

REVIEWER'S REPORT

Manuscript No.: **IJAR- 59378**

Title: Predicting African timber strength from density: accuracy and robustness of simple models.

Recommendation:

Accept as it is

Accept after Minor revision.....

Accept after Major revision

Do not accept (*Reasons below*)

Rating	Excel.	Good	Fair	Poor
Originality		✓		
Techn. Quality		✓		
Clarity			✓	
Significance			✓	

Reviewer Name: **Dr. ANAPANA GOPAL**

Reviewer's Comment for Publication.

General Comments

The manuscript investigates the relationship between wood density at 12% moisture content and two mechanical properties—bending strength and compression strength parallel to grain—for 13 African timber references obtained from CIRAD/Tropix. The study compares proportional, affine and power models against a mean-only baseline and uses leave-one-out, leave-one-genus-out, influence diagnostics and corpus-deletion sensitivity analyses.

The manuscript has several strengths. The authors are appropriately cautious about the distinction between reference-level mean prediction and specimen-level prediction, explicitly acknowledge the limitations of the documentary dataset, and avoid presenting the equations as structural design values. The analytical framework presented in Figure 1 is also useful in clearly distinguishing documentary prediction from future independent validation.

However, the study is based on only 13 documentary observations representing 12 genera, with unknown source sample sizes, unknown within-reference variance and potentially heterogeneous testing protocols. Therefore, the reported validation statistics should be interpreted very cautiously. Several aspects of the data-selection procedure, model validation, statistical uncertainty, biological comparability and practical application require further clarification.

Overall, the manuscript is potentially publishable after **minor revision**, particularly if the authors strengthen the methodological transparency and moderate the interpretation of predictive accuracy.

REVIEWER'S REPORT

Minor Comments

1. The Extremely Small Sample Size Must Be Central to the Statistical Interpretation

The entire analysis is based on only **13 reference means**. This is a very small dataset for comparing three functional models, conducting cross-validation, calculating influence diagnostics and assessing model stability.

Although the manuscript acknowledges this limitation, the Results and Abstract repeatedly emphasize large percentage improvements such as **73.2% and 84.4%**.

With such a small dataset, these percentages should not be interpreted as general predictive accuracy for African timber.

The authors should:

- emphasize the exploratory/case-study nature of the analysis;
- report uncertainty around model coefficients where statistically defensible;
- avoid language implying broad validation;
- distinguish clearly between **within-corpus predictive performance** and external predictive validity.

The manuscript already recognizes this issue in the Discussion, but the qualification should be consistently applied throughout the Abstract, Results, Discussion and Conclusion.

2. The Documentary References Are Not Fully Independent Biological Observations

The authors correctly state that the 13 references represent 12 genera and that two entries belong to **Guibourtia**. More importantly, some commercial references may encompass multiple species.

This creates an important issue: the 13 observations are not necessarily equivalent biological sampling units.

The manuscript should explain more explicitly:

- whether each CIRAD sheet represents a single species or a commercial group;
- whether any commercial designations combine several species;
- whether multiple references originate from the same testing campaign;
- whether any references share provenance or source populations;
- whether the strength data were generated under comparable standards.

The current leave-one-genus-out analysis does not completely solve this problem because **commercial group identity, species identity, genus identity and testing-source identity are different levels of dependence**.

3. Selection of the 13 Timbers Needs Greater Reproducibility

The authors state that the selection was based on the intersection between the African timber scope of the Wood Handbook and CIRAD/Tropix sheets containing all required properties.

This is reasonable, but the manuscript should provide a clearer selection flow.

A table or supplementary flow diagram should show:

1. Total African timber references initially considered.
2. References identified from the Wood Handbook.
3. References available in Tropix.
4. References excluded because one or more required properties were unavailable.
5. Final 13 references.

REVIEWER'S REPORT

6. Exact reasons for every exclusion.

This would make the documentary corpus genuinely auditable.

4. The Source Data Should Be Provided in a Reproducible Supplement

The manuscript states that exact sheet URLs, botanical names and content dates are supplied with the data.

However, the main manuscript should either include this information in a supplementary dataset or provide a clear data repository.

At minimum, the supplementary file should contain:

- timber reference name;
- accepted botanical name;
- commercial designation;
- genus;
- density source;
- bending strength source;
- compression strength source;
- moisture condition;
- source URL;
- source date/version;
- any available testing standard;
- sample size if available.

A machine-readable dataset would substantially improve reproducibility.

5. Testing Protocols Across CIRAD References May Not Be Comparable

The manuscript itself notes that a common 12% moisture reference does **not establish identical testing protocols**.

This is a minor issue because differences in:

- specimen dimensions;
- loading configuration;
- grain orientation;
- testing standard;
- conditioning procedure;
- provenance;
- defect criteria;
- failure definition

could contribute to differences in reported strength independently of density.

The authors should provide as much information as available about the testing methodologies behind the 13 source values.

If these data are unavailable, this should be presented as a minor source of **between-reference heterogeneity**, rather than merely a minor limitation.

6. The Conversion of CIRAD Density Values Requires Stronger Documentation

The manuscript converts the dimensionless density ρ_{12} into mass density using:

$$\rho_{12} = 1000d_{12} \text{ kg/m}^3$$

and explicitly distinguishes this from basic density.

This conversion needs a more explicit explanation of the exact density convention used by the source database.

REVIEWER'S REPORT

The authors should provide:

- the original definition of (d_{12}) ;
- the source reference for the conversion;
- confirmation that the reported values correspond to apparent density at 12% moisture;
- an example calculation using one of the source values.

This is important because the entire prediction framework depends on the consistency of the density definition.

7. The Model Selection Procedure Needs More Statistical Justification

The study compares:

- M0: mean-only;
- M1: proportional;
- M2: affine;
- M3: power model.

The authors appropriately avoid declaring a universally superior functional form. Nevertheless, the manuscript would benefit from a clearer statistical rationale for choosing these three models.

Please explain why these particular functional forms were selected rather than, for example:

- logarithmic;
- quadratic;
- segmented;
- robust regression;
- allometric alternatives.

More importantly, the authors should explain that the purpose is **comparison of simple candidate relationships**, rather than formal model selection from a large candidate set.

8. Cross-Validation With 13 Observations Does Not Establish External Predictive Accuracy

The use of LOO and LOGO is methodologically thoughtful. However, both procedures remain internal to the same documentary compilation. The manuscript acknowledges this appropriately.

The terminology should therefore consistently use:

internal cross-validation

rather than simply “validation” when referring to the main performance estimates.

The Abstract currently says “internal leave-one-out and leave-one-genus-out validation,” which is relatively careful, but the Results and Conclusion should maintain the same distinction.

The strongest statement supported by the data is:

Density predicts the observed variation among the selected documentary reference means within this corpus.

It does not establish performance for:

- individual trees;
- individual boards;
- individual logs;
- commercial lots;

REVIEWER'S REPORT

- specific countries;
- Guinean timber;
- independent African timber datasets.

The manuscript generally acknowledges this, but the wording should be systematically checked.

9. LOGO Provides Limited Additional Protection Against Dependence

The authors state that LOGO removes all observations belonging to a genus, but only **Guibourtia has two references**.

Consequently, LOGO is only marginally different from LOO for most folds.

The manuscript should explicitly state that:

The genus-level grouping structure is weak because only one genus has more than one reference.

Therefore, the reported LOGO performance should not be interpreted as strong evidence of generalization to unseen genera.

This is especially important because the Abstract highlights the LOGO RMSE values as key performance metrics.

10. Confidence Intervals or Other Measures of Parameter Uncertainty Should Be Considered

The study reports remarkably stable proportional coefficients, with CV values of 0.93% and 0.53% across deletion scenarios.

However, these CVs describe **deterministic corpus-deletion sensitivity**, not statistical uncertainty.

The manuscript correctly states this, but readers may nevertheless interpret “0.93%” and “0.53%” as evidence of very high coefficient precision.

The authors should consider providing appropriate uncertainty estimates for the coefficients, or explicitly state in the table and Results that:

“The reported CVs are sensitivity measures and should not be interpreted as confidence intervals or standard errors.”

This distinction is particularly important because the manuscript itself notes that the deletion scenarios overlap.

11. The Influence of Azobé Requires More Careful Interpretation

The manuscript identifies azobé as influential, with a Cook's distance of **1.074 for compression under M1** and 1.033 under M2.

The authors conclude that its influence does not invalidate the observation.

That is reasonable, but the analysis should go further.

Because azobé represents the highest density and highest strength observation, it has substantial leverage on the fitted relationship. The authors should report:

- correlation with and without azobé;
- slope change;
- prediction error change;
- model fit change;
- whether the relationship remains approximately linear/proportional after removal.

REVIEWER'S REPORT

The existing deletion analysis helps, but the discussion could more clearly demonstrate how much of the apparent density–strength relationship is driven by the density extremes.

12. The Use of a Zero-Intercept Model Requires Additional Justification

M1 assumes:

$$f = c \rho$$

with a zero intercept.

The authors correctly state that the zero intercept is an empirical constraint within the observed density range and should not be extrapolated to zero density.

Nevertheless, the manuscript should explain why this model is retained as a principal illustrative model despite M2 and M3 having greater flexibility.

The argument should be based on:

- parsimony;
- comparable predictive performance;
- interpretability;
- stability;
- avoidance of unnecessary parameters.

This would strengthen the justification for using M1 in the practical example.

13. The Reported Percentage Improvement Should Be Presented Alongside Absolute Error

The manuscript states that M1 reduced LOGO RMSE by 73.2% for bending and 84.4% for compression.

These percentages are impressive but can be misleading because they are calculated relative to a mean-only baseline.

The manuscript should emphasize the actual errors:

- **8.96 MPa bending**
- **2.97 MPa compression**

and perhaps report normalized errors relative to the mean response or range.

The absolute RMSE should remain the primary measure, with percentage reduction presented as a secondary descriptive statistic.

14. MAPE Has Limitations That Should Be Acknowledged

MAPE values of 8.00% for bending and 4.74% for compression are reported.

Because all observed strengths are positive, MAPE is calculable, but it can still disproportionately emphasize relative error at lower-strength observations.

The authors should either:

- provide an additional scale-free metric such as normalized RMSE; or
- briefly explain why MAPE is appropriate for this dataset.

This would provide a more balanced assessment of prediction performance.

15. The Manuscript Should Distinguish Model Error From Source Measurement Error

The reported RMSE measures differences between predicted and **published reference means**.

It does not measure:

- laboratory measurement error;

REVIEWER'S REPORT

- specimen-to-specimen variation;
- provenance variation;
- species-level variation;
- uncertainty in published mean values.

The authors recognize this distinction in Section 4.2.

This is one of the strongest conceptual aspects of the manuscript, but it should be emphasized earlier—particularly in the Abstract and Results.

A reader should immediately understand that:

8.96 MPa is documentary-reference prediction error, not timber strength measurement uncertainty.

16. The Worked Example Could Be Misinterpreted as a Practical Prediction

The worked example calculates a density of 700 kg/m³ and predicts:

- 100.6 MPa bending;
- 59.8 MPa compression.

The manuscript clearly states that this is a constructed example, not an experiment, which is commendable.

Nevertheless, the example could still be misinterpreted by non-specialist readers as a practical strength prediction for a timber piece.

Consider changing the wording to:

“illustrative reference-mean estimate”

throughout this section and visually highlighting that the values are **not characteristic or design strengths**.

17. The Structural Design Section Is Useful but Somewhat Disconnected

The discussion of Eurocode characteristic and design strength is technically relevant because it prevents inappropriate use of the equations.

However, the section is relatively detailed compared with the empirical dataset.

The authors could shorten this section and focus on the central point:

predicted reference means cannot be substituted for characteristic or design strength.

If retained, the applicability of the cited Eurocode edition and the jurisdictional context should be made explicit.

18. The Link to Guinea Needs More Empirical Support

The manuscript identifies **Lophira alata**, **Triplochiton scleroxylon** and **Terminalia superba** as taxa documented in Guinea and proposes local testing based on the documentary relationships.

This is a reasonable research recommendation.

However, the manuscript should avoid implying that the current equations have already been demonstrated to work for Guinean timber.

The correct interpretation is:

the selected references provide hypotheses and candidate density ranges for future Guinean testing.

The authors already acknowledge this distinction and should preserve it consistently.

19. The Study Needs a Clearer Data-Quality Hierarchy

The manuscript combines information from:

- CIRAD/Tropix;

International Journal of Advanced Research

Publisher's Name: Jana Publication and Research LLP

www.journalijar.com

REVIEWER'S REPORT

- Wood Handbook;
- published scientific literature;
- standards.

The numerical model, however, uses CIRAD values only. This should be made visually clearer.

I recommend a table identifying:

Information	Source	Used in model?
Density	CIRAD/Tropix	Yes
Bending strength	CIRAD/Tropix	Yes
Compression strength	CIRAD/Tropix	Yes
Wood Handbook	Selection/reference framework	No
Literature	Background	No
Eurocode	Design-context discussion	No

This would eliminate possible confusion about which sources constitute the actual training dataset.

20. The Manuscript Would Benefit From an Explicit “Limitations” Subsection

Limitations are currently distributed throughout the Discussion. The manuscript would be stronger with a dedicated subsection covering:

1. $n = 13$ reference means;
2. unknown source sample sizes;
3. unknown within-reference variance;
4. possible commercial multi-species groups;
5. unknown testing-protocol differences;
6. lack of independent external validation;
7. weak genus replication;
8. possible source-campaign dependence;
9. inability to estimate specimen-level prediction error;
10. limited geographic transferability.

