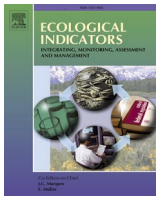
Ecological Indicators 166 (2024) 112567

Contents lists available at ScienceDirect

Ecological Indicators

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



Original Articles

Plant size inequality: How to analyse competing spatial dissimilarity indices?

Arne Pommerening [a,*], Aila Särkkä [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] Department of Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, 412 96 Gothenburg, Sweden

ARTICLE INFO

Keywords:
Biodiversity
Plant size differentiation
Size inequality
Mark correlation function
Monte Carlo testing
Negative/positive autocorrelation

ABSTRACT

Worldwide biodiversity loss is perceived as a major threat and is likely to be enforced by ongoing climate change. Continued monitoring of biodiversity can assist in compensating for decreasing biodiversity by goal-oriented conservation management. Plant size diversity, also termed size inequality and size hierarchy, has often been neglected in studies of biodiversity, but plays a crucial role in many ecosystems, e.g. in forests. Several competing diversity indices and other characteristics of spatial plant size inequality have been proposed but to date no protocols or guidelines exist for evaluating their relative merits. In our study, we proposed a broad framework for such an analysis. In order to put this framework to a test, we revisited the dissimilarity coefficient, a somewhat ignored but promising spatial size inequality index published at the end of the 1990s, identified its nearest index competitors and analysed their performances in different mathematical-statistical contexts. We learned that the dissimilarity coefficient is more sensitive in statistical significance tests relating to three very different summary characteristics than its closest competitor, the size differentiation index. In addition the dissimilarity coefficient has a more solid foundation than the differentiation index, since it is based on the well-known coefficient of variation. The dissimilarity coefficient can be recommended in a wide range of plant diversity applications. The principle of the dissimilarity coefficient can also be used to define an effective mark correlation function. We recommend our evaluation framework for use in similar analyses of competing diversity characteristics.

1. Introduction

With ongoing climate change the continued monitoring of biodiversity is crucial to mitigate the loss of species and size diversity in terrestrial and aquatic ecosystems (Banks-Leite et al., 2020). At the same time plant diversity usually also safeguards ecosystem resilience. In this context, spatial plant diversity indices are important, because they can be measured with comparatively little effort whilst acting as surrogate measures of biodiversity (Pommerening and Grabarnik, 2019).

In the past, structural diversity of plant communities has mainly been studied in terms of species diversity (Gaston and Spicer, 2004; Wang et al., 2020). Contrary to its importance spatial size diversity, also referred to as *size hierarchy* and *size inequality* (Weiner and Solbrig, 1984), has not been considered much in research. Consequently the focus in our study was on size inequality.

There are many ways to quantify spatial plant diversity and new characteristics are published all the time whilst quite a few older approaches are nearly forgotten although they clearly have their merits (Magurran, 2004). Usually new and old characteristics are listed or referred to in publications side by side without an attempt to identify differences in performance. Also, many new and old approaches seemingly quantify similar aspects of plant diversity, e.g. *location diversity* (*dispersion*), *species diversity* or *size diversity* (Pommerening and Grabarnik, 2019), i.e. they appear to compete with one another. This begs the question of how such competing characteristics can be analysed in a meaningful way so that advantages and disadvantages become apparent. To date there is not much of a protocol or consistent methodology for carrying out such comparative analyses.

Torquato et al. (2002) and Crawford et al. (2003) proposed the intriguing method of reconstruction for evaluating the merits of competing spatial characteristics. Reconstruction draws on stochastic optimisation techniques that generate realisations of spatial plant patterns. This approach is based on the hypothesis that a given characteristic carries more information about a pattern under study if it is able to

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

https://doi.org/10.1016/j.ecolind.2024.112567
Received 18 June 2024; Received in revised form 26 July 2024; Accepted 29 August 2024
Available online 10 September 2024

contribute to a more precise reconstruction of this pattern than another, competing characteristic. We have successfully used the reconstruction method ourselves in a variety of contexts (Pommerening and Stoyan, 2008; Pommerening et al., 2019; Wang et al., 2021). Reconstruction assumes that in addition to competing indices or measures another or several other characteristics exist which are so accurate that they can be applied independently for judging the quality of reconstruction. This assumption, however, may not be true or it may be hard to find such characteristics. In the past, mark correlation functions as introduced in Section 3.3 were often used for that purpose in the context of plant size inequality, since they are known to be comparatively accurate. However, residual doubts naturally remain.

Pommerening et al. (2024) analysed how competing tree diversity indices performed in terms of bias and sampling error when estimating them through distance sampling. This is another way of analysing comparative properties of competing diversity characteristics. The sampling performance of competing diversity indices helps understand how they interact with specific sampling designs and provides decision support in practical sampling applications in terms of which index to prefer from a list of competitors (Pommerening et al., 2024).

In this study, we decided to explore alternative avenues. As study object we identified the *dissimilarity coefficient*, a spatial size inequality index which was briefly published by Ali (1997) and Hagner and Nyquist (1998) and then largely disappeared from literature records. The objective of this paper is to (1) fully characterise the dissimilarity coefficient, (2) to explore how it can be applied, to (3) identify its competitor(s) and (4) to establish the relative merits of the dissimilarity coefficient and its nearest competitor through a combination of simulations and data applications.

2. Materials and methods

2.1. Dissimilarity coefficient

The dissimilarity coefficient is a spatial measure of size inequality and was first proposed by Ali (1997) and by Hagner and Nyquist (1998). The index is based on the coefficient of variation of the pairs of sizes of subject plant i and neighbouring plants j, i.e.

$$V_i = \frac{1}{k}\sum_{j=1}^{k} \frac{s(m_i, m_j)}{\overline{m}} \tag{1}$$

In Eq. (1), $s(m_i, m_j)$ is the standard deviation of plant sizes m_i and m_j. Denominator $\overline{m} = 0.5 \times (m_i + m_j)$ is the arithmetic mean of the two size marks of a given pair. The number of nearest neighbours is denoted by k and originally Ali (1997) and Hagner and Nyquist (1998) defined $k = 1$, but from a general multivariate perspective, any other number such as $k = 4$ is also possible. The dissimilarity coefficient used the *size difference construction principle* of spatial diversity indices and test functions (Illian et al., 2008; Pommerening and Grabarnik, 2019). Variance computed from a pair of plant sizes can be written as

$$\begin{aligned} s^2(m_i, m_j) &= \frac{(m_i - \overline{m})^2 + (m_j - \overline{m})^2}{2-1} = (m_i - \overline{m})^2 + (m_j - \overline{m})^2 \\ &= \left(m_i - \frac{(m_i + m_j)}{2}\right)^2 + \left(m_j - \frac{(m_i + m_j)}{2}\right)^2 \\ &= \left(\frac{m_i - m_j}{2}\right)^2 + \left(\frac{m_i - m_j}{2}\right)^2 \\ &= \frac{1}{4}(m_i - m_j)^2 + \frac{1}{4}(m_i - m_j)^2 \\ &= \frac{1}{2}(m_i - m_j)^2 \end{aligned} \tag{2}$$

Hence

$$s(m_i, m_j) = \sqrt{\frac{1}{2}(m_i - m_j)^2} = \frac{|m_i - m_j|}{\sqrt{2}} \tag{3}$$

