: Examples for Analysis and Discussion, in: Volume 4: 24th Design for Manufacturing and the Life Cycle Conference; 13th International Conference on Micro- and Nanosystems. American Society of Mechanical Engineers. https://doi.org/10.1115/DETC2019-97182

Hasan, M., Haseli, Y., Karadogan, E., 2018. Correlations to Predict Elemental Compositions and Heating Value of Torrefied Biomass. Energies (Basel) 11, 2443. https://doi.org/10.3390/en11092443

Hassan, E.M., Steele, P.H., Ingram, L., 2009. Characterization of Fast Pyrolysis Bio-oils Produced from Pretreated Pine Wood. Appl Biochem Biotechnol 154, 3–13. https://doi.org/10.1007/s12010-008-8445-3

Heo, H.S., Park, H.J., Dong, J.-I., Park, S.H., Kim, S., Suh, D.J., Suh, Y.-W., Kim, S.-S., Park, Y.-K., 2010a. Fast pyrolysis of rice husk under different reaction conditions. Journal of Industrial and Engineering Chemistry 16, 27–31. https://doi.org/10.1016/j.jiec.2010.01.026

Heo, H.S., Park, H.J., Yim, J.-H., Sohn, J.M., Park, J., Kim, S.-S., Ryu, C., Jeon, J.-K., Park, Y.-K., 2010b. Influence of operation variables on fast pyrolysis of Miscanthus sinensis var. purpurascens. Bioresour Technol 101, 3672–3677. https://doi.org/10.1016/j.biortech.2009.12.078

Hernandez-Mena, L.E., Pecora, A.A.B., Beraldo, A.L., 2014. Slow Pyrolysis of Bamboo Biomass: Analysis of Biochar Properties. Chemical Engineering Transactions2 37, 115–120. https://doi.org/10.3303/CET1437020

Hu, Z., Zheng, Y., Yan, F., Xiao, B., Liu, S., 2013. Bio-oil production through pyrolysis of blue-green algae blooms (BGAB): Product distribution and bio-oil characterization. Energy 52, 119–125. https://doi.org/10.1016/j.energy.2013.01.059

Hukkerikar, A.S., Kalakul, S., Sarup, B., Young, D.M., Sin, G., Gani, R., 2012a. Estimation of environment-related properties of chemicals for design of sustainable processes: Development of group-contribution+ (GC +) property models and uncertainty analysis. J Chem Inf Model 52, 2823–2839. https://doi.org/10.1021/ci300350r

Hukkerikar, A.S., Sarup, B., ten Kate, A., Abildskov, J., Sin, G., Gani, R., 2012b. Group-contribution + (GC +) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. Fluid Phase Equilib 321, 25–43. https://doi.org/10.1016/j.fluid.2012.02.010

IBM Cloud Education, 2020a. Supervised Learning [WWW Document]. IBM Cloud Learn Hub. URL https://www.ibm.com/cloud/learn/supervised-learning

IBM Cloud Education, 2020b. Unsupervised Learning [WWW Document]. IBM Cloud Learn Hub. URL https://www.ibm.com/cloud/learn/unsupervised-learning

IBM Cloud Education, 2020c. Machine Learning [WWW Document]. IBM Cloud Learn Hub. URL https://www.ibm.com/cloud/learn/machine-learning#toc-reinforcem-oVSqcdnf

IDSS, 1999. User's Guide ROSE 2 Rough Set Data Explorer.

Islam, M.N., Islam, M.N., Beg, M.R.A., 2004. The fuel properties of pyrolysis liquid derived from urban solid wastes in Bangladesh. Bioresour Technol 92, 181–186. https://doi.org/10.1016/j.biortech.2003.08.009

Islam, M.R., Parveen, M., Haniu, H., 2010. Properties of sugarcane waste-derived bio-oils obtained by fixed-bed fire-tube heating pyrolysis. Bioresour Technol 101, 4162–4168. https://doi.org/10.1016/j.biortech.2009.12.137

Jackson, J.D., Weis, D.C., Visco Jr, D.P., 2008. Potential Glucocorticoid Receptor Ligands with Pulmonary Selectivity Using I-QSAR with the Signature Molecular Descriptor. Chem Biol Drug Des 72, 540–550. https://doi.org/10.1111/j.1747-0285.2008.00732.x

Jahirul, M., Rasul, M., Chowdhury, A., Ashwath, N., 2012. Biofuels Production through Biomass Pyrolysis —A Technological Review. Energies (Basel) 5, 4952–5001. https://doi.org/10.3390/en5124952

Jalowka, J.W., Daubert, T.E., 1986. Group contribution method to predict critical temperature and pressure of hydrocarbons. Industrial & Engineering Chemistry Process Design and Development 25, 139–142. https://doi.org/10.1021/i200032a021

Jenkins, R.W., Sutton, A.D., Robichaud, D.J., 2016. Pyrolysis of Biomass for Aviation Fuel, in: Biofuels for Aviation. Elsevier, pp. 191–215. https://doi.org/10.1016/B978-0-12-804568-8.00008-1

Jiang, X., Ellis, N., 2010. Upgrading Bio-oil through Emulsification with Biodiesel: Mixture Production. Energy & Fuels 24, 1358–1364. https://doi.org/10.1021/ef9010669

Joback, K.G., Reid, R.C., 1987. Estimation of Pure-Component Properties from Group-Contributions. Chem Eng Commun 57, 233–243. https://doi.org/10.1080/00986448708960487

Kabir, G., Mohd Din, A.T., Hameed, B.H., 2017. Pyrolysis of oil palm mesocarp fiber and palm frond in a slow-heating fixed-bed reactor: A comparative study. Bioresour Technol 241, 563–572. https://doi.org/10.1016/j.biortech.2017.05.180

Kalaivani, R., Suresh, M.V., Srinivasan, N., 2017. A Study of Rough Sets Theory and It's Application Over Various Field. Journal of Applied Science and Engineering Methodologies 3, 447–455.

Kandaramath Hari, T., Yaakob, Z., Binitha, N.N., 2015. Aviation biofuel from renewable resources: Routes, opportunities and challenges. Renewable and Sustainable Energy Reviews 42, 1234–1244. https://doi.org/10.1016/j.rser.2014.10.095

Kang, B.-S., Lee, K.H., Park, H.J., Park, Y.-K., Kim, J.-S., 2006. Fast pyrolysis of radiata pine in a bench scale plant with a fluidized bed: Influence of a char separation system and reaction conditions on the production of bio-oil. J Anal Appl Pyrolysis 76, 32–37. https://doi.org/10.1016/j.jaap.2005.06.012

Kaniapan, S., Hassan, S., Ya, H., Patma Nesan, K., Azeem, M., 2021. The Utilisation of Palm Oil and Oil Palm Residues and the Related Challenges as a Sustainable Alternative in Biofuel, Bioenergy, and Transportation Sector: A Review. Sustainability 13, 3110. https://doi.org/10.3390/su13063110

Kasar, P., Ahmaruzzaman, Md., 2021. Correlative HHV prediction from proximate and ultimate analysis of char obtained from co-cracking of residual fuel oil with plastics. Korean Journal of Chemical Engineering. https://doi.org/10.1007/s11814-021-0790-8

Khor, S.Y., Liam, K.Y., Loh, W.X., Tan, C.Y., Ng, L.Y., Hassim, M.H., Ng, D.K.S., Chemmangattuvalappil, N.G., 2017. Computer Aided Molecular Design for alternative sustainable solvent to extract oil from palm pressed fibre. Process Safety and Environmental Protection 106, 211–223. https://doi.org/10.1016/j.psep.2017.01.006

Khosravanipour Mostafazadeh, A., Solomatnikova, O., Drogui, P., Tyagi, R.D., 2018. A review of recent research and developments in fast pyrolysis and bio-oil upgrading. Biomass Convers Biorefin 8, 739–773. https://doi.org/10.1007/s13399-018-0320-z

Kim, K.H., Eom, I.Y., Lee, S.M., Choi, D., Yeo, H., Choi, I.-G., Choi, J.W., 2011. Investigation of physicochemical properties of biooils produced from yellow poplar wood (Liriodendron tulipifera) at various temperatures and residence times. J Anal Appl Pyrolysis 92, 2–9. https://doi.org/10.1016/j.jaap.2011.04.002

Kim, S.-J., Jung, S.-H., Kim, J.-S., 2010. Fast pyrolysis of palm kernel shells: Influence of operation parameters on the bio-oil yield and the yield of phenol and phenolic compounds. Bioresour Technol 101, 9294–9300. https://doi.org/10.1016/j.biortech.2010.06.110

Klein, D.J., 2002. Topological Indices and Related Descriptors in QSAR and QSPR. J Chem Inf Comput Sci 42, 1507–1507. https://doi.org/10.1021/ci010441h

Kostetskyy, P., Broadbelt, L.J., 2020. Progress in Modeling of Biomass Fast Pyrolysis: A Review. Energy & Fuels 34, 15195–15216. https://doi.org/10.1021/acs.energyfuels.0c02295

Kotsiantis, S.B., Zaharakis, I.D., Pintelas, P.E., 2006. Machine learning: a review of classification and combining techniques. Artif Intell Rev 26, 159–190. https://doi.org/10.1007/s10462-007-9052-3

Kurnia, J.C., Jangam, S. v., Akhtar, S., Sasmito, A.P., Mujumdar, A.S., 2016. Advances in biofuel production from oil palm and palm oil processing wastes: A review. Biofuel Research Journal 3, 332–346. https://doi.org/10.18331/BRJ2016.3.1.3

Lai, Y.Y., Yik, K.C.H., Hau, H.P., Chow, C.P., Chemmangattuvalappil, N.G., Ng, L.Y., 2019. Enterprise Decision-making Framework for Chemical Product Design in Integrated Biorefineries. Process Integration and Optimization for Sustainability 3, 25–42. https://doi.org/10.1007/s41660-018-0037-2

Lee, K.-H., Kang, B.-S., Park, Y.-K., Kim, J.-S., 2005. Influence of Reaction Temperature, Pretreatment, and a Char Removal System on the Production of Bio-oil from Rice Straw by Fast Pyrolysis, Using a Fluidized Bed. Energy & Fuels 19, 2179–2184. https://doi.org/10.1021/ef050015o

Leng, L., Zhang, W., Peng, H., Li, H., Jiang, S., Huang, H., 2020. Nitrogen in bio-oil produced from hydrothermal liquefaction of biomass: A review. Chemical Engineering Journal 401, 126030. https://doi.org/10.1016/j.cej.2020.126030

Leng, L., Zhou, W., 2018. Chemical compositions and wastewater properties of aqueous phase (wastewater) produced from the hydrothermal treatment of wet biomass: A review. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 40, 2648–2659. https://doi.org/10.1080/15567036.2018.1495780

León, M., Silva, J., Carrasco, S., Barrientos, N., 2020. Design, Cost Estimation and Sensitivity Analysis for a Production Process of Activated Carbon from Waste Nutshells by Physical Activation. Processes 8, 945. https://doi.org/10.3390/pr8080945

Li, R., Zhong, Z., Jin, B., Zheng, A., 2012. Selection of Temperature for Bio-oil Production from Pyrolysis of Algae from Lake Blooms. Energy & Fuels 26, 2996–3002. https://doi.org/10.1021/ef300180r

Li, Y., Wang, T., Liang, W., Wu, C., Ma, L., Zhang, Q., Zhang, X., Jiang, T., 2010. Ultrasonic Preparation of Emulsions Derived from Aqueous Bio-oil Fraction and 0# Diesel and Combustion Characteristics in Diesel

Generator. Energy & Fuels 24, 1987–1995. https://doi.org/10.1021/ef9010934

Liao, Y.-Y., Wang, Z.-T., Chen, J.-W., Han, S.-K., Wang, L.-S., Lu, G.-Y., Zhao, T.-N., 1996. The Prediction of Soil Sorption Coefficients of Heterocyclic Nitrogen Compounds by Octanol/Water Partition Coefficient, Water Solubility, and by Molecular Connectivity Indices. Bull Environ Contam Toxicol 56, 711–716. https://doi.org/10.1007/s001289900104

Lin, B.-J., Chen, W.-H., Budzianowski, W.M., Hsieh, C.-T., Lin, P.-H., 2016. Emulsification analysis of bio-oil and diesel under various combinations of emulsifiers. Appl Energy 178, 746–757. https://doi.org/10.1016/j.apenergy.2016.06.104

Lipnick, R.L., 1989. Structure-Activity relationships in environmental toxicology and chemistry : Narcosis, electrophile and proelectrophile toxicity mechanisms: Application of SAR and QSAR. Environ Toxicol Chem 8, 1–12. https://doi.org/10.1002/etc.5620080101

Liu, K., Zhao, W., Guo, T., Lei, Q., Guan, Y., Wang, D., Cui, M., Fu, S., Zhao, J., Zong, Z., Wei, X., 2021. Emulsification and Performance Measurement of Bio-oil with Diesel. Waste Biomass Valorization 12, 2933–2944. https://doi.org/10.1007/s12649-019-00917-1

Liu, Y., Zhai, Y., Li, S., Liu, Xiangmin, Liu, Xiaoping, Wang, B., Qiu, Z., Li, C., 2020. Production of bio-oil with low oxygen and nitrogen contents by combined hydrothermal pretreatment and pyrolysis of sewage sludge. Energy 203, 117829. https://doi.org/10.1016/j.energy.2020.117829

Lu, X., Tao, S., Hu, H., Dawson, R.W., 2000. Estimation of bioconcentration factors of nonionic organic compounds in fish by molecular connectivity indices and polarity correction factors. Chemosphere 41, 1675–1688. https://doi.org/10.1016/S0045-6535(00)00050-3

Luo, G., Chandler, D.S., Anjos, L.C.A., Eng, R.J., Jia, P., Resende, F.L.P., 2017. Pyrolysis of whole wood chips and rods in a novel ablative reactor. Fuel 194, 229–238. https://doi.org/10.1016/j.fuel.2017.01.010

Luo, Z., Wang, S., Liao, Y., Zhou, J., Gu, Y., Cen, K., 2004. Research on biomass fast pyrolysis for liquid fuel. Biomass Bioenergy 26, 455–462. https://doi.org/10.1016/j.biombioe.2003.04.001

Madathil, D., V, R.P., Nair, M.G., Jamasb, T., Thakur, T., 2021. Consumer-focused solar-grid net zero energy buildings: A multi-objective weighted sum optimization and application for India. Sustain Prod Consum 27, 2101–2111. https://doi.org/10.1016/j.spc.2021.05.012

Mah, A.X.Y., Chin, H.H., Neoh, J.Q., Aboagwa, O.A., Thangalazhy-Gopakumar, S., Chemmangattuvalappil, N.G., 2019. Design of bio-oil additives via computer-aided molecular design tools and phase stability analysis on final blends. Comput Chem Eng 123, 257–271. https://doi.org/10.1016/j.compchemeng.2019.01.008

Mahajan, P., Kandwal, R., Vijay, R., 2012. Rough Set Approach in Machine Learning: A Review. Int J Comput Appl 56, 1–13. https://doi.org/10.5120/8924-2996

Majhi, A., Sharma, Y.K., Naik, D. V., Chauhan, R., 2015. The Production and Evaluation of Bio-oil Obtained from the Jatropha Curcas Cake. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 37, 1782–1789. https://doi.org/10.1080/15567036.2011.645120

Majidi, M., Nojavan, S., Nourani Esfetanaj, N., Najafi-Ghalelou, A., Zare, K., 2017. A multi-objective model for optimal operation of a battery/PV/fuel cell/grid hybrid energy system using weighted sum technique and fuzzy satisfying approach considering responsible load management. Solar Energy 144, 79–89. https://doi.org/10.1016/j.solener.2017.01.009

Makibar, J., Fernandez-Akarregi, A.R., Amutio, M., Lopez, G., Olazar, M., 2015. Performance of a conical spouted bed pilot plant for bio-oil production by poplar flash pyrolysis. Fuel Processing Technology 137, 283–289. https://doi.org/10.1016/j.fuproc.2015.03.011

Maliutina, K., Tahmasebi, A., Yu, J., Saltykov, S.N., 2017. Comparative study on flash pyrolysis characteristics of microalgal and lignocellulosic biomass in entrained-flow reactor. Energy Convers Manag 151, 426–438. https://doi.org/10.1016/j.enconman.2017.09.013

Manara, P., Bezergianni, S., Pfisterer, U., 2018. Study on phase behavior and properties of binary blends of bio-oil/fossil-based refinery intermediates: A step toward bio-oil refinery integration. Energy Convers Manag 165, 304–315. https://doi.org/10.1016/j.enconman.2018.01.023

Manuela Pavan, A.W. and T.N. 2005, Pavan, M., Worth, A., Netzeva, T., Manuela Pavan, A.W. and T.N. 2005, 2005. Preliminary Analysis of an

Aquatic Toxicity Dataset and Assessment of QSAR Models for Narcosis. Italy.

Marler, R.T., Arora, J.S., 2010. The weighted sum method for multi-objective optimization: new insights. Structural and Multidisciplinary Optimization 41, 853–862. https://doi.org/10.1007/s00158-009-0460-7

Marrero, J., Gani, R., 2002. Group-contribution-based estimation of octanol/water partition coefficient and aqueous solubility. Ind Eng Chem Res 41, 6623–6633. https://doi.org/10.1021/ie0205290

Marrero, J., Gani, R., 2001. Group-contribution based estimation of pure component properties. Fluid Phase Equilib 183–184, 183–208. https://doi.org/10.1016/S0378-3812(01)00431-9

Martin, S., 2012. Lattice Enumeration for Inverse Molecular Design Using the Signature Descriptor. J Chem Inf Model 52, 1787–1797. https://doi.org/10.1021/ci3001748

Mason, T.J., Paniwnyk, L., Lorimer, J.P., 1996. The uses of ultrasound in food technology. Ultrason Sonochem 3, S253–S260. https://doi.org/10.1016/S1350-4177(96)00034-X

Mata, T.M., Martins, A.A., Caetano, Nidia.S., 2010. Microalgae for biodiesel production and other applications: A review. Renewable and Sustainable Energy Reviews 14, 217–232. https://doi.org/10.1016/j.rser.2009.07.020

Matamba, T., Tahmasebi, A., Khoshk Rish, S., Yu, J., 2020. Promotion Effects of Pressure on Polycyclic Aromatic Hydrocarbons and H 2 Formation

during Flash Pyrolysis of Palm Kernel Shell. Energy & Fuels 34, 3346–3356. https://doi.org/10.1021/acs.energyfuels.9b04409

Matel, E., Vahdatikhaki, F., Hosseinyalamdary, S., Evers, T., Voordijk, H., 2019. An artificial neural network approach for cost estimation of engineering services. International Journal of Construction Management 1–14. https://doi.org/10.1080/15623599.2019.1692400

McCalmont, J.P., Hastings, A., McNamara, N.P., Richter, G.M., Robson, P., Donnison, I.S., Clifton-Brown, J., 2017. Environmental costs and benefits of growing Miscanthus for bioenergy in the UK. GCB Bioenergy 9, 489–507. https://doi.org/10.1111/gcbb.12294

Mei, Y., Liu, R., Wu, W., Zhang, L., 2016. Effect of Hot Vapor Filter Temperature on Mass Yield, Energy Balance, and Properties of Products of the Fast Pyrolysis of Pine Sawdust. Energy & Fuels 30, 10458–10469. https://doi.org/10.1021/acs.energyfuels.6b01877

Mekhilef, S., Siga, S., Saidur, R., 2011. A review on palm oil biodiesel as a source of renewable fuel. Renewable and Sustainable Energy Reviews 15, 1937–1949. https://doi.org/10.1016/j.rser.2010.12.012

MIDA, 2021. Malaysia aims 31% RE capacity by 2025. Malaysian Investment Development Authority.

Midhun Prasad, K., Murugavelh, S., 2020. Experimental investigation and kinetics of tomato peel pyrolysis: Performance, combustion and emission characteristics of bio-oil blends in diesel engine. J Clean Prod 254, 120115. https://doi.org/10.1016/j.jclepro.2020.120115

Minkin, V.I., 1999. Glossary of terms used in theoretical organic chemistry. Pure and Applied Chemistry 71, 1919–1981. https://doi.org/10.1351/pac199971101919

Mohammed, I.Y., Abakr, Y.A., Kazi, F.K., Yusuf, S., Alshareef, I., Chin, S.A., 2015. Pyrolysis of Napier Grass in a Fixed Bed Reactor: Effect of Operating Conditions on Product Yields and Characteristics. Bioresources 10, 6457–6478. https://doi.org/10.15376/biores.10.4.6457-6478

Mortensen, P.M., Grunwaldt, J.-D., Jensen, P.A., Knudsen, K.G., Jensen, A.D., 2011. A review of catalytic upgrading of bio-oil to engine fuels. Appl Catal A Gen 407, 1–19. https://doi.org/10.1016/j.apcata.2011.08.046

MPOB, 2022. Number and Capacities of Palm Oil Sectors December 2021. Kuala Lumpur.

MPOB, 2021. Production of Palm Oil Product 2021. Kuala Lumpur.

MPOB, 2020. Environmental Impact [WWW Document]. Official Palm Oil Information Source. URL http://www.palmoilworld.org/environment.html (accessed 6.6.20).

MPOC, 2022. Malaysian Palm Oil Industry [WWW Document]. Malaysian Palm Oil Council. URL https://mpoc.org.my/malaysian-palm-oil-industry/ (accessed 5.9.22).

MPOC, 2020. The Oil Palm Tree [WWW Document]. Malaysian Palm Oil Council (MPOC). URL http://mpoc.org.my/the-oil-palm-tree/ (accessed 6.6.20).

Mullen, C.A., Boateng, A.A., Goldberg, N.M., Lima, I.M., Laird, D.A., Hicks, K.B., 2010. Bio-oil and bio-char production from corn cobs and stover by fast pyrolysis. Biomass Bioenergy 34, 67–74. https://doi.org/10.1016/j.biombioe.2009.09.012

Müller-Langer, F., Klemm, M., Schneider, J., 2017. Biofuels Production Processess and Technologies, in: Riazi, M.R., Chiaramonti, D. (Eds.), Biofuels Production and Processing Technology. CRC Press, Boca Raton, pp. 153–182.

Mwiti, D., 2021. 10 Real Life Applications of Reinforcement Learning [WWW Document]. Neptule Blog. URL https://neptune.ai/blog/reinforcement-learning-applications

Mythili, R., Venkatachalam, P., Subramanian, P., Uma, D., 2013. Characterization of bioresidues for biooil production through pyrolysis. Bioresour Technol 138, 71–78. https://doi.org/10.1016/j.biortech.2013.03.161

Nanda, S., Mohammad, J., Reddy, S.N., Kozinski, J.A., Dalai, A.K., 2014. Pathways of lignocellulosic biomass conversion to renewable fuels. Biomass Convers Biorefin 4, 157–191. https://doi.org/10.1007/s13399-013-0097-z

Naqvi, S.R., Uemura, Y., Yusup, S.B., 2014. Catalytic pyrolysis of paddy husk in a drop type pyrolyzer for bio-oil production: The role of temperature and catalyst. J Anal Appl Pyrolysis 106, 57–62. https://doi.org/10.1016/j.jaap.2013.12.009

National Biodiesel Board, 2020. Environmental & Safety Information [WWW Document]. URL https://www.biodiesel.org/docs/default-source/fact-sheets/environment-and-safety.pdf?sfvrsn=a5c4d5c5_4 (accessed 6.23.20).

Neoh, J.Q., Chin, H.H., Mah, A.X.Y., Aboagwa, O.A., Thangalazhy-Gopakumar, S., Chemmangattuvalappil, N.G., 2019. Design of bio-oil additives using mathematical optimisation tools considering blend functionality and sustainability aspects. Sustain Prod Consum 19, 53–63. https://doi.org/10.1016/j.spc.2019.03.005

Ng, L.Y., Andiappan, V., Chemmangattuvalappil, N.G., Ng, D.K.S., 2015a. A systematic methodology for optimal mixture design in an integrated biorefinery. Comput Chem Eng 81, 288–309. https://doi.org/10.1016/j.compchemeng.2015.04.032

Ng, L.Y., Chemmangattuvalappil, N.G., Ng, D.K.S., 2015b. Robust chemical product design via fuzzy optimisation approach. Comput Chem Eng 83, 186–202. https://doi.org/10.1016/j.compchemeng.2015.01.007

Ng, L.Y., Chong, F.K., Chemmangattuvalappil, N.G., 2015c. Challenges and opportunities in computer-aided molecular design. Comput Chem Eng 81, 115–129. https://doi.org/10.1016/j.compchemeng.2015.03.009

Nikolić, S., Trinajstić, N., Mihalić, Z., 1995. The Wiener Index: Development and Application. Croatica Chemica Acta1 1, 105–129.

NIST, 2018. NIST Standard Reference Database Number 69 [WWW Document]. NIST Chemistry WebBook. https://doi.org/10.18434/T4D303

Nurul Islam, M., Nurul Islam, M., Rafiqul Alam Beg, M., Rofiqul Islam, M., 2005. Pyrolytic oil from fixed bed pyrolysis of municipal solid waste and its characterization. Renew Energy 30, 413–420. https://doi.org/10.1016/j.renene.2004.05.002

Nurul Islam, M., Zailani, R., Nasir Ani, F., 1999. Pyrolytic oil from fluidised bed pyrolysis of oil palm shell and itscharacterisation. Renew Energy 17, 73–84. https://doi.org/10.1016/S0960-1481(98)00112-8

Oasmaa, A., Kuoppala, E., 2003. Fast Pyrolysis of Forestry Residue. 3. Storage Stability of Liquid Fuel. Energy & Fuels 17, 1075–1084. https://doi.org/10.1021/ef030011o

Oasmaa, A., Solantausta, Y., Arpiainen, V., Kuoppala, E., Sipilä, K., 2010. Fast Pyrolysis Bio-Oils from Wood and Agricultural Residues. Energy & Fuels 24, 1380–1388. https://doi.org/10.1021/ef901107f

Obernberger, I., Thek, G., 2004. Physical characterisation and chemical composition of densified biomass fuels with regard to their combustion behaviour. Biomass Bioenergy 27, 653–669. https://doi.org/10.1016/j.biombioe.2003.07.006

OJEU, 2009. Renewable Energy Directive (2009/28/EC).

Olivier, C., Bernhard, S., Alexander, Z., 2006. Introduction to Semi-Supervised Learning, in: Semi-Supervised Learning. The MIT Press, pp. 1–12. https://doi.org/10.7551/mitpress/9780262033589.003.0001

Omar, R., Idris, A., Yunus, R., Khalid, K., Aida Isma, M.I., 2011. Characterization of empty fruit bunch for microwave-assisted pyrolysis. Fuel 90, 1536–1544. https://doi.org/10.1016/j.fuel.2011.01.023

Omar, S., Alsamaq, S., Yang, Y., Wang, J., 2019. Production of renewable fuels by blending bio-oil with alcohols and upgrading under supercritical conditions. Front Chem Sci Eng 13, 702–717. https://doi.org/10.1007/s11705-019-1861-9

Ooi, J., Ng, D.K.S., Chemmangattuvalappil, N.G., 2018. Optimal molecular design towards an environmental friendly solvent recovery process. Comput Chem Eng 117, 391–409. https://doi.org/10.1016/j.compchemeng.2018.06.008

Ooi, Y.J., Aung, K.N.G., Chong, J.W., Tan, R.R., Aviso, K.B., Chemmangattuvalappil, N.G., 2022. Design of fragrance molecules using computer-aided molecular design with machine learning. Comput Chem Eng 157, 107585. https://doi.org/10.1016/j.compchemeng.2021.107585

Ostberg, G., Bergenstahl, B., Hulden, M., 1995. Influence of emulsifier on the formation of alkyd emulsions. Colloids Surf A Physicochem Eng Asp 94, 161–171. https://doi.org/10.1016/0927-7757(94)02977-6

Pacheco, R., Silva, C., 2019. Global Warming Potential of Biomass-to-Ethanol: Review and Sensitivity Analysis through a Case Study. Energies (Basel) 12, 2535. https://doi.org/10.3390/en12132535

Park, J., Lee, Y., Ryu, C., Park, Y.-K., 2014. Slow pyrolysis of rice straw: Analysis of products properties, carbon and energy yields. Bioresour Technol 155, 63–70. https://doi.org/10.1016/j.biortech.2013.12.084

Park, J.-W., Heo, J., Ly, H.V., Kim, J., Lim, H., Kim, S.-S., 2019. Fast pyrolysis of acid-washed oil palm empty fruit bunch for bio-oil production in a bubbling fluidized-bed reactor. Energy 179, 517–527. https://doi.org/10.1016/j.energy.2019.04.211

Park, J.Y., Kim, J.-K., Oh, C.-H., Park, J.-W., Kwon, E.E., 2019. Production of bio-oil from fast pyrolysis of biomass using a pilot-scale circulating fluidized bed reactor and its characterization. J Environ Manage 234, 138–144. https://doi.org/10.1016/j.jenvman.2018.12.104

Pattiya, A., 2011. Bio-oil production via fast pyrolysis of biomass residues from cassava plants in a fluidised-bed reactor. Bioresour Technol 102, 1959–1967. https://doi.org/10.1016/j.biortech.2010.08.117

Pattiya, A., Sukkasi, S., Goodwin, V., 2012. Fast pyrolysis of sugarcane and cassava residues in a free-fall reactor. Energy 44, 1067–1077. https://doi.org/10.1016/j.energy.2012.04.035

Pattiya, A., Suttibak, S., 2012. Production of bio-oil via fast pyrolysis of agricultural residues from cassava plantations in a fluidised-bed reactor with a hot vapour filtration unit. J Anal Appl Pyrolysis 95, 227–235. https://doi.org/10.1016/j.jaap.2012.02.010

Pawlak, Z., 2002. Rough set theory and its applications. Journal of Telecomunications and Information Technoogy 3, 7–10.

Pawlak, Z., 1982. Rough sets. International Journal of Computer & Information Sciences 11, 341–356. https://doi.org/10.1007/BF01001956

Peng, X., Wen, J., Li, Z., Yang, G., Zhou, C., Reid, A., Hepburn, D.M., Judd, M.D., Siew, W.H., 2017. Rough set theory applied to pattern recognition of Partial Discharge in noise affected cable data. IEEE Transactions on Dielectrics and Electrical Insulation 24, 147–156. https://doi.org/10.1109/TDEI.2016.006060

Perkin Elmer, 2022. Elite-5ms Capillary Column - 60 m x 0.25 mm I.D. x 1.00 µm [WWW Document]. Perkin Elmer. URL https://www.perkinelmer.com/uk/product/col-elite-5ms-60-m-1-0-m-25-mm-n9316287 (accessed 9.22.22).

Phan, B.M.Q., Duong, L.T., Nguyen, V.D., Tran, T.B., Nguyen, M.H.H., Nguyen, L.H., Nguyen, D.A., Luu, L.C., 2014. Evaluation of the production potential of bio-oil from Vietnamese biomass resources by fast pyrolysis. Biomass Bioenergy 62, 74–81. https://doi.org/10.1016/j.biombioe.2014.01.012

Pidtasang, B., Udomsap, P., Sukkasi, S., Chollacoop, N., Pattiya, A., 2013. Influence of alcohol addition on properties of bio-oil produced from fast pyrolysis of eucalyptus bark in a free-fall reactor. Journal of Industrial and Engineering Chemistry 19, 1851–1857. https://doi.org/10.1016/j.jiec.2013.02.031

Pięta, P., Szmuc, T., Kluza, K., 2019. Comparative Overview of Rough Set Toolkit Systems for Data Analysis. MATEC Web of Conferences 252, 03019. https://doi.org/10.1051/matecconf/201925203019

Pirouz, B., Khorram, E., 2016. A Computational Approach based on the Epsilon-Constraint Method in Multi-Objective Optimization Problems. Adv Appl Stat 49, 453–483. https://doi.org/10.17654/AS049060453

Pittman, C.U., Mohan, D., Eseyin, A., Li, Q., Ingram, L., Hassan, E.-B.M., Mitchell, B., Guo, H., Steele, P.H., 2012. Characterization of Bio-oils Produced from Fast Pyrolysis of Corn Stalks in an Auger Reactor. Energy & Fuels 26, 3816–3825. https://doi.org/10.1021/ef3003922

Pleanjai, S., Gheewala, S., Garivait, S., 2007. Environmental Evaluation of Biodiesel Production from Palm Oil in a Life Cycle Perspectie. Asian J Energy Environ 8.

Pratas, M.J., Freitas, S.V.D., Oliveira, M.B., Monteiro, S.C., Lima, A.S., Coutinho, J.A.P., 2011. Biodiesel Density: Experimental Measurements and Prediction Models. Energy & Fuels 25, 2333–2340. https://doi.org/10.1021/ef2002124

Prausnitz, J.M., 1969. Molecular Thermodynamics of Fluid-Phase Equilibria. Prentice-Hall, Michigan.

Predki, B., Słowiński, R., Stefanowski, J., Susmaga, R., Wilk, S., 1998. ROSE - Software Implementation of the Rough Set Theory, in: Polkowski, L., Skowron, A. (Eds.), Rough Sets and Current Trends in Computing.

Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 605–608. https://doi.org/10.1007/3-540-69115-4_85

Prędki, B., Wilk, S., 1999. Rough set based data exploration using ROSE system, in: Ras, Z.W., Skowron, A. (Eds.), Foundation of Intelligent Systems. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 172–180. https://doi.org/10.1007/BFb0095102

Qian, Y., Liang, J., Dang, C., 2008. Converse approximation and rule extraction from decision tables in rough set theory. Computers & Mathematics with Applications 55, 1754–1765. https://doi.org/10.1016/j.camwa.2007.08.031

Qureshi, S.S., Nizamuddin, S., Baloch, H.A., Siddiqui, M.T.H., Mubarak, N.M., Griffin, G.J., 2019. An overview of OPS from oil palm industry as feedstock for bio-oil production. Biomass Convers Biorefin 9, 827–841. https://doi.org/10.1007/s13399-019-00381-w

Radhakrishnapany, K.T., Wong, C.Y., Tan, F.K., Chong, J.W., Tan, R.R., Aviso, K.B., Janairo, J.I.B., Chemmangattuvalappil, N.G., 2020. Design of fragrant molecules through the incorporation of rough sets into computer-aided molecular design. Mol Syst Des Eng 5, 1391–1416. https://doi.org/10.1039/D0ME00067A

Randic, M., Randie, M., Randic, M., 1975. Characterization of molecular branching. J Am Chem Soc 97, 6609–6615. https://doi.org/10.1021/ja00856a001

Rofiqul Islam, M., Haniu, H., Rafiqul Alam Beg, M., 2008. Liquid fuels and chemicals from pyrolysis of motorcycle tire waste: Product yields, compositions and related properties. Fuel 87, 3112–3122. https://doi.org/10.1016/j.fuel.2008.04.036

Roganov, G.N., Pisarev, P.N., Emel'yanenko, V.N., Verevkin, S.P., 2005. Measurement and Prediction of Thermochemical Properties. Improved Benson-Type Increments for the Estimation of Enthalpies of Vaporization and Standard Enthalpies of Formation of Aliphatic Alcohols. J Chem Eng Data 50, 1114–1124. https://doi.org/10.1021/je049561m

Rogers, J.G., Brammer, J.G., 2012. Estimation of the production cost of fast pyrolysis bio-oil. Biomass Bioenergy 36, 208–217. https://doi.org/10.1016/j.biombioe.2011.10.028

Ronsse, F., van Hecke, S., Dickinson, D., Prins, W., 2013. Production and characterization of slow pyrolysis biochar: influence of feedstock type and pyrolysis conditions. GCB Bioenergy 5, 104–115. https://doi.org/10.1111/gcbb.12018

Rosillo-Calle, F., Groot, P. de, Hemstock, S.L., Woods, J., 2008. The Biomass Assessment Handbook: Bioenergy for a Sustainable Environment. Earthscan, London.

Ross, S., 2019. 4 Countries that Product the Most Food. Investopedia.

Rout, T., Pradhan, D., Singh, R.K., Kumari, N., 2016. Exhaustive study of products obtained from coconut shell pyrolysis. J Environ Chem Eng 4, 3696–3705. https://doi.org/10.1016/j.jece.2016.02.024

RTS, 2021. What is Waste-to-Energy? [WWW Document]. Recycle Track Systems. URL https://www.rts.com/blog/what-is-waste-to-energy/ (accessed 9.4.22).