The manuscript already acknowledges many of these issues, so consolidating them would improve clarity.

Minor Comments

1. The title is appropriate, but “African timber strength” could be interpreted as representing all African timber. Consider **“Selected African timber references”** if the authors want maximum precision.
2. Define p_{12} at its first occurrence.
3. Define $\{f_m\}$ and $\{f_{c,0}\}$ consistently in the Abstract and Methods.
4. Use consistent notation for compression strength: $\{f_{c,0}\}$ throughout.
5. Clarify whether “African mahogany” is a commercial group or a species-level reference.
6. The manuscript states that commercial references, genera and biological species are not interchangeable sampling units; this important point should also be emphasized in the statistical interpretation.
7. Table 1 should preferably include botanical species names, where available.
8. Provide the exact source/version date for each Tropix record in supplementary material.

REVIEWER'S REPORT

9. Explain why the Wood Handbook was used for selection rather than as an independent validation source.
10. Table 2 should clearly indicate that RMSE values are from **internal** validation.
11. Figure 2 is useful, but the $\pm 10\%$ and $\pm 20\%$ bands should be explicitly described as descriptive bands, as currently stated in the caption.
12. Figure 4 should state clearly in the axis/caption that the displayed ranges arise from deterministic deletion scenarios rather than confidence intervals.
13. Figure 5 contains only 13 observations; the text appropriately recognizes that this is insufficient for a reliable variance model.
14. Table 6 should clearly label CV as **descriptive deletion-scenario CV** if CV values are discussed in the text.
15. Table 7 should explain the interpretation of Cook's distance for readers who are not familiar with regression diagnostics.
16. Avoid the phrase "stable" without qualification; use **"stable across the examined deletion scenarios."**
17. Similarly, use **"predictive advantage within this corpus"** rather than simply "predictive advantage."
18. The conclusion is generally appropriately cautious and should be retained in this style.
19. The reference list should be checked carefully for consistency in DOI, URL, volume, issue and page formatting.
20. Ensure all cited references appear in the reference list and all listed references are cited in the text.

Recommended Additional Analyses / Revisions

If the underlying data permit, I recommend adding:

A. Sensitivity to Biological Grouping

Repeat the analysis after:

- excluding the second *Guibourtia* observation;
- treating commercial groups separately;
- reporting results with and without potential multi-species commercial references.

B. Additional Model Diagnostics

Report:

- $\sqrt{R^2}$ as a descriptive measure;
- confidence intervals for slopes where appropriate;
- residual standard error;
- normalized RMSE;
- correlation between observed and predicted values.

These should be clearly labelled as descriptive because of the small sample.

C. Source-Heterogeneity Assessment

Provide a supplementary table documenting available information on:

- test standard;
- specimen dimensions;
- moisture conditioning;
- provenance;
- sample size;

REVIEWER'S REPORT

- reported mean.

D. Independent Validation Plan

The manuscript's proposed next step—paired density and strength measurements on traceable local timber—is appropriate.

This could be described more concretely as:

training dataset → fixed equation → independent regional validation dataset → assessment of prediction error → recalibration only if justified.

Overall Assessment

This is a **well-conceived and unusually transparent exploratory modelling study**. A major strength is that the authors do not simply fit a density–strength regression; they attempt to examine internal predictive performance, sensitivity to corpus composition and influence of extreme observations. The manuscript also makes a valuable distinction between **reference-mean prediction and specimen-level strength estimation**, which substantially improves the scientific integrity of the work. Nevertheless, the evidence base remains very limited: only 13 documentary means, 12 genera, unknown original sample sizes and no genuinely independent validation dataset. Consequently, the manuscript should be positioned explicitly as a **documentary, exploratory modelling study**, not as a validated predictive method for African timber generally.

The central finding—that density contains substantial information about the variation in the selected reference means—is plausible and potentially useful. The practical equations are also clearly presented. However, their application to individual timber pieces or regional timber lots remains an empirical question, as the authors themselves acknowledge.

Final Recommendation

MINOR REVISION

The manuscript has sufficient methodological interest and potential value to justify revision. The authors should particularly address dataset independence, source-data reproducibility, testing-protocol heterogeneity, limitations of the very small sample, interpretation of LOGO validation, coefficient uncertainty, influence of extreme observations, and the distinction between documentary prediction and real-world timber strength before publication.