We decided to give all computations in Eq. (2), since some steps were missing in the original publications, which may impair the appreciation of this index (Ali, 1997; Hagner and Nyquist, 1998), and the details provided emphasise the solid mathematical-statistical basis of the diversity index. Using Eqs. (1) and (3), the coefficient of variation computed from a single pair of plant sizes (indicated by upper index (1)) is then given as

$$\begin{aligned} V_i^{(1)} &= \frac{s(m_i, m_j)}{\overline{m}} = \frac{\frac{|m_i - m_j|}{\sqrt{2}}}{\overline{m}} = \frac{\frac{|m_i - m_j|}{\sqrt{2}}}{\frac{(m_i + m_j)}{2}} = \frac{2 \times \frac{|m_i - m_j|}{\sqrt{2}}}{m_i + m_j} = \frac{(\sqrt{2})^2 \times \frac{|m_i - m_j|}{\sqrt{2}}}{m_i + m_j} \\ &= \frac{\sqrt{2} \times |m_i - m_j|}{m_i + m_j} \end{aligned} \tag{4}$$

In general, for k nearest neighbours, i.e. for k pairs of plants, Eq. (1) can now be written as

$$V_i = \frac{1}{k}\sum_{j=1}^{k} \frac{s(m_i, m_j)}{\overline{m}} = \frac{1}{k}\sum_{j=1}^{k} \frac{\sqrt{2} \times |m_i - m_j|}{m_i + m_j} \tag{5}$$

Index values of V_i lie between 0 and 1 and Hagner and Nyquist (1998) interpreted the term $\sqrt{2}$ as scale factor leading to a better spread of index values between these two boundaries. When one plant size of a pair of plants is zero, i.e. $V_i^{(1)} = \sqrt{2}$, and when the difference between the sizes of pairs of plants is large relative to their sum, i.e. when the ratio $\frac{|m_i - m_j|}{m_i + m_j}$ exceeds 0.70, the corresponding $V_i^{(1)}$ (Eq. (4) is larger than 1. This is probably the reason why Hagner and Nyquist (1998) advocated a simplified version of Eq. (5), where $\sqrt{2}$ is replaced by 1, i.e.

$$V_i' = \frac{1}{k}\sum_{j=1}^{k} \frac{|m_i - m_j|}{m_i + m_j} \tag{6}$$

Individual values of $V_i'^{(1)}$ are always smaller than those of $V_i^{(1)}$ and they have the advantage of never exceeding 1. In our experience when using tree stem diameters as size variables, cases of $V_i^{(1)} > 1$ are rare and only occur in individual pairs but usually not in V_i. However, if in any application many plant sizes have the value of zero or the differences of neighbouring sizes are often large, V_i' should possibly be preferred to V_i, since index values between 0 and 1 are easier to interpret and more straightforward to compare with other indices.

Incidentally, the fracture term of Eq. (6) is also known from animal movement ecology where distances between moving animals i and j are used for size variable m in sociality and interaction indices (Fronville et al. 2023).

2.2. Interpretation and properties of the dissimilarity coefficient

V_i and V_i' are close to zero when both plants of a pair are of similar size. When one plant is double the size of the other in a pair, $V_i^{(1)} = 0.4714$ and $V_i'^{(1)} = \frac{1}{3}$. V_i and V_i' tend to $\sqrt{2}$ and one, respectively, as the difference between plant sizes increases. V_i and V_i' are scale invariant, i.e. if plant sizes m are multiplied by a constant, the values of V_i and V_i' remain the same.

2.3. Application of the dissimilarity coefficient

After calculating individual-plant indices V_i, the arithmetic mean for a whole plant population can be computed as an estimation of population dissimilarity:

$$\overline{V} = \frac{1}{N}\sum_{i=1}^{N} \frac{\sqrt{2}}{k}\sum_{j=1}^{k} \frac{|m_i - m_j|}{m_i + m_j} \tag{7}$$

The number of all plants in a population is denoted as N. In analogy to

Eq. (7), $\overline{V}$ can be calculated, however, $\sqrt{2}$ then needs to be replaced by 1 as before. Either m_i or m_j of a given pair of plants can be 0, as long as not both of them are 0. When computing mean size dissimilarity for a particular species population, the values of V_i of the whole population (regardless of species) are simply a subset for that species:

$$\overline{V}_a = \frac{1}{N_a}\sum_{i=1}^{N_a}\frac{\sqrt{2}}{k}\sum_{j=1}^{k}\frac{|m_i - m_j|}{m_i + m_j} \tag{8}$$

Here, index a denotes a theoretical species a occurring in a plant population alongside others. For example, N_a is the number of plants of species a. This definition usually implies that all plants i must be of species a, but their k nearest neighbours j can be of any species (Pommerening and Grabarnik, 2019). The neighbourhood definition implied in Eq. (8) relies on "natural" neighbourhoods defined by spatial proximity alone. Theoretically another option of calculating conspecific means is to subset for species a in such a way that also the k neighbours of a given plant i have to be of the same species a (*intratype analysis*; Lotwick and Silverman, 1982). This can be achieved by excluding all plants that are not of species a from the conspecific analysis. Such an intratype strategy would, however, define a very different neighbourhood and address different eco-physiological questions.

Apart from population means it is also possible to consider density distributions of individual-plant V_i or V_i' as shown by Pommerening et al. (2020). Another option for analysis is to calculate $\overline{V}$ and $\overline{V}'$ separately for several numbers of nearest neighbours k and then to compose a function

$$V(r) = \begin{cases} \overline{V}(k) & \text{for } r = \bar{r}_k, k = 1,2,3,\cdots, \\ \overline{V}(k) + \dfrac{\overline{V}(k+1) - \overline{V}(k)}{\bar{r}_{k+1} - \bar{r}_k} \times (r - \bar{r}_k) & \text{for } \bar{r}_k < r < \bar{r}_{k+1}, k = 1,2,3,\cdots \end{cases} \tag{9}$$

In Eq. (9), $\bar{r}_k$ is the population mean distance $\bar{r}_k = \frac{1}{N}\sum_{i=1}^{N} r_{ik}$ between any individual i and its kth nearest neighbour. $\overline{V}(k)$ is the set of means of V_i (Eq. (7)) calculated for increasing k. Essentially, Eq. (9) assembles $\overline{V}$ calculated iteratively from Eq. (7) with $k = 1,2,3,\cdots$ in a continuous curve with a linear interpolation between the different means. Since different numbers of k can involve quite different distances between plants i and neighbours k, k is translated to distance r in Eq. (9) for better comparison between plant populations. A similar function can also be defined for conspecific plant communities, i.e. for the specimens of only one species. Function $V(r)$ describes how size dissimilarity changes with increasing neighbourhood.

In many situations when population means are calculated, it is recommended to apply methods of spatial edge-bias compensation (Pommerening and Stoyan, 2006). This is particular important for small numbers of plants and for large k.

2.4. Expected dissimilarity coefficient

The expected dissimilarity coefficient, $\mathbf{E}V$, is independent of the number of nearest neighbours, k. This measure is based on a plant population's size distribution and describes the mean dissimilarity coefficient when all plant sizes are spatially completely (independently) dispersed without any spatial correlation. It is useful to consider $\mathbf{E}V$, because this quantity can serve as a reference to understand to what degree ecological processes have influenced the spatial plant pattern so that observed $\overline{V}$ differs from $\mathbf{E}V$. Expected size dissimilarity can be calculated for any plant population as

$$\mathbf{E}V = \frac{\sqrt{2}}{N(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{|m_i - m_j|}{m_i + m_j} \tag{10}$$

Again, $\mathbf{E}V'$ can be computed in analogy to Eq. (10) by simply replacing $\sqrt{2}$ by 1. $\mathbf{E}V$ and $\mathbf{E}V'$ can be approximated by random labelling simulations (Illian et al, 2008), i.e. by a randomisation or permutation of plant sizes. For example, when permuting plant sizes, i.e. when randomly assigning them to the fixed locations of plants of an observed population, e.g. 9999 times, thus simulating complete spatial independence, the arithmetic means of the population means of 9999 simulations, $\overline{\overline{V}}$ or $\overline{\overline{V}}'$, are very close to $\mathbf{E}V$ or $\mathbf{E}V'$. Thus considering random labelling simulations offers another way to understand the meaning of Eq. (10).

