

## Allometry for estimating aboveground tree biomass in tropical and subtropical eucalypt woodlands: towards general predictive equations

Richard J. Williams<sup>A,B,H</sup>, Ayalsew Zerihun<sup>C,D</sup>, Kelvin D. Montagu<sup>C,D</sup>, Madonna Hoffman<sup>E,D</sup>,  
Lindsay B. Hutley<sup>F,B</sup> and Xiaoyong Chen<sup>F,G</sup>

<sup>A</sup>CSIRO Sustainable Ecosystems, PMB 44 Winnellie, NT 0810, Australia.

<sup>B</sup>Cooperative Research Centre for Tropical Savanna Management, Charles Darwin University, Darwin, NT 0909, Australia.

<sup>C</sup>NSW Department of Primary Industries, Science & Research, PO Box 100, Beecroft, NSW 2119, Australia.

<sup>D</sup>Cooperative Research Centre for Greenhouse Accounting, GPO Box 475, ACT 2601, Australia.

<sup>E</sup>Queensland Department of Primary Industries and Fisheries, PO Box 6014, Rockhampton, Qld 4702, Australia.

<sup>F</sup>School of Science and Primary Industries, Charles Darwin University, Darwin, NT 0909, Australia.

<sup>G</sup>Current address: Department of Earth and Environmental Science, Okanagan University College,  
3333 College Way Kelowna, British Columbia V1V 1V7, Canada.

<sup>H</sup>Corresponding author. Email: dick.williams@csiro.au

**Abstract.** A fundamental tool in carbon accounting is tree-based allometry, whereby easily measured variables can be used to estimate aboveground biomass (AGB). To explore the potential of general allometry we combined raw datasets from 14 different woodland species, mainly eucalypts, from 11 sites across the Northern Territory, Queensland and New South Wales. Access to the raw data allowed two predictor variables, tree diameter (at 1.3-m height; D) and tree height (H), to be used singly or in various combinations to produce eight candidate models. Following natural log (ln) transformation, the data, consisting of 220 individual trees, were re-analysed in two steps: first as 20 species–site-specific AGB equations and, second, as a single general AGB equation. For each of the eight models, a comparison of the species–site-specific with the general equations was made with the Akaike information criterion (AIC). Further model evaluation was undertaken by a leave-one-out cross-validation technique. For each of the model forms, the species–site-specific equations performed better than the general equation. However, the best performing general equation,  $\ln(\text{AGB}) = -2.0596 + 2.1561 \ln(D) + 0.1362 (\ln(H))^2$ , was only marginally inferior to the species–site-specific equations. For the best general equation, back-transformed predicted *v.* observed values (on a linear scale) were highly concordant, with a slope of 0.99. The only major deviation from this relationship was due to seven large, hollow trees (more than 35% loss of cross-sectional stem area at 1.3 m) at a single species–site combination. Our best-performing general model exhibited remarkable stability across species and sites, when compared with the species–site equations. We conclude that there is encouraging evidence that general predictive equations can be developed across sites and species for Australia's woodlands. This simplifies the conversion of long-term inventory measurements into AGB estimates and allows more resources to be focused on the extension of such inventories.

### Introduction

The woodlands of northern Australia cover about a quarter of the continent. Because of the large area, changes in structure owing to clearing, fire, thickening and drought can significantly contribute to continental/national-scale accounts of carbon fluxes (Burrows *et al.* 2002; Chen *et al.* 2003; Williams *et al.* 1997, 2004). As a consequence, robust dynamic estimates of carbon stored in terrestrial landscapes, particularly forests and woodlands, are required at international, national and regional scales (Scholes and Noble 2001; Körner 2003; Grace 2004).

Tree-based allometry, which uses easily measured variables such as trunk diameter and tree height, to estimate carbon stored in aboveground biomass (AGB), is a fundamental tool in carbon accounting (Parresol 1999; Burrows *et al.* 2000, 2002; Snowdon *et al.* 2002). Tree-based allometry combined with stand-based inventory is probably the most accurate technique to detect AGB change in the short–medium term (*c.* 5 years, Burrows *et al.* 2000). Furthermore, allometry and inventory data are often used to validate model outputs and remotely sensed spatial predictions of AGB (e.g. Gower *et al.* 1999).

For regional estimates of AGB, generalised allometric relations are required to simplify the conversion of inventory measurements to estimates of biomass (Jenkins *et al.* 2003; Wirth *et al.* 2004). Published allometric relationships are typically developed from sampling at one site and may not be suitable for making estimates beyond the stand or plot from which they were developed. The development of general allometric equations, especially from published sources, is often constrained because of the different forms of equations (e.g. natural or base 10 logarithmic transformations; power functions), differences in measurement techniques (e.g. diameter at 30 cm *v.* diameter at breast height) and use of different predictor variables.

The use of allometric relationships for estimating AGB is common in temperate forests (e.g. Hingston *et al.* 1981; Husch *et al.* 1982; Ter-Mikaelian and Korzukhin 1997; Keith *et al.* 1999) and in the humid tropics (Brown *et al.* 1989; Phillips *et al.* 1998; Chave *et al.* 2003). There have also been a number of allometric studies in Australia's northern eucalypt woodlands, including O'Grady *et al.* (2000) and Werner and Murphy (2001) for the mesic savannas in the Northern Territory (NT), Ostendorf *et al.* (2004) for the eucalypt rainforest ecotone in the humid tropics and Burrows *et al.* (2000, 2002) and Grigg and Mulligan (1999) for a number of species in eastern and central Queensland.

Eamus *et al.* (2000), in summarising allometric relationships for a number of species in the tropical region of northern Australia, concluded that within a site a common allometric relationship for different species existed but variation in allometry was observed across sites. In contrast, Montagu *et al.* (2005) provided evidence that general allometry could be developed across contrasting sites for a single eucalypt species. Thus, there is a need to further explore the issue of species-site generality of allometric equations that predict AGB in savannas and other woodland ecosystems. One way to resolve these issues is to undertake analyses of multiple raw datasets, preferably from a wide range of species and sites.

The aim of this paper, therefore, is to further explore the potential of using a general allometric relationship to predict AGB in Australia's savannas. We combined raw datasets from 14 different tree species, mainly eucalypts, from 11 woodland sites across NT, Queensland and New South Wales (NSW). In each raw dataset, two predictor variables (tree diameter at 1.3-m height, D, and tree height, H) were measured for all individual trees. Thus, these variables could be combined to determine the best allometric model for predicting AGB.

## Materials and methods

### Data sources

We used data sets, both published and unpublished, from multiple species and sites across Australia (Table 1). Six sites were located

in the wet-dry tropical region of the NT, between  $\sim 12^\circ$  and  $17^\circ$ S. (Humpty Doo, Howard Springs, Katherine, Katherine Research Station, Manbulloo, and Kidman Springs). Tree basal area ranged from  $\sim 4$  to  $10 \text{ m}^2 \text{ ha}^{-1}$ . The NT species are all very common, and dominant or co-dominant in most of the major savanna types mapped by Wilson *et al.* (1990) in the northern part of the NT (i.e. north of  $\sim 17^\circ$ S). Four sites were located in central and eastern Queensland (between  $22^\circ$ S and  $151^\circ$ E) and one in north-western NSW (Oakvale, near Bourke:  $30.92^\circ$ S,  $146.50^\circ$ E). At the Queensland and NSW woodland sites, rainfall ranged from  $\sim 350$  to  $1100 \text{ mm year}^{-1}$ , with tree basal area  $\sim 8$ – $16 \text{ m}^2 \text{ ha}^{-1}$ ; the eucalypt species were either dominant or co-dominant components of the plant community.

### Biomass estimation

Aboveground tree biomass was estimated by direct-harvest techniques following the national carbon accounting system (NCAS) protocols (Snowdon *et al.* 2000, 2002). The approaches were essentially the same across the NT, NSW and Queensland sites. A sample of 5–20 trees of each species was collected, which spanned the range of diameters in the stand. Variables determined prior to sampling were basal diameter, diameter at 30 cm, diameter at breast height (1.3 m; D, cm) and tree height (H, m). The extent of tree hollows was recorded for some species (e.g. *E. populnea* F.Muell.) and for all species from Katherine and Kidman Springs in the NT. For these trees, the diameter of any hollows in the trunk was measured at 1.3 m above ground level.

Trees were felled and samples were separated into the following component parts: trunk, branches, leaves (Burrows *et al.* 2000; Montagu *et al.* 2005). The fresh mass of each component was measured. Immediately following fresh-mass determination, subsamples of each biomass component were taken to determine the dry-mass correction factor. For the trunks and large branches, 2–4 disks (each 200–2000 g) per tree were taken. Similar-sized subsamples (on the basis of fresh weight) for dry-mass determination were collected from the smaller branches and foliage of each tree. Subsamples were returned to the laboratory and dried at  $65$ – $80^\circ$ C until they reached a constant mass. Thus, dry-mass correction factors were calculated for each component of each tree separately.

### Data analysis

#### Data screening

Of the initial total of 260 individual trees, 36 observations from Howard Springs did not have tree-height information and four data points were from very small trees ( $D \leq 2.5$  cm); these data were excluded from further analyses. The analyses reported here were thus based on a core dataset of 220 observations from 14 species (12 *Eucalyptus* species; *sensu lat.* as per Brooker and Kleinig (1994), plus *Terminalia ferdinandiana* Exell and *Erythrophleum chlorostachys* (F.Muell.) Baill.) across 11 sites (Table 1). The number of individuals per species and site varied from 5 to 20 (Table 1).

#### Model formulation and selection

The core dataset contained two basic measures of tree size, D and H, which have been routinely used as predictor variables for AGB. Commonly, AGB estimation models have been formulated by using the natural logarithm (ln) of either variable alone or in combination. In some cases, combinations have involved second-order polynomials of these variables. Here, instead of evaluating all possible forms of model specification involving these variables, we selected eight of the more commonly used formulations as candidate models for evaluation with our dataset (loosely, the models for evaluation were selected

**Table 1. Site, species information and descriptive statistics on tree data used in this analysis**

Eucalypt terminology follows Brooker and Kleinig (1994); full species names given in text; the non-eucalypt species are *Erythrophileum chlorostachys* (F.Muell.) Baill. (Caesalpiniaceae) and *Terminalia ferdinandiana* Exell (Combretaceae). The abbreviations for species–site combinations as per Figs 1 and 6 are indicated, along with references to previously published studies for these species. MAR, mean annual rainfall; D, diameter at breast height (1.3 m); H, tree height; AGB, aboveground biomass per tree. For each of D, H and AGB, maximum, minimum and median values indicated

Species	Location (lat., long.)	Abbr. & Ref	MAR (mm)	N	D (cm)			H (m)			AGB (kg)		
					Min.	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.
<i>E. bleeseri</i>	Humpty Doo (12.5°S 131.3°E)	E. ble – Hdo <sup>A</sup>	1600	8	4.6	23	10.2	7.7	16	11.4	6.8	337.4	49.0
<i>E. crebra</i>	Kiaaroo (23.08°S 149.33°E)	E. cre – Kro <sup>B</sup>	–	18	3.1	52.2	26.3	3.8	26.5	17.1	1.9	3104.0	492.4
<i>E. foelscheana</i>	Katherine Research Station (14.5°S 132.4°E)	E. fol – Kat	950	20	4.3	48.1	14.7	4.5	14.9	8.7	5.2	1289.9	85.2
<i>E. melanophloia</i>	Mitchell (c. 23.1°S 147.0°E)	E. mel – Mit <sup>C</sup>	–	5	5.9	46	15.4	4.2	15.2	11.5	10.1	858.0	139.5
	Summerdell (23.75°S 146.0°E)	E. mel – Sum <sup>B</sup>	–	18	4.9	46.6	24.5	4.6	16.9	13.3	3.9	1871.8	318.5
<i>E. miniata</i>	Humpty Doo (12.5°S 131.3°E)	E. min – Hdo <sup>A</sup>	1600	8	6.8	21.6	14.2	6.8	14.7	12.5	8.5	197.0	66.4
	Katherine (14.7°S 132.7°E)	E. min – Kat <sup>A</sup>	950	6	5.8	29.8	13.3	6.2	13.0	9.3	8.4	428.0	73.6
<i>E. patellaris</i>	Katherine Research Station (14.5°S 132.4°E)	E. pat – Kat RS	950	9	3.9	38	10.8	4.7	13.0	9.2	2.5	599.0	53.5
<i>E. populnea</i>	Roma (25.75°S 148.41°E)	E. pop – Inj <sup>D</sup>	602	9	3	39.3	16.4	3.6	18.8	11.4	1.8	1179.0	103.6
	Rocky (23.17°S 150.56°E)	E. pop – Roc <sup>D</sup>	1103	9	4.3	61.1	27.7	4.2	22.2	15.7	5.7	2432.1	430.4
	Mit (c. 23.10°S 147.0°E)	E. pop – Mit <sup>C</sup>	–	7	3.5	54.7	30.2	4.5	17.8	17.4	3.1	2531.6	662.2
	Oakvale (30.92°S 146.50°E)	E. pop – Oak <sup>D</sup>	367	10	13.1	86	44.3	9.2	19.7	14.8	60.2	3695.7	890.1
<i>E. porrecta</i>	Humpty Doo (12.5°S 131.3°E)	E. por – Hdo <sup>A</sup>	1600	8	7.5	24.9	13.8	6.7	13.4	8.4	14.1	226.0	64.1
<i>E. pruinosa</i>	Kidman Springs (16.1°S 131.9°E)	E. pru – Kid Sp	650	12	2.8	42	12.3	2.9	10.8	7.0	2.4	978.5	64.8
<i>E. tectifica</i>	Manbulloo (14.6°S 132.2°E)	E. tec – Man	930	11	6	41.2	17.6	5.8	13.6	9.6	10.7	779.0	165.1
<i>E. terminalis</i>	Kidman Springs (16.1°S 131.9°E)	E. ter – Kid Sp	650	20	2.8	50	16.1	2.9	13.4	7.3	1.6	1575.0	101.4
<i>E. tetradonta</i>	Humpty Doo (12.5°S 131.3°E)	E. tet – Hdo <sup>A</sup>	1600	13	4.1	20.8	8.5	5.3	16.1	10.6	3.4	232.3	22.4
	Katherine (14.7°S, 132.7°E)	E. tet – Kat <sup>A</sup>	950	8	6	17.4	13.4	5.4	13.7	10.9	8.8	121.4	60.2
<i>Er. chlorostachys</i>	Humpty Doo (12.5°S 131.3°E)	Er. ch – Hdo <sup>A</sup>	1600	13	4.6	14.7	7.3	3.0	9.5	5.2	3.1	110.1	10.3
<i>T. ferdinandiana</i>	Humpty Doo (12.5°S 131.3°E)	T. fer – Hdo <sup>A</sup>	1600	8	4.6	16.7	9.4	3.6	9.4	6.9	3.2	136.6	22.5

<sup>A</sup>O'Grady *et al.* (2000).

<sup>B</sup>Burrows *et al.* (2000).

<sup>C</sup>Burrows *et al.* (2001).

<sup>D</sup>Zerihun *et al.* (2006).

*a priori*; see Montagu *et al.* 2005). The model forms evaluated are listed below:

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(H), \quad (1)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D \times H^2), \quad (2)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D \times H), \quad (3)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D)^2 + \beta_2 \ln(H), \quad (4)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D), \quad (5)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D^2 \times H), \quad (6)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D) + \beta_2 \ln(H), \text{ and} \quad (7)$$

$$\ln(\text{AGB}) = \beta_0 + \beta_1 \ln(D) + \beta_2 (\ln(H))^2, \quad (8)$$

where AGB = aboveground biomass (kg tree<sup>-1</sup>), H = tree height (m), D = diameter at breast (1.3 m) height (cm) and  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  are parameter estimates.

Selection of the best-performing models from these eight candidate models was based on the Akaike information criterion (AIC; for details see Burnham and Anderson 2002). Briefly, consider that there is a true, but unknown, model that underlies the data being analysed. When the true model is unknown, a set of candidate models is fitted to the data to identify that which best fits the data. AIC values of fitted candidate models provide estimates of expected relative distance (i.e. how far) the candidate models are from the true model that underlies the observed data. In addition, aptly, the AIC value is an estimate of the expected loss of information when a fitted model is used to approximate the true model (Burnham and Anderson 2002). The candidate model with the smallest AIC is considered closest to the true model and, therefore, selected as the best approximation. In practice, candidate models are fitted to the same dataset, and comparisons are carried out relative to the model with the smallest AIC. Generally, when the difference between the AIC value of a candidate model and the model with smallest AIC exceeds 10, then the candidate model with the higher AIC is said to be not supported by the data. If differences are between 0 and 10, the data provide empirical support that the model with the lower value is approximating the true model, but closer scrutiny may be necessary before selecting one model over another. Fitting of data to the eight candidate models specified above was carried out by the maximum likelihood procedure (by using S-PLUS Version 6.0, Insightful Corp., Seattle, WA). All analyses were undertaken on ln-transformed data to stabilise variance.

There were 20 unique species–site groups. One of our primary questions was whether one general model could fit these data or whether there was a need for species–site-specific models. For each of the eight models, the performance of general *v.* species–site-specific models was also evaluated, using AIC. Model performance was further evaluated by examining the patterns of residuals arising from species–site-specific models and those arising from the non-specific form of the same model.

The approach employed here is useful for identifying best-performing models in a given dataset. However, it might not indicate predictive performance of each model when tested against new data that have not been used in model building. Thus, for these models, a leave-one-out cross-validation was carried out to evaluate how each of these models would perform when applied to ‘new data’ not used during parameter estimation.

Finally, for each model, the performance of the general model—without site or species effects, but including diameter

and or height as variates—was evaluated by comparing observed values with predicted results. The various models were used to predict AGB for each individual tree, following back-transformation of the log-transformed values with the correction factor of Baskerville (1972).

## Results

### *Variation in tree size*

Across the data set tree AGB varied by more than three orders of magnitude from ~2 to 3696 kg tree<sup>-1</sup> (Fig. 1). There was variation among sites and species in tree size. The minimum diameter was set at >2.5 cm (see above); maximum diameter varied from ~15 cm for *Erythrophleum chlorostachys* at Humpty Doo to 86 cm for *E. populnea* at Oakvale. Tree height also varied among sites and species; minimum heights ranged from 2.9 to 9.2 m, whereas maximum tree heights in the sample data ranged from 9.4 to 26.5 m (Table 1).

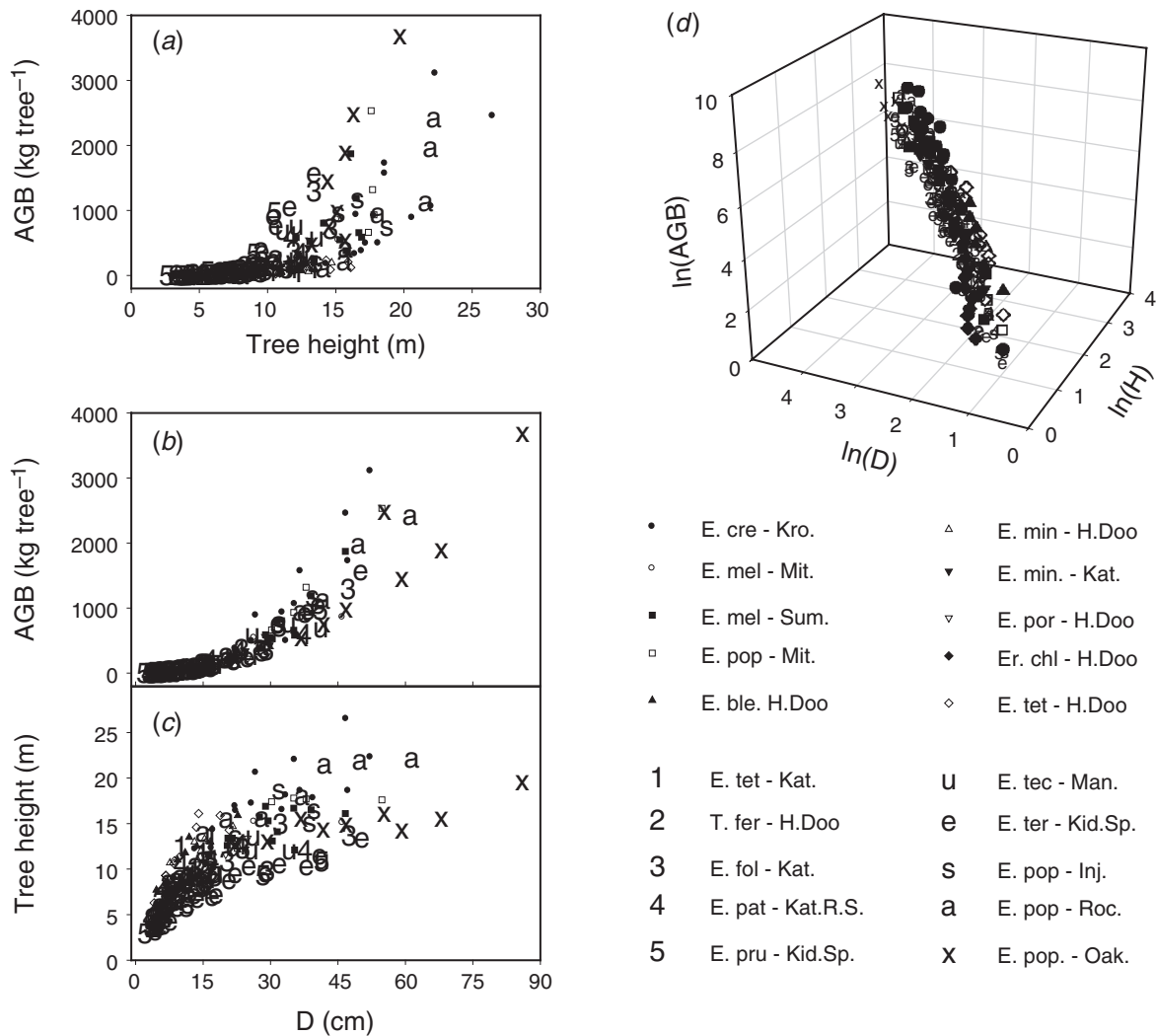
### *Site and species effects on the relationships between tree dimensions*

The overlaid plot of tree biomass against tree height for each species–site combination showed considerable scatter (Fig. 1a). *E. populnea* trees from Oakvale had the highest AGB for a given H. In contrast, *E. crebra* F.Muell. from Kiauroo generally had lowest AGB for a given H. Aboveground biomass as a function of tree D showed a smaller degree of scatter than AGB *v.* H (cf. Fig. 1a, b). *E. crebra* from Kiauroo had the highest AGB biomass for a given D, whereas *E. populnea* from Oakvale had the lowest (Fig. 1b). The overlaid plot of tree height *v.* D showed that *E. crebra* and *E. populnea* generally were tallest and shortest, respectively, for a given D (Fig. 1c). When AGB was examined as a function of both H and D, on a ln–ln scale, there was a convergence in the relationship across sites and species except for large *E. populnea* trees from Oakvale (Fig. 1d).

### *Comparative model performance*

The eight candidate models that related ln(AGB) to predictor variables ln(D), ln(H), or linear and quadratic combinations of these two variables are presented in Table 2. Of these, Model 1, which used height alone as a predictor gave the poorest fit as indicated by the AIC and the root mean square error (RMSE). The predictor ln(D) alone (Model 5) was superior to ln(H). However, the best models involved ln(D) in some combination with ln(H), viz. ln(D) + (ln(H))<sup>2</sup>, closely followed by ln(D) + ln(H).

Comparative plots of predicted against observed data (log–log scale) are shown in Fig. 2. Consistent with the results in Table 2, the agreement between predicted and observed values was poor for height as the only predictor. Agreement increased with decreasing values of RMSE (Table 2), with the best based on the model that included both ln(D) and (ln(H))<sup>2</sup> (Fig. 2).



**Fig. 1.** Relationship between aboveground biomass (AGB) and tree size variables (diameter at breast height, D, and height above ground, H) for various species and sites. (a) AGB v. H, (b) AGB v. D, (c) H v. D, (d) ln(AGB) as a function of ln(D) and ln(H). Sites and species pooled. See Table 1 for abbreviations of species and sites.

The comparison of predicted results with observed values, following back transformation, allowed more detailed examination of species–site effects. This is shown for our best-performing model (Model 8) and the most commonly used model form (Model 5) in Fig. 3. Over the entire range of tree sizes (Fig. 4), predicted values of biomass from Model 8a were between 0.56 and 1.76 times the measured AGB value (Model 8b = 0.54–1.69 times), whereas for Model 5a the range was slightly wider, at 0.45–2.1 times the measured AGB (Model 5b = 0.46–1.89). This relative difference was independent of tree size. Consequently, the absolute difference between measured and predicted AGB increased with increasing tree size, particularly for termite-hollowed *E. populnea* trees from Oakvale and Rockhampton. At Oakvale, for example, more than 35% of trunk biomass was missing in trees with  $D \geq 30$  cm. When data were

re-analysed by excluding hollow trees from Oakvale (Models 5b and 8b, Table 2), there was a remarkable consistency across sites and species in the pattern of observed and predicted values. The intercepts are not significantly different from zero, and the slopes (which should be 1.0 for constancy) are  $>0.99$  for  $\ln(D) + (\ln(H))^2$  as predictors and  $>0.95$  for  $\ln(D)$  alone as a predictor).

For the eight models, comparison of the distribution of residuals from species–site-specific models (‘specific models’) and residuals arising from model parameterisation based on the entire data (‘general models’) are shown in Fig. 5. Generally, these diagnostic plots show that for all models, the spread of residuals from the specific models was narrower than that from the general models. Nonetheless, for Models 5–7 differences between the specific and the general parameterisations were

**Table 2. Selected models, their parameters with standard errors (in parenthesis) and fit statistics**

Sample size was 220 except for Models 5*b* and 8*b*  $n = 213$ . Note that although Models 5*b* and 8*b* can be compared by using AIC, neither can be compared with the remaining 8 models because of differences in the datasets

Model	Inputs	Intercept	ln(D) (cm)	ln(H) (m)	ln(DH) (m <sup>2</sup> )	ln(D <sup>2</sup> H) (m <sup>3</sup> )	(ln(H)) <sup>2</sup> (m <sup>2</sup> )	(ln(D)) <sup>2</sup> (cm <sup>2</sup> )	ln(DH <sup>2</sup> ) (m <sup>3</sup> )	RMSE	AIC
1	ln(H)	-3.5413 (0.2732)		3.5337 (0.1188)						0.8299	548.3
2	ln(DH <sup>2</sup> )	1.5794 (0.0586)							1.1051 (0.0193)	0.4656	294.0
3	ln(DH)	3.9254 (0.0235)			1.5571 (0.0192)					0.3348	147.1
4	ln(H) + (ln(D)) <sup>2</sup>	-0.6266 (0.1300)		1.0475 (0.0815)				0.3497 (0.0096)		0.3109	118.3
5 <i>a</i>	ln(D)	-2.2111 (0.0679)	2.4831 (0.0245)							0.2696	53.5
5 <i>b</i> <sup>A</sup>	ln(D)	-2.3046 (0.0671)	2.5243 (0.0246)							0.2570	41.5
6	ln(D <sup>2</sup> H)	5.9821 (0.0226)				0.9659 (0.0090)				0.2549	30.2
7	ln(D) + ln(H)	-2.6392 (0.0800)	2.1735 (0.0438)	0.5574 (0.0689)						0.2362	-2.6
8 <i>a</i>	ln(D) + (ln(H)) <sup>2</sup>	-2.0596 (0.0611)	2.1561 (0.0430)				0.1362 (0.0156)			0.2319	-10.6
8 <i>b</i> <sup>A</sup>	ln(D) + (ln(H)) <sup>2</sup>	-2.1432 (0.0617)	2.2143 (0.0435)				0.1251 (0.0152)			0.2242	-8.6

<sup>A</sup>Parameters of these models were estimated by excluding *E. populnea* trees from Oakvale whose hollow area exceeded 25% of the stem cross-sectional area at breast height.

relatively small, suggesting the general model performed comparatively well.

The distribution of residuals arising from three of the general models as applied to all 20 possible site-specific combinations (a test of the level of systematic variation) is illustrated in Fig. 6. Not surprisingly, for the general model with H as the only predictor, 17 of the 20 species–site combinations exhibited systematic over- or under-estimation. The distribution of residuals across species and sites was markedly more stable with D as the only predictor. Importantly, there was only a modest further improvement when (ln(H))<sup>2</sup> was added to ln(D).

#### Model cross-validation

The evaluation of the performance of the candidate models in predicting new responses that were not used in model development, via leave-one-out cross-validation, gave averaged mean square errors of prediction that reflected the order by AIC. The mean square error was highest for the model with H alone as the predictor and lowest for model that included ln(D) and (ln(H))<sup>2</sup> as predictor variables (Table 3).

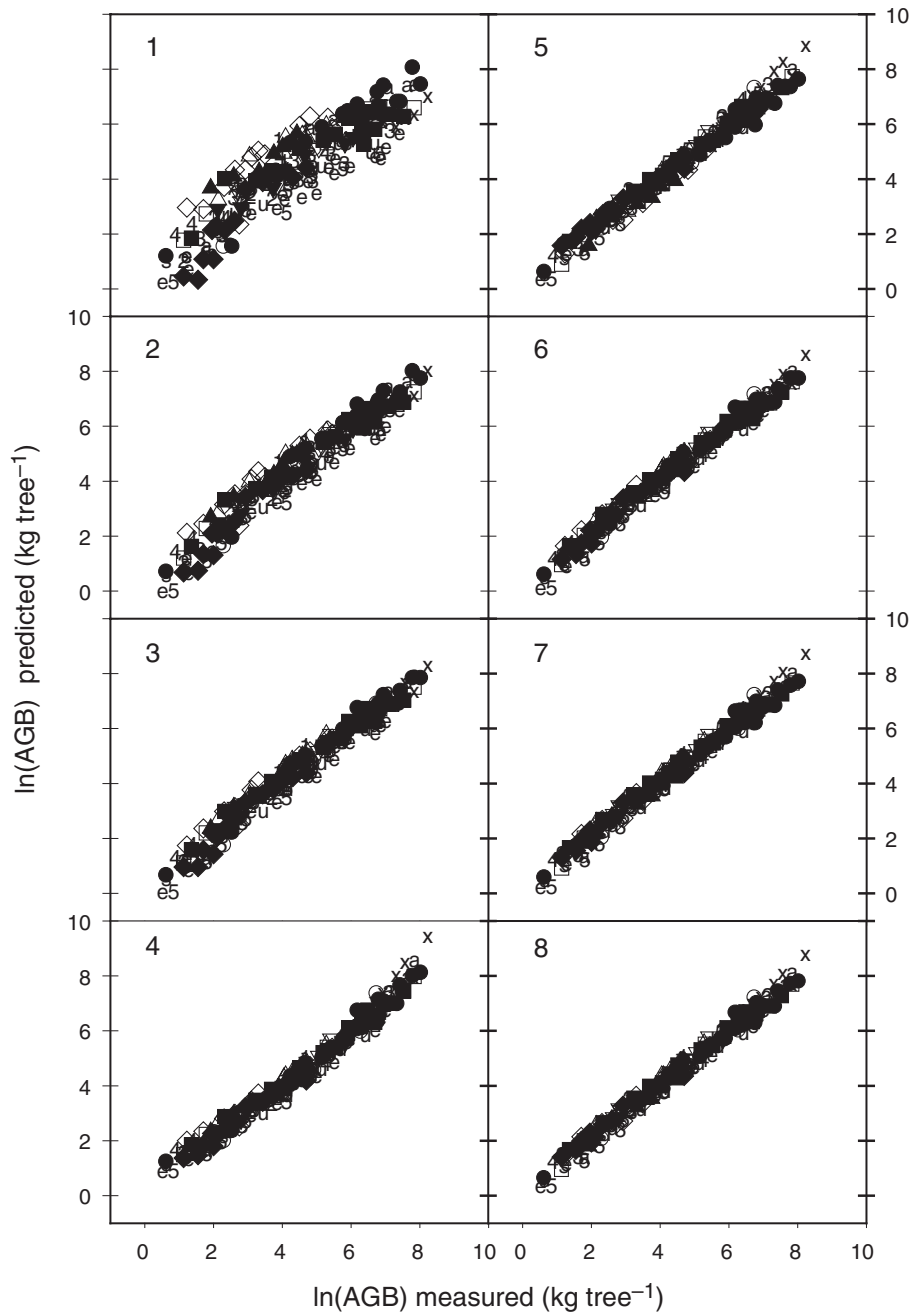
## Discussion

### Allometric generality

The principal objective of this analysis was to explore the potential of developing, for the savanna, general allometry

that is not species or site specific. The 14 species included in this analysis account for most of the aboveground tree biomass in tropical and sub-tropical woodlands of Australia (O'Grady *et al.* 2000; Burrows *et al.* 2002). We combined raw data from multiple sources, allowing formal tests of species and site, identification of idiosyncratic sites and/or species and individuals, and comparisons of differing model forms, a process not possible when relying on published equations alone (Keith *et al.* 1999; Jenkins *et al.* 2003).

Despite the wide range of species and sites (Fig. 1, Table 1), combining data to produce a general allometric relationship led to only a small decrease in the amount of variation in AGB accounted for, compared with species–site-specific allometry. Thus, although we were able to statistically detect species–site differences, largely because of the power of testing a large number of samples, the amount of variation accounted for by species–site factors was less than 0.5%. Consequently, when our best general equations are compared with species–site-specific equations it is difficult to observe the difference (Fig. 5). This conclusion is strengthened by the fact that we analysed data from species that are phylogenetically distinct, representative of their respective regions at landscape scales, and drawn from different locations and environments in northern Australia. We therefore conclude that there is excellent potential for the use of general allometry across the eucalypt woodlands of northern and eastern Australia.

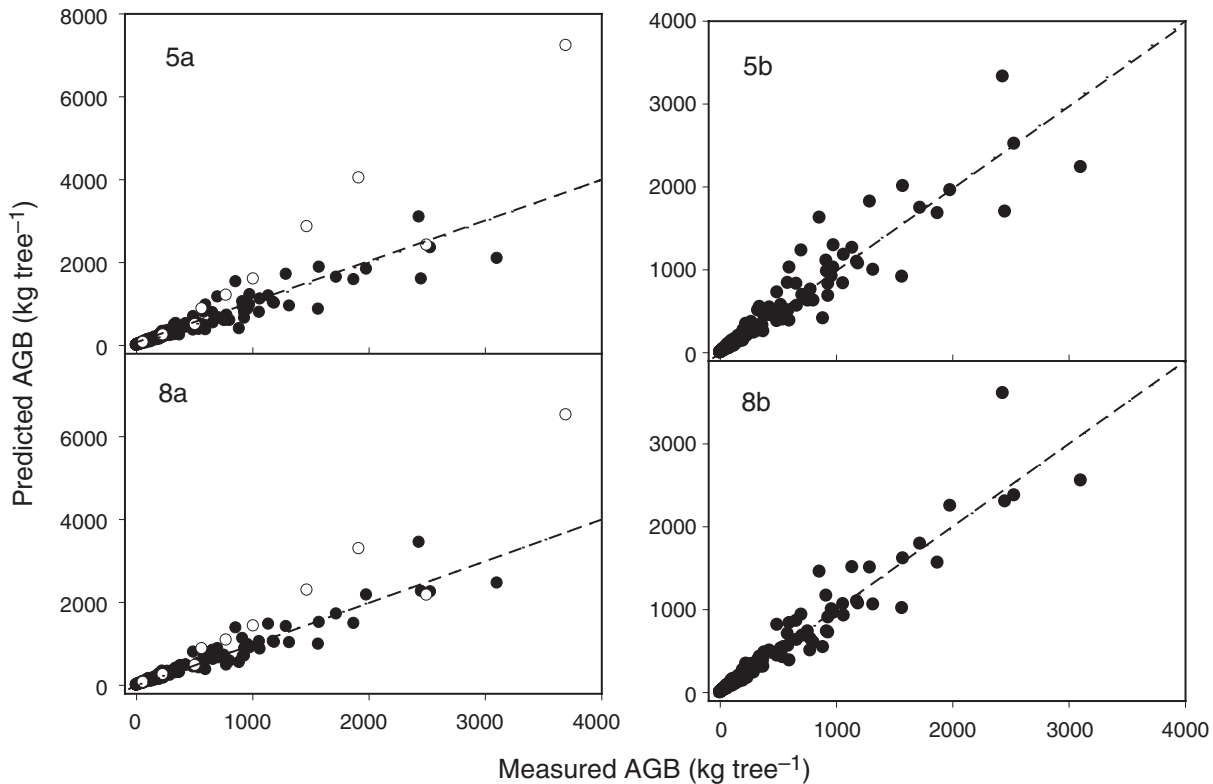


**Fig. 2.** Comparative plots of measured and predicted aboveground biomass (AGB) from the eight candidate models. Numbers in panels (1–8) correspond to models shown in Table 2. Symbols as for Fig. 1.

Below we discuss the most appropriate model for general allometry, deviations from the general equation and the advantages in using general allometry.

We examined a number of candidate models to identify one that is empirically best supported by our data (Table 2). There was consistency in the ranking of the models, with the model form  $\ln(\text{AGB}) = \ln(D) + (\ln(H))^2$  identified as the best for this dataset, on the basis of the model selection

and validation methods employed (Tables 2, 3). The simple, most commonly used model form, based on  $D$  alone, also performed well. In both model forms, there was very good agreement between predicted and observed values on both  $\ln$ - $\ln$  and linear scales (Figs 2, 3, respectively). A further examination of these general equations also indicated a marked stability of residuals, with a low level of systematic under- or over-estimation (Fig. 6). The greatest



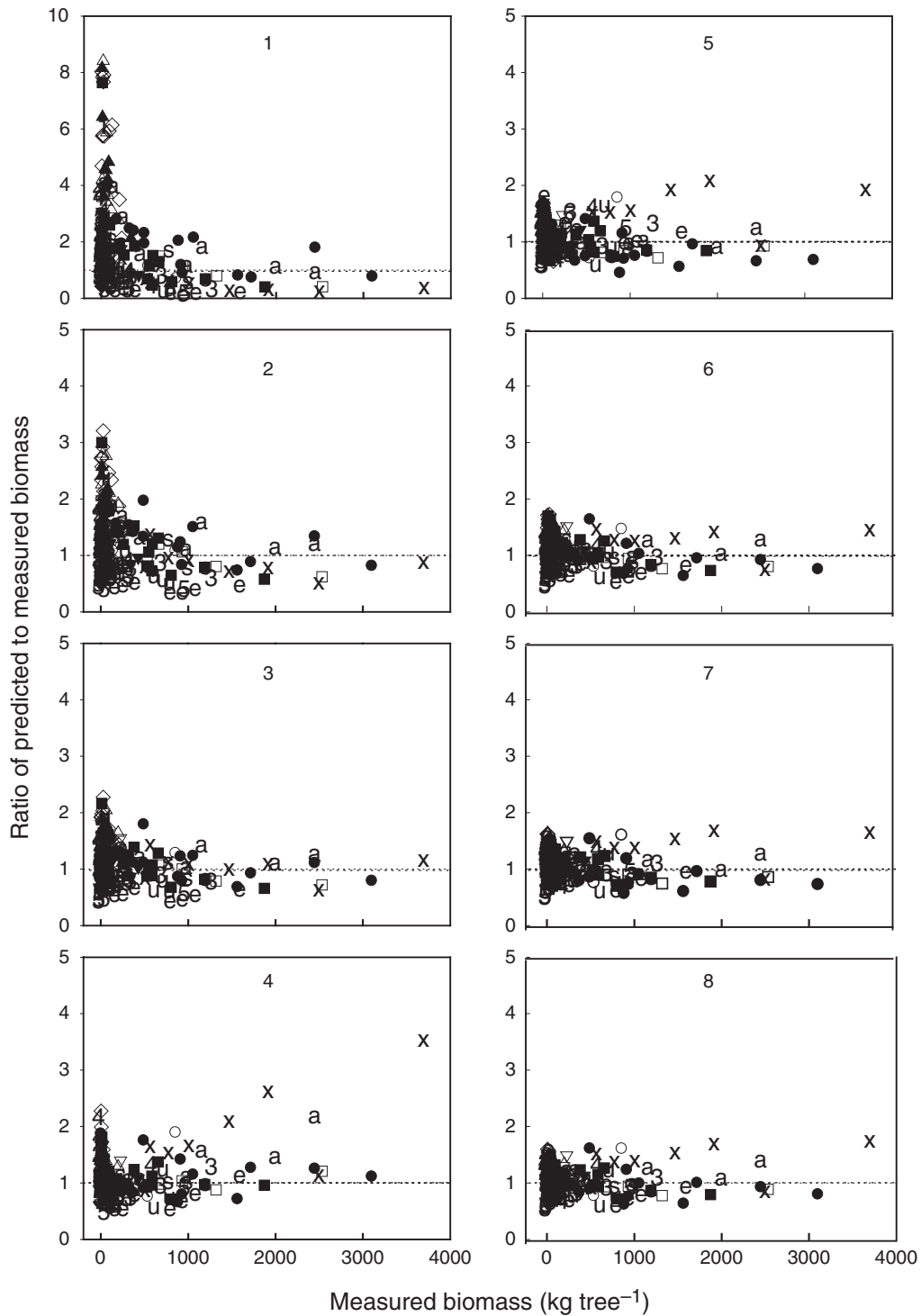
**Fig. 3.** Back-transformed plots of measured *v.* predicted aboveground biomass (AGB; on linear scale) for the most commonly used model form and the best-performed model identified here (see Table 2). Plots for Models 5*a* and 8*a* are based on the entire dataset with observations from Oakvale (open symbols). In plots for Models 5*b* and 8*b*, parameters were re-estimated after excluding hollow trees from Oakvale in which hollow area at breast height exceeded 25% of the stem cross-sectional area at that height. The dashed line in each panel is the 1 : 1 line.

deviation from the general equation was at the species–site combination of *E. populnea* at Oakvale, which could be attributed to the high degree of hollowness in the larger trees. For seven of the largest trees more than 35% of the cross-section area at 1.3 m was lost because of termite activity. Although woodland trees typically have some degree of hollowness, the extent of hollowness for these trees was exceptional. Thus, the general equation over-predicted AGB for these trees (Figs 3, 4, 6). The removal of these seven trees from the analysis (Models 5*b* and 8*b*; Table 2) further increased the agreement between the observed and predicted AGB, such that the slopes of the relationships were 0.99 for the best-performing model and 0.95 for the model with  $\ln(D)$  alone as the predictor (Fig. 3). Thus, we recommend that some measure of hollowness, e.g. coring of standing trees, be undertaken to determine whether the stand contains trees that are significantly hollowed. Although more detailed examination of the effect of trunk hollowing on allometric relationships is warranted, from our limited dataset we suggest that, as a rule of thumb, the general equation should not be used where more than ~20% of the trunk cross-sectional area is missing.

The development of general allometric relationships produces a number of benefits. First, it greatly simplifies

the conversion of inventory measurements to regional estimates of biomass carbon stocks and fluxes. Decisions regarding which species–site equations to use are avoided. Second, general models allow increases in sample numbers that are used to generate allometric equations, compared with species–site-specific relationships. The high cost of destructively harvesting trees, and in some cases restricted access to certain species and tree sizes, combine to limit the number of trees sampled. The data used in this study were typical of published allometry, with unique species–site combinations having between 5 and 20 samples (Table 1). With such sample numbers the parameter and regression variance estimates are inherently unstable. Thus, parameter estimates can be quite sensitive to single data points. Furthermore, such sample numbers produce uncertain regression variance estimates. When log–log equations are used, this has a subsequent impact on bias-correction factors relying on regression variance (Baskerville 1972; Beauchamp and Olson 1973). Combining data to produce general allometric relationships overcomes many of these limitations by increasing sample size, thereby producing stable parameter and regression variance estimates.





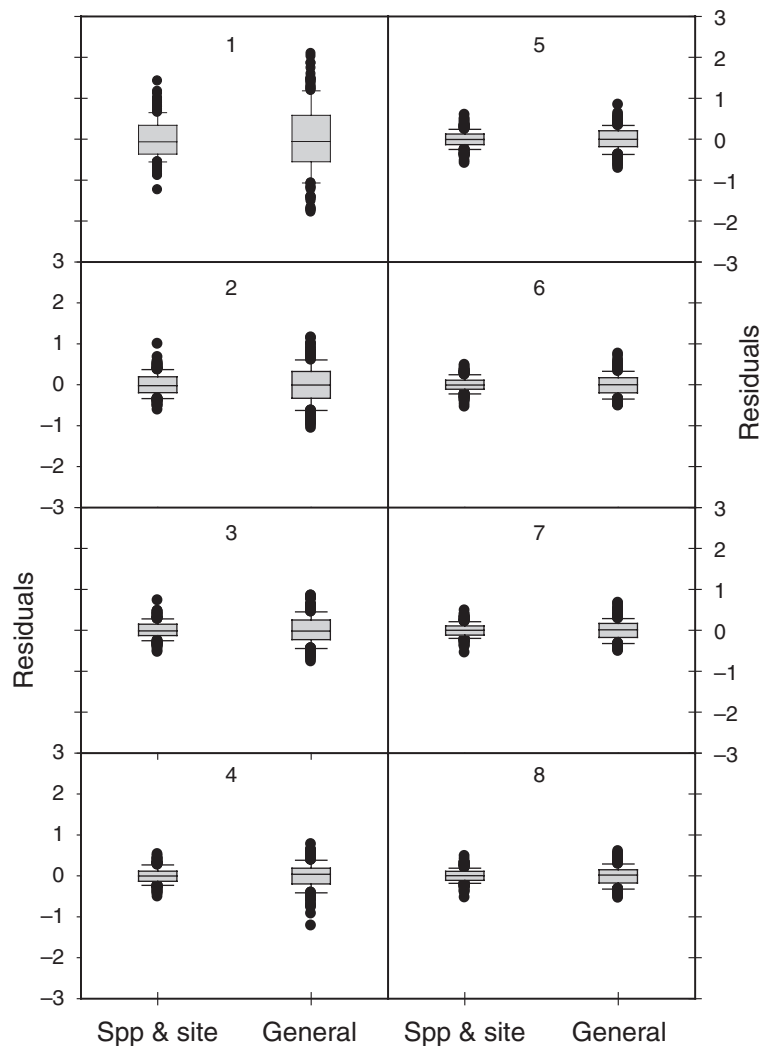
**Fig. 4.** Ratios of predicted to measured AGB displayed as a function of measured AGB for all candidate models, on linear scale. Numbers in panels (1–8) correspond to models shown in Table 2.

*Using the equations to derive estimates*

*Tree height—is it a surrogate for species or site differences?*

The evidence from this study is that although D accounts for more than 97% of the AGB variation, the addition

of height does improve the performance of the general allometric model. Globally, the evidence regarding the value of including height in allometric equations is equivocal, with neutral, positive and negative results reported. Madgwick and Satoo (1975) concluded that height added little

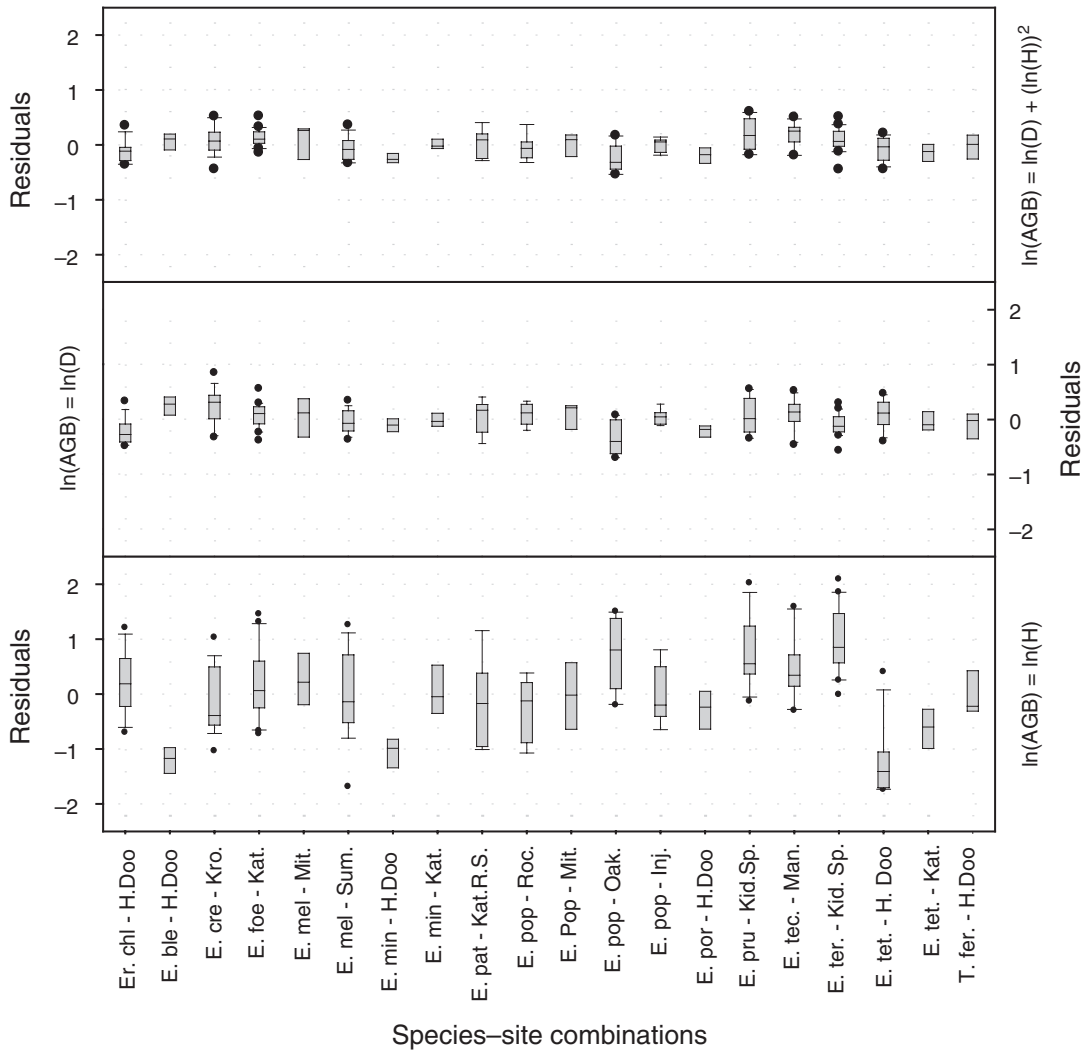


**Fig. 5.** Comparison of residuals *v.* fits for the different models. The graphs are plotted in decreasing order of AIC. The model with the largest AIC value is plotted first; that with the smallest AIC value is plotted last. The numbers on the plot correspond to the models in Tables 2 and 3.

to the predictive capacity of diameter-based allometric equations at a site. In contrast, inclusion of tree height in allometric relationships may account for variations in the allometric relationship across contrasting sites (Schmitt and Grigal 1981). By comparison, incorporating height into the allometric equation decreased the performance of a general equation across sites in one eucalypt species (*E. pilularis* Sm; not analysed in this study), compared with the relationship based on *D* alone (Montagu *et al.* 2005). Given the mix of species and sites in this study, it is not possible to tease out whether height is acting as a surrogate for species or site, or some combination of both.

Height alone did not improve model performance. Performance was improved only when height was specified in some models as a second order polynomial. Even

where inclusion of height as a predictor improved model performance, there are practical and statistical reasons for using diameter alone as the predictor variable when making estimates of AGB from inventory measurements (Montagu *et al.* 2005). Typically, the errors associated with diameter measurement are less than 3% and this feature makes it a more appropriate predictor (Gregoire *et al.* 1989). This is an important feature, since parameter estimates are often derived without taking into account measurement errors in predictor variables. By contrast, measurements of tree height in mature stands are usually only within 10–15% of the true height (Brown *et al.* 1995); they are also time-consuming, and thus costly. Therefore, predictive equations that require tree-height measurements may be limited in their application. However, in open savanna woodlands (as opposed to forest),



**Fig. 6.** Detailed plots of the residuals from three of the general models drawn from the 20 possible species–site combinations. The plot of residuals allows identification of systematic under- or over-estimation for given species–site combinations. The three models shown are the following: residuals from our best-performing model,  $\ln(D) + \ln(H)^2$  (top panel); residuals from the model with  $\ln(D)$  as the only predictor (the most commonly used model form) (middle panel); residuals from the model with  $\ln(H)$  as the only predictor (bottom panel). See Table 1 for abbreviations of species and sites.

where tree cover is typically 20–40%, and trees <20 m tall, measuring tree height may be a practical option for improving predictive capacity. The final choice of model form will depend on the cost of measuring additional predictor variables, the resulting improvements in AGB estimates, and the benefits of measuring fewer predictor variables at additional locations. In terms of regional carbon accounting, it may be that, per unit of effort and expense, evaluating diameter at more sites, as opposed to measuring both diameter and height at fewer sites, will reduce uncertainties associated with deriving regional estimates. Further simulation studies, preferably based on combined datasets, are warranted to test such scenarios.

*The application of allometry in carbon accounting in savannas*

We regard it as axiomatic that tree-based allometry will continue to be important in estimating changes in terrestrial biomass carbon stocks. Allometry provides the means to calibrate local and regional estimates of AGB, and for providing the basis of future estimates in stocks and fluxes. Allometry will continue to have important applications in savannas, because predictive, mechanistic models of the impact of vegetation changes on the carbon budget need baseline estimates of change, to account for landscape-scale changes in land management.

**Table 3. Average predictive errors of the different models from leave-one-out cross validation**

Model no.	Predictors	Averaged mean square error of prediction
1	ln(H)	0.6987
2	ln(DH <sup>2</sup> )	0.2203
3	ln(DH)	0.1140
4	ln(H) + (ln(D)) <sup>2</sup>	0.1005
5	ln(D)	0.0743
6	ln(D <sup>2</sup> H)	0.0661
7	ln(D) + ln(H)	0.0574
8	ln(D) + (ln(H)) <sup>2</sup>	0.0538

There are numerous long-term inventory studies based on changes in diameter of permanently marked trees at multiple reference sites across northern Australia (TRAPS in Queensland, Burrows *et al.* 2000; Kapalga in the NT, Williams *et al.* 1999, 2003; Cook *et al.* 2005; Munmarlary, Russell-Smith *et al.* 2003) that will continue to be important meso-scale sampling points for the monitoring of change in stand structure, and hence the derivation of regional estimates of carbon, in the coming decades. For example, the mesic savannas of the NT, Western Australia and Queensland are subject to frequent fire (Williams *et al.* 2002), and the impact of fire varies considerably with tree size and fire intensity (Williams *et al.* 1999, 2003). The post-fire trajectory of biomass accumulation or depletion may also depend on stand structure (G. D. Cook, A. Liedloff and R. J. Williams, unpubl. data). Alternatively, in many savanna areas, woody vegetation is thickening (Dyer *et al.* 2001; Burrows *et al.* 2002). In these cases, stand-based allometry will be integral to predicting biomass changes as a function of stand structure and land use.

We have developed a general allometric relationship across 14 woodlands species, from different regions of northern Australia, with little loss in accuracy, compared with 20 species-site-specific equations. The small reduction in accuracy of general allometry is more than countered by the simplification of regional estimates of aboveground estimates and the greater certainty in equation parameters and variance estimates. Our results indicate that in addition to *D*, height improved model performance. However, we were unable to determine whether height was acting as a surrogate for species or site, or a combination of both.

Accounting for the inputs required to manage landscapes for optimising carbon sequestration, and minimising greenhouse gas emissions, will require accurate estimates of changes in stand structure and survival and growth of individual trees. Tree-based allometric equations applied to stand-based inventories will continue to be an important component of carbon accounting at local, regional and national scales.

## Acknowledgments

The Tropical Savanna Management CRC and the Greenhouse Accounting CRC, the Australian Greenhouse Office and the Bureau of Rural Resources all supported the work financially. There were many helpers in the field. We thank all landholders whose support enabled the study. NT Department of Business Industry and Rural Development, in particular Don and Judy Cherry of Kidman Springs Research Station, provided resources and support in the NT. We thank Garry Cook, Harold Hopman and two anonymous referees for their comments on the manuscript.

## References

- Baskerville GL (1972) The use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research* **2**, 49–53.
- Beauchamp JJ, Olson JS (1973) Correction for bias in regression estimates after logarithmic transformation. *Ecology* **54**, 1403–1407.
- Brooker MIH, Kleinig DA (1994) 'Field guide to the eucalypts. 3. Northern Australia.' (Inkata Press: Sydney)
- Brown IF, Martinelli LA, Thomas WW, Moreira MZ, Ferreira CAC, Victoria RA (1995) Uncertainty in the biomass of Amazonian forests—an example from Rondonia, Brazil. *Forest Ecology and Management* **75**, 175–189. doi: 10.1016/0378-1127(94)03512-U
- Brown S, Gillespie AJR, Lugo AE (1989) Biomass estimation methods for tropical forests with applications to forest inventory data. *Forest Science* **35**, 881–902.
- Burnham KP, Anderson DR (2002) 'Model selection and multimodel inference: a practical information-theoretic approach.' 2nd edn. (Springer: New York)
- Burrows WH, Hoffmann MB, Compton JF, Back PV, Tait LJ (2000) Allometric relationships and community biomass estimates for some dominant eucalypts in central Queensland woodlands. *Australian Journal of Botany* **48**, 707–714. doi: 10.1071/BT99066
- Burrows WH, Hoffmann MB, Compton JF, Back PV (2001) Allometric relationships and community biomass stocks in white cypress pine (*Callitris glaucophylla*) and associated eucalypts of the Carnarvon area—south central Queensland. Technical Report No. 33. (Australian Greenhouse Office: Canberra)
- Burrows WH, Henry BK, Back PV, Hoffmann MB, Tait LJ, Anderson ER, Menke N, Danahar T, Carter JO, McKeon GM (2002) Growth and carbon stock change in eucalypt woodlands in north east Australia: ecological and greenhouse sink implications. *Global Change Biology* **8**, 769–784. doi: 10.1046/j.1365-2486.2002.00515.x
- Chave J, Condit R, Lao S, Caspersen JP, Foster RB, Hubbell SP (2003) Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama. *Journal of Ecology* **91**, 240–252. doi: 10.1046/j.1365-2745.2003.00757.x
- Cook GD, Liedloff AC, Eager RW, Chen X, Williams RJ, O'Grady AP, Hutley LB (2005) The estimation of carbon budgets of frequently burnt tree stands in savannas of northern Australia using allometric analysis and isotopic discrimination. *Australian Journal of Botany* **53**, 621–630.
- Chen X, Hutley LB, Eamus D (2003) Carbon balance of a tropical savanna in northern Australia. *Oecologia* **137**, 405–416. doi: 10.1007/s00442-003-1358-5
- Dyer R, Jacklyn P, Partridge I, Russell-Smith J, Williams RJ (2001) 'Savanna burning: understanding and using fire in northern Australia.' (Tropical Savannas Cooperative Research Centre: Darwin)

- Eamus D, McGuinness K, Burrows W (2000) Review of allometric relationships for estimating woody biomass for Queensland, the Northern Territory and Western Australia. National Carbon Accounting System Technical Report No. 5a. (Australian Greenhouse Office: Canberra)
- Gower ST, Kcharik CJ, Norman JM (1999) Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. *Remote Sensing of Environment* **70**, 29–51. doi: 10.1016/S0034-4257(99)00056-5
- Grace J (2004) Understanding and managing the global carbon cycle. *Journal of Ecology* **92**, 189–202. doi: 10.1111/j.0022-0477.2004.00874.x
- Gregoire TG, Zedaker SM, Nicholas NS (1989) Modeling relative error in stem basal area estimates. *Canadian Journal of Forest Research* **20**, 496–502.
- Grigg AH, Mulligan DR (1999) Biometric relationships for estimating standing biomass, litterfall and litter accumulation of *Acacia salicina* on mined land in central Queensland. *Australian Journal of Botany* **47**, 807–816. doi: 10.1071/BT97122
- Hingston FJ, Dimmock GM, Turton AG (1981) Nutrient distribution in a jarrah (*Eucalyptus marginata* Donn ex Sm.) ecosystem in south-west Western Australia. *Forest Ecology and Management* **3**, 183–207. doi: 10.1016/0378-1127(80)90015-8
- Husch B, Miller CI, Beers TW (1982) 'Forest mensuration.' 3rd edn. (John Wiley & Sons: New York)
- Jenkins JC, Chohnacky DC, Heath LS, Birdsey RA (2003) National-scale biomass estimators for United States tree species. *Forest Science* **49**, 12–35.
- Keith H, Barrett D, Kennan R (1999) Review of allometric relationships for woody biomass for NSW, ACT, Vic, Tas, SA. National Carbon Accounting Technical Report No. 5b. (Australian Greenhouse Office: Canberra)
- Körner Ch (2003) Slow in, rapid out—carbon flux studies and Kyoto targets. *Science* **300**, 1242–1243. doi: 10.1126/science.1084460
- Madgwick HAI, Satoo T (1975) On estimating the aboveground weights of tree stands. *Ecology* **56**, 1446–1450.
- Montagu KD, Duttmer K, Barton CVM, Cowie AL (2005) Developing general allometric relationships for regional estimates of carbon sequestration—an example using *Eucalyptus pilularis* from seven contrasting sites. *Forest Ecology and Management* **204**, 113–127.
- O'Grady AP, Chen X, Eamus D, Hutley LB (2000) Composition, leaf area index and standing biomass of eucalypt open forests near Darwin in the Northern Territory Australia. *Australian Journal of Botany* **48**, 629–638.
- Ostendorf B, Bradford MG, Hilbert DW (2004) Regional analysis of forest biomass at the rainforest/sclerophyll boundary, Queensland, Australia. *Tropical Ecology* **45**, 31–41.
- Parresol BR (1999) Assessing tree and stand biomass: a review with examples and critical comparisons. *Forest Science* **45**, 573–593.
- Phillips OL, Malh Y, Higuchi N, Laurance WF, Nunez PV, Vasquez RM, Laurance SG, Ferreira LV, Stern M, Brown S, Grace J (1998) Changes in the carbon balance of tropical forests: evidence from long-term plots. *Science* **282**, 439–442. doi: 10.1126/science.282.5388.439
- Russell-Smith J, Whitehead PJ, Cook GD, Hoare JL (2003) Response of *Eucalyptus*-dominated savanna to frequent fires: lessons from Munmarlary, 1973–1996. *Ecological Monographs* **73**, 349–375.
- Scholes R, Noble I (2001) Storing carbon on land. *Science* **294**, 1012–1013. doi: 10.1126/science.1065307
- Schmitt MDC, Grigal DF (1981) Generalized biomass estimation equations for *Betula papyrifera* Marsh. *Canadian Journal of Forest Research* **11**, 837–840.
- Snowdon P, Eamus D, Gibbons P, Khanna H, Raison J, Kirschbaum M (2000) Synthesis of allometrics review of root biomass and design of future woody biomass sampling strategies. National Carbon Accounting Technical Report No. 17. (Australian Greenhouse Office: Canberra)
- Snowdon P, Raison J, Keith H, Ritson P, Grierson P, Adams M, Montagu K, Bi H, Burrows W, Eamus D (2002) Protocol for sampling tree and stand biomass. National Carbon Accounting Technical Report No. 31. (Australian Greenhouse Office: Canberra)
- Ter-Mikaelian MT, Korzukhin MD (1997) Biomass equations for sixty-five North American tree species. *Forest Ecology and Management* **97**, 1–24. doi: 10.1016/S0378-1127(97)00019-4
- Werner PA, Murphy PG (2001) Size-specific biomass allocation and water content of above- and below-ground components of three *Eucalyptus* species in a northern Australia savanna. *Australian Journal of Botany* **49**, 155–167. doi: 10.1071/BT99026
- Williams RJ, Cook GD, Ludwig JL, Tongway D (1997) Torch, trees, teeth and tussocks: disturbance in the tropical savannas of the Northern Territory (Australia). In 'Frontiers in ecology. Building the links'. (Eds N Klomp, I Lunt) pp. 55–66. (Elsevier: Oxford)
- Williams RJ, Cook GD, Gill AM, Moore PHR (1999) Fire regime, fire intensity and tree survival in a tropical savanna in northern Australia. *Australian Journal of Ecology* **24**, 50–59. doi: 10.1046/j.1442-9993.1999.00946.x
- Williams RJ, Griffiths AD, Allan G (2002) Fire regimes and biodiversity in the wet-dry tropical savanna landscapes of northern Australia. In 'Flammable Australia: the fire regimes and biodiversity of a continent'. (Eds RA Bradstock, JA Williams, AM Gill) pp. 281–304. (Cambridge University Press: Cambridge)
- Williams RJ, Muller WJ, Wahren C-H, Setterfield SA, Cusack J (2003) Vegetation. In 'Fire in tropical savannas. The Kapalga fire experiment'. (Eds AN Andersen, GD Cook, RJ Williams) pp. 79–106. (Springer-Verlag: New York)
- Williams RJ, Hutley LB, Cook GD, Russell-Smith J, Edwards A, Chen X (2004) Assessing the carbon sequestration potential of mesic savannas in the Northern Territory, Australia: approaches, uncertainties and potential impacts of fire. *Functional Plant Biology* **31**, 415–422. doi: 10.1071/FP03215
- Wilson BA, Brocklehurst PS, Clark MJ, Dickinson KJM (1990) Vegetation survey of the Northern Territory, Australia. Technical Report No. 49. (Conservation Commission of the Northern Territory: Darwin)
- Wirth C, Schumacher J, Schulze E-D (2004) Generic biomass functions for Norway spruce in Central Europe—a meta-analysis approach toward prediction and uncertainty estimation. *Tree Physiology* **24**, 121–139.
- Zerihun A, Montagu KD, Hoffmann MB, Bray SG (2006) Patterns of below ground and above ground biomass in *Eucalyptus populnea* woodland communities of northeast Australia along a rainfall gradient. *Ecosystems* (in press).

Manuscript received 17 September 2004, accepted 4 July 2005