Fuel 320 (2022) 123944

Contents lists available at ScienceDirect

Fuel

journal homepage: www.elsevier.com/locate/fuel

Predictability of higher heating value of biomass feedstocks via proximate and ultimate analyses – A comprehensive study of artificial neural network applications

Fatih Güleç [a,*], Direnc Pekaslan [b], Orla Williams [a], Edward Lester [a]

[a] Advanced Materials Research Group, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, UK
[b] IMA and LUCID Research Group, School of Computer Science, University of Nottingham, NG8 1BB Nottingham, UK

ARTICLE INFO

Keywords:
Biomass
Higher heating value
Artificial intelligence (AI)
Artificial neural network (ANN)
Proximate analysis
Ultimate analysis

ABSTRACT

Higher heating value (HHV) is a key characteristic for the assessment and selection of biomass feedstocks as a fuel source. The HHV is usually measured using an adiabatic oxygen bomb calorimeter; however, this method can be time consuming and expensive. In response, researchers have attempted to use artificial neural network (ANN) systems to predict HHV using proximate and ultimate analysis data, but these efforts were hampered by varying case specific approaches and methodologies. Based on the complex ANN structures, a clear state of the art ANN understanding must be required for the prediction of biomass HHV. This study provides a comprehensive ANN application for HHV prediction in terms of how the activation functions, algorithms, hidden layers, dataset, and randomisation of the dataset affects the prediction of HHV of biomass feedstocks. In this paper we present a comparative qualitative and quantitative analysis of thirteen different algorithms, four different activation functions (logsig, tansig, poslin, purelin) with a wide range of hidden layer (3–15) for ANN models, used to predict the HHV of the biomass feedstocks. ANN models trained by the combination of ultimate-proximate analyses (UAPA) datasets provided more accurate predictions than the models trained by ultimate analysis or proximate analysis datasets. Regardless of the used datasets, sigmoidal activation functions (tansig and logsig) provide better prediction results than linear activation function (poslin and purelin). Furthermore, as training activation functions, "Levenberg-Marquardt (lm)" and "Bayesian Regularization (br)" algorithms provide the best HHV prediction. The best average correlation coefficients of 30 randomised run were observed with tansig as 0.962 and 0.876 for the ANN model developed by the UAPA dataset with a relatively high confidence levels of ~96% for training and ~92% for testing.

1. Introduction

Biomass and bio-based fuels such as bio-oil, biogas and biochar are potential energy sources of low carbon energy [1,2] to generate heat and power. Since biomass-based power generation is defined as a low carbon energy source [3], CO_2 emissions from biomass and/or bio-based fuels combustion can be considered to be approximately neutral in terms of mass balance [4–6]. Biomass has a lower energy content than coal [7,8], and thus more fuel is required for the same thermal output. The differences in physical and chemical properties of biomasses complicate the application of the biomasses in terms of the process application and product standardisation [9,10]. A variety of characterisation and experiment studies are required to assess the applicability of different kinds of biomass feedstocks for their potential application in energy conversion processes such as combustion, pyrolysis, gasification, liquefaction, hydrothermal processing [11–19]. Higher heating value (HHV, also termed gross calorific value or gross energy) is a crucial property for the thermal conversion of biomasses [20–23]. HHV refers to the heat released when a unit mass of the fuel is completely combusted and generated water in a condensed state [24,25]. The HHV of fuel sources such as biomasses can be measured experimentally using an adiabatic oxygen bomb calorimeter [20–22], which is the most common, accurate and simple method [20,21,26–28]. However, the method is time-consuming with high consumables costs [29,30], and therefore, is not always available to all researchers seeking to measure the HHV of biomasses. Numerous empirical correlations have been proposed in the

* Corresponding author.
E-mail address: Fatih.Gulec1@nottingham.ac.uk (F. Güleç).

https://doi.org/10.1016/j.fuel.2022.123944
Received 18 January 2022; Received in revised form 10 March 2022; Accepted 20 March 2022
Available online 28 March 2022

Table 1
Ultimate and Proximate analysis datasets for different kind of biomasses [54,55].

Biomass Samples	Ultimate analysis (wt.%)					Proximate analysis (wt.%)			HHV
	N	C	S	H	O	Ash	VM	FC	kJ/g
Almond shell	0.30	46.35	0.22	5.67	47.46	2.20	82.00	15.80	18.28
Almond tree branches	0.65	47.35	0.16	6.36	45.47	5.40	75.60	19.00	18.35
Almond tree leaves	2.85	43.25	0.34	5.50	48.06	9.30	87.19	3.50	17.56
American oak acorn	0.60	44.68	0.18	5.98	48.55	3.20	74.00	22.80	17.37
Apple tree branches	0.81	46.24	0.39	11.55	41.01	5.00	74.00	21.00	17.82
Apple tree leaves	1.61	44.45	0.23	6.15	47.56	12.00	71.90	16.10	17.51
Barley grain	1.79	41.59	0.35	6.08	50.18	3.00	76.90	20.10	16.52
Barley straw	1.64	40.69	0.23	6.95	50.50	6.10	77.90	16.00	17.37
Bean husk	0.66	39.66	0.31	5.38	53.98	8.00	74.00	18.20	15.11
Beetroot pellets	1.19	38.94	0.51	5.23	54.13	9.00	76.00	15.00	15.10
Black poplar bark	0.42	43.25	0.34	6.33	49.66	8.00	71.00	20.80	17.41
Black poplar leaves	1.03	58.30	0.35	8.41	31.92	7.80	71.20	21.00	18.17
Black poplar wood	0.18	46.19	0.56	5.70	47.36	1.50	86.00	12.30	18.39
Briquette	1.24	46.74	0.10	6.39	45.52	0.80	85.00	14.20	18.50
Building wastes chips	0.08	47.26	0.17	6.45	46.04	0.80	86.00	13.20	18.28
Cherry stone	0.43	48.57	0.19	6.21	44.60	0.87	85.00	14.10	19.07
Cherry tree branches	0.52	46.42	0.17	6.21	46.68	4.40	74.00	21.50	19.36
Cherry tree leaves	1.49	45.52	0.19	6.25	46.55	7.40	71.00	21.60	17.73
Chestnut shell	0.42	42.31	0.33	5.17	51.77	3.90	67.00	29.10	14.31
Chestnut tree chips	0.23	45.30	0.17	6.10	48.20	1.30	78.20	20.50	17.49
Chestnut tree leaves	2.21	47.82	0.27	6.24	43.46	4.90	72.41	22.70	18.76
Chestnut tree shaving	0.12	45.88	0.27	5.00	48.73	0.40	79.00	20.60	17.62
Cocoa beans husk	2.64	43.25	0.29	5.89	47.93	9.96	69.00	21.00	17.31
Coconut shell	0.15	47.93	0.24	6.05	45.63	1.40	79.20	19.40	18.88
Coffee husk	2.53	45.06	0.48	6.42	45.51	5.80	76.20	18.00	18.33
Corncob	0.22	44.78	0.21	6.02	48.77	2.40	83.00	14.60	17.69
Cypress fruit	0.35	27.81	0.18	5.70	65.96	4.70	71.80	23.50	20.17
Date stone	1.03	43.37	0.32	6.23	49.05	1.40	82.00	16.60	18.15
Eucalyptus bark	1.69	46.53	0.30	5.87	45.61	6.20	77.00	16.80	16.24
Eucalyptus chips	0.14	44.77	0.15	6.33	48.60	1.90	79.00	19.10	16.49
Eucalyptus fruit	1.14	46.81	0.39	5.81	45.84	4.70	73.60	21.70	18.52
Feijoa leaves	1.23	45.28	0.20	6.03	47.25	6.70	71.20	22.10	17.81
Gorse	1.49	43.49	0.33	5.53	49.16	5.00	45.20	50.20	18.60
Grapevine branches	0.76	45.00	0.46	6.95	46.83	7.60	71.50	20.90	16.82
Grapevine waste	1.35	35.74	0.30	5.95	56.67	13.30	73.00	13.70	16.47
Hazelnut and alder chips	0.40	45.47	0.20	5.94	47.99	5.00	77.00	18.00	17.56
Hazelnut shell	0.27	47.80	0.16	6.14	45.64	2.20	77.00	20.80	18.87
Hazelnut tree leaves	2.05	45.14	0.31	6.79	45.71	8.00	79.00	13.40	17.87
Holm oak branch chips	0.76	45.65	1.99	5.75	45.84	7.40	74.90	17.70	17.18
Horse chestnut burr	0.45	53.38	0.23	7.16	38.77	5.40	70.00	24.60	17.17
Horse chestnut tree br.	1.05	43.71	0.43	6.27	48.54	6.90	73.50	19.60	17.47
Kiwi branches	1.06	46.41	2.44	6.09	43.99	4.50	74.00	21.50	17.81
Lemon rind	1.08	42.95	0.42	6.56	48.98	9.70	73.20	17.10	17.18
Lemon tree branches	0.54	54.74	0.33	5.72	38.68	4.70	76.70	18.60	17.56
Maize grain	1.17	40.96	0.23	6.92	50.71	2.10	78.90	19.10	16.43
Mimosa branches	0.75	45.81	0.17	6.19	47.08	4.00	75.00	21.00	17.75
Medlar tree branches	0.52	44.36	0.18	6.17	48.77	8.40	74.00	17.60	17.65
Miscanthus	0.10	47.09	0.10	6.30	46.42	9.60	79.00	11.40	18.07
Nectarine stone	0.50	48.57	0.23	6.22	44.48	1.10	76.00	22.90	19.56
Oak acorn	0.80	41.84	0.25	6.82	50.28	2.60	75.10	22.30	16.17
Oak tree branches	2.87	48.26	0.33	6.28	42.26	4.20	78.40	17.40	17.72
Oak tree leaves	3.04	46.90	0.38	5.47	44.20	3.80	72.00	24.20	18.31
Oak tree pruning	0.73	37.89	0.21	5.94	55.23	4.30	77.00	18.70	17.59
Oats and vetch	0.92	41.69	0.29	5.82	51.27	7.33	72.00	20.70	16.66
Oats bran	2.17	44.01	0.29	7.17	46.36	4.15	77.00	18.90	18.06
Olive stone	1.81	46.55	0.11	6.33	45.20	1.40	78.30	20.40	17.88
Olive tree pruning	1.47	45.36	0.28	5.47	47.42	13.00	78.00	9.00	17.34
Orange tree branches	0.56	45.76	0.21	6.12	47.34	4.50	79.00	16.90	16.31
Orange tree leaves	2.59	41.11	0.40	5.28	50.62	15.40	73.20	11.40	16.17
Pea husk	1.24	39.62	1.82	6.54	50.78	4.50	83.00	12.50	15.46
Pea plant waste	0.90	44.06	0.39	4.73	49.91	5.80	78.00	15.90	17.35
Peach stone	3.94	40.72	0.30	6.96	48.07	0.50	75.60	23.90	19.59
Peach tree leaves	2.03	59.59	0.77	9.76	27.86	10.20	75.00	14.70	18.34
Peanut shell	1.05	49.35	0.24	6.40	42.96	2.50	81.00	16.50	20.09
Pepper plant waste	3.66	36.56	0.83	5.27	53.67	22.90	73.10	4.00	13.66
Pine and eucalyptus chips	1.59	45.90	0.19	6.30	46.03	3.60	71.60	24.80	16.99
Pine chips	0.09	48.15	0.28	5.59	45.90	0.60	81.60	17.80	19.43
Pine kernel shell	0.31	47.91	0.60	4.90	46.28	2.70	77.60	19.70	18.89
Pine pellets	0.28	46.83	0.31	5.30	47.28	1.30	83.50	15.20	18.84
Pine shaving	0.07	48.67	0.26	5.08	45.92	0.80	85.00	14.20	19.79
Pineapple leaf	0.40	42.26	0.27	4.81	52.27	3.20	75.00	21.80	18.15
Pinecone heart	0.29	42.22	0.84	5.06	51.59	3.50	66.00	30.50	16.44
Pinecone leaf	0.27	47.65	0.44	5.43	46.21	1.30	80.00	18.70	18.63

(*continued on next page*)

Table 1 (*continued*)

