Application of Multi-Output Extra Trees Algorithm for Enhanced Biomass Prediction in Tropical Forest Ecosystems

Juan Chen
Zhumadian Preschool Education College, Zhumadian 463000, China
E-mail: czqjuan@126.com

Keywords: Biomass forecasting, extra trees (ET), aboveground biomass (AGB), belowground biomass (BGB), GridSearchCV

Received: July 2, 2025

For carbon accounting and ecosystem monitoring, one must be able to correctly forecast biomass in tropical forests. The ability of Extra Trees to concurrently forecast biomass is investigated in this research. Estimated aboveground biomass (AGB), belowground biomass (BGB), and total biomass (TB) depend on 175 trees gathered from 27 Central Vietnamese sites. Among the predicting variables were tree structural measurements, site features, and categorical ecological markers. One-hot encoding categorical data, normalizing numerical characteristics, and removing missing values variables were among the preprocessing procedures. The dataset was split into training and testing sets (80/20) using a fixed random seed, and GridSearchCV revealed the best model hyperparameters. Utilizing a set random seed (80/20), the data was separated into training and testing groups; then GridSearchCV was used to find the best model parameters. The performance of ET at baseline was examined in light of linear regression, Long Short-Term Memory (LSTM), Bagging, and Gradient Boosting (GB). With an R^2 value of 0.9975 and RMSE for ET of 0.9975, the score on the held-out test set was 4.89 which Bagging (R^2 = 0. 9948, RMSE = 7. 08) and linear regression were superior than LSTM (R^2 = 0. 9813, RMSE = 13. 41). 0. 9849, RMSE = 12. 06). These results support the greater accuracy and robustness of ET for biomass estimation, so emphasizing its importance for producing dependable Plans for forest carbon stock assessments and conservation.

Povzetek: Študija preuči Extra Trees za sočasno napoved nadzemne, podzemne in skupne biomase tropskih gozdov iz meritev dreves in lokacijskih značilnosti ter pokaže, da ta ansambelski pristop dosega bolj robustne napovedi kot primerjalni modeli za potrebe ogljičnega knjigovodstva..

1 Introduction

Biomass prediction [1] is essential for an understanding of carbon storage and the general condition of tropical woodlands. Field data combined with allometric models—which tend to focus on either aboveground or belowground biomass—had formerly been used to estimate biomass. Aboveground biomass, comprising trunks, branches, and leaves of trees, can be measured relatively more easily compared to the belowground biomass, which consists of roots and soil organic matter. However, since precise estimates of both components are important for complete carbon accounting and informed decisions regarding forest management and conservation, traditional methods fall short due to various limitations. This has led to a growing call for more integrated and reliable approaches [2], [3], [4], [5].

Recent advances in deep learning have improved the accuracy of biomass prediction in tropical forests. In our comparison, the deep model is an LSTM configured for multi-output regression, jointly predicting AGB, BGB, and TB. We include this sequential architecture to test whether deep representation capacity adds value beyond tree ensembles; however, given the cross-sectional nature of our features, LSTM serves as a stress-test rather than the natural inductive bias for these data. This approach enhances both the reliability and efficiency of biomass estimation. This approach offers an integrated view of forest biomass whereas the methodology enhances the dependability and accuracy of biomass estimation. The inclusion of the multi-output deep learning model into biomass prediction in this chapter represents a significant step forward in our ability to monitor and control Far more efficient tropical forests [6], [7], [8]. In recent years, the integration of advanced Machine Learning techniques, including deep learning methods, into ecological research has gained increasing attention. Indeed, a number of studies have so far explored the potential of deep learning models to achieve increases in the accuracy of biomass predictions from various forest ecosystems [9]. For example, the application of convolutional neural networks and recurrent neural networks has been explored to deal with spatial and temporal data, respectively, with the aim of enhancing biomass estimate accuracy from remote sensing data. Another equally attractive option includes multivariate models, which predict several interdependent variables simultaneously and have gained popularity in ecological studies by modeling more holistic interactions

within a complex forest ecosystem. For instance, Li et al. [10]illustrated that the integration of Landsat 8 and Sentinel-1A data using the XGBoost algorithm had better performance in estimating aboveground biomass for subtropical forests. This method outperformed traditional approaches and has the potential to reduce estimation errors. Sharma et al. [11] investigated UAV-based multispectral data and Machine Learning for the estimation of oat biomass. Their study showed promising correlations between derived vegetation indices and biomass at some locations; however, model accuracy varied greatly across sites. The study also underlined that airborne remote sensing is able to efficiently support biomass predictions, although model consistency and the integration of more biophysical parameters have to be considered in further studies. Zhang et al. [12] evaluated Machine Learning regression algorithms for forest aboveground biomass estimation from satellite data, finding that ensemble methods, particularly CatBoost, outperformed nonensemble methods across various sampling scenarios, emphasizing the need for improved estimation accuracy at extreme biomass levels in future research. Ma et al. [13] utilized airborne LiDAR data to assess forest aboveground biomass (AGB), revealing that certain LiDAR-derived features consistently ranked as sensitive indicators for estimating total and component AGB, with random forest showing superior performance compared to support vector regression.

Furthermore, Li at al. [14] innovatively integrated random forest and least squares models to enhance forest biomass prediction accuracy, emphasizing the importance of canopy height, soil organic matter, and the red-edge chlorophyll vegetation index as critical variables for precise estimation of carbon storage in subtropical forests. In their research [3], Huy et al. looked at how well multi-output deep learning (MODL) models predicted biomass components in forests and discovered that Approaches covering a range of tropical forest types and environmental gradients produced considerably less mistakes than traditional methods. Machine learning algorithms, especially in support vector regression, performed better than conventional biomass predictions in Lebanon's cedar tree stands, according to a recent study [15] is quite important. elements of efficient forest management and carbon-bracketing techniques. The accuracy of RF models in estimating soil CO2 emissions from different restoration habitats was as Canteral et al. [16] have shown. This shows their ongoing capacity to reduce the doubts about the dynamics of the carbon cycle [17]. Sesnie et al. showed how well ensemble Machine Learning models worked by using data from many satellite sensors that foretold tree biomass. Diversity indexes of seasonally dry tropical woods that provide pertinent data for more comprehensive biodiversity mapping projects. As discussed by Li et al. [18], Machine Learning algorithms have shown significant improvements in biomass prediction accuracy in forest ecosystems.

While this study explores the potential of Multi-Output Deep Learning (MODL) models, such as Long Short-Term Memory (LSTM), for biomass prediction, we focus on the Extra Trees (ET) algorithm for its superior performance in small, noisy datasets typical of ecological research. While LSTM has been shown to be effective for sequential data, we demonstrate that in this non-sequential, cross-sectional scenario, ET outperforms MODL due to its robustness in capturing nonlinear interactions and preventing overfitting.

Building on these considerations, Table 1 provides a comparative overview of prior biomass prediction studies, highlighting the diversity of applied methodologies, ranging from traditional allometric approaches to advanced machine learning frameworks. The table emphasizes the variation in datasets, predictor variables, and evaluation metrics, illustrating the fragmented landscape of biomass modeling. This contextualization underscores the rationale for testing both MODL architectures and ensemble learners in the present work.

Table 1: Summary of prior biomass prediction studies.

Study	Models	Dataset	Metrics	Key findings
Li et al. (2020) [14]	XGBoost, RF	Landsat 8 + Sentinel-1A; subtropical forests	R^2, RMSE	XGBoost outperformed traditional methods, improving AGB estimation accuracy.
Sharma et al. (2022) [15]	ML regressors with UAV multispectral data	Oat biomass, multisite	Correlation, RMSE	Strong correlations found, but performance varied across sites.
Zhang et al. (2020) [16]	8 ML regression algorithms (CatBoost, SVR, RF, etc.)	Satellite-based Forest AGB	R^2, RMSE	Ensemble models (CatBoost) outperformed non-ensemble methods.
Ma et al. (2023) [17]	Random Forest, SVR	Airborne LiDAR, subtropical forests	R^2, feature sensitivity	Random Forest superior; LiDAR-derived features critical predictors.
Huy et al. (2024) [5]	Multi-output Deep Learning (MODL)	175 destructively sampled trees, Vietnam	RMSE, MAE	MODL improved multi-output biomass

				prediction vs traditional WNSUR.
Present study	Extra Trees (ET), GB, Bagging, LSTM	175 trees, 27 plots, Vietnam	R^2, RMSE, MAE, VAF	ET outperforms baselines; robust under small, noisy ecological datasets.

This study will test the ET algorithm's efficiency in enhancing the accuracy and reliability of simultaneous biomass estimates of aboveground, belowground, and total biomass in tropical forests.

Gaps and challenges. Prior biomass-estimation studies have been characterized by (i) limited sample sizes resulting from destructive sampling, (ii) collinearity among structural and site variables, and (iii) heterogeneous protocols that complicate generalization. In addition, many works have compared single-output models or emphasized cost or emissions proxies without jointly estimating AGB, BGB, and TB or reporting cross-validated variability. This study's contribution. Biomass prediction is cast as a multi-output regression problem (AGB, BGB, TB), with ET, GB, Bagging, and LSTM benchmarked under a uniform pipeline with leakage-safe cross-validation. A baseline linear regression model is provided for clarity, and model interpretability is added via feature importance and SHAP analysis to ground ecological plausibility. The results show that ET achieves the strongest test performance for this dataset—an empirical conclusion rather than an a priori assumption. These models were selected to represent complementary methodological families. GB provides strong boosting performance on nonlinear ecological data; ET offers variance reduction and robustness on small, noisy datasets; Bagging improves stability via bootstrap aggregation; and LSTM serves as a deep-learning comparator with sequence-modeling capacity. Together, they cover both ensemble and neural approaches, enabling a balanced evaluation of methodological suitability for tropical biomass prediction.

By incorporating a new optimization approach to couple with the ET algorithm, this study aims to overcome the deficiencies of traditional methods and enhance predictive performance. It does so by introducing, for the first time, the use of ET to predict biomass, comparing it with established Machine Learning techniques such as Gradient Boosting, Bagging, and LSTM networks. Until now, ET has been recently applied in ecological modeling; this study will offer fresh viewpoints on its capacity for more exact and reliable biomass estimation supports environmental monitoring and helps forest management. We benchmark ET against Gradient Boosting (GB), Bagging, and a Long Short-Term Memory (LSTM) network. GB provides a strong boosted-tree baseline that incrementally corrects residuals; Bagging isolates the effect of variance reduction via bootstrap aggregation; and LSTM offers a deep-learning comparator with sequence-model capacity. This panel spans complementary inductive biases for nonlinear tabular data, enabling a fair assessment of whether multi-output ET is preferable for tropical biomass estimation.

1.1 Research questions and hypotheses

This study addresses the following research questions:

- RQ1: Can multi-output ensemble methods (Extra Trees, Gradient Boosting, Bagging) and deep learning models (LSTM) provide accurate simultaneous predictions of AGB, BGB, and TB in small, destructively sampled tropical forest datasets?
- RQ2: How does the ET algorithm perform relative to other machine learning methods in terms of predictive accuracy, variance reduction, and robustness to noisy ecological features?
- RQ3: How do machine learning approaches compare to a traditional linear regression baseline for biomass prediction?

Based on prior literature, we hypothesize that (H1) ET will outperform other methods due to its strong variance-reduction properties; (H2) ensemble methods in general will provide more stable results than LSTM in small cross-sectional datasets; and (H3) all machine learning approaches will outperform linear regression due to their ability to capture nonlinear interactions.

