Ecological Informatics 81 (2024) 102579

Contents lists available at ScienceDirect

Ecological Informatics

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

Monitoring spatial tree diversity indices using relascope sampling: Estimators, interactions and limitations

Arne Pommerening [a,*], Hubert Sterba [b]

[a] Swedish University of Agricultural Sciences SLU, Faculty of Forest Sciences, Department of Forest Ecology and Management, Skogsmarksgränd 17, SE-901 83 Umeå, Sweden
[b] Institute of Forest Growth, Department of Forest and Soil Sciences, BOKU University of Natural Resources and Life Sciences, Peter-Jordan-Straße 82, A-1190 Vienna, Austria

ARTICLE INFO

Keywords:
Climate change
Biodiversity
Resilience
Edge effects
Forest inventory
Sampling simulation
Continuous cover forestry

ABSTRACT

With ongoing climate change the monitoring of tree diversity has become very important for avoiding or at least decelerating the loss of biodiversity and for maintaining forest ecosystem resilience. As part of such monitoring, spatial indices of tree diversity which are calculated for individual trees often serve as surrogates for more direct measures of biodiversity. Mainly for its efficiency and ease of application, relascope sampling is a widespread method applied in forest inventory of many countries and thus has often been suggested as a data source for the monitoring of tree diversity. Since the interaction between sampling design and spatial diversity indices is not always clear to data analysts, we reviewed existing estimators and experimentally examined a new one, conducted extensive sampling simulations using different indices and estimators and additionally analysed the data from a large-scale forest inventory in Austria. We found that both forest structure and index algorithm greatly influence the sampling error. The largest source of sampling error was the index variance and contrary to our expectation not so much the bias due to spatial effects. For diversity indices related to distances, it has turned out to be best to apply estimators that include spatial edge correction methods. For all other indices an estimator performed better that included information on both the sample trees and their nearest neighbours, as it much reduced overall index variance. However, if possible the plus-sampling edge correction method should be applied.

1. Introduction

Ongoing human-induced climate change is currently unfolding at a pace that is considerably faster than what we know from past events of natural climate alterations (The Royal Society and the US National Academy of Sciences, 2020). Such an unprecedented rate can potentially imply a substantial loss of species that are unable to adapt quickly (McElwee, 2021; Román-Palacios and Wiens, 2020). Loss of species is a major concern in itself, however, research has also shown that biodiversity is crucial to ecosystem resilience and thus to sustaining terrestrial and marine ecosystems and habitats (Fischer et al., 2006; Matias et al., 2013; Oliver et al., 2015; Yachi and Loreau, 1999). As a pre-requisite for stopping or at least decelerating the loss of biodiversity it is important to have efficient monitoring systems in place that provide the necessary information to be able to interpret ongoing trends correctly (Hoffmann, 2022; Kühl et al., 2020).

In forest ecosystems, individual-tree indices of spatial diversity are often considered surrogates for more direct measures of biodiversity, since forest ecosystems with complex structures usually provide more habitats that can accommodate a greater diversity of plant and animal species than forests with low tree diversity (Aguirre et al., 2003; Gadow, 1993; Neumann and Starlinger, 2001; Pommerening and Grabarnik, 2019). Such local measures of spatial tree diversity are based on the nearest-neighbour principle and are particularly informative, since they supply information about *tree species*, *tree size* and *tree location diversity*, see Table 1. They also have the advantage that they are comparatively easy to estimate from data collected in existing forest inventories that were carried out for other purposes and this can even be achieved retrospectively by analysing inventory data that have been collected some time ago (Motz et al., 2010; Sterba, 2008).

The estimation of spatial tree diversity indices from simple, circular or rectangular *sample plots* (also referred to as fixed-area plots) is

---

* Corresponding author.
E-mail address: arne.pommerening@slu.se (A. Pommerening).

https://doi.org/10.1016/j.ecoinf.2024.102579
Received 20 December 2023; Received in revised form 24 March 2024; Accepted 25 March 2024

comparatively straightforward. All trees whose stem-centre coordinates occur inside such plots are considered *sample trees*, i.e. they are included in the sample. Tree diversity indices can be calculated for each sample tree and arithmetic means of these sample-tree indices constitute the index estimate of each sample point which forms the centre of the sample plot. Naturally, trees outside the sample plots have zero *inclusion probabilities* and are therefore not included in the sample. The probability that a given individual is included in a sample is proportional to the sample plot area and the fixed-area sample plot constitutes the *inclusion zone* which is the same for all sample trees of a given fixed-area sample plot (Gregoire and Valentine, 2008; Mandallaz, 2008).

Sampling tree diversity indices based on fixed-area plots usually carries a low sampling error (Motz et al., 2010). An important source of bias much discussed in spatial statistics (Illian et al., 2008) are edge effects which arise from the fact that some of the nearest neighbours of sample trees are just outside the sample plot but according to the sampling design have not been accounted for. Such nearest neighbours are essential for the calculation of spatial tree diversity indices (Pommerening and Grabarnik, 2019). One important method for mitigating edge effects is *plus-sampling* (Hui and Albert, 2004; Mauro et al., 2017; Motz et al., 2010), where off-plot tree neighbours of sample trees are measured as well. Information about such off-plot neighbour trees is not included in the sample but merely used to compute the spatial diversity indices. Often the measurement of off-plot neighbours is not possible in forest inventories primarily designed for other purposes, because this would constitute a considerable additional effort or the data available for the index calculation have already been collected in the past without considering plus-sampling and it is not possible to re-visit the sample plots for collecting additional information on tree neighbours.

The estimation of spatial tree diversity indices is less straightforward in *nested sample plots* (Gadow et al., 2021; Häbel et al., 2019; Roesch, 2008), where tree information from more than one sample plot is collected or the sample trees relating to the same sample point have different inclusion zones. Nested plots typically share the same sample point as plot centre but have different sizes which imply different inclusion probabilities associated with each tree depending on the plot the tree is assigned to according to the sampling design. In contrast to fixed-area plots, the design of nested sample plots also implies that not all trees are sampled between the sample plot centre and the radius of the largest plot. Since they are often part of the original spatial neighbourhoods, these unselected trees are a second source of sampling error in addition to edge effects. Typical examples of nested plots are *concentric circular sample plots*. The set of individual-tree inclusion zones associated with *relascope sampling* (also referred to as angle count (Barabesi and Fattorini, 1999; Bitterlich, 1948) or point sampling (Grosenbaugh, 1958)) form a certain analogy to nested plots: In relascope sampling, each tree's inclusion zone radius is proportional to its stem diameter and a viewing angle expressed by the basal-area factor. Therefore relascope sampling can be described as a sampling method where each tree of a sample has an individual inclusion zone or 'sample plot'. Thus relascope sampling can be interpreted in terms of sampling that is based on circular sample plots.

The application of nested plots and relascope sampling has turned out to be more efficient in forests with complex forest structure where tree sizes can vary considerably even at close proximity and small-sized trees can occur in great numbers whilst large trees are comparatively rare (Mandallaz and Lanz, 2001). These sampling designs have therefore often been introduced in countries where forest management followed the principles of continuous cover forestry (Pommerening, 2023) which promotes diverse forest structure. Both concentric circular sample and relascope sample plots are common in large-scale forest inventories, e.g. in regional and forest estate inventories but also in national forest inventories (NFI). Particularly at this spatial scale, plus-sampling is often argued not to be affordable. Therefore the existence of different inclusion probabilities for sample trees, the design-based exclusion of trees and the ignorance of off-plot neighbours of sample trees can lead to considerable sampling errors thus decreasing the quality of the monitoring information. However, nested plots and relascope sampling are a monitoring reality in many countries and an important source of tree diversity information. Many data analysts process nested-plot and relascope-sampling data for estimating measures of tree diversity and are unaware of the interaction between indices and sampling design. Our study therefore established the associated possibilities and limitations.

