FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 11, November, 2025, pp 441 – 448
DOI: https://doi.org/10.33003/fjs-2025-0911-4009

PERFORMANCE EVALUATION OF TRAINING FUNCTIONS IN NEURAL NETWORKS FOR OPTIMAL BIOMASS ENERGY CONTENT PREDICTION

***Aisha Saeed Bello, Usman Alhaji Dodo**

Department of Electrical and Computer Engineering, Faculty of Engineering, Baze University, Abuja, Nigeria.

*Corresponding authors' email: aisha6956@bazeuniversity.edu.ng

ABSTRACT

This study compares the performance of ten training functions in artificial neural networks for predicting the higher heating value (HHV) of biomass using ultimate analysis. A 5-10-1 feed-forward back-propagation neural network architecture was implemented, with data normalized and divided into 70% training and 30% testing. Model accuracy was assessed using the coefficient of determination (R^2), mean squared error (MSE), and mean absolute error (MAE). Results obtained showed that the Bayesian Regularization algorithm (trainbr, M3) outperformed other models, achieving an R^2 of 0.8358, MSE of 0.001432, and MAE of 0.000321 in the training phase, and an R^2 of 0.9451, MSE of 0.003077, and MAE of 0.001225 in the testing phase. The Levenberg-Marquardt algorithm (trainlm, M1) followed closely, while the Gradient Descent algorithms (traingd and traingdm) gave the weakest results with low or negative R^2 values. The findings demonstrate that the trainbr function provides superior predictive reliability for biomass HHV.

Keywords: Artificial neural networks, Biomass, Higher heating value, Transfer functions, Ultimate analysis

INTRODUCTION

Fossil fuels (coal, oil, and natural gas) have long dominated global energy consumption. However, their finite nature and harmful environmental impact make them unsustainable. As a result, there has been a progressive shift toward renewable energy sources, with biomass evolving as a significant substitute due to its abundance and environmental benefits (Msheliza & Dodo, 2025). Biomass encompasses products from agriculture and forestry, along with waste materials from the wood processing sector, and it is commonly used for generating electricity, providing heat, and producing fuels for transportation (Kujawska et al., 2023).

An essential element in biomass energy applications is the heating value, which defines biomass's energy content and practicality in combustion systems. Heating values of fuels are typically expressed in two variations: lower (net) heating value and higher (gross) heating value (Agha et al., 2025; Aghel et al., 2023). HHV, also known as the gross calorific value, indicates the total amount of energy released during complete combustion, including the latent heat of water vapor produced. An accurate estimation of HHV is crucial for evaluating the efficiency of biomass fuels and improving their utilization in energy production.

HHV of biomass is experimentally measured using a bomb calorimeter, specifically an oxygen bomb calorimeter (Agha et al., 2025; Dodo et al., 2022). This method is labor-intensive, time-consuming, and requires specialized expertise for sample preparation. As a result, there is growing interest among researchers in fast and cost-effective methods for estimating the HHV of biomass in waste-to-energy systems, particularly for those with limited or no access to bomb calorimeters. This approach leverages machine learning (ML) techniques such as ANN, support vector regression (SVR), random forest (RF), Gaussian process regression (GPR), and adaptive neuro-fuzzy inference system (ANFIS) to evaluate the multi-dimensional relationships among biomass properties, including proximate analysis, ultimate analysis, and physical composition. (Adeleke et al., 2024). Among various ML techniques, ANNs have been widely recognized for their ability to capture non-linear dependencies and complex interactions in data (Abba et al., 2020; Dodo et al., 2023).

There are several techniques to implement ML, specifically, ANN, for predicting the HHV of biomass using ultimate analysis. Thus, they vary based on network architecture, training functions, and input-output data processing methods. For example, Kujawska et al. (2023) compared two ML models, linear Regression (LR) and Multivariate Adaptive Regression Splines (MARS) method, to improve input selection of ANN-based biomass Prediction. The best ANN models had three input neurons and nine hidden layer neurons, which achieved a high accuracy of R = 0.988, RMSE = 0.3. Singh et al. (2018) developed an ANN model that effectively predicted the calorific value of municipal solid waste (MSW) using carbon, hydrogen, oxygen, nitrogen, sulfur, phosphorus, potassium, and ash content as input parameters, although external factors like seasonal variations, waste segregation, and moisture content were not considered, potentially affecting the calorific value of the prediction. Liou et al. (2024) developed and analyzed an artificial intelligence (AI) model, which successfully predicted nitrogen oxide (NO_x) emissions in the Taichung thermal power plant. The best results were obtained using eight specific input features. Tahir et al. (2023) used ML models to predict the calorific value of urban waste-derived fuels; the models surpassed the performance of empirical models, with (R^2) varied between 0.78 and 0.80, signifying enhanced predictive accuracy compared to prior models. (Jayapal et al., 2025) developed an ANN model (9-6-6-1) which successfully predicted biomass using proximate and ultimate analysis data and compared its performance to empirical data. The model achieved an R^2 of 0.81, an MAE of 0.77 MJ/kg, and outperformed 54 correlations. Brandić et al. (2023) compared ML methods (ANN, SVM, RF, Polynomial Regression) for HHV prediction using proximate analysis. ANN had the best result amongst the 4 with a strong R^2 of 0.92 and RMSE of 1.33. Adeleke et al. (2024) evaluated the performance of Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost) for the prediction of HHV using proximate and ultimate analysis. XGBoost outperformed the other models with an R^2 of 0.968 in the training phase and an R^2 of 0.731, and an RMSE of 0.355 in the testing phase. Balsora et al. (2022) ANN-3 and ANN-4 achieved R^2 of approximately 0.99 and MAE <0.