2 Methodology and dataset

A leakage-safe pipeline is implemented, consisting of the following stages: data curation → preprocessing (missing-data handling, encoding, scaling) → train/validation selection with K-fold cross-validation applied within the training set only → model selection and tuning (ET, GB, Bagging, LSTM) → final training on the full training set using the best hyperparameters → held-out test evaluation. All three targets (AGB, BGB, TB) are predicted jointly in a multi-output framework, and robustness is quantified through cross-validation statistics and confidence intervals on the test-set metrics

The contribution is framed around ET (ET)—an ensemble of fully randomized decision trees (a variant of Random Forest) in which randomized split thresholds and feature subsampling are employed to reduce variance and capture nonlinear feature interactions in tabular ecological data. ET is deployed in a multi-output regression setting for AGB, BGB (belowground biomass), and TB (total biomass) in order to exploit cross-target correlations. On the small-to-medium tropical forest dataset, ET is shown to yield the most reliable test-set performance among the compared learners by effectively handling noisy measurements, collinearity, and heterogeneous predictor scales. ET is particularly suitable here because (i) destructive-sampling forest datasets are typically small-n with complex, nonlinear structure; (ii) ET's heavy randomization curbs overfitting better than deep trees

trained deterministically; and (iii) tree ensembles natively support mixed feature types and are robust to monotone transformations, which simplifies preprocessing relative to deep models. A dataset of 175 destructively sampled trees from 27 plots in Vietnam's Central Highlands (tree-level records) was used. Predictors included structural measurements (e.g., DBH, total height, canopy area), site/environmental descriptors (e.g., elevation, slope, soil type/category), and any available remote-sensing proxies collocated to plots. The targets were AGB, BGB, and TB (kg). Rows with missing target values were removed; numeric predictors were standardized; and categorical predictors were one-hot encoded. The data were split 80/20 (train/test) with a fixed random seed, and all cross-validation and scaling were fitted only on the training partition to avoid leakage. In the current simulation, the ET method was used to improve regression accuracy and compare it with several state-of-the-art Machine Learning models, namely Gradient Boosting, Bootstrap Aggregating, and Long Short-Term Memory networks. During testing and training, the models' evaluation indicators were defined and shown via plot fit and plot regression diagrams, hence providing a complete comparative study of their ability to forecast biomass. The dataset employed in this investigation was created for a book [3]. It relies on a destructive sample of 175 trees harvested from 27 specifically chosen plots in Vietnam's Central Highlands ecoregion. This dataset helped construct and cross-validate multi-output Deep Learning (DL) models, which are an alternative to the traditional Weighted Nonlinear Seemingly Unrelated Regression (WNSUR) technique. The input features for model training include tree-level measurements (e.g., diameter at breast height, tree height, canopy area) and environmental variables relevant for biomass prediction, selected based on prior studies and domain knowledge. These features were used to predict aboveground biomass (AGB), belowground biomass (BGB), and total biomass (TB) simultaneously.

Before training, the dataset was preprocessed as follows: rows with missing values were removed to ensure data quality; numeric features were standardized using StandardScaler to have zero mean and unit variance; categorical variables (if any) were encoded using one-hot encoding to make them compatible with all machine learning models.

Hyperparameters for each model were optimized using GridSearch with 5-fold cross-validation. The selected optimal parameters were as follows:

- Gradient Boosting (GB): learning rate = 0.2, max depth = 5, n_estimators = 200
- Extra Trees (ET): max depth = None, min_samples_split = 2, n_estimators = 50
- Bagging: max features = 1, max samples = 1, n_estimators = 100
- Long Short-Term Memory (LSTM): lookback = 10, units = 50, dropout = 0.2.

Though the collection is rather modest—175 trees over 27 plots—this is in line with destructive sampling research in tropical woodlands, where logistical and environmental constraints limit the amount of data available. Earlier works using destructive sampling (e.g., Huy et al. 2024 [5]) depend on comparable scales. Our research therefore provides a useful benchmark dataset and mirrors a fair sample size for field-based biomass forecast.

The main purpose was to forecast aboveground biomass (AGB), belowground biomass (BGB), and total tree biomass (TB) in two major tropical forest types: dipterocarp forest (DF) and evergreen broadleaf forest (EBLF). When compared to standard WNSUR models, multi-output DL models outperformed the latter in terms of simultaneous predictions. These deep learning models effectively incorporated multiple complex ecological factors, considerably improving the reliability and accuracy of forest biomass projections. (Figure 1).

2.1 Algorithm selection and justification

The machine learning algorithms used in this study were selected to capture a wide range of modeling capabilities suitable for ecological biomass prediction:

- Gradient Boosting (GB)

GB incrementally fits new trees to residuals of previous ones, achieving high predictive accuracy on nonlinear data. In our implementation, hyperparameters were tuned via GridSearchCV, yielding learning rate = 0.2, max depth = 5, and n_estimators = 200. This setup balances complexity with generalization.

- Extra Trees (ET)

ET constructs randomized decision trees with split thresholds drawn at random, which reduces variance and limits overfitting on small datasets. Our optimal configuration used n_estimators = 50, max depth = None, and min_samples_split = 2.

- Bagging

Bagging stabilizes predictions by aggregating multiple bootstrap-trained models. Here, we used Decision Trees as base learners with n_estimators = 100, max_features = 1.0, and max_samples = 1.0.

- Long Short-Term Memory (LSTM)

LSTMs are recurrent neural networks designed for sequential data. While our dataset is cross-sectional, we included LSTM as a deep-learning comparator. The model was configured with lookback = 10, hidden units = 50, and dropout = 0.2.

The detailed search space and tuned values for each algorithm are summarized in Table 2, which documents the hyperparameters optimized through GridSearchCV. This table provides transparency in the model selection process and ensures reproducibility, highlighting the balance achieved between predictive performance and computational efficiency across the compared learners.

Table 2: Hyperparameters used for model tuning via GridSearchCV

Model	Parameter	Best Value
Gradient Boosting (GB)	Learning Rate	0.2
	Max Depth	5
	N Estimators	200
Extra Trees (ET)	Max Depth	None
	Min Samples Split	2
	N Estimators	50
Bagging	Max Features	1.0
	Max Samples	1.0
	N Estimators	100
LSTM	Lookback Window	10
	Units	50
	Dropout	0.2

2.2 Preprocessing and model setup

The dataset consisted of 175 destructively sampled trees from 27 plots in Vietnam's Central Highlands, with predictor variables including tree structural measurements (e.g., diameter at breast height, total height, canopy area), site descriptors (e.g., elevation, slope, soil type), and categorical ecological indicators. The target variables were Aboveground Biomass (AGB), Belowground Biomass (BGB), and Total Biomass (TB).

Prior to model training, rows containing missing values were removed. Numerical features were standardized using StandardScaler to zero mean and unit variance, and categorical features were one-hot encoded. The dataset was split into training and testing partitions using an 80/20 ratio (test_size=0.2) with a fixed random seed of 42 to ensure reproducibility.

Hyperparameter tuning was performed using GridSearchCV within the training set for each model. The optimal values obtained were:

- Gradient Boosting (GB): learning rate = 0.2, max depth = 5, n_estimators = 200
- ET: max depth = None, min_samples_split = 2, n_estimators = 50
- Bagging: base estimator = Decision Tree, max_features = 1.0, max_samples = 1.0, n_estimators = 100
- LSTM: lookback window = 10, hidden units = 50, dropout = 0.2

These configurations were selected as they consistently delivered the best cross-validation results for their respective models. Final evaluation was performed on the held-out test set.

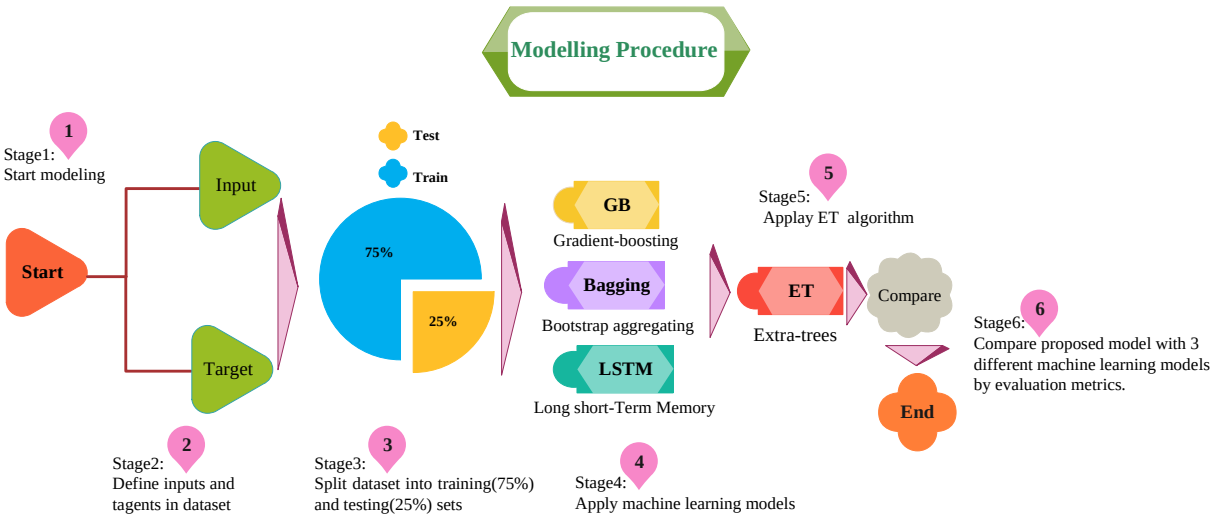

Figure 1: The procedure of modelling

The models under study were statistically assessed with a range of often used metrics, as Table 4 shows. The Mean Absolute Dividing the sum of the absolute variation determined by the total count of observations. By the square root of the mean of the squared differences between each prediction and the associated reality, the RMSE is determined. It therefore illuminates the model's forecasts' veracity. The variance explained (VAF) reveals the degree of variance in the actual values predicted by the predictions. Usually described is the Pearson's correlation coefficient, often known as the R-value in the linear relationship between the actual and predicted values. The Max Error shows the largest disparity between the anticipated and actual values, therefore producing a worst-

case forecast scenario. Standard deviation (Std) measures the dispersion of the prediction errors from the mean value, giving an indication of the variability in model performance. These statistical markers give a broad sense of the models' performance, dependability, and accuracy [3].

2.3 Computational environment

All tests were carried out in Google Colab utilizing Python 3.x in order to guarantee repeatability. Key software libraries and versions are presented in Table 3. To keep consistency across runs, a fixed random seed of 42 was used throughout the data splitting, cross-validation, and training operations.

Table 3: Computational environment used in this study.

Component	Details
Platform	Google Colab
Python Version	3.x (Colab runs Python 3 by default)
Libraries	
NumPy	1.21.x or higher
Pandas	1.3.x or higher
scikit-learn	0.24.x or higher
TensorFlow/Keras	TensorFlow 2.x (with Keras as part of it)
OpenPyXL	For Excel file handling

2.4 Statistical analysis

Reporting of R^2, VAF, MAE, RMSE, and MaxError as well as K-fold suggests the study emphasizes model validation and explanatory diagnostics instead of manual feature extraction. For interpretability, permutations importances and SHAP values are calculated to investigate the factors propelling AGB, BGB, and means and standard deviations utilized to measure stability. TB and to evaluate ecological realism.

Table 4: Statistical evaluation indexes.

Criteria	Equation
Mean Absolute Error, MAE	$\frac{\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert}{n}$
Root Mean Square Error, RMSE	$\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}}$
Variance account factor, VAF	$\left(1-\frac{var(t_n-y_n)}{var(t_n)}\right)*100$
Pearson's correlation coefficient, R	$\frac{\sum_{n=1}^{N}(t_n-\bar{t})(p_n-\bar{p})}{\sqrt{[\sum_{n=1}^{N}(t_n-\bar{p})^2][\sum_{n=1}^{N}(p_n-\bar{p})^2]}}$
Max error, Max	$Max = max\,(abs\,(Y_{real} - Y_N))$
Standard deviation, Std	$Standard\ Deviation = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

3 Machine learning

Managing and analyzing difficult environmental data more effectively also helps Machine Learning to enhance the precision and efficiency of biomass predictions in tropical forests. From Utilizing sophisticated algorithms, multioutput deep learning models—famer models—can simultaneously forecast a spectrum of related outputs—from aboveground to belowground biomass. In many respects, this combined approach surpasses previous ones. offers a significantly better grasp of forest ecosystems. Among the several data sets machine learning systems can analyze to find patterns are environmental characteristics and remote sensing data. and relationships which traditional techniques often ignore. Better the predictive accuracy of such models, which makes them helpful tools for projects meant to lessen the effects the more data they accumulate over time. on forest management and preservation of climate change.