Biomass Samples	Ultimate analysis (wt.%)					Proximate analysis (wt.%)			HHV
	N	C	S	H	O	Ash	VM	FC	kJ/g
Pistachio shell	0.11	44.69	0.18	5.16	49.87	1.30	82.50	16.20	17.35
Plum stone	0.87	48.22	0.17	6.60	44.14	1.80	77.00	21.20	19.14
Pomegranate peel	0.69	42.19	0.33	5.11	51.68	6.80	68.00	25.20	15.17
Potato plant waste	1.13	38.33	0.44	5.07	55.03	15.80	69.00	14.70	15.07
Rice husk	0.21	26.69	0.17	2.88	70.05	13.70	73.00	13.30	15.90
Rye grain	1.20	41.11	0.21	6.76	50.72	1.80	78.90	19.30	16.14
Rye straw	1.16	40.18	0.32	6.85	51.48	3.20	79.90	16.90	17.11
Sainfoin	1.80	41.68	0.57	5.90	50.05	9.20	73.00	17.80	16.41
Sawdust	0.53	45.34	1.07	6.02	47.05	1.60	81.00	17.40	18.02
Sorghum	0.73	40.79	0.23	4.38	53.87	17.00	62.00	21.00	11.87
Soya grain	1.16	44.42	0.24	6.33	47.86	4.80	77.00	18.20	16.71
Straw pellets (grass)	0.56	47.89	0.17	5.51	45.87	9.80	79.00	11.20	16.58
Sunflower Seed husk	0.38	45.33	0.24	5.91	48.14	1.90	80.00	18.10	18.00
Tomato plant waste	1.19	36.63	1.48	0.68	60.01	16.20	78.00	6.00	14.15
Triticale	1.23	42.14	0.76	5.80	50.07	6.20	75.00	18.80	16.65
Vegetal coal	0.65	79.34	0.30	2.74	16.97	5.90	26.00	68.10	29.71
Vine orujillo	1.91	44.15	0.58	5.31	48.04	12.70	79.00	8.30	17.74
Vine shoot chips	0.61	40.15	0.31	5.02	53.91	9.70	66.00	24.30	14.63
Vine shoot waste	0.63	34.60	0.24	5.61	58.91	4.10	64.00	31.90	13.29
Walnut shell	0.22	46.97	0.10	6.27	46.44	2.30	79.00	18.70	18.38
Wheat bran	2.34	42.74	0.31	6.62	47.98	3.50	78.00	18.50	17.37
Wheat grain	0.24	49.22	0.26	6.52	43.76	2.80	80.00	17.20	16.33
Wheat straw	1.18	45.58	0.59	6.04	46.60	5.30	76.00	18.20	17.34
Wood chips	0.13	42.20	0.27	5.51	51.88	1.50	68.60	29.90	15.16
Wood pellets	0.60	46.79	0.32	6.13	46.15	1.30	82.00	17.10	18.22
Wood sawdust	0.12	45.97	0.24	5.13	48.53	0.60	83.00	16.40	18.21

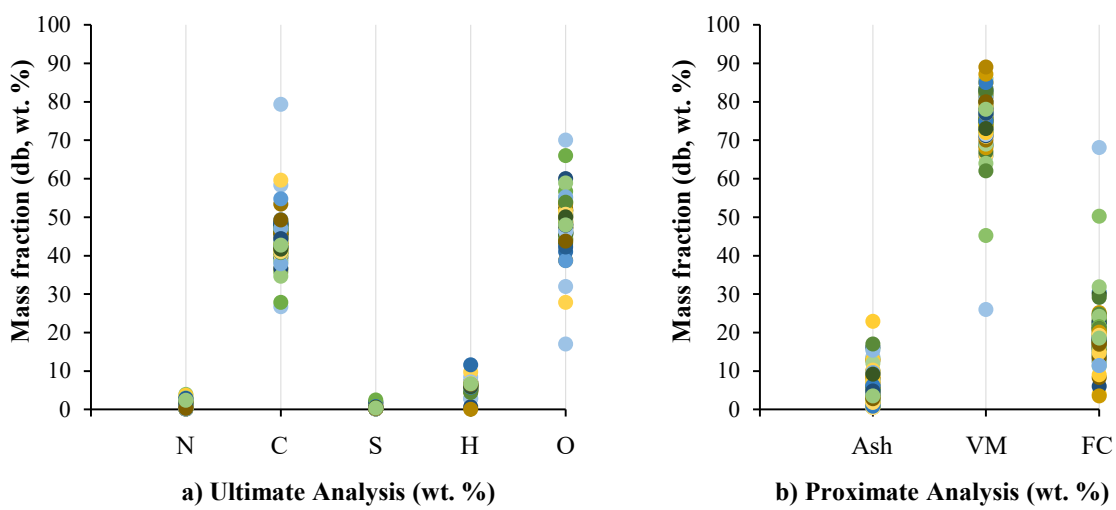

Fig. 1. Data distributions of a) ultimate and b) proximate analysis (mass fractions were presented on a dry basis (db)).

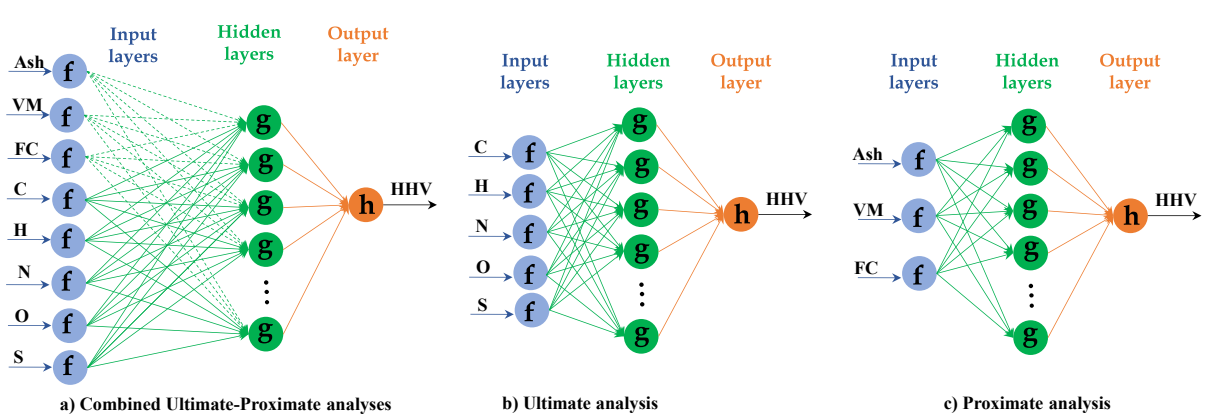

Fig. 2. The schematic topological architecture of three layers ANN models; input layers (a-combined Ultimate-Proximate analyses, b-Ultimate analysis and c-Proximate analysis datasets), hidden layers (3–14) and output layer (HHV).

Table 2
Artificial neural network structures.

Optimisation Algorithms	Activation Functions	Hidden Layers	Input Layers
1. Levenberg-Marquardt (lm) 2. Bayesian Regularization (br) 3. Scaled Conjugate Gradient(scg), 4. BFGS Quasi Newton (bfg) 5. Conjugate Gradient Powell/Beale Restarts (cgb) 6. Fletcher-Powell Conjugate Gradient (cgf) 7. Polak-Ribiére Conjugate Gradient (cgp) 8. Gradient Descent (gd) 9. Gaussian Discriminant Analysis (gda) 10. Gradient Descent Momentum (gdm) 11. Variable Learning Rate Gradient Dscent (gdx) 12. One Step Secant (oss) 13. Resilient Backpropagation (rp)	**1.** Log sigmoid transfer function (logsig) **2.** Tan sigmoid transfer function (tansig) **3.** Pure linear transfer function (purelin)**4.** Positive linear t ansfer function (poslin)	From 3 up to 15 nodes each trial	8 inputs nodes (PA + UA) 5 inputs nodes (UA) 3 inputs nodes (PA)

literature to estimate the HHV using the proximate analysis (PA); Moisture (M), Volatile matter (VM), Fixed carbon (FC), Ash (A), and ultimate analysis (UA); Carbon (C), Hydrogen (H), Oxygen (O), Nitrogen (N), Sulphur (S) for a variety of solid fuels such as coal [31–34], solid wastes [35–38], and biomass [20,35,39,40]. As UA and PA are the main characteristics of biomass fuel, predicting HHV from these characterisations eliminate the requirement of bomb calorimeter with an ANN model. These correlations are mainly based on linear regression even though biomass feedstocks show clear non-linear relationships between HHV and PA [41]. Furthermore, Ghugare et al. [29] demonstrated that linear models may not be the most appropriate method for the accurate prediction of biomass HHV, and non-linear models need to be investigated. The reliability of linear regression-based models is therefore relatively low and are inadequate in predicting the HHV for different types of biomass feedstocks [41]. Due to the variety of PA/UA values, the exact form of the PA/UA based nonlinear empirical model is unknown [29] and is beyond the realm of possibility without computational models.

More recently, the application of artificial intelligence (AI) for the prediction of experiment has gained attention due to its exclusively data-driven nonlinear modelling formalisms. Artificial neural network (ANN) is one of the most commonly used methods to analyse data that contains inherent dependency and non-linear relationships. Over the last few decades, ANN models have been thoroughly statistically tested in a wide range of technology and it has been demonstrated better prediction capacity over the counterpart ML models [42–44] such as logistic regression [45], SVM [46,47], ANFIS [48] or MLP [49]. The application of genetic programming (GP) and multilayer perceptron (MLP) neural network were investigated for the prediction of HHV of biomass feedstocks by Ghugare et al. [19]. Both models provided better prediction of HHV compared to existing linear and/or nonlinear simulations. Furthermore, Akkaya [41] demonstrated a high precision with adaptive neuro-fuzzy inference system (ANFIS) models based on fuzzy inference systems in which computational algorithms works in a collaborative way with expert knowledge and experimental results. Artificial neural networks (ANN) models are also emerging as an advanced tool for

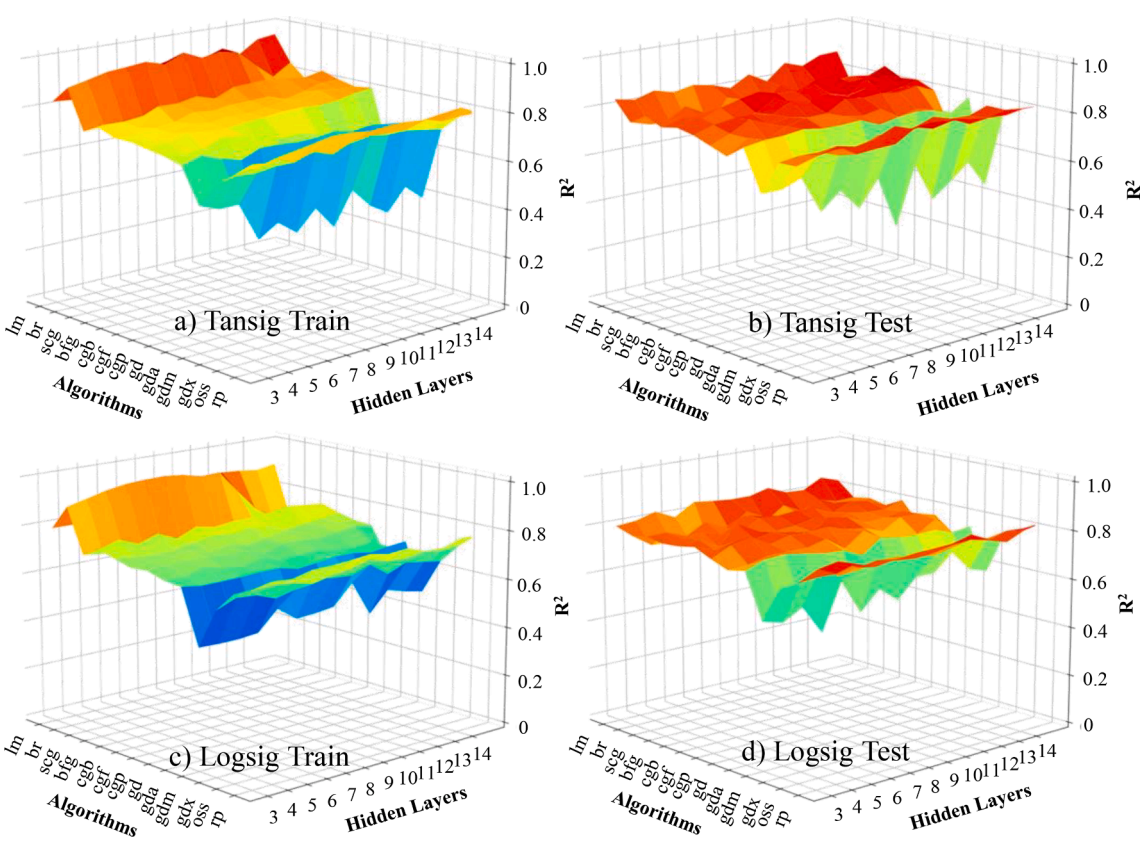

Fig. 3. Prediction results of the ANN models developed by sigmoidal activation functions, a) Tansig-train, b) Tansig-test, c) Logsig-train, and d) Logsig-test using ultimate-proximate analyses dataset.

