RESEARCH ARTICLE   **Open Access**

New Zealand Journal of Forestry Science

A comparison between traditional ordinary least-squares regression and three methods for enforcing additivity in biomass equations using a sample of *Pinus radiata* trees

Mohan KC[1,2*], Euan G. Mason[2], Horacio E. Bown[3], Grace Jones[2,4]

[1]Ministry of Industry, Tourism, Forests and Environment, Bagmati Province, Division Forest Office, Lalitpur, 44700, Nepal
[2]School of Forestry, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
[3]Faculty of Forestry, University of Chile, PO Box 9206, Santiago, Chile
[4]Department of Forestry and Wood Technology, Linnaeus University, PG Vejdes väg, 351 95 Växjö, Sweden

*Corresponding author: mohankc.forestry@gmail.com

Abstract

Background: Additivity has long been recognised as a desirable property of systems of equations to predict the biomass of components and the whole tree. However, most tree biomass studies report biomass equations fitted using traditional ordinary least-squares regression. Therefore, we aimed to develop models to estimate components, subtotals and above-ground total biomass for a *Pinus radiata* D.Don biomass dataset using traditional linear and nonlinear ordinary least-squares regressions, and to contrast these equations with the additive procedures of biomass estimation.

Methods: A total of 24 ten-year-old trees were felled to assess above-ground biomass. Two broad procedures were implemented for biomass modelling: (a) independent; and (b) additive. For the independent procedure, traditional linear models (LINOLS) with scaled power transformations and y-intercepts and nonlinear power models (NLINOLS) without y-intercepts were compared. The best linear (transformed) models from the independent procedure were further tested in three different additive structures (LINADD1, LINADD2, and LINADD3). All models were evaluated using goodness-of-fit statistics, standard errors of estimates, and residual plots.

Results: The LINOLS with scaled power transformations and y-intercepts performed better for all components, subtotals and total above-ground biomass in contrast to NLINOLS that lacked y-intercepts. The additive model (LINADD3) in a joint generalised linear least-squares regression, also called seemingly unrelated regression (SUR), provided the best goodness-of-fit statistics and residual plots for four out of six components (stem, branch, new foliage and old foliage), two out of three subtotals (foliage and crown), and above-ground total biomass compared to other methods. However, bark, cone and bole biomass were better predicted by the LINOLS method.

Conclusions: SUR was the best method to predict biomass for the 24-tree dataset because it provided the best goodness-of-fit statistics with unbiased estimates for 7 out of 10 biomass components. This study may assist silviculturists and forest managers to overcome one of the main problems when using biomass equations fitted independently for each tree component, which is that the sum of the biomasses of the predicted tree components does not necessarily add to the total biomass, as the additive biomass models do.

Keywords: above-ground; additive; biomass; linear; nonlinear; radiata pine; seemingly unrelated regression

Introduction

Forests play a vital role in the carbon cycle to mitigate climate change by accumulating and sequestering atmospheric carbon dioxide (CO_2) (Houghton 1991).

Pinus radiata D.Don, native to California, is a widely planted commercial tree species in the Southern Hemisphere, including New Zealand, Australia, Chile, Spain and South Africa (Lavery & Mead 2000; Mead

2013). It is grown primarily for timber production, as this species is versatile, fast-growing, and has a wide range of end uses (Lavery & Mead 2000; Lewis et al. 1993; Rogers 2002; Sutton 1999; Toro & Gessel 1999). The global plantation area of *P. radiata* is now more than 4.2 million hectares (Mead 2013). In New Zealand, it is the predominant planted species, and accounts for about 90% of the total 1.7 million hectares of forest plantations (Nixon et al. 2017). Plantation forests in New Zealand have not only been recognised as providing financial returns from traditional wood products, but also as providing environmental services by accumulating biomass and sequestering a substantial amount of carbon. To quantify such benefits, a precise biomass model with a required level of accuracy is essential.

Tree biomass estimation is required by scientists and practitioners alike as a surrogate of ecosystem production, product outturn and carbon accounting, among others. Biomass modelling is important for estimating carbon sequestration of forest ecosystems, as individual tree biomass or its components is aggregated to yield the stand biomass (Zheng et al. 2015). Allometric models are commonly used to assess the biomass accumulated in forests. Allometric relationships can be developed from destructive sampling by using several forms of regression equations. Generally, biomass equations are fitted in linear form using logarithmic transformation of B and D of the form, $B = aD^b$, where B is the biomass of the tree, or its components, and D is the diameter of the tree (Baskerville 1972; Beauchamp & Olson 1973; Canadell et al. 1988; Santa Regina et al. 1997; Sprugel 1983). Clutter et al. (1983) explained various linear and nonlinear additive regression models to estimate the biomass of an individual tree or its components. The additivity of biomass equations has long been recognised as a desirable property, so that predictions of tree components added together equals predictions of total tree biomass (Cunia & Briggs 1984, 1985; Parresol 1999, 2001). Three procedures for forcing additivity have been proposed (Cunia & Briggs 1985; Parresol 1999): (a) adding the best regression functions of the components' biomass to determine the total biomass regression function; (b) using the same independent variables for each component; and (c) using joint generalised least-squares regression, also known as seemingly unrelated regression (SUR), in which statistical dependencies among sample data are accounted for by forcing constraints on the regression coefficients. These three procedures have been extensively applied for estimating tree biomass around the world (Canga et al. 2013; Návar, González et al. 2004). The additive procedure in SUR has been mostly used for biomass modelling of single species (Cunia & Briggs 1984, 1985; Green & Reed 1985; Parresol 1999; Zheng et al. 2015).

Over the last 50 years, a substantial number of biomass studies for *P. radiata* have been undertaken in New Zealand (Beets & Madgwick 1988; Beets et al. 2007; Beets & Pollock 1987; Cromer et al. 1985; Madgwick 1983, 1985, 1994; Madgwick et al. 1977; Mead et al. 1984; Moore 2010; Webber & Madgwick 1983; Will

1964). Previous studies for *P. radiata* in New Zealand aimed to find the best biomass equations using various functional linear and nonlinear forms, with models generally fitted separately for each individual biomass component and for the whole tree. Separately calculated biomass equations ignore inherent correlation among the component equations measured on the same tree (Kozak 1970; Parresol 1999). Simultaneous fits with related equations using additive procedures have greater statistical efficiency, as they take into account statistical dependencies among biomass components in parameter estimation recorded from the same tree (Bi et al. 2010; Bi et al. 2004; Carvalho & Parresol 2003; Parresol 1999, 2001). Two country specific systems of additive biomass equations were developed for *P. radiata* using routinely measured stand variables from Australia and New Zealand (Bi et al. 2010). It has been noted that prediction accuracy varies across methodological differences and uncertainties associated over a range of stand variables (Bi et al. 2010; Moore 2010). As there is uncertainty about how to better meet additivity requirements, this study was undertaken to compare traditional linear and nonlinear ordinary least-squares regressions, and additive procedures in the estimation of tree component and total biomass for a dataset composed of 24 trees of *P. radiata*.

Methods

Study site and experiment
This study was carried out in the Canterbury region of New Zealand, planted with *P. radiata* in 2005, in a forestry trial designed to test the effect of stocking, genetics, fertiliser application, and follow-up weed control treatment on productivity and wood quality (Mason 2008). The site is located at latitude 43° 37.2′ S and longitude 172° 20.4′ E, and about 45 m above sea level on a flat landscape. The site has a mean annual air temperature between 11 and 13 °C with a monthly minimum (July) of −2 to +4 °C and a monthly maximum (January) of 20 to 23 °C (Macara 2016). Annual rainfall is about 618 mm with a monthly range between 38 and 68 mm (Macara 2016). The experiment consisted of 48 permanent plots with a randomised complete block split-split design, with the arrangement of factors within four complete blocks (Mason 2008). During the summer of 2015 to 2016, 24 ten-year-old trees of *P. radiata* were harvested and measured from six plots of the trial, and within each plot four trees were felled. These plots consisted of three levels of stocking (625, 1250 and 2500 stems ha⁻¹), two levels of follow-up weed control treatment (herbicide and no chemical treatment) and two clones (1 and 2).

Biomass data
Trees were felled at ground level. The over-bark diameter of each tree at breast height was recorded at 1.4 m. Total tree height was measured from ground level to the tip of the tree bole. For each tree, the components were separated into stem, branch, bark, foliage, and cones. Needles and twigs less than 1 cm in diameter were

considered foliage, and this was separated into "new" and "old" foliage. The total fresh mass of all components including subsamples were measured immediately after felling, using a portable balance. All the cones and small branches were weighed separately. The logs were separated into small pieces and weighed fresh in the field. A subsample of stem discs with bark (cut at the 1.4 m section and every 2 m upwards in the stem) and subsamples of all other components, were weighed to determine fresh weight in the field. Subsamples were dried in an oven at 70 °C until constant mass was achieved, and then this weight was recorded. Dry mass of each component was calculated as the fresh mass recorded in the field for that component multiplied by the ratio of subsample dry to fresh mass (Eq. 1):

$$Y_i = \frac{DW_i}{FW_i} \times TFW_i \qquad (1)$$

where Y is the total dry mass (kg), DW and FW refers to the sub sampled dry and fresh mass (kg) respectively, TFW is the total fresh mass (kg), and i is the tree component such as stem, bark, branch, new foliage, old foliage and cones.

Descriptive statistics of the trees including components, sub-total and above-ground total biomass are shown in Table 1. The notations and definitions used in this manuscript are explained in the Abbreviations section.

Variance stabilisation
Biomass data generally exhibit non-constant variance in model residuals (Parresol 1993, 2001). When developing predictive equations, variance can be stabilised either by providing a weight function or by using transformations (Parresol 1993, 2001). Curvilinearity and heterogeneity in variance of all linear models were reduced by transforming the response as well as explanatory variables using scaled-power transformations (Eq. 2), widely known as Box-Cox transformations (Box & Cox 1964). The predicted values of these models were back transformed to the original form using Eq. 3. A similar variance stabilisation technique was implemented by Zheng (2015) while using the additive procedure of biomass modelling for *Quercus variabilis* in northern China.

$$Y^{(\lambda)} = \begin{cases} \dfrac{Y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(Y), & \lambda = 0 \end{cases} \qquad (2)$$

$$Y' = \left((\lambda \times Y^{(\lambda)}) + 1 \right)^{\frac{1}{\lambda}} \qquad (3)$$

where $Y^{(\lambda)}$ is the transformed variable, and λ is a coefficient of the transformed variable that varies normally between −3 and +5 (Cook & Weisberg 2009), Y' is the back-transformed variable. A λ term is chosen to make the frequency distribution of each variable as close to normal as possible, thus promoting linear relationships and stabilising variance.

Model assessment and evaluation
In this study, a dataset consisting of 24 trees was used to evaluate the fitting bias, precision, and validity of models using the following goodness-of-fit statistics: root mean square error (RMSE), mean absolute bias (MAB), mean prediction error (MPE), residual standard error (RSE), coefficient of variation (CV), coefficient of determination (R²), index of agreement (IOA), and Akaike information criterion (AIC). Models were considered better with small AIC, RMSE, MAB, MPE, RSE, and CV of the residuals,

TABLE 1: Descriptive statistics for the 24-felled trees used for developing regression models, and components, sub-total and above-ground total biomass.

	[1]**Description**	**Max**	**Min**	**Mean**	[2]**SD**	[3]**CI (*P*= 95%)**
Tree variables	DBH (cm)	28	8.2	18.68	5.46	2.30
	H (m)	13.77	8.85	11.66	1.19	0.50
	CrL (m)	6	0.2	3.35	2.07	0.87
Components	Stem (kg tree⁻¹)	118.46	9.29	60.46	32.24	13.62
	Branch (kg tree⁻¹)	62.68	0.28	17.42	20.92	8.83
	OF (kg tree⁻¹)	34.84	0.94	13.18	11.63	4.91
	Bark (kg tree⁻¹)	11.92	0.65	5.18	3.19	1.35
	Cone (kg tree⁻¹)	16.96	0.05	3.65	3.87	1.63
	NF (kg tree⁻¹)	9.79	0.29	3.37	2.86	1.21
Subtotals	Bole (kg tree⁻¹)	128.69	9.94	65.64	35.33	14.92
	Crown (kg tree⁻¹)	123.63	3.12	37.60	37.29	15.75
	Foliage (kg tree⁻¹)	61.59	2.84	20.19	17.45	7.37
Total	AGT (kg tree⁻¹)	241.65	13.07	103.24	71.45	30.17

[1] Abbreviation details provided at the end of the text; [2]SD = standard deviation, [3]CI = confidence interval

and large R^2 and IOA. The interpretation of these fitting statistics can be found in Von Gadow and Hui (2001) and Goicoa et al. (2011). In addition, model performance was assessed by residual plots and histograms of residuals.

Modelling procedure

In this research, two procedures were implemented to estimate components, subtotals and above-ground total biomass: (1) independent; and (2) additive. All models were fitted to estimate biomass in terms of kg tree^{-1}.

Independent procedure for biomass estimation

In this procedure, biomass equations were fitted independently using traditional linear ordinary least-squares regressions with scaled power transformations and *y*-intercepts (denoted as, **LINOLS**; Eq. 4) and nonlinear ordinary least-squares power equations that lacked *y*-intercepts (denoted as, **NLINOLS**; Eq. 5). The mathematical specifications of these models are as follows (Parresol 1999; Zeng 2011; Zianis et al. 2005).

$$f_l(X_l, \beta_l) = \beta_{lo} + \beta_{l1}X_1 + \beta_{l2}X_2 + \dots.\beta_{lp}X_p + \varepsilon_l \quad (4)$$

$$f_l(X_l, \beta_l) = \beta_{lo}X_1^{\beta_{l1}}X_2^{\beta_{l2}} \dots \dots. X_p^{\beta_{lp}} + \varepsilon_l \quad (5)$$

where $f_l(X_l, \beta_l)$ is the regression function for the above ground biomass or one of its components, X_l are tree dimension variables such as D, H and CrL ($l = 1, 2, \dots, p$) while β_l denote the regression coefficients.

Each component equation contained its own independent variables. All components, subtotals and AGT biomass equations were fitted separately using the *lm* and *nls* function of R statistical software (R Core Team, 2018), for linear and nonlinear regressions, respectively.

Additive procedure of biomass estimation

In this procedure, biomass equations were fitted based on three additive procedures, described and compared by Parresol (1999, 2001). The additivity requirement to estimate total tree biomass is ensured by (a) adding the separately calculated best regression functions of each component, (b) using the same independent variables for each component, and (c) using joint generalised least-squares methods, also known as seemingly unrelated regression (SUR). In SUR, statistical dependencies among components are forced by constraining regression coefficients (Cunia & Briggs 1985; Parresol 1999). In
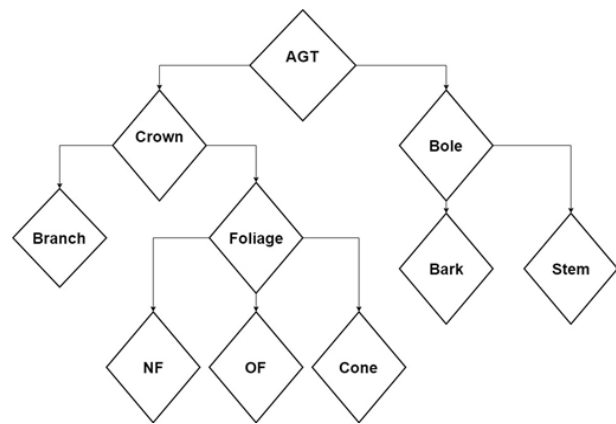

FIGURE 1: A statistical framework showing model structure with four restrictions (foliage, crown, bole and AGT) for biomass additivity.

this study, four restrictions were provided for the SUR model: (1) foliage; (2) crown; (3) bole; and (4) AGT, as illustrated in Figure 1. For example, foliage biomass is the sum of NF, OF and cone biomass (Eq. 6). Mathematically, the additive system of biomass equations in additive error terms with cross-equation correlation is specified in Eq. 6 where $\hat{Y}_i$ represents the predicted biomass of a given component and $f_i(X_i, \beta_i)$ is a regression function for the biomass component, (i = cone, new foliage, old foliage, branch, bark and stem, foliage, crown, bole and AGT biomass). The residual is ε_{ij} for the ith equation and j is an index for component. All additive biomass equations were fitted in the R statistical software (R Core Team 2018) using the *systemfit* package (Henningsen & Hamann 2007).

In the first additive procedure, the additivity was ensured by adding individually calculated best regression functions of each component to give a total biomass regression function (Cunia & Briggs 1985; Parresol 1999). The best regression functions obtained from the independent procedure of biomass modelling that were fitted separately for each component given in Table A.1 were used. The additive structure of this model, denoted as **LINADD1,** is specified in Eq. 7. In the second additive procedure, additivity was implemented by using the same explanatory variables for each component. For this, the most frequent independent variable (D) was selected from the best linear regression function as it was best fitted for stem, bark, foliage, bole, and AGT (Table A.1). Using D as an independent variable

$$\hat{Y}_{fol} = f_{co}(X_{co}, \beta_{co}) + f_{nf}(X_{nf}, \beta_{nf}) + f_{of}(X_{of}, \beta_{of}) + \varepsilon_{fol}$$

$$\hat{Y}_{cr} = f_{co}(X_{co}, \beta_{co}) + f_{nf}(X_{nf}, \beta_{nf}) + f_{of}(X_{of}, \beta_{of}) + f_{br}(X_{br}, \beta_{br}) + \varepsilon_{cr}$$

$$\hat{Y}_{bol} = f_{ba}(X_{ba}, \beta_{ba}) + f_{st}(X_{st}, \beta_{st}) + \varepsilon_{bol} \quad (6)$$

$$\hat{Y}_{AGT} = f_{co}(X_{co}, \beta_{co}) + f_{nf}(X_{nf}, \beta_{nf}) + f_{of}(X_{of}, \beta_{of}) + f_{br}(X_{br}, \beta_{br}) + f_{ba}(X_{ba}, \beta_{ba})$$

$$+ f_{st}(X_{st}, \beta_{st}) + \varepsilon_{AGT}$$

$$\hat{Y}_{co} = \beta_{10} + \beta_{11}D^2 + \varepsilon_{co}$$

$$\hat{Y}_{nf} = \beta_{20} + \beta_{21}D + \beta_{22}CrL + \varepsilon_{nf}$$

$$\hat{Y}_{of} = \beta_{30} + \beta_{31}D + \beta_{32}CrL^2 + \varepsilon_{of}$$

$$\hat{Y}_{br} = \beta_{40} + \beta_{41}D + \beta_{42}H + \varepsilon_{br}$$

$$\hat{Y}_{ba} = \beta_{50} + \beta_{51}D + \varepsilon_{ba}$$

$$\hat{Y}_{st} = \beta_{60} + \beta_{61}D + \varepsilon_{st}$$

$$\hat{Y}_{fol} = \left(\hat{Y}_{co} + \hat{Y}_{nf} + \hat{Y}_{of}\right)$$

$$= (\beta_{10} + \beta_{20} + \beta_{30}) + \beta_{11}D^2 + (\beta_{21} + \beta_{31})D + \beta_{22}CrL + \beta_{32}CrL^2 + \varepsilon_{fol}$$

$$\hat{Y}_{cr} = \left(\hat{Y}_{fol} + \hat{Y}_{br}\right)$$

$$= (\beta_{10} + \beta_{20} + \beta_{30} + \beta_{40}) + \beta_{11}D^2 + (\beta_{21} + \beta_{31} + \beta_{41})D + \beta_{22}CrL$$

$$+ \beta_{32}CrL^2 + \beta_{42}H + \varepsilon_{cr}$$

$$\hat{Y}_{bol} = \left(\hat{Y}_{ba} + \hat{Y}_{st}\right) = (\beta_{50} + \beta_{60}) + (\beta_{51} + \beta_{61})D + \varepsilon_{bol}$$

$$\hat{Y}_{AGT} = \left(\hat{Y}_{cr} + \hat{Y}_{bol}\right)$$

$$= (\beta_{10} + \beta_{20} + \beta_{30} + \beta_{40} + \beta_{50} + \beta_{60}) + \beta_{11}D^2$$

$$+ (\beta_{21} + \beta_{31} + \beta_{41} + \beta_{51} + \beta_{61})D + \beta_{22}CrL + \beta_{32}CrL^2 + \beta_{42}H + \varepsilon_{AGT}$$

(7)

$$\hat{Y}_{co} = \beta_{10} + \beta_{11}D + \varepsilon_{co}$$

$$\hat{Y}_{nf} = \beta_{20} + \beta_{21}D + \varepsilon_{nf}$$

$$\hat{Y}_{of} = \beta_{30} + \beta_{31}D + \varepsilon_{of}$$

$$\hat{Y}_{br} = \beta_{40} + \beta_{41}D + \varepsilon_{br}$$

$$\hat{Y}_{ba} = \beta_{50} + \beta_{51}D + \varepsilon_{ba}$$

$$\hat{Y}_{st} = \beta_{60} + \beta_{61}D + \varepsilon_{st}$$

$$\hat{Y}_{fol} = (\hat{Y}_{co} + \hat{Y}_{nf} + \hat{Y}_{of}) = (\beta_{10} + \beta_{20} + \beta_{30}) + (\beta_{11} + \beta_{21} + \beta_{31})D + \varepsilon_{fol}$$

$$\hat{Y}_{cr} = (\hat{Y}_{fol} + \hat{Y}_{br}) = (\beta_{10} + \beta_{20} + \beta_{30} + \beta_{40}) + (\beta_{11} + \beta_{21} + \beta_{31} + \beta_{41})D + \varepsilon_{cr}$$

$$\hat{Y}_{bol} = (\hat{Y}_{ba} + \hat{Y}_{st}) = (\beta_{50} + \beta_{60}) + (\beta_{51} + \beta_{61})D + \varepsilon_{bol}$$

$$\hat{Y}_{AGT} = \left(\hat{Y}_{cr} + \hat{Y}_{bol}\right)$$

$$= (\beta_{10} + \beta_{20} + \beta_{30} + \beta_{40} + \beta_{50} + \beta_{60}) + (\beta_{11} + \beta_{21} + \beta_{31} + \beta_{41} + \beta_{51} + \beta_{61})D$$

$$+ \varepsilon_{AGT}$$

(8)

$$\hat{Y}_{co} = \beta_{10} + \beta_{11}D^2 + \varepsilon_{co}$$

$$\hat{Y}_{nf} = \beta_{20} + \beta_{21}D + \beta_{22}CrL + \varepsilon_{nf}$$

$$\hat{Y}_{of} = \beta_{30} + \beta_{31}D + \beta_{32}CrL^2 + \varepsilon_{of}$$

$$\hat{Y}_{br} = \beta_{40} + \beta_{41}D + \beta_{42}H + \varepsilon_{br}$$

$$\hat{Y}_{ba} = \beta_{50} + \beta_{51}D^2H + \varepsilon_{ba}$$

$$\hat{Y}_{st} = \beta_{60} + \beta_{61}D + \varepsilon_{st}$$

$$\hat{Y}_{fol} = \left(\hat{Y}_{co} + \hat{Y}_{nf} + \hat{Y}_{of}\right) = (\beta_{10} + \beta_{20} + \beta_{30}) + \beta_{11}D^2 + (\beta_{21} + \beta_{31})D + \beta_{22}CrL + \beta_{32}CrL^2 + \varepsilon_{fol}$$

$$\hat{Y}_{cr} = \left(\hat{Y}_{fol} + \hat{Y}_{br}\right)$$

$$= (\beta_{10} + \beta_{20} + \beta_{30} + \beta_{40}) + \beta_{11}D^2 + (\beta_{21} + \beta_{31} + \beta_{41})D + \beta_{22}CrL + \beta_{32}CrL^2 + \beta_{42}H$$

$$+ \varepsilon_{cr}$$

$$\hat{Y}_{bol} = \left(\hat{Y}_{ba} + \hat{Y}_{st}\right) = (\beta_{50} + \beta_{60}) + \beta_{51}D^2H + \beta_{61}D + \varepsilon_{bol}$$

$$\hat{Y}_{AGT} = \left(\hat{Y}_{cr} + \hat{Y}_{bol}\right)$$

$$= (\beta_{10} + \beta_{20} + \beta_{30} + \beta_{40} + \beta_{50} + \beta_{60}) + \beta_{11}D^2 + (\beta_{21} + \beta_{31} + \beta_{41} + \beta_{61})D + \beta_{22}CrL$$

$$+ \beta_{32}CrL^2 + \beta_{42}H + \beta_{51}D^2H + \varepsilon_{AGT}$$

$$(9)$$

for all components, the additive structure of the model, denoted as **LINADD2**, is specified in Eq. 8.

In the third additive procedure, we used different explanatory variables in a joint generalised linear least-squares regression, known as SUR (Cunia & Briggs 1985; Parresol 1999). For this, best-fitted explanatory variables from the independent procedure of biomass modelling were used for stem, cone, branch, NF, and OF (Table A.1). We used the second-best regression D^2H as an independent variable for bark (data not shown). The additive structure of the model, denoted as **LINADD3,** is specified in Eq. 9.

Results

Comparison of fitted equations for components, subtotals and AGT

Tested LINOLS and NLINOLS equations with their best-fit results are given in Table A.1, and fitted statistics with their regression estimates are presented in Table A.2. We attempted to take into account follow-up herbicide, stocking, and clone factors into all models as dummy variables. These were found to be non-significant ($P>0.05$) so were discarded from all subsequent modelling. In comparison, LINOLS provided relatively higher R^2 values than NLINOLS for all, except for branch and cone biomass (Table A.2). However, plotting residuals with predicted values and with other variables

demonstrated that NLINOLS regression was unsuitable for these two components (data not shown). Therefore, overall, the best fitted LINOLS model according to goodness-of-fit statistics and residual plots were Eq. (i) for stem, bark, foliage, bole and AGT biomass, Eq. (ii) for cone biomass, Eq. (iii) for branch biomass, Eq. (viii) for NF and crown biomass, and Eq. (ix) for OF biomass (Table A.1). Finally, these selected LINOLS models were further tested in the additive process of biomass estimations.

The estimated coefficients for six components, three subtotals, and AGT using four methods (LINOLS, LINADD1, LINADD2 and LINADD3) are presented in Table 2 and their goodness of fit statistics are given in Table 3. The distribution of residuals with predicted values of the fitted best models for six components, three subtotals, and AGT are given in Fig. 2.

The LINADD3 fitted in SUR was considered best to predict stem (Eq. 10), branch (Eq. 11), NF (Eq. 13), OF (Eq. 14) biomass as it provided the better-fitting statistics when compared to the other three equations (Table 3). For stem, LINADD3 simultaneously decreased the RMSE, RSE, and CV by 0.1%, MPE by 0.2% while R^2 increased by 0.005%, compared to the other three equations. For branches, LINADD3 provided a marginal decrease in the goodness-of-fit statistics (e.g. RMSE, RSE, and CV by 1%) in contrast to LINOLS and LINADD1. For NF, LINADD3 model recorded a decrease in fitting statistics (e.g. RMSE by 3.7%, RSE by 1.43%), in contrast to the other three

TABLE 2: Regression model for each biomass component across modelling techniques. Table shows parameter estimates, their standard error between parentheses and significance indicated as: ns, non-significant, *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$. Box-Cox transformation values (λ) are also presented.

Components	Parameter estimates	Methods				λ value
		LINOLS	*LINADD1*	*LINADD2*	*LINADD3*	
Cone	β_{10}	−2.301 ** (0.769)	−2.225 ** (0.767)	−5.240 ** (1.483)	−2.229 ** (0.767)	$\lambda_d = 0.34$ $\lambda_{co} = 0.27$
	β_{11}	0.132 *** (0.029)	0.129 *** (0.029)	1.262 *** (0.296)	0.129 *** (0.029)	
NF	β_{20}	−4.959 *** (0.375)	−4.959 *** (0.375)	−4.959 *** (0.375)	−4.213 *** (0.663)	$\lambda_d = 0.34$ $\lambda_{crl} = 1.45$ $\lambda_{nf} = 0.07$
	β_{21}	1.184 *** (0.075)	1.184 *** (0.075)	1.184 ns (0.075)	1.059 *** (0.117)	
	β_{22}				−0.035 ns (0.026)	
OF	β_{30}	−4.089 *** (0.331)	−3.209 *** (0.336)	−4.089 *** (0.331)	−3.105 *** (0.350)	$\lambda_d = 0.34$ $\lambda_{crl} = 1.45$ $\lambda_{of} = 0.01$
	β_{31}	1.266 *** (0.066)	1.119 *** (0.063)	1.266 *** (0.066)	1.101 *** (0.066)	
	β_{32}		−0.006 *** (0.001)		−0.007 *** (0.002)	
Branch	β_{40}	−5.803 *** (0.666)	−5.786 *** (0.619)	−6.486 *** (0.681)	−5.758 *** (0.619)	$\lambda_d = 0.34$ $\lambda_{crl} = 1.67$ $\lambda_{br} = 0.04$
	β_{41}	1.981 *** (0.149)	1.987 *** (0.128)	(1.760) *** (0.136)	1.973 *** (0.128)	
	β_{42}	−0.049 * (0.019)	−0.051 *** (0.009)		−0.049 *** (0.009)	
Bark	β_{50}	−4.983 *** (0.373)	−4.983 *** (0.373)	−4.983 *** (0.373)	−3.340 *** (0.279)	$\lambda_{d2h} = 0.3$ $\lambda_{ba} = 0.36$
	β_{51}	1.416 *** (0.074)	1.416 *** (0.074)	0.373 *** (0.074)	0.146 *** (0.007)	
Stem	β_{60}	−6.691 *** (0.493)	−6.691 *** (0.493)	−6.691 *** (0.493)	−6.655 *** (0.492)	$\lambda_d = 0.34$ $\lambda_{st} = 0.38$
	β_{61}	3.264 *** (0.099)	3.264 *** (0.099)	3.264 *** (0.099)	3.257 *** (0.098)	
Foliage	β_{70}	−3.259 *** (0.393)				$\lambda_d = 0.34$ $\lambda_{fol} = 0.03$
	β_{71}	1.213 *** (0.079)				
Crown	β_{80}	−0.731 ns (0.510)				$\lambda_d = 0.34$ $\lambda_{crl} = 1.45$ $\lambda_{cr} = −0.13$
	β_{81}	0.693 *** (0.089)				
	β_{82}	−0.046 * (0.021)				
Bole	β_{90}	−7.002 *** (0.510)				$\lambda_d = 0.34$ $\lambda_{bol} = 0.38$
	β_{91}	3.403 *** (0.102)				
AGT	β_{100}	−1.182 *** (0.244)				$\lambda_d = 0.34$ $\lambda_{AGT} = 0.08$
	β_{101}	1.309 *** (0.049)				

Note: The λ value shown in the table indicates that the variables were subjected to a scaled power transformation. The estimated parameter values for each technique are presented in power-transformed scale.

TABLE 3: Goodness-of-fit statistics for given biomass components using four methods of modelling.

Biomass	Model	RMSE	MAB	MPE	RSE	CV	R²	IOA	RANK
Stem	LINOLS	4.864	3.391	23.660	5.080	8.046	0.976	0.994	3
	LINADD1	4.864	3.391	23.660	5.080	8.046	0.976	0.994	4
	LINADD2	4.864	3.391	23.659	5.080	8.046	0.976	0.994	2
	LINADD3	4.859	3.406	23.611	5.075	8.038	0.976	0.994	1
Branch	LINOLS	8.034	4.935	64.538	8.588	46.124	0.846	0.962	2
	LINADD1	8.046	4.97	64.733	8.601	46.194	0.846	0.962	3
	LINADD2	11.168	6.658	124.713	11.664	64.118	0.703	0.909	4
	LINADD3	7.958	4.848	63.322	8.507	45.688	0.849	0.962	1
Bark	LINOLS	0.880	0.604	0.774	0.919	16.987	0.921	0.979	1
	LINADD1	0.879	0.604	0.774	0.919	16.987	0.921	0.979	3
	LINADD2	0.879	0.604	0.774	0.919	16.987	0.921	0.979	1
	LINADD3	0.899	0.613	0.809	0.939	17.374	0.917	0.979	4
NF	LINOLS	0.896	0.609	0.803	0.936	26.627	0.898	0.973	2
	LINADD1	0.896	0.609	0.803	0.936	26.627	0.898	0.973	4
	LINADD2	0.896	0.609	0.803	0.936	26.627	0.898	0.973	3
	LINADD3	0.863	0.599	0.745	0.923	25.642	0.905	0.975	1
OF	LINOLS	2.879	2.109	8.289	3.007	21.852	0.936	0.985	3
	LINADD1	2.486	1.758	6.181	2.658	18.869	0.952	0.989	2
	LINADD2	2.879	2.109	8.289	3.007	21.852	0.936	0.985	3
	LINADD3	2.436	1.727	5.935	2.604	18.490	0.954	0.989	1
Cone	LINOLS	2.743	1.777	7.522	2.865	75.229	0.475	0.781	1
	LINADD1	2.760	1.783	7.619	2.883	75.710	0.468	0.774	3
	LINADD2	2.811	1.785	7.904	2.936	77.113	0.449	0.753	4
	LINADD3	2.759	1.783	7.613	2.882	75.681	0.469	0.775	2
Foliage	LINOLS	4.561	3.179	20.800	4.764	22.592	0.929	0.981	2
	LINADD1	4.509	3.194	20.327	4.939	22.334	0.930	0.982	2
	LINADD2	4.648	3.307	21.606	4.855	23.026	0.926	0.981	4
	LINADD3	4.478	3.173	20.048	4.905	22.180	0.931	0.982	1
Crown	LINOLS	15.359	9.019	235.915	16.043	40.845	0.823	0.955	4
	LINADD1	10.146	6.633	102.932	11.403	26.979	0.923	0.981	2
	LINADD2	14.044	8.408	197.219	14.668	37.346	0.852	0.959	3
	LINADD3	10.023	6.570	100.466	11.265	26.655	0.925	0.981	1
Bole	LINOLS	5.287	3.765	27.951	5.522	8.055	0.977	0.994	1
	LINADD1	5.295	3.762	28.035	5.530	8.067	0.987	0.994	4
	LINADD2	5.295	3.762	28.035	5.530	8.067	0.987	0.994	2
	LINADD3	5.293	3.799	28.014	5.658	8.064	0.987	0.994	3
AGT	LINOLS	17.131	11.419	293.497	17.894	16.594	0.940	0.985	3
	LINADD1	14.704	9.702	216.201	16.526	14.242	0.956	0.989	2
	LINADD2	17.405	11.171	302.933	18.179	16.859	0.938	0.984	4
	LINADD3	14.667	9.708	215.132	16.485	14.207	0.956	0.989	1

Note: RANK indicates the model's performance in comparison. For example, a model in RANK 1 is a best and RANK 4 is a worst in terms of goodness-of-fit statistics, and residual plots.
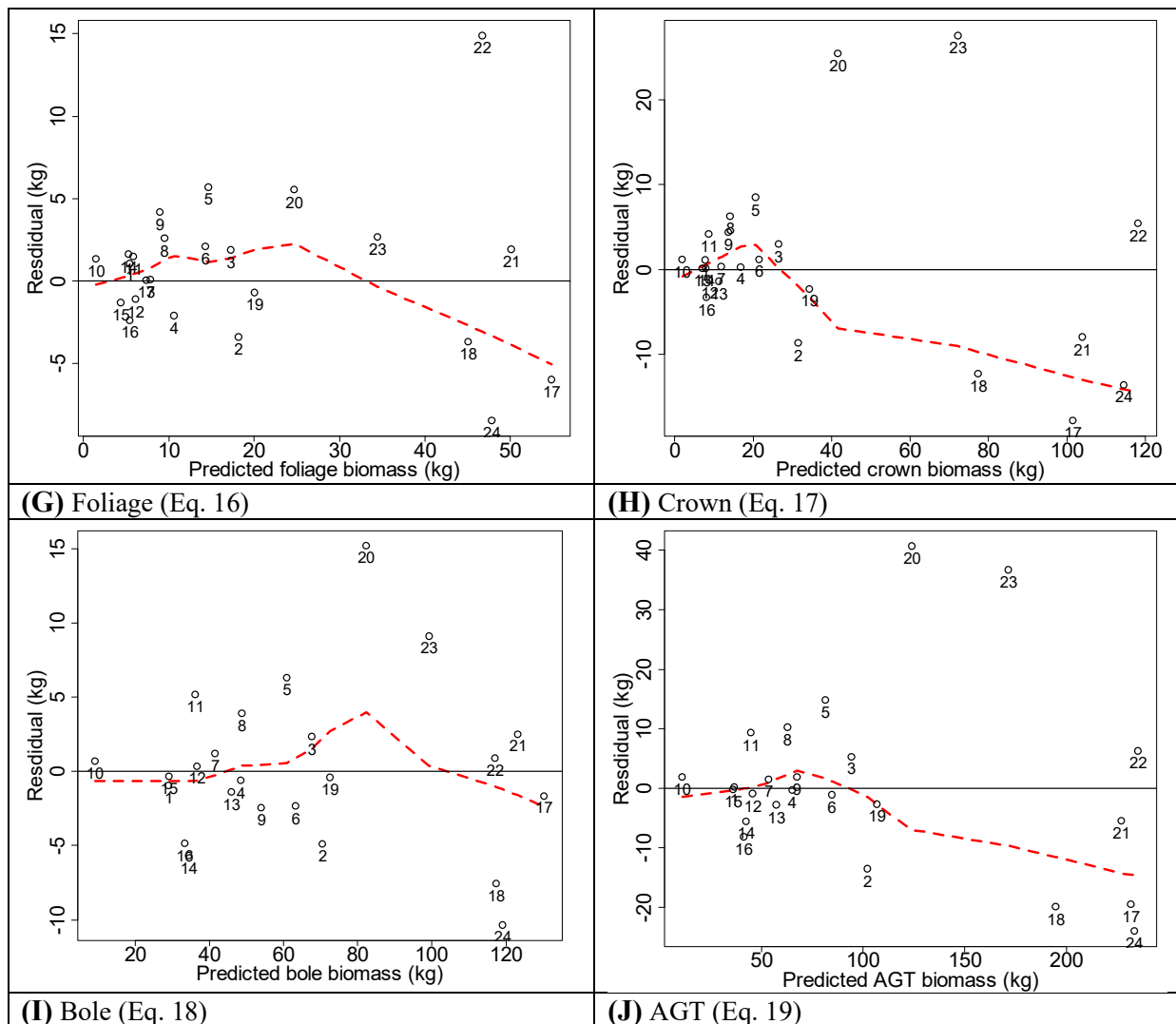
(A) Stem (Eq.10)

(B) Branch (Eq.11)

(C) Bark (Eq. 12)

(D) NF (Eq. 13)

(E) OF (Eq. 14)

(F) Cone (Eq. 15)
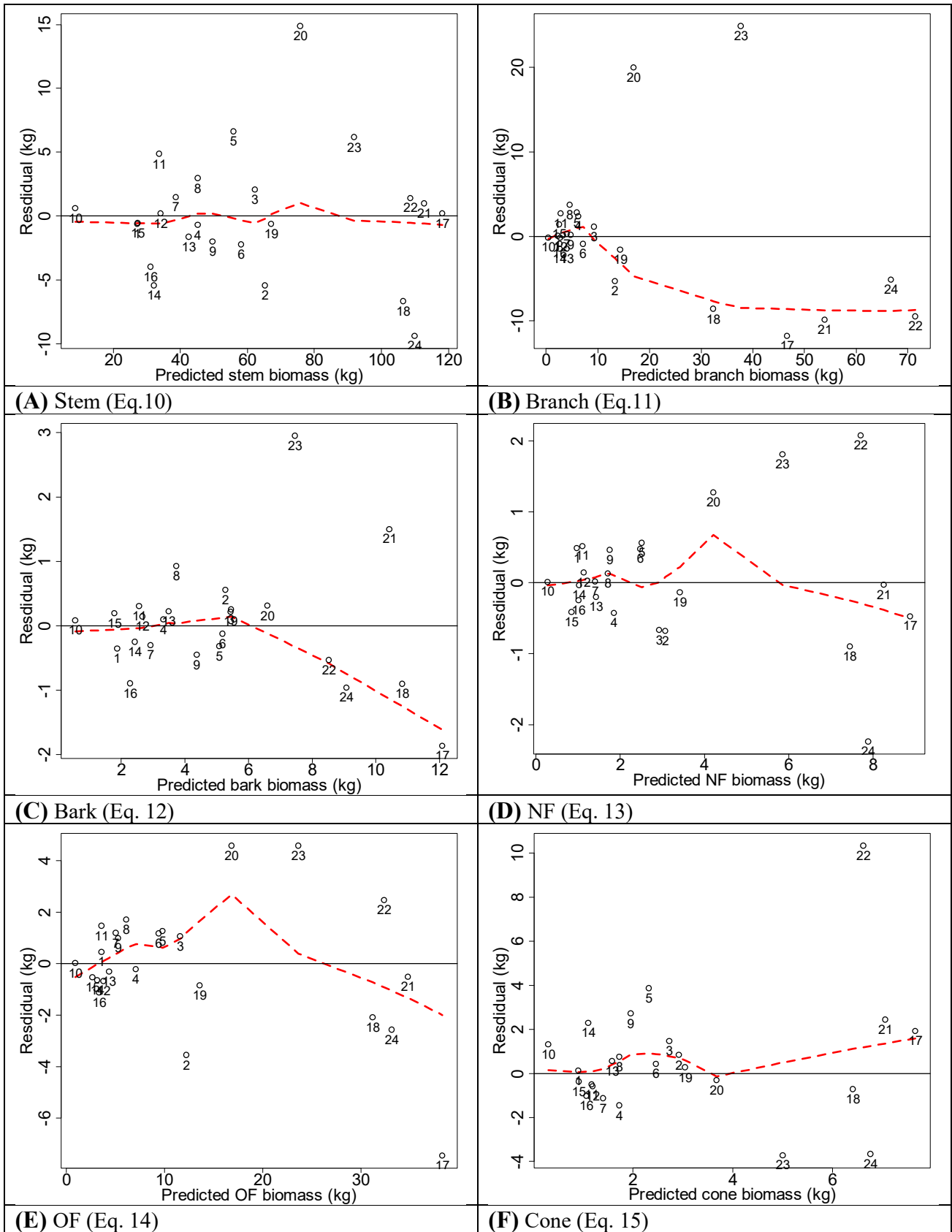
FIGURE 2: Residuals vs predicted values of biomass for the selected best model for each biomass compponent (A–J). The solid black horizontal line across zero represent baseline and the dotted red line is LOESS curve.

equations. For OF, the LINADD3 provided a decrease in RMSE (by 10.9%), MAB (by 12.7%), MPE (by 20.3%), RSE (by 9.6%) and CV (by 10.9%), and an increase in R^2 (by 1.2%) and IOA (by 0.3%), in contrast to the other three equations (Table 3). The LINOLS was considered the best to predict bark (Eq. 12), cone (Eq. 15) and bole (Eq. 12) biomass exhibiting relatively good fit statistics as compared to SUR models (Table 3). For bark, LINOLS simultaneously decreased RMSE, MAB, MPE, RSE and CV by 0.7%, 0.5%, 1.5%, 0.7%, and 0.7%, respectively, and it increased R^2 by 0.1% in contrast to the other three additive SUR equations (Table 3). For cone, the LINOLS equation provided a compatible decrease in RMSE, RSE, and CV by 1.2%; MAB decreased by 0.4%, MPE decreased by 2.4%, R^2 increased by 3% and IOA increased by 1.9%, in contrast to the other three additive equations tested. For bole, the LINOLS indicated there was a marginal decrease in RMSE (0.1%), MPE (0.2%), RSE (0.1%) and CV (0.1%), compared to three additive equations (Table 3).

The additive equation fitted in SUR was considered the best to predict foliage (Eq. 16) and crown (Eq. 17)

biomass as LINADD3 provided better fit statistics (Table 3). For foliage, LINADD3 provided a decrease in RMSE, MAB, MPE and CV by 2.1%, 1.6%, 4.1% and 2.1%, respectively, and R^2 by 0.3% and IOA by 0.1% compared to other three equations (Table 3). For crown, using LINADD3, average fitting statistics decreased by 21.5% for RMSE, 16.7% for MAB, 36.3% for MPE, 18.1% for RSE, and 21.5% for CV; and increased R^2 by 7% and IOA by 1.7%, in contrast to other three equations. The LINADD3 fitted in SUR (Eq. 19) was considered the best to predict AGT biomass as it provided a decrease in RMSE by 10.1%, MAB by 9.3%, MPE by 18.7, RSE by 5.8% and CV by 10.1%; and an increase in R^2 by 1.2% and IOA by 0.3%, in contrast to the other three methods.

Discussion
In this study, follow-up herbicide, stocking, and clone factors fitted into models as dummy variables were all non-significant. The possible reason for this insignificance could be that the site or silvicultural effects

Biomass	Fitted formula	Equation
Stem	$$\hat{Y}_{st} = \left(\lambda_{st} \times \left(\beta_{0_{st}} + \beta_{1_{st}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right) \right) + 1 \right)^{\frac{1}{\lambda_{st}}}$$	10
Branches	$$\hat{Y}_{br} = \left(\lambda_{br} \times \left(\beta_{0_{br}} + \beta_{1_{br}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right) + \beta_{2_{br}} \left(\frac{H^{\lambda_h} - 1}{\lambda_h} \right) \right) + 1 \right)^{\frac{1}{\lambda_{br}}}$$	11
Bark	$$\hat{Y}_{ba} = \left(\lambda_{ba} \times \left(\beta_{0_{ba}} + \beta_{1_{ba}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right) \right) + 1 \right)^{\frac{1}{\lambda_{ba}}}$$	12
New foliage	$$\hat{Y}_{nf} = \left(\lambda_{nf} \times \left(\beta_{0_{nf}} + \beta_{1_{nf}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right) + \beta_{2_{nf}} \left(\frac{CrL^{\lambda_{crl}} - 1}{\lambda_{crl}} \right) \right) + 1 \right)^{\frac{1}{\lambda_{nf}}}$$	13
Old foliage	$$\hat{Y}_{of} = \left(\lambda_{of} \times \left(\beta_{0_{of}} + \beta_{1_{of}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right) + \beta_{2_{of}} \left(\frac{CrL^{\lambda_{crl}} - 1}{\lambda_{crl}} \right)^2 \right) + 1 \right)^{\frac{1}{\lambda_{of}}}$$	14
Cone	$$\hat{Y}_{co} = \left(\lambda_{co} \times \left(\beta_{0_{co}} + \beta_{1_{co}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right)^2 \right) + 1 \right)^{\frac{1}{\lambda_{co}}}$$	15
Foliage	$$\hat{\hat{Y}}_{fol} = \left(\hat{Y}_{co} + \hat{Y}_{nf} + \hat{Y}_{of} \right)$$	16
Crown	$$\hat{Y}_{cr} = \left(\hat{Y}_{fol} + \hat{Y}_{br} \right)$$	17
Bole	$$\hat{Y}_{bol} = \left(\lambda_{bol} \times \left(\beta_{0_{bol}} + \beta_{1_{bol}} \left(\frac{D^{\lambda_d} - 1}{\lambda_d} \right) \right) + 1 \right)^{\frac{1}{\lambda_{bol}}}$$	18
Above-ground total	$$\hat{Y}_{AGT} = \left(\hat{Y}_{cr} + \hat{Y}_{bol} \right)$$	19

were explained by the tree size variables because the silvicultural treatments may simply speed up/slow the growth of the trees (Beets & Pollock 1987; Moore 2010), as opposed to changing their allometry. The biomass of *P. radiata* plantation changes with a wide range of climatic, edaphic, silvicultural and genetic factors (Beets & Pollock 1987; Bi et al. 2010; Moore 2010). Another reason of this lack of significance could be the small sample size (Duncanson et al. 2015) as the parameterisation of allometric equations depends significantly on the size of sample.

In this study, the best model of each component selected from independent procedures provided the logical base equations for further tests in the additivity of biomass equations. The same approach was also used by other researchers (Magalhães & Seifert 2015; Návar et al. 2002) to utilise additive properties in biomass modelling. Biomass additivity reduces the discrepancy between the sum of predicted values for components and those for a total tree (Kozak 1970), and it has long been documented as a desirable property of systems of equations to predict total tree biomass (Bi et al. 2004; Parresol 2001). Three procedures were implemented for the additivity in the biomass model (Parresol 1999, 2001): (1) using a separately calculated best linear function of the biomass of the components (best linear functions were D, D and H, D, D and CrL, D and CrL^2, and D^2 for stem, branch, bark, NF, OF, and cone biomass, respectively); (2) using the most frequently observed predictor (D) as the same independent variable for all components; and (3) using different independent variables for each component by forcing four linear restrictions on the regression coefficients, the SUR technique. The additivity of biomass equations to predict biomass of components, subtotal, and AGT has been explained in some other studies (Carvalho &

Parresol 2003; Magalhães & Seifert 2015; Návar et al. 2002; Návar, Méndez et al. 2004; Parresol 1999, 2001; Zhao et al. 2015).

A linear SUR model (LINADD3) of this study provided better results than LINOLS, LINADD1, and LINADD2, in terms of goodness-of-fit statistics, standard error of estimates and residual plots. LINADD3 fitted in SUR was superior to the other two additive models since it considered the correlation between each component equation, and provided greater statistical efficiencies (Carvalho & Parresol 2003). In contrast to our results, a study reported that the additive model (denoted as CON) that used the same independent variable for all components, similar to our LINADD2 model, was statistically superior to the linear and nonlinear SUR model with the different independent variables in parameter restriction (Magalhães & Seifert 2015). However, the authors (Magalhães & Seifert 2015) indicated that the CON method had the limitation that it did not take into account the correlations among plant parts.

Applying SUR to the system of additive equations with the same explanatory variables for each component does not provide precise estimation of biomass (Srivastava & Giles 1987). Therefore, the LINADD3, a SUR model that consisted of different explanatory variables for each component is consistent with that of Srivastava and Giles (1987), which was more effective than the other two additive models. Model LINADD1 also consisted of the same explanatory variables for two-component equations such as stem and bark, and LINADD2 consisted of the same independent variable for all component equations. Therefore, LINADD1 and LINADD2 were not effective compared to LINADD3. In addition, the individually calculated best equations for each component (LINOLS) provided the least efficient biomass estimates for all components except for bark, cone and bole biomass, compared with the linear SUR model (LINADD3). Researchers recommended using SUR to estimate biomass as it provides greater statistical efficiency than separately calculated equations for each component (Bi et al. 2010; Bi et al. 2004; Kozak 1970; Návar et al. 2002; Návar, Méndez et al. 2004; Parresol 2001).

Although this study focused on testing different procedures to fit biomass equations and highlighted the importance of SUR, there could be concerns over the applicability of the resulting models given the size of the dataset from the trial. While applying these models, it is advisable to consider that the dataset used was relatively small with only 24 trees of the same age class, with D (8.2 to 28 cm), H (8.85 to 13.77 m), and CrL (0.2 to 6 m). Small sample sizes may provide biased estimates, as the allometric parameters are sample size dependent (Duncanson et al. 2015). In addition, the models developed were based on only two out of five clones, selected across a range of treatment plots. Extrapolation should be avoided as uncertain prediction errors are expected from the selected models. The site characteristics (Duncanson et al. 2015; Madgwick 1994; Mason 2008), growth inputs (Dong et al. 2015), tree age

(Beets & Madgwick 1988; Beets & Pollock 1987), and nutrients and silvicultural practices (Beets & Madgwick 1988; Madgwick 1985; Mead et al. 1984) may influence the biomass allometry.

Conclusions

Two procedures for biomass modelling were compared, namely, independent and additive. For the independent procedure of biomass modelling, LINOLS models with scaled power transformations and *y*-intercepts and NLINOLS power models that lacked *y*-intercepts were compared for six components, three subtotals, and AGT biomass. The LINOLS models with scaled power transformations and *y*-intercepts provided superior results in contrast to NLINOLS power models without y-intercepts. The best-fitted component equations from LINOLS models were further tested in an additive procedure. A system of additive equations in SUR with different independent variables for each component (LINADD3) showed better performance than LINOLS, LINADD1, and LINADD2.

Besides, the linear SUR model provided comparatively unbiased estimates for stem (Eq. 10), branch (Eq. 11), NF (Eq. 13), OF (Eq. 14), foliage (Eq. 16), crown (Eq. 17), and AGT (Eq. 19), while LINOLS showed comparatively better fitting statistics for bark (Eq. 12), cone (Eq. 15) and bole biomass (Eq. 18) for the dataset of this study. Since seven out of ten biomass components were well fitted with SUR that provided lower variance by taking account of the existence of correlations among residuals of the component equations, we suggest that SUR could be a superior method for fitting biomass equations.

List of terms and abbreviations

D	Diameter at breast height (cm)
H	Total tree height (m)
D^2H	Product of H and D^2 ($cm^2.m$)
CrL	Crown length (m)
NF	New foliage
OF	Old foliage
Foliage	Sum of cone, NF and OF biomass (kg)
Crown	Sum of foliage and branch biomass (kg)
Bole	Sum of stem and bark biomass (kg)
Components	Cone, NF, OF, bark, branch, stem
Subtotal	Foliage, crown and bole
AGT	Sum of all components (above-ground total biomass) in kg
λ	Variable-specific transformation coefficient
β	Variable-specific parameter estimate
$\hat{Y}_{co}, \hat{Y}_{nf}, \hat{Y}_{of}, \hat{Y}_{br},$ $\hat{Y}_{ba}, \hat{Y}_{st}, \hat{Y}_{fol}, \hat{Y}_{cr},$ $\hat{Y}_{bol}, \hat{Y}_{AGT}$	Predicted biomass in kg for cone, NF, OF, branch, bark, stem, foliage, crown, bole and AGT, respectively.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Acknowledgements

References

Baskerville, G. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research, 2*(1), 49-53. https://doi.org/10.1139/x72-009

Beauchamp, J.J., & Olson, J.S. (1973). Corrections for bias in regression estimates after logarithmic transformation. *Ecology, 54*(6), 1403-1407. https://doi.org/10.2307/1934208

Beets, P., & Madgwick, H. (1988). Above-ground dry matter and nutrient content of Pinus radiata as affected by lupin, fertiliser, thinning, and stand age. *New Zealand Journal of Forestry Science, 18*(1), 43-64.

Beets, P., Pearce, S., Oliver, G., & Clinton, P. (2007). Root/shoot ratios for deriving below-ground biomass of Pinus radiata stands. *New Zealand Journal of Forestry Science, 37*(2), 267.

Beets, P., & Pollock, D. (1987). Accumulation and partitioning of dry matter in *Pinus radiata* as related to stand age and thinning. *New Zealand Journal of Forestry Science*, 17(2), 246-271.

Bi, H., Long, Y., Turner, J., Lei, Y., Snowdon, P., Li, Y., Harpur, R., Zerihun, A., & Ximenes, F. (2010). Additive prediction of aboveground biomass for *Pinus radiata* D. Don plantations. *Forest Ecology and Management, 259*(12), 2301-2314. https://doi.org/10.1016/j.foreco.2010.03.003

Bi, H., Turner, J., & Lambert, M.J. (2004). Additive biomass equations for native eucalypt forest trees of temperate Australia. *Trees, 18*(4), 467-479. https://doi.org/10.1007/s00468-004-0333-z

Box, G.E., & Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological),* 211-252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Canadell, J., Riba, M., & Andres, P. (1988). Biomass equations for *Quercus ilex* L. in the Montseny Massif, Northeastern Spain. *Forestry: An International Journal of Forest Research, 61*(2), 137-147. https://doi.org/10.1093/forestry/61.2.137

Canga, L., Aranda, U., Khouri, A., & Obregón, C. (2013). Above-ground biomass equations for *Pinus radiata* D. Don in Asturias. *Forest Systems, 22*(3), 408-415. https://doi.org/10.5424/fs/2013223-04143

Carvalho, J.P., & Parresol, B.R. (2003). Additivity in tree biomass components of Pyrenean oak (*Quercus pyrenaica* Willd.). *Forest Ecology and Management, 179*(1-3), 269-276. https://doi.org/10.1016/S0378-1127(02)00549-2

Clutter, J.L., Fortson, J.C., Pienaar, L.V., Brister, G.H., & Bailey, R.L. (1983). *Timber management: a quantitative approach.* New York: John Wiley & Sons, Inc.

Cook, R., & Weisberg, S. (2009). *Applied regression including computing and graphics.* New York: John Wiley & Sons.

Cromer, R., Barr, N., Williams, E., & McNaught, A. (1985). Response to fertiliser in a *Pinus radiata* plantation. 1. Above-ground biomass and wood density. *New Zealand Journal of Forestry Science, 15*(1), 59-70.

Cunia, T., & Briggs, R. (1984). Forcing additivity of biomass tables: some empirical results. *Canadian Journal of Forest Research*, *14*(3), 376-384. https://doi.org/10.1139/x84-067

Cunia, T., & Briggs, R. (1985). Forcing additivity of biomass tables: use of the generalized least squares method. *Canadian Journal of Forest Research, 15*(1), 23-28. https://doi.org/10.1139/x85-006

Dong, L., Zhang, L., & Li, F. (2015). Developing additive systems of biomass equations for nine hardwood species in Northeast China. *Trees, 29*(4), 1149-1163. https://doi.org/10.1007/s00468-015-1196-1

Duncanson, L., Rourke, O., & Dubayah, R. (2015). Small sample sizes yield biased allometric equations in temperate forests. *Scientific reports, 5*, 17153. https://doi.org/10.1038/srep17153

Goicoa, T., Militino, A., & Ugarte, M. (2011). Modelling aboveground tree biomass while achieving the additivity property. *Environmental and Ecological Statistics, 18*(2), 367-384. https://doi.org/10.1007/s10651-010-0137-9

Green, E.J., & Reed, D.D. (1985). Compatible tree volume and taper functions for pitch pine. *Northern Journal of Applied Forestry, 2*(1), 14-16. https://doi.org/10.1093/njaf/2.1.14

Henningsen, A., & Hamann, J. (2007). systemfit: A Package for Estimating Systems of Simultaneous Equations in R. *Journal of Statistical Software, 23*(4), 1-40. https://doi.org/10.18637/jss.v023.i04

Houghton, R.A. (1991). Tropical deforestation and atmospheric carbon dioxide. *Climatic Change, 19*(1), 99-118. https://doi.org/10.1007/BF00142217

Kozak, A. (1970). Methods for ensuring additivity of biomass components by regression analysis. *The Forestry Chronicle, 46*(5), 402-405. https://doi.org/10.5558/tfc46402-5

Lavery, P.B., & Mead, D.J. (2000). *Pinus radiata*: A narrow endemic from North America takes on the world.

In: D.M. Richardson (Ed.), *Ecology and Biogeography of* Pinus (pp. 432-449). Cambridge UK: Cambridge University Press.

Lewis, N.B., Ferguson, I.S., Sutton, W., Donald, D., & Lisboa, H. (1993). *Management of radiata pine*. Melbourne: Inkata Press Pty Ltd/Butterworth-Heinemann.

Macara, G.R. (2016). *The climate and weather of Canterbury.* NIWA Science and Technology Series, Taihoro Nukurangi. Retrieved from https://www. niwa.co.nz/static/web/canterbury_climatology_ second_ed_niwa.pdf.

Madgwick, H. (1983). Estimation of the oven-dry weight of stems, needles, and branches of individual *Pinus radiata* trees. *New Zealand Journal of Forestry Science, 13*(1), 108-109.

Madgwick, H. (1985). Dry matter and nutrient relationships in stands of *Pinus radiata. New Zealand Journal of Forestry Science, 15*(3), 324-336.

Madgwick, H. (1994). Pinus radiata: *biomass, form and growth*. Rotorua NZ: HAI Madgwick.

Madgwick, H., Jackson, D.S., & Knight, P. (1977). Above-ground dry matter, energy, and nutrient contents of trees in an age series of *Pinus radiata* plantations. *New Zealand Journal of Forestry Science, 7*(3), 445-468.

Magalhães, T.M., & Seifert, T. (2015). Biomass modelling of *Androstachys johnsonii* Prain: A comparison of three methods to enforce additivity. *International Journal of Forestry Research:* 878402. https://doi. org/10.1155/2015/878402

Mason, E. (2008). Influences of silviculture, genetics and environment on radiata pine corewood properties: Results from recent studies and a future direction. *New Zealand Journal of Forestry, 53*(2), 26-31.

Mead, D. (2013). *Sustainable management of* Pinus radiata *plantations*. FAO Forestry Paper No. 170. Rome: FAO.

Mead, D., Draper, D., & Madgwick, H. (1984). Dry matter production of a young stand of *Pinus radiata*: Some effects of nitrogen fertiliser and thinning. *New Zealand Journal of Forestry Science, 14*(1), 97-108.

Moore, J. (2010). Allometric equations to predict the total above-ground biomass of radiata pine trees. *Annals of Forest Science, 67*(8), 806. https://doi. org/10.1051/forest/2010042

Návar, J., González, N., Graciano, L., Dale, V., & Parresol, B. (2004). Additive biomass equations for pine species of forest plantations of Durango, Mexico. *Madera y Bosques, 10*(2).

Návar, J., Méndez, E., & Dale, V. (2002). Estimating stand biomass in the Tamaulipan thornscrub of northeastern Mexico. *Annals of Forest Science, 59*(8), 813-821. https://doi.org/10.1051/forest:2002079

Návar, J., Méndez, E., Nájera, A., Graciano, J., Dale, V., & Parresol, B. (2004). Biomass equations for shrub species of Tamaulipan thornscrub of North-eastern Mexico. *Journal of Arid Environments, 59*(4), 657-674. https://doi.org/10.1016/j.jaridenv.2004.02.010

Nixon, C., Gamperle, D., Pambudi, D., & Clough, P. (2017). *Plantation forestry statistics: Contribution of forestry to New Zealand*. Wellington, NZ.

Parresol, B. (1993). Modeling multiplicative error

variance: an example predicting tree diameter from stump dimensions in bald cypress. *Forest Science, 39*(4), 670-679.

Parresol, B. (1999). Assessing tree and stand biomass: a review with examples and critical comparisons. *Forest Science, 45*(4), 573-593.

Parresol, B. (2001). Additivity of nonlinear biomass equations. *Canadian Journal of Forest Research, 31*(5), 865-878. https://doi.org/10.1139/x00-202

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https:// www.R-project.org/.

Rogers, D.L. (2002). In situ genetic conservation of Monterey pine (*Pinus radiata* D. Don): Information and recommendations. *New Zealand Journal of Forestry Science, 32*(23), 395-398.

Santa Regina, I., Tarazona, T., & Calvo, R. (1997). Aboveground biomass in a beech forest and a Scots pine plantation in the Sierra de la Demanda area of Northern Spain. *Annals of Forest Science, 54*(3), 261-269. https://doi.org/10.1051/forest:19970304

Sprugel, D. (1983). Correcting for bias in log-transformed allometric equations. *Ecology, 64*(1), 209-210. https://doi.org/10.2307/1937343

Srivastava, V.K., & Giles, D.E. (1987). *Seemingly unrelated regression equations models: estimation and inference* (Vol. 80): CRC Press.

Sutton, W.R.J. (1999). The need for planted forests and the example of radiata pine. *New Forests, 17*(1), 95-110. https://doi.org/10.1023/A:1006567221005

Toro, J., & Gessel, S. (1999). Radiata pine plantations in Chile. *New Forests, 18*(1), 33-44. https://doi. org/10.1023/A:1006597823190

Von Gadow, K., & Hui, G. (2001). *Modelling forest development* (Vol. 57): Springer Science & Business Media.

Webber, B., & Madgwick, H. (1983). Biomass and nutrient content of 29-year-old *Pinus radiata* stand. *New Zealand Journal of Forestry Science, 13*(2), 222-228.

Will, G. (1964). Dry matter production and nutrient uptake by *Pinus radiata* in New Zealand. *The Commonwealth Forestry Review*, 57-70.

Zeng, W. (2011). *Methodology on modeling of single-tree biomass equations for national biomass estimation in China*. Doctoral Dissertation. Chinese Academy of Forestry, Beijing.

Zhao, D., Kane, M., Markewitz, D., Teskey, R., & Clutter, M. (2015). Additive tree biomass equations for midrotation loblolly pine plantations. *Forest Science, 61*(4), 613-623. https://doi.org/10.5849/ forsci.14-193

Zheng, C., Mason, E., Jia, L., Wei, S., Sun, C., & Duan, J. (2015). A single-tree additive biomass model of *Quercus variabilis* Blume forests in North China. *Trees, 29*(3), 705-716. https://doi.org/10.1007/s00468-014-1148-1

Zianis, D., Muukkonen, P., Mäkipää, R., & Mencuccini, M. (2005). Biomass and stem volume equations for tree species in Europe. *Silva Fennica Monographs* No. 4, 63 p.

APPENDIX

TABLE A.1: Tested linear and nonlinear ordinary least-squares equations with their best-fit results.

	LINOLS	
Equation	**Tested models**	**Fitted best with**
i	$\hat{Y} = \beta_0 + \beta_1 D$	Stem, Bark, Foliage, Bole, AGT
ii	$\hat{Y} = \beta_0 + \beta_1 D^2$	Cone
iii	$\hat{Y} = \beta_0 + \beta_1 D + \beta_2 H$	Branch
iv	$\hat{Y} = \beta_0 + \beta_1 D^2 H$	
v	$\hat{Y} = \beta_0 + \beta_1 D^2 H + \beta_2 H^2$	
vi	$\hat{Y} = \beta_0 + \beta_1 D^2 H + \beta_2 H^4$	
vii	$\hat{Y} = \beta_0 + \beta_1 D^2 H + \beta_2 D^2$	
viii	$\hat{Y} = \beta_0 + \beta_1 D + \beta_2 CrL$	NF,Crown
ix	$\hat{Y} = \beta_0 + \beta_1 D + \beta_2 CrL^2$	OF
	NLINOLS	
Equation	**Tested models**	**Fitted best with**
i	$\hat{Y} = \beta_0 D^{\beta_1}$	Stem, NF, OF, Cone, Foliage, Bole, AGT
ii	$\hat{Y} = \beta_0 H^{\beta_1}$	
iii	$\hat{Y} = \beta_0 (D \times H)^{\beta_1}$	
iv	$\hat{Y} = \beta_0 (D^2 H)^{\beta_1}$	Bark
v	$\hat{Y} = \beta_0 D^{\beta_1} H^{\beta_2}$	Branch, Crown
vi	$\hat{Y} = \beta_0 (D \times CrL)^{\beta_1}$	
vii	$\hat{Y} = \beta_0 (D \times H^2)^{\beta_1}$	

TABLE A.2: Goodness-of-fit statistics across best LINOLS and NLINOLS regression models for each biomass component. Table shows parameter estimates, their standard error between parentheses and significance indicated as: ns, non significant, *, *P*<0.05; **, *P*<0.01; ***, *P*<0.001.

Component	Model	RMSE	MAB	MPE	RSE	CV%	R^2	IOA	β_0	β_1	β_2
LINOLS best models											
Stem	i	4.86	3.39	23.66	5.08	8.05	0.98	0.99	−6.691*** (0.493)	3.264*** (0.0986)	
Branch	iii	8.03	4.93	64.54	8.59	46.12	0.85	0.96	−5.803*** (0.666)	1.981*** (0.149)	−0.049* (0.0196)
Bark	i	0.88	0.6	0.77	0.92	16.99	0.94	0.98	−4.983*** (0.373)	1.416*** (0.074)	
NF	viii	0.9	0.61	0.8	0.94	26.63	0.92	0.97	−3.803*** (0.792)	0.992*** (0.137)	−0.054* (0.033)
OF	ix	2.88	2.11	8.29	3.01	21.85	0.96	0.98	−2.87 *** (0.441)	1.064*** (0.079)	−0.009** (0.0024)
Cone	ii	2.74	1.78	7.52	2.86	75.23	0.48	0.78	−2.301** (0.769)	0.132*** (0.029)	
Foliage	i	4.56	3.18	20.8	4.76	22.59	0.94	0.98	−3.259*** (0.393)	1.213*** (0.079)	
Crown	viii	15.36	9.02	235.92	16.04	40.85	0.82	0.96	−0.731 ns (0.51)	0.693*** (0.089)	−0.046* (0.021)
Bole	i	5.29	3.77	27.95	5.52	8.05	0.98	0.99	−7.003*** (0.51)	3.403*** (0.102)	
AGB	i	17.13	11.42	293.5	17.89	16.59	0.96	0.98	−1.182*** (0.244)	1.309*** (0.049)	
NLINOLS best models											
Stem	i	5.27	3.88	27.72	5.5	8.71	0.97	0.99	0.292*** (0.069)	1.803*** (0.075)	
Branch	v	6.87	4.3	47.18	7.34	39.44	0.89	0.97	0.517 ns (1.126)	3.791*** (0.474)	−3.199*** (0.645)
Bark	iv	0.9	0.6	0.8	0.94	17.31	0.92	0.98	0.0021 ns (0.001)	0.931*** (0.071)	
NF	i	0.88	0.61	0.77	0.92	26.14	0.9	0.97	0.001 ns (0.001)	2.868*** (0.269)	
OF	i	2.42	1.79	5.86	2.53	18.37	0.95	0.99	0.001 ns (0.001)	3.044*** (0.195)	
Cone	i	2.64	1.71	6.98	2.76	72.46	0.51	0.82	0.001 ns (0.002)	2.861** (0.762)	
Foliage	i	4.41	3.03	19.42	4.6	21.83	0.93	0.98	0.003 ns (0.002)	2.985*** (0.23)	
Crown	v	8.56	5.73	73.19	9.15	22.75	0.95	0.99	0.156 ns (0.19)	3.456*** (0.264)	−1.987*** (0.379)
Bole	i	5.76	4.25	33.15	6.01	8.77	0.97	0.99	0.302*** (0.072)	1.819*** (0.076)	
AGB	i	16.34	11.24	266.87	17.06	15.82	0.95	0.99	0.095* (0.045)	2.347*** (0.149)	

Note: The estimated parameter values are presented in power-transformed scale. Model i, ii, iii and so on refers to the best-fitted equations given in Table A.2.