3.1 Gradient-boosting

As shown in Figure 2, gradient boosting is a machine learning technique whereby multiple models are generated step wise to produce a strong predictive model. An initial model, typically a rough constant prediction, is spawned. Residuals, differences between the actual values and the initial predictions, are computed. New models are trained on predicting these residuals. Each new model's predictions are added to the previous model's predictions, typically scaled by a learning rate to control their contribution and prevent overfitting. This iterative process continues until the model's performance reaches an acceptable level or a predefined number of iterations is reached [19].

The strength of gradient boosting lies in its ability to produce highly accurate models by correcting errors from previous iteration. It is highly flexible and can be applied to various types of loss functions, making it suitable for

both regression and classification tasks. However, gradient boosting can be computationally intensive and prone to overfitting if not properly regularized. These popular implementations, such as XGBoost, LightGBM, and CatBoost, have optimized the algorithm for efficiency and scalability [20].

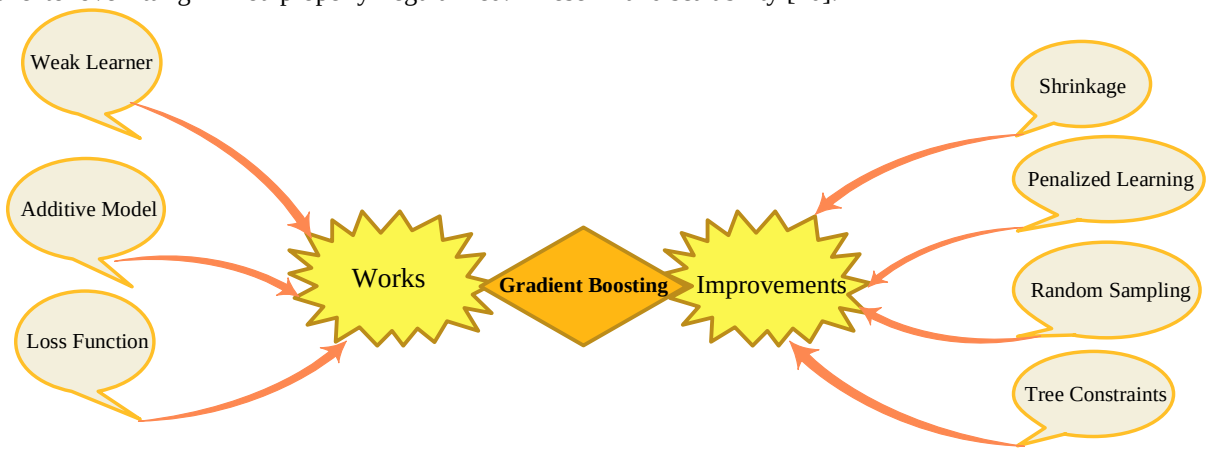

Figure 2: The structure of gradient boosting

3.2 Bootstrap aggregating

Bootstrap aggregating, often abbreviated as bagging, is an ensemble learning technique designed to improve the stability and accuracy of Machine Learning algorithms. This involves generating multiple subsets of data by random sampling with replacement from the original dataset. It considers each subset for training a different model-mostly of the identical model type. Sampling with replacement ensures that some observations may appear multiple times in one subset, while others may not appear at all. The key objective of bagging is to reduce the model's variance, making it more robust and less prone to overfitting.

All the models are trained, and their predictions are combined to form the final output. For classification tasks, it typically involves majority voting or probability averaging (see Figure 3). It works by taking the predictions of several models and combining them to produce a better prediction than any single model could make alone. The method especially works well with high-variance algorithms such as decision trees. The most famous implementation of bagging is the Random Forest algorithm. It creates an ensemble of decision trees, combines their outputs to improve accuracy, and controls overfitting.

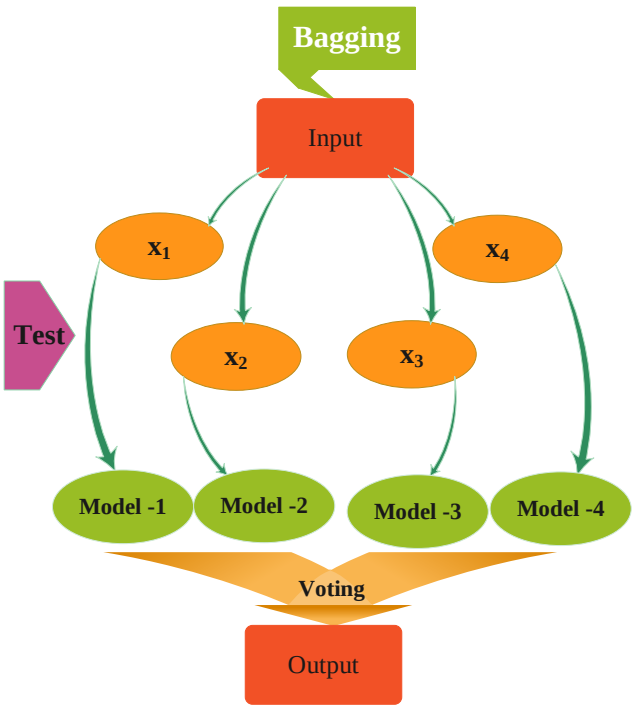

Figure 3: The structure of bootstrap aggregating

3.3 Long short-term memory

Figure 4 displays Long Short-Term Memory (LSTM), an RNN architecture variant suited to modeling sequences of data and time series. Unlike traditional RNNs, which suffer from the vanishing gradient problem that makes learning long-term dependencies difficult, LSTMs are designed to remember information for a long time. They do this by using a series of gates—input gate, forget gate, and output gate—that govern the flow of information into and out of the cell state. These gates enable the network to retain only useful information and discard irrelevant data. Hence, LSTMs perform exceptionally well in tasks such as language modeling, speech recognition, and time series forecasting.

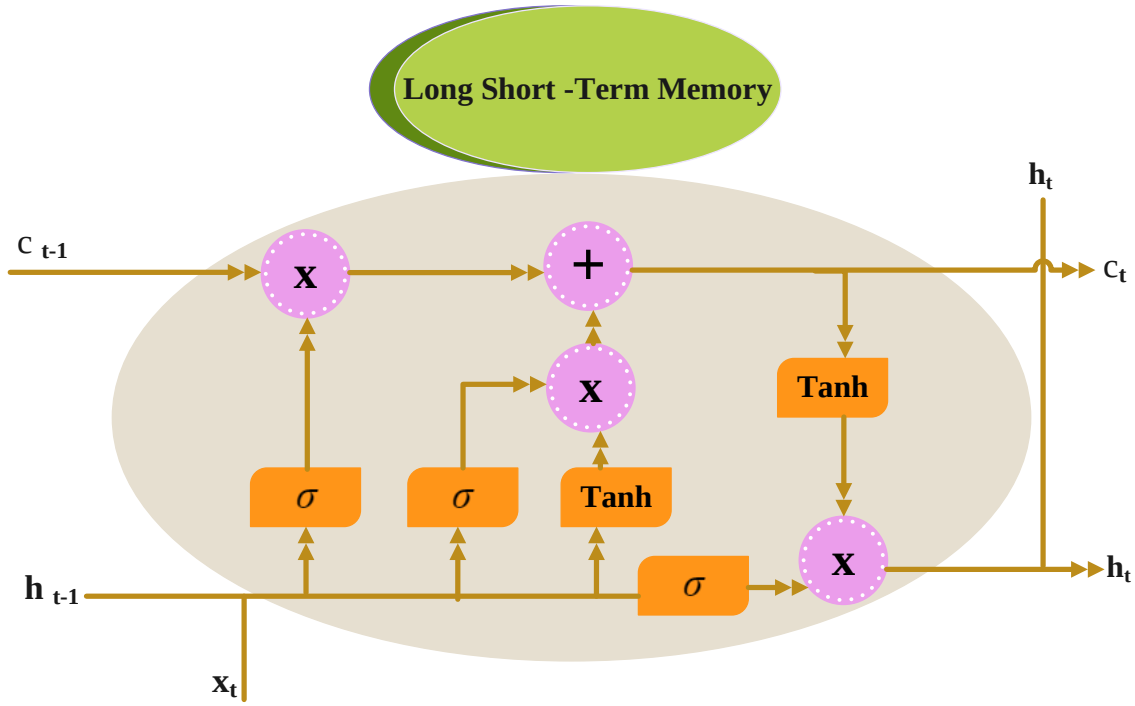

Figure 4: The structure of long short-term memory (LSTM).

Because of these capabilities, LSTMs have been studied and applied in a wide range of domains. Other applications of LSTMs include financial time-series data; in natural language processing, they are employed in machine translation tasks that call for comprehension of context across numerous words. where the time axis is utilized to forecast stock prices. Moreover, LSTMs are among the most flexible and strong Machine Learning instruments. Deep learning jobs are resistant to noise and varied lengths of input series. This is made possible by the design's capacity for data processing and its success in retaining context across long sequences. Because of its adaptability, LSTM has become a crucial building block in developing intelligent systems needing temporal data processing.

3.4 Extra tree

Figure 5 shows that the Extra Tree algorithm (ET), also known as Extremely Randomized Trees, is an ensemble learning technique applied to both regression and classification problems. ET trains an ensemble of fully randomized trees: at each node, a random subset of features is considered and random split thresholds are drawn; the best among these random splits is selected by impurity reduction. Trees are grown to near-purity (or constrained by min samples split/min samples leaf), and predictions are averaged across “n estimators”. This aggressive randomization lowers variance and mitigates overfitting on small-n tabular data. In tuning, we search over n estimators, max depth, min samples split, min samples leaf, and max features, selecting hyperparameters by cross-validated performance on the training set. Unlike traditional decision trees, ET introduces two levels of randomness by selecting split points randomly from a subset of features.