Rudin, C., Ertekin, Ş., 2018. Learning customized and optimized lists of rules with mathematical programming. Math Program Comput 10, 659–702. https://doi.org/10.1007/s12532-018-0143-8

Sakthivel, R., Ramesh, K., Purnachandran, R., Mohamed Shameer, P., 2018. A review on the properties, performance and emission aspects of the third generation biodiesels. Renewable and Sustainable Energy Reviews 82, 2970–2992. https://doi.org/10.1016/j.rser.2017.10.037

Salary Explorer, 2021. Factory and Manufacturing Average Salaries in Malaysia 2021 [WWW Document]. Salary Explorer.

Salehi, E., Abedi, J., Harding, T., 2009. Bio-oil from Sawdust: Pyrolysis of Sawdust in a Fixed-Bed System. Energy & Fuels 23, 3767–3772. https://doi.org/10.1021/ef900112b

Sara Muznik, 2018. 9 reasons why we better move away from waste-to-energy, and embrace zero waste instead [WWW Document]. Zero Waste Europe. URL https://zerowasteeurope.eu/2018/02/9-reasons-why-we-better-move-away-from-waste-to-energy-and-embrace-zero-waste-instead/ (accessed 9.21.22).

Schultz, T.W., Cronin, M.T.D., Netzeva, T.I., Aptula, A.O., 2002. Structure-toxicity relationships for aliphatic chemicals evaluated with Tetrahymena

pyriformis. Chem Res Toxicol 15, 1602–1609. https://doi.org/10.1021/tx025589p

Shahzadah Nayyar Jehan, Mirzosaid Sultonov, 2019. Green Investment Policy Initiatives in Japan, in: Proceedings of the 15th International Scientific and Practical Conference of the Russian Society for Ecological Economics Strategies and Tools. Stavropol, Russia, pp. 1–9.

Sharma, A., Kumar, S., 2020. Bayesian rough set based information retrieval. Journal of Statistics and Management Systems 23, 1147–1158. https://doi.org/10.1080/09720510.2020.1799575

Silver, D., Hassabis, D., 2017. AlphaGo Zero: Starting from scratch [WWW Document]. DeepMind. URL https://deepmind.com/blog/article/alphago-zero-starting-scratch

Simon, A., Deo, M.S., Venkatesan, S., Babu, D.R.R., 2015. An Overview of Machine Learning and its Applications. International Journal of Electrical Sciences and Engineering 1, 22–24.

Singh, J., Trivedi, J., 2017. Raw Materials for Biofuels Production, in: Riazi, M.R., Chiaramonti, D. (Eds.), Biofuels Production and Processing Technology2. CRC Press, Boca Raton, pp. 127–152.

Singh, V.K., Soni, A.B., Kumar, S., Singh, R.K., 2014. Pyrolysis of sal seed to liquid product. Bioresour Technol 151, 432–435. https://doi.org/10.1016/j.biortech.2013.10.087

Sinha, M., Achenie, L.E.K., 2001. Systematic design of blanket wash solvents with recovery considerations. Advances in Environmental Research 5, 239–249. https://doi.org/10.1016/S1093-0191(00)00058-7

Sinha, R., Kumar, S., Singh, R.K., 2013. Production of biofuel and biochar by thermal pyrolysis of linseed seed. Biomass Convers Biorefin 3, 327–335. https://doi.org/10.1007/s13399-013-0076-4

Sınağ, A., Gülbay, S., Uskan, B., Uçar, S., Özgürler, S.B., 2010. Production and characterization of pyrolytic oils by pyrolysis of waste machinery oil. J Hazard Mater 173, 420–426. https://doi.org/10.1016/j.jhazmat.2009.08.100

Solikhah, M.D., Pratiwi, F.T., Heryana, Y., Wimada, A.R., Karuana, F., Raksodewanto, A., Kismanto, A., 2018. Characterization of Bio-Oil from Fast Pyrolysis of Palm Frond and Empty Fruit Bunch. IOP Conf Ser Mater Sci Eng 349, 012035. https://doi.org/10.1088/1757-899X/349/1/012035

Sotoudehnia, F., Baba Rabiu, A., Alayat, A., McDonald, A.G., 2020. Characterization of bio-oil and biochar from pyrolysis of waste corrugated cardboard. J Anal Appl Pyrolysis 145, 104722. https://doi.org/10.1016/j.jaap.2019.104722

Street, C., Woody, J., Ardila, J., Bagajewicz, M., 2008. Product Design: A Case Study of Slow-Release Carpet Deodorizers/Disinfectants. Ind Eng Chem Res 47, 1192–1200. https://doi.org/10.1021/ie0710622

Sudlow, K., Woolf, A.A., 1995. Stabilisation of tetrahedrane by fluorination. J Fluor Chem 71, 31–37. https://doi.org/10.1016/0022-1139(94)03148-S

Summers, S., Yang, S., Watson, J., Zhang, Y., 2022. Diesel blends produced via emulsification of hydrothermal liquefaction biocrude from food waste. Fuel 324, 124817. https://doi.org/10.1016/j.fuel.2022.124817

Suraj, Z., 2004. An Introduction to Rough Set Theory and Its Applications: A tutorial, in: 1st International Computer Engineering Conference New Technologies for the Information Society. Cairo, Egypt.

Sutrisno, B., Hidayat, A., 2018. Pyrolysis of palm empty fruit bunch: Yields and analysis of bio-oil. MATEC Web of Conferences 154, 01036. https://doi.org/10.1051/matecconf/201815401036

Sutton, R.S., 1992. Introduction: The Challenge of Reinforcement Learning, in: Reinforcement Learning. Springer US, Boston, MA, pp. 1–3. https://doi.org/10.1007/978-1-4615-3618-5_1

Taha, A., Ahmed, E., Ismaiel, A., Ashokkumar, M., Xu, X., Pan, S., Hu, H., 2020. Ultrasonic emulsification: An overview on the preparation of different emulsifiers-stabilized emulsions. Trends Food Sci Technol 105, 363–377. https://doi.org/10.1016/j.tifs.2020.09.024

Talebian-Kiakalaieh, A., Amin, N.A.S., Mazaheri, H., 2013. A review on novel processes of biodiesel production from waste cooking oil. Appl Energy 104, 683–710. https://doi.org/10.1016/j.apenergy.2012.11.061

Tang, Q., Chen, Y., Yang, H., Liu, M., Xiao, H., Wang, S., Chen, H., Raza Naqvi, S., 2021. Machine learning prediction of pyrolytic gas yield and compositions with feature reduction methods: Effects of pyrolysis

conditions and biomass characteristics. Bioresour Technol 339, 125581. https://doi.org/10.1016/j.biortech.2021.125581

Ten, J.Y., Hassim, M.H., Ng, D.K.S., Chemmangattuvalappil, N.G., 2017. A molecular design methodology by the simultaneous optimisation of performance, safety and health aspects. Chem Eng Sci 159, 140–153. https://doi.org/10.1016/j.ces.2016.03.026

TNB, 2021. Pricing & Tariffs - Industrial Tariffs [WWW Document]. Tenaga Nasional Berhad.

Toldrá-Reig, F., Mora, L., Toldrá, F., 2020. Trends in Biodiesel Production from Animal Fat Waste. Applied Sciences 10, 3644. https://doi.org/10.3390/app10103644

Trinh, T.N., Jensen, P.A., Dam-Johansen, K., Knudsen, N.O., Sørensen, H.R., Hvilsted, S., 2013. Comparison of Lignin, Macroalgae, Wood, and Straw Fast Pyrolysis. Energy & Fuels 27, 1399–1409. https://doi.org/10.1021/ef301927y

Tsai, W.T., Lee, M.K., Chang, Y.M., 2006. Fast pyrolysis of rice straw, sugarcane bagasse and coconut shell in an induction-heating reactor. J Anal Appl Pyrolysis 76, 230–237. https://doi.org/10.1016/j.jaap.2005.11.007

Tsekos, C., Tandurella, S., de Jong, W., 2021. Estimation of lignocellulosic biomass pyrolysis product yields using artificial neural networks. J Anal Appl Pyrolysis 157. https://doi.org/10.1016/j.jaap.2021.105180

van Dyk, B., Nieuwoudt, I., 2002. A Computer-aided Molecular Design of Solvents for Distillation Processes. International Conference on Distillation and Absorption 24.

van Engelen, J.E., Hoos, H.H., 2020. A survey on semi-supervised learning. Mach Learn 109, 373–440. https://doi.org/10.1007/s10994-019-05855-6

Vargas-Moreno, J.M., Callejón-Ferre, A.J., Pérez-Alonso, J., Velázquez-Martí, B., 2012. A review of the mathematical models for predicting the heating value of biomass materials. Renewable and Sustainable Energy Reviews 16, 3065–3083. https://doi.org/10.1016/j.rser.2012.02.054

Varma, A.K., Thakur, L.S., Shankar, R., Mondal, P., 2019. Pyrolysis of wood sawdust: Effects of process parameters on products yield and characterization of products. Waste Management 89, 224–235. https://doi.org/10.1016/j.wasman.2019.04.016

Vasu, H., Wong, C.F., Vijiaretnam, N.R., Chong, Y.Y., Thangalazhy-Gopakumar, S., Gan, S., Lee, L.Y., Ng, H.K., 2020. Insight into Co-pyrolysis of Palm Kernel Shell (PKS) with Palm Oil Sludge (POS): Effect on Bio-oil Yield and Properties. Waste Biomass Valorization 11, 5877–5889. https://doi.org/10.1007/s12649-019-00852-1

Venderbosch, R., Prins, W., 2010. Fast pyrolysis technology development. Biofuels, Bioproducts and Biorefining 4, 178–208. https://doi.org/10.1002/bbb.205

Venkatasubramanian, V., Chan, K., Caruthers, J.M., 1994. Computer-aided molecular design using genetic algorithms. Comput Chem Eng 18, 833–844. https://doi.org/10.1016/0098-1354(93)E0023-3

Visco, D.P., Chen, J.J., 2016. The Signature Molecular Descriptor in Molecular Design: Past and Current Applications. Computer Aided Chemical Engineering 39, 315–343. https://doi.org/10.1016/B978-0-444-63683-6.00011-3

Visco, D.P., Pophale, R.S., Rintoul, M.D., Faulon, J.-L., 2002. Developing a methodology for an inverse quantitative structure-activity relationship using the signature molecular descriptor. J Mol Graph Model 20, 429–438. https://doi.org/10.1016/S1093-3263(01)00144-9

Walters, R.N., 2002. Molar group contributions to the heat of combustion. Fire Mater 26, 131–145. https://doi.org/10.1002/fam.802

Wang, H., Olhofer, M., Jin, Y., 2017. A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges. Complex & Intelligent Systems 3, 233–245. https://doi.org/10.1007/s40747-017-0053-9

Wang, K., Brown, R.C., Homsy, S., Martinez, L., Sidhu, S.S., 2013. Fast pyrolysis of microalgae remnants in a fluidized bed reactor for bio-oil and biochar production. Bioresour Technol 127, 494–499. https://doi.org/10.1016/j.biortech.2012.08.016

Weerachanchai, P., Tangsathitkulchai, C., Tangsathitkulchai, M., 2011. Characterization of products from slow pyrolysis of palm kernel cake and

cassava pulp residue. Korean Journal of Chemical Engineering 28, 2262–2274. https://doi.org/10.1007/s11814-011-0116-3

Weis, D.C., Visco, D.P., 2010. Computer-aided molecular design using the Signature molecular descriptor: Application to solvent selection. Comput Chem Eng 34, 1018–1029. https://doi.org/10.1016/j.compchemeng.2009.10.017

Whitnack, C., Heller, A., Frow, M.T., Kerr, S., Bagajewicz, M.J., 2009. Financial risk management in the design of products under uncertainty. Comput Chem Eng 33, 1056–1066. https://doi.org/10.1016/j.compchemeng.2008.09.018

Wiener, H., 1947. Structural Determination of Paraffin Boiling Points. J Am Chem Soc 69, 17–20. https://doi.org/10.1021/ja01193a005

Williams, P.A., 2001. Food Emulsions: Principles, Practice, and Techniques. Int J Food Sci Technol 36, 223–224. https://doi.org/10.1046/j.1365-2621.2001.00459.x

Wilson, A., 2019. A Brief Introduction to Supervised Learning [WWW Document]. Towards Data Science. URL https://towardsdatascience.com/a-brief-introduction-to-supervised-learning-54a3e3932590

Wood, R.W., Loomis, A.L., 1927. XXXVIII. *The physical and biological effects of high-frequency sound-waves of great intensity*. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 4, 417–436. https://doi.org/10.1080/14786440908564348

World Nuclear Association, 2021. Heat Values of Various Fuels [WWW Document]. World Nuclear Association.

Wright, M.M., Daugaard, D.E., Satrio, J.A., Brown, R.C., 2010. Techno-economic analysis of biomass fast pyrolysis to transportation fuels. Fuel 89, S2–S10. https://doi.org/10.1016/j.fuel.2010.07.029

Wu, S.-R., Chang, C.-C., Chang, Y.-H., Wan, H.-P., 2016. Comparison of oil-tea shell and Douglas-fir sawdust for the production of bio-oils and chars in a fluidized-bed fast pyrolysis system. Fuel 175, 57–63. https://doi.org/10.1016/j.fuel.2016.02.008

Yaakob, Z., Mohammad, M., Alherbawi, M., Alam, Z., Sopian, K., 2013. Overview of the production of biodiesel from Waste cooking oil. Renewable and Sustainable Energy Reviews 18, 184–193. https://doi.org/10.1016/j.rser.2012.10.016

Yakub, M.I., Abdalla, A.Y., Feroz, K.K., Suzana, Y., Ibraheem, A., Chin, S.A., 2015. Pyrolysis of Oil Palm Residues in a Fixed Bed Tubular Reactor. Journal of Power and Energy Engineering 03, 185–193. https://doi.org/10.4236/jpee.2015.34026

Yang, H., Yao, J., Chen, G., Ma, W., Yan, B., Qi, Y., 2014. Overview of Upgrading of Pyrolysis Oil of Biomass. Energy Procedia 61, 1306–1309. https://doi.org/10.1016/j.egypro.2014.11.1087

Yang, L., Liu, S., Tsoka, S., Papageorgiou, L.G., 2015. Sample re-weighting hyper box classifier for multi-class data classification. Comput Ind Eng 85, 44–56. https://doi.org/10.1016/j.cie.2015.02.022

Yin, Q., Wang, S., Li, X., Guo, Z., Gu, Y., 2010. Review of Bio-oil Upgrading Technologies and Experimental Study on Emulsification of Bio-oil and Diesel, in: 2010 International Conference on Optoelectronics and Image Processing. IEEE, pp. 343–347. https://doi.org/10.1109/ICOIP.2010.112

Yunus, N.A., Zaki, N.M., Wan Alwi, S.R., 2018. Design of Solvents for Palm Oil Recovery using Computer Aided Approach. Chem Eng Trans 63, 583–588. https://doi.org/10.3303/CET1863098

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, H., Xiao, R., Huang, H., Xiao, G., 2009. Comparison of non-catalytic and catalytic fast pyrolysis of corncob in a fluidized bed reactor. Bioresour Technol 100, 1428–1434. https://doi.org/10.1016/j.biortech.2008.08.031

Zhang, L., Cignitti, S., Gani, R., 2015. Generic mathematical programming formulation and solution for computer-aided molecular design. Comput Chem Eng 78, 79–84. https://doi.org/10.1016/j.compchemeng.2015.04.022

Zhang, L., Mao, H., Liu, Q., Gani, R., 2020. Chemical product design – recent advances and perspectives. Curr Opin Chem Eng 27, 22–34. https://doi.org/10.1016/j.coche.2019.10.005

Zhang, M., Yewe-Siang Lee Shee We, M., Wu, H., 2018. Direct emulsification of crude glycerol and bio-oil without addition of surfactant via ultrasound and mechanical agitation. Fuel 227, 183–189. https://doi.org/10.1016/j.fuel.2018.04.099

Zhang, Q., Xie, Q., Wang, G., 2016. A survey on rough set theory and its applications. CAAI Trans Intell Technol 1, 323–333. https://doi.org/10.1016/j.trit.2016.11.001

Zhang, S., Yang, X., Zhang, H., Chu, C., Zheng, K., Ju, M., Liu, L., 2019. Liquefaction of Biomass and Upgrading of Bio-Oil: A Review. Molecules 24, 2250. https://doi.org/10.3390/molecules24122250

Zhang, Z., Pittman, C., Sui, S., Sun, J., Wang, Q., 2013. Catalytic Upgrading of Bio-Oil by Reacting with Olefins and Alcohols over Solid Acids: Reaction Paths via Model Compound Studies. Energies (Basel) 6, 1568–1589. https://doi.org/10.3390/en6031568

Zheng, G., Zhang, F., Zheng, Z., Xiang, Y., Yuan, N.J., Xie, X., Li, Z., 2018. DRN, in: Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18. ACM Press, New York, New York, USA, pp. 167–176. https://doi.org/10.1145/3178876.3185994

Zheng, J.-L., 2008. Pyrolysis oil from fast pyrolysis of maize stalk. J Anal Appl Pyrolysis 83, 205–212. https://doi.org/10.1016/j.jaap.2008.08.005

Zhou, L., Yang, H., Wu, H., Wang, M., Cheng, D., 2013. Catalytic pyrolysis of rice husk by mixing with zinc oxide: Characterization of bio-oil and its rheological behavior. Fuel Processing Technology 106, 385–391. https://doi.org/10.1016/j.fuproc.2012.09.003

Zhou, T., McBride, K., Linke, S., Song, Z., Sundmacher, K., 2020. Computer-aided solvent selection and design for efficient chemical processes. Curr Opin Chem Eng 27, 35–44. https://doi.org/10.1016/j.coche.2019.10.007

Zhu, X., Yang, G., 2021. Study on HHV prediction of municipal solid wastes: A machine learning approach. Int J Energy Res. https://doi.org/10.1002/er.7327

Zimmermann, H.-J., 1978. Fuzzy programming and linear programming with several objective functions. Fuzzy Sets Syst 1, 45–55. https://doi.org/10.1016/0165-0114(78)90031-3

Zimmermann, H.-J., 1976. Description and Optimization of Fuzzy Systems. Int J Gen Syst 2, 209–215. https://doi.org/10.1080/03081077608547470

Zimmermann, H.-J., Zysno, P., 1983. Decisions and evaluations by hierarchical aggregation of information. Fuzzy Sets Syst 10, 243–260. https://doi.org/10.1016/S0165-0114(83)80118-3

APPENDIX A: CASE STUDY IN CHAPTER 4

MATHEMATICAL FORMULATION FOR HEIGHT 1

```
MODEL:
!SGN1 = Signature at Height 1;
!NoS = Number of Signature at Height 1;
!MW = Molar Weight;
!Tm = Melting Point;
!Tb = Boiling Point;
!Fp = Flash Point;
!DelS = Total Solubility Parameter;
!DelD, DelP, DelH = Hansen Solubility Parameter;
!Tc = Critical Temperature;
!Pc = Critical Pressure;
!Acc = Accentric Factor;
!CN = Cetane Number;
!Tait1, Tait2 = Auto Ignition Temperature;
!HHV = Higher Heating Value;
!Val =Valency;
!KOW = Octanol/Water Partition Coefficient;
!NoDB = No. of Double Bond;
!NoTB = No. of Triple Bond;
!CI0 = 0TH Order Connectivity Index;

!SGN1 = S1 S2 S3 S4 S5 S6 S7 S8 S9 S10 S11 S12 S13 S14 S15 S16
S17 S18 S19 S20 S21 S22 S23 S24;
!SGN = C1(C) C1(O) C2(=C) C2(CC) C2(CO) C3(=CC) C3(=CO) C3(=OC)
C3(=OO) C3(CCC) C3(CCO) C4(=C=C) C4(=C=O) C4(=CCC) C4(=CCO)
C4(=OCC) C4(=OCO) C4(=-NC) C4(CCCC) C4(CCCO) N3(=-C) O1(C)
O2(=C) O2(CC);

SETS:
SGN1: NoS, MW, Tm, Tb, Fp, DelS, DelD, DelP, DelH, Tc, Pc, Acc,
CN, Tait1, Tait2, HHV, Val1, Val2, Val3, Val4, Carbon, Oxygen,
Nitrogen, KOW, NoDB, NoTB, MMW, MV, DV, ORDER0;
ENDSETS

DATA:
SGN1 = S1 S2 S3 S4 S5 S6 S7 S8 S9 S10 S11 S12 S13 S14 S15 S16
S17 S18 S19 S20 S21 S22 S23 S24;
MW = 15 31 26 14 14 25 25 29 45 13 13 24 24 24 24 42 58 38 12
12 38 17 45 29;
Tm = 0.6953 1.3643 0.7662 0.2515 0.2515 0.1732 0.1732 3.0186
2.0223 -0.373 -0.373 0.3928 0.3928 0.3928 0.3928 2.5232 1.6329
3.3807 0.0256 0.0256 3.3807 2.7888 7.4042 -0.4446;
Tb = 0.8491 1.7703 1.3621 0.7141 0.7141 1.2971 1.2971 2.5388
2.5972 0.2924 0.2924 1.2739 1.2739 1.2739 1.2739 2.6761 2.9850
2.8870 -0.0671 -0.0671 2.8870 2.5670 5.1108 0.8924;
Fp = 33.0909 53.2444 6.0196 11.4107 11.4107 -15.4082 -15.4082
72.5327 66.8687 -17.7415 -17.7416 -44.0014 -44.0014 -44.0014 -
44.0014 63.7628 46.4475 43.3055 -36.6949 -36.6949 43.3055
87.6576 128.2650 -19.6811;
DelS = -1.8029 -1.3922 -0.83 -0.1323 -0.1323 1.6192 1.6192
1.4631 -0.1553 1.0139 1.0139 2.7550 2.7550 2Å.7550 2.7550
0.7468 1.1235 4.0573 1.2449 1.2449 4.0573 3.0524 3.6819 0.5968;
```

```
DelD = 7.5982 7.6317 0.1956 -0.0023 -0.0023 -7.0086 -7.0086
7.8411 7.9230 -7.5390 -7.5390 -14.6160 -14.6160 -14.6160 -
14.6160 0.5371 0.5148 -7.7575 -15.6455 -15.6455 -7.7575 8.0503
8.4172 -7.6174;
DelP = 2.3037 3.2154 2.3059 -0.1664 -0.1664 1.2790 1.2790
7.8726 4.8158 -3.3851 -3.3851 -1.4590 -1.4590 -1.4590 -1.4590
1.2706 0.2595 -0.8832 -5.1979 -5.1979 -0.8832 5.2379 3.1400 -
2.7093;
DelH = 2.2105 3.3464 1.0623 -0.2150 -0.2150 -0.1204 -0.1204
5.3761 6.8448 -2.6826 -2.6826 -2.9995 -2.9995 -2.9995 -2.9995 -
0.0788 2.7824 0.4757 -6.4821 -6.4821 0.4757 11.8005 7.5917 -
2.1534;
Tc = 1.0898 5.6587 7.3554 3.4604 4.4604 10.0135 10.0135 11.2208
11.5492 4.6659 4.6659 13.5316 13.5316 13.5316 13.5316 15.1201
15.6198 13.0005 6.6169 6.6169 13.0005 10.1672 30.6307 5.4851;
Pc = 0.0100 0.0133 0.0176 0.0101 0.0101 0.0170 0.0170 0.0045
0.0083 0.0107 0.0107 0.0209 0.0209 0.0209 0.0209 0.0178 0.0223
0.0291 0.0075 0.0075 0.0291 -0.0071 0.0060 0.0263;
Acc = 0.0017 0.0055 0.0020 0.0019 0.0019 0.0036 0.0036 0.0079
0.0086 0.0029 0.0029 0.0007 0.0007 0.0007 0.0007 0.008 0.0109
0.0069 0.0008 0.0008 0.0069 0.018 0.0206 0.003;
CN = 3.47 15.253 -2.623 4.879 4.879 -3.646 -3.646 20.867 -6.28
-3.7 -3.7 0 0 0 0 -8.401 -12.851 0 -7.296 -7.296 0 -13.769 -
21.342 -2.696;
Tait1 = -0.3516 -0.1637 0.0794 0.1009 0.1009 0.5499 0.5499
0.4254 -0.3396 -0.5829 0.5829 0.8132 0.8132 0.8132 0.8132
0.1214 0.1089 0.2795 0.6668 0.6668 0.2795 -0.1516 0.0375
0.3381;
Tait2 = -57.8605 -64.8653 -44.1695 2.6047 2.6047 -66.0852 -
66.0852 -30.4259 -62.1878 79.2115 79.2115 11.7928 11.7928
11.7928 11.7928 94.507 110.1176 2.3359 95.9781 95.9781 2.3359
32.3056 120.0741 82.8063;
HHV = 775 643 1161 670 670 971 971 449 317 518 518 781 781 781
781 777 782 548 431 431 548 -108 302 386;
Val1 = 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0;
Val2 = 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1;
Val3 = 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0;
Val4 = 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0 0 0;
Carbon = 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0;
Oxygen = 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1;
Nitrogen= 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0;
KOW = 0.1152 -0.3882 0.3122 0.4594 0.4594 0.6 0.6026 -0.9687 -
0.9687 0.43 0.43 0.5563 0.5563 0.5563 0.5563 0.2663 -0.2896 -
0.0882 0.8143 0.8143 -0.0882 1.3365 -1.2021 0.1109;
NoDB = 0 0 1 0 0 1 1 1 1 0 0 2 2 1 1 1 1 0 0 0 0 0 1 0;
NoTB = 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0;
MMW = 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
12 14 16 16 16;
MV = 0.02614 0.03274 0.02697 0.01641 0.01641 0.0161 0.0161
0.02002 0.02667 0.00711 0.00711 0.00296 0.00296 0.00296 0.00296
0 0.03567 0 -0.0038 -0.0038 0 0.00551 0.02232 0.01799;
DV = -1.0278 -0.6902 0.3612 0.2125 0.2125 1.4719 1.4719 -0.076
-0.1208 1.318 1.318 0 0 0 0 2.2504 1.0292 -0.2253 2.8147 2.8147
-0.2253 1.3057 1.143 3.6344;
ORDER0 = 1 1 0.7071 0.7071 0.7071 0.5774 0.5774 0.5774 0.5774
0.5774 0.5774 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.4472
0.4082 0.4082;
ENDDATA

GROUP = @SUM(SGN1:NoS);
```

```
!Structural Constraint
!Handshaking Dilemma;
!Valency of one;
SUMVal1 = @SUM(SGN1:NoS*Val1);
!Valency of two;
SUMVal2 = @SUM(SGN1:NoS*Val2);
!Valency of three;
SUMVal3 = @SUM(SGN1:NoS*Val3);
!Valency of four;
SUMVal4 = @SUM(SGN1:NoS*Val4);
!One double bond;
DBond=@SUM(SGN1:NoS*NoDB);
!One triple bond;
TBond=@SUM(SGN1:NoS*NoTB);

SUMVal1+2*SUMVal2+3*SUMVal3+4*SUMVal4=((GROUP+(0.5*(DBond))+TBo
nd)-1)*2;

!Compliment Signature;
!C-O = O-C;
NoS(2)+NoS(5)+NoS(7)+NoS(9)+NoS(11)+NoS(15)+NoS(17)+NoS(20)=NoS
(22)+(2*NoS(24));
!C=O = O=C;
NoS(8)+NoS(9)+NoS(13)+NoS(16)+NoS(17)=NoS(23);
!C=N = N=C;
NoS(18)=NoS(21);

!Signature with same edges;
!C-C;
2*H1=(NoS(1)+(2*NoS(4))+NoS(5)+NoS(6)+NoS(8)+(3*NoS(10))+(2*NoS
(11))+(2*NoS(14))+NoS(15)+(2*NoS(16))+NoS(17)+NoS(18)+(4*NoS(19
))+(3*NoS(20)));

!C=C;
2*H2=(NoS(3)+NoS(6)+NoS(7)+(NoS(12))+NoS(13)+NoS(14)+NoS(15));

H1>0;@GIN(H1);
H2>0;@GIN(H2);

!Properties Constraint;
!Normal Melting Point;
!unit = K;
SUMTm = @SUM(SGN1:NoS*Tm);
Tmelt = @LOG(SUMTm)*147.450;
Tmelt <= 298.15;

!Normal Boiling Point;
!unit = K;
SUMTb = @SUM(SGN1:NoS*Tb);
Tboil = @LOG(SUMTb)*222.543;
Tboil >= 400;

!Flash Point;
!unit = K;
SUMFp = @SUM(SGN1:NoS*Fp);
Tflash = SUMFp+150.0218;
!Total Solubility Parameter;
!unit = MPa^0.5;
SUMDelS = @SUM(SGN1:NoS*DelS);
Sol = SUMDelS+20.7339;
```

```
!Hansen Solubility Parameter;
!unit = MPa^0.5;
SUMDelD = @SUM(SGN1:NoS*DelD);
SUMDelP = @SUM(SGN1:NoS*DelP);
SUMDelH = @SUM(SGN1:NoS*DelH);
SUMD = @SQR(SUMDelD);
SUMP = @SQR(SUMDelP);
SUMH = @SQR(SUMDelH);
Sp = @SQRT(SUMD+SUMP+SUMH);

!Molar Volume;
SUMMV = @SUM(SGN1:NoS*MV);
MolarVolume = (SUMMV+0.01211);

!Density;
Density = (SUMMW/MolarVolume);
Density > 750;
Density < 1000;

!Critical Temperature;
!unit = K;
SUMTc = @SUM(SGN1:NoS*Tc);
Tcrit = @LOG(SUMTc)*181.6738;

!Critical Pressure;
!unit = bar;
SUMPc = @SUM(SGN1:NoS*Pc);
SUMPca = SUMPc+0.1155;
SUMPcb = 1/SUMPca;
Pcrit = ((SUMPcb^2)+0.0519);

!Acentric Factor;
SUMAcc = @SUM(SGN1:NoS*Acc);
AccFac = ((@LOG(SUMAcc+1.0039))^(1/0.0447))*0.9132;

!Vapour Pressure;
!unit = atm;
Tr = Tboil/Tcrit;
@FREE(Z0);@FREE(Z1);@FREE(lnPr);
Z0 = 5.92714-(6.09648/Tr)-
(1.28862*(@LOG(Tr)))+(0.169347*(Tr^6));
Z1 = 15.2518-(15.6875/Tr)-
(13.4721*(@LOG(Tr)))+(0.43577*(Tr^6));
lnPr = Z0+(AccFac*Z1);
Psat = @EXP(lnPr)*Pcrit;
Psat > 0.01;

!Auto Ignition Temperature;
!unit = K;
SUMTait1 = @SUM(SGN1:NoS*Tait1);
SUMTait2 = @SUM(SGN1:NoS*Tait2);
Tauto = (71.2584*(10^-(SUMTait1)))+525.93+SUMTait2;
Tauto >= 500;

!Higher Heating Value;
!unit = MJ/kg;
MAX = HHVs;
SUMHHV = @SUM(SGN1:NoS*HHV);
SUMMW = @SUM(SGN1:NoS*MW);
```

```
HHVs = SUMHHV/SUMMW;
HHVs > 30;

!Dynamic Viscosity;
SUMDV = @SUM(SGN1:NoS*DV);
DynVis = @EXP(SUMDV);

!Kinematics Viscosity;
KV = DynVis/Density*1000;
KV < 5;
KV > 1;

!Environmental Properties;
!0th Order Connectivity Index;
CI0=@SUM(SGN1:NoS*ORDER0);

!Octanol/Water Partition Coefficient;
SUMKOW = @SUM(SGN1:NoS*KOW);
logKow = SUMKOW+0.7520;
expkow = @exp(SUMKOW);

!MMW;
SUMMMW = @SUM(SGN1:NoS*MMW);

!No of Atoms;
NoC = @SUM(SGN1:NoS*Carbon);
NoO = @SUM(SGN1:NoS*Oxygen);
NoN = @SUM(SGN1:NoS*Nitrogen);

!Acute Toxicity, LC50;
LC50a = NoC*(0.246449);
LC50b = NoO*(-0.21041);
LC50c = NoN*(-0.02395);
LC50d = LC50a+LC50b+LC50c;
LC50e = 0.253883*CI0;
LC50 = -LC50d-LC50e+2.9717;
@FREE(LC50a);@FREE(LC50b);@FREE(LC50c);@FREE(LC50d);@FREE(LC50e
);@FREE(LC50);
LC50 > 1.041;

!Acute Toxicity, Aquatic, EC50;
EC50a = 0.95*logKow;
EC50 = EC50a+1.32;
@FREE(EC50a);@FREE(EC50);
EC50 > 1.041;

!Acute Toxicity, Oral, LD50;
LD50a = NoC*(0.052218);
LD50b = NoO*(0.004962);
LD50c = LD50a+LD50b;
LD50d = 0.018103*CI0;
LD50 = -1.9372-(0.0016*SUMMW)-LD50c-LD50d;
@FREE(LD50a);@FREE(LD50b);@FREE(LD50c);@FREE(LD50d);@FREE(LD50)
;
!LD50 > 3.7;

!Ozone Depletion Potential;
ODPa = NoC*(-1.59188); !-1.59188;
ODPb = NoO*(-3.62951); !-3.62951;
ODPc = ODPa+ODPb;
```

```
ODPd = 8.967731*CI0;
ODP = ODPd+ODPc+3.298083; !logODP;
@FREE(ODPa);@FREE(ODPb);@FREE(ODPc);@FREE(ODPd);@FREE(ODP);

!Global Warming Potential;
GWPa = NoO*(0.424024);
GWPb = NoC*(-3.05993);
GWPc = GWPa+GWPb;
GWPd = (-0.01877)*CI0;
GWP = GWPc+GWPd-0.52073; !logGWP;
@FREE(GWPa);@FREE(GWPb);@FREE(GWPc);@FREE(GWPd);@FREE(GWP);

!Photochemical Oxidation Potential;
PCOa = NoO*(0.126631);
PCOb = NoN*(0.09317);
PCOc = NoC*(-0.03181);
PCOd = PCOa+PCOb+PCOc;
PCOe = -0.10486*CI0;
PCO = -PCOd-PCOe+0.25708; !-logPCO;
@FREE(PCOa);@FREE(PCOb);@FREE(PCOc);@FREE(PCOd);@FREE(PCOe);@FR
EE(PCO);

!Relative Toxicity, IGC50;
logIGCa = 0.723*0.14*logKow;
logIGCb = 1.79*0.031;
logIGC = logIGCa-logIGCb;
@FREE(logIGCa);@FREE(logIGCb);@FREE(logIGC);

!Bioconcentration Factor;
logBCF = logKow-1.32;
@FREE(logBCF);
logBCF <= 3;

!Soil/water Partition Coefficient, Koc;
logKoca = 1.03*logKow;
logKoc = logKoca-0.61;
@FREE(logKoca);@FREE(logKoc);
!logKoc < 4.5;
logKoc > 3;

@FOR(SGN1:@GIN(NoS));
```

