

# Comparison of sampling methods for estimation of nearest-neighbor index values

Francisco Mauro, Zane Haxtema, and Hailemariam Temesgen

**Abstract:** Neighborhood-based indices such as mingling index and diameter differentiation are a set of diversity measures that are based on the relationship between a reference tree and a certain number of nearest neighbors (i.e., trees to which it has the lowest horizontal distance). Using stem-mapped data from eight headwater sites, we compared the relative bias and relative root mean square error (relative to the true mean of each site) of several different methods of choosing reference trees for calculation of diameter differentiation ( $\bar{D}$ ) and species mingling ( $\bar{M}$ ) index. Indices were defined using two, three, and four neighbors and methods for selection of the reference tree were random selection of a tree in a fixed-radius plot (FI), random selection of a tree in a variable-radius plot (VA), azimuth selection method (AZ), and nearest tree selection (NT). In general, the relative bias was lower than  $\pm 2.5\%$  for  $\bar{D}$  and lower than  $\pm 10\%$  for  $\bar{M}$  regardless of the method. The FI method consistently had the lowest relative bias and relative root mean squared error. The NT and AZ methods were second in terms of relative root mean squared error for  $\bar{D}$  and  $\bar{M}$ , respectively. Simplicity of these two methods might outweigh their slightly worse performance.

**Key words:** species mingling, diameter differentiation, headwater stream, density management studies, riparian zones.

**Résumé :** Les indices basés sur le voisinage, tels que l'indice de mélange et la différenciation de diamètre, constituent un groupe de mesures de diversité basées sur la relation entre un arbre de référence et un certain nombre de plus proches voisins (c.-à-d. les arbres qui sont à la plus courte distance horizontale). À l'aide de données de tiges géoréférencées provenant de huit stations d'amont, nous avons comparé le biais relatif et l'erreur quadratique moyenne relative (par rapport à la moyenne vraie de chaque station), selon plusieurs méthodes différentes de choix des arbres de référence, pour le calcul de la différenciation de diamètre ( $\bar{D}$ ) et de l'indice de mélange des espèces ( $\bar{M}$ ). Les indices ont été définis en utilisant deux, trois et quatre voisins et les méthodes de sélection de l'arbre de référence qui ont été utilisées sont la sélection aléatoire d'un arbre dans une placette à rayon fixe (FI), la sélection aléatoire d'un arbre dans une placette à rayon variable (VA), la méthode de sélection par azimuth (AZ) et la sélection de l'arbre le plus proche (AP). En général, le biais relatif était inférieur à  $\pm 2,5\%$  pour  $\bar{D}$  et inférieur à  $\pm 10\%$  pour  $\bar{M}$  indépendamment de la méthode. La méthode FI a toujours eu le biais relatif le plus faible et la plus faible erreur quadratique moyenne relative. Les méthodes AP et AZ venaient au second rang en termes d'erreur quadratique moyenne relative pour  $\bar{D}$  et  $\bar{M}$  respectivement. La simplicité de ces deux méthodes pourrait l'emporter sur leur performance légèrement moins bonne. [Traduit par la Rédaction]

**Mots-clés :** mélange d'espèces, différenciation de diamètre, bassin hydrographique primaire, études de gestion de la densité, zones ripariennes.

## Introduction

According to Gadow et al. (2012), the term forest structure “usually refers to the way in which the attributes of trees are distributed within a forest ecosystem”. Tree growth and mortality shape the forest structure and can be influenced by multiple factors, ranging from human intervention to biotic and abiotic events and processes. But the forest structure itself modifies the driving factors of tree growth, mortality, and regeneration (Gadow et al. 2012). Silvicultural interventions (Franklin and Spies 1991) modify forest structure pursuing an enhancement in productivity and (or) the ecological functions of the forest at hand. A proper quantification of forest structure is therefore necessary for a better understanding of forest history and dynamics and for guiding the forest management.

Different forest structures are better suited for different management objectives (i.e., production, recreation, or conservation of ecological values). To obtain those forest structures that best

suit specific objectives, management systems try to mimic the disturbance regimes that generate the desired structures. Certain forest management systems and their associated structures are therefore preferred to others depending on the management goals. On one end of the gradient of forest management systems, even-aged forestry using clearcuts, shade-intolerant species, and short rotation periods is usually the selected method for production objectives. This management system tries to imitate high-intensity disturbances such as forest fires and has certain advantages for harvest operations; however, it does not completely imitate the disturbance regime that it is meant to mimic (García-Abril et al. 2013, p. 289). For example, trees and snags remaining after a fire produce conditions that facilitate or accelerate the regeneration, and a complete tree removal fails at producing such conditions. In addition, this system produces oversimplified forest structures with a highly homogeneous composition, similar to those obtained in the first stages of the ecological succession. Such simplification in structure results in a

Received 5 June 2016. Accepted 2 December 2016.

F. Mauro, Z. Haxtema, and H. Temesgen. Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, OR 97331-5706, USA.

**Corresponding author:** Hailemariam Temesgen (email: [hailemariam.temesgen@oregonstate.edu](mailto:hailemariam.temesgen@oregonstate.edu)).

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.elsevier.com/locate/permissions).

larger susceptibility to biotic and abiotic damages and in a lower resilience.

Management systems known as “close-to-nature” silviculture or continuous cover forestry advocate a closer adaptation to natural processes and lead to more complex uneven-aged forest structures, similar to those found in mature forests (Schütz et al. 2012). This type of management appeared in the nineteenth century as a response to the problems associated with intense even-aged management and had a revival in the last decades of the twentieth century. Uneven-aged management systems are better suited for scenarios where multiple objectives have to be considered. They result in larger diversity (structural and biological), are better suited for production of high-quality wood, and are able to outperform even-aged management systems in terms of economic balance (Howard and Temesgen 1997; Schütz 1997). A thorough review on the history of continuous cover forestry and comparisons between even-aged and uneven-aged management systems can be found in García-Abril et al. (2013, pp. 289–319).

Forest structure, as mentioned before, is a broad concept that is not restricted to a single attribute or to a single scale. Forest structure is a multidimensional and multiscale property. Different descriptors have been proposed to characterize its different facets at different scales. We will restrict the scope of this manuscript to quantifying forest structure at the stand level which, is usually linked to the concept of  $\alpha$ -diversity (i.e., fine-scale diversity) proposed by Whittaker (1960).

To quantify stand-level structure, standard aggregated measures of structure such as stand density, basal area, and volume can provide the necessary information to characterize stands for even-aged forest management. However, they do not suffice to fully characterize forest structure (Pretzsch 1997) and contain no information on the spatial arrangement of live trees within a stand (Pommerening 2002), which are key features for uneven-aged management systems. Standard aggregated measures and even more sophisticated presentations such as a stand table can tell us “what is out there” but not “how what is out there is arranged”. Zenner and Hibbs (2000) presented a hypothetical example where two different stands have identical tree densities, basal areas, and diameter distributions and yet have very different forest structures as a result of differing management regimes. Spatial functions, such as Ripley’s  $K$  (Ripley 1977) or the pair correlation function (Pommerening 2002), describe the spatial pattern of trees across a continuous range of intertree distances. However, they require knowledge of the coordinate positions of all trees within a plot, limiting their use in practical forest inventory applications. Furthermore, they can be difficult to explain to forest managers who may lack advanced statistical training.

In response to the need for a set of analytical tools to quantify forest structure at the stand level, nearest-neighbor indices (NNI) have received much attention from the forest research community in recent years. By focusing on the relationship between a given tree (the “reference” tree) and a fixed number of its nearest neighbors (generally three or four trees), NNI are capable of quantifying structural complexity at a very fine scale (e.g., Haxtema et al. 2012). Designed as a framework for characterizing structure without the need for expensive stem mapping (Gadow and Pogoda 2000), NNI have been promoted for their potential use in operational inventory programs (Pommerening 2006) and have been widely used in descriptive studies to quantify the structure of stem-mapped plots (e.g., Aguirre et al. 2003; Mason et al. 2007; Saunders and Wagner 2008; Haxtema et al. 2012). Neighborhood-based indices can be seen as fine-scale measurements of structural diversity, while conventional indices such as stand table measure diversity at the stand level and other diversity measures such as Shannon’s and Simpson’s indices require proportional data (i.e., number of individuals of each species and total number of individuals).

Tree size variability, species composition, and spatial distribution of trees have been typically considered as the main features that describe forest structure at the stand level (Pommerening 2002; Aguirre et al. 2003) and different NNI have been proposed to measure those features. A comprehensive review of available indices, not restricted to NNI, to analyze forest structure can be found in (Pommerening 2002). The two commonly used NNI are differentiation index  $D_t$  developed by Gadow (1993) and species mingling index  $M_t$  developed by Fuldner (1995). Below, we will briefly describe these indices and their uses.

The differentiation index  $D_t$  describes the degree of interspersion of trees of different sizes. While this index can describe the differentiation of any size-related variable (e.g., diameter at breast height (DBH), height, or volume), only diameter differentiation (the differentiation of DBH) will be considered here. Higher values imply higher differentiation (i.e., more fine-scale variability in size); lower values imply lower differentiation. For the  $i$ th reference tree and its  $j$ th nearest neighbor, diameter differentiation  $D_t$  is estimated as

$$(1) \quad D_t = \frac{1}{n} \sum_{j=1}^n \left[ 1 - \frac{\min(\text{dbh}_i, \text{dbh}_j)}{\max(\text{dbh}_i, \text{dbh}_j)} \right]$$

where  $\min(\text{dbh}_i, \text{dbh}_j)$  is the smaller DBH of the two trees and  $\max(\text{dbh}_i, \text{dbh}_j)$  is the larger.

The  $D_t$  index can be interpreted as follows (Pommerening 2002). (1)  $D_t$  index values from 0.0 to 0.3 are evidence of small differentiation. The tree with the smaller DBH is 70% or more of the neighboring tree’s DBH. (2)  $D_t$  index values from 0.3 to 0.5 are evidence of “average” differentiation. The tree with the smaller DBH is 50%–70% of the neighboring tree’s DBH. (3)  $D_t$  index values from 0.5 to 0.7 are evidence of large differentiation. The tree with the smaller DBH is 30%–50% of the neighboring tree’s DBH. (4)  $D_t$  index values from 0.7 to 1.0 are evidence of very large differentiation. The tree with the smallest DBH is less than 30% of the neighboring tree’s DBH.