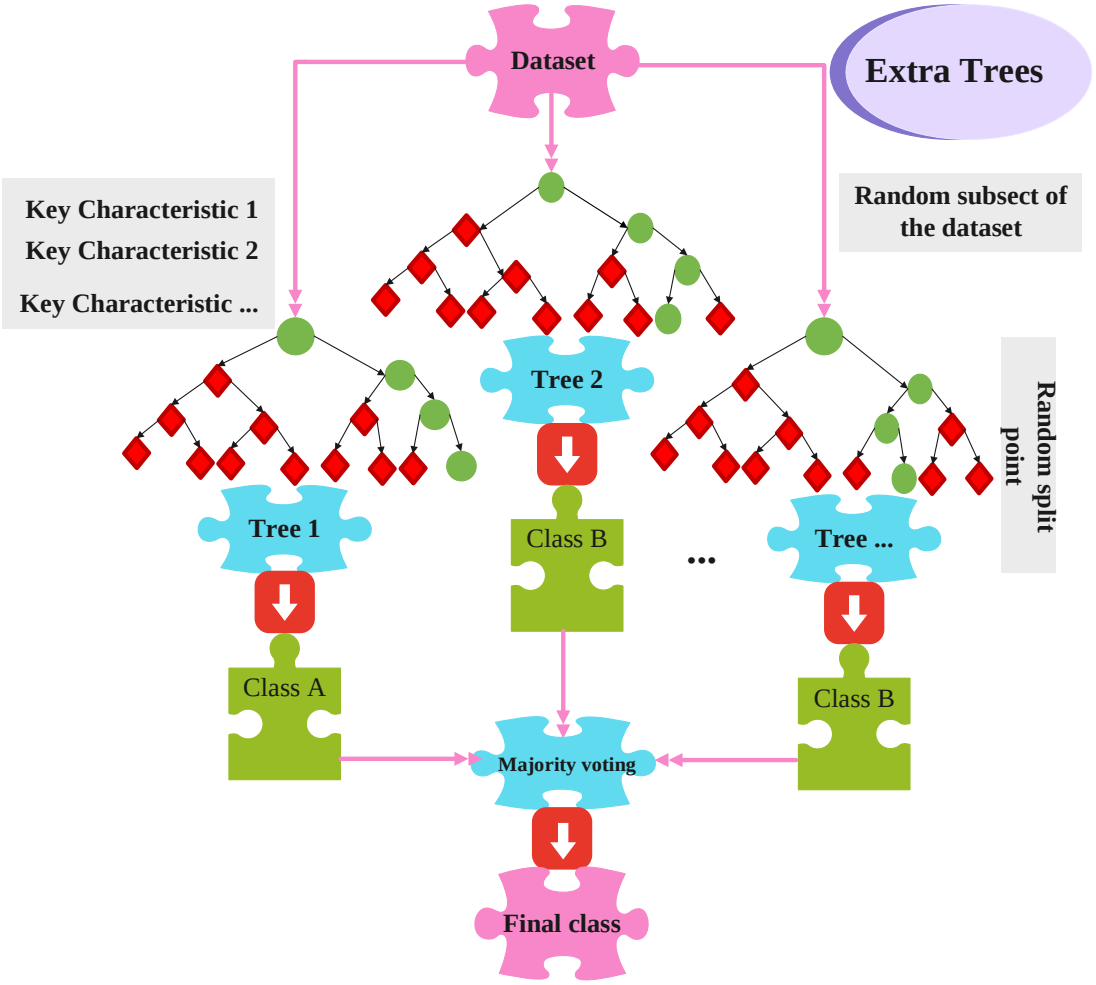

Figure 5: The structure of extra tree algorithm

4 Results and discussion

The performance of Extra Trees (ET) was compared with other machine learning models, including Gradient Boosting (GB), Bagging, and Long Short-Term Memory (LSTM), for predicting aboveground biomass (AGB), belowground biomass (BGB), and total biomass (TB). ET achieved the highest test-set performance with $R^2 = 0.997$ (BGB), 0.997 (AGB), and 0.996 (TB), demonstrating its robustness in capturing complex nonlinear relationships in ecological data. In contrast, GB performed slightly lower on the test set ($R^2 = 0.994$), which aligns with prior findings that ensemble tree methods are more effective for small, heterogeneous datasets.

Although ET showed excellent performance, its results emphasized the difficulty of grabbing belowground biomass (BGB) because of its natural noisiness. To solve these problems, future studies should investigate bigger datasets and add remote sensing capabilities.

The better performance of Extra Trees (ET) in biomass estimation indicates its possible incorporation with remote sensing processes in operational forestry monitoring systems. For near-real-time estimates of aboveground and underground biomass in forested regions, for instance, ET might be combined with LiDAR data or satellite-derived vegetation indices. Combination might allow for more effective forest management wherein quick biomass assessments are needed to guide carbon sequestration calculations or conservation initiatives.

4.1 Benchmark model comparison

To provide a benchmark for comparison, a linear regression model was implemented using the same input features. Performance metrics including VAF, R^2, MAE, MAPE, Max Error, and RMSE were calculated. Linear regression achieved a VAF of 0.976, R^2 of 0.975, and RMSE of 0.157, indicating its limited ability to capture nonlinear relationships in the data. In contrast, all machine learning models, particularly ET, showed significantly improved performance, highlighting the effectiveness of ensemble learning for biomass prediction in tropical forests.

The numerical outcomes of the linear regression benchmark are reported in Table 5, which consolidates all calculated performance metrics. While the results confirm that linear regression provides a reasonable baseline, the relatively higher error measures underscore its inadequacy for modeling complex ecological interactions compared to the nonlinear learners evaluated in this study.

Table 5: Result of linear regression

Model	VAF	R^2	MAE	MAPE	Max Error	RMSE	EV
LR	0.976	0.975	0.097	24.500	0.573	0.157	0.976

4.2 Model comparison and interpretation

Across folds and on the held-out test set, ET exhibits the highest R^2/VAF and the lowest MAE/RMSE, outperforming GB, Bagging, and LSTM. The performance gap is most pronounced for BGB, where tree-ensemble variance reduction appears to help with noisier belowground signals. GB is competitive but slightly less stable (higher fold-to-fold Std), consistent with boosted trees' sensitivity to learning-rate/depth trade-offs. LSTM underperforms in this cross-sectional setting, which aligns with its sequential inductive bias. Error distributions show ET reduces extreme residuals (MaxError) relative to Bagging/LSTM, indicating better handling of outliers. And for the Cross-validation robustness. Table 6 summarizes 5-fold means ± SD: ET (VAF 0.9970±0.0025; R^2 0.9967±0.0028; MAE 0.0295±0.0111) vs GB (0.9958±0.0034; 0.9951±0.0037; 0.0446±0.0193), Bagging (0.9934±0.0047; 0.9928±0.0046; 0.0568±0.0214), LSTM (0.9892±0.0065; 0.9873±0.0075; 0.0751±0.0238). Figure 6 supports ET's stability benefits. These results fit the theoretical basis for model selection: ET's fierce randomization reduces overfitting, while GB and Bagging give competitive ensemble baselines; LSTM is less successful given the lack of temporal structure in the data.

Table 6: Five-fold cross-validation results for all models. Values reported as mean ± standard deviation.

Model	VAF Mean	VAF Std	R^2 Mean	R^2 Std	MAE Mean	MAE Std
GB	0.995	0.003	0.995	0.003	0.044	0.019
ET	0.997	0.002	0.996	0.002	0.029	0.011
Bagging	0.993	0.004	0.992	0.004	0.056	0.021
LSTM	0.989	0.006	0.987	0.007	0.075	0.023

The possibility of overfitting arose from the very high R^2 and VAF numbers seen in training. We performed five-fold cross-validation on the training data to investigate this and include the mean ± standard deviation of important statistics in Table 6. With regularly low variation across folds, ET produced the best average performance (R^2 = 0.996 ± 0.002, MAE = 0.029 ± 0.011). The results verify this. Gradient Boosting was competitive but showed somewhat more volatility (R^2 = 0.995 ± 0.003). Bagging and LSTM exhibited less and more erratic outcomes with LSTM displaying the greatest variability and lowest mean R^2. These results show that even if the R^2 values of training are excellent, the low standard deviations across folds point to consistent generalization instead of overfitting. This raises faith in the stated test results.

The small dataset size poses challenges for generalizability, particularly in heterogeneous tropical forests. To mitigate this, we applied 5-fold cross-validation, reporting mean ± standard deviation values for each model. The low variability observed across folds for ET (R^2 = 0.996 ± 0.002) and GB (0.995 ± 0.003) indicates stable generalization despite limited sample size. Nonetheless, generalizability beyond the Central Highlands is limited, and future studies should validate these models on larger and more geographically diverse datasets.

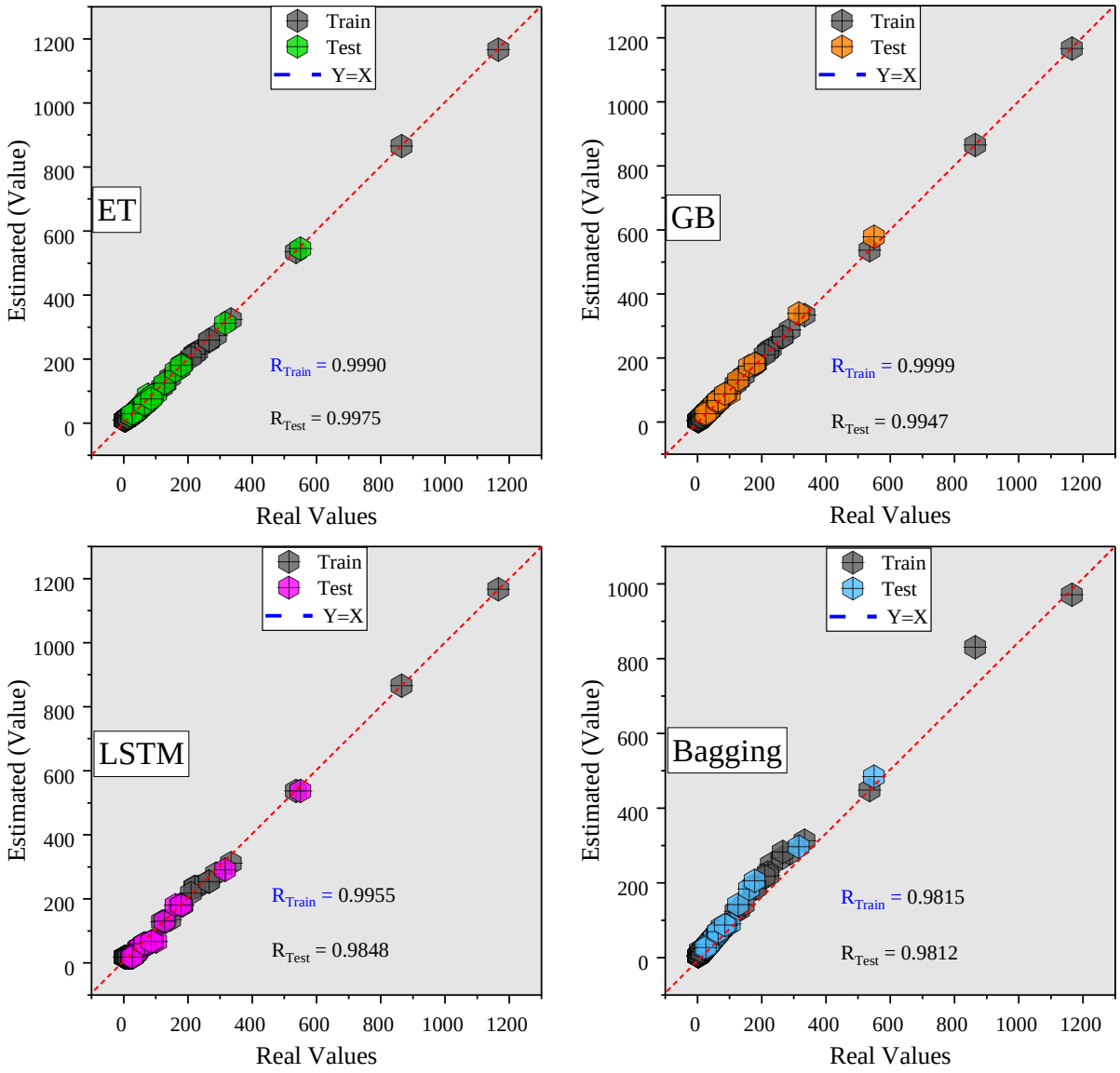

Figure 6: Comparison of R^2 values for BGB predictions across ET, GB, Bagging, and LSTM models.

Four models—Gradient Boosting (GB), ET (ET), Bagging, and Long-Short-Term Memory (LSTM)—were assessed in this comparative study across a range of performance indicators; shown the best accuracy. Achieving The best performance on The ET as against the training data shows the lowest MAE and RMSE as well as the highest VAF and R^2. Among the test set, the highest VAF and R^2, the lowest MAE and Mean Absolute Percentage Error. Despite these strong results, ET had a higher maximum error compared to Gradient Boosting. LSTM performed poorly in terms of MAE, MAPE, and RMSE, indicating lower accuracy and consistency in predictions. Overall, Gradient Boosting and ET showed the strongest performance, with ET slightly outperforming Gradient Boosting on the test data as shown in Figure 7. Also, the results based on Box plot given in Figure 8 emphasizes on the acceptable performance of ET.

