

Allometric equations and height-diameter models for estimating above-and below-ground biomass of *Colophospermum mopane* J.Léonard. in Mozambique

Sá Nogueira Lisboa

sanogueiralisboa@gmail.com

Eduardo Mondlane University

Severino Macôo

High Polytechnic Institute of Gaza

Almeida Siteo

Eduardo Mondlane University

Article

Keywords:

Posted Date: April 11th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4125867/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on January 28th, 2025.

See the published version at <https://doi.org/10.1038/s41598-025-87297-w>.

Allometric equations and height-diameter models for
estimating above-and below-ground biomass of
Colophospermum mopane J.Léonard. in Mozambique.

Sá Nogueira Lisboa^{1,2*}, Severino Macôo^{3,4}, Almeida Siteo¹

¹Department of Forest Engineering, Faculty of Agronomy and Forest Engineering, Eduardo Mondlane
University, P.O. Box 257, Maputo, Mozambique.

²N'Lab, Nitidae, Agostinho Neto Avenue, Maputo, Mozambique.

³Agriculture Division, High Polytechnic Institute of Gaza, Campus Politécnico de Lionde, P. O. Box 1,
Chokwe, Gaza, Mozambique.

⁴Department of Forest Science, Federal University of Lavras, Brazil.

(*) Corresponding author

Abstract

This study aimed to develop species-specific biomass allometric equations (BAE) for
the estimation of above-ground biomass (AGB) and below-ground biomass (BGB) of
Colophospermum mopane, J.Léonard species in Mozambique and to evaluate the
feasibility of including total height estimated by the height-diameter (H-D) model in
the equations to improve the accuracy of biomass estimation. We applied a destructive
method and felled 120 trees of *C. mopane* in Mabalane and Tambara districts. We
measured breast height diameter, total height (H), and dry weight of each component
(roots, stem, branches, and leaves). We fitted three BAE and fifteen H-D models using
OLS and non-linear least squares regression. A validation procedure checked all
models for applicability. We conducted a forest inventory using 78 temporary clusters,

each consisting of four plots measuring 100 m x 20 m, to estimate the average AGB and BGB of *C. mopane* using the best-fitting BAE and H-D model. We also compared the best BAE of *C. mopane* with existing, locally developed species-specific BAEs. The study found that the model, including the interaction of D^2 and H , performed better than the model with separate D and H variables. However, models including D and H performed similarly to models with D alone. The Power model and HossfeldIV performed best among the models with 2 and 3 parameters, respectively. Integrating the H-D model into the BAE for AGB and estimation can improve the efficiency and reliability of biomass estimation and stand volume by forest inventory data. Overall, this research is timely as there is an urgent need to improve the assessment of biomass and carbon emissions in mopane woodland (MW) in southern Africa, given their spatial extent, and to support climate goals and the sustainable management of these ecosystems.

Introduction

Global land use change is increasing significantly, and recent studies suggest it is four times greater than previously estimated ¹. The land use change has important implications for the global carbon cycle, as the atmospheric concentration of carbon dioxide (CO_2) continues to rise ^{2,3}. Land use and land cover change (LULCC) have been identified as a significant contributor to anthropogenic CO_2 emissions, accounting for approximately 206 ± 57 GtC over the past 268 years ³. Given the crucial role of LULCC in CO_2 emissions, extensive research has been undertaken to quantify the global

carbon budget, focusing on the impacts of human activities on soil and plant biomass in terrestrial ecosystems and address climate change challenges ^{4,5}.

Seasonally dry tropical woodlands play a significant role in climate change mitigation, as recognized by the Paris Agreement under the United Nations Framework Convention on Climate Change (UNFCCC). These woodlands cover 40% of the tropical forest area and contain substantial carbon stocks ^{6,7}. However, they face unprecedented deforestation and widespread forest degradation ⁸. The destruction of these woodlands eliminates important carbon sinks that mitigate climate change and undermines future climate stabilization efforts ^{9,10}. The Paris Agreement encourages voluntary efforts in developing countries to contribute to climate change mitigation by reducing greenhouse gas emissions and enhancing carbon removals from forest activities, known as REDD+ (Reducing Emissions from Deforestation and Forest Degradation, conservation, sustainable management of forests, and enhancement of forest carbon stocks). Developing countries participating in REDD+ must develop forest reference levels (emission levels) based on guidance provided by the Conference of the Parties (COP) 17 of the UNFCCC ¹¹.

In the case of Mozambique, for example, it submitted a proposal for Forest Reference Emission Levels (FREL) to the UNFCCC in 2018 to consolidate its REDD+ process. However, the reporting of forest carbon for tree species and forest categories in their FREL proposal mainly relied on standard functions and parameters from the IPCC and few local generalize and species-specific allometric equations. In the context of REDD+, it is crucial to improve the accuracy of emissions and removal estimates by developing country-specific allometric equations ¹². In this context, countries like

Mozambique would benefit from developing their equations to estimate forest carbon and improve the precision of their emissions and removal assessments. Using country-specific equations would contribute to a more accurate accounting of carbon stocks and emissions, enhancing the effectiveness of climate change mitigation efforts under the REDD+ mechanism.

Allometric regression models are widely used in converting national forest inventory data into stand biomass and forest carbon stock estimates. These models are mathematical functions that relate the dry mass of trees to various tree dimensions, including diameter at breast height, total height, wood density, basal area, and other dendrometry variables ^{13,14}. Wood density in allometric models has been shown to improve the goodness of fit ¹⁵. Total tree height is another valuable variable collected in forest inventories for biomass modeling. While the improvement may be slight, adding tree height as a predictor variable can enhance biomass estimation ¹⁶.

Moreover, tree height information is crucial for computing tree volumes, growth, yield simulations, forest dynamics prediction, and site index ¹⁷. However, directly measuring tree height is often costly, time-consuming, and labor-intensive compared to measuring tree diameter ¹⁸. Many forest inventories use height-diameter (H-D) models to predict tree heights instead of direct measurements ¹⁷. These H-D models can then be combined with biomass allometric equations, where tree height is a predictor variable, to estimate biomass ^{16,19}. By utilizing H-D models and including wood density in allometric equations, forest inventory practitioners can improve the efficiency and accuracy of biomass estimation and facilitate the assessment of forest carbon stocks.

In sub-Saharan Africa, many allometric equations have been developed to estimate carbon stocks. However, more generalized biomass models are available than species-specific models ²⁰. This is also true in Mozambique, a country known for its diverse vegetation ecosystems and high tree species diversity. Many generalized allometric models have been developed for miombo woodlands in Mozambique ^{16,21,22}, while fewer species-specific models are available ²³⁻²⁵. As a result, no allometric equation was developed explicitly for the mopane tree species, which is prevalent in southern Africa. Therefore, the actual carbon stock stored in mopane woodland (MW) remains unknown.

It is argued that species-specific allometric equations are preferred because different tree species can vary significantly in tree architecture and wood density ^{20,26}. Therefore, using generalized biomass models, especially when applying pantropical equations, can lead to overestimating biomass and carbon stocks ^{27,28}. In the context of REDD+ and carbon stock estimation, it is crucial to have species-specific biomass models. These models provide more accurate estimates of carbon stocks and are considered preferable over-generalized models ^{29,30}. Assuming that MW covers a significant area of southern African countries ³¹, developing species-specific allometric equations for mopane would help improve the accuracy of carbon stock estimates in Mozambique and other MW regions.

This study aims to develop accurate allometric equations for estimating below- and above-ground tree biomass (BGB and AGB) of the *C. mopane* species in Mozambique. Specifically, the study aims to achieve the following: (1) Develop allometric equations for stem dry mass, crown dry mass (branches + leaves), root dry

mass, and total dry mass (stem plus crown) of *C. mopane* species; (2) Develop a H-D model for *C. mopane* species; and (3) Estimate AGB and BGB using the selected allometric equations and H-D model, using forest inventory data. The outcomes of this study will provide valuable tools for assessing carbon stocks and biomass in mopane woodlands, ultimately supporting effective climate change mitigation efforts and sustainable forest management practices in the region.

Material and Methods

Site description

The first study area is in Mabalane District, which is in Gaza province in southern Mozambique (S 23° 53', 32° 74' E). The climate in this area is classified as semi-arid, with an average annual rainfall of 505 mm/year and an average annual temperature of 24 °C. The region experiences distinct dry and wet seasons, with most precipitation between October and April ³². The vegetation in the Mabalane District primarily consists of Mopane woodlands (*Colophospermum mopane* (Benth.) J. Léonard), with occasional patches of *Androstachys johnsonii* Prain, *Combretum* spp., and *Boscia albitrunca* (Burch.) Gilg & Gilg-Ben.³². The second study site is in Tambara District, situated in the central region of Mozambique in Manica province (**Figure 1**). The coordinates of the site range from 16°25' to 17°23' S and 33°54' to 34°36' E. Tambara District has a tropical dry climate, with average annual temperatures of 27.6 °C and an average annual rainfall of 850 mm. The dry season in this region typically occurs from March to September. The vegetation in Tambara District is characterized by three major types: mopane woodland, dry deciduous miombo woodland, and shrub

savannah. Baobab and acacia trees are also present in various parts of the district. Therefore, it is expected that these two Mopane woodland sites (Mabalane and Tambara) will exhibit different climatic, soil, and biotic influences on the growth and AGB of *C. mopane*.

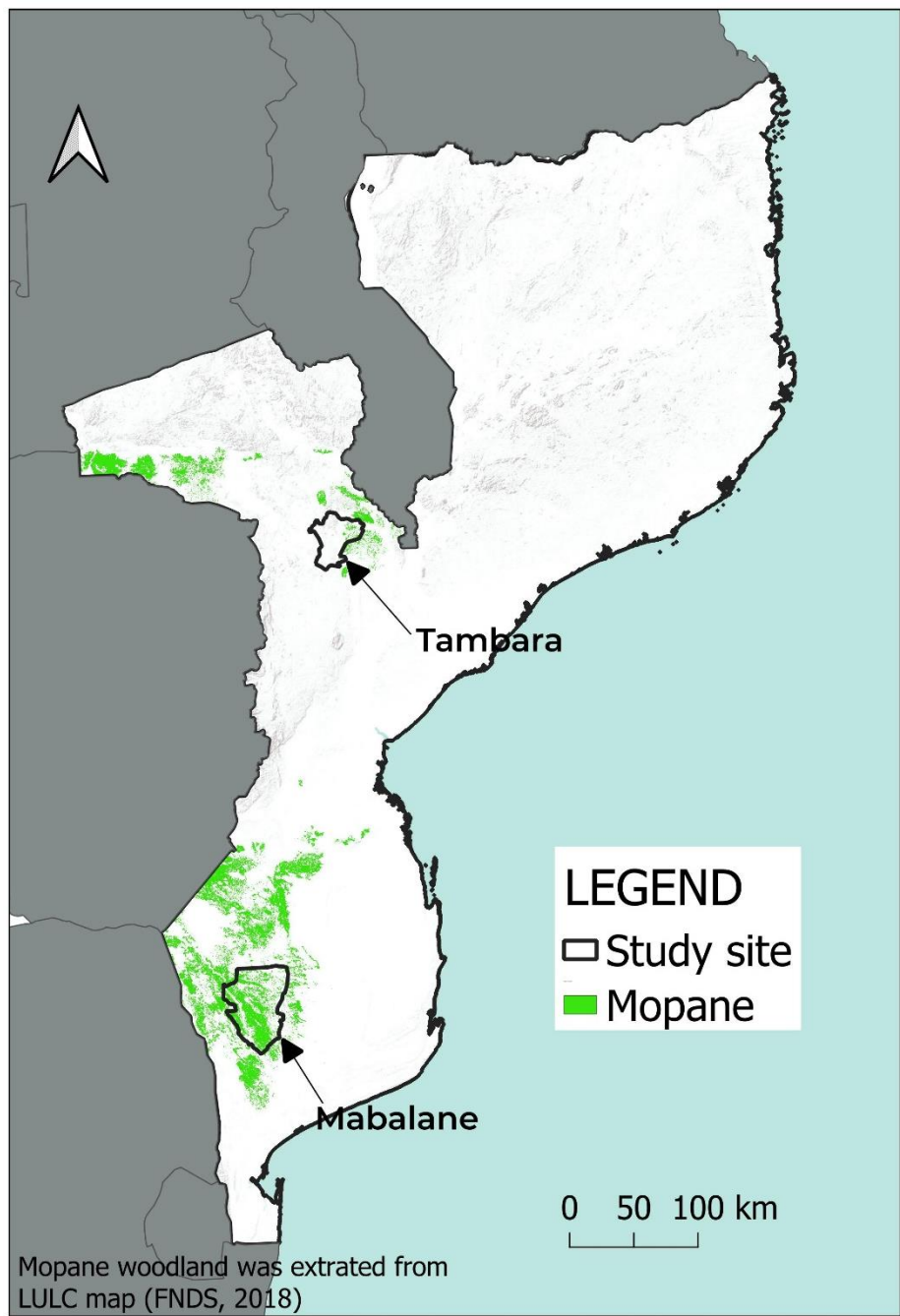


Figure 1: Location of the study areas and the spatial extent of the mopane woodland in Mozambique.

Species description

Colophospermum mopane J. Léonard, commonly known as mopane, is a dominant leguminous tree or multi-stemmed shrub belonging to the Caesalpinioideae subfamily. One notable characteristic of *C. mopane* is its ability to form ectomycorrhizal associations, potentially providing the species with an adaptive advantage to better access nutrients, water, and other resources ^{33,34}. It is native to hot, low-lying regions of subtropical southern Africa, occurring between latitudes 9°S and 25°S ^{35,36}. Mopane woodlands (MW) cover approximately 30-35% of the savannahs in southern Africa, which amounts to about 1.5 million km² of land, representing over a quarter of the region's total land area ^{31,36}. As a result, MW is characterized by low compositional and structural diversity ³⁵. The species composition of the non-mopane species in the MW is similar to that of the miombo woodlands ³⁷. MW is typically associated with regions that experience down to moderate rainfall, high temperatures, low altitudes, and various soil types ³⁶. Its distribution spans from the hot, dry valley bottoms and adjacent plains of southern Angola and northern Namibia, through Botswana and Zimbabwe to central and southern Mozambique, as well as the Luangwa valley in Zambia, central Malawi, and northern South Africa ³¹. The total area covered by MW in southern Africa is estimated to be 555,000 km² ³¹. According to the National Forest Inventory of Mozambique, mopane woodlands can store approximately 44.51 ± 3.86 Mg ha⁻¹ (95% CI) of AGB, with about 85% coming from *C. mopane* ³⁸. The soil organic carbon content in mopane woodlands is estimated to be around 19 ± 10 ($\pm$ SE) Mg ha⁻¹ ³². MW covers approximately 80,435 hectares, accounting for about 2.7% of the total

forest area in Mozambique, making it the second-largest woodland type in the country after the miombo woodland (covering approximately 2.9 million hectares) ³⁷.

Forest inventory and destructive sampling

We conducted a conventional forest inventory in the Tambara and Mabalane districts. We established 78 clusters, each consisting of four plots measuring 100 m x 20 m (0.2 hectares). Within each plot, we measured the D at 1.3 m of all trees and shrubs with a D of 5 cm or greater. To obtain data for developing the biomass allometric equations (BAE) and H-D model within the landscape, we randomly selected 120 individuals of *C. mopane* from different D classes. We then employed the destructive sampling methodology following the procedure described in Picard et al. ³⁹. The D of the selected trees was measured at 1.3 m using a diameter tape, while the total tree length (tree height) was measured on the fallen tree using a conventional measuring tape. Each sampled tree was then divided into three components: the stem as the merchantable section, branches with a minimum diameter of 3 cm, and twigs and leaves. The fresh weight of each component was immediately determined in the field using a mechanical weighing scale with a maximum capacity of 200 kg and a precision of 100 g. Representative sub-samples weighing between 200-3000 g were taken from each component, and their fresh weight was measured in the field using a digital weighing scale with a maximum capacity of 3 kg and a precision of 0.5 g. These sub-samples were later sent to the laboratory at Eduardo Mondlane University for further analysis following Lisboa et al. ¹⁶.

Excavation and root removals

Excavation and root removals were conducted to obtain information on the BGB and root characteristics. The process involved exposing the primary roots and proximal parts of each lateral root in the first soil horizons using tools such as a pickaxe, a daba (a type of hoe with a short handle), and a machete. The roots were excavated from their insertion zone towards their distal ends. Fine roots with a diameter of less than 2 mm were not sampled during the excavation, as defined by Jackson et al.⁴⁰. To remove the lateral roots, the soil around the tree was excavated, and the roots were pulled out using a backhoe. As a result, the AGB part of the tree became unstable. The stump (with the root attached) was pushed and overturned using a backhoe to extract the root system safely. Once the tree was cut and toppled, it was separated into the AGB and BGB portions. The separation was made precisely at the point where the two pieces were connected when the tree was standing. This allowed for the accurate measurement and analysis of the BGB and root characteristics.

Laboratory measurements

In the laboratory, the stem, branches, and foliage sub-samples were dried at 75 °C in an oven until reaching a constant weight. The dry weight of each sub-sample was then determined using a digital weighing scale with a maximum capacity of 3 kg and a precision of 0.5 g. We used the dry-to-fresh weight ratio of the sub-samples for each component to determine the dry weight of each tree component. This ratio was calculated by dividing the dry weight of the sub-sample by its corresponding fresh weight. We multiplied the dry-to-fresh weight ratio by the fresh weight of each tree

component to obtain the dry weight of each component (stem, branches, and foliage). Finally, the dry weights of the tree components were summed together to get the total dry weight of the tree. **Figure 2** shows the workflow of the process of developing the BAEs and H-D models.

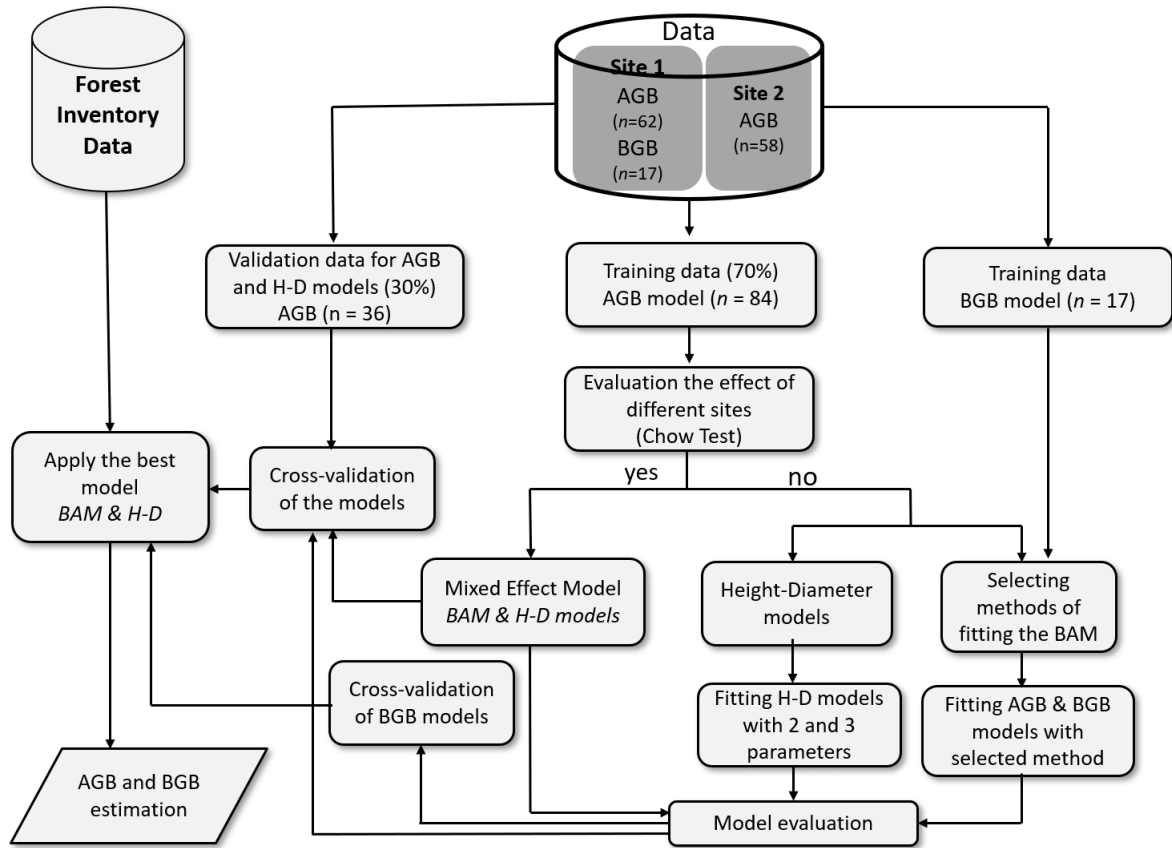


Figure 2: Methodological framework

Biomass allometric equation and height–diameter regression models

We selected three candidate biomass allometric equations from the literature ^{16,41,42}. We chose these equations because they had a small number of parameters, ranging from 1 to 3 and were elected as the best models in those studies. One of the candidate models (Model 1) only used D as a predictor variable, while the other two

models (Models 2 and 3) used both D and total height (H) as predictor variables (**Table 1**). We fitted these models for the stem, crown, total dry weight, and root biomass. We labeled the biomass allometric models for the total AGB as TM1, TM2, and TM3 for models 1, 2, and 3, respectively. We labeled the models SM1, SM2, and SM3 for the stem biomass for models 1, 2, and 3, respectively. The crown biomass models were labeled CM1, CM2, and CM3 for models 1, 2, and 3, respectively. Similarly, the root biomass models were labeled RM1, RM2, and RM3 for models 1, 2, and 3, respectively. Additionally, we have fifteen H-D models in this study, following Mehtatalo and Kansanen ⁴³. All the H-D models were non-linear and consisted of either two (seven models) or three (height models) parameters. **Table 2** overviews the BAE and H-D models we fitted in this study.

Table 1: Alternative models were fitted to predict total AGB, stem biomass, crown biomass, BGB (models 1 to 3), and total height (models 4 to 10) of *C. mopane*'s species in Mozambique.

Model	Equation	Reference
<i>Biomass Allometric Model for AGB and BGB</i>		
1	$\ln(tDW) = \ln(a) + b(\ln D) + \varepsilon$	Lisboa et al. ¹⁶
2	$\ln(tDW) = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	Magalhães ⁴²
3	$\ln(tDW) = \ln(a) + b\ln(D^2H) + \varepsilon$	Mugasha et al. ⁴²

where tDW is the total (stem, branches, and foliage) dry weight of individual tree (kg tree⁻¹); D is the diameter at breast height (cm); H is the total height (m); $\ln$ is natural logarithm with base e ; and a , b and c are the regression parameters.

243

244 **Table 2:** The applied H-D functions. The references give the original reference and the first
 245 use in H-D modeling ⁴³ when applicable.

Number	Function name	Equation
<i>2-parameters models</i>		
1	Naslund	$H(D) = 1.3 + D^2/(a + bD)^2 + \varepsilon$
2	Curtis	$H(D) = 1.3 + a(D/1+D)^b + \varepsilon$
3	Michailoff	$H(D) = 1.3 + a \exp(-bD^{-1}) + \varepsilon$
4	Meyer	$H(D) = 1.3 + a(1 - \exp(-bD))$
5	Power	$H(D) = 1.3 + aD^b + \varepsilon$
6	Michaelis-Menten	$H(D) = 1.3 + aD/(b + D) + \varepsilon$
7	Wykoff	$H(D) = 1.3 + \exp(a + b/D + 1) + \varepsilon$
<i>3-parameters models</i>		
8	Prodan	$H(D) = 1.3 + D^2/(a + bD + c D^2) + \varepsilon$
9	Logistic	$H(D) = 1.3 + a/(1 + b \exp(-cD)) + \varepsilon$
10	Weibull	$H(D) = 1.3 + a(1 - \exp(-bD^c)) + \varepsilon$
11	Gomperz	$H(D) = 1.3 + a \exp(-b \exp(-cD)) + \varepsilon$
12	Sibbesen	$H(D) = 1.3 + aD^{bD^{-c}} + \varepsilon$
13	Korf	$H(D) = 1.3 + a \exp(-bD^{-c}) + \varepsilon$
14	Ratkowsky	$H(D) = 1.3 + a \exp(-b/D+c) + \varepsilon$
15	HossfeldIV	$H(D) = 1.3 + a/(1+(1/bD^c)) + \varepsilon$

246 where D is the diameter at breast height (cm), H is the total height (m); a , b , and c are
 247 the regression parameters.

248

249 **Data analysis**

250 *Test of equality of regression equations for sites*

251 We utilized a regression stability test to assess the equality of regression
252 equations for the different sites (Mabalane and Tambara). Based on the approach
253 described by Cienciala et al. ⁴⁴ and Umunay et al. ¹⁴, this test involved comparing pairs
254 of linear regressions derived from the two datasets. Using Model 1 (TM1) from **Table**
255 **2**, we performed log-linear ordinary least square (OLS) regressions separately for each
256 site (Mabalane and Tambara) and the combined sample data. We recorded the sum
257 squares of residuals for each regression. To conduct the regression stability test, known
258 as the Chow test, we assessed whether the regression coefficients differed significantly
259 for the two-site data sets. This test examined whether one regression line provided a
260 better fit for a split data set. The Chow F (FCH) statistic was calculated using the
261 residual sum squares obtained from each subsample. The Chow F statistic is calculated
262 as follows:

$$263 \quad F_{CH} = \frac{(RSC - RSC_1 - RSC_2)(n - 2m)}{(RSC_1 + RSC_2)m} \quad (1)$$

264

265 Where RSC = pooled (combined) residual sum square; RSC_1 = residual sum square for
266 site 1, RSC_2 = residual sum square for site 2, $n = (n_1 + n_2)$ is the sample size, and m is
267 the number of parameters in the model. The Chow test was then compared with the
268 F-distribution quintile at $F_{0.95}(m, n-2m)$ for homoscedasticity and with $F_{0.95}(m, r)$ for
269 heteroscedasticity where (Eq.(2)):

$$r = \frac{[(n_1-m)_1^2 + (n_2-m)_2^2]^2}{(n_1-m)_1^4 + (n_2-m)_2^4} \quad (2)$$

We then compared the calculated Chow F-statistic with the critical value to determine whether there was a significant difference between the regression coefficients for the two data sets. This test allows us to assess the equality of the regression equations for the different sites and decide whether to separate or combine the data for further analysis.

Analytical methods for Biomass Allometric Equation and height-diameter model

To avoid arbitrary choices, we conducted a research review to identify the most used methods for fitting biomass allometric models (BAM). We performed these analyses to ensure that we employed reliable and widely accepted analytical methods, as suggested by Sileshi ⁴⁵. We tested several analytical methods for fitting the BAM using Model 1 from **Table 2**. The methods we examined were: 1. Log-linear ordinary least squares (OLS): This is a commonly used method for fitting linear regression models; (2) Robust regression: We tested various robust regression methods, including Huber losses, LMS (Least Median Squares), LTS (Least Trimmed Squares), LAD (Least Absolute Deviations), S-estimator, and MM-estimator. After comparing the results, we selected the MM-estimator as the best estimate for our data and used it as the robust regression estimator; (3) Non-linear regression: We used the ‘*nlstools*’ package ⁴⁶ to fit the non-linear regression model; (4) Mixed-effect regression analysis: We tested the weighted mixed effect method using the ‘*nlme*’ package ⁴⁷. For each method, we estimated the model parameters, confidence intervals for each parameter, error of each

parameter (PRSE), and mean absolute percentage error of the model (MAPE), as proposed by Sileshi ⁴⁵. These measures allowed us to assess the accuracy and precision of each method. To select the best method for fitting the candidate BAMs from **Table 1**, we considered the methods with lower AIC, MAPE, and PRSE. We used these criteria to determine the method that provided the best fit for our data. We employed the same analysis to fit the below-ground biomass model (root biomass), ensuring consistency in the analytical methods used. For the H-D model, we used the non-linear model instead of comparing several methods as we did for the biomass allometric equation. Although some authors have criticized the fixed-effect prediction of a mixed-effects model for poor predictive performance, nowadays, it is seldom problematic to fit the H-D model using a non-linear mixed-effect model, and most non-linear models have been demonstrated to provide a good fit ¹⁷.

Fitting biomass allometric model and H-D functions

We performed model diagnostics to assess the adequacy of the BAM, including residual plots and influence statistics. We used global statistics metrics such as kurtosis, skewness, link function, and heteroscedasticity, which were calculated using the '*gvlma*' package ⁴⁸. Furthermore, we tested for collinearity using the variance inflation factor (VIF). Sileshi ⁴⁵ suggested that a VIF value higher than 5 indicates significant collinearity. In such cases, robust regression is considered efficient and appropriate. On the other hand, OLS regression is suitable when the predictors are few, not collinear, and have a well-understood relationship with the response variables. The model becomes unreliable if the percent relative standard errors (PRSE)

exceed 25% for one or more parameters. We employed the log OLS regression approach for the BAE. Conversely, the root biomass equations were fitted using robust regression with the MM-estimator from the 'MASS' package ⁴⁹. Simultaneously, we used the R-package 'lmfor' to fit and evaluate the height-diameter (H-D) models, following the protocol outlined by Mehtatalo and Kansanen ⁴³. These functions utilize the fitting routines of the 'nlme' package within the implemented functions. Default arguments and starting values for the parameters were used, making it convenient and efficient for height modeling ¹⁷. The initial parameter values were obtained from a linear model fit of the linearized H-D data.

Model evaluation and validation.

We employed several techniques to evaluate and validate the selected BAE and H-D equations. First, we selected the candidate equations based on the lowest mean absolute percentage error (MAPE, Eq. (3)) value, ensuring that the PRSE was lower than 25% for all parameters ⁴⁵. Additionally, we calculated Akaike's information criterion (AIC, Eq.(4)). Still, we used Akaike weights (AICw, Eq.(5)) ⁵⁰ for model selection as they provide more informative comparisons between models ⁴⁵. We also used model precision (V, Eq.(6)) statistics ⁵¹. We considered the model with the lowest MAPE, V, AIC, and highest AICw the best-fitting model. For further evaluation, we examined the plots of the residuals, which can provide information about the suitability and appropriateness of the selected models ⁵². We employed repeated K-fold cross-validation with 10-fold and three repetitions to validate the BAE and the H-D models. This cross-validation method assesses the performance of the models by

338 dividing the data into training and testing subsets and repeating the process multiple
 339 times. We evaluated the validation performance using metrics such as model efficiency
 340 (EF, Eq.(7))⁵³, error analysis (Eq.(8)), and comparisons between predicted and
 341 observed values using the Wilcoxon test. We employed these evaluation and
 342 validation techniques to assess the selected BAE and H-D models' accuracy, reliability,
 343 and predictive capabilities.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|M_o - M_p|}{M_o} \quad (3)$$

$$AIC_i = -2\log L_i + 2V_i \quad (4)$$

$$AIC_w = \frac{\exp\left\{-\frac{1}{2}\Delta_i(AIC)\right\}}{\sum_{k=1}^k \exp\left\{-\frac{1}{2}\Delta_k(AIC)\right\}} \quad (5)$$

$$V_i = \frac{\sum(M_o - M_p)^2 - (\sum(M_o - M_p))^2/n}{n(n-1)} \quad (6)$$

$$EF = 1 - \frac{\sum_{i=1}^n (M_o - M_p)^2}{\sum_{i=1}^n (M_o - \bar{M})^2} \quad (7)$$

$$Error = \sum_{i=1}^n (M_o - \bar{M}) \quad (8)$$

where n is the number of observations, M_o is observed data, M_p is predicted data, $\bar{M}$ is the mean of observed data, and L_i , the maximum likelihood for the candidate model i , is determined by adjusting the V_i free parameters to maximize the probability that the candidate model has generated the observed data.

Biomass estimation using allometric equation and tree height-diameter equation.

After fitting the BAE and the H-D models, we used the best-performing BAE (AGB and BGB), and the H-D model to estimate the total biomass of the *C. mopane* species. These equations were selected based on their performance in terms of accuracy and reliability, as determined through the evaluation and validation processes described earlier. We applied these selected equations to forest inventory data collected in two study areas: Mabalane and Tambara. Additionally, we assessed the uncertainty associated with the estimation of AGB, applying the model with D and H as second variables and the model with only D. We used H estimated by the selected best fitting H-D model. This uncertainty analysis provides insights into the reliability and precision of biomass estimation and H-D modeling results.

Plant guideline statement.

The collection of plant material complied with the guidelines of the Eduardo Mondlane University, Maputo, Mozambique.

Results

Site effects and biomass accumulation

The comparison between the Mabalane and Tambara sites revealed notable differences in tree characteristics and biomass accumulation. The Mabalane site exhibited more large-sized trees, with a maximum D of 109 cm, while Tambara had a maximum D of 44 cm (**Supplementary Table S1**). **Figure 3** illustrates the power dependence of tDW on tree D (Pearson coefficient, $r = 0.93$, $P < 0.001$) and H ($r = 0.78$, $P < 0.001$) in all sites. Similar power dependence relationships were observed for stem biomass on D ($r = 0.95$) and H ($r = 0.80$, $P < 0.001$), crown biomass on D ($r = 0.90$, $P < 0.001$) and H ($r = 0.73$, $P < 0.001$), as well as root biomass on D ($r = 0.94$, $P < 0.001$) and H ($r = 0.69$). The estimated value of the test criterion FCH, as calculated according to Eq.(1), was 0.46. Notably, despite the discrepancy in sample size between Mabalane ($n = 62$) and Tambara ($n = 58$) and the presence of more large-sized trees in Mabalane, the applied Chow test confirmed that the dependence of tDW (Model 2) on both independent variables (D and H) could be treated as identical for all sites. Furthermore, since the tabular reference value of $F_{0.95}(m, r)$ was 19.38, the experimental data from the two sites could be approximated with the standard regression equation at a confidence level of 95%. These findings highlight the influence of site characteristics on tree size and biomass accumulation.

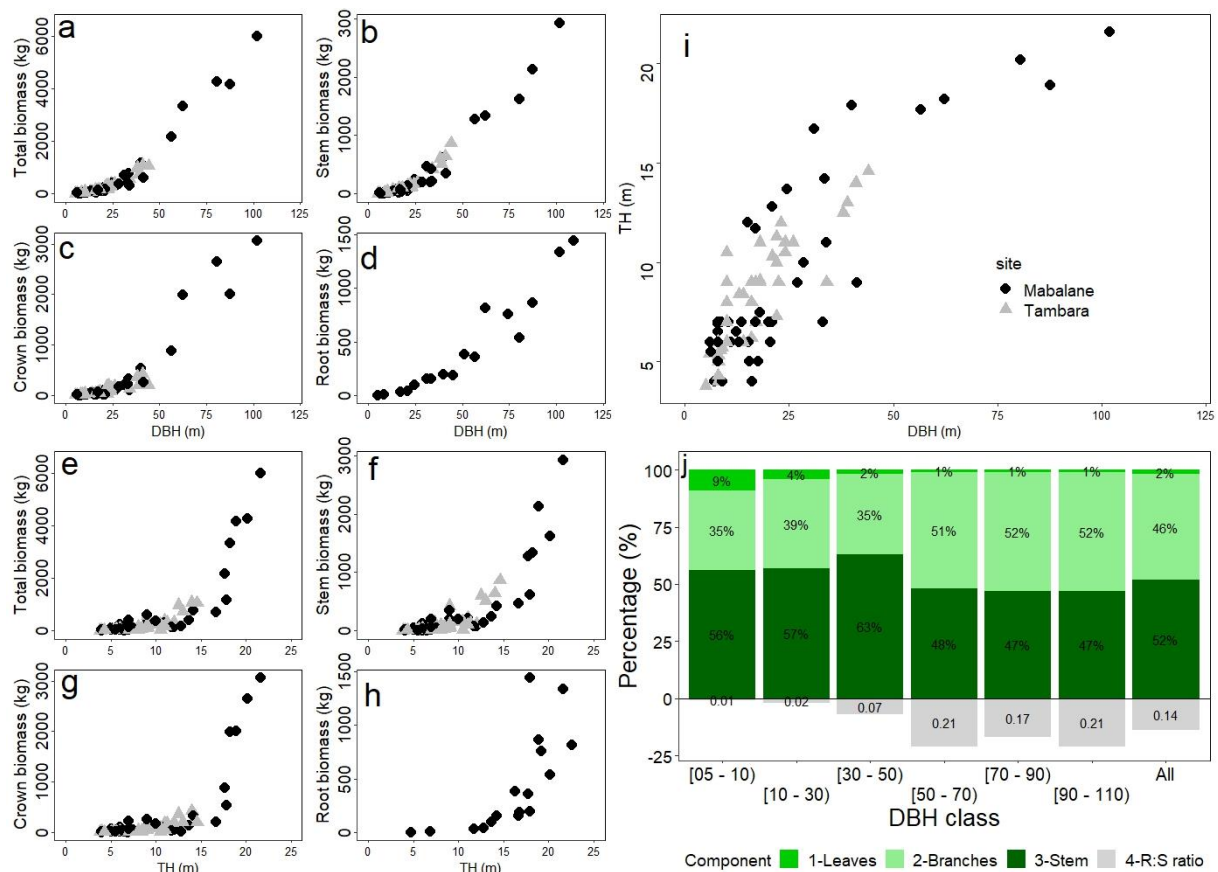


Figure 3: Individual sites plotted tree breast height diameter against a) total above-ground biomass, b) stem biomass, c) crown biomass, and d) below-ground biomass; tree height against e) total biomass, f) stem biomass, g) crown biomass and h) root biomass, i) DBH (D) against tree height and j) AGB proportion and the root-to-shoot ratio of *C. mopane* in Mozambique.

The root-to-shoot ratio and the accumulation of total tree biomass into different tree components along diameter classes are depicted in Figure 3j. The analysis reveals a relationship between D size and biomass allocation. The pooled data showed a root-to-shoot ratio of 0.14. Examining the root-to-shoot ratio across the D range, Figure 3j indicates that larger trees ($D > 50$ cm) exhibited significantly higher root-shoot ratios, ranging from 0.17 to 0.21. Furthermore, the figure illustrates that large mopane trees

($D > 50$ cm) stored almost the same amount of carbon in the stem and lower amounts in the branches. Conversely, small to mid-sized mopane trees ($D < 50$ cm) stored more carbon in the stem than in branches and leaves. Additionally, the percentage of leaf biomass declined rapidly with increasing tree diameter, suggesting that a higher proportion of biomass was accumulated in the stem (52%) and branches (46%). In large-sized trees ($D > 50$ cm), the biomass accumulation was notably higher in the branches compared to the stems. This observation aligns with the understanding that *C. mopane* trees, particularly when they reach a specific diameter (often large-sized), tend to have the heartwood deteriorate and become hollow. These findings shed light on the biomass allocation patterns and carbon storage dynamics in *C. mopane* trees, highlighting the influence of tree diameter on the distribution of biomass among different tree components.

Selection of methods of biomass model fitting

The different estimation methods tested in this study exhibited varying performances when applied to the AGB and BGB models. Ordinary Least Squares (OLS) and Robust regression with MM-estimator demonstrated statistically similar performance across all statistical parameters (**Table 4**). Both methods had comparable confidence intervals for the model parameters (intercept and exponent), similar PRSE, and the same bias, with a MAPE of 6.3% for both AGB and BGB models (MAPE = 4.8%). In contrast, non-linear and mixed effects regression displayed higher bias levels, exceeding 20% for both AGB and BGB models. Non-linear regression exhibited a high bias (MAPE = 24%) with the AGB model, while mixed effects regression performed

poorly (MAPE = 99%) with the BGB models. We selected the OLS regression method over the MM-estimator to fit the AGB and BGB models based on these results. This decision is in line with the recommendation made by Sileshi ⁴⁵, who stated that OLS regression is suitable when the predictors are few, not collinear, and have a well-understood relationship with the response variables. Considering the characteristics of our analysis, the OLS method demonstrated exemplary performance in fitting the AGB and BGB models which overestimates cancel out with underestimates (Supplementary Figure S1). By selecting the OLS regression method, we ensure the accuracy and reliability of the biomass estimation for both AGB and BGB components.

Table 4: Statistical performance of the model $tDW = aD^b$ for regression methods selection, i.e., log-linear ordinary least square (OLS), robust regression with the MM estimator, and non-linear.

Method of fit	Goodness of fit statistics	Parameters	Tree components			
			Total AGB (n = 84)	Stem biomass	Crown biomass	BGB
OLS	PRSE (%)	<i>b1</i>	9.67	8.94	8.62	14.15
		<i>b2</i>	2.57	3.11	3.53	3.72
	AIC (kg)		910.20	843.10	865.28	163.73
	AICw (kg)		0.594	0.533	0.572	0.501
	MAPE (%)		28.02	36.24	42.84	19.80
Robust regression with MM-estimator	PRSE (%)	<i>b1</i>	8.42	8.81	7.89	15.03
		<i>b2</i>	2.36	3.12	3.29	3.95
	AIC (kg)		933.97	866.52	837.65	163.70
	AICw (kg)		0.406	0.467	0.428	0.499

NLS	MAPE (%)		27.88	36.51	43.46	19.71
	PRSE (%)	<i>b1</i>	21.74	25.46	29.00	85.11
		<i>b2</i>	2.42	3.03	3.06	10.09
	AIC (kg)		888.43	797.11	813.14	163.45
	AICw (kg)		0.00	0.00	0.00	0.00
	MAPE (%)		75.32	88.71	102.95	22.54

432

433 Biomass allometric equations

434 Tables 5 and 6 present the parameter estimates and model performance results
435 for all BAE. All the models had significantly different parameters from zero, with PRSE
436 values below 25% for all models (**Table 5**). Among the models, the parameter *c*
437 associated with the *H* variable exhibited the highest PRSE values, exceeding 20% in all
438 models (SM2, CM2, RM2, and TM2). The root model (RM2) had an exceptionally high
439 PRSE value, exceeding 100%. However, the parameter *c* associated with the height
440 variable was slightly high but still significantly greater than zero (t-test, $P < 0.001$) for
441 the stem model (SM2), crown model (CM2), and total tree model (TM2). The first
442 model (SM1, CM1, and TM1) displayed a PRSE for the parameter three times lower
443 than the *b* parameter of the *D* variable. This trend was observed in the other models as
444 well, including the stem biomass models (SM2 and SM3), crown biomass models (CM2
445 and CM3), and tDW models (TM2 and TM3). However, in the root model (RM1),
446 parameter *a* was almost four times lower than parameter *b*. These findings highlight
447 the different parameter estimates and model performance for the various biomass
448 allometric equations. Despite some variations in parameter values, all the models

demonstrated statistically significant relationships between the predictors (D and height) and the biomass components.

Table 5: Parameter estimation and its root standard error in percent of the allometric equations for biomass.

Equation	Model	Parameters			PRSE(%)		
		$a \pm \text{SE}$	$b \pm \text{SE}$	$c \pm \text{SE}$	a	b	c
		(95% CI)	(95% CI)	(95% CI)			
Stem biomass							
$\ln(\text{tDW}) = \ln(a) + b(\ln D) + \varepsilon$	SM1	-2.251 ± 0.201 (-2.651 to -1.851)	2.258 ± 0.070 (2.118 to 2.398)		8.93	3.10	
$\ln(\text{tDW}) = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	SM2	-2.819 ± 0.230 (-3.277 to -2.360)	1.866 ± 0.115 (1.637 to 2.095)	0.797 ± 0.194 (0.411 to 1.183)	8.16	6.16	24.34
$\ln(\text{tDW}) = \ln(a) + b\ln(D^2H) + \varepsilon$	SM3	-2.882 ± 0.200 (-3.279 to -2.485)	0.904 ± 0.026 (0.853 to 0.955)		6.94	2.88	
Crown biomass							
$\ln(\text{tDW}) = \ln(a) + b(\ln D) + \varepsilon$	CM1	-2.723 ± 0.235 (-3.190 to -2.256)	2.327 ± 0.082 (2.163 to 2.490)		8.63	3.52	
$\ln(\text{tDW}) = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	CM2	-3.145 ± 0.285 (-3.715 to -2.581)	2.034 ± 0.142 (1.750 to 2.317)	0.597 ± 0.240 (0.119 to 1.074)	9.06	6.98	40.20
$\ln(\text{tDW}) = \ln(a) + b\ln(D^2H) + \varepsilon$	CM3	-3.343 ± 0.249 (-3.839 to -2.848)	0.927 ± 0.0319 (0.865 to 0.991)		7.45	3.44	
Root biomass							
$\ln(\text{rDW}) = \ln(a) + b(\ln D) + \varepsilon$	RM1	-1.927 ± 0.2728 (-2.509 to -1.3459)	1.964 ± 0.073 (1.808 to 2.120)		14.17	3.72	
$\ln(\text{rDW}) = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	RM2	-2.346 ± 0.570 (-3.570 to -1.123)	1.797 ± 0.212 (1.342 to 2.252)	0.380 ± 0.453 (-0.591 to 1.350)	24.30	11.80	119.21
$\ln(\text{rDW}) = \ln(a) + b\ln(D^2H) + \varepsilon$	RM3	-2.797 ± 0.306 (-3.450 to -2.144)	0.803 ± 0.030 (0.739 to 0.867)		10.94	3.74	
Tree total biomass							
$\ln \text{DW} = \ln(a) + b(\ln D) + \varepsilon$	TM1	-1.740 ± 0.168	2.289 ± 0.059		9.66	2.58	

		(-2.074 to -1.405)	(2.172 to 2.406)				
$\ln DW = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	TM2	-2.280 ± 0.187	1.917 ± 0.093	0.757 ± 0.157	8.20	4.855	20.740
		(-2.651 to -1.909)	(1.731 to 2.102)	(0.445 to 1.070)			
$\ln DW = \ln(a) + b \ln(D^2H) + \varepsilon$	TM3	-2.373 ± 0.162	0.916 ± 0.021		6.827	2.293	
		(-2.696 to -2.050)	(0.875 to 0.957)				

453

454 The results of the goodness of fit and validation for all the allometric biomass
455 models are presented in Table 6. The total biomass models (TM1, TM2, and TM3)
456 performed well, fitting the data well with acceptable precision. TM1, which only used
457 D as a predictor, and TM3, which included interactive D^2H as a predictor, had the same
458 MAPE value of 5.8%. TM2 and TM3 had the same R-squared value of 0.96, while TM1
459 had the lowest R-squared value of 0.95. However, based on the AICw, TM3, which
460 used interactive D^2H as a predictor, provided the best fit for the dataset. All the models
461 demonstrated high predictive accuracy for total biomass, with V values ranging from
462 0.003 kg (TM2) to 0.004 kg for TM1 and TM3. These results indicate that the biomass
463 models incorporating D and H as predictors represented the dataset well regarding
464 goodness-of-fit parameters. However, the AICw showed that the TM3 model had the
465 highest predictive power, with a higher EF and lower error for the validation data
466 compared to TM1 and TM2.

467 The allometric models for stem biomass outperformed the crown biomass based
468 on AIC, AICw, V, MAPE, R-squared, and MSA. The model incorporating D^2H (labeled
469 as SM3 and CM3 for stem and crown biomass, respectively) provided the best fit for
470 both stem and crown biomass. This model also performed better on the validation

data, with a higher EF and significantly lower error (Wilconx test, $P < 0.001$) than the other allometric models for stem and crown biomass (SM1, SM2, CM1, and CM2). The allometric models for root biomass performed well regarding AIC, MAPE, MSA, and R-squared but had the lowest model accuracy. Among the root biomass models, RM1 was the best-fitting model with a higher AICw than RM2 and RM3. The residual distribution of each model showed that Model 3 (D^2H variable) was accurate but not precise for all components, including total AGB and root biomass (BGB), as the scatter points deviated slightly from the zero line. On the other hand, Models 1 (with D only) and 2 with H as the second variable were both inaccurate and imprecise, as shown in Table 6 and residuals distribution (**Supplementary Figure S2**).

Table 6: The statistical performance of the biomass allometric equation in terms of predictive accuracy and predictive ability tested with data not included in the model fit.

The goodness of fit statistics								Validation statistics		
(predictive accuracy)								(predictive ability)		
Equation	Model	AIC	AICw	V(kg)	MAPE	R-	MSA	EF (%)	Error	P value for
		(kg)			(%)	square			(kg)	Wilcoxon test
Stem biomass										
$\ln fDW = \ln(a) + b(\ln D) + \varepsilon$	SM1	95.29	0.00	0.00204	10.28	92.63	30.60	93.04	3.27	0.665
$\ln fDW = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	SM2	81.41	0.28	0.00169	9.06	93.75	23.60	90.52	-15.78	0.539
$\ln fDW = \ln(a) + b\ln(D^2H) + \varepsilon$	SM3	79.73	0.72	0.00170	9.06	93.80	23.77	90.85	-15.91	0.532
Crown biomass										
$\ln fDW = \ln(a) + b(\ln D) + \varepsilon$	CM1	121.30	0.061	0.00278	12.51	90.63	36.84	97.19	0.46	0.942
$\ln fDW = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	CM2	117.13	0.444	0.00259	12.05	91.19	36.42	97.11	-12.67	0.915

$\ln tDW = \ln(a) + b \ln(D^2H) + \varepsilon$	CM3	117.11	0.495	0.00265	12.23	91.08	31.73	97.47	-10.63	0.879
Root biomass										
$\ln tDW = \ln(a) + b(\ln D) + \varepsilon$	RM1	5.57	0.461	0.00357	4.77	97.83	16.61	-	-	-
$\ln tDW = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	RM2	6.74	0.122	0.00340	4.39	97.78	16.87	-	-	-
$\ln tDW = \ln(a) + b \ln(D^2H) + \varepsilon$	RM3	5.78	0.417	0.00361	4.67	97.80	18.93	-	-	-
Tree total biomass										
$\ln tDW = \ln(a) + b(\ln D) + \varepsilon$	TM1	65.23	0.00	0.00143	6.65	94.81	25.22	95.98	-15.54	0.757
$\ln tDW = \ln(a) + b(\ln D) + c(\ln H) + \varepsilon$	TM2	46.06	0.36	0.00111	5.82	95.91	17.12	94.10	-50.22	0.714
$\ln tDW = \ln(a) + b \ln(D^2H) + \varepsilon$	TM3	45.12	0.64	0.00112	5.83	95.91	17.22	94.11	-53.81	0.681

Comparison with local tree-specific-allometric equation

To compare the performance of the best-fitted model from this study, we selected tree-specific BAE developed for native tree species in Mozambique, such as *Millettia stuhlmannii*, *Azelaia quansensis*, *Pterocarpus angolensis*²³, and *Androstachys johnsonii*⁴². We applied these equations to estimate AGB for *C. mopane* and compared the results. The relative mean percentage error (RMPE) depicted in **Figure 4** revealed that the Afzelia and Millettia models significantly overestimated the AGB of *C. mopane* (Wilcox test, $V > 900$, $p < 0.001$). Similarly, the Pterocarpus and Androstachys models significantly overestimated the AGB (Wilcox test, $V > 600$, $p < 0.001$). Consequently, the mean AGB estimated from these equations differed substantially from the observed AGB of *C. mopane* (Wilcox test, V ranged from 348 to 2737, $P < 0.001$). These results indicate that although these models were developed for tree species within the same country, they were not suitable for accurately estimating the AGB of *C. mopane*. However, the best-fitted model from this study resulted in a lower underestimation of

AGB, with an RMPE of 5.6%. Therefore, the findings suggest that the best-fitted model from this study is more appropriate for estimating the AGB of *C. mopane* compared to the tree-specific BAE developed for other native tree species in Mozambique.

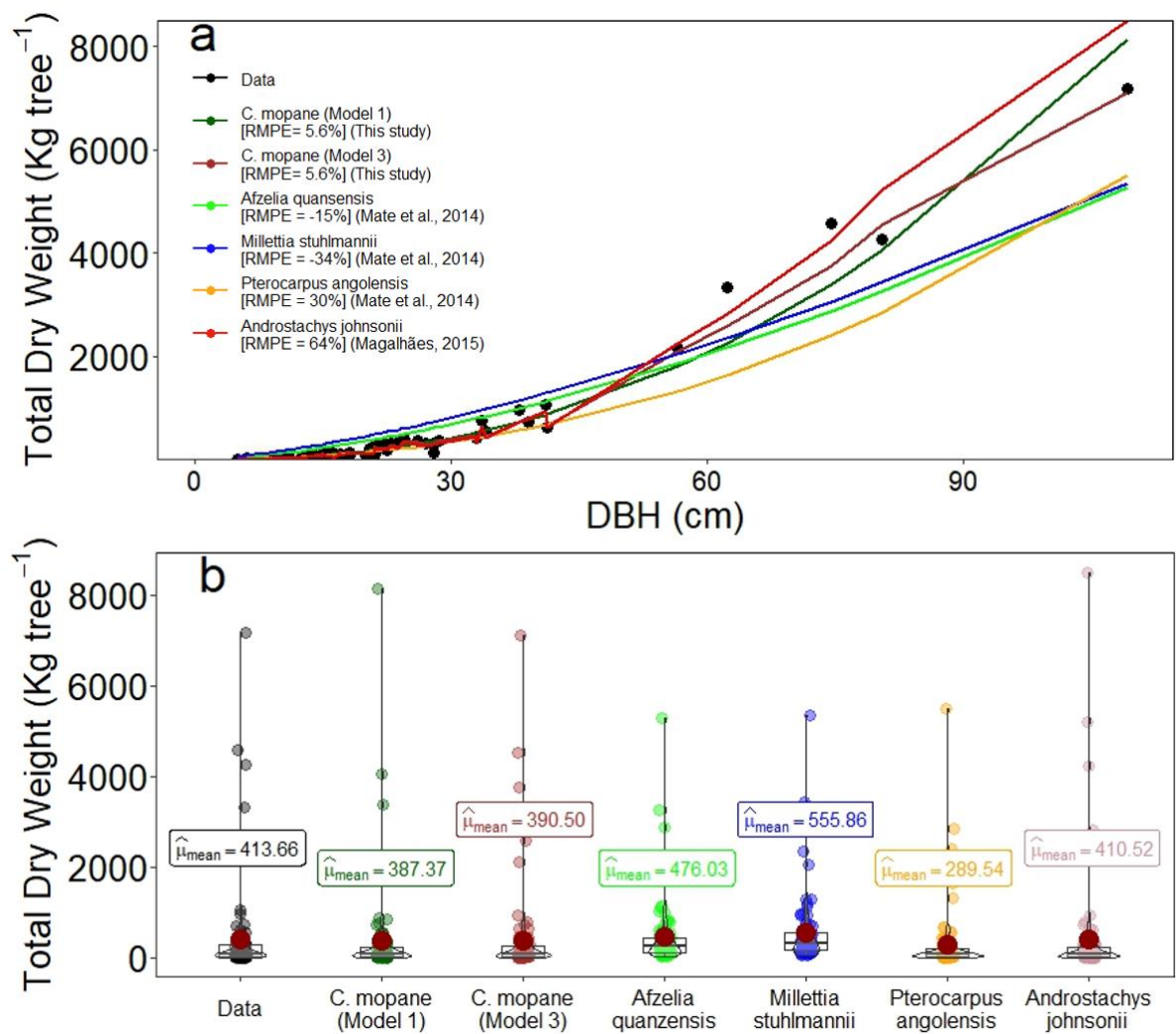


Figure 4: Comparison of predictive accuracy of the biomass allometric Model 1 (TM1) and 3 (TM3) developed in this study against existing species-specific biomass allometric equation in Mozambique, as compared with observed above-ground total dry weight (tDW).

Height-Diameter relationship

The goodness-of-fit of the candidate models, both 2- and 3-parameters, is depicted in **Table 7**. Among the seven 2-parameter models (Meyer, Power, and Michaelis-Menten) fit the data well (**Supplementary Figure S3**). The HossefeldIV and Weibull models also offered a good fit. The other models only work well for small-sized trees ($D < 40$ cm) and show a flat trend for trees above 40 cm of D . In the mixed-effects model fitting, all the 2-parameter models estimated their parameters (a and b) to significantly differ from zero (**Table 7**). The Power model had the lowest PRSE values for its parameters, followed by the Meyer model. On the other hand, for the 3-parameter models, all parameters scored a PRSE very higher than 25%. As result only one parameter of each model was significantly different from zero, suggesting that although these models fitted the trend of the data well, only the c parameter for HossefeldIV and Weibull models and a parameter for Sibbesen model explain the variability in the data. According to the AICw, the best 2-parameter model was the Power model (Table 8). Among the 3-parameter models, the HossfeldIV performed the best. The overestimation and underestimation of the model either with 2 or 3 parameters neither, was statistically significant (Wilcox test, $p > 0.05$). **Supplementary Figure S4** displays the residuals dispersion against the D of fitting data.

Table 7: Parameter estimation and their root standard errors in percent of height-diameter models with 2 and 3 parameters.

Model	Parameters	PRSE(%)
-------	------------	---------

	$a \pm \text{SE}$ (95% CI)	$b \pm \text{SE}$ (95% CI)	$c \pm \text{SE}$ (95% CI)	b_0	b_2	b_3
<i>Models with 2 parameters</i>						
Power	1.225 ± 0.157 (0.917 to 1.533)	0.616 ± 0.042 (0.533 to 0.699)	-	12.82	6.82	
Michaelis-Menten	23.222 ± 3.397 (16.564 to 29.879)	36.894 ± 8.104 (21.010 to 52.777)	-	14.63	21.97	
Meyer	16.300 ± 2.012 (12.356 to 20.244)	0.035 ± 0.006 (0.023 to 0.047)	-	12.34	17.14	
<i>Models with 3 parameters</i>						
HossfeldIV	355.104 ± 3489.245 (-6483.816 to 7194.024)	0.003 ± 0.032 (-0.060 to 0.067)	0.632 ± 0.167 (0.306 to 0.959)	982.60	1066.67	26.42
Sibbesen	1.182 ± 0.579 (0.047 to 2.317)	0.641 ± 0.327 (-0.00052 to 1.282)	0.006 ± 0.081 (-0.152 to 0.165)	48.98	51.01	1350
Weibull	92.751 ± 333.356 (-572.386 to 757.888)	0.012 ± 0.042 (-0.070 to 0.095)	0.654 ± 0.155 (0.351 to 0.957)	365.88	340.34	23.64

Table 8: The statistical performance of the height-diameter models in terms of predictive accuracy and predictive ability tested with data not included in the model fit.

Models	Goodness of fit statistics (Predictive accuracy)					Validation statistics (Predictive ability)		
	AIC (m)	AICw	RSE (m)	MAPE (%)	Adj.R-square (%)	EF (%)	Error (m)	P-value for Wilcoxon test
<i>Models with 2 parameters</i>								
Power	352.14	0.874	0.546	19.81	75.61	83.37	-0.15	0.510
Michaelis-Menten	356.42	0.103	0.504	20.41	73.18	81.05	-0.03	0.562
Meyer	359.41	0.023	0.495	20.66	71.11	79.69	0.04	0.592

<i>Models with 3 parameters</i>								
HossfeldIV	354.39	0.495	0.550	19.81	75.33	79.57	0.71	0.951
Sibbesen	354.39	0.494	0.550	19.82	75.32	83.41	-0.17	0.510
Weibull	361.24	0.011	1.23	19.77	75.35	83.41	0.08	0.656

Estimating Above – and below-ground biomass of mopane woodland and uncertainty.

Table 9 presents the results of estimating AGB and BGB using the best-fitting and H-D models. Model 3 of BAE emerged as the top performer, accurately estimating both AGB and BGB for mopane woodland from validation data. In contrast, the power model was the most suitable among the H-D models. Consequently, Model 3 was employed to estimate AGB and BGB, while the power H-D model was used to estimate H, which is a component of Model 3. The power model estimated mean H to be 8.0 m, and there was no significant difference in AGB estimation between Model 3 and Model 1. The same trend was observed for BGB. Both Model 1 and Model 3 estimated a total biomass (AGB+BGB) of 15.0 Mg ha⁻¹, with an uncertainty of 8%. Although the differences were not statistically insignificant, Model 3 exhibited higher biomass estimation uncertainty than Model 1. The biomass estimation for other species, using the equation proposed by Guedes et al. (2018), amounted to 6.0 Mgha⁻¹. This result suggests that mopane species account for more than half of Mozambique's total biomass in mopane forests. The average stem size of *C. mopane* species (D = 18.6 cm) aligns with that of other species (D = 18.01 cm) in mopane forests, but the stem density of *C. mopane* species (85 N ha⁻¹) is twice that of different species (40 N ha⁻¹). These findings shed light on the higher biomass reported for *C. mopane*.

553

554 **Table 9:** The AGB of *C. mopane* was estimated using the best-fitting biomass allometric
 555 equations and the Power height-diameter model selected in this study, estimated from
 556 forestry inventory data.

	Mean (95% of CI)	Minimum	Maximum	Uncertainty from estimation (%)
<i>C. mopane</i>				
Dg (quadratic diameter, cm)	18.16 (18.14 – 18.18)	5.0	107.0	5.78
TH (estimated with Power model, m)	8.01 (7.91 – 8.11)	4.6	23.09	5.79
Stem density (N ha ⁻¹)	85 (84 – 87)	4	265	7.83
Basal area (m ² ha ⁻¹)	2.19 (2.15 – 2.23)	0.01	5.64	8.72
<i>Model 1</i>				
AGB (Mg ha ⁻¹)	12.08 (11.95 – 12.21)	0.05	32.30	4.78
BGB (Mg ha ⁻¹)	3.63 (3.60 – 3.66)	0.02	9.33	3.75
Total (Mg ha ⁻¹)	15.71 (15.41 – 16.01)	0.08	41.63	8.56
<i>Model 3</i>				
AGB (Mg ha ⁻¹)	12.40 (12.26 – 12.54)	0.05	34.08	5.19
BGB (Mg ha ⁻¹)	3.09 (3.05 – 3.13)	0.02	7.96	5.15
Total (Mg ha ⁻¹)	15.48 (15.18 – 15.78)	0.07	41.33	8.84
<i>Other species (Guedes et al., 2018)</i>				
Mean D (cm)	18.01	5	95.00	2.36

	(17.97 – 18.05)			
Stem density (N ha ⁻¹)	40 (40 – 41)	1.25	182.5	10.14
Basal area (m ² ha ⁻¹)	1.01 (0.99 – 1.03)	0.01	3.78	10.58
AGB (Mg ha ⁻¹)	6.28 (6.12 – 6.44)	0.05	25.50	11.15
<i>Mopane woodland</i>				
Total Biomass (AGBcm + BGBcm + AGBosp.)	21.6 (21.3 – 22.22)	0.05	41.33	9.58

Discussion

Biomass Allometric Equation and Biomass estimation

This study aimed to develop species-specific allometric equations to estimate the AGB and BGB of *C. mopane* species in Mozambique. In addition, the study investigated the feasibility of integrating tree height data from the H-D model into the BAE to improve biomass estimation in MW. Including estimated H in BAE for biomass estimation is controversial in ecological research. While some studies argue that including *H* improves accuracy^{15,54}, others suggest that it may lead to an underestimation of carbon storage⁵⁵ or may not improve accuracy at all^{16,56}. In the case of *C. mopane* species in Mozambique, this study found that the model that included the interaction of D^2 and *H* performed better than the model that used separate *D* and *H* variables, suggesting a non-linear relationship between *D* and *H*. However, models that included both *D* and *H* performed similarly to models that used only diameter, suggesting that including *H* as an explanatory variable did not significantly improve the biomass estimation accuracy for *C. mopane* species. This

result is consistent with previous studies ^{16,56}. The results of AGB estimation showed that the inclusion of *H* data from the H-D model in Model 3 for AGB estimation resulted in higher error propagation compared to Model 1, which used only *D* as an explanatory variable. The results also suggest that Model 1 may be a more appropriate option for estimating AGB with lower uncertainty. However, Model 3 could be used interchangeably with Model 1 to estimate the overall AGB of *C. mopane* species as it resulted in similar levels of uncertainty as Model 3. In this study, the allometric equation from Guedes et al.²² was applied to estimate the AGB of non-mopane species, as some analyses ⁵⁷ suggest that there are common plant species between MW and miombo woodland. Woollen et al.⁵⁸ applied a BAE developed for miombo woodlands to estimate AGB in mopane woodlands, including mopane species. Guedes' model produced AGB estimates similar to those from Woollen et al. ⁵⁸, suggesting that Guedes' model could complement the BAE developed for *C. mopane* in this study to estimate AGB of non-mopane species in MW.

Model comparison with existing models.

This study serves as a critical evaluation of the applicability of existing species-specific ABE to *C. mopane* species. By examining the discrepancies between the ABEs for *M. stuhlmannii*, *P. angolensis* and *A. quanzensis* ⁵⁹ and *A. johnsonii* ⁴² in estimating the AGB of *C. mopane*, we uncovered a significant underestimation that is likely due to differences in wood density and architectural features between species. Conversely, our study revealed that the BAE for *A. johnsonii* overestimated AGB in *C. mopane*. The divergent estimates of these species-specific equations highlight the complexity of biomass estimation and the need to account for differences in tree architecture and

wood density when developing predictive models ⁶⁰. Furthermore, these results emphasize the importance of considering species-specific traits when developing biomass estimation models to improve accuracy and reliability. Despite the challenges posed by the diversity of tree species ²⁶, the precision and predictive power achieved through the development of species-specific equations justify the necessary investment in effort and resources for sustainable forest management and carbon sequestration initiatives.

Height-Diameter model

A comprehensive analysis was performed to compare 20 non-linear height-diameter models. Although they showed a high goodness of fit, several models were discarded based on the graphical evaluation, indicating a lack of fit to our dataset. Among the 2- and 3-parameter model categories examined, several models were selected for each group, and the power model was identified as the optimal 2-parameter model for predicting the tree height of *C. mopane* species. The Power model was also selected as the best H-D model in the evergreen mountain forests in Mozambique¹⁶. This result may indicate that the power model effectively represents the H-D relationship of generalized species and specie-specific data. The HossfeldIV model proved to be the best among the 3-parametre models. Nonetheless, the problems with model convergence, especially for the 3-parameter models, highlight the difficulties that arise with the increasing complexity of non-linear H-D models, a phenomenon also observed in previous studies ^{17,55}. The convergence emphasizes the importance of selecting models that balance complexity and practical utility. This study's careful and rigorous approach emphasizes the importance of methodological

precision in developing accurate height-diameter models tailored to the specific characteristics of individual species and forest environments. Furthermore, the developed height-diameter model can improve the predictive ability of vegetation parameters in *C. mopane* species, especially in the context of anthropogenic disturbances. This model not only has the potential to streamline national forest inventories by saving costs and time but also serves as a valuable tool to improve forest management strategies and conservation efforts.

Implication on sustainable management of mopane woodlands

The biomass stocks estimated in our study are consistent with those reported in previous studies on MW⁵⁸. The study showed that the proportion of stem biomass decreased significantly in trees with a $D > 50$ cm. This was mainly expected because the stems of *C. mopane* become hollow with increasing age due to a natural process of decay and decomposition. This hollowing process is a common phenomenon in older mopane trees and is a natural part of their life cycle. The importance of including BGB in the estimation of total biomass is crucial to avoid underestimating carbon stocks in the mopane ecosystem. The development of customized allometric equations for the estimation of AGB and BGB of *C. mopane* species in Mozambique is an essential contribution to the existing species-specific ABE^{42,51,59,61} and improves biomass assessment procedures. In addition, the study highlights the effectiveness of the H-D model in accurately estimating H , thus improving the precision and efficiency of estimating stand volume and biomass from forest inventory data in MW. The implications of these results go beyond academic research and offer practical applications. For instance, charcoal production in the MW is a selective process⁵⁸, with

mainly large trees being felled in specific areas such as the Mabalane and Tambara districts^{62,63}. This selective extraction of trees for charcoal production can lead to deforestation and forest degradation, which greatly concerns governments, forestry institutions, and non-governmental organizations. To address these challenges, Mozambique has developed an operational Nationally Determined Contribution (NDC) plan, which includes a Nationally Appropriate Mitigation Action (NAMA) to promote sustainability in the charcoal value chain⁶⁴. By integrating the results of our study into forest management strategies, policymakers, and conservationists can transition to more sustainable practices that support ecological conservation and climate change mitigation efforts as outlined in the NAMA. This holistic approach to biomass assessment not only improves carbon stock estimates for initiatives such as the REDD+ mechanism but also provides valuable insights for sustainable management practices in MW under charcoal production.

Conclusion

This study provides valuable insights for the development of species-specific allometric equations for estimating AGB and BGB of *C. mopane* species in Mozambique. The study demonstrates the importance of considering the interaction between D and TH in the ABE. The model that includes the quadratic variables for D and TH performs better than the model with separate variables for D and TH. However, the inclusion of TH as an explanatory variable did not lead to a significant improvement in the accuracy of the biomass estimation. The power model showed the best performance and revealed the potential of the H-D model as an alternative for estimating TH from forest inventory data in MW. The results highlight the need for species-specific

allometric equations to account for the unique characteristics and growth patterns of individual tree species in southern Africa and Mozambique in particular, as the existing species-specific models cannot be generalized. This highlight is particularly important for Mozambique's efforts to accurately estimate biomass for the REDD+ mechanism and meet climate targets. These results have practical implications for the sustainable management of MW and can help to accurately assess carbon emissions and climate change mitigation efforts.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Winkler, K., Fuchs, R., Rounsevell, M. & Herold, M. Global land use changes are four times greater than previously estimated. *Nat. Commun.* **12**, 1–10 (2021).
2. Houghton, R. A. *et al.* Carbon emissions from land use and land-cover change. *Biogeosciences* **9**, 5125–5142 (2012).
3. Gasser, T. *et al.* Historical CO₂ emissions from land use and land cover change and their uncertainty. *Biogeosciences* **17**, 4075–4101 (2020).
4. Gumbo, D. *et al.* How have carbon stocks in central and southern Africa's miombo woodlands changed over the last 50 years? A systematic map of the evidence. *Environ. Evid.* **7**, 1–19 (2018).
5. IPCC. Climate Change and Land: an IPCC special report. *Clim. Chang. L. an IPCC Spec. Rep. Clim. Chang. Desertif. L. Degrad. Sustain. L. Manag. food Secur. Greenh. gas fluxes Terr. Ecosyst.* 1–864 (2019).
6. Grainger, A. Constraints on modelling the deforestation and degradation of tropical open woodlands. *Glob. Ecol. Biogeogr.* **8**, 179–190 (1999).
7. Ciais, P. *et al.* The carbon balance of Africa: synthesis of recent research studies. *Philos. Trans. Ser. A* **369**, 2038–2057 (2011).
8. McNicol, I. M., Ryan, C. M. & Mitchard, E. T. A. Carbon losses from deforestation and widespread degradation offset by extensive growth in African woodlands. *Nat. Commun.* **9**, 1–11 (2018).
9. Millar, C. I., Stephenson, N. L. & Stephens, S. L. Climate change and forests of the future: Managing in the face of uncertainty. *Ecol. Appl.* **17**, 2145–2151 (2007).
10. Njana, M. A., Mbilinyi, B. & Eliakimu, Z. The role of forests in the mitigation of global climate change: Empirical evidence from Tanzania. *Environ. Challenges* **4**, 100170 (2021).
11. Herold, M. *et al.* The Role and Need for Space-Based Forest Biomass-Related Measurements in Environmental Management and Policy. *Surv. Geophys.* **40**, 757–778 (2019).
12. Henry, M. *et al.* An overview of existing and promising technologies for national forest monitoring. *Ann. For. Sci.* **72**, 779–788 (2015).
13. Nam, V. T., Van Kuijk, M. & Anten, N. P. R. Allometric equations for aboveground and

- belowground biomass estimations in an evergreen forest in Vietnam. *PLoS One* **11**, 6–9 (2016).
14. Umunay, P. M., Gregoire, T. G. & Ashton, M. S. Estimating biomass and carbon for Gilbertiodendron dewevrei (De Wild) Leonard, a dominant canopy tree of African tropical Rainforest: Implications for policies on carbon sequestration. *For. Ecol. Manage.* **404**, 31–44 (2017).
15. Chave, J. *et al.* Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **20**, 3177–3190 (2014).
16. Lisboa, S. N., Guedes, B. S., Ribeiro, N. & Siteo, A. Biomass allometric equation and expansion factor for a mountain moist evergreen forest in Mozambique. *Carbon Balance Manag.* **13**, (2018).
17. Mehtätalo, L., de-Miguel, S. & Gregoire, T. G. Modeling height-diameter curves for prediction. *Can. J. For. Res.* **45**, 826–837 (2015).
18. Patrício, M. S., Dias, C. R. G. & Nunes, L. Mixed-effects generalized height-diameter model: A tool for forestry management of young sweet chestnut stands. *For. Ecol. Manage.* **514**, 120209 (2022).
19. Chave, J. *et al.* Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **145**, 87–99 (2005).
20. Henry, M. *et al.* Estimating Tree Biomass of Sub-Saharan African Forests: a Review of Available Allometric Equations. *Silva Fenn.* **45**, (2011).
21. Ryan, C. M., Williams, M. & Grace, J. Above- and Belowground Carbon Stocks in a Miombo Woodland Landscape of Mozambique. *Biotropica* 1–10 (2010).
22. Guedes, B. S., Siteo, A. A. & Olsson, B. A. Allometric models for managing lowland miombo woodlands of the Beira corridor in Mozambique. *Glob. Ecol. Conserv.* **13**, e00374 (2018).
23. Mate, R., Johansson, T. & Siteo, A. Biomass equations for tropical forest tree species in mozambique. *Forests* **5**, 535–556 (2014).
24. Magalhães, T. M. & Seifert, T. Tree component biomass expansion factors and root-to-shoot ratio of Lebombo ironwood: measurement uncertainty. *Carbon Balance Manag.* **10**, 1–14 (2015).
25. Magalhães, T. M. & Seifert, T. Biomass Modelling of *Androstachys johnsonii* Prain : A Comparison of Three Methods to Enforce Additivity. *Int. J. For. Res.* **2015**, 17 (2015).
26. Ketterings, Q. M., Coe, R., Van Noordwijk, M., Ambagau, Y. & Palm, C. A. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *For. Ecol. Manage.* **146**, 199–209 (2001).
27. Fayolle, A., Doucet, J. L., Gillet, J. F., Bourland, N. & Lejeune, P. Tree allometry in Central Africa: Testing the validity of pantropical multi-species allometric equations for estimating biomass and carbon stocks. *For. Ecol. Manage.* **305**, 29–37 (2013).
28. Daba, D. E. & Soromessa, T. The accuracy of species-specific allometric equations for estimating aboveground biomass in tropical moist montane forests: Case study of *Albizia grandibracteata* and *Trichilia dregeana*. *Carbon Balance Manag.* **14**, 1–13 (2019).
29. Chave, J. *et al.* Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **145**, 87–99 (2005).
30. Pati, P. K., Kaushik, P., Khan, M. L. & Khare, P. K. Allometric equations for biomass and carbon stock estimation of small diameter woody species from tropical dry deciduous forests : Support to REDD +. *Trees, For. People* **9**, 100289 (2022).
31. Smith, H. E. *et al.* Urban energy transitions and rural income generation: Sustainable opportunities for rural development through charcoal production. *World Dev.* **113**, (2019).
32. Lisboa, S. N. *et al.* Effect of charcoal production and woodland type on soil organic carbon and total nitrogen in drylands of southern Mozambique. *For. Ecol. Manage.* **457**, (2020).
33. van Voorthuizen, E. G. The Mopane Tree. *Botsw. Notes Rec.* **8**, 223–230 (1976).
34. Henning, A. C. A STUDY OF EDAPHIC FACTORS INFLUENCING THE GROWTH OF COLOPUOSPERMUM MOPANE (ex BENTH) KIRK ex J, LEONARD. (University of the Witwatersrand Johannesburg, 1976).
35. Stevens, N., Swemmer, A. M., Ezzy, L. & Erasmus, B. F. N. Investigating potential determinants of the distribution limits of a savanna woody plant: *Colophospermum mopane*. *J. Veg. Sci.* **25**, 363–373 (2014).
36. Makhado, R. A., Mapaure, I., Potgieter, M. J., Luus-Powell, W. J. & Saidi, A. T. Factors influencing the adaptaton and distributon of *Colophospermum mopane* in southern Africa’s mopane savannas - A review. *Bothalia* **44**, 1–9 (2014).

37. Ribeiro, N. S., Chirwa, P. W., Katerere, Y. & Grundy, I. M. *Miombo Woodlands in a changing environment: Securing the resilience and sustainability of people and woodlands. Miombo Woodlands in a Changing Environment: Securing the Resilience and Sustainability of People and Woodlands* (2020). doi:10.1007/978-3-030-50104-4.
38. Government of Mozambique. Republic of Mozambique Ministry of Land, E. and R. D. M. *Mozambique's Forest Reference Emission Level for Reducing Emissions from Deforestation in Natural Forests*. (2018).
39. Picard N.; Saint-André L.; Henry M. *Manual for building tree volume and biomass allometric equations: From field measurement to prediction*. Cirad (2012).
40. Jackson, R. B. *et al.* A global analysis of root distributions for terrestrial biomes. *Oecologia* **108**, 389–411 (1996).
41. Mugasha, W. A. *et al.* Allometric models for prediction of above- and belowground biomass of trees in the miombo woodlands of Tanzania. *For. Ecol. Manage.* **310**, 87–101 (2013).
42. Magalhães, T. M. Allometric equations for estimating belowground biomass of *Androstachys johnsonii* Prain. *Carbon Balance Manag.* **10**, (2015).
43. Mehtatalo, M. L. *Package 'lmfor'*. (2022).
44. Cienciala, E., Černý, M., Tatarinov, F., Apltauer, J. & Exnerová, Z. Biomass functions applicable to Scots pine. *Trees - Struct. Funct.* **20**, 483–495 (2006).
45. Sileshi, G. W. A critical review of forest biomass estimation models, common mistakes and corrective measures. *For. Ecol. Manage.* **329**, 237–254 (2014).
46. Baty, F. *et al.* A Toolbox for Nonlinear Regression in R : The Package nlstools. *J. Stat. Softw.* **66**, 1–21 (2015).
47. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team. 'nlme' Title Linear and Nonlinear Mixed Effects Models. *R package version 3.1-164*, <https://CRAN.R-project.org/package=nlme>. (2022).
48. Pena, E. A. & Slate, E. Global Validation of Linear Models Assumptions. 16 (2022).
49. Ripley, B., Bates, D. M., Hornik, K., Gebhardt, A. & Firth, D. Package 'MASS'. 165 (2024).
50. Wagenmakers, E. J. & Farrell, S. AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–196 (2004).
51. Magalhães, T. M., Cossa, V. N., Guedes, B. S. & Fanheiro, A. S. M. Species-specific biomass allometric models and expansion factors for indigenous and planted forests of the Mozambique highlands. *J. For. Res.* **32**, 1047–1065 (2021).
52. Mascaro, J., Litton, C. M., Hughes, R. F., Uowolo, A. & Schnitzer, S. A. Minimizing Bias in Biomass Allometry: Model Selection and Log-Transformation of Data. *Biotropica* **43**, 649–653 (2011).
53. Soares, P., Tomé, M., Skovsgaard, J. P. & Vanclay, J. K. Evaluating a growth model for forest management using continuous forest inventory data. *For. Ecol. Manage.* **71**, 251–265 (1995).
54. Djomo, A. N., Ibrahima, A., Saborowski, J. & Gravenhorst, G. Allometric equations for biomass estimations in Cameroon and pan moist tropical equations including biomass data from Africa. *For. Ecol. Manage.* **260**, 1873–1885 (2010).
55. Feldpausch, T. R. *et al.* Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* **9**, 3381–3403 (2012).
56. Istrefi, E., Toromani, E., Çollaku, N. & Thaçi, B. Allometric biomass equations for young trees of four broadleaved species in Albania. *New Zeal. J. For. Sci.* **49**, 1–14 (2019).
57. Traerup, S. L. M. Informal networks and resilience to climate change impacts : A collective approach to index insurance. *Glob. Environ. Chang.* **22**, 255–267 (2012).
58. Woollen, E. *et al.* Charcoal production in the Mopane woodlands of Mozambique : what are the trade-offs with other ecosystem services ? *Philos. Trans. R. Soc. B* **371**, (2016).
59. Mate, R., Johansson, T. & Siteo, A. Biomass equations for tropical forest tree species in mozambique. *Forests* **5**, 535–556 (2014).
60. Ramírez-Ramírez, G., Ramírez y Avilés, L., Solorio-Sánchez, F. J., Navarro-Alberto, J. A. & Dupuy-Rada, J. M. Shifts in tree allometry in a tropical dry forest: implications for above-ground biomass estimation. *Bot. Sci.* **97**, 167–179 (2019).
61. Magalhães, T. M., Seifert, T., Florestal, D. D. E. & Mondlane, U. E. Estimation of Tree Biomass , Carbon Stocks , and Error Propagation in Mecrusse Woodlands. 471–488 (2015).

62. Sedano, F. *et al.* Monitoring intra and inter annual dynamics of forest degradation from charcoal production in Southern Africa with Sentinel – 2 imagery. *Int. J. Appl. Earth Obs. Geoinf.* **92**, 102184 (2020).
63. Silva, J. A. & Sedano, F. Lessons from Mozambique ' s Charcoal Trade Africa ' s Urban Revolu3on & The Future of Forests : Tete , Mozambique. (2016).
64. Lisboa, S. N., Mate, R. & Siteo, A. Applying the ICAT Sustainable Development Methodology to Assess the Impacts of Promoting a Greater Sustainability of the Charcoal Value Chain in Mozambique. *Sustain.* **12**, 1–28 (2020).

Acknowledgements

This work was financially supported by the Oliver Tambo Africa Research Chair Initiative supported by the South African National Research Foundation (NRF), the South African Department of Science and Innovation (DSI), the International Development Research Center of Canada (IDRC), the Oliver & Adelaide Tambo Foundation (OATF), and the National Research Fund of Mozambique (FNI). The views expressed in this article do not necessarily represent those of the Oliver Tambo Africa Research Chairs Initiative, its partners, and their Directorates.

Author contributions

S.N.L.: Conceptualization, Methodology, Investigation, analysis, Writing – original draft, Writing & editing. S.M.: Data, Reviewing & editing, A.S.: Fund acquisition, Supervision, Reviewing & editing.

Competing interests

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterialAllometricEquation.docx](#)