



**VICTORIA UNIVERSITY**  
MELBOURNE AUSTRALIA

*A hybrid SVR-PSO model to predict a CFD-based optimised bubbling fluidised bed pyrolysis reactor*

This is the Accepted version of the following publication

Jalalifar, Salman, Masoudi, Mojtaba, Abbassi, Rouzbeh, Garaniya, Vikram, Ghiji, Mohammadmahdi and Salehi, Fatemeh (2020) A hybrid SVR-PSO model to predict a CFD-based optimised bubbling fluidised bed pyrolysis reactor. *Energy*, 191. ISSN 0360-5442

The publisher's official version can be found at  
<https://www.sciencedirect.com/science/article/pii/S0360544219321097>  
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/40873/>

# A Hybrid SVR-PSO Model to Predict a CFD-based Optimised Bubbling Fluidised Bed Pyrolysis Reactor

Salman Jalalifar <sup>1</sup>, Mojtaba Masoudi <sup>2</sup>, Rouzbeh Abbassi <sup>\*3</sup>, Vikram Garaniya <sup>1</sup>, Mohammadmahdi Ghiji <sup>4</sup>, Fatemeh Salehi <sup>3</sup>

<sup>1</sup> Australian Maritime College, College of Sciences and Engineering, University of Tasmania, Launceston, Tasmania, Australia

<sup>2</sup> Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.

<sup>3</sup> School of Engineering, Faculty of Science and Engineering, Macquarie University, Sydney, NSW, Australia

<sup>4</sup> Institute of Sustainable Industries and Liveable Cities, Victoria University, Victoria, Australia

\* Corresponding Author E-mail address: Rouzbeh.Abbassi@mq.edu.au

## Abstract

Comprehensive scrutiny is necessary to achieve an optimised set of operating conditions for a pyrolysis reactor to attain the maximum amount of the desired product. To reach this goal, a computational fluid dynamic (CFD) model is developed for biomass fast pyrolysis process and it is validated using the experiment of a standard lab-scale bubbling fluidised bed reactor. This is followed by a detailed CFD parametric study. Key influencing parameters investigated are operating temperature, biomass flow rate, biomass and sand particle sizes, carrier gas velocity, biomass injector location, and pre-treatment temperature. Machine learning algorithms (MLAs) are then employed to predict the optimised conditions that lead to the maximum bio-oil yield. For this purpose, support vector regression with particle swarm optimisation algorithm (SVR-PSO) is developed and applied to the CFD datasets to predict the optimum values of parameters. The maximum bio-oil yield is then computed using the optimum values of the parameters. The CFD simulation is also performed using the optimum parameters obtained by the SVR-PSO. The CFD results and the values predicted by the MLA for the product yields are finally compared where a good agreement is achieved.

**Keywords:** Support Vector Regression (SVR), Particle Swarm Optimisation (PSO), computational fluid dynamic (CFD) simulation, bubbling fluidised bed reactor, fast pyrolysis process

## 1. Introduction

Environmental concern about global warming and climate change, as well as energy crisis such as the depletion of fossil fuels, have encouraged researchers to develop eco-friendly energy sources [1]. An alternative sustainable source of energy is biomass which significantly produces lower sulfur dioxides (SO<sub>2</sub>) and particulate matter (PM) as well as less carbon dioxide (CO<sub>2</sub>) due to its carbon neutrality [2]. Biomass can be obtained from various natural sources such as agricultural products and waste [3-6], forest residue [7], land and aquatic animals and it has a high conversion capability to other forms of energy, i.e. thermochemical conversions such as direct combustion, gasification, and pyrolysis process [8]. Products generated by pyrolysis processes have some benefits over other thermochemical conversion processes in term of their applications. For instance, bio-oil as liquid fuel can be a source of high-value chemicals, solid biochar is a source for adsorbent, soil amendment, and/or catalyst, while biogases are commonly employed for energy recovery [9-12]. There are three main types of pyrolysis processes; slow, fast, and flash pyrolysis. Solid biochar is the main product

of slow pyrolysis [13] whereas, fast and flash pyrolysis are favourable for the production of liquid bio-oil. Co-generation of heat and power in boilers, supplementary fuel in gas turbines and diesel engines are some of the applications of pyrolysis bio-oil [10, 14].

Pyrolysis process is the oxygen-free thermal decomposition of biomass. Generally, fast pyrolysis process occurs in a few seconds at moderate temperatures (about 500 °C). During the past years, fast pyrolysis process has been studied by many researchers experimentally [3, 15-18] and numerically using Computational Fluid Dynamics [19-24]. Although experimental studies are highly valuable, they are expensive, time-consuming and limited due to the shortage of measurements in severe reactor conditions (high temperature and pressure). Alternatively, CFD can provide a detailed understanding of complex physical phenomena inside the reactor with a relatively lower cost. Many types of research have focused on CFD simulations of the pyrolysis process [24-28] and studied the effects of different operating factors on the product yields [25]. A parametric study has confirmed that the optimum temperature for the production of bio-oil is in the range of 500–525 °C. Shorter residence time at higher temperatures reduces chances for the secondary crack of condensable vapours to non-condensable gases which is favourable to achieve higher bio-oil yields. Two options are proposed for having shorter residence time, these options are higher nitrogen velocities and higher biomass injector heights.

Another important parameter is the particle sizes. The intraparticle temperature gradient is higher for larger biomass particle sizes, leading to accumulation of unreacted biomass inside the reactor. The use of larger sand particles requires higher carrier gas velocity for efficient particle fluidisation which results in a higher rate of bio-oil productions. Niemelä et al. [29] suggest that while using large scale CFD simulation, considering the whole size distribution results in more accurate devolatilisation schemes. To address the effect of the biomass size on the devolatilisation scheme, ground and sieved biomass particles with three different size ranges are considered; small (112-125 µm), medium (500-600 µm) and large (800-1000 µm) fractions. While faster devolatilisation is predicted for small particles, medium and large particles show similar mass loss behaviour. The difference between the chemical kinetics of small particles and medium/large ones are explained by the compensation of the internal heat transfer resistance which is neglected in the isothermal assumption.

Preheating the biomass up to the temperature of 400 K also leads to increased bio-oil yields. A higher rate of the bio-oil yield production is expected from the feedstocks that have higher contents of cellulose and hemicellulose. The highest amount of bio-oil is expected from pyrolysis of pure cellulose feedstock [25, 26]. When lignin is considered as a feedstock, the temperature has a minor effect on the bio-oil yield for values higher than 530 °C. The molecular weight of bio-oil is a key feature for the upgrading of the oil to chemicals and/or fuels and it is managed with pressure when different types of lignins with various molecular weight are utilised [30].

In the previous work by Jalalifar et al. [25], effects of various operating parameters on product yields variations were separately studied by varying only one parameter while other parameters were maintained matching the base condition [25]. However, in reality, operating parameters are not independent and considering a variation of all the parameters simultaneously is essential to effectively optimise the process. Such CFD analysis requires a high number of simulations which is computationally expensive. Alternatively, machine learning algorithms (MLAs) can be applied to predict the optimised condition by considering a set of parameters to achieve the maximum favourable product.

The substantial advancement of MLAs has spanned the application of this data-driven approach in science and engineering. Recently, MLAs have been introduced in gasification and pyrolysis processes. Çepelioğullar et al. [26] employed MLAs to predict the thermal behaviour of refuse-derived fuel (RDF) as heterogeneous fuel at the heating rates where the experimental data was not available. Mutlu and Yucel [31] considered two types of machine learning classifiers to estimate the gas composition and its calorific value obtained by gasification of woody biomass in a downdraft gasifier. The non-linear relationship between kinetic parameters and biomass components has encouraged the application of MLAs for analysing such complex relationships [32]. In a study by Sunphorka et al. [33], MLA was applied to obtain the correlation between biomass components, including cellulose, hemicellulose, and lignin, and the kinetic parameters such as activation energy ( $E_a$ ), pre-exponential factor ( $k_0$ ) and reaction order ( $n$ ). They found the highest values of all kinetic parameters occurred for cellulose. Hough et al. [34] presented MLAs and decision trees to reduce the computational cost of using detailed kinetic models in simulations. Sun et al. [35] applied MLAs to study effects of operating temperature, biomass particle size, and space velocity on predicting the yield of the gas products from pyrolysis of pine sawdust which is a typical industrial biomass waste.