The objectives of this paper are (1) to explain in detail the principles of estimating tree diversity indices from relascope sampling data, since this is not well documented in the literature, and (2) based on extensive sampling simulations to identify factors influencing the estimation of different diversity indices by using alternative, heuristic estimators. Finally (3) we validated all estimators using data from a real forest estate inventory in Austria and provided conclusions based on our findings.

2. Materials and methods

2.1. *Spatial diversity indices*

We selected five different tree diversity indices measuring the diversity of tree location, tree species and tree size diversity (Table 1). These indices are often considered in the monitoring of tree diversity (Aguirre et al., 2003; Gadow, 1993; Neumann and Starlinger, 2001; Pommerening and Grabarnik, 2019), but they were also chosen to represent a range of different approaches to constructing such indices: One index included a ratio (Eq. 1 in Table 1), others used indicator functions (Eqs. 2, 3 and 5 in Table 1) and one was a combination of different approaches (Eq. 3 in Table 1). To include this range of index construction methods was important in order to understand how they interact with the relascope inventory design.

The aggregation index of Clark and Evans (1954) is one of the earliest

Table 1
Definitions of the spatial tree diversity indices used in this study.

Eq.	Name	Diversity of	Formula	Variable definitions
(1)	Aggregation index (Clark and Evans, 1954)	Location	$R = \frac{\bar{r}}{\mathbf{E}r}$	$\bar{r}$ – observed mean arithmetic distance between sample tree i and its first nearest neighbour; $\mathbf{E}r = 1/\left(2 \times \sqrt{N/A}\right)$ with N – number of trees and A – sample plot area.
(2)	Species mingling (Gadow, 1993)	Species	$M_i = \frac{1}{k}\sum_{j=1}^{k}\mathbf{1}\left(\text{species}_i \neq \text{species}_j\right)$	$\mathbf{1}(A) = 1$, if A is true, otherwise $\mathbf{1}(A) = 0$
(3)	Weighted species mingling (Hui et al., 2011)	Species	$M_i^{'} = \frac{1}{k \times c}\sum_{j=1}^{k}\mathbf{1}\left(\text{species}_i \neq \text{species}_j\right) \times s_i$	s_i – species richness among the k nearest neighbours of tree i; $c = min(S, k+1)$, where S – total species richness (either in the total population or at a specific sample point)
(4)	Size differentiation (Gadow, 1993)	Size	$T_i = 1 - \frac{1}{k}\sum_{j=1}^{k}\frac{min(m_i, m_j)}{max(m_i, m_j)}$	m_i – size of subject tree i, m_j – size of neighbour j
(5)	Size dominance (Aguirre et al., 2003)	Size	$U_i = \frac{1}{k}\sum_{j=1}^{k}\mathbf{1}(\mathrm{m}_i > \mathrm{m}_j)$	See previous definitions

diversity indices quantifying the spatial pattern of plant dispersion. The index is very efficient and popular. Observed mean first-neighbour distance, $\bar{r}$, is divided by the mean distance of a tree pattern where the tree locations are completely random, $\mathbf{E}r$, (Eq. 1 in Table 1). Usually, the interpretation of R values is as follows: $R > 1$, if the pattern has a tendency to regularity, $R = 1$, if it is completely random and $R < 1$, if there is clustering in a pattern. We also monitored the estimation of $\bar{r}$ separately of R in order to better understand the behaviour of the R estimations.

Gadow (1993) defined *spatial species mingling* as the mean heterospecific fraction of plants among the k nearest neighbours of a subject plant i (Eq. 2 in Table 1). Due to the discrete nature of outcomes for a given k, there are only $k+1$ possible values M_i can take, i.e. $0/k, 1/k, \ldots, k/k$, where the number in the numerator denotes the number of neighbours with a species different from that of tree i. All index values lie between 0 and 1.

Hui et al. (2008) and Hui et al. (2011) proposed the *species richness-weighted mingling index* $M_i^{'}$ (Eq. 3 in Table 1) as an extension of the original mingling index by Gadow (1993) by merging the concept of species mingling with the concept of species richness. Accordingly, each M_i (from Eq. 2) is multiplied by the species richness s_i among the k nearest neighbours. Wang et al. (2021) amended the index definition by introducing term c in Eq. (3) to ensure that the maximum number of species that are theoretically possible in a group of $k+1$ trees is limited by the number of species present in the forest stand or in the monitoring area studied. Values of $M_i^{'}$ are generally smaller and take a larger range of different values than those of M_i.

Gadow (1993) defined *size differentiation* (Eq. 4 in Table 1) as the mean ratio of smaller-sized and larger-sized marks of the k nearest neighbours subtracted from one. Size differentiation produces continuous results between 0 and 1 and T_i increases with increasing average size difference between neighbouring trees.

The *size dominance index* (Eq. 5 in Table 1) was introduced by Hui et al. (1998) and Aguirre et al. (2003) and gives the proportion of the k nearest neighbours dominated by tree i. The index draws on similarities with the construction of the mingling index, thus transforming a continuous variable into a binary one and in analogy to the species mingling index also produces only $k+1$ possible U_i values.

All diversity indices were estimated for $k = 1$ nearest neighbour and those of Eqs. (2), (3), (4) and (5) were additionally estimated for $k = 4$ neighbours to better understand the effect of k on the estimators.

2.2. Sampling design and estimators

2.2.1. General, non-spatial estimators

The standard, unbiased inventory estimator used in the context of both simple, circular sample plots and nested plots includes so-called Horvitz-Thompson (HT) weights (Cochran, 1977; Gregoire and Valentine, 2008; Horvitz and Thompson, 1952; Mandallaz, 2008). Assuming circular sample plots, for a sample tree i this weight is typically defined as

$$w_i = \frac{1}{r_i^2 \times \pi} \tag{6}$$

and is often expressed as the reciprocal of hectare, i.e. ha^{-1}. r_i is the radius of the inclusion zone that tree i is related to according to the sampling design. The product $r_i^2 \times \pi$ is the area of this zone. As a consequence, a tree in a hypothetical, small inclusion zone carries a larger weight w_i than the same tree in a large sample plot and corresponding inclusion zone. The general idea of HT weights is to give rarer events larger weights than the more common events (Pommerening and Stoyan, 2006). w_i is the reciprocal of the probability that tree i is included in the sample.

Due to our validation data set (see Section 2.3.3) in this study, our interest is mainly in the estimation of sample point means rather than in the estimation of means for the entire sampling area. Using HT weights, the general, non-spatial sample-point estimator, $\widehat{\bar{y}}$, of an arbitrary characteristic y can be written as

$$\widehat{\bar{y}} = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} y_i \times w_i. \tag{7}$$

In Eq. (7), n is the number of trees included in the sample at a given sample point, i.e. the number of sample trees. In the remainder of this paper, we referred to Eq. (7) as the *inventory estimator.*

2.2.2. NN1 and NN2 estimators

Unlike traditional forestry characteristics such as tree stem diameter (usually measured at 1.3 m above ground level) or basal area (cross-sectional surface area at 1.3 m above ground level), tree diversity indices as used in this study are spatially explicit and their calculation involves not only characteristics of sample tree i but also characteristics of neighbouring trees j. This implies (1) that edge effects potentially occur and need to be accounted for. Also, (2) the question arises whether the simple HT weight of Eq. (6) is sufficient or whether the HT weight(s) of the nearest neighbours involved in the index calculation need to be considered as well.