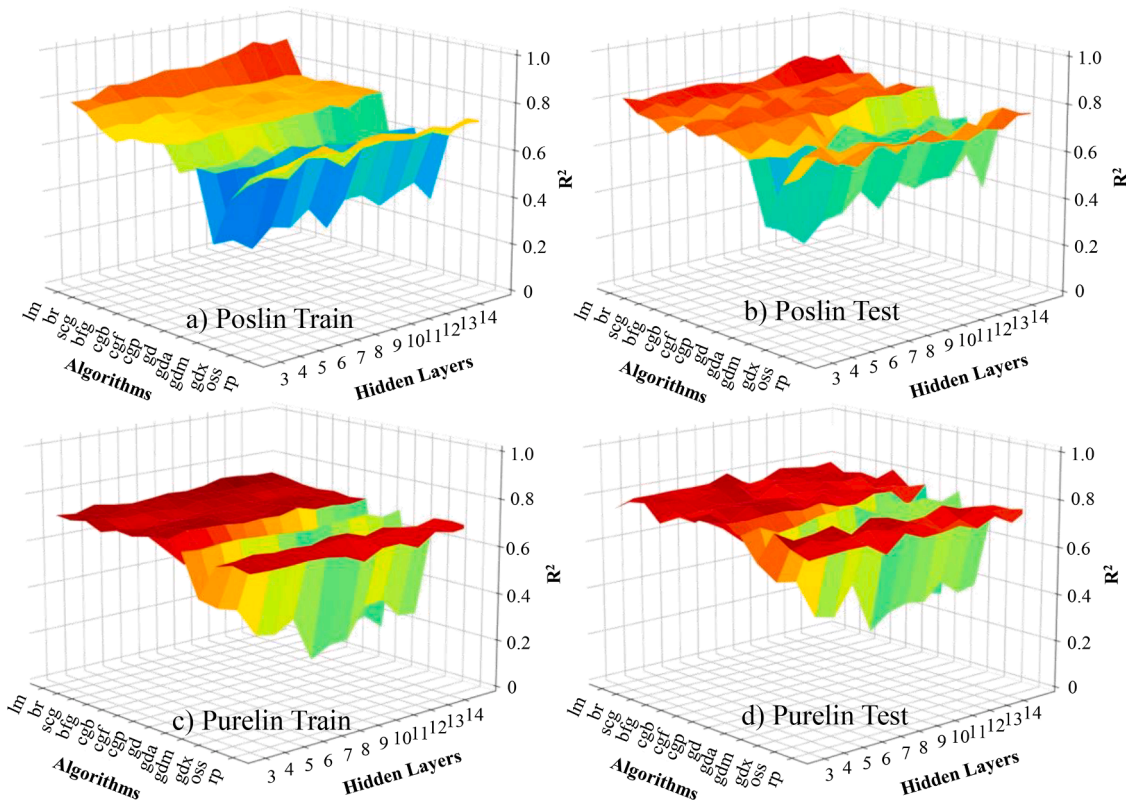

Fig. 4. Prediction results of the ANN models developed by linear activation functions, a) Poslin-train, b) Poslin-test, c) Purelin-train, and d) Purelin-test using ultimate-proximate analyses dataset.

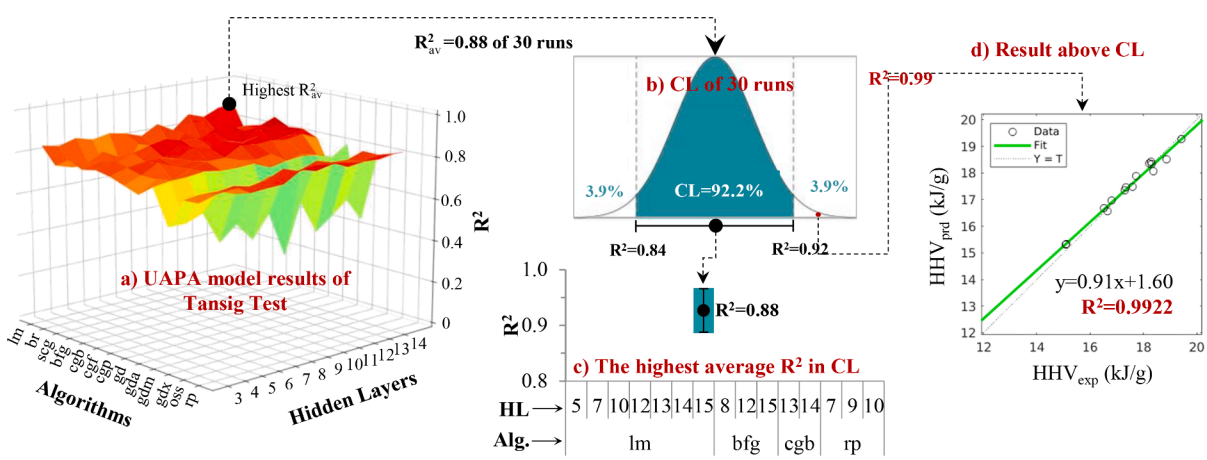

Fig. 5. a) Predicted average correlation coefficients (Tansig-test, also presented in Fig. 3b), b) the visualisation of confidence level of 30 random ANN runs, c) the highest 15 average R^2 values - each was determined by 30 random ANN runs, and d) the results above the confidence level (HL: Hidden Layer, Alg: Algorithms).

biomass feedstock [20,50,51] and bio/hydro-chars [21,52]. Jakšić et al. [20] demonstrated that ANN can predict the calorific values of different biomass feedstocks and any mixture of biomass feedstocks using the PA results. Pattanayak et al. [22] also developed three ANN models based on PA, UA, and a combination of UA-PA datasets to predict the HHV of a wide range of bamboo biomasses. Aladejare et al. [50] developed an optimised ANN model using particle swarm optimisation (ANN-PSO) to predict the HHV of a variety of different solid fuels such as coal, lignite, industrial waste, biomasses varying from agricultural wastes to forest residues using their UA and PA results. The ANN-PSO models predict the HHV of various solid biomasses better than the multilinear regression models as reflected in the statistical analysis conducted to validate the models. ANN models also provided better HHV prediction for the hydrochars [52] and biochars [21] produced by thermal conversion of biomass feedstocks. The ANN models usually perform better HHV predictions for both biomass and biobased products compared to the linear/nonlinear regression correlations [21,22,50,53]. With ANN, biomass feedstock HHV prediction has been shown to outperform HHV prediction using PA and/or UA datasets. Current literature is focusing on specifically tailored ANN models or integrating these models with other the sate-of-the art algorithms on a specifically chosen on set of biomasses. Although the current approaches report high accuracy in modelling HHV, the used ANN models are frequently unfeasible as studies are often focus on a case-specific model such as a specified custom ANN model structure based on a selected specific dataset which often ultimately presenting a selected highest correlation coefficient

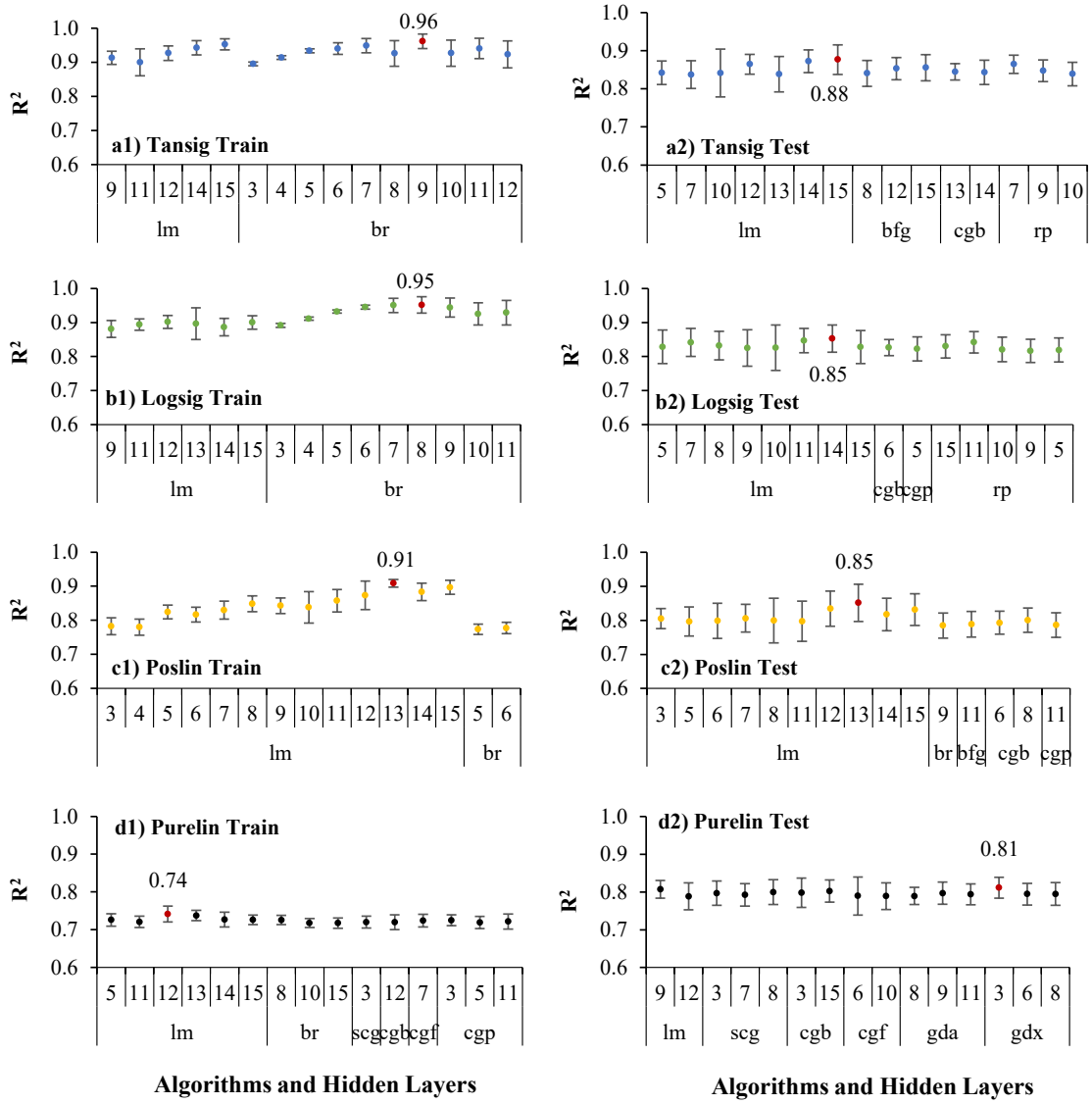

Fig. 6. The best fifteen average correlation coefficients at the ANN models developed by the combination of ultimate-proximate analyses dataset. (Numbers in x-axis represent the hidden layers, the bars represent the confidence intervals of the average correlation coefficient).

(R^2) value in the experiment results. Consequently, the applicability of these models is not generalisable due to the case-specific model structures. Furthermore, there is not a clear and detailed ANN study for the prediction of biomass HHV based on the algorithms, activation functions, and hidden layers.

In this study, a detailed state-of-the-art ANN model was created to predict HHV of biomass based on PA and UA datasets of different biomass feedstocks. Whilst some literature exists on the application of ANN for the prediction of HHV of various biomass fuels, there is a need for a detailed understanding of the importance of data and randomisations in training and testing data partitioning. This research covers the comparative qualitative and quantitative analysis of thirteen different algorithms, four different sigmoidal and linear activation functions (logsig, tansig, poslin, purelin), a wide range of hidden layers (3–15) for training the ANN model which predicts the HHV of the biomass feedstocks characterised by the PA, UA, and combination of UA-PA.

2. Material and methods

2.1. Biomass feedstocks dataset

As dataset, a hundred data (both PA and UA of biomass feedstocks, Table 1) were collected from the literature [54,55]. This data was used to develop a correlation between PA–UA and HHV of biomass feedstocks. The dataset consists of a wide range of biomass feedstocks such as commercial fuels (17 data points), industrial wastes (32 data points), forest wastes (including branches and leaves, 38 data points), energy crops and cereals (13 data points), which all can build an ANN model valid in a wide range of biomass feedstocks to predicts the HHV. Additionally, the statistical data distributions of UA and PA in mass fractions are presented in Fig. 1.

2.2. Artificial neural network structures

In the ANN model constructions, three main representative models are built based on the experimental setting. The first model's input layer is designed to receive 8 inputs (C, H, O, N, and S (ash free distribution) and Ash, volatile matter (VM), fixed carbon (FC) on dry basis), as shown

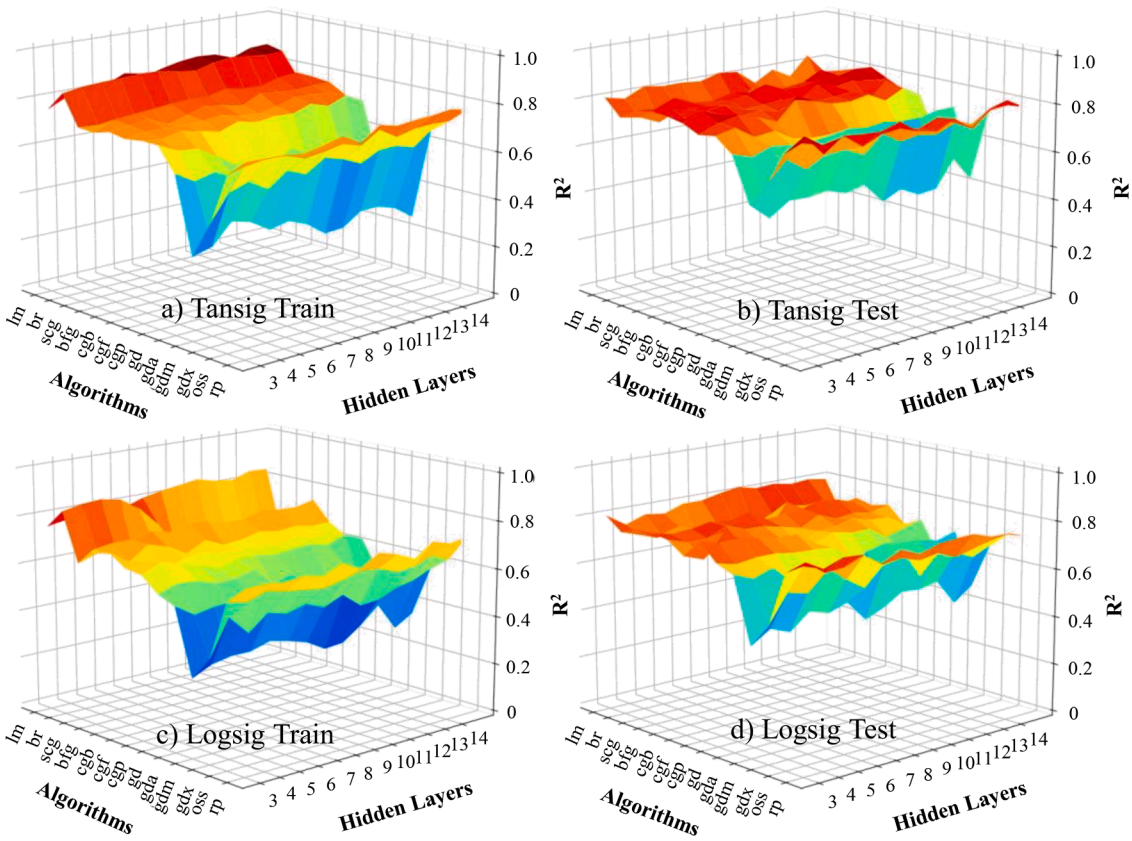

Fig. 7. Prediction results of the ANN models developed by sigmoidal activation functions, a) Tansig-train, b) Tansig-test, c) Logsig-train, and d) Logsig-test using ultimate analysis dataset.

in Fig. 2a. The second models input layer is designed to receive 5 inputs (C, H, O, N, and S (ash free distribution)), as shown in the Fig. 2b. Finally, the third model input layer is designed to receive 3 inputs (Ash, volatile matter (VM), fixed carbon (FC) on dry basis), as shown in Fig. 2c.

