Transactions on Engineering and Computing Sciences - Vol. 12, No. 06
Publication Date: December 25, 2024
DOI:10.14738/tmlai.1206.17881.
Quartey, G. A., Dadzie, P. K., Asante-Okyere, S., & Eshun, J. F. (2024). Estimating Young Modulus of Elasticity of *Terminalia catappa*: A Machine Learning Approach. Transactions on Engineering and Computing Sciences, 12(06). 21-28.

Estimating Young Modulus of Elasticity of *Terminalia catappa*: A Machine Learning Approach

Gladys Ama Quartey
ORCID: 0000-0002-7073-6342
Department of Interior Design and Technology
Faculty of Built and Natural Environment,
Takoradi technical University. P. O. Box 256, Takoradi

Peter Kessels Dadzie
Department of Interior Design and Materials Technology
Kumasi technical University P. O. Box 854, Kumasi

Solomon Asante-Okyere
Department of Petroleum and Natural Gas Engineering
University of Mines and Technology. P. O. Box 237 Tarkwa

John Frank Eshun
Department of Interior Design and Technology
Takoradi Technical University. P. O. Box 256 Takoradi

ABSTRACT

The purpose of this research was to evaluate the potential of Magnetic Resonance Spectroscopy (MRS) in estimating Young's modulus of elasticity of wood species. To do so, *Terminalia catappa*, a wood species of common occurrence was chosen and its mechanical properties such as bending strength, compression parallel to the grain, and shear parallel to the grain properties were determined using testing methods for small and clear specimens of wood with the British (BS 373, 1957) and American Society of Testing Materials' specifications (ASTM D143, 1983s. The results showed that at 18% moisture content the wood has a density of 520 kg/m^3 with a mean modulus of rupture of 86.04 Mpa, compressive strength parallel to the grain of 42.02 Mpa, modulus of elasticity of 10,500 Mpa, and shear strength parallel to the grain of 16.42 N/mm^2. This dataset was used on machine learning approaches such as decision tree and random forest. The estimated value of Young's modulus using the machine learning models varies between 1000 to 13000 MPa. The obtained results indicated that the use of Magnetic Resonance Spectroscopy (MRS) is an efficient tool for predicting Wood-Young's modulus. This research paves the way for further investigations on the application of MRS and machine learning for predicting a wider range of wood properties. By employing machine learning techniques such as decision trees and random forests, researchers can develop robust models for estimating Young's modulus in other wood species. This approach allows for leveraging large datasets that encompass various influencing factors, ultimately leading to more accurate predictions compared to traditional methods.

Keywords: *Terminalia catappa*, Magnetic Resonance Spectroscopy (MRS), Decision Tree, Random Forest, Young's modulus, Mean Absolute Percentage Error (MAPE)

INTRODUCTION

Terminalia catappa, a tropical hardwood renowned for its durability and aesthetic appeal, finds diverse applications in construction, furniture making, and even traditional medicine. However, a deeper understanding of its mechanical properties, particularly Young's modulus of elasticity, remains crucial for optimizing its usage. Traditional methods for determining Young's modulus, using the testing of small clear samples, are time-consuming, destructive, and often limited by sample availability.

This research proposes a novel approach to estimate Young's modulus of *Terminalia catappa* wood using machine learning. By leveraging a dataset of wood characteristics like density, moisture content, and grain orientation, coupled with existing experimental data, the aim is to develop an accurate and efficient predictive model. This machine learning model can not only expedite the estimation process but also offer valuable insights into the factors influencing the elastic behavior of this valuable hardwood. This research holds significant implications for the sustainable and optimized utilization of *Terminalia catappa* wood across various industries.

Terminalia catappa is a tree very easily recognised because of its pagoda-like structure [1]. According to [2] *Terminalia catappa* is a tall deciduous and erect tree reaching 15-25 m in height, trunk 1-1.5 m in diameter, and often buttressed at the base. The branches are in whorls of nearly horizontal, slightly ascending, and are spaced 1-2 m apart in tiers, or stories, up the trunk. The pagoda-like habit becomes less noticeable as the branches elongate and droop at the tips [1]. It grows best in moist tropical climates is admirably adapted to sandy and rocky coasts and flourishes on oolitic limestone. It has many uses including medicinal and construction.

