


RESEARCH

Open Access



Application of gene expression programming, artificial neural network and multilinear regression in predicting hydrochar physicochemical properties

Jibril Abdulsalam^{1*} , Abiodun Ismail Lawal^{2,3}, Ramadimetja Lizah Setsepu¹, Moshood Onifade⁴ and Samson Bada¹

Abstract

Globally, the provision of energy is becoming an absolute necessity. Biomass resources are abundant and have been described as a potential alternative source of energy. However, it is important to assess the fuel characteristics of the various available biomass sources. Soft computing techniques are presented in this study to predict the mass yield (MY), energy yield (EY), and higher heating value (HHV) of hydrothermally carbonized biomass using Gene Expression Programming (GEP), multiple-input single output-artificial neural network (MISO-ANN), and Multilinear regression (MLR). The three techniques were compared using statistical performance metrics. The coefficient of determination (R^2), mean absolute error (MAE) and mean bias error (MBE) were used to evaluate the performance of the models. The MISO-ANN with 5-10 to 10-1 and 5-15-15-1 network architectures provided the most satisfactory performance of the three proposed models ($R^2 = 0.976, 0.955, 0.996$; MAE = 2.24, 2.11, 0.93; MBE = 0.16, 0.37, 0.12) for MY, EY and HHV, respectively. The GEP technique's ability to predict hydrochar properties based on the input parameters was found to be satisfactory, while MLR provided an unsatisfactory predictive model. Sensitivity analysis was conducted, and the analysis revealed that volatile matter (VM) and temperature (Temp) have more influence on the MY, EY, and HHV.

Keywords: Artificial neural network, Biomass, Gene expression programming, Higher heating value, Hydrochars, Hydrothermal carbonization

Introduction

The increasing energy demand has led to the need to find alternative energy sources that are affordable, widely available, and environmentally friendly. Biomass is a biological and sustainable material originated from plants and animals, along with their waste and residues (Krylova and Zaitchenko 2018). Biomass is the most available renewable energy source, with a contribution

of about 50% of the total global renewable energy as of 2018, and providing energy to billions of people and stimulating economic growth (Pradhan et al. 2018). The studies by (Tekin et al. 2014) and (Rousset et al. 2012) reported that biomass is a potential alternative renewable energy source for power generation as a result of its low emissions, low ash, and total sulphur content. (Saba et al. 2017) and (Perlack et al. 2011) reported that biomass greenhouse gas emission status is zero to net negative as carbon dioxide is absorbed by plants during photosynthesis.

The most generally used thermochemical pre-treatment techniques include pyrolysis, gasification, torrefaction, and hydrothermal carbonization (HTC) (Wang et al.

*Correspondence: jibril.abdulsalam@wits.ac.za

¹ DSI/NRF Clean Coal Technology Research Group, School of Chemical and Metallurgical Engineering, Faculty of Engineering and the Built Environment, University of the Witwatersrand, WITS 2050, Johannesburg, South Africa

Full list of author information is available at the end of the article

2018; Kambo and Dutta 2015). The studies by (Kubacki et al. 2012) and (Makwavela et al. 2017) stated that the co-combustion of biomass and coal allowed coal to ignite and burnout at lower temperatures because of the interactions with the early combustion of biomass volatile matter. The study concluded that the emission reductions reported were due to an improved reaction between coal and biomass volatiles in a hot oxidizing atmosphere. A number of studies have been carried out on the pre-treatment of biomass by various researchers (Safarian et al. 2019; Zhang and Pang 2019; Kambo and Dutta 2015). The type of feedstock and the preferred end product determines the type of pre-treatment method to be used (Kambo and Dutta 2015).

The hydrochar utilized in the study was produced using the HTC method, which is generally considered to be a more effective technique (Danso-Boateng 2015). Biomass for HTC treatment does not require drying before treatment and thus uses less energy. In fact, unlike the conventional biological treatment technique, the presence of toxic compounds in the biomass does not affect HTC. HTC treatment typically takes place at relatively low temperatures (180–260 °C) and under internally generated pressure from an enclosed reactor, which decreases the oxygen and hydrogen content of the starting material by dehydration and decarboxylation (Libra et al. 2011). The HTC treatment converts the wet biomass into a hydrochar, a solid substance with improved carbon content. Hydrochar has a heating value higher than the feedstock and a chemical structure similar to that of coal (Mumme et al. 2011). The process is controlled by process parameters such as temperature and residence time, which define the intensity of biomass treatment (Wiedner et al. 2013; Xu et al. 2013).

The temperature has a significant influence on the HTC process. It is the key determinant of the water properties, which leads to ionic reactions in the subcritical region. A rise in temperature alters the viscosity of the water, making it easier to penetrate the pores of the material and thus further degrade the biomass (Funke and Ziegler 2010). With an increase in temperature, the disintegration of solid residues increases, and this further leads to an increase in the yield of solids to gas products. In most of the studies reviewed (Wang et al. 2018; Kim et al. 2014; Parshetti et al. 2013; Hwang et al. 2012; Sevilla and Fuertes 2009), an increase in temperature has been reported to result in lower mass yields, with an increase in HHV of hydrochars, suitable for power generation. A study conducted by (Wang et al. 2018) also reported on the significance of residence time on the severity of the HTC process. The residence time at a given temperature influences the degree of decomposition of the feedstock,

but the minimum impact on hydrochar mass yield compared to temperature.

The investigation conducted by (Zhu et al. 2018) at different temperatures and residence time has shown that the above-mentioned parameters do influence the properties of cornstalk hydrochar. The author reported that the fuel mass yield decreased from 70.57 to 33.40% with an increase in temperature. While the residence time tends to have a lower influence on the mass yield relative to the temperature. The energy content of the raw cornstalk was increased from 16.35 to 26.31 MJ/kg. A similar result was also observed for the HTC treatment of other biomass such as biogas sludge, barley, and maize silage, starch, municipal solid waste, and sewage sludge (Seyed-sadr et al. 2018; Kim et al. 2014; Parshetti et al. 2013; Hwang et al. 2012; Sevilla and Fuertes 2009). Therefore, it can be concluded that temperature and residence time do influence the properties of the hydrochar as well as the raw biomass type. With the understanding of the relationship between temperature and residence time on the mass yield of hydrochar, the physiochemical properties of the hydrochar under the set conditions were used to predict its yield theoretically.

Empirical and semi-empirical correlations have been reported in the literature for the estimation of biomass fuel HHV based on their proximate, ultimate, and chemical analyses (Saldarriaga et al. 2015; Saidur et al. 2011; Sheng and Azevedo 2005). (Vargas-Moreno et al. 2012) reported on mathematical models used to predict biomass HHV and assessed the performances of the prediction models. The study reported that the R^2 remained as high as 0.748 for biomass in the 15 univariate and multivariate prediction equations reviewed. Artificial neural networks and, in particular, feed-forward artificial neural networks (FANNs) have been widely used to develop process models over the last 10 years, and their use in industry has evolved rapidly (Onifade et al. 2019; Aladejare et al. 2020; Majumder et al. 2008; Hansen and Meservy 1996; Wasserman 1993). (Aladejare et al. 2020) and (Estiati et al. 2016) used neural networks and regression analysis to predict the HHV of coal and biomass-based fuels from their proximate and ultimate analyses using both experimental and existing data from the literature. The results obtained by these authors show that the adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network–particle swarm optimization (ANN-PSO) models perform better than the MLR models as reflected in the statistical analysis conducted to assess the performance of the models.

There is limited study in the literature that assesses the influence of temperature, residence time, and the composition of biomass sources on hydrochar properties. The aim of this study is, therefore, to predict the mass

yield, energy yield, and higher heating value of hydrochars using the HTC process conditions (temperature and residence time), and the biomass proximate analysis results. The experimental data from this study and the data obtained from the literature were utilized in the linear and non-linear empirical models proposed to predict these properties for different biomass sources (as provided in Table 1). The performance of the proposed models was compared using the R^2 , mean absolute error, and mean bias errors.

Experimental method and models overview

Sample characterization and data generation

To develop the proposed models, data from the proximate analysis, HTC process conditions (temperature and residence time), hydrochar properties (MY, EY, and HHV) relating to a number of biomass species were used. The woody biomass (*Searsia lancea*) used in this study were harvested from a phytoremediation trial site polluted with groundwater from gold and the uranium tailings dam at AngloGold Ashanti Limited' West Wits and Vaal River mining operations in South Africa. The different biomass components were milled in a Retsch SM 200 mill to -1 mm and -212 μ m size fractions. The -1 mm fraction was used for the hydrothermal carbonization and the -212 μ m fraction for the physicochemical characterization. The proximate analysis for these samples was performed based on the ASTM D5142, with approximately 1 g used to calculate the fixed carbon, moisture, ash, and volatile matter contents. The fixed carbon is expressed as the subtraction of the sum of moisture, volatile matter, and ash contents from 100%. The bomb calorimeter (Leco AC500), in accordance with the ASTM

articles on several biomass species were used to obtain predictive models. The summary of the statistics of the dataset obtained from the experimental tests and the literature is presented in Table 1 and the details of the data used in the model development are presented in Table 2. Summary of the statistics in Table 1 shows that volatile matter, Ash content, fixed carbon, and residence time do not follow normal distributions based on their respective skewness. To enable the general application of the proposed models, the data set was trained, tested, and validated using GEP, MISO-ANN, and MLR and compared with one another.

Hydrothermal carbonization

The woody *Searsia lancea* tree species was carbonized in a laboratory-scale high-pressure Berghof BR-1500 reactor. For each experiment, the reactor was loaded with 100 g of air-dried sample and 800 ml of deionized water, with the reactor pressurized at 20 bar using nitrogen. The hydrothermal test was conducted at different reaction temperatures of 200, 250, and 280 °C and residence time of 30, 60, and 90 min. The mixture was stirred with the reactor agitated at 200 rpm and was sustained for the entire experiment. After the holding time, the reactor was allowed to cool to room temperature. The solid hydrochar was collected via filtration and allowed to dry in an oven at 105 °C for 24 h. The results from the nine (9) hydrochar samples produced under the set conditions are depicted in Table 2. The mass yield and energy yield of hydrochars were calculated using the following equations:

$$MY = \frac{M_{HC}}{M_R} \times 100\%, \quad (1)$$

where MY is mass yield, M_{HC} is the mass of hydrochar and M_R is the mass of the raw sample:

$$EY = \left(MY \times \frac{HHV_{HC}}{HHV_R} \right) \times 100\%, \quad (2)$$

where EY is energy yield, HHV_{HC} and HHV_R is the higher heating value of the hydrochars and raw samples, respectively.

Overview of GEP and ANN

Gene expression programming (GEP)

GEP is an evolutionary-based algorithm that explores the genotype from the genetic algorithm (GA) and phenotype from genetic programming (GP). Like a living organism, the GEP utilizes a simple chromosome with a fixed length for keeping and transmitting genetic information and complex tree structures for learning and adaptation by changing size, shapes, and composition.