Later, a comparative qualitative and quantitative analysis were carried out on each of the ANN models. Overall, as shown in Table 2, thirteen optimisation algorithms, four activation functions, and from three to fifteen hidden layers were used for training each ANN model, which predict the biomass HHV with different input sets. In total, extensive number of experiments ($13 \times 4 \times 12 \times 3 = 1872$) are implemented to test and report the ability of the state-of-the-art ANNs with different structure and algorithms. Furthermore, in order to capture the statistical insight, each of these experiments are repeated 30 times. A detailed explanation of the used optimisation algorithms has been previously reported [56,57]. As shown in Table 2, the activation functions are selected based on two main structures as linear (poslin and purelin) and sigmoidal (logsig and tansig).

Log sigmoid transfer function (logsig)

$$z = f(x) = logsig\,(x) = \frac{1}{1+e^{-x}} \tag{1}$$

Tan sigmoid transfer function (tansig)

$$z = f(x) = tansig\,(x) = \frac{2}{(1+e^{-2x})-1} \tag{2}$$

Pure linear transfer function (purelin)

$$z = f(x) = purelin\,(x) = (x, 0) \tag{3}$$

Positive linear transfer function (poslin)

$$z = f(x) = poslin\,(x) = \begin{cases} 0, & x \leqslant 0 \\ x, & x \geqslant 1 \end{cases} \tag{4}$$

Due to the stochastic nature of optimisation process in ANNs, as a common practice the weights of each ANN are initialised randomly for each experiment. Also, in order to minimise any bias in the model and increase the chance of having a generalisable model, the randomness is used in the selection of training, validation and testing datasets. For training, validation and testing the model, the dataset is divided into 70:15:15 ratios. Accordingly, the ANN model randomly selected of those sets in test and validation analysis. In the ANN models, the state-of-the-art instruments have been used to avoid overfitting. As mentioned throughout the paper, the random data partitioning has been applied which maximise the diverse coverage of selected samples for both training and testing sets. This minimises the chance of having a biased model on dataset and increasing the chances of having a representative dataset for training. In addition to randomisation and the statistical procedures, the validation technique has also been applied to avoid the overfitting of the models. Furthermore, extensive experiments have been carried out to deduce the structure which statistically to be optimum model in predictions. The ANN network model is formed using MATLAB Deep Learning Toolbox [58,59]. The training settings of the 3 main ANN models were selected as follows: number of input nodes 8 for combined UA-PA, 5 for UA, and 3 for PA, number of hidden layers are between 3 and 15, number of output node is 1 (HHV), as presented in Fig. 2. The results of 30 randomised run were presented with a confidence interval of minimum and the maximum coefficient of correlation values (R^2).

2.3. Statistical analysis of the developed ANN models

The developed ANN models randomly choose the training data from

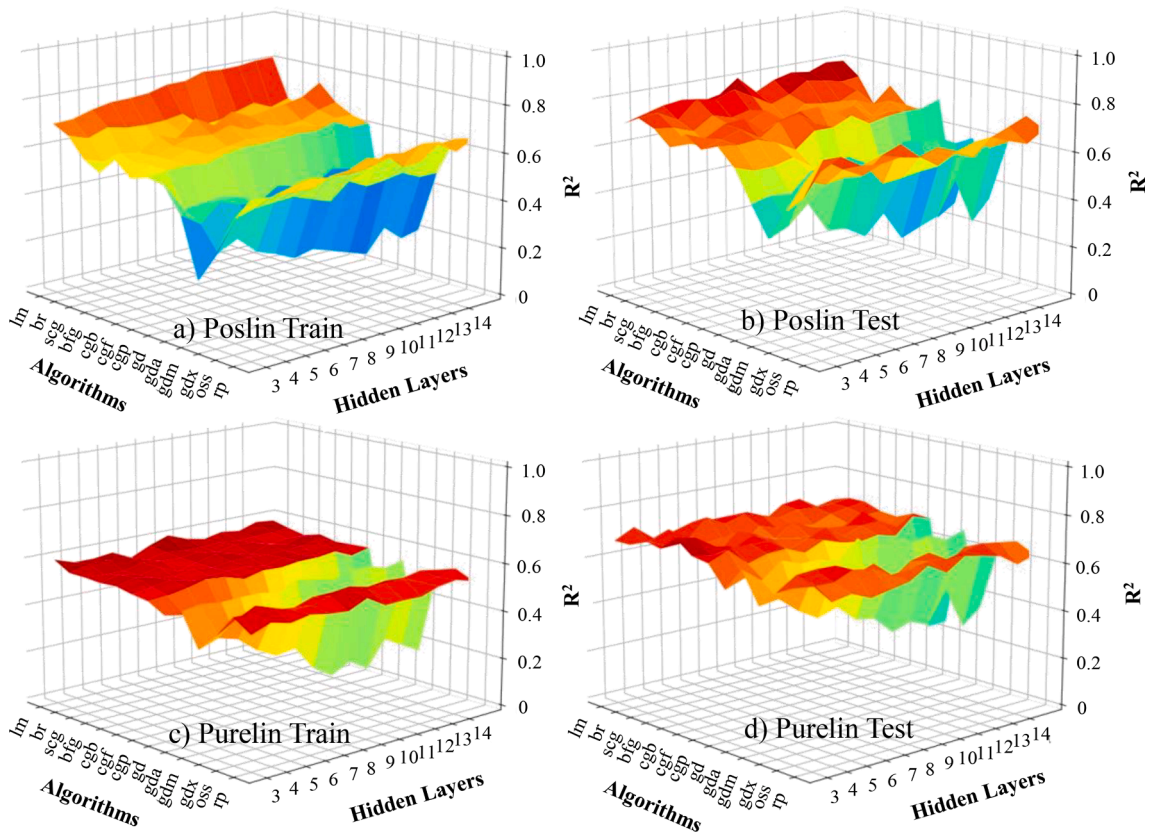

Fig. 8. Prediction results of the ANN models developed by linear activation functions, a) Poslin-train, b) Poslin-test, c) Purelin-train, and d) Purelin-test using ultimate analysis dataset.

the dataset to make the model more comprehensive for other types of biomass feedstocks. Therefore, the model shows a range of predicted data depending on the run number instead of a specific number as experimental data points. Statistical evaluation of the results produced by ANN models are very important in demonstrating the applicability of these ANN models in this specific subject. Confidence interval (CI) (Equation (5)) is a useful statistical method used in these kinds of applications, as they include a population value with a certain degree of confidence. As CI shows how the average results could be produced in the same confidence range and eliminate the highest and lowest points in the application. CI is expressed as a percentage, and it represents how much the prediction results will match with the results produced by experimental results in multiple cycles. The other most commonly used statistical values are correlation coefficient (R^2, Equation (6)) and mean square error (MSE, Equation (7)), which are also reported.

$$CI = \bar{x} \pm z\frac{s}{\sqrt{n}} \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N_{total}} (y_{i,cal} - y_{i,exp})^2}{\sum_{i=1}^{N_{total}} (y_{i,exp} - \bar{y}_{exp})^2} \tag{6}$$

$$MSE = \frac{1}{N_{total}} \sum_{i=1}^{N_{total}} (y_{i,cal} - y_{i,exp})^2 \tag{7}$$

Where CI is confidence interval, x is sample mean, z is confidence level value, s is the sample standard deviation, n is the sample size. R^2 and MSE are the correlation coefficient and mean square error, respectively. N_{total} is the number of the total samples. i is the sample index. $y_{i,cal}$ and $y_{i,exp}$ are the predicted and measured HHVs for the ith sample, and y_{exp} is the average measured HHV of the entire training or application samples.

3. Results and discussions

In order to demonstrate the potential applicability of ANN for the prediction of biomass HHV feedstocks, detailed ANN models were developed and presented with the models developed by thirteen different training functions, four different activation functions and up to fifteen hidden layers using a combination of UA-PA datasets. The ANN models was assessed using the datasets of PA (VM, FC, Ash, dry basis), UA (C, H, O, N, and S) and combination of UA-PA to predict HHV. The prediction results of ANN models are evaluated and presented with the statistical analysis and data interpretation.

3.1. HHV prediction via the combination of ultimate and proximate analyses

In order to predict the HHV of the biomass feedstocks, the first ANN model was developed using eight inputs of combine UA-PA dataset (C, H, O, N, S, and VM, FC, Ash) and one output as HHV with a wide range of hidden layers, training functions and activation functions. The relationship of hidden layers, training functions, and correlation coefficients (R^2, between the experimental HHV and predicted HHV) are presented in Fig. 3 and Fig. 4 based on the ANN models build by sigmoidal and linear activation functions, respectively.

Each correlation coefficient presented in Figs. 3 and 4 is the average of 30 runs of randomly selected train (85 data points) and test (15 data points) data with the defined activation functions, algorithms, and hidden layers. Although there is a slight difference, the train and test results show relatively similar trends for different activation functions; sigmoidal transfer functions (tansig and logsig) in Fig. 3 and linear transfer functions (poslin and purelin) in Fig. 4.

Fig. 3 and Fig. 4 show that the training functions could be categorised into three groups:

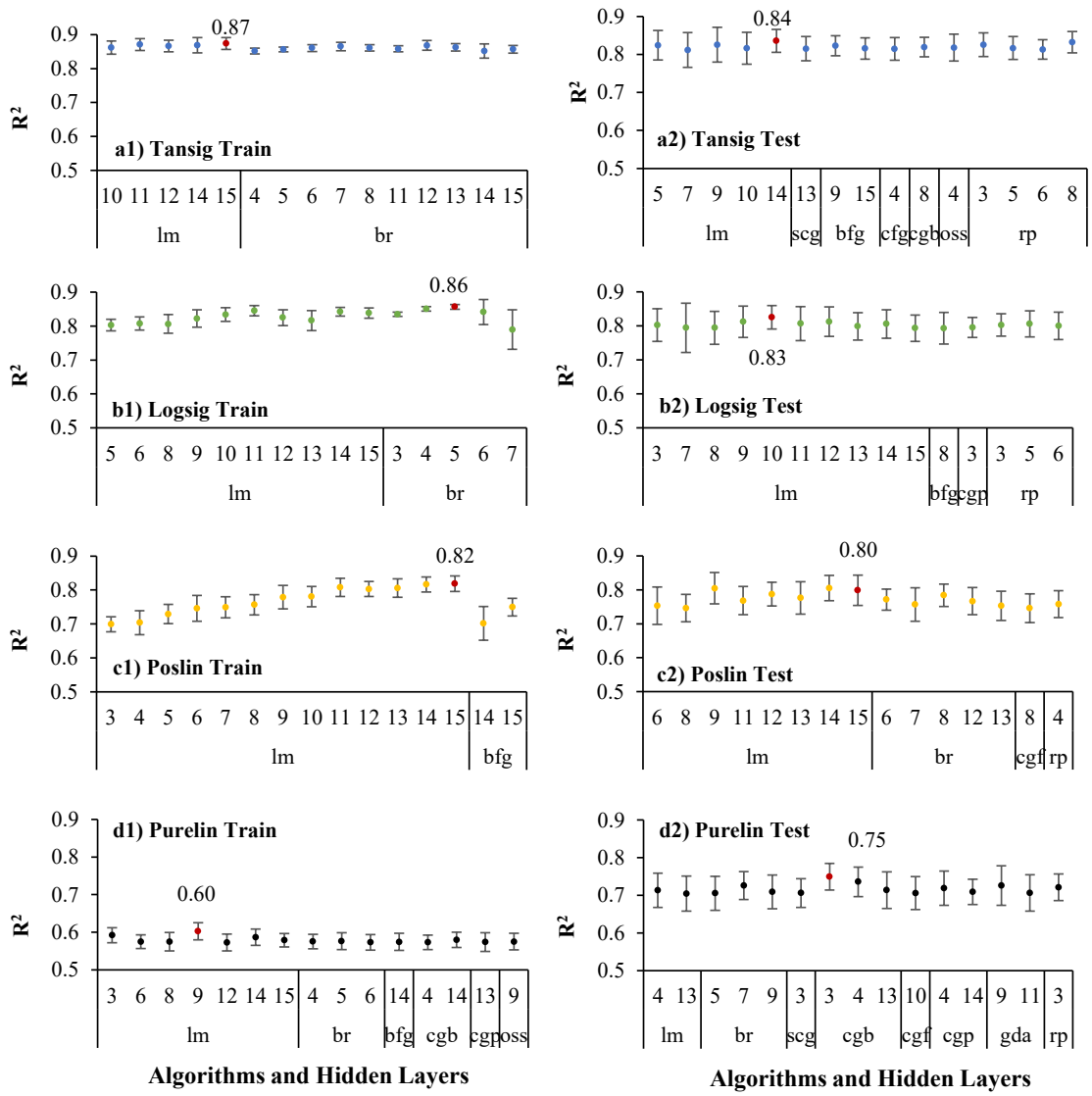

Fig. 9. The highest fifteen average correlation coefficients at the ANN models developed by ultimate analysis dataset. (Numbers in x-axis represent the hidden layers, the bars represent the confidence intervals of the average correlation coefficient).