Mean population and expected plant size dissimilarity can be combined in a size segregation index (Pommerening and Uria-Diez, 2017), Υ, providing more detailed information on the difference between observed and expected spatial size inequality patterns:

$$\Upsilon = 1 - \frac{\overline{V}}{\mathbf{E}V} \tag{11}$$

If plant sizes are independently dispersed without any spatial correlation, $\Upsilon = 0$. If the sizes of neighbouring plants are always of similar size, $\Upsilon \approx 1$, i.e. there is a spatial attraction of similar sizes leading to a segregation or clustering of sizes. If all neighbours tend to have sizes quite different of that of plant i, Υ is negative and tends towards -1 in the extreme case (spatial attraction or aggregation of different sizes).

Based on the definition of conspecific means given in Section 2.3, expected size dissimilarity of species community a is given as

$$\mathbf{E}V_a = \frac{\sqrt{2}}{N_a(N-1)}\sum_{i=1}^{N_a}\sum_{j\neq i}^{N}\frac{|m_i - m_j|}{m_i + m_j} \tag{12}$$

Similar to the analyses of the whole population regardless of species, it is possible to define a conspecific size segregation index Υ_a for any species community a. Functions of distance r similar to Eq. (9), i.e. $\Upsilon(r)$ and $\Upsilon_a(r)$, can be constructed from both Υ and Υ_a (Wang et al., 2020) and they describe how the effect of spatial correlation between plant sizes declines with distance.

2.5. Mark correlation functions

Mark correlation functions are second-order characteristics. They depend on a distance variable r and quantify structural properties of spatial plant patterns based on the probability that two plant locations are r distance away from each other (Illian et al., 2008; Pommerening and Grabarnik, 2019). This allows second-order characteristics to be related to various ecological scales and to account for short, medium and long-range plant size interactions. Due to this scale dependency second-order characteristics are often considered less ambiguous and more precise than spatial diversity indices, if the data they are applied to are of sufficient quality and quantity. The mark correlation function quantifies the similarity and dissimilarity of pairs of plant sizes at a given distance r (Penttinen et al., 1992; Illian et al., 2008; Pommerening and Grabarnik, 2019). The general form of the estimator of the mark correlation functions is:

$$k_t(r) = \frac{1}{\mathbf{E}t}\sum_{\xi_i,\xi_j\in W}^{\neq}\frac{t\big(m(\xi_i), m(\xi_j)\big) \times k_h(\|\xi_i - \xi_j\| - r)}{2\pi r \times A\left(W_{\xi_i} \cap W_{\xi_j}\right)} \tag{13}$$

Here ξ_i and ξ_j are two arbitrary plant locations of the spatial plant pattern in observation window W. k_h is a kernel function, $A\left(W_{\xi_i} \cap W_{\xi_j}\right)$ is the area of intersection of W_{ξ_i} and W_{ξ_j}, see Illian et al. (2008), relating to the translation edge correction (Ohser and Stoyan, 1981). Function k_t is usually standardised using an expected value of the test function of plant sizes $m(\xi_i)$ and $m\left(\xi_j\right)$, $\mathbf{E}t$ (similar to Eqs. (10) and (12), corresponding to test function t.

The key element of any mark correlation function is test function $t\left(m(\xi_i), m\left(\xi_j\right)\right)$ in the numerator of Eq. (13), quantifying the similarity

or dissimilarity of plant sizes m at locations ξ_i and ξ_j. This test function can be defined in a way similar to spatial inequality indices. In this study we selected the following options:

$$t_1\big(m(\xi_i), m(\xi_j)\big) = \sqrt{2} \times \frac{|m(\xi_i) - m(\xi_j)|}{m(\xi_i) + m(\xi_j)} \tag{14}$$

$$t_2\big(m(\xi_i), m(\xi_j)\big) = 1 - \frac{\min\big(m(\xi_i), m(\xi_j)\big)}{\max\big(m(\xi_i), m(\xi_j)\big)} \tag{15}$$

$$t_3\big(m(\xi_i), m(\xi_j)\big) = \frac{1}{2}\big(m(\xi_i) - m(\xi_j)\big)^2 \tag{16}$$

Obviously, test function t_1 reflects the principle of the dissimilarity coefficient (Eq. (4), test function t_2 that of the differentiation index (Eq. (16) whilst we decided to use the mark variogram $\gamma(r)$ (Wälder and Stoyan, 1996; Stoyan and Wälder, 2000; Pommerening and Särkkä, 2013; defined by test function t_3) as a reference due to its prominence in the literature. We calculated the mark correlation functions using the R spatstat package (Baddeley et al., 2016; R Development Core Team, 2023).

We carried out 2499 random labelling simulations using the mark variogram (including test function t_3) and the mark correlation function based on test functions t_1 and t_2. Then we applied a simultaneous global envelope test (Myllymäki et al., 2017) to the results using the R GET package whilst correcting for multiple testing at the same time.

2.6. Competing spatial inequality indices

Considering the nearest competitors of spatial diversity indices is a way commonly used in spatial statistics for identifying comparative properties and behaviour (Crawford et al., 2003; Torquato, 2002). There are quite a few alternative indices of spatial size inequality. Among these, close competitors of a given index can be identified by similar construction principles and similar purpose. Some indices of spatial size inequality are further removed from the concept of size dissimilarity so that they can be excluded from this study. These include the size dominance index (Aguirre et al., 2003), the hyperbolic tangent index (Pommerening et al., 2020) and the size dominance-differentiation index (Albert, 1999). These three indices are representatives of a subgroup of size inequality indices, the *dominance indices*, which have the special purpose of indicating whether some plants dominate their local neighbourhoods or whether they are rather dominated by their neighbours. They share some similarities with competition indices (Burkhart and Tomé, 2012; Weiskittel et al, 2011). This is contrasted by the subgroup of size diversity indices quantifying the spatial diversity of plants in local neighbourhoods. Size dominance indices tend to be closely correlated with plant growth rates, size diversity indices usually do not have this tendency.

An early measure of spatial size inequality in plants was the size differentiation index introduced by Gadow (1993) as the mean of the ratio of smaller and larger plant sizes m of the k nearest neighbours subtracted from one (Eq. (17). This ratio is based on the *size ratio construction principle* (Illian et al, 2008; Pommerening and Grabarnik, 2019):

$$T_i = 1 - \frac{1}{k}\sum_{j=1}^{k} \frac{\min(m_i, m_j)}{\max(m_i, m_j)} \tag{17}$$

As with the dissimilarity coefficient and other indices, m can be any quantifiable plant size measure, e.g. biomass, weight, carbon, height, stem diameter among others. The value of T_i increases with increasing average size difference between neighbouring trees. $T_i = 0$ implies that neighbouring trees have equal size (Pommerening and Grabarnik, 2019). Occasional sizes $m = 0$ are possible to process with the size differentiation index, as long as no division by zero occurs. T_i takes values between 0 and 1.