071.bimochemcial inputs contributed about 38% to accuracy. Güleç et al. (2022) systematically studied how the ANN structure, consisting of activation functions, algorithms, dataset composition, and hidden layers. The combined dataset had the best predictions with an R^2 of 0.962 in training and 0.876 in testing.

Despite a growing emphasis on ML for biomass HHV prediction, persistent limitations remain in the literature, particularly concerning the systematic comparison of neural network training functions and comprehensive model evaluation. ANN models face issues such as data dependency, overfitting, and sensitivity to input variables, compromising their robustness and accuracy. The lack of standardized architectures and limited comparisons of training functions within the feedforward backpropagation neural network (FBNN) further complicates performance evaluation. Although Abdollahi et al. (2024), Brandić et al. (2023), and Msheliza & Dodo (2025) have demonstrated that ANN can accurately predict HHV, there has been limited attention to compare multiple training functions within the ANN model. Thus, there is a lack of clear consensus on the most effective training function under differing experimental conditions.

This study aimed to fill this gap by leveraging a comprehensive biomass ultimate analysis data for reliable and generalizable HHV prediction. To achieve this aim, a suite of ten training functions in feedforward neural networks was employed for the HHV prediction models development, and their performances were compared using statistical indices, namely, determination coefficient (R^2), mean squared error (MSE), and mean absolute error (MAE). The study is envisaged to contribute to the understanding of how different training functions influence the accuracy of feedforward backpropagation neural networks in biomass HHV prediction. This also provides a cost-effective alternative to experimental procedures, which are costly and labor-intensive, thereby supporting sustainable energy generation and reducing reliance on fossil fuels.

MATERIALS AND METHODS

This study compared the performance of 10 different training functions in FBNN for biomass HHV prediction using ultimate analysis variables (carbon (C), hydrogen (H), nitrogen (N), sulphur (S), and oxygen (O)) The training functions considered include, Levenberg–Marquardt (trainlm), Scaled Conjugate Gradient (trainscg), Bayesian Regularization (trainbr), Gradient Descent (traingd), BFGS Quasi-Newton (trainbfg), Conjugate Gradient with Powell–Beale Restarts (traincgb), Resilient Backpropagation (trainrp), One-Step Secant (trainoss), Gradient Descent with Momentum (traingdm), and Random Order Weight/Bias Rule (trainr). The training functions are algorithms that adjust the network's weights and biases to minimize errors. They figure out the network's training time, final performance, and memory usage. The network utilised in this study followed a 5-10-1 (input-hidden-output layer) topology, shown in Figure 1. Furthermore, the training functions are designated as models M1-M9 as shown in Table 1.

To improve the ANN model's training efficiency and pattern recognition, input and output variables were normalized with min-max scaling to a range of 0 to 1 using equation (1).

$$X_{norm} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

X_{norm} is the normalized experimental data, X_{min} is the minimum value, while X_{max} is the maximum value of the experimental datasets.

Table 1: Model Designation

Model designation	Training Functions
M1	trainlm
M2	trainscg
M3	trainbr
M4	traingd
M5	trainbfg
M6	traincgb
M7	trainrp
M8	trainoss
M9	traingdm
M10	Trainr