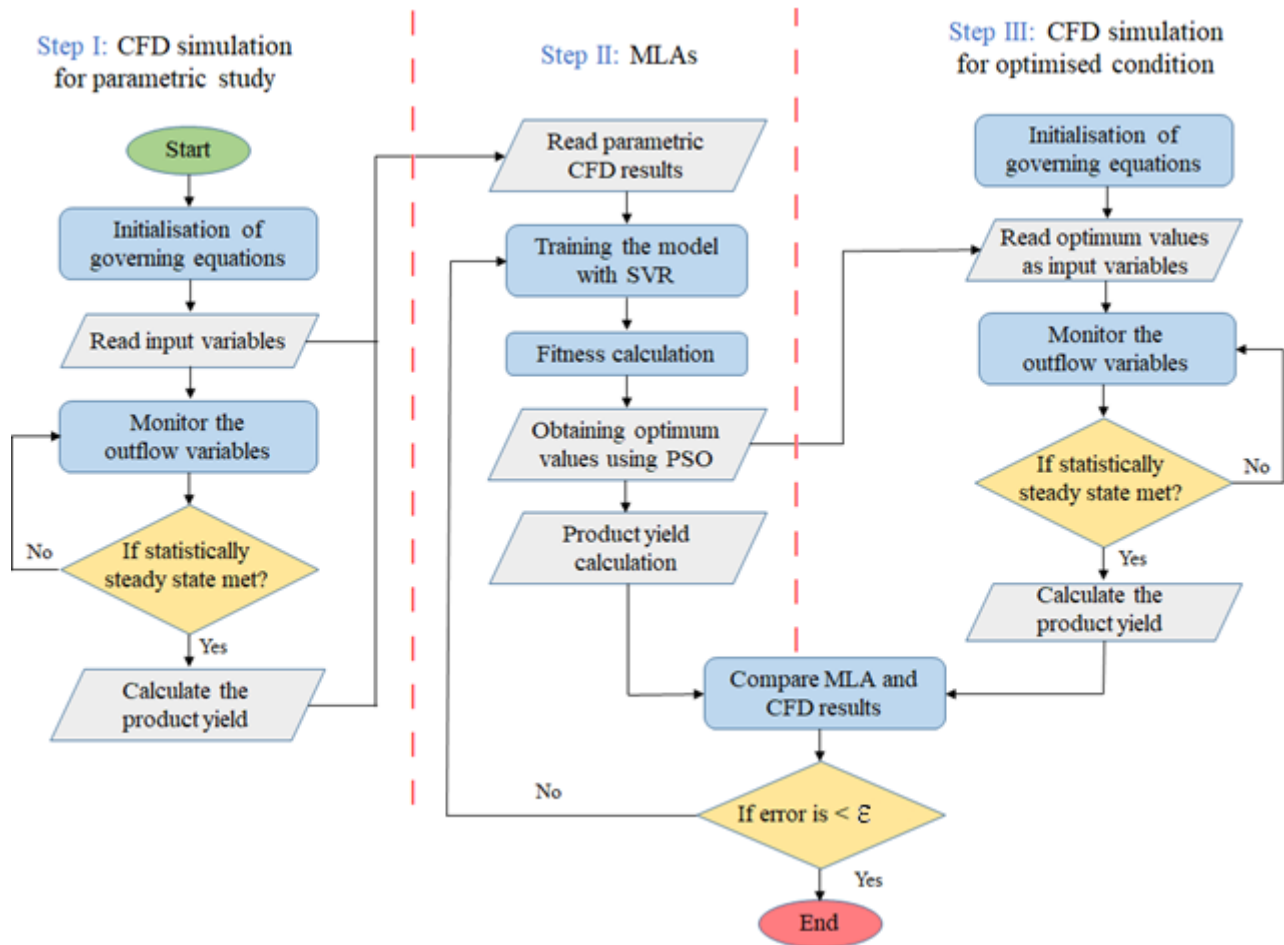
Machine learning techniques use various regressions such as Linear Regression (LR), Least Absolute Shrinkage and Selection Operator (LASSO), Least Angle Regression (LARS) and SVR to analyse the relationship between a dependent variable and a set of independent variables [36, 37]. While LR, LASSO and LARS are linear methods and suitable for linear data, SVR performs more effectively for non-linear datasets. In the present study, SVR is applied because our CFD dataset showed a non-linear complex trend. Particle Swarm Optimisation (PSO) algorithm is also employed to find an optimum value after data modelling by SVR. The implementation of the PSO algorithm is relatively easy and it requires a fewer number of particles to be tuned [38]. In this study, data materials are small and not diverse, which may reduce the generalisation of the regressor [39]. However, the use of SVR model diminishes this issue due to its appealing property and high capacity for generalisation with a relatively small number of training data. SVR minimises the generalised error bound and includes the combination of the training error and a regularisation term. A regularised term controls the complexity of the hypothesis space that achieves a generalised performance instead of only minimising the observed training error. Moreover, unlike Simple Linear Regression (SLR) in which the results depend on Gauss-Markov assumptions, SVR is a non-parametric technique. Hence, the model output does not rely on distributions of the underlying dependent and independent variables but it is affected by kernel functions [40, 41].

This paper first presents the CFD model for a 2-D standard lab-scale bubbling fluidised bed reactor. A parametric study is then conducted to address the effect of the most important influential parameters on the product yields in the process. It is followed by the development of a hybrid SVR-PSO model to predict an optimised set of parameters that achieve the maximum bio-oil yield. The model is trained using the datasets provided by the CFD analysis. The set of parameters attained by the SVR-PSO model are then applied to the CFD simulation to compute new values of product yields. Finally, the values obtained from MLAs and CFD are statistically compared to analyse the accuracy of the SVR-PSO model.

## 2. Methodology

As illustrated in Fig.1, the overall methodology can be divided into three steps. In the first step, the CFD code reads the input data such as initial mass and volume fractions, temperature, pressure, and

inlet velocity. All the flow variables including velocity, pressure, temperature, and the volume fraction of each phase are then computed by solving the governing equations. Similarly, the species mass fractions are obtained using the chemical reactions and species transport equations. Once the statistically steady state condition is achieved for the outflow variables, including the outflow temperature and the product outflux, the yield of the products is calculated. In the second step, the MLA model takes the simulation parameters and the product yields obtained using the CFD (totalling 82 datasets) as the input. The MLA model is initially trained by the CFD data, and then the SVR is applied to estimate the values for points which are not available in the input CFD datasets. This is followed by fitting the best function to the given data. In the next sub-step, the PSO algorithm is employed to find an optimum set of parameters based on the defined fitness function. When the optimum set of parameters for having maximum condensable yield is obtained, all the product yields will be recalculated. In the third step, the obtained optimum values for a set of parameters are given to the CFD code as the input. The CFD simulations will run until the statistically steady state condition is achieved and consequently, the product yield is computed. Finally, the obtained product yields using the MLA and the CFD are evaluated and compared. The optimum values are determined if the deviation between MLA and CFD results is less than 5%. Details of the CFD simulation and the MLA model are described in the following sections.



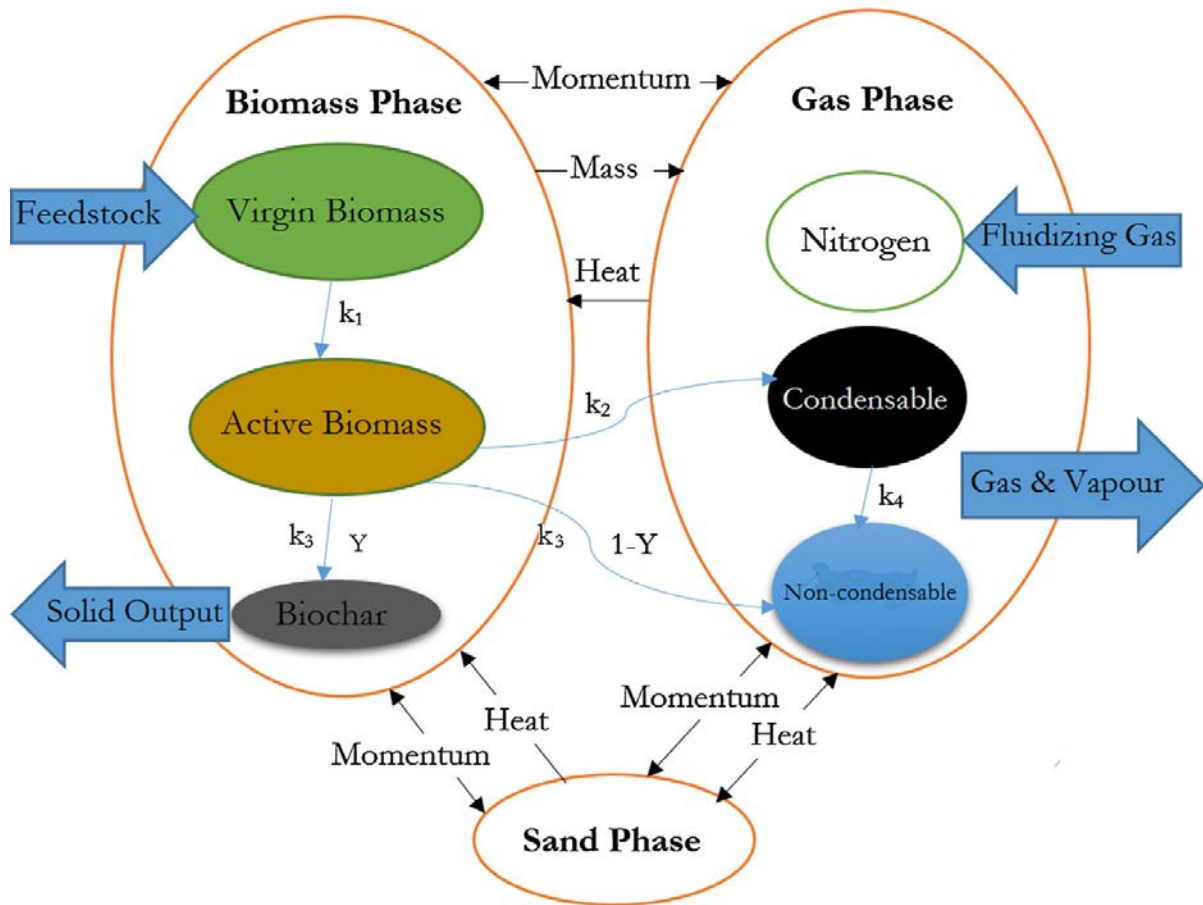
**Fig.1.** Developed methodology for obtaining an optimised solution of the pyrolysis reactor

## 2.1. Computational Fluid Dynamics

A numerical analysis of a pyrolysis process includes modelling of reacting multi-phase flow dynamics that solves the fundamental governing equations including the continuity, momentum, energy, and species mass fraction. The eulerian-granular approach is adopted for modelling the multi-phase flow dynamics. A coupled Multi-Fluid Model (MFM) and a chemical solver are employed to describe fast pyrolysis of biomass. When the thermal decomposition process which a single biomass particle undertake in the absence of oxygen is known, the chemical reactions are considered in the phase interactions and coupled with the MFM model. The MFM and chemical kinetics are briefly discussed in the following sections.