In order to tackle problem (1), Hanisch (1984) and Pommerening and Stoyan (2006) introduced two alternative weights that share similarities with HT weights and account for possible edge effects, NN1 and NN2, which were originally proposed in the context of point process statistics and therefore for data obtained from mapping forests. In both approaches, the tree diversity characteristic of tree i only contributes to the sample point estimate, if distance c_i between tree i and its nearest neighbour is smaller than distance s_i between tree i and the sample-plot or inclusion-zone boundary (see Fig. 1). Checking up on the relationship between the two distances is handled by an indicator function, $\mathbf{1}(c_i < s_i)$, returning the value of 1, if the condition in brackets is true and 0 otherwise. A simple alternative to Eq. (6) is to define a weight that only considers sample trees fulfilling condition $c_i < s_i$:

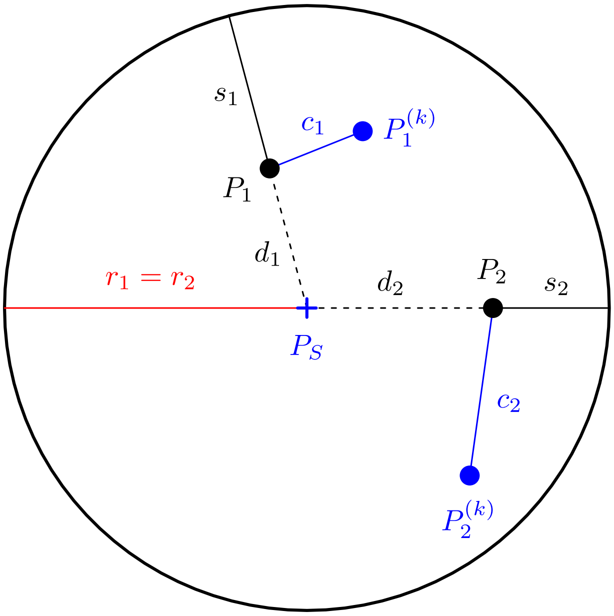

Fig. 1. Illustration of the working principle of the indicator function $\mathbf{1}(c_i < s_i)$. P_S is the sample point coinciding with the centre of the circular sample plot. Points P_1 and P_2 represent locations of trees included in a relascope sample and for simplicity we assume that they have the same stem diameter. For point P_1, $\mathbf{1}(c_i < s_i) = 1$, since its distance s_1 to the sample plot boundary is larger than distance c_1 to the nearest neighbour $P_1^{(k)}$ and for P_2, $\mathbf{1}(c_i < s_i) = 0$, since $s_2 < c_2$.

$$w_i^{(2)} = \mathbf{1}(c_i < s_i) \times w_i = \frac{\mathbf{1}(c_i < s_i)}{r_i^2 \times \pi} \tag{8}$$

$w_i^{(2)}$ is the weight of the NN2 estimator, which is not ratio-unbiased (Hanisch, 1984; Illian et al., 2008; Pommerening and Stoyan, 2006), i.e. although intuitively appealing the estimator is not supported by statistical theory. An alternative to the NN2 estimator is the more sophisticated NN1 estimator. This estimator not only excludes some sample trees in the same way as NN2, but also modifies the individual-tree inclusion zones and as a consequence the NN2 weights. The inclusion zone radius, r_i, is reduced by distance, c_i, between tree i and its nearest neighbour (Eq. 9). This reduction results in smaller individual-tree inclusion zones and larger weights, $w_i^{(1)}$, of those sample trees that meet condition $c_i < s_i$, particularly if c_i is large. These potentially large weights, $w_i^{(1)}$, are supposed to compensate for the data loss incurred due to the condition of the indicator function (Pommerening and Stoyan, 2006). Weight $w_i^{(1)}$ is thus defined as

$$w_i^{(1)} = \frac{\mathbf{1}(c_i < s_i)}{(r_i - c_i)^2 \times \pi}. \tag{9}$$

For estimating the sample plot means of spatial diversity indices, both weights, $w_i^{(1)}$ and $w_i^{(2)}$, are used in the same way as w_i in Eq. (7). Eq. (9) leads to ratio-unbiased estimations (Hanisch, 1984; Illian et al., 2008; Pommerening and Stoyan, 2006), i.e. the NN1 estimator has a better statistical foundation than the NN2 estimator.

When applying the NN1 and NN2 estimators, we expected only a partial bias reduction, since some small sample-tree neighbours $P_i^{(k)}$ (Fig. 1) not very far from the sample point are neighbours in nature but may not have been included in the relascope sample for design reasons. In that case, as an alternative other neighbours are considered in the NN1 and NN2 calculations, which potentially contribute to a bias.

2.2.3. *NN3 estimator*

We also studied the nearest-neighbour estimator suggested by Mauro et al. (2017), i.e.

$$\hat{\hat{y}} = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{k} w_{i,j}} \sum_{i=1}^{n}\sum_{j=1}^{k} y_i \times w_{i,j}. \tag{10}$$

This estimator only involves the weights of the nearest neighbours of sample trees. In Eq. (10), k is the number of nearest neighbours considered and $w_{i,j}$ is the weight (Eq. 6) of the jth neighbour tree of sample tree i. Since in monitoring practice only sample data are available, these neighbour trees j are sample trees and, as mentioned in Section 2.2.2, may therefore partly differ from the Euclidean nearest neighbours observed in nature.

2.2.4. *NN4 estimator*

As an explorative line of investigation and as a logical “compromise” between and combination of the inventory estimator (Eq. 7) and Eq. (10), we designed a fourth nearest-neighbour estimator that not only includes the weights of the nearest neighbours of sample trees but also those of the sample trees. This NN4 estimator we defined as

$$\hat{\hat{y}} = \frac{1}{2}\left(\frac{1}{\sum_{i=1}^{n} w_i}\sum_{i=1}^{n} y_i \times w_i + \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{k} w_{i,j}}\sum_{i=1}^{n}\sum_{j=1}^{k} y_i \times w_{i,j}\right). \tag{11}$$

In Eq. (11), k is again the number of nearest neighbours considered and as in Eq. (10), $w_{i,j}$ is the weight (Eq. 6) of the jth neighbour tree of sample tree i.

For calculating the aggregation index by Clark and Evans (1954), $\bar{r}$ (Eq. 1 in Table 1) is estimated according to Eqs. (7)–(11) and tree density $\lambda = N/A$ is estimated in a non-spatially explicit way for each sample point as $\widehat{\lambda} = \frac{1}{\sum_{i=1}^{n} w_i}$.

Whilst NN1 and NN2 address the edge-bias problem, the NN3 and NN4 estimators take care of the fact that information of sample tree neighbours also contributes to the index estimation. According to the sampling design, these neighbours are accounted for in inclusion zones different from that of tree i. As a consequence, these neighbours typically have weights, $w_{i,j}$, that differ from those of sample tree i and are ignored by the estimators of Eqs. (8) and (9).

2.3. *Data*

For simulating the approximate equivalence of large estate inventories, two fully mapped large forest areas were identified in Austria, i.e. Hirschlacke and Rosalia forest areas. In addition, we used data from a real large-scale forest inventory in Kammer forest for validation.

2.3.1. *Hirschlacke forest area*

The Hirschlacke forest area (48.64 N, 13.96 E) is situated in the northwestern corner of Austria, near the borders to the Czech Republic in the North and to Germany in the West. When Hirschlacke was established as a research forest in 1977, the trees were 120 years old and dominated by Norway spruce (*Picea abies* L. Karst.). The forest also includes smaller proportions of silver fir (*Abies alba* Mill.) and beech (*Fagus sylvatica* L.). After 1977, the Hirschlacke forest area was treated by a combination of local, individual-based crown thinnings and target diameter harvesting (Reininger, 2000) for achieving a transformation to a complex-structured forest. After 40 years of treatment and observation, the Hirschlacke forest area is approaching a structure reminiscent of that which is typical of single-tree selection systems (Sterba, 2004). For the sampling simulation study, a 153 × 166 m large rectangular sub-plot and the data of the 2017 survey were used.