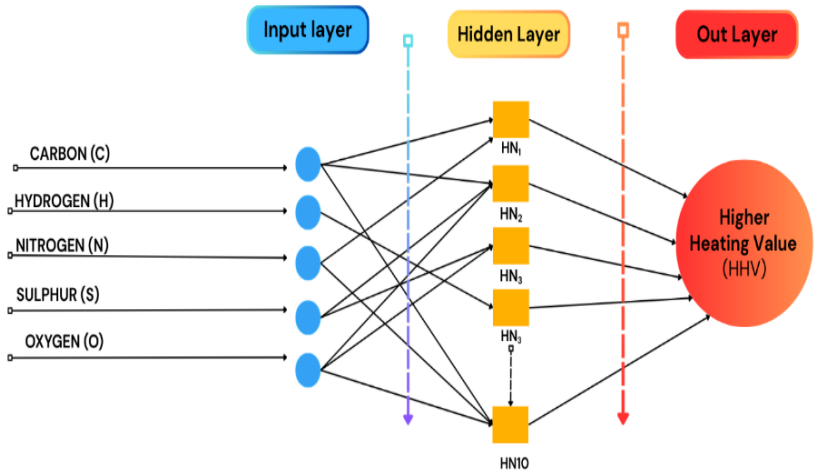


Figure 1: A 5-10-1 FBNN Topology

Training Functions of FBNN

Additive momentum (Traingdm)

By responding to the error surface and local gradient, a quicker convergence occurs with gradient descent that introduces momentum. Together with the thresholds of change, an additive value proportionate to the prior weights is added, and based on the backpropagation method, a new set of weights and a threshold are formed (Dodo et al., 2024). It helps the algorithm accelerate in the right direction and dampens oscillations. It is mathematically expressed in equation (2).

$$v_t = \gamma v_t - 1 + \alpha \nabla L(w_t) \quad (2)$$

Where V_t is the velocity at time t and γ is the momentum coefficient.

Gradient Desent (Traingd)

It updates the weights in the opposite direction of the gradient of the loss function with respect to the weights. It is an adaptive learning rate that aims to sustain a high learning step size while maintaining stability (Nguyen et al., 2021). It is mathematically expressed in equation (3).

$$W_{new} = W_{old} - \alpha \nabla L(W) \quad (3)$$

Where W represents the weights, a is the learning rate, and $\nabla L(W)$ is the gradient of the loss function.

Conjugate Gradient (Traincgp)

Iterations are used in most conjugate gradient backpropagations to modify the step size (Dodo et al., 2024). They speed up training by using 'conjugate" search directions. These directions are chosen to minimize interference with earlier steps, leading to a faster convergence. These can be expressed mathematically in equations (4) and (5).

$$d_k = -\nabla L(W_k) + \beta_k d_k - 1 \quad (4)$$

$$\beta_k = \frac{\|\nabla L(w_k)\|2}{\|\nabla L(w_k - 1)\|} \quad (5)$$

Levenberg-Marquart (Trainlm)

The Levenberg-Marquardt (LM) algorithm, often known as the dampened, is an approach of the least squares that is suited for problems involving nonlinear least squares (Dodo et al., 2024). It combines the stability of gradient descent and the fast convergence of the Gauss-Newton method. The equation describing this algorithm is expressed as:

$$\nabla W = -\left[J^T J + \mu I\right] - 1 J^T e \quad (6)$$

Where J^T is the Jacobian matrix of the network errors with respect to the weights, e is the vector of errors, and μ is a scalar that controls the transition between gradient descent and Gauss-Newton.

Quasi-Newton (Trainbfg)

The advantage of the quasi-Newton method is that it is computational and inexpensive because it does not need many operations to evaluate the Hessian matrix and calculate the corresponding inverse. Each iteration is the approximation of the inverse Hessian Matrix. It is computed using only information on the first derivatives of the loss function, expressed in equation (7) (Nguyen et al., 2021):

$$W_{k+1} = W_k - H_k^{-1} \nabla L(W_k) \quad (7)$$

Where H_k approximates the Hessian matrix.

Conjugate Gradient Backpropagation (Traincgb, Traincgp, Traincgf, and Trainscg)

These are search and conjugate gradient algorithms to achieve much faster convergence. Iterations are used in most conjugate gradient backpropagations to modify the step size (Dodo et al., 2024). Trainscg uses a "scaling" factor to decide the step size at each iteration. This makes it a very efficient algorithm for large-scale problems.

Resilient Backpropagation (Trainrp and Trainscg)

This training function focuses on the sign of the gradient rather than its size. It uses a separate learning rate for each weight, and this rate is based on whether the gradient changes sign. As shown in equation (8), this method attains the best convergence without parameter selection (Nguyen et al., 2021).

$$\Delta W_{ij}(t) = -\eta_{ij}(t) \times sign\left(\frac{\partial E}{\partial W_{ij}}\right) \quad (8)$$

Where η_{ij} *is* the learning rate for a specific weight.

One-step Secant (Trainoss)

It uses an approximation of the Hessian matrix based on the secant method. It is calculated just with the loss function's first derivatives' information. The loss function's second partial derivatives make up the Hessian matrix (Nguyen et al., 2021).