### 2.1.1. Multi-fluid model

The MFM treats all phases as inter-penetrating continua. In the present study, the MFM consists of the gaseous phase as the primary phase and two solid phases as the secondary phases (see Fig. 2). The conservation equations (mass, momentum, energy, and species) for each separate phase are solved using phase volume fraction. The governing equations also consider the stress-strain tensor in the momentum equation, the conductive heat flux in energy equations, and the diffusive flux in species transport equations. In addition to the aforementioned equations, models for phase interaction terms are required. This includes the empirical correlations for drag force and heat transfer as well as the Kinetic Theory of Granular Flow (KTGF) [42], which provides granular temperature, pressure, and viscosity. The reader is referred to the literature for further details [22, 43-45].



**Fig. 2.** Chemical reactions and exchange of mass, momentum, and heat between phases.



### 2.1.2. Chemical kinetics

Development of the detailed decomposition scheme of the biomass pyrolysis process is very challenging. It is common to adopt methods such as a lumped global kinetics which is able to address the devolatilisation and the secondary cracking due to involving many unknown elementary reactions [46], and formation of compounds for a specific feedstock [47]. Shafizadeh and Chin [48] first proposed a single-component single-step reaction scheme for thermal deterioration of wood, yet their model was unable to predict the secondary cracking [49]. Later it was proposed that biomass is first devolatilised to reach the activated state [50]. Then they proposed single-component multistep reaction kinetics to address this issue [48]. Later, multicomponent single-step reaction kinetics was proposed to account for the effects of biomass components such as cellulose, hemicellulose and lignin [51]. Ultimately, multicomponent multistep reaction kinetics, which is believed to be the most practicable method, was introduced by researchers such as Ward and Braslaw [52], Koufopoulos et al. [53, 54], Orfao et al. [55], and Miller and Bellan. [56]. More sophisticated schemes are also available in the literature that considers more reactions in the mechanism [57, 58].

In this study, lignocellulosic biomass is considered for the feedstock. Thus, a superimposed reactions method, based on the multicomponent multistep reaction kinetics, is a suitable approach for the fast pyrolysis of the biomass. The lignocellulosic biomass can be expressed by Equation (1).

$$\text{Biomass} = \alpha \text{ Cellulose} + \beta \text{ Hemicellulose} + \gamma \text{ Lignin}, \quad (1)$$

where  $(\alpha, \beta, \gamma)$  represents the initial mass composition of biomass. As shown in Fig. 2, in the first reaction step, virgin biomass is converted into the active biomass and then reacts and produces condensable vapours, non-condensable gases and biochar. At higher temperatures, condensable may react and form non-condensable as a result of the secondary reaction. Eleven different species are involved in the process. The gas phase consists of condensable vapours, non-condensable gases, and nitrogen (three species). The biomass phase is made up of virgin cellulose, virgin hemicellulose, virgin lignin, active cellulose, active hemicellulose, active lignin, and biochar (seven species) while the sand phase includes one species. The sand phase and nitrogen are inert and hence they do not participate in the chemical reactions. However, the amount of nitrogen will affect the partial pressure of the gas phase and the reaction rate constants. The reaction rates are computed using Equation (2).

$$k_i = A_i \exp[-E_{a_i}/(RT)], \quad (2)$$

where  $k_i$  is the rate constant for reaction “ $i$ ”, and  $A_i$  and  $E_{a_i}$  are the associated Arrhenius constant and activation energies, respectively, “ $T$ ” is the temperature and “ $R$ ” is the gas constant. Table 1 illustrates the values of the kinetic parameters of the reaction scheme. The formation ratio for the char component,  $Y$ , is 0.35, 0.6, and 0.75 for cellulose, hemicellulose, and lignin, respectively [45]. Table 2 provides the thermo-physical properties of the species. The density of gaseous species and the viscosity of the solid species are computed using the incompressible ideal gas model and the granular models, respectively.

**Table 1.** Pre-exponential factors and activation energies for the biomass component [45].

Components	Reaction	A(s <sup>-1</sup> )	E(MJ/kmol)	Heat release, Δh (MJ/kmole)
Cellulose	k <sub>1C</sub>	2.8×10 <sup>19</sup>	242.4	0
	k <sub>2C</sub>	3.28×10 <sup>14</sup>	196.5	41.35
	k <sub>3C</sub>	1.3×10 <sup>10</sup>	150.5	-3.24
Hemicellulose	k <sub>1H</sub>	2.1×10 <sup>16</sup>	186.7	0
	k <sub>2H</sub>	8.75×10 <sup>15</sup>	202.4	33.69
	k <sub>3H</sub>	2.6×10 <sup>11</sup>	145.7	-2.64
Lignin	k <sub>1L</sub>	9.6×10 <sup>8</sup>	107.6	0
	k <sub>2L</sub>	1.5×10 <sup>9</sup>	143.8	53.09
	k <sub>3L</sub>	7.7×10 <sup>6</sup>	111.4	-4.16
Tar	k <sub>4</sub>	4.25×10 <sup>6</sup>	108.0	-4.2

**Table 2.** Thermo-physical properties of species [59].

Species	Density $\rho$ (kg/m <sup>3</sup> )	Particle diameter $d_s$ (m)	Molecular weight (g/mol)	Heat capacity $C_p$ (J/kg K)	Dynamic viscosity $\mu$ (kg/ms)	Thermal conductivity $k$ (J/kg K)
Condensable	-	-	100	2500	3×10 <sup>-5</sup>	2.577×10 <sup>-2</sup>
Non-condensable	-	-	30	1100	3×10 <sup>-5</sup>	2.577×10 <sup>-2</sup>
N <sub>2</sub>	-	-	28	1121	3.58×10 <sup>-5</sup>	5.63×10 <sup>-2</sup>
Biomass	400	4×10 <sup>-4</sup>	*	2300	-	0.3
Biochar	2333	4×10 <sup>-4</sup>	12.01	1100	-	0.1
Sand	2649	5.2×10 <sup>-4</sup>	60.08	800	-	0.27

\*Molecular weight of the biomass components are 162.14, 132.11, and 208.21 (g/mol) for cellulose, hemicellulose, and lignin, respectively.

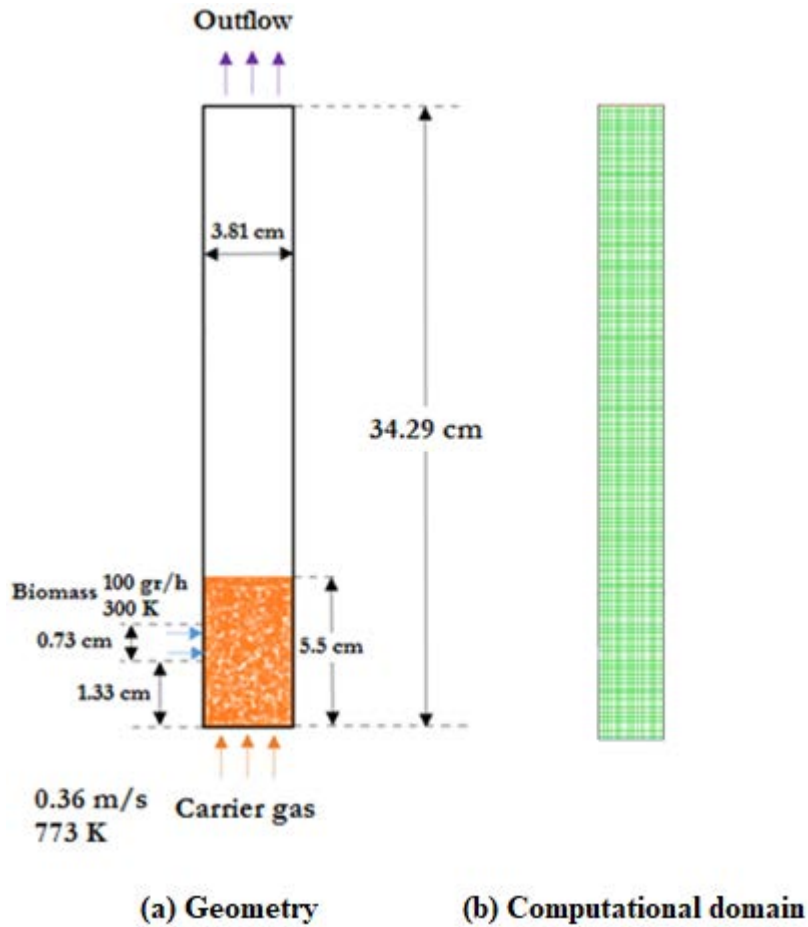
### 2.1.3. The CFD validation

The experiment of Xue et al. [39] focusing on a lab-scale bubbling fluidised bed reactor is selected for case study validation. Any pre-processing activities, such as the preparation of the feedstock, biomass grinding and dehydration, and post-processing activities including biochar removal, bio-oil catches and upgrading, are beyond the scope of the current work as it mainly focuses on the modelling of the reactor itself. Fig. 3 shows the 2-D geometry and the computational domain. The selected initial and boundary conditions are illustrated in Table 3.