Table 1 Statistics of the inputs and output data used

Input variables	Min	Max	Std	Skewness	Kurtosis
VM (%)	34.3	89.74	12.69501	− 1.05705	0.90247
Ash (%)	0.25	49.85	14.04995	1.20789	0.72747
FC (%)	2.66	54	8.38191	2.87274	12.34524
Temperature (°C)	140	300	33.71556	0.08694	− 0.63807
RT (min)	5	480	81.73551	1.62185	3.68928
Output variables					
Mass Yield (%)	26.7	99.2	15.89677	0.46028	0.10385
Energy Yield (%)	37.11	112.55	14.4083	0.22913	0.08753
HHV (MJ/kg)	9.8	34.53	5.29801	− 0.24796	− 0.23794

VM Volatile matter, FC fixed carbon, HHV high heating value, RT residence time

D5865-04 standard, was used to estimate the HHV of the samples.

One hundred and fifteen (115) data points; 9 from the experimental investigation and 106 from published

Table 2 Experimental data and data from Literature

S/N	Samples	VM (%)	Ash (%)	FC (%)	Temp (°C)	RT (Min)	MY (%)	EY (%)	HHV (MJ/kg)	References
1.	Searsia lancea	75.67	4.26	20.07	200	30	67.25	79.12	20.27	Present study
2.		75.67	4.26	20.07	200	60	61.91	83.11	23.13	
3.		75.67	4.26	20.07	200	90	61.78	84.12	23.46	
4.		75.67	4.26	20.07	250	30	49.49	79.13	27.55	
5.		75.67	4.26	20.07	250	60	47.11	76.31	27.91	
6.		75.67	4.26	20.07	250	90	46.36	75.15	27.93	
7.		75.67	4.26	20.07	280	30	35.61	59.87	28.97	
8.		75.67	4.26	20.07	280	60	34.89	59.43	29.35	
9.		75.67	4.26	20.07	280	90	41.66	71.84	29.71	
10.	Wanlip sewage sludge	68.56	27.54	3.9	140	240	74.04	72.75	21.45	(Danso-Boateng 2015)
11.		68.56	27.54	3.9	170	240	66.17	66.75	22.02	
12.		68.56	27.54	3.9	190	240	61.27	62.4	22.65	
13.		68.56	27.54	3.9	200	240	60.17	62.15	23.13	
14.	Seaweed	63.29	24.28	12.43	180	30	46.9	53.09	13.81	(Patel 2019)
15.		63.29	24.28	12.43	180	60	45.65	58.97	15.76	
16.		63.29	24.28	12.43	180	120	45.3	61.75	16.63	
17.		63.29	24.28	12.43	200	30	45.95	61.73	16.39	
18.		63.29	24.28	12.43	200	60	41.8	60.23	17.58	
19.		63.29	24.28	12.43	200	120	39.7	58.96	18.12	
20.		63.29	24.28	12.43	220	30	36.1	53.11	17.55	
21.		63.29	24.28	12.43	220	60	34.65	53.51	18.27	
22.		63.29	24.28	12.43	220	120	30.5	49.65	18.93	

Table 2 (continued)

S/N	Samples	VM (%)	Ash (%)	FC (%)	Temp (°C)	RT (Min)	MY (%)	EY (%)	HHV (MJ/kg)	References
23.	Coffee cherry parchment	76.31	6.80	16.88	180	180	55	66	21.83	(Sliakova 2018)
24.		76.31	6.80	16.88	200	180	54	69	23.12	
25.		76.31	6.80	16.88	220	180	46	63	24.92	
26.		76.31	6.80	16.88	240	180	34	54	28.83	
27.	Coffee wood	82.70	1.83	15.47	180	180	68	75	21.93	
28.		82.70	1.83	15.47	200	180	66	75	22.53	
29.		82.70	1.83	15.47	220	180	57	72	25.09	
30.		82.70	1.83	15.47	240	180	46	60	26.02	
31.	Eucalyptus wood	84.96	0.25	14.79	180	180	72	78	21.4	
32.		84.96	0.25	14.79	200	180	60	74	24.07	
33.		84.96	0.25	14.79	220	180	57	72	24.77	
34.		84.96	0.25	14.79	240	180	52	66	24.97	
35.	Bamboo	83.51	1.16	15.32	180	180	62	70	22.29	
36.		83.51	1.16	15.32	200	180	54	65	23.45	
37.		83.51	1.16	15.32	220	180	53	66	24.2	
38.		83.51	1.16	15.32	240	180	52	71	26.54	
39.	Miscanthus	87.5	0.8	11.7	190	5	83.5	91.17	20.2	(Kambo and Dutta 2014)
40.		87.5	0.8	11.7	225	5	66.9	78.11	21.6	
41.		87.5	0.8	11.7	260	5	47.8	66.92	25.9	
42.		85.18	0.68	14.14	175	30	81.96	86.09	21.91	
43.	Spruce	85.18	0.68	14.14	200	10	77.63	82.48	22.16	(Bach et al. 2016)
44.		85.18	0.68	14.14	200	30	73.61	80.98	22.95	
45.		85.18	0.68	14.14	200	60	69.54	76.98	23.09	
46.		85.18	0.68	14.14	225	30	67.63	76.75	23.67	
47.	Birch	89.74	0.64	9.63	175	30	78.54	80.44	20.03	
48.		89.74	0.64	9.63	200	10	64.22	66.94	20.39	
49.		89.74	0.64	9.63	200	30	62.17	66.98	21.09	
50.		89.74	0.64	9.63	200	60	60.25	65.97	21.42	
51.		89.74	0.64	9.63	225	30	56.91	64.28	22.1	

Table 2 (continued)

S/N	Samples	VM (%)	Ash (%)	FC (%)	Temp (°C)	RT (Min)	MY (%)	EY (%)	HHV (MJ/kg)	References
52.	Sewage Sludge	34.3	11.1	54	200	30	97	100.8	29.8	(Saba et al. 2017)
53.		34.3	11.1	54	230	30	99.2	101.8	29	
54.		34.3	11.1	54	260	30	96.1	101.5	28.1	
55.	Miscanthus	84.5	1.3	12.8	200	30	68.8	71	20.4	
56.		84.5	1.3	12.8	230	30	57.2	71.2	24.6	
57.		84.5	1.3	12.8	260	30	46.2	67	28.6	
58.	Empty Fruit Bunch	77.91	4.55	17.54	220	240	56.6	72.7	27.43	(Stemann et al. 2013)
59.		77.91	4.55	17.54	220	240	58.6	77	28.03	
60.	Olive Mill Waste	72	2	26	180	180	81.70	95.06	26.32	(Gao et al. 2019)
61.		72	2	26	220	180	70.80	88.14	28.16	
62.		72	2	26	250	180	57.60	82.12	32.25	
63.	Municipal Solid Waste	76	5	20	220	120	63.90	82.46	28.35	
64.		75.6	4.9	19.5	250	180	54.9	74.22	29.7	
65.	Opuntia Ficus Indica	66.8	12.9	20.3	220	180	37.1	52.03	19.45	
66.		66.8	12.9	20.3	250	180	26.7	39.83	20.69	
67.	Gwangju sewage sludge	72.33	17.05	10.62	180	30	93.13	101.51	22.47	(Lee et al. 2019)
68.		72.33	17.05	10.62	210	30	85.84	97.86	23.42	
69.		72.33	17.05	10.62	240	30	67.93	75.40	22.86	
70.		72.33	17.05	10.62	270	30	40.78	37.11	18.64	
71.	Chlorella sp	80.73	6.67	12.6	180	30	98.23	105.11	24.73	
72.		80.73	6.67	12.6	210	30	97.03	112.55	26.68	
73.		80.73	6.67	12.6	240	30	81.3	93.50	26.55	
74.		80.73	6.67	12.6	270	30	60.59	67.86	25.82	
75.	Mixture (sewage Sludge + Chlorella sp)	74.55	14.32	11.13	180	30	93.97	102.43	23.07	
76.		74.55	14.32	11.13	210	30	87.68	100.83	24.29	
77.		74.55	14.32	11.13	240	30	64.38	73.39	24.15	
78.		74.55	14.32	11.13	270	30	45.98	46.90	21.69	

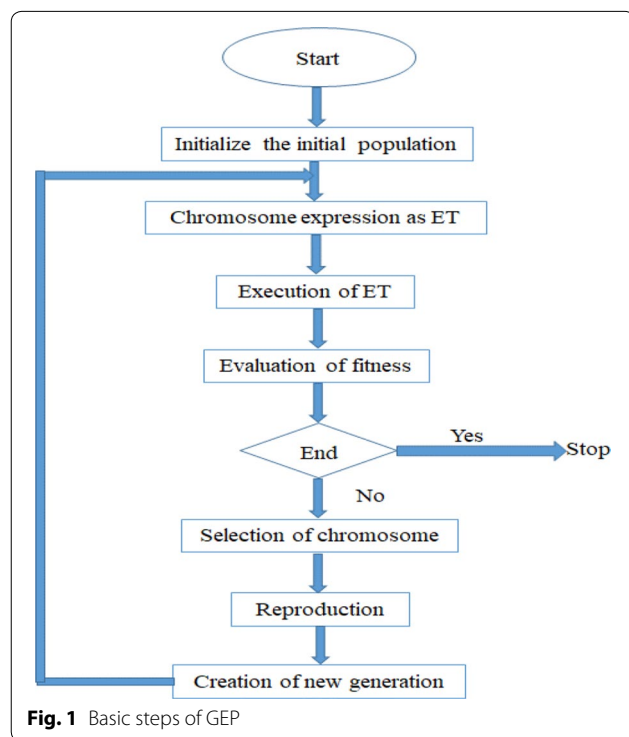
Table 2 (continued)

S/N	Samples	VM (%)	Ash (%)	FC (%)	Temp (°C)	RT (Min)	MY (%)	EY (%)	HHV (MJ/kg)	References
79.	Changsha sewage sludge	47.49	49.85	2.66	180	30	66.18	65.15	10.8	(Peng et al. 2016)
80.		47.49	49.85	2.66	220	30	60.58	55.61	10.07	
81.		47.49	49.85	2.66	260	30	56.23	60.43	11.79	
82.		47.49	49.85	2.66	300	30	53	54.26	11.23	
83.		47.49	49.85	2.66	260	60	58.51	64.32	12.06	
84.		47.49	49.85	2.66	260	90	59.37	61.59	11.38	
85.		47.49	49.85	2.66	260	360	63.32	56.57	9.8	
86.	Opuntia ficus-indica cladodes	47.49	49.85	2.66	260	480	66.19	65.53	10.86	(Volpe et al. 2018)
87.		70.62	15.15	14.23	180	30	66.05	80	16.8	
88.		70.62	15.15	14.23	180	60	64.90	81	17.31	
89.		70.62	15.15	14.23	180	180	64.21	83	17.93	
90.		70.62	15.15	14.23	220	30	61.91	84	18.82	
91.		70.62	15.15	14.23	220	60	59.35	81	18.93	
92.		70.62	15.15	14.23	220	180	59.10	83	19.48	
93.		70.62	15.15	14.23	250	30	53.19	78	20.34	
94.		70.62	15.15	14.23	250	60	52.76	80	21.03	
95.		70.62	15.15	14.23	250	180	51.42	83	22.39	

Table 2 (continued)