2.3.2. *Rosalia forest area*

The Rosalia forest area (48.72 N, 16.17 E) is situated in the Rosalia Mountains in the eastern part of Austria at the northern border between the federal states Lower Austria and Burgenland. The area is part of the university forest of the University of Natural Resources and Life Sciences, BOKU, and includes a mixture of mainly Norway spruce, Scots pine (*Pinus sylvestris* L.), larch (*Larix decidua* Mill.) and beech. Before 1980 the forest area was managed as part of an agricultural estate with irregular, selective single-tree harvests which gradually gave rise to a complex CCF structure. Later on the ownership changed and the forest structure was simplified (Sterba, 2004). When the monitoring data were analysed in 1997, the forest was 98 years old with only a small cohort of young trees. For the simulation study, a 170 × 164 m large rectangular sub-plot and the data of the 1997 survey were used.

2.3.3. *Kammer forest area*

In 2006, a large-scale forest inventory based on relascope sampling with basal area factor 4 $\mathrm{m}^2\ \mathrm{ha}^{-1}$ was performed in the Kammer forest area (47.93 N, 13.60 E), a 750-ha forest district which is owned by the same Prämonstratenser Chorherrenstift Schlägl in Upper Austria that also owns the Hirschlacke forest area. A detailed description of the sites and the location of the forest area are given in Sterba (2008). Transformation to CCF had started here 40 years ago, i.e. the forest already has a considerable degree of complexity. Similar to the Hirschlacke forest area, Kammer is dominated by Norway spruce but also involves silver fir and beech. The inventory included a total of 245 sample points, but only 169 were used in this study, as they included more than three sample trees as required by the diversity-index estimators NN1-NN4. In the spirit of plus-sampling, the first neighbour of each sample tree was additionally measured irrespective whether this neighbour was included in the sample or not. These data were used for a validation of the results obtained from the sampling simulations.

The highest densities both in terms of the number of trees and basal area per hectare occur in the Hirschlacke forest area (Table 2, Fig. 2). The Kammer forest area has the lowest basal area but the second highest number of trees per hectare. Quadratic mean diameter is by far the

Table 2

Population basal area, G, number of trees per hectare, N, quadratic mean diameter, d_g, nearest-neighbour distance, $\bar{r}$, aggregation index, R, species mingling, $\overline{M}^{(1)}$, stem-diameter differentiation, $\overline{T}^{(1)}$, and stem-diameter dominance, $\overline{U}^{(1)}$, (see Table 1) in the three study areas Hirschlacke, Rosalia and Kammer. For Kammer forest, the diversity indices were calculated using plus-sampling for the first Euclidean neighbour which was included in the inventory.

Forest	G [m^2 ha^{-1}]	N [ha^{-1}]	d_g [cm]	$\bar{r}$ [m]	R	$\overline{M}^{(1)}$	$\overline{T}^{(1)}$	$\overline{U}^{(1)}$
Hirschlacke	38.7	938.3	22.9	1.6	0.96	0.17	0.37	0.51
Rosalia	36.5	375.9	35.2	3.1	1.21	0.46	0.32	0.55
Kammer	31.3	557.9	26.7	2.1	1.35	0.38	0.34	0.47

largest at Rosalia and, owing to the high number of small, regenerated trees (Fig. 2A), the smallest at Hirschlacke. Tree density is also reflected well by the average distance between a tree and its first nearest neighbour, $\bar{r}$. The quantity is largest at Rosalia and smallest at Hirschlacke. Apparently the tree locations form the most regular spatial pattern at Kammer forest area and there is a moderate cluster formation at Hirschlacke. Species mingling is moderately high in Rosalia and Kammer forests, but very low at Hirschlacke, mostly due to the dominance of Norway spruce. Stem-diameter differentiation is highest at Hirschlacke, but size dominance has a maximum at Rosalia and a minimum in the Kammer forest area.

The stem-diameter distributions show the strong influence of Norway spruce at Rosalia (Fig. 2B) and Kammer (Fig. 2C). European beech is also very prominent in these two forests so that both forest areas are clearly mixed in terms of species proportions. By contrast, the Hirschlacke site is almost entirely made up of Norway spruce, which explains the previously mentioned low species mingling value in Table 2. Hirschlacke (Fig. 2A) shows clear signs of approaching a negative exponential diameter distribution which is typical of selection forests (Pommerening, 2023; Schütz, 2001). A similar trend seems to exist for Kammer (Fig. 2B), however, since the stem-diameter distribution is estimated from the inventory data of a forest estate, the underlying samples were drawn from multiple forest stands across a very large area.

2.4. Sampling simulation

To study the behaviour of the estimators explained in the Section 2.2 and the interaction between the relascope sampling design and the tree diversity indices (Section 2.1), we carried out sampling simulations in the fully mapped forest areas described in the Section 2.3. The sampling simulator simulated systematic sampling grids at a random angle to the plot boundaries and at a distance of 25 m between grid points. These grid points were the sample points where relascope samples were simulated. Based on the definition of the basal area factor (van Laar and Akça, 2007), trees were included in the sample, if

$$\|P_S - P_i\| < r_i = \frac{50}{\sqrt{v}} \times d_i. \tag{12}$$

In Eq. (12), P_S is an arbitrary sample point, whilst P_i is the location of sample tree i (see Fig. 1). The expression $\|P_S - P_i\|$ denotes the Euclidean distance between points P_S and P_i. The individual inclusion zone radius of tree i is r_i and depends on basal area factor v and on the tree's stem diameter d_i. The basal area factor defines the sample size at each sample point, i.e. more trees are included with smaller values of v than with larger ones (Bitterlich, 1984). We ran independent simulations for $v = 1, 1.5, 2, 2.5, \ldots, 5$.

For each basal area factor, we simulated as many independent individual sampling grids as necessary to obtain at least 100,000 sample points in each of the two forest areas, i.e. Hirschlacke and Rosalia. For each of the nine basal area factors and approximately 100,000 sample points we estimated the tree diversity indices of Table 1 according to the standard inventory estimator of Eq. (7) and additionally by using the NN1 (Eq. 9), NN2 (Eq. 8), NN3 (Eq. 10) and NN4 (Eq. 11) estimators. We compared the results of these estimations with those obtained from applying again the standard inventory estimator of Eq. (7), however, this time using plus-sampling. The results of the estimations that included plus-sampling were treated as a reference. One should note that the estimation of spatial diversity indices including plus-sampling can still be biased, because some trees around the sample point, particularly small trees at larger distance from the sample point, may not have been included in the sample due to the selective sampling design (cf. Eq. 12) and therefore the spatial pattern formed by the sample trees may systematically misrepresent the underlying spatial tree pattern in nature. Still, this comparison is very meaningful for tree diversity monitoring, as edge effects are often perceived as the most important source of bias in the estimation of spatial tree-diversity indices from forest inventory data and plus-sampling is currently the standard recommendation for addressing this problem. Accordingly, plus-sampling was also applied in the Kammer validation data set (see Section 2.3.3).

For evaluating the simulation results and as an expression of sampling error we quantified *relative root mean squared error* (rRMSE*)* defined as

$$\text{rRMSE} = \frac{\sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(\widehat{y}_i - y_i)^2 + \left(\frac{1}{m}\sum_{i=1}^{m}(\widehat{y}_i - y_i)\right)^2}}{\bar{y}}, \tag{13}$$

where $\widehat{y}_i$ is the ith estimated diversity index per sample point, y_i is the corresponding ith estimation based on plus-sampling, m is the number of simulated sample points ($m \approx 100,000$) and $\bar{y}$ is the mean diversity index based on plus-sampling. In addition we also calculated the relative bias as

$$\text{rBias} = \frac{\sum_{i=1}^{m}(\widehat{y}_i - y_i)}{m\bar{y}}, \tag{14}$$

For this analysis we used our own R code (R Development Core Team, 2023) and the spatstat R package (Baddeley et al., 2016).