The CFD simulations are conducted using ANSYS FLUENT V18.00 and run on a high-performance computing (HPC) system supported by Tasmanian Partnership for Advanced Computing. The computations are carried out using Xeon (R) E5-2620 processor with 14 CPUs. After the steady state condition is achieved, the data is collected over 20 seconds of physical time which is about 336 CPU-hours for each second to compute the average yield products. The computational time to achieve the steady state condition varies between 2352 CPU-hours to 4707 CPU-hours, depending on the values of influential parameters. It was found that the initial transient period is longer for cases with higher temperatures.



The maximum Reynolds number in a studied range of parameters (carrier gas velocity and temperature) is  $Re=595.7$ . In addition, Gao et.al [60] verified that not only both laminar and turbulent models are appropriate for predicting the hydrodynamics of gas-solid flow behaviour in fluidized bed reactor, but also laminar model presents a better hydrodynamics prediction than the turbulent model. Therefore, the multi-phase laminar flow is modelled using the Eulerian-Granular approach. Both energy and species transport equations are also solved to compute chemical reactions and heat transfer between phases. The solution includes two fractional steps. Firstly, the multiphase species are solved only spatially by setting the reaction terms to zero. Secondly, the reactions terms in each cell are integrated by applying a stiff Ordinary Differential Equation (ODE) solver. A second-order implicit method is adopted for the time discretisation. The least-square cell-based is used for the pressure-based solver that uses SIMPLE-algorithm for the pressure-velocity coupling. The second-order (upwind) method is adopted for the discretisation of the convection term in the momentum, energy and species equations. The QUICK algorithm is applied for the volume fraction calculations. Finally, a hybrid initialisation is applied for each phase, particularly for the sand phase by considering the initial packing limit. To avoid numerical instability, the simulations start with a time step size of  $1 \times 10^{-4}$  s. Then, depending on the mesh size, the time step is adjusted. The sum of the species mass fraction values must be equal to 1. Therefore, the mass fraction for Nth species is determined as one minus the sum of the N-1 solved species mass fractions. To minimise the numerical error, it is more efficient if the N<sup>th</sup> species is selected as the species with the overall largest mass fraction.

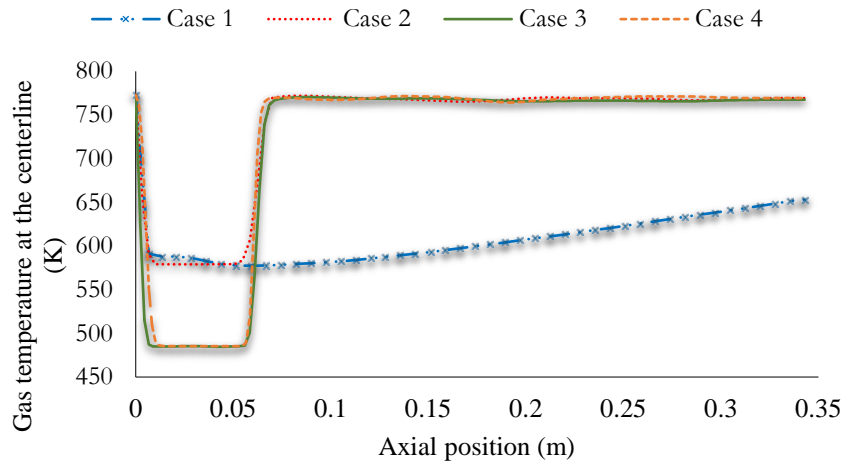


**Fig 3.** (a) Schematic geometry and (b) computational domain of the 2-D model of bubbling fluidised bed reactor

**Table 3.** Boundary and initial conditions

<b>Biomass Inlet</b>	
Temperature:	300 K
Diameter:	0.4 mm
Feed rate:	100 g/h
<b>Nitrogen Inlet</b>	
Velocity:	0.36 m/s
Temperature:	773 K
<b>Outlet</b>	
Outflow boundary condition	
<b>Wall</b>	
No-slip velocity condition for solid walls	
The constant temperature of 800 K up to a height of 8 cm.	
<b>Initial conditions</b>	
Sand diameter:	0.52 mm
Initial packing limit height:	5.5 cm
Sand Porosity:	0.41
Bed temperature:	773 K
Biomass initial mass fraction:	$(\alpha, \beta, \gamma) = (0.41, 0.32, 0.27)$

Four different mesh resolutions are considered for the grid dependency study. The maximum sand particle size is 1 mm and the mesh size is 7.62, 3.81, 2.54, and 1.91 mm, for cases 1 to 4, respectively. The size of the cells, even the smallest ones are still larger than the size of the particles. For Cases 1 to 4, the number of cells is 225, 910, 2055, 3640, respectively. Fig. 4 shows that along the centreline of the reactor at statistically steady state condition. It is worth noting that the feedstock that has been used for the mesh dependence study is pure cellulose. It can be seen that the temperature initially changes as the mesh resolution increases while further mesh refinement marginally changes the results, confirming Case 3 with 2055 cells provide an acceptable level of accuracy. Temperature distribution along the centreline of the reactor at statistically steady state condition. It is worth noting that the feedstock ben used for the mesh dependence study is pure cellulose. It can be seen that the temperature initially changes as the mesh resolution increases while further mesh refinement marginally changes the results, confirming Case 3 with 2055 cells provide an acceptable level of accuracy.



**Fig. 4.** The gas temperature along the axial direction for different grid sizes. case 1: 225 cells, case 2: 910 cells, case 3: 2055 cells, and case 4: 3640 cells.

The product yields are obtained by the time integration at the reactor outlet after the statistically steady state condition is achieved. For instance, the condensable yield is computed by using Equation (9).

$$\eta_{condensable} = \frac{\int_{t_{ss}}^{t_{ss}+\Delta t} \int_{outlet} (\varepsilon_g \rho_g U_g Y_c) dA dt}{\int_{t_{ss}}^{t_{ss}+\Delta t} \int_{outlet} [\varepsilon_g \rho_g U_g (Y_c + Y_{nc}) + \varepsilon_b \rho_b U_b (Y_{bc} + Y_{ub})] dA dt} \quad (9)$$

where  $\varepsilon$ ,  $\rho$ ,  $U$  are the volume fraction, the density, and the velocity of each phase, respectively.  $t_{ss}$  represents the time when statistically steady state condition is achieved and  $\Delta t=20$  s is taken here.  $Y$  refers to the species mass fraction and the subscript  $g$ ,  $c$ ,  $nc$ ,  $bc$ ,  $ub$  denote gas phase, condensable, non-condensable, biochar, and unreacted biomass, respectively. The computed results for product yields are compared to the measurements [44] in Table 4. The predicted results demonstrate good agreement with the experimental data for product yield. The percentage of the discrepancies between the experiment and the simulation results for condensable, non-condensable, biochar, and operating temperature are 13, 17.6, 13.8, and 1.4, respectively.

**Table 4.** Comparison of product yield for red oak pyrolysis (wt %) between simulation and experiment

Components	Condensable	Non-condensable	Biochar	Unreacted biomass
Experiment [44]	71.7±1.4	20.5±1.3	13±1.5	-
Current study	62.4	16.9	11.2	9.5
Percentage of difference	13.0	17.6	13.8	-

## 2.2. Machine Learning Algorithms

The present study aims to apply machine learning techniques to determine the relationship between different parameters affecting the product yields in the pyrolysis process. The SVR algorithm is adopted here due to its ability with complex and non-linear systems. By training the SVR with the data obtained from the CFD simulations, the model is developed which has a capacity to predict the product yields when a new set of data are available. The concept of Genetic Algorithms (GAs) can then be employed for the optimisation. GAs are adaptive heuristic search algorithms based on the idea of natural selection and genetics. Among GAs, the PSO algorithm contains a more efficient search ability to solve multidimensional optimisation [61], enabling us to use an intelligent method for estimation of the desired values. In this section, the SVR and the PSO are first described and then the methodology developed for integrating them is presented.

### 2.2.1. SVR model

Various regressions including Linear Regression (LR), Least Absolute Shrinkage and Selection Operator (LASSO), Least Angle Regression (LARS) and SVR can be used to analyse the relationship between a dependent variable and a set of independent variables [36, 37]. While LR, LASSO and LARS are linear methods and suitable for linear data, SVR performs more effectively for non-linear datasets and is applied in this paper since the CFD dataset showed a non-linear complex trend. Particle Swarm Optimisation (PSO) algorithm is also employed to find an optimum value after data modelling by SVR.

Support Vector Machines (SVMs) are used for non-linear modelling by transforming the original space into a higher dimensional feature space. SVMs have been developed based on statistical learning theory and have been designed as a classification technique by Vapnik and colleagues [47, 62]. SVR later applies the margin concept to solve regression problems with the help of e-insensitive loss functions.