This research explores the potential of Magnetic Resonance Spectroscopy (MRS) in estimating Young's modulus of elasticity in wood. The study focuses on *Terminalia catappa*, a commonly occurring wood species, and investigates the correlation between MRS data and traditional mechanical properties data. This review will provide context for the research by examining existing literature on wood mechanical properties and testing methods. Young's modulus is a fundamental material property that represents the stiffness of a material. It is crucial for understanding a wood's structural integrity, its response to stress, and its suitability for different applications. Traditional Testing Methods that is standard methods for determining wood mechanical properties, like those outlined in [3] and [4], rely on destructive tests on small, clear specimens. These methods are accurate but time-consuming and require specialised equipment.

Magnetic Resonance Spectroscopy (MRS) in wood science is used for wood characterization [5]. MRS is a non-destructive technique used to study the chemical composition and structure of materials. In wood science, it has been employed to investigate, identify and quantify different chemical components, such as cellulose, hemicellulose, and lignin. It is also used in structural analysis to examine the molecular arrangement and changes in the wood cell wall. In moisture content to determine the amount of water present in the wood. Additionally, MRS has potential for predicting mechanical Properties. While MRS is established in wood characterisation, its application in predicting mechanical properties remains relatively unexplored. This research seeks to bridge this gap by employing machine learning for wood property prediction using machine learning applications and algorithms, like Decision Trees and Random Forest, which are increasingly utilised to analyse complex datasets and predict material properties. This is so because machine learning offers the potential to reduce the need for destructive testing. Predictive models trained on MRS data can potentially replace traditional tests thereby improve efficiency by helping analyse large data sets and develop predictive models quickly and efficiently. It can also uncover complex relationships by identifying hidden correlations between MRS data and mechanical properties.

The machine learning application, Decision Train refers to the process of training a decision tree model on a dataset. The decision tree learns rules and relationships from the training data, effectively building a branching structure to classify or predict outcomes. Once the decision tree is trained, there is the use of a separate "test set" of data (never seen by the model before) to evaluate its performance. This helps us gauge how well the model generalises to new, unseen data.

Random Forest implies using a random forest model for training. A random forest is an ensemble method that combines multiple decision trees. Each tree is trained on a random subset of the data and features. This helps reduce overfitting and improve generalization. As with decision trees, there is the use of a separate test set to evaluate the random forest model's performance on unseen data.

There is limited previous research on MRS and wood mechanical properties: While some research explores the correlation between MRS and specific wood properties, like density or moisture content, there is a scarcity of studies directly investigating the relationship between MRS data and Young's modulus. This makes this research a novel approach where it pioneers the use of MRS to predict Young's modulus in wood, a critical mechanical property. If successful, this approach could revolutionize wood property evaluation, making it faster, more efficient, and less destructive.

In summary, this research holds significant promise for advancing our understanding of wood properties and for developing innovative, non-destructive methods for predicting their mechanical behavior. The study's findings, demonstrating the potential of MRS in estimating Young's modulus, highlight the need for further investigation into the power of this technique in wood science. This research could ultimately lead to the development of more efficient and sustainable approaches to wood utilization.

The applicability of other machining learning models such as artificial neural networks in wood science has already been evaluated by several authors ([6], [7], [8], [9], [10], [11], which denotes its potentiality.

Estimating the Young's Modulus of wood is crucial for various applications, including structural design, wood processing, and quality control. Traditional methods often involve destructive testing, which is time-consuming and expensive. Machine learning offers a promising alternative for non-destructive estimation of Young's Modulus using easily accessible data like wood properties and image analysis.

MATERIALS AND METHODS

To estimate the Young's modulus using machine learning, a dataset containing the following wood samples from *Terminalia catappa* trees, from different locations and tree portions (base, middle, top) to capture variability [12], mechanical test data on the wood samples, including static bending tests to determine the modulus of elasticity (MOE) which is equivalent to the Young's modulus [12] and predictor variables that could potentially influence the MOE, such as wood density, and chemical composition [13]. The data was preprocessed by cleaning the dataset by handling missing values and outliers. It is then normalised or standardised to ensure all features contribute equally to the model.