i) algorithms ("lm" and "br") provide good HHV prediction
ii) algorithms ("scg", "bfg", "cgb", "cgf", "oss", and "rp") provide fair HHV prediction
iii) algorithms ("gd", "gda", "gdm" and "gdx") failed to predict the HHV

Regardless of activation function, the higher correlation coefficients were observed with the "lm" and "br" algorithms at the higher hidden layers of 11–15 in training (Fig. 3a, 3c, 4a, and 4c). Additionally, category-ii algorithms also show higher correlation coefficient in the test ANN models as demonstrated in Fig. 3c-d and 4c-d. Sigmoidal activation functions (tansig and logsig) provided better prediction results (Fig. 3) than linear activation function (poslin and purelin, Fig. 4). Furthermore, in the sigmoidal group, the tansig activation function (Fig. 3a-b) demonstrated better predictions than the logsig activation function (Fig. 3c-d) in both train and test groups. The order of activation functions was tansig > logsig > poslin ≫ purelin.

In order to elaborate the reported results in the statistical view, Fig. 5 demonstrates a representative Fig. of average correlation coefficient, the confidence level of this average correlation coefficient, and the highest correlation coefficient above the confidence level. Fig. 5a shows the average corelation coefficient of UA-PA Tansig Test models. Specifically, the average correlation coefficient of the experiment Tansig-Test with "lm" algorithm, and 15 hidden layer provides a confidence level of 88% (see black point in Fig. 5a). However, as mentioned this is the average of 30 repeated implementation. All 30 repeated implementations are visualised in Fig. 5b, where the confidence interval of this particular experiment is shown to be between 0.84 and 0.92. The ANN model run for Tansig-Test with "lm" algorithm, and 15 hidden layer provides a confidence level of 92.2% (lead to ~0.04 confidence intervals) with an average correlation coefficient of $R^2 = 0.88$ (Fig. 5b). The highest fifteen average correlation coefficient for the activation function presented with the related algorithms and hidden layers as in Fig. 5c. These average correlation coefficients with the confidence level demonstrate how ANN model can be applicable for the prediction of HHV. An average correlation coefficient with a higher confidence level demonstrates the reliability of the prediction of biomass HHV using the developed ANN models. However, some correlation coefficients predicted in these randomised 30 ANN runs could be above (and/or below) the confidence level, i.e. $R^2 = 0.9922$ in the same ANN model for Tansig-Test with "lm" algorithm, and 15 hidden layers (Fig. 5d). Presenting the statistical analysis, particularly regarding confidence intervals and confidence

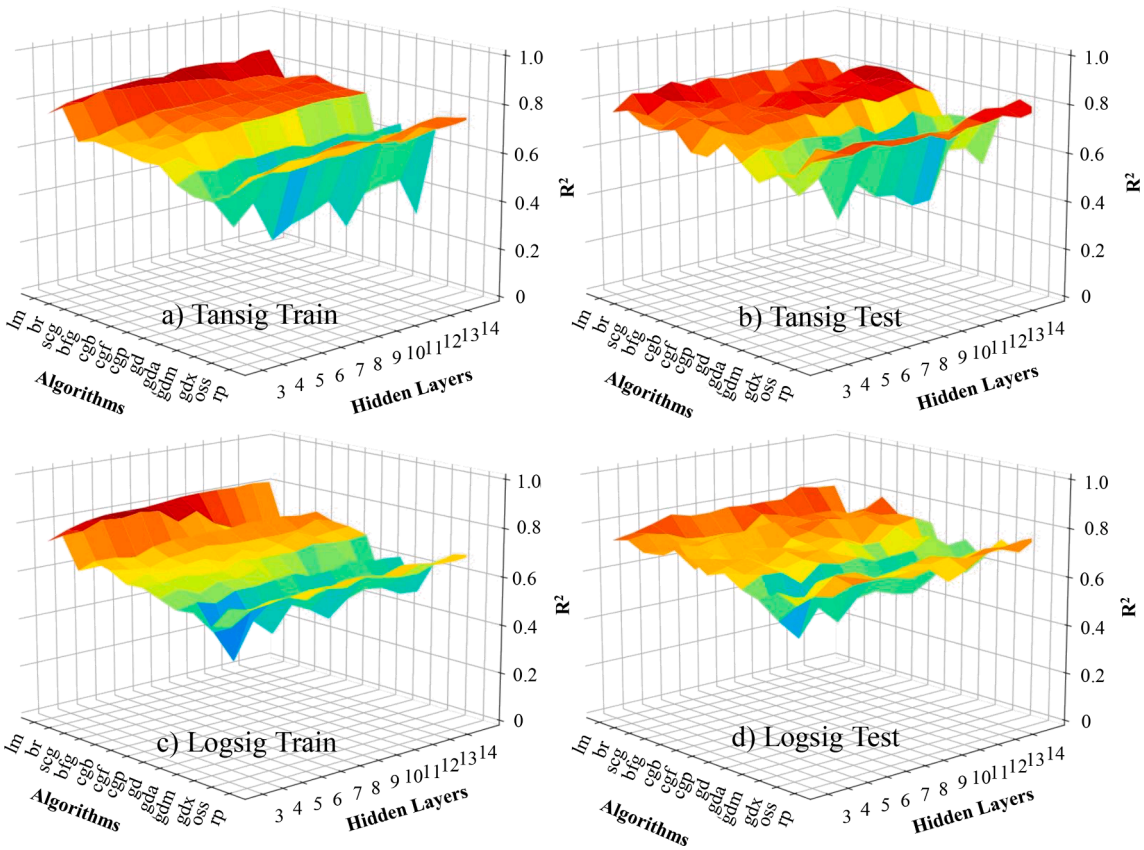

Fig. 10. Prediction results of the ANN models developed by sigmoidal activation functions, a) Tansig-train, b) Tansig-test, c) Logsig-train, and d) Logsig-test using proximate analysis dataset.

level are therefore critical in these kind of ANN predictions in order to justify the applicability of ANN models for the prediction of targeted values. As in the statistical point of view, the highest correlation coefficient does not validate the ANN models if it is not in the confidence level. Further details of the importance of statistical analysis are presented in Section 3.5.

In order to demonstrate the best model structure, the best 15 correlation coefficients (each is the average of 30 runs) are presented with the ANN structure of algorithms and hidden layers in Fig. 6. The "lm" and "br" algorithm provides relatively higher prediction values for both sigmoidal and linear activation functions at a wide range of hidden layers. In terms of the average correlation coefficients calculated on a set of training data, the best performance (higher R^2), had the "br" training function for the sigmoidal activation functions (0.96 for Tansig and 0.95 for Logsig) and "lm" for the linear activation functions (0.91 for Poslin and 0.74 for Purelin). As for the correlation coefficient calculated on a set of testing data, "lm" training function has the best performance for Tansig, Logsig, and Poslin activation functions for the hidden layers higher than 5. In general, training results provide a higher confidence level (lower confidence intervals) than the testing results regardless of the activation functions, which could be attributed to the datasets used for training and test work. The increase in the test datasets would increase the confidence level and potentially the average correlation coefficient.

3.2. HHV prediction via ultimate analysis

The second ANN model was developed using five inputs of UA (C, H, O, N, S) and one output as HHV with a wide range of hidden layers, training functions and activation functions. The relationship between hidden layers, training functions, and correlation coefficients (R^2) between the experimental and predicted data are presented in Figs. 7 and 8 based on the ANN models built using sigmoidal and linear activation functions, respectively. Figs. 7 and 8 show that UA dataset also generate three similar categories of training functions as defined in Section 3.1.

Regardless of activation function, category-i ("lm" and "br") showed relatively higher average correlation coefficients in the ANN models developed by UA, which is similar to the ANN models developed by combined UA-PA (in Section 3.1). Furthermore, category-iii algorithms such as "gd", "gda", "gdm" and "gdx" failed to predict the HHV of biomasses as these algorithms showed the lowest average correlation coefficient regardless of activation function. The hidden layers up to 15 has insignificant effects on the prediction of HHV unlike algorithms and activation functions. As similar to the results provided in Section 3.1, sigmoidal activation functions (tansig and logsig) provided higher average correlation coefficients (Fig. 7) than linear activation function (poslin and purelin, Fig. 8). In the ANN models developed by UA, the activation functions could be followed as similar as ANN models developed by combined UA-PA; tansig > logsig > poslin ≫ purelin.

The best 15 average correlation coefficients (each is the average of 30 randomise runs) are presented in Fig. 9 with the models (algorithms and hidden layers) and confidence intervals. The "lm" algorithm provides relatively higher prediction values for both sigmoidal and linear activation functions at a wide range of hidden layers. Following "lm", the other training functions such as "br" and "bfg," also provide relatively good prediction in training and testing (Fig. 9).

The best average correlation coefficients for train and test were observed with tansig as (R^2 = 0.87 and 0.84), followed by logsig (R^2 = 0.86 and 0.83) and poslin (R^2 = 0.82 and 0.80), which provides a reasonable level of HHV prediction. However, purelin provides relatively poor average correlation coefficients (R^2 = 0.60 and 0.75) for

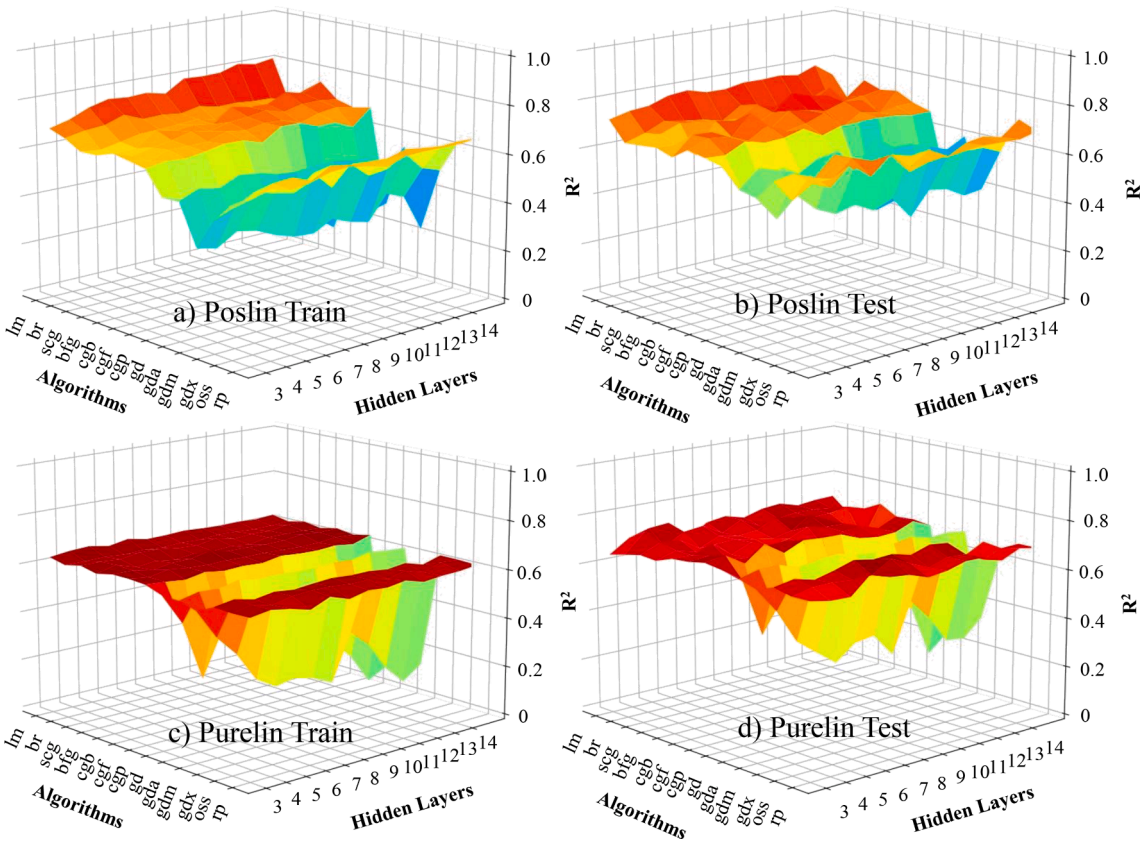

Fig. 11. Prediction results of the ANN models developed by linear activation functions, a) Poslin-train, b) Poslin-test, c) Purelin-train, and d) Purelin-test using proximate analysis dataset.

train and test results (in Fig. 8-c and -d). Additionally, the confidence intervals of the average correlation coefficients for training results usually lower than that for testing results, as presented with the error bars in Fig. 9.