The SVR model considers  $x_1 \dots x_n \in \mathbb{R}$  as a set of training input, and  $y \in \mathbb{R}$  as a training output. The linear regression can be presented in the high dimensional space using the inner product function of the mapping inputs into a high-dimensional plane. The inner product of two vectors  $x$  and  $\omega$  is defined by Equation (3).

$$f(x) = \langle \omega, x \rangle + b, \quad (3)$$

where  $\omega$  is a weight vector and  $b$  is a scalar bias. The objective of the SVR is to minimise the regression risk by considering equations (4) and (5).

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (4)$$

$$s.t. \begin{cases} y_i - (\langle \omega, x \rangle + b) \leq \varepsilon + \xi_i \\ \langle \omega, x \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (5)$$

where  $C > 0$  is the penalty or regularisation term that controls the trade-off between the model complexity and data training error,  $l$  is the number of samples, and  $\xi_i$  and  $\xi_i^*$  indicate the upper and lower training errors, respectively. Equation (4) is solved by a dual optimisation that uses the Lagrange function expressed as equation (6).

$$L = \frac{1}{2} \omega^T \omega + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l \alpha_i (\xi + \xi_i - y_i + \langle \omega, x \rangle + b) - \sum_{i=1}^l \alpha_i^* (\xi + \xi_i^* + y_i - \langle \omega, x \rangle - b) - \sum_{i=1}^l (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (6)$$

where  $\eta_i, \xi_i, \eta_i^*, \xi_i^*$  indicate the Lagrange multipliers which are non-negative real numbers.

### 2.2.2. PSO algorithm

The PSO algorithm is a population-based search optimisation technique developed by Kennedy and Eberhart [63]. There are many similarities between the PSO method and other evolutionary computation techniques such as Genetic Algorithms (GA). However, the PSO model updates each candidate solution using a velocity rather than the evolutionary operators like selection, crossover and mutation. Compared to the GA algorithm, the PSO method is relatively easier to implement and requires fewer number of parameters to be adjusted. It has also demonstrated good convergence and better performance.

The PSO algorithm is inspired by social behaviour observed in nature such as bird flocking and fish schooling. Initially, PSO, like other GAs, starts with a population of random particles (solutions). It then organises a swarm moving in the search space looking for the best solution. In an iteration, each particle is updated by following two “best” values known as *pbest* and *gbest*. *pbest* is the best solution (fitness) which is achieved so far while *gbest* is a global best obtained so far by any particle in the population. After finding these two best values, velocity and positions of the particles are updated according to equations (7) and (8).

$$V[i] = V[i] + c_1 * rand() * (pbest[i] - present[i]) + c_2 * rand() * (gbest[i] - present[i]) \quad (7)$$

$$Present[i] = present[i] + V[i] \quad (8)$$

where  $V[i]$  is the particle velocity,  $present[i]$  is the current particle (solution),  $rand()$  is a random number between (0,1), and  $c_1$  and  $c_2$  are the learning factors.

### 2.2.3. SVR-PSO combined model

The SVR-PSO strategy is employed to obtain optimal values of the considered parameters to maximise the condensable yield. The SVR-PSO algorithm includes the following steps:

1. The SVR is initially trained by the CFD datasets to predict the condensable yield.
2. The initial particle swarm is created and the fitness function is evaluated. Lower and upper bands are adopted for each parameter to reduce the search space. Eventually, the next generation is produced until the stopping criterion is satisfied.
3. Finally, the SVR-PSO gives the condensable for the optimal values of the dataset.

## 3. Results and discussions

This section presents the results obtained using both CFD and MLAs. The validated numerical simulations are followed by the parametric study. The optimised operating parameters and the relevant yields obtained using the SVR-PSO model are then presented. Finally, the optimised set of parameters is integrated into the CFD simulation, and the results are compared with the MLA outputs.

### 3.1. Parametric study

Seven different influential parameters are considered here; operating temperature, biomass flow rate, biomass and sand particle sizes, carrier gas velocity, biomass injector location and pre-treatment temperature. For the parametric study, a single influential parameter varies in the given range while other parameters are set as the base condition. Table 5 shows the variation range considered for each parameter whereas the base values are the same as those in the experiment [39] (see Table 3).

The Thermogravimetric analysis (TGA) by Yang et al. [64] reveals that the three components of lignocellulosic biomass show different behaviours during the pyrolysis process. Hemicellulose decomposes easily and most of the weight loss occurs in the temperature range of 220–315°C. The maximum loss rate (0.95 wt.%/°C) occurs at 268°C, and even at 900°C, about 20% of solid residue is left. Cellulose pyrolysis arises at a higher temperature range of 315–400°C with the maximum weight loss rate (2.84 wt.%/°C) reached at 355°C. At higher temperature (>400°C), almost all cellulose content is decomposed thermally whereas only a very small percentage of solid residues is left (~6.5 wt.%/°C). Among the three components, lignin has the highest resistance to degradation. Although it pyrolyses gradually under the whole temperature range from ambient to 900°C, the mass-loss rate is as low as (<0.14 wt.%/°C) and the solid residue left is the highest (~45.7 wt.%). The TGA results, therefore, justify why the temperature range of 400–675 °C is chosen for parametric study. In addition, as long as the biomass preheating temperature is under 220 °C, no reaction will occur outside the reactor bed that we don't have any control over it. Hence, for example, the effect of preheating in the temperature range of 300–400 K is investigated. The specified range for sand and biomass particle sizes are selected based on the experimental test of the lab-scale bubbling fluidised bed reactor [22, 44]. The adopted carrier gas velocity primarily depends on particle sizes. Larger sand particle sizes

are more resistant to fluidisation [65]. Hence, a minimum fluidisation velocity (MFV) is defined as a function of particle size to ensure the fluidisation occurs. On the other hand, higher carrier gas velocities shorten the residence time of condensable vapour. This reduces the possibility of the secondary crack of condensable vapours to non-condensable gases and consequently increases the bio-oil yield. However, there are upper limits for the carrier gas velocities. If the nitrogen velocity is exceeded by a certain limit, which is called maximum effective velocity (MEV), the sand particles will be thrown out of the reactor bed by the carrier gas. It has a negative effect on the mixing of sand and biomass particles. It has been demonstrated that for enhanced heat transfer it is much more efficient when the cold virgin biomass particles are surrounded by hot sand particles of larger size [65]. High carrier gas velocities may also push the small and low-density biomass particles out of the reactor bed. Hence, it is important to effectively select the carrier gas velocity, otherwise adequate time will not be provided for mixing the sand and biomass particles, the biomass particles will leave the reactor exit unreacted, and the most importantly the sand particles that play an important role as a heat carrier will be lost. The higher the position of the biomass injector, the lower the residence time of condensable vapours, resulting in fewer secondary reactions. However, the upper limit of the biomass injector location is restricted by the initial packing limit of the sand particles. The initial packing limit is 55 mm, and the biomass injector diameter is 7.3 mm, therefore the maximum allowable biomass injector height is defined as  $55-7.3=47.7$ . That is why 47 mm is selected for the upper limit of the biomass injector height. Study of biomass feed rate is also required to find its optimum value for reducing heat energy consumption. Certainly, it is expected to have fewer product yields when the biomass feed rate increases. However, the objective is to find out how much feed can be added to the biomass without the need to consume excessive heat to have maximum product yields.

**Table 5.** Range of operating parameters' variation for parametric study

Operating Conditions	Range of variation
Operating Temperature (°C)	400-675
Sand Particle Size (mm)	0.4-1
Nitrogen Velocity (m/s)	0.3-1
Biomass Particle Size (mm)	0.2-1
Biomass feed rate (kg/hr)	0.1-1.3
Pre-treatment Temperature (K)	300-400
Biomass Injector Height (mm)	12-47

