

# ARTICLE

# Sector subsampling for basal area ratio estimation: an alternative to big BAF sampling<sup>1</sup>

Xiao Dai, Mark J. Ducey, John A. Kershaw, Jr., and Haozhou Wang

**Abstract:** Big basal area factor (big BAF) sampling is a widely used subsampling method to select measure-trees. Several studies have shown big BAF sampling to be an efficient sampling scheme. In this study, we use sector sampling (Smith et al. 2008, For. Sci. **54**: 67–76) as an alternative subsample selection method. Based on some simulated mapped stands derived from three balsam fir (*Abies balsamea* (L.) Mill.) spacing trials in western Newfoundland, we show that sector subsampling is comparable to big BAF sampling in terms of estimated mean basal area ratios and their associated standard errors. Differences between big BAF sampling and sector sampling methods showed less than 1% difference across the three sites. As with big BAF sampling, changes in sample intensity had no significant (p < 0.05) effects on the accuracy of estimating mean biomass to basal area ratios and the resulting estimated mean biomasses per unit area.

Key words: sector sampling, subsampling, basal area ratio estimation, sample efficiency, big BAF sampling, biomass estimation.

**Résumé :** L'échantillonnage par grand facteur de prisme (GFP) est une méthode de sous-échantillonnage largement utilisée pour sélectionner les arbres-études. Plusieurs études ont montré que l'échantillonnage par GFP est un plan d'échantillonnage efficace. Dans cette étude, nous utilisons l'échantillonnage sectoriel (Smith et al. 2008, For. Sci. **54**: 67–76) comme méthode alternative de sélection des sous-échantillons. En se basant sur des peuplements cartographiés simulés issus de trois essais de dégagement du sapin baumier (*Abies balsamea* (L.) Mill.) dans l'ouest de Terre-Neuve, nous montrons que le sous-échantillonnage sectoriel est comparable à un échantillonnage par GFP en termes d'estimation des ratios moyens de surface terrière et de leurs erreurs-types associées. Les comparaisons entre les méthodes d'échantillonnage par GFP et par secteur montrent une différence de moins de 1 % entre les trois sites. Comme pour l'échantillonnage par GFP, les changements de densité de l'échantillon n'ont aucun effet significatif (p < 0,05) sur l'exactitude de l'estimation des ratios de la biomasse moyenne sur la surface terrière et sur les estimations de biomasses moyennes par unité de surface qui en résultent. [Traduit par la Rédaction]

*Mots-clés* : échantillonnage sectoriel, sous-échantillonnage, estimation du rapport de surface terrière, efficacité d'échantillonnage, échantillonnage par grand facteur de prisme, estimation de la biomasse.

# Introduction

Many forest-level attributes such as volume, biomass, and carbon rely on individual tree measurements and allometric models (Kershaw et al. 2016). Direct measurement of these attributes is often impractical, expensive, destructive, and therefore limited to research efforts to estimate allometric relationships (Jenkins et al. 2003; Ketterings et al. 2001). For volume and biomass estimation, allometric models that include both diameter at breast height (DBH) and total height (HT) are often more accurate and applicable to a wider range of stand conditions and ages than models that only use DBH (Honer 1967; Kershaw et al. 2016; Lambert et al. 2005; Ung et al. 2008); however, measurement of height is costly relative to counting sample trees and measuring DBH (Iles 2003; Lynch 2017; Yang et al. 2017).

Selecting a subsample of plot trees has long been used in forest inventory (Iles 2003; Marshall et al. 2004). Unfortunately, many selection methods were ad hoc or haphazard at best (Iles 2003) with the potential of introducing selection bias at the subsampling stage. An easily implemented solution to this is big basal area factor (BAF) sampling. Big BAF sampling is a form of horizontal point sampling (HPS) that utilizes two angle gauges: a small one to count "in" trees and a larger one to select trees to measure (Iles 2003; Marshall et al. 2004; Yang et al. 2017). The "measure-trees" are used to estimate the ratio of the tree attribute of interest (volume, biomass, carbon content, and so on) to tree basal area ratio of the *i*th tree,  $X_i$  is the attribute of interest for the *i*th tree, and BA<sub>i</sub> is the basal area (cross-sectional area) of the *i*th tree. The estimated mean XBAR across all measure-trees is multiplied by the estimated mean basal area per unit area determined from the count trees to obtain the estimated mean per unit area estimate of the attribute(s) of interest.

Because plot-to-plot variability in tree counts is often much greater than variability in XBAR among measure-trees, sampling effort is concentrated on selecting more count plots (Iles 2003; Marshall et al. 2004). Yang et al. (2019) developed methods for optimizing small and large BAF choices for volume estimation, and Chen et al. (2019) generalized those results for volume,

Received 24 November 2020. Accepted 2 April 2021.

X. Dai, J.A. Kershaw, Jr., and H. Wang. Faculty of Forestry and Environmental Management, University of New Brunswick, P.O. Box 4400, Fredericton, NB E3B 5A3, Canada.

M.J. Ducey. Department of Natural Resources and the Environment, University of New Hampshire, 114 James Hall Durham, NH 03824, USA.

Corresponding author: Xiao Dai (email: xdai1@unb.ca).

<sup>&</sup>lt;sup>1</sup>This article is part of the special issue "Advances in forest mensuration and biometrics, featuring papers presented at the 2020 Western Mensurationists Conference".

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from copyright.com.

**Fig. 1.** Map showing location of Newfoundland in eastern North America and locations of early spacing trials on the island of Newfoundland. Map was drawn using the world database from the maps package in R (Brownrigg 2018).



biomass, and carbon content. Chen et al. (2020) further applied big BAF sampling on a forest-level scale to correct LiDAR-derived enhanced forest inventory estimates to develop baseline carbon estimates for a carbon offset project. The ability to estimate biomass and, subsequently, carbon in a statistically accurate and cost-effective manner is important for developing improved forest management projects for monetized carbon offset projects (Chen et al. 2019, 2020).

While big BAF sampling is very efficient (in terms of its costvariance tradeoff) and logistically easy to implement in the field, it is just one of many potential subsampling schemes that could be applied to the problem of efficiently selecting trees to measure (Iles 2003, pp. 562-567). Selection based on probability proportional to prediction and systematic list sampling are other potential methods commonly used (Iles 2003). Sector sampling (Iles and Smith 2006; Smith and Iles 2012) is another potential subsampling scheme based on randomly selected sectors radiating from plot centers. Originally developed to sample small or irregular-shaped forest areas (Iles and Smith 2006), sector sampling, though not widely applied, has the potential to be a very efficient sampling scheme in certain situations. The sector orientation is randomly selected and all trees within the sector radiating from plot center to the boundary of the area of interest are measured (Iles and Smith 2006; Smith and Iles 2012). The angle of the sector is usually predetermined and all trees within a sector will be sampled with equal probability.

Sector sampling is commonly applied to small areas regardless of boundary shapes or vegetation types (Iles and Smith 2006; Smith and Iles 2012). This approach leads to unbiased estimates when appropriate randomization procedures and estimating procedures are used, because all trees within each sector are selected from the vertex point to the edge of the angle border (Lynch 2006) and can be scaled to forest-level parameters. If sector azimuths are chosen at random, tract totals and mean tree attributes can be estimated using a simple expansion factor approach (Iles and Smith 2006). Per unit area estimates require a ratio of means approach to account for different sector sizes (Smith et al. 2008; Smith and Iles 2012). The simplicity of implementing sector sampling in small areas makes it a potentially ideal alternative to big BAF sampling in some sampling situations. To our knowledge, no one has explored the potential of sector sampling as an alternative subsampling scheme similar to big BAF sampling.

The specific objectives of this study were to (*i*) estimate the efficiency of sector sample selection in comparison with big BAF selection for estimating aboveground biomass and (*ii*) determine the effects of sample intensity across the three different subsampling selection methods used in this study.

### Materials and methods

#### Study sites

Data from three early spacing trials located on western Newfoundland (NL), Canada, were used in this study (Fig. 1). The spacing trials were established in the early 1980s by the government of Newfoundland in cooperation with the Canadian Forest Service (Donnelly et al. 1986). The sites were dominated by balsam fir (Abies balsamea L.) with minor components of black spruce (Picea mariana (Mill.) Britton, Sterns & Poggenb.) and white birch (Betula papyrifera Marshall). There were five spacing treatments: control or no spacing (S00); 1.2 m spacing (S12); 1.8 m spacing (S18); 2.4 m spacing (S24); and 3.0 m spacing (S30). The treatments were arranged in a randomized complete block design with three replicates per site  $(3 \times 3 \times 5 = 45$  permanent sample plots were used in this study). Each treatment was applied to a 0.25 ha block (50 m  $\times$  50 m), and a circular permanent sample plot (PSP) was established near the center of each block. The PSP size varied such that there were approximately 100 trees per plot at the time of establishment.

## Simulation of subsampling protocols

Each NL plot was expanded to a 1 ha "mapped" plot using simulation of spatial locations and random sampling of individual trees from each PSP's tree list (Fig. 2A). For control plots (S00), locations were randomly assigned by generating random  $\{x, y\}$  coordinates. For thinned plots, the 1 ha area was divided into cells based on average spacing. For example, when applied to the 2.4 m spacing in simulations, the 1 ha area was divided into 2.4 m × 2.4 m cells. Within each cell, a tree was randomly located by generating a random  $\{x, y\}$  coordinate within the bounds of the cell, and a tree was randomly drawn from the tree list for that PSP with replacement.

A 2 M count BAF (i.e., each treed tallied =  $2 \text{ m}^2 \cdot \text{ha}^{-1}$  basal area) was simulated to select count ("IN") trees for basal area estimation (Fig. 2B). Three different subsample selection methods were used to select trees for height measurement (Fig. 2B). The first

2

Fig. 2. Simulation of (A) 1 ha spacing plots and (B) count tree selection and the three measure-tree subsample selection methods.



used a big BAF approach. Measure-trees were selected using a large BAF. We tested five different big BAFs: 20, 30, 40, 50, and 60 M. The second method used sector sampling to subsample count trees (SectorIN). Five different sector intensities, expressed as percentages of the full compass (360°), were used in this study: 10%, 7%, 5%, 4%, and 3%. The midpoint azimuth of the sector was randomly oriented and all count ("IN") trees within the sector were selected for measurement (DBH and HT). The third method used sectors to select trees, but independent of whether they were count trees. All trees within the sector and within a specified distance from plot center (11.28 m was used in these simulations) were selected for measurement (DBH and HT). Again, we used five different sector intensities: 1.2%, 1.0%, 0.7%, 0.5%, and 0.4%. The big BAF and two different sets of sector intensities were selected to give approximately the same number of measure-trees across the samples (Table 1). Three points were randomly selected within each 1 ha simulated plot. At each sample point, the count trees and measure-trees were determined. The walk-through method (Ducey et al. 2004) was used to account for any boundary overlap. The simulations were repeated 100 times, and convergence of the statistical results reported here were checked by examining the cumulative standard deviations of estimated mean biomass over simulation order compared with bootstrapped simulation order.

#### **Biomass estimation**

Individual tree biomass was estimated for each tree in the simulated plots using the Canadian national biomass models (Lambert



(B) Measure-Tree Subsample Selection

et al. 2005). We used eq. 3 from table 4 in Lambert et al. (2005), which utilized both DBH and HT. Total tree biomass (BM<sub>i</sub>) was estimated by summing the biomass estimates of the separate components (wood, bark, branches, and foliage). For the simulated 1 ha plots, "true" field biomass per hectare (FBM;  $t \cdot ha^{-1}$ ) was obtained by summing the biomass estimates across all trees and dividing by 1000 kg·t<sup>-1</sup>:

(1) 
$$FBM = \frac{\sum_{i=1}^{n} BM_i}{1000}$$

The mean biomass to basal area ratio ( $\overline{BBAR}$ ; kg·m<sup>-2</sup>) was estimated from the measure-trees and used to estimate mean BM for the simulated subsampling methods ( $\overline{SBM}$ ) using estimated  $\overline{BA}$ :

(2) 
$$\overline{\text{SBM}} = \frac{\overline{\text{BBAR}} \times \overline{\text{BA}}}{1000}$$

where

(3) 
$$\overline{BA} = \frac{\sum_{j=1}^{p} BA_j}{p} = \frac{\sum_{j=1}^{p} BAF \times Count_j}{p}$$

and p is the number of count plots; BA<sub>j</sub> is the BA per hectare on the *j*th plot; and Count<sub>j</sub> is the number of "in" trees on the *j*th plot. The estimation of BBAR depended on which subsampling method was employed. Big BAF selection and SectorIN selection

	Sample intensity	Sample selection method						
Study site		Big BAF		SectorIN		SectorDST		
		Level	Size	Level	Size	Level	Size	
Cormack	1	20	100 (7.9)	10	101 (8.2)	1.2	107 (9.6)	
			[79, 119]		[76, 122]		[87, 137]	
	2	30	67 (7.1)	7	71 (7.5)	1.0	90 (9.3)	
			[46, 82]		[56, 91]		[71, 117]	
	3	40	50 (5.8)	5	50 (6.0)	0.7	61 (7.3)	
			[38, 63]		[36, 66]		[41, 79]	
	4	50	40 (5.9)	4	41 (6.0)	0.5	45 (6.8)	
			[26, 53]		[27, 57]		[29, 61]	
	5	60	34 (4.7)	3	30 (4.6)	0.4	36 (6.1)	
			[22, 43]		[21, 41]		[25, 53]	
Pasadena	1	20	154 (8.4)	10	152 (9.8)	1.2	87 (7.7)	
			[137, 177]		[130, 184]		[68, 108]	
	2	30	102 (7.5)	7	107 (8.5)	1.0	72 (7.8)	
			[83, 118]		[82, 127]		[55, 95]	
	3	40	77 (6.7)	5	77 (6.8)	0.7	50 (6.7)	
			[63, 92]		[62, 94]		[32, 65]	
	4	50	62 (6.6)	4	61 (6.8)	0.5	35 (6.19)	
			[50, 78]		[45, 77]		[22, 52]	
	5	60	51 (5.0)	3	46 (5.8)	0.4	29 (4.9)	
			[39, 67]		[33, 59]		[20, 41]	
Roddickton	1	20	116 (8.9)	10	114 (9.2)	1.2	115 (9.9)	
			[88, 134]		[93, 145]		[94, 134]	
	2	30	77 (7.8)	7	80 (7.9)	1.0	95 (9.4)	
			[59, 96]		[61, 99]		[74, 123]	
	3	40	58 (6.2)	5	57 (6.4)	0.7	68 (6.8)	
			[44, 79]		[35, 73]		[50, 84]	
	4	50	46 (5.7)	4	46 (5.5)	0.5	46 (6.5)	
			[32, 59]		[34, 61]		[30, 63]	
	5	60	39 (5.6)	3	34 (5.7)	0.4	38 (6.3)	
			[25, 52]		[21, 53]		[26, 56]	

Table 1. Mean number of measure-trees, standard errors (in parentheses), and ranges [in brackets] by sample selection method, study site, and sample intensity.

Note: Factor levels represent metric BAFs for big BAF sampling and sector widths as a percentage of full circle (360°) for sector subsampling.

were both variable probability methods and the mean ratio approach (Kershaw et al. 2016) was used:

(4a) 
$$\overline{\text{BBAR}} = \frac{\sum_{i=1}^{m} \text{BBAR}_{i}}{m} = \frac{\sum_{i=1}^{m} \left(\frac{\text{BM}_{i}}{\text{BA}_{i}}\right)}{m}$$

For the SectorDST method, measure-trees were selected with equal probability; therefore, we used a ratio of means approach (Kershaw et al. 2016):

(4b) 
$$\overline{\text{BBAR}} = \frac{\sum_{i=1}^{m} \text{BM}_i}{\sum_{i=1}^{m} \text{BA}_i} = \frac{\sum_{i=1}^{m} \text{BM}_i}{\sum_{i=1}^{m} 0.00007854 \times (\text{DBH}_i)^2}$$

Percent standard error for SBM was estimated from two independent variables based on Bruce's formula (Chen et al. 2020; Hsu et al. 2020; Iles 2003; Marshall et al. 2004; Yang et al. 2017):

(5) 
$$\% e(\overline{SBM}) = \sqrt{\% se(\overline{BA})^2 + \% se(\overline{BBAR})^2}$$

where %se() is the estimated standard error as a percentage of the estimated mean. Gove et al. (2020) showed the relationship between Bruce's formula and the Delta method. While Bruce's formula does involve an assumption of independence between BA and BBAR, Gove et al. (2020) found in simulations that the impact of non-independence was negligible for the forest conditions that they used in their study. For  $\$se(\overline{BA})$  and  $\$se(\overline{BBAR})$ for big BAF and SectorIN selection, we used the estimator derived from simple random sampling to estimate standard error (se):

(6a) 
$$\operatorname{se}(\overline{x}) = \frac{s}{\sqrt{n}} = \sqrt{\frac{\sum x^2 - (\sum x)^2/n}{n(n-1)}}$$

where x is either BA for the sample point or BBAR for the individual measure-tree, s is the estimated standard deviation, and n is sample size (number of sample points or number of measure-trees). For SectorDST selection, we used (Kershaw et al. 2016):

(6b) 
$$\operatorname{se}(\overline{\operatorname{BBAR}}) = \sqrt{\left(\frac{\overline{\operatorname{BBAR}}^2}{n(n-1)}\right)\left(\frac{\sum \mathrm{BM}^2}{\overline{\mathrm{BM}}^2} + \frac{\sum \mathrm{BA}^2}{\overline{\mathrm{BA}}^2} - \frac{2\sum \mathrm{BM} \cdot \mathrm{BA}}{\overline{\mathrm{BM}} \cdot \overline{\mathrm{BA}}}\right)}$$

where BM and  $\overline{BM}$  are the individual biomass estimates and the estimated mean biomass for the measure-trees, respectively; BA and BA are the basal areas (cross-sectional areas) and estimated mean basal area (cross-sectional area of the measure-trees), respectively; and BBAR is the estimated mean biomass to basal area ratio from eq. 4b.

4

Dai et al.

)

		Big	Big BAF		SectorIN		SectorDST	
Selection method	Sample intensity	BBAR (kg⋅m <sup>-2</sup> )	Biomass (t∙ha <sup>-1</sup> )	BBAR (kg⋅m <sup>-2</sup> )	Biomass (t∙ha <sup>-1</sup> )	BBAR (kg⋅m <sup>-2</sup> )	Biomass (t·ha <sup>-1</sup> )	
Cormack	1	3162 (43.0)	140.8 (4.0)	3157 (45.2)	140.6 (4.4)	3164 (55.0)	140.9 (4.4)	
		[3066, 3268]	[131.9, 150.9]	[3004, 3305]	[129.9, 152.7]	[2957, 3328]	[129.3, 150.9]	
	2	3151 (56.0)	140.3 (4.4)	3162 (48.3)	140.7 (3.6)	3156 (66.4)	140.5 (4.5)	
		[3010, 3286]	[128.3, 150.0]	[3008, 3262]	[131.8, 149.5]	[2962, 3322]	[128.5, 151.4]	
	3	3153 (58.1)	140.4 (4.3)	3155 (65.9)	140.5 (4.6)	3147 (81.6)	140.1 (5.1)	
		[2975, 3285]	[127.3, 151.8]	[2998, 3324]	[130.3, 151.2]	[2917, 3304]	[125.2, 153.0]	
	4	3155 (68.9)	140.8 (4.5)	3147 (69.6)	140.5 (4.5)	3146 (83.3)	140.4 (5.3)	
		[2934, 3282]	[128.1, 154.3]	[2956, 3285]	[127.4, 151.8]	[2904, 3381]	[128.9, 162.1]	
	5	3154 (71.2)	140.4 (4.3)	3151 (89.7)	140.3 (5.3)	3143 (103.7)	139.9 (5.0)	
		[2983, 3310]	[128.2, 152.2]	[2899, 3405]	[128.2, 155.5]	[2773, 3353]	[127.4, 153.4]	
Pasadena	1	3205 (17.5)	218.7 (3.8)	3205 (20.0)	218.7 (3.9)	3201 (28.7)	218.5 (4.2)	
		[3164, 3256]	[209.3, 226.9]	[3149, 3261]	[208.4, 225.9]	[3127, 3268]	[207.5, 227.3]	
	2	3204 (25.5)	218.3 (3.8)	3204 (24.5)	218.4 (4.1)	3199 (32.0)	218.0 (4.44)	
		[3148, 3272]	[209.0, 226.7]	[3138, 3252]	[208.3, 227.3]	[3120, 3314]	[207.6, 230.3]	
	3	3205 (28.1)	219.0 (3.9)	3205 (27.9)	219.1 (3.7)	3198 (43.1)	218.5 (4.6)	
		[3132, 3258]	[208.6, 229.9]	[3117, 3272]	[208.8, 227.3]	[3098, 3307]	[207.1, 231.6]	
	4	3207 (28.7)	218.8 (3.4)	3209 (32.2)	218.9 (3.8)	3195 (49.7)	217.9 (4.4)	
		[3141, 3270]	[210.3, 227.1]	[3136, 3282]	[208.0, 230.2]	[3083, 3365]	[208.2, 232.8]	
	5	3208 (35.3)	218.5 (3.9)	3209 (33.1)	218.5 (3.7)	3207 (55.6)	218.4 (5.0)	
		[3123, 3297]	[210.2, 230.4]	[3118, 3286]	[209.7, 231.2]	[3052, 3339]	[205.5, 228.7]	
Roddickton	1	2924 (26.3)	150.2 (3.7)	2926 (24.7)	150.3 (3.5)	2924 (37.9)	150.1 (3.9)	
		[2860, 3003]	[141.4, 160.0]	[2876, 3015]	[140.7, 159.9]	[2835, 3021]	[141.7, 160.3]	
	2	2930 (32.5)	150.7 (3.5)	2928 (35.8)	150.6 (3.5)	2933 (45.6)	150.9 (3.8)	
		[2848, 3000]	[142.9, 159.5]	[2852, 3022]	[142.4, 159.7]	[2834, 3059]	[143.5, 160.7]	
	3	2928 (41.9)	150.0 (3.6)	2921 (42.1)	149.6 (3.7)	2917 (52.0)	149.4 (4.1)	
		[2816, 3043]	[139.5, 159.4]	[2785, 3037]	[138.4, 157.4]	[2747, 3069]	[138.7, 160.1]	
	4	2927 (42.3)	150.1 (3.7)	2926 (51.7)	150.0 (3.9)	2929 (68.3)	150.2 (4.9)	
		[2802, 3013]	[141.9, 162.5]	[2767, 3027]	[139.3, 157.5]	[2773, 3078]	[137.8, 164.4]	
	5	2918 (50.2)	149.8 (4.0)	2923 (60.0)	150.1 (4.4)	2919 (67.8)	149.9 (4.8)	
		[2784, 3033]	[140.8, 159.5]	[2784, 3084]	[141.1, 160.9]	[2709, 3059]	[137.0, 160.4]	

**Table 2.** Estimated means, standard errors (in parentheses), and ranges [in brackets] for BBAR ( $kg \cdot m^{-2}$ ) and biomass (tonnes (t)·ha<sup>-1</sup>) by study site, measure-tree selection method, and sample intensity for the western Newfoundland (NL) spacing trials.

Note: Measure-tree sample intensities are defined in Table 1.

The three different measure-tree selection methods were compared on the basis of  $\overline{\text{BBAR}}$ , %se( $\overline{\text{BBAR}}$ ), mean biomass estimates and their distributions, equivalence tests, and rank correlations by study site and sample intensity. All simulations and analyses were conducted in the R statistical language (R Core Team 2019).

# Results

## Estimated mean biomass to basal area ratios

Estimated  $\overline{\text{BBARs}}$  (kg·m<sup>-2</sup>) from the 100 sample simulations using the simulated 1 ha plots did not vary substantially by study site, measure-tree selection method, or sample intensity (Table 2; Fig. 3). Overall estimated  $\overline{\text{BBAR}}$  (mean of the 100 simulated sample means) ranged from about 3000 kg·m<sup>-2</sup> on Roddickton to 3200 kg·m<sup>-2</sup> on Cormack and Pasadena (Table 2). Cormack had greater variability in estimated  $\overline{\text{BBAR}}$  than the other two sites (Figs. 3 and 4). Comparisons within study sites but across both measure-tree subsample selection method and measure-tree sample intensity were much closer, with differences consistently well below 1% of the estimated  $\overline{\text{BBARs}}$  (Table 2).

While overall estimated  $\overline{BBAR}$  did not change across the range of sample intensities, the range of estimated  $\overline{BBARs}$  (Fig. 3) and the associated standard errors (Fig. 4) of estimated  $\overline{BBARs}$  increased with decreasing sample intensity (Table 2). There were slight biases observed between the overall estimated  $\overline{BBARs}$  and the "true" population  $\overline{\text{BBAR}}$ s (Fig. 3) as calculated using all trees across the 1 ha simulated plots on each study site. Here, bias is assessed relative to the mean computed from a large number of repeated simulations, which stands in for the unknown population mean. The bias for Roddickton was about twice that observed on the other two sites ( $\approx 6\%$  on Roddickton and  $\approx 3\%$  on Cormack and Pasadena); however, this was the result of a single tree measurement with a potential error in measured height (almost double all other trees). When the observation for this tree was eliminated from both the population estimate and the measure-tree estimates, the bias was reduced to just under 3%, as observed on the other two sites.

### **Biomass estimates**

Overall estimated mean biomass (t·ha<sup>-1</sup>) varied by site, reflecting inherent differences in site productivity across the three sites (Table 2; Supplementary Fig. S1<sup>2</sup>) with Pasadena > Roddickton > Cormack. While overall mean biomass varied across the three sites, estimated mean biomasses such as estimated mean BBARs did not vary substantially by measure-tree selection method or sample intensity (Table 2). Similar to estimated BBAR, the range of estimates and the associated standard errors for biomasss increased with decreasing sample intensity (Table 2; Supplementary Figs. S1 and S2<sup>2</sup>); however, while measure-tree sample sizes Fig. 3. Distribution of BBARs by study site, measure-tree subsample selection method, and measure-tree subsample intensity for the Newfoundland (NL) spacing trials. Horizontal grey bars are the overall mean BBARs generated from each simulation method under different measure intensities. Dashed black line is the "true" population BBAR based on all trees in the 1 ha simulated plots within each study site. (For Measure BAF, intensity is expressed in terms of metric basal area factors, m<sup>2</sup>·ha<sup>-1</sup> per tree tallied; for SectorIN and SectorDST, intensity is expressed in terms of percentage of full circle.)



**Decreasing Sample Intensity** 

typically decreased by 70% (Table 1), standard errors only increased by 15% or less (Table 2).

For the 100 sample simulations conducted on each site, the nominal 95% confidence intervals included the "true" population means for all replicate samples across all sites × measure-tree selection methods  $\times$  measure-tree sample intensities (Supplementary Figs. S3-S5<sup>2</sup>). The "true" population mean was calculated as the sum of the individual trees on each simulated 1 ha plot and averaged across the 15 plots within each spacing trial. Based on the trends observed for cumulative standard deviation of the simulation means, 100 simulations were adequate for convergence of the statistical results reported here (Supplementary Figs. S6–S8<sup>2</sup>).

At the largest measure-tree sample intensities, the correspondences between estimated mean biomass among the three measure-tree subsample selection methods were quite good (Supplementary Figs. S9-S11<sup>2</sup>; Table 3). As sample intensity decreased, the relationships became increasingly noisy (Supplementary Figs. S9–S11<sup>2</sup>), especially for Cormack (Supplementary Fig. S9<sup>2</sup>). Spearman's rank correlation decreased with decreasing sample intensity, while the minimum detectable negligible differences increased (Table 3). Even though minimum detectable, non-negligible differences (MDNDs) increased with decreasing sample intensity, when expressed as a percentage of standard deviation (Table 3), most were less than 25%, which is generally considered sufficient to conclude that the two samples are statistically equivalent (Robinson and Froese 2004).

# Discussion

Big BAF sampling is a well-established and increasingly used method for selecting a subsample of measure-trees in a variety of forest inventory applications (Chen et al. 2019; Corrin 1998; Desmarais 2002; Iles 2003; Marshall et al. 2004; Yang et al. 2017). Because variability in counts of "in" trees between sample points is generally greater than the variability in the volume to basal area ratio (Marshall et al. 2004; Yang et al. 2017) or biomass to basal area ratio (Chen et al. 2019), big BAF sampling places inventory effort on establishing more count plots than measuring sample trees (Iles 2003; Marshall et al. 2004). The alternative selection methods proposed in this study using sector sampling were comparable to big BAF measure-tree selection in terms of both average BBAR (Table 2; Fig. 3) and percent standard error of mean BBAR (Fig. 4). Differences in mean BBARs across the three measure-tree selection methods averaged less than 0.2% across the three spacing trials and five sample intensities (Table 3). Because estimated biomass per hectare is simply BBAR multiplied by average BA (which was constant for all three sample selection methods and measure-tree sample intensities), biomass per hectare did not vary substantially among the measure-tree selection methods as well (Tables 2 and 3).

Big BAF selection is a subsampling protocol (Marshall et al. 2004). The trees selected using the big angle gauge are a subset of the trees counted using the small angle gauge (Iles 2003; Marshall et al. 2004). Similarly, the SectorIN selection method proposed here subsamples count trees (Fig. 2); however, the SectorDST method potentially selects a mixture of count and non-count trees (Fig. 2). Both big BAF and SectorIN select sample trees with variable probability and the mean ratio (eq. 4a) and its associated standard error (eq. 6a) were used. The SectorDST method selects sample trees with equal probability; therefore, the ratio of means estimator (eq. 4b) and its associated standard error (eq. 6b) were used. Despite the different selection probabilities and associated

6

Dai et al.

**Fig. 4.** Distribution of BBAR errors by study site, measure-tree subsample selection method, and measure-tree subsample intensity for the Newfoundland (NL) spacing trials. Dark grey bars are the mean errors generated from each method under different measure intensities. (For Measure BAF, intensity is expressed in terms of metric basal area factors,  $m^2 \cdot ha^{-1}$  per tree tallied; for SectorIN and SectorDST, intensity is expressed in terms of percentage of full circle.)

4 Cormack 4  $\sim$  $\sim$ ŝ 2 Errors of Simulated Biomass to Basal Area Ratio 0 C 30 40 50 60 10 Ż 5 4 ż 1.2 1.0 0.7 0.5 0.4 20 4 Pasadena 4 4 m m m  $\sim$ 2 2 0 0 0 30 40 50 60 ż 5 4 ż 1.2 1.0 0.7 0.5 0.4 20 10 4 Roddickton 4 4  $\sim$ m m 2 0 0 0 ż 5 4 30 40 50 60 10 3 1.2 1.0 0.7 0.5 0.4 20 Measure BAF SectorIN intensity (%) SectorDST intensity (%)

**Decreasing Sample Intensity** 

**Table 3.** Comparisons of minimum detectable, non-negligible differences (MDND) and Spearman's rank correlation coefficients ( $\gamma$ ) of estimated mean biomass by study site, measure-tree selection method, and measure-tree sample intensity for the western Newfoundland spacing trials (MDNDs are expressed in tonnes per hectare (t-ha<sup>-1</sup>), with percentage of standard deviation of differences in parentheses, and represents the minimum percent difference required to reject the null hypothesis of the equivalence test: there IS a significant difference).

	Sample intensity <sup>a</sup>	Measure-tree selection method						
Study site		Big BAF vs. SectorIN		Big BAF vs. SectorDST		SectorIN vs. SectorDST		
		MDND	γ	MDND	γ	MDND	γ	
Cormack	1	0.65 (24)	0.83	0.64 (20)	0.70	0.83 (26)	0.77	
	2	0.94 (31)	0.72	0.88 (21)	0.57	0.89 (24)	0.60	
	3	0.76 (19)	0.60	1.01 (23)	0.55	1.15 (24)	0.52	
	4	0.99 (25)	0.58	1.17 (25)	0.46	0.86 (18)	0.46	
	5	0.96 (19)	0.39	1.37 (26)	0.42	1.42 (23)	0.25	
Pasadena	1	0.28 (17)	0.91	0.62 (27)	0.83	0.66 (26)	0.80	
	2	0.44 (19)	0.81	0.74 (27)	0.81	0.83 (29)	0.79	
	3	0.48 (18)	0.73	1.09 (30)	0.63	1.09 (32)	0.71	
	4	0.66 (22)	0.68	1.43 (39)	0.59	1.66 (42)	0.50	
	5	0.62 (20)	0.64	0.78 (17)	0.57	0.89 (20)	0.52	
Roddickton	1	0.34 (21)	0.88	0.43 (19)	0.79	0.50 (22)	0.78	
	2	0.49 (21)	0.76	0.59 (22)	0.74	0.75 (25)	0.67	
	3	0.80 (28)	0.71	1.14 (33)	0.59	0.84 (23)	0.57	
	4	0.60 (18)	0.61	0.78 (19)	0.52	0.88 (20)	0.53	
	5	0.86 (24)	0.63	0.76 (18)	0.48	1.06 (20)	0.34	

<sup>a</sup>Sample intensities are defined in Table 1.

8

mean and error estimates, all three selection methods had comparable results (Table 2; Figs. 3 and 4).

Sector sampling was originally developed to efficiently sample small, irregular areas (Smith and Iles 2012). Because randomness enters the selection process through the azimuth of the sector plot, sector sampling has several interesting properties: (*i*) placement of the sampling point (sector vertex) can be subjective provided that the randomization of the sector azimuth is independent of sector placement; and (*ii*) because selection probabilities are based on sector angle, there are no boundary overlap issues (Iles and Smith 2006; Smith and Iles 2012). Despite these advantages, sector sampling has seen limited application in the field. Sector sampling does not easily scale to larger areas because of the difficulty in tracing sectors over long distances in the field. The applications developed here are a novel use of sector sampling and could be employed as a useful method of subsampling measure-trees on fixed-area plots.

As Lynch (2006) observed, the change in sample size results in variance changes. Here, sample intensity did not influence the means of either BBAR or the associated biomass per hectare estimates; however, variability was greatly impacted (Table 2; Fig. 3; Supplementary Fig. S1<sup>2</sup>). This is similar to results reported by Yang et al. (2017) and Chen et al. (2019) and is consistent with the sampling theory as proposed by Marshall et al. (2004). As Yang et al. (2017) demonstrated, the big BAF can vary from 10 M to almost 100 M without impacting the overall errors of estimates or required sample sizes for estimating volume. The sector sampling methods developed here were comparable to the big BAF approach, showing that sector subsampling is a viable alternative to big BAF sampling (Table 2; Fig. 4).

Yang et al. (2017) developed methods for analyzing the costerror surface based on the Fairfield Smith equation (Lynch 2017). Chen et al. (2019) generalized this approach for estimating carbon content in eastern North America and developed methods to optimize choice of small and big BAF. Given the resulting standard errors and the associated numbers of trees requiring measurement, there is no reason to expect that the relationships for sector subsampling would deviate substantially from the relationships established for big BAF selection; however, this needs further study.

While we applied sector subsampling with horizontal point sampling, it could just as easily be applied using fixed-area plots; however, the use of fixed-area plots would require measurement of all tree diameters on both the sector trees and the plot trees to estimate basal area, which would reduce the efficiency of the approach presented here. Hsu (2019) presented some alternative measures to basal area for ratio estimation using spherical photos. Results from Hsu (2019) show that the strength of the correlation between the two variables and the variation in the denominator variable (i.e., the covariate) drive the efficiency of the sampling strategy. With fixed-area plots, stand average canopy height or stand density might be more cost-effective variables to measure in the field than basal area. Biomass to stand height or biomass to density ratios could be estimated and used in place of the biomass to basal area ratios as used here (Iles (2003) demonstrates this for a number of different stand variables). As Hsu (2019) showed for other metrics, these ratios may be more variable, requiring more sample plots and more measurements within a sample plot. The use of alternative ratios requires more study to evaluate the efficiency of these estimators. The Fairfield Smith equation (Chen et al. 2019; Lynch 2017; Yang et al. 2017) may be an approach for evaluating the cost-error surfaces and developing optimal sample designs for these alternative ratios. As Wang (2019) proposed, spherical images were an effective medium for extracting forest attributes. With no restriction on plot shapes, photo plots selected from spherical images combined with sector subsampling may be a promising approach for estimating forest attributes as well.

# Acknowledgements

We are grateful to the Newfoundland and Labrador Department of Fisheries, Forestry and Agriculture for the use of their spacing trial data and allowing us access to acquire images on their study sites. This work was partially funded by a Contribution Agreement with the Government of Newfoundland and Labrador, the Natural Sciences and Engineering Research Council of Canada, Discovery Grant Program (RGPIN04280), and the New Brunswick Innovation Foundation Research Assistantship Initiative (RAI2017-032). The corresponding author is grateful for an Accelerated Master's Award from the New Brunswick Department of Postsecondary Education and Labour.

# References

- Brownrigg, R. 2018. maps: Draw geographical maps. R package version 3.3.0 (original S code by Richard A. Becker, Allan R. Wilks with enhancements by Thomas P Minka and Alex Deckmyn). Available from https://cran. r-project.org/package=maps [last accessed 8 April 2021].
- Chen, Y., Yang, T.-R., Hsu, Y.-H., Kershaw, J.A., and Prest, D. 2019. Application of big BAF sampling for estimating carbon on small woodlots. For. Ecosyst. 6: 13. doi:10.1186/s40663-019-0172-4.
- Chen, Y., Kershaw, J.A., Hsu, Y.-H., and Yang, T.-R. 2020. Carbon estimation using sampling to correct LiDAR-assisted enhanced forest inventory estimates. For. Chron. 96: 9–19. doi:10.5558/tfc2020-003.
- Corrin, D. 1998. BIG BAF: a very efficient sampling method for cruising timber. John Bell Associates.
- Desmarais, K.M. 2002. Using BigBAF sampling in a New England mixedwood forest. John Bell Associates. Inventory Cruis. Newsl. 58.
- Donnelly, J.G., Lavigne, M.B., and van Nostrand, R.S. 1986. Precommercial thinning spacing trials established between 1979 and 1985. Newfound-land Forestry Centre, St. John's, Newfoundland.
- Ducey, M.J., Gove, J.H., and Valentine, H.T. 2004. A walkthrough solution to the boundary overlap problem. For. Sci. 50: 427–435. doi:10.1093/forestscience/ 50.4.427.
- Gove, J.H., Gregoire, T.G., Ducey, M.J., and Lynch, T.B. 2020. A note on the estimation of variance for big BAF sampling. For. Ecosyst. 7: 62. doi:10.1186/ s40663-020-00272-x.
- Honer, T. 1967. Standard volume tables and merchantable conversion factors for the commercial tree species of central and eastern Canada. Information Report FMR-X-5, Forest Management Research and Services Institute, Ottawa, Ontario.
- Hsu, Y.-H. 2019. Applications of variable probability sampling using remotely sensed covariates. M.Sc. Forestry thesis, University of New Brunswick, Fredericton, N.B., Canada.
- Hsu, Y.-H., Chen, Y., Yang, T.-R., Kershaw, J.A., and Ducey, M.J. 2020. Sample strategies for bias correction of regional LiDAR-assisted forest inventory estimates on small woodlots. Ann. For. Sci. 77: 75. doi:10.1007/s13595-020-00976-8.
- Iles, K. 2003. A sampler of inventory topics. Kim Isles & Associates Ltd., Nanimo, B.C., Canada.
- Iles, K., and Smith, N.J. 2006. A new type of sample plot that is particularly useful for sampling small clusters of objects. For. Sci. 52: 148–154. doi:10.1093/ forestscience/52.2.148.
- Jenkins, J.C., Chojnacky, D.C., Heath, L.S., and Birdsey, R.A. 2003. Nationalscale biomass estimators for United States tree species. For. Sci. 49: 12–35. doi:10.1093/forestscience/491.12.
- Kershaw, J.A., Ducey, M.J., Beers, T.W., and Husch, B. 2016. Forest mensuration. 5th ed. John Wiley & Sons.
- Ketterings, Q.M., Coe, R., van Noordwijk, M., Ambagau', Y., and Palm, C.A. 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. For. Ecol. Manage. 146: 199–209. doi:10.1016/S0378-1127(00)00460-6.
- Lambert, M.-C., Ung, C.-H., and Raulier, F. 2005. Canadian national tree aboveground biomass equations. Can. J. For. Res. 35(8): 1996–2018. doi:10.1139/x05-112.
- Lynch, T.B. 2006. Variance reduction for sector sampling. For. Sci. 52: 251– 261. doi:10.1093/forestscience/52.3.251.
- Lynch, T.B. 2017. Optimal plot size or point sample factor for a fixed total cost using the Fairfield Smith relation of plot size to variance. Forestry, 90: 211–218. doi:10.1093/forestry/cpw038.
- Marshall, D.D., Iles, K., and Bell, J.F. 2004. Using a large-angle gauge to select trees for measurement in variable plot sampling. Can. J. For. Res. 34(4): 840– 845. doi:10.1139/x03-240.
- R Core Team. 2019. R: a language and environment for statistical computing. The R Foundation for Statistical Computing, c/o Institute for Statistics and Mathematics, Vienna, Austria. Available from www.R-project.org.
- Robinson, A.P., and Froese, R.E. 2004. Model validation using equivalence tests. Ecol. Model. 176: 349–358. doi:10.1016/j.ecolmodel.2004.01.013.

Dai et al.

- Smith, N.J., and Iles, K. 2012. Sector sampling synthesis and applications. Forests, 3: 114–126. doi:10.3390/f3010114.
- Smith, N.J., Iles, K., and Raynor, K. 2008. Investigation of some sector sampling statistical properties. For. Sci. 54: 67–76. doi:10.1093/forestscience/ 54.1.67.
- Ung, C.-H., Bernier, P., and Guo, X.-J. 2008. Canadian national biomass equations: new parameter estimates that include British Columbia data. Can. J. For. Res. 38(5): 1123–1132. doi:10.1139/X07-224.
- Wang, H. 2019. Estimating forest attributes from spherical images. M.Sc. Forestry thesis, University of New Brunswick, Fredericton, N.B., Canada.
- Yang, T.-R., Hsu, Y.-H., Kershaw, J.A., McGarrigle, E., and Dan, K. 2017. Big BAF sampling in mixed species forest structures of northeastern North America: influence of count and measure BAF under cost constraints. For. Int. J. For. Res. **90**: 649–660. doi:10.1093/forestry/cpx020.Yang, T.-R., Kershaw, J.A., Weiskittel, A.R., Lam, T.Y., and McGarrigle, E.
- Yang, T.-R., Kershaw, J.A., Weiskittel, A.R., Lam, T.Y., and McGarrigle, E. 2019. Influence of sample selection method and estimation technique on sample size requirements for wall-to-wall estimation of volume using airborne LiDAR. For. Int. J. For. Res. 92: 311–323. doi:10.1093/forestry/cpz014.