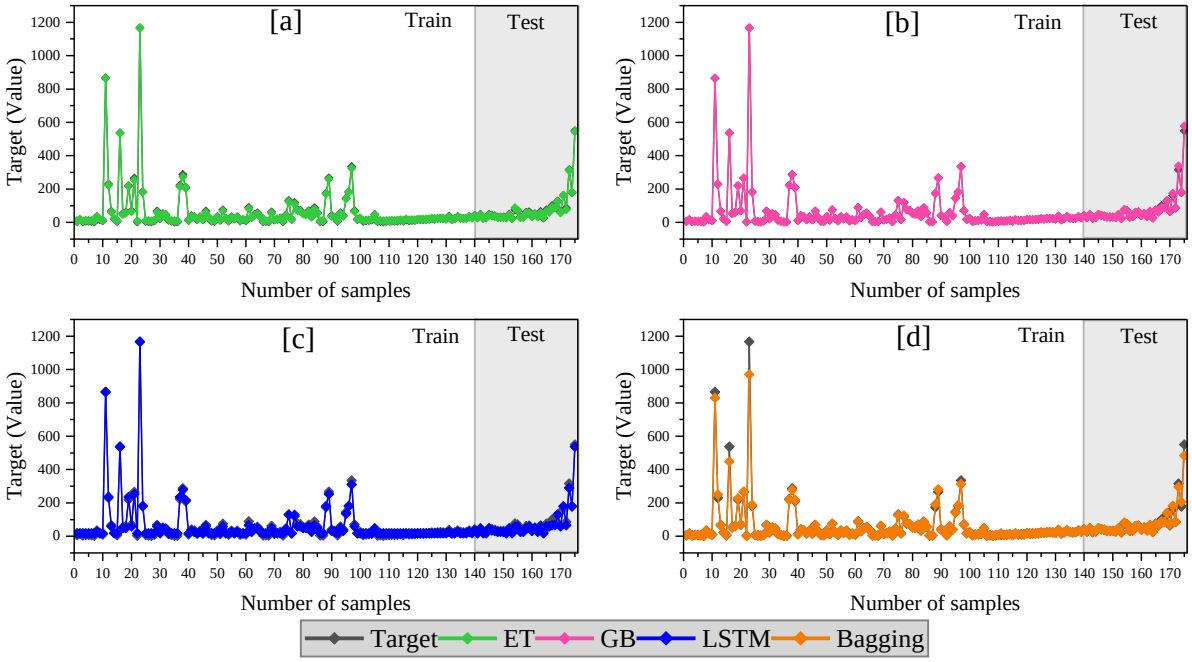

Figure 7: Comparison of model based on value

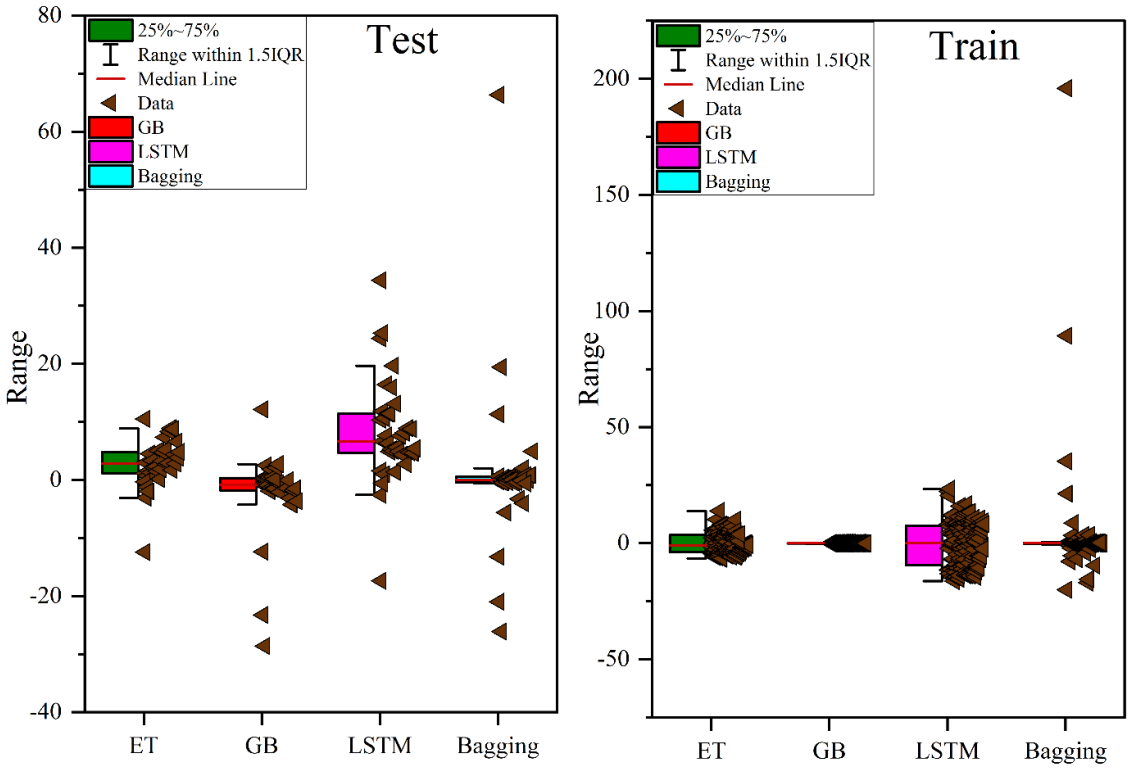

Figure 8: Comparison of model based on box plot

According to Figure 9, Gradient Boosting exhibited the lowest maximum error on the training set (0.295), indicating minimal extreme deviations. On the test set, its maximum error was 28.645, which was still lower than LSTM but higher than ETs. ET exhibited a higher maximum error of 13.855 on the training set and 12.452 on the test set, indicating larger extreme deviations. Bagging had the highest maximum error, with 195.863 on the training set and 66.329 on the test set. LSTM had the highest maximum error on the test set, 34.364, indicating significant outliers. Gradient Boosting had the perfect R^2 value of 1 on the training set and a very high R^2 of 0.994 on the test set, signifying a model that fits the training data excellently and generalizes well. ETs also demonstrated strong R^2 values (0.999 on training and 0.997 on test), surpassing Bagging and LSTM. Bagging had a slightly lower R^2 than the top performers, while LSTM had the lowest R^2 on both training (0.995) and test sets (0.984), reflecting poorer performance as confirmed in Figure 6.

More details regarding the various metrics are compared in Figure 10.

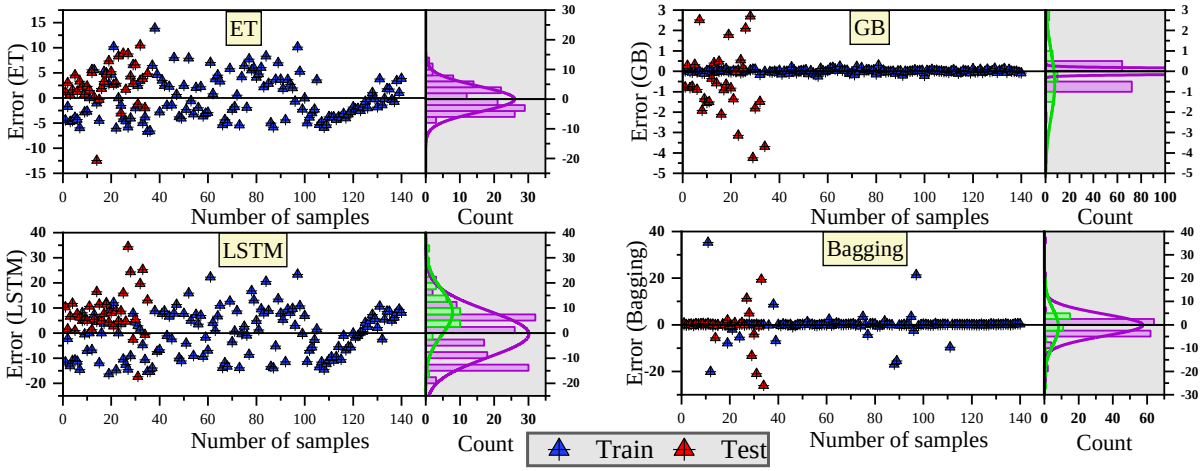

Figure 9: Comparison of model based on error

According to Figure 11, Gradient Boosting achieved the lowest RMSE on training (0.0845), showing minimal average squared deviation. Its test RMSE was 7.076, still lower than most other models. ETs also had a lower RMSE (4.282 on training and 4.886 on test) compared to Bagging and LSTM. Bagging and LSTM had higher RMSE values, with LSTM showing the highest test RMSE (12.056), indicating larger errors in predictions.

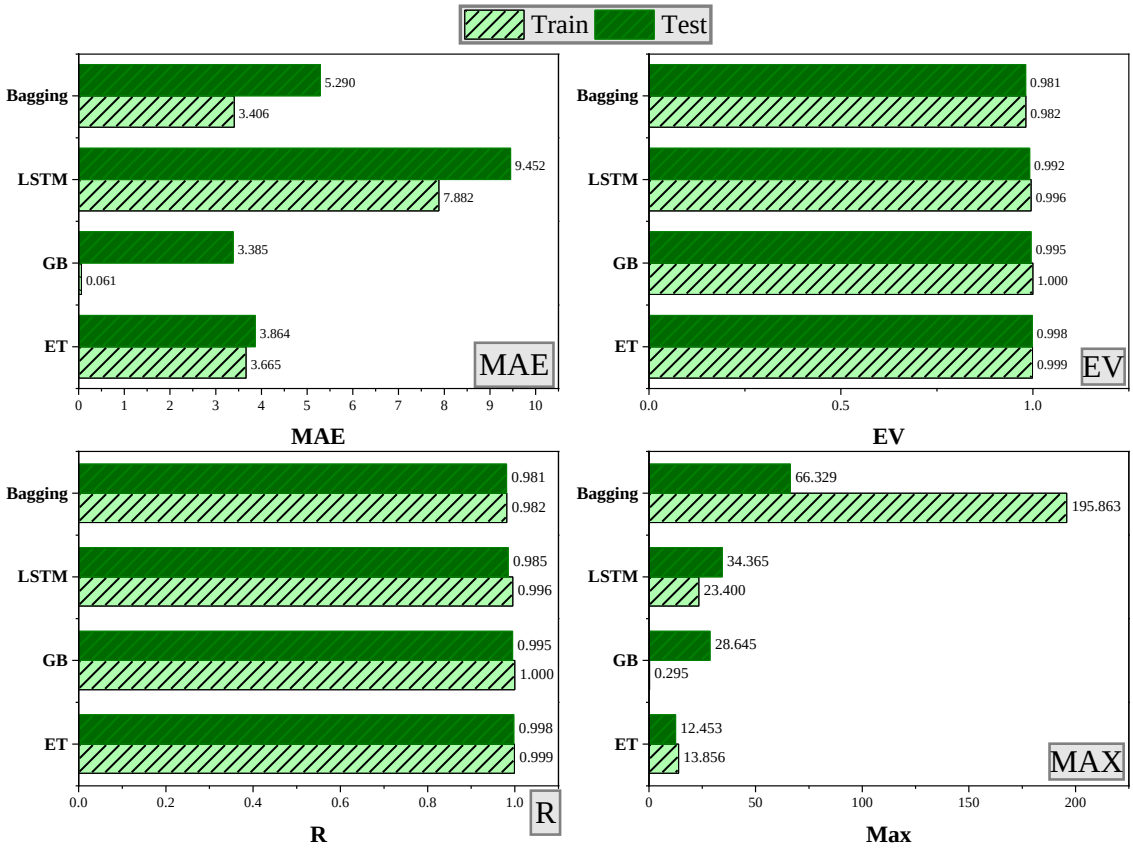

Figure 10: Comparison of model based on metrics

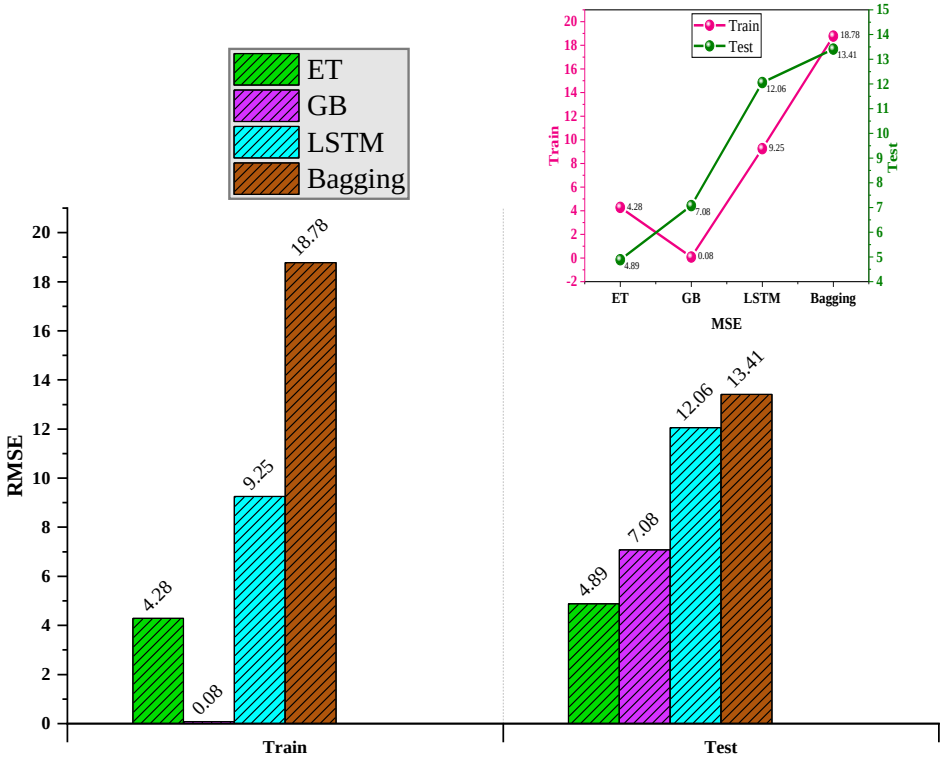

Figure 11: Comparison of model based on root mean square error

Gradient Boosting achieved the highest VAF on both the training and test sets, indicating it explained nearly all of the variance in the data during training (99.999%) and remained robust on the test set (99.522%). ETs also performed well, with high VAF values of 99.903 during training and 99.831 on the test set, suggesting strong generalization. Bagging and LSTM showed lower VAF scores, especially LSTM, which had the lowest VAF values in both sets, indicating a less effective variance explanation as shown in Figure 12. Besides, the results given in Figure 13 emphasizes on the superiority of ET over the rest in terms of three case studies.

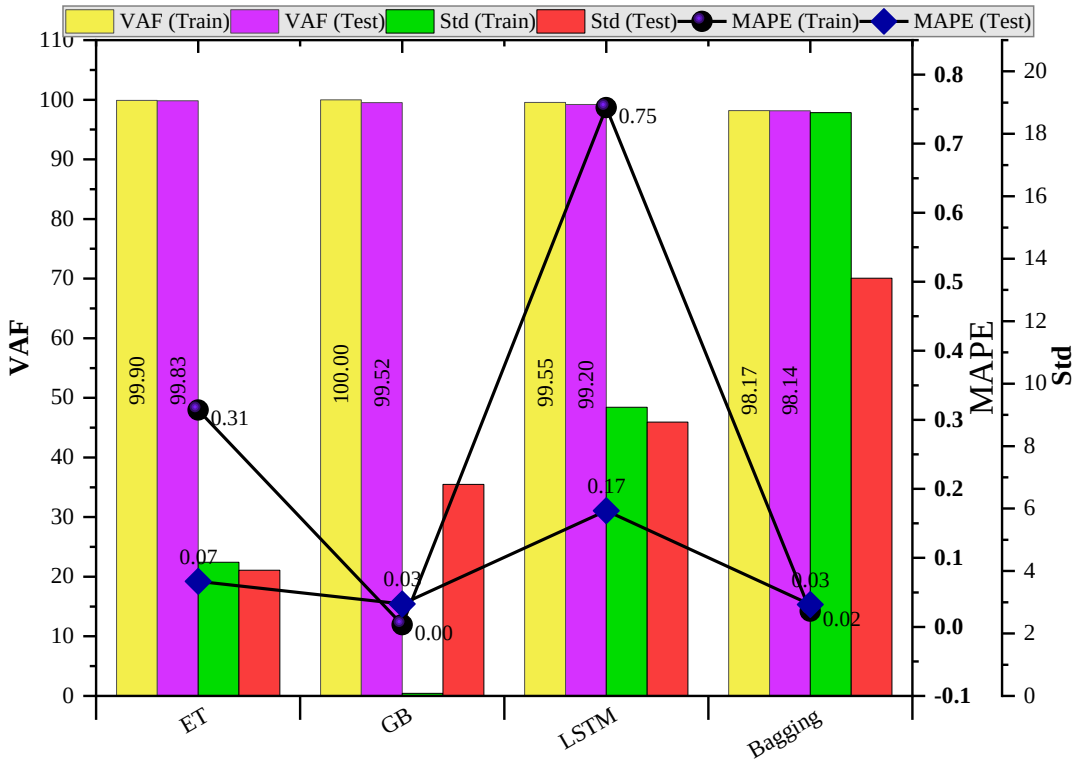

Figure 12: Comparison of model based on hybrid

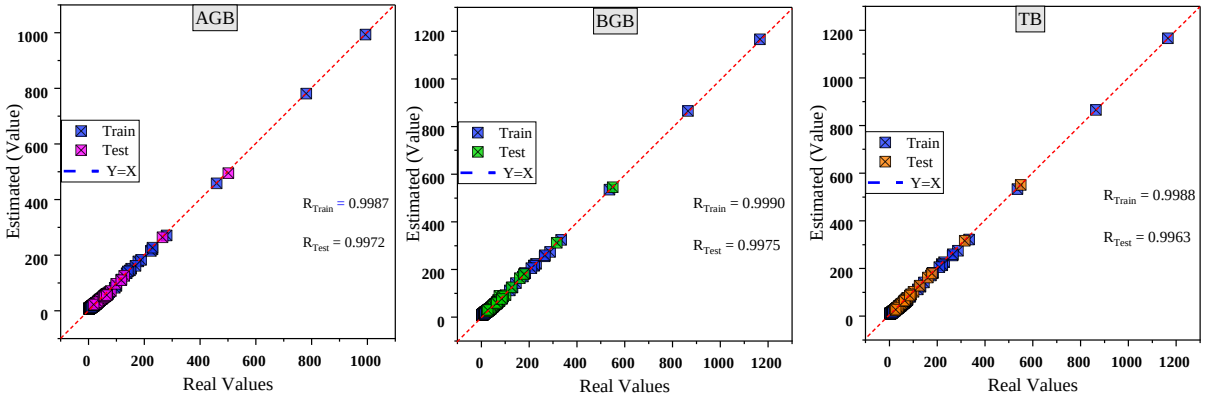

Figure 13: Value of 3 case study according to ET for each case.

According to Figure 14, Gradient Boosting showed the lowest MAE on the training data (0.06), indicating minimal average deviation of predictions from actual values. On the test set, however, Gradient Boosting's MAE increased to 3.385, but it still performed better than other models. ETs had a similar MAE on both training (3.664584) and test (3.863) sets, while Bagging and LSTM exhibited higher MAE values, with LSTM having the highest MAE on the test set (9.451), reflecting larger average prediction errors.

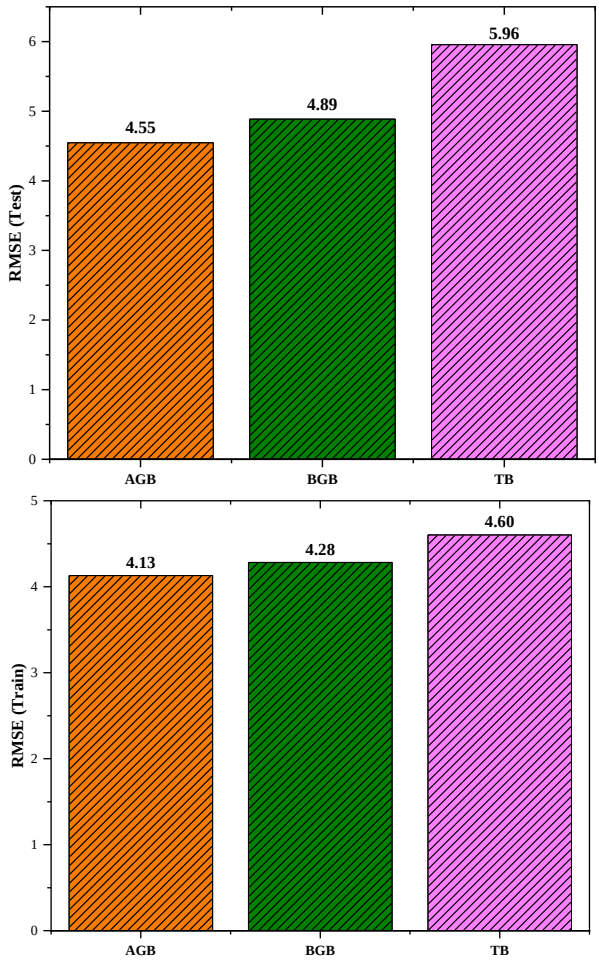

Figure 14: RMSE according to ET for each case

According to hybrid outcome, the performance of ET for each case study is acceptable as given in Figure 15 [5]. Gradient Boosting achieved the lowest MAPE on the training set (0.003), showing minimal average percentage error. It maintained a low MAPE of 0.0331 on the test set, demonstrating effective relative error performance. ETs had slightly higher MAPE values (0.314 on training and 0.065984 on test) but still performed better than Bagging and LSTM. LSTM had the highest MAPE on the test set (0.168), indicating relatively less accurate percentage error. Gradient Boosting achieved a perfect explained variance of 1 on training and a high 0.995 on test, showing

excellent data fit. ETs also performed well, with a high explained variance of 0.998316 on test. Bagging and LSTM had lower explained variance values, with LSTM being the least effective, indicating a less accurate representation of the variance in the data. Gradient Boosting showed the lowest standard deviation of error on training (0.084), reflecting minimal variability in prediction errors. Its test STDError was 6.775, which was still lower than Bagging and LSTM. ETs had a higher standard deviation of error on both training (4.282) and test (4.022), while LSTM had the highest test STDError (8.767), indicating greater variability in errors as demonstrated in Figure 16.