MATHEMATICAL FORMULATION FOR HEIGHT 2

```
MODEL:
!SGN2 = Signature at height 2;
!NoS = Number of Signature height 2;
!MW = Molar Weight;
!Tm = Melting Point;
!Tb = Boiling Point;
!Fp = Flash Point;
!DelS = Total Solubility Parameter;
!DelD, DelP, DelH = Hansen Solubility Parameter;
!Tc = Critical Temperature;
!Pc = Critical Pressure;
```

```
!Acc = Accentric Factor;
!CN = Cetane Number;
!Tait1, Tait2 = Auto Ignition Temperature;
!HHV = Higher Heating Value;
!Val =Valency;
!KOW = Octanol/Water Partition Coefficient;
!CI0 = 0TH Order Connectivity Index;
!CI1 = 1ST Order Connectivity Index;


SETS:
SGN2: NoS, MW, Tm, Tb, Fp, DelS, DelD, DelP, DelH, Tc, Pc, Acc,
CN, Tait1, Tait2, HHV, Carbon, Oxygen, Hydrogen, KOW, MMW,
ORDER0, ORDER1, MV, DV, VAL1, VAL2, VAL3;
ENDSETS


DATA:
SGN2 = M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 M13 M14 M15 M16
M17;
MW = 15 15 15 14 14 14 14 14 14 14 14 14 13 13 13 17 17;
Tm = 0.6953 0.6953 0.6953 0.2515 0.2515 0.2515 0.2515 0.2515
0.2515 0.2515 0.2515 0.2515 -0.373 -0.373 -0.373 2.7888 2.7888;
Tb = 0.8491 0.8491 0.8491 0.7141 0.7141 0.7141 0.7141 0.7141
0.7141 0.7141 0.7141 0.7141 0.2924 0.2924 0.2924 2.567 2.567;
Fp = 33.0909 33.0909 33.0909 11.4107 11.4107 11.4107 11.4107
11.4107 11.4107 11.4107 11.4107 11.4107 -17.7416 -17.7416 -
17.7416 87.6576 87.6576;
DelS = -1.8029 -1.8029 -1.8029 -0.1323 -0.1323 -0.1323 -0.1323
-0.1323 -0.1323 -0.1323 -0.1323 -0.1323 1.0139 1.0139 1.0139
3.0524 3.0524;
DelD = 7.5983 7.5983 7.5983 -0.0023 -0.0023 -0.0023 -0.0023 -
0.0023 -0.0023 -0.0023 -0.0023 -0.0023 -7.539 -7.539 -7.539
8.0503 8.0503;
DelP = 2.3037 2.3037 2.3037 -0.1664 -0.1664 -0.1664 -0.1664 -
0.1664 -0.1664 -0.1664 -0.1664 -0.1664 -3.3851 -3.3851 -3.3851
5.2379 5.2379;
DelH = 2.2105 2.2105 2.2105 -0.215 -0.215 -0.215 -0.215 -0.215
-0.215 -0.215 -0.215 -0.215 -2.6826 -2.6826 -2.6826 11.8005
11.8005;
Tc = 1.0898 1.0898 1.0898 3.4604 3.4604 3.4604 3.4604 3.4604
3.4604 3.4604 3.4604 3.4604 4.6659 4.6659 4.6659 10.1672
10.1672;
Pc = 0.01 0.01 0.01 0.0101 0.0101 0.0101 0.0101 0.0101 0.0101
0.0101 0.0101 0.0101 0.0107 0.0107 0.0107 -0.0071 -0.0071;
Acc = 0.0017 0.0017 0.0017 0.0019 0.0019 0.0019 0.0019 0.0019
0.0019 0.0019 0.0019 0.0019 0.0029 0.0029 0.0029 0.018 0.018;
CN = 3.49 3.49 3.49 4.879 4.879 4.879 4.879 4.879 4.879 4.879
4.879 4.879 -3.7 -3.7 -3.7 -13.769 -13.769;
Tait1 = -0.3516 -0.3516 -0.3516 0.1009 0.1009 0.1009 0.1009
0.1009 0.1009 0.1009 0.1009 0.1009 0.5829 0.5829 0.2859 -0.1516
-0.1516;
Tait2 = -57.8605 -57.8605 -57.8605 2.6047 2.6047 2.6047 2.6047
2.6047 2.6047 2.6047 2.6047 2.6047 79.2115 79.2115 79.2115
32.3056 32.3056;
HHV = 775 775 775 670 670 670 670 670 670 670 670 670 518 518
518 -108 -108;
Carbon = 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0;
Oxygen = 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1;
Hydrogen = 3 3 3 2 2 2 2 2 2 2 2 2 1 1 1 1 1;
```

```
KOW = 0.1152 0.1152 0.1152 0.4594 0.4594 0.4594 0.4594 0.4594
0.4594 0.4594 0.4594 0.4594 0.43 0.43 0.43 -1.3365 -1.3365;
ORDER0 = 1 1 1 0.7071 0.7071 0.7071 0.7071 0.7071 0.7071 0.7071
0.7071 0.7071 0.5774 0.5774 0.5774 0.4472 0.4471;
ORDER1 = 0.5774 0.7071 0.7071 1.2071 1.2071 1.1154 1 1 0.9082
0.9082 1.4142 1.2071 1.4129 1.2438 1.0747 0.3162 0.2582;
MMW = 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 16 16;
MV = 0.02614 0.02614 0.02614 0.01641 0.01641 0.01641 0.01641
0.01641 0.01641 0.01641 0.01641 0.01641 0.01641 0.00711 0.00711 0.00711
0.00551 0.00551;
DV = -1.0278 -1.0278 -1.0278 0.2125 0.2125 0.2125 0.2125 0.2125
0.2125 0.2125 0.2125 0.2125 1.318 1.318 1.318 1.3057 1.3057;
Val1 = 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1;
Val2 = 0 0 0 1 1 1 1 1 1 1 1 0 0 0 0 0;
Val3 = 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0;

ENDDATA

GROUP = @SUM(SGN2:NoS);
GROUP > 5;
GROUP < 30;

NoS(1)=NoS(13)*2+NoS(14);
NoS(2)=NoS(4)+NoS(5)+NoS(6);
NoS(3)=NoS(11);
NoS(6)+NoS(9)+NoS(10)=NoS(14)+NoS(15)*2;
NoS(12)=NoS(5)+NoS(8)+NoS(10);
NoS(16)=NoS(11)+NoS(12);
NoS(17)=NoS(13)+NoS(14)+NoS(15);

2*NoS(7) < NoS(4)+NoS(7)+NoS(8)+NoS(9);

NoS(4)+2*NoS(7)+NoS(8)+NoS(9)=2*H1;
H1>0;@GIN(H1);

!Structural Constraint
!Handshaking Dilemma;
!Valency of one;
SUMVal1 = @SUM(SGN2:NoS*Val1);
!Valency of two;
SUMVal2 = @SUM(SGN2:NoS*Val2);
!Valency of three;
SUMVal3 = @SUM(SGN2:NoS*Val3);

SUMVal1+2*SUMVal2+3*SUMVal3=(GROUP-1)*2;

!Properties Constraint;
!Normal Melting Point;
!unit = K;
SUMTm = @SUM(SGN2:NoS*Tm);
Tmelt = @LOG(SUMTm)*147.450;
Tmelt <= 298.15;

!Normal Boiling Point;
!unit = K;
SUMTb = @SUM(SGN2:NoS*Tb);
Tboil = @LOG(SUMTb)*222.543;
Tboil >= 400;

!Flash Point;
```

```
!unit = K;
SUMFp = @SUM(SGN2:NoS*Fp);
Tflash = SUMFp+150.0218;
!Tflash >= 403.15;

!Total Solubility Parameter;
!unit = MPa^0.5;
SUMDelS = @SUM(SGN2:NoS*DelS);
Sol = SUMDelS+20.7339;

!Hansen Solubility Parameter;
!unit = MPa^0.5;
SUMDelD = @SUM(SGN2:NoS*DelD);
SUMDelP = @SUM(SGN2:NoS*DelP);
SUMDelH = @SUM(SGN2:NoS*DelH);
SUMD = @SQR(SUMDelD);
SUMP = @SQR(SUMDelP);
SUMH = @SQR(SUMDelH);
Sp = @SQRT(SUMD+SUMP+SUMH);

!Molar Volume;
SUMMV = @SUM(SGN2:NoS*MV);
MolarVolume = (SUMMV+0.01211);

!Density;
Density = (SUMMW/MolarVolume);
Density > 750;
Density < 1000;

!Critical Temperature;
!unit = K;
SUMTc = @SUM(SGN2:NoS*Tc);
Tcrit = @LOG(SUMTc)*181.6738;

!Critical Pressure;
!unit = bar;
SUMPc = @SUM(SGN2:NoS*Pc);
SUMPca = SUMPc+0.1155;
SUMPcb = 1/SUMPca;
Pcrit = ((SUMPcb^2)+0.0519);

!Acentric Factor;
SUMAcc = @SUM(SGN2:NoS*Acc);
AccFac = ((@LOG(SUMAcc+1.0039))^(1/0.0447))*0.9132;

!Vapour Pressure;
!unit = atm;
Tr = Tboil/Tcrit;
@FREE(Z0);@FREE(Z1);@FREE(lnPr);
Z0 = 5.92714-(6.09648/Tr)-
(1.28862*(@LOG(Tr)))+(0.169347*(Tr^6));
Z1 = 15.2518-(15.6875/Tr)-
(13.4721*(@LOG(Tr)))+(0.43577*(Tr^6));
lnPr = Z0+(AccFac*Z1);
Psat = @EXP(lnPr)*Pcrit;
Psat > 0.01;

!Auto Ignition Temperature;
!unit = K;
SUMTait1 = @SUM(SGN2:NoS*Tait1);
```

```
SUMTait2 = @SUM(SGN2:NoS*Tait2);
Tauto = (71.2584*(10^-(SUMTait1)))+525.93+SUMTait2;
!Tauto >= 500;

!Higher Heating Value;
!unit = MJ/kg;
MAX = HHVs;
SUMHHV = @SUM(SGN2:NoS*HHV);
SUMMW = @SUM(SGN2:NoS*MW);
HHVs = SUMHHV/SUMMW;
HHVs > 30;

!Dynamic Viscosity;
SUMDV = @SUM(SGN2:NoS*DV);
DynVis = @EXP(SUMDV);

!Kinematics Viscosity;
KV = DynVis/Density*1000;
KV < 5;
KV > 1;

!Environmental Properties;
!0th Order Connectivity Index;
CI0 = @SUM(SGN2:NoS*ORDER0);
CI1 = @SUM(SGN2:NoS*ORDER1);

!Octanol/Water Partition Coefficient;
SUMKOW = @SUM(SGN2:NoS*KOW);
logKow = SUMKOW+0.7520;
expkow = @exp(SUMKOW);

!MMW;
SUMMMW = @SUM(SGN2:NoS*MMW);

!No of Atoms;
NoC = @SUM(SGN2:NoS*Carbon);
NoO = @SUM(SGN2:NoS*Oxygen);
NoH = @SUM(SGN2:NoS*Hydrogen);

!Connectivity Index;
CI0 = @SUM(SGN2:NoS*ORDER0);
CI1 = @SUM(SGN2:NoS*ORDER1);

!Acute Toxicity, LC50;
LC50a = NoC*(0.246449);
LC50b = NoO*(-0.21041);
LC50c = NoH*(-0.01047);
LC50d = LC50a+LC50b+LC50c;
LC50e = 0.243883*(CI0);
LC50f = 2*0.024195*(CI1);
LC50 = -LC50d-LC50e-LC50f+2.9717;
@FREE(LC50a);@FREE(LC50b);@FREE(LC50c);@FREE(LC50d);@FREE(LC50e
);@FREE(LC50f);@FREE(LC50);
!LC50 > 1.041;

!Acute Toxicity, Aquatic, EC50;
EC50a = 0.95*logKow;
EC50 = EC50a+1.32;
@FREE(EC50a);@FREE(EC50);
EC50 > 1.041;
```

```
!Acute Toxicity, Oral, LD50;
LD50a = NoC*(-0.052218);
LD50b = NoO*(0.004962);
LD50c = NoH*(-0.01047);
LD50d = LD50a+LD50b+LD50c;
LD50e = 0.018103*CI0;
LD50f = 2*(-0.02677)*CI1;
LD50 = 1.9372+0.0016*SUMMW-LD50d-LD50e-LD50f;
@FREE(LD50a);@FREE(LD50b);@FREE(LD50c);@FREE(LD50d);@FREE(LD50e
);@FREE(LD50f);@FREE(LD50);
!LD50 > 3.7;

!Ozone Depletion Potential;
ODPa = NoC*(-1.59188); !-1.59188;
ODPb = NoO*(-3.62951); !-3.62951;
ODPc = NoH*(-3.37822);
ODPd = ODPa+ODPb+ODPc;
ODPe = 8.967731*CI0;
ODPf = 2*0.22623*CI1;
ODP = ODPd+ODPe+ODPf+3.298083; !logODP;
@FREE(ODPa);@FREE(ODPb);@FREE(ODPc);@FREE(ODPd);@FREE(ODPe);@FR
EE(ODPf);@FREE(ODP);

!Global Warming Potential;
GWPa = NoO*(0.424024);
GWPb = NoC*(-3.05993);
GWPc = NoH*(1.67752);
GWPd = GWPa+GWPb+GWPc;
GWPe = (-0.01877)*CI0;
GWPf = 2*(-1.52848)*CI1;
GWP = GWPd+GWPe+GWPf-0.52073; !logGWP;
@FREE(GWPa);@FREE(GWPb);@FREE(GWPc);@FREE(GWPd);@FREE(GWPe);@FR
EE(GWPf);@FREE(GWP);

!Photochemical Oxidation Potential;
PCOa = NoO*(0.126631);
PCOb = NoH*(0.079188);
PCOc = NoC*(-0.03181);
PCOd = PCOa+PCOb+PCOc;
PCOe = -0.10486*CI0;
PCOf = 2*0.005087*CI1;
PCO = -PCOd-PCOe-PCOf+0.25708; !logPCO;
@FREE(PCOa);@FREE(PCOb);@FREE(PCOc);@FREE(PCOd);@FREE(PCOe);@FR
EE(PCOf);@FREE(PCO);

!Relative Toxicity, IGC50;
logIGCa = 0.723*0.14*logKow;
logIGCb = 1.79*0.031;
logIGC = logIGCa-logIGCb;
@FREE(logIGCa);@FREE(logIGCb);@FREE(logIGC);

!Bioconcentration Factor;
logBCF = logKow-1.32;
@FREE(logBCF);
logBCF <= 3;

!Soil/water Partition Coefficient, Koc;
logKoca = 1.03*logKow;
logKoc = logKoca-0.61;
```

```
@FREE(logKoca);@FREE(logKoc);
!logKoc < 4.5;
!LogKoc > 3;

@FOR(SGN2:@GIN(NoS));
```

MATHEMATICAL FORMULATION FOR HEIGHT 3

```
MODEL:
!SGN3 = Signature at Height 3;
!NoS = Number of Signature;
!MW = Molar Weight;
!Tm = Melting Point;
!Tb = Boiling Point;
!Fp = Flash Point;
!DelS = Total Solubility Parameter;
!DelD, DelP, DelH = Hansen Solubility Parameter;
!Tc = Critical Temperature;
!Pc = Critical Pressure;
!Acc = Accentric Factor;
!CN = Cetane Number;
!Tait1, Tait2 = Auto Ignition Temperature;
!HHV = Higher Heating Value;
!Val =Valency;
!KOW = Octanol/Water Partition Coefficient;
!CI0 = 0TH Order Connectivity Index;
!CI1 = 1ST Order Connectivity Index;
!CI2 = 2ND Order Connectivity Index;
!WI = WIENER INDEX;
!DV = DYNAMIC VISCOSITY;
!KV = KINEMATIC VISCOSITY;

SETS:
SGN3: NoS, MW1, MW2, Tm1, Tm2, Tb1, Tb2, Fp1, Fp2, DelS1,
DelS2, DelD1, DelD2, DelP1, DelP2, DelH1, DelH2, Tc1, Tc2, Pc1,
Pc2, Acc1, Acc2, CN1, CN2, Tait11, Tait12, Tait21, Tait22,
KOW1, KOW2, HHV, Carbon, Oxygen, Hydrogen, MMW, ORDER0, ORDER1,
ORDER2, ORDER3, Val1, Val2, Val3, WI, MV1, MV2, DV1, DV2, CH,
CH2, CH3;
ENDSETS

DATA:
SGN3 = T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 T11 T12 T13 T14;
MW1 = 15 15 14 14 14 14 14 14 14 14 14 14 13 17;
MW2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 30;
Tm1 = 0.6953 0.6953 0.2515 0.2515 0.2515 0.2515 0.2515 0.2515
0.2515 0.2515 0.2515 0.2515 -0.373 2.7888;
Tm2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.3489;
Tb1 = 0.8491 0.8491 0.7141 0.7141 0.7141 0.7141 0.7141 0.7141
0.7141 0.7141 0.7141 0.7141 0.2925 2.567;
Tb2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.2825;
Fp1 = 33.0909 33.0909 11.4107 11.4107 11.4107 11.4107 11.4107
11.4107 11.4107 11.4107 11.4107 11.4107 -17.7416 87.6576;
Fp2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 3.2443;
DelS1 = -1.8029 -1.8029 -0.1323 -0.1323 -0.1323 -0.1323 -0.1323
-0.1323 -0.1323 -0.1323 -0.1323 -0.1323 1.0139 3.0524;
DelS2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.4165;
```

```
DelD1 = 7.5983 7.5983 -0.0023 -0.0023 -0.0023 -0.0023 -0.0023 -
0.0023 -0.0023 -0.0023 -0.0023 -0.0023 -7.539 8.0503;
DelD2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0394;
DelP1 = 2.3037 2.3037 -0.1664 -0.1664 -0.1664 -0.1664 -0.1664 -
0.1664 -0.1664 -0.1664 -0.1664 -0.1664 -3.3851 5.2379;
DelP2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 0.1176;
DelH1 = 2.2105 2.2105 -0.215 -0.215 -0.215 -0.215 -0.215 -0.215
-0.215 -0.215 -0.215 -0.215 -2.6826 11.8005;
DelH2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 0.6216;
Tc1 = 1.0898 1.0898 3.4604 3.4604 3.4604 3.4604 3.4604 3.4604
3.4604 3.4604 3.4604 3.4604 4.6659 10.1672;
Tc2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.5693;
Pc1 = 0.01 0.01 0.0101 0.0101 0.0101 0.0101 0.0101 0.0101
0.0101 0.0101 0.0101 0.0101 0.0107 -0.0071;
Pc2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 0.001;
Acc1 = 0.0017 0.0017 0.0019 0.0019 0.0019 0.0019 0.0019 0.0019
0.0019 0.0019 0.0019 0.0019 0.0029 0.018;
Acc2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0004;
CN1 = 3.49 3.49 4.879 4.879 4.879 4.879 4.879 4.879 4.879 4.879
4.879 4.879 -3.7 -13.769;
CN2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.507;
Tait11 = -0.3516 -0.3516 0.1009 0.1009 0.1009 0.1009 0.1009
0.1009 0.1009 0.1009 0.1009 0.1009 0.5829 -0.1516;
Tait12 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.2087;
Tait21 = -57.8605 -57.8605 2.6047 2.6047 2.6047 2.6047 2.6047
2.6047 2.6047 2.6047 2.6047 2.6047 79.2115 32.3056;
Tait22 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -46.5762;
HHV = 775 775 670 670 670 670 670 670 670 670 670 670 518 -108;
CARBON = 1 1 1 1 1 1 1 1 1 1 1 1 1 0;
OXYGEN = 0 0 0 0 0 0 0 0 0 0 0 0 0 1;
HYDROGEN= 3 3 2 2 2 2 2 2 2 2 2 2 1 1;
ORDER0 = 1 1 0.7071 0.7071 0.7071 0.7071 0.7071 0.7071 0.7071
0.7071 0.7071 0.7071 0.5774 0.4472;
ORDER1 = 0.5774 0.7071 1.2071 1.2071 1 1 1 1 1 1 0.9082 0.9082
1.2438 0.2582;
ORDER2 = 1.2438 1.2071 0 1.2071 0 0 1.5607 0 1.4142 1.3493 0
1.5866 1.0068 0.5562;
ORDER3 = 1.0068 1.2071 0 1.9571 0 0 1.8536 0 2.0577 2.1219 0
1.6660 1.9553 0.4502;
MMW = 12 12 12 12 12 12 12 12 12 12 12 12 12 16;
Val1 = 1 1 0 0 0 0 0 0 0 0 0 0 0 1;
Val2 = 0 0 1 1 1 1 1 1 1 1 1 1 0 0;
Val3 = 0 0 0 0 0 0 0 0 0 0 0 0 1 0;
MV1 = 0.0261 0.0261 0.0164 0.0164 0.0164 0.0164 0.0164 0.0164
0.0164 0.0164 0.0164 0.0164 0.0071 0.0055;
MV2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.0009;
DV1 = -1.0278 -1.0278 0.2125 0.2125 0.2125 0.2125 0.2125 0.2125
0.2125 0.2125 0.2125 0.2125 1.3180 1.3057;
DV2 = 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.2116;
CH = 0 0 0 0 0 0 0 0 0 0 0 0 1 0;
CH2 = 0 0 1 1 1 1 1 1 1 1 1 1 0 0;
CH3 = 1 1 0 0 0 0 0 0 0 0 0 0 0 0;
ENDDATA

NoCH = @SUM(SGN3:NoS*CH);
NoCH2 = @SUM(SGN3:NoS*CH2);
NoCH3 = @SUM(SGN3:NoS*CH3);
SUMCgrp = NoCH2+NoCH3-NoCH;
MAX = SUMCgrp;
```

```
GROUP = @SUM(SGN3:NoS);
GROUP > 5;
GROUP < 30;

NoS(14) > 0;

NoS(1) = NoS(13);
NoS(2) = NoS(3)+NoS(4)+NoS(5);
NoS(4) = 2*NoS(6)+NoS(7)+NoS(8);
NoS(11) = NoS(5);
NoS(12) = NoS(8)+NoS(10);
NoS(13) = NoS(11)+NoS(12);
NoS(14) = NoS(13);

NoS(7)+2*NoS(9)+NoS(10)=2*H1;
@GIN(H1); H1>0;

2*NoS(9) < NoS(7)+NoS(9)+NoS(10);

!Edges constraint:
!Valency of one;
SUMVal1 = @SUM(SGN3:NoS*Val1);
!Valency of two;
SUMVal2 = @SUM(SGN3:NoS*Val2);
!Valency of three;
SUMVal3 = @SUM(SGN3:NoS*Val3);
SUMVal1+2*SUMVal2+3*SUMVal3=(GROUP-1)*2;

!Properties Constraint;
!Normal Melting Point;
!unit = K;
SUMTm1 = @SUM(SGN3:NoS*Tm1);
SUMTm2 = @SUM(SGN3:NoS*Tm2);
@FREE(SUMTm2);
SUMTm = SUMTm1+SUMTm2;
SUMTm <=7.55;

!Normal Boiling Point;
!unit = K;
SUMTb1 = @SUM(SGN3:NoS*Tb1);
SUMTb2 = @SUM(SGN3:NoS*Tb2);
@FREE(SUMTb2);
SUMTb = SUMTb1+SUMTb2;
SUMTb >= 6.034;

!Flash Point;
!unit = K;
SUMFp1 = @SUM(SGN3:NoS*Fp1);
SUMFp2 = @SUM(SGN3:NoS*Fp2);
@FREE(SUMFp2);
SUMFp = SUMFp1+SUMFp2;
Tflash = SUMFp+150.0218;

!Molar Volume;
SUMMV1 = @SUM(SGN3:NoS*MV1);
SUMMV2 = @SUM(SGN3:NoS*MV2);
@FREE(SUMMV2);
SUMMV = SUMMV1+SUMMV2;
MV = (SUMMV+0.01211);
```

```
!Density;
Density = (SUMMW/MV);

!Vapour Pressure;
!unit = mmhg;
Psata = (Tboil/298.15)^1.7;
Psatb = 2.7*Psata;
Psat = 5.58-Psatb;
Psat < 76;

!Higher Heating Value;
!unit = MJ/kg;
SUMHHV = @SUM(SGN3:NoS*HHV);
SUMMW = @SUM(SGN3:NoS*MW1);
HHVs = SUMHHV/SUMMW;

!Dynamic Viscosity;
SUMDV1 = @SUM(SGN3:NoS*DV1);
SUMDV2 = @SUM(SGN3:NoS*DV2);
@FREE(SUMDV2);
DV = SUMDV1+SUMDV2;
DV <=1.44;

!MMW;
SUMMMW = @SUM(SGN3:NoS*MMW);

!No of Atoms;
NoC = @SUM(SGN3:NoS*Carbon);
NoO = @SUM(SGN3:NoS*Oxygen);
NoH = @SUM(SGN3:NoS*Hydrogen);

!Connectivity Index;
CI0 = @SUM(SGN3:NoS*ORDER0);
CI1 = @SUM(SGN3:NoS*ORDER1);
CI2 = @SUM(SGN3:NoS*ORDER2);
CI3 = @SUM(SGN3:NoS*ORDER3);

!Octanol/Water Partition Coefficient;
!logKow = 1.297+0.539*(CI1);
logKow = 1.267*CI1+0.612*CI3-0.976*CI3-2.13;
Kow = 10^logKow;

!Acute Toxicity, LC50;
LC50a=0.81*logKow+1.744;
LC50b=1/LC50a;
LC50=10^(LC50b)*SUMMW*1000;
@FREE(LC50a);@FREE(LC50b);@FREE(LC50);
LC50 > 11;

!Acute Toxicity, Aquatic, EC50;
EC50a = 0.95*logKow;
EC50 = EC50a+1.32;
@FREE(EC50a);@FREE(EC50);
EC50 > 2;

!Acute Toxicity, Oral, LD50;
LD50a=0.805*logKow;
LD50b=0.0807*Kow+1;
LD50c=@LOG(LD50b);
LD50d=0.971*LD50c;
```

```
LD50e=LD50a-LD50d+0.984;
LD50f=1/LD50e;
LD50=10^(LD50f)*SUMMW*1000;
LD50>100;
@FREE(LD50a);@FREE(LD50b);@FREE(LD50c);@FREE(LD50d);@FREE(LD50e
);@FREE(LD50f);@FREE(LD50);

!Relative Toxicity, IGC50;
logIGCa = 0.723*0.14*logKow;
logIGCb = 1.79*0.031;
logIGC = logIGCa-logIGCb;
@FREE(logIGCa);@FREE(logIGCb);@FREE(logIGC);

!Bioconcentration Factor;
logBCF=0.032+0.636*logKow;
logBCF <= 3;

!Soil/water Partition Coefficient, Koc;
logKoc=0.59*CI1-0.97;
logKoc < 4;

!Ozone Depletion Potential;
ODPa = NoC*(-1.59188);
ODPb = NoO*(-3.62951);
ODPc = NoH*(-3.37822);
ODPd = ODPa+ODPb+ODPc;
ODPe = 8.967731*CI0;
ODPf = 2*0.22623*CI1;
ODP = ODPd+ODPe+ODPf+3.298083; !logODP;
@FREE(ODPa);@FREE(ODPb);@FREE(ODPc);@FREE(ODPd);@FREE(ODPe);@FR
EE(ODPf);@FREE(ODP);
ODP < 0;

!Global Warming Potential;
GWPa = NoO*(0.424024);
GWPb = NoC*(-3.05993);
GWPc = NoH*(1.67752);
GWPd = GWPa+GWPb+GWPc;
GWPe = (-0.01877)*CI0;
GWPf = 2*(-1.52848)*CI1;
GWP = GWPd+GWPe+GWPf-0.52073; !logGWP;
@FREE(GWPa);@FREE(GWPb);@FREE(GWPc);@FREE(GWPd);@FREE(GWPe);@FR
EE(GWPf);@FREE(GWP);
GWP < 0;

!Photochemical Oxidation Potential;
PCOa = NoO*(0.126631);
PCOb = NoH*(0.079188);
PCOc = NoC*(-0.03181);
PCOd = PCOa+PCOb+PCOc;
PCOe = -0.10486*CI0;
PCOf = 2*0.005087*CI1;
PCO = -PCOd-PCOe-PCOf+0.25708; !logPCO;
@FREE(PCOa);@FREE(PCOb);@FREE(PCOc);@FREE(PCOd);@FREE(PCOe);@FR
EE(PCOf);@FREE(PCO);
PCO < 0;

@FOR(SGN3:@GIN(NoS));
```

APPENDIX B: CASE STUDY IN CHAPTER 5

MATHEMATICAL FORMULATION FOR MOO MODEL

```
MODEL:
!FUZZY MOO MODEL
!SOLV = SOLVENT CANDIDATE
!RATIO = RATIO OF SOLVENT:BIO-OIL
!HHV = HIGHER HEATING VALUES
!HHVC = HIGHER HEATING VALUES (CUSTOMER)
!HHVS = HHV OF SOLVENTS
!HHVB = HHV OF BIOOIL
!AP = PRODUCT PRICE
!AC = PRODUCT PRICE (CUSTOMER)
!TP = ANNUAL PRODUCTION/DEMAND
!TC = ANNUAL PRODUCTION/DEMAND (CUSTOMER)
!P = ELASTICITY OF SUBSTITUTION
!ALPHA = MARKETING BUDEGET
!BETA = OPTIMISATION POINT
!Y = TOTAL MARKET SIZE
!PM = PRICING MODEL;

SETS:
SOLV: NoS, HHVS, SOLVCOST;
ENDSETS

DATA:
SOLV = S1 S2 S3 S4;
HHVS = 40.89 40 38.92 37.5;
SOLVCOST= 2640 3000 395802.5 100985.2;
ENDDATA

NoSolv = @SUM(SOLV:NoS);
NoSolv = 1;
@FOR(SOLV:@GIN(NoS));

nos(1)=0;

!PRICING MODEL;
PM1 = ALPHA/BETA;
PM2 = (Y-(AP*TP))/AC;
PM3 = 1-P;

AP*TP = AC*(TP^P)*(PM1^P)*(PM2^PM3);

AC = 895.285; !USD/tonne blend;
TP = 36000; !tonne/yr;
TC = 27000; !tonne/yr;
Y = 500000000; !$/yr;
Y > AC*TC+AP*TP;
ALPHA = 0.85;
BETA = 27.55/HHV;
!HHVC = 27.55 MJ/kg;
```

```
BETA < 0.99;
P = 0.10;

!PRODUCT PRICE;
AP = PROFIT+PRODUCTION+RAWMATERIAL;
RAWMATERIAL = @SUM(SOLV:NoS*SOLVCOST)*RATIO;
PRODUCTION = 80.30706734*(1-RATIO);
!BIOOILCOST = 70.67458694 USD/tonne bio-oil;
!MAX = PROFIT; !USD/tonne bio-oil blend;

RATIO < 0.5;

!HHV OF BLEND;
HHVSolv = RATIO*@SUM(SOLV:NoS*HHVS);
HHV = HHVSolv+HHVB*(1-(RATIO));
HHVB = 19; !MJ/kg;
!MAX = HHV;

!FUZZY;
LAMDA1 = (PROFIT-4626.725)/(5033.319-4626.725);
LAMDA2 = (HHV-27.82828)/(29.945-27.82828);
LAMDA <= LAMDA1;
LAMDA <= LAMDA2;
MAX = LAMDA;
```

APPENDIX C: CASE STUDY IN CHAPTER 6

Table C-1 Pyrolysis bio-oil database

No	Biomass Type	Carbon Content (%)	Hydrogen Content (%)	Oxygen Content (%)	Nitrogen Content (%)	Sulphur Content (%)	Volatile Matter (%)	Ash Content (%)	Moisture Content (%)	Fixed Carbon (%)	HHV of Biomass (MJ/kg)	Pyrolysis Temperature (ºC)	HHV of Bio-Oil (MJ/kg)		pH of Bio-Oil		Ref
													Actual	Decision	Actual	Decision	
1	EFB	49.10	5.80	41.10	0.50	-	65.00	1.50	15.20	18.30	16.40	500	13.20	1	-	-	1
2	Sawdust	43.20	7.90	45.10	1.00	0.60	70.00	5.90	9.60	14.50	17.60	500	15.90	1	-	-	1
3	Giant Miscanthus	51.30	6.20	42.20	0.40	-	67.50	1.50	11.20	19.80	19.20	500	17.00	1	-	-	1
4	Cherry Stones	51.08	6.49	42.03	0.38	0.02	73.90	0.20	-	25.90	22.50	600	13.90	1	2.6	2	2
5	Acacia cincinnata phyllodes	49.90	6.04	41.79	2.13	0.14	70.16	2.48	7.45	19.91	19.72	500	30.65	3	3.1	3	3
6	Acacia cincinnata trunk	46.44	5.98	47.35	0.23	-	74.40	1.13	8.12	16.35	18.74	500	24.36	2	2.5	2	3
7	Acacia holosericea phyllodes	51.05	6.53	40.22	2.20	-	70.15	1.58	6.68	21.59	19.11	500	28.13	2	3.2	3	3
8	Acacia holosericea trunk	44.03	5.67	50.05	0.25	-	68.98	1.31	9.75	19.96	17.67	500	23.46	2	2.8	2	3
9	Coconut shell	64.23	6.89	27.61	0.77	0.50	75.50	3.20	10.10	11.20	20.15	575	19.75	1	-	-	4
10	Oil palm trunk	54.22	7.10	38.36	0.06	0.26	82.00	2.97	49.27	15.03	17.73	600	19.24	1	3.2	3	5
11	Oil palm trunk	59.77	7.28	32.61	0.06	0.28	82.05	2.07	72.90	15.88	17.91	600	21.76	2	3.2	3	5
12	Oil palm trunk	60.17	7.93	31.56	0.06	0.28	80.11	5.53	77.52	14.36	17.33	600	20.15	2	3.2	3	5
13	Oil palm fronds	45.05	5.86	48.82	0.23	0.04	80.88	3.61	11.71	15.51	17.55	600	21.92	2	3.1	3	5
14	Oil palm fronds	50.68	6.12	40.67	0.90	1.63	79.69	6.51	31.92	13.80	18.60	600	21.22	2	2.9	2	5
15	Oil Palm EFB	45.26	5.13	47.29	1.82	0.50	75.01	4.76	8.44	20.23	16.07	600	21.61	2	3.2	3	5
16	Sal seed	68.69	23.68	33.12	1.23	3.28	71.54	6.00	8.46	16.00	21.64	600	23.75	2	3.7	3	6

Table C-1 Pyrolysis bio-oil database (continued)

17	Napier grass stem	48.60	6.01	44.10	0.99	0.32	81.50	1.75	75.30	16.70	18.10	450	25.30	2	2.4	2	7
18	Napier grass stem	48.60	6.01	44.10	0.99	0.32	81.50	1.75	75.30	16.70	18.10	550	25.40	2	3.0	2	7
19	Napier grass stem	48.60	6.01	44.10	0.99	0.32	81.50	1.75	75.30	16.70	18.10	600	28.50	2	3.0	2	7
20	Napier grass stem	48.60	6.01	44.10	0.99	0.32	81.50	1.75	75.30	16.70	18.10	650	28.90	2	2.8	2	7
21	Bagasse	48.58	5.97	38.94	0.20	0.05	69.00	1.26	4.00	29.70	19.20	600	19.91	1	4.5	4	8
22	Jatropha Curcas Cake	45.50	6.70	45.33	2.47	-	18.33	6.93	7.17	-	17.05	550	24.91	2	-	-	9
23	Corn stalk	48.72	5.64	34.68	0.91	0.09	73.46	4.49	5.47	16.58	17.50	450	21.38	2	3.3	3	10
24	Linseed	64.99	9.94	21.26	3.58	0.23	84.90	4.10	6.90	4.10	22.06	550	33.83	3	3.1	3	11
25	Sewage sludge	20.34	3.39	15.05	3.00	0.44	41.14	57.78	-	1.08	-	550	32.97	3	-	-	12
26	Pearl Millet	43.66	6.30	49.54	0.45	0.06	61.29	34.51	3.61	0.59	17.53	400	23.38	2	-	-	13
27	Sida cordifolia	41.38	6.44	50.89	1.15	0.14	65.15	10.31	6.22	18.32	16.38	400	17.15	1	-	-	13
28	Date tree mixture waste	51.88	6.56	41.56	-	-	72.51	7.54	11.63	8.32	12.66	500	24.35	2	3.0	3	14
29	Date seed	70.92	10.45	18.63	-	-	81.13	5.29	6.02	7.55	20.27	500	29.06	2	3.0	3	14
30	Eucalyptus bark	38.70	4.50	54.90	0.30	1.60	76.10	4.20	10.70	19.70	15.70	400	12.45	1	2.8	2	15
31	Eucalyptus bark	38.70	4.50	54.90	0.30	1.60	76.10	4.20	10.70	19.70	15.70	450	13.89	1	2.8	2	15
32	Eucalyptus bark	38.70	4.50	54.90	0.30	1.60	76.10	4.20	10.70	19.70	15.70	500	12.23	1	2.9	2	15
33	Eucalyptus bark	38.70	4.50	54.90	0.30	1.60	76.10	4.20	10.70	19.70	15.70	550	12.77	1	2.4	2	15
34	Pine Sawdust	47.21	6.25	44.40	0.05	0.21	73.52	1.88	7.84	16.76	19.41	500	23.83	2	3.6	3	16
35	Pine wood	51.30	5.83	35.99	0.07	0.01	78.54	0.46	6.34	14.66	18.60	450	22.49	2	3.1	3	17
36	Corrugated cardboard	43.24	5.80	-	0.12	-	-	4.00	-	13.10	18.00	350	21.20	2	3.2	3	18
37	Corrugated cardboard	43.24	5.80	-	0.12	-	-	4.00	-	13.10	18.00	400	21.50	2	3.3	3	18
38	Corrugated cardboard	43.24	5.80	-	0.12	-	-	4.00	-	13.10	18.00	450	21.70	2	4.1	4	18
39	Palm kernel shell	44.56	5.22	49.77	0.40	0.05	82.50	6.70	9.40	1.40	15.60	490	17.90	1	3.3	3	19
40	Rice straw	46.17	6.29	43.55	3.99	-	72.91	16.81	-	-	-	470	17.07	1	3.8	3	20