Based on the test function of the mark variogram, Pommerening et al. (2011) introduced the size variogram index. This index uses the same size difference construction principle as the dissimilarity coefficient:

$$V_i^* = \frac{1}{2k\widehat{\sigma}_m^2}\sum_{j=1}^{k}\big(m_i - m_j\big)^2 \tag{18}$$

Dividing by size variance, $\widehat{\sigma}_m^2$, estimated from all plant sizes of a given population, is a normalisation easing the interpretation of index values and the comparison between plants and plant communities. The smaller V_i^* the more similar the plant sizes considered are. In this case either both plants are small or both plants are large, no difference is made between these two scenarios (Pommerening et al., 2020). Index V_i^* often takes values larger than 1. Sizes $m = 0$ are acceptable when using this index as long as the variance estimation is not zero.

Finally inspired by the test function of the mark correlation function (Illian et al., 2008; see Section 2.5), Davies and Pommerening (2008) introduced the size correlation index. Realising the *size-product construction principle* (Illian et al, 2008; Pommerening and Grabarnik, 2019), this index compares the mean product of the sizes of a subject plant and its k nearest neighbours with the squared arithmetic mean size, $\widetilde{m}^2$, of all plants in a given plant population:

$$C_i = \frac{m_i\sum_{j=1}^{k} m_j}{k \times \widetilde{m}^2} \tag{19}$$

Values of C_i larger than 1 indicate positive correlation which can be the result of similar-sized large plants at close proximity whilst $C_i < 1$ indicates negative correlation which typically is the result of pairs of plants with large and very small sizes and pairs of plants with small sizes only (Pommerening et al., 2020). Index C_i often takes values larger than 1. Occasional size marks $m = 0$ are acceptable when using this index, but can lead to $C_i = 0$ when $m_i = 0$.

All of these three closest competing indices can be used in ways similar to those of applying the dissimilarity coefficient, which are described in Sections 2.3-2.5.

2.7. Test sensitivity

Monte Carlo tests are a way of testing spatial diversity against a well-defined null hypothesis. In the context of this study the null hypothesis is plant size independence, i.e. a situation where there is no spatial correlation between tree sizes.

We devised an experiment by (1) simulating 1000 Poisson point processes with an overall density $\lambda = 0.035$ trees/m^2 in an observation window of size 120 × 120 m so that spatial edge effects are minor. The points simulated by a Poisson point process are spatially uncorrelated and can represent plant locations (Illian et al., 2008; Pommerening and Grabarnik, 2019). In this study, we assumed tree populations and for each tree location, local tree density was determined. Depending on a chosen quantile q of local tree density, λ^{loc}, (2) the stem diameter of each tree was either sampled from a three-parameter Weibull diameter distribution describing comparatively large trees ($\alpha = 25.1$, $\beta = 22.8$, $\gamma = 2.5$) or from a separate Weibull diameter distribution of comparatively small trees ($\alpha = 3.0$, $\beta = 3.8$, $\gamma = 2.5$): Mimicking natural patterns, small trees were simulated where local density was high and large trees where local density was low. Thus through this method of density *dependent marking*, spatial correlations of tree sizes were introduced to the plant pattern and quantile q of local tree density, λ^{loc}, is the key factor in this process: With very small values of q, only small trees are simulated and with large values of q only large trees are simulated.

In a third step, (3) for each Poisson point pattern with dependent marking, 9999 so-called random labelling simulations were carried out. In each of them, all stem diameters were randomly re-allocated (permuted) to fixed tree locations, as described in Section 2.4. These

simulations typically produce patterns without spatial correlation of tree sizes, i.e. patterns supporting the null hypothesis. The simulation results can be used to calculate the p-values for a two-sided test (Baddeley et al., 2016; Pommerening and Grabarnik, 2019). To establish sensitivity, we studied the p-values for very small ($\lambda^{\text{loc}} \leq 10$ %) and very large values ($\lambda^{\text{loc}} \geq 90$ %) of quantile q of local density to see which of the competitive indices indicated significant spatial inequality patterns in this statistical boundary region. We calculated local density using the spatstat function density, where parameter sigma determining kernel density was set to the value of 2.0 (Baddeley et al., 2016; R Development Core Team, 2023).

2.8. Example data

Clocaenog Forest lies on the southern side of the Denbigh moors, a relatively high dissected plateau. The forest stand (Tyfiant Coed plot 1 at 53° 04′ 29″ N, 003° 25′ 48″ W) included in this study is situated at an altitude of 395 m asl. The underlying solid geology is Silurian made up of slates, shales and grits. The soil is generally fine textured and often quite stony. Podzolic brown earth predominates where site drainage is sufficient. The climate is relatively harsh with cool temperatures and high rainfall. Rainfall is in excess of 1300 mm. The site was planted with Sitka spruce (*Picea sitchensis* (Bong.) Carr.) and lodgepole pine (*Pinus contorta* Dougl. ex Loud.) in 1951, but only Sitka spruce survived. Clocaenog forest area has size 100 × 100 m and was surveyed in 2002 (Pommerening et al., 2024).

Beech (*Fagus sylvatica* L.) plot 41–700 (0.25 ha in size) is part of the Swiss thinning trial at Embrach (longitude: 8° 10′22.13′′, latitude: 47° 22′18.32′′). The plot is in an even-aged beech forest which was established between 1891 and 1905 and re-measured every 5–10 years until 1991. However, spatial information is only available from 1940 onwards and the 1961 survey data were used in this study. A minor species proportion includes oak (*Quercus robur* L.). The plot is located at 590 m a.s.l with a mean annual temperature of 8.3 °C and a mean annual precipitation of 1030 mm (Pommerening and Särkkä, 2013).

Knysna Forest is part of the southernmost patches of the Afromontane forest in South Africa located south of the mountains between Humansdorp and Mossel Bay. Established in the Diepwalle State Forest in 1937, the Knysna Forest represents the largest indigenous forest complex in South Africa. The forest has been taken out of forest management in 1954 and is located to the north of the southern coastal town of Knysna (at about 33° 57′S, 23° 11′E). The forest area involves 25 different species, the most frequent ones include ironwood (*Olea capensis* L. subsp. *macrocarpa*), kamassi (*Gonioma kamassi* E. MEY.) and real yellowwood (*Podocarpus latifolius* (THUNB.) R. BR. ex MIRB.). The study area is situated at 517 m asl and the average annual maximum temperature for the region is 19.2 °C whilst the average minimum is 11.1 °C. The mean annual precipitation may vary between 700 and 1230 mm (Gadow et al., 2016). A large sub-plot of size 116 × 116 m measured in 1972 was included in this study (Pommerening et al., 2024).

Species richness, S, is comparatively high at Knysna Forest and very low for the other two sites. The highest tree density both in terms of number of trees per hectare and basal area occurs at the Embrach monitoring plot followed by Knysna (Table 1). The size range in terms of stem diameter is largest at Knysna and lowest at Clocaenog. Dissimilarity coefficient (Eq. (5)) and diameter differentiation (Eq. (17)) are very similar and so are the corresponding size segregation indices (Eq. (11)). Size diversity is highest in the Embrach and lowest in the Clocaenog monitoring plot. For all three sites and both size inequality indices, the corresponding size segregation indices yielded values near zero, which indicates a trend towards spatial independence, i.e. spatial size correlation is weak. However, at Embrach, there is a stronger trend towards an attraction of similar sizes, whilst at Clocaenog and Knysna there is a weak trend towards and attraction of different sizes (Table 1).