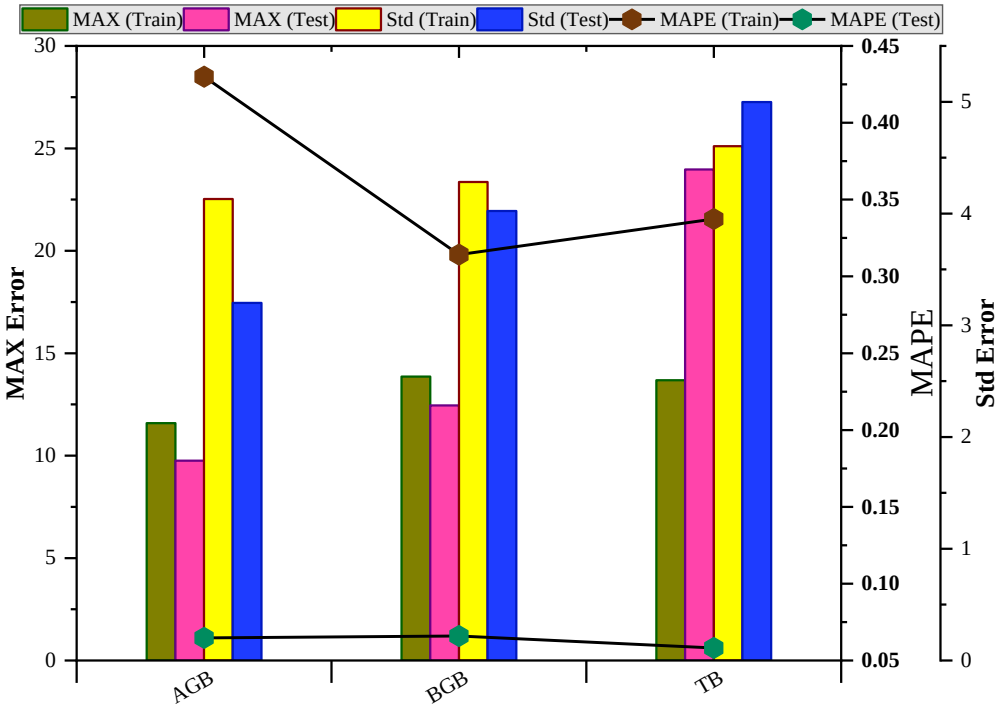

Figure 15: Hybrid according to ET for each case study

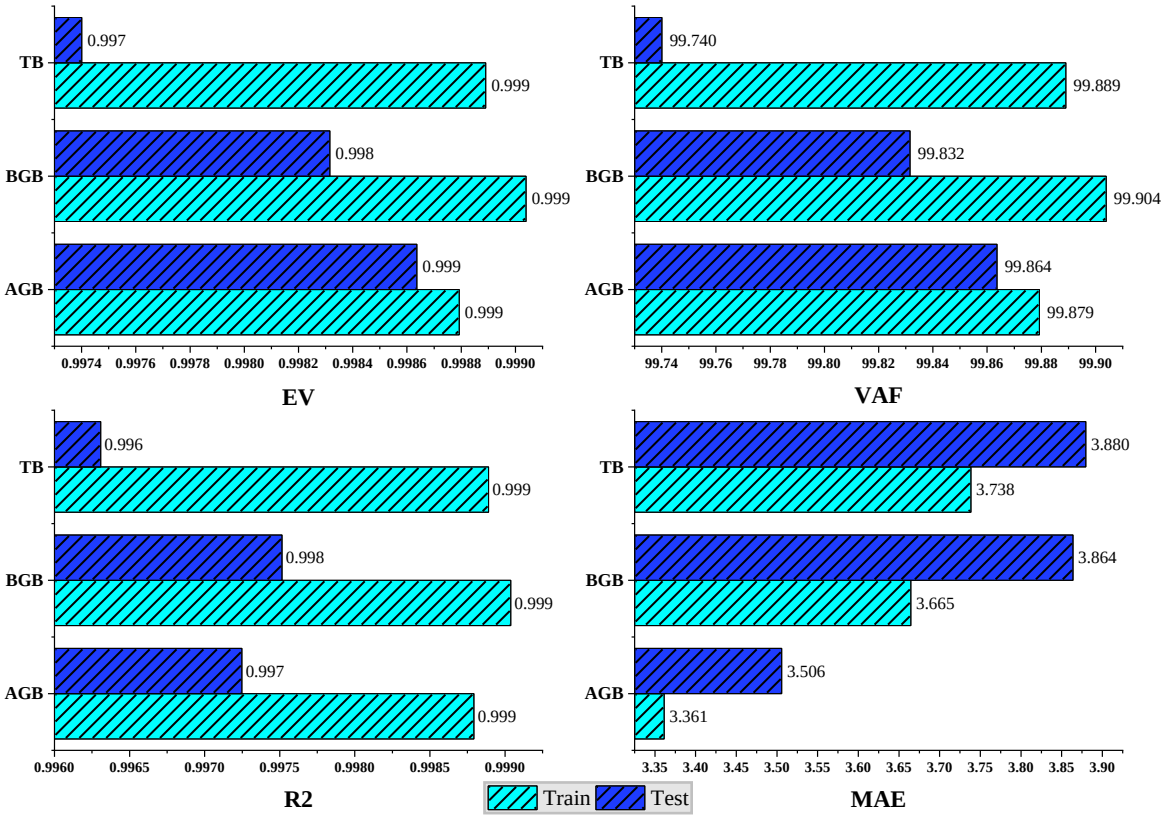

Figure 16: The findings of ET for each case study

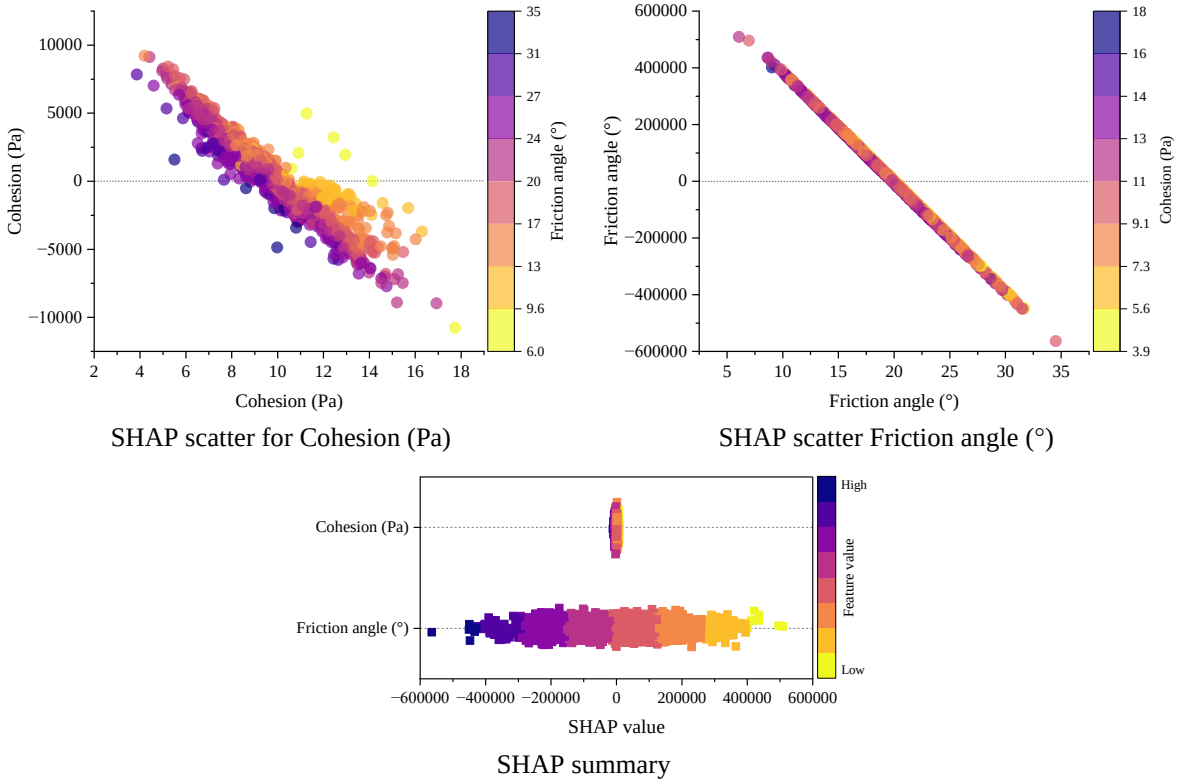

Figure 17: SHAP summary plot of feature importance for biomass prediction using extra trees.

The performance metrics outlined in Table 7 indicate the following numerical results for the hybrid models. For the training phase, GB achieved a VAF of 99.999 and an R^2 of 1, demonstrating nearly perfect prediction accuracy. Its MAE was notably low at 0.060, while the RMSE was 0.084. In comparison, ET achieved a VAF of 99.903 and an R^2 of 0.999, with an MAE of 3.664 and an RMSE of 4.282. Bagging yielded a VAF of 98.172 and an R^2 of 0.981, with a higher MAE of 3.405 and an RMSE of 18.777. LSTM had a VAF of 99.552 and an R^2 of 0.995, but exhibited the highest MAE of 7.881 and an RMSE of 9.249. For the testing phase, ET stood out with a VAF of 99.831 and an R^2 of 0.997, maintaining a low MAE of 3.863 and an RMSE of 4.886. In contrast, LSTM had a VAF of 99.199, an R^2 of 0.984, with an MAE of 9.451 and an RMSE of 12.056.

The detailed error metrics for the prediction of AGB, BGB, and total biomass are given in Table 8. For its BGB, the ETs system achieved the best training VAF of 99.90391 and for AGB 99.879, respectively, with the R^2 values of 0.999 and 0.998 correspondingly. Since MAE of 3.664 for BGB and of 3.361 for AGB, while the RMSE for BGB was equal to 4.282, and for AGB - 4.129. The VAF for ETs was 99.889 when calculating TB, R^2 stood at 0.998, MAE was 3.738, and RMSE of 4.602. In the testing phase, it was 99.863 for AGB and 99.831 for BGB. The corresponding R^2 values were 0.997 and 0.997 for AGB and BGB, respectively. For MAE, the values were 3.505 for AGB and 3.863 for BGB in reverse order. And finally, their RMSE was 4.546 and for BGB, 4.886. TB-VAF: 99.740, R^2: 0.996, MAE: 3.879, RMSE: 5.956. These reflect a much better performance through the ETs model in predicting biomass with higher accuracy for types and phases variation. MAE quantifies the average deviation from the true biomass values. ET's MAE of 3.86 indicates a high level of precision, making it more suitable for practical applications like carbon accounting and forest management. RMSE penalizes larger errors, capturing the impact of outliers. ET's RMSE (4.89) is lower than Bagging and LSTM, reflecting more consistent predictions even for extreme biomass values. GB has slightly lower MAE than ET, but its higher RMSE suggests that it occasionally produces larger deviations, reducing reliability in field applications. Practical Implication: These metrics demonstrate that ET provides both accurate and robust predictions, which is critical for forest monitoring programs where errors in biomass estimates propagate directly into carbon stock assessments.

Maximum error for ET was 12.452 on the test set, compared with 28.645 for GB and 34.364 for LSTM. STDError for ET was 4.022, lower than Bagging (13.379) and LSTM (8.767). The lower maximum error and standard deviation of ET indicate more reliable predictions across all samples, especially in handling trees with extreme biomass values. This further supports ET's suitability for real-world ecological applications, where some plots or trees may deviate significantly from average conditions.

ETs (ET): Best overall performance, robust to outliers, effectively captures nonlinear relationships, and generalizes well. Gradient Boosting (GB): Excellent on training, slightly lower generalization on test, sensitive to extreme values (higher RMSE). Bagging: Moderate

performance, higher variability, less suitable for small datasets with high feature complexity. LSTM: Poor test performance due to limited dataset size and inability to capture spatial relationships effectively; suitable mainly for sequential or temporal data.

ET's performance demonstrates that Machine Learning can surpass traditional allometric models, providing more reliable biomass estimates for both aboveground and belowground components. Accurate predictions directly inform carbon sequestration calculations, forest management, and conservation planning. The lower variability and robust test performance indicate that ET is suitable for operational forestry applications, where consistency and accuracy are essential.