Bayesian Regularization(Trainbr)

It generalizes an existing network by reducing the combination of squared errors and weights (Dodo et al., 2024). It trains the network using the Levenberg-Marquardt algorithm, but with changes in the performance function that include a penalty for large weights. It is described by equation (9).

$$F = \beta E_D + \alpha E_W \quad (9)$$

Where E_D *is* the sum-squared error, E_W is the sum of squared weights, and β and α are parameters that are optimized during training.

Transfer Functions

A logsig (log-sigmoid) transfer function was used in the hidden and output layers, respectively layer to evaluate activation behavior. It was selected because it introduces non-linearity, which ensures the outputs remain bounded between 0 and 1 (Msheliza & Dodo, 2025). Mathematically, it is expressed in equation (10).

$$F(x) = \log sig(x) = \left(1 - e^{-x}\right)^{-1} \quad (10)$$

Performance Evaluation

To access the predictive performance of the neural network models, the dataset was divided into two subsets: 70% for training and 30% for testing. This is done to prevent overfitting and ensure that the model's performance reflects its true predictive capability rather than memorization of the input data. The predictive accuracy of the models was assessed using three statistical performance indices represented in equations (11) - (13) the coefficient of determination (R^2); which highlights the overall fit of the model, mean absolute error (MAE); indicates the average prediction error, and mean squared error (MSE); captures the sensitivity of the model to larger deviations. By combining these three evaluation metrics, a balanced assessment of neural network performance was achieved.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(HHV_{e,i} - HHV_{p,i}\right)^2 \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(HHV_{p,i} - HHV_{e,i}\right)^2}{\sum_{i=1}^{n}\left(HHV_{e,i} - \overline{HHV}_e\right)^2} \quad (12)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|HHV_{e,i} - HHV_{p,i}\right| \quad (13)$$

$HHV_{p,i}$, $HHV_{e,i}$, $\overline{HHV}_e$, and n denote predicted HHV, experimental HHV, the mean of experimental HHV, and the number of data instances, respectively.

RESULTS AND DISCUSSION

The results of the performance analysis of the models (M-M9) using the statistical indices, R^2, MSE, and MAE, respectively, are presented in Table 2. For an in-depth analysis and comparison of prediction model performances, at least two evaluation metrics are recommended by researchers. In this way, the constraints of a particular metric that can result in ineffective judgment can be mitigated (Dodo et al., 2024). As such, the evaluation metrics used were R^2 (coefficient of determination), MSE (mean squared error), and MAE (mean absolute error) to compare the predicted and experimental HHVs. Lower MSE and MAE, and higher R^2, respectively, indicate a strong relationship between the predicted and experimental HHVs, while higher MSE and MAE, and lower R^2 suggest a significant dispersion between the predicted and experimental HHVs.

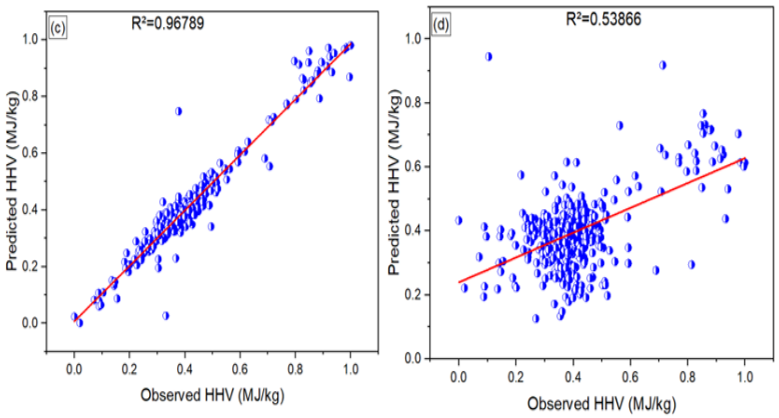
Table 2: R^2, MSE, and MAE Values for Models M1 to M10