The species mingling index  $M_t$  developed by Gadow and Hui (1999) provides a quantitative assessment of the degree of interspersion between trees of different species. Higher mingling values imply greater intermixing of species throughout a stand; lower values imply a tendency towards species-specific clumping and segregation. For a given reference tree, mingling is calculated as

$$(2) \quad M_t = \frac{1}{n} \sum_{j=1}^n v_j$$

where  $n$  is the number of neighbors considered and  $v_j$  is 0 if the  $j$ th neighbor is of the same species as the reference tree and 1 otherwise.

Previous research considered the distribution of the tree-level NNI values as the fundamental unit of information for forest structure characterization (Pommerening 2002). However, for comparative studies analyzing different edge-bias-corrections methods, the arithmetic mean of all individual NNI values has been considered as the parameter of interest, and the performance of different methods has been evaluated using their accuracy in estimating NNI population means (Pommerening and Stoyan 2006). We aim at providing references about the best method to select reference trees for NNI calculation by considering a broad range of alternative sampling protocols, sites, and measuring options. At the expense of losing interpretability in ecological terms, we will restrict our analysis to the estimation of mean NNI values because analyzing the effect of multiple factors in distributions of NNI would make the results intractable.

**Table 1.** Summary statistics for each stem map.

Site	$n_n$						Percentage of total density											N
	2		3		4		DF	WH	WR	GF	PY	RA	BM	BC	GC	PD	SX	
	$\bar{M}$	$\bar{D}$	$\bar{M}$	$\bar{D}$	$\bar{M}$	$\bar{D}$												
BL13	0.14	0.40	0.16	0.40	0.17	0.39	85.2				0.5		11.4			2.9		210
KM17	0.43	0.27	0.43	0.27	0.45	0.27	30.1	60.3	0.8			8.8						239
KM18	0.48	0.39	0.50	0.39	0.51	0.39	30.8	51.2	11.7			6.3						383
KM19	0.47	0.38	0.47	0.37	0.47	0.37	49.6	30.8	17.3			2.1		0.3				341
KM21	0.41	0.34	0.43	0.34	0.45	0.34	36.4	40.4	8.4			14.7						225
OM36	0.22	0.36	0.22	0.35	0.23	0.35	73.6	26.4										239
TH46	0.35	0.27	0.35	0.26	0.36	0.26	83.5	3.3	1.2	2.9		6.6	0.4	0.4	0.8	0.4	0.4	242
TH75	0.37	0.37	0.40	0.38	0.41	0.38	65.2	4.6	3.5			2.7	23.7	0.3				371

**Note:** DF, Douglas-fir; WH, western hemlock; RC, western red cedar; GF, grand fir; PY, Pacific yew; RA, red alder; BM, bigleaf maple; BC, black cottonwood; GC, golden chinquapin; PD, Pacific dogwood; SX, *Salix* sp.  $\bar{M}$  and  $\bar{D}$  are the average value for the species mingling and diameter differentiation indices, respectively. N is the total number of trees at each site.

Sampling all trees in relatively large plots of fixed radius is a common alternative in forest inventories; however, this system will be suboptimal in terms of efficiency (precision) for NNI estimation. When relatively large plots (e.g., permanent plots) are installed, some similarity between trees within a plot can be expected. Because of similarities between tree attributes within permanent plots, difficulties to cover the range of different structural variations that might appear in one stand can be expected. Thus, large plots will show some statistical inefficiency to characterize forest structure. In many cases, temporary sample points are used if the cost of sample point establishment is small relative to the cost of travel. Neighborhood-based indices were specifically designed to be easy to measure in the field, so it may be more statistically efficient to distribute temporary sampling points and reference trees throughout a stand rather than selecting them in permanent plots. Even though this alternative seems to be the best for NNI estimation, minimal guidance exists for practitioners seeking to integrate these indices into sampling protocols, and little is known about the best method of selecting reference trees when commonly used systematic designs, where sampling points are located on the nodes of a regular grid, are used. Instead of substituting standard measures of structure like basal area, volume, or dominant height, NNI are meant to complement those measures. Similarly, it is important to note that the sampling for NNI is not supposed to substitute standard sampling procedures. It aims at providing specific information about NNI in a fast and reliable way. Aiming at providing references for future inventories where NNI are measured, we consider four different methods for selection of reference trees.

Two alternative methods based on plots can be employed to select reference trees. These methods use either fixed-radius (FI) or variable-radius (VA) plots to select a relatively small group of candidate trees. Selected trees are listed and then a number of candidates, ranging from one to the total number of trees selected in the plot, are randomly drawn and considered as reference trees.

Gadow and Pogoda (2000) suggested that reference trees be selected with what we refer to as the “nearest-tree” (NT) selection method, whereby the nearest tree to a given sample point is selected as the reference tree. Under this method, trees are selected with probability proportional to the area in which they are the nearest tree. Thus, selection probabilities will be unequal in all but the most uniformly spaced populations. While Iles (2009) described an unbiased method for estimating inclusion areas from field measurements, the added effort required to unbiasedly implement this method may tempt the sampler to avoid it due to the high effort required for its implementation. Nonetheless, if the bias incurred by treating a nearest-tree sample as an equal-probability sample is demonstrated to be small in a given forest type, the NT selection method may be acceptable.

A method sometimes used in forest inventory is to subsample from a fixed plot by selecting the first tree whose center is encountered in a clockwise sweep from north. This has been termed the “azimuth” (AZ) method by Iles (1979, p. 23). As with the NT selection method, inclusion areas will not be equal in most cases, and a satisfactory method has not yet been devised for measuring or estimating inclusion areas in the field. However, it may be reasonable to ignore the bias incurred from this sampling method if it is demonstrated to be small.

### Objective

The objective of this study is to compare the statistical performance of the FI, VA, NT, and AZ methods of reference tree selection for estimation of different NNI in headwater riparian forests of western Oregon. We selected diameter differentiation index (Pommerening 2002, after Fuldner 1995) and species mingling index (Aguirre et al. 2003, after Gadow 1993) as descriptors of tree size and species composition variability.

### Materials and methods

#### Description of the stem maps

Selected sampling methods were evaluated using data from eight sites located in stem-mapped research plots installed as part of the United States Bureau of Land Management Density Management Study. Methods used to collect data are described in detail in Marquardt et al. (2010) and Eskelson et al. (2011) and briefly summarized here. Plots were located in the Coast Range and foothills of the Cascade Range of western Oregon. Each 0.518 ha plot (72 m × 72 m) was located such that a headwater stream flowed through the approximate centerline of the plot. The coordinate position, DBH, and species of all trees greater than 6.9 m tall or 7.6 cm in DBH were recorded.

All stem maps were composed of second-growth forest. Species composition was conifer-dominated, with either Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) or western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) comprising the majority of total stem density on any given site. Western red cedar (*Thuja plicata* Donn ex D. Don) was present on six stem maps and some stem maps contained Pacific yew (*Taxus brevifolia* Nutt.) or grand fir (*Abies grandis* (Douglas ex D. Don) Lindl.). The hardwood species red alder (*Alnus rubra* Bong.), bigleaf maple (*Acer macrophyllum* Pursh), and black cottonwood (*Populus balsamifera* ssp. *trichocarpa* (Torr. & A. Gray ex Hook.) Brayshaw) were a minor component of some stem maps (Marquardt et al. 2012). Species composition and average values for the species mingling and diameter differentiation indices computed with two, three, and four neighbors in each site are provided in Table 1.

**Table 2.** *t* test *p* values for each index and combination of stem map, method, number of neighbors, and number of trees per plot.

Index:		Species mingling									Diameter differentiation								
$n_i$ :		3			6			9			3			6			9		
Site	$n_i$ :	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4	2	3	4
BL13	NT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.81	0.55	0.03	0.81	0.55	0.03	0.81	0.55	0.03
	FI	0.31	0.77	0.48	0.74	0.97	0.77	0.48	0.52	0.43	0.82	0.52	0.60	0.50	0.64	0.99	0.48	0.45	0.93
	AZ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.02	0.00	0.04	0.02	0.00	0.04	0.02	0.00
	VA	0.14	0.30	0.01	0.29	0.14	0.00	0.72	0.07	0.00	0.33	0.40	0.28	0.28	0.15	0.38	0.21	0.30	0.34
KM17	NT	0.43	0.05	0.06	0.43	0.05	0.06	0.43	0.05	0.06	0.20	0.10	0.09	0.20	0.10	0.09	0.20	0.10	0.09
	FI	0.07	0.04	0.01	0.26	0.22	0.10	0.31	0.24	0.29	0.77	0.86	0.94	0.33	0.61	0.72	0.62	0.96	0.91
	AZ	0.03	0.16	0.45	0.03	0.16	0.45	0.03	0.16	0.45	0.78	0.04	0.00	0.78	0.04	0.00	0.78	0.04	0.00
	VA	0.47	0.35	0.46	0.54	0.26	0.24	0.00	0.01	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KM18	NT	0.04	0.05	0.09	0.04	0.05	0.09	0.04	0.05	0.09	0.01	0.06	0.36	0.01	0.06	0.36	0.01	0.06	0.36
	FI	0.53	0.11	0.50	0.77	0.71	0.70	0.39	0.12	0.21	0.36	0.38	0.58	0.67	0.40	0.33	0.76	0.97	0.93
	AZ	0.11	0.30	0.40	0.11	0.30	0.40	0.11	0.30	0.40	0.90	0.38	0.53	0.90	0.38	0.53	0.90	0.38	0.53
	VA	0.64	0.89	0.65	0.76	0.96	0.40	0.88	0.96	0.31	0.25	0.13	0.03	0.37	0.13	0.04	0.27	0.30	0.48
KM19	NT	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00
	FI	0.00	0.00	0.00	0.33	0.13	0.03	0.00	0.00	0.00	0.18	0.23	0.16	0.92	0.48	0.39	0.50	0.92	0.98
	AZ	0.00	0.06	0.29	0.00	0.06	0.29	0.00	0.06	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	VA	0.02	0.02	0.05	0.00	0.02	0.11	0.00	0.00	0.01	0.08	0.11	0.28	0.06	0.02	0.09	0.11	0.04	0.08
KM21	NT	0.09	0.02	0.00	0.09	0.02	0.00	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	FI	0.29	0.61	0.82	0.86	0.04	0.01	0.32	0.40	0.70	0.90	0.52	0.67	0.97	0.62	0.12	0.93	0.41	0.77
	AZ	0.95	0.46	0.30	0.95	0.46	0.30	0.95	0.46	0.30	0.36	0.14	0.56	0.36	0.14	0.56	0.36	0.14	0.56
	VA	0.09	0.76	0.97	0.04	0.84	0.55	0.01	0.41	0.18	0.03	0.55	0.48	0.00	0.00	0.00	0.00	0.00	0.00
OM36	NT	0.00	0.25	0.90	0.00	0.25	0.90	0.00	0.25	0.90	0.75	0.88	0.67	0.75	0.88	0.67	0.75	0.88	0.67
	FI	0.55	0.48	0.28	0.29	0.51	0.33	0.83	0.94	0.66	0.38	0.24	0.09	0.06	0.15	0.10	0.91	0.99	0.68
	AZ	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.53	0.13	0.61	0.53	0.13	0.61	0.53	0.13	0.61
	VA	0.59	0.94	0.79	0.47	0.33	0.19	0.96	0.17	0.07	0.07	0.04	0.16	0.97	0.83	0.61	0.03	0.01	0.00
TH46	NT	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	FI	0.08	0.19	0.14	0.86	0.90	0.61	0.62	0.97	0.67	0.19	0.18	0.00	0.17	0.15	0.06	0.01	0.01	0.01
	AZ	0.82	0.90	0.39	0.82	0.90	0.39	0.82	0.90	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	VA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	0.33	0.08	0.01	0.02	0.00	0.01	0.02
TH75	NT	0.37	0.88	0.77	0.37	0.88	0.77	0.37	0.88	0.77	0.81	0.29	0.01	0.81	0.29	0.01	0.81	0.29	0.01
	FI	0.84	0.57	0.79	0.12	0.17	0.22	0.24	0.29	0.23	0.90	0.57	0.67	0.26	0.08	0.18	0.75	0.15	0.10
	AZ	0.50	0.86	0.86	0.50	0.86	0.86	0.50	0.86	0.86	0.29	0.06	0.02	0.29	0.06	0.02	0.29	0.06	0.02
	VA	0.09	0.00	0.00	0.04	0.00	0.00	0.33	0.06	0.00	0.42	0.20	0.04	0.03	0.03	0.00	0.56	0.25	0.09