3. Results

3.1. Index correlations

We applied the dissimilarity coefficient and the competing indices to the highly diverse data of the South African Knysna Forest (see Section 2.8). As stated in Section 2.1, the values of V_i' are smaller than those of V_i, a tendency that increases with increasing size dissimilarity, and V_i', values also have a shorter range (Fig. 1).

It is also clear that the strongest relationship of the dissimilarity coefficient with competing indices (listed in Section 2.6) exists for size differentiation index T_i (Eq. (17); Fig. 1A), even though the relationship is nonlinear and weakly hyperbolic. The values of the original index including $\sqrt{2}$ as expansion factor, V_i (Eq. (5)), are apparently more similar to T_i than the simplified index values V_i' (Eq. (6)).

For the other two competing indices, size variogram index V_i^* (Eq. (18)) and size correlation index C_i (Eq. (19)), the relationships with the size dissimilarity coefficient are much noisier than for the size differentiation index. For both indices, variance markedly increases with increasing values of V_i^* and C_i. The relationship with the size variogram index (Fig. 1B) is stronger than that with the size correlation index (Fig. 1C). We also computed all three indices for a range of other data and always found similar relationships between the individual-tree indices (not shown).

Accordingly we have identified size differentiation index T_i as the main competitor of the size dissimilarity coefficient and only proceeded with these two indices in the following analysis.

3.2. Monte Carlo test performance

We tested the performance of mean population size dissimilarity, $\overline{V}$, and mean size differentiation, $\overline{T}$, to detect significant trends in simulated plant patterns with spatial correlations between neighbouring plants (dependent marking). Here the boundary regions defined by quantile q of local tree density were examined to see whether the use of the two indices would lead to a different behaviour of the corresponding p-values (Fig. 2).

In both boundary regions, the curve associated with the size dissimilarity index was always lower than that related to the size differentiation index, particularly for q values leading to increasing p. This suggests that under the same conditions the size dissimilarity index is more sensitive than the size differentiation index. The difference in p-values is largest where p for the dissimilarity index is around 0.5.

Table 1
Area, species richness, S, number of trees, N, basal area, G, minimum stem diameter, $d_{\min}$, maximum stem diameter, $d_{\max}$, mean size dissimilarity, $\overline{V}$, mean size differentiation, $\overline{T}$, and the corresponding size segregation indices, $\Upsilon^{(V)}$ and $\Upsilon^{(T)}$, in the three sample monitoring plots Clocaenog, Embrach and Knysna. The spatial indices were calculated from all trees in the plots using $k = 4$ nearest neighbours and the NN1 spatial edge correction method (Pommerening and Stoyan, 2006).

Forest	Area [ha]	S	N[ha^{-1}]	G[m^2/ha]	$d_{\min}$[cm]	$d_{\max}$[cm]	$\overline{V}$	$\overline{T}$	$\Upsilon^{(V)}$	$\Upsilon^{(T)}$
Clocaenog	1.02	1	281	29.3	20.4	55.5	0.14	0.17	−0.01	−0.01
Embrach	0.25	2	600	32.4	6.7	51.0	0.45	0.45	−0.08	−0.07
Knysna	1.35	20	562	26.0	5.5	66.7	0.42	0.42	0.02	0.02

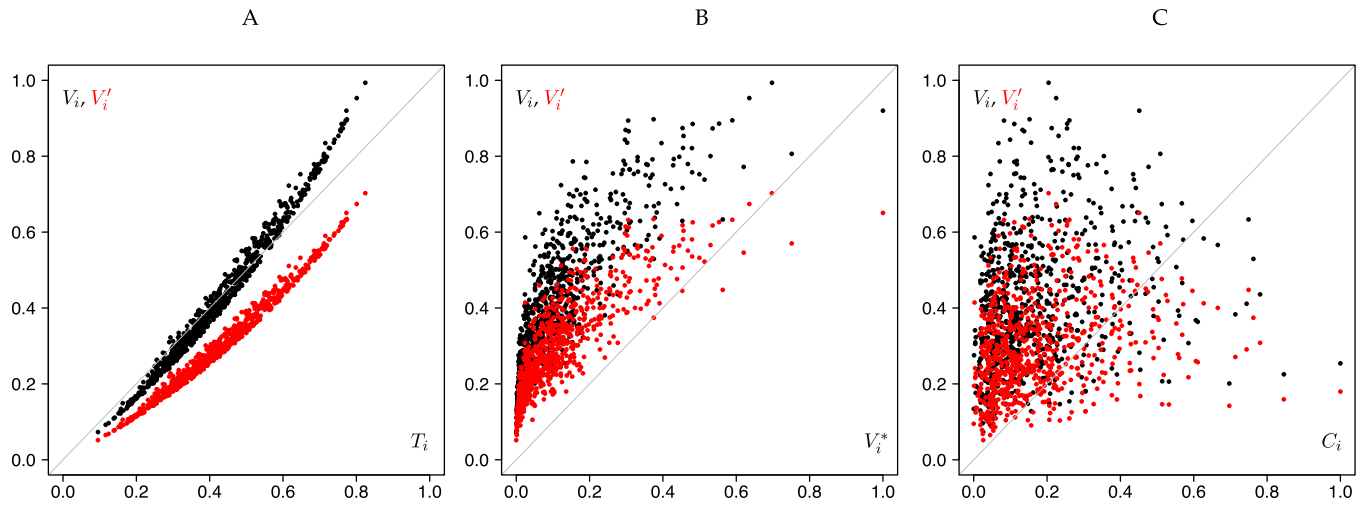


Fig. 1. Scatterplots of individual-plant size dissimilarity indices V_i (black) or V_i' (red) of Eqs. (5) and (6) over (A) the size differentiation index T_i (Eq. (17)), (B) the variogram index V_i^* (Eq. (18)) and (C) the mark correlation index C_i (Eq. (19)) applied to mapped trees at Knysna Forest. For better comparison, the values of the latter two indices were transformed to values between 0 and 1 using the common formula $x_i' = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}$, where x_i represents the index values V_i^* and C_i. Mean population characteristics are $\overline{V} = 0.42$, $\overline{V}' = 0.29$ and $\overline{T} = 0.42$. (For the interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

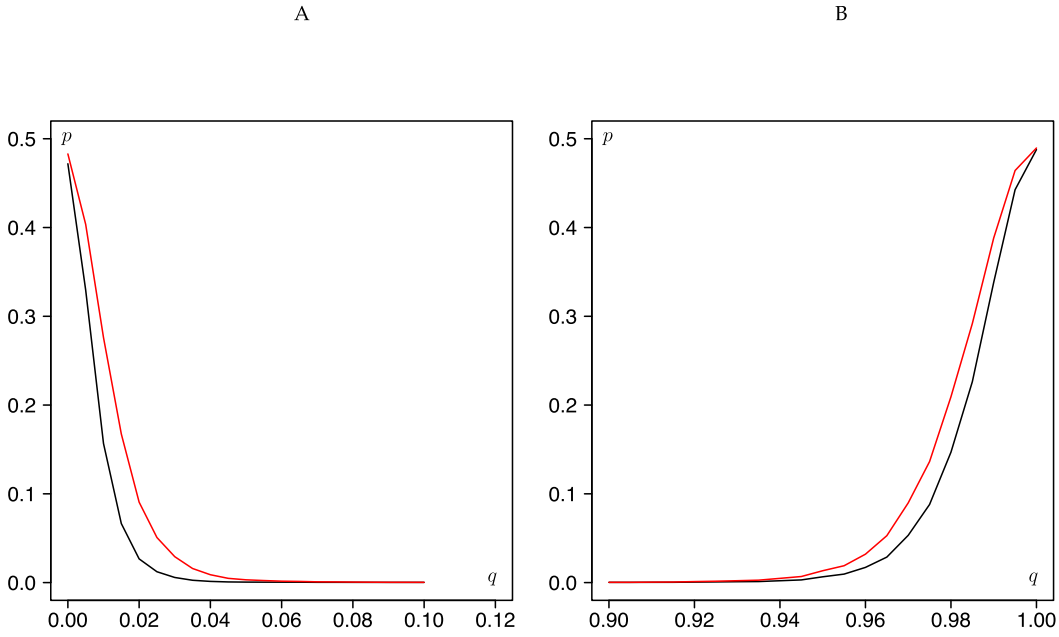


Fig. 2. Results of the Monte Carlo random labelling test applied to density-dependent size marking of 1000 replications of a Poisson point process. The p-value of the two-sided test is denoted as p for $\overline{V}$ (Eq. (7); black) and $\overline{T}$ (Eq. (17); red). The quantile of local tree density is q. A: $q \leq 10$ % of local tree density, B: $q \geq 90$ % of local tree density. (For the interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

3.3. Mark correlation functions

Competing diversity indices can also be used as test functions in mark correlation functions, see Section 2.5. We applied the resulting mark correlation functions $\delta(r)$ (using test function t_1) and $\tau(r)$ (using test function t_2) to the Embrach (Fig. 3A) and Clocaenog data (Fig. 3B; see Section 2.8). As another reference we computed the mark variogram, $\gamma(r)$, using test function t_3.

All three mark correlation functions indicated the same trends in both forests. For the beech forest at Embrach (Fig. 3A) all three mark correlation functions had curves that were larger than 1 up to a distance of $r \approx 7.5$ m, a phenomenon referred to as *negative autocorrelation* (Pommerening and Särkkä, 2013). In this distance range, tree sizes markedly differed at short distances. This was followed by a depression around $r \approx 10$ m where tree sizes of pairs of trees tended to be similar, which is termed *positive autocorrelation* (Illian et al., 2008; Pommerening and Grabarnik, 2019). For larger r there were random fluctuations around 1, which is to be expected.

For the Sitka spruce forest at Clocaenog (Fig. 3B) all three mark correlation functions initially indicated mainly negative autocorrelation at $r \approx 2.5$ m, which was weaker here than at Embrach (Fig. 3A). This first maximum was followed by a depression showing positive autocorrelation at $r \approx 7$ m. After that two smaller maxima indicating negative autocorrelation occurred for larger r followed by random fluctuations around 1.

A comparison of the curves in the two forests shows that $\delta(r)$ and $\tau(r)$

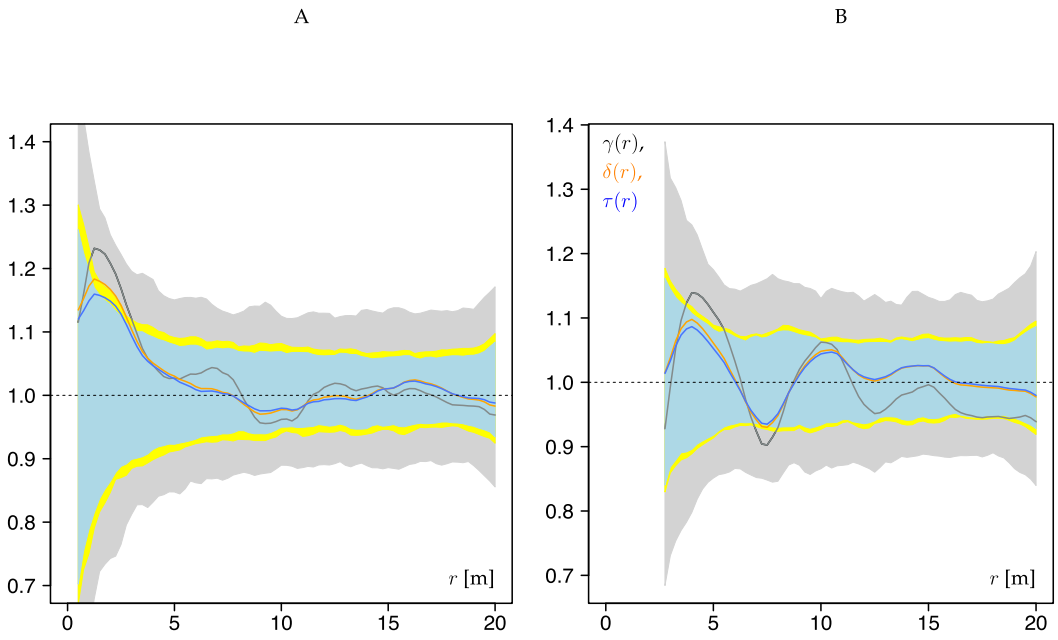


Fig. 3. Mark variogram, $\gamma(r)$, using test function t_3 (Eq. (16)), mark correlation function, $\delta(r)$, (Eq. (13)) using test function t_1 (Eq. (14)) and mark correlation function, $\tau(r)$, (Eq. (13)) using test function t_2 (Eq. (15)) with bandwidth $h = 1.0$ m for all three functions along with the corresponding envelopes obtained from global envelope tests (Myllymäki et al., 2017) based on 2499 random labelling simulations applied to the (A) Embrach and (B) Clocaenog mapped tree data. (For the interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

were always very close, while $\gamma(r)$ often produced greater deviations from the other two mark correlation functions. When comparing $\delta(r)$ and $\tau(r)$, it also becomes clear that $\delta(r)$ tended to produce larger values in local maxima and minima than $\tau(r)$.

When considering the random-labelling envelopes it is evident that those of $\delta(r)$ and $\tau(r)$ were more similar and narrower than those of $\gamma(r)$. Although the mark variogram produced the largest and smallest function values, none of them were significant for any r neither at Embrach nor at Clocaenog. This was different for $\delta(r)$ and $\tau(r)$ at Embrach (Fig. 3A), because the functions' first peaks at around $r \approx 2$ m were partly outside the corresponding envelopes and therefore significant. It is also interesting to note that more values of $\delta(r)$ were significant than values of $\tau(r)$.

For Clocaenog Forest (Fig. 3B), only few instances of $\delta(r)$ were significant in the first local minimum at $r \approx 7$ m whilst $\tau(r)$ came very close to significance at this distance. Also here, more values of $\delta(r)$ were significant than values of $\tau(r)$. Overall the results suggest that $\delta(r)$ using the principle of the dissimilarity coefficient as test function is more sensitive than the other two correlation functions.

3.4. Dissimilarity, differentiation and size segregation functions

When comparing functions $\overline{V}(r)$ and $\overline{T}(r)$ applied to the two forests, it is evident that size diversity is markedly larger and more variable at Embrach (Fig. 4A) than at Clocaenog Forest (Fig. 4B). While at Embrach both functions decline with increasing numbers of neighbours and distance, they remain fairly constant at Clocaenog Forest. At Embrach Forest, $\overline{V}(r) > \overline{T}(r)$ particularly at short distances, but both curves were quite similar (and significant up to $r \approx 2.5$ m), whilst at Clocaenog Forest, $\overline{V}(r) < \overline{T}(r)$ and there was a marked difference between the two curves and their envelopes did not overlap.