S/N	Samples	VM (%)	Ash (%)	FC (%)	Temp (°C)	RT (Min)	MY (%)	EY (%)	HHV (MJ/kg)	References
96.	Filamentous algae (<i>H. reticulatum</i>)	77.16	13.25	9.59	150	60	72	85.68	18.29	(Park et al. 2018)
97.		77.16	13.25	9.59	180	60	69	83.49	18.59	
98.		77.16	13.25	9.59	210	60	60	81.6	20.9	
99.		77.16	13.25	9.59	240	60	48	72.96	23.32	
100.		77.16	13.25	9.59	270	60	37	61.05	25.4	
101.	microalgae (<i>Chlorella Vulgaris</i>)	85.81	4.55	9.64	150	60	71	73.84	24.78	
102.		85.81	4.55	9.64	180	60	66	75.24	27.12	
103.		85.81	4.55	9.64	210	60	66	82.5	29.82	
104.		85.81	4.55	9.64	240	60	59	80.24	32.48	
105.		85.81	4.55	9.64	270	60	50	72.5	34.53	
106.	WWTP sewage sludge	66.87	26.06	7.07	150	60	94	98.7	17.29	
107.		66.87	26.06	7.07	180	60	91	96.46	17.48	
108.		66.87	26.06	7.07	210	60	88	97.68	18.3	
109.		66.87	26.06	7.07	240	60	67	69.68	17.16	
110.		66.87	26.06	7.07	270	60	51	51	16.41	
111.	Swine manure	55.3	36.1	8.6	200	60	67	76.22	12.4	(Xiong et al. 2019)
112.		55.3	36.1	8.6	220	60	57.4	67.41	12.8	
113.		55.3	36.1	8.6	240	60	56.6	71.14	13.7	
114.		55.3	36.1	8.6	260	60	62	90.44	15.9	
115.		55.3	36.1	8.6	280	60	50.4	73.98	16	

The key advantage of the GEP model is the ability to present its output in the form of an expression tree and a simple relationship between the model parameters and the targeted output. Unlike many optimization algorithms that require prior suggestion of the relationship between the model parameters and the output parameter. Hence, the rigors required in many optimization algorithms in establishing the model parameter combination that will give optimum results have been solved. In GA and GP, mutation and crossover operators are the common means of reproduction between them, which operate based on their respective algorithms which could increase the computational resources (Güven and Aytekin 2009; Teodorescu and Sherwood 2008). The GEP proposed by (Ferreira 2001) explores the merit of GA and GP, however, overcomes the demerits of both the GA and GP. It utilizes two entities which are the chromosomes and the expression trees. Instead of applying its operators on the expression tree directly, it operates on the chromosome which reduces the computational resources (Güven and Aytekin 2009; Teodorescu and Sherwood 2008). The flowsheet in Fig. 1 shows the steps involved in GEP.



Artificial neural networks (ANN)

ANN belongs to the family of artificial intelligence which imitates the functionality of the human brains. It explores how the human brain receives, process, and transform information. There are different types of ANNs, but the multilayer neural network is the most used. Mainly, in a supervised ANN, the input parameters are supplied with the targeted output (Jain et al. 1996). The input parameters will be multiplied with the connecting weights and their summation together with the bias will be fed into the transfer function at the hidden layer. The output of the hidden layer will be multiplied by another weight connecting the hidden layer to the output layer and its summation will be added to the bias and then fed into the transfer function at the output layer to obtain the final predicted output. The transfer function at the hidden layer is usually non-linear, while that at the output layer could be linear or non-linear. The flow chart explaining the steps involved in the ANN training is illustrated in Fig. 2.

Description of experimental data

Table 2 shows the proximate analysis of the biomass feedstock (*Searsia lancea*). The results of the proximate analysis test were presented on a moisture-free basis (dried-basis). Volatile matter, ash content, and fixed carbon are influential constituents of fuel materials used to ascertain its quality. The content of volatile matter significantly influences the process of combustion (Mierzwa-Hersztek et al. 2019; Sadiku et al. 2016). In addition, (Brewer et al. 2014) and (Holtmeyer et al. 2013) reported that material with higher volatile matter could be advantageous for combustion processes, because it is easier to ignite, lower temperature of complete burnout, and a stable flame. A high volatile matter of 75.67% was obtained for the *Searsia lancea*, making the material a potential feedstock for combustion. Ash content of 4.26% was obtained for the feedstock. (He et al. 2018) reported that, with lower ash content, there might be a decrease in fouling and slagging. The fixed carbon content of any material indicates the fuel's heating value (Sadiku et al. 2016). For our biomass feedstock, the fixed carbon content of 20.07% was obtained.

The mass yields of the hydrochars from *Searsia lancea* calculated using Eq. (1) decreases as the temperature increases at each residence time, reaching yields as low as 34.89% at 250 °C. The reduction in mass yield is a result

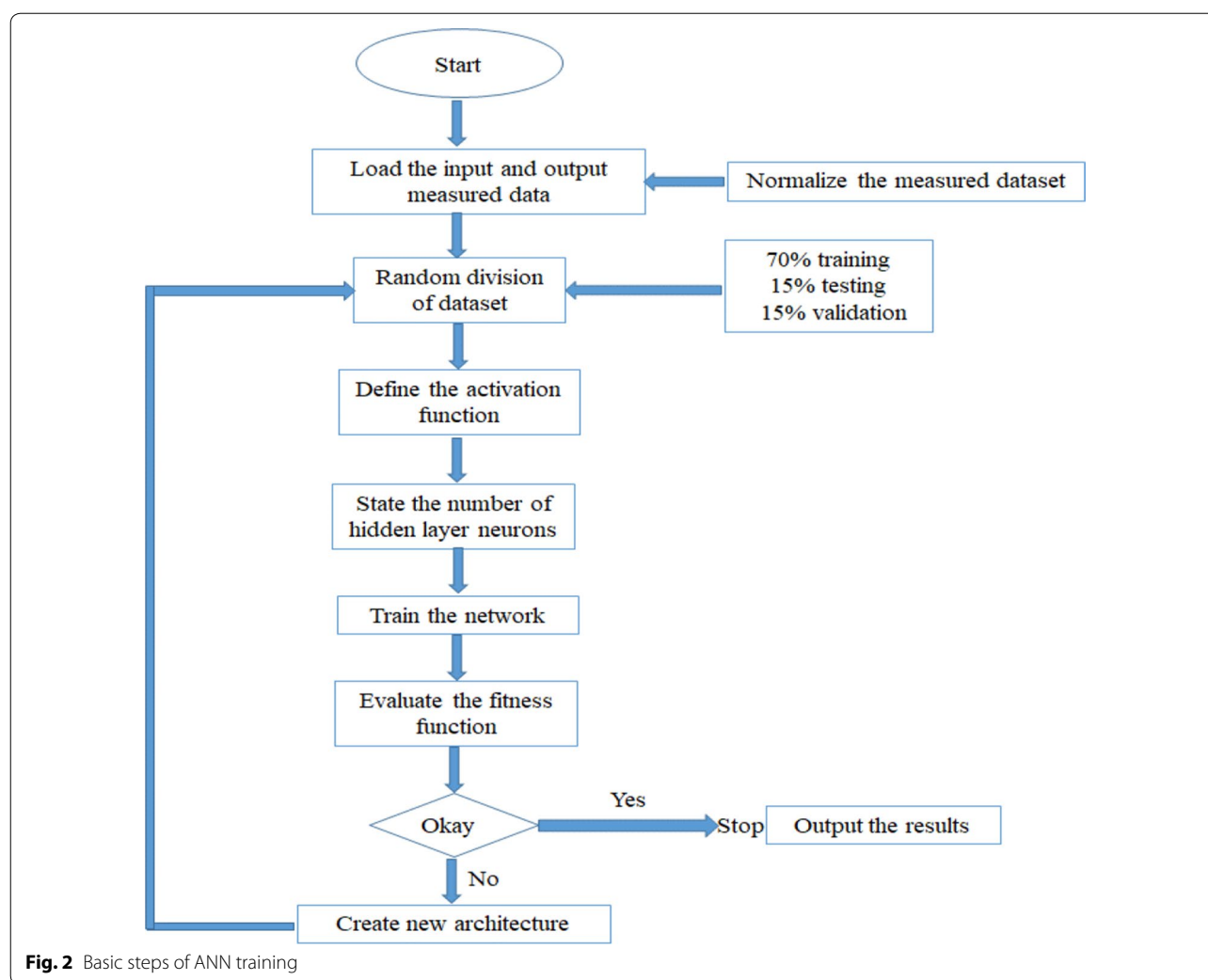


Fig. 2 Basic steps of ANN training

of decarboxylation and dehydration reactions during the HTC process resulting in the decomposition of the biomass feedstock (Saba et al. 2017; Reza et al. 2014). For each time interval, it is observed that the energy yield of the hydrochars decreases with an increase in temperature. This can be attributed to the influence of the mass yield in the calculation of the energy yield (Eq. 2). The HHV of *Searsia lancea* (17.23 MJ/kg) increased after the HTC process to as high as 29.71 MJ/kg at 280 °C and as low as 20.27 MJ/kg at 200 °C. It is observed that HHV increases with an increase in temperature and residence time.

Development of models using soft computing and regression analyses

GEP model

In developing the GEP model, the dataset used in the training and testing of the ANN model was also used. However, instead of normalizing the dataset within the range of -1 and 1 , the dataset was normalized within 0 and 1 in the GEP model. The purpose is to ensure dimensional linearity and forestall overfitting. The GEP model was implemented in GeneXproTools 5.0. After loading the data into the software, the number of chromosomes, the head size, the number of genes, and

linking function were set to 30, 8, and 5 for the respective MY and HHV, while 30, 8, and 6 were used for the EY. The linking function used for MY was Average (Avg2), while addition was used for the HHV, and Minimum (Min) was used for the EY. The fitness function adopted for the MY, EY, and HHV was the root-mean-squared error (RMSE). For the genetic operators, the

trained for the respective MY, EY, and HHV. The final expression trees obtained and the mathematical interpretation of each of the trees are presented in Figs. 3, 4, 5 and Eqs. (3), (5), and (7):

$$MY = 72.5y_5 + 26.7, \quad (3)$$

where y_5 is given in Eq. (4e):

$$y_1 = \tan^{-1} \left\{ 1 / [\max ((\text{Ash} - 0.25) / 49.6, (\text{VM} - 34.3) / 55.44) - 0.0771(\text{FC} - 2.66)] \right\} \quad (4a)$$

$$y_2 = 0.5 \left\{ y_1 + \left[1 - (0.5(-7.2301 + (\text{VM} - 34.3) / 55.44))^2 \right] \times \dots \right. \\ \left. (\min ((\text{VM} - 34.3) / 55.44, (\text{Temp} - 140) / 160) (\text{Temp} - 140) / 160 - 0.1076) \right\} \quad (4b)$$

$$y_3 = 0.5(y_2 + 3.3587) \quad (4c)$$

$$y_4 = 0.5 \left\{ y_3 + \left[1 / \left(\frac{0.5((\text{Temp} - 140) / 160 + (\text{VM} - 34.3) / 55.44) \times \dots}{((\text{VM} - 34.3) / 55.44)^{1/3} + \dots} \right) \right]^2 \right\} \quad (4d)$$

$$y_5 = 0.5 \left[y_4 - 135.4044((\text{RT} - 5)(\text{Ash} - 0.25) / 23560)^2 (\text{FC} - 2.66) / 51.34 \right] \quad (4e)$$

strategy used was the Optimal Evolution. The functional operators (e.g., addition, subtraction, division, multiplication, hyperbolic tangent, etc.) were also selected for the respective MY, EY, and HHV. The maximum fitness was used as the stop condition. The model was then

$$EY = 75.44x_6 + 37.11, \quad (5)$$

where x_6 is given in Eq. (6f):

$$x_1 = 0.7833 / (2(\text{FC} - 0.25) / 49.6 + 1.7742(\text{RT} - 5) / 475) \quad (6a)$$

$$x_2 = \min \left\{ x_1, \left[1 - 0.5 \left((\text{FC} - 5) / 475 + ((\text{VM} - 34.3) / 55.44)^2 \right) \right] \right\} \quad (6b)$$

$$x_3 = \min \left\{ x_2, \left[\tan^{-1} \left(\tan^{-1} ((\text{Ash} - 0.25) / 49.6 + 0.5(0.06505 + (\text{FC} - 2.66) / 51.34)) \right) \right]^{1/3} \right\} \quad (6c)$$

$$x_4 = \min \left\{ x_3, \left[\frac{\left(0.5((\text{RT} - 5) / 475 + (\text{Ash} - 0.25) / 49.6) - ((\text{VM} - 34.3) / 55.44)^2 \right)}{\left(0.5(((\text{Ash} - 0.25) / 49.6)^2 - 0.3248) \right)} \right]^2 \right\} \quad (6d)$$

$$x_5 = \min \left\{ x_4, \left[\frac{0.5((\text{Temp} - 140) / 160 + (\text{Ash} - 0.25) / 49.6) (\text{Temp} - 140) / 160}{((\text{FC} - 2.66) / 51.34 - (\text{VM} - 34.3) / 55.44) - ((\text{RT} - 5) / 475 + 2(\text{Temp} - 140) / 160 - 4.07501) / 4} \right] \right\} \quad (6e)$$

$$x_6 = \min \left\{ x_5, \left[((RT - 5)/475 + ((Temp - 140)/160 + (FC - 2.66)/51.34)/4)^2 + 0.6882 \right] \right\} \quad (6f)$$

$$HHV = 24.73 \sum_{i=1}^n z_i + 9.8. \quad (7)$$

The z_i in Eq. (7) can be computed using Eq. (8):

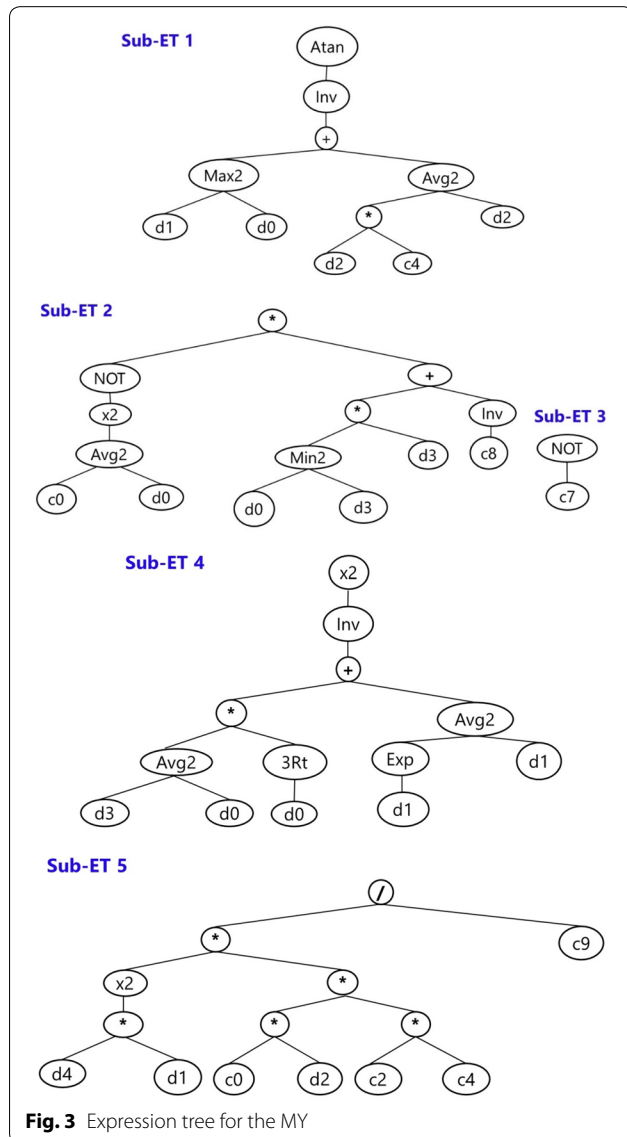
$$z_1 = -\text{EXP}(-4.6508 + \text{VM}/55.44 - \text{Ash}/49.6) \quad (8a)$$

$$z_2 = \text{EXP}(0.5(-12.8458 + \text{Temp}/160 + \text{VM}/55.44 + \text{EXP}(-1.4937 + \text{Temp}/160 + \dots \text{VM}/55.44))) \quad (8b)$$

$$z_3 = \tanh \left\{ 2/[2.2321 + (0.8649 + 2\text{Ash}/49.6 - \text{Temp}/160)(\text{Temp} - 140)^{1/3}/160^{1/3}] \right\} \quad (8c)$$

$$z_4 = 1 / \left\{ \begin{array}{l} -3.0552(-0.6705 + \text{FC}/51.34 + \text{VM}/55.44) \\ -\tanh[(\text{VM} - 34.3)/55.44] - \dots \\ \text{EXP}[-13.2680(\text{FC} - 2.66)/51.34] \end{array} \right\} \quad (8d)$$

$$z_5 = (\text{FC} - 2.66) \tanh(0.5((\text{RT} - 5)/475 + ((\text{Temp} - 140)/ - 332.0686)(-2.9276 + \dots \text{VM}/55.44))/51.34. \quad (8e)$$



Artificial neural network

A MISO-ANN is proposed in this study for the prediction of EY, MY, and HHV. To achieve this, single hidden-layer and double hidden-layers were tried for each of the EY, MY, and HHV as presented in Tables 3, 4. The optimum networks obtained for each of the EY, MY, and HHV are presented in Figs. 7 and 8. In developing the MISO-ANN models, the number of neurons in the input, hidden, and output layers are to be defined and the respective transfer functions at the hidden and output layers are to be defined. Therefore, in this study, there are five neurons in the input layers comprising VM, Ash, FC, Temp, and RT. For the hidden layer, several neurons ranging from 3 to 15 were tried for the MISO-ANN with single hidden layer architecture, while for the MISO-ANN with double hidden layer architecture, the neurons combinations tried ranged 5–3 to 15–15 for each of the targeted variables. The transfer function adopted for the network with a single hidden layer is a hyperbolic tangent for the hidden and output layers, respectively. For the double hidden layer, hyperbolic tangent was used in the first and second hidden layers, while purlin was used for the output layer. Feedforward Backpropagation training algorithm with Levenberg–Marquardt training function was used for the training of the network. One hundred and fifteen (115) datasets were used for model development, divided into 70% for training, 15% each for testing, and validation, respectively (Fig. 2). The datasets were

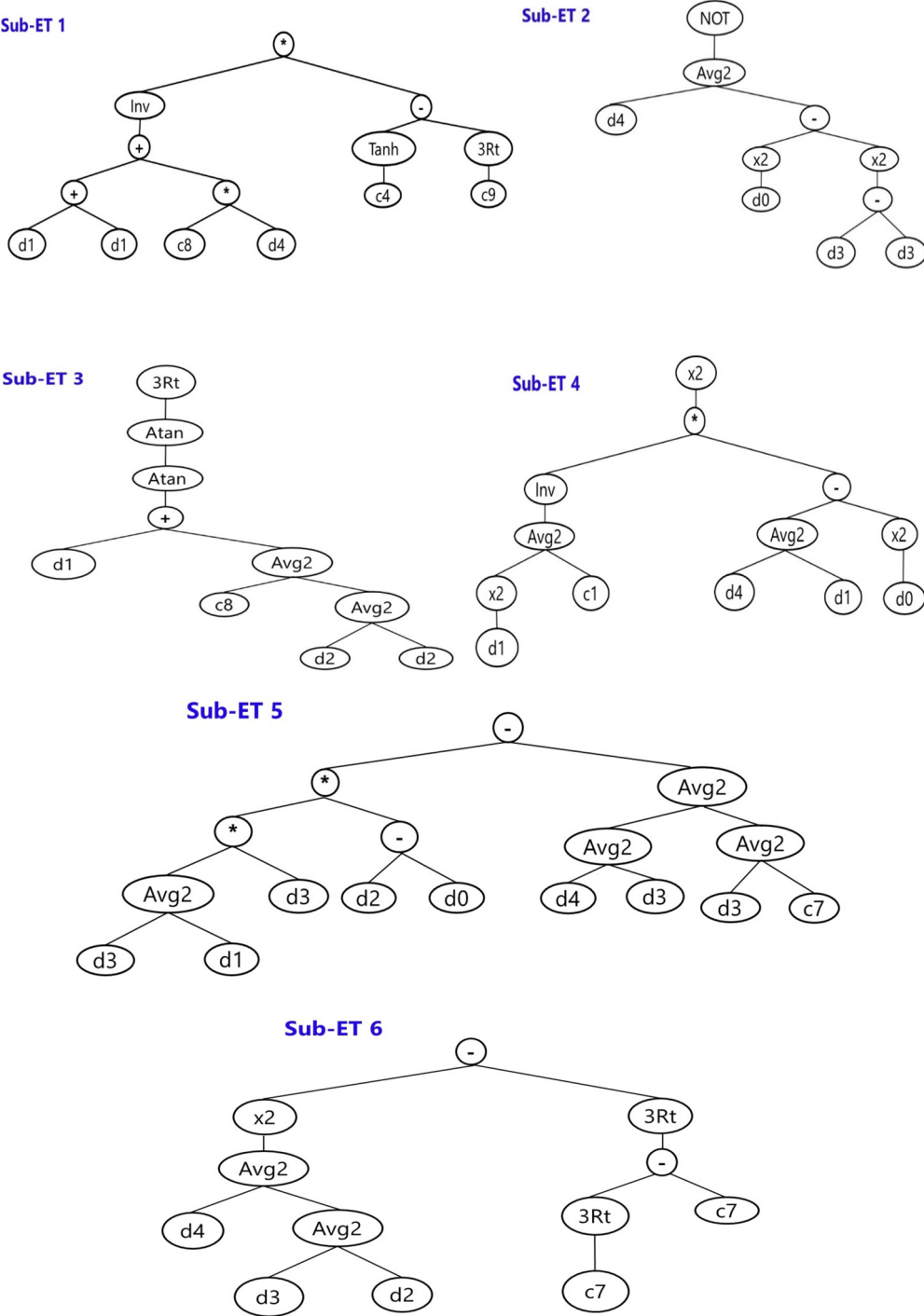


Fig. 4 Expression tree for the EY

normalized to within the range of -1 and 1 to forestall overfitting and ensure dimensional uniformity. The performance of each of the trained networks using the normalized datasets was evaluated using R^2 , RMSE, ME, and standard deviation (std). The obtained outputs for various combinations of neurons are presented in Tables 3, 4. The best network for the MY prediction is 5-10 to 10-, while 5-15 to 15-1 is the best network for EY and 5-15 to 15-1 for the HHV as bolded in Tables 3, 4 and presented in Figs. 6 and 7.

Multiple linear regression analysis

Regression analysis is commonly used in establishing the relationship between the regressor and the targeted variable. When it involves a relationship between the targeted variable and a regressor, it is known as linear regression analysis. However, for more than one regressor, it is known as multiple linear regression analysis. MLR has been used by researchers (Said et al. 2020a, b; Onifade et al. 2019) for prediction purposes. MLR is also adopted in this study, to enable the comparison between GEP and ANN models. MLR model was developed for each of the three predicted parameters: MY, EY, and HHV. The MLR analysis was performed in the Microsoft Excel software Add-ins using the same datasets used in GEP and ANN models. The obtained MLR models are as presented in Eqs. (9) to (11):

$$MY = 557.5078 - 4.4502VM - 4.3243Ash - 4.0474FC - 0.2495Temp - 0.0409RT \quad (9)$$

$$EY = -230.2695 + 3.2966VM + 3.2934Ash + 3.9157FC - 0.1372Temp - 0.0337RT \quad (10)$$

$$HHV = -73.8668 + 0.89395VM + 0.5879Ash + 0.9914FC + 0.0398Temp + 0.0058RT. \quad (11)$$

Results and discussion

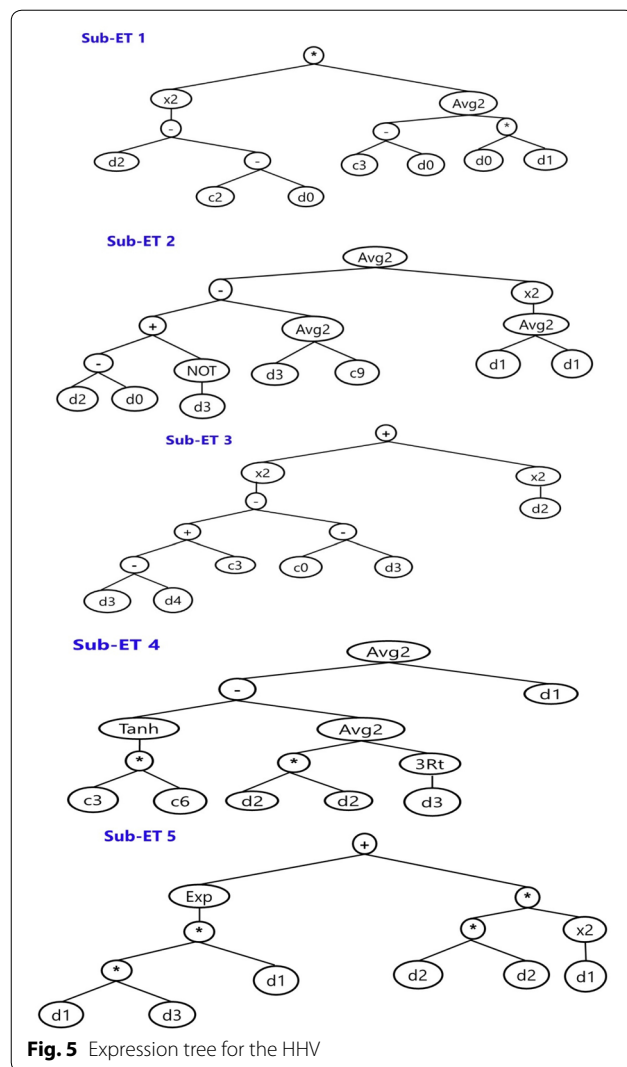
Models comparison

The accuracy of the proposed models using GEP, MISO-ANN, and MLR methods are compared with the laboratory-measured values using the testing and validation datasets. For the MY, the outcome of the comparison is presented in Fig. 8. For the training datasets, the points predicted with MISO-ANN fall largely within the 3% error line, while many of the points predicted by the GEP

and MLR fall outside the error line. This hitherto gave rise to R^2 of 0.981 obtained for the MISO-ANN, while the R^2 recorded for both the GEP and MLR models are 0.691 and 0.463, respectively, for the testing datasets. For the validation data points, however, the R^2 values recorded for the MISO-ANN are 0.976, while those of GEP and MLR are 0.548 and 0.154, respectively. The MISO-ANN predictions are generally closer to the experimentally measured values among the three proposed models. The performance of MISO-ANN can be attributed to its ability to handle complex non-linearity between the model parameters (Gevrey et al. 2003). The outcome of the MISO-ANN is consistent with most of the previous studies that compared the performance of ANN with the regression-based models in predicting the HHV of solid fuels (Onifade et al. 2019; Ghugare et al. 2017; Uzun et al. 2017; Patel et al. 2007). Aside from the HHV of solid fuel, many authors have found that the ANN provides a more reliable predictive model than the regression-based model (Lawal 2020; Lawal et al. 2020; Said et al. 2020a; Saadat et al. 2014; Khandelwal and Singh 2010).

The outcome of the comparison of the predictive ability of the three proposed models GEP, MISO-ANN, and MLR are also tested for the hydrochar property EY as presented in Fig. 9. The majority of the data points predicted with MISO-ANN fall within the 3% error lines, while many of the predicted data points using GEP and

MLR fall outside the error line. As a result of this, the resulting performance indicator R^2 of MISO-ANN for the testing and validation datasets are 0.965 and 0.955, respectively, while that of the GEP are 0.622 and 0.419. For the MLR model, the R^2 values for the respective testing and validation datasets are 0.219 and 0.205. Again, the MISO-ANN model outperforms the GEP and MLR models.



Similarly, the accuracy of the proposed models (GEP, MISO-ANN, and MLR) for predicting HHV is also evaluated as presented in Fig. 10. All the predicted data points using the MISO-ANN fall within the 3% error lines for both the testing and validation data points. The majority of the predicted data points by GEP and MLR models fall outside the error lines in Fig. 10. As a result of the presence of the data points predicted using the MISO-ANN model within the error band, the performance of the MISO-ANN is excellent as the R^2 of 0.999 is obtained for

testing data points (Fig. 10a), while 0.996 was recorded for the validation data points (Fig. 10b). The R^2 values obtained for the GEP models are 0.810 and 0.717, respectively, while that of the MLR models are 0.788 and 0.643 for the respective testing and validation data points. The low R^2 values observed in GEP and MLR models for the MY, EY, and HHV predictions can be attributed to their respective predicted data points that fall outside the error bands. Hence, the MISO-ANN can give reliable predictions of the MY, EY, and HHV follow by the GEP model, while MLR may not be reliable.

Error analysis

To enable the selection of the best performing model for predicting the MY, EY, and HHV, mean absolute error and mean bias errors were evaluated for each of the three techniques used in developing the proposed models as presented in Eqs. (12) and (13):

$$MAE = \frac{\sum_{i=1}^n \left| \frac{P_i - M_i}{M_i} \right|}{n} \times 100\% \quad (12)$$

$$MBE = \frac{\sum_{i=1}^n \left[\frac{P_i - M_i}{M_i} \right]}{n} \times 100\%. \quad (13)$$

The obtained results from the conducted analyses using Eqs. (12) and (13) are presented in Table 5. From Table 5, the MAE of 2.24, 2.11, and 0.93% were obtained for MY, EY, and HHV, respectively, using the MISO-ANN model, while the MBE obtained are 0.16, 0.37, and 0.12% for MY, EY, and HHV, respectively. Hence, the best model for the prediction of MY, EY, and HHV is the MISO-ANN model follow by the GEP model, while the MLR will overestimate the values of the MY, EY, and HHV based on the MBE values in Table 5.

Sensitivity analysis

The sensitivity analysis helps in providing useful information on the contributions of each of the input parameters on the output predicted by the model. Various techniques have been proposed to perform this task but the Cosine Amplitude method (CAM) (Yang and Zhang 1997) as presented in Eq. (14) is adopted in this study:

Table 3 Selection of the optimum MISO-ANN network for MY and EY

		MY				EY			
		Tr	Te	Val	Ov	Tr	Te	Val	Ov
5-3-1	R^2	0.821	0.834	0.847	0.824	0.695	0.644	0.563	0.668
	RMSE	0.183	0.197	0.184	0.185	0.219	0.195	0.251	0.221
	ME	0.004	− 0.014	0.044	0.007	− 0.003	− 0.06	0.046	− 0.004
	std	0.184	0.202	0.184	0.186	0.22	0.191	0.255	0.221
5-5-1	R^2	0.935	0.897	0.921	0.928	0.83	0.817	0.794	0.824
	RMSE	0.107	0.128	0.155	0.118	0.154	0.132	0.212	0.161
	ME	0.01	− 0.031	0.002	0.003	0.005	− 0.009	− 0.024	− 0.002
	std	0.107	0.128	0.16	0.119	0.154	0.135	0.217	0.161
5-5 to 3-1	R^2	0.975	0.921	0.862	0.95	0.822	0.872	0.809	0.827
	RMSE	0.069	0.147	0.148	0.099	0.161	0.177	0.145	0.161
	ME	0.001	0.009	0.03	0.006	0.005	− 0.008	0.059	0.011
	std	0.069	0.152	0.149	0.099	0.162	0.182	0.137	0.162
5-5 to 5-1	R^2	0.977	0.979	0.933	0.972	0.923	0.938	0.916	0.923
	RMSE	0.07	0.064	0.1	0.074	0.109	0.097	0.103	0.107
	ME	0.004	0.018	− 0.02	0.002	0.009	0.003	− 0.02	0.004
	std	0.07	0.063	0.101	0.075	0.109	0.1	0.105	0.107
5-5 to 6-1	R^2	0.97	0.979	0.97	0.972	0.944	0.944	0.893	0.938
	RMSE	0.074	0.075	0.078	0.074	0.088	0.109	0.118	0.096
	ME	− 0.004	− 0.029	− 0.011	− 0.009	− 0.004	0.008	− 0.009	− 0.003
	std	0.074	0.071	0.08	0.074	0.088	0.112	0.121	0.097
5-5 to 10-1	R^2	0.985	0.985	0.975	0.984	0.967	0.99	0.935	0.968
	RMSE	0.052	0.066	0.072	0.058	0.07	0.041	0.088	0.07
	ME	0.002	0.016	0.025	0.008	− 0.017	0.005	− 0.013	− 0.013
	std	0.053	0.066	0.07	0.058	0.069	0.042	0.09	0.069
5-7-1	R^2	0.97	0.947	0.915	0.962	0.911	0.947	0.889	0.913
	RMSE	0.083	0.084	0.105	0.087	0.122	0.088	0.099	0.115
	ME	0.005	0.001	0.034	0.008	− 0.006	0.02	0.001	− 0.001
	std	0.083	0.087	0.011	0.087	0.123	0.088	0.102	0.115
5-10-1	R^2	0.964	0.976	0.967	0.966	0.905	0.958	0.934	0.916
	RMSE	0.08	0.08	0.098	0.083	0.115	0.089	0.115	0.112
	ME	0.003	0.012	0.033	0.009	− 0.005	0.015	0.03	0.004
	std	0.08	0.082	0.095	0.083	0.116	0.09	0.114	0.112
5-10 to 10-1	R^2	0.989	0.984	0.977	0.986	0.967	0.98	0.905	0.964
	RMSE	0.049	0.063	0.058	0.053	0.075	0.049	0.089	0.074
	ME	0.001	0.017	− 0.024	− 0.001	− 0.006	− 0.008	− 0.007	− 0.006
	std	0.049	0.062	0.055	0.053	0.075	0.049	0.091	0.074
5-15-1	R^2	0.975	0.967	0.986	0.975	0.945	0.87	0.947	0.936
	RMSE	0.07	0.078	0.063	0.07	0.095	0.113	0.095	0.098
	ME	− 0.002	− 0.03	0.01	− 0.005	0.015	− 0.02	0.021	0.011
	std	0.07	0.075	0.064	0.07	0.094	0.115	0.095	0.098
5-15-15-1	R^2	0.984	0.965	0.983	0.981	0.977	0.966	0.955	0.972
	RMSE	0.057	0.081	0.057	0.061	0.059	0.078	0.079	0.065
	ME	− 0.001	0.011	− 0.019	− 0.002	− 0.006	0.012	− 0.024	− 0.006
	std	0.057	0.083	0.056	0.061	0.059	0.079	0.077	0.065

^a Tr training, Te testing, Val validation, Ov overall

Table 4 Selection of the optimum MISO-ANN network for HHV

		HHV			
		Tr	Te	Val	Ov
5-3-1	R^2	0.897	0.937	0.935	0.91
	RMSE	0.134	0.095	0.137	0.129
	ME	− 0.011	0.022	− 0.008	− 0.006
	std	0.134	0.095	0.141	0.13
5-5-1	R^2	0.969	0.982	0.971	0.971
	RMSE	0.08	0.052	0.066	0.075
	ME	0.01	0.011	− 0.005	0.079
	std	0.08	0.052	0.068	0.075
5-5-3-1	R^2	0.982	0.994	0.991	0.985
	RMSE	0.058	0.038	0.044	0.054
	ME	− 0.001	0.017	0.003	0.003
	std	0.059	0.035	0.045	0.054
5-5-5-1	R^2	0.989	0.995	0.988	0.99
	RMSE	0.047	0.038	0.043	0.045
	ME	0.001	0.002	0.009	0.002
	std	0.047	0.04	0.044	0.045
5-5-6-1	R^2	0.989	0.98	0.984	0.986
	RMSE	0.049	0.065	0.049	0.051
	ME	0.008	− 0.022	0.004	0.003
	std	0.048	0.063	0.05	0.051
5-5-10-1	R^2	0.995	0.991	0.998	0.995
	RMSE	0.033	0.045	0.023	0.033
	ME	− 0.002	0.018	0.008	0.003
	std	0.032	0.043	0.022	0.033
5-7-1	R^2	0.966	0.924	0.981	0.965
	RMSE	0.083	0.093	0.064	0.082
	ME	− 0.006	− 0.001	− 0.034	− 0.009
	std	0.083	0.096	0.056	0.082
5-10-1	R^2	0.988	0.989	0.991	0.988
	RMSE	0.052	0.047	0.033	0.049
	ME	0.003	− 0.023	0.008	− 0.001
	std	0.052	0.043	0.033	0.049
5-10-10-1	R^2	0.994	0.987	0.993	0.992
	RMSE	0.035	0.057	0.039	0.04
	ME	0.003	− 0.016	− 0.01	− 0.003
	std	0.035	0.056	0.039	0.04
5-15-1	R^2	0.994	0.997	0.988	0.994
	RMSE	0.036	0.031	0.043	0.036
	ME	0.006	− 0.015	0.022	0.005
	std	0.036	0.028	0.038	0.036
5-15-15-1	R^2	0.997	1	0.997	0.997
	RMSE	0.025	0.014	0.032	0.025
	ME	− 0.004	− 0.006	0.012	− 0.002
	std	0.025	0.014	0.031	0.025

$$R_{ij} = \frac{\sum_{k=1}^n (r_m \times P)}{\sqrt{\sum_{k=1}^n r_m^2 \sum_{i=1}^n P^2}}, \quad (14)$$

where R_{ij} stands for the strength of the input parameter, r_m represents the model regressors, P is the predicted output, n is the data points number.

The MISO-ANN model which is adjudged the best out of the three proposed models based on the previous analysis conducted in this study is used to perform the sensitivity analysis. The output obtained is presented in the Pareto chart (PC) shown in Figs. 11, 12, 13. The VM and Temp have the highest influence on the MY, EY, and HHV as presented in Figs. 11, 12, 13. The order of the influence in all the figures is VM > Temp > FC > RT > Ash. In addition, based on the Pareto chart analysis, the Ash and RT in that order should not be ignored when predicting MY, EY, and HHV.

Conclusion

Mass yield, energy yield, and higher heating value are important hydrochar properties required for the analysis and design of any bioenergy systems. In the present study, Gene Expression Programming, multiple-input single output-artificial neural network, and Multilinear regression were applied to predict MY, EY, and HHV of hydrochars using the composition of biomass source from proximate analysis and HTC process conditions (temperature and residence time). Based on R^2 values and error analysis, MISO-ANN with 5-10 to 10-1 and 5-15 to 15-1 network architectures presented the best performance among the proposed models with $R^2 = 0.976$, 0.955, 0.996; MAE = 2.24, 2.11, 0.93; MBE = 0.16, 0.37, 0.12 for the respective MY, EY, and HHV. GEP has been shown to provide satisfactory predictive alternative to MISO-ANN with $R^2 = 0.691$, 0.622, 0.810; MAE = 12.38, 10.31, 8.58; MBE = 2.95, 0.64, 0.78 for MY, EY and HHV, respectively. From the sensitivity analysis, volatile matter and temperature were found to be the most influencing input variables. This study demonstrated the ability of GEP to satisfactorily model hydrochar properties based on biomass composition and HTC process conditions. Although the accuracy of the GEP models was slightly lower than that of the MISO-ANN models, the GEP models provided much more accurate predictions than the MLR models, which proved unsatisfactory.

Table 5 Error analysis

Models	GEP		MLR		MISO-ANN	
	MAE (%)	MBE (%)	MAE (%)	MBE (%)	MAE (%)	MBE (%)
MY	12.38	2.95	18.10	5.80	2.24	0.16
EY	10.31	0.64	13.20	4.01	2.11	0.37
HHV	8.58	0.78	10.75	1.23	0.93	0.12

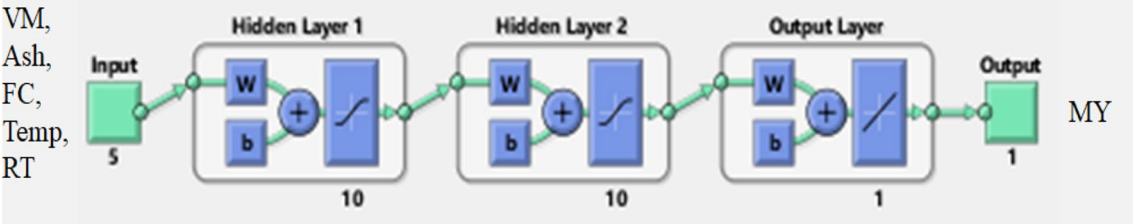


Fig. 6 Optimum MISO-ANN architecture for MY (5–10–10–1)

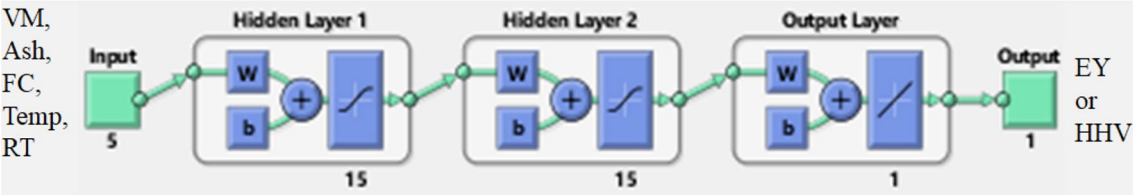


Fig. 7 Optimum MISO-ANN architecture for either EY or HHV (5–15–15–1)

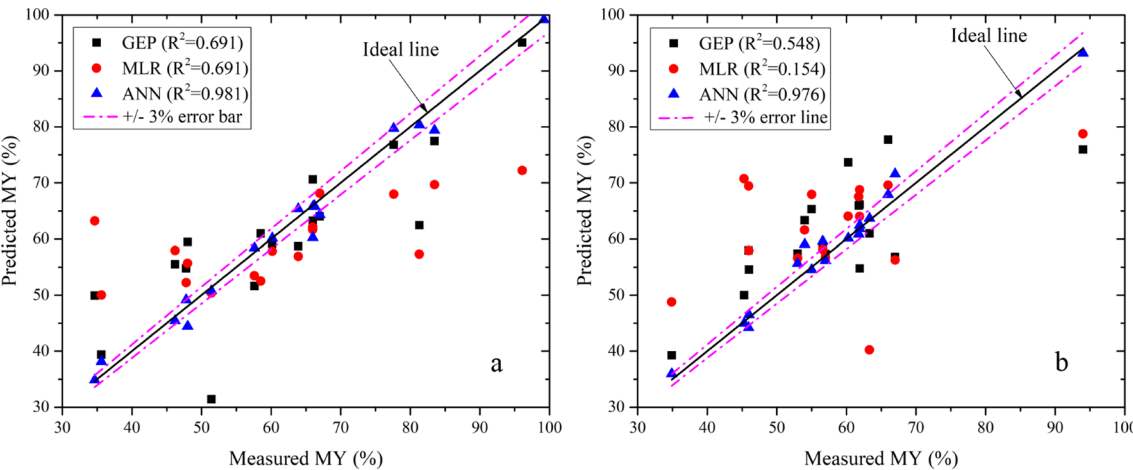
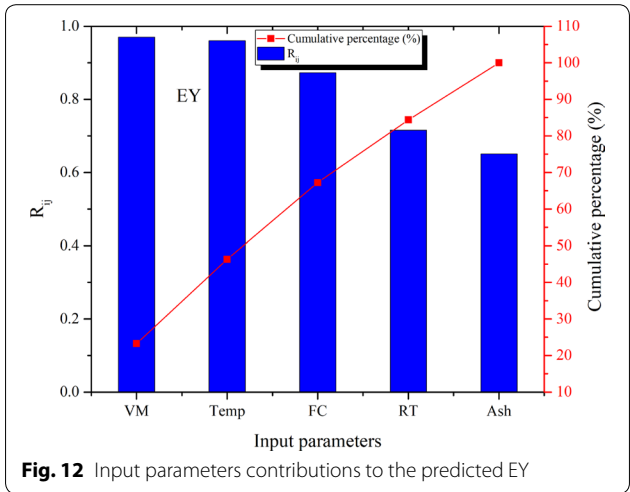
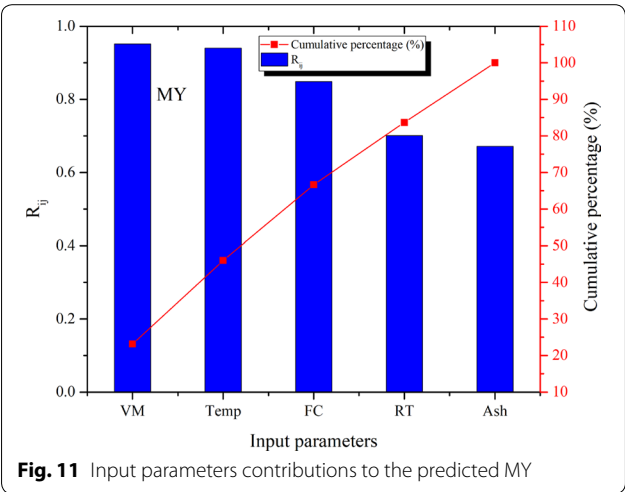
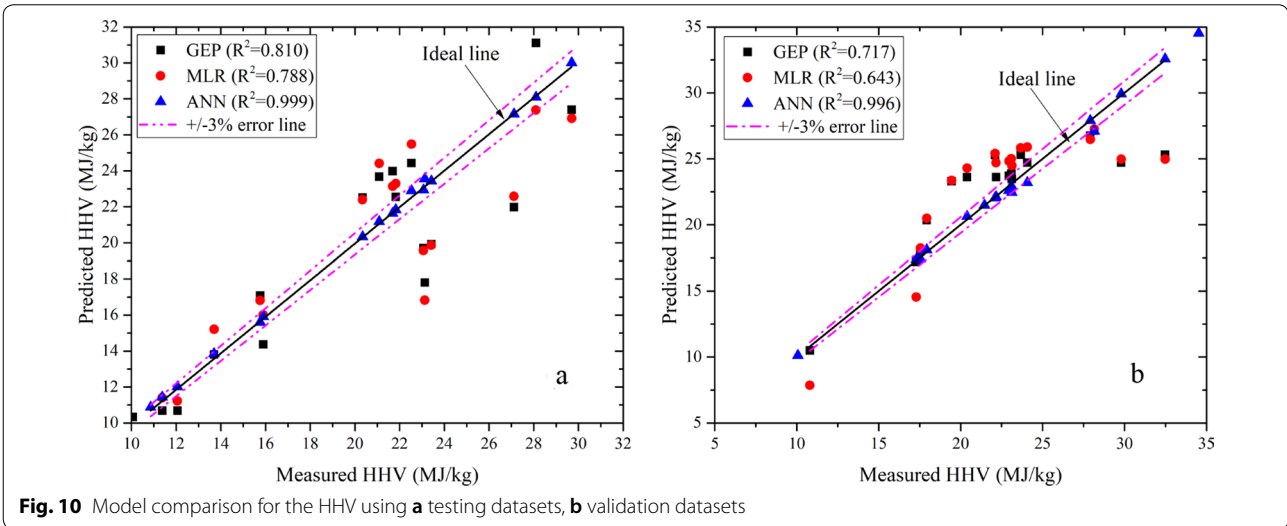
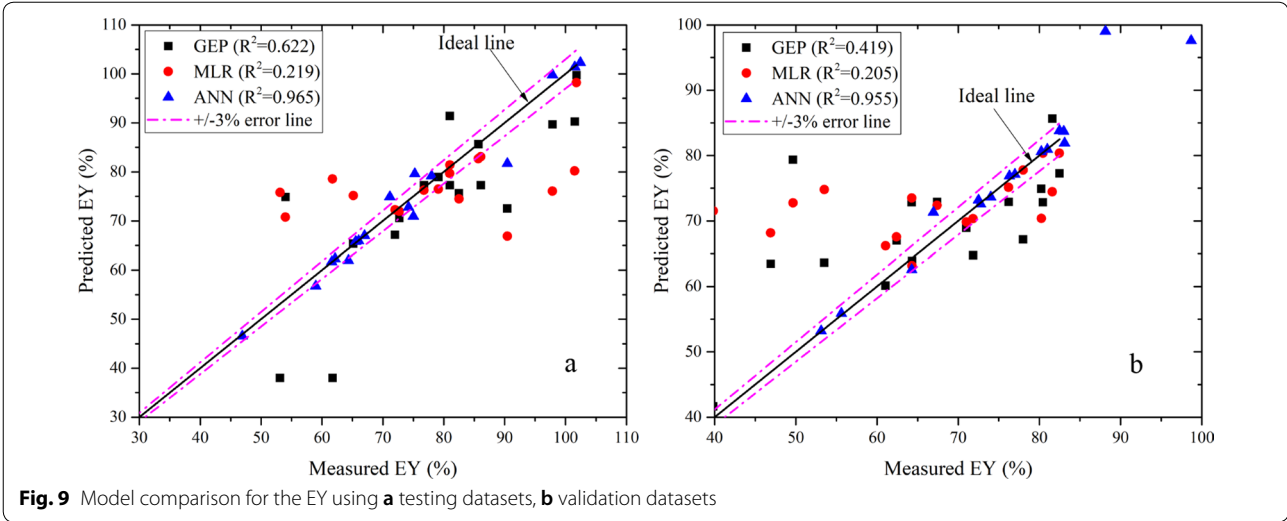


Fig. 8 Model comparison for the MY using a testing datasets, b validation datasets



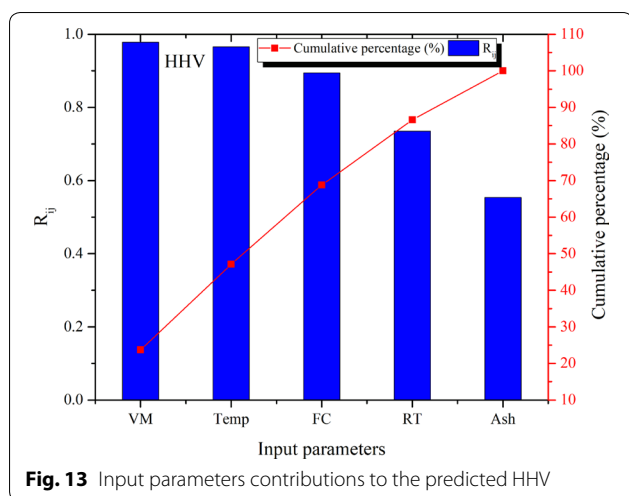


Fig. 13 Input parameters contributions to the predicted HHV

Abbreviations

ANFIS: Adaptive neuro-fuzzy inference system; ANN: Artificial neural network; ANN-PSO: Artificial neural network-particle swarm optimization; ASTM: American Society for Testing and Materials; CAM: Cosine amplitude method; EY: Energy yield; FC: Fixed carbon; FANNs: Feed-forward artificial neural networks; GA: Genetic Algorithm; GEP: Gene expression programming; GP: Genetic Programming; HHV: Higher heating value; HHV_{HC} : Higher heating value of hydrochar; HHV_R : Higher heating value of raw biomass; HTC: Hydrothermal carbonization; MAE: Mean absolute error; MBE: Mean bias errors; ME: Mean error; M_{HC} : Mass of hydrochar; M_R : Mass of raw biomass; MISO-ANN: Multiple-input single output-artificial neural network; MLR: Multilinear regression; MY: Mass yield; PC: Pareto chart; R^2 : Coefficient of determination; RMSE: Root Mean Square Error; RT: Residence time; TGA: Thermogravimetric analysis; VM: Volatile matter.

Acknowledgements

This work was supported by the National Research Foundation (NRF), South Africa (Grant Number: 86421). The opinions, observations, and findings are those of the authors and cannot be attributed to the NRF.

Author contributions

JA: Conceptualization, methodology, Data analysis, Experimental investigation, and writing the original draft. AL: Model development, software, data analysis, and validation. RS: Experimental investigation and data collection. MO: Reviewed and edited the original draft. SB: provided resources, supervision, and project administration. All authors read and approved the final manuscript.

Funding

We acknowledge the funding support from the National Research Foundation (NRF), South Africa for the experimental part of this study.

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Ethics approval and consent to participate

Not Applicable.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ DSI/NRF Clean Coal Technology Research Group, School of Chemical and Metallurgical Engineering, Faculty of Engineering and the Built Environment, University of the Witwatersrand, WITS 2050, Johannesburg, South Africa. ² Department of Energy Resources Engineering, Inha University Yong-Hyun Dong, Nam Ku, Incheon, Korea. ³ Department of Mining Engineering, Federal University of Technology, Akure, Nigeria. ⁴ Department of Mining and Metallurgical Engineering, University of Namibia, Windhoek, Namibia.

Received: 4 September 2020 Accepted: 19 November 2020

Published online: 27 November 2020

References

- Aladejare AE, Onifade M, Lawal AI (2020) Application of metaheuristic based artificial neural network and multilinear regression for the prediction of higher heating values of fuels. *Int J Coal Prep Util*:1–22
- Bach Q-V, Tran K-Q, Skreiberg Ø (2016) Hydrothermal pretreatment of fresh forest residues: effects of feedstock pre-drying. *Biomass Bioenerg* 85:76–83
- Brewer CE, Chuang VJ, Masiello CA, Gonnermann H, Gao X, Dugan B, Driver LE, Panzacchi P, Zygourakis K, Davies CA (2014) New approaches to measuring biochar density and porosity. *Biomass Bioenerg* 66:176–185
- Danso-Boateng E (2015) Biomass hydrothermal carbonisation for sustainable engineering. Doctoral dissertation, Loughborough University
- Estati I, Freire FB, Freire JT, Aguado R, Olazar M (2016) Fitting performance of artificial neural networks and empirical correlations to estimate higher heating values of biomass. *Fuel* 180:377–383
- Ferreira C (2001) Gene expression programming: a new adaptive algorithm for solving problems. <https://arxiv.org/abs/cs/0102027>. Accessed 22 Jul 2020
- Funke A, Ziegler F (2010) Hydrothermal carbonization of biomass: a summary and discussion of chemical mechanisms for process engineering. *Biofuel Bioprod Bior* 4(2):160–177
- Gao L, Volpe M, Lucian M, Fiori L, Goldfarb JL (2019) Does hydrothermal carbonization as a biomass pretreatment reduce fuel segregation of coal-biomass blends during oxidation? *Energy Convers. Manage* 181:93–104
- Gevrey M, Dimopoulos I, Lek S (2003) Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol Model* 160(3):249–264
- Ghugare SB, Tiwary S, Tambe SS (2017) Computational intelligence based models for prediction of elemental composition of solid biomass fuels from proximate analysis. *Int J Sys Ass Eng Mgt* 8(4):2083–2096
- Guvén A, Aytekin A (2009) New approach for stage-discharge relationship: gene-expression programming. *J Hydrol Eng* 14(8):812–820
- Hansen JV, Meservy RD (1996) Learning experiments with genetic optimization of a generalized regression neural network. *Decis Support Syst* 18(3–4):317–325
- He X, Liu Z, Niu W, Yang L, Zhou T, Qin D, Niu Z, Yuan Q (2018) Effects of pyrolysis temperature on the physicochemical properties of gas and biochar obtained from pyrolysis of crop residues. *Energy* 143:746–756
- Holtmeyer ML, Li G, Kumfer BM, Li S, Axelbaum RL (2013) The impact of biomass cofiring on volatile flame length. *Energy Fuel* 27(12):7762–7771
- Hwang I-H, Aoyama H, Matsuto T, Nakagishi T, Matsuo T (2012) Recovery of solid fuel from municipal solid waste by hydrothermal treatment using subcritical water. *J Waste Manag* 32(3):410–416
- Jain AK, Mao J, Mohiuddin KM (1996) Artificial neural networks: a tutorial. *Comput J* 29(3):31–44
- Kambo HS, Dutta A (2014) Strength, storage, and combustion characteristics of densified lignocellulosic biomass produced via torrefaction and hydrothermal carbonization. *Appl Energy* 135:182–191
- Kambo HS, Dutta A (2015) A comparative review of biochar and hydrochar in terms of production, physico-chemical properties, and applications. *Renew Sust Energ Rev* 45:359–378

- Khandelwal M, Singh T (2010) Prediction of macerals contents of Indian coals from proximate and ultimate analyses using artificial neural networks. *Fuel* 89(5):1101–1109
- Kim D, Lee K, Park KY (2014) Hydrothermal carbonization of anaerobically digested sludge for solid fuel production and energy recovery. *Fuel* 130:120–125
- Krylova AY, Zaitchenko V (2018) Hydrothermal carbonization of biomass: a review. *Solid Fuel Chem* 52(2):91–103
- Kubacki ML, Ross AB, Jones JM, Williams A (2012) Small-scale co-utilisation of coal and biomass. *Fuel* 101:84–89
- Lawal AI (2020) An artificial neural network-based mathematical model for the prediction of blast-induced ground vibration in granite quarries in Ibadan, Oyo State Nigeria. *Sci Afr* 8:e00413
- Lawal AI, Aladejare AE, Onifade M, Bada S, Idris MA (2020) Predictions of elemental composition of coal and biomass from their proximate analyses using ANFIS, ANN, and MLR. *Int J Coal Sci Technol*. <https://doi.org/10.1007/s40789-020-00346-9>
- Lee J, Sohn D, Lee K, Park KY (2019) Solid fuel production through hydrothermal carbonization of sewage sludge and microalgae *Chlorella* sp. from wastewater treatment plant. *Chemosphere* 230:157–163
- Libra JA, Ro KS, Kammann C, Funke A, Berge ND, Neubauer Y, Titirici M-M, Fühner C, Bens O, Kern J (2011) Hydrothermal carbonization of biomass residuals: a comparative review of the chemistry, processes, and applications of wet and dry pyrolysis. *Biofuels* 2(1):71–106
- Majumder AK, Jain R, Banerjee P, Barnwal J (2008) Development of a new proximate analysis based correlation to predict calorific value of coal. *Fuel* 87(13–14):3077–3081
- Makwarela M, Bada S, Falcon R (2017) Co-firing combustion characteristics of different ages of *Bambusabalcooa* relative to a high ash coal. *Renew Energy* 105:656–664
- Mierzwa-Hersztel M, Gondek K, Jewiarz M, Dziedzic K (2019) Assessment of energy parameters of biomass and biochars, leachability of heavy metals, and phytotoxicity of their ashes. *J Mater Cycles Waste* 21(4):786–800
- Mumme J, Eckervogt L, Pielert J, Diakité M, Rupp F, Kern J (2011) Hydrothermal carbonization of anaerobically digested maize silage. *BioresourTechnol* 102(19):9255–9260
- Onifade M, Lawal AI, Aladejare AE, Bada S, Idris MA (2019) Prediction of gross calorific value of solid fuels from their proximate analysis using soft computing and regression analysis. *Int J Coal Prep Util*:1–15
- Park KY, Lee K, Kim D (2018) Characterized hydrochar of algal biomass for producing solid fuel through hydrothermal carbonization. *BioresourTechnol* 258:119–124
- Parshetti GK, Hoekman SK, Balasubramanian R (2013) Chemical, structural and combustion characteristics of carbonaceous products obtained by hydrothermal carbonization of palm empty fruit bunches. *BioresourTechnol* 135:683–689
- Patel N (2019) Hydrothermal Carbonization (HTC) of Marine Seaweed (Macroalgae) for Producing Hydro-Char. Masters thesis, Dalhousie University
- Patel SU, Kumar BJ, Badhe YP, Sharma B, Saha S, Biswas S, Chaudhury A, Tambe SS, Kulkarni BD (2007) Estimation of gross calorific value of coals using artificial neural networks. *Fuel* 86(3):334–344
- Peng C, Zhai Y, Zhu Y, Xu B, Wang T, Li C, Zeng G (2016) Production of char from sewage sludge employing hydrothermal carbonization: char properties, combustion behavior, and thermal characteristics. *Fuel* 176:110–118
- Perlack RD, Eaton LM, Turhollow Jr AF, Langholtz MH, Brandt CC, Downing ME, Graham RL, Wright LL, Kavkewitz JM, Shamey AM (2011) US billion-ton update: biomass supply for a bioenergy and bioproducts industry. https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1015&context=abe_eng_reports. Accessed 16 Jul 2020
- Pradhan P, Mahajani SM, Arora A (2018) Production and utilization of fuel pellets from biomass: a review. *Fuel Process Technol* 181:215–232
- Reza MT, Uddin MH, Lynam JG, Hoekman SK, Coronella CJ (2014) Hydrothermal carbonization of loblolly pine: reaction chemistry and water balance. *Biomass Convers Biorefin* 4(4):311–321
- Rousset P, Macedo L, Commandré J-M, Moreira A (2012) Biomass torrefaction under different oxygen concentrations and its effect on the composition of the solid by-product. *J Anal Appl Pyrolysis* 96:86–91
- Saadat M, Khandelwal M, Monjezi M (2014) An ANN-based approach to predict blast-induced ground vibration of Gol-E-Gohar iron ore mine. *Iran J Rock Mech Geotech Eng* 6(1):67–76
- Saba A, Saha P, Reza MT (2017) Co-Hydrothermal Carbonization of coal-biomass blend: Influence of temperature on solid fuel properties. *Fuel Process Technol* 167:711–720
- Sadiku NA, Oluyeye AO, Sadiku IB (2016) Analysis of the calorific and fuel value index of bamboo as a source of renewable biomass feedstock for energy generation in Nigeria. *Lignocellulose* 5(1):34–49
- Safarian S, Unnþórsson R, Richter C (2019) A review of biomass gasification modelling. *Renew SustEnergy Rev* 110:378–391
- Said KO, Onifade M, Lawal AI, Githiria JM (2020a) An artificial intelligence-based model for the prediction of spontaneous combustion liability of coal based on its proximate analysis. *Combust Sci Technol*:1–18
- Said KO, Onifade M, Lawal AI, Githiria JM (2020b) Computational intelligence-based models for predicting the spontaneous combustion liability of coal. *Int J Coal Prep Util*:1–25
- Saidur R, Abdelaziz E, Demirbas A, Hossain M, Mekhilef S (2011) A review on biomass as a fuel for boilers. *Renew SustEnergy Rev* 15(5):2262–2289
- Saldarriaga JF, Aguado R, Pablos A, Amutio M, Olazar M, Bilbao J (2015) Fast characterization of biomass fuels by thermogravimetric analysis (TGA). *Fuel* 140:744–751
- Sevilla M, Fuertes AB (2009) The production of carbon materials by hydrothermal carbonization of cellulose. *Carbon* 47(9):2281–2289
- Seyedsadr S, Al Afif R, Pfeifer C (2018) Hydrothermal carbonization of agricultural residues: a case study of the farm residues-based biogas plants. *Carbon Resour Convers* 1(1):81–85
- Sheng C, Azevedo J (2005) Estimating the higher heating value of biomass fuels from basic analysis data. *Biomass Bioenerg* 28(5):499–507
- Silakova M (2018) Hydrothermal carbonization of the tropical biomass.
- Stemann J, Erlach B, Ziegler F (2013) Hydrothermal carbonisation of empty palm oil fruit bunches: laboratory trials, plant simulation, carbon avoidance, and economic feasibility. *Waste Biomass Valor* 4(3):441–454
- Tekin K, Karagöz S, Bektaş S (2014) A review of hydrothermal biomass processing. *Renew SustEnergy Rev* 40:673–687
- Teodorescu L, Sherwood D (2008) High energy physics event selection with gene expression programming. *Comput Phys Commun* 178(6):409–419
- Uzun H, Yıldız Z, Goldfarb JL, Ceylan S (2017) Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. *BioresourTechnol* 234:122–130
- Vargas-Moreno J, Callejón-Ferre A, Pérez-Alonso J, Velázquez-Martí B (2012) A review of the mathematical models for predicting the heating value of biomass materials. *Renew SustEnergy Rev* 16(5):3065–3083
- Volpe M, Goldfarb JL, Fiori L (2018) Hydrothermal carbonization of *Opuntia ficus-indica* cladodes: role of process parameters on hydrochar properties. *BioresourTechnol* 247:310–318
- Wang T, Zhai Y, Zhu Y, Li C, Zeng G (2018) A review of the hydrothermal carbonization of biomass waste for hydrochar formation: Process conditions, fundamentals, and physicochemical properties. *Renew SustEnergy Rev* 90:223–247
- Wasserman PD (1993) *Advanced methods in neural computing*. Wiley, New York
- Wiedner K, Naisse C, Rumpel C, Pozzi A, Wiecek P, Glaser B (2013) Chemical modification of biomass residues during hydrothermal carbonization—What makes the difference, temperature or feedstock? *Org Geochem* 54:91–100
- Xiong J-B, Pan Z-Q, Xiao X-F, Huang H-J, Lai F-Y, Wang J-X, Chen S-W (2019) Study on the hydrothermal carbonization of swine manure: The effect of process parameters on the yield/properties of hydrochar and process water. *J Anal Appl Pyrolysis* 144:104692
- Xu Q, Qian Q, Quek A, Ai N, Zeng G, Wang J (2013) Hydrothermal carbonization of macroalgae and the effects of experimental parameters on the properties of hydrochars. *ACS Sustain Chem Eng* 1(9):1092–1101
- Yang Y, Zhang Q (1997) A hierarchical analysis for rock engineering using artificial neural networks. *Rock Mech Rock Eng* 30(4):207–222

- Zhang Z, Pang S (2019) Experimental investigation of tar formation and producer gas composition in biomass steam gasification in a 100 kW dual fluidised bed gasifier. *Renew Energy* 132:416–424
- Zhu Y, Si Y, Wang X, Zhang W, Shao J, Yang H, Chen H (2018) Characterization of hydrochar pellets from hydrothermal carbonization of agricultural residues. *Energ Fuel* 32(11):11538–11546

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