3.3. HHV prediction via proximate analysis

Although the calorific value of a solid fuels usually relies on the combustible portion of the fuel, non-combustible portions such as "ash" can also contribute the calorific value due to a potential catalytic reaction during combustion [60,61]. Therefore, in order to predict the HHV of the biomass feedstocks, a second ANN model was developed using three inputs of PA (VM, FC, Ash) and one output as HHV with a wide range of hidden layers, training functions and activation functions. The effects of hidden layers, training functions, and correlation coefficients (between the experimental HHV and predicted HHV) are presented in Figs. 10 and 11 based on the ANN models build by sigmoidal and linear activation functions, respectively, using the PA dataset.

Similar to the previous models, the train and test results show relatively similar trends for different activation functions; sigmoidal transfer functions (tansig and logsig) in Fig. 10 and linear transfer functions (poslin and purelin) in Fig. 11. As previously defined with the models developed by combined UA-PA (section 3.1) and UA (section 3.2), the training functions could be divided into three categories, algorithms have better HHV prediction ("lm" and "br"), fair HHV prediction ("scg", "bfg", "cgb", "cgf", "oss", and "rp"), and failed to predict the HHV ("gd", "gda", "gdm" and "gdx") in the model developed by PA. Regardless of activation function, the higher correlation coefficients were observed with the "lm" and "br" algorithms at the higher hidden layers in training (Fig. 10a, 10c, 11a and 11c) while category-i and -ii algorithms provide similar correlation coefficient in testing (Fig. 10b, 10d, 11b, and 11d). Sigmoidal activation functions (Tansig and Logsig, Fig. 10) provided better prediction results than linear activation function (Poslin and Purelin, in Fig. 11). Among the sigmoidal activation functions, tansig shows the best correlation coefficients in both train and test results (Fig. 10a and 10b) with the PA dataset, which is similar with the results provided by combined UA-PA (section 3.1) and UA (section 3.2) datasets. The alignment of activation functions is also similar as previous sections; tansig > logsig > poslin ≫ purelin.

The best 15 average correlation coefficients (each is the average of 30 randomise runs) are presented in Fig. 12 with the models (algorithms and hidden layers) and confidence intervals. The "lm" is the only algorithm provides higher correlation coefficients for the activation functions of tansig, logsig, and poslin as shown in Fig. 12a, 12b, and 12c. The best average correlation coefficient of 30 randomise runs was observed as $R^2 = 0.85$ and 0.83 for tansig train and test, $R^2 = 0.82$ for logsig train and test, and $R^2 = 0.81$ and 0.79 for poslin train and test in Fig. 12. Besides "lm", the other algorithms such as "br" and "bfg" provide relatively good prediction in training (Fig. 12-a1, b1, c1) and "br", "bfg", "scg", "cgb", "cfg", and "rp" are the other potential algorithms showing fair prediction in testing (Fig. 12-a2, b2, c2). Additionally, the train results provide better confidence levels (lower confidence intervals) than the test results for all four activation functions, as demonstrated in Fig. 12. The ANN model developed by the combination of UA-PA provide higher average correlation coefficients (Fig. 6) than the ANN models developed by only UA (Fig. 9) and only PA (Fig. 12) datasets.

3.4. Comparison of the ANN models

The best average correlation coefficients between the experimental and predicted HHV for each activation functions (tansig, logsig, poslin and purelin) and developed ANN models are presented in Table 3 with the algorithms and hidden layers. Regardless of the dataset, tansig and logsig activation functions provides the best average correlation

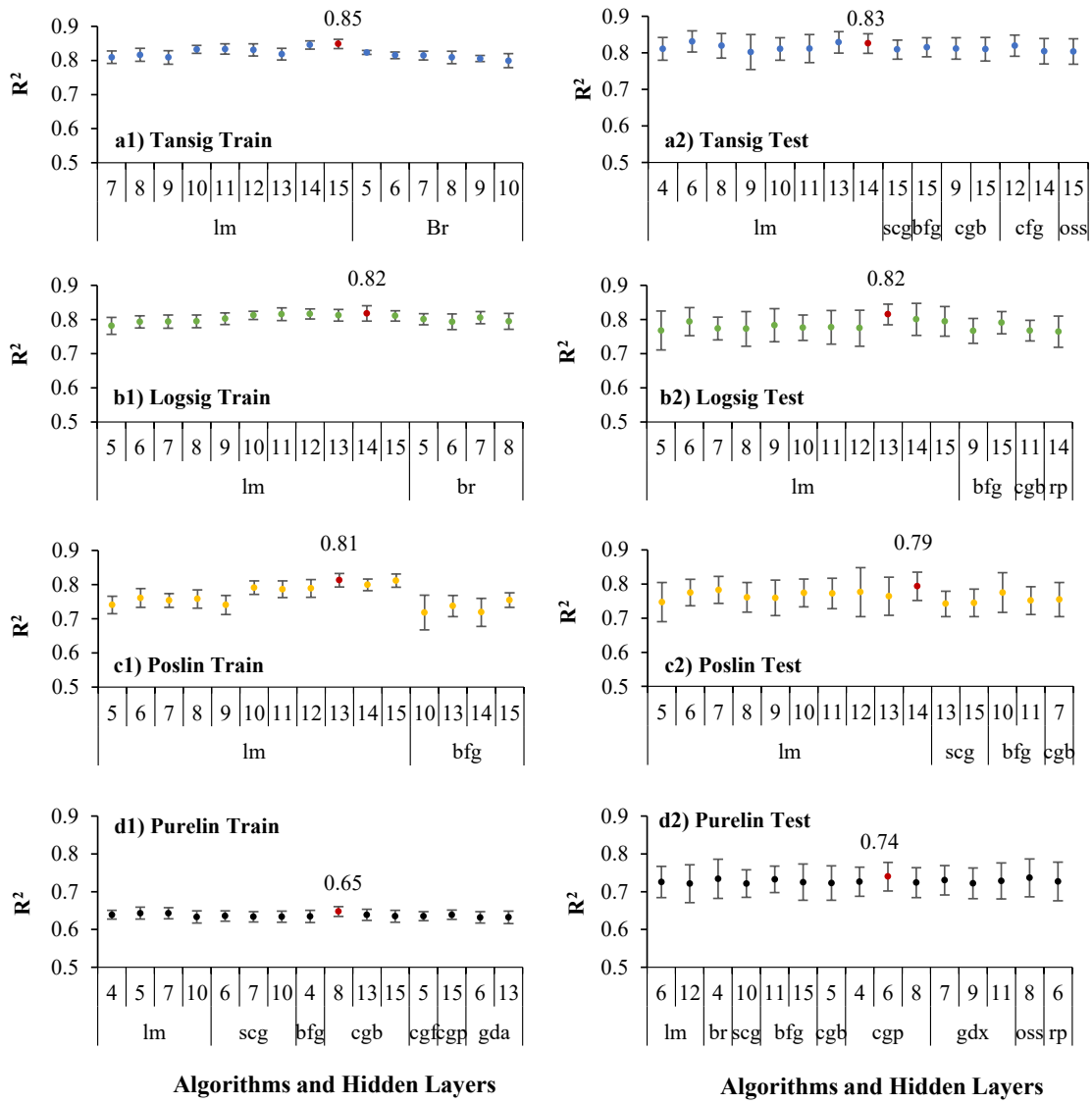

Fig. 12. The best fifteen average correlation coefficients at the ANN models developed by proximate analysis dataset. (Numbers in x-axis represent the hidden layers, the bars represent the confidence intervals of the average correlation coefficient).

coefficients for train and test results. Following tansig and logsig activation functions, poslin also provided reasonable level of average correlation coefficients. Among 14 algorithms, "lm" and "br" are defined the best algorithms for the train and test models in the activation functions of tansig, logsig and purelin. While "lm" provides better results with a high hidden layer (13–15) in the models developed via low number inputs (Model-3, three inputs of PA), "br" shows the better results with a lower hidden layer (5–9) in the models developed via high number inputs (Model-1, eight inputs of ultimate-proximate analyses). The ANN model developed by the combination of ultimate-proximate dataset provide the highest average correlation coefficients (0.965 for train and 0.876 for test) and the corresponding MSE (0.15 and 0.69) are the lowest among the three ANNs models. Furthermore, the results are provided relatively high confidence levels, higher than 95% for the training and 89% for the testing results regardless of activation function. Confidence level (CL) demonstrates how confident the average correlation coefficient results from the 30 runs of ANN model with the same algorithm, hidden layer, activation function and randomly selected train and test data.

3.5. Importance of statistical analysis

There are two concerns in the ANN applications for the prediction of HHV using ultimate and proximate analyses datasets. The first is data selection, as the datasets were usually divided into two different data groups and used as separated train data and test data in the ANN applications [21,22,26,29,41,62]. The clarification "how" and "why" the datasets divided as test data and train data are not provided. The only clarification on the data separation as train and test was "to ensure a homogeneous distribution of data (to select the training versus testing sets) data were put in increasing order of HHV, and every third dataset was selected for testing" [25]. However, the data partition (specifically being homogenous) into two categories (train and test) can cause bias for the developed models, as homogenously divided data can *trick* the model in the training phase. Consequently, the biased model could provide relatively higher prediction value since the test data is *similar* to the train data. The second concern is presenting the ANN results without a right statistical analysis – in some literature, the results were justified based on the highest correlation coefficient and lowest mean square error [22,28,63]. However, due to the non-deterministic behaviours of ANNs, the selected highest correlation coefficient does not validate the

Table 3
The best average correlation coefficients and corresponding MSE for the ANN models and activation functions.

Activation Functions	Train results					
	Alg.	HL	(R^2)*	CL**	CI†	MSE+
ANN model 1: Combination of ultimate-proximate analyses dataset						
Tansig	br	9	0.962	95.7	0.02	0.15
Logsig	br	8	0.951	95.2	0.02	0.19
Poslin	lm	13	0.908	97.8	0.01	0.37
Purelin	lm	12	0.741	95.8	0.02	0.99
ANN model 2: Ultimate analysis dataset						
Tansig	lm	15	0.873	96.5	0.02	0.60
Logsig	br	5	0.856	98.6	0.01	0.57
Poslin	lm	15	0.818	95.4	0.02	0.75
Purelin	lm	9	0.602	95.4	0.02	1.37
ANN model 3: Proximate analysis dataset						
Tansig	lm	15	0.848	97.3	0.01	0.64
Logsig	lm	14	0.818	95.5	0.02	0.74
Poslin	lm	13	0.812	95.9	0.02	0.75
Purelin	cgb	8	0.647	97.4	0.01	1.28
Activation Functions	**Test results**					
	Alg.	**HL**	**(R^2)***	**CL****	**CI†**	**MSE+**
ANN model 1: Combination of ultimate-proximate analyses dataset						
Tansig	lm	15	0.876	92.2	0.04	0.69
Logsig	lm	14	0.852	92.0	0.04	0.62
Poslin	lm	13	0.851	89.1	0.05	0.84
Purelin	gdx	3	0.811	94.4	0.03	0.87
ANN model 2: Ultimate analysis dataset						
Tansig	lm	14	0.835	93.9	0.03	0.75
Logsig	lm	10	0.825	93.0	0.03	0.64
Poslin	lm	14	0.805	92.5	0.04	0.74
Purelin	cgb	3	0.749	92.9	0.04	1.41
ANN model 3: Proximate analysis dataset						
Tansig	lm	14	0.831	94.2	0.03	0.85
Logsig	lm	13	0.815	93.9	0.03	0.79
Poslin	lm	14	0.793	91.7	0.04	0.88
Purelin	cgp	6	0.739	92.4	0.04	0.95

* Average R^2 of 30 runs (the red labelled results in Figs. 6, 9, 12). + Mean square error. **Confidence level (%) of 30 runs. † Confidence intervals of 30 runs.

ANN models in the statistical point of view (as the clarified example can be seen in Fig. 5). This section therefore elaborates some results showing how the highest correlation coefficient may actually be misleading for validation level of the ANN models (as showed in Figs. 13–15).

Fig. 13 shows the highest correlation coefficients of HHV prediction with the ANN models developed by combined UA-PA datasets. The highest correlation coefficients were observed as 0.933 and 0.992 for tansig train and test, 0.944 and 0.982 for logsig train and test, and 0.927 and 0.977 for poslin train and test, respectively. Figs. 14 and 15 present the highest correlation coefficients of HHV prediction with the ANN models developed by UA and PA datasets, respectively. In the ANN model developed by UA dataset, the highest correlation coefficients were observed as ~0.94 for tansig, ~0.90 for logsig, and ~0.88 for poslin activation functions, which are slightly higher than the correlation coefficients observed by the ANN models developed by PA (~0.86 for tansig, ~0.83 for logsig, and ~0.82 for poslin activation functions, Fig. 15).