The corresponding size segregation functions are very similar for both forests, but more similar at Clocaenog Forest (Fig. 4B). At Clocaenog Forest, the size segregation functions for both size diversity indices largely suggest independent tree sizes with nearly no spatial correlation. At Embrach, both size segregation functions, similar to the mark correlations functions (see Section 3.3), indicate a marked attraction of different sizes, which is significant for $r \approx 2.5$ m. This trend is slightly stronger and more significant for the size segregation function which is based on the dissimilarity coefficient.

4. Discussion

Analysing the comparative properties of competing spatial plant diversity indices is not an easy task and to date there are no firm protocols or guidelines for such work. Our study attempted to introduce a broad framework or protocol for such an analysis that examines a number of different aspects of the characteristic under study. This was possible by analysing the behaviour of the two competing indices in (1) Monte Carlo tests (Baddeley et al., 2016), (2) in mark correlation functions (Illian et al., 2008) and (3) in size segregation functions including the associated global envelope tests (Myllymäki et al., 2017). Each of these analyses highlighted different properties of the dissimilarity coefficient.

Our study identified size diversity rather than size dominance as the main objective of analysis rather than size dominance. Given this, the results in Fig. 1A highlighted that the dissimilarity coefficient by Ali (1997) and Hagner and Nyquist (1998) shares much similarity with the size differentiation index by Gadow (1993). Ali (1997) even explicitly referred to the size differentiation index as competitor, but did not carry out any comparative analysis. As Section 2.1 has demonstrated, the statistical reasoning of the dissimilarity coefficient likely is better than that of the differentiation index.

The Monte Carlo test simulations revealed a greater sensitivity of the size dissimilarity coefficient compared to the size differentiation index. Under 'normal' conditions between $0.10 < q < 0.90$, both indices performed equally well, but in the boundary regions of $q < 0.10$ and $q > 0.90$ the dissimilarity coefficient clearly turned out to be more sensitive. This could mean that the statistical power (i.e. 1 – probability of type II error) associated with the dissimilarity coefficient is greater than that related to the differentiation index, but this is notoriously difficult to determine, since true size inequality is hard to define. However, we know that the simulations were clearly based on dependent marking, i.e. the point process model used indeed included a correlation between

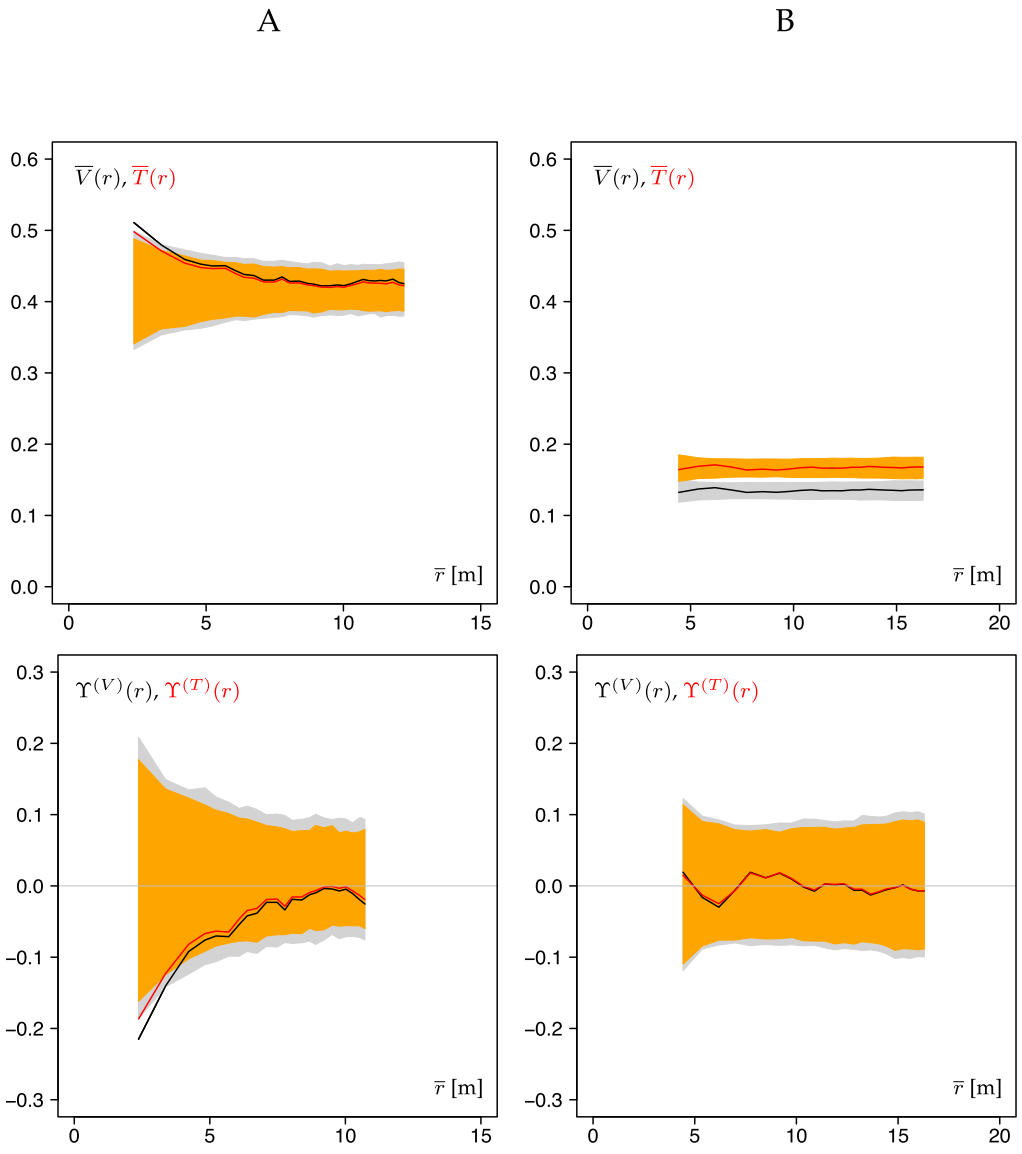


Fig. 4. Size dissimilarity and size differentiation functions, $\overline{V}(r)$ and $\overline{T}(r)$, respectively, along with their size segregation variants, $\Upsilon^{(V)}(r)$ and $\Upsilon^{(T)}(r)$, computed based on Eqs. (9) and (11) for (A) Embrach and (B) Clocaenog mapped tree data. (For the interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

space and size differences. Therefore it is likely that the curves in Fig. 2 describe the type II error.

When using mark correlation functions with three different test functions, it was interesting to see how some of the corresponding function curves indicated the trends involved more strongly than others (Fig. 3). On first sight, the mark variogram based on test function t_3 (Eq. (16)), appeared to show the spatial trends more clearly, but none of the $\gamma(r)$ curves were statistically significant according to the global envelope test. Mark correlation function $\delta(r)$ using test function t_1 (Eq. (14)), based on the dissimilarity coefficient, more strongly indicated spatial inequality trends than mark correlation function $\tau(r)$ using test function t_2 (Eq. (15)), based on the size differentiation index. In addition, $\delta(r)$ was significant for more instances of r than $\tau(r)$. This outcome supports the view that $\delta(r)$ is potentially a powerful characteristic that owes its strength much to test function t_1, which again is based on the dissimilarity coefficient.