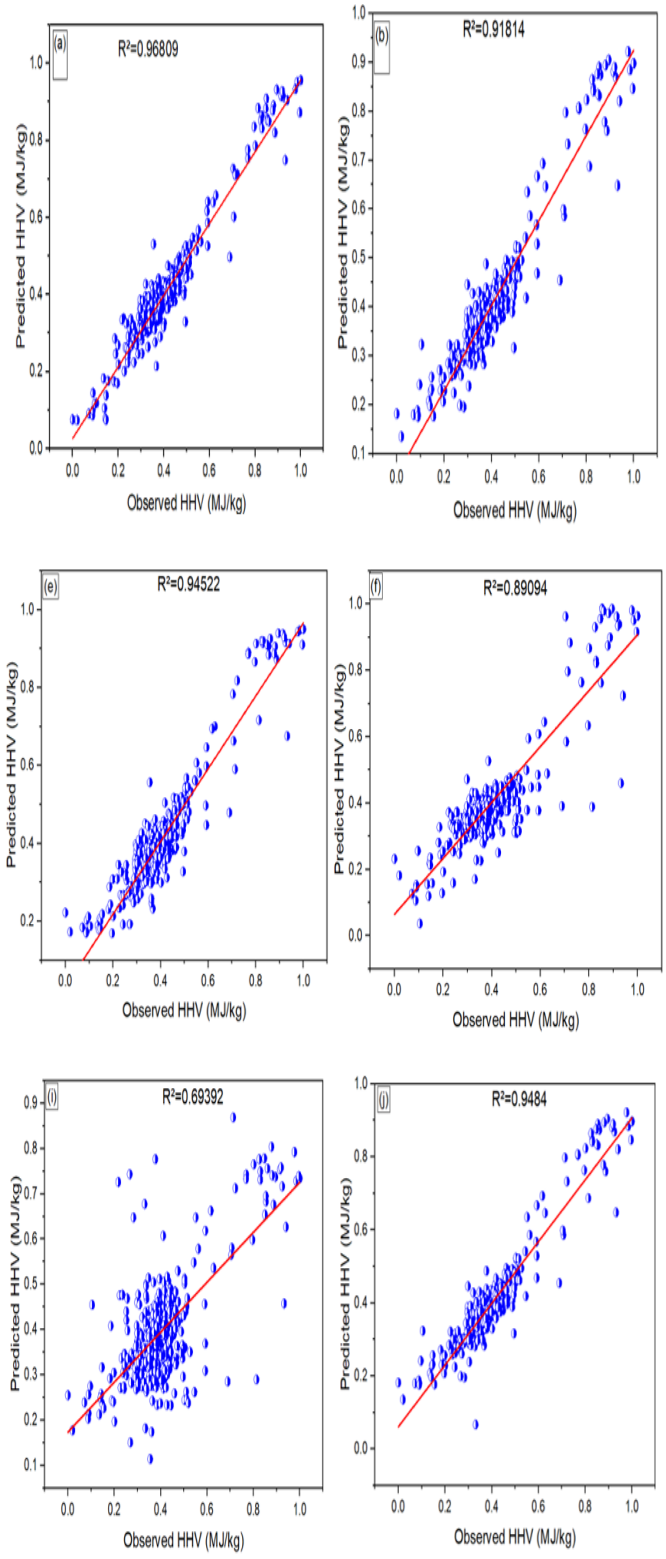
Model	Training Phase			Testing Phase		
	R^2	MSE	MAE	R^2	MSE	MAE
M1	0.871242	0.001123	0.000447	0.934967	0.003646	0.002584
M2	0.770048	0.002006	0.001633	0.860208	0.007836	-0.01365
M3	0.835832	0.001432	0.000321	0.945112	0.003077	0.001225
M4	-0.35935	0.01186	0.003424	0.578989	0.023601	0.018519
M5	0.800243	0.001743	-0.00456	0.901124	0.005543	-0.00751
M6	0.62758	0.003249	-0.00298	0.851586	0.00832	0.001359
M7	0.60975	0.003405	0.001255	0.865735	0.007527	0.010495
M8	0.701673	0.002603	-0.00409	0.902218	0.005481	0.004614
M9	0.055295	0.008242	0.004562	0.680699	0.017899	0.013638
M10	0.804767	0.001703	0.001871	0.911139	0.004981	-0.00085

As shown in Table 2, the trainbr (M3) and trainlm (M1) show the best prediction performances, evidenced by high R^2 and low MSE and MAE values. The lowest performing training function is M4 (traingd), M9 (traingdm), with M4 exhibiting a negative R^2 of -0.35935 during the testing phase and M9 having a very low R^2 of 0.05529, which was the second lowest of all models. Model M3 achieved the best results with R^2 of 0.96789, MSE of 0.001432, and 0.000321. Model M1 (trainlm) performed exceptionally well in the testing phase with an R^2 of 0.934967, MSE of 0.00365, and MAE of 0.00258. The training function is well-known for its efficiency and rapid convergence. M10 (trainrp) followed closely, with R^2 values of 0.80477 (training) and 0.91114 (testing). The model demonstrated stability across training and testing phases, with low mean bias error (MAE) and relatively small error variance, making it a reliable training function for biomass HHV prediction.

Models M5 (trainbfg) and M8 (trainoss) produced satisfactory results in the testing phase with R^2 values of 0.90112 and 0.90222, respectively. Although not as strong as M3, M1, and M10, they maintained consistent performance with moderate error values and good generalization ability. Models M2 (trainscg), M6 (traincgb), and M7 (trainrp) produced acceptable but less competitive results, with testing R^2 values ranging between 0.85159 and 0.86574. These models had higher error magnitudes compared to the top-performing ones, but still demonstrated stable predictions without severe bias.

The Scatter plots of predicted HHV values against actual values depicted in Figure 2 further illustrate the differences in model performance. The comparison between the experimental and predicted HHV was visualized using scattered plots, illustrating the model's prediction performance and how closely the clustered points were along the line of perfect fit (y=x).




(a)
R²=0.96809
(b)
R²=0.91814
(e)
R²=0.94522
(f)
R²=0.89094
(i)
R²=0.69392
(j)
R²=0.9484
Predicted HHV (MJ/kg)
Observed HHV (MJ/kg)

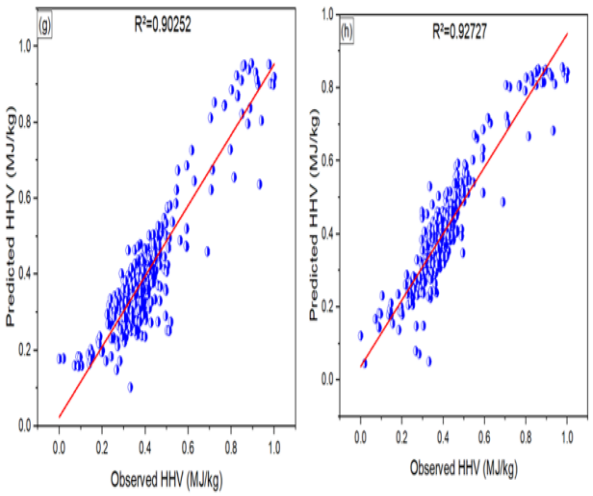


Figure 2: Scatter Plots for Various Models (a) M1; (b) M2; (c) M3; (d) M4; (e) M5; (f) M6; (g) M7; (h) M8; (i) M9; (j) M10

Figure 2 shows M3 with an R^2 of 0.96789 and M1 with an R^2 of 0.96809, implying strong alignment and low deviations from the line of best fit. In contrast, M4, having an R^2 of 0.53866, and M9, having an R^2 of 0.69392, displayed wide scatter from the line of equality, implying poor agreement between predicted and experimental HHVs. Neither of these models used descent-based training functions, providing the visual evidence that supports the statistical results highlighting the unsuitability of traingd and traingdm for biomass HHV prediction.

Another method used to visually identify the optimal model and algorithm for HHV prediction was Rader plots, where MSE and MAE, respectively, in both training and testing phases are shown in Figures 3 and 4. The MSE and MAE, respectively, measure the average squared and absolute average difference between the split predicted and actual values. Lower values indicate higher accuracy and are represented by smaller areas on the plot. During the training, the values for the models ranged from approximately 0-0.014, while the testing ranged from 0-0.03. The points on the plot remained low, confirming that models generalized well without significant degradation in accuracy. The small gap between the training and testing MAE values shows that models did not overfit and retained stability to unseen data. Thus, in the Figure, both the training and testing phases indicate that M3 (trainbr) forms the smallest shape, located nearest to the center, implying the lowest MAE during the training phase (0.0003) and 0.0012 at the testing phase. Models M1, M2, and M10 also have very small areas, while the plots for M4 and M9 are significantly larger, indicating high error rates.

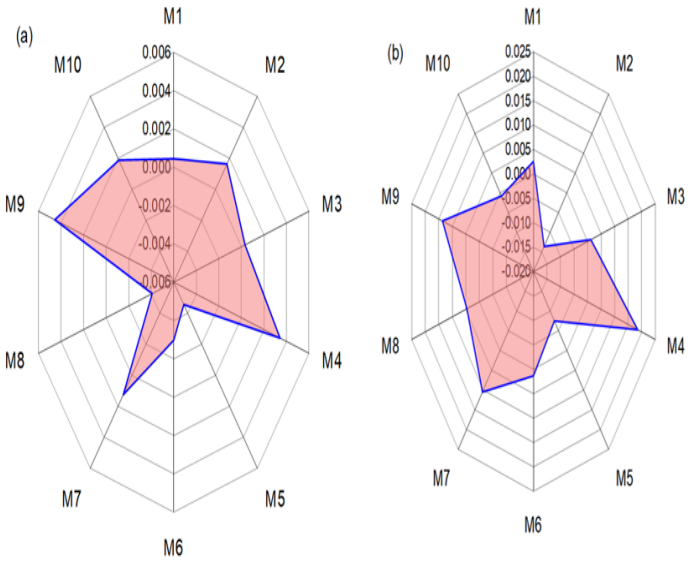


Figure 3: Radar Plots Using MSE (a) Training Phase (b) Testing Phase

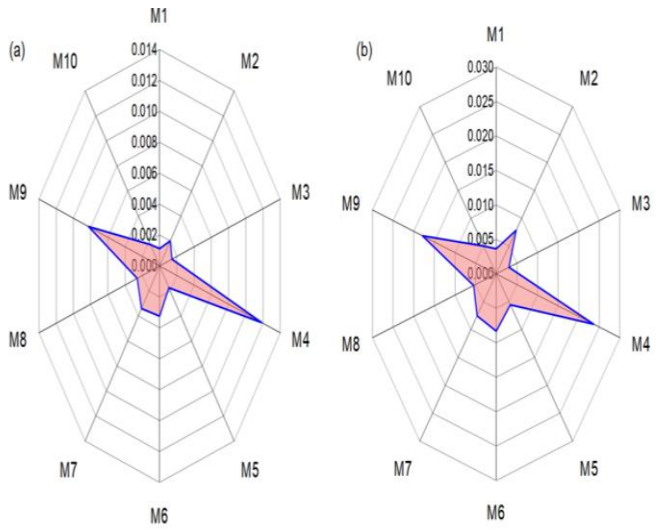


Figure 4: Radar Plots Using MAE (a) Training Phase (b) Testing Phase

Figure 4 shows the radar plot of MAE, which measures the average magnitude of the errors in the set of predictions. A lower MAE value indicated a more accurate model. The plot having a higher MAE is represented by a larger area extending farther from the center. Conversely, a smaller area closer to the center corresponds to a low MAE. Again, the M4 and M9, which use gradient descent with momentum, stand out with a more marginal shape deviating from the center, which confirms their poor performance and inability to make accurate predictions. Overall, the ranking of model performance is M3, M1, M10, M8, M5, M7, M2, M6, M9, and M4, in descending order of accuracy.

CONCLUSION

This study focused on the performance comparison of ten different training functions in a feedforward backpropagation neural network with a topology of 5-10-1 to predict the higher heating value of biomass based on ultimate analysis. This was achieved by varying the training functions while maintaining a consistent FFBN architecture consisting of five inputs, ten hidden layers, and the logsig activation function throughout. The results highlighted Bayesian regularization (trainbr) and Levenberg-Marquardt (trainlm) as the best training algorithms with the highest R^2 and low MSE and MAE, respectively. Thus, the optimal performance was attributed to M3 (trainbr), which exhibited the highest R^2 value of 0.96789, along with the lowest MSE of 0.00143 and MAE of 0.000032 during the training phase. The next best-performing model was M1 (trainlm), with an R^2 of 0.96809 and low MAE and MSE values of 0.00112 and 0.00045, respectively, in the training phase. On the other hand, the lowest performing model, M4, resulted in a negative R^2 in the training phase of -0.35935 and exhibited high MAE and MSE values of 0.01186 and 0.00032. Future research may consider a larger and more diverse biomass dataset, comprising proximate and ultimate analysis data, to improve prediction.

REFERENCES

Abba, S. I., Linh, N. T. T., Abdullahi, J., Ali, S. I. A., Pham, Q. B., Abdulkadir, R. A., Costache, R., Nam, V. T., & Anh, D. T. (2020). Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration. *IEEE Access*, *8*(September), 157218–157237. https://doi.org/10.1109/ACCESS.2020.3017743

Abdollahi, S. A., Basem, A., Alizadeh, A., Jasim, D. J., Ahmed, M., Sultan, A. J., Ranjbar, S. F., & Maleki, H. (2024). Combining artificial intelligence and computational fluid dynamics for optimal design of laterally perforated finned heat sinks. *Results in Engineering*, *21*(March), 102002. https://doi.org/10.1016/j.rineng.2024.102002

Adeleke, A. A., Adedigba, A., Adeshina, S. A., Ikubanni, P. P., Lawal, M. S., Olosho, A. I., Yakubu, H. S., Ogedengbe, T. S., Nzerem, P., & Okolie, J. A. (2024). Comparative studies of machine learning models for predicting higher heating values of biomass. *Digital Chemical Engineering*, *12*(100159), 10. https://doi.org/10.1016/j.dche.2024.100159

Agha, D. C., Dodo, U. A., & Hussein, M. A. (2025). Analysis of the Electricity Generation Potentials of Plastic Wastes on a University Campus. *Journal of Science Technology and Education*, *13*(1), 7–15.

