Kufa Journal of Engineering
Vol. 17, No. 3, July 2026, P.P. 772-788
Article history: Received 16 December 2025, last revised 29 January 2026, accepted 13 February 2026

PREDICTION OF HIGHER HEATING VALUE OF BIOMASS: STATISTICAL AND EMPIRICAL MODELLING USING MULTIVARIATE REGRESSION AND RESPONSE SURFACE METHODOLOGY

Anthony O. Onokwai [1,2], Dawood A. Desai[3], and Udochukwu B. Akuru[4]

[1] **Department of Mechanical and Mechatronics Engineering, Tshwane University of Technology, Pretoria 0183, South Africa, Email:onokwaiao@tut.ac.za.**

[2] **Department of Mechanical Engineering, Pan-Atlantic University, Lekki, 105101, Nigeria, Email:onokwaianthony@gmail.com.**

[3] **Department of Mechanical and Mechatronics Engineering, Tshwane University of Technology, Pretoria 0183, South Africa, Email:desaida@tut.ac.za.**

[4] **Department of Electrical Engineering, Tshwane University of Technology, Pretoria 0183, South Africa, Email:akuruub@tut.ac.za.**

https://doi.org/10.30572/2018/KJE/170345

ABSTRACT

This study developed experimental and statistical predictive models to determine the higher heating value (HHV) of lignocellulosic biomass using proximate, ultimate, and structural composition analysis data. Multiple linear regression (MLR) and response surface methodology (RSM) were utilized to establish a mathematical relationship between important compositional parameters and HHV. Results indicate that carbon, hydrogen, fixed carbon, and lignin contents positively influenced HHV, whereas increased oxygen content decreases energy density. The RSM model gives a better prediction accuracy as compared to the MLR model because it takes into consideration the interaction and nonlinear effects of biomass constituents. The lignin inclusion as a structural parameter contributes greatly to the strength of the model. The suggested method is fast and economical compared to bomb calorimetry and can be applied to screen biomass, select feedstock, as well as optimize combustion, gasification, and pyrolysis systems.

KEYWORDS

Energy density, Higher heating value, Lignocellulosic biomass, Multiple linear regression, Response surface methodology.

1. INTRODUCTION

Biomass is a widely distributed, available renewable resource and a practical feedstock for low-carbon energy and power generation (Jasim and Khalil, 2025; Mia et al., 2025). Through thermochemical processes such as combustion, gasification, and pyrolysis, lignocellulosic materials can be converted into heat, syngas, bio-oil, or biochar, depending on process conditions and material properties (Onoroh, Ogbonnaya and Agberegha, 2023; Wang and Wu, 2023). A central fuel-quality parameter for all these processes is the higher heating value (HHV), which represents the maximum energy released during complete combustion and strongly influences system efficiency and design (Maksimuk et al., 2020) Standard HHV determination uses bomb calorimetry. Although accurate, calorimetry is time-consuming, requires specialized equipment, and is unsuitable for rapid screening across diverse biomass families (Daskin et al., 2024). Because HHV is composition-driven, empirical prediction using routine proximate, ultimate, and structural analyses offers a faster, lower-cost alternative (Dashti et al., 2019) (Mondal and Rafizul, 2025). Classical correlations such as Dulong-type and Channiwala–Parikh models are useful but often rely on narrow variable sets and can generalize poorly for heterogeneous lignocellulosic feeds (Mondal et al., 2023; Ardila, Figueroa and Maciel, 2024). Recent studies, therefore, apply statistical and machine learning approaches, including Multiple linear regression (MLR) and more flexible non-linear models, to raise predictive accuracy and robustness (Özkan et al., 2019; Dargahi et al., 2023; Ozule et al., 2025). Response surface methodology (RSM) is particularly attractive because it includes interaction and quadratic terms that capture combined compositional effects typical of biomass systems (Chen and Chen, 2025) . Yet, integrated RSM prediction using proximate, ultimate, and structural variables together, with lignin explicitly represented, remains limited in the open literature, especially for Nigerian lignocellulosic resources (Onokwai et al., 2022).

Despite the extensive amount of work that has been conducted to predict HHV of biomass, most of the current models rely on proximate or ultimate analysis, and as a result, are limited by their ability to identify the relationship between physicochemical and structural properties of lignocellulosic materials. Classical empirical correlations do not consider nonlinear effects and interaction, which decrease prediction accuracy for heterogeneous feedstocks (Maksimuk et al., 2020; Qian et al., 2020; Ibikunle et al., 2025). Furthermore, despite some studies on structural parameters, the explicit and simultaneous interaction between proximate, ultimate, and structural variables, especially lignin, within a single modelling system has not been established in detail. This research is novel as it formulates and compares the models of the integrated multiple linear regression and the response surface methodology, where the fixed carbon, elemental composition, and lignin are compared using the same experimental data, thus enhancing the reliability and practical applicability of HHV prediction.

2. MATERIALS AND METHODS

2.1. Sources of Materials and Pretreatment

Fifteen lignocellulosic biomasses; switchgrass, sunflower straw, lemon grass, siam weed, moringa oleifera, afara wood, iroko wood, sawdust, mansonia wood, pine chips, baobab wood, meliana wood, almond shell, corn stover, sugarcane straw, oil palm fibre, rice husk, and coconut shell were collected in Ota, Ogun State, Nigeria (6.6985° N, 3.2342° E). Samples were washed, air-dried for two days, then oven-dried at 105 ± 2 °C to constant mass. Dried materials were milled (Retsch SM100) and sieved to 0.2–0.4 mm. Subsamples were stored in desiccators and analyzed within four weeks.

2.2. Characterizations of Biomass

Samples were analyzed in triplicate for proximate, ultimate, structural composition, and HHV; results are mean ± SD (dry basis). Methods followed ASTM and NREL protocols. Proximate analysis measured MC (ASTM D3173-02), VM (ASTM D3175-20; 900 ± 10 °C for 7 min, oxygen-free), ash (ASTM D3174-02; 750 ± 10 °C), and FC by difference using Eq.1 (Tu et al., 2021). Ultimate analysis used LECO CHNS-628 for C, H, N (ASTM D5373-16) and S (ASTM D4239-11); O was by difference as shown in Eq.2 (Adeleke et al., 2021). Structural fractions were determined by NREL/TP-510-42618: solvent extraction, two-step acid hydrolysis, HPLC (Shimadzu LC-20AD; HPX-87H; 0.005 M H_2SO_4; 0.6 mL min^{-1}). Lignin was quantified by Klason method (ASTM D1106-21; Sluiter et al., 2008). HHV was measured by bomb calorimetry (ASTM D5865-13), and LHV was estimated from HHV using Eq.3. (Gómez-Vásquez, Fernández-Ballesteros and Camargo-Trillos, 2022).

$$FC\ (\%) = 100 - (MC + VM + Ash) \qquad (1)$$

$$O\ (\%) = 100 - (C + H + N + S + Ash) \qquad (2)$$

$$LHV\ \left(\frac{MJ}{kg}\right) = HHV - 2.442 \times (9H) \qquad (3)$$

Where FC is a fixed carbon, MC is the moisture content, and VM is the volatile matter. O, C, H, N, and S are associated with oxygen, carbon, hydrogen, nitrogen, and sulphur content, respectively. H = the hydrogen content in wt.%. The lower and higher heating values are LHV and HHV, respectively.

