Ecological Indicators 121 (2021) 106995

Contents lists available at ScienceDirect

Ecological Indicators

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

Unravelling the mechanisms of spatial correlation between species and size diversity in forest ecosystems

Arne Pommerening [a], Gongqiao Zhang [b,*], Xiaohong Zhang [c,*]

[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] Research Institute of Forestry, Chinese Academy of Forestry, Key Laboratory of Tree Breeding and Cultivation of National Forestry and Grassland Administration, Box 1958, Beijing 100091, China

[c] Research Institute of Forest Resource Information Techniques (IFRIT), Chinese Academy of Forestry, Key Laboratory of Forest Management and Growth Modelling, NFGA, Beijing 100091, China

ARTICLE INFO

Keywords:
Insurance hypothesis
Plant-plant interactions
Point process statistics
Mark mingling function
Mark variogram
Random-labelling test

ABSTRACT

With ongoing climate change at global scale we are currently losing biodiversity at an unprecedented rate. The insurance hypothesis and associated research, however, suggest that biodiversity has a major stabilising effect in ecosystems. In this situation, it is crucial to develop a better understanding of natural processes of maintaining biodiversity for employing them in conservation practice. In forest ecosystems, spatial species and size diversity are important aspects of α-diversity at woodland community and species population level. Both aspects of spatial diversity stem from complex relationships between tree interaction, disturbances and subsequent waves of colonisation by tree seedlings of various species. Using point process statistics, particularly the mark mingling function and the mark variogram, we studied the processes causing spatial correlations of species and size diversity. We found that spatial species dispersal and conspecific size distributions are key drivers of spatial species-size correlations and that a combination of simple random size-labelling techniques applied to mark variograms is instrumental in efficiently diagnosing them. If size ranges differ between species, spatial size diversity is largely a function of spatial species mingling. The existence of these correlations is crucial to conservation because they imply that conservation efforts can be rationalised: It is possible to focus on only one of the two tree diversity aspects. Interestingly, in multi-species forest ecosystems, although general species diversity is high, spatial species-size correlations can be diluted, because some of the many species populations may have similar size distributions.

1. Introduction

A growing body of empirical studies has found that more diverse animal and plant communities are more stable, i.e. exhibit smaller fluctuations over time (Valone and Barber, 2008). The *insurance hypothesis*, for example, involves correlation relationships among species and suggests that species that might be functionally redundant in a given ecosystem increase in numbers in more favourable conditions to compensate for the reduction in performance of the dominant species and thus provide "insurance" for community productivity. Through niche complementarity biodiversity promotes greater insurance when communities are made up of species that are better performers (i.e. specialists) in different, localised environments and thus complement each other (Yachi and Loreau, 1999; Matias et al., 2013). This concept was extended by the *spatial insurance hypothesis* predicting that functional complementarity of species across space and time insures the system against the impact of environmental fluctuation (Loreau et al., 2003). Conversely, lower levels of species richness may compromise the insurance functions of biodiversity (Leary and Petchey, 2009). Similar theories and hypotheses include *statistical averaging* or the *portfolio effect* and *compensatory dynamics* (Shanafelt et al., 2015) and put more weight on statistical mechanisms.

There has always been a natural coming and going of species in natural history. However, the current loss of biodiversity due to human interference, particularly due to climate change, is currently happening at an unprecedented rate and may largely be irreversible (Leary and Petchey, 2009). Therefore it is crucial to mitigate biodiversity losses through appropriate conservation management. Monitoring

* Corresponding authors.
E-mail addresses: zhanggongqiao@126.com (G. Zhang), zhangxh@ifrit.ac.cn (X. Zhang).

https://doi.org/10.1016/j.ecolind.2020.106995
Received 16 June 2020; Received in revised form 18 September 2020; Accepted 20 September 2020
Available online 19 October 2020

biodiversity is a necessary pre-requisite for goal-oriented conservation management and the corresponding statistical characteristics for measuring biodiversity need to be well chosen (Krebs, 1999; Magurran, 2004).

In forest ecosystems, species and size diversity are crucial aspects of α-diversity at community level. In the past, highest priority was assigned to species diversity (Gaston and Spicer, 2004). Ford (1975) and Weiner and Solbrig (1984), however, have pointed out that size inequality (as they termed size diversity) is equally important and often develops in natural ecosystems as a consequence of the interplay between tree interaction, disturbances and colonisation by tree seedlings.