Aghel, B., Yahya, S. I., Rezaei, A., & Alobaid, F. (2023). A Dynamic Recurrent Neural Network for Predicting Higher Heating Value of Biomass. *International Journal of Molecular Sciences*, *24*(6), 1–13. https://doi.org/10.3390/ijms24065780

Balsora, H. K., Kartik, A., Dua, V., Joshi, J. B., Kataria, G., Sharma, A., & Chakinala, A. G. (2022). Machine learning approach for the prediction of biomass pyrolysis kinetics from preliminary analysis. *Journal of Environmental Chemical Engineering*, *10*(3), 108025. https://doi.org/10.1016/j.jece.2022.108025

Brandić, I., Antonović, A., Pezo, L., Matin, B., Krička, T., Jurišić, V., Špelić, K., Kontek, M., Kukuruzović, J., Grubor, M., & Matin, A. (2023). Energy Potentials of Agricultural Biomass and the Possibility of Modelling Using RFR and SVM Models. *Energies*. https://doi.org/10.3390/en16020690

Brandić, I., Pezo, L., Bilandžija, N., Peter, A., Šurić, J., & Voća, N. (2023). Comparison of Different Machine Learning Models for Modelling the Higher Heating Value of Biomass. *Mathematics*, *11*(9), 2098. https://doi.org/10.3390/MATH11092098/S1

Dodo, U. A, Dodo, M. A., Shehu, A. F., & Badamasi, Y. A. (2023). Performance Analysis of Intelligent Computational

Algorithms for Biomass Higher Heating Value Prediction. *Nigerian Journal of Technological Development*, *20*(4), 44–52.

Dodo, U. A., Ashigwuike, E. C., & Emechebe, J. N. (2022). Optimization of Standalone Hybrid Power System Incorporating Waste-to-electricity Plant: A Case Study in Nigeria. *2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON)*, 1–5. https://doi.org/10.1109/NIGERCON54645.2022.9803081

Dodo, U. A., Dodo, M. A., Belgore, A. T., Husein, M. A., Ashigwuike, E. C., Mohammed, A. S., & Abba, S. I. (2024). Comparative study of different training algorithms in backpropagation neural networks for generalized biomass higher heating value prediction. *Green Energy and Resources*, *2*(1), 100060. https://doi.org/10.1016/j.gerr.2024.100060

Güleç, F., Pekaslan, D., Williams, O., & Lester, E. (2022). Predictability of higher heating value of biomass feedstocks via proximate and ultimate analyses – A comprehensive study of artificial neural network applications. *Fuel*, *320*(123944), 1–16. https://doi.org/10.1016/j.fuel.2022.123944

Jayapal, A., Ordonez Morales, F., Ishtiaq, M., Kim, S. Y., & Reddy, N. G. S. (2025). Modeling the Higher Heating Value of Spanish Biomass via Neural Networks and Analytical Equations. *Energies*, *18*(15), 4067. https://doi.org/10.3390/EN18154067/S1

Kujawska, J., Kulisz, M., Oleszczuk, P., & Cel, W. (2023). Improved Prediction of the Higher Heating Value of Biomass Using an Artificial Neural Network Model Based on the Selection of Input Parameters. *Energies*, *16*(10), 1–16. https://doi.org/10.3390/en16104162

Liou, J. L., Liao, K. C., Wen, H. T., & Wu, H. Y. (2024). A study on nitrogen oxide emission prediction in the Taichung thermal power plant using an artificial intelligence (AI) model. *International Journal of Hydrogen Energy*. https://doi.org/10.1016/j.ijhydene.2024.03.120

Msheliza, S. A., & Dodo, U. A. (2025). Performance Comparison of Different Activation Functions in Neural Networks for Biomass Energy Content Prediction. *FUDMA Journal of Sciences*, *9*(4), 285 – 294. https://doi.org/https://doi.org/10.33003/fjs-2025-0904-3493

Nguyen, T. A., Ly, H. B., Mai, H. V. T., & Tran, V. Q. (2021). On the Training Algorithms for Artificial Neural Networks in Predicting the Shear Strength of Deep Beams. *Complexity*, *2021*. https://doi.org/10.1155/2021/5548988

Singh, D., Satija, A., & Hussain, A. (2018). Predicting the calorific value of municipal solid waste of Ghaziabad City, Uttar Pradesh, India, using an artificial neural network approach. In *Advances in Intelligent Systems and Computing* (Vol. 584, pp. 495–503). https://doi.org/10.1007/978-981-10-5699-4_46

Tahir, J., Ahmad, R., & Tian, Z. (2023). Calorific value prediction models of processed refuse-derived fuel 3 using ultimate analysis. *Biofuels*. https://doi.org/10.1080/17597269.2022.2116771