3. Results

3.1. Sampling simulations

In terms of general patterns and trends, the results obtained from the extensive sampling simulations are similar in both forests with the exception that the relative root mean square error (rRMSE) had a tendency to be generally larger in the Hirschlacke forest area (Fig. 3) than at Rosalia (Fig. 4) which can be attributed to the difference in tree location diversity as expressed by the aggregation index, see Table 2. This outcome emphasises that the sampling error strongly depends on forest structure and therefore can differ from forest to forest.

All rRMSE values are comparatively large highlighting the general difficulty to estimate spatial diversity indices from relascope sample data.

Generally, as expected, the sampling error decreases with increasing sample size. The rate of decrease much depends on the forest and the diversity index. For example, all estimators show a strong response to sample size (indicated by the curvature of the graphs) for diversity index $M^{(1)}$ in Fig. 4 whilst for indices $M^{(4)}$ (Eq. 2 in Table 1) and $M^{(4)}$ (Eq. 3 in Table 1) in Fig. 3 we can only see a weak response of the five estimators to sample size, i.e. the curves are almost horizontal lines.

It was interesting to observe that in both forests for the same di-

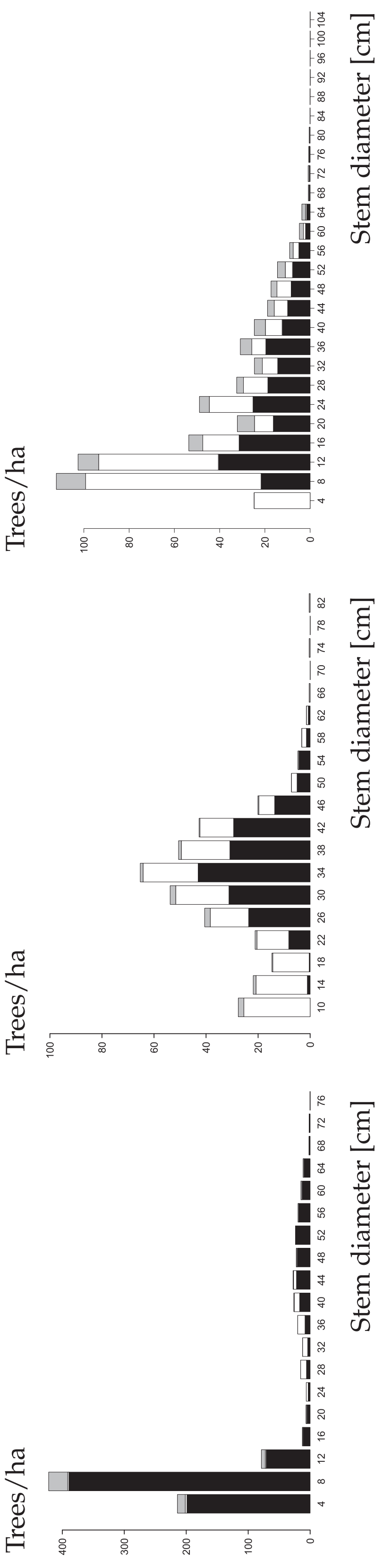

Fig. 2. Empirical stem-diameter distributions of the Hirschlacke (A), Rosalia (B) and Kammer (C) forests using 4-cm diameter classes. Black – Norway spruce, white – European beech and grey – other tree species.

versity indices with $k = 4$ rRMSE was usually lower than with $k = 1$ nearest neighbours. This finding was unexpected, since with increasing k the spatial neighbourhood and associated estimation errors potentially increase, which would also increase the problems of unobserved neighbours.

Diversity indices U (Eq. 5 in Table 1) and T (Eq. 4 in Table 1) are generally associated with a lower rRMSE than M (Eq. 2 in Table 1). The combination of M and species richness in index M' (Eq. 3 in Table 1) apparently leads to a marked increase of the sampling error. This is plausibly related to the variability of total species richness, S, estimated at each sample point.

Only for the mean distance between a sample tree and its first nearest neighbour, $\bar{r}$, and for R (Eq. 1 in Table 1), the NN1 (Eq. 9) and NN2 (Eq. 8) estimators lead to a lower rRMSE than NN3 and NN4. In these two cases but also generally across all diversity indices the NN2 estimator results in a lower sampling error than the NN1 estimator. The rRMSE values related to $\bar{r}$ and R are nearly identical at Rosalia, whilst the difference between the two quantities is slightly greater at Hirschlacke. This seems to suggest that, in contrast to the other composite index, M', the influence of estimated $\mathbf{E}\, r$ (Table 1, Eq. 1) on the sampling error of the aggregation index is comparatively low. The NN1 and NN2 estimators are apparently poor choices for any diversity index not directly related to distances.

Generally speaking the NN4 estimator (Eq. 11) that considers the weights of both sample trees and of the neighbours of sample trees is in most cases an improvement over the NN3 estimator (Eq. 10), although the improvement of NN4 over the inventory estimator (Eq. 7) is low except for diversity index U. Explorative test simulations have shown that averaging as in Eq. (11) retains a comparatively large influence of the weights of the sample trees as opposed to those of their neighbours and this appears to keep the overall sampling error low. The inventory estimator is generally the best choice for the M and M' indices, whilst T is situated halfway between U and M in terms of the best choice of estimator, where the inventory estimator, NN3 and NN4 largely show quite similar performances.

3.2. Forest inventory analysis

The analysis of the large-scale forest inventory in the Kammer forest area partly confirmed the results of the sampling simulation (see Section 3.1) and partly offered new insights. Given an average sample size of 7.8 trees per sample point (as a result of the application of basal area factor 4 m^2 ha^{-1}) the rRMSE values are lower than those in the Hirschlacke forest area but slightly larger or approximately the same compared to those obtained from the Rosalia simulations (Figs. 3, 4 and 5A).

Similar to the simulation results we can clearly see that in terms of rRMSE the NN1 and NN2 estimators are the best choice for $\bar{r}$ and for R, whilst for $M^{(1)}$, $T^{(1)}$ and $U^{(1)}$ the inventory, NN3 or NN4 estimators result in the lowest rRMSE values (Fig. 5A). The good performance of the inventory estimator for $M^{(1)}$ confirms the simulation results, whilst the poor performance of NN3 and NN4 for $U^{(1)}$ is surprising and difficult to explain.

Of particular interest are the rBias results (Fig. 5B). For $\bar{r}$ and for R all estimators overestimate the indices using plus-sampling, whilst for $M^{(1)}$, $T^{(1)}$ and $U^{(1)}$ all estimators underestimate the indices based on plus-sampling. Across all five diversity indices it is always the NN2 estimator followed by the NN1 estimator that leads to the lowest rBias values. By contrast, the NN3 and NN4 estimators often result in the largest values of rBias. This result was not so clear from the rBias results obtained from the sampling simulations (not shown). These results confirm that the NN1 and NN2 estimators have a better theoretical foundation whilst the inventory, the NN3 and the NN4 estimators apparently help reduce the overall variation.