Note: Highlighted cells indicate significant differences after performing the Bonferroni correction.

**Parameter of interest**

The parameters of interest were the average values of the diameter differentiation  $\bar{D}$  and species mingling index  $\bar{M}$  (Table 1) for each stem map. Because trees near a plot edge may have immediate neighbors that are outside the plot, bias can result from the calculation of index values when only trees inside the plot are considered (Pommerening 2006). Considering Pommerening and Stoyan (2006), edge effects were accounted for by using a toroidal edge correction (Ripley 1979) on each site.

Both, the diameter differentiation index and the species mingling index depend on the number of neighbors ( $n_n$ ) considered to compute each index. We considered three different numbers of neighbors,  $n_n = 2, 3$ , and 4, as these are sensible values to use in operational forestry applications. If  $n_n$  changes, the index  $y_i$ , associated with each tree ( $t$ ), and the site averages ( $\bar{Y}$ ) change too. To assess the importance of the change in  $\bar{Y}$ , we tested by means of ANOVA whether or not increasing the number of neighbors from two to four resulted in different values for the average values of both indices. We accounted for the effect of multiple comparisons and performed a Bonferroni correction (Dunn 1961) by dividing the significance level  $\alpha = 0.05$  by the number of comparisons performed.

**Reference tree selection methods**

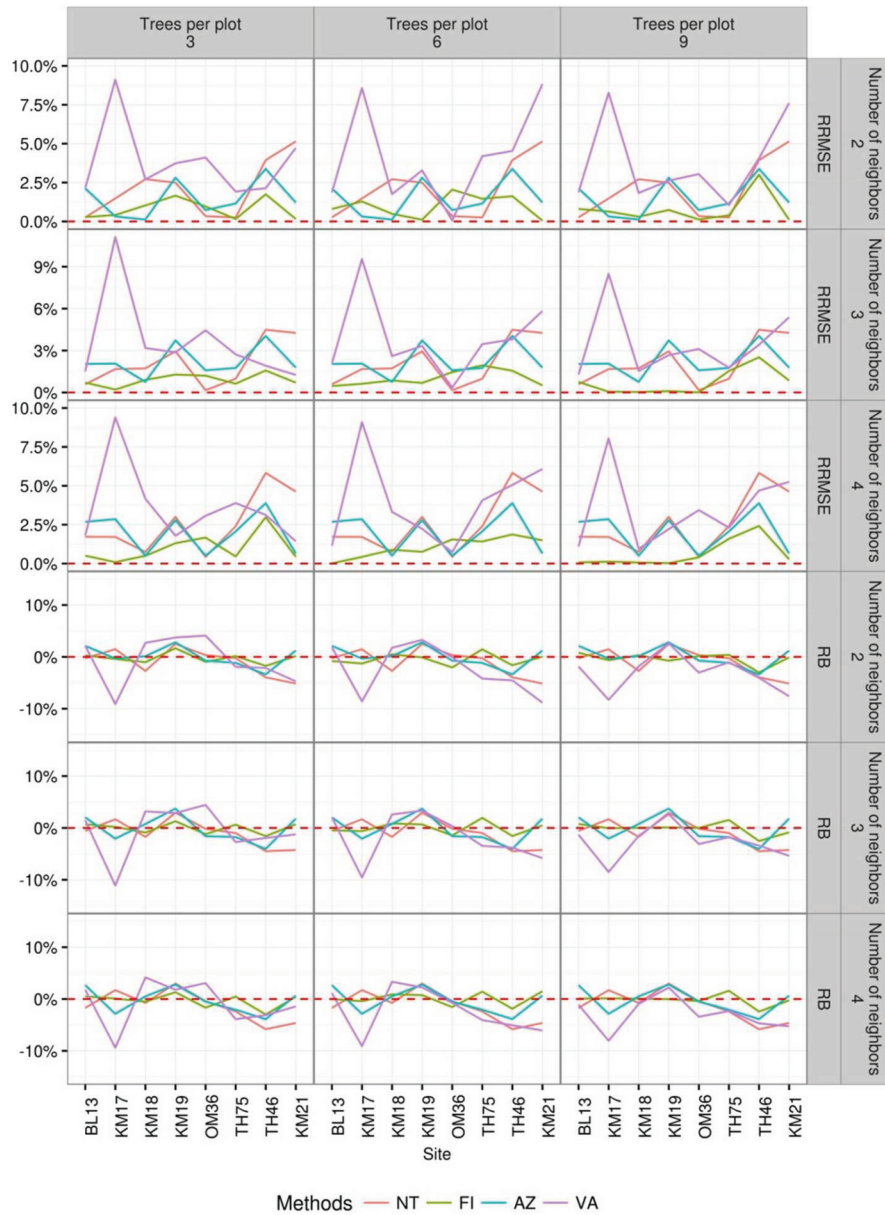
The four methods compared in this study result in different reference tree selection probabilities. Selection probabilities are fundamental to derive estimators for different sampling protocols. An unbiased estimator of the population mean for general sample designs of fixed size  $n$  (Horvitz and Thompson 1952) is

given by  $\hat{Y}_{HT} = \frac{\sum_{t=1}^n w_t y_t}{N}$ , where subindex HT stands for Horvitz and Thompson and  $w_t$  is the reciprocal of the probability of selecting tree  $t$ . For equal probability sample designs, this estimator becomes the sample mean. To make the methods comparable, and to take into account the particular characteristics of each selection method, several factors were considered and, when possible, approximated sampling weights were obtained for each tree.

First, to make the methods comparable, it was necessary to keep the number of measurements equal between selection methods. Since the NT and AZ methods only select one tree per plot, only one tree was selected from the FI and VA plot methods. Then, the selection protocol resulting for each of these methods can be described as follows. For the FI method, a tree was subsampled from a fixed plot by drawing a random number from 1 to  $F$  where  $F$  is the number of trees in fixed the plot. When using the VA method to capture trees, a tree was subsampled by randomly selecting one of  $V$  trees captured in the variable-radius plot with probability proportional to its basal area. Weights were then  $w_{t,FI} = F$  and  $w_{t,VA} = V_t \frac{BAF}{BA_t}$ , where the term BAF is the basal area factor used in the variable plot and  $BA_t$  is the basal area of the selected tree. It is important to note that both weights are approximations to the real weight of the trees for the described sampling protocols. For the NT and AZ methods, it was impossible to obtain approximated weights for the selected reference trees and they were all equally weighted with a weight of 1 in subsequent estimations.

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by Oregon State University on 10/04/17 For personal use only.

Fig. 1. Relative root mean squared error (RRMSE) and relative bias (RB) for estimation of the overall species mingling index.



**Simulation of alternatives methods for reference tree selection**

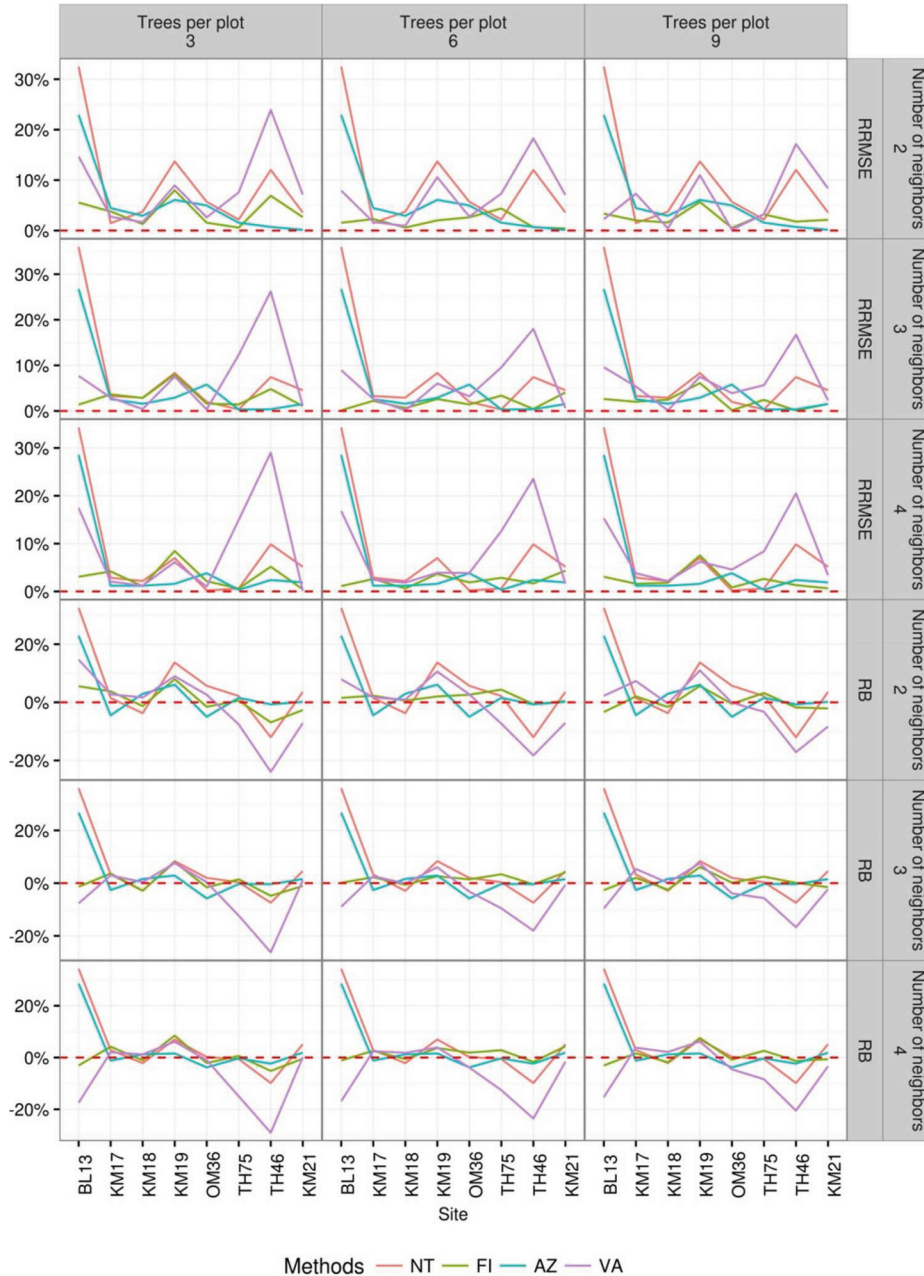
A simulation algorithm was written in R (R Core Team 2016) to simulate the field measurement performed using each of the described methods. Four sampling points were simulated at each of 500 repetitions in each site. Since systematic sampling is far more common than random sampling in most forestry sampling applications, this method of sample point placement was simulated. The first point was randomly located within the lower left quadrant of the stem map using the pseudo-random number generator function in R. Subsequent points were placed on a 36 m × 36 m grid so that every quadrant of the stand map had a plot included. The fixed-plot radius and variable-plot BAFs were scaled to capture an average of three, six, and nine trees ( $n_t$ ). Finally, the four methods for selection of reference trees were applied in each sampling point. This process was iterated 500 times, as we empirically observed that rate of change for the accuracy statistics described in next section was negligible after 500 replications.

Since the probability of selecting a tree as a reference varies with the number of trees captured at a sample point, it is necessary to factor the number of trees per sample point into the weighted sample estimate produced by each selection process (Iles 2003, p. 562). For each sampling method, each neighborhood-based index was estimated at iteration  $k$  as a weighted average of the sampled values in an attempt to obtain estimators resembling the Horvitz and Thompson (1952) estimator:

$$(3) \quad \hat{Y}_{k,j,m,n_t,n_n} = \frac{\sum_{t=1}^4 w_{t,k,j,m,n_t} y_{t,k,j,n_n}}{\sum_{t=1}^4 w_{t,k,j,m,n_t}}$$

Note that for each method ( $m$ ) and site ( $j$ ),  $\hat{Y}_{k,j,m,n_t,n_n}$  always depends on  $n_n$  and the iteration  $k$ . For the FI and VA methods, the selected reference trees depend on  $n_t$  but not for the NT and AZ methods.

Fig. 2. Relative root mean squared error (RRMSE) and relative bias (RB) for estimation of the overall diameter differentiation index.



For generality, subindex  $n_t$  was left on  $\hat{Y}_{k,j,m,n_t,n_t}$ . Four reference trees are selected for each site, iteration,  $n_t$ , and method. Trees are indicated with subindex  $t$ . The approximated weight for the selected  $t$ th tree selected at iteration  $k$  in site  $j$  by the  $m$ th method depends on  $n_t$ , which results in reference tree weights indexed as  $w_{t,k,j,m,n_t}$ . Weights  $w_{t,k,j,m,n_t}$  were assigned as explained in the Reference tree selection methods section. Finally, the index value for the  $t$ th tree selected at iteration  $k$  in site  $j$  by the  $m$ th method depends on  $n_t$ . Therefore, reference tree values for the neighborhood-based indices are indicated as  $y_{t,k,j,n_t}$ .

The sampling methods were evaluated for estimation of species-specific index values for Douglas-fir, western hemlock, western red cedar, and red alder, the four species that occurred on at least six stem maps. Species-specific estimates for both indices were calculated by considering only reference trees of each species in eq. 3.

**Accuracy assessment**

The sampling methods were compared using relative bias (RB) and relative root mean square error (RRMSE) (e.g., Temesgen 2003; Temesgen et al. 2011; Poudel et al. 2015). For each site, sampling method, number of neighbors for the indices definition, and average number of trees in the fixed- and variable-radius plots, RB was computed as

$$(4) \quad RB_{j,m,n_t,n_t}(\%) = 100 \frac{\sum_{k=1}^{500} [\hat{Y}_{k,j,m,n_t,n_t} - \bar{Y}_{j,n_t}]}{500 \bar{Y}_{j,n_t}}$$

where  $\bar{Y}_{j,n_t}$  is the true arithmetic mean index value of the population for site  $j$  when neighborhood-based indices are calculated based on  $n_t$  neighbors. We tested by means of  $t$  tests whether or

**Table 3.** ANOVA for relative bias (RB) and relative root mean squared error (RRMSE) for estimation of the overall diameter differentiation and species mingling indices.

Factor	RB diameter differentiation					RRMSE diameter differentiation				
	df	Sum of squares	Mean square	F	Pr(F)	df	Sum of squares	Mean square	F	Pr(F)
Factor $n_n$	2	3.0E-04	1.5E-04	1.8E-01	8.3E-01	2	2.0E-04	1.0E-04	3.2E-01	7.3E-01
Factor $n_t$	2	1.6E-03	7.9E-04	9.6E-01	3.8E-01	2	1.6E-04	7.8E-05	2.5E-01	7.8E-01
Factor <i>method</i> ( $m$ )	3	1.6E-02	5.4E-03	6.6E+00	2.7E-04*	3	3.1E-02	1.0E-02	3.2E+01	2E-16*
Interaction $n_n \times n_t$	4	6.0E-05	1.5E-05	1.9E-02	1.0E+00	4	1.0E-05	2.0E-06	8.0E-03	1.0E+00
Interaction $n_n \times m$	6	4.7E-04	7.8E-05	9.5E-02	1.0E+00	6	9.4E-04	1.6E-04	5.0E-01	8.1E-01
Interaction $n_t \times m$	6	4.6E-03	7.7E-04	9.4E-01	4.7E-01	6	1.8E-04	3.0E-05	9.4E-02	1.0E+00
Interaction $n_t \times n_t \times m$	12	1.1E-04	9.0E-06	1.1E-02	1.0E+00	12	2.0E-05	2.0E-06	6.0E-03	1.0E+00
Residuals	252	2.1E-01	8.2E-04			252	7.9E-02	3.2E-04		

Factor	RB species mingling					RRMSE species mingling				
	df	Sum of squares	Mean square	F	Pr(F)	df	Sum of squares	Mean square	F	Pr(F)
Factor $n_n$	2	9.4E-03	4.7E-03	4.9E-01	6.1E-01	2	2.0E-03	9.9E-04	1.6E-01	8.6E-01
Factor $n_t$	2	6.0E-04	3.0E-04	3.2E-02	9.7E-01	2	2.4E-03	1.2E-03	1.9E-01	8.3E-01
Factor <i>method</i> ( $m$ )	3	3.1E-01	1.0E-01	1.1E+01	1.1E-06*	3	1.5E-01	4.8E-02	7.6E+00	6.9E-05*
Interaction $n_n \times n_t$	4	4.0E-04	1.0E-04	1.0E-02	1.0E+00	4	3.0E-04	9.0E-05	1.3E-02	1.0E+00
Interaction $n_n \times m$	6	2.1E-02	3.5E-03	3.7E-01	9.0E-01	6	5.8E-03	9.7E-04	1.5E-01	9.9E-01
Interaction $n_t \times m$	6	2.7E-03	4.6E-04	4.8E-02	1.0E+00	6	2.7E-03	4.4E-04	7.0E-02	1.0E+00
Interaction $n_t \times n_t \times m$	12	5.0E-04	4.0E-05	4.0E-03	1.0E+00	12	6.0E-04	5.0E-05	8.0E-03	1.0E+00
Residuals	252	2.4E+00	9.5E-03			252	1.6E+00	6.4E-03		

**Note:** The factor method is indicated as  $m$  and  $n_n$  and  $n_t$  indicate the number of neighbors used to compute the indexes and the average number of trees in the fixed- and variable-radius plots. Factors and interactions that were significant in the ANOVA, after performing a multiple tests Bonferroni correction, are indicated with an asterisk. The significance level was 0.05.

**Table 4.** ANOVA for relative bias (RB) and relative root mean squared error (RRMSE) when estimating diameter differentiation and species mingling indices by species and method.

Factor	RB species mingling					RRMSE species mingling				
	df	Sum of squares	Mean square	F	Pr(F)	df	Sum of squares	Mean square	F	Pr(F)
Factor $n_n$	2	7.0E-03	3.6E-03	6.2E-01	5.4E-01	2	1.6E-02	8.0E-03	2.4E+00	8.9E-02
Factor $n_t$	2	1.5E-02	7.7E-03	1.3E+00	2.7E-01	2	7.4E-03	3.7E-03	1.1E+00	3.3E-01
Factor <i>method</i> ( $m$ )	3	3.4E-02	1.1E-02	1.9E+00	1.2E-01	3	6.6E-01	2.2E-01	6.6E+01	2.0E-16*
Factor <i>species</i> ( $sp$ )	3	2.7E-01	9.0E-02	1.5E+01	1.2E-09*	3	3.0E-01	9.8E-02	3.0E+01	2.0E-16*
Interaction $n_n \times n_t$	4	0.0E+00	1.1E-04	1.8E-02	1.0E+00	4	2.0E-04	6.0E-05	1.9E-02	1.0E+00
Interaction $n_n \times me$	6	4.0E-03	6.5E-04	1.1E-01	1.0E+00	6	3.5E-03	5.8E-04	1.7E-01	9.8E-01
Interaction $n_n \times me$	6	1.8E-02	3.0E-03	5.1E-01	8.1E-01	6	1.6E-02	2.7E-03	8.1E-01	5.6E-01
Interaction $n_n \times sp$	6	5.0E-02	8.3E-03	1.4E+00	2.1E-01	6	3.3E-02	5.5E-03	1.7E+00	1.3E-01
Interaction $n_t \times sp$	6	2.3E-02	3.8E-03	6.5E-01	6.9E-01	6	1.9E-02	3.2E-03	9.7E-01	4.5E-01
Interaction $me \times sp$	9	1.5E+00	1.6E-01	2.7E+01	2.0E-16*	9	2.0E-01	2.2E-02	6.7E+00	2.8E-09*
Interaction $n_n \times n_t \times m$	12	1.0E-03	8.0E-05	1.3E-02	1.0E+00	12	8.0E-04	7.0E-05	2.1E-02	1.0E+00
Interaction $n_n \times n_t \times sp$	12	1.0E-03	1.0E-04	1.7E-02	1.0E+00	12	7.0E-04	6.0E-05	1.7E-02	1.0E+00
Interaction $n_n \times me \times sp$	18	6.0E-02	3.4E-03	5.7E-01	9.2E-01	18	5.2E-02	2.9E-03	8.8E-01	6.0E-01
Interaction $n_t \times me \times sp$	18	1.1E-01	6.1E-03	1.0E+00	4.2E-01	18	6.5E-02	3.6E-03	1.1E+00	3.5E-01
Interaction $n_t \times n_t \times me \times sp$	36	3.0E-03	9.0E-05	1.6E-02	1.0E+00	36	3.7E-03	1.0E-04	3.1E-02	1.0E+00
Residuals	822	4.8E+00	5.9E-03			822	2.7E+00	3.3E-03		

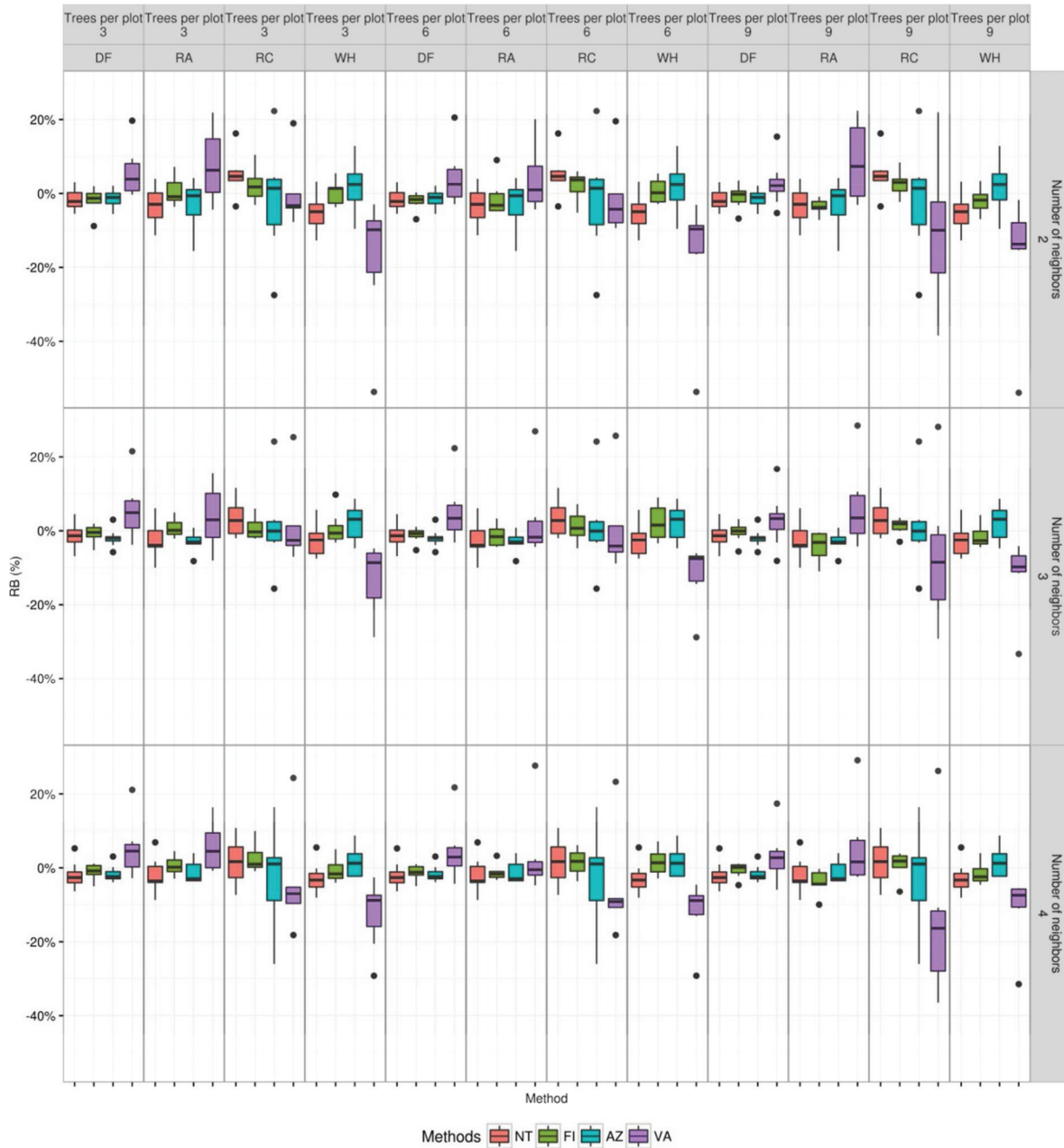
  

Factor	RB diameter differentiation					RRMSE diameter differentiation				
	df	Sum of squares	Mean square	F	Pr(F)	df	Sum of squares	Mean square	F	Pr(F)
Factor $n_n$	2	1.7E-02	8.7E-03	7.4E-01	4.8E-01	2	7.0E-03	3.6E-03	6.2E-01	5.4E-01
Factor $n_t$	2	3.0E-03	1.6E-03	1.4E-01	8.7E-01	2	1.5E-02	7.7E-03	1.3E+00	2.7E-01
Factor <i>method</i> ( $m$ )	3	8.2E-01	2.7E-01	2.3E+01	1.9E-14*	3	3.4E-02	1.1E-02	1.9E+00	1.2E-01
Factor <i>species</i> ( $sp$ )	3	5.0E-01	1.7E-01	1.4E+01	5.6E-09*	3	2.7E-01	9.0E-02	1.5E+01	1.2E-09*
Interaction $n_n \times n_t$	4	0.0E+00	1.0E-05	1.0E-03	1.0E+00	4	0.0E+00	1.1E-04	1.8E-02	1.0E+00
Interaction $n_n \times m$	6	6.5E-02	1.1E-02	9.2E-01	4.8E-01	6	4.0E-03	6.5E-04	1.1E-01	1.0E+00
Interaction $n_t \times m$	6	5.0E-03	7.5E-04	6.4E-02	1.0E+00	6	1.8E-02	3.0E-03	5.1E-01	8.1E-01
Interaction $n_n \times sp$	6	4.9E-02	8.1E-03	6.9E-01	6.6E-01	6	5.0E-02	3.8E-03	1.4E+00	2.1E-01
Interaction $n_t \times sp$	6	9.0E-03	1.5E-03	1.2E-01	9.9E-01	6	2.3E-02	3.8E-03	6.5E-01	6.9E-01
Interaction $me \times sp$	9	2.2E-01	2.4E-02	2.1E+00	3.1E-02	9	1.5E+00	1.6E-01	2.7E+01	2.0E-16*
Interaction $n_n \times n_t \times m$	12	1.0E-03	6.0E-05	5.0E-03	1.0E+00	12	1.0E-03	8.0E-05	1.3E-02	1.0E+00
Interaction $n_n \times n_t \times sp$	12	2.0E-03	2.1E-04	1.7E-02	1.0E+00	12	1.0E-03	1.0E-04	1.7E-02	1.0E+00
Interaction $n_n \times m \times sp$	18	7.5E-02	4.2E-03	3.5E-01	9.9E-01	18	6.0E-02	3.4E-03	5.7E-01	9.2E-01
Interaction $n_t \times m \times sp$	18	9.1E-02	5.0E-03	4.3E-01	9.8E-01	18	1.1E-01	6.1E-03	1.0E+00	4.2E-01
Interaction $n_t \times n_t \times m \times sp$	36	7.0E-03	1.9E-04	1.6E-02	1.0E+00	36	3.0E-03	9.0E-05	1.6E-02	1.0E+00
Residuals	822	9.7E+00	1.2E-02			822	4.8E+00	5.9E-03		

**Note:** The factor method is indicated as  $m$ , the factor species is denoted as  $sp$ , and  $n_n$  and  $n_t$  indicate the number of neighbors used to compute the indexes and the average number of trees in the fixed- and variable-radius plots. Factors and interactions that were significant in the ANOVA, after performing a multiple tests Bonferroni correction, are indicated with an asterisk. The significance level was 0.05.

Can. J. For. Res. Downloaded from www.nrcresearchpress.com by Oregon State University on 10/04/17  
For personal use only.

**Fig. 3.** Variability of relative bias (RB) across sites by species and methods for the species mingling index. Species: DF, Douglas-fir; RA, red alder; RC, western red cedar; WH, western hemlock.



not for each combination of: site, method,  $n_n$ , and  $n_t$ , the average of the difference  $\hat{Y}_{k,j,m,n_t,n_n} - \bar{Y}$  (i.e., the absolute bias) was significantly different from zero.

The RRMSE was computed as

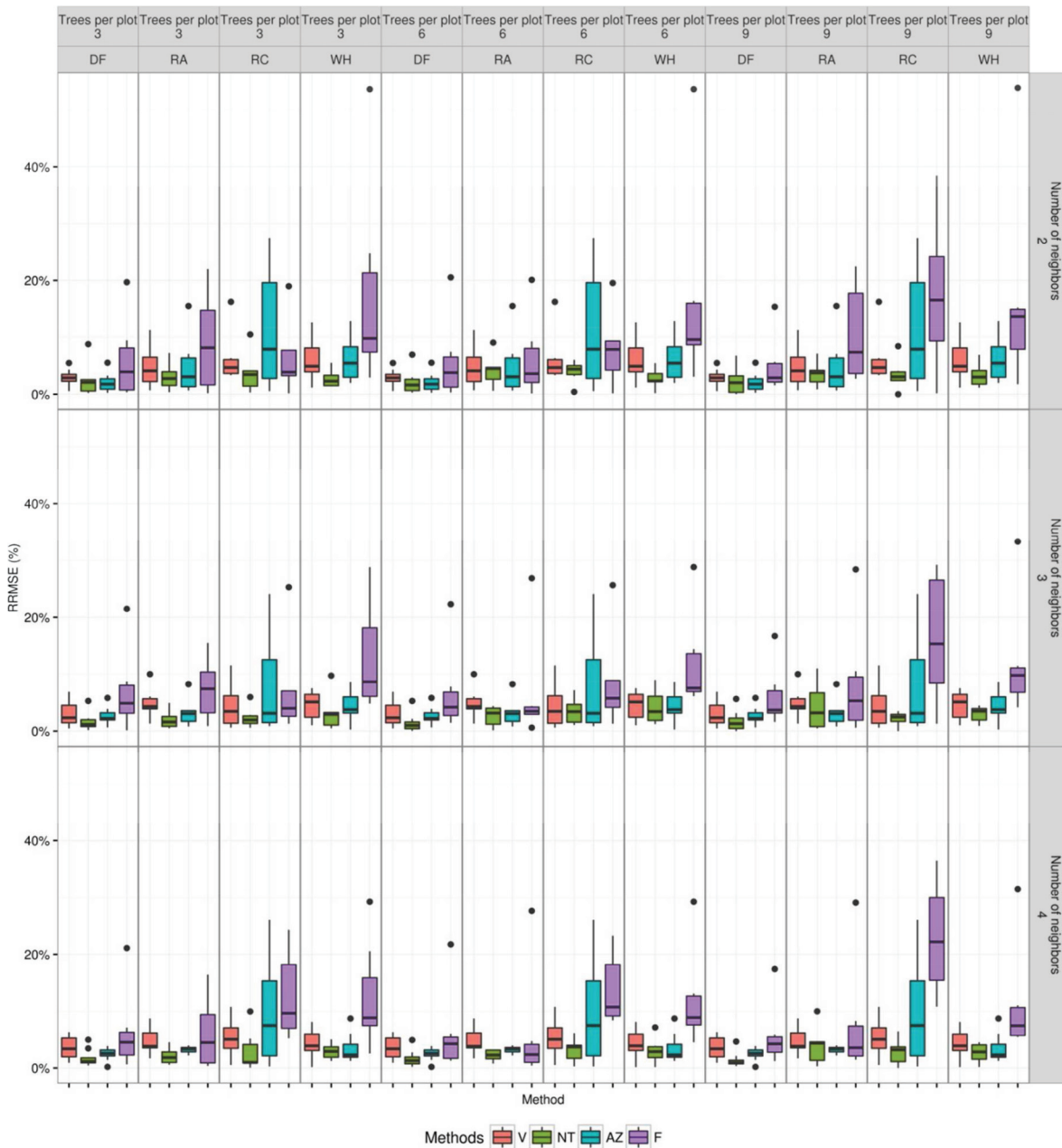
$$(5) \quad RRMSE_{j,m,n_t,n_n}(\%) = \frac{100}{\bar{Y}} \sqrt{\frac{\sum_{k=1}^{500} [\hat{Y}_{k,j,m,n_t,n_n} - \bar{Y}]^2}{500}}$$

The RB and the RRMSE for each stem map and species were computed similarly, but only considering those iterations where at least one reference tree of the considered species was selected. Then the denominator and the summation limits in eqs. 4 and 5 were changed to  $N_{iter,sp}$ , where  $N_{iter,sp}$  is the number of iterations

for which at least one reference tree belonged to the species of interest. We did not test for each species whether or not for each combination of stem maps, method,  $n_n$ , and  $n_t$ , RB was significantly different from zero because the species compositions of the stem maps are relatively uneven. This would result in a very different number of observations for each species and therefore different values for the power of the  $t$  test, making interpretations of further comparisons difficult. Finally, factorial ANOVA was used to assess if significant differences were present for RB and RRMSE across methods and across methods and species when changing both  $n_n$  and  $n_t$ . We accounted for the effect of multiple hypothesis testing by using a Bonferroni correction, and when factors were significant, we performed a post hoc analysis by computing Tukey's honest significant differences.



**Fig. 4.** Variability of relative root mean squared error (RRMSE) across sites by species and methods for the species mingling index. Species: DF, Douglas-fir; RA, red alder; RC, western red cedar; WH, western hemlock.



**Results and discussion**

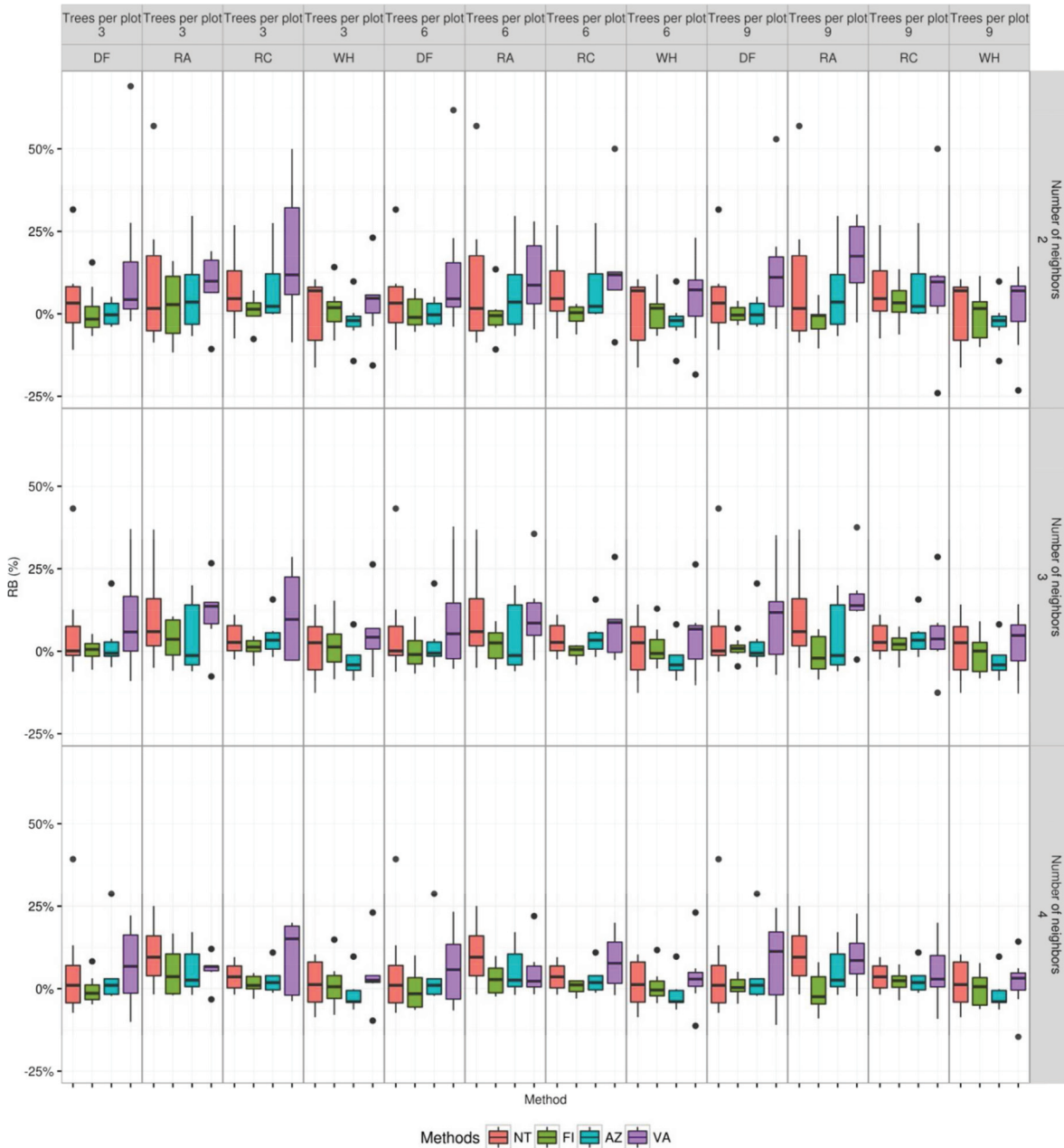
The statistical scope of inference of this study is limited to forests adjoining headwater streams within the Density Management Study sites. However, the general principles underlying the sampling methods that we examined may be applicable to a wide array of forest types.

Reducing the number of neighbors from four to two in the definition of both indices had no significant effect on their average values ( $p$  value = 0.99 for the diameter differentiation index and  $p$  value = 0.94 for the species mingling index). This result has important practical consequences for future inventories. By reducing the number of trees, the mean value of the studied indices is not affected, but less neighbors need to be measured for each reference tree. This result partially supports previous research from Sterba (2008) who measured one nearest neighbor for all trees selected with a variable-plot sample in the Austrian Alps and

concluded that estimates of the diameter differentiation index could be obtained with “only little additional field work”. Finally, stand-level distribution of NNI is more meaningful than the means studied here. The number of neighbors to use for the index computation was not a significant factor for the average NNI in a stand. Leaving this factor out of the analysis in further studies might be a good compromise solution to approach the more complex problem of estimating stand-level distributions of NNI.

For the fixed-radius plot method, the  $t$  test for the average of the differences  $\hat{Y}_{k,j,m,n_t,n_n} - \bar{Y}$  for the species mingling index was significantly different from zero in only three of the 72 possible combinations of stem map,  $n_n$ , and  $n_t$  (Table 2). For the diameter differentiation index, the average difference was, in all cases, not significantly different from zero. This result supports the idea that the approximated weights used for this method are not perfect but effective to prevent biased estimates of the two consid-

**Fig. 5.** Variability of relative bias (RB) across sites by species and methods for the diameter differentiation index. Species: DF, Douglas-fir; RA, red alder; RC, western red cedar; WH, western hemlock.

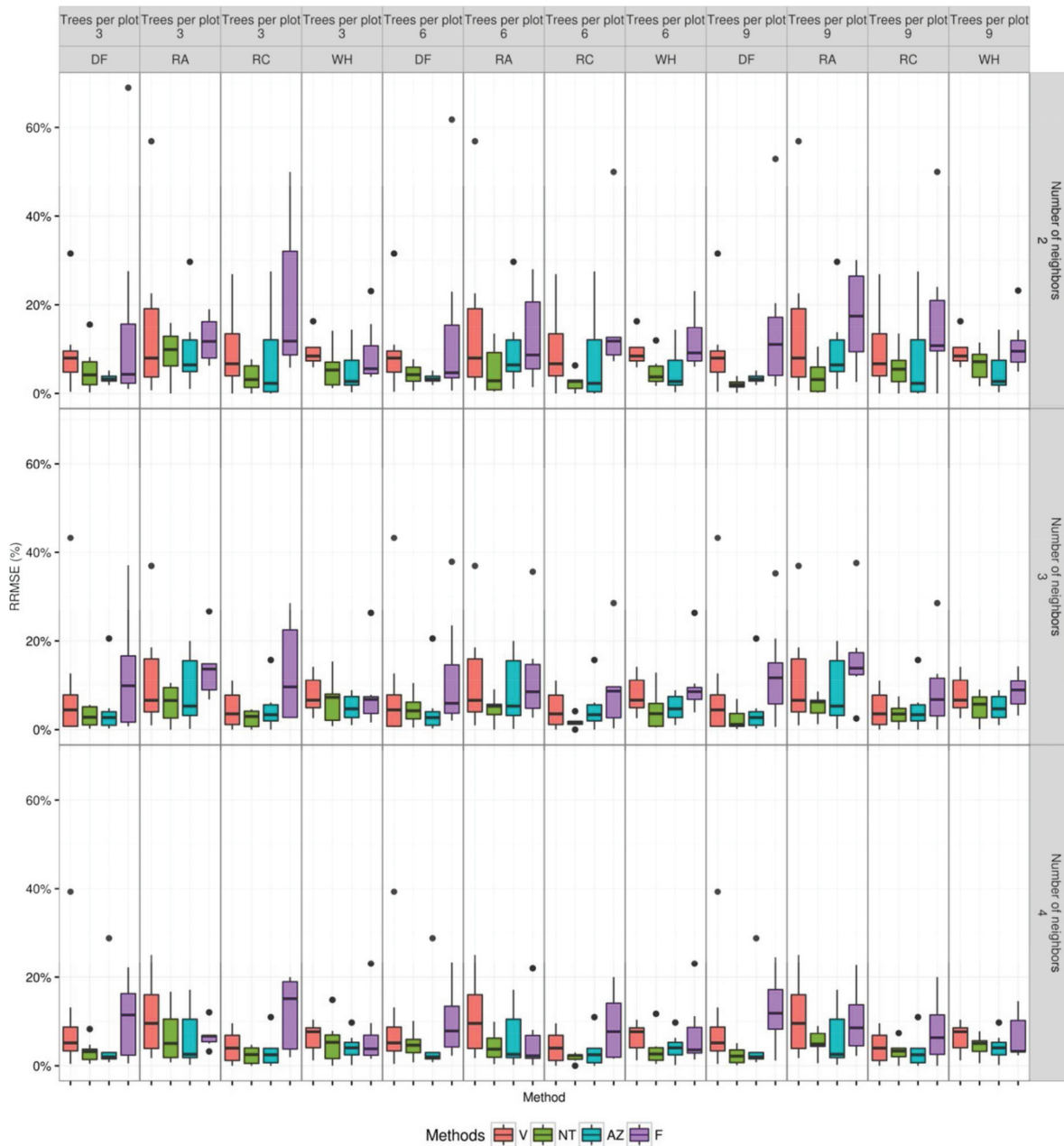


ered indices. For the variable-radius plot method, the number of combinations of stem maps,  $n_n$ , and  $n_t$  for which a significant bias was found was six and 14 out of 72 possible combinations for the species mingling and diameter differentiation index, respectively. For the nearest-tree and the AZ method, the number of combinations of stem maps,  $n_n$ , and  $n_t$  for which a significant bias was found was 16 and 16 for the species mingling and eight and six for diameter differentiation index, respectively. Despite trying to correct for bias by incorporating approximated weights, the VA method provided biased results in a number of combinations of  $n_n$  and  $n_t$ , similar to the NT and AZ methods (Table 2).

In general, RB was in the range of (-10%,10%) for the species mingling index and in the range of (-2.5%,2.5%) for the diameter differentiation index; however, for certain combinations of stem map, method,  $n_n$ , and  $n_t$ , the absolute value of RB exceeded 10% and 5% for each index, respectively (Figs. 1 and 2). The RRMSE was

in general below 10% for the species mingling index and below 4% for the diameter differentiation index, while for certain combinations of stem maps, methods,  $n_n$ , and  $n_t$ , RRMSE exceeded 20% and 5% for each index, respectively (Figs. 1 and 2). The ANOVA revealed that differences in RRMSE and RB among methods were significant for both variables (Table 3). Pairwise comparisons among methods by means of Tukey's honest significant differences tests showed that the FI method had significantly lower RRMSE than any other method, except when compared to the AZ method for species mingling estimation. The between-site variability in RRMSE and RB for the remaining methods was high for both indices, making it difficult to draw general conclusions about the performance of the sampling methods (Table 3). Neither  $n_n$  nor  $n_t$  had a significant effect on RRMSE or RB and interaction terms were not significant either.

**Fig. 6.** Variability of relative root mean squared error (RRMSE) across sites by species and methods for the diameter differentiation index. Species: DF, Douglas-fir; RA, red alder; RC, western red cedar; WH, western hemlock.



Species-specific results were similar to those obtained for all species together. The factor species was significant for the RB and RRMSE of both indices. The method was again a significant factor for both error measures and indices, and pairwise comparisons among methods by means of Tukey's honest significant differences tests showed that fixed-radius plots had significantly lower RRMSE than any other method (Table 4).

For the diameter differentiation index, the following sequence displays the methods tested in this study sorted by increasing species-specific RRMSE: FI < (AZ,NT) < VA. The symbol "<" indicates that a method on the left is superior in terms of RRMSE (i.e., lower RRMSE) to the method or group of methods (grouped with a parenthesis) on the right. For the species mingling index, the sequence of methods sorted by their species-specific RRMSE is similar to the previous; however, in this case, the RRMSE of the FI method was not significantly smaller than the RRMSE of the AZ

method. The variability of RB and RRMSE by methods and species is shown in Figs. 3 and 4 for the species mingling index. The variability of RB and RRMSE for the diameter differentiation index estimates is shown in Figs. 5 and 6. The FI method was fairly stable in terms of RRMSE and RB among sites for both indices. The variability of RRMSE and RB for the AZ and NT methods was important for certain combinations of species and index. For example, when considering the species mingling index for Douglas-fir, the variability of RB and RRMSE for the NT method was consistently larger than the variability for the FI and AZ methods. Similarly, the variability of RB and RRMSE for diameter differentiation of western red cedar was larger than that observed for the same species with the FI and NT methods.

While all sampling methods examined had low RRMSE and RB, fixed-plot sampling performed consistently better than the other methods, which suggests that this method should be preferred for

further applications when only precision is considered. This method is relatively simple and does not require an important amount of extra field measurements when compared to the NT or AZ method, and its implementation in the field is straightforward. All methods are designed to select a single reference tree in each sampling point; however, the AZ and NT methods do not require a first step to determine what trees are included in a plot. Both the NT and AZ methods select trees with probability proportional to the area of an associated polygon, the size and shape of which are dictated by the sampling system and positions of the trees. The selection bias inherent in the NT method has been a topic of discussion in natural resources sampling but the AZ method has been little studied. The simplicity of the AZ and NT selection protocols comes at the expense of (1) a tendency to select more isolated trees and (2) a difficult quantification of inclusion probabilities for each tree. However, for both methods, the RRMSE and RB, while dependent on the site, were relatively low (Figs. 1 and 2). It is therefore reasonable to think that the simplicity of these methods might outweigh their worse performance when compared to the FI method. Unfortunately, precise information about the time and effort required to implement each method in the field was not available for this study. Therefore, a proper evaluation of the operational implications of each method was impossible.

In the only other comparison of sampling methods for estimation of NNI published in English, Kint et al. (2004) evaluated two sampling methods, “distance sampling” and “plot sampling”. Distance sampling is the random selection of trees from a tree list, while plot sampling is the selection of a given number of trees that are nearest to a sample point. Distance sampling is a design-unbiased sampling method, while plot sampling is not design-unbiased (for a treatment of the calculation of inclusion areas for this sampling design, see Kleinn and Vilčko (2006)). Kint et al. (2004) found that plot sampling had substantially greater bias than distance sampling even with relatively large sample sizes. However, they concluded that plot sampling was sometimes a better candidate for estimation of diameter differentiation, while distance sampling was always a better candidate for estimation of species mingling.

One assumption of this study has been that a sampler wishes to estimate the arithmetic mean index value for a population. However, Gadow and Pogoda (2000) defended the NT approach by stating that “the structural attributes of the reference trees do not represent the relative share of the number of trees but proportions of the forest area”. Under this framework, the sampler would actually be interested in estimating the average value that is weighted by the area  $a_i$  where each tree would be the nearest tree (i.e., the area of the Thiessen polygon of each tree). This weighted average will obviously be different from the arithmetic average  $\bar{Y}$  introduced in eq. 3 in all but the most uniform of stands. Selection probabilities for the NT method are proportional to  $a_i$ , so the Horvitz–Thompson estimator can be computed because the terms  $a_i$  cancel out. In addition, a strong correlation of the selection probabilities (proportional to  $a_i$ ) with the variable of interest,  $a_i y_i$ , is very likely and would ensure that this method is nearly optimal. Other choices of weights may also be considered. If the index values of large trees are considered to be more meaningful in determining forest structure than those of small trees, weighting by basal area would be appropriate and variable-plot sampling would probably be the most efficient system. Regardless of what definition of NNI is adopted by the scientific community, a consistent definition will be helpful in facilitating comparisons between different research studies and moving NNI from the world of academia to the world of operational forest inventory.

## Conclusions

While NNI have attracted much attention from the forest research community, little work has been done to determine the most efficient (lowest RRMSE) sampling system for estimating index values. This study was performed to fill that gap by evaluating the performance of four sampling methods for estimating the species mingling and diameter differentiation methods in headwater riparian forests of western Oregon. Of the sampling methods considered, fixed-plot sampling consistently had the lowest bias and RRMSE. Fixed-plot sampling would likely produce low bias in any forest type. The bias of the other sampling methods was unpredictable and may be substantial under some conditions; however, in most situations, it was moderately low.

The appropriate sampling method to use is dependent on the attribute to be estimated. This study assumed that the sampler is interested in estimating the arithmetic mean index value for a population. Results would differ if a weighted mean were to be estimated.

Finally, for further studies aiming at analyzing the performance of different methods for estimation of NNI distributions, keeping the number of neighbors constant seems to be a reasonably good starting point.

## Acknowledgements

The authors wish to thank Vincent Kint for his support in earlier versions of the manuscript and two anonymous reviewers for the helpful and constructive comments provided.

## References

- Aguirre, O., Hui, G., Gadow, K.V., and Jiménez, J. 2003. An analysis of spatial forest structure using neighbourhood-based variables. *For. Ecol. Manage.* **183**(1–3): 137–145. doi:10.1016/S0378-1127(03)00102-6.
- Dunn, O.J. 1961. Multiple comparisons among means. *J. Am. Stat. Assoc.* **56**(293): 52–64. doi:10.1080/01621459.1961.10482090.
- Eskelson, B.N.I., Anderson, P.D., Hagar, J.C., and Temesgen, H. 2011. Geostatistical modeling of riparian forest microclimate and its implications for sampling. *Can. J. For. Res.* **41**(5): 974–985. doi:10.1139/x11-015.
- Franklin, J.F., and Spies, T.A. 1991. Composition, function, and structure of old-growth Douglas-fir forests. In *Wildlife and vegetation of unmanaged Douglas-fir forests*. Gen. Tech. Rep. PNW-GTR-285. Edited by L.F. Ruggiero, K.B. Aubry, A.B. Carey, and M.H. Huff. USDA For. Serv. Pacific Northwest Research Station, Portland, Ore. pp. 71–80.
- Földner, K. 1995. Strukturbeschreibung von Buchen-Edellaubholz-Mischwäldern [Describing stand structure in mixed beech–deciduous stands]. Ph.D. thesis, University of Göttingen, Göttingen, Germany.
- Gadow, K.V. 1993. Zur Bestandesbeschreibung in der Forsteinrichtung. *Forst Holz*, **48**(21): 602–606.
- Gadow, K.V., and Hui, G. 1999. Individual tree growth. In *Modelling forest growth*. Springer Netherlands, Dordrecht. pp. 129–179. doi:10.1007/978-94-011-4816-0\_5.
- Gadow, K.V., and Pogoda, P. 2000. Assessing forest structure and diversity. In *Man and forest*. Edited by R.K. Kohli, H.P. Singh, S.P. Vij, K.K. Dhir, D.R. Batish, and D.K. Khurana. Panjab University Publication, Chandigarh, India. pp. 1–8.
- Gadow, K.V., Zhang, C.Y., Wehenkel, C., Pommerening, A., Corral-Rivas, J., Korol, M., Myklush, S., Hui, G.Y., Kiviste, A., and Zhao, X.H. 2012. Forest structure and diversity. In *Continuous cover forestry*. Edited by T. Pukkala and K.V. Gadow. Springer, Dordrecht, Netherlands. pp. 29–83. doi:10.1007/978-94-007-2202-6\_2.
- García-Abril, A., Núñez, M.V., Grande, M.A., Velarde, M.D., Martínez-Obispo, P., and Rodríguez-Solano, R. 2013. Landscape indicators for sustainable forest management. In *Quantitative techniques in participatory forest management*. Edited by E. Martínez-Falero, S. Martín-Fernández, and A.D. García-Abril. CRC Press, New York. pp. 263–366. doi:10.1201/b15366-7.
- Haxtema, Z., Temesgen, H., and Marquardt, T. 2012. Evaluation of  $n$ -tree distance sampling for inventory of headwater riparian forests of western Oregon. *West. J. Appl. For.* **27**(3): 109–117. doi:10.5849/wjaf.10-035.
- Horvitz, D.G., and Thompson, D.J. 1952. A generalization of sampling without replacement from a finite universe. *J. Am. Stat. Assoc.* **47**(260): 663–685. doi:10.1080/01621459.1952.10483446.
- Howard, A.F., and Temesgen, H. 1997. Potential financial returns from alternative silvicultural prescriptions in second-growth stands of coastal British Columbia. *Can. J. For. Res.* **27**(9): 1483–1495. doi:10.1139/x97-114.
- Iles, K. 1979. Systems for the selection of truly random samples from tree populations and the extension of variable plot sampling to the third dimension. Ph.D. thesis, The University of British Columbia, Vancouver, B.C.

- Iles, K. 2003. A sampler of inventory topics: a practical discussion for resource samplers, concentrating on forest inventory techniques. Kim Iles & Associates, Limited, Nanaimo, B.C.
- Iles, K. 2009. "Nearest-tree" estimations — a discussion of their geometry. *MCFNS*, **1**(2): 47–51.
- Kint, V., Robert, D.W., and Noël, L. 2004. Evaluation of sampling methods for the estimation of structural indices in forest stands. *Ecol. Model.* **180**(4): 461–476. doi:10.1016/j.ecolmodel.2004.04.032.
- Kleinn, C., and Vilčko, F. 2006. Design-unbiased estimation for point-to-tree distance sampling. *Can. J. For. Res.* **36**(6): 1407–1414. doi:10.1139/x06-038.
- Marquardt, T., Temesgen, H., and Anderson, P.D. 2010. Accuracy and suitability of selected sampling methods within conifer dominated riparian zones. *For. Ecol. Manage.* **260**(3): 313–320. doi:10.1016/j.foreco.2010.04.014.
- Marquardt, T., Temesgen, H., Anderson, P.D., and Eskelson, B. 2012. Evaluation of sampling methods to quantify abundance of hardwoods and snags within conifer-dominated riparian zones. *Ann. For. Sci.* **69**(7): 821–828. doi:10.1007/s13595-012-0204-5.
- Mason, W.L., Connolly, T., Pommerening, A., and Edwards, C. 2007. Spatial structure of semi-natural and plantation stands of Scots pine (*Pinus sylvestris* L.) in northern Scotland. *Forestry*, **80**(5): 567–586. doi:10.1093/forestry/cpm038.
- Pommerening, A. 2002. Approaches to quantifying forest structures. *Forestry*, **75**(3): 305–324. doi:10.1093/forestry/75.3.305.
- Pommerening, A. 2006. Evaluating structural indices by reversing forest structural analysis. *For. Ecol. Manage.* **224**(3): 266–277. doi:10.1016/j.foreco.2005.12.039.
- Pommerening, A., and Stoyan, D. 2006. Edge-correction needs in estimating indices of spatial forest structure. *Can. J. For. Res.* **36**(7): 1723–1739. doi:10.1139/x06-060.
- Poudel, K.P., Temesgen, H., and Gray, A.N. 2015. Evaluation of sampling strategies to estimate crown biomass. *For. Ecosyst.* **2**: 1. doi:10.1186/s40663-014-0025-0.
- Pretzsch, H. 1997. Analysis and modeling of spatial stand structures. Methodological considerations based on mixed beech–larch stands in Lower Saxony. *For. Ecol. Manage.* **97**(3): 237–253. doi:10.1016/S0378-1127(97)00069-8.
- R Core Team. 2016. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna, Austria. Available from <https://www.R-project.org/>.
- Ripley, B.D. 1977. Modelling spatial patterns. *J. R. Stat. Soc. B Methodol.* **39**(2): 172–212.
- Ripley, B.D. 1979. Tests of 'randomness' for spatial point patterns. *J. R. Stat. Soc. B Methodol.* **41**(3): 368–374.
- Saunders, M.R., and Wagner, R.G. 2008. Long-term spatial and structural dynamics in Acadian mixedwood stands managed under various silvicultural systems. *Can. J. For. Res.* **38**(3): 498–517. doi:10.1139/X07-155.
- Schütz, J.P. 1997. Sylviculture 2, la gestion des forêts irrégulières et mélangées. Les Presses polytechniques et universitaires romandes, Lausanne, Switzerland.
- Schütz, J.-P., Pukkala, T., Donoso, P.J., and Gadov, K.V. 2012. Historical emergence and current application of CCF. In *Continuous cover forestry*. Edited by T. Pukkala and K.V. Gadov. Springer, Dordrecht, Netherlands. pp. 1–28. doi:10.1007/978-94-007-2202-6\_1.
- Sterba, H. 2008. Diversity indices based on angle count sampling and their interrelationships when used in forest inventories. *Forestry*, **81**(5): 587–597. doi:10.1093/forestry/cpn010.
- Temesgen, H. 2003. Evaluation of sampling alternatives to quantify tree leaf area. *Can. J. For. Res.* **33**(1): 82–95. doi:10.1139/x02-159.
- Temesgen, H., Monleon, V., Weiskittel, A., and Wilson, D. 2011. Sampling strategies for efficient estimation of tree foliage biomass. *For. Sci.* **57**(2): 153–163.
- Whittaker, R.H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.* **30**(3): 279–338. doi:10.2307/1943563.
- Zenner, E.K., and Hibbs, D.E. 2000. A new method for modeling the heterogeneity of forest structure. *For. Ecol. Manage.* **129**(1–3): 75–87. doi:10.1016/S0378-1127(99)00140-1.