Similar trends as for the mark correlation functions we obtained from the $\overline{V}(r)$ and $\overline{T}(r)$ functions and from the corresponding size segregation functions $\Upsilon(r)$ (Fig. 4). These are functions which are based on mean dissimilarity and mean differentiation using different numbers of nearest neighbours k. As expected, the difference between mean dissimilarity and mean differentiation were often small, however, $\overline{V}(r)$ and $\Upsilon^{(V)}(r)$ showed greater sensitivity and often led to statistical significance for more instances of r than $\overline{T}(r)$ and $\Upsilon^{(T)}(r)$.

The strategy of using a diversity index under study in different functional contexts has clearly helped us to identify subtle differences in the performance of this index.

5. Conclusions

When size diversity rather than size dominance is the study objective, the size dissimilarity index by Ali (1997) and Hagner and Nyquist (1998) definitely is a size inequality index that merits attention, since it is statistically more sensitive and most likely also has greater statistical power than its closest competitor, the size differentiation index by Gadow (1993). The statistical rationale of the size dissimilarity coefficient also has a solid foundation and the characteristic can also be used in a context where occasional sizes are zero. When many of neighbouring plant sizes are zero, the simplified version of the index (Eq. (6))

is recommended. In addition, the principle of the dissimilarity coefficient can also be recommended as a test function of the mark correlation function. Our new method of analysing competing diversity indices has revealed interesting properties of the dissimilarity coefficient.

Funding

This research was supported by the Swedish government research council for sustainable development (Formas) grant #2023–00994.

CRediT authorship contribution statement

Arne Pommerening: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Aila Särkkä:** Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

We thank Graham Durrheim (South African National Parks, Sedgefield, South Africa) and Andreas Zingg (Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland) for the permission to include the Knysna and the Embrach data, respectively, in this study. We are also grateful to Pavel Grabarnik (Russian Academy of Sciences, Pushchino, Russia) for helpful discussions and Mari Myllymäki supported us in the application of the global envelope test and her `GET` R package.

References

Aguirre, O., Hui, G.Y., Gadow, K., Jiménez, J., 2003. An analysis of spatial forest structure using neighbourhood-based variables. Forest Ecology and Management 183, 137–145.

Albert, M., 1999. Analyse der eingriffsbedingten Strukturveränderung und Durchforstungsmodellierung in Mischbeständen. [Analysis of thinning-induced changes in stand structure and modelling of thinnings in mixed-species stands.] Göttingen University, Hainholz Verlag Göttingen. PhD thesis.

Ali, A.A., 1997. Describing tree size diversity. Swedish University of Agricultural Sciences, Umeå. MSc thesis.

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

Banks-Leite, C., Ewers, R.M., Folkard-Tapp, H., Fraser, A., 2020. Countering the effect of habitat loss, fragmentation, and degradation through habitat restoration. One Earth 3, 672–676.

Burkhart, H.E., Tomé, M., 2012. Modeling forest trees and stands. Springer, New York.

Crawford, J., Torquato, S., Stillinger, F.H., 2003. Aspects of correlation function realizability. Journal of Chemical Physics 119, 7065–7074.

Davies, O., Pommerening, A., 2008. The contribution of structural indices to the modelling of Sitka spruce (*Picea sitchensis*) and birch (*Betula* spp.) crowns. Forest Ecology and Management 256, 68–77.

Fronville, T., Blaum, N., Kramer-Schadt, S., Schlägel, U., Radchuk, V., 2024. Performance of five statistical methods to infer interactions among moving individuals in a predator-prey system. Methods in Ecology and Evolution 15, 1097–1112.

Gadow, K. v., Zhang, G., Durrheim, G., Drew, D., Seydack, A., 2016. Diversity and production in an Afromontane Forest. Forest Ecosystems 3, 15.

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

Gaston, K.J., Spicer, J.I., 2004. Biodiversity. An introduction, 2nd edition. Blackwell Publishing, Oxford.

Hagner, M., Nyquist, H., 1998. A coefficient for describing size variation among neighboring trees. Journal of Agricultural, Biological, and Environmental Statistics 3, 62–74.

Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical analysis and modelling of spatial point patterns. John Wiley & Sons, Chichester.

Lotwick, H.W., Silverman, B.W., 1982. Methods for analysing spatial processes of several types of points. Journal of the Royal Statistical Society Series B 44, 406–413.

Magurran, A.E., 2004. Measuring biological diversity. Blackwell Publishing, Malden.

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., Hahn, U., 2017. Global envelope tests for spatial processes. Journal of the Royal Statistical Society: Series B (statistical Methodology) 79, 381–404.

Ohser, J., Stoyan, D., 1981. On the second-order and orientation analysis of planar stationary point processes. Biometrical Journal 23, 523–533.

Penttinen, A., Stoyan, D., Henttonen, H.M., 1992. Marked point process in forest statistics. Forest Science 38, 806–824.

Pommerening, A., Grabarnik, P., 2019. Individual-based methods in forest ecology and management. Springer Nature, Cham.

Pommerening, A., Särkkä, A., 2013. What mark variograms tell about spatial plant interactions. Ecological Modelling 251, 64–72.

Pommerening, A., Sterba, H., Eskelson, B.N.I., 2024. Distance and T-square sampling for spatial measures of tree diversity. Ecological Indicators 163, 111995.

Pommerening, A., Stoyan, D., 2006. Edge-correction needs in estimating indices of spatial forest structure. Canadian Journal of Forest Research 36, 1723–1739.

Pommerening, A., Stoyan, D., 2008. Reconstructing spatial tree point patterns from nearest neighbour summary statistics measured in small subwindows. Canadian Journal of Forest Research 38, 1110–1122.

Pommerening, A., Svensson, A., Zhao, Z., Wang, H., Myllymäki, M., 2019. Spatial species diversity in temperate species-rich forest ecosystems: Revisiting and extending the concept of spatial species mingling. Ecological Indicators 105, 116–125.

Pommerening, A., Szmyt, J., Zhang, G., 2020. A new nearest-neighbour index for monitoring spatial size diversity: The hyperbolic tangent index. Ecological Modelling 435, 109232.

Pommerening, A., Uria-Diez, J., 2017. Do large forest trees tend towards high species mingling? Ecological Informatics 42, 139–147.

Pommerening, A., Gonçalves, A.C., Rodríguez-Soalleiro, R., 2011. Species mingling and diameter differentiation as second-order characteristics. Allgemeine Forst- Und Jagdzeitung 182, 115–129.

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.

Stoyan, D., Wälder, O., 2000. On variograms in point process statistics, II: Models of markings and ecological interpretation. Biometrical Journal 42, 171–187.

Torquato, S., 2002. Random heterogeneous materials. Microstructure and macroscopic properties, Springer, New York.

Wälder, O., Stoyan, D., 1996. On variograms in point process statistics. Biometrical Journal 38, 895–905.

Wang, H., Zhao, Z., Myllymäki, M., Pommerening, A., 2020. Spatial size diversity in natural and planted forest ecosystems: Revisiting and extending the concept of spatial size inequality. Ecological Informatics 57, 101054.

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

Weiner, J., Solbrig, O.T., 1984. The meaning and measurement of size hierarchies in plant populations. Oecologia 61, 334–336.

Weiskittel, A.R., Hann, D.W., Kerschaw, J.A., Vanclay, J.K., 2011. Forest growth and yield modeling. Wiley Blackwell, Chichester.