Regardless of dataset, the order of the importance of the activation function is similar to the average correlation coefficients presented in Section 3.1–3.3 as tansig > logsig > poslin. However, the level of correlation coefficients presented in this section are relatively higher than the average correlation coefficient presented in Table 2. The highest correlation coefficients are therefore clearly out of the confidence level of ANN models and misguiding the validity of these models and prediction results, as the highest correlation coefficient is only one of the many randomise runs (one of the 30 runs in this study). However, once the models were trained with randomly selected train and test datasets in a dataset group for each run, the models provide a range of prediction data for these dataset groups. From the statistical point of view, it is possible to do an adequate justification whether ANN models be applicable for the prediction of HHV with the average correlation coefficients in a range of confidence levels.

3.6. General summary of ANN models

As a summary of this research, a clear state-of-the-art ANN model was developed for the prediction of HHV of biomass feedstocks using UA and/or PA datasets. In order to provide a comparative qualitative and quantitative analysis, thirteen different algorithms (lm, br, scg, bfg, cgb, cgf, cgp, gd, gda, gdm, gdx, oss, rp), four different activation functions (logsig, tansig, poslin, purelin) with a wide range of hidden layer (3–15) for training and testing the ANN model were investigated to predict the HHV of the biomass feedstocks characterised by the PA and UA datasets. The research shows that a better ANN model for the prediction of HHV of biomass feedstocks could be build based on the following decisions.

- **Dataset**: Combine UAPA (eight inputs) > UA (five inputs) > PA (three inputs)
- **Data points**: (train and test) must be seleced randomly to generelise the model
- **Activation function selection**: tansig > logsig > poslin ≫ purelin
- **Algorithm selection:** "lm" and "br" algorithms provide better HHV prediction
- **Hidden layers**: High hidden layers (13–15) for "lm" and lower hidden layers (5–9) for "br" algorithms

Using this comprehensive ANN findings, it is possible to build ANN models providing better prediction for a specific group of biomass feedstocks with the minimum dataset or for a wide range of biomass feedstocks with extensive datasets of combined UA-PA, UA and/or PA of biomass feedstocks. These experimental results have relied on the collected dataset which contains 100 samples of UA and PA. In the future, we will seek to expand the dataset by including different biomass as well as further machine learning models explorations. In particular, as the ANNs are known as black-box models, it poses a challenge to interpret results. Thus, while the ANN deliver good performance, providing an insight into the decision process is also an important asset for understanding the relation between inputs and outputs, such as the level of effect of "FC" on the given output "HHV".

As the initial findings emphasize that the zero centred sigmoidal activation functions tend to produce better results in comparison to the linear activation functions which might be an indication of the necessity of having negative values in the output of each neuron and the range from −1 to 1. Even though these sigmoidal functions may pose risk for the well-known vanishing gradient problems, the used ANN structures have no multiple hidden layers which naturally avoid such issues. Furthermore, the experiment results show the number of hidden layers has no critical effect on the produced correlation coefficient values. Yet, the phenomenon of "the more data the better" model has been followed in the dataset selection process where the more inputs (UA-PA with 8 neurons in the input layers) led to more accurate predictions.

4. Conclusions

In this study, a detailed state-of-the-art ANN model was developed to produce a comprehensive understanding to predict biomass feedstocks HHV basing on PA and UA datasets of different biomass feedstocks. ANN models trained by the combination of ultimate-proximate analyses datasets provided better prediction than the other ANN models trained separately by UA or PA datasets. Regardless of the dataset, sigmoidal activation functions provide better prediction results than linear activation function as tansig > logsig > poslin ≫ purelin. The highest average correlation coefficients were observed with the Levenberg-Marquardt (lm) and Bayesian Regularisation (br) algorithms regardless of activation function. While "lm" provides better results with a high

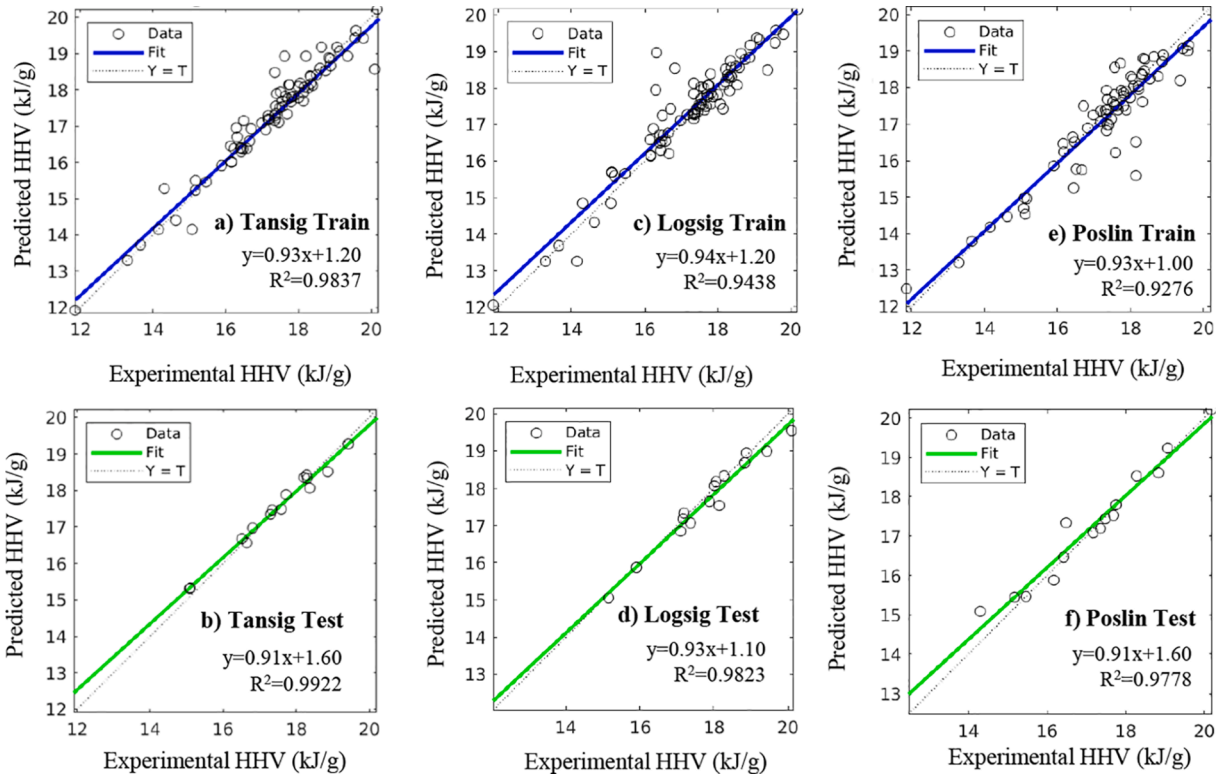

Fig. 13. The highest correlation coefficients of HHV prediction with the ANN model developed by combined ultimate-proximate analyses datasets (one of the highest training and testing results in 30 run).

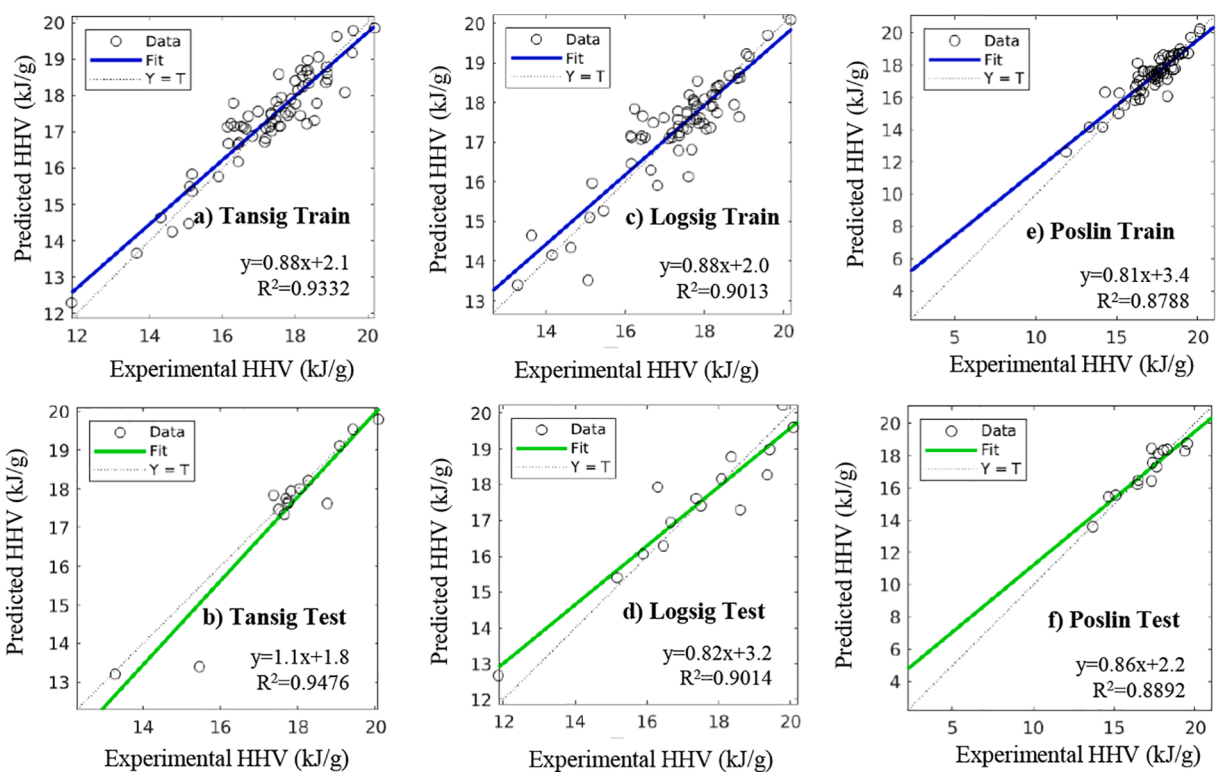

Fig. 14. The highest correlation coefficients of HHV prediction with the ANN model developed by ultimate analysis datasets (one of the highest training and testing results in 30 runs).

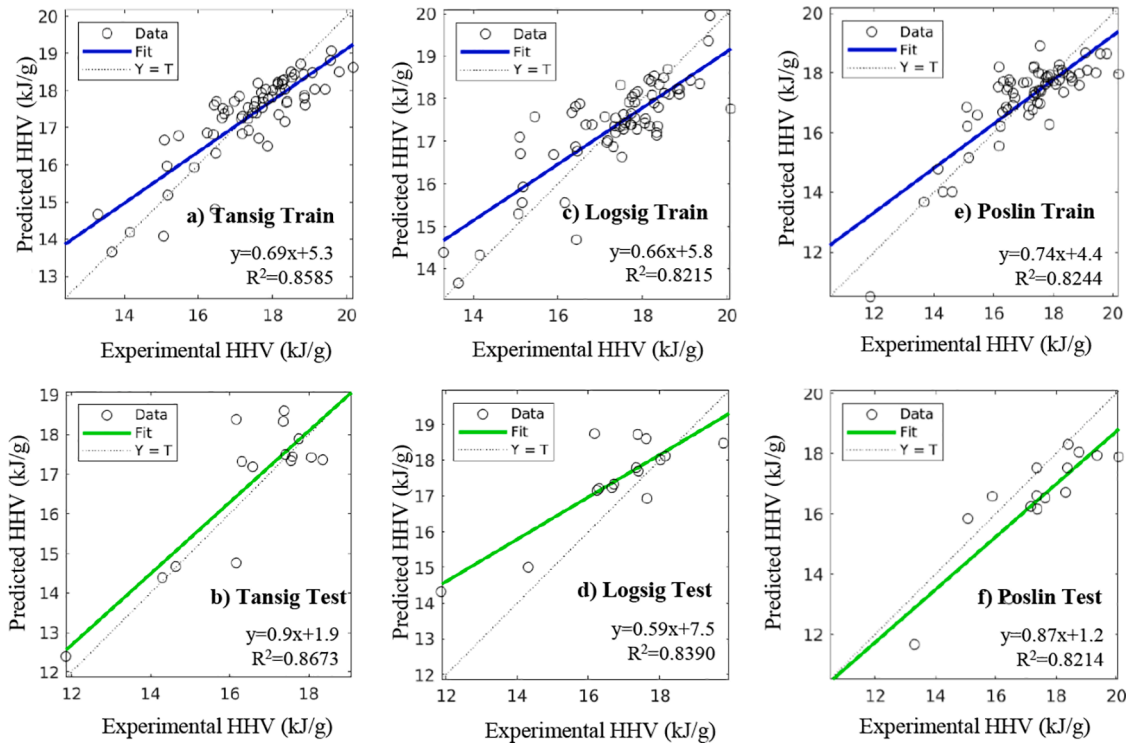

Fig. 15. The highest correlation coefficients of HHV prediction with the ANN model developed by proximate analysis datasets (one of the highest training and testing results in 30 runs).

hidden layer (13–15) in the models developed via low number inputs (three inputs of PA), "br" shows the better results with a lower hidden layer (5–9) in the models developed via high number inputs (eight inputs of ultimate-proximate analyses). The best average correlation coefficients provided by tansig for train and test are 0.848 and 0.831 by PA dataset (with "lm" and hidden layers of 14–15), 0.873 and 0.835 by UA dataset (with "lm" and hidden layers of 14–15), and 0.962 and 0.876 by the combination of ultimate-proximate analyses datasets (with "br" and "lm" and hidden layers of 9 and 15). Although the highest correlation coefficients were observed as 0.992 for tansig and 0.982 for logsig in the ANN models developed by the combination of ultimate-PA, the average correlation coefficients (of 30 randomised run) demonstrate the real ability of ANN models for the prediction of HHV. As the best average correlation coefficients (for train and test) were observed with tansig as 0.962 and 0.876 for the ANN model developed by the combination of ultimate-proximate analyses dataset with a relatively high confidence level of ~96% for training and ~92% for testing. This study demonstrated how important to randomise the train and test data selection and present the results within a confidence level for understanding the real potential of ANN models in the application of HHV prediction. Furthermore, this study provides a clear understanding on how the activation functions, algorithms, hidden layers, dataset, and randomisation of the dataset effects the prediction of HHV of biomass feedstocks for the further studies.