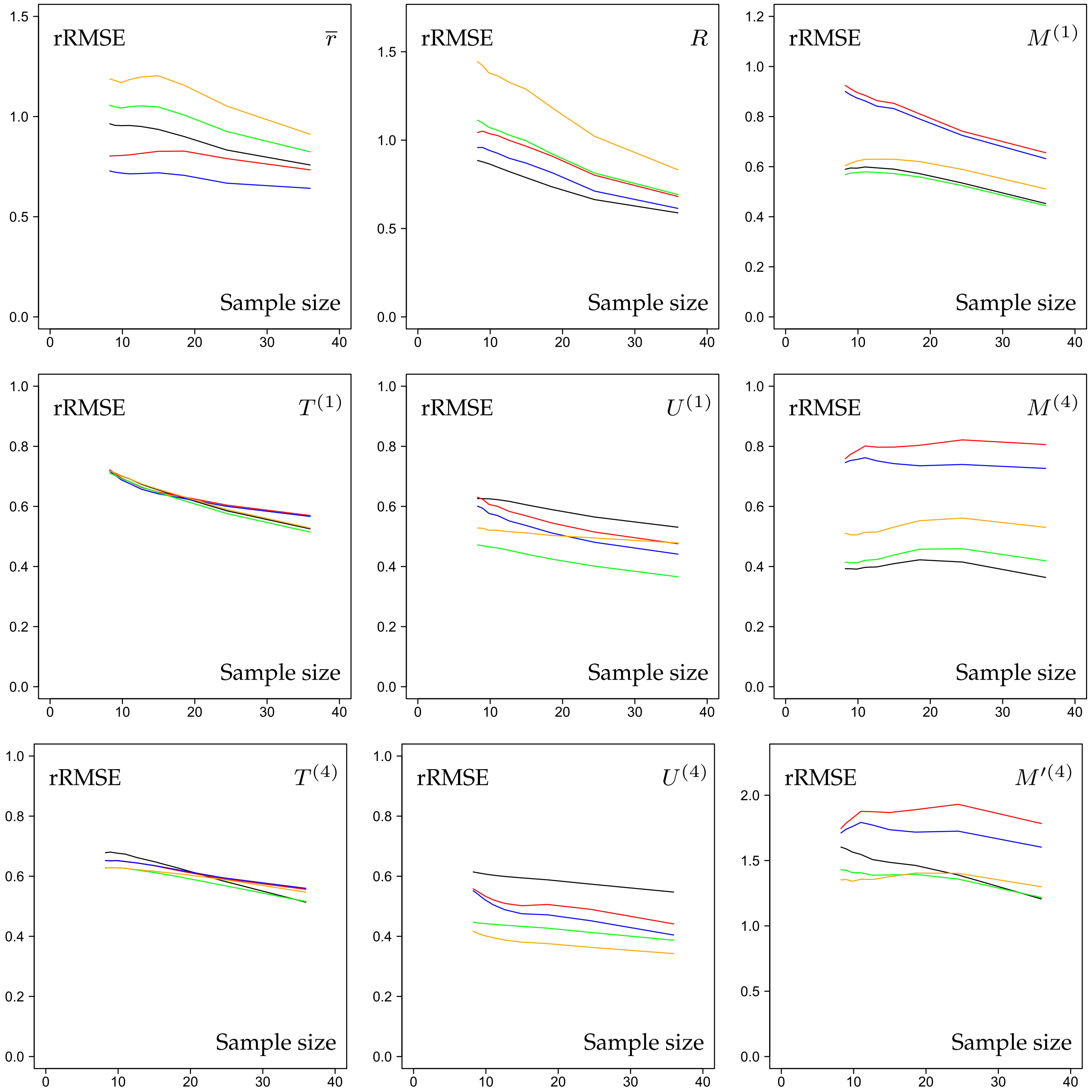

Fig. 3. Relative root mean square error (rRMSE, Eq. 13) of the diversity indices of Table 1 over sample size, i.e. the mean number of sample trees per relascope sample plot in the Hirschlacke forest. Black – inventory estimator (Eq. 7), red – NN1 estimator (Eq. 9), blue – NN2 estimator (Eq. 8), orange – NN3 estimator (Eq. 10), green – NN4 estimator (Eq. 11). The upper index in round brackets indicates the number of nearest neighbours, *k*, used to estimate the corresponding diversity index. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion and conclusions

In times of climate change, monitoring of tree diversity is more important than ever for mitigating the loss of diversity and resilience in forest ecosystems (Fischer et al., 2006; Hoffmann, 2022; Kühl et al., 2020; McElwee, 2021; Oliver et al., 2015; Román-Palacios and Wiens, 2020; The Royal Society and the US National Academy of Sciences, 2020). Spatial indices of tree diversity often serve as surrogates for more direct measures of biodiversity, but also provide strategic information on forest structure which is crucial to sustainable forest management (Aguirre et al., 2003; Gadow, 1993; Neumann and Starlinger, 2001; Pommerening and Grabarnik, 2019). To save costs, the monitoring of tree diversity is often combined with existing programmes or systems of forest inventory (Motz et al., 2010; Sterba, 2008) which raises the question of appropriate estimators, particularly when the measures of tree diversity are spatially explicit and the forestry characteristics estimated in the same surveys are not.

Our study confirmed that estimating spatial diversity indices from relascope sample data leads to varied outcomes, since each tree included in the sample – subject to its size – has a separate and unique, circular inclusion zone. To add more complication, spatial tree diversity indices are based on the nearest-neighbour principle and following standard forest inventory protocols the neighbours of sample trees are usually not explicitly recorded in the sample. Furthermore, forest structure clearly impacts the sampling error, which was also reported by Häbel et al. (2019), and so does the algorithm of the diversity index. If possible, simple circular sample plots should therefore be preferred when monitoring spatial diversity indices.

However, often the data analysist involved in monitoring has no choice or there are convincing arguments for using relascope samples. For example, as mentioned before, the sampling of tree diversity indices is often added to existing forest inventories (Motz et al., 2010; Sterba,

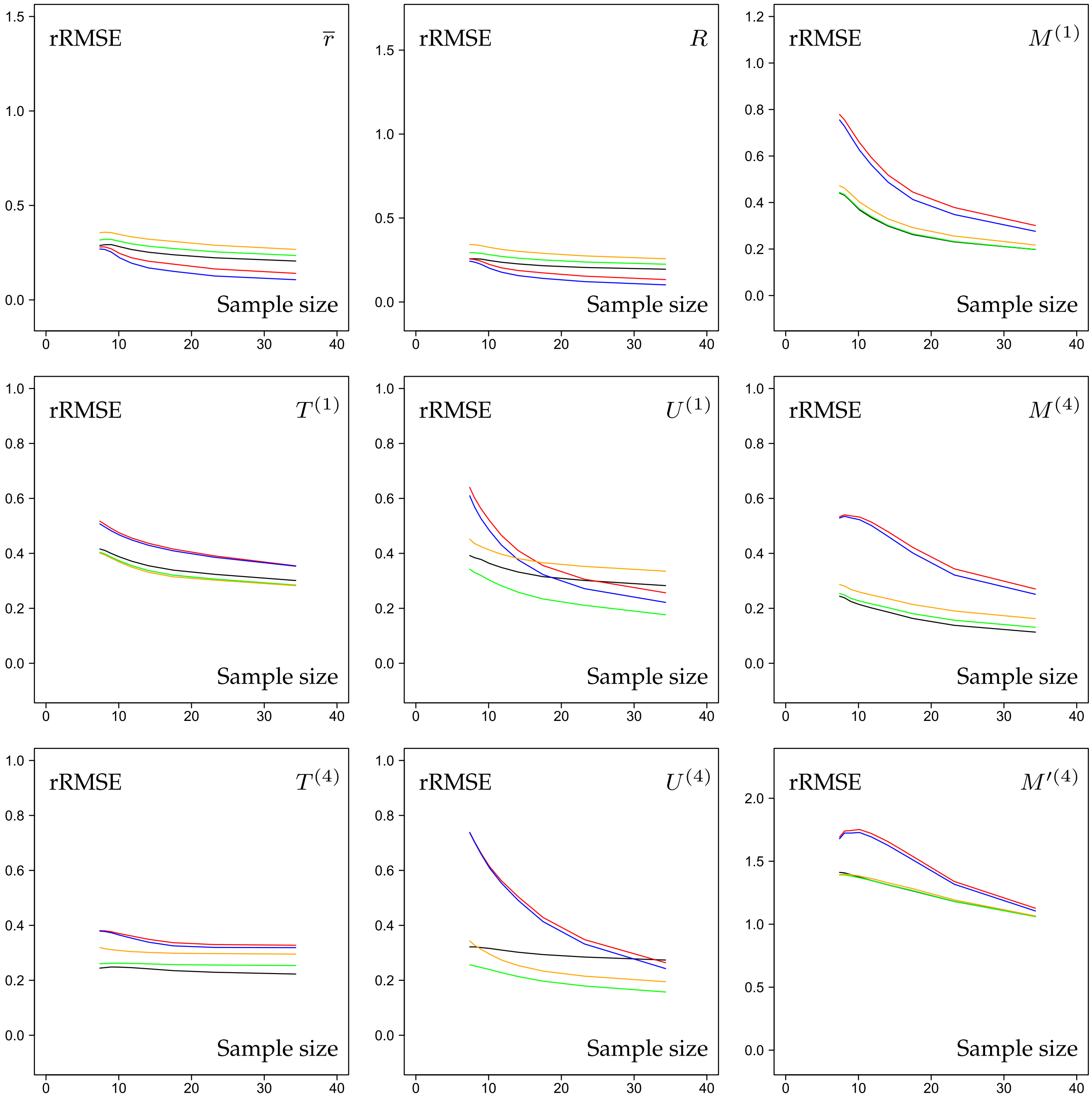

Fig. 4. Relative root mean square error (rRMSE, Eq. 13) of the diversity indices of Table 1 over sample size, i.e. the mean number of sample trees per relascope sample plot in the Rosalia forest. Black – inventory estimator (Eq. 7), red – NN1 estimator (Eq. 9), blue – NN2 estimator (Eq. 8), orange – NN3 estimator (Eq. 10), green – NN4 estimator (Eq. 11). The upper index in round brackets indicates the number of nearest neighbours, *k*, used to estimate the corresponding diversity index. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2008) and nested sample plot methods as well as relascope sampling provide a number of efficiency benefits that simple circular sample plots cannot offer (Häbel et al., 2019). In that case it is good practice to sample off-plot nearest neighbours of sample trees as well (plus-sampling) or alternatively to determine the diversity indices in-situ, i.e. in the field, thus automatically involving off-plot nearest neighbours (Motz et al., 2010). For logistic reasons or when existing (historic) inventory data are re-analysed for diversity indices, plus-sampling is often not possible. In such situations it is generally good advice to estimate only simple diversity indices which exclusively rely on a single source of information, since the sampling errors of components of diversity indices potentially add up to become very large when combined, such as in the case of weighted species mingling index, M' (Eq. 3 in Table 1). In our study, small and large numbers of neighbours, *k*, did not make a big difference in terms of the sampling error. There was even a case, where the rRMSE values associated with $T^{(4)}$ (Eq. 4 in Table 1) were lower than those related to $T^{(1)}$ (Fig. 4). However, it may be prudent not to extend *k* beyond 4, particularly when the basal area factor is 4 or larger.

NN1 and NN2 estimators (Pommerening and Stoyan, 2006) are generally associated with a low bias, although they were not originally designed for relascope sampling. Our results suggest that they can be considered safe options, since they are edge-correction methods that are more firmly based on statistical theory. Therefore, for indices directly related to distances, according to our results the NN2 estimator should be used. Unfortunately both NN1 and NN2 estimators have the disadvantage that they exclude the diversity index values of some sample trees. Since the number of sample trees per sample point if often small, such exclusions increase the index variance and thus the rRMSE in relascope-based forest inventories. Our simulation results highlighted that the variance between sample plots constitutes the largest contribution to sampling errors. This variance is likely to increase in forests managed for continuous cover forestry or similar concepts

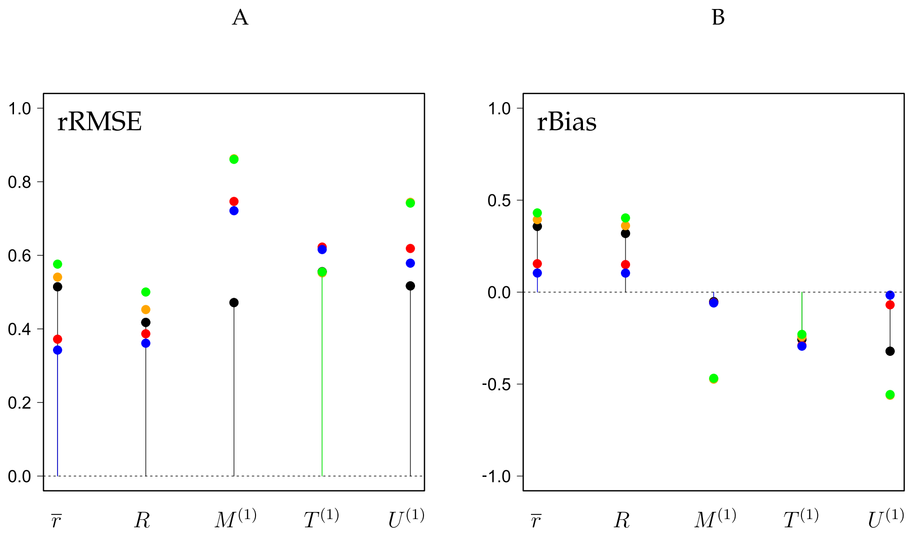

Fig. 5. Relative root mean square error (rRMSE, Eq. 13, A) and relative Bias (rBias, Eq. 14, B) of the diversity indices of Table 1 (only the first neighbour of sample trees could be considered) in the Kammer large-scale forest inventory from 2006. Black – inventory estimator (Eq. 7), red – NN1 estimator (Eq. 9), blue – NN2 estimator (Eq. 8), orange – NN3 estimator (Eq. 10), green – NN4 estimator (Eq. 11). The vertical lines lead to the lowest and to the rRMSE or rBias values of the inventory estimator. The basal area factor was $v = 4$ m^2 ha^{-1} and the mean number of sample trees per sample point was 7.8. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Pommerening, 2023), as structural complexity is an important objective of this type of forest management. Compared to variance, the contribution of bias to the values of rRMSE is rather small. This is where the NN3 and NN4 estimators can help, since they are effective in reducing the overall index variation. They have performed well with all diversity indices other than distance and the aggregation index. In this context, it is particularly interesting that the NN4 estimator, which is based on the weights of both sample trees and of the neighbours of sample trees, has performed best with most diversity indices other than distance and aggregation index. Given the general aim to reduce the index variance, this finding seems to suggest that the information provided by sample-tree neighbours complements that of the sample trees and is with most indices more important than spatial edge-bias concerns that are explicitly corrected by NN1 and NN2. This is interesting and strategic information prompting future research to continue along these lines and to optimise estimators which result in low bias and low rRMSE values by balancing the weights of sample trees with those of their neighbours.

Author contributions

AP and HS conceived the project idea and designed the methodology. HS secured and curated the data used in this project. AP and HS jointly carried out the data analysis. AP and HS wrote the first manuscript draft.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or non-profit sectors.

CRediT authorship contribution statement

Arne Pommerening: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Hubert Sterba:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors did not use any AI or AI-assisted technologies.

Declaration of competing interest

The authors declare that they have no known financial interests or personal relationships that could have influenced the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Mari Myllymäki (LUKE, Finland) for inspiring and helpful discussions on the matter of this article.

References

Aguirre, O., Hui, G.Y., Gadow, K., Jiménez, J., 2003. An analysis of spatial forest structure using neighbourhood-based variables. For. Ecol. Manag. 183, 137–145.

Baddeley, A., Rubak, E., Turner, R., 2016. Spatial point patterns. Methodology and applications with R. CRC Press, Boca Raton, p. 810.

Barabesi, L., Fattorini, L., 1999. Bitterlich sampling estimators of diversity indexes. Stat. Appl. 11, 453–463.

Bitterlich, W., 1948. Die Winkelzählprobe. [Relascope sampling.] Allgemeine Forst- und Holzwirtschaftliche Zeitun, 60, pp. 41–42.

Bitterlich, W., 1984. The Relascope Idea. Commonwealth Agric, Bureaux, p. 242.

Clark, P.J., Evans, F.C., 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. Ecology 35, 445–453.

Cochran, W.G., 1977. Sampling techniques, 3rd edition. John Wiley & Sons, New York, p. 428.

Fischer, J., Lindenmayer, D.B., Manning, A.D., 2006. Biodiversity, ecosystem function, and resilience: ten guiding principles for commodity production landscapes. Front. Ecol. Environ. 4, 80–86.

Gadow, K.V., 1993. Zur Bestandesbeschreibung in der Forsteinrichtung. [New variables for describing stands of trees.]. Forst und Holz 48, 602–606.

Gadow, K.v., Álvarez González, J.G., Zhang, C., Pukkala, T., Zhao, X., 2021. Sustaining forest ecosystems. In: Managing Forest Ecosystems, 37. Cham, Springer Nature, p. 419.

Gregoire, T.G., Valentine, H.T., 2008. Sampling Strategies for Natural Resources and the Environment. Chapman & Hall/CRC, Boca Raton, p. 474.

Grosenbaugh, L.R., 1958. Point sampling and line-sampling; probability theory, geometric implications, synthesis. In: Occasional Paper, Southern Forest Experiment Station No. 145, p. 27.

Häbel, H., Kuronen, M., Henttonen, H., Kangas, A., Myllymäki, M., 2019. The effect of spatial structure of forests on the precision and costs of plot-level forest resource estimation. Forest Ecosyst. 6, 8.

Hanisch, K.-H., 1984. Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point processes. Math. Operationsforschung Statistik-Stat. 15, 409–412.

Hoffmann, S., 2022. Challenges and opportunities of area-based conservation in reaching biodiversity and sustainability goals. Biodivers. Conserv. 31, 325–352.

Horvitz, D.G., Thompson, D.J., 1952. A generalization of sampling without replacement from a finite universe. J. Am. Stat. Assoc. 47, 663–685.

Hui, G., Albert, M., 2004. Stichprobensimulationen zur Schätzung nachbarschaftsbezogener Strukturparameter in Waldbeständen [Estimating neighborhood-based stand structure parameters using sampling simulation. Allgemeine Forst- und Jagdzeitung 175, 199–209.

Hui, G.Y., Albert, M., Gadow, K.V., 1998. Das Umgebungsmaß als Parameter zur Nachbildung von Bestandesstrukturen. [Diameter dominance as a parameter for simulating forest structure.]. Forstwissenschaftliches Centralblatt 117, 258–266.

Hui, G., Hu, Y., Zhao, Z., 2008. Evaluating tree species segregation based on neighbourhood spatial relationships. J. Beijing Forest. Univ. 30, 131–134.

Hui, G., Zhao, X., Zhao, Z., Gadow, K.V., 2011. Evaluating tree species spatial diversity based on neighborhood relationships. For. Sci. 57, 292–300.

Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. John Wiley & Sons, Chichester, p. 534.

Kühl, H.S., Bowler, D.E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, D., Eisenhauer, N., Fernández, N., Guerra, C.A., Henle, K., Herbinger, I., Isaac, N.J.B., Jansen, F., König-Ries, B., Kühn, I., Nilsen, E.B., Pe'er, G., Richter, A., Schulte, R., Settle, J., van Dam, N.M., Voigt, M., Wägele, W.J., Wirth, C., Bonn, A., 2020. Effective biodiversity monitoring needs a culture of integration. One Earth 3, 462–474.

Mandallaz, D., 2008. Sampling Techniques for Forest Inventories. Chapman & Hall/CRC, Boca Raton, p. 256.

Mandallaz, D., Lanz, A., 2001. Further results for optimal sampling schemes based on the anticipated variance. Can. J. For. Res. 31, 1845–1853.

Matias, M.G., Combe, M., Barbera, C., Mouquet, N., 2013. Ecological strategies shape the insurance potential of biodiversity. Front. Microbiol. 3, 432.

Mauro, F., Haxtema, Z., Temesgen, H., 2017. Comparison of sampling methods for estimation of nearest-neighbor index values. Can. J. For. Res. 47, 703–715.

McElwee, P., 2021. Climate change and biodiversity loss: two sides of the same coin. In: Current History November 2021, pp. 295–300.

Motz, K., Sterba, H., Pommerening, A., 2010. Sampling measures of tree diversity. For. Ecol. Manag. 260, 1985–1996.

Neumann, M., Starlinger, F., 2001. The significance of different indices for stand structure and diversity in forests. For. Ecol. Manag. 145, 91–106.

Oliver, T.H., Heard, M.S., Isaac, N.J.B., Roy, D.B., Procter, D., Eigenbrod, F., Freckleton, R., Hector, A., Orme, C.D.L., Petchey, O.L., Proença, V., Raffaelli, D., Suttle, K.B., Mace, G.M., Martín-López, B., Woodcock, B.A., Bullock, J.M., 2015. Biodiversity and resilience of ecosystem functions. Trends Ecol. Evol. 30, 673–684.

Pommerening, A., 2023. Continuous Cover Forestry. Theories, Concepts & Implementation. John Wiley & Sons, Chichester, p. 416.

Pommerening, A., Grabarnik, P., 2019. Individual-Based Methods in Forest Ecology and Management. Springer Nature, Cham, p. 411.

Pommerening, A., Stoyan, D., 2006. Edge-correction needs in estimating indices of spatial forest structure. Can. J. For. Res. 36, 1723–1739.

R Development Core Team, 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org.

Reininger, H., 2000. Das Plenterprinzip. [The selection principle.]. Leopold Stocker Verlag, Graz, p. 238.

Roesch, F.A., 2008. An alternative view of continuous forest inventories. For. Sci. 54, 455–464.

Román-Palacios, C., Wiens, J.J., 2020. Recent responses to climate change reveal the drivers of species extinction and survival. Proc. Natl. Acad. Sci. USA 117, 4211–4217.

Schütz, J.-Ph., 2001. Der Plenterwald und weitere Formen strukturierter und gemischterWälder [The selection forest and other forms of structured and mixed forests]. Parey Buchverlag, Berlin, p. 207.

Sterba, H., 2004. Equilibrium curves and growth models to deal with forests in transition to uneven-aged structure – application in two sample stands. Silva Fennica 38, 413–423.

Sterba, H., 2008. Diversity indices based on angle count sampling and their interrelationships when used in forest inventories. Forestry 81, 587–597.

The Royal Society and the US National Academy of Sciences, 2020. Climate change. In: Evidence & causes. Update 2020. An overview from the Royal Society and the US National Academy of Sciences. Report, p. 24.

van Laar, A., Akça, A., 2007. Forest mensuration. In: Managing Forest Ecosystems, 13. Springer, Dordrecht, p. 383.

Wang, H., Zhang, X., Hu, Y., Pommerening, A., 2021. Spatial patterns of correlation between conspecific species and size diversity in forest ecosystems. Ecol. Model. 457, 109678.

Yachi, S., Loreau, M., 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. Proc. Natl. Acad. Sci. USA 96, 1463–1468.