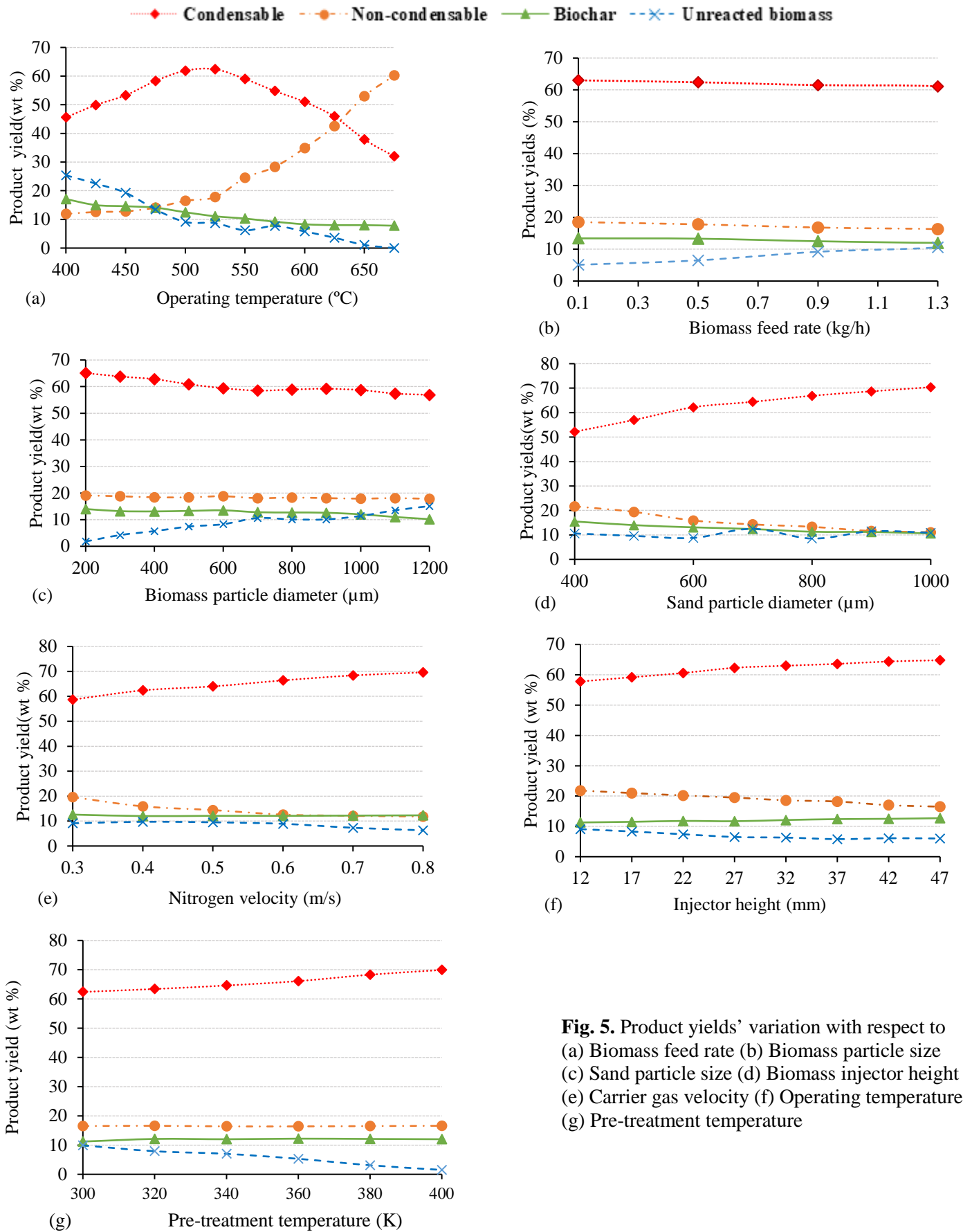
A summary of the CFD parametric study is shown in Fig 5, presenting the product yields versus the influential parameters. Fig 5(a) shows the effects of the operating temperature. The two main factors affecting the amount of product yields due to temperature rise are the devolatilisation and the secondary crack. Increasing the temperature accelerates the devolatilisation of biomass as well as the primary reaction rates. However, at temperatures above 525 °C, the secondary crack occurs which converts condensable vapours to non-condensable gases. Consequently, by increasing the operating temperature, non-condensable yield uniformly rises whereas unreacted biomass, condensable and biochar yield decrease. The influence of biomass flow rate on product yields is illustrated in Fig. 5(b). At higher biomass feed rates, the supplied heat is insufficient for biomass particles to react. As a result, although the changes are minor the product yields, particularly the condensable yield reduce as the biomass feed rate increases.



Effects of the biomass and sand particle sizes are depicted in Fig 5(c) and 5(d), respectively. It can be seen that the condensable yield slightly decreases as the biomass particle size increases. The use of coarse biomass particles is acceptable if the provided heat is sufficient for biomass particles and they are surrounded by sand particles of the same size or larger. Despite the increased costs for grinding, it is more efficient to use smaller biomass particles due to lower intraparticle temperature gradients. On the other hand, the use of very fine biomass particles may cause the carrier gas to throw the unreacted biomass out of the reactor bed. The feeder size limits the biomass particle size since coarse particles may clog the feeder. Larger sand particles necessitate higher carrier gas velocity to achieve an effective fluidisation. Thus, the residence time reduces and it consequently minimises the secondary crack of condensable vapours to non-condensable gases. The results confirm that the efficiency increases if biomass particles are surrounded by larger sand particles since it enhances the heat transfer supported by hot sand particles. Consequently, various parameters are important to determine the optimum size for sand and biomass particles. As shown in Fig. 5(e), the higher the carrier gas velocity, the higher the condensable yield, and the lower the non-condensable yield due to shorter residence time and less chance of secondary crack. The minimum fluidisation velocity, which is a function of the sand particle size, determines the lower limits for nitrogen velocity. On the other hand, the carrier gas velocity can be increased if the unreacted biomass and sand particles are not thrown out of the reactor. Therefore, the optimal nitrogen velocity is primarily a function of the particles size. Fig. 5(f) shows the effects of the biomass injector location. As the elevation of the injector increases, the residence time becomes shorter, and hence there is a lesser chance of the secondary crack of condensable vapours to non-condensable gases. Consequently, the yield of condensable vapours is higher. It is worth noting that the optimal height for the biomass injector also depends on how high the sand particles are initially packed.

Finally, the effects of biomass preheat are illustrated in Fig. 5(g). Preheating the biomass accelerates the time taken to reach to the active state. By providing the same amount of heat, the primary reaction rates are larger due to the shorter time interval. Therefore, a larger proportion of biomass is converted to the products whereas biochar yield remains stable. Consequently, more product yields are expected in the gaseous phase including condensable and non-condensable components. In the meantime, higher condensable but lower non-condensable yields are expected as the time taken to reach the active state is faster, and hence the secondary crack is lower at the shorter residence time, becoming less dominant at temperatures below 525 °C. Thus, considering all these phenomena, preheating has a favourable effect on condensable yields whereas non-condensable yields remain unaffected. It may be concluded that preheating maximises the bio-oil yield as long as the virgin biomass is preheated below the threshold temperature for initiating the reaction.

As mentioned, the described parametric study only analysed one parameter at the time while others were kept constant - same as the experimental values. The parametric study is extended by variations of different parameters, simultaneously. For this purpose, 23 other CFD simulations are conducted where the influential parameters are randomly selected in the given ranges. The new simulated data are added to the CFD dataset which will be used as an input for the MLA modelling. However, unrelated data points, are first removed from the CFD dataset. As discussed in Fig. 5, the temperature ranges can be filtered and only the temperatures from 450 to 550 °C are considered. Similarly, it is not necessary to consider the values for biomass flow rate higher than 1 kg/h. A summary of the final CFD dataset is presented in Table 6 shows the range of influential parameters as well as product yields which will be the input for the MLA modelling.



**Fig. 5.** Product yields' variation with respect to (a) Biomass feed rate (b) Biomass particle size (c) Sand particle size (d) Biomass injector height (e) Carrier gas velocity (f) Operating temperature (g) Pre-treatment temperature

**Table 6.** Dataset for the MLA modelling

Variables	Range of variation
Operating Temperature (°C)	450-550
Sand Particle Size (mm)	0.4-1
Nitrogen Velocity (m/s)	0.3-1
Biomass Particle Size (mm)	0.2-1
Biomass feed rate (kg/hr)	0.1-1
Pre-treatment Temperature (K)	300-400
Biomass Injector Height (mm)	12-47
Condensable yield	53.3-71
Non-condensable yield	10.7-24.5
Biochar yield	10.3-17.1
Unreacted biomass	2.3-19

### 3.2. Machine Learning Algorithms

To obtain the optimised condition, the MLA analysis is conducted using MATLAB 2015b Linux Mint operating system with 32G RAM, Intel (R) Core (TM) i7-6700K CPU and NVIDIA GeForce GTX 1070 graphics card. The datasets are provided by the CFD results. 70% of the datasets are allocated for the training purpose while 30% of the datasets are used for the prediction purpose. The training dataset builds the SVR model and it requires the allocation of the trade-off for different parameters. In this study, three common kernel functions, Linear, Polynomial and Gaussian (radial basis function) are considered (see Table 7). The predicted data are evaluated with the statistical parameters consisting of the mean square error (MSE), the root mean square error (RMSE), and the correlation coefficient ( $R$ ) as defined

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (10)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (11)$$

$$MSE = \left( \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right) / \left( \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \right), \quad (12)$$

where  $n$  is the number of data points.  $x_i$ ,  $y_i$ , are the observed and the predicted values while  $\bar{x}$ ,  $\bar{y}$  are the mean values for the observed and the predicted data, respectively. As can be seen in Table 8, the proposed model has a diverse performance when using different kernels. The results demonstrate that the Gaussian kernel has a better performance compared to other kernels by maximising  $R$  and minimising  $MSE$  and  $RMSE$  for the test data.