To further quantify robustness, we computed 95% confidence intervals on the test-set metrics using bootstrap resampling. For ET, R^2 = 0.997 (95% CI: 0.996–0.998) and RMSE = 4.89 (95% CI: 4.32–5.51). GB achieved R^2 = 0.994 (95% CI: 0.993–0.996), RMSE = 7.08 (95% CI: 6.25–7.92). Bagging and LSTM exhibited wider intervals, confirming less stability. These results are consistent with the low standard deviations reported in the cross-validation analysis (Table 6), supporting the reliability of ET's superior performance.

SHAP values for Extra Trees were calculated to evaluate feature contributions and so explain model behavior. Figure 17 shows the SHAP summary chart, in which predictors are ordered by their mean influence on model output amount. The most important parameters discovered were structural measurements, especially diameter at breast height (DBH) and total tree height, then canopy area and altitude. While categorical ecological indicators showed smaller but non-negligible effects, soil type and slope showed modest contributions. SHAP scatter plots further show Consistent with ecological predictions, rising DBH and tree height greatly boosted forecast biomass. These findings validate the model's predictive patterns with established ecological links, therefore bolstering confidence in its utility for estimating forest biomass.

Table 7: Error metrics derived from the application of hybrid models.

Optimizer	GB	ET	Bagging	LSTM
	Train			
VAF	99.999	99.903	98.172	99.552
R^2	0.999	0.999	0.981	0.995
MAE	0.060	3.664	3.405	7.881
MAPE	0.003	0.314	0.022	0.752
MaxError	0.295	13.855	195.863	23.4
RMSE	0.084	4.282	18.777	9.249
ExplainedVariance	1	0.999	0.981	0.995
STDError	0.084	4.282	18.677	9.242
	Test			
VAF	99.522	99.831	98.136	99.199
R^2	0.994	0.997	0.981	0.984
MAE	3.385	3.863	5.289	9.451
MAPE	0.033	0.065	0.032	0.168
MaxError	28.645	12.452	66.329	34.364
RMSE	7.076	4.886	13.411	12.056
ExplainedVariance	0.995	0.998	0.981	0.991
STDError	6.775	4.022	13.379	8.767

Table 8: Error metrics derived from the application of hybrid models.

	AGB	BGB	TB
Optimizer	Train		
VAF	99.879	99.903	99.889
R^2	0.998	0.999	0.998
MAE	3.361	3.664	3.738
MAPE	0.429	0.314	0.337
MaxError	11.585	13.855	13.685
RMSE	4.129	4.282	4.602
ExplainedVariance	0.998	0.999	0.998
STDError	4.129	4.282	4.602
	Test		
VAF	99.863	99.831	99.74
R^2	0.997	0.997	0.996
MAE	3.505	3.863	3.879
MAPE	0.064	0.065	0.058

MaxError	9.755	12.452	23.966
RMSE	4.546	4.886	5.956
ExplainedVariance	0.998	0.998	0.997
STDError	3.199	4.022	4.997

4.3 Discussion

The results demonstrate that ETs (ET) achieved superior performance compared to Gradient Boosting, Bagging, LSTM, and a linear regression baseline. These outcomes are consistent with prior studies reporting the effectiveness of ensemble tree models for ecological prediction tasks. In particular, ET's strength lies in its ability to handle small, destructively sampled datasets by leveraging aggressive randomization in both feature selection and split thresholds, which reduces overfitting and improves generalization. This characteristic is critical in tropical forest applications where datasets are often limited in size and contain noisy, heterogeneous measurements. Compared with state-of-the-art approaches based on remote sensing and deep learning, ET provides competitive or superior accuracy despite relying on smaller field-based datasets. This suggests that ET is well-suited for scenarios where destructive sampling restricts sample size, and where nonlinear relationships between tree structure, site descriptors, and biomass must be captured reliably.

Nonetheless, some performance gaps were observed, particularly in belowground biomass (BGB) estimation, which is inherently noisier and less directly observable. These gaps may be explained by collinearity among predictors, class imbalance between tree size categories, or sparse distribution of extreme values in the dataset. Future research could address these challenges by enlarging sample sizes, integrating additional remote sensing features, or applying hybrid ensemble–deep learning frameworks.

Overall, this study reinforces the value of ET as a robust, variance-reducing ensemble model for ecological prediction, while highlighting areas where ecological data properties still limit predictive performance

4.4 Computational cost comparison

While Extra Trees (ET) outperformed Gradient Boosting (GB), Bagging, and Long Short-Term Memory (LSTM) in accuracy, it comes with a computational cost. As shown in Table 9, ET required 1.142 seconds and 8.42 MB of memory on average, making it more computationally intensive than GB (0.973 s, 0 MB) and Bagging (0.250 s, 2.84 MB). LSTM (0.446 s, 0 MB) was also faster, but less effective in biomass estimation. These findings highlight that ET offers the best performance, albeit at a higher computational cost, which must be considered when deploying in real-time or field-based forestry applications.

Table 9: Computational cost comparison of models.

Model	Time (s)	Memory (MB)
GB	0.973	0
ET	1.142	8.421
Bagging	0.250	2.835
LSTM	0.446	0

5 Conclusion

Indeed, according to the evidence provided, the ET method has yielded promising results regarding biomass estimation in tropical forests compared with other Machine Learning models: GB, Bagging, and LSTM networks. Particularly, the ET got amazing training VAF results of 99. 903 for BGB, 99. 879 for AGB, and 99. 889 for TB. Its R-squared values 0.999 were for as a result, it was the best-fit model with a Mean Absolute Error of 3. 664 for ET, 0. 998 for BGB and 0. 998 for TB. The Root Mean Square Errors of 4. 282, 4. 129, and 4. 602 for the related biomass are given in respectively. Respectively values for BGB, AGB, and TB are 3. 361 and 3. 738. Because the ET model precisely and consistently estimates biomass, as demonstrated by the emphasis of these findings is its capacity as a tool for measuring biomass; a device that could be used to enhance carbon accounting and forest management techniques.

To provide more accurate forecasts, future studies should focus on combining ETs with other sophisticated Machine Learning prediction methods and their hybrid models. Increasing the dataset more would let coverage of a far wider spectrum of forest kinds and environmental conditions, therefore improving the generalizing ability and stability of the model. ET's coupling with remote sensing methods and high-resolution photographs might improve the accuracy of forest biomass distribution and changes in structure. Additionally, creating a real-time ET-based monitoring system would make dynamic woodland management easier so that to preserve the Adopt more Efficient ways to slow down climate change.

Future research should explore the integration of ET with high-resolution remote sensing data, such as satellite imagery (Landsat, Sentinel) or UAV-based LiDAR, to improve biomass prediction accuracy across larger forested regions. By incorporating these technologies, ET can provide spatially explicit, real-time biomass estimates, which could support dynamic forest management systems and help track carbon stock changes over time.

References

[1] W. Chen, X. Li, and L. Wang, *Remote Sensing Intelligent Interpretation for Mine Geological Environment: From Land Use and Land Cover Perspective*. Springer Nature, 2022.

[2] D. B. Gemeda, "mplementation of Machine Learning Models for Transmission Grid Monitoring and Blackout Prevention," *Dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT)*, 2024.

[3] B. Huy, N. Q. Truong, K. P. Poudel, H. Temesgen, and N. Q. Khiem, "Multi-output deep learning models for enhanced reliability of simultaneous tree above-and below-ground biomass predictions in tropical forests of Vietnam," *Comput Electron Agric*, vol. 222, p. 109080, 2024. https://doi.org/10.1016/j.compag.2024.109080

[4] V. N. Manawat, "Technical Efficiency of Wildfire Detection and Machine Learning Predictions in Alberta, Canada," 2024. https://doi.org/10.7939/r3-y6bq-8a23

[5] M. F. Oliveira *et al.*, "Predicting below and above-ground peanut biomass and maturity using multi-target regression," *Comput Electron Agric*, vol. 218, p. 108647, 2024. https://doi.org/10.1016/j.compag.2024.108647

[6] J. Saini, M. Dutta, G. Marques, J. Saini, M. Dutta, and G. Marques, "Predicting indoor air quality: integrating IoT with artificial intelligence," *Internet of Things for Indoor Air Quality Monitoring*, pp. 51–67, 2021. https://doi.org/10.1007/978-3-030-82216-3_4

[7] M. V. Sebt, Y. Sadati-Keneti, M. Rahbari, Z. Gholipour, and H. Mehri, "Regression Method in Data Mining: A Systematic Literature Review," *Archives of Computational Methods in Engineering*, pp. 1–20, 2024. https://doi.org/10.1007/s11831-024-10088-5

[8] C. Silva-Perez, A. Marino, and I. Cameron, "Monitoring Agricultural Fields Using Sentinel-1 and Temperature Data in Peru: Case Study of Asparagus (Asparagus officinalis L.)," *Remote Sens (Basel)*, vol. 12, no. 12, p. 1993, 2020.

[9] C. Xu, B. Wang, and J. Chen, "Forest carbon sink in China: Linked drivers and long short-term memory network-based prediction," *J Clean Prod*, vol. 359, p. 132085, 2022. https://doi.org/10.1016/j.jclepro.2022.132085

[10] Y. Li, M. Li, C. Li, and Z. Liu, "Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms," *Sci Rep*, vol. 10, no. 1, p. 9952, 2020. https://doi.org/10.1038/s41598-020-67024-3

[11] P. Sharma, L. Leigh, J. Chang, M. Maimaitijiang, and M. Caffé, "Above-ground biomass estimation in oats using UAV remote sensing and machine learning," *Sensors*, vol. 22, no. 2, p. 601, 2022.

[12] Y. Zhang, J. Ma, S. Liang, X. Li, and M. Li, "An evaluation of eight machine learning regression algorithms for forest aboveground biomass estimation from multiple satellite data products," *Remote Sens (Basel)*, vol. 12, no. 24, p. 4015, 2020.

[13] J. Ma, W. Zhang, Y. Ji, J. Huang, G. Huang, and L. Wang, "Total and component forest aboveground biomass inversion via LiDAR-derived features and machine learning algorithms," *Front Plant Sci*, vol. 14, p. 1258521, 2023. https://doi.org/10.3389/fpls.2023.1258521

[14] G. Li, C. Li, G. Jia, Z. Han, Y. Huang, and W. Hu, "Estimating the Vertical Distribution of Biomass in Subtropical Tree Species Using an Integrated Random Forest and Least Squares Machine Learning Mode," *Forests*, vol. 15, no. 6, p. 992, 2024.

[15] M. J. Diamantopoulou, A. Çömez, R. Özçelik, and S. T. Güner, "Exploring machine learning modeling approaches for biomass and carbon dioxide weight estimation in Lebanon cedar trees," *iForest-Biogeosciences and Forestry*, vol. 17, no. 1, p. 19, 2024. https://doi.org/10.3832/ifor4328-016

[16] K. F. F. Canteral *et al.*, "Machine learning for prediction of soil CO2 emission in tropical forests in the Brazilian Cerrado," *Environmental Science and Pollution Research*, vol. 30, no. 21, pp. 61052–61071, 2023. https://doi.org/10.1007/s11356-023-26824-6

[17] S. E. Sesnie, C. I. Espinosa, A. K. Jara-Guerrero, and M. F. Tapia-Armijos, "Ensemble machine learning for mapping tree species alpha-diversity using multi-source satellite data in an Ecuadorian seasonally dry forest," *Remote Sens (Basel)*, vol. 15, no. 3, p. 583, 2023.

[18] Y. Li, M. Li, C. Li, and Z. Liu, "Forest aboveground biomass estimation using Landsat 8 and sentinel-1a data with machine learning algorithms.," 2020. https://doi.org/10.1038/s41598-020-67024-3

[19] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Front Neurorobot*, vol. 7, p. 21, 2013. https://doi.org/10.3389/fnbot.2013.00021

[20] K. Ramalingam *et al.*, "Light gradient boosting-based prediction of quality of life among oral cancer-treated patients," *BMC Oral Health*, vol. 24, no. 1, p. 349, 2024. https://doi.org/10.1186/s12903-024-04050-x