Table C-1 Pyrolysis bio-oil database (continued)

41	Rice straw	46.17	6.29	43.55	3.99	-	72.91	16.81	-	-	-	480	17.01	1	3.8	3	20
42	Rice straw	46.17	6.29	43.55	3.99	-	72.91	16.81	-	-	-	490	17.15	1	3.7	3	20
43	Rice straw	46.17	6.29	43.55	3.99	-	72.91	16.81	-	-	-	500	17.37	1	2.9	2	20
44	Rice straw	46.17	6.29	43.55	3.99	-	72.91	16.81	-	-	-	510	16.95	1	3.2	3	20
45	Rice straw	46.17	6.29	43.55	3.99	-	72.91	16.81	-	-	-	515	16.76	1	3.6	3	20
46	Rice husk	50.02	6.11	38.71	5.17	-	70.85	11.57	-	-	-	500	18.53	1	3.7	3	20
47	Bagasse	48.40	5.96	41.30	4.34	-	83.90	2.09	-	-	-	500	18.99	1	3.4	3	20
48	Corn cob	47.07	6.30	43.79	2.84	-	76.61	2.15	-	-	-	500	19.36	1	3.1	3	20
49	Corn cob	39.76	5.15	43.68	0.30	-	71.80	2.41	8.64	17.15	16.19	550	18.80	1	2.8	2	21
50	Tomato peel	55.00	7.90	34.00	2.80	0.30	78.32	4.87	4.67	12.34	22.50	600	33.04	3	-	-	22
51	Camellia oleifera shell	46.46	6.03	45.11	0.32	-	74.12	1.55	-	24.33	-	400	21.37	2	2.1	2	23
52	Camellia oleifera shell	46.46	6.03	45.11	0.32	-	74.12	1.55	-	24.33	-	500	23.77	2	2.4	2	23
53	Camellia oleifera shell	46.46	6.03	45.11	0.32	-	74.12	1.55	-	24.33	-	600	21.75	2	2.9	2	23
54	Stillingia oil	77.49	7.37	12.57	0.94	-	-	-	-	-	-	400	37.46	3	4.4	4	23
55	Stillingia oil	77.49	7.37	12.57	0.94	-	-	-	-	-	-	500	36.66	3	3.2	3	23
56	Stillingia oil	77.49	7.37	12.57	0.94	-	-	-	-	-	-	600	39.68	3	4.7	4	23
57	Blue-green algae blooms	42.26	6.27	43.07	7.88	0.52	70.13	6.15	9.29	14.14	16.20	500	31.90	3	-	-	24
58	Chlorella vulgaris	42.51	6.77	27.95	6.64	-	66.56	15.64	6.18	11.62	16.80	500	24.57	2	-	-	25
59	Arecanut stalk	47.12	5.95	43.54	3.39	-	68.70	5.90	9.80	15.60	-	450	22.75	2	-	-	26
60	Cotton stalk	46.77	5.76	41.71	5.75	-	62.07	8.20	10.20	19.53	-	450	19.45	1	-	-	26
61	Redgram stalk	44.11	5.69	41.97	8.23	-	72.09	10.62	5.91	11.38	-	450	24.26	2	-	-	26
62	Soybean stalk	43.61	5.45	44.66	6.28	-	58.43	13.59	11.84	16.14	-	450	16.87	1	-	-	26
63	Paddy straw	41.62	5.39	45.12	7.87	-	65.47	15.50	8.66	10.37	-	450	19.37	1	-	-	26
64	Arecanut husk	47.89	5.93	43.10	3.08	-	62.97	5.60	12.50	18.93	-	450	24.75	2	-	-	26
65	Jatropha husk	42.82	5.43	46.34	5.41	-	60.70	14.30	11.30	13.70	-	450	22.70	2	-	-	26
66	Rice husk	47.79	6.04	44.16	2.02	-	73.80	4.57	5.83	15.80	-	450	24.44	2	-	-	26
67	Wheat husk	48.29	6.22	44.87	0.62	-	68.10	1.60	13.90	16.40	-	450	32.34	3	-	-	26

Table C-1 Pyrolysis bio-oil database (continued)

68	Cashew nut shells	46.02	5.83	42.71	5.44	-	70.44	7.90	6.93	14.73	-	450	28.90	2	-	-	26
69	Coconut shell	48.88	6.05	43.82	1.25	-	71.99	3.82	4.40	19.79	-	450	20.45	2	-	-	26
70	Tamarind shell	45.00	5.73	44.13	5.14	-	68.56	9.56	8.44	13.44	-	450	19.79	1	-	-	26
71	Maize cob	48.54	6.11	44.62	0.73	-	64.80	3.31	15.20	16.69	-	450	30.79	3	-	-	26
72	Sugarcane bagasse	45.72	5.84	42.88	5.56	-	70.04	8.02	8.50	13.44	-	450	21.02	2	-	-	26
73	Coir pith	47.25	5.97	43.74	3.04	-	69.35	5.56	9.65	15.44	-	450	24.95	2	-	-	26
74	Sawdust	46.81	5.96	43.79	3.44	-	73.00	5.95	7.01	14.04	-	450	24.32	2	-	-	26
75	Cumbu Napier grass	49.00	6.11	44.44	0.45	-	71.43	3.03	7.07	18.47	-	450	24.95	2	-	-	26
76	Blue buffel grass	43.92	5.71	42.30	8.07	-	69.47	10.46	10.14	9.93	-	450	33.58	3	-	-	26
77	Polyalthia longifolia	47.12	5.95	43.54	3.39	-	61.64	13.11	8.86	16.39	-	450	20.19	2	-	-	26
78	Melia dubia wood	46.77	5.76	41.71	5.75	-	65.82	7.61	9.80	16.77	-	450	32.15	3	-	-	26
79	Prosopis wood	44.11	5.69	41.97	8.23	-	71.25	5.87	7.77	15.11	-	450	20.78	2	-	-	26
80	Samanea saman fruit	43.61	5.45	44.66	6.28	-	75.27	4.30	3.47	16.96	-	450	28.24	2	-	-	26
81	Chlorella	41.62	5.39	45.12	7.87	-	62.20	12.56	9.05	16.19	-	450	25.97	2	-	-	26
82	Parthenium hysterophorus	47.89	5.93	43.10	3.08	-	69.14	14.47	6.08	10.31	-	450	32.43	3	-	-	26
83	Pinfed computer paper	42.82	5.43	46.34	5.41	-	79.14	5.36	8.26	7.24	-	450	29.39	2	-	-	26
84	Washed rice straw	40.90	8.50	49.40	0.01	0.13	70.22	10.28	8.18	11.33	20.00	412	17.30	1	4.2	4	27
85	Algae from lake blooms	50.78	7.30	38.48	3.45	-	68.52	4.35	7.52	19.61	19.95	300	19.19	1	-	-	28
86	Algae from lake blooms	50.78	7.30	38.48	3.45	-	68.52	4.35	7.52	19.61	19.95	400	20.26	2	-	-	28
87	Algae from lake blooms	50.78	7.30	38.48	3.45	-	68.52	4.35	7.52	19.61	19.95	500	21.25	2	-	-	28
88	Algae from lake blooms	50.78	7.30	38.48	3.45	-	68.52	4.35	7.52	19.61	19.95	600	21.42	2	-	-	28
89	Algae from lake blooms	50.78	7.30	38.48	3.45	-	68.52	4.35	7.52	19.61	19.95	700	21.57	2	-	-	28
90	Pine Sawdust	55.02	6.08	38.81	0.09	-	75.38	1.10	7.80	15.72	20.17	500	18.10	1	-	-	28

Table C-1 Pyrolysis bio-oil database (continued)

91	Woods	51.30	5.70	40.50	0.21	-	84.30	2.70	9.10	13.00	19.00	550	24.00	2	3.2	3	29
92	Straw	45.70	6.00	40.90	1.40	-	75.80	5.80	7.90	18.50	18.40	550	23.70	2	3.8	3	29
93	Lignin	57.80	5.70	23.60	1.20	-	61.20	12.10	4.70	26.70	22.80	550	29.70	2	3.9	3	29
94	Ulva lactuca	33.60	5.10	28.20	3.30	-	67.10	29.10	8.90	6.50	12.90	550	25.70	2	4.3	4	29
95	Almond shell	50.50	6.58	42.65	0.21	0.01	80.28	0.55	3.30	15.87	18.20	300	13.70	1	5.5	5	30
96	Almond shell	50.50	6.58	42.65	0.21	0.01	80.28	0.55	3.30	15.87	18.20	400	14.10	1	5.5	5	30
97	Almond shell	50.50	6.58	42.65	0.21	0.01	80.28	0.55	3.30	15.87	18.20	500	12.40	1	5.5	5	30
98	Almond shell	50.50	6.58	42.65	0.21	0.01	80.28	0.55	3.30	15.87	18.20	600	12.00	1	5.5	5	30
99	Almond shell	50.50	6.58	42.65	0.21	0.01	80.28	0.55	3.30	15.87	18.20	700	11.60	1	5.5	5	30
100	Almond shell	50.50	6.58	42.65	0.21	0.01	80.28	0.55	3.30	15.87	18.20	800	11.30	1	5.5	5	30
101	Cassava rhizome	51.40	7.30	40.20	1.00	0.10	81.50	3.60	1.80	14.90	20.30	475	24.80	2	3.3	3	31
102	Cassava stalk	55.80	8.10	34.70	1.30	0.10	81.20	5.20	2.40	13.80	21.50	469	24.30	2	3.4	3	31
103	Forest residue	51.40	6.00	40.00	0.50	-	76.70	2.10	9.00	8.10	20.80	520	15.50	1	2.6	2	32
104	Rice husk	48.59	8.27	40.47	2.67	-	71.47	7.45	2.69	18.39	17.21	550	24.74	2	4.2	4	33
105	Palm kernel shell	45.10	5.10	49.20	0.56	0.04	74.00	3.00	12.00	23.00	-	550	23.48	2	2.5	2	34
106	Palm shell	49.74	5.32	44.86	0.08	0.16	67.20	2.10	11.00	19.70	-	500	6.58	0	2.5	2	35
107	Motorcycle tire	75.50	6.75	15.50	0.81	1.44	57.50	20.12	1.53	20.85	29.18	475	42.00	4	4.4	4	36
108	Jute stick	49.79	6.02	41.37	0.19	0.05	77.00	0.62	3.00	22.00	19.70	500	18.25	1	4.0	4	37
109	Maize stalk	49.10	6.10	43.70	0.70	0.11	71.95	8.33	7.67	12.05	15.07	500	19.60	1	3.2	3	38
110	Rice Straw	50.93	6.04	41.61	0.83	0.23	76.85	9.54	13.61	-	16.35	400	18.43	1	2.9	2	39
111	Rice Straw	50.93	6.04	41.61	0.83	0.23	76.85	9.54	13.61	-	16.35	500	20.40	2	2.8	2	39
112	Rice Straw	50.93	6.04	41.61	0.83	0.23	76.85	9.54	13.61	-	16.35	700	5.06	0	3.3	3	39
113	Sugarcane bagasse	58.14	6.05	34.57	0.69	0.19	79.59	4.34	16.07	-	18.61	400	3.79	0	2.1	2	39
114	Sugarcane bagasse	58.14	6.05	34.57	0.69	0.19	79.59	4.34	16.07	-	18.61	500	4.43	0	2.3	2	39
115	Sugarcane bagasse	58.14	6.05	34.57	0.69	0.19	79.59	4.34	16.07	-	18.61	700	3.72	0	2.0	2	39
116	Coconut shell	63.45	6.73	28.27	0.43	0.17	85.36	3.38	11.26	-	22.83	400	7.75	0	3.2	3	39

Table C-1 Pyrolysis bio-oil database (continued)

117	Coconut shell	63.45	6.73	28.27	0.43	0.17	85.36	3.38	11.26	-	22.83	500	8.09	0	3.1	3	39
118	Coconut shell	63.45	6.73	28.27	0.43	0.17	85.36	3.38	11.26	-	22.83	700	10.32	1	3.3	3	39
119	Paddy husk	42.78	5.77	51.18	0.34	-	74.53	11.98	1.37	12.11	15.75	350	8.24	0	2.1	2	40
120	Paddy husk	42.78	5.77	51.18	0.34	-	74.53	11.98	1.37	12.11	15.75	400	8.76	0	2.2	2	40
121	Paddy husk	42.78	5.77	51.18	0.34	-	74.53	11.98	1.37	12.11	15.75	450	10.15	1	2.7	2	40
122	Paddy husk	42.78	5.77	51.18	0.34	-	74.53	11.98	1.37	12.11	15.75	500	9.14	0	2.5	2	40
123	Paddy husk	42.78	5.77	51.18	0.34	-	74.53	11.98	1.37	12.11	15.75	550	8.56	0	2.3	2	40
124	Paddy husk	42.78	5.77	51.18	0.34	-	74.53	11.98	1.37	12.11	15.75	600	8.05	0	2.2	2	40
125	Pine Sawdust	50.50	6.40	43.00	0.10	-	84.00	0.10	9.50	-	20.40	500	16.90	1	2.7	2	41
126	Brown forest residue	51.10	5.90	43.00	0.50	-	73.20	3.80	4.90	-	20.50	500	16.90	1	3.2	3	41
127	Green forest residue	51.40	6.00	42.00	0.50	-	76.70	2.10	8.10	-	20.80	500	16.70	1	-	-	41
128	Eucalyptus crandis	50.10	6.00	44.00	0.10	-	82.70	0.40	7.60	-	19.90	500	17.30	1	2.2	2	41
129	Barley straw	48.80	5.90	45.00	0.80	-	73.90	5.80	8.80	-	18.50	500	11.10	1	3.7	3	41
130	Timothy hay	47.60	6.10	46.00	0.70	-	76.60	3.20	3.90	-	19.10	500	13.30	1	3.4	3	41
131	RCG	45.70	5.60	48.00	0.90	-	81.30	1.10	10.90	-	19.70	500	16.00	1	3.6	3	41
132	Sewage sludge	40.60	7.10	41.20	7.70	3.30	54.20	37.20	5.60	8.60	11.10	450	18.40	1	8.5	6	42
133	Sewage sludge	40.60	7.10	41.20	7.70	3.30	54.20	37.20	5.60	8.60	11.10	500	18.80	1	8.5	6	42
134	Sewage sludge	40.60	7.10	41.20	7.70	3.30	54.20	37.20	5.60	8.60	11.10	600	17.60	1	8.5	6	42
135	Waste tyre	86.40	8.00	3.40	0.50	1.70	62.20	7.10	1.30	29.40	40.00	450	42.60	4	-	-	43
136	Waste tyre	86.40	8.00	3.40	0.50	1.70	62.20	7.10	1.30	29.40	40.00	475	42.90	4	-	-	43
137	Waste tyre	86.40	8.00	3.40	0.50	1.70	62.20	7.10	1.30	29.40	40.00	500	42.10	4	-	-	43
138	Waste tyre	86.40	8.00	3.40	0.50	1.70	62.20	7.10	1.30	29.40	40.00	525	42.40	4	-	-	43
139	Waste tyre	86.40	8.00	3.40	0.50	1.70	62.20	7.10	1.30	29.40	40.00	560	42.10	4	-	-	43
140	Waste tyre	86.40	8.00	3.40	0.50	1.70	62.20	7.10	1.30	29.40	40.00	600	41.20	4	-	-	43
141	Waste tyre	86.39	6.91	1.42	0.46	0.12	64.91	4.70	0.85	29.54	16.60	440	43.20	4	-	-	44
142	Plastic	83.93	12.84	0.80	-	-	96.88	2.43	0.41	0.28	37.00	550	43.50	4	5.5	5	45

Table C-1 Pyrolysis bio-oil database (continued)

143	Scrap tyre	80.30	5.18	10.33	-	-	62.70	4.17	0.82	32.31	33.30	450	41.50	4	4.3	4	45
144	Wastepaper	39.71	7.14	53.15	-	-	76.31	6.03	6.51	11.15	17.10	450	13.19	1	1.5	1	45
145	Sewage sludge	27.70	4.40	21.90	3.90	0.80	46.10	41.30	6.60	6.00	11.90	530	22.80	2	7.9	6	46
146	Willow	47.78	5.90	46.01	0.31	-	-	1.30	7.80	-	18.90	500	18.40	1	2.7	2	47
147	Switchgrass	44.77	5.79	49.13	0.31	-	-	4.30	8.30	-	17.30	500	16.40	1	2.9	2	47
148	Washed switchgrass	47.14	6.08	46.71	0.07	-	-	3.40	46.10	-	18.70	500	16.00	1	2.8	2	47
149	Reed canary grass	45.36	5.81	48.59	0.34	-	-	5.10	7.90	-	17.60	500	17.10	1	3.0	3	47
150	Straw	44.94	5.75	48.84	0.47	-	-	6.30	9.10	-	17.30	500	13.60	1	3.5	3	47
151	Dactylis glomerata	42.96	5.70	49.44	1.90	-	-	7.50	49.30	-	16.80	500	17.60	1	3.1	3	47
152	Festuca arundinacea	42.22	5.64	50.65	1.50	-	-	7.30	50.30	-	16.40	500	16.70	1	3.2	3	47
153	Washed festuca arundinacea	45.20	5.98	47.95	0.87	-	-	4.40	47.90	-	17.80	500	21.70	2	3.0	3	47
154	Lolium perenne	43.12	5.80	49.80	1.28	-	-	6.20	49.40	-	17.00	500	15.80	1	3.2	3	47
155	Sugarcane leaves	48.90	6.50	44.40	0.20	-	79.00	5.70	6.70	8.60	18.40	450	26.70	2	3.3	3	48
156	Sugarcane tops	49.00	6.60	43.80	0.60	-	74.90	6.00	6.60	12.50	18.30	450	29.30	2	3.2	3	48
157	Cassava stalk	48.80	6.70	43.40	1.10	-	69.70	7.10	8.50	14.70	18.10	450	19.00	1	3.3	3	48
158	Cassava rhizome	49.50	6.50	42.90	1.10	-	65.00	11.20	8.80	15.00	17.10	450	26.70	2	3.5	3	48
159	beetle-killed lodgepole pine	49.62	6.37	43.29	0.33	-	82.29	0.26	7.30	10.15	20.34	600	16.86	1	-	-	49
160	Poplar	44.20	6.30	48.80	0.70	-	75.40	0.50	9.30	14.80	17.40	425	16.30	1	-	-	50
161	Poplar	44.20	6.30	48.80	0.70	-	75.40	0.50	9.30	14.80	17.40	435	16.10	1	-	-	50
162	Poplar	44.20	6.30	48.80	0.70	-	75.40	0.50	9.30	14.80	17.40	455	17.40	1	-	-	50
163	Poplar	44.20	6.30	48.80	0.70	-	75.40	0.50	9.30	14.80	17.40	485	17.20	1	-	-	50
164	Poplar	44.20	6.30	48.80	0.70	-	75.40	0.50	9.30	14.80	17.40	505	18.20	1	-	-	50
165	Poplar	44.20	6.30	48.80	0.70	-	75.40	0.50	9.30	14.80	17.40	525	18.40	1	-	-	50

Table C-1 Pyrolysis bio-oil database (continued)

166	Forest pinewood waste	49.33	6.06	44.57	0.04	-	73.40	0.50	9.40	16.70	19.80	500	14.60	1	2.8	2	51
167	Wheat straw	43.40	5.40	50.10	1.10	-	74.40	7.70	1.50	16.40	-	400	11.70	1	-	-	52
168	Acid treated wheat straw	44.40	5.60	44.60	5.40	-	80.50	6.90	2.00	10.60	-	400	17.00	1	-	-	52
169	Maize stalk	40.72	5.68	47.51	0.92	0.12	62.94	4.20	1.14	30.24	15.87	490	17.80	1	-	-	53
170	Palm shell	49.74	5.32	44.86	0.08	0.16	67.20	2.10	11.00	19.70	18.86	600	11.94	1	2.5	2	54
171	Hybrid poplar	49.40	6.00	43.10	0.23	0.05	-	-	-	-	19.74	500	24.30	2	-	-	55
172	Hybrid poplar	49.80	5.90	42.50	0.26	0.06	-	-	-	-	19.96	500	23.30	2	-	-	55
173	Switchgrass	46.90	5.80	42.00	0.58	0.11	-	-	-	-	19.53	500	23.80	2	-	-	55
174	Switchgrass	47.30	5.90	41.10	0.44	0.09	-	-	-	-	19.50	500	24.00	2	-	-	55
175	Corn stover	46.00	5.90	41.40	0.88	0.12	-	-	-	-	18.62	500	24.70	2	-	-	55
176	Corn stover	46.00	5.40	39.20	0.64	0.07	-	-	-	-	18.23	500	23.50	2	-	-	55
177	Corn stover	46.50	6.10	40.10	0.73	0.09	-	-	-	-	17.90	500	22.50	2	-	-	55
178	Yellow poplar wood	48.80	6.50	44.50	0.20	-	-	-	-	-	17.90	400	15.10	1	1.9	1	56
179	Wastepaper	39.71	7.14	53.15	-	-	76.31	6.03	6.51	11.15	-	450	13.10	1	1.5	1	57
180	Cassava stalk	51.10	6.90	41.30	0.70	0.10	79.90	6.00	15.50	14.10	17.60	500	12.70	1	2.7	2	58
181	Cassava rhizome	51.60	6.70	40.50	1.30	0.10	77.70	4.10	8.30	18.20	3.70	500	15.80	1	2.5	2	58
182	Oil palm shell	55.35	6.43	38.01	0.37	-	68.80	2.30	8.40	20.30	-	500	22.10	2	2.7	2	59
183	Corn cob	47.35	5.90	38.07	0.69	0.18	-	-	-	-	17.80	500	19.50	1	-	-	60
184	Corn stover	46.60	4.99	40.05	0.79	0.22	-	-	-	-	18.30	500	22.10	2	-	-	60
185	Rice husk	53.50	7.00	38.10	1.40	-	78.80	11.90	9.30	-	-	450	24.80	2	-	-	61
186	Rice husk	40.00	5.03	29.75	0.53	0.13	60.98	12.26	12.30	14.46	14.57	510	13.36	1	3.4	3	62
187	Radiata pine	44.80	5.90	46.20	0.10	-	-	-	-	-	18.00	502	23.00	2	2.5	2	63
188	Douglas fir	48.07	6.91	44.86	0.06	0.10	76.63	0.24	12.03	11.10	18.12	480	18.30	1	1.9	1	64
189	Oil tea shell	52.10	7.61	38.63	0.69	0.25	69.25	1.11	11.09	18.55	18.16	460	18.69	1	2.7	2	64
190	Rice Straw	48.75	5.98	43.28	1.99	-	60.84	22.55	7.30	16.61	13.45	300	10.10	1	-	-	65
191	Rice Straw	48.75	5.98	43.28	1.99	-	60.84	22.55	7.30	16.61	13.45	400	12.00	1	-	-	65

Table C-1 Pyrolysis bio-oil database (continued)

192	Rice Straw		5.98	43.28	1.99	-	60.84	22.55	7.30	16.61	13.45	500	12.30	1	-	-	65
193	Rice Straw	48.75	5.98	43.28	1.99	-	60.84	22.55	7.30	16.61	13.45	600	11.80	1	-	-	65
194	Rice Straw	48.75	5.98	43.28	1.99	-	60.84	22.55	7.30	16.61	13.45	700	11.90	1	-	-	65
195	Miscanthus	48.00	6.00	45.90	0.10	-	74.90	1.40	8.00	15.70	17.50	450	22.50	2	-	-	66
196	Sugarcane bagasse	48.67	6.70	44.10	0.45	0.08	75.85	4.14	-	20.01	16.80	475	23.50	2	3.9	3	67
197	Sawdust	49.30	6.39	44.19	0.12	-	-	-	-	-	19.05	500	10.93	1	1.9	1	68
198	Waste machinery oil	83.00	13.00	2.80	0.11	0.80	-	-	-	-	44.54	400	45.70	4	-	-	69
199	Waste machinery oil	83.00	13.00	2.80	0.11	0.80	-	-	-	-	44.54	600	45.70	4	-	-	69
200	Waste machinery oil	83.00	13.00	2.80	0.11	0.80	-	-	-	-	44.54	800	45.71	4	-	-	69
201	Oil palm mesocarp fibre	45.38	10.59	42.04	1.32	0.67	66.84	1.40	4.75	27.01	17.00	550	23.00	2	3.0	3	70
202	Palm frond	41.00	6.74	51.24	0.67	0.35	70.33	5.87	4.83	18.97	16.00	550	21.00	2	3.0	3	70
203	Poplar wood	48.36	5.93	43.94	0.52	0.12	78.26	1.13	-	20.61	18.23	400	11.85	1	3.1	3	71
204	Poplar wood	48.36	5.93	43.94	0.52	0.12	78.26	1.13	-	20.61	18.23	450	12.73	1	3.3	3	71
205	Poplar wood	48.36	5.93	43.94	0.52	0.12	78.26	1.13	-	20.61	18.23	500	14.10	1	3.7	3	71
206	Poplar wood	48.36	5.93	43.94	0.52	0.12	78.26	1.13	-	20.61	18.23	550	13.96	1	3.8	3	71
207	Poplar wood	48.36	5.93	43.94	0.52	0.12	78.26	1.13	-	20.61	18.23	600	13.98	1	3.8	3	71

1. (J. Y. Park et al., 2019)
2. (González et al., 2003)
3. (Ahmed et al., 2018)
4. (Rout et al., 2016)
5. (Yakub et al., 2015)
6. (Singh et al., 2014)
7. (Mohammed et al., 2015)
8. (Asadullah et al., 2007)
9. (Majhi et al., 2015)
10. (Pittman et al., 2012)
11. (Sinha et al., 2013)
12. (Liu et al., 2020)
13. (Boubacar Laougé et al., 2020)
14. (Bharath et al., 2020)
15. (Pidtasang et al., 2013)
16. (Mei et al., 2016)
17. (Hassan et al., 2009)
18. (Sotoudehnia et al., 2020)
19. (Kim et al., 2010)
20. (Phan et al., 2014)
21. (Zhang et al., 2009)
22. (Midhun Prasad and Murugavelh, 2020)
23. (Dai et al., 2019)
24. (Hu et al., 2013)
25. (Wang et al., 2013)
26. (Mythili et al., 2013)
27. (Lee et al., 2005)
28. (Li et al., 2012)
29. (Trinh et al., 2013)
30. (González et al., 2005)
31. (Pattiya and Suttibak, 2012)
32. (Oasmaa and Kuoppala, 2003)
33. (Zhou et al., 2013)
34. (Asadullah et al., 2013)
35. (Abnisa et al., 2011a)
36. (Rofiqul Islam et al., 2008)
37. (Asadullah et al., 2008)
38. (Zheng, 2008)
39. (Tsai et al., 2006)
40. (Naqvi et al., 2014)
41. (Oasmaa et al., 2010)
42. (Alvarez et al., 2016)
43. (Cunliffe and Williams, 1998)
44. (Edwin Raj et al., 2013)
45. (Islam et al., 2004)
46. (Fonts et al., 2009)
47. (Fahmi et al., 2008)
48. (Pattiya et al., 2012)
49. (Luo et al., 2017)
50. (Makibar et al., 2015)
51. (Amutio et al., 2012)
52. (Ding et al., 2012)
53. (Ali et al., 2016)
54. (Abnisa et al., 2011b)
55. (Agblevor et al., 1995)
56. (Kim et al., 2011)
57. (Nurul Islam et al., 2005)
58. (Pattiya, 2011)
59. (Nurul Islam et al., 1999)
60. (Mullen et al., 2010)
61. (Heo et al., 2010a)
62. (Guo et al., 2011)
63. (Kang et al., 2006)
64. (Wu et al., 2016)
65. (Park et al., 2014)
66. (Heo et al., 2010b)
67. (Islam et al., 2010)
68. (Salehi et al., 2009)
69. (Sınağ et al., 2010)
70. (Kabir et al., 2017)
71. (Chen et al., 2016)

Table C-2 Decision rules generated for Case 1: Reduct 1

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Hydrogen (%)	Temperature (ºC)				
1	≥ 56.55	< 6.75	< 502.5	0	1.38	20.00	100
2	< 42.90	5.75 - 6.15	≥ 495.0	0	2.07	30.00	100
3	49.65 - 49.75	-	< 502.5	0	0.69	10.00	100
4	-	< 6.15	≥ 650.0	0	1.38	20.00	100
5	42.65 - 42.90	-	< 412.5	0	1.38	20.00	100
6	≥ 44.1	6.25 - 6.45	-	1	8.28	16.44	100
7	49.5 - 51.75	-	502.5 - 650	1	3.45	6.85	100
8	≥ 61.85	< 7.2	≥ 502.5	1	1.38	2.74	100
9	47.35 - 51.75	6.55 - 7	-	1	6.21	12.33	100
10	47.35 - 49.35	< 6.15	≥ 452.5	1	7.59	15.07	100
11	44.1 - 45.85	-	477.5 - 502.5	1	3.45	6.85	100
12	≥ 49.5	6.05 - 6.25	< 502.5	1	2.07	4.11	100
13	48.05 - 48.5	< 5.95	-	1	3.45	6.85	100
14	< 43.3	≥ 6.85	-	1	3.45	6.85	100
15	< 43.1	5.45 - 5.75	-	1	1.38	2.74	100
16	49.5 - 49.85	≥ 5.75	-	1	1.38	2.74	100
17	46.65 - 51	-	< 412.5	1	4.14	8.22	100
18	46.1 - 46.95	< 5.85	-	1	0.69	1.37	100
19	≥ 51.35	< 6.05	-	1	1.38	2.74	100
20	33.2 - 41.1	-	-	1	6.90	13.70	100
21	43.3 - 43.5	-	-	1	0.69	1.37	100
22	52 - 56.55	-	-	1	2.07	4.11	100
23	44.1 - 45	-	-	1	5.52	10.96	100
24	< 42.9	≥ 5.45	412.5 - 452.5	1	2.07	4.11	100
25	< 48.05	5.95 - 6.05	-	2	6.90	22.22	100
26	45 - 46.1	-	≥ 502.5	2	4.14	13.33	100
27	≥ 48.5	< 6.15	445 - 452.5	2	3.45	11.11	100
28	43.1 - 43.3	< 7.2	-	2	0.69	2.22	100
29	43.65 - 44.1	-	-	2	1.38	4.44	100
30	45.85 - 46.1	-	-	2	2.07	6.67	100
31	50.65 - 51.2	< 6.65	452.5 - 540	2	1.38	4.44	100
32	43.3 - 61.85	≥ 7.2	≥ 472.5	2	4.83	15.56	100
33	≥ 50.65	6.45 - 6.65	< 540	2	1.38	4.44	100
34	46.95 - 48.05	-	< 452.5	2	4.14	13.33	100
35	45 - 45.85	-	< 452.5	2	0.69	2.22	100
36	42.65 - 42.9	< 5.45	-	2	0.69	2.22	100
37	49.35 - 49.5	-	-	2	0.69	2.22	100
38	< 73.2	≥ 10.2	-	2	2.07	6.67	100
39	42.4 - 42.65	-	-	2	0.69	2.22	100
40	24 - 33.2	-	-	2	0.69	2.22	100
41	76.5 - 80.25	-	-	3	2.07	37.50	100
42	49.85 - 66.85	-	540 - 562.5	3	0.69	12.50	100
43	49.85 - 49.95	-	-	3	0.69	12.50	100
44	48.2 - 48.35	-	-	3	0.69	12.50	100
45	42.25 - 42.4	-	-	3	0.69	12.50	100
46	< 24	-	-	3	0.69	12.50	100
47	≥ 80.25	-	-	4	5.52	88.89	100
48	73.2 - 76.5	-	-	4	0.69	11.11	100

Table C -3 Decision rules generated for Case 1: Reduct 2

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Oxygen (%)	Temperature (ºC)				
1	≥ 42.65	≥ 50.95	≥ 495	0	2.07	30.00	100
2	58.95 - 63.85	-	< 502.5	0	1.38	20.00	100
3	≥ 42.65	≥ 50.95	< 412.5	0	1.38	20.00	100
4	-	50.85 - 58.95	≥ 650	0	1.38	20.00	100
5	≥ 49.65	≥ 44.8	< 502.5	0	0.69	10.00	100
6	48.35 - 49.35	≥ 41.9	≥ 452.5	1	4.83	9.59	100
7	42.9 - 46.3	41.9 - 49.95	452.5 - 522.5	1	8.97	17.81	100
8	≥ 49.75	≥ 41.9	-	1	6.21	12.33	100
9	≥ 48.2	38.55 - 40.1	-	1	3.45	6.85	100
10	51.2 - 56.55	< 40.7	≥ 477.5	1	2.76	5.48	100
11	49.5 - 49.85	< 43.45	-	1	1.38	2.74	100
12	46.65 - 48.5	-	≥ 452.5	1	5.52	10.96	100
13	≥ 48.35	≥ 41.75	≥ 502.5	1	5.52	10.96	100
14	< 51	< 44.65	< 437.5	1	4.83	9.59	100
15	61.85 - 64.6	-	≥ 502.5	1	1.38	2.74	100
16	33.2 - 41.1	-	-	1	6.90	13.70	100
17	< 43.5	47.7 - 50.95	-	1	2.07	4.11	100
18	-	≥ 47.7	412 - 452.5	1	4.14	8.22	100
19	48.05 - 48.85	41.75 - 44	-	1	5.52	10.96	100
20	-	41.65 - 41.75	-	1	0.69	1.37	100
21	51 - 51.2	≥ 40.35	-	1	1.38	2.74	100
22	45 - 48.05	≥ 41.9	< 452.5	2	6.90	22.22	100
23	≥ 50.85	41.35 - 41.65	452.5 - 502.5	2	1.38	4.44	100
24	47.7 - 51.5	< 38.55	≥ 375	2	2.76	8.89	100
25	42.4 - 46.65	< 50.4	≥ 527.5	2	4.83	15.56	100
26	42.4 - 48.65	39.05 - 41.5	-	2	2.76	8.89	100
27	≥ 48.5	≥ 43.45	445 - 477.5	2	2.07	6.67	100
28	42.4 - 61.85	< 33.85		2	2.76	8.89	100
29	43.65 - 44.1	-	-	2	1.38	4.44	100
30	40.1 - 40.35	-	-	2	1.38	4.44	100
31	66.85 - 73.2	-	-	2	1.38	4.44	100
32	46.15 - 47.7	-	-	2	2.07	6.67	100
33	49.35 - 49.5	-	-	2	0.69	2.22	100
34	24 - 33.2	-	-	2	0.69	2.22	100
35	-	8 - 15.3	-	3	2.76	50.00	100
36	49.85 - 49.95	-	-	3	0.69	12.50	100
37	64.6 - 66.85	-	-	3	0.69	12.50	100
38	48.2 - 48.35	-	-	3	0.69	12.50	100
39	42.25 - 42.4	-	-	3	0.69	12.50	100
40	≥ 80.25	-	-	4	5.52	88.89	100
41	73.2 - 76.5	-	-	4	0.69	11.11	100

Table C-4 Decision rules generated for Case 1: Reduct 3

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Hydrogen (%)	Oxygen (%)	Temperature (ºC)				
1	-	50.95 - 52.2	≥ 495	0	2.07	30.00	100
2	< 6.15	-	≥ 650	0	1.38	20.00	100
3	≥ 5.25	≥ 50.95	< 412.5	0	1.38	20.00	100
4	-	28.15 - 29.05	< 502.5	0	1.38	20.00	100
5	5.25 - 5.35	-	< 502.5	0	0.69	10.00	100
6	-	47.7 - 49.95	452.5 - 522.5	1	5.52	10.96	100
7	6.05 - 7.95	41.65 - 43.05	-	1	5.52	10.96	100
8	6.05 - 6.45	≥ 43.2	≥ 452.5	1	8.28	16.44	100
9	5.95 - 7.2	37.05 - 40.1	-	1	3.45	6.85	100
10	6.55 - 6.75	29.05 - 41.05	-	1	0.69	1.37	100
11	5.35 - 5.95	43.2 - 46.15	≥ 452.5	1	3.45	6.85	100
12	5.95 - 6.05	< 41.5	-	1	0.69	1.37	100
13	-	≥ 50.4	437.5 - 452.5	1	2.76	5.48	100
14	6.85 - 7.2	≥ 24.75	-	1	5.52	10.96	100
15	< 7.35	41.9 - 42.1	-	1	1.38	2.74	100
16	5.55 - 5.65	≥44	-	1	2.07	4.11	100
17	≥ 6.65	42.1 - 45.2	-	1	2.07	4.11	100
18	< 6.75	37.05 - 39.05	-	1	2.76	5.48	100
19	-	43.2 - 44	≥ 452.5	1	8.28	16.44	100
20	-	29.05 - 44.65	< 437.5	1	4.83	9.59	100
21	5.65 - 5.95	43.2 - 44.95	-	1	3.45	6.85	100
22	-	37.05 - 41.75	< 472.5	1	2.76	5.48	100
23	4.45 - 5.35	< 44.95	≥ 502.5	1	2.07	4.11	100
24	-	47.7 - 49	< 437.5	1	1.38	2.74	100
25	< 6.75	< 29.05	≥ 650	1	0.69	1.37	100
26	< 5.45	≥ 49.65	-	1	4.14	8.22	100
27	5.95 - 6.05	-	412.5 - 452.5	2	4.83	15.56	100
28	-	41.5 - 41.65	412.5 - 562.5	2	1.38	4.44	100
29	< 5.95	41.9 - 43.2	-	2	2.07	6.67	100
30	-	45.05 - 50.4	≥ 527.5	2	3.45	11.11	100
31	5.95 - 6.15	≥ 43.75	< 452.5	2	4.83	15.56	100
32	≥ 7.2	21.6 - 41.5	≥ 472.5	2	4.83	15.56	100
33	6.75 - 6.85	≥ 17.05	-	2	0.69	2.22	100
34	< 5.95	38.45 - 42.1	≥ 495	2	1.38	4.44	100
35	≥ 5.45	49.45 - 50.4	-	2	1.38	4.44	100
36	-	21.6 - 43.2	527.5 - 562.5	2	2.76	8.89	100
37	< 6.05	43.05 - 43.2	-	2	1.38	4.44	100
38	-	46.15 - 47.7	-	2	2.07	6.67	100
39	-	40.1 - 40.35	-	2	1.38	4.44	100
40	10.2 - 11.7	-	-	2	1.38	4.40	100
41	-	8 - 15.2	-	3	2.76	50.00	100
42	9.1 - 10.2	-	-	3	0.69	12.50	100
43	6.25 - 6.35	< 43.2	-	3	0.69	12.50	100
44	6.15 - 6.25	≥ 43.05	-	3	0.69	12.50	100
45	-	41.74 - 41.9	-	3	0.69	12.50	100
46	≥ 6.74	< 8	-	4	5.52	88.89	100
47	-	15.3 - 17.05	-	4	0.69	11.11	100

Table C-5 Decision rules generated for Case 1: Reduct 4

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Nitrogen (%)	Temperature (ºC)				
1	56.55 - 63.85	-	< 502.5	0	1.38	20.00	100
2	42.65 - 42.9	-	≥ 495	0	2.07	30.00	100
3	42.65 - 42.9	-	< 412.5	0	1.38	20.00	100
4	≥ 50.85	≥ 0.45	≥ 650	0	1.38	20.00	100
5	49.65 - 49.75	-	< 502.5	0	0.69	10.00	100
6	49.75 - 56.55	0.05 - 0.75	≥ 477.5	1	8.97	17.81	100
7	< 45.85	0.35 - 6.05	452.5 - 522.5	1	6.90	13.70	100
8	46.65 - 49.35	< 2.05	≥ 452.5	1	8.28	16.44	100
9	51.5 - 56.55	≥ 0.15	-	1	1.38	2.74	100
10	-	1.05 - 1.15	-	1	1.38	2.74	100
11	46.1 - 64.6	≥ 3.55	-	1	4.14	8.22	100
12	44.1 - 51	≥ 0.35	< 437.5	1	4.83	9.59	100
13	61.85 - 64.6	-	≥ 502.5	1	1.38	2.74	100
14	33.2 - 41.1	-	-	1	6.90	13.70	100
15	49.5 - 50.65	0.15 - 0.55	-	1	4.83	9.59	100
16	48.05 - 48.85	< 0.55	-	1	5.52	10.96	100
17	44.1 - 45.45	0.25 - 0.35	-	1	0.69	1.37	100
18	-	2.75 - 2.9	-	1	0.69	1.37	100
19	49.5 - 49.75	-	≥ 502.5	1	1.38	2.74	100
20	< 42.9	< 0.35	412.5 - 452.5	1	2.76	5.48	100
21	45 - 48.75	0.85 - 3.55	< 452.5	2	5.52	17.78	100
22	45 - 46.65	-	≥ 527.5	2	4.83	15.56	100
23	49.05 - 66.85	0.75 - 1.05	412.5 - 502.5	2	1.38	4.44	100
24	45 - 46.1	≥ 2.15	-	2	2.07	6.67	100
25	45.85 - 46.65	< 2.15	-	2	3.45	11.11	100
26	50.65 - 51.2	≥ 1.15	≥ 375	2	2.07	6.67	100
27	-	1.15 - 2.75	≥ 527.5	2	4.14	13.33	100
28	< 43.65	3.05 - 5.5	-	2	1.38	4.44	100
29	42.4 - 58.95	< 0.15	< 452.5	2	2.07	6.67	100
30	48.85 - 49.05	-	-	2	1.38	4.44	100
31	58.95 - 73.2	< 0.25	-	2	2.07	6.67	100
32	43.65 - 44.1	-	-	2	1.38	4.44	100
33	≥ 49.5	< 0.05	< 502.5	2	1.38	4.44	100
34	≥ 42.4	≥ 6.45	-	2	0.69	2.22	100
35	49.35 - 49.5	-	-	2	0.69	2.22	100
36	≥ 42.25	≥ 7.8	-	3	0.69	12.50	100
37	≥ 76.5	≥ 0.85	-	3	2.07	37.50	100
38	≥ 64.6	≥ 2.05	-	3	0.69	12.50	100
39	-	2.05 - 2.15	-	3	0.69	12.50	100
40	48.2 - 48.35	-	-	3	0.69	12.50	100
41	< 24	-	-	3	0.69	12.50	100
42	≥ 80.25	-	-	4	5.52	88.89	100
43	73.2 - 76.5	-	-	4	0.69	11.11	100

Table C-6 Decision rules generated for Case 1: Reduct 5

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Oxygen (%)	Nitrogen (%)	Temperature (ºC)				
1	50.95 - 52.2	-	≥ 495	0	2.07	30.00	100
2	-	0.45 - 0.85	≥ 650	0	1.38	20.00	100
3	50.95 - 52.2	-	< 412.5	0	1.38	20.00	100
4	< 29.05	0.35 - 0.45	< 502.5	0	1.38	20.00	100
5	≥ 44.8	< 0.15	495 - 502.5	0	0.69	10.00	100
6	39.6 - 44	0.25 - 0.75	-	1	8.97	17.81	100
7	≥ 43.2	≥ 0.35	452.5 - 522.5	1	13.10	26.03	100
8	24.75 - 30.7	-	≥ 502.5	1	2.07	4.11	100
9	37.05 - 41.35	≥ 3.55	-	1	2.07	4.11	100
10	41.65 - 43.05	< 1.15	-	1	6.21	12.33	100
11	41.5 - 44.95	< 0.35	≥ 502.5	1	4.14	8.22	100
12	< 48.7	0.25 - 0.35	437.5 - 502.5	1	1.38	2.74	100
13	≥ 45.95	< 1.05	437.5 - 495	1	4.83	9.59	100
14	39.05 - 41.75	≥ 5.7	-	1	2.07	4.11	100
15	41.5 - 44.95	< 0.25	< 495	1	2.07	4.11	100
16	-	1.05 - 1.15	-	1	1.38	2.74	100
17	≥ 37.05	≥ 0.55	< 437.5	1	4.83	9.59	100
18	37.05 - 39.05	< 0.75	-	1	3.45	6.85	100
19	≥ 52.2	-	-	1	4.14	8.22	100
20	40.7 - 41.5	< 0.25	-	1	0.69	1.37	100
21	< 40.7	1.25 - 1.35	-	1	0.69	1.37	100
22	≥ 41.65	< 0.05	-	1	2.07	4.11	100
23	≥ 43.45	0.85 - 3.95	412.5 - 477.5	2	4.83	15.56	100
24	45.05 - 50.4	< 0.55	< 452.5	2	2.07	6.67	100
25	39.05 - 41.5	1.35 - 7.15	-	2	2.07	6.67	100
26	21.6 - 38.55	≥ 0.85	-	2	4.14	13.33	100
27	45.2 - 50.4	< 0.25	-	2	2.07	6.67	100
28	41.5 - 41.65	-	452 - 502.5	2	1.38	4.44	100
29	45.05 - 45.2	< 0.35	-	2	1.38	4.44	100
30	≥ 40.35	0.85 - 2.75	≥ 502.5	2	3.45	11.11	100
31	41.9 - 43.2	-	445 - 477.5	2	2.07	6.67	100
32	45.2 - 49	-	445 - 452.5	2	1.38	4.44	100
33	49 - 50.4	< 1.05	≥ 495	2	1.38	4.44	100
34	39.05 - 44.45	0.75 - 1.05	445 - 502.5	2	2.76	8.89	100
35	43.05 - 44	< 0.25	-	2	0.69	2.22	100
36	44 - 44.45	-	-	2	2.07	6.67	100
37	39.05 - 39.6	-	-	2	0.69	2.22	100
38	17.05 - 37.05	< 0.15	-	2	2.76	8.89	100
39	< 0.4	-	-	2	0.69	2.22	100
40	< 21.6	≥ 0.85	-	3	3.45	62.50	100
41	41.75 - 41.9	-	-	3	0.69	12.50	100
42	< 43.2	≥ 7.8	-	3	0.69	12.50	100
43	44.8 - 44.95	≥ 0.55	-	3	0.69	12.50	100
44	0.4 - 8	-	-	4	5.52	88.89	100
45	15.3 - 17.05	-	-	4	0.69	11.11	199

Table C-7 Decision rules generated for Case 2: Reduct 1

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Oxygen (%)	Temperature (ºC)				
1	≥ 47.85	≥ 44.45	< 485	1	2.22	66.67	100
2	≥ 39.2	≥ 52.2	-	1	1.11	33.33	100
3	≥ 41.9	≥ 50.5	-	2	5.56	17.86	100
4	45.95 - 47.15	≥ 44.25	-	2	4.44	14.29	100
5	50.6 - 51.2	< 42.35	< 625	2	3.33	10.71	100
6	≥ 48.5	43.95 - 44.95	≥ 495	2	6.67	21.43	100
7	≥ 57.95	≥ 33.85	-	2	2.22	7.14	100
8	45.95 - 46.3	-	495 - 505	2	1.11	3.57	100
9	44 - 44.95	-	-	2	1.11	3.57	100
10	51.35 - 51.75	-	≥ 495	2	2.22	7.14	100
11	-	≥ 54.05	-	2	2.22	7.14	100
12	< 48.65	43.95 - 44.25	-	2	4.44	14.29	100
13	39.75 - 40.2	-	-	2	1.11	3.57	100
14	49.9 = 51.35	< 40.6	-	3	3.33	6.82	100
15	46.3 - 73.2	< 33.85	-	3	8.89	18.18	100
16	46.3 - 49.6	< 43.95	< 575	3	10.00	20.45	100
17	44 - 46.3	-	< 495	3	2.22	4.55	100
18	≥ 76.5	-	460 - 505	3	1.11	2.27	100
19	42.9 - 46.3	-	≥ 505	3	5.56	11.36	100
20	48.65 - 51.75	-	406 - 477.5	3	8.89	18.18	100
21	51.75 - 57.95	-	-	3	3.33	6.82	100
22	42.9 = 43.15	-	-	3	2.22	4.55	100
23	47.15 - 47.85	-	-	3	2.22	4.55	100
24	44.95 - 45.95	-	-	3	5.56	11.36	100
25	≥ 50.8	≥ 42.8	-	3	1.11	2.27	100
26	≥ 47.15	≥ 44.95	-	3	2.22	4.45	100
27	40.95 - 41.9	-	-	3	1.11	2.27	100
28	41.5 - 41.8	-	≥ 505	3	1.11	2.27	100
29	-	< 0.4	< 375	3	1.11	2.27	100
30	-	< 28.25	406 - 477.5	4	3.33	37.05	100
31	≥ 73.2	-	≥ 575	4	1.11	12.50	100
32	< 40.95	-	<431	4	1.11	12.50	100
33	48.5 - 49.9	38.8 - 41.5	-	4	2.22	25.00	100
34	< 36.15	-	-	4	1.11	12.50	100
35	50.3 - 50.6	-	-	5	5.56	83.33	100
36	≥ 82.1	-	-	5	1.11	16.67	100
37	40.2 - 40.75	-	-	8	1.11	100.00	100

Table C-8 Decision rules generated for Case 2: Reduct 2

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Hydrogen (%)	Oxygen (%)	Temperature (ºC)				
1	6.4 - 7.2	≥ 44.45	< 485	1	3.33	100.00	100
2	5.95 - 6.05	≥ 43.75	-	2	8.89	28.57	100
3	5.95 - 6.4	41.5 - 43.65	495 - 505	2	2.22	7.14	100
4	≥ 5.95	46.35 - 50.5	≥ 495	2	2.22	7.14	100
5	-	48.95 - 49.25	-	2	1.11	3.57	100
6	6.05 - 6.55	39.45 - 42.35	-	2	2.22	7.14	100
7	< 5.85	≥ 50.5	-	2	7.78	25.00	100
8	≥ 5.75	39.45 - 40.6	≥ 495	2	2.22	7.14	100
9	6.05 - 6.2	< 34.65	-	2	2.22	7.14	100
10	5.15 - 5.45	≥ 33.85	-	2	2.22	7.14	100
11	-	44.55 - 44.75	-	2	1.11	3.57	100

Table C-8 Decision rules generated for Case 2: Reduct 2 (continued)

12	≥ 6.2	42.8 - 44.45	< 495	3	7.78	15.91	100
13	≥ 6.05	41.35 - 41.8	-	3	1.11	2.27	100
14	≥ 6.05	43.65 - 44.45	-	3	6.67	13.64	100
15	5.45 - 5.95	38.8 - 48.95	-	3	8.89	18.18	100
16	≥ 6.2	17.05 - 38.8	-	3	8.89	18.18	100
17	-	40.8 - 41.8	≥ 505	3	2.22	4.55	100
18	< 6.2	49.25 - 50.5	-	3	2.22	4.55	100
19	-	34.65 - 40.35	< 505	3	4.44	9.09	100
20	≥ 6.05	≥ 42.8	≥ 505	3	3.33	6.82	100
21	-	45.55 - 46.35	-	3	1.11	2.27	100
22	≥ 7	< 25.9	< 505	3	2.22	4.55	100
23	5.85 - 6.05	< 41.35	< 505	3	1.11	2.27	100
24	-	< 44.95	< 375	3	1.11	2.27	100
25	5.45 - 5.75	-	-	3	5.56	11.36	100
26	< 5.15	-	≥ 575	3	1.11	2.27	100
27	5.75 - 6.85	< 28.25	≥ 406	4	2.22	25.00	100
28	-	41.35 - 41.5	-	4	1.11	12.50	100
29	< 5.25	< 28.25	-	4	2.22	25.00	100
30	7.95 - 9.2	-	-	4	1.11	12.50	100
31	-	38.8 - 39.45	-	4	1.11	12.50	100
32	-	< 14.05	≥ 575	4	1.11	12.50	100
33	-	42.35 - 42.8	-	5	5.56	83.33	100
34	11.54 - 18.25	-	-	5	1.11	16.67	100
35	-	41.05 - 41.25	-	8	1.11	100.00	100

Table C-9 Decision rules generated for Case 2: Reduct 3

No	Rules				Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Hydrogen (%)	Sulphur (%)	Temperature (°C)				
1	-	-	-	477.5 - 485	1	1.11	33.33	100
2	47.85 - 48.85	-	-	< 406	1	1.11	33.33	100
3	39.2 - 39.75	-	-	-	1	1.11	33.33	100
4	45.95 - 47.15	< 6.2	-	-	2	4.44	14.29	100
5	-	< 6.2	≥ 0.15	< 575	2	6.67	21.43	100
6	-	6.05 - 6.55	-	≥ 575	2	3.33	10.71	100
7	≥ 50.05	5.95 - 6.2	< 0.15	-	2	2.22	7.14	100
8	45.95 - 46.3	-	-	495 - 505	2	1.11	3.57	100
9	36.15 - 42.9	< 6.2	-	-	2	8.89	28.57	100
10	48.5 - 49.75	-	≥ 0.15	-	2	5.56	17.86	100
11	49.2 - 49.4	-	-	-	2	1.11	3.57	100
12	44.95 - 51.75	≥ 6.65	-	≥ 495	2	1.11	3.57	100
13	44 - 44.95	-	-	-	2	1.11	3.57	100
14	≥ 42.9	≥ 6.2	< 0.15	406 - 477.5	3	7.78	15.91	100
15	51.75 - 73.2	≥ 6.2	-	-	3	10.00	20.45	100
16	≥ 44.95	5.45 - 5.95	-	-	3	12.22	25.00	100
17	43.15 - 47.15	-	-	≥ 505	3	5.56	11.36	100
18	46.3 - 47.85	≥ 6.2	-	-	3	2.22	4.55	100

Table C-9 Decision rules generated for Case 2: Reduct 3 (continued)

19	42.9 - 43.15	-	-	-	3	2.22	4.55	100
20	47.15 - 49.2	-	-	485 - 505	3	6.67	13.64	100
21	≥ 51.75	≥ 6.65	-	495 - 505	3	3.33	6.82	100
22	≥ 50.05	< 6.05	-	≥ 535	3	3.33	6.82	100
23	41.9 - 46.3	≥ 5.85	-	> 495	3	2.22	4.55	100
24	49.4 - 50.05	≥ 6.05	-	-	3	2.22	4.55	100
25	≥ 42.9	-	-	< 375	3	1.11	2.27	100
26	-	-	0.35 - 0.95	-	3	2.22	4.55	100
27	≥ 48.5	< 6.05	< 0.15	≥ 575	4	1.11	12.50	100
28	< 44	≥ 7.35	-	-	4	1.11	12.50	100
29	≥ 73.2	< 6.85	-	-	4	2.22	25.00	100
30	49.75 - 49.9	-	-	-	4	1.11	12.50	100
31	≥ 73.2	-	-	≥ 575	4	1.11	12.50	100
32	43.15 - 44	-	-	≥ 406	4	1.11	12.50	100
33	< 36.15	-	-	-	4	1.11	12.50	100
34	50.3 - 50.6	-	-	-	5	5.56	83.33	100
35	≥ 82.1	-	-	-	5	1.11	16.67	100
36	< 40.75	-	≥ 2.45	-	8	1.11	100.00	100

Table C-10 Decision rules generated for Case 2: Reduct 4

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Oxygen (%)	Nitrogen (%)	Temperature (ºC)				
1	44.45 - 44.95	-	< 485	1	2.22	66.67	100
2	52.2 - 54.05	-	-	1	1.11	33.33	100
3	≥ 48.95	0.23 - 0.35	-	2	8.89	28.57	100
4	44.55 - 47.7	< 0.35	≥ 495	2	5.56	17.86	100
5	33.85 - 44.25	0.25 - 1.05	517.5 - 675	2	7.78	25.00	100
6	40.6 - 44.25	0.75 - 1.05	< 505	2	2.22	7.14	100
7	34.65 - 43.65	0.85 - 4.15	495 - 505	2	2.22	7.14	100
8	< 34.65	0.65 - 0.75	-	2	2.22	7.14	100
9	≥ 44.55	-	< 406	2	3.33	10.71	100
10	41.8 - 44.05	< 0.15	-	2	1.11	3.57	100
11	-	-	505 - 517.5	3	2.22	4.55	100
12	41.25 - 43.95	≥ 0.85	< 495	3	4.44	9.09	100
13	28.25 - 40.35	≥ 0.75	-	3	4.44	9.09	100
14	40.8 - 41.35	-	≥ 477.5	3	2.22	4.55	100
15	-	0.35 - 0.65	< 505	3	7.78	15.91	100
16	≥ 44.25	≥ 0.55	-	3	7.78	15.91	100
17	43.65 - 43.95	≥ 0.35	-	3	5.56	11.36	100
18	11.45 - 38.8	< 0.45	-	3	7.78	15.91	100
19	11.45 - 25.9	-	477.5 - 575	3	4.44	9.09	100
20	40.1 - 41.8	< 0.85	≥ 505	3	2.22	4.55	100
21	44.05 - 44.45	< 0.35	-	3	2.22	4.55	100
22	47.7 - 48.95	-	-	3	3.33	6.82	100
23	11.45 - 41.8	< 0.15	-	3	5.56	11.36	100
24	-	< 0.4	< 375	3	1.11	2.27	100
25	< 28.25	-	406 - 477.5	4	3.33	37.50	100
26	25.9 - 28.25	-	-	4	1.11	12.50	100
27	41.35 - 41.5	-	-	4	1.11	12.50	100
28	38.8 - 39.45	-	-	4	1.11	12.50	100
29	-	< 0.05	< 431	4	1.11	12.50	100

Table C-10 Decision rules generated for Case 2: Reduct 4 (continued)

30	< 14.05	-	≥ 575	4	1.11	12.50	100
31	42.35 - 42.8	-	-	5	5.56	83.33	100
32	0.4 - 5.55	-	-	5	1.11	16.67	100
33	-	≥ 6.45	-	8	1.11	100.00	100

Table C-11 Decision rules generated for Case 2: Reduct 5

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Nitrogen (%)	Temperature (ºC)				
1	47.85 - 48.85	< 0.25	< 485	1	2.22	66.67	100
2	39.2 - 39.75	-	-	1	1.11	33.33	100
3	< 44.95	0.25 - 0.45	-	2	10.00	32.14	100
4	47.15 - 51.75	≥ 0.25	505 - 675	2	6.67	21.43	100
5	≥ 48.65	1.25 - 1.35	-	2	1.11	3.57	100
6	48.5 - 50.3	< 0.15	-	2	3.33	10.71	100
7	45.95 - 47.15	< 0.35	-	2	4.44	14.29	100
8	51.5 - 58.95	0.65 - 1.05	-	2	2.22	7.14	100
9	48.5 - 48.65	≥ 0.75	-	2	4.44	14.29	100
10	50.05 - 51	≥ 0.75	< 505	2	1.11	3.57	100
11	-	3.8 - 4.15	495 - 505	2	1.11	3.57	100
12	40.95 - 48.5	≥ 0.65	≥ 505	3	5.56	11.36	100
13	< 49.2	0.35 - 0.95	-	3	11.11	22.73	100
14	58.95 - 73.2	-	-	3	7.78	15.91	100
15	50.8 - 57.95	< 0.25	-	3	4.44	9.09	100
16	46.3 - 50.05	≥ 1.05	-	3	5.56	11.36	100
17	50.8 - 57.95	≥ 0.45	≥ 535	3	2.22	4.55	100
18	45.95 - 46.3	-	< 495	3	2.22	4.55	100
19	51 - 51.5	-	< 505	3	3.33	6.82	100
20	44.95 - 45.95	-	-	3	5.56	11.36	100
21	76.5 - 78.0	-	< 505	3	1.11	2.27	100
22	47.15 - 47.85	-	-	3	2.22	4.55	100
23	42.9 - 43.15	-	-	3	2.22	4.55	100
24	48.85 - 49.2	-	-	3	3.33	6.82	100
25	≥ 42.9	-	< 375	3	1.11	2.27	100
26	40.75 - 44	< 0.25	≥ 406	4	2.22	25.00	100
27	≥ 73.2	-	< 477.5	4	2.22	25.00	100
28	49.75 - 49.9	-	-	4	1.11	12.50	100
29	≥ 73.2	-	≥ 575	4	1.11	12.50	100
30	48.5 - 48.65	< 0.25	-	4	1.11	12.50	100
31	< 36.15	-	-	4	1.11	12.50	100
32	50.3 - 50.6	-	-	5	5.56	83.33	100
33	≥ 82.1	-	-	5	1.11	16.67	100
34	-	≥ 6.45	-	8	1.11	100.00	100

Table C-12 Decision rules generated for Case 2: Reduct 6

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Oxygen (%)	Sulphur (%)	Temperature (ºC)				
1	44.45 - 44.95	-	< 485	1	2.22	66.67	100
2	52.2 - 54.05	-	-	1	1.11	33.33	100
3	45.05 - 45.55	-	-	2	2.22	7.14	100
4	43.95 - 44.75	< 0.05	≥ 495	2	2.22	7.14	100
5	43.5 - 46.35	-	≥ 517.5	2	5.56	17.86	100
6	46.35 - 47.7	-	< 517.5	2	2.22	7.14	100
7	39.45 - 40.8	≥ 0.05	≥495	2	2.22	7.14	100
8	48.95 - 49.25	-	-	2	1.11	3.57	100
9	50.5 - 52.2	< 0.05	-	2	5.56	17.86	100
10	43.5 - 43.65	-	495 - 505	2	1.11	3.57	100
11	≥ 52.2	≥ 0.15	-	2	2.22	7.14	100
12	41.5 - 44.25	≥ 0.15	< 505	2	2.22	7.14	100
13	41.8 - 42.35	-	-	2	1.11	3.57	100
14	33.85 - 34.65	-	-	2	2.22	7.14	100
15	-	-	517.5 - 535	2	1.11	3.57	100
16	42.8 - 44.45	< 0.25	< 495	3	8.89	18.18	100
17	44.95 - 45.05	-	-	3	1.11	2.27	100
18	43.65 - 43.95	-	-	3	5.56	11.36	100
19	40.1 - 41.8	< 0.05	-	3	4.44	9.09	100
20	47.7 - 48.95	-	-	3	3.33	6.82	100
21	17.05 - 33.85	≥ 0.05	-	3	6.67	13.64	100
22	49.25 - 50.5	< 0.05	-	3	2.22	4.55	100
23	11.45 - 40.35	< 0.15	< 505	3	6.67	13.64	100
24	42.8 - 43.65	-	≥ 505	3	2.22	4.55	100
25	41.8 - 44.45	0.15 - 0.25	-	3	1.11	2.27	100
26	40.1 - 41.8	< 0.25	≥ 535	3	3.33	6.82	100
27	-	0.35 - 0.95	-	3	2.22	4.55	100
28	17.05 - 25.9	-	-	3	3.33	6.82	100
29	42.8 - 43.5	-	-	3	3.33	6.82	100
30	34.65 - 38.8	-	-	3	4.44	9.09	100
31	45.55 - 46.35	-	-	3	1.11	2.27	100
32	< 0.4	-	< 375	3	1.11	2.28	100
33	< 28.25	-	406 - 477.5	4	3.33	37.50	100
34	< 41.35	< 0.15	≥ 575	4	2.22	25.00	100
35	≥ 49.25	0.05 - 0.15	-	4	1.11	12.50	100
36	41.35 - 41.5	-	-	4	1.11	12.50	100
37	25.9 - 28.25	-	-	4	1.11	12.50	100
38	42.35 - 42.8	-	-	5	5.56	83.33	100
39	0.4 - 5.55	-	-	5	1.11	16.67	100
40	≥ 41.05	≥ 2.45	-	8	1.11	100.00	100

Table C-13 Decision rules generated for Case 3: Reduct 1

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Ash Content (%)	Temperature (ºC)				
1	≥ 74.45	≥ 11.6	≥ 537.5	0	1.33	50.00	100
2	74.45 - 74.7	-	< 412.5	0	1.33	50.00	100
3	64.95 - 70.1	1.45 - 10.75	< 502.5	1	8.00	17.14	100
4	< 82.05	2 - 3.4	≥ 502.5	1	6.67	14.29	100
5	-	< 0.85	≥ 452.5	1	13.33	28.57	100
6	75.2 - 78	-	< 502.5	1	12.00	25.71	100
7	79.8 - 80.6	-	< 555	1	4.00	8.57	100
8	< 61	-	-	1	5.33	11.43	100
9	75.95 - 78	-	-	1	10.67	22.86	100
10	-	≥ 11.6	412.5 - 452.5	1	1.33	2.86	100
11	70.1 - 75.95	3.4 - 8.9	-	2	9.33	24.14	100
12	78.4 - 81.75	1.45 - 5.85	-	2	8.00	20.69	100
13	78.4 - 84.6	1.45 - 2.85	-	2	4.00	10.34	100
14	61 - 68.75	≥ 10.75	-	2	6.67	17.24	100
15	≥ 68	0.85 - 2	-	2	8.00	20.69	100
16	78.4 - 79.8	-	-	2	4.00	10.34	100
17	68 - 68.75	-	≥ 502.5	2	1.33	3.45	100
18	65.9 - 67.15	-	-	2	4.00	10.34	100
19	-	4.85 - 5.05	-	3	1.33	33.33	100
20	84.6 - 90.9	-	-	3	1.33	33.33	100
21	70.1 - 70.25	≥ 2.45	-	3	1.33	33.33	100
22	61.75 - 64.95	-	-	4	6.67	83.33	100
23	≥ 90.0	-	-	4	1.33	16.67	100

Table C-14 Decision rules generated for Case 3: Reduct 2

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	HHV (MJ/kg)	Temperature (ºC)				
1	-	15.75 - 15.9	≥ 537.5	0	1.33	50.00	100
2	74.45 - 74.7	-	< 412.5	0	1.33	50.00	100
3	75.2 - 78	-	< 537.5	1	16.00	34.29	100
4	75.2 = 80	< 18.15	-	1	16.00	34.29	100
5	< 70.1	17.8 - 19.75	-	1	4.00	8.57	100
6	-	16.15 - 16.5	-	1	4.00	8.57	100
7	81.75 - 82.05	-	-	1	1.33	2.86	100
8	70.1 - 75.65	19.75 - 20.55	-	1	2.67	5.71	100
9	≥ 75.2	< 18.25	< 502.5	1	14.67	31.43	100
10	82.2 - 83.3	-	-	1	1.33	2.86	100
11	-	17.55 - 17.65	< 502.5	1	2.67	5.71	100
12	-	-	≥ 675	1	4.00	8.57	100
13	< 61	-	-	1	5.33	11.43	100
14	-	< 15.9	412.5 - 452.5	1	2.67	5.71	100
15	-	-	< 350	1	4.00	8.57	100
16	68 - 81.75	16.7 - 20.1	537.5 - 675	2	9.33	24.14	100
17	70.1 - 74.45	16.7 - 19.15	-	2	4.00	10.34	100
18	< 67.15	16.7 - 28.05	-	2	6.67	17.24	100
19	73.45 - 75.2	≥ 16.7	-	2	5.33	13.79	100
20	80.6 - 84.6	-	< 555	2	4.00	10.34	100
21	65.9 - 72.95	< 16.15	-	2	4.00	10.34	100
22	69.35 - 75.2	-	≥ 587.5	2	2.67	6.90	100
23	-	18.25 - 18.8	-	2	8.00	20.69	100
24	≥ 82.05	< 18	-	2	1.33	3.45	100
25	68.75 - 69.35	-	-	2	1.33	3.45	100

Table C-14 Decision rules generated for Case 3: Reduct 2 (continued)

26	70.1 - 90.9	≥ 21.85	-	3	2.67	66.67	100
27	< 70.25	19.55 - 19.75	-	3	1.33	33.33	100
28	-	≥ 28.05	-	4	6.67	83.33	100
29	61.75 - 64.95	-	-	4	6.67	83.33	100

Table C-15 Decision rules generated for Case 3: Reduct 3

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	HHV (MJ/kg)	Temperature (ºC)				
1	11.6 - 12.05	-	≥ 537.5	0	1.33	50.00	100
2	< 12.05	< 15.9	< 412.5	0	1.33	50.00	100
3	12.05 - 25.85	< 16.5	-	1	2.67	5.71	100
4	1.55 - 4.55	18.15 - 18.95	-	1	1.33	2.86	100
5	< 4.55	20.1 - 21.15	-	1	4.00	8.57	100
6	< 4.3	< 16.5	-	1	8.00	17.14	100
7	< 0.85	< 18.25	-	1	13.33	28.57	100
8	5.85 - 7.3	< 18.15	< 502.5	1	5.33	11.43	100
9	-	< 16.5	412.5 - 452.5	1	2.67	5.71	100
10	2.85 - 3.4	-	-	1	2.67	5.71	100
11	< 1.55	-	-	1	5.33	11.43	100
12	≥ 35.85	-	-	1	2.67	5.71	100
13	< 10.75	-	< 412.5	1	4.00	8.57	100
14	≥ 3.4	16.7 - 21.85	≥ 537.5	2	9.33	24.14	100
15	0.85 - 6.25	17.8 - 18.8	-	2	8.00	20.69	100
16	≥ 10.75	≥ 15.9	-	2	5.33	13.79	100
17	≥ 1.55	18.95 - 19.55	-	2	4.00	10.34	100
18	4.3 - 8.9	< 16.15	-	2	4.00	10.34	100
19	0.85 - 1.45	-	-	2	5.33	13.79	100
20	5.05 - 5.85	-	-	2	6.67	17.24	100
21	25.85 - 35.85	-	-	2	2.67	6.90	100
22	< 0.55	18.15 - 18.8	-	2	1.33	3.45	100
23	2.45 - 5.05	≥ 21.85	≥ 490	3	2.67	66.67	100
24	≥ 2.45	19.55 - 19.75	-	3	1.33	33.33	100
25	-	≥ 28.05	-	4	6.67	83.33	100
26	-	-	437.5 - 445	4	1.33	16.67	100

Table C-16 Decision rules generated for Case 3: Reduct 4

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Fixed Carbon (%)	Temperature (ºC)				
1	-	11.85 - 12.2	≥ 537.5	0	1.33	50.00	100
2	74.45 - 74.7	-	< 412.5	0	1.33	50.00	100
3	75.2 - 78	< 17	-	1	12.00	25.71	100
4	≥ 80.6	14.25 - 15.25	-	1	1.33	2.86	100
5	75.95 - 80.6	≥ 14.75	-	1	12.00	25.71	100
6	65.1 - 73.45	11.85 - 19.85	< 502.5	1	8.00	17.14	100
7	-	17 - 18.35	-	1	5.33	11.43	100
8	≥ 74.7	9.4 - 12.2	-	1	5.33	11.43	100
9	< 61	-	-	1	5.33	11.43	100
10	-	13.95 - 14.25	-	1	1.33	2.86	100
11	68 - 74.7	-	412.5 - 452.5	1	2.67	5.71	100
12	64.95 - 68	-	≥ 567.5	1	1.33	2.86	100
13	-	15.25 - 17	502.5 - 675	2	5.33	13.79	100
14	73.45 - 75.2	≥ 12.4	-	2	6.67	17.24	100
15	78.4 - 81.75	< 13.95	-	2	5.33	13.79	100
16	-	18.35 - 19.65	≥ 412.5	2	5.33	13.79	100

Table C-16 Decision rules generated for Case 3: Reduct 4 (continued)

17	< 72.95	19.95 - 28.2	-	2	5.33	13.79	100
18	≥ 80	11.4 - 14.75	-	2	4.00	10.34	100
19	< 72.95	< 8.45	-	2	4.00	10.34	100
20	61 - 67.15	< 15.25	-	2	5.33	13.79	100
21	78.4 - 79.8	-	-	2	4.00	10.34	100
22	12.2 - 12.4	-	-	3	1.33	33.33	100
23	84.6 - 90.9	-	-	3	1.33	33.33	100
24	-	19.85 - 19.95	-	3	1.33	33.33	100
25	-	≥ 28.2	-	4	6.67	83.33	100
26	≥ 90.9	-	-	4	1.33	16.67	100

Table C-17 Decision rules generated for Case 3: Reduct 5

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	Fixed Carbon (%)	Temperature (ºC)				
1	11.6 - 12.05	-	≥ 537.5	0	1.33	50.00	100
2	< 12.05	< 12.2	< 412.5	0	1.33	50.00	100
3	≥ 0.85	13.95 - 14.75	< 537.5	1	4.00	8.75	100
4	≥ 11.6	-	412.5 - 452.5	1	1.33	2.86	100
5	< 4.55	17 - 19.85	< 555	1	10.67	22.86	100
6	< 0.85	≥ 14.75	-	1	14.67	31.43	100
7	< 3.4	7.85 - 11.4	-	1	5.33	11.43	100
8	≥ 19.1	≥ 7.85	-	1	5.33	11.43	100
9	-	15.8 - 19.85	< 452.5	1	6.67	14.29	100
10	2.25 - 3.4	-	≥ 555	1	2.67	5.71	100
11	5.95 - 6.25	< 11.4	-	1	1.33	2.86	100
12	2 - 2.25	≥ 15.95	-	1	1.33	2.86	100
13	7.3 - 11.6	< 15.25	-	2	2.67	6.90	100
14	0.85 - 6.8	15.25 - 17	-	2	9.33	24.14	100
15	-	12.4 - 13.95	-	2	5.33	13.79	100
16	4.3 - 5.95	≥ 14.6	≥ 445	2	5.33	13.79	100
17	≥ 7.3	≥ 18.35	-	2	2.67	6.90	100
18	< 5.85	14.25 - 14.75	-	2	2.67	6.90	100
19	0.85 - 2	≥ 19.95	-	2	4.00	10.34	100
20	4.3 - 35.85	< 9.4	-	2	6.67	17.24	100
21	12.05 - 19.1	-	-	2	2.67	6.90	100
22	3.85 - 5.05	< 12.4	-	3	2.67	66.67	100
23	-	≥ 28.2	-	4	6.67	83.33	100
24	-	< 0.45	-	4	1.33	16.67	100

Table C-18 Decision rules generated for Case 3: Reduct 6

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moisture Content (%)	Temperature (ºC)				
1	1.35 - 1.9	≥ 537.5	0	1.33	50.00	100
2	< 1.9	< 412.5	0	1.33	50.00	100
3	9.95 - 11.4	-	1	8.00	17.14	100
4	8.95 - 9.7	< 537.5	1	10.67	22.86	100
5	≥ 5.8	< 412.5	1	4.00	8.57	100
6	< 5.8	477.5 - 502.5	1	4.00	8.57	100
7	5.2 - 7.4	≥ 555	1	4.00	8.57	100
8	≥ 11.85	< 502.5	1	4.00	8.57	100
9	8.2 - 8.7	< 555	1	4.00	8.57	100
10	6.4 - 6.55	-	1	1.33	2.86	100
11	40.6 - 61.1	-	1	1.33	2.86	100
12	-	≥ 675	1	4.00	8.57	100
13	1.35 - 1.9	412.5 - 452.5	1	1.33	2.86	100

Table C-18 Decision rules generated for Case 3: Reduct 6 (continued)

14	-	< 350	1	4.00	8.57	100
15	7.4 - 8.55	≥ 502.5	2	5.33	13.79	100
16	8.7 - 8.95	-	2	2.67	6.90	100
17	5.8 - 6.4	≥ 412.5	2	4.00	10.34	100
18	7.65 - 8.2	-	2	5.33	13.79	100
19	6.55 - 6.8	-	2	4.00	10.34	100
20	1.9 - 5.2	537.5 - 555	2	5.33	13.79	100
21	≥ 61.1	-	2	4.00	10.34	100
22	8.7 - 9.7	≥ 537.5	2	2.67	6.90	100
23	11.4 - 11.85	-	2	2.67	6.90	100
24	3.45 - 4.15	-	2	1.33	3.45	100
25	9.7 - 9.95	-	2	1.33	3.45	100
26	23.7 - 40.6	-	2	1.33	3.45	100
27	1.9 - 2.85	-	2	2.67	6.90	100
28	7.4 - 7.65	490 - 502.5	3	1.33	33.33	100
29	6.8 - 7.1	-	3	1.33	33.33	100
30	4.15 - 4.75	≥ 587.5	3	1.33	33.33	100
31	< 1.35	-	4	8.00	100.00	100

Table C-19 Decision rules generated for Case 3: Reduct 7

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Fixed Carbon (%)	HHV (MJ/kg)	Temperature (ºC)				
1	-	15.75 - 15.9	≥ 537.5	0	1.33	50.00	100
2	11.85 - 12.2	-	< 412.5	0	1.33	50.00	100
3	7.85 - 14.9	17.35 - 18.15	-	1	12.00	25.71	100
4	< 15.25	17.55 - 18.15	-	1	6.67	14.29	100
5	15.8 - 19.85	< 18.25	< 502.5	1	10.67	22.86	100
6	7.85 - 16.75	19.15 - 21.15	-	1	5.33	11.43	100
7	19.65 - 19.85	-	-	1	6.67	14.29	100
8	< 12.2	< 17.15	412.5 - 452.5	1	2.67	5.71	100
9	21.8 - 24.35	-	-	1	1.33	2.86	100
10	-	16.15 - 16.5	-	1	4.00	8.57	100
11	-	-	≥ 675	1	4.00	8.57	100
12	-	< 11.9	-	1	4.00	8.57	100
13	-	-	< 350	1	4.00	8.57	100
14	< 19.65	18.25 - 19.55	-	2	10.67	27.59	100
15	≥ 15.25	16.7 - 18.15	-	2	9.33	24.14	100
16	12.4 - 28.2	≥ 19.9	≥ 445	2	5.33	13.79	100
17	11.4 - 15.25	16.15 - 17.35	-	2	4.00	10.34	100
18	< 8.45	< 20.55	-	2	5.33	13.79	100
19	-	15.9 - 16.15	-	2	2.67	6.90	100
20	19.95 - 21.8	-	-	2	4.00	10.34	100
21	2.34 - 19.95	≥ 21.85	-	3	2.67	66.67	100
22	19.85 - 19.95	-	-	3	1.33	33.33	100
23	≥ 28.2	-	-	4	6.67	83.33	100
24	< 0.45	-	-	4	1.33	16.67	100

Table C-20 Decision rules generated for Case 4: Reduct 1

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Moisture Content (%)				
1	< 76.67	≥ 11.85	1	2.00	100.00	100
2	71.65 - 81.35	8 - 11.4	2	14.00	38.89	100
3	81.35 - 81.75	≥ 9.6	2	8.00	22.22	100
4	-	1.1 - 1.45	2	10.00	27.78	100
5	< 69.75	≥ 9.6	2	4.00	11.11	100
6	≥ 59.35	4.35 - 8	3	18.00	50.00	100
7	80.6 - 90.0	< 62.3	3	16.00	44.44	100
8	< 71.65	5.8 - 8.85	3	6.00	16.67	100
9	-	≥ 76.4	3	2.00	5.56	100
10	< 71.65	< 4.35	4	8.00	66.67	100
11	76.86 - 77.35	-	4	2.00	16.67	100
12	66.05 - 68.05	-	4	2.00	16.67	100
13	-	3.15 - 3.65	5	8.00	80.00	100
14	≥ 90.0	-	5	2.00	20.00	100
15	< 55.85	-	8	4.00	100.00	100

Table C-21 Decision rules generated for Case 4: Reduct 2

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	Moisture Content (%)				
1	< 0.35	-	1	2.00	100.00	100
2	0.35 - 2.55	≥ 8	2	20.00	55.56	100
3	≥ 4.35	< 1.45	2	10.00	27.78	100
4	3.85 - 4.35	≥ 8	2	6.00	16.67	100
5	4.35 - 16.1	≥ 4.35	3	20.00	55.56	100
6	2.55 - 3.85	-	3	8.00	22.22	100
7	-	5.8 - 8	3	12.00	33.33	100
8	-	1.65 - 2.55	3	4.00	11.11	100
9	4.15 - 4.35	< 3.15	4	2.00	16.67	100
10	≥ 7.1	1.45 - 4.35	4	4.00	33.33	100
11	16.1 - 33.15	-	4	4.00	33.33	100
12	< 0.85	< 3.15	4	2.00	16.67	100
13	-	3.15 - 3.65	5	8.00	80.00	100
14	-	< 0.6	5	2.00	20.00	100
15	≥ 33.15	-	8	4.00	100.00	100

Table C-22 Decision rules generated for Case 4: Reduct 3

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Fixed Carbon (%)				
1	-	9.85 - 11.6	1	2.00	100.00	100
2	73.95 - 75.15	-	2	12.00	33.33	100
3	≥ 75.95	16.65 - 20.45	2	14.00	38.89	100
4	71.65 - 73.45	-	2	4.00	11.11	100
5	< 69.75	18.55 - 20.45	2	4.00	11.11	100
6	-	7.85 - 8.35	2	2.00	5.56	100
7	78.1 - 90.9	< 15.7	3	22.00	61.11	100
8	< 75.15	21.25 - 28.2	3	4.00	11.11	100
9	69.75 - 75.95	≥ 18.45	3	6.00	16.67	100
10	< 73.95	9.85 - 16.65	3	6.00	16.67	100
11	-	≥ 28.2	4	4.00	33.33	100
12	≥ 70.9	≥ 20.45	4	2.00	16.67	100
13	70.9 - 71.65	≥ 18.3	4	2.00	16.67	100
14	66.05 - 68.05	-	4	2.00	16.67	100
15	55.85 - 59.35	-	4	2.00	16.67	100

Table C-22 Decision rules generated for Case 4: Reduct 3 (continued)

16	80.2 - 80.6	-	5	8.00	80.00	100
17	≥ 90.9	-	5	2.00	20.00	100
18	< 55.85	-	8	4.00	100.00	100

Table C-23 Decision rules generated for Case 4: Reduct 4

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Temperature (ºC)				
1	-	477.5 - 485	1	2.00	100.00	100
2	73.95 - 75.15	-	2	12.00	33.33	100
3	71.65 - 76.85	495 - 535	2	10.00	27.78	100
4	68.05 - 69.75	< 510	2	4.00	11.11	100
5	81.35 - 81.75	≥ 495	2	6.00	16.67	100
6	77.35 - 78.1	-	2	2.00	5.56	100
7	75.95 - 76.35	-	2	4.00	11.11	100
8	≥ 81.35	< 455	2	2.00	5.56	100
9	71.65 - 73.45	-	2	4.00	11.11	100
10	81.75 - 90.9	-	3	8.00	22.22	100
11	80.6 - 81.35	-	3	6.00	16.67	100
12	69.75 - 70.9	-	3	4.00	11.11	100
13	75.15 - 80.2	425 - 477.5	3	4.00	11.11	100
14	75.15 - 80.2	≥ 535	3	4.00	11.11	100
15	63.85 - 66.05	-	3	2.00	5.56	100
16	69.75 - 73.45	≥ 575	3	2.00	5.56	100
17	≥ 73.95	455 - 477.5	3	4.00	11.11	100
18	73.45 - 73.95	-	3	2.00	5.56	100
19	59.35 - 61.95	-	3	2.00	5.56	100
20	61.95 - 69.15	≥ 510	4	4.00	33.33	100
21	70.9 - 71.65	< 575	4	2.00	16.67	100
22	55.85 - 63.85	< 477.5	4	4.00	33.33	100
23	76.85 - 77.35	-	4	2.00	16.67	100
24	80.2 - 80.6	-	5	8.00	80.00	100
25	≥ 90.9	-	5	2.00	20.00	100
26	< 55.85	-	8	4.00	100.00	100

Table C-24 Decision rules generated for Case 4: Reduct 5

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	HHV (MJ/kg)				
1	76.35 - 76.65	-	1	2.00	100.00	100
2	68.05 - 78.1	< 15.9	2	16.00	44.44	100
3	71.65 - 73.45	-	2	4.00	11.00	100
4	76.65 - 81.75	17.65 - 18.15	2	8.00	22.22	100
5	< 75.15	17.65 - 18.85	2	6.00	16.67	100
6	20.55 - 21.15	-	2	2.00	5.56	100
7	≥ 78.1	18.3 - 26	3	14.00	38.89	100
8	≥ 73.45	15.9 - 17.9	3	8.00	22.22	100
9	69.75 - 70.9	-	3	4.00	11.11	100
10	75.15 - 75.95	-	3	2.00	5.56	100
11	-	21.15 - 26	3	8.00	22.22	100
12	≥ 78.1	< 17.15	3	2.00	5.56	100
13	63.85 - 66.05	-	3	2.00	5.56	100
14	68.05 - 69.15	≥ 19.15	4	2.00	16.67	100
15	< 71.65	≥ 26	4	2.00	16.67	100
16	70.9 - 71.65	< 17.25	4	2.00	16.67	100
17	76.85 - 77.35	-	4	2.00	16.67	100
18	-	12 - 14.25	4	2.00	16.67	100
19	80.2 - 80.6	-	5	8.00	80.00	100

Table C-24 Decision rules generated for Case 4: Reduct 5 (continued)

20	≥ 90.9	-	5	2.00	20.00	100
21	< 55.85	-	8	4.00	100.00	100

Table C-25 Decision rules generated for Case 4: Reduct 6

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	HHV (MJ/kg)				
1	< 0.35	-	1	2.00	100.00	100
2	0.35 - 2.55	< 18.15	2	12.00	33.33	100
3	-	15.65 - 15.9	2	14.00	38.89	100
4	< 2.25	>- 19.75	2	4.00	11.11	100
5	0.85 - 1.2	-	2	4.00	11.11	100
6	-	< 7.4	2	2.00	5.56	100
7	2.55 - 7.1	15.9 - 26	3	26.00	72.22	100
8	≥ 1.45	18.3 - 19.15	3	8.00	22.22	100
9	4.35 - 11.6	< 17.15	3	6.00	16.67	100
10	4.35 - 16.1	≥ 18.3	3	12.00	33.33	100
11	0.35 - 0.55	< 18.65	3	2.00	5.56	100
12	-	19.15 - 19.75	4	4.00	33.33	100
13	≥ 4.15	≥ 26	4	4.00	33.33	100
14	16.1 - 33.15	-	4	4.00	33.33	100
15	7.1 - 9.35	-	4	2.00	16.67	100
16	0.55 - 0.85	< 18.3	5	8.00	80.00	100
17	-	≥ 35.15	5	2.00	20.00	100
18	≥ 33.15	-	8	4.00	100.00	100

Table C-26 Decision rules generated for Case 4: Reduct 7

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moisture Content (%)	HHV (MJ/kg)				
1	11.85 - 30.65	-	1	2.00	100.00	100
2	≥ 62.3	≥ 17.65	2	8.00	22.22	100
3	-	15.65 - 15.9	2	14.00	38.89	100
4	≥ 8.55	≥ 19.75	2	4.00	11.11	100
5	9.6 - 11.4	-	2	8.00	22.22	100
6	8 - 8.7	< 18.85	2	6.00	16.67	100
7	4.35 - 8	≥ 14.25	3	18.00	50.00	100
8	≥ 11.4	< 17.9	3	6.00	16.67	100
9	≥ 1.65	≥ 21.15	3	8.00	22.22	100
10	8.7 - 9.6	14.25 - 19.15	3	6.00	16.67	100
11	-	20.05 - 20.55	3	4.00	11.11	100
12	2.55 - 4.35	≥ 19.15	4	4.00	33.33	100
13	≥ 0.6	≥ 26	4	4.00	33.33	100
14	-	17.15 - 17.25	4	2.00	16.67	100
15	≥ 8.85	< 14.25	4	2.00	16.67	100
16	3.15 - 3.65	-	5	8.00	80.00	100
17	< 0.6	-	5	2.00	20.00	100
18	< 5.8	< 12	8	4.00	100.00	100

Table C-27 Decision rules generated for Case 4: Reduct 8

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moisture Content (%)	Fixed Carbon (%)				
1	11.85 - 30.65	-	1	2.00	100.00	100
2	≥ 8	≥ 16.2	2	24.00	66.67	100
3	1.1 -1.45	-	2	10.00	27.78	100
4	-	7.85 - 8.35	2	2.00	5.56	100
5	-	12.55 - 15.7	3	16.00	44.44	100
6	5.8 - 8	-	3	12.00	33.33	100

Table C-27 Decision rules generated for Case 4: Reduct 8 (continued)

7	4.35 - 5.55	-	3	6.00	16.67	100
8	8.4 - 8.55	-	3	2.00	5.56	100
9	≥ 1.65	< 5.3	3	4.00	11.11	100
10	< 4.35	≥ 20.45	4	8.00	66.67	199
11	2.55 - 3.15	-	4	4.00	33.33	100
12	-	5.3 - 7.05	4	2.00	16.67	100
13	3.15 - 3.65	-	5	8.00	80.00	100
14	< 0.6	-	5	2.00	20.00	100
15	5.55 - 5.85	-	8	4.00	100.00	100

Table C-28 Decision rules generated for Case 4: Reduct 9

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Fixed Carbon (%)	HHV (MJ/kg)				
1	9.85 - 11.6	-	1	2.00	100.00	100
2	≥ 11.6	< 15.9	2	16.00	44.44	100
3	16.2 - 18.3	≥ 17.65	2	12.00	33.33	100
4	≥ 16.65	17.65 - 18.3	2	12.00	33.33	100
5	16.65 - 18.3	-	2	14.00	38.89	100
6	7.85 - 8.35	-	2	2.00	5.56	100
7	8.35 - 16.2	≥ 18.3	3	12.00	33.33	100
8	12.55 - 15.7	-	3	16.00	44.44	100
9	21.25 - 21.8	-	3	2.00	5.56	100
10	18.8 - 19.35	-	3	2.00	5.56	100
11	24.35 - 28.2	-	3	2.00	5.56	100
12	0.85 - 7.85	≥ 14.25	3	6.00	16.67	100
13	-	18.3 - 18.65	3	6.00	16.67	100
14	-	17.25 - 17.65	3	6.00	16.67	100
15	≥ 28.2	-	4	4.00	33.33	100
16	21.8 - 24.35	-	4	2.00	16.67	100
17	-	17.15 - 17.25	4	2.00	16.67	100
18	≥ 5.3	≥ 26	4	4.00	33.33	100
19	5.3 - 7.05	-	4	2.00	16.67	100
20	< 15.95	18.15 - 18.3	5	8.00	80.00	100
21	< 0.85	-	5	2.00	20.00	100
22	< 9.85	< 12	8	4.00	100.00	100

Table C-29 Decision rules generated for Case 4: Reduct 10

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	Fixed Carbon (%)				
1	< 0.35	-	1	2.00	100.00	100
2	< 4.35	16.2 - 20.45	2	24.00	66.67	24.00
3	-	11.6 - 12.55	2	10.00	27.78	10.00
4	< 2.25	< 8.35	2	2.00	5.56	2.00
5	2.55 - 11.6	< 16.65	3	26.00	72.22	100
6	4.35 - 16.1	≥ 18.45	3	6.00	16.67	100
7	1.45 - 1.7	-	3	2.00	5.56	100
8	0.35 - 0.55	< 15.7	3	2.00	5.56	100
9	≥ 16.1	≥ 20.45	4	2.00	16.67	100
10	< 4.35	≥ 21.8	4	6.00	50.00	100
11	≥ 9.35	< 7.05	4	2.00	16.67	100
12	18.3 - 18.45	-	4	2.00	16.67	100
13	0.55 - 0.85	< 15.95	5	8.00	80.00	100
14	< 0.85	-	5	2.00	20.00	100
15	≥ 33.15	-	8	4.00	100.00	100

Table C-30 Decision rules generated for Case 4: Reduct 11

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moist Content (%)	Temperature (ºC)				
1	11.85 - 30.65	-	1	2.00	100.00	100
2	≥ 8	495 - 535	2	12.00	33.33	100
3	62.3 - 76.4	-	2	8.00	22.22	100
4	1.1 - 1.45	-	2	10.00	27.78	100
5	9.6 - 11.4	-	2	8.00	22.22	100
6	8.55 - 8.7	-	2	2.00	5.56	100
7	5.8 - 8	-	3	12.00	33.33	100
8	5.8 - 62.3	≥ 575	3	6.00	16.67	100
9	8.7 - 9.6	< 495	3	4.00	11.11	100
10	4.35 - 5.55	-	3	6.00	16.67	100
11	1.65 - 9.05	455 - 477.5	3	4.00	11.11	100
12	≥ 76.4	-	3	2.00	5.56	100
13	9.05 - 9.25	-	3	2.00	5.56	100
14	2.55 - 3.15	-	4	4.00	33.33	100
15	1.45 - 1.65	-	4	2.00	16.67	100
16	8.85 - 8.95	-	4	2.00	16.67	100
17	3.65 - 4.35	-	4	2.00	16.67	100
18	0.6 - 1.1	-	4	2.00	16.67	100
19	3.15 - 3.65	-	5	8.00	80.00	100
20	< 0.6	-	5	2.00	20.00	100
21	5.55 - 5.8	-	8	4.00	100.00	100

Table C-31 Decision rules generated for Case 4: Reduct 12

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	HHV (MJ/kg)	Temperature (ºC)				
1	-	477.5 - 485	1	2.00	100.00	100
2	15.65 - 15.9	-	2	14.00	38.89	100
3	17.9 - 18.15	≥ 495	2	6.00	16.67	100
4	19.75 - 20.05	-	2	2.00	5.56	100
5	-	455 - 464.5	2	2.00	5.56	100
6	17.65 - 17.9	< 510	2	2.00	5.56	100
7	17.9 - 18.15	< 455	2	2.00	5.56	100
8	18.65 - 18.85	-	2	2.00	5.56	100
9	16.1 - 16.65	-	2	2.00	5.56	100
10	20.55 - 21.15	-	2	2.00	5.56	100
11	< 7.4	-	2	2.00	5.56	100
12	18.85 - 19.15	-	3	4.00	11.11	100
13	15.9 - 17.9	≥ 575	3	6.00	16.67	100
14	18.3 - 18.65	-	3	6.00	16.67	100
15	21.15 - 26	-	3	8.00	22.22	100
16	15.9 - 17.65	< 455	3	4.00	11.11	100
17	20.05 - 20.55	-	3	4.00	11.11	100
18	15.9 - 16.1	-	3	2.00	5.56	100
19	14.25 - 15.65	-	3	2.00	5.56	100
20	19.15 - 19.75	-	4	4.00	33.33	100
21	26 - 35.15	-	4	4.00	33.33	100
22	17.15 - 17.25	-	4	2.00	16.67	100
23	12 - 14.25	-	4	2.00	16.67	100
24	18.15 - 18.3	≥ 495	5	4.00	40.00	100
25	≥ 18.15	< 425	5	4.00	40.00	100
26	≥ 35.15	-	5	2.00	20.00	100
27	7.4 - 12	-	8	4.00	100.00	100

Table C-32 Summary of analysed validation results for Case 1: Reduct 1

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Hydrogen (%)	Temperature (ºC)				
1	≥ 56.55	< 6.75	< 502.5	0	3.23	100.00	66.67
6	≥ 44.1	6.25 - 6.45	-	1	17.74	55.00	61.11
10	47.35 - 49.35	< 6.15	≥ 452.5				
20	33.2 - 41.1	-	-				
25	< 48.05	5.95 - 6.05	-	2	3.23	6.56	100.00
47	≥ 80.25	-	-	4	4.84	75.00	100.00

Table C-33 Summary of analysed validation results for Case 1: Reduct 2

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Oxygen (%)	Temperature (ºC)				
7	42.9 - 46.3	41.9 - 49.95	452.5 - 522.5	1	20.97	65.00	72.22
8	≥ 49.75	≥ 41.9	-				
16	33.2 - 41.1	-	-				
19	48.05 - 48.85	41.75 - 44	-				
27	≥ 48.5	≥ 43.45	445 - 477.5	2	14.52	29.03	64.29
32	46.15 - 47.7	-	-				
40	≥ 80.25	-	-	4	4.84	75.00	100.00

Table C-34 Summary of analysed validation results for Case 1: Reduct 3

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Hydrogen (%)	Oxygen (%)	Temperature (ºC)				
6	-	47.7 - 49.95	452.5 - 522.5	1	16.13	50.00	83.33
8	6.05 - 6.45	≥ 43.2	≥ 452.5				
19	-	43.2 - 44	≥ 452.5				
30	-	45.05 - 50.4	≥ 527.5	2	14.52	29.03	69.23
35	≥ 5.45	49.45 - 50.4	-				
37	< 6.05	43.05 - 43.2	-				
47	-	15.3 - 17.05	-	4	4.84	75.00	100.00

Table C-35 Summary of analysed validation results for Case 1: Reduct 4

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Nitrogen (%)	Temperature (ºC)				
1	56.55 - 63.85	-	< 502.5	0	3.23	100.00	100.00
8	46.65 - 49.35	< 2.05	≥ 452.5	1	17.74	55.00	52.38
11	46.1 - 64.6	≥ 3.55	-				
14	33.2 - 41.1	-	-				
25	45.85 - 46.65	< 2.15	-	2	4.84	9.68	100.00
36	≥ 42.25	≥ 7.8	-	3	1.61	20.00	33.33
42	≥ 80.25	-	-	4	4.84	75.00	100.00

Table C-36 Summary of analysed validation results for Case 1: Reduct 5

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Oxygen (%)	Nitrogen (%)	Temperature (ºC)				
7	≥ 43.2	≥ 0.35	452.5 - 522.5	1	12.90	40.00	72.73
9	37.05 - 41.35	≥ 3.55	-				
17	≥ 37.05	≥ 0.55	< 437.5				
26	21.6 - 38.55	≥ 0.85	-	2	20.97	41.94	68.42
30	≥ 40.35	0.85 - 2.75	≥ 502.5				
36	44 - 44.45	-	-				
44	0.4 - 8	-	-	4	4.84	75.00	100.00

Table C-37 Summary of analysed validation results for Case 2: Reduct 1

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Oxygen (%)	Temperature (ºC)				
2	≥ 39.2	≥ 52.2	-	1	2.70	50.00	100.00
5	50.6 - 51.2	< 42.35	< 625	2	13.51	33.33	71.43
6	≥ 48.5	43.95 - 44.95	≥ 495				
11	-	≥ 54.05	-				
16	46.3 - 49.6	< 43.95	< 575	3	16.22	40.00	75.00
17	44 - 46.3	-	< 495				
24	44.95 - 45.95	-	-				
33	48.5 - 49.9	38.8 - 41.5	-	4	2.70	50.00	100.00
35	50.3 - 50.6	-	-	5	2.70	100.00	50.00
37	40.2 - 40.75	-	-	8	5.41	100.00	100.00

Table C-38 Summary of analysed validation results for Case 2: Reduct 2

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Hydrogen (%)	Oxygen (%)	Temperature (ºC)				
1	6.4 - 7.2	≥ 44.45	< 485	1	2.70	50.00	100.00
7	< 5.85	≥ 50.5	-	2	8.11	20.00	75.00
15	5.45 - 5.95	38.8 - 48.95	-	3	18.92	46.67	58.33
16	≥ 6.2	17.05 - 38.8	-				
25	5.45 - 5.75	-	-				
30	7.95 - 9.2	-	-	4	2.70	50.00	50.00
33	-	42.35 - 42.8	-	5	2.70	100.00	100.00
35	-	41.05 - 41.25	-	8	5.41	100.00	100.00

Table C-39 Summary of analysed validation results for Case 2: Reduct 3

No	Rules				Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Hydrogen (%)	Sulphur (%)	Temperature (ºC)				
3	39.2 - 39.75	-	-	-	1	2.70	50.00	100.00
5		< 6.2	≥ 0.15	< 575	2	16.22	40.00	66.67
9	36.15 - 42.9	< 6.2	-	-				
15	51.75 - 73.2	≥ 6.2	-	-	3	13.51	33.33	62.50
16	≥ 44.95	5.45 - 5.95	-	-				
34	50.3 - 50.6	-	-	-	5	2.70	100.00	50.00
36	< 40.75	-	≥ 2.45	-	8	5.41	100.00	100.00

Table C-40 Summary of analysed validation results for Case 2: Reduct 4

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Oxygen (%)	Nitrogen (%)	Temperature (ºC)				
2	52.2 - 54.05	-	-	1	2.70	50.00	100.00
3	≥ 48.95	0.23 - 0.35	-	2	21.62	53.33	100.00
4	44.55 - 47.7	< 0.35	≥ 495				
15	-	0.35 - 0.65	< 505	3	18.92	46.67	77.78
16	≥ 44.25	≥ 0.55	-				
17	43.65 - 43.95	≥ 0.35	-				
31	42.35 - 42.8	-	-	5	2.70	100.00	100.00
33	-	≥ 6.45	-	8	5.41	100.00	100.00

Table C-41 Summary of analysed validation results for Case 2: Reduct 5

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Carbon (%)	Nitrogen (%)	Temperature (ºC)				
2	39.2 - 39.75	-	-	1	2.70	50.00	100.00
3	< 44.95	0.25 - 0.45	-	2	16.22	40.00	85.71
8	51.5 - 58.95	0.65 - 1.05	-				
13	< 49.2	0.35 - 0.95	-	3	21.62	53.33	88.89
20	44.95 - 45.95	-	-				
27	≥ 73.2	-	< 477.5	4	2.70	50.00	100.00
32	50.3 - 50.6	-	-	5	2.70	100.00	50.00
34	-	≥ 6.45	-	8	5.41	100.00	100.00

Table C-42 Summary of analysed validation results for Case 2: Reduct 6

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Oxygen (%)	Sulphur (%)	Temperature (ºC)				
2	52.2 - 54.05	-	-	1	2.70	50.00	100.00
9	50.5 - 52.2	< 0.05	-	2	8.11	20.00	75.00
11	≥ 52.2	≥ 0.15	-				
16	42.8 - 44.45	< 0.25	< 495	3	16.22	40.00	85.71
18	43.65 - 43.95	-	-				
21	17.05 - 33.85	≥ 0.05	-				
38	42.35 - 42.8	-	-	5	2.70	100.00	100.00
40	≥ 41.05	≥ 2.45	-	8	5.41	100.00	100.00

Table C-43 Summary of analysed validation results for Case 3: Reduct 1

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Ash Content (%)	Temperature (ºC)				
1	≥ 74.45	≥ 11.6	≥ 537.5	0	6.25	66.67	100.00
2	74.45 - 74.7	-	< 412.5				
5	-	< 0.85	≥ 452.5	1	28.13	60.00	81.82
6	75.2 - 78	-	< 502.5				
8	< 61	-	-				
12	78.4 - 81.75	1.45 - 5.85	-	2	12.50	44.44	57.14
13	78.4 - 84.6	1.45 - 2.85	-				
15	≥ 68	0.85 - 2	-				
21	70.1 - 70.25	≥ 2.45	-	3	3.13	100.00	50.00
22	61.75 - 64.95	-	-	4	9.38	75.00	75.00

Table C-44 Summary of analysed validation results for Case 3: Reduct 2

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	HHV (MJ/kg)	Temperature (ºC)				
1	-	15.75 - 15.9	≥ 537.5	0	6.25	66.67	100.00
2	74.45 - 74.7	-	< 412.5				
3	75.2 - 78	-	< 537.5	1	34.38	73.33	78.57
4	75.2 = 80	< 18.15	-				
9	≥ 75.2	< 18.25	< 502.5				
13	< 61	-	-				
16	68 - 81.75	16.7 - 20.1	537.5 - 675	2	12.50	44.44	66.67
20	80.6 - 84.6	-	< 555				
28	-	≥ 28.05	-	4	12.50	100.00	100.00

Table C-45 Summary of analysed validation results for Case 3: Reduct 3

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	HHV (MJ/kg)	Temperature (ºC)				
1	11.6 - 12.05	-	≥ 537.5	0	3.13	33.33	100.00
3	12.05 - 25.85	< 16.5	-	1	40.63	86.67	92.86
7	< 0.85	< 18.25	-				
11	< 1.55	-	-				
13	< 10.75	-	< 412.5				
15	0.85 - 6.25	17.8 - 18.8	-	2	9.38	33.33	75.00
25	-	≥ 28.05	-	4	12.50	100.00	100.00

Table C-46 Summary of analysed validation results for Case 3: Reduct 4

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Fixed Carbon (%)	Temperature (ºC)				
1	-	11.85 - 12.2	≥ 537.5	0	6.25	66.67	100.00
2	74.45 - 74.7	-	< 412.5				
3	75.2 - 78	< 17	-	1	31.25	66.67	83.33
5	75.95 - 80.6	≥ 14.75	-				
9	< 61	-	-				
13	-	15.25 - 17	502.5 - 675	2	12.50	44.44	57.14
16	-	18.35 - 19.65	≥ 412.5				
25	-	≥ 28.2	-	4	9.38	75.00	60.00

Table C-47 Summary of analysed validation results for Case 3: Reduct 5

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	Fixed Carbon (%)	Temperature (ºC)				
1	11.6 - 12.05	-	≥ 537.5	0	3.13	33.33	100.00
5	< 4.55	17 - 19.85	< 555	1	28.13	60.00	75.00
6	< 0.85	≥ 14.75	-				
8	≥ 19.1	≥ 7.85	-				
14	0.85 - 6.8	15.25 - 17	-	2	18.75	66.67	85.71
16	4.3 - 5.95	≥ 14.6	≥ 445				
23	-	≥ 28.2	-	4	9.38	75.00	60.00

Table C-48 Summary of analysed validation results for Case 3: Reduct 6

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moisture Content (%)	Temperature (ºC)				
1	1.35 - 1.9	≥ 537.5	0	6.25	66.67	100.00
2	< 1.9	< 412.5				
3	9.95 - 11.4	-	1	12.50	26.67	80.00
5	≥ 5.8	< 412.5				
15	7.4 - 8.55	≥ 502.5	2	12.50	44.44	100.00
21	≥ 61.1	-				
31	< 1.35	-	4	9.38	75.00	75.00

Table C-49 Summary of analysed validation results for Case 3: Reduct 7

No	Rules			Decision	Strength (%)	Coverage (%)	Certainty (%)
	Fixed Carbon (%)	HHV (MJ/kg)	Temperature (ºC)				
1	-	15.75 - 15.9	≥ 537.5	0	6.25	66.67	100.00
2	11.85 - 12.2	-	< 412.5				
5	15.8 - 19.85	< 18.25	< 502.5	1	21.88	46.67	70.00
6	7.85 - 16.75	19.15 - 21.15	-				
15	≥ 15.25	16.7 - 18.15	-	2	21.88	77.78	77.78
16	12.4 - 28.2	≥ 19.9	≥ 445				
17	11.4 - 15.25	16.15 - 17.35	-				
23	≥ 28.2	-	-	4	9.38	75.00	60.00

Table C-50 Summary of analysed validation results for Case 4: Reduct 1

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Moisture Content (%)				
2	71.65 - 81.35	8 - 11.4	2	13.64	50.00	60.00
5	< 69.75	≥ 9.6				
6	≥ 59.35	4.35 - 8	3	27.27	60.00	66.67
8	< 71.65	5.8 - 8.85				
13	-	3.15 - 3.65	5	9.09	100.00	100.00
15	< 55.85	-	8	9.09	100.00	100.00

Table C-51 Summary of analysed validation results for Case 4: Reduct 2

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	Moisture Content (%)				
2	0.35 - 2.55	≥ 8	2	18.18	66.67	80.00
3	≥ 4.35	< 1.45				
4	3.85 - 4.35	≥ 8				
5	4.35 - 16.1	≥ 4.35	3	36.36	80.00	61.54
7	-	5.8 - 8				
14	-	< 0.6	5	9.09	100.00	100.00
15	≥ 33.15	-	8	9.09	100.00	100.00

Table C-52 Summary of analysed validation results for Case 4: Reduct 3

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Fixed Carbon (%)				
1	-	9.85 - 11.6	1	4.55	100.00	50.00
2	73.95 - 75.15	-	2	18.18	66.67	66.67
3	≥ 75.95	16.65 - 20.45				
5	< 69.75	18.55 - 20.45				
8	< 75.15	21.25 - 28.2	3	27.27	60.00	85.71
9	69.75 - 75.95	≥ 18.45				
10	< 73.95	9.85 - 16.65				
16	80.2 - 80.6	-	5	9.09	100.00	100.00
18	< 55.85	-	8	9.09	100.00	100.00

Table C-53 Summary of analysed validation results for Case 4: Reduct 4

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	Temperature (ºC)				
2	73.95 - 75.15	-	2	13.64	50.00	50.00
7	75.95 - 76.35	-				
10	81.75 - 90.9	-	3	13.64	30.00	100.00
18	73.45 - 73.95	-				
19	59.35 - 61.95	-				
24	80.2 - 80.6	-	5	9..09	100.00	100.00
26	< 55.85	-	8	9..09	100.00	100.00

Table C-54 Summary of analysed validation results for Case 4: Reduct 5

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Volatile Matter (%)	HHV (MJ/kg)				
2	68.05 - 78.1	< 15.9	2	13.63	50.00	60.00
8	≥ 73.45	15.9 - 17.9	3	9.09	20.00	40.00
9	69.75 - 70.9	-				
19	80.2 - 80.6	-	5	9..09	100.00	100.00
21	< 55.85	-	8	9..09	100.00	100.00