These are complex ecological processes that are often hard to disentangle. Truly spatially explicit studies of the insurance hypothesis are still rare (Loreau et al., 2003), however, over the last fifty years point process statistics has contributed sophisticated measures that allow more refined ecological analyses involving spatial information. In point process statistics, plant locations are represented by points and so-called marks can provide additional information on plants such as species and size. This field of spatial statistics has produced a number of so-called second-order characteristics that consider pairs of points separated by distance r. These characteristics are functions of distance r and allow quantifying species and size diversity related to a spatial scale. In addition the point process statistics community has produced a number of models that allow the simulation of spatial tree patterns involving different species and sizes, which greatly helps to understand correlations between spatial species and size diversity (Illian et al., 2008; Wiegand and Moloney, 2014; Pommerening and Grabarnik, 2019).

Hence the objectives of this study were

(1) assisted by simulations to discuss what ecological processes lead to correlations between spatial tree species and size diversity and
(2) using diverse example forest data from China to identify methods from point process statistics that efficiently indicate different patterns of correlated spatial species and size diversity. These methods render the results of interaction processes traceable in field data.

2. Materials and methods

Based on spatial simulations we studied the processes involved in size-species correlations and then traced them in data from species-rich forest ecosystems in China. For this we used combinations of characteristics and tests from point process statistics.

2.1. Spatial simulations

We simulated the processes leading to different spatial patterns of species and size diversity by applying a combination of different point process models. First, we simulated moderately clustered tree locations by using the Matérn cluster process model (Matérn, 1960) with the intensity of cluster centres $\lambda_p = 0.0031$, cluster radius $R = 18$ m and a mean number of trees per cluster of $\bar{c} = 12$ trees. Next, we assigned approximately half of the points to a theoretical species 1 and the other half to a theoretical species 2. Following this we simulated stem diameters (as an example size characteristic) from two very different two-parameter Weibull distributions (Nagel and Biging, 1995): The shape and scale parameters of species 1 were 6 and 20, those of species 2 were set to 6 and 60. In addition the smallest possible stem diameters were 5 cm and 20 cm, respectively. These settings led to simulated stem-diameter distributions that were distinctively different for the two species: Whilst both distributions were similarly bell-shaped, the size distribution of species 1 always occupied the lower size range and that of species 2 the larger size range with little overlap between them (Fig. 1). As a result the two species had rather small conspecific size diversity, however, heterospecific size diversity was large.

Finally we modified the initial random allocation of species and size

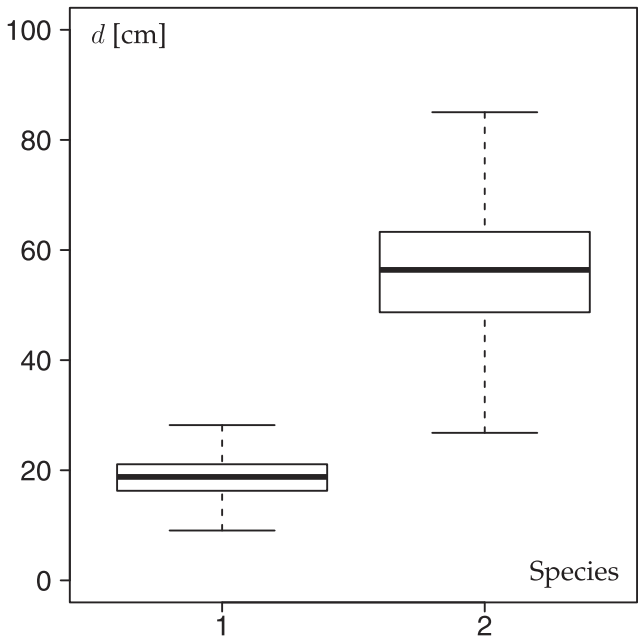

Fig. 1. Empirical stem-diameter distributions of species 1 ($\bar{n} = 188$) and 2 ($\bar{n} = 191$) simulated from 100 Weibull density distributions as explained in Section 2.1. $\bar{n}$ is mean species abundance, d is stem diameter at breast height (1.3 m above ground level).

to tree locations by using the spatial *partial construction technique* as detailed in Pommerening et al. (2019): We defined two variants of the species segregation function $\Psi(k)$ to simulate situations where (a) species 1 and 2 intimately mingled among the seven nearest neighbours and where (b) species 1 and 2 did not mingle much at this neighbourhood scale and as a consequence occurred in segregated clusters. Here k denotes the nearest kth neighbour. We modelled these two variants of $\Psi(k)$ using the power function $a_0 \times k^{a_1}$ and applied parameters $a_0 = -0.697$ and $a_1 = -0.429$ to represent situation (a) and parameters $a_0 = 0.697$ and $a_1 = -0.429$ to represent situation (b), i.e. the two curves applied were reflections of each other with respect to the horizontal line through zero indicating independent species marks, see Fig. 2. The parameters were carefully chosen so that species segregation values were high in absolute terms for $k = 1, \cdots, 7$ and the decline of the species segregation function towards 0 was very gradