Biomass and Bioenergy 130 (2019) 105353

Contents lists available at ScienceDirect

Biomass and Bioenergy

journal homepage: www.elsevier.com/locate/biombioe

Research paper

Developing biomass estimation models for above-ground compartments in *Eucalyptus dunnii* and *Corymbia citriodora* plantations

Lina Garcia_Florez*, Jerome K. Vanclay, Kevin Glencross, J. Doland Nichols

Forest Research Centre, School of Environment, Science and Engineering, Southern Cross University, Lismore, 2480, Australia

ARTICLE INFO

Keywords:
Destructive sampling
Regression analysis
Model quality
Tree biomass
Hardwood species

ABSTRACT

Biomass has been widely studied in terms of ecosystem ecology, timber production profitability, bioenergy (biofuels) and greenhouse gas emission reduction mechanisms. However, uncertainty in biomass estimation is still a current concern. In this study, direct and indirect methods were used to develop species-specific biomass estimation models (BEMs) for stem, bark, branch and crown compartments in 16-year old plantations of *Eucalyptus dunnii* and *Corymbia citriodora*. A total of 93 trees were destructively sampled. An analysis of covariance (ANCOVA) assessed the effect of species on biomass prediction. Our results indicated that equations developed by using parameters or predictors such as diameter at breast height (DBH), height (H), wood density (*p*) and branch diameter were generally significant ($p < 0.05$) and their regression lines fitted well the data ($R^2 > 0.84$). After a rigorous process that included testing hypotheses, checking diagnostic statistics, assessing model coefficients and model functionality, the most suitable stem BEMs corresponded to those ones derived from the compound variable DBH^2Hp. The most reliable branch and crown BEMs used DBH and branch diameter respectively as single variable (simple linear models). Bark BEMs differ between species as DBH was the best predictor for *E. dunnii* whilst the compound variable DBH H predicted better for *C. citriodora*. The BEMs with multiple predictors, and in particular polynomial models, produced wider confidence intervals, unreliable coefficients, multicollinearity and higher proportion of outliers and leverage points. In conclusion, appropriate model diagnosis can reduce pitfalls and ensure selection of valid BEMs.

1. Introduction

Biomass, and particularly above-ground biomass (AGB), is one of the most measured variables in vegetation systems, widely used to understand ecological and management processes in ecosystems, and to quantify timber products in the forestry industry [1,2]. Biomass has been also studied to derive estimates of quantities of biofuels and biogas as part of renewable energy alternatives [3], and to quantify carbon stocks under the greenhouse gas (GHG) emission reduction mechanisms [4–6]. Tree biomass estimation includes direct and indirect methods. Direct methods can involve field measurements and destructive procedures whilst indirect methods include statistical modelling (regression analyses) and remote sensing techniques [7–11]. Undertaking field measurements in conjunction with destructive sampling is a crucial step to yield precise data to develop biomass estimation models (BEMs) [7,15,16]. Destructive sampling consists of harvesting trees, and separating and weighing their compartments, being an expensive and time consuming procedure. However, it has been considered the most accurate method of estimating biomass [3,12–14]. Data produced from destructive sampling is used in regression analyses to develop biomass estimation models (BEMs) [15,16] and is also widely used to validate results from existing BEMs and biomass estimates obtained from remote sensing analyses [12,23].

The use of BEMs is the basis for estimating tree biomass, by using easy-to-measure parameters (e.g. DBH, height) [13,17,18]. These BEMs are mathematical expressions that quantify proportionality relationships between dimensions of tree characteristics as responses to ontogenic development [19–21]. Developing BEMs becomes a key process to predict biomass of individual trees sampled in forest inventory plots [21]. Tree biomass can be used to obtain plot level estimates, then to extrapolate mean values to a whole population [6,12]. These BEMs can be 1) generalised or 2) specific equations (species-, site-, age-specific) [13]. The use of generalised equations on different species and ecosystem types can be problematic, yielding around 15–25% error on biomass estimations [1,6,17,21,22]. Therefore, species-specific equations are preferable to produce reliable estimates [5,12,14,22,23].

Reducing uncertainty and bias during the development of BEMs can improve the accuracy of biomass estimates [24,25]. Undertaking model

* Corresponding author. Southern Cross University, PO_Box 157, Lismore, 2480, Australia.
E-mail address: consultoriasforestales@yahoo.es (L. Garcia_Florez).

https://doi.org/10.1016/j.biombioe.2019.105353
Received 10 February 2018; Received in revised form 31 May 2019; Accepted 13 September 2019
Available online 08 October 2019
0961-9534/

diagnosis and the assessment of appropriate statistics to select suitable equations are required to reduce sources of errors, and therefore, produce statistically reliable and biologically realistic models [26]. Generalised as well as species-specific BEMs have been developed for some hardwood species belonging to *Eucalyptus* and *Corymbia* genera [16,18,22,27,28], but not to the same extent as for softwood species. In addition, existing equations commonly focus on the prediction of stem biomass given that this is the tree compartment that is merchantable, and little information is available for the remaining compartments, and therefore total AGB [3,17,29]. The stem can represent a large proportion of the total tree biomass, however, the inclusion of other components such as bark and in particular crown, allows assessing the manner in which trees allocate biomass to each compartment. The biomass partitioning reflects the distribution of the net primary productivity (NPP) in response to the different species genetic characteristics, physiological processes and environmental conditions [12,30].

In Australia, the hardwood plantation state is dominated by several eucalypts species [31], and two of them, Dunn's White Gum (*Eucalyptus dunnii* Maiden) and Spotted Gum (*Corymbia citriodora* subsp. *Variegata* (F.Muell.) A.R.Bean & M.W.McDonald, hereafter referred to as *C. citriodora*) are of particular interest for this research. In subtropical regions of Australia, *E. dunnii* has been primarily planted for sawlog and pulp production whilst *C. citriodora* has been widely used for structural timber [31]. Although *E. dunnii* is naturally distributed in Australia [32], this species has been successfully grown in China, Africa and South America (Brazil and Argentina) to produce timber suitable for a range of uses [33]. With a wider distribution range, *C. citriodora* occurs naturally in Australia and New Zealand, and has also been established for timber and essential oil production in the Mediterranean area, southern China, southern Africa, South-East Asia, South America and North America [32,34,35]. Research is increasing in fields such as the production of biofuels and electricity from thinning residues (low-value and defective stems) [36], and the estimation of biomass and sequestered carbon [6] for these species and other subtropical eucalypts.

This research was carried out in 16-year old hardwood plantations of *E. dunnii* and *C. citriodora* established in 2000/2001 in north-eastern New South Wales (NSW), Australia. In the study area, thinning practices undertaken at eight years after planting to reduce competition between trees and increase yield of high-quality timber products, resulted in stands of residual stocking densities of 300 (heavily thinned) and 600 stems ha^{-1} (moderately thinned). Unthinned areas (1200 stems ha^{-1}) represented the “Control”. As no BEMs are available for these two hardwood species, this study aimed to 1) develop equations from harvested data to predict biomass in stem, bark, branch and crown compartments and 2) select the most suitable BEMs for each species and tree compartments through a rigorous process of testing hypotheses, checking diagnostic statistics, confirming reliability of model coefficients, and assessing model functionality.

2. Methods

2.1. Study area

The study site comprised of even-aged, monospecific planted stands of *E. dunnii* and *C. citriodora* located 50 km south-west of Lismore in north-eastern NSW, Australia, latitude 29°03′ South, longitude 153°05’ East (Fig. 1). Climate statistics from 1995 to 2017 from the weather station Casino Airport AWS located at 17 km from the study site indicated 1037 mm average annual rainfall, and a maximum of 143 mm per month (late summer) and minimum of 32 mm per month (late winter). Annual mean minimum and maximum temperatures are 13.3 °C (6.6 °C in the coldest month in July) and 26 °C (30.3 °C in the warmest month in January) respectively [37]. The planted area is distributed in land between 37 and 79 m above sea level, with parent material dominated by lithic and quartz sandstones, siltstones and conglomerates [38]. Soils are Ultisols (Kurosols based on the Australian Soil Classification) characterised by poor drainage, strongly leached, relatively acid (pH $<$ 5.5) and low fertility [39].

2.2. Experiment site

Seedlings of *E. dunnii* and *C. citriodora* were established during 2000/2001 at a planting density of approximately 1200 trees ha^{-1}. In 2008, at 8 years after establishment of plantation, a thinning experiment was implemented by extracting small-sized and defective trees to reduce competition, and increase the growth and timber quality of residual trees [40]. The thinning experiment followed a randomised complete block design, comprising a total of 24 plots of 750 m^2 (25 × 30 m), 12 plots located in *E. dunnii* and 12 plots in *C. citriodora* stands. Each 12-plot set per species covered three thinning regimes (four replications per regime) of designated stocking densities: 1) heavily thinned to a residual stocking density of 300 stems ha^{-1}, 2) moderately thinned to 600 stems ha^{-1}, and 3) unthinned areas (1200 stems ha^{-1}). A two tree-row buffer treated under same thinning regime separated each plot from the other.

2.3. Field measurements

In 2015, stem and crown variables were measured for 1031 standing trees in 24 plots. Tree measurements included DBH (cm), total height (m) and crown depth (lowest living branch to tip of crown, in m) (Table 1). All trees inside each plot were stratified into three to four diameter classes, and one tree per class was randomly selected to represent the diameter range from small to large trees. A total of 93 trees, taken from non-border rows to avoid “edge effect”, were selected for destructive sampling. Each tree was cut at ground level, felled and divided into stem and crown sections. Total length from stem base to crown tip or height (m), and crown depth (m) were measured. Then, each stem was divided into sections following specified lengths from tree base towards crown, at 0.3, 1.3, 2.5 m, and each 2-m length afterwards. Last section was determined once upper diameter reached $<$ 10 cm. From each section, a sample disc (2–3 cm thick) was collected from the bottom, and its diameter was measured. Thirteen to fifteen discs were collected per tree. Disc diameter and DBH were measured using diameter tapes.

From these 93 trees, 50 trees were selected to measure and sample their crown elements, where all branches were numbered and each branch diameter (cm) measured at its base. For the purpose of this study, a branch is considered to be a part of the tree which grows out from the trunk and that is composed by leaves (and twigs) and woody material. In total, 1135 branches were measured from the 50 trees using a diameter tape. Once branch diameter measurements were completed for each tree, the crown was divided into upper and lower sections, and one to two branches were collected from each section. A total of 172 branches were collected, and leaves and twigs were separated from the woody material. The woody material was cut into pieces and classified into coarse ($\geq$ 2 cm diameter) and fine woody material ($<$ 2 cm diameter). Fresh weight of each branch component was recorded on-site using 50-kg spring weighing scales (0.1-kg precision).

2.4. Laboratory procedures and biomass calculations from collected material

Stem: total fresh weight of discs was recorded using an electronic digital scale (0.1 g precision). Bark was removed from discs and remaining wood was weighed. Under bark diameter (UBD) of each disc was measured. Each stem section volume was calculated by the Smalian's formulae (sections assumed to have a cylindrical shape) using the UBDs of the bottom and top discs, and each section length. Upper stem section volume was calculated by the cone formulae. One wedge-shaped piece was obtained from each disc, saturated with water and its volume measured by the water displacement technique. Wedges were

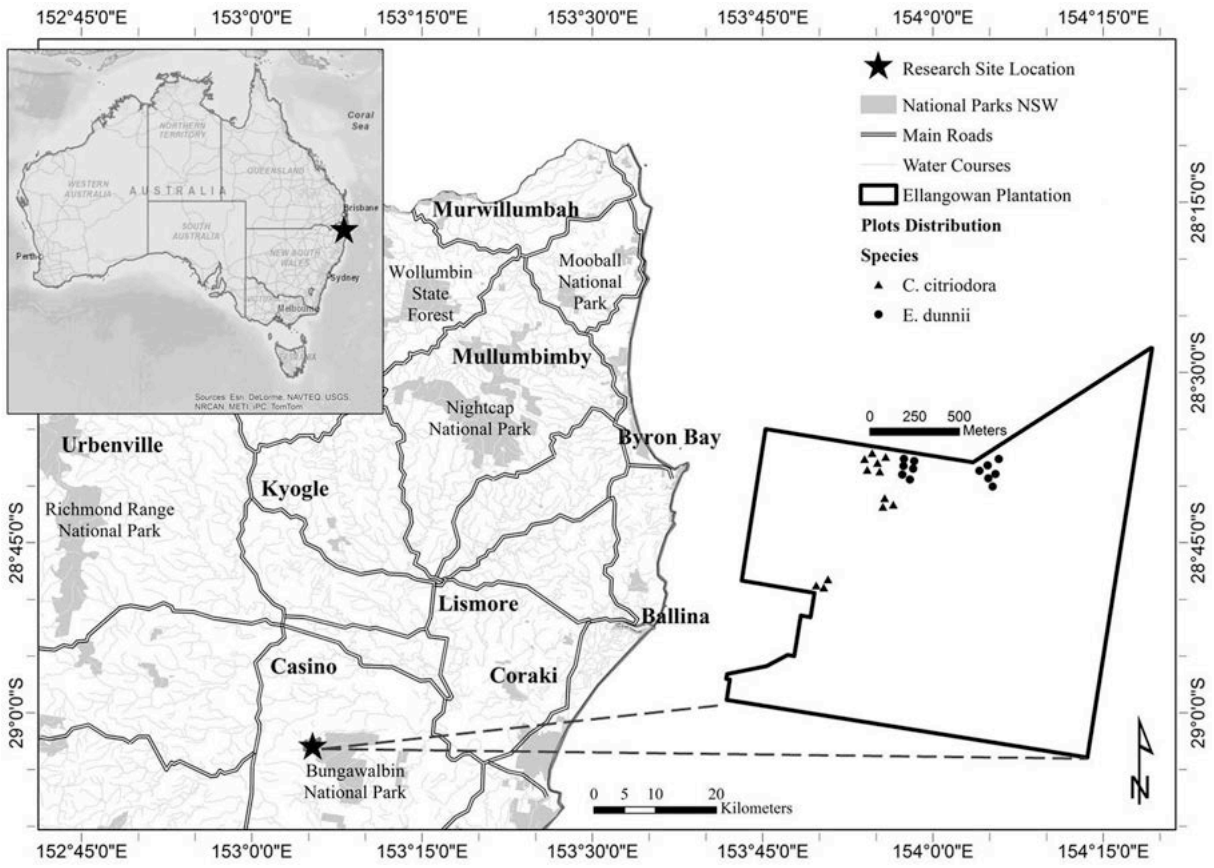

Fig. 1. Location of the hardwood plantation area in north-eastern NSW Australia and the 24 study plots in *E. dunnii* and *C. citriodora* stands.

oven-dried (103 °C) until constant dry weight achieved. Wedge basic density (g cm^{-3}) was calculated by dividing its dry weight by its volume. Wood density of each section was calculated as the average density between the bottom and top discs. The average of the wood densities of all stem sections of the tree was also calculated to obtain the tree wood density. Dry weight of each stem section was calculated by multiplying its volume by its wood density, and total stem biomass per tree was calculated by summing the dry weights of all its sections. **Bark:** bark material removed from each disc was weighed, oven-dried (80 °C) to a constant weight and then re-weighed. The dry weight of each bottom and top disc was divided by the respective disc thickness to calculate its dry weight per cm, and then averaged. The dry weight of bark for each section was calculated by multiplying the average bark dry weight by the respective section length. Total bark biomass per tree was calculated by summing the dry weights of all sections.

Branch: for each collected branch component such as leaves (and twigs), coarse woody material and fine woody material were separately fresh weighed, and from each component a 200-g sample was oven-dried (80 °C) until constant weight was achieved. The percentage of dry content for each sample was calculated by dividing the sample dry weight by its fresh weight, and was used to calculate the total dry weight of each branch component. Total biomass of each collected branch was calculated by summing the dry weights of its leaves (and twigs), coarse woody material and fine woody material. **Crown:** biomass calculations of all collected branches as well as their branch diameter measurements were used as data input to develop equations or branch BEMs to estimate the biomass of the remaining branches (only-measured branches). Total crown biomass per tree was calculated by summing the measured biomass of the collected branches and the modelled biomass of the only-measured branches in each tree.

Table 1
Summary statistics of the trees that were measured in the stands in the 16-year old plantations of *Eucalyptus dunnii* and *Corymbia citriodora*.

Species/Thinning regime	Tree density (stems ha^{-1})	Number of trees measured per plot	DBH (cm)			Height (m)			Average crown depth (m)	Average wood density[a] (kg m^{-3})
			Min	Mean	Max	Min	Mean	Max		
E. dunnii										
Heavily thinned	300	91	12.7	23.5	30.9	10.6	22.8	26.4	11.2	0.56
Moderately thinned	543	167	11.0	21.2	28.8	14.6	22.3	28.2	8.9	0.56
Unthinned	883	285	6.8	16.7	27.6	4.8	19.4	26.0	8.3	0.57
C. citriodora										
Heavily thinned	260	82	17.2	25.3	35.4	16.4	23.9	30.4	16.5	0.68
Moderately thinned	510	156	9.8	20.4	34.8	9.6	21.2	29.2	11.2	0.65
Unthinned	940	298	3.0	15.4	31.7	2.8	16.3	27.4	9.6	0.69

[a] Data obtained from the destructive sampling of 93 trees.

2.5. Data analysis

2.5.1. Data exploration, variable selection and model fitting

Identification of potential BEMs included exploratory analysis of the relationships between response and predictor variables. Response variables were stem, bark, branch and crown biomass measured from each harvested tree. Predictor variables were the respective DBH, height, wood density, branch diameter and crown depth measurements from each harvested tree. Data exploration consisted in identifying the presence of linear function between variables by plotting each response variable ("y" axis) against a predictor, either single or combined variable ("x" axis). Linear least-square regression lines were added to the graph to assess the variation between regression lines and distribution of plotted points to identify multiplicative, heteroscedastic and non-normal errors. If these errors were common during the data exploration, then log-linear regressions were used for predicting biomass. Logarithmic transformations were applied to both response and predictor variables to stabilise residual variance (additive residuals), to ensure symmetry of residuals (normal distribution), and to render linear relationships between variables. Normal probability and Kernel density distributions (which reveals multimodal distribution) of response variables were also calculated to confirm the necessity of applying logarithm transformations.

Selection of predictors to be included in BEMs was achieved by calculating their significance (F-test, $p = 0.05$) at predicting biomass. Only significant predictors were used to develop BEMs by fitting linear, multiple and polynomial regressions commonly cited in the forestry literature for tree biomass estimations (Table 2) [4,7,8,14,41]. The first type model to determine relationships between response and predictor variables, corresponds to the power-law model (Equation Ia) where logarithm transformation renders the linear relation (Equation Ib). In the logarithm form of the power-law model, "Y" is the response variable or biomass, "X" is the predictor variable, "a" is the intercept, and "b" is the slope. The power-law is the most popular and commonly used model as the power relationship is the principle for the allometric scaling theory which applies to several tree characteristics [25,42]. Allometry consists of a relative change in the dimensions of one tree characteristic that results in a proportional relative change in the dimensions of other tree characteristic [25]. The preference for the use of power-law models in studies developing BEMs also relies on the simplicity for their replication and validation, the higher certainty and reliability at estimating their 'a" and "b" parameters, and their easy interpretation in a biological context [26].

A second type of model corresponds to linear models with multiple predictors (more than two), where a typical example is Equation II, where "a", "b", "c", "d" are regression coefficients, and "X_1", "X_2", …"X_n" are predictors. A third type of model includes those with polynomial terms (e.g. quadratic) often framed as Equation III. A last type of models (Type IV) corresponds to those with compound predictors such as Equation IV, where "$X_1X_2X_3$" is a single term resulting from the multiplicative effect of several variables.

$$Y = aX^b \quad \text{(Equation Ia)}$$

$$\ln Y = \ln(a) + b(\ln X) \quad \text{(Equation Ib)}$$

$$\ln Y = \ln(a) + b(\ln X_1) + c(\ln X_2) + d(\ln X_3) \quad \text{(Equation II)}$$

$$\ln Y = \ln(a) + b(\ln X_1) + c(\ln X_2)^2 \quad \text{(Equation III)}$$

$$\ln Y = \ln(a) + b(\ln(X_1X_2X_3)) \quad \text{(Equation IV)}$$

Stem BEMs were fitted by using DBH, height (H) and wood density (p) data from 80 trees. Bark BEMs also used DBH and height (H) data from 80 trees. Branch BEMs were developed from 150 branches of 50 trees by using branch diameter as predictor. Crown BEMs were developed once the total branch biomass in each of the 50 trees was estimated, and by using DBH and H as predictors. Model fitting consisted in estimating regression coefficients (a, b, c …), residual standard error (RSE), coefficient of determination (R^2) and p-value for all candidate

Table 2
Formulae of the linear, multiple, polynomial and compound predictor models used to develop the biomass estimation models (BEMs) for the respective tree compartments.

Compartment	Model type	Equation and predictor variables
Stem	Type I	Eq (1).ln(B) = ln(a) + $b$ln (DBH)
		Eq (2).ln(B) = ln(a) + $b$ln(H)
		Eq (3).ln(B) = ln(a) + $b$ln(ρ)
	Type II	Eq (4).ln(B) = ln(a) + $b$ln (DBH3) + $c$ln(H)
		Eq (5).ln(B) = ln(a) + $b$ln (DBH) + $c$ln(ρ)
		Eq (6).ln(B) = ln(a) + $b$ln(H) + $c$ln(ρ)
		Eq (7).ln(B) = ln(a) + $b$ln (DBH) + $c$ln(H) + $d$ln(ρ)
	Type III	Eq (8).ln(B) = ln(a) + $b$ln (DBH) + $c$ln $(DBH)^2$
	Type IV	Eq (9).ln(B) = ln(a) + $b$ln (DBH^2H)
		Eq (10).ln(B) = ln(a) + $b$ln $(DBH^2Hρ)$
Bark	Type I	Eq (11).ln(B) = ln(a) + $b$ln (DBH)
	Type II	Eq (12).ln(B) = ln(a) + $b$ln (DBH) + $c$ln(H)
	Type IV	Eq (13).ln(B) = ln(a) + $b$ln (DBH^2H)
Branch	Type I	Eq (14).ln(B) = ln(a) + $b$ln (BD)
	Type III	Eq (15).ln(B) = ln(a) + $b$ln (BD) + $c$ln $(BD)^2$
Crown	Type I	Eq (16).ln(B) = ln(a) + $b$ln (DBH)
	Type II	Eq (17).ln(B) = ln(a) + $b$ln (DBH) + $c$ln (CD)
		Eq (18).ln(B) = ln(a) + $b$ln (DBH) + $c$ln(H)
		Eq (19).ln(B) = ln(a) + $b$ln (DBH) + $c$ln (CD) + $d$ln(H)

Where.
B is the biomass of each compartment (kg)
a, b, c, d are regression coefficients
DBH is tree diameter at breast height (cm)
H is tree height (m)
ρ is tree wood density (g cm^{-3})
BD is branch diameter (cm)
CD is crown depth (m)

equations. Data exploration, variable selection, model fitting as well as all statistical analyses hereafter described, were undertaken by using R software 3.2.5 (R Foundation for Statistical Computing).

As stem, bark, branch and crown biomass data was collected from a population characterised by the two species *E. dunnii* and *C. citriodora*, an analysis of covariance (ANCOVA) was undertaken to assess the effect of species on the predictive ability of each candidate model. ANCOVA analysis split the relationship between predictors (or co-variates, i.e. DBH, H, *p*, BD) and response variables (stem, bark, branch and crown biomass) into one regression line for each species, and assessed differences in Y-intercepts and slopes among lines [25]. The interaction between predictor and response variables was tested (at $p = 0.05$) whilst controlling this species effect. The assessment of the interaction to determine whether the relationship between biomass production and predictors was comparable between species was undertaken through an analysis of variance (ANOVA).

2.5.2. Model assessment and selection

Candidate BEMs for each compartment were derived from several model types and different combination of predictors. To better understand factors affecting the predictive ability of BEMs, this study included the assessment of additional statistics as recommended by Sileshi [26]. The process to appropriately select the best BEM included: 1) testing hypotheses, 2) checking diagnostic statistics, 3) assessing reliability of model coefficients, and 4) assessing model functionality. For each BEM, two main **hypotheses** were assessed: *i*) linearity, indicating systematic change of response variable with predictors, and *ii*) errors or residuals being *a*) independent, *b*) normally distributed and *c*) of constant variance. Satisfaction of hypotheses *i* and *iia* were verified as per the process described in 2.5.1 (by plotting response variables against predictors once log-transformed). Hypothesis *iib* was checked through the residuals distribution by plotting standardized residuals versus theoretical quantiles (known as the Quantile-Quantile graph). Hypothesis *iic*, also known as homoscedasticity check, was visually explored by plotting residuals in function to fitted (predicted) biomass. Presence of multimodal distribution of residuals was checked by graphical exploration of normal and kernel distributions.

Diagnostic statistics check consisted in assessing regression coefficients, residual standard errors (RSE), coefficients of determination (R^2) and *p*-values. **Reliability of the coefficients** was assessed throughout the 95% confidence intervals (95%CIs), percent relative standard error (PRSE), variance inflation factor (VIF) and partial least square regressions (PLS regression) (see formula in Appendix A). Confidence intervals are often used to identify the coefficient reliability to reject or accept hypothesis, particularly when calculated from small samples. The PRSE indicates whether the standard error is small or large relative to the coefficient, and therefore reliable or not. Collinearity of predictors, which can have significant impacts on coefficients such as inflated standard errors and erratic signs, was assessed through VIF calculation. The percentage of variance explained by predictors, also known as the proportion of their contribution to the model fit was calculated by PLS regressions. In this study, and following guidance in Sileshi [26], unreliable coefficients were indicated by unrealistic 95%CIs, coefficients with PRSE > 30% and VIF > 5, and PLS regression with small (< 10%) to nil (0%) contribution.

Model functionality consisted in checking statistics such as leverage points, outliers, and prediction errors. Leverage points are values of predictors that are distant from the main data arrangement, whilst outliers correspond to values with same distant behaviour belonging to response variable. Leverage points and outliers can negatively affect the certainty of coefficients by increasing their standard errors, and therefore reducing the accuracy of fitted BEMs. Leverage points were identified from the distribution of studentized residuals, where values exceeding −2 or +2 indicated severe heteroscedasticity. Outliers corresponded to predicted observations with Cook's Distance values greater than 0.5. Accurate and functional BEMs included models with less than 10–15% leverage points and outliers. Prediction errors were calculated by using the mean absolute percentage error (MAPE), where BEMs with MAPE > 10% were considered to be non-functional and unreliable. All BEMs that did not fit with the hypotheses testing, diagnostic statistics and model functionality were classified as unreliable and excluded from further analyses. The remaining models were included in the selection of final BEMs by comparing their Akaike Information Criterion (AIC) and the Akaike weight (AICw) values [12] where suitable BEMs correspond to those with lower AIC value, and most importantly, the highest AICw [26,43].

2.5.3. Model validation

Assessment of the accuracy of the selected BEMs was undertaken by applying a K-Fold cross-validation. This method can be applied to independent datasets, and consists of dividing the original dataset into K subsets or folds of “training” and “test” data. Each fold is of approximately equal size, where one of the folds is used for validation (training data) for a model that is to be fitted on the remaining folds (test data). As number of folds is important to produce balanced proportion between bias and variance, a 5-fold cross-validation was applied for stem, bark and branch BEMs. Conversely, because of a smaller sample size for the crown compartment compared to the others, a 3-fold cross-validation was applied for these BEMs. The 5-fold cross-validation divided the data in five subsets, where four of them were used to “train” the model and predict biomass estimates, which were then compared with the measured biomass from the 5th fold. This procedure was repeated five times, leaving out each of the folds in turn, and was performed by using the cv. lm () function (DAAG package) in R Software. Similar procedure applied for the 3-fold cross-validation. For each fold, variation in R^2 and MAPE as compared, and the significant difference of the mean of weighted residuals between on-ground measured and estimated biomass was tested by applying a two-tailed *t*-test ($p = 0.05$).

3. Results

3.1. Data exploration, variable selection and model fitting

Positive strong relationships were observed between single predictors DBH and H, and the corresponding stem and bark biomass estimates, whilst ρ by itself did not show a relationship. Relationships were stronger when using combined variables such as DBH^2H and $DBH^2H\rho$. Similarly, strong relationships were also found between branch biomass and BD as well as between crown biomass and DBH and H. Initial exploration showed that our data had no linear relationship as the variance of biomass for each compartment increased as the dimensions of predictors did, suggesting multiplicative, heteroscedastic and log-normal error. However, linear relationships and constant variance of residuals were observed after applying logarithm transformations to both response and predictor variables (graphs not shown). During model fitting process, all selected variables, except ρ itself ($p = 0.909$ for *E. dunnii* and $p = 0.517$ for *C. citriodora*), were significant for F-test ($p < 0.05$), and the mean of weighted residuals were not significantly different from zero. For each species, nine equations were developed for stem biomass, three for bark biomass and two for branch biomass. Branch biomass estimates were used to build four crown biomass equations. In total, 19 BEMs per species were developed.

ANCOVA analysis indicated that species had a significant effect ($p < 0.001$) on stem, bark, branch, and crown biomass predictions as there were significant differences in intercepts (coefficient *a*) between the regression lines of each BEM (see Appendix B). Most predictors, except for BD^2 in Equation (15), and CD and H in equations (17)–(19), had an effect on biomass ($p < 0.05$) indicating the mean slopes (coefficients) associated with these predictors were significantly different from “0”. ANOVA results clearly demonstrated significant interactions ($p < 0.001$) between species and the predictors $DBH^2H\rho$, DBH and BD, and DBH and H used in equations (1), (4), (5), (7)–(10),

Table 3
Residual standard errors (RSEs) and regression coefficients for the biomass estimations models (BEMs) developed for *E. dunnii*. *P*-values were < 0.001 for all equations.

Compartment/Equation	Sample Size	RSE	R2	Coefficients			
				ln*a*	*b*	*c*	*d*
Stem	40						
Eq (1a).		0.142	0.967	−1.70338	2.243379		
Eq (2a).		0.250	0.899	−4.91323	3.195144		
Eq (4a).		0.141	0.969	−2.1917	1.9763	0.4144	
Eq (5a).		0.137	0.971	−1.2054	2.2483	0.8922	
Eq (6a).		0.252	0.900	−4.6186	3.1968	0.5218	
Eq (7a).		0.136	0.972	−1.6654	2.008	0.3726	0.8548
Eq (8a).		0.143	0.968	−2.7004	2.9892	−0.136	
Eq (9a).		0.143	0.968	−2.67895	0.8457		
Eq (10a).		0.137	0.970	−2.20872	0.84761		
Bark	40						
Eq (11a).		0.188	0.937	−3.16931	2.10418		
Eq (12a).		0.190	0.937	−3.3901	1.9834	0.1873	
Eq (13a).		0.193	0.934	−4.07318	0.79201		
Branch	75						
Eq (14a).		0.187	0.957	−2.61184	2.34891		
Eq (15a).		0.188	0.957	−2.68083	2.53037	−0.09965	
Crown	24						
Eq (16a).		0.223	0.879	−1.5309	1.5563		
Eq (17a).		0.208	0.900	−1.9014	1.4778	0.2744	
Eq (18a).		0.227	0.881	−1.2258	1.7523	−0.2842	
Eq (19a).		0.210	0.903	−1.5364	1.716	0.281	−0.3484

and (16)-(19), indicating different slopes between the regression lines of each BEM. Conversely, interactions were not significant between species and associated predictors in equations (2), (6), and (11)-(15). As indicated by ANCOVA, the goodness-of-fit of BEMs is significantly affected by species. Thus, to include the species interaction, separate regression lines (equations) per species were fitted, whilst appropriate sample sizes were maintained. The separate equations were named and numbered following the format in Table 2 for *E. dunnii* (Eq (1a) to Eq (19a)) and *C. citriodora* (Eq (1b) to Eq (19b)). Candidate BEMs and their regression coefficients are presented in Table 3 (*E. dunnii*) and Table 4 (*C. citriodora*). Most of BEMs had $R^2 > 0.84$. Equations (3a) and (3b) are not presented as they were not significant ($P > 0.05$) and had small power fitting (or model predictive ability) ($R^2 < 0.01$).

Table 4
Residual standard errors (RSEs) and regression coefficients for the biomass estimations models (BEMs) developed for *C. citriodora*. *P*-values were < 0.001 for all equations.

Compartment/Equation	Sample Size	RSE	R^2	Coefficients			
				ln*a*	*b*	*c*	*d*
Stem	40						
Eq (1b).		0.118	0.988	−2.41135	2.519862		
Eq (2b).		0.298	0.920	−5.1239	3.3469		
Eq (4b).		0.114	0.989	−2.8148	2.2374	0.4071	
Eq (5b).		0.112	0.989	−2.13151	2.5132	0.65202	
Eq (6b).		0.274	0.934	−4.358	3.3527	1.97	
Eq (7b).		0.101	0.991	−2.6032	2.1177	0.5669	0.862
Eq (8b).		0.119	0.988	−2.56173	2.6385	−0.02232	
Eq (9b).		0.122	0.987	−3.30641	0.93195		
Eq (10b).		0.102	0.991	−2.91783	0.93012		
Bark	40						
Eq (11b).		0.184	0.959	−3.2719	2.15058		
Eq (12b).		0.179	0.963	−3.8546	1.7426	0.5879	
Eq (13b).		0.178	0.962	−4.04977	0.79689		
Branch	75						
Eq (14b).		0.147	0.981	−2.40071	2.32456		
Eq (15b).		0.147	0.981	−2.42214	2.38612	−0.03348	
Crown	26						
Eq (16b).		0.211	0.838	0.94724	1.01407		
Eq (17b).		0.208	0.849	0.9273	0.8154	0.2526	
Eq (18b).		0.206	0.852	1.6175	1.5345	−0.7274	
Eq (19b).		0.198	0.870	1.7352	1.3943	0.3194	−0.8825

Table 5
Final biomass estimation models (BEMs) for stem, bark, branches and crown for *E. dunnii* and *C. citriodora*.

Species	Equation
E. dunnii	Eq (10a).ln (Stem Biomass) = −2.20872 + 0.84761ln (DBH^{2}Hρ)
	Eq (11a).ln (Bark Biomass) = −3.16931 + 2.10418ln (DBH)
	Eq (14a).ln (Branch Biomass) = −2.61184 + 2.34891ln (BD)
	Eq (16a).ln (Crown Biomass) = −1.5309 + 1.5563ln (DBH)
C. citriodora	Eq (10b).ln (Stem Biomass) = −2.91783 + 0.93012ln (DBH^{2}Hρ)
	Eq (13b).ln (Bark Biomass) = −4.04977 + 0.79689ln (DBH^{2}H)
	Eq (14b).ln (Branch Biomass) = −2.40071 + 2.32456ln (BD)
	Eq (16b).ln (Crown Biomass) = 0.94724 + 1.01407ln (DBH)

3.2. Model assessment and selection

Data exploration and testing of hypotheses confirmed conditions of linearity between variables (hypothesis *i*) and independency of residuals (hypothesis *iia*). Residuals were normally distributed (hypothesis *iib*) as indicated by the Quantile-Quantile graph of standardized residuals versus theoretical quantiles, and constant variance or homoscedasticity (hypothesis *iic*) was demonstrated by graphing residuals in function of fitted biomass estimates (graphs for individual BEMs are not shown). The main statistics to assess reliability of coefficients for all BEMs developed for *E. dunnii* and *C. citriodora* species are presented in Appendix C and Appendix D respectively. Confidence intervals of *E. dunnii* BEMs were highly variable, particularly between the stem BEMs. Equations (4a)-(8a), (12a), (15a), and (17a)-(19a) are unreliable having at least one of its coefficients with PRSE > 50%. Equation (4a), (7a)-(8a), (12a), (15a), (18a) and (19a) had VIFs > 5 indicating that the variance of coefficients increased because of multicollinearity. The PLS regression indicated that additional variables included in multiple linear (4a)-(7a), (12a), (17a)-(19a) and polynomial (15a) BEMs had small (< 22%) to nil (0%) contribution to explain model variance.

For *C. citriodora* BEMs, confidence intervals were also variable for stem BEMs (Appendix D). Equations (4b), (8b) and (12b), (15b), and (17b)-(19b) are unreliable (coefficients with PRSE > 50%). Equations (5b) and (7b) had coefficients with PRSE > 30%, requiring further analysis of diagnostic statistics to confirm their reliability. Multicollinearity was present in equations (4b), (7b), (8b), (12b), (15b), (18b) and (19b) (VIFs > 5), and additional predictors contributed less than 1% to explain variance in BEMs (4b)–(8b), (12b), (15b), and (17b)–(19b). For both species, candidate BEMs had not more than 10% of outliers (Appendix E and Appendix F respectively). Equation (4a), (4b) to (7a), (7b) had more than 20% leverage points whilst equations (16a)-(19a) and (17b)-(19b) had around 11–13% leverage points. All BEMs had a MAPE < 10%.

After testing hypotheses, assessing reliability of coefficients and checking model functionality, selection of best *E. dunnii* and *C. citriodora* stem BEMs was limited to the candidate equation (1a,b), (2a,b), (9 a,b) and (10a,b). Equation (11a,b) and (13a,b) were the final nominees for bark biomass. Equation 14(a,b) and 16(a,b) were the most suitable branch and crown BEMs respectively. In both *E. dunnii* and *C. citriodora* species, equations (10a) and (10b) had low AIC (−41.7 and −64.8) and the highest AICw (0.71 and 1) compared to remaining stem BEMs (Appendix E and F respectively). For bark biomass estimations in *E. dunnii*, equation (11a) had the minimum AIC (−16.1) and the highest AICw (0.72) whereas for *C. citriodora*, equation (13b) had the minimum AIC (−20.6) and the highest AICw (0.81).

The final BEMs are presented in Table 5. Statistic diagnostic graphs such as logarithm transformations to variables, normality tests and homoscedasticity checks for the selected BEMs are presented in Figs. 2 and 3. The selected BEMs were developed to be used in the calculation of biomass (in tonnes) of the stem, bark and crown compartments of individual trees of *E. dunnii* and *C. citriodora*. The summation of biomass of these compartments results in the total tree AGB (in tonnes), and this calculation can be repeated for each tree using their respective tree parameters as predictors to obtain the total biomass of all trees in a given plot (in tonnes per plot area). Plot biomass estimates can be extrapolated to stand level estimates (in tonnes per hectare).

3.3. Model validation

The 5-fold and 3-fold cross-validation showed that residuals of all tested folds were not significantly different from zero, indicating that stem, bark and crown BEMs for both species were valid (Table 6). Branch BEMs had only one fold (Fold 4 in *E. dunnii*, and Fold 5 in *C. citriodora*) for which the equation was not valid as per the two-tailed *t*-test (t-statistic > Student critical t-value) at p = 0.05. However, as four of the folds produced valid estimates, it is assumed in overall, than the equations are valid. The coefficient of determination R^2 was more variable between folds for the bark and crown equations than for stem and branch. However, relatively high goodness-of-fit was obtained for all tested folds. The MAPE results showed that prediction errors for the tested folds oscillated between 0.4 and 1.6% for stem, 0.7–2.6% for bark, 1.4–3.2% for branch and 1.5–4.1% for crown BEMs.

4. Discussion

For both species, most of BEMs (except equations (3a) and (3b), adequately described the relationships between response and predictor variables. Each set of candidate BEMs had relatively high R^2 (> 0.84) with little variation between compartments. In addition, the RSE values were very consistent among the BEMs for bark, branch and crown biomass, and more variable for the BEMs for stem biomass. Based on R^2 and RSE values, most of the BEMs would seem to be suitable as their regression lines fitted approximately well the measured biomass, and with relatively low errors. However, as stated in Piccard [25] and Sileshi [26], limiting the assessment of model quality to the use of these two statistics, particularly when BEMs share similar R^2 and RSE values, represents not just a poor criterion but questionable approach.

Based on the calculated PRSE values, equations (4a,b), (5a,b), (6a,b) and (7a,b) for stem biomass which represented linear models with multiple predictors had at least one unreliable coefficient. The power fitting represented by the coefficient of determination (R^2) of these equations increased at including additional predictors but the contribution of these predictors at explaining the variance in the regressions was insignificant with respect to others (PLS < 1% and VIF > 5). Moreover, all BEMs had leverage points proportions considered to influence the goodness-of-fit statistics of the regression and therefore to reduce accuracy of biomass predictions. These results contradict other research's findings such as Chave et al., Henry et al. and Xu et al. [3,19,44], who reported that equations using more than one predictor and in different combinations significantly (p < 0.05) contributed to more accurate stem biomass estimates than single-variable equations. Similar findings were reported for polynomial BEMs developed in study sites such as tropical and subtropical forests in America, Asia and Oceania that were published as best models [19,44]. These studies limited the assessment of model quality to few diagnostic statistics, making difficult to interpret whether differences are attributed to the nature of models or to incomplete statistical reports during model fitting.

Our results indicated that polynomial BEMs 8a and 8b had unreliable and highly collinear coefficients, with poor contribution of additional predictors. Conversely, stem BEMs using DBH and H as single variable (Model Type I) as well as the combined variables DBH^{2}H and DBH^{2}Hρ (Model Type II), had reliable coefficients and did not present any issues of multicollinearity, leverage points, outliers and prediction errors. Similar results were found in *Eucalyptus nitens* plantations in Northern Spain [16], and *E. grandis*, *E. microcorys*, *E. saligna* and *E.*

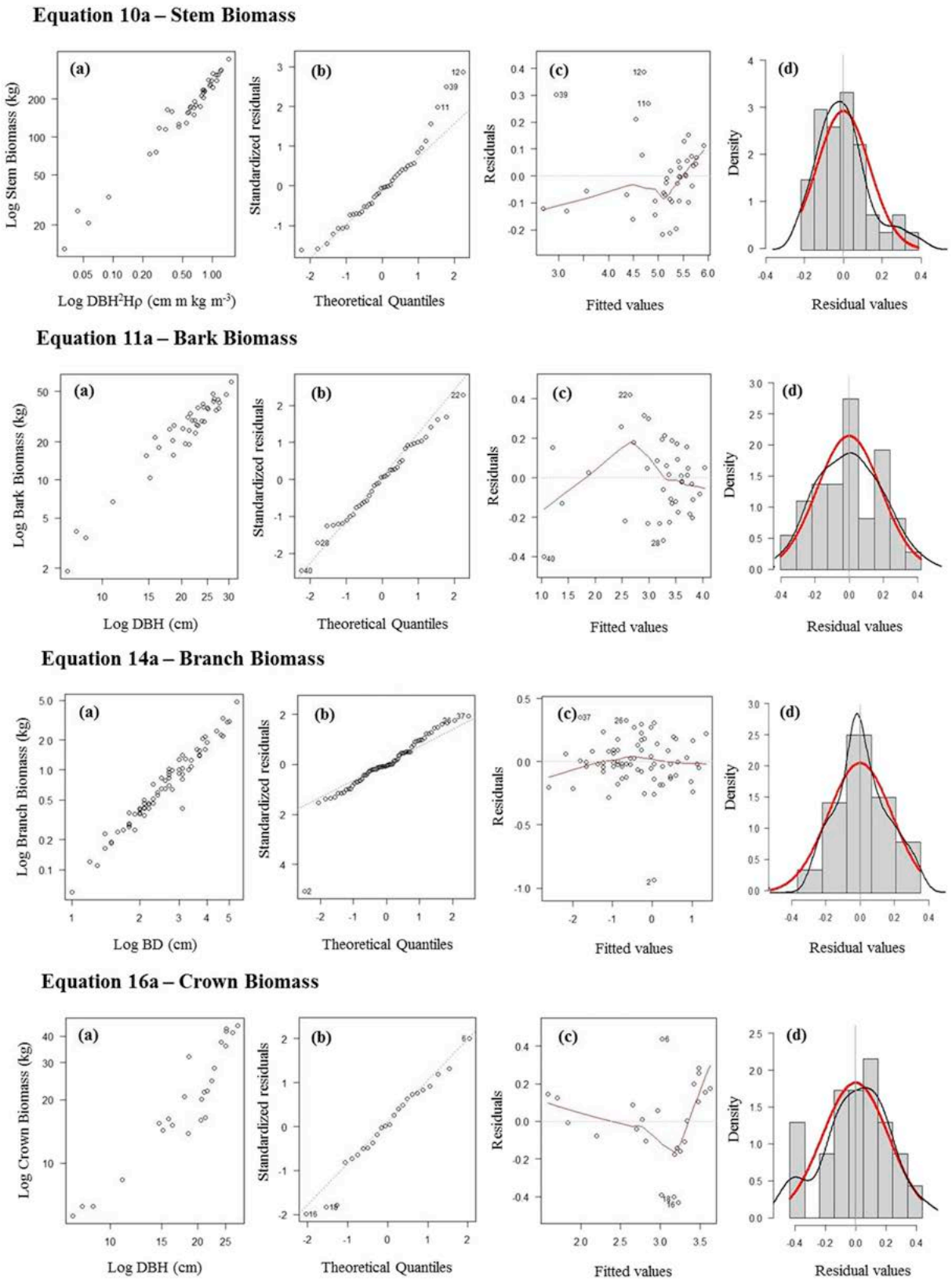

Fig. 2. Diagnostics for biomass estimation models for stem, bark, branches and crown of *E. dunnii*: a) Log transformation to variables, b) Normal Quantile-Quantile graph for normality check, c) Residuals in function of fitted values for homoscedasticity check, and d) Normal and Kernel distribution of residuals.

nitens plantations in Australia [6] at estimating either total tree AGB or any of its compartments from the single variable DBH. Other studies also reported that the use of combined variables ($DBH^2H\rho$) effectively predicted accurate stem and total biomass estimations in different vegetation types [2,5,8,19]. The importance of including H and ρ in conjunction with DBH as combined variable in stem BEMs relied on the premise that these predictors are sensitive to the influence of species characteristics and site factors [12,45], and therefore can impact on

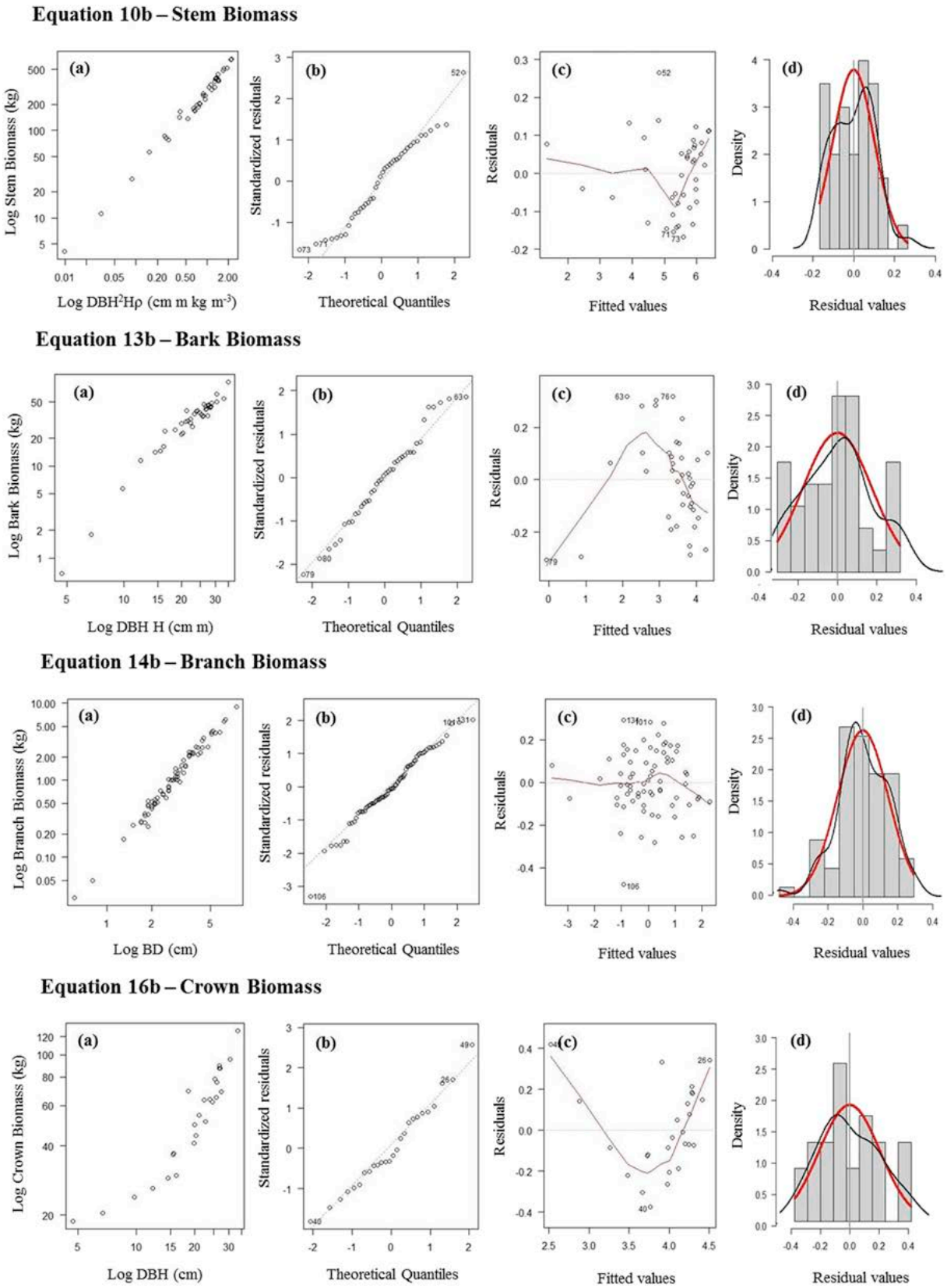

Fig. 3. Diagnostics for biomass estimation models for stem, bark, branches and crown of *C. citriodora*: a) Log transformation to variables, b) Normal Quantile-Quantile graph for normality check, c) Residuals in function of fitted values for homoscedasticity check, and d) Normal and Kernel distribution of residuals.

biomass estimations.

The multiple linear bark BEMs (12a and 12b) as well as the polynomial branch BEMs (15a and 15b) had unreliable and highly collinear coefficients, and their additional predictors poorly contributed to model fit. In this study, DBH was the best predictor of bark biomass in *E. dunnii* similar to reported for some North American hardwood and softwood

Table 6
Cross-validation for the final biomass estimation models (BEMs). Coefficient of determination (R^2) and mean absolute percentage error (MAPE) (in %) for each fold are presented. Values in bold correspond to the folds for which the tested equation was not valid as per the two-tailed *t*-test (p = 0.05).

Species	Biomass Equation	Fold 1		Fold 2		Fold 3		Fold 4		Fold 5	
		R^2	MAPE	R^2	MAPE	R^2	MAPE	R^2	MAPE	R^2	MAPE
E. dunnii	Eq (10a)	0.94	1.2	0.97	0.8	0.92	1.3	0.96	0.7	0.93	0.9
	Eq (11a)	0.62	2.2	0.76	2.5	0.88	2.3	0.93	0.7	0.75	1.7
	Eq (14a)	0.92	3.1	0.71	3.2	0.89	2.5	0.81	3.0	0.93	2.6
	Eq (16a)	0.59	2.2	0.78	3.0	0.72	4.1	–	–	–	–
C. citriodora	Eq (10b)	0.98	0.4	0.95	1.6	0.97	1.4	0.98	0.8	0.97	1.0
	Eq (13b)	0.61	2.0	0.66	2.2	0.89	1.9	0.86	1.6	0.90	1.3
	Eq (14b)	0.92	2.5	0.93	2.9	0.82	3.2	0.90	2.7	0.95	1.4
	Eq (16b)	0.87	2.5	0.75	2.4	0.69	1.5	–	–	–	–

species [46]. Conversely, DBH combined with H were the best predictors in bark BEMs for *C. citriodora* similar to reported in Bernardo [47]. On the other hand, branch diameter alone was the best predictor in branch BEMs. However, due to practical difficulty in undertaking branch measurements, some studies have developed high predictive branch BEMs ($R^2 > 0.90$) from DBH itself [41,45] or DBH combined with H [47].

Multiple linear BEMs for crown biomass (17a,b), (18a,b) and (19a,b) had coefficients with significant large errors and presence of multicollinearity. Also, the inclusion of additional predictors resulted in poor contribution to the model fit and large proportion of leverage points. These results were consistent with studies in other types of vegetation such as tropical lowland Dipterocarp and mixed secondary forests that found that multiple predictors did not significantly increase the predictive ability compared to biomass equations that used DBH alone [12,23]. The equations with the single predictor DBH (16a and 16b) had small proportion of leverage points which did not affect model coefficients reliability, and resulted in the best BEMs for crown biomass for *E. dunnii* and *C. citriodora* respectively. Our results also clearly demonstrated that not necessarily the model with the minimum AIC was the best BEM, making the AICw a better criterion for model selection based on the relative importance of its predictors. For instance, equation (17a) (*E. dunnii*) and (19b) (*C. citriodora*) for crown biomass which had the minimum AIC were unreliable, whilst equations (16a) and (16b) which had the higher AICw were the best BEMs. Similar findings were found in Sileshi [26] who calculated the AICw values of several BEMs published in the literature and that demonstrated that models with the highest R^2 and minimum AIC were not necessarily the best models as claimed by the studies because they had very low likelihood (small AICw). .

Biomass estimation models for stem, bark, branch and crown differed between species as described by statistically different intercepts and slopes in the separately fitted regressions. Different intercepts indicated differences in biomass amounts, whilst different slopes indicated differences in biomass production rates in each tree compartment. This suggests that the effect of species on the production of total AGB could be related and greatly dependent upon tree architecture. For instance, *C. citriodora* stands had larger trees, and in particular larger crowns, compared to *E. dunnii* stands. Furthermore, *C. citriodora* had higher stem wood density than *E. dunnii*. These results illustrate how species characteristics (e.g. height, crown depth, stem wood density) can induce specific tree form and growth, and therefore influence the amount and distribution of AGB [12,17,23]. Chave [19] suggested tree form responds to effective biomass allocation into compartments to optimise photosynthetic production.

Our findings support the preference for developing species-specific BEMs to accurately predict AGB and reflect the differences in proportional relationships between compartments that occur because of the variability in tree form and growth characteristics between species. Similarly, Basuki [12], Ketterings [23], Marshall [5], Montagu [17], Nelson [14], and Paul [22] found that species- and site-specific BEMs produced more accurate estimates. Validation of equations indicated predicted biomass in compartments differed up to 1.6% (stem), 2.6% (bark) and 3.2% (branch), and 4.1% (crown) from true measured biomass, confirming that the selected equations have better capacity to produce reliable biomass estimates. Overestimation is a common pattern when validating models to estimate above-ground biomass [19]. Conversely, our results indicated underestimated biomass estimates.

5. Conclusions

Strong relationships were distinctive between stem, bark, branch and crown biomass estimates and their respective predictors. Because *E. dunnii* and *C. citriodora* have different growth characteristics and therefore varying proportional relationships, separate species-specific BEMs were developed. Variable confidence intervals, unreliable coefficients, multicollinearity, poor contribution of predictors in multiple linear models, and the presence of outliers and leverage points were found to affect the goodness-of-fit statistics in most of the candidate BEMs. The rigorous process of testing hypotheses, checking diagnostic statistics, assessing reliability of model coefficients and assessing model functionality allowed for the selection of the most reliable BEMs which were those ones derived from either single (simple linear - Type I) or combined variables (compound predictor - Type IV), where DBH was the best predictor variable. On the other hand, multiple linear (Type II) and polynomial (Type III) BEMs generally produced unacceptable levels of errors in coefficients and variables, and their use is not recommended. The selected BEMs can be used by forest managers to obtain reliable biomass estimates for timber production profitability, bioenergy or carbon stocks in plantations, as long as the predictors fall within ranges of the study population. For BEMs for stem biomass, and if wood density is not available, we recommend the use of our calculated species-level average density (0.56 for *E. dunnii*, 0.66 for *C. citriodora*).

Acknowledgements

We would like to acknowledge the Forest Research Centre at Southern Cross University, the Subtropical Farm Forestry Association SFAA, and the Hurford Forests company for providing support to undertake this research. We also thank researcher Mia Cassidy, and technical officers Ron Cox and Peter Bligh-Jones, who assisted during data collection in 2015/2016.

Appendix A. Formula of assessed parameters

Parameter	Abbreviation	Formula	Source
Percent relative standard error	PRSE	100x (SE/|θ|), where SE is the standard error and "θ" the coefficients	Barde M and Barde P (2012).
Variance inflation factor	VIF	$\frac{1}{1-R^2}$	Hair et al. (2006)
Mean absolute percentage error	MAPE	$\frac{100}{n} = \sum_{i=1}^{n} \frac{\|Mo - Mp\|}{Mo}$ Where Mo and Mp are the observed and predicted biomass of the ith tree, and n is the number of trees	Yao et al. (2013)

Cited references in Appendix A.
Barde MP, Barde PJ. (2012). What to use to express the variability of data: standard deviation or standard error of mean? Perspect. Clin. Res. 3, 113–116.
Hair JF, Anderson R, Tatham RL, Black WC. (2006). Multivariate Data Analysis. Upper Saddle River, NJ: Prentice Hall.
Yao X, Fu B, Lu Y, Sun F, Wang S, Liu M. (2013). Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment. PLoS ONE, 8(1), e54660.

Appendix B. Analysis of Covariance (ANCOVA) for all significant candidate models. Values of interaction ANOVA (Analysis of Variance) in grey were not significant (*p*-value > 0.05)

Model	Species effect	Predictors (Covariates)	Interaction ANOVA
Stem			
Eq (1). $\ln(B) = \ln(a) + b\ln(DBH)$	$1.155e^{-12}$	$< 2.2e^{-16}$	0.000878
Eq (2). $\ln(B) = \ln(a) + b\ln(H)$	0.000128	$< 2.2e^{-16}$	0.532127
Eq (4). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(H)$	$6.58e^{-13}$	$< 2.2e^{-16}$	0.001717
		0.009832	0.984479
Eq (5). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(\rho)$	$2.666e^{-13}$	$< 2.2e^{-16}$	0.000898
		0.002438	0.644333
Eq (6). $\ln(B) = \ln(a) + b\ln(H) + c\ln(\rho)$	$6.899e^{-05}$	$< 2.2e^{-16}$	0.515136
		0.008934	0.188460
Eq (7). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(H) + d\ln(\rho)$	$6.952e^{-14}$	$< 2.2e^{-16}$	0.001868
		0.006111	0.585257
		0.000479	0.988661
Eq (8). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(DBH)^2$	$2.643e^{-13}$	$< 2.2e^{-16}$	0.000304
		0.014978	NA
Eq (9). $\ln(B) = \ln(a) + b\ln(DBH^2H)$	$2e^{-12}$	$< 2.2e^{-16}$	0.005695
Eq (10). $\ln(B) = \ln(a) + b\ln(DBH^2H\rho)$	$5.625e^{-14}$	$< 2.2e^{-16}$	0.003779
Bark			
Eq (11). $\ln(B) = \ln(a) + b\ln(DBH)$	0.000995	$< 2.2e^{-16}$	0.684739
Eq (12). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(H)$	0.000918	$< 2.2e^{-16}$	0.826576
		0.094620	0.457895
Eq (13). $\ln(B) = \ln(a) + b\ln(DBH^2H)$	0.000944	$< 2.2e^{-16}$	0.908646
Branch			
Eq (14). $\ln(B) = \ln(a) + b\ln(BD)$	$< 2e^{-16}$	$< 2e^{-16}$	0.7226
Eq (15). $\ln(B) = \ln(a) + b\ln(BD) + c\ln(BD)^2$	$< 2e^{-16}$	$< 2e^{-16}$	0.7326
		0.3652	0.6095
Crown			
Eq (16). $\ln(B) = \ln(a) + b\ln(DBH)$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	0.000858
Eq (17). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(CD)$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$8.012e^{-05}$
		0.3164	0.9261
Eq (18). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(H)$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	0.001104
		0.110231	0.525045
Eq (19). $\ln(B) = \ln(a) + b\ln(DBH) + c\ln(CD) + d\ln(H)$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$7.018e^{-05}$
		0.30653	0.95946
		0.06762	0.42246

Appendix C. Confidence intervals (95% CIs), percent relative standard error (PRSE), variance inflation factor (VIF) and partial least square regressions (PLS regressions) for *Eucalyptus dunnii* biomass estimation models (BEMs). Values in grey indicate unreliable coefficients

Compartment/Equation	Coefficient	95%CIs	PRSE (%)	VIF	PLS (%)
Stem					
Eq (1a).	*lna*	(-2.114, −1.293)	11.91		
	b	(2.108, 2.379)	2.98	1	0.967
Eq (2a).	*lna*	(-6.015, −3.811)	11.08		
	b	(2.843, 3.547)	5.45	1	0.899

Eq (4a).	*lna*	(-3.060, −1.323)	19.55		
	b	(1.536, 2.417)	11.00	10.72	0.966
	c	(-0.236, 1.065)	77.49	10.72	0.003
Eq (5a).	*lna*	(-1.848, - 0.563)	26.30		
	b	(2.118, 2.379)	2.87	1	0.967
	c	(-0.014, 1.798)	50.12	1	0.003
Eq (6a).	*lna*	(-6.076, −3.161)	15.57		
	b	(2.841, 3.552)	5.49	1	0.899
	c	(-1.146, 2.190)	157.76	1	0.001
Eq (7a).	*lna*	(-2.672, −0.658)	29.81		
	b	(1.581, 2.435)	10.48	10.79	0.966
	c	(-0.258, 1.003)	83.41	10.78	0.006
	d	(-0.049, 1.759)	52.14	1.006	0
Eq (8a).	*lna*	(-5.171, −0.230)	45.15		
	b	(1.162, 4.816)	30.17	180	0.749
	c	(-0.468, 0.196)	120.59	180	0.219
Eq (9a).	*lna*	(-3.149, −2.209)	8.67		
	b	(0.795, 0.897)	2.99	1	0.967
Eq (10a).	*lna*	(-2.633, −1.785)	9.48		
	b	(0.798, 0.897)	2.87	1	0.970
Bark					
Eq (11a).	*lna*	(-3.713, −2.625)	8.48		
	b	(1.925, 2.284)	4.21	1	0.937
Eq (12a.	*lna*	(-4.563, −2.217)	17.07		
	b	(1.389, 2.578)	14.80	10.72	0.932
	c	(-0.691, 1.066)	231.55	10.72	0.005
Eq (13a).	*lna*	(-4.710, −3.437)	7.72		
	b	(0.723, 0.861)	4.32	1	0.934
Branch					
Eq (14a).	*lna*	(-2.727, −2.497)	2.21		
	b	(2.232, 2.466)	2.49	1	0.957
Eq (15a).	*lna*	(-2.893, −2.469)	3.96		
	b	(2.049, 3.012)	9.55	16.99	0.957
	c	(-0.356, 0.157)	129.17	16.99	0
Crown					
Eq (16a).	*lna*	(-2.274, −0.788)	23.41		
	b	(1.301, 1.812)	7.92	1	0.879
Eq (17a).	*lna*	(-2.687, −1.116)	19.87		
	b	(1.227, 1.729)	8.17	1.11	0.876
	c	(0.002, 0.547)	47.81	1.11	0.024
Eq (18a).	*lna*	(-2.564, 0.112)	52.50		
	b	(0.997, 2.507)	20.72	8.42	0.858
	c	(-1.312, 0.744)	173.96	8.42	0.023
Eq (19a).	*lna*	(-2.817, −0.256)	39.96		
	b	(1.014, 2.418)	19.62	8.44	0.874
	c	(0.004, 0.558)	47.26	1.11	0.008
	d	(-1.306, 0.609)	131.75	8.46	0.021

Appendix D. Confidence intervals (95% CIs), percent relative standard error (PRSE), variance inflation factor (VIF) and partial least square (PLS) regressions for *Corymbia citriodora* biomass estimation models (BEMs). Values in grey indicate unreliable coefficients

Compartment/Equation	Coefficient	95%CIs	PRSE (%)	VIF	PLS (%)
Stem					
Eq (1b).	*lna*	(-2.699, −2.124)	5.89		
	b	(2.427, 2.613)	1.82	1	0.988
Eq (2b).	*lna*	(-6.138, −4.110)	9.78		
	b	(3.023, 3.671)	4.78	1	0.920
Eq (4b).	*lna*	(-3.311, −2.319)	8.70		
	b	(1.936, 2.539)	6.66	11.3	0.984
	c	(-0.008, 0.823)	50.36	11.3	0.004
Eq (5b).	*lna*	(-2.500, −1.763)	8.54		
	b	(2.425, 2.602)	1.74	1	0.988
	c	(0.074, 1.231)	43.78	1	0.001
Eq (6b).	*lna*	(-5.442, −3.274)	12.27		
	b	(3.054, 3.651)	4.39	1	0.921
	c	(0.554, 3.387)	35.49	1	0.013
Eq (7b).	*lna*	(-3.067, −2.140)	8.78		
	b	(1.837, 2.398)	6.53	12.2	0.985
	c	(0.182, 0.952)	33.50	12.2	0.007
	d	(0.318, 1.407)	31.15	1.1	0
Eq (8b).	*lna*	(-3.652, −1.471)	21.01		
	b	(1.804, 3.473)	15.61	78.4	0.987
	c	(-0.178, 0.134)	344.98	78.4	0

Eq (9b).	*lna*	(-3.637, −2.976)	4.94		
	b	(0.896, 0.968)	1.89	1	0.987
Eq (10b).	*lna*	(-3.183, −2.652)	4.50		
	b	(0.900, 0.960)	1.58	1	0.991
Bark					
Eq (11b).	*lna*	(-3.721, −2.823)	6.78		
	b	(2.005, 2.296)	3.34	1	0.959
Eq (12b).	*lna*	(-4.636, −3.073)	10.00		
	b	(1.267, 2.218)	13.46	11.3	0.961
	c	(-0.066, 1.242)	54.91	11.3	0.002
Eq (13b).	*lna*	(-4.532, −3.567)	5.89		
	b	(0.745, 0.849)	3.21	1	0.962
Branch					
Eq (14b).	*lna*	(-2.486, −2.316)	1.77		
	b	(2.248, 2.401)	1.65	1	0.981
Eq (15b).	*lna*	(-2.529, −2.315)	2.21		
	b	(2.186, 2.586)	4.20	6.8	0.970
	c	(-0.134, 0.067)	150.48	6.8	0.011
Crown					
Eq (16b).	*lna*	(0.386, 1.508)	28.70		
	b	(0.826, 1.202)	8.99	1	0.838
Eq (17b).	*lna*	(0.372, 1.482)	28.93		
	b	(0.451, 1.180)	21.60	3.8	0.837
	c	(-0.146, 0.651)	76.25	3.8	0.011
Eq (18b).	*lna*	(0.547, 2.689)	31.99		
	b	(0.797, 2.272)	23.23	16.1	0.813
	c	(-1.726, 0.271)	66.33	16.1	0.039
Eq (19b).	*lna*	(0.695, 2.776)	28.92		
	b	(0.664, 2.124)	25.25	17	0.826
	c	(-0.067, 0.706)	58.36	4	0.025
	d	(-1.862, 0.097)	53.50	16.7	0.018

Appendix E. Leverage points, outliers, mean absolute percentage error (MAPE), Akaike Information Criteria (AIC) and its weight (AICw) for *Eucalyptus dunnii* biomass estimation models (BEMs). Values in grey indicate unreliable and uncertain models. Bold figures indicate the highest AICw values in each compartment, and therefore the selected equations. AICw values were not calculated (NA) for unsuitable BEMs

Compartment/Equation	Leverage points (%)	Outliers (%)	MAPE (%)	AIC	AICw
Stem					
Eq (1a).	10	5	4.3	−38.6	0.15
Eq (2a).	10	0	8.1	6.6	0
Eq (4a).	27.5	2.5	4.2	−38.3	NA
Eq (5a).	22.5	2.5	4	−40.7	NA
Eq (6a).	22.5	0	7.9	8.2	NA
Eq (7a).	35	0	3.9	−40.2	NA
Eq (8a).	10	5	4.4	−37.3	NA
Eq (9a).	10	2.5	4.3	−38.4	0.14
Eq (10a).	10	2.5	4	−41.7	**0.71**
Bark					
Eq (11a).	10	2.5	6.1	−16.1	**0.72**
Eq (12a).	10	2.5	6.1	−14.3	NA
Eq (13a).	10	2.5	6.3	−14.2	0.28
Branch					
Eq (14a).	9	0	10	−34.6	**0.67**
Eq (15a).	12	0	0	−33.2	0.33
Crown					
Eq (16a).	13	0	4.2	0.06	**1**
Eq (17a).	13	0	3.8	−2.48	NA
Eq (18a).	13	0	4.3	1.68	NA
Eq (19a).	4	0	3.7	1.16	NA

Appendix F. Leverage points, outliers, mean absolute percentage error (MAPE), Akaike Information Criteria (AIC) and its weight (AICw) for *Corymbia citriodora* biomass estimation models (BEMs). Values in grey indicate unreliable and uncertain models. Bold figures indicate the highest AICw values in each compartment, and therefore the selected equations. AICw values were not calculated (NA) for unsuitable BEMs

Compartment/Equation	Leverage points (%)	Outliers (%)	MAPE (%)	AIC	AICw
Stem					
Eq (1b).	7.5	0	3.9	−53.7	0
Eq (2b).	5	0	9.8	20.7	0
Eq (4b).	20	0	3.6	−55.7	NA
Eq (5b).	25	0	3.4	−56.9	NA
Eq (6b).	27.5	0	8.8	14.9	NA
Eq (7b).	30	0	3.2	−63.8	NA
Eq (8b).	7.5	0	3.9	−51.8	NA
Eq (9b).	10	0	4	−51	0
Eq (10b).	7.5	0	3.4	−64.8	**1**
Bark					
Eq (11b).	7.5	2.5	5.7	−17.9	0.19
Eq (12b).	10	2.5	5.6	−19.4	NA
Eq (13b).	7.5	2.5	5.7	−20.6	**0.81**
Branch					
Eq (14b).	6.7	0	8.7	−71.3	**0.68**
Eq (15b).	6.7	1.3	0	−69.7	0.32
Crown					
Eq (16b).	7.7	3.6	4.5	−3.2	**1**
Eq (17b).	11.5	7.7	4	−3	NA
Eq (18b).	11.5	7.7	4.1	−3.6	NA
Eq (19b).	11.5	7.7	3.5	−4.9	NA

References

[1] J. Chave, R. Condit, S. Aguilar, A. Hernandez, S. Lao, R. Perez, Error propagation and scaling for tropical forest biomass estimates, Philos. Trans. R. Soc. Biol. Sci. 359 (1443) (2004) 409–420.

[2] M. Hunter, M. Keller, D. Victoria, D. Morton, Tree height and tropical forest biomass estimation, Biogeosciences 10 (12) (2013) 8385–8399.

[3] M. Henry, N. Picard, C. Trotta, R.J. Manlay, R. Valentini, M. Bernoux, L. Saint-André, Estimating Tree Biomass of Sub-saharan African Forests: a Review of Available Allometric Equations, Finnish Society of Forest Science, 2011.

[4] D. Eamus, K. McGuinness, W. Burrows, Review of Allometric Woody Biomass for Queensland, the Northern Territory and Western Australia, (2000).

[5] A.R. Marshall, S. Willcock, P.J. Platts, J.C. Lovett, A. Balmford, N.D. Burgess, J.E. Latham, P.K.T. Munishi, R. Salter, D.D. Shirima, S.L. Lewis, Measuring and modelling above-ground carbon and tree allometry along a tropical elevation gradient, Biol. Conserv. 154 (2012) 20–33 https://doi.org/10.1016/j.biocon.2012.03.017.

[6] A. Specht, P.W. West, Estimation of biomass and sequestered carbon on farm forest plantations in northern New South Wales, Australia, Biomass Bioenergy 25 (4) (2003) 363–379 https://doi.org/10.1016/S0961-9534(03)00050-3.

[7] T.s.M. Araújo, N. Higuchi, J.A. de Carvalho-Júnior, Comparison of formulae for biomass content determination in a tropical rain forest site in the state of Pará, Brazil, For. Ecol. Manag. 117 (1) (1999) 43–52.

[8] S. Brown, A.J.R. Gillespie, A.E. Lugo, Biomass estimation methods for tropical forests with applications to forest inventory data, For. Sci. 35 (Number 4) (1989) 881–902 822.

[9] B. Chave, Riéra, M.-A. Dubois, Estimation of biomass in a neotropical forest of French Guiana: spatial and temporal variability, J. Trop. Ecol. 17 (01) (2001) 79–96.

[10] R.M. Gifford, Carbon Contents of Above-Ground Tissues of Forest and Woodland Trees: Australian Greenhouse Office, (2000).

[11] K. Paul, Jacobsen, V. Koul, P. Leppert, J. Smith, Predicting growth and sequestration of carbon by plantations growing in regions of low-rainfall in southern Australia, For. Ecol. Manag. 254 (2) (2008) 205–216 https://doi.org/10.1016/j.foreco.2007.08.003.

[12] T.M. Basuki, P.E. van Laake, A.K. Skidmore, Y.A. Hussin, Allometric equations for estimating the above-ground biomass in tropical lowland Dipterocarp forests, For. Ecol. Manag. 257 (8) (2009) 1684–1694 https://doi.org/10.1016/j.foreco.2009.01.027.

[13] P. Muukkonen, Generalized allometric volume and biomass equations for some tree species in Europe, Eur. J. For. Res. 126 (2) (2007) 157–166, https://doi.org/10.1007/s10342-007-0168-4.

[14] B.W. Nelson, R. Mesquita, J.L.G. Pereira, S. Garcia Aquino de Souza, G. Teixeira Batista, L. Bovino Couto, Allometric regressions for improved estimate of secondary forest biomass in the central Amazon, For. Ecol. Manag. 117 (1–3) (1999) 149–167 https://doi.org/10.1016/S0378-1127(98)00475-7.

[15] M. Neumann, A. Moreno, V. Mues, S. Härkönen, M. Mura, O. Bouriaud, M. Lang, Wouter M. Achten, A. Thivolle-Cazat, K. Bronisz, Comparison of carbon estimation methods for European forests, For. Ecol. Manag. 361 (2016) 397–420.

[16] C. Pérez-Cruzado, R. Rodríguez-Soalleiro, Improvement in accuracy of aboveground biomass estimation in Eucalyptus nitens plantations: effect of bole sampling intensity and explanatory variables, For. Ecol. Manag. 261 (11) (2011) 2016–2028 https://doi.org/10.1016/j.foreco.2011.02.028.

[17] K.D. Montagu, K. Düttmer, C.V.M. Barton, A.L. Cowie, Developing general allometric relationships for regional estimates of carbon sequestration—an example using Eucalyptus pilularis from seven contrasting sites, For. Ecol. Manag. 204 (1) (2005) 115–129 https://doi.org/10.1016/j.foreco.2004.09.003.

[18] K. Senelwa, R.E.H. Sims, Tree biomass equations for short rotation eucalypts grown in New Zealand, Biomass Bioenergy 13 (3) (1997) 133–140 https://doi.org/10.1016/S0961-9534(97)00026-3.

[19] J. Chave, C. Andalo, S. Brown, M. Cairns, J. Chambers, D. Eamus, H. Fölster, F. Fromard, N. Higuchi, T. Kira, Tree allometry and improved estimation of carbon stocks and balance in tropical forests, Oecologia 145 (1) (2005) 87–99.

[20] R.H. Razakamanarivo, A. Razakavololona, M.-A. Razafindrakoto, G. Vieilledent, A. Albrecht, Below-ground biomass production and allometric relationships of eucalyptus coppice plantation in the central highlands of Madagascar, Biomass Bioenergy 45 (2012) 1–10 https://doi.org/10.1016/j.biombioe.2011.01.020.

[21] P. Snowdon, Protocol for Sampling Tree and Stand Biomass, Australian Greenhouse Office, 2002.

[22] K.I. Paul, S.H. Roxburgh, J.R. England, P. Ritson, T. Hobbs, K. Brooksbank, R.J. Raison, J.S. Larmour, S. Murphy, J. Norris, Development and testing of allometric equations for estimating above-ground biomass of mixed -species environmental plantings, For. Ecol. Manag. 310 (2013) 483–494.

[23] Q.M. Ketterings, R. Coe, M. van Noordwijk, Y. Ambagau, C.A. Palm, Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests, For. Ecol. Manag. 146 (1–3) (2001) 199–209 https://doi.org/10.1016/S0378-1127(00)00460-6.

[24] C. Pérez-Cruzado, L. Fehrmann, P. Magdon, I. Cañellas, H. Sixto, C. Kleinn, On the site-level suitability of biomass models, Environ. Model. Softw 73 (2015) 14–26 https://doi.org/10.1016/j.envsoft.2015.07.019.

[25] N. Piccard, L. Saint André, M. Henry, Manual for Building Tree Volume and Biomass Allometric Equations: from Field Measurement to Prediction, Food and Agricultural Organization of the United Nations, Rome, and Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, 2012.

[26] G.W. Sileshi, A critical review of forest biomass estimation models, common mistakes and corrective measures, For. Ecol. Manag. 329 (2014) 237–254, https://doi.org/10.1016/j.foreco.2014.06.026.

[27] A. Zerihun, K.D. Montagu, M.B. Hoffmann, S.G. Bray, Patterns of below- and aboveground biomass in Eucalyptus populnea woodland communities of northeast Australia along a rainfall gradient, Ecosystems 9 (4) (2006) 501–515, https://doi.

org/10.1007/s10021-005-0155-x.
[28] M. Zewdie, M. Olsson, T. Verwijst, Above-ground biomass production and allometric relations of Eucalyptus globulus Labill. coppice plantations along a chronosequence in the central highlands of Ethiopia, Biomass Bioenergy 33 (3) (2009) 421–428 https://doi.org/10.1016/j.biombioe.2008.08.007.
[29] S. Bastien-Henri, A. Park, M. Ashton, C. Messier, Biomass distribution among tropical tree species grown under differing regional climates, For. Ecol. Manag. 260 (3) (2010) 403–410 https://doi.org/10.1016/j.foreco.2010.04.035.
[30] N. António, M. Tomé, J. Tomé, P. Soares, L. Fontes, Effect of tree, stand, and site variables on the allometry of Eucalyptus globulus tree biomass, Can. J. For. Res. 37 (5) (2007) 895–906, https://doi.org/10.1139/X06-276.
[31] R. Downhan, M. Gavran, Australian Plantation Statistics 2018 Update, Research by the Australian Bureau of Agricultural and Resource Economics and Sciences, 2018.
[32] D.J. Boland, M.W. McDonald, Forest Trees of Australia, fifth ed., CSIRO Publishing, Collingwood, Vic, 2006.
[33] T. Jovanovic, R. Arnold, T. Booth, Determining the climatic suitability of Eucalyptus dunnii for plantations in Australia, China and Central and South America, N. For. 19 (3) (2000) 215–226, https://doi.org/10.1023/A:1006662718206.
[34] J.A. Duke, Handbook of Energy Crops, Purdue University, Center for New Crops & Plants Products, Lafayette, 1983.
[35] E.A. Weiss, Essential Oil Crops, CAB International, Wallingford, United Kingdom, 1997, pp. 272–278.
[36] J. Cummins, C. Skennar, L. Capill, M. Cassidy, G. Palmer, Using small hardwood logs to produce liquid fuels and electricity, Aust. For. 79 (3) (2016) 189–195, https://doi.org/10.1080/00049158.2016.1179434.
[37] Australian Government Bureau of Meteorology, Climate statistics for Australian locations, http://www.bom.gov.au/, (2017).
[38] D.T. Morand, The world reference base for soils (WRB) and soil taxonomy: an appraisal of their application to the soils of the northern rivers of New South Wales, Soil Res. 51 (3) (2013) 167–181, https://doi.org/10.1071/SR12144.
[39] R.F. Isbell, The Australian soil classification, Clayton South, Vic.: National Committee on Soil Terrain, second ed., CSIRO Publishing, 2016.
[40] K.S. Glencross, G. Palmer, M.-C. Pelletier, J.D. Nichols, Growth Response to Thinning in Two Subtropical Hardwood Species, (2011).
[41] H. Keith, R.J. Raison, K.L. Jacobsen, Allocation of carbon in a mature eucalypt forest and some effects of soil phosphorus availability, An International Journal on Plant-Soil Relationships 196 (1) (1997) 81–99, https://doi.org/10.1023/A:1004286030345.
[42] G. West, J. Brown, B. Enquist, A general model for the structure and allometry of plant vascular systems, Nature 400 (6745) (1999) 664–667, https://doi.org/10.1038/23251.
[43] K.P. Burnham, D.R. Anderson, Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach, Springer Science & Business Media, 2002.
[44] Y. Xu, J. Zhang, S.B. Franklin, J. Liang, P. Ding, Y. Luo, L. Zhijun, D. Bao, M. Jiang, Improving allometry models to estimate the above- and belowground biomass of subtropical forest, China, Ecosphere 6 (12) (2015) 1–15, https://doi.org/10.1890/ES15-00198.1.
[45] S. Kuyah, J. Dietz, C. Muthuri, M. van Noordwijk, H. Neufeldt, Allometry and partitioning of above- and below-ground biomass in farmed eucalyptus species dominant in Western Kenyan agricultural landscapes, Biomass Bioenergy 55 (2013) 276–284 https://doi.org/10.1016/j.biombioe.2013.02.011.
[46] J.C. Jenkins, D.C. Chojnacky, L.S. Heath, R.A. Birdsey, Comprehensive Database of Diameter-Based Biomass Regressions for North American Tree Species, (2004).
[47] A.L. Bernardo, M.G. Reis, G.G. Reis, R.B. Harrison, D.J. Firme, Effect of spacing on growth and biomass distribution in Eucalyptus camaldulensis, E. pellita and E. urophylla plantations in southeastern Brazil, For. Ecol. Manag. 104 (1) (1998) 1–13.