**Table 7.** Kernel functions

Kernel	Function	Kernel parameter
Linear	$K(x_i, x_j) = (x_i, x_j)$	-
Polynomial	$K(x_i, x_j) = ((x_i, x_j) + 1)^d$	$d$
Gaussian (RBF)	$K(x_i, x_j) = \exp\left(-\frac{\ x_i, x_j\ ^2}{2\sigma^2}\right)$	$\sigma$

**Table 8.** Effect of different kernels on the SVR-PSO performance. Errors calculated on condensable parameter

Kernel	MSE		R		RMSE	
	Train	Test	Train	Test	Train	Test
Linear	17.1622	345.2164	0.0826	0.0594	4.1427	18.58
Polynomial	9.2192	930.86	0.0634	0.0224	3.0363	30.51
Gaussian (RBF)	2.0195	4.321	0.9771	0.9213	1.4211	2.21

After determining the type of Gaussian function, different aspects of designer-determined parameters were analysed for the SVR model. The first parameter is “C” that shows the trade-off between the training error and the flatness of the solution. The larger values of “C” result in lower final training error. However, too much increment of “C” leads to the risk of losing the generalisation properties of the classifier, since it tries to fit in the best possible way for all the training points. If “C” is small, then the classifier is flat. A value for “C” should be found that keeps the training error small but also generalises well. The second parameter is “ε” that controls the width of the ε-insensitive zone, used to fit the training data. The value of “ε” can affect the number of support vectors used to construct the regression function. Increasing the “ε” values, reduces the number of support vectors. A large enough value of “ε” will lead to a constant regression function. And finally, the value of “σ” that controls the width of the Gaussian function. If “σ” is very small, the radius area of influence of the support vectors only includes the support vector itself and no amount of regularisation with “C” will be able to prevent overfitting. When “σ” is too large, the model is too constrained and cannot capture the complexity of the data. Table 9 shows two optimum parameters that were obtained from the model. It can be seen that these values of parameters show the best correlation coefficient and minimum values of error. The obtained optimised values for operating parameters by adjusting the SVR model with C=5.26, ε= 0.52, σ=1.99 are illustrated in Table 10. Based on the obtained set of parameters, the predicted values by MLAs and the ones calculated by CFD for the product yields are presented in Table 11. It is worth noting that in this study our favourable product is condensable, and the discrepancy for condensable yields between the predicted results by MLAs and CFD is 3.7 %. The unreacted biomass is computed as the leftover percentage of the three by-products.

**Table 9.** Design parameters of SVR model.

	MSE		R		RMSE	
	Train	Test	Train	Test	Train	Test
$C=5.26, \epsilon=0.52, \sigma=1.99$	1.9511	3.11	0.9971	0.9444	1.8731	1.3
$C=5.11, \epsilon=0.51, \sigma=175.4$	2.0195	4.321	0.9771	0.9213	1.4211	2.21

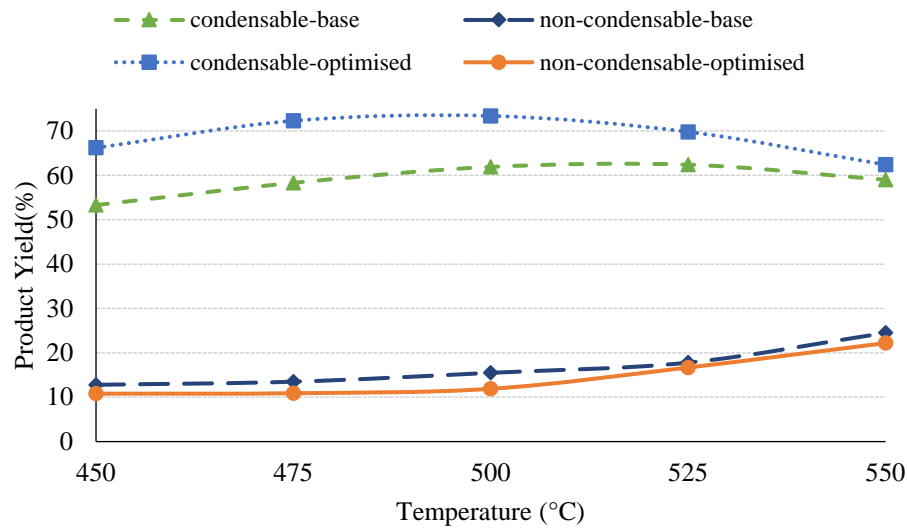
**Table 10.** Base and optimised parameters

Operating parameter	Base	Optimised
Operating Temperature (°C)	500	498
Sand Particle Size (mm)	0.52	0.832
Nitrogen Velocity (m/s)	0.36	0.76
Biomass Particle Size (mm)	0.4	0.463
Biomass feed rate (kg/hr)	0.1	0.15
Pretreatment Temperature (K)	300	400
Biomass Injector Height (mm)	17	47

Fig. 6 shows the CFD results for condensable and non-condensable yields for two different datasets; the base parameters that are used in the experiment [39] and the optimised values obtained by MLAs. These values are presented in Table 11. As illustrated in Fig. 6, using optimised values of the parameters for CFD simulation lead to more condensable and lower non-condensable yield in the temperature range of 450-550 °C. This means that, in addition to the temperature, which is the most influential parameter in the products yields, impact of other parameters can also be of great value in minimising the secondary crack of condensable vapours to non-condensable gases and finally optimising the process. The most important factor to maximise bio-oil production is to keep the operating temperature at about 500 °C. For this purpose, heat transfer from carrier gas to sand particles was facilitated. The heat must then be transferred from hot sand particles to the cold virgin biomass particles. Things that have been done to improve the heat transfer are using larger hot sand particles than cold virgin biomass particles to increase heat transfer area, preheating the biomass and effective mixing of sand and biomass particles. Nitrogen has the role of carrier gas in addition to mixing the solid particles. Increasing the carrier gas velocity up to a certain extent which is called maximum effective velocity (MEV) reduces the condensable vapour residence time and secondary reactions. Exceeding the MEV has a negative effect on the product yields since it causes the sand particles and unreacted biomass to be thrown out of the reactor bed. Another option that has been taken to reduce the residence time is moving the biomass injector to a higher level.

**Table 11.** Comparison of the results between MLAs and CFD

Product yield (%)	MLAs	CFD	Discrepancy (%)
Condensable yield	70.7	73.4	3.7
Non-condensable	11.9	11.9	0
Biochar yield	12.2	12.4	1.6
Unreacted biomass	5.2	2.3	55.8

**Fig. 6.** Condensable and non-condensable yields for the base and optimised parameters

#### 4. Conclusion

- The CFD results are compared and validated with experimental data where a good agreement is witnessed.
- A parametric study is performed to address the effect of seven operating factors on the product yields. The parameters that are considered are operating temperature, biomass flow rate, biomass and sand particle size, carrier gas velocity, biomass injector location and pre-treatment temperature.
- Created dataset by parametric CFD results is then utilised as an input variable in a developed SVR-PSO algorithm to predict and optimise the bio-oil yield considering the operating conditions. The optimum set of operating parameters received by the developed algorithm are used as an input in CFD simulation validating the obtained results received from MLAs.
- The optimum operating conditions achieved by MLAs are for the simulated case study: operating temperature = 498 °C, biomass pre-treatment temperature = 400 K, nitrogen velocity



= 0.76 m/s, sand particle size = 832  $\mu\text{m}$ , biomass particle size = 463  $\mu\text{m}$ , biomass feed rate = 0.15 kg/h, and biomass injector height = 47 mm. Based on the optimum set of parameters, the predicted values for condensable yield by MLAs and those calculated by CFD are 70.7 and 73.4 respectively with only 3.7% discrepancy.

- The CFD results for condensable and non-condensable yields showed that application of optimised parameters raises the condensable yield and reduces the non-condensable yield in other temperature ranges.

## 5. References

1. Blin, J., et al., *Biodegradability of biomass pyrolysis oils: Comparison to conventional petroleum fuels and alternatives fuels in current use*. Fuel, 2007. **86**(17-18): p. 2679-2686.
2. Panwar, N., R. Kothari, and V. Tyagi, *Thermo chemical conversion of biomass–Eco friendly energy routes*. Renewable and Sustainable Energy Reviews, 2012. **16**(4): p. 1801-1816.
3. Tsai, W., M. Lee, and Y. Chang, *Fast pyrolysis of rice straw, sugarcane bagasse and coconut shell in an induction-heating reactor*. Journal of analytical and applied pyrolysis, 2006. **76**(1-2): p. 230-237.
4. Lajili, M., et al., *Fast pyrolysis and steam gasification of pellets prepared from olive oil mill residues*. 2018. **150**: p. 61-68.
5. Cai, W., L. Dai, and R.J.E. Liu, *Catalytic fast pyrolysis of rice husk for bio-oil production*. 2018. **154**: p. 477-487.
6. Yang, S., M. Wu, and C.J.E. Wu, *Application of biomass fast pyrolysis part I: Pyrolysis characteristics and products*. 2014. **66**: p. 162-171.
7. Kim, S.-S., et al., *Pyrolysis kinetics and decomposition characteristics of pine trees*. Bioresource technology, 2010. **101**(24): p. 9797-9802.
8. Panwar, N. and N. Rathore, *Potential of surplus biomass gasifier based power generation: A case study of an Indian state Rajasthan*. Mitigation and adaptation strategies for global change, 2009. **14**(8): p. 711.
9. Goyal, H., D. Seal, and R. Saxena, *Bio-fuels from thermochemical conversion of renewable resources: a review*. Renewable and sustainable energy reviews, 2008. **12**(2): p. 504-517.
10. Balat, M., et al., *Main routes for the thermo-conversion of biomass into fuels and chemicals. Part I: Pyrolysis systems*. Energy Conversion and Management, 2009. **50**(12): p. 3147-3157.
11. Chen, J., D. Fang, and F.J.A.e. Duan, *Pore characteristics and fractal properties of biochar obtained from the pyrolysis of coarse wood in a fluidized-bed reactor*. 2018. **218**: p. 54-65.
12. Wang, W., et al., *Formate-assisted analytical pyrolysis of kraft lignin to phenols*. 2019.
13. Zhang, H., et al., *Effect of feedstock and pyrolysis temperature on properties of biochar governing end use efficacy*. 2017. **105**: p. 136-146.
14. DEMİRBAŞ, A., *Hydrocarbons from pyrolysis and hydrolysis processes of biomass*. Energy sources, 2003. **25**(1): p. 67-75.
15. Anca-Couce, A., P. Sommersacher, and R. Scharler, *Online experiments and modelling with a detailed reaction scheme of single particle biomass pyrolysis*. Journal of analytical and applied pyrolysis, 2017. **127**: p. 411-425.
16. Rezaei, H., S. Sokhansanj, and C.J. Lim, *Minimum fluidization velocity of ground chip and ground pellet particles of woody biomass*. Chemical Engineering and Processing-Process Intensification, 2018. **124**: p. 222-234.
17. Guizani, C., et al., *Biomass fast pyrolysis in a drop tube reactor for bio oil production: Experiments and modeling*. Fuel, 2017. **207**: p. 71-84.