For the model development, decision tree regression and random forest regression were used. The decision trees were used to model the relationship between input features (e.g., density, moisture) and Young's modulus. This is because decision trees can capture non-linear relationships effectively. A random forest model was implemented to improve prediction accuracy by averaging multiple decision trees. This method helps mitigate overfitting commonly seen in single decision trees by introducing randomness in feature selection. For model training and validation, the dataset was split into training and testing sets (e.g., 80% training, 20% testing). both models were trained (decision tree and random forest) using the training set. The model performance was validated using metrics such as Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and R-squared values on the testing set. After training, the feature importance scores provided by the random forest model will be analysed to understand which factors most significantly influence Young's modulus in *Terminalia catappa.* The trained models were used to predict Young's modulus for new samples based on their features and results interpreted to establish correlations between mechanical properties and environmental factors. Statistical Analysis was conducted to validate the significance of findings, ensuring that predictions align with empirical data obtained from physical testing methods.

RESULTS AND DISCUSSION

Table 1: Young modulus of *Terminalia catappa* using MARS, Decision Tree, and Random Forest

Measured	MARS	Decision Tree	Random Forest
8030	7930.548	9604	9378.609
4862	4635.688	4379.75	4671.83
5376	5096.847	4379.75	4636.803
4744	4395.911	4379.75	4304.634
9986	9051.212	9604	9914.009
4921	5081.319	4379.75	4629.001
12850	12584.7	12290	11173.91
962.8	980.5729	1034.838	1087.168
4778	4523.367	4379.75	4379.538

12680	12743.02	12290	12167.9
4876	4434.73	4379.75	4450.768
3908	4213.43	4379.75	6763.994
4554	4445.764	4379.75	4364.119
12990	11135.99	10745.71	10716.51
4816	4750.799	4379.75	4631.501
1080	1252.742	1034.838	1037.356
1002	1151.661	1034.838	1021.702

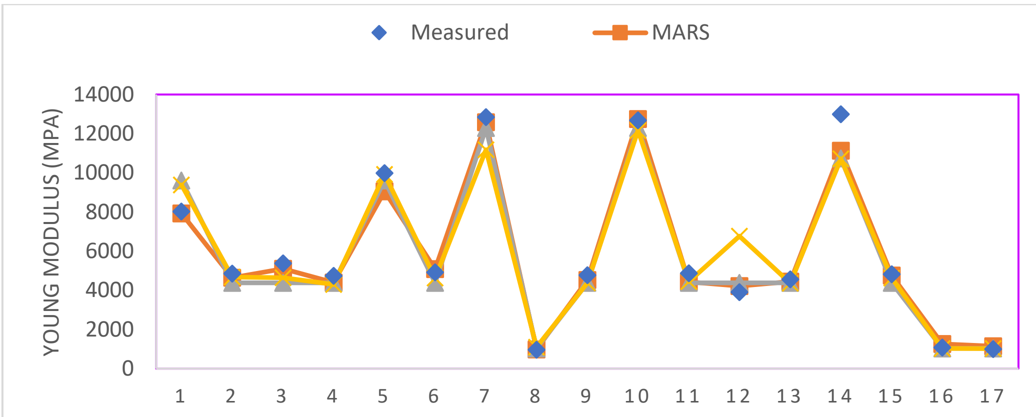

Figure 1: Graph of the measured against the machine learning

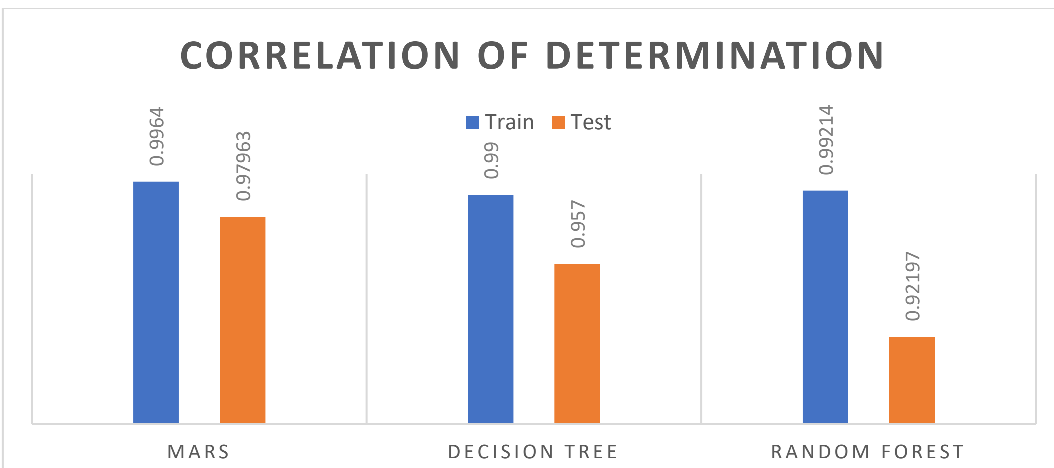

Figure 2: Correlation of determination between MARS, Decision Tree and Random Forest

Table 2: MARS, Decision Tree and Random Forest models compared

MAPE	Train	Test
MARS	0.06609	0.06269
Decision Tree	0.07666	0.0904
Random Forest	0.07646	0.11878

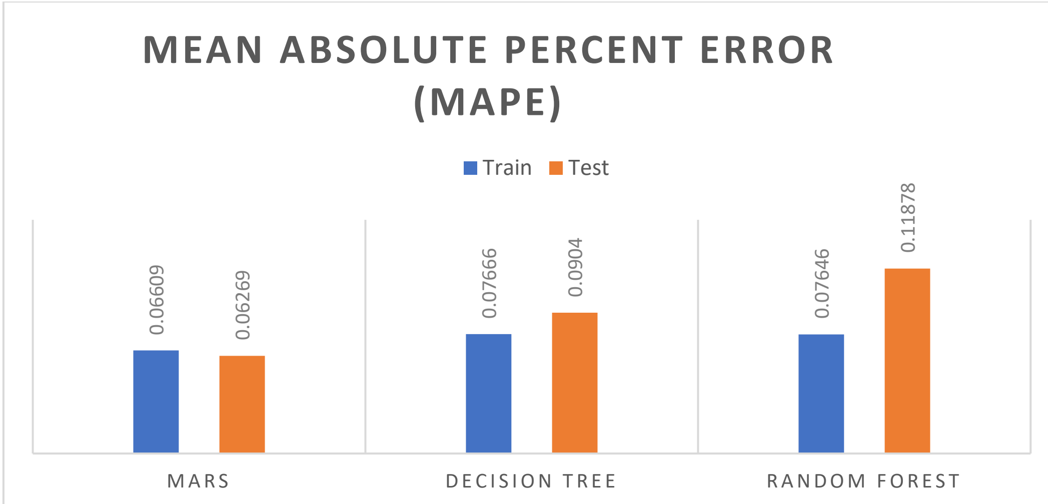

Figure 3: Mean Absolute Percent Error for MARS, Decision Tree and Random Forest

Table 1 shows the Mean Absolute Percentage Error (MAPE) for different machine learning models on training and testing data. MAPE is a metric that measures the average percentage error of the model's predictions. A lower MAPE means your model is more accurate. The closer the MAPE is to 0, the better.

Analysing the results:

- MARS (Multivariate Adaptive Regression Splines):
    - The MARS model performed consistently well on both training (0.06609) and testing (0.06269) data.
    - The slight improvement in the test set suggests that the model is generalizing well to unseen data.
- Decision Tree:
    - The decision tree model showed a noticeable increase in MAPE on the test set (0.0904) compared to the training set (0.07666). This indicates potential overfitting: the model might be too closely tuned to the training data and struggling to predict accurately on new data.
- Random Forest:
    - The random forest model exhibited the largest difference in MAPE between training (0.07646) and testing (0.11878). This signifies significant overfitting. The model appears to be performing much worse on unseen data, which is a major concern.

MARS seems to be the best-performing model based on its low MAPE and consistency between training and testing. Decision Tree and Random Forest models are likely overfitting the training data, leading to poor performance on the test set.

Figure 1 shows the Young's modulus of a material measured by different methods. The measured values are shown in blue diamonds, the MARS values in orange squares, and a third set of values in gray. The values vary between 0 and 14000 MPa. The overall trend is a decline in Young's modulus with an increasing sample number. It's worth noting that there is a high degree of variability between the methods at each sample number. It would be interesting to know what the gray values represent. Figure 2 shows the correlation of determination for three different machine learning models: MARS, Decision Tree, and Random Forest. The correlation of determination is a measure of how well the model fits the data. A higher correlation of determination indicates a better fit. Figure 2 shows that the MARS model has the highest correlation of determination for both the training and test data. This indicates that the MARS model is the best fit for the data. The Decision Tree model has a lower correlation of determination, but it is still a good fit for the data. The Random Forest model has the lowest correlation of determination.

Table 2 suggests that MARS is the best model to use for this dataset, followed by Decision Tree and Random Forest. Figure 3 shows the Mean Absolute Percent Error (MAPE) for three different machine learning models: MARS, Decision Tree, and Random Forest. The chart compares the MAPE of each model on both the training and test data sets. The MAPE is a measure of the accuracy of a model's predictions. A lower MAPE indicates a more accurate model.

CONCLUSION

MARS has the lowest MAPE on both the training and test data sets. This suggests that MARS is the most accurate model overall. Decision Tree has a higher MAPE than MARS on both data sets. This suggests that the Decision Tree is less accurate than MARS. Random Forest has the highest MAPE on both data sets. This suggests that Random Forest is the least accurate model of the three. Overall, the chart suggests that MARS is the most accurate model for this dataset, followed by Decision Tree and then Random Forest. However, It is important to note that these results may not be generalisable to other datasets. This research paves the way for further investigations on the application of MRS and machine learning for predicting a wider range of wood properties. By employing machine learning techniques such as decision trees and random forests, researchers can develop robust models for estimating Young's modulus in *Terminalia catappa* and other wood species. This approach allows for leveraging large datasets that encompass various influencing factors, ultimately leading to more accurate predictions compared to traditional methods alone.

References

[1] Thomson, L.A.J., and B. Evans. 2006.Terminalia catappa (tropical almond), ver. 2.2. In: Elevitch, C.R. (ed.). Species Profiles for Pacific Island Agroforestry. Permanent Agriculture Resources (PAR), Hōlualoa, Hawai'i.

[2] Orwa C, A Mutua, Kindt R, Jamnadass R, S Anthony. 2009 Agroforestree Database: a tree reference and selection guide version 4.0 (http://www.worldagroforestry.org/sites/treedbs/treedatabases.asp)

[3] British Standard BS 373:1957, Methods of Testing Small Clear Specimens of Timber. British Standards Institution

[4] American Society for Testing Materiaals (1983). Standard Method of Testing Small Clear Specimens of Timber. ASTM D-143-52, Philadelphia, PA

[5] Balcom J. B. and Zhung M. 2022 Magnetic Resonance Studies of Water in Wood Materials Book Editor(s):Sabina Haber-Pohlmeier, Bernhard Blümich, Luisa Ciobanu. Wiley Online Library https://doi.org/10.1002/9783527827244.ch15

[6] Tiryaki, S.; Hamzacebi, C. 2014. Predicting modulus of rupture (MOR) and modulus of elasticity (MOE) of heat-treated woods by artificial neural networks. Measurement 49(3): 266–274.

[7] Tiryaki, S.; Aydın, A. 2014. An artificial neural network model for predicting compression strength of heat-treated woods and comparison with a multiple linear regression model. Construction and Building Materials 62(13): 102-108.

[8] Okan, O.T.; Deniz, I.; Tiryaki, S. 2015. Application of artificial neural networks for predicting tensile index and brightness in bleaching pulp. Maderas-Cienc Tecnol 17(3): 571-584.

[9] Melo, R.R.; Miguel, E.P. 2016. Use of artificial neural networks in predicting particleboard quality parameters. Revista Árvore 40(5): 949-958. Haykin, S. 2001. Redes neurais: princípios e prática. Porto Alegre: Bookman, 900p.

[10] Tiryaki, S.; Bardak, S.; Aydın, A.; Nemli, G. 2016.Analysıs of volumetrıc swellıng and shrınkage of heat treated woods: experımental and artıfıcıal neural network modelıng approach. Maderas-Cienc Tecnol 18(3): 477–492.

[11] Bardak S.; Tiryaki S.; Nemli, G.; Aydın A. 2016. Investigation and neural network prediction of wood bonding quality based on pressing conditions. International Journal of Adhesion and Adhesives 68(5): 115–123.

[12] Quartey, G. (2022) Mechanical Properties of Terminalia catappa from Ghana. Materials Sciences and Applications, 13, 334-341. doi: 10.4236/msa.2022.135018.

[13] Quartey, G. A., Eshun, J. F., & Marfo, E. D. (2022). Proximate Analysis of the Fuel Energy Potential of Guarea Cedrata and Terminalia Catappa. European Journal of Applied Sciences, 10(4). 848-858.

Introduction to Evaluating Regression Models
https://www.sciencedirect.com/science/article/abs/pii/S1364032115013258
Coursera https://www.coursera.org/articles/decisiontreemachinelearning