CRediT authorship contribution statement

Fatih Güleç: Conceptualization, Formal analysis, Investigation, Validation, Visualization, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. **Direnc Pekaslan:** Conceptualization, Formal analysis, Investigation, Validation, Visualization, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. **Orla Williams:** Conceptualization, Methodology, Writing – review & editing. **Edward Lester:** Conceptualization, Methodology, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded and supported by the EPSRC, BBSRC and UK Supergen Bioenergy Hub [Grant number EP/S000771/1].

References

[1] Kumar M, Oyedun AO, Kumar A. A review on the current status of various hydrothermal technologies on biomass feedstock. Renew Sustain Energy Rev 2018; 81:1742–70.
[2] Kostas ET, Williams OSA, Duran-Jimenez G, Tapper AJ, Cooper M, Meehan R, et al. Microwave pyrolysis of Laminaria digitata to produce unique seaweed-derived bio-oils. Biomass Bioenergy 2019;125:41–9.
[3] Daioglou V, Doelman JC, Wicke B, Faaij A, van Vuuren DP. Integrated assessment of biomass supply and demand in climate change mitigation scenarios. Global Environ Change 2019;54:88–101.
[4] Lester E, Avila C, Pang CH, Williams O, Perkins J, Gaddipatti S, et al. A proposed biomass char classification system. Fuel 2018;232:845–54.
[5] Shen Y. A review on hydrothermal carbonization of biomass and plastic wastes to energy products. Biomass Bioenergy 2020;134:105479.
[6] Sharma R, Jasrotia K, Singh N, Ghosh P, srivastava S, Sharma NR, et al. A comprehensive review on hydrothermal carbonization of biomass and its applications. Chem Africa 2020;3(1):1–19.
[7] Vassilev SV, Vassileva CG, Vassilev VS. Advantages and disadvantages of composition and properties of biomass in comparison with coal: an overview. Fuel 2015;158:330–50.
[8] Demirbas A. Combustion characteristics of different biomass fuels. Prog Energy Combust Sci 2004;30(2):219–30.
[9] Guo M, Song W. The growing US bioeconomy: drivers, development and constraints. New Biotechnol 2019;49:48–57.

[10] Güleç F, Şimşek EH, Sarı HT. Prediction of biomass pyrolysis mechanisms and kinetics–application of Kalman filter. Chem Eng Technol 2021.
[11] Ruiz HA, Rodríguez-Jasso RM, Fernandes BD, Vicente AA, Teixeira JA. Hydrothermal processing, as an alternative for upgrading agriculture residues and marine biomass according to the biorefinery concept: a review. Renew Sustain Energy Rev 2013;21:35–51.
[12] Güleç F, Özdemir G. Investigation of drying characteristics of Cherry Laurel (Laurocerasus officinalis Roemer) fruits. Akademik ziraat dergisi 2017;6(1):73–80.
[13] Şimşek EH, Güleç F, Akçadağ FS. Understanding the liquefaction mechanism of Beypazarı lignite in tetralin with ultraviolet irradiation using discrete time models. Fuel Process Technol 2020;198:106227.
[14] Elliott DC, Biller P, Ross AB, Schmidt AJ, Jones SB. Hydrothermal liquefaction of biomass: developments from batch to continuous process. Bioresour Technol 2015;178:147–56.
[15] Heidari M, Norouzi O, Salaudeen S, Acharya B, Dutta A. Prediction of hydrothermal carbonization with respect to the biomass components and severity factor. Energy Fuels 2019;33(10):9916–24.
[16] Arellanoa O, Flores M, Guerra J, Hidalgo A, Rojas D, Strubinger A. Hydrothermal carbonization (HTC) of corncob and characterization of the obtained hydrochar. Chem Eng 2016;50.
[17] Cherad R, Onwudili J, Biller P, Williams P, Ross A. Hydrogen production from the catalytic supercritical water gasification of process water generated from hydrothermal liquefaction of microalgae. Fuel 2016;166:24–8.
[18] Abdoulmoumine N, Adhikari S, Kulkarni A, Chattanathan S. A review on biomass gasification syngas cleanup. Appl Energy 2015;155:294–307.
[19] Dimitriadis A, Bezergianni S. Hydrothermal liquefaction of various biomass and waste feedstocks for biocrude production: a state of the art review. Renew Sustain Energy Rev 2017;68:113–25.
[20] Jakšić OM, Jakšić Z, Guha K, Silva AG. Comparing artificial neural network algorithms for prediction of higher heating value for different types of biomass. 2021.
[21] Çakman G, Gheni S, Ceylan S. Prediction of higher heating value of biochars using proximate analysis by artificial neural network. Biomass Convers Biorefin 2021:1–9.
[22] Pattanayak S, Loha C, Hauchhum L, Sailo L. Application of MLP-ANN models for estimating the higher heating value of bamboo biomass. Biomass Convers Biorefin 2020:1–10.
[23] Ighalo JO, Adeniyi AG, Marques G. Application of artificial neural networks in predicting biomass higher heating value: an early appraisal. Energy Sources Part A 2020:1–8.
[24] Friedl A, Padouvas E, Rotter H, Varmuza K. Prediction of heating values of biomass fuel from elemental composition. Anal Chim Acta 2005;544(1–2):191–8.
[25] Uzun H, Yıldız Z, Goldfarb JL, Ceylan S. Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. Bioresour Technol 2017;234:122–30.
[26] Ceylan Z, Pekel E, Ceylan S, Bulkan S. Biomass higher heating value prediction analysis by ANFIS, PSO-ANFIS and GA-ANFIS. Global Nest J 2018;20(3):589–97.
[27] Dashti A, Noushabadi AS, Raji M, Razmi A, Ceylan S, Mohammadi AH. Estimation of biomass higher heating value (HHV) based on the proximate analysis: smart modeling and correlation. Fuel 2019;257:115931.
[28] Hosseinpour S, Aghbashlo M, Tabatabaei M, Mehrpooya M. Estimation of biomass higher heating value (HHV) based on the proximate analysis by using iterative neural network-adapted partial least squares (INNPLS). Energy 2017;138:473–9.
[29] Ghugare S, Tiwary S, Elangovan V, Tambe S. Prediction of higher heating value of solid biomass fuels using artificial intelligence formalisms. Bioenergy Res 2014;7(2):681–92.
[30] Samadi SH, Ghobadian B, Nosrati M. Prediction of higher heating value of biomass materials based on proximate analysis using gradient boosted regression trees method. Energy Sources Part A 2021;43(6):672–81.
[31] King TN, Attwood DH. Predicting specific energies for brown coal samples. Fuel; (United Kingdom) 1980;59(8):602–3.
[32] Ringen S, Lanum J, Miknis FP. Calculating heating values from elemental compositions of fossil fuels. Fuel; (United Kingdom) 1979;58(1):69–71.
[33] Tan P, Zhang C, Xia J, Fang Q-Y, Chen G. Estimation of higher heating value of coal based on proximate analysis using support vector regression. Fuel Process Technol 2015;138:298–304.
[34] Güleç F, Meredith W, Sun C-G, Snape CE. Selective low temperature chemical looping combustion of higher alkanes with Cu-and Mn-oxides. Energy 2019;173:658–66.
[35] Górnicki K, Kaleta A, Winiczenko R, Jewiarz M. Prediction of higher heating value of oat grain and straw biomass. E3S Web Conf 2020;154:01003.
[36] Kathiravale S, Yunus MNM, Sopian K, Samsuddin A, Rahman R. Modeling the heating value of Municipal Solid Waste. Fuel 2003;82(9):1119–25.
[37] Khan MA, Abu-Ghararah ZH. New approach for estimating energy content of municipal solid waste. J Environ Eng 1991;117(3):376–80.
[38] Gulec F, Simsek EH, Karaduman A. Disproportionation kinetics of 2-methylnaphthalene in the presence of Zr/ZSM5 zeolite catalysts. J Faculty Eng Archit Gazi Univ 2016;31(3):610–21.
[39] Yin C-Y. Prediction of higher heating values of biomass from proximate and ultimate analyses. Fuel 2011;90(3):1128–32.
[40] Demirbaş A, Demirbaş AH. Estimating the calorific values of lignocellulosic fuels. Energy Explor Exploit 2004;22(2):135–43.
[41] Akkaya E. ANFIS based prediction model for biomass heating value using proximate analysis components. Fuel 2016;180:687–93.
[42] Song X, Mitnitski A, Cox J, Rockwood K. Comparison of machine learning techniques with classical statistical models in predicting health outcomes. *MEDINFO 2004*. IOS Press; 2004:736-40.
[43] Gevrey M, Dimopoulos I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model 2003;160(3):249–64.
[44] Dobbelaere MR, Plehiers PP, Van de Vijver R, Stevens CV, Van Geem KM. Machine learning in chemical engineering: strengths, weaknesses, opportunities, and threats. Engineering 2021;7(9):1201–11.
[45] Eftekhar B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. BMC Med Inf Decis Making 2005;5(1):1–8.
[46] Moraes R, Valiati JF, Neto WPG. Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Syst Appl 2013;40(2):621–33.
[47] Hu Y, Zhang S, Yu J, Liu J, Zheng S. SELDI-TOF-MS: the proteomics and bioinformatics approaches in the diagnosis of breast cancer. The Breast 2005;14(4):250–5.
[48] Kiran TR, Rajput S. An effectiveness model for an indirect evaporative cooling (IEC) system: Comparison of artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS) and fuzzy inference system (FIS) approach. Appl Soft Comput 2011;11(4):3525–33.
[49] Şahin M, Kaya Y, Uyar M. Comparison of ANN and MLR models for estimating solar radiation in Turkey using NOAA/AVHRR data. Adv Space Res 2013;51(5):891–904.
[50] Aladejare AE, Onifade M, Lawal AI. Application of metaheuristic based artificial neural network and multilinear regression for the prediction of higher heating values of fuels. Int J Coal Prepar Utiliz 2020:1–22.
[51] Khalida B, Mohamed Z, Belaid S, Samir HO, Sobhi K, Midane S. Prediction of higher heating value HHV of date palm biomass fuel using artificial intelligence method. In: *2019 8th International Conference on Renewable Energy Research and Applications (ICRERA)*. IEEE; 2019. p. 59–62.
[52] Abdulsalam J, Lawal AI, Setsepu RL, Onifade M, Bada S. Application of gene expression programming, artificial neural network and multilinear regression in predicting hydrochar physicochemical properties. Bioresour Bioprocess 2020;7(1):1–22.
[53] Xing J, Luo K, Wang H, Gao Z, Fan J. A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. Energy 2019;188:116077.
[54] García R, Pizarro C, Lavín AG, Bueno JL. Spanish biofuels heating value estimation. Part I: Ultimate analysis data. Fuel 2014;117:1130–8.
[55] García R, Pizarro C, Lavín AG, Bueno JL. Spanish biofuels heating value estimation. Part II: Proximate analysis data. Fuel 2014;117:1139–47.
[56] Livingstone DJ. Artificial neural networks: methods and applications. Springer; 2008.
[57] Gurney K. An introduction to neural networks. CRC Press; 2018.
[58] Kim P. In: MATLAB deep learning. Berkeley, CA: Apress; 2017. p. 1–18.
[59] Paluszek M, Thomas S. Practical Matlab deep learning: a project-based approach. first ed. Berkeley, CA: Apress; 2020.
[60] Güleç F, Riesco LMG, Williams O, Kostas ET, Samson A, Lester E. Hydrothermal conversion of different lignocellulosic biomass feedstocks–effect of the process conditions on hydrochar structures. Fuel 2021;302:121166.
[61] Güleç F, Meredith W, Sun C-G, Snape CE. Demonstrating the applicability of chemical looping combustion for the regeneration of fluid catalytic cracking catalysts. Chem Eng J 2020;389:124492.
[62] Dai Z, Chen Z, Selmi A, Jermsittiparsert K, Denić NM, Nešić Z. Machine learning prediction of higher heating value of biomass. Biomass Convers Biorefin 2021:1–9.
[63] Keybondorian E, Zanbouri H, Bemani A, Hamule T. Application of MLP-ANN strategy to predict higher heating value of biomass in terms of proximate analysis. Energy Sources Part A 2017;39(22):2105–11.