18. Park, J.-W., et al., *Fast pyrolysis of acid-washed oil palm empty fruit bunch for bio-oil production in a bubbling fluidized-bed reactor*. 2019.
19. Hu, C., et al., *CFD-DEM Investigation on the Biomass Fast Pyrolysis: The Influences of Shrinkage Patterns and Operating Parameters*. Industrial & Engineering Chemistry Research, 2018.
20. Liu, B., et al., *CFD modelling of particle shrinkage in a fluidized bed for biomass fast pyrolysis with quadrature method of moment*. Fuel Processing Technology, 2017. **164**: p. 51-68.
21. Cardoso, J., et al., *Improved numerical approaches to predict hydrodynamics in a pilot-scale bubbling fluidized bed biomass reactor: A numerical study with experimental validation*. Energy Conversion and Management, 2018. **156**: p. 53-67.
22. Xiong, Q., S. Aramideh, and S.-C. Kong, *Modeling effects of operating conditions on biomass fast pyrolysis in bubbling fluidized bed reactors*. Energy & Fuels, 2013. **27**(10): p. 5948-5956.
23. Mellin, P., E. Kantarelis, and W. Yang, *Computational fluid dynamics modeling of biomass fast pyrolysis in a fluidized bed reactor, using a comprehensive chemistry scheme*. Fuel, 2014. **117**: p. 704-715.
24. Xiong, Q., et al., *Coupling DAEM and CFD for simulating biomass fast pyrolysis in fluidized beds*. Journal of Analytical and Applied Pyrolysis, 2016. **117**: p. 176-181.
25. Jalalifar, S., et al., *Parametric analysis of pyrolysis process on the product yields in a bubbling fluidized bed reactor*. Fuel, 2018. **234**: p. 616-625.
26. Jalalifar, S., et al. *Numerical modelling of a fast pyrolysis process in a bubbling fluidized bed reactor*. in *IOP Conference Series: Earth and Environmental Science*. 2017. IOP Publishing.
27. Blanco, A. and F. Chejne, *Modeling and simulation of biomass fast pyrolysis in a fluidized bed reactor*. Journal of analytical and applied pyrolysis, 2016. **118**: p. 105-114.
28. Hejazi, B., et al., *Coupled reactor and particle model of biomass drying and pyrolysis in a bubbling fluidized bed reactor*. Journal of Analytical and Applied Pyrolysis, 2016. **121**: p. 213-229.
29. Niemelä, N.P., et al., *CFD based reactivity parameter determination for biomass particles of multiple size ranges in high heating rate devolatilization*. 2017. **128**: p. 676-687.
30. Marathe, P., R. Westerhof, and S.J.A.e. Kersten, *Fast pyrolysis of lignins with different molecular weight: Experiments and modelling*. 2019. **236**: p. 1125-1137.
31. Mutlu, A.Y. and O.J.E. Yucel, *An artificial intelligence based approach to predicting syngas composition for downdraft biomass gasification*. 2018. **165**: p. 895-901.
32. Saleem, M. and I. Ali, *Machine Learning Based Prediction of Pyrolytic Conversion for Red Sea Seaweed*.
33. Sunphorka, S., B. Chalermssinsuwan, and P.J.F. Piumsomboon, *Artificial neural network model for the prediction of kinetic parameters of biomass pyrolysis from its constituents*. 2017. **193**: p. 142-158.
34. Hough, B.R., et al., *Application of machine learning to pyrolysis reaction networks: Reducing model solution time to enable process optimization*. 2017. **104**: p. 56-63.
35. Sun, Y., et al., *Pyrolysis products from industrial waste biomass based on a neural network model*. 2016. **120**: p. 94-102.
36. Tibshirani, R., *Regression shrinkage and selection via the lasso*. Journal of the Royal Statistical Society. Series B (Methodological), 1996: p. 267-288.
37. Efron, B., et al., *Least angle regression*. The Annals of statistics, 2004. **32**(2): p. 407-499.
38. Kennedy, J., *Particle swarm optimization*, in *Encyclopedia of machine learning*. 2011, Springer. p. 760-766.
39. Zhang, Y. and C. Ling, *A strategy to apply machine learning to small datasets in materials science*. npj Computational Materials, 2018. **4**(1): p. 25.

40. Pal, M. and P. Mather, *Support vector machines for classification in remote sensing*. International Journal of Remote Sensing, 2005. **26**(5): p. 1007-1011.
41. Bishop, C., *Pattern Recognition and Machine Learning (Information Science and Statistics)*, chapter 3, pages 138--147. 2006, Springer-Verlag New York, Inc.
42. Gidaspow, D., *Multiphase flow and fluidization: continuum and kinetic theory descriptions*. 1994: Academic press.
43. Xiong, Q., et al., *BIOTC: an open-source CFD code for simulating biomass fast pyrolysis*. Computer Physics Communications, 2014. **185**(6): p. 1739-1746.
44. Xue, Q., et al., *Experimental validation and CFD modeling study of biomass fast pyrolysis in fluidized-bed reactors*. Fuel, 2012. **97**: p. 757-769.
45. Xue, Q., T. Heindel, and R. Fox, *A CFD model for biomass fast pyrolysis in fluidized-bed reactors*. Chemical Engineering Science, 2011. **66**(11): p. 2440-2452.
46. Demirbaş, A., *Mechanisms of liquefaction and pyrolysis reactions of biomass*. Energy conversion and management, 2000. **41**(6): p. 633-646.
47. Van de Velden, M., et al., *Fundamentals, kinetics and endothermicity of the biomass pyrolysis reaction*. Renewable energy, 2010. **35**(1): p. 232-242.
48. Shafizadeh, F. and P.P. Chin. *Thermal deterioration of wood*. in ACS Symposium Series American Chemical Society. 1977.
49. Han, J., H.J.R. Kim, and s.e. reviews, *The reduction and control technology of tar during biomass gasification/pyrolysis: an overview*. 2008. **12**(2): p. 397-416.
50. Authier, O., et al., *Solid Pyrolysis Modelling by a Lagrangian and Dimensionless Approach-Application to Cellulose Fast Pyrolysis*. 2010. **8**(1).
51. Yang, H., et al., *In-depth investigation of biomass pyrolysis based on three major components: hemicellulose, cellulose and lignin*. 2006. **20**(1): p. 388-393.
52. Ward, S., J.J.C. Braslaw, and flame, *Experimental weight loss kinetics of wood pyrolysis under vacuum*. 1985. **61**(3): p. 261-269.
53. Koufopoulos, C., A. Lucchesi, and G.J.T.C.J.o.C.E. Maschio, *Kinetic modelling of the pyrolysis of biomass and biomass components*. 1989. **67**(1): p. 75-84.
54. Koufopoulos, C., et al., *Modelling of the pyrolysis of biomass particles. Studies on kinetics, thermal and heat transfer effects*. 1991. **69**(4): p. 907-915.
55. Orfao, J., F. Antunes, and J.L.J.F. Figueiredo, *Pyrolysis kinetics of lignocellulosic materials—three independent reactions model*. 1999. **78**(3): p. 349-358.
56. Miller, R.S., J.J.C.s. Bellan, and technology, *A generalized biomass pyrolysis model based on superimposed cellulose, hemicellulose and lignin kinetics*. 1997. **126**(1-6): p. 97-137.
57. Ranzi, E., et al., *Chemical kinetics of biomass pyrolysis*. 2008. **22**(6): p. 4292-4300.
58. Matta, J., et al., *Comparison of multi-component kinetic relations on bubbling fluidized-bed woody biomass fast pyrolysis reactor model performance*. 2017. **210**: p. 625-638.
59. Lathouwers, D. and J. Bellan, *Yield optimization and scaling of fluidized beds for tar production from biomass*. Energy & Fuels, 2001. **15**(5): p. 1247-1262.
60. Gao, J., et al., *CFD modeling and validation of the turbulent fluidized bed of FCC particles*. 2009. **55**(7): p. 1680-1694.
61. Chuang, L.-Y., et al., *Correlation-based gene selection and classification using Taguchi-BPSO*. Methods of information in medicine, 2010. **49**(03): p. 254-268.
62. Vapnik, V., *The nature of statistical learning theory*. 2013: Springer science & business media.
63. Kennedy, R. J. and Eberhart, *Particle swarm optimization*. in *Proceedings of IEEE International Conference on Neural Networks IV*, pages. 1995.
64. Yang, H., et al., *Characteristics of hemicellulose, cellulose and lignin pyrolysis*. 2007. **86**(12-13): p. 1781-1788.

65. Jalalifar, S., et al., *Parametric analysis of pyrolysis process on the product yields in a bubbling fluidized bed reactor*. 2018. **234**: p. 616-625.

## Appendix

Matlab Code will be added later