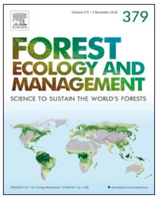
Forest Ecology and Management 412 (2018) 70–79

Contents lists available at ScienceDirect

Forest Ecology and Management

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



Validation of allometric biomass models: How to have confidence in the application of existing models

Keryn I. Paul[a,*], Philip J. Radtke[b], Stephen H. Roxburgh[a], John Larmour[a], Robert Waterworth[c], Don Butler[d], Kim Brooksbank[e], Fabiano Ximenes[f]

[a] *CSIRO Agriculture and CSIRO Land and Water, GPO Box 1700, ACT 2601, Australia*
[b] *Virginia Tech, Forest Resources & Environmental Conservation, 319E Cheatham Hall, Blacksburg, VA 24061, United States*
[c] *Mullion Group, 2a Fitzroy Rd, Forrest, ACT 2603, Australia*
[d] *Queensland Department of Science, Information Technology, Innovation and the Arts (DSITIA), Mt Coot-tha Road, Toowong, QLD 4066, Australia*
[e] *Department of Agriculture and Food, Western Australia, Private Mail Bag 50, Esperance, WA 6450, Australia*
[f] *Department of Primary Industries, Locked Bag 5123, Parramatta, NSW 2150, Australia*

ARTICLE INFO

Keywords:
Allometry
Above ground biomass
Bias
Carbon sequestration
Eucalyptus
Verification

ABSTRACT

The development of biomass estimation models is highly resource intensive as it generally entails harvesting (or excavating) trees of a range of sizes to determine dry weight of above-ground (or below-ground) biomass. To maximise the cost effectiveness of such sampling, guidance is required on whether an allometric model that already exists is suitable for a new site or species, or whether further sampling and model development is necessary.

With the aim to provide such guidance, we collated 12 pairs of well-sampled ($N > 50$) data sets of the same species at two sites, or two species at the same site. These provided case studies for: (i) assessing alternative statistical approaches to validate the application of a model developed using one data set to predict biomass of independent data from another site or species, and (ii) applying scenario analyses to explore the impact of sample size on uncertainty of validation, e.g. minimising type I and type II errors.

Our results indicate that although an allometric model for a given species or plant functional type may be applied across multiple sites, validation will be important when an existing generic multi-site and multi-species model is applied to a new species. Results obtained demonstrated that an independent sample size of $N \le 15$ frequently (37–46% of the time) provides insufficient power to avoid incorrectly accepting "validation" (type II errors). Hence, to ensure a useful outcome from resources spent in sampling biomass, it is recommended that at least 50 trees be sampled for each species. An equivalence test may then be applied to determine if the minimum detectable negligible difference between the existing model and the new independent data is < 25% (or whichever threshold is deemed acceptable). If so, the new data set may then be combined with existing data to refine a generalised model, which may then be applied with confidence. If not, then the resources expended need not be wasted as the sample size is sufficient to develop a new model suitable for application to the specific species sampled.

1. Introduction

Management of woody vegetation (e.g. reforestation or reducing rates of deforestation and degradation) can contribute to regulation of atmospheric carbon for mitigation of climate change. Carbon markets are expanding to incentivise such management (e.g. Carbonmarketdata.com, 2017). However, currently the economic viability for many vegetation management projects is low, especially for reforestation of degraded land of low productivity, and thus relatively low carbon sequestration (Maraseni and Cockfield, 2015; Rooney and Paul, 2017). To maximise participation in emerging carbon markets, costs of participation need to be minimised.

One of the largest costs associated with such reforestation projects, particularly those entailing management of mixed-species, is the sampling of biomass to either directly quantify carbon sequestered, or to calibrate models which may then be applied for this purpose. This sampling generally entails a two-stage process; first, felling of a range of representative individual trees to develop allometric models of biomass

---

[*] Corresponding author.
E-mail address: Keryn.Paul@csiro.au (K.I. Paul).

https://doi.org/10.1016/j.foreco.2018.01.016
Received 22 September 2017; Received in revised form 21 December 2017; Accepted 11 January 2018

Table 1

Paired data sets for testing validation of AGB allometric models. Data sets are described in terms of their location (latitude S, and longitude E, decimal degrees), mean annual precipitation (MAP, mm yr^{-1}), mean annual temperature (MAT, °C), sample size (*N*), range in sizes of trees sampled (minimum and maximum observed D_{130}), distribution of sizes in trees sampled (D_{130} quartiles; including the 1st quartile, median, and 3rd quartile), and the sample size used to measure the moisture content correction (*N* MC). Plant functional types (as defined by Paul et al. (2016) of species sampled was Eucalypts$^{\wedge}$ or Multi-stemmed trees^{+}. Datasets from Clyde, Flat Rock and Mogo have been previously described by Ximenes et al. (2006), while all other datasets were described by Paul et al. (2016). Photos of these sites are provided in Fig. S1.

Test	Species	Site	Location (°)	MAP (mm)	MAT (°C)	N	D_{130} range (cm)	D_{130} quartiles (cm)	*N* MC
Species	*Eucalyptus spathulata*$^{\wedge}$	Biddulph	−33.72, 119.71	438	15.9	63	3.4, 41.4	13.9, 23.3, 31, 1	16
		Moir	−34.28, 118.18	419	15.8	205	2.3, 41.3	6.0, 8.9, 12.6	41
	Eucalyptus populnea$^{\wedge}$	Boxvale	−24.41, 148.62	621	21.7	57	1.8, 37.4	7.7, 12.9, 19.7	13
		Wallal	−26.61, 146.26	456	21.1	55	1.3, 37.6	8.7, 16.1, 23.0	13
	Eucalyptus melliodora$^{\wedge}$	Sanaja	−34.87, 149.26	695	13.0	61	1.5, 46.3	9.3, 19.5, 28.7	23
		Tallawangra	−32.59, 149.43	732	15.3	60	1.2, 47.0	10.3, 19.6, 29.8	15
	Acaica harpophylla^{+}	BrigalowRS	−24.82, 149.79	658	21.5	50	3.0, 36.2	8.6, 16.1, 24.6	14
		Broadacres	−27.42, 149.94	589	19.9	50	0.9, 37.6	7.2, 16.3, 25.6	14
	Eucalyptus crebra$^{\wedge}$	Dykehead	−25.70, 151.00	628	20.6	54	2.1, 43.8	10.8, 18.2, 27.3	13
		Hinchcliff	−23.55, 150.52	797	22.1	60	1.4, 44.5	7.2, 13.6, 25.8	15
	Corymbia maculata$^{\wedge}$	Clyde	−35.45, 150.20	1173	15.6	74	11.0, 71.8	25.7, 38.0, 49.6	25
		Flat Rock	−35.42, 150.30	1226	14.6	60	10.1, 70.2	12.8, 31.2, 50.2	15
	Corymbia maculata$^{\wedge}$	Clyde	−35.45, 150.20	1173	15.6	74	11.0, 71.8	25.7, 38.0, 49.6	25
		Mogo	−35.73, 150.07	1090	15.7	166	10.0, 71.2	17.6, 24.5, 34.4	21
	Corymbia maculata$^{\wedge}$	Flat Rock	−35.42, 150.30	1226	14.6	60	10.1, 70.2	12.8, 31.2, 50.2	15
		Mogo	−35.73, 150.07	1090	15.7	166	10.0, 71.2	17.6, 24.5, 34.4	21
Site	Moir	*Eucalyptus spathulata*$^{\wedge}$	−34.28, 118.18	419	15.8	186	2.3, 15.9	5.8, 8.1, 11.3	41
		Eucalyptus occidentalis$^{\wedge}$	−34.28, 118.18	419	15.8	84	2.7, 15.6	5.3, 6.8, 8.3	23
	Wallal	*Acaica aneura*$^{+}$	−26.61, 146.26	456	21.1	55	0.7, 29.3	6.3, 9.8, 19.8	11
		Eucalyptus populnea$^{\wedge}$	−26.61, 146.26	456	21.1	51	1.3, 30.3	8.7, 14.5, 21.1	13
	Clyde	*Corymbia maculata*$^{\wedge}$	−35.45, 150.20	1173	15.6	76	11.0, 83.0	25.8, 39.3, 50.3	25
		Eucalyptus muelleriana$^{\wedge}$	−35.45, 150.20	1173	15.6	59	10.2, 84.4	14.5, 18.1, 31.6	12
	Flat Rock	*Corymbia maculata*$^{\wedge}$	−35.42, 150.30	1226	14.6	55	10.2, 68.0	13.2, 27.2, 48.7	15
		Eucalyptus saligna$^{\wedge}$	−35.42, 150.30	1226	14.6	52	10.2, 67.0	12.9, 18.9, 26.1	14

based on variables such as stem diameter (e.g. generally measured at 130 cm height above the ground, D_{130}), and second, applying these models to area-based inventories of D_{130} to estimate the stand-level biomass (e.g. Picard et al., 2012; Sileshi, 2014). It is the first stage that is particularly resource-intensive.

The impact of sample size on the accuracy of allometry parameters has been demonstrated in Sileshi (2014). To satisfy regulatory requirements and/or to provide market confidence, allometric models of above-ground biomass (AGB) or below-ground biomass (BGB) must be based on sufficient sample sizes of representative trees or shrubs to ensure predictions are repeatable, e.g. target precision of prediction of biomass (e.g. Australian Government, 2015). Using computer resampling experiments of species of trees or shrubs sampled from both planted and natural stands where the data sets maximum stem diameter was 16–79 cm, Roxburgh et al. (2015) found that to achieve this, typically > 50 individual tree measurements of AGB are required from a given site, with individuals being selected based on random stratified (size-class) sampling. They found that even larger sample sizes are required for multi-site and multi-species models (where the data sets maximum stem diameter was 32–101 cm), due to their larger inherent variability than less generalised models. As inherent variability is relatively high for allometry-predicted BGB cf. AGB (Paul et al., 2017b cf. Paul et al., 2016), the required samples size will be substantial when excavating roots to develop allometric models of BGB. In addition to sampling for fresh weights of tree components, for at least each plant functional type category of Eucalypts, Multi-stemmed trees, Shrubs and other Hardwood trees (Paul et al., 2016) at each measurement site, subsampling is required to determine corrections for moisture contents of these components (Paul et al., 2017a). Hence, for projects that have multiple species or plant functional types, and/or multiple sites, undertaking this resources intensive sampling and sub-sampling is unlikely to be economically viable (Temesgen et al., 2015).

To increase the economic viability of field sampling for biomass, some carbon markets have allowed for the use of an independent data set to provide validation of an existing allometric model, thereby minimising the requirement for extensive biomass sampling to develop new models (e.g. Australian Government, 2014). Indeed, there is currently a vast library of existing allometric models of AGB that could potentially be validated and applied (e.g. Henry et al., 2013). It is often suggested that an independent AGB sample of 10 trees may be applied to statistically justify the use of an existing AGB allometric model (e.g. Mugasha et al., 2012). However, there is a paucity of information providing guidance on what is a sufficient sample size for independent validation of an existing model.

Alternative approaches may be applied when using independent data sets to validate existing allometric models. These include testing null hypotheses that datasets either have: (i) *similar* allometry, e.g. conventional tests such as the t-tests (e.g. Fisher, 1971), or (ii) *dissimilar* allometry, i.e. tests of equivalence (e.g. Robinson and Froese, 2004; Robinson et al., 2005). Conventional hypothesis testing leads to some well-known problems in validating model predictions (Reynolds, 1984). One is that analysts sometimes incorrectly interpret a result of failing to reject a null hypothesis as evidence that the null is actually true. Another is the relationship between increased sample sizes and the power of a test to reject the null hypothesis. Simply put, if the mean difference between modelled and observed values is hypothesized to be zero, increasing the sample size will penalize the model being tested (McBride et al., 2014). There is therefore merit in exploring whether testing for *dissimilar* allometry may provide greater confidence in statistical validation results than the more traditional statistical tests for *similar* allometry. Indeed, in the context of model validation, the goal is to learn whether model predictions are in some sense equivalent – in actuality whether they are sufficiently close – to observations gathered independent of the model.

In order to inform guidelines for estimation of stand-level biomass that support operational application in vegetation management projects in carbon markets, the aim of this study was to provide recommendations on cost-effective validation of existing allometric models. There were two objectives to achieve this aim: (i) explore alternative statistical approaches that utilise independent data sets of AGB to validate