Table C-55 Summary of analysed validation results for Case 4: Reduct 6

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	HHV (MJ/kg)				
3	-	15.65 - 15.9	2	13.64	50.00	100.00
7	2.55 - 7.1	15.9 - 26	3	22.72	50.00	62.50
9	4.35 - 11.6	< 17.15				
16	0.55 - 0.85	< 18.3	5	9..09	100.00	100.00
18	≥ 33.15	-	8	9..09	100.00	100.00

Table C-56 Summary of analysed validation results for Case 4: Reduct 7

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moisture Content (%)	HHV (MJ/kg)				
3	-	15.65 - 15.9	2	18.18	66.67	100.00
5	9.6 - 11.4	-				
7	4.35 - 8	≥ 14.25	3	31.82	70.00	77.78
8	≥ 11.4	< 17.9				
16	3.15 - 3.65	-	5	9.09	100.00	100.00
18	< 5.8	< 12	8	4.55	50.00	100.00

Table C-57 Summary of analysed validation results for Case 4: Reduct 8

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moisture Content (%)	Fixed Carbon (%)				
2	≥ 8	≥ 16.2	2	18.18	66.67	80.00
3	1.1 -1.45	-				
5	-	12.55 - 15.7	3	27.27	60.00	60.00
6	5.8 - 8	-				
13	3.15 - 3.65	-	5	9.09	100.00	100.00
15	5.55 - 5.85	-	8	4.55	50.00	100.00

Table C-58 Summary of analysed validation results for Case 4: Reduct 9

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Fixed Carbon (%)	HHV (MJ/kg)				
1	9.85 - 11.6	-	1	4.55	100.00	50.00
2	≥ 11.6	< 15.9	2	13.64	50.00	60.00
7	8.35 - 16.2	≥ 18.3	3	13.64	30.00	50.00
8	12.55 - 15.7	-				
20	< 15.95	18.15 - 18.3	5	9.09	100.00	66.67
22	< 9.85	< 12	8	9.09	100.00	100.00

Table C-59 Summary of analysed validation results for Case 4: Reduct 10

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Ash Content (%)	Fixed Carbon (%)				
2	< 4.35	16.2 - 20.45	2	18.18	66.67	50.00
3	-	11.6 - 12.55				
5	2.55 - 11.6	< 16.65	3	22.73	50.00	55.56
6	4.35 - 16.1	≥ 18.45				
13	0.55 - 0.85	< 15.95	5	9.09	100.00	100.00
14	< 0.85	-				
15	≥ 33.15	-	8	9.09	100.00	100.00

Table C-60 Summary of analysed validation results for Case 4: Reduct 11

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	Moist Content (%)	Temperature (ºC)				
2	≥ 8	495 - 535	2	22.72	83.33	71.43
4	1.1 - 1.45	-				
5	9.6 - 11.4	-				
7	5.8 - 8	-	3	27.27	60.00	60.00
8	5.8 - 62.3	≥ 575				
10	4.35 - 5.55	-				
19	3.15 - 3.65	-	5	9.09	100.00	100.00
21	5.55 - 5.8	-	8	4.55	100.00	100.00

Table C-61 Summary of analysed validation results for Case 4: Reduct 12

No	Rules		Decision	Strength (%)	Coverage (%)	Certainty (%)
	HHV (MJ/kg)	Temperature (ºC)				
2	15.65 - 15.9	-	2	13.64	50.00	100.00
13	15.9 - 17.9	≥ 575	3	18.18	40.00	80.00
14	18.3 - 18.65	-				
19	14.25 - 15.65	-				
24	18.15 - 18.3	≥ 495	5	9.09	100.00	100.00
27	7.4 - 12	-	8	9.09	100.00	100.00

APPENDIX D: CASE STUDY IN CHAPTER 7

GC-MS Results:

Table D-1 GC-MS result for diesel

RT	Area	Area %	Match	R. Match	Compound Name
6.66	4764726	0.83	727	891	Decane
8.63	3377294	0.59	913	929	Benzene, 1,2,4-trimethyl-
8.74	6088798	1.05	902	909	Decane
9.62	7907015	1.37	854	892	Undecane
10.46	10066590	1.74	909	923	Dodecane
11.66	8871268	1.54	824	881	Dodecane
12.03	11981998	2.08	890	920	Dodecane
12.79	3657356	0.63	758	807	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
12.87	3376610	0.58	814	877	Octane, 2,3,6-trimethyl-
13.07	7251438	1.26	661	695	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.31	11016673	1.91	896	918	Tridecane
13.48	15309139	2.65	886	899	Tridecane
13.65	3701805	0.64	852	874	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
13.84	4828131	0.84	882	893	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.16	6140765	1.06	721	747	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.24	3446818	0.60	697	824	Tridecane, 2-methyl-
14.35	3321366	0.58	750	894	Tridecane, 2-methyl-
14.50	4292493	0.74	886	908	Dodecane, 2,6,10-trimethyl-
14.83	21484084	3.72	904	905	Tetradecane
15.50	3537746	0.61	728	811	Cyclohexane, octyl-
15.56	6629260	1.15	790	869	Heptadecane, 2,6,10,14-tetramethyl-
15.60	9352007	1.62	904	910	Heptadecane, 2,6,10,14-tetramethyl-
15.78	3109651	0.54	687	906	Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene
16.07	14535260	2.52	883	887	Heptadecane, 2,6,10,14-tetramethyl-
16.11	24649662	4.27	911	925	Pentadecane
16.87	4386367	0.76	837	868	Pentadecane, 2,6,10-trimethyl-
17.30	34462092	5.97	907	910	Hexadecane
17.84	10888193	1.89	913	925	Pentadecane, 2,6,10-trimethyl-
18.02	3866066	0.67	704	749	Octadecane, 6-methyl-
18.45	32954976	5.71	890	890	Heptadecane
18.48	31911574	5.53	875	875	Pentadecane, 2,6,10,14-tetramethyl-
19.53	25991256	4.50	905	907	Eicosane
19.60	6671488	1.16	892	904	Hexadecane, 2,6,10,14-tetramethyl-
20.56	24322290	4.21	920	922	Eicosane
20.81	60354504	10.46	932	932	Hexadecanoic acid, methyl ester
21.54	20945470	3.63	921	921	Eicosane
22.41	5606499	0.97	869	876	9,12-Octadecadienoic acid (Z,Z)-, methyl ester
22.48	64651952	11.20	871	871	10-Octadecenoic acid, methyl ester
22.71	5711366	0.99	868	871	Methyl stearate
23.37	14354714	2.49	901	902	Eicosane
24.24	11442078	1.98	892	895	Eicosane
25.20	7828885	1.36	887	925	Octadecane, 2-methyl-
26.38	5132010	0.89	866	902	Octadecane, 2-methyl-
27.85	3066197	0.53	847	892	Nonadecane, 2-methyl-

Table D-2 GC-MS result for bio-oil

RT	Area	Area %	Match	R. Match	Compound Name
3.78	23,610,478	6.66	895	904	Acetic acid
3.81	53,489,560	15.10	682	798	Formic acid hydrazide
3.98	19,378,148	5.47	572	876	Formamide
4.32	5,279,309	1.49	523	879	Hydrogen azide
4.37	10,380,873	2.93	453	758	Formamide
4.49	2,257,410	0.64	447	778	Dimethylamine
4.52	2,848,451	0.80	497	900	Silane, methyl-
4.55	1,676,396	0.47	415	729	Threo-3-bromo-2-pentanol
4.58	1,520,537	0.43	470	849	Formic acid hydrazide
4.69	1,819,441	0.51	447	865	Methane, nitroso-
7.82	13,879,391	3.92	707	781	2-Aminopyrimidine-1-oxide
7.97	4,149,613	1.17	598	818	2-Thiophenecarboxylic acid, 4-cyanophenyl ester
8.10	1,655,791	0.47	472	797	3,4-Furandiol, tetrahydro-, trans-
8.13	1,560,378	0.44	463	763	1H-Imidazole-4-carboxaldehyde
8.20	1,664,658	0.47	440	735	Pyrazole, 1,4-dimethyl-
10.54	135,165,856	38.15	916	929	Phenol
10.84	50,939,704	14.38	786	827	Phenol, 2-methoxy-
11.08	10,704,461	3.02	420	742	Pyrrole, 4-ethyl-2-methyl-
11.25	2,395,276	0.68	420	753	Benzene, (1,1-dimethylethoxy)-
14.97	2,901,670	0.82	791	855	Phenol, 2,6-dimethoxy-
16.02	5,338,457	1.51	572	704	1,2,4-Trimethoxybenzene
16.04	1,672,589	0.47	366	741	Hexyl methyl ethylphosphonate

Table D-3 GC-MS result for solvent-extracted bio-oil at 20 wt.% 2-Heptanol

RT	**Area**	**Area %**	**Match**	**R. Match**	**Compound Name**
2.20	6494234	1.20	520	719	2-Chloroethyl methyl sulfoxide
2.36	2783374	0.51	888	893	Acetone
3.56	11026101	2.03	923	924	Acetic acid
3.68	4486446	0.83	841	841	Propane, 2,2-dimethoxy-
7.17	4308301	0.79	899	935	Furfural
8.67	325703456	60.04	913	913	2-Heptanol
9.16	4160975	0.77	632	747	6-Oxa-bicyclo[3.1.0]hexan-3-one
9.56	3929289	0.72	723	779	3-Aminopyrazine 1-oxide`
10.39	32933824	6.07	936	951	Phenol
12.00	6435717	1.19	862	880	Phenol, 2-methoxy-
13.64	4077414	0.75	904	913	Creosol
14.91	3158687	0.58	833	895	Phenol, 4-ethyl-2-methoxy-
15.94	5803105	1.07	855	874	Phenol, 2,6-dimethoxy-
17.20	1948840	0.36	710	799	1,2,4-Trimethoxybenzene
17.31	2785705	0.51	550	785	Phenol, 2-methoxy-4-(1-propenyl)-
18.16	2693023	0.50	675	738	5-tert-Butylpyrogallol
22.96	119719952	22.07	744	801	n-Hexadecanoic acid

Table D-4 GC-MS result for solvent-extracted bio-oil at 20 wt.% 2-Octanol

RT	**Area**	**Area %**	**Match**	**R. Match**	**Compound Name**
2.08	7682798	2.98	731	804	12-Methyl-E,E-2,13-octadecadien-1-ol
2.36	3869319	1.50	916	935	Acetone
2.51	2258961	0.88	892	921	Acetic acid, methyl ester
2.67	2592088	1.01	463	837	Dimethyl ether
3.39	7338069	2.85	879	881	Acetic acid
3.66	5649017	2.19	869	869	Propane, 2,2-dimethoxy-
4.82	280943	0.11	726	864	Propanoic acid
7.13	1876945	0.73	881	924	Furfural
7.66	379382	0.15	752	831	2-Furanmethanol
9.53	1961175	0.76	733	765	2-Furanethanol, á-methoxy-(S)-
10.30	8926521	3.47	915	929	Phenol
10.50	200574848	77.86	935	935	2-Octanol
11.50	391092	0.15	858	895	Phenol, 3-methyl-
11.87	338471	0.13	848	890	p-Cresol
11.95	2688778	1.04	922	927	Phenol, 2-methoxy-
13.60	1477343	0.57	915	917	Creosol

Table D-4 GC-MS result for solvent-extracted bio-oil at 20 wt.% 2-Octanol (continued)

14.86	1086927	0.42	885	904	Phenol, 4-ethyl-2-methoxy-
15.41	293199	0.11	714	825	2-Methoxy-4-vinylphenol
15.88	1754060	0.68	902	904	Phenol, 2,6-dimethoxy-
17.13	759555	0.29	759	786	1,2,4-Trimethoxybenzene
17.26	343745	0.13	767	856	Phenol, 2-methoxy-4-(1-propenyl)-
18.10	622398	0.24	741	761	5-tert-Butylpyrogallol
20.23	521853	0.20	701	807	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
22.89	3932021	1.53	783	830	n-Hexadecanoic acid

Table D-5 GC-MS result for solvent-extracted bio-oil at 20 wt.% 2-Octanone

RT	Area	Area %	Match	R. Match	Compound Name
2.36	2273250	0.44	905	913	Acetone
3.43	7810739	1.52	874	877	Acetic acid
3.69	4892537	0.95	826	826	Propane, 2,2-dimethoxy-
7.16	4853063	0.94	917	948	Furfural
9.54	3967056	0.77	730	762	2-Furanethanol, á-methoxy-(S)-
10.25	395568512	77.01	886	888	2-Octanone
10.96	5622837	1.09	697	740	2-Cyclopenten-1-one, 2-hydroxy-3-methyl-
11.53	40842140	7.95	629	830	1-Octene, 2-methoxy-
11.98	23471146	4.57	908	918	Phenol, 2-methoxy-
12.55	1696148	0.33	779	893	1-Octene, 2-methoxy-
13.63	3062003	0.60	879	884	Creosol
14.88	2823767	0.55	883	901	Phenol, 4-ethyl-2-methoxy-
15.91	4078483	0.79	894	900	Phenol, 2,6-dimethoxy-
17.17	1584365	0.31	752	794	1,2,4-Trimethoxybenzene
18.13	1648356	0.32	684	744	5-tert-Butylpyrogallol
22.92	9488370	1.85	761	809	n-Hexadecanoic acid

Table D-6 GC-MS result for aqueous bio-oil at 20 wt.% 2-Heptanol

RT	Area	Area %	Match	R. Match	Compound Name
2.36	2151111	5.64	866	878	Acetone
2.74	650889	1.71	922	951	Formic acid
3.54	15783130	41.36	912	912	Acetic acid
3.68	796959	2.09	844	848	Propane, 2,2-dimethoxy-
3.97	2175034	5.70	817	828	Acetic acid, methyl ester
7.16	1401427	3.67	848	912	Furfural
8.58	3153796	8.27	926	926	2-Heptanol
9.55	811863	2.13	727	769	2-Furanethanol, á-methoxy-(S)-
10.35	1752068	4.59	859	899	Phenol
15.91	700685	1.84	720	832	Formic acid, 2,6-dimethoxyphenyl ester
22.93	8781275	23.01	762	823	n-Hexadecanoic acid

Table D-7 GC-MS result for aqueous bio-oil at 20 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.37	986154	5.53	889	891	Acetone
2.52	288513	1.62	864	884	Acetic acid, methyl ester
3.38	8365044	46.93	925	926	Acetic acid
3.68	533389	2.99	851	854	Propane, 2,2-dimethoxy-
3.92	908668	5.10	918	920	2-Propanone, 1-hydroxy-
7.13	496223	2.78	840	918	Furfural
9.54	582499	3.27	717	746	2-Furanethanol, á-methoxy-(S)-
10.25	682635	3.83	910	931	Phenol
10.43	670346	3.76	866	874	2-Octanol
22.89	4310376	24.18	755	803	n-Hexadecanoic acid

Table D-8 GC-MS result for aqueous bio-oil at 20 wt.% 2-Octanone

RT	Area	Area %	Match	R. Match	Compound Name
2.34	2195318	8.65	910	918	Acetone
2.49	1419559	5.59	929	938	Acetic acid, methyl ester
2.64	1637866	6.45	905	948	Formic acid
3.14	322614	1.27	755	847	Ethyl Acetate
3.43	9961028	39.24	914	915	Acetic acid
3.63	668861	2.63	888	900	Propane, 2,2-dimethoxy-
3.88	1011279	3.98	817	837	Acetic acid, methyl ester
7.10	513504	2.02	787	899	Furfural
9.50	296345	1.17	660	739	2-Furanethanol, á-methoxy-(S)-
10.16	569885	2.25	841	861	2-Octanone
22.87	2398099	9.45	834	864	n-Hexadecanoic acid
22.93	4390097	17.29	761	907	n-Hexadecanoic acid

Table D-9 GC-MS result for solvent-extracted bio-oil at 5 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.51	1700201	0.47	951	951	Acetic acid, methyl ester
3.33	18413408	5.14	898	899	Acetic acid
3.95	1602364	0.45	896	896	2-Propanone, 1-hydroxy-
4.58	687629	0.19	840	857	Propanoic acid
5.56	665373	0.19	828	836	2,2-Dimethoxybutane
7.11	7834933	2.19	895	944	Furfural
7.60	1819761	0.51	824	833	2-Furanmethanol
7.79	1537144	0.43	867	870	2-Propanone, 1-(acetyloxy)-
8.58	1256356	0.35	869	871	2-Cyclopenten-1-one, 2-methyl-
8.99	682256	0.19	803	821	1,2-Cyclopentanedione
9.52	8912101	2.49	739	767	2-Furanethanol, á-methoxy-(S)-
10.17	60152964	16.79	940	949	Phenol
10.44	182750944	51.01	929	929	2-Octanol
10.86	1103643	0.31	813	827	1,2-Cyclopentanedione, 3-methyl-
11.40	3171448	0.89	769	871	Phenol, 2-methyl-
11.77	2546675	0.71	885	900	Phenol, 3-methyl-
11.91	11090160	3.10	937	940	Phenol, 2-methoxy-
12.31	1506373	0.42	854	903	Benzofuran, 2-methyl-
12.74	676405	0.19	770	830	Phenol, 2-ethyl
12.93	1252031	0.35	842	886	Phenol, 2,4-dimethyl-
13.22	1144328	0.32	799	819	Phenol, 3-ethyl-
13.36	620527	0.17	846	869	2-Methoxy-6-methylphenol
13.57	7611967	2.12	913	913	Creosol
14.64	889453	0.25	787	847	1,2-Benzenediol, 3-methoxy-
14.82	5973841	1.67	916	927	Phenol, 4-ethyl-2-methoxy-
15.36	1871105	0.52	856	892	2-Methoxy-4-vinylphenol
15.83	9749076	2.72	923	925	Phenol, 2,6-dimethoxy-
15.93	827850	0.23	792	824	Eugenol
16.06	995897	0.28	854	891	Phenol, 2-methoxy-4-propyl-
17.08	5804166	1.62	813	831	1,2,4-Trimethoxybenzene
17.20	2880618	0.80	903	932	Phenol, 2-methoxy-4-(1-propenyl)-
17.34	1934736	0.54	869	898	Methylparaben
18.06	3912643	1.09	779	781	5-tert-Butylpyrogallol
18.48	855655	0.24	809	858	Dodecanoic acid
18.98	630579	0.18	894	910	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
19.57	610053	0.17	826	841	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
20.17	2595604	0.72	896	902	Phenol, 2,6-dimethoxy-4-(2-propenyl)-

Table D-10 GC-MS result for solvent-extracted bio-oil at 10 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.38	5575829	1.32	944	947	Acetone
2.53	1151977	0.27	885	885	Acetic acid, methyl ester
3.25	17971238	4.25	926	926	Acetic acid
3.68	20727876	4.90	898	898	Propane, 2,2-dimethoxy-
3.91	807865	0.19	887	887	2-Propanone, 1-hydroxy-
4.53	710359	0.17	818	837	Propanoic acid
5.55	665636	0.16	855	860	2,2-Dimethoxybutane
5.77	547283	0.13	860	868	1-Hydroxy-2-butanone
7.10	7066023	1.67	914	952	Furfural
7.59	1695218	0.40	867	872	2-Furanmethanol
7.78	1136185	0.27	843	850	2-Propanone, 1-(acetyloxy)-
8.01	502757	0.12	867	878	Furan, tetrahydro-2,5-dimethoxy-
8.35	505688	0.12	716	759	Furan, tetrahydro-2,5-dimethoxy-
8.57	937406	0.22	800	836	2-Cyclopenten-1-one, 2-methyl-
8.98	533716	0.13	853	869	1,2-Cyclopentanedione
9.51	6776889	1.60	735	762	2-Furanethanol, á-methoxy-(S)-
10.16	44337096	10.49	945	952	Phenol
10.44	257655696	60.94	921	921	2-Octanol
10.85	993134	0.23	852	866	1,2-Cyclopentanedione, 3-methyl-
11.39	2371494	0.56	742	854	Phenol, 2-methyl-
11.76	1952138	0.46	902	917	Phenol, 3-methyl-
11.91	8120913	1.92	950	953	Phenol, 2-methoxy-
12.31	1022119	0.24	794	867	Benzofuran, 2-methyl-
12.74	538752	0.13	733	825	Phenol, 2-ethyl-
12.94	985232	0.23	838	886	Phenol, 2,4-dimethyl-
13.21	1183984	0.28	786	816	Phenol, 4-ethyl-
13.56	6336124	1.50	928	928	Creosol
13.85	1458602	0.35	705	895	Catechol
14.82	4458922	1.05	919	929	Phenol, 4-ethyl-2-methoxy-
15.36	1279827	0.30	866	898	2-Methoxy-4-vinylphenol
15.82	7276674	1.72	935	937	Phenol, 2,6-dimethoxy-
15.92	582689	0.14	809	827	Eugenol
16.06	918910	0.22	786	838	Phenol, 2-methoxy-4-propyl-
17.08	4348289	1.03	795	811	1,2,4-Trimethoxybenzene
17.20	2156547	0.51	896	927	Phenol, 2-methoxy-4-(1-propenyl)-
17.35	1395405	0.33	847	891	Methylparaben
18.05	2906264	0.69	788	790	5-tert-Butylpyrogallol
18.97	527894	0.12	858	877	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
19.57	525810	0.12	832	850	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
20.17	2138779	0.51	895	899	Phenol, 2,6-dimethoxy-4-(2-propenyl)-

Table D-11 GC-MS result for solvent-extracted bio-oil at 15 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.36	831144	0.17	861	903	Acetone
2.51	1086555	0.22	965	965	Acetic acid, methyl ester
3.30	17585224	3.59	910	910	Acetic acid
3.69	2281176	0.47	866	866	Propane, 2,2-dimethoxy-
3.95	1019051	0.21	880	892	2-Propanone, 1-hydroxy-
4.56	669664	0.14	841	860	Propanoic acid
5.56	578042	0.12	827	836	2,2-Dimethoxybutane
5.78	461978	0.09	819	849	1-Hydroxy-2-butanone
7.10	4118189	0.84	860	925	Furfural
7.58	1415417	0.29	860	866	Phenol Decreasing.
7.78	825180	0.17	877	892	2-Propanone, 1-(acetyloxy)-
8.01	636122	0.13	857	865	Furan, tetrahydro-2,5-dimethoxy-
8.33	631739	0.13	768	808	Furan, tetrahydro-2,5-dimethoxy-
8.57	459352	0.09	873	890	2-Cyclopenten-1-one, 2-methyl-
8.98	479529	0.10	817	839	1,2-Cyclopentanedione
9.50	10792296	2.20	729	758	2-Furanethanol, á-methoxy-(S)-
10.16	41015516	8.37	941	948	Phenol
10.44	362932224	74.06	923	923	2-Octanol
10.84	782826	0.16	843	857	1,2-Cyclopentanedione, 3-methyl-

Table D-11 GC-MS result for solvent-extracted bio-oil at 15 wt.% 2-Octanol (continued)

11.37	2224289	0.45	784	877	Phenol, 2-methyl-
11.75	1681502	0.34	905	913	Phenol, 3-methyl-
11.90	7064702	1.44	934	937	Phenol, 2-methoxy-
12.30	960275	0.20	801	862	Benzofuran, 2-methyl-
12.74	408683	0.08	707	797	Phenol, 2-ethyl-
12.92	754697	0.15	832	881	Phenol, 2,4-dimethyl-
13.22	861695	0.18	775	797	Phenol, 4-ethyl-
13.55	5431032	1.11	933	933	Creosol
14.64	885633	0.18	801	879	1,2-Benzenediol, 3-methoxy-
14.81	3582011	0.73	899	909	Phenol, 4-ethyl-2-methoxy-
15.35	1058732	0.22	831	871	2-Methoxy-4-vinylphenol
15.82	6122154	1.25	922	923	Phenol, 2,6-dimethoxy-
15.92	438710	0.09	682	747	Eugenol
16.06	627269	0.13	777	828	Phenol, 2-methoxy-4-propyl-
17.07	3216597	0.66	803	821	1,2,4-Trimethoxybenzene
17.20	1423097	0.29	895	925	Phenol, 2-methoxy-4-(1-propenyl)-
17.33	1030020	0.21	877	908	Methylparaben
18.05	1980700	0.40	759	761	5-tert-Butylpyrogallol
18.47	395873	0.08	763	835	Dodecanoic acid
20.17	1333962	0.27	873	879	Phenol, 2,6-dimethoxy-4-(2-propenyl)-

Table D-12 GC-MS result for solvent-extracted bio-oil at 20 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.36	1036868	0.17	924	939	Acetone
2.51	1181150	0.20	932	932	Acetic acid, methyl ester
3.05	593224	0.10	888	940	2-Butanone
3.31	18833852	3.11	906	907	Acetic acid
3.68	2565026	0.42	873	873	Propane, 2,2-dimethoxy-
3.95	1214705	0.20	925	932	2-Propanone, 1-hydroxy-
4.56	759806	0.13	807	827	Propanoic acid
5.56	711035	0.12	855	863	2,2-Dimethoxybutane
7.11	6866566	1.13	895	942	Furfural
7.60	1558394	0.26	887	891	2-Furanmethanol
7.79	1079251	0.18	882	890	2-Propanone, 1-(acetyloxy)-
8.03	606308	0.10	848	864	Furan, tetrahydro-2,5-dimethoxy-
8.36	610574	0.10	720	753	Furan, tetrahydro-2,5-dimethoxy-
8.59	581021	0.10	830	839	2-Cyclopenten-1-one, 2-methyl-
8.99	641527	0.11	836	854	1,2-Cyclopentanedione
9.52	7541849	1.25	727	756	2-Furanethanol, á-methoxy-(S)-
10.20	42741452	7.06	949	958	Phenol
10.46	463974464	76.62	905	905	2-Octanol
10.86	983210	0.16	852	863	1,2-Cyclopentanedione, 3-methyl-
11.39	2389709	0.39	781	892	Phenol, 2-methyl-
11.77	1800304	0.30	877	891	Phenol, 3-methyl-
11.91	8127146	1.34	942	945	Phenol, 2-methoxy-
12.32	960121	0.16	757	838	Benzofuran, 2-methyl-
12.93	851368	0.14	828	871	Phenol, 2,4-dimethyl-
13.22	715484	0.12	765	797	Phenol, 4-ethyl-
13.56	5985779	0.99	926	926	Creosol
13.84	1566947	0.26	746	882	Catechol
14.64	1391343	0.23	869	906	1,2-Benzenediol, 3-methoxy-
14.82	3882634	0.64	892	904	Phenol, 4-ethyl-2-methoxy-
15.36	1233866	0.20	859	892	2-Methoxy-4-vinylphenol
15.83	7135183	1.18	923	924	Phenol, 2,6-dimethoxy-
15.92	543176	0.09	804	825	Eugenol
16.06	880941	0.15	771	823	Phenol, 2-methoxy-4-propyl-
17.08	4269336	0.71	800	816	1,2,4-Trimethoxybenzene
17.20	2040526	0.34	905	930	Phenol, 2-methoxy-4-(1-propenyl)-
17.34	1499281	0.25	898	926	Methylparaben
18.05	2782874	0.46	769	770	5-tert-Butylpyrogallol
18.48	930041	0.15	817	856	Dodecanoic acid
18.98	540463	0.09	836	836	Phenol, 2,6-dimethoxy-4-(2-propenyl)-
20.17	1939874	0.32	888	894	Phenol, 2,6-dimethoxy-4-(2-propenyl)-

Table D-13 GC-MS result for aqueous bio-oil at 5 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.36	2180242	3.91	926	932	Acetone
2.51	1243155	2.23	914	916	Acetic acid, methyl ester
2.63	903785	1.62	849	863	Formic acid
3.05	426105	0.76	782	899	2-Butanone
3.37	26271484	47.08	911	914	Acetic acid
3.68	1625069	2.91	885	886	Propane, 2,2-dimethoxy-
3.94	2119883	3.80	921	925	2-Propanone, 1-hydroxy-
4.56	359425	0.64	768	835	Propanoic acid
5.80	570014	1.02	820	856	1-Hydroxy-2-butanone
7.11	1783035	3.20	838	910	Furfural
7.60	671422	1.20	754	777	2-Furanmethanol
7.80	507379	0.91	857	867	2-Propanone, 1-(acetyloxy)-
8.02	273108	0.49	803	831	Furan, tetrahydro-2,5-dimethoxy-
8.35	237657	0.43	773	806	Furan, tetrahydro-2,5-dimethoxy-
8.58	216433	0.39	715	786	2-Cyclopenten-1-one, 2-methyl-
8.69	541990	0.97	690	789	2(5H)-Furanone
8.98	370146	0.66	786	832	1,2-Cyclopentanedione
9.52	2460609	4.41	709	737	2-Furanethanol, á-methoxy-(S)-
10.12	7178304	12.86	940	949	Phenol
10.39	2589869	4.64	795	802	2-Octanol
10.84	289683	0.52	742	796	1,2-Cyclopentanedione, 3-methyl-
11.39	235797	0.42	728	819	Phenol, 2-methyl-
11.76	156172	0.28	733	802	Phenol, 3-methyl-
11.91	927109	1.66	933	936	Phenol, 2-methoxy-
13.56	558932	1.00	851	861	Creosol
14.81	136056	0.24	767	824	Phenol, 4-ethyl-2-methoxy-
15.82	757290	1.36	874	878	Phenol, 2,6-dimethoxy-
17.08	215954	0.39	765	800	1,2,4-Trimethoxybenzene

Table D-14 GC-MS result for aqueous bio-oil at 10 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.36	2736490	3.50	936	940	Acetone
2.51	1049966	1.34	936	937	Acetic acid, methyl ester
2.63	1365091	1.75	851	861	Formic acid
3.05	233719	0.30	873	970	2-Butanone
3.39	37695752	48.26	912	915	Acetic acid
3.68	3119368	3.99	889	889	Propane, 2,2-dimethoxy-
3.94	3737610	4.79	935	937	2-Propanone, 1-hydroxy-
4.56	654318	0.84	791	827	Propanoic acid
4.70	163099	0.21	712	823	Acetoin
5.80	988818	1.27	869	879	1-Hydroxy-2-butanone
6.22	212220	0.27	816	957	Hydrogen azide
7.12	3061843	3.92	856	918	Furfural
7.45	477918	0.61	702	771	Ethane, 1,1,1-trimethoxy-
7.60	662617	0.85	767	773	2-Furanmethanol
7.80	651507	0.83	932	938	2-Propanone, 1-(acetyloxy)-
8.03	294185	0.38	819	841	Furan, tetrahydro-2,5-dimethoxy-
8.36	287434	0.37	810	833	Furan, tetrahydro-2,5-dimethoxy-
8.58	308056	0.39	633	697	2-Cyclopenten-1-one, 2-methyl-
8.99	652547	0.84	833	848	1,2-Cyclopentanedione
9.52	2796574	3.58	725	755	2-Furanethanol, á-methoxy-(S)-
10.12	7819134	10.01	941	949	Phenol
10.39	3457680	4.43	898	914	2-Octanol
10.84	568829	0.73	849	862	1,2-Cyclopentanedione, 3-methyl-
11.39	319510	0.41	744	847	Phenol, 2-methyl-
11.77	174000	0.22	756	826	Phenol, 3-methyl-
11.91	1297648	1.66	930	935	Phenol, 2-methoxy-
13.56	537979	0.69	847	856	Creosol
13.84	323131	0.41	742	847	Catechol
14.63	346307	0.44	827	871	1,2-Benzenediol, 3-methoxy-
14.82	173785	0.22	768	825	Phenol, 4-ethyl-2-methoxy-

Table D-14 GC-MS result for aqueous bio-oil at 10 wt.% 2-Octanol (continued)

15.83	1358244	1.74	891	891	Phenol, 2,6-dimethoxy-
17.08	408722	0.52	745	765	1,2,4-Trimethoxybenzene
18.05	175539	0.22	728	765	5-tert-Butylpyrogallol

Table D-15 GC-MS result for aqueous bio-oil at 15 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.35	1221293	1.81	927	936	Acetone
2.51	903794	1.34	901	902	Acetic acid, methyl ester
2.63	1329733	1.97	877	881	Formic acid
2.96	937629	1.39	854	888	n-Hexane
3.05	209496	0.31	860	905	2-Butanone
3.37	36172740	53.52	904	906	Acetic acid
3.68	1022357	1.51	871	873	Propane, 2,2-dimethoxy-
3.94	3638652	5.38	909	909	2-Propanone, 1-hydroxy-
4.53	565807	0.84	786	848	Propanoic acid
5.79	943089	1.40	841	853	1-Hydroxy-2-butanone
6.21	177390	0.26	725	794	Propanoic acid, 2-oxo-, methyl ester
7.11	2856701	4.23	846	905	Furfural
7.60	655089	0.97	748	755	2-Furanmethanol
7.79	631183	0.93	882	885	2-Propanone, 1-(acetyloxy)-
8.03	258415	0.38	808	837	Furan, tetrahydro-2,5-dimethoxy-
8.36	224952	0.33	787	812	Furan, tetrahydro-2,5-dimethoxy-
8.69	739669	1.09	725	781	2(5H)-Furanone
8.97	584485	0.86	836	850	1,2-Cyclopentanedione
9.51	2118057	3.13	745	774	2-Furanethanol, á-methoxy-(S)-
10.12	5193435	7.68	925	933	Phenol
10.38	2682123	3.97	875	891	2-Octanol
10.83	461158	0.68	814	834	1,2-Cyclopentanedione, 3-methyl-
11.38	181931	0.27	676	797	Phenol, 2-methyl-
11.75	144494	0.21	678	786	Phenol, 3-methyl-
11.90	926825	1.37	888	894	Phenol, 2-methoxy-
13.56	386937	0.57	842	859	Creosol
13.84	359089	0.53	725	846	Catechol
14.63	192151	0.28	794	836	1,2-Benzenediol, 3-methoxy-
15.82	1011242	1.50	898	899	Phenol, 2,6-dimethoxy-
17.08	312018	0.46	746	776	1,2,4-Trimethoxybenzene
17.89	411027	0.61	700	756	Phenol, 2,4-bis(1,1-dimethylethyl)-
18.16	129194	0.19	712	827	2-Propanone, 1-(4-hydroxy-3-methoxyphenyl)-

Table D-16 GC-MS result for aqueous bio-oil at 20 wt.% 2-Octanol

RT	Area	Area %	Match	R. Match	Compound Name
2.36	1175764	2.76	938	949	Acetone
2.51	823804	1.93	944	945	Acetic acid, methyl ester
2.64	896617	2.10	887	896	Formic acid
3.05	294719	0.69	799	908	2-Butanone
3.36	23273764	54.63	906	907	Acetic acid
3.68	649008	1.52	874	875	Propane, 2,2-dimethoxy-
3.94	2449434	5.75	908	908	2-Propanone, 1-hydroxy-
4.53	354868	0.83	745	813	Propanoic acid
5.80	505238	1.19	867	882	1-Hydroxy-2-butanone
7.12	1426440	3.35	838	921	Furfural
7.60	285356	0.67	737	749	2-Furanmethanol
7.79	462643	1.09	898	902	2-Propanone, 1-(acetyloxy)-
8.03	213358	0.50	812	842	Furan, tetrahydro-2,5-dimethoxy-
8.35	191333	0.45	816	849	Furan, tetrahydro-2,5-dimethoxy-
8.69	483874	1.14	696	812	2(5H)-Furanone
8.98	386911	0.91	798	820	1,2-Cyclopentanedione
9.51	2164956	5.08	747	776	2-Furanethanol, á-methoxy-(S)-
10.12	2502129	5.87	917	925	Phenol
10.39	1713722	4.02	895	904	2-Octanol
10.84	303466	0.71	866	900	1,2-Cyclopentanedione, 3-methyl-
11.90	519645	1.22	890	900	Phenol, 2-methoxy-
13.56	197807	0.46	762	795	Creosol

Table D-16 GC-MS result for aqueous bio-oil at 20 wt.% 2-Octanol (continued)

14.64	131088	0.31	720	812	1,2-Benzenediol, 3-methoxy-
15.83	603780	1.42	894	896	Phenol, 2,6-dimethoxy-
17.07	165009	0.39	739	788	1,2,4-Trimethoxybenzene
17.89	346601	0.81	665	718	Phenol, 2,4-bis(1,1-dimethylethyl)-
18.05	82641	0.19	674	688	5-tert-Butylpyrogallol

Table D-17 GC-MS result for 2-Octanol-extracted bio-oil/diesel emulsion with 5 wt.% surfactant

RT	Area	Area %	Match	R. Match	Compound Name
6.84	2377735	0.49	913	914	Nonane
8.75	4762455	0.98	898	917	Nonane
8.81	23301776	4.81	881	881	2-Octanol
9.14	2292709	0.47	672	785	Benzene, 1-ethyl-3-methyl-
10.47	10266583	2.12	907	915	Dodecane
10.76	2327237	0.48	815	842	Bicyclo[4.1.0]heptan-3-one, 4,7,7-trimethyl-, [1R-(1à,4á,6à)]-
12.03	13319191	2.75	879	905	Dodecane
12.22	3099370	0.64	823	914	Undecane, 2,6-dimethyl-
12.67	3083208	0.64	837	877	Cyclohexane, hexyl-
13.44	2656870	0.55	812	871	Cyclohexane, hexyl-
13.48	18079244	3.74	891	902	Tridecane
13.84	5584727	1.15	878	887	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.16	7552321	1.56	729	757	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.35	3689285	0.76	760	895	Tridecane, 2-methyl-
14.51	5832992	1.21	888	909	Dodecane, 2,6,10-trimethyl-
14.60	2488609	0.51	848	930	Decahydro-4,4,8,9,10-pentamethylnaphthalene
14.84	27486169	5.68	899	906	Tetradecane
15.19	2864039	0.59	731	759	Benzene, 1-cyclohexyl-3-methyl-
15.25	3047349	0.63	850	880	Naphthalene, decahydro-1,6-dimethyl-4-(1-methylethyl)-
15.31	2420731	0.50	795	886	Decahydro-4,4,8,9,10-pentamethylnaphthalene
15.61	11782807	2.43	921	925	Heptadecane, 2,6,10,14-tetramethyl-
15.65	3891676	0.80	791	850	Tetradecane, 2-methyl-
15.79	2791581	0.58	737	902	Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene
16.11	27990230	5.78	912	916	Pentadecane
16.80	2632423	0.54	685	797	Tridecane, 5-cyclohexyl-
16.88	3392498	0.70	833	874	Nonadecane, 2-methyl-
16.97	2230648	0.46	707	845	Pentadecane, 3-methyl-
17.31	25173290	5.20	907	908	Hexadecane
17.84	8470381	1.75	903	918	Pentadecane, 2,6,10-trimethyl-
18.03	2806627	0.58	722	723	Dodecane, 5,8-diethyl-
18.45	23601046	4.88	900	901	Heptadecane
18.49	24140940	4.99	900	902	Pentadecane, 2,6,10,14-tetramethyl-
19.53	19512562	4.03	904	906	Eicosane
19.60	4962631	1.03	874	888	Hexadecane, 2,6,10,14-tetramethyl-
20.56	18650692	3.85	924	925	Eicosane
20.80	46524908	9.61	933	934	Hexadecanoic acid, methyl ester
21.54	16316416	3.37	929	929	Eicosane
22.41	4981332	1.03	886	904	Methyl 9-cis,11-trans-octadecadienoate
22.48	50934064	10.52	864	864	10-Octadecenoic acid, methyl ester
22.71	4328913	0.89	867	868	Methyl stearate
23.37	11129382	2.30	904	904	Eicosane
24.24	8705618	1.80	892	919	Octadecane, 2-methyl-
25.20	6117803	1.26	882	915	Octadecane, 2-methyl-
26.38	3842216	0.79	857	890	Octadecane, 2-methyl-
27.85	2533328	0.52	839	870	Eicosane, 2-methyl-

Table D-18 GC-MS result for 2-Octanol-extracted bio-oil/diesel emulsion with 10 wt.% surfactant

RT	Area	Area %	Match	R. Match	Compound Name
2.02	52440952	6.27	919	936	Methylene chloride
2.37	3973076	0.48	861	873	2-Butanone, 3-methyl-
2.44	22296304	2.67	865	875	Butane, 2-chloro-2-methyl-
6.84	4292034	0.51	910	911	Nonane
8.64	3879391	0.46	922	936	Benzene, 1,2,4-trimethyl-
8.75	8288107	0.99	890	892	Decane
8.81	42239568	5.05	915	915	2-Octanol
10.46	17197392	2.06	900	909	Dodecane
12.03	21527342	2.57	870	894	Dodecane
12.22	4942757	0.59	842	927	Undecane, 2,6-dimethyl-
12.66	4847549	0.58	786	827	Cyclohexane, hexyl-
13.07	13492168	1.61	668	703	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.43	4892415	0.59	818	888	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.48	29412920	3.52	900	909	Tridecane
13.84	8863481	1.06	867	880	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.16	11603838	1.39	729	752	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.35	5798440	0.69	755	891	Tridecane, 2-methyl-
14.50	11181634	1.34	894	915	Dodecane, 2,6,10-trimethyl-
14.60	3934859	0.47	859	928	Decahydro-4,4,8,9,10-pentamethylnaphthalene
14.84	40820144	4.88	895	903	Tetradecane
15.19	4269697	0.51	691	736	Benzene, 1-cyclohexyl-3-methyl-
15.25	4929519	0.59	846	889	Naphthalene, decahydro-1,6-dimethyl-4-(1-methylethyl)-
15.31	3864995	0.46	793	863	Decahydro-4,4,8,9,10-pentamethylnaphthalene
15.47	7706275	0.92	765	785	Naphthalene, 1,2,3,4-tetrahydro-1,4,6-trimethyl-
15.51	5987878	0.72	705	800	Dodecane, 2-cyclohexyl-
15.61	18481376	2.21	898	901	Heptadecane, 2,6,10,14-tetramethyl-
15.65	5741210	0.69	814	863	Tetradecane, 2-methyl-
15.79	4317461	0.52	741	923	Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene
16.11	43466488	5.20	907	912	Pentadecane
16.80	3946869	0.47	687	797	Tridecane, 5-cyclohexyl-
16.87	5133339	0.61	837	882	Nonadecane, 2-methyl-
17.31	37812332	4.52	905	906	Hexadecane
17.84	12078066	1.44	897	911	Pentadecane, 2,6,10-trimethyl-
18.03	4003182	0.48	733	734	Dodecane, 5,8-diethyl-
18.45	35853680	4.29	898	899	Heptadecane
18.49	35222980	4.21	900	901	Pentadecane, 2,6,10,14-tetramethyl-
19.53	28785568	3.44	905	908	Eicosane
19.60	7135223	0.85	874	887	Hexadecane, 2,6,10,14-tetramethyl-
20.56	27272448	3.26	914	916	Eicosane
20.80	68693848	8.21	924	924	Hexadecanoic acid, methyl ester
21.54	22794510	2.73	914	914	Eicosane
22.41	10334490	1.24	917	918	9,12-Octadecadienoic acid (Z,Z)-, methyl ester
22.48	75578016	9.04	873	873	10-Octadecenoic acid, methyl ester
22.70	6298197	0.75	885	886	Methyl stearate
23.37	15333134	1.83	898	900	Eicosane
24.23	11885138	1.42	897	901	Eicosane
25.20	8283170	0.99	892	899	Eicosane
26.38	5136400	0.61	864	901	Octadecane, 2-methyl-

Table D-19 GC-MS result for 2-Octanol-extracted bio-oil/diesel emulsion with 15 wt.% surfactant

RT	Area	Area %	Match	R. Match	Compound Name
2.02	44263856	6.65	937	950	Methylene chloride
2.37	3997591	0.60	765	847	4-Penten-2-one
2.44	18030432	2.71	865	870	Butane, 2-chloro-2-methyl-
6.85	2931974	0.44	903	904	Nonane
8.76	5713589	0.86	898	900	Decane
8.82	33220936	4.99	930	931	2-Octanol
10.48	12615753	1.89	900	909	Dodecane
12.04	16649342	2.50	875	901	Dodecane
12.23	3820683	0.57	830	923	Undecane, 2,6-dimethyl-
12.67	3262975	0.49	770	841	Cyclohexane, hexyl-
13.09	10405504	1.56	670	705	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.45	3884430	0.58	833	888	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.49	22667078	3.40	896	905	Tridecane
13.85	6790051	1.02	885	898	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.17	9239339	1.39	722	751	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.36	4467436	0.67	730	895	Tridecane, 2-methyl-
14.52	6937604	1.04	896	919	Dodecane, 2,6,10-trimethyl-
14.61	2922195	0.44	835	930	Decahydro-4,4,8,9,10-pentamethylnaphthalene
14.85	18522552	2.78	903	909	Tetradecane
15.20	3335119	0.50	703	737	Benzene, 1-cyclohexyl-3-methyl-
15.26	3795771	0.57	805	877	Naphthalene, decahydro-1,6-dimethyl-4-(1-methylethyl)-
15.32	2816059	0.42	800	882	Decahydro-4,4,8,9,10-pentamethylnaphthalene
15.49	5647624	0.85	753	776	Naphthalene, 1,2,3,4-tetrahydro-1,4,6-trimethyl-
15.52	5072962	0.76	682	787	Dodecane, 2-cyclohexyl-
15.62	14309189	2.15	909	914	Heptadecane, 2,6,10,14-tetramethyl-
15.66	4760264	0.72	801	870	Tetradecane, 2-methyl-
15.80	3344756	0.50	739	920	Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene
16.12	34228256	5.14	912	918	Pentadecane
16.81	3146227	0.47	698	811	Tridecane, 5-cyclohexyl-
16.89	4247777	0.64	837	887	Nonadecane, 2-methyl-
16.98	2920350	0.44	708	839	Pentadecane, 3-methyl-
17.32	30786396	4.62	908	908	Hexadecane
17.86	10566637	1.59	913	926	Pentadecane, 2,6,10-trimethyl-
18.04	3286197	0.49	701	702	Dodecane, 5,8-diethyl-
18.46	28611286	4.30	906	906	Heptadecane
18.50	29776462	4.47	898	899	Pentadecane, 2,6,10,14-tetramethyl-
19.54	23777910	3.57	913	915	Eicosane
19.61	6058795	0.91	890	902	Hexadecane, 2,6,10,14-tetramethyl-
20.57	22790984	3.42	921	923	Eicosane
20.81	57413524	8.62	928	929	Hexadecanoic acid, methyl ester
21.55	19575840	2.94	923	923	Eicosane
22.42	9137961	1.37	922	923	9,12-Octadecadienoic acid (Z,Z)-, methyl ester
22.49	64838108	9.74	876	876	10-Octadecenoic acid, methyl ester
22.72	5700727	0.86	899	900	Methyl stearate
23.38	13170710	1.98	904	906	Eicosane
24.24	10352500	1.56	895	901	Eicosane
25.21	7212878	1.08	889	910	Eicosane
26.39	4716700	0.71	883	916	Octadecane, 2-methyl-

Table D-20 GC-MS result for 2-Heptanol-extracted bio-oil/diesel emulsion with 15 wt.% surfactant

RT	Area	Area %	Match	R. Match	Compound Name
2.02	51496436	6.23	937	944	Methylene chloride
2.37	4257518	0.52	742	754	2-Butanone, 3-methyl-
2.45	20179920	2.44	880	886	Butane, 2-chloro-2-methyl-
2.87	4137444	0.50	782	823	Heptane
6.84	3936807	0.48	916	918	Nonane
6.92	45229180	5.48	931	931	2-Heptanol
8.64	3666900	0.44	923	930	Benzene, 1,2,4-trimethyl-
8.75	8845418	1.07	893	895	Decane
10.47	16930434	2.05	914	924	Dodecane
10.76	3626360	0.44	804	931	Bicyclo[4.1.0]heptan-3-one, 4,7,7-trimethyl-, [1R-(1à,4á,6à)]-
12.03	21743066	2.63	883	912	Dodecane
12.22	4933774	0.60	842	928	Undecane, 2,6-dimethyl-
12.66	4947015	0.60	846	887	Cyclohexane, hexyl-
13.08	13570991	1.64	665	702	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.43	4923951	0.60	818	883	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.48	29384058	3.56	894	905	Tridecane
13.84	8742345	1.06	879	891	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.17	11708993	1.42	757	784	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.35	5770080	0.70	763	891	Tridecane, 2-methyl-
14.51	8873529	1.07	902	924	Dodecane, 2,6,10-trimethyl-
14.60	3667294	0.44	850	934	Decahydro-4,4,8,9,10-pentamethylnaphthalene
14.84	42248232	5.11	902	909	Tetradecane
15.19	4244397	0.51	718	748	Benzene, 1-cyclohexyl-3-methyl-
15.25	4757321	0.58	833	889	Naphthalene, decahydro-1,6-dimethyl-4-(1-methylethyl)-
15.61	18056284	2.19	892	894	Heptadecane, 2,6,10,14-tetramethyl-
15.65	6102686	0.74	813	865	Tetradecane, 2-methyl-
15.79	4142459	0.50	743	920	Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene
16.11	42384308	5.13	905	910	Pentadecane
16.80	3908476	0.47	700	806	Tridecane, 5-cyclohexyl-
16.88	5144111	0.62	840	882	Nonadecane, 2-methyl-
17.31	37650092	4.56	892	892	Hexadecane
17.85	12718891	1.54	911	928	Pentadecane, 2,6,10-trimethyl-
18.03	3903135	0.47	714	714	Dodecane, 5,8-diethyl-
18.45	34065692	4.12	897	898	Heptadecane
18.49	36602956	4.43	906	908	Pentadecane, 2,6,10,14-tetramethyl-
19.53	28584220	3.46	897	900	Eicosane
19.60	7318629	0.89	866	876	Hexadecane, 2,6,10,14-tetramethyl-
20.56	27080978	3.28	929	931	Eicosane
20.81	69407128	8.40	925	925	Hexadecanoic acid, methyl ester
21.54	23133574	2.80	915	915	Eicosane
22.41	11265037	1.36	900	901	9,12-Octadecadienoic acid (Z,Z)-, methyl ester
22.48	75782896	9.17	867	867	10-Octadecenoic acid, methyl ester
22.71	6737016	0.82	886	888	Methyl stearate
23.37	15109996	1.83	902	904	Eicosane
24.24	11694040	1.42	894	896	Eicosane
25.20	8337055	1.01	885	891	Eicosane
26.38	5130198	0.62	883	916	Octadecane, 2-methyl-

Table D-21 GC-MS result for 2-Octanone-extracted bio-oil/diesel emulsion with 15 wt.% surfactant

RT	Area	Area %	Match	R. Match	Compound Name
2.02	43778416	6.90	928	938	Methylene chloride
2.44	16355701	2.58	824	843	Butane, 2-chloro-2-methyl-
6.85	2586800	0.41	899	901	Nonane
8.57	37653492	5.94	884	890	2-Octanone
8.76	5842040	0.92	920	923	Decane
10.47	11703097	1.84	889	897	Dodecane
10.77	2559587	0.40	806	834	Bicyclo[4.1.0]heptan-3-one, 4,7,7-trimethyl-, [1R-(1à,4á,6à)]-
12.04	15515375	2.45	893	925	Dodecane
12.23	3592275	0.57	829	927	Undecane, 2,6-dimethyl-
13.08	9693642	1.53	666	703	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.44	3477058	0.55	819	886	Naphthalene, 1,2,3,4-tetrahydro-6-methyl-
13.49	21557338	3.40	891	901	Tridecane
13.85	6459849	1.02	872	884	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.17	8663058	1.37	735	762	Naphthalene, 1,2,3,4-tetrahydro-2,6-dimethyl-
14.36	4136242	0.65	746	898	Tridecane, 2-methyl-
14.51	6564826	1.03	882	902	Dodecane, 2,6,10-trimethyl-
14.60	2801287	0.44	858	940	Decahydro-4,4,8,9,10-pentamethylnaphthalene
14.84	30939102	4.88	898	905	Tetradecane
15.19	3149452	0.50	741	770	Benzene, 1-cyclohexyl-3-methyl-
15.25	3539538	0.56	853	885	Naphthalene, decahydro-1,6-dimethyl-4-(1-methylethyl)-
15.31	2777331	0.44	793	865	Decahydro-4,4,8,9,10-pentamethylnaphthalene
15.61	13696896	2.16	903	905	Heptadecane, 2,6,10,14-tetramethyl-
15.65	4475668	0.71	824	873	Tetradecane, 2-methyl-
15.79	3226932	0.51	709	901	Decahydro-8a-ethyl-1,1,4a,6-tetramethylnaphthalene
16.11	31942096	5.04	909	914	Pentadecane
16.81	3071849	0.48	699	802	Tridecane, 5-cyclohexyl-
16.88	3991231	0.63	842	888	Nonadecane, 2-methyl-
16.97	2694827	0.42	670	813	Pentadecane, 3-methyl-
17.32	28460650	4.49	922	922	Hexadecane
17.85	9847972	1.55	909	922	Pentadecane, 2,6,10-trimethyl-
18.04	3101111	0.49	723	724	Dodecane, 5,8-diethyl-
18.46	27257290	4.30	895	896	Heptadecane
18.49	27319780	4.31	898	899	Pentadecane, 2,6,10,14-tetramethyl-
19.54	22080012	3.48	907	910	Eicosane
19.61	5636411	0.89	861	872	Hexadecane, 2,6,10,14-tetramethyl-
20.56	21033532	3.32	911	913	Eicosane
20.81	54182156	8.54	923	923	Hexadecanoic acid, methyl ester
21.54	17900946	2.82	918	918	Eicosane
22.42	9036623	1.42	922	923	9,12-Octadecadienoic acid (Z,Z)-, methyl ester
22.48	60843740	9.59	869	869	10-Octadecenoic acid, methyl ester
22.71	5423941	0.86	882	884	Methyl stearate
23.38	12211397	1.93	905	907	Eicosane
24.24	9636901	1.52	894	898	Eicosane
25.21	6791801	1.07	885	890	Eicosane
26.39	4310198	0.68	874	911	Octadecane, 2-methyl-
27.85	2837216	0.45	830	881	Nonadecane, 2-methyl-

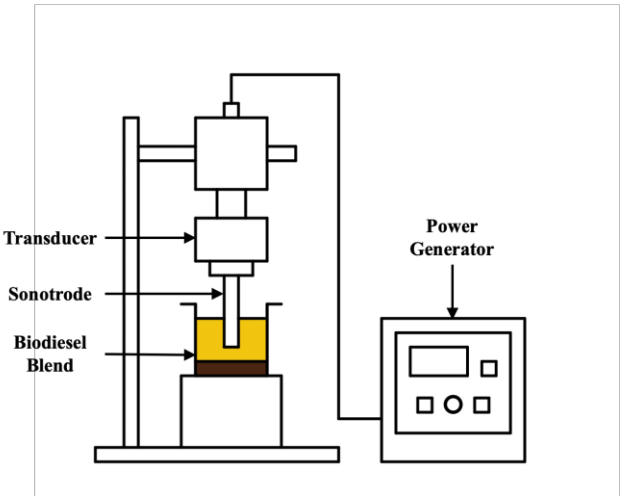


Figure D-7 Schematic representation of ultrasonic emulsification setup

FT-IR Spectra:

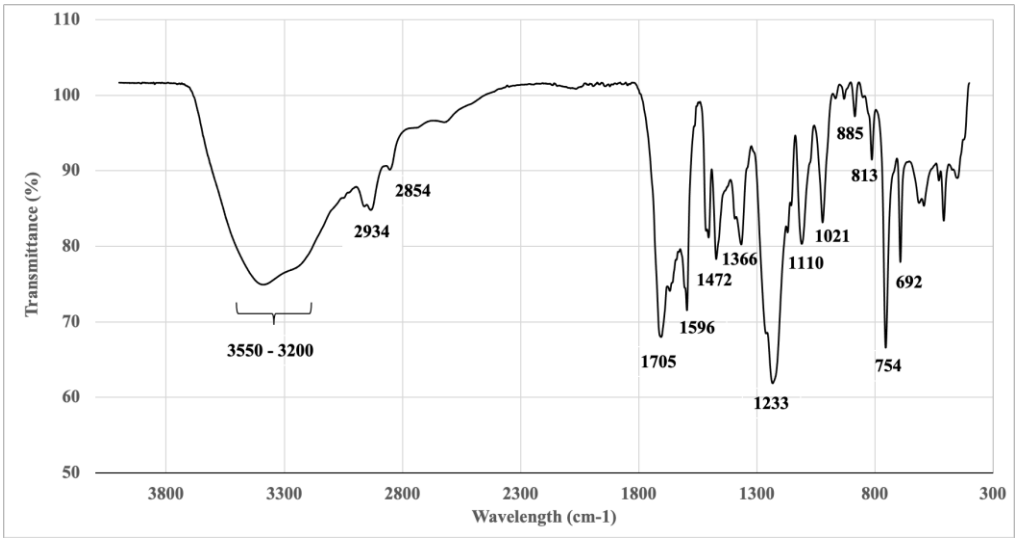


Figure D-8 FT-IR spectra of PKS bio-oil

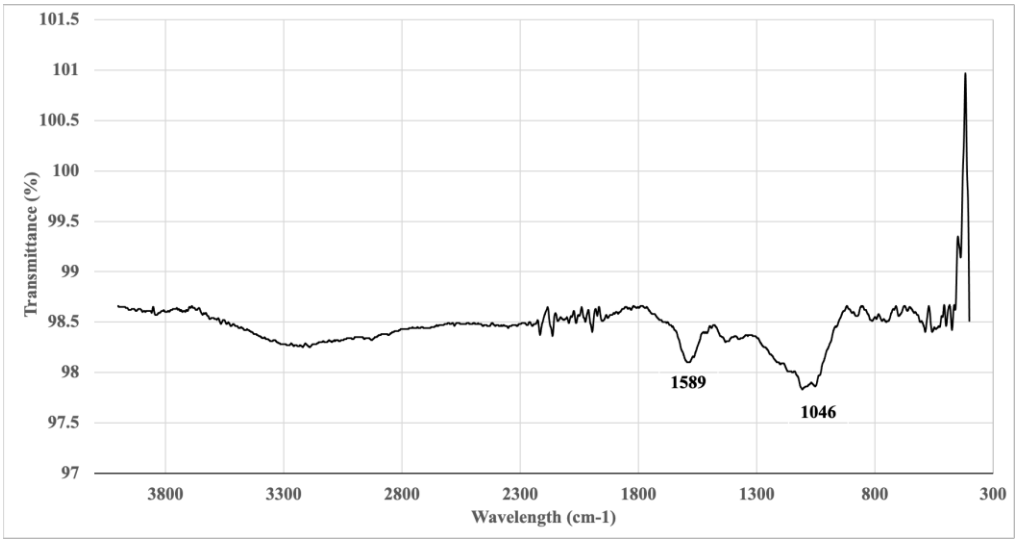


Figure D-9 FT-IR spectra of biochar

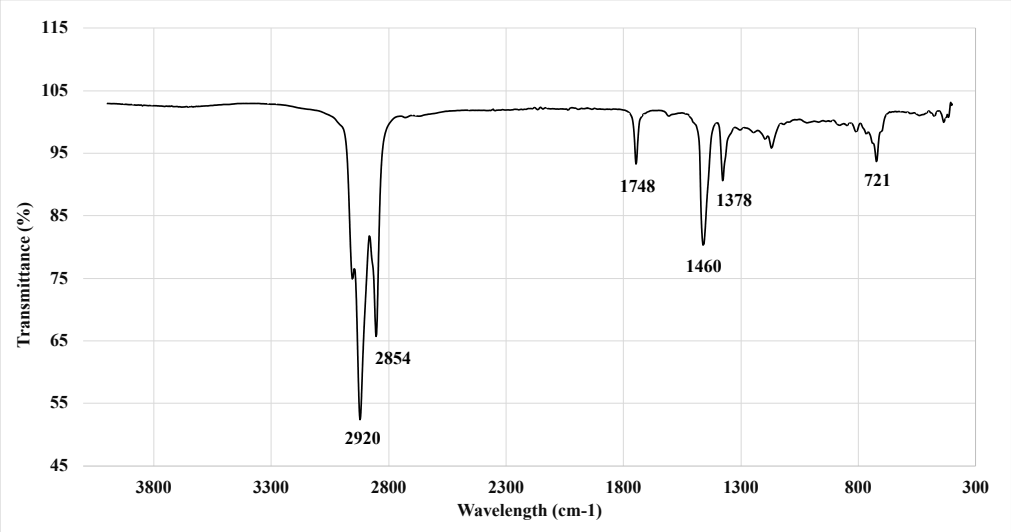


Figure D-10 FT-IR spectra of Diesel

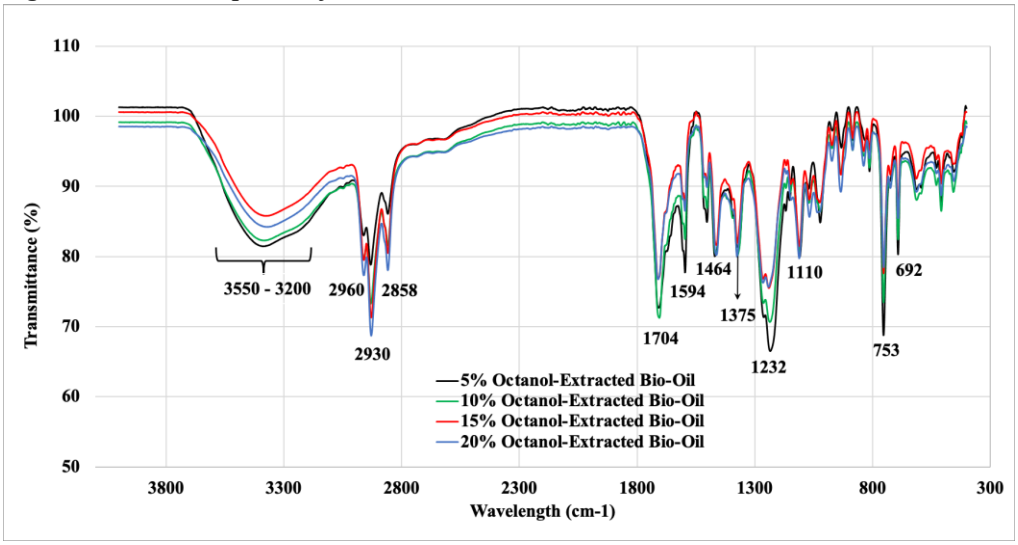


Figure D-11 FT-IR spectra of 2-Octanol-extracted bio-oil at different 2-Octanol mixing ratio

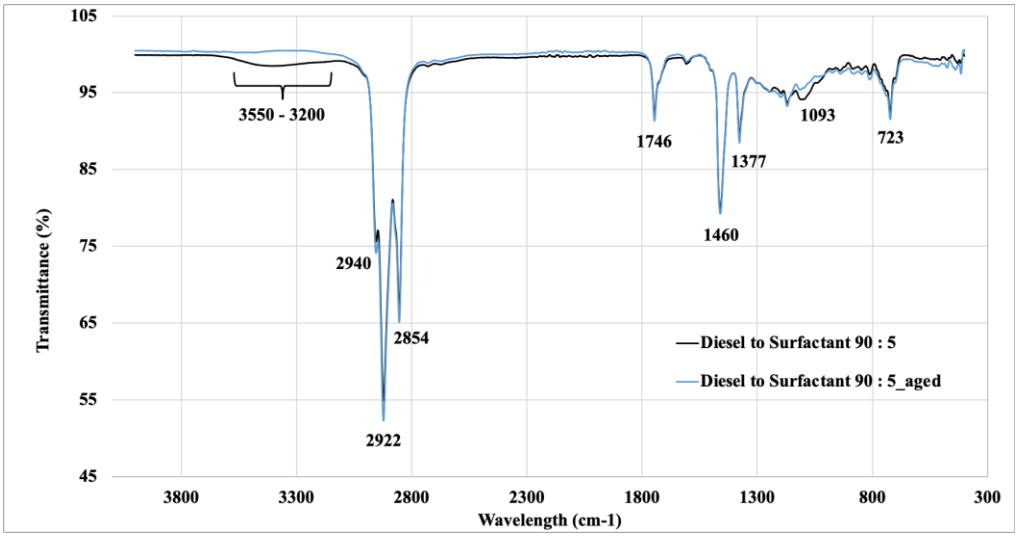


Figure D-12 FT-IR spectra of 2-Octanol-extracted bio-oil/diesel emulsion at diesel to surfactant ratio of 90:5, before and after accelerated aging

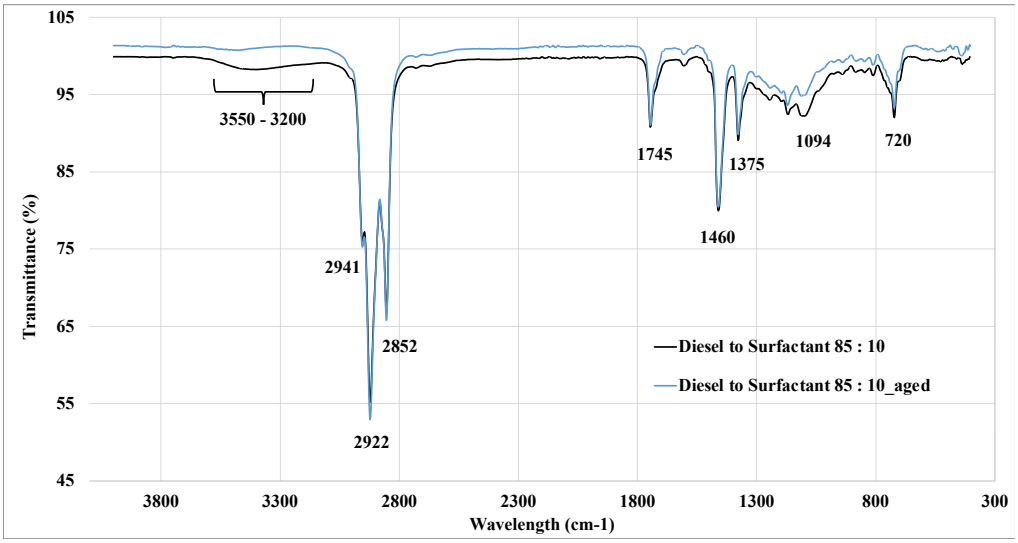


Figure D-13 FT-IR spectra of 2-Octanol-extracted bio-oil/diesel emulsion at diesel to surfactant ratio of 85:10, before and after accelerated aging

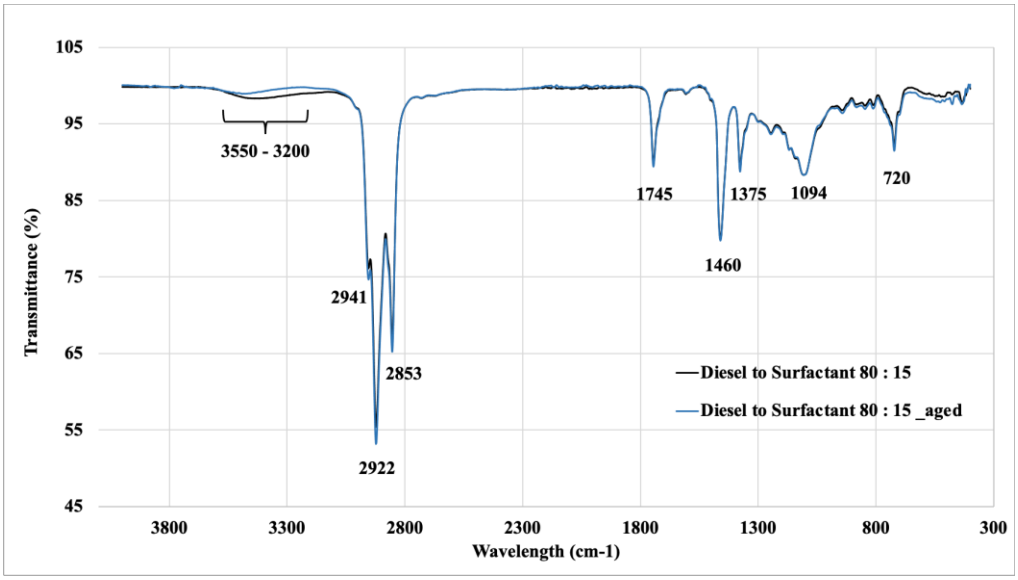
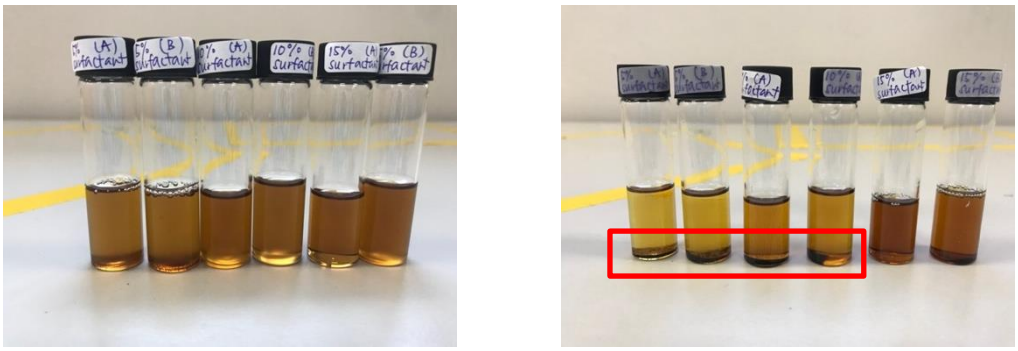
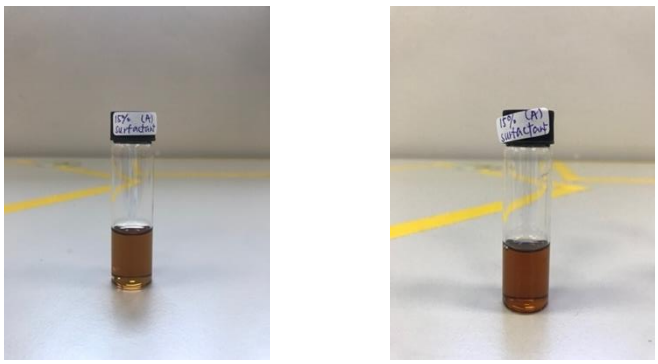


Figure D-14 FT-IR spectra of 2-Octanol-extracted bio-oil/diesel emulsion at diesel to surfactant ratio of 80:15, before and after accelerated aging



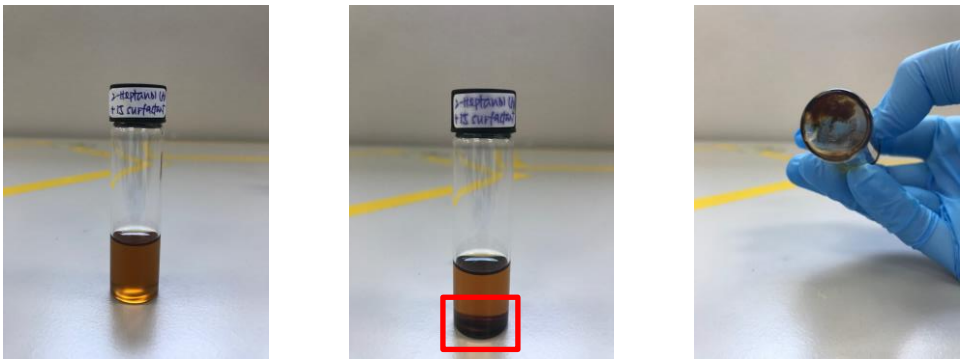
(a) (b)

Figure D-15 Solvent-extracted bio-oil/diesel emulsion (a) before and; (b) after accelerated aging at different diesel to surfactant ratio, from left: 90:5(A); 90:5(B); 85:10(A); 85:10(B); 80:15(A); 80:15(B)



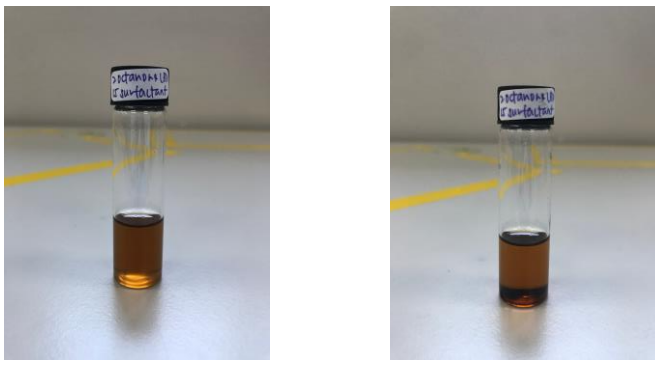
(a) (b)

Figure D-16 2-Octanol-extracted bio-oil/diesel emulsion (a) before accelerated aging; (b) after accelerated aging



(a) (b) (c)

Figure D-17 2-Heptanol-extracted bio-oil/diesel emulsion (a) before accelerated aging; (b) after accelerated aging; (c) bio-oil sediment after accelerated aging



(a) (b)

Figure D-18 2-Octanone-extracted bio-oil/diesel emulsion (a) before accelerated aging; (b) after accelerated aging