2.3. Data Modelling Procedures

The prediction of HHV was performed using a statistical modelling framework that combined multivariate linear regression (MLR) and response surface methodology (RSM). The modelling process involved data normalization, correlation analysis, regression model formulation, and response surface optimization. All statistical analyses were carried out using Design Expert (version 13) and R software (version 4.3.1) to ensure reproducibility and statistical consistency. HHV was modelled using MLR and RSM. Steps were data cleaning, normalization, correlation/VIF screening, MLR fitting, CCD-based RSM optimization, and validation. All analyses were run using Design Expert (version 13) and R software (version 4.3.1). Despite a relatively small sample size of biomass analysed (n = 15), the dataset has a large compositional space that cuts across a number of different lignocellulosic feedstocks, which is essential to multivariate modelling. The sample size is also consistent with RSM, which is found on central composite design and provides sufficient information to estimate quadratic and interaction effects. Variable screening and cross-validation provided model robustness. However, in future research, the size of the dataset could be increased to enhance the model's generalisability.

2.3.1. Normalizations

Proximate, ultimate, and structural data were checked for errors/outliers. Min-max and z-score normalization were both analyzed in the model development. Min-max scaling was utilized due to its ability to maintain the physical meaning and relative scale of the compositional variables, resulting in more consistent coefficients and sensitivity analysis in regression based models (Jadoon, Díaz and Rodríguez, 2025). With the limited dataset, z-score scaling also introduced an increased variability in the coefficients with no clear robustness advantage. This method thus ensures numerical stability while maintaining model interpretability Eq.4.

$$x_{norm} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{4}$$

In z-score normalization, each variable was standardized to a zero mean and unit variance as expressed by Eq.5

$$x_{norm} = \frac{X_i - \bar{X}}{s} \tag{5}$$

2.3.2. Correlation Analysis and Variable Selection

Pearson correlations identified strong HHV predictors ($|r| \geq 0.70$). Multicollinearity was handled with VIF; variables with VIF > 10 were removed or combined. In the variable selection process, volatile matter and ash were included during data preprocessing in the initial screening of correlation and multicollinearity. Predefined selection criteria were used to select variables, ensuring model parsimony and numerical stability. Volatile matter and the ash did not meet these criteria compared to retained predictors and were thus omitted in the final models, with their effect incorporated implicitly in correlated compositional variables.

2.3.3. Multivariate Linear Regression (MLR)

MLR related HHV to selected predictors via least squares; terms with $p < 0.05$ were retained. Model adequacy used R^2, adjusted R^2, F-tests, and residual checks in Eq.6 (Santiago, Guo and Sigman, 2018):

$$HHV = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \qquad (6)$$

Where β_0 is the intercept, β_1 are the regression coefficients corresponding to each predictor x_1, and ε represents the random error term.

2.3.4. Response Surface Methodology (RSM)

RSM captured curvature and interactions. FC, C, H, O, and lignin were fitted using CCD to a quadratic polynomial in Eq.7 (Reji and Kumar, 2023; Chen and Chen, 2025):

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{j>1}^{k} \beta_{ij} x_i x_j + \epsilon_i \qquad (7)$$

where β_0 is the intercept, β_i are the linear coefficients, β_{ii} are the quadratic coefficients, and β_{ij} are the interaction coefficients.

2.4. Model Validation

Models were assessed with R^2, adjusted R^2, root mean square error (RMSE) Eq.8, and mean absolute error (MAE) Eq.9. High R^2 and, with low RMSE/MAE/PRESS, indicate strong predictive performance (Chicco, Warrens and Jurman, 2021; Dargahi et al., 2023). Triplicate experimental measurements were conducted and expressed as mean values to minimize the uncertainty associated with random measurements. Regression residuals, error measures, and cross-validation outcomes capture the impacts of experimental variability on model predictions. Although the explicit uncertainty propagation analysis was not fully carried out, this study employed validation procedures that offer a realistic evaluation of the predictive reliability. Further research can broaden the models by including direct uncertainty propagation techniques.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(HHV_{exp,i} - HHV_{pred,i}\right)^2} \qquad (8)$$

$$MAE = \frac{1}{n}\sum_{i-1}^{n}\left[HHV_{exp,i} - HHV_{pred,i}\right] \qquad (9)$$

3. RESULTS

3.1. Proximate Analysis and Its Influence on Heating Value

Table 1 presents the proximate analysis of fifteen lignocellulosic biomass samples, such as MC, VM, FC, and Ash. The MC (5.2–10.8 wt%) showed a negative relation with HHV because energy is spent in evaporating water; MC ≤ 5 wt% is preferable for biomass conversion. VM (63.4–78.6 wt%) was high across samples and generally supported HHV by supplying flammable volatiles, also favouring thermochemical conversion processes like bio-oil formation in pyrolysis. FC (12.5–22.3 wt%) correlated positively with HHV as the solid combustible fraction and was a strong predictor. Ash (1.2–6.8 wt%) reduced HHV and may cause slagging/fouling; Ash < 5 wt% is more desirable (Sakulkit et al., 2020). The positive effect of fixed carbon on HHV indicates its contribution as the solid aromatic fraction that directly causes char formation and sustained heat release during combustion, while an increase in volatile matter leads to rapid devolatilization with a low net energy density. The high ash content lowers HHV by diluting combustible matter and raising the heat losses due to mineral matter.

Table 1. Proximate Analysis against Heating Value

SAMPL	MC (%)	VM (%)	FC (%)	Ash (%)	Exp HHV (MJ/kg)
Grass/Leaves					
Switchgrass	2.7 ±0.3	79.8 ± 0.03	12.8 ±0.01	4.7 ±0.02	16.4± 0.1
Sunflower straw	2.2 ±0.1	72.7 ± 0.03	17.0 ± 0.02	8.2 ±0.1	17.2 ±0.1
Lemon grass	0.4±0.04	66.5±0.20	24.2 ± 0.40	9.9 ± 0.02	17.5±0.1
Siam weed	0.1±0.02	77.3 ± 1.11	19.1 ±0.10	1.5 ± 0.02	20.2±0.2
Moringa oleifera	0.3±0.03	33.4 ± 0.03	45.7 ±0.14	20.2 ±0.02	16.3±0.1
Woody/Forestry					
Afara wood	2.1 ±0.1	88.0 ± 0.03	9.0 ± 0.02	0.9 ±0.01	22.6± 0.3
Iroko wood	4.4 ±0.1	76.7 ± 0.03	15.9 ±0.02	2.94 ±0.1	18.9± 0.2
Sawdust	3.5 ±0.3	81.5 ±0.03	13.6 ±0.01	1.5 ±0.01	20.1± 0.1
Mansonia wood	1.8 ±0.1	81.7 ± 0.03	14.21 ± 0.02	2.31 ±0.11	19.6± 0.1
Pine chips	1.3 ±0.1	71.7 ± 0.03	22.33 ± 0.02	4.7 ±0.01	19.5± 0.2
Baobab wood	0.2±0.04	82.3 ± 0.30	14.1 ±0.12	1.2 ±0.02	21.9 ±0.5
Meliana wood	0.4±0.05	82.80± 0.15	14.79 ± 0.05	2.06± 0.01	20.4±0.2
Agro-wastes					
Almond shell	3.4±0.1	74.27 ±0.03	19.2 ±0.02	3.2 ±0.01	19.1± 0.1
Corn stover	2.9 ±0.3	68.9 ± 0.03	22.5 ±0.01	5.8 ±0.02	17.3± 0.1
Sugarcane straw	2.5 ±0.3	81.2 ± 0.03	13.4 ± 0.01	2.9 ± 0.02	16.7 ±0.1
Oil Palm Fibre	5.6 ±0.1	74.4 ± 0.03	16.5 ±0.02	3.5 ±0.01	18.7± 0.1
Rice husk	0.3±0.01	62.6 ± 1.15	17.4 ±0.10	21.0± 0.01	16.5±0.3
Coconut shell	0.3±0.01	78.5 ± 1.25	19.1 ± 0.50	1.9 ± 0.01	19.4 ±0.3

3.2. Ultimate Analysis and Its Influence on Heating Value

Table 2 presents the elemental composition of the biomass samples. These elements influence the energy content, combustion chemistry, and emission characteristics of the biomass samples. The C (41.8–54.6 wt%) and H (4.9–6.6 wt%) contributed positively to HHV via high combustion enthalpy. O (36.5–48.2 wt%) contributed negatively by diluting combustibles. N (0.3–1.2 wt%) and S (<0.5 wt%) had little effect on HHV but inform emissions; low values indicate cleaner fuels (Moulay and Jang, 2025). An increase in carbon and hydrogen content increased the HHV due to greater bond energies and exothermic oxidation reactions, while the HHV decreased with an increase in oxygen content due to partial oxidation and lower effective fuel value. The small effect of nitrogen and sulphur on HHV is due to their low concentrations, and as they are more relevant to the formation of emissions, as opposed to the energy release.

Table 2. Ultimate Analysis against Heating Value

SAMPLE	C (%)	H (%)	N (%)	O (%)	S (%)	Exp HHV (MJ/kg)
Grass/Leaves						
Switchgrass	49.2 ±0.3	6.3± 0.02	0.31±0.01	44.1 ±0.2	0. 09 ± 0.03	16.4± 0.1
Sunflower straw	48.9 ±0.3	5.6± 0.03	0.9 ± 0.01	44.2± 0.2	0.2 ± 0.01	17.2 ±0.1
Lemon grass	43.3± 0.2	5.2± 0.03	1.8 ± 0.01	49.7± 0.2	0.04± 0.004	17.5± 0.1
Siam weed	42.6± 0.3	5.5± 0.01	1.6 ± 0.01	49.2± 0.2	0.20 ± 0.01	20.2± 0.2
Moringa olefeira	46.4± 0.2	5.0± 0.02	3.0 ± 0.01	43.8± 0.1	0.5 ± 0.02	16.3± 0.1
Woody/Forestry						
Afara wood	48.9 ±0.1	6.0± 0.02	0. 1 ± 0.02	44.9±0.1	0.02 ± 0.01	22.6± 0.3
Iroko wood	48.8 ±0.1	5.9± 0.01	0.4±0.02	44.9 ±0.1	0. 02 ± 0.01	18.9± 0.2
Sawdust	49.3 ±0.3	6.3 ±0.03	0.4± 0.01	43.9 ±0.2	0.03 ±0.001	20.1± 0.1
Mansonia wood	50.9 ±0.1	5.6± 0.04	0.02 ±0.02	43.3 ±0.1	0.01 ±0.001	19.6± 0.1
Pine chips	49.4 ±0.1	5.8± 0.03	0.4± 0.02	37.3±0.1	0.05 ± 0.01	19.5± 0.2
Baobab wood	47.6 ±0.3	6.8± 0.03	0.7 ± 0.01	44.6 ± 0.2	0.32 ± 0.01	21.9±0.5
Meliana wood	45.8 ±0.1	5.6± 0.03	1.3 ± 0.02	35.9 ±0.1	0.11± 0.001	20.4± 0.2
Agro-wastes						
Almond shell	50.0 ±0.1	5.6 ±0.03	0.5 ± 0.02	43.7±0.1	0.07 ± 0.01	19.1± 0.1
Corn stover	47.8 ±0.3	5.7± 0.03	0.6 ± 0.01	45.8 ±0.2	0.10 ± 0.03	17.3± 0.1
Sugarcane straw	40.3 ±0.2	6.1± 0.01	2.0 ± 0.02	26.7± 0.2	0.40 ± 0.01	16.7 ±0.1
Oil Palm Fibre	47.4 ±0.1	6.0± 0.06	0.9 ± 0.02	48.9 ±0.1	0.23 ± 0.01	18.7± 0.1
Rice husk	42.4± 0.1	5.3± 0.03	0.3 ± 0.02	52.1± 0.3	0.02± 0.002	16.5± 0.3
Coconut shell	47.4 ± 0.1	6.1 ± 0.03	0.2± 0.01	44.2± 0.1	0.03 ± 0.003	19.4 ± 0.3

3.3. Structural Composition Analysis and Its Influence on Heating Value

Cellulose (32.4–46.7%) contributes mainly through volatile release during decomposition. Hemicellulose (20.1–31.5%) decomposes earlier, is more oxygenated, and weakly supports HHV. Lignin (17.8–28.9%) is aromatic, thermally stable, and yields energy-dense char; it shows the strongest positive relation with HHV and remains a dominant predictor.

Table 3. Structural Composition Analyses against Heating Value

SAMPLE	Cellulose %	Hemicellulose %	Lignin %	Exp HHV (MJ/kg)
Grass/Leaves				
Switchgrass	44.7±1.1	31.5 ± 0.2	21.7 ± 0.3	16.4± 0.1
Sunflower straw	37.7±0.2	25.9± 0.2	16.4 ±0.3	17.2 ±0.1
Lemon grass	38.2 ± 0.1	26.5 ± 0.1	21.9 ± 0.4	17.5 ± 0.1
Siam weed	29.6 ± 0.1	25.7 ± 0.1	30.3 ± 0.2	20.2 ± 0.2
Moringa oleifera	14.1 ± 0.1	12.3 ± 0.1	9.1 ± 0.1	16.3 ± 0.1
Woody/Forestry				
Afara wood	46.8±0.1	30.7 ± 0.4	22.5±0.1	22.6± 0.3
Iroko wood	44.7±1.2	21.8 ±0.2	35.5 ± 0.1	18.9± 0.2
Sawdust	40.4± 0.1	30.6 ±0.2	29 ±0.2	20.1± 0.1
Mansonia wood	45.9 ± 0.4	33.6± 0.2	20.5 ±0.3	19.6± 0.1
Pine chips	43.2±0.3	23.2 ± 0.1	26.4 ±0.1	19.5± 0.2
Baobab wood	44.4 ± 0.2	28.8 ± 0. 1	23.4 ± 0.1	21.9 ±0.5
Meliana wood	44.6 ± 0.2	29.6 ± 0.2	24.3 ± 0.3	20.4 ± 0.2
Agro-wastes				
Almond shell	39.8 ± 0.1	29.5 ± 0.1	30.6 ±0.2	19.1± 0.1
Corn stover	41.6 ± 0.2	30.0 ±0.2	18.4 ±0.2	17.3± 0.1
Sugarcane straw	42.7 ± 0.9	26.2 ± 0.2	30.0 ±0.3	16.7 ±0.1
Oil Palm Fibre	39.4±0.2	35.7 ± 0.1	24.7 ±0.1	18.7± 0.1
Rice husk	35.3 ± 0.5	22.2 ± 0.2	18.5 ± 0.4	16.5 ± 0.3
Coconut shell	33.1 ± 0.1	26.6 ± 0.2	38.8 ± 0.1	19.4 ± 0.3

3.4. Experimental HHV

Fig.1 shows HHV of 16.8–23.7 MJ kg^{-1}. Higher HHVs align with greater FC, lignin, C, and H; lower HHVs align with higher O and ash. MC and Ash showed negative effects consistent with established trends (Yin, 2011) (Qian et al., 2020). The HHV spread is moderate, supporting relatively uniform fuel potential despite compositional differences. This enhances co-firing and co-gasification without any major changes in thermal performance.

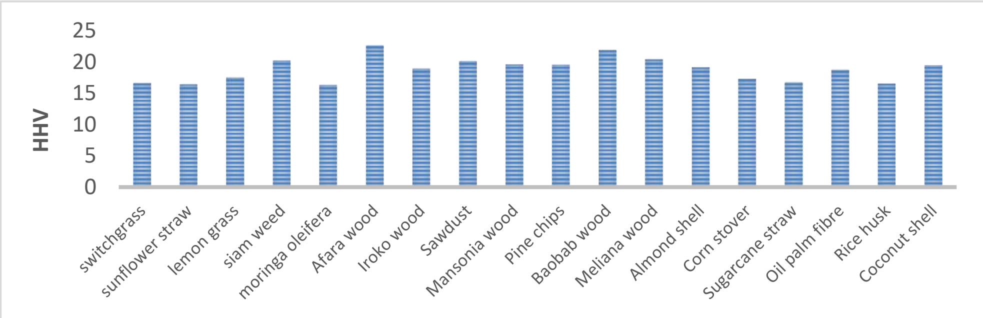

Fig. 1. HHV for Various Lignocellulosic Biomass Samples

3.5. Model Development

The HHV was modelled using key compositional predictors from proximate, ultimate, and structural analyses. The FC, C, H, O, and Li contents were selected, and both MLR and RSM were fitted against experimental HHV.

3.5.1. MLR Model

The MLR model Eq.7 related HHV linearly to FC, C, H, O, and Li. The analysis of variance (ANOVA) and adequacy statistics for this model are presented in Table 4. The ANOVA confirmed significance ($F = 36.24$; $p < 0.0001$), with all predictors significant at 95% confidence. The FC, C, and Li were the strongest positive contributors, while O had a significant negative effect, reflecting energy dilution from higher oxidation. Model fit was strong ($R^2 = 0.977$; $R^2adj = 0.968$; $R^2pred = 0.956$), with low errors (RMSE = 0.138 MJ kg^{-1}; MBD = 0.013 MJ kg^{-1}; MRE = 0.84%). Adequate precision (8.92) exceeded the minimum threshold, and lack-of-fit was not significant ($p = 0.563$).

Although effective for direct linear effects, MLR cannot capture curvature or interactions that may limit proportional HHV gains at high FC or Li.

$$HHV\left(\frac{MJ}{kg}\right) = 3.218 + 0.165FC + 0.348C + 0.528H - 0.114O + 0.236L \quad (7)$$

All independent variables are in weight percent (wt%).

Table 4. ANOVA and adequacy statistics for the MLR model predicting HHV

Source	Sum of Squares	DF	Mean Square	F-value	P-values	Remark
Model	24.285	5	4.857	36.24	< 0.0001	Significant
A-Fixed Carbon (FC)	6.412	1	6.412	47.81	0.0002	Significant
B-Carbon (C)	5.936	1	5.936	44.27	0.0003	Significant

Source	Sum of Squares	DF	Mean Square	F-value	P-values	Remark
C-Hydrogen (H)	3.187	1	3.187	23.79	0.0011	Significant
D-Oxygen (O)	2.814	1	2.814	20.74	0.0018	Significant
E-Lignin (L)	4.275	1	4.275	31.89	0.0006	Significant
Residual	1.073	9	0.119			
Lack of Fit	0.423	4	0.106	0.84	0.563	not significant
Pure Error	0.650	5	0.130			
Cor Total	25.358	14				

$R^2 = 0.977$; Adjusted $R^2 = 0.968$; Predicted $R^2 = 0.956$; RMSE = 0.138 $MJkg^{-1}$; MBD = 0.013 $MJkg^{-1}$; MRE = 0.84%; PRESS = 1.28; Adeq Precision = 8.92

3.5.2. Response Surface Methodology Modelling

The RSM model Eq.8 incorporated linear, interaction, and squared terms to capture nonlinear behaviour. FC, C, and Li remained dominant positive drivers; negative FC^2, C^2, and Li^2 terms indicate diminishing HHV returns beyond optimal levels. Likewise, the significant positive interactions (FC×C, FC×H, FC×O, FC×Li, C×H, C×Li) show synergistic effects between carbonaceous and structural fractions. The ANOVA Table 5 confirmed strong significance ($F = 96.31$; $p < 0.0001$) and a non-significant lack-of-fit ($p = 0.618$). The Fit and prediction were excellent ($R^2 = 0.993$; $R^2adj = 0.991$; $R^2pred = 0.987$), with very low error (RMSE = 0.062 MJ kg^{-1}) and high signal strength (Adeq Precision = 21.34), outperforming MLR.

$$HHV = 21.49 + 0.36FC + 0.35C + 0.23H - 0.19O + 0.25Li + 0.07(FC \times C) + 0.08(FC \times H) + 0.09(FC \times O) + 0.06(FC \times Li) + 0.06(C \times H) + 0.04(C \times Li) - 0.13FC^2 - 0.12C^2 - 0.14Li^2 \quad (8)$$

All independent variables are in weight percent (wt%).

Table 5. ANOVA and adequacy statistics for the RSM model predicting HHV

Source	Sum of Squares	DF	Mean Square	F-value	P-values	Remark
Model	24.967	14	1.783	96.31	<0.0001	Significant
Fixed Carbon (FC)	3.485	1	3.485	188.42	<0.0001	Significant
Carbon (C)	2.836	1	2.836	153.28	<0.0001	Significant
Hydrogen (H)	1.257	1	1.257	67.93	0.0001	Significant
Oxygen (O)	1.019	1	1.019	54.99	0.0002	Significant
Lignin (L)	2.211	1	2.211	119.42	<0.0001	Significant
FC × C	0.364	1	0.364	19.66	0.0020	Significant
FC × H	0.279	1	0.279	15.21	0.0036	Significant
FC × O	0.210	1	0.210	11.54	0.0068	Significant
FC × L	0.243	1	0.243	13.54	0.0048	Significant
C × H	0.172	1	0.172	9.62	0.0097	Significant
C × L	0.198	1	0.198	10.81	0.0079	Significant
FC^2	0.412	1	0.412	17.92	0.0025	Significant
C^2	0.379	1	0.379	15.84	0.0032	Significant
L^2	0.420	1	0.420	18.32	0.0023	Significant
Residual (Error)	0.164	9	0.018			
Lack of Fit	0.055	4	0.014	0.71	0.618	not significant
Pure Error	0.109	5	0.022			
Cor Total	25.131	23				

$R^2 = 0.993$; Adjusted $R^2 = 0.991$; Predicted $R^2 = 0.987$; RMSE = 0.062 $MJkg^{-1}$; MBD = 0.04 $MJkg^{-1}$; MRE = 0.37%; PRESS = 0.37; Adeq Precision = 21.34

3.6. Model Validation and Comparison

Predicted HHVs matched experiments closely Table 6. The RSM residuals (0.01–0.04 MJ kg^{-1}) were smaller than MLR residuals (0.05–0.22 MJ kg^{-1}). RSM also showed higher explanatory power and lower error (R^2 = 0.993; RMSE = 0.062 MJ kg^{-1}) than MLR (R^2 = 0.975; RMSE = 0.137 MJ kg^{-1}), confirming superior precision and stability Table 7. Fig. 2 shows a near-perfect overlap between experimental and RSM predictions, whereas MLR deviates slightly at higher HHV. Parity and residual plots Fig. 3 further confirm unbiased, tightly clustered RSM errors versus wider MLR spread.

Table 6. Comparison of experimental and predicted HHV values using MLR and RSM

Run	FC (wt%)	C (wt%)	H (wt%)	O (wt%)	Li (wt%)	Exp. HHV (MJ kg^{-1})	Pred. HHV (MLR)	Pred. HHV (RSM)	Residual (MLR)	Residual (RSM)
1	18.5	45.8	5.8	36.1	20.3	18.62	18.57	18.61	0.05	0.01
2	20.3	46.9	6.0	35.0	23.4	19.42	19.33	19.41	0.09	0.01
3	22.7	47.5	6.2	33.8	25.6	20.18	20.05	20.14	0.13	0.04
4	24.1	48.2	6.4	32.6	27.1	20.75	20.56	20.70	0.19	0.05
5	25.4	48.5	6.3	32.0	28.5	21.10	20.91	21.07	0.19	0.03
6	26.0	49.0	6.5	31.8	29.3	21.42	21.20	21.38	0.22	0.04
7	27.3	49.6	6.6	31.0	30.0	21.65	21.54	21.61	0.11	0.04
8	28.5	50.1	6.8	30.3	31.5	21.93	21.80	21.89	0.13	0.04
9	29.6	50.7	6.9	29.9	32.7	22.15	22.00	22.11	0.15	0.04
10	30.8	51.3	7.0	29.1	33.8	22.42	22.29	22.40	0.13	0.02
11	31.5	51.8	7.1	28.6	34.6	22.56	22.44	22.53	0.12	0.03
12	32.1	52.0	7.2	28.2	35.0	22.68	22.53	22.65	0.15	0.03
13	33.0	52.3	7.3	27.8	35.5	22.81	22.66	22.77	0.15	0.04
14	33.8	52.8	7.4	27.2	36.1	22.93	22.80	22.89	0.13	0.04
15	34.2	53.0	7.5	26.9	36.5	23.02	22.90	22.99	0.12	0.03

Table 7. Model Performance Metrics

Metric	MLR	RSM
Coefficient of determination (R^2)	0.975	0.993
Adjusted R^2	0.969	0.991
Predicted R^2	0.957	0.987
Root Mean Square Error (RMSE, MJ kg^{-1})	0.137	0.062
Mean Bias Deviation (MBD, MJ kg^{-1})	0.012	0.004
Mean Relative Error (MRE, %)	0.83	0.37
Predicted Residual Sum of Squares (PRESS)	1.28	0.37
Adequate Precision	8.92	21.34

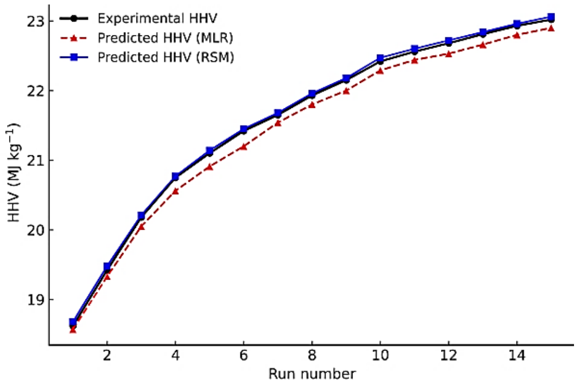

Fig. 2. Comparison of experimental and predicted HHV values using MLR and RSM models

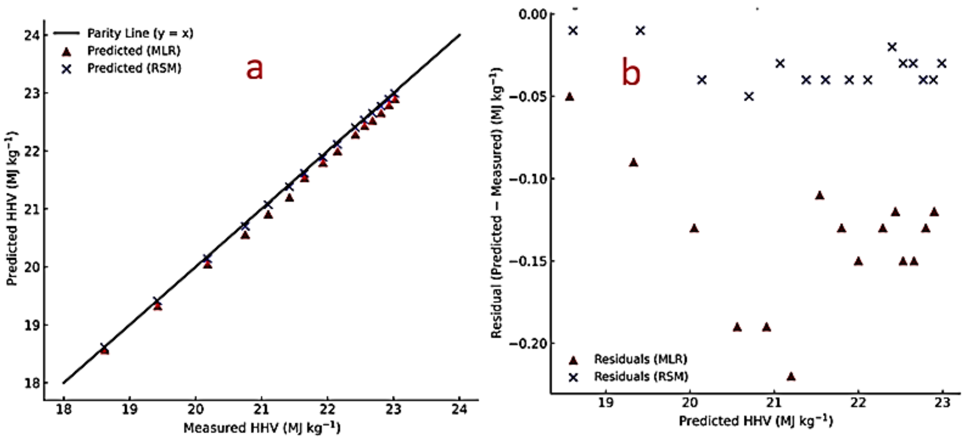

Fig. 3. (a) Parity plot comparing experimental and predicted HHV values using MLR and RSM models (b)Residual distribution plots for MLR and RSM models

Compared with published correlations Table 8, the present approach improves accuracy by integrating proximate (FC), ultimate (C, H, O), and structural (Li) data in one framework. The addition of lignin, which is a highly aromatic and high-energy polymer, strengthens prediction by capturing both the physicochemical and structural composition influences on HHV, while RSM's quadratic and interaction terms better represent biomass compositional coupling.

Table 8. Validation and comparison with published correlation for HHV of lignocellulose biomass based on ultimate and structural composition analysis

Correlation Equation	based on	R^2	MAE	MBE	Authors
HHV = 17.79 + 0.03FC + 0.01VM − 0.16 Ash	Proximate analysis	0.94	3.52	0.27	(Xing et al., 2019)
HHV = −30.4FC + 65.3Ash^2 + 55.4FC − 48.5Ash + 9.591	Proximate analysis	0.94	2.59	0.08	(Qian et al., 2020)
HHV = 3.918+0.276C+0.016H	Ultimate analysis	0.90	3.86	0.84	(Xing et al., 2019)
HHV=−0.004(−19.98FC$^{1.226}$ − 1.029×10^{-13}VM$^{8.07}$+0.1026Ash$^{2.4}$ ^{2}−1.207×10−7(FC×ASH$^{4.67}$) + 0.023(FC×VM×Ash) − 0.251(VM/ Ash)) − 0.048(FC/VM) + 15.72	Proximate analysis	0.96	-	-	(Dashti et al., 2019)
HHV = 32.90C+162.70H−16.2O − 954.40S + 1.41	Ultimate analysis	0.95	1.12	0.05	(Qian et al., 2020)
HHV = 0.405C-0.004Li	Ultimate/structural composition analysis	0.98	2.66	0.11	(Maksimuk et al., 2020)
HHV = 4.91+0.26N+0.11C +0.61H+0.39S+0.021O	Ultimate analysis	0.94	-	-	(Özyuğuran, Yaman and Küçükbayrak, 2018)
HHV = 1.3849+85.0807C-28.9675H-666.125N+11.6296S	Ultimate analysis	0.98	-	-	(Ibikunle et al., 2025)
HHV = 8.2081+0.4975Li-0.0608Ex	Structural composition analysis	0.81	-	-	(Maksimuk et al., 2020)

* FC = Fixed carbon; VM = Volatile Matter; C = Carbon; H = Hydrogen; O = Oxygen; N = Nitrogen; S = Sulphur; Li = Lignin; Ex = Extractive

3.6.1. Effect of the two most important parameters on the HHV

Fig.4 presents the 3D surface and contour plots, showing the effect of FC and C on the HHV. The HHV increases with FC and C, then plateaus, showing diminishing gains at high levels due to quadratic effects Fig. 4(a). Elliptical contours Fig. 4(b) confirm a strong FC×C interaction. The predicted optimum HHV (21.96 MJ kg^{-1}) occurs at FC = 27.85 wt% and C = 49.95 wt%, with desirability 0.997. This reaction pattern shows that a concurrent increase of FC and C allows attainment of maximum heat production until the optimum composition range is attained.

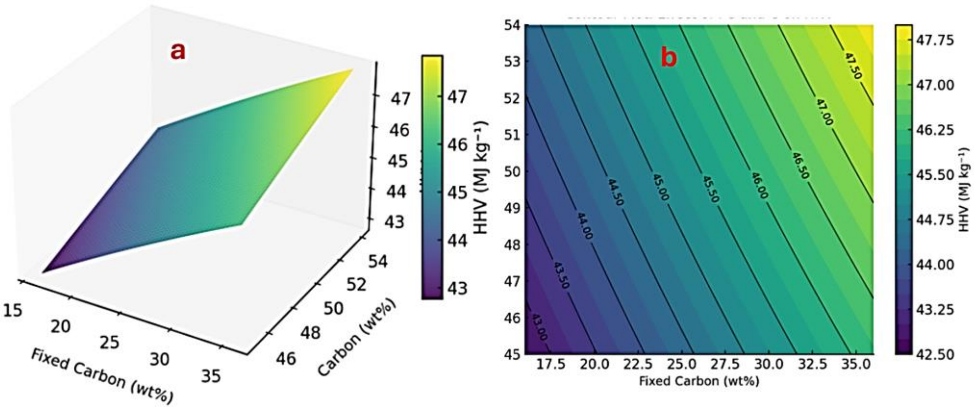

Fig. 4. (a) 3D surface plot: Effect of fixed carbon (wt%) and carbon on HHV (b)2D contour plot: Effect of fixed carbon (wt%) and carbon on HHV

In Fig. 5(a), the HHV rises with both FC and Li to an optimum (21.96 MJ kg-1) at FC = 27.85 wt% and Li = 30.65 wt%, after which negative quadratic terms reduce marginal gains. Well-defined elliptical contours Fig. 5(b) confirm a significant and positive FC x Li interaction term. This trend indicates that the Li and FC jointly enhance the density of biomass energy till structural balance is obtained.

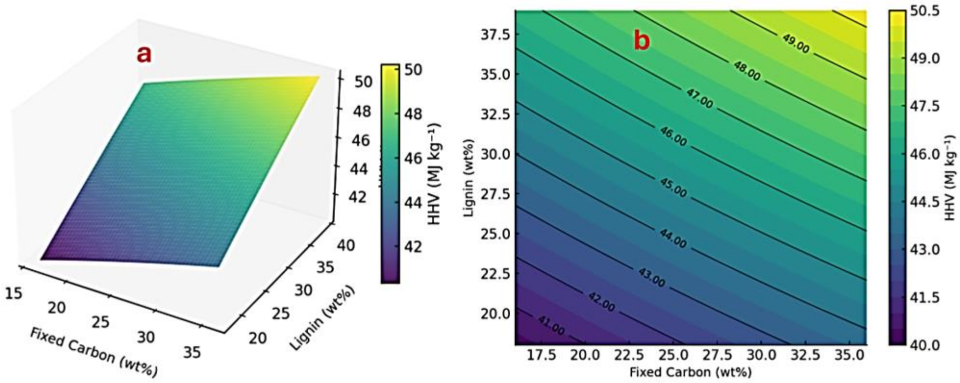

Fig. 5. (a) 3D surface plot: Effect of fixed carbon (wt%) and lignin on HHV (b)2D contour plot: Effect of fixed carbon (wt%) and lignin on HHV

The two most significant interaction effects of C and Li on HHV, depicted in Fig. 6(a), showed a smooth, gradually curved response surface. HHV increases smoothly with C and Li, with weaker interaction than FC-based pairs. Curvature reflects C^2 and Li^2 terms Fig. 6(b), confirming compositional balance, where the highest energy output is achieved when C and Li are in balanced proportions. An optimum HHV (21.96 MJ kg-1) was attained at C = 49.95 wt.% and Li = 30.65 wt.%. Beyond these levels, the response surface becomes a little flatter. This tendency depicts that an optimal proportion of aromatic and carbonaceous materials is required to enhance the best thermal performance in the feedstock.

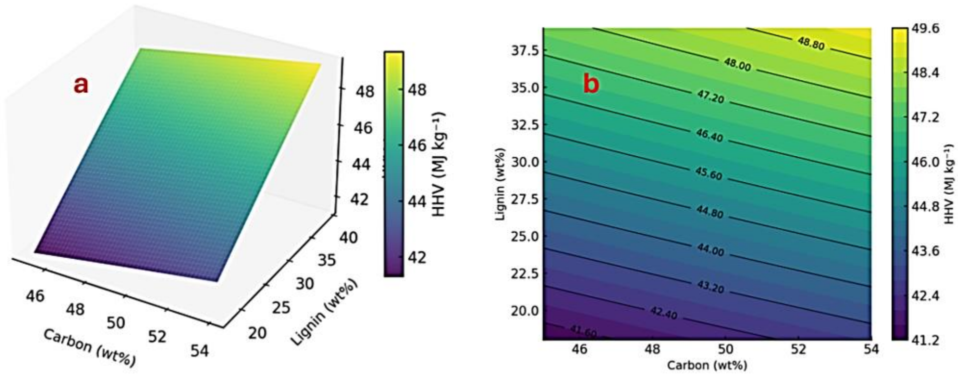

Fig. 6. (a) 3D surface plot: Effect of carbon (wt%) and lignin on HHV (b)2D contour plot: Effect of carbon (wt%) and lignin on HHV

3.6.2. Effect Implications for Thermochemical Conversion

The models aid the appropriate selection of biomass for thermochemical conversion processes because its HHV is well predicted from FC, C, H, O, and Li. High FC, C, and Li correspond to higher HHV and improved combustion efficiency; high O or VM lowers HHV and can reduce combustion stability. For gasification, higher predicted HHV implies favorable carbon-to-oxygen ratios and higher-quality syngas with reduced coking risk. For pyrolysis, Li- and FC-rich biomasses are expected to yield higher-energy char and improved bio-oil heating value, helping optimize operating conditions and blending strategies.

4. CONCLUSIONS

The MLR and RSM were successfully developed for HHV prediction in lignocellulosic biomass. RSM outperformed MLR and classical correlations, showing near-complete variance explanation ($R^2 = 0.993$) and low error (RMSE = 0.062 MJ kg^{-1}). HHV increases mainly with FC, C, H, and Li, while O reduces energy density; interaction and quadratic effects highlight the need for compositional balance. The RSM model offers a rapid, reliable alternative to exhaustive calorimetry and can guide feedstock selection and optimization in combustion, gasification, and pyrolysis. The models are a scientifically based and data-driven framework that evaluates the energy potential of biomaterials, which contributes significantly to the advancement of cleaner, efficient, and sustainable technologies in the manufacture of bioenergy. Despite the high predictive ability of the developed model, the study is constrained by the size of the sample and biomass types that are subjected to analysis. Future studies could consider increasing the dataset to include more biomass categories while utilizing explicit uncertainty propagation and advanced machine learning to make the model more generalizable and robust.

ACKNOWLEDGMENTS
The authors appreciate the support provided by Tshwane University of Technology, South Africa, for this study.

5. REFERENCES

Adeleke, A. A., Odusote, J. K., Ikubanni, P. P., Agboola, O. O., Balogun, A. O. and Lasode, O. A. (2021) 'Tumbling strength and reactivity characteristics of hybrid fuel briquette of coal and biomass wastes blends', Alexandria Engineering Journal, 60(5), pp. 4619-4625, available: https://doi.org/10.1016/j.aej.2021.03.069.

Ardila, Y. C., Figueroa, J. E. J. and Maciel, M. R. W. (2024) 'Mathematical models for predicting the higher heating value and ultimate analysis of biomass', Industrial Crops and Products, 208, p. 117777, available: https://doi.org/10.1016/j.indcrop.2023.117777.

Chen, H. and Chen, C. (2025) 'A study of the response surface methodology model with regression analysis in three fields of engineering', Applied System Innovation, 8(4), p. 99, available: https://doi.org/10.3390/asi8040099.

Chicco, D., Warrens, M. J. and Jurman, G. (2021) 'The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation', PeerJ Computer Science, 7, p. e623, available: https://doi.org/10.7717/peerj-cs.623.

Dargahi, P., Nasseri, S., Hadi, M., Nodehi, R. N. and Mahvi, A. H. (2023) 'Prediction models for groundwater quality parameters using a multiple linear regression (MLR): a case study of Kermanshah, Iran', Journal of Environmental Health Science & Engineering, 21(1), pp. 63-71, available: https://doi.org/10.1007/s40201-022-00836-9.

Dashti, A., Noushabadi, A. S., Raji, M., Razmi, A., Ceylan, S. and Mohammadi, A. H. (2019) 'Estimation of biomass higher heating value (HHV) based on the proximate analysis: Smart modeling and correlation', Fuel, 257, p. 115931, available: https://doi.org/10.1016/j.fuel.2019.115931.

Daskin, M., Erdogan, A., Fatih, G. and Okolie, J. (2024) 'Generalizability of empirical correlations for predicting higher heating values of biomass', Energy Sources, Part a: Recovery, Utilization, and Environmental Effects, 46, pp. 5434-5450, available: https://doi.org/10.1080/15567036.2024.2332472.

Gómez-Vásquez, R., Fernández-Ballesteros, E. and Camargo-Trillos, D. (2022) 'Biogenic nanoporous oxides recovery from by-products of bioenergy production: Rice husks and corncob biochars', Biomass and Bioenergy, 161, p. 106455, available: https://doi.org/10.1016/j.biombioe.2022.106455.

Ibikunle, R. A., Lukman, A. F., Titiladunayo, I. F., Akeju, E. A. and Dahunsi, S. O. (2025) 'Modeling and robust prediction of high heating values of municipal solid waste based on ultimate analysis', Energy Sources, Part a: Recovery, Utilization, and Environmental Effects, 47(1), pp. 2239-2256, available: https://doi.org/10.1080/15567036.2020.1841343.

Jadoon, U. K., Díaz, I. and Rodríguez, M. (2025) 'Comparative analysis of aspen plus simulation strategies for woody biomass air gasification processes', Biomass and Bioenergy, 194, p. 107626, available: https://doi.org/10.1016/j.biombioe.2025.107626.

Jasim, I. N. and Khalil, A. H. (2025) 'Enhanced biogas generation from lignocellulosic biomass with the implementation of a chemical pre-treatment procedure', Kufa Journal of Engineering, 16(4), pp. 148-160, available: https://doi.org/10.30572/2018/KJE/160408.

Maksimuk, Y., Antonava, Z., Krouk, V., Korsakova, A. and Kursevich, V. (2020) 'Prediction of higher heating value based on elemental composition for lignin and other fuels', Fuel, 263, p. 116727, available: https://doi.org/10.1016/j.fuel.2019.116727.

Mia, S., Ahmed, F., Khan, I., Kabir, M. E., Roni, M. H., Cobra, K., Khatun, A. A. and Mahmud, S. (2025) 'Integrating renewable energy with internet of things (IoT): Pathways to a smart green planet', Kufa Journal of Engineering, 16(1), pp. 361-404, available: https://doi.org/10.30572/2018/KJE/160121.

Mondal, C., Pal, S. K., Samanta, B., Dutta, D. and Raj, S. (2023) 'Analysis and significance of prediction models for higher heating value of coal: an updated review', Journal of Thermal Analysis and Calorimetry, 148(15), pp. 7521-7538, available: https://doi.org/10.1007/s10973-023-12272-4.

Mondal, S. and Rafizul, I. M. (2025) 'Predicting calorific value through proximate analysis of municipal solid waste using soft computing system', Discover Applied Sciences, 7(3), p. 212, available: https://doi.org/10.1007/s42452-025-06643-9.

Moulay, I. and Jang, K. (2025) 'Sustainable approaches to NOx emissions: Capture and utilization technologies', Next Energy, 8, p. 100356, available: https://doi.org/10.1016/j.nxener.2025.100356.

Onokwai, A. O., Okokpujie, I. P., Ajisegiri, E. S., Oki, M., Adeoye, A. O. and Akinlabi, E. T. (2022) 'Characterization of lignocellulosic biomass samples in omu-aran metropolis, kwara state, nigeria, as potential fuel for pyrolysis yields.', International Journal of Renewable Energy Development, 11(4) available: https://doi.org/10.14710/ijred.2022.45549.

Onoroh, F., Ogbonnaya, M. and Agberegha, L. O. (2023) 'Performance evaluation of hot air thermoelectric generator using biomass energy source', Kufa Journal of Engineering, 14(4), pp. 69-85, available: https://doi.org/10.30572/2018/KJE/140406.

Özkan, K., Işık, Ş., Günkaya, Z., Özkan, A. and Banar, M. (2019) 'A heating value estimation of refuse derived fuel using the genetic programming model', Waste Management, 100, pp. 327-335, available: https://doi.org/10.1016/j.wasman.2019.09.035.

Ozule, C. P., Oluwo, A., Alozie, N. S., Rajan, J., Jose, S., Oke, S. A. and Nwankiti, U. S. (2025) 'A grey wolf optimization approach for evaluating the engine responses of various biodiesel blends in an internal combustion engine', Kufa Journal of Engineering, 16(1), pp. 298-323, available: https://doi.org/10.30572/2018/KJE/160118.

Özyuğuran, A., Yaman, S. and Küçükbayrak, S. (2018) 'Prediction of calorific value of biomass based on elemental analysis', International Advanced Researches and Engineering Journal, 2(3), pp. 254-260, available: https://dergipark.org.tr/en/download/article-file/593749 (Accessed: 31 October 2025).

Qian, C., Li, Q., Zhang, Z., Wang, X., Hu, J. and Cao, W. (2020) 'Prediction of higher heating values of biochar from proximate and ultimate analysis', Fuel, 265, p. 116925, available: https://doi.org/10.1016/j.fuel.2019.116925.

Reji, M. and Kumar, R. (2023) 'Response surface methodology (RSM): An overview to analyze multivariate data', Indian Journal of Microbiology Research, 9, pp. 241-248, available: https://doi.org/10.18231/j.ijmr.2022.042.

Sakulkit, P., Palamanit, A., Dejchanchaiwong, R. and Reubroycharoen, P. (2020) 'Characteristics of pyrolysis products from pyrolysis and co-pyrolysis of rubber wood and oil palm trunk biomass for biofuel and value-added applications', Journal of Environmental Chemical Engineering, 8, p. 104561, available: https://doi.org/10.1016/j.jece.2020.104561.

Santiago, B. C., Guo, J. and Sigman, M. S. (2018) 'Predictive and mechanistic multivariate linear regression models for reaction development', Chemical Science, 9(9), pp. 2398-2412, available: https://doi.org/10.1039/C7SC04679K.

Tu, R., Sun, Y., Wu, Y., Fan, X., Cheng, S., Jiang, E. and Xu, X. (2021) 'A new index for hydrochar based on fixed carbon content to predict its structural properties and thermal behavior', Energy, 229, p. 120572, available: https://doi.org/10.1016/j.energy.2021.120572.

Wang, Y. and Wu, J. J. (2023) 'Thermochemical conversion of biomass: Potential future prospects', Renewable and Sustainable Energy Reviews, 187, p. 113754, available: https://doi.org/10.1016/j.rser.2023.113754.

Xing, J., Luo, K., Wang, H., Gao, Z. and Fan, J. (2019) 'A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches', Energy, 188, p. 116077, available: https://doi.org/10.1016/j.energy.2019.116077.

Yin, C.-Y. (2011) 'Prediction of higher heating values of biomass from proximate and ultimate analyses', Fuel, 90(3), pp. 1128-1132, available: https://doi.org/10.1016/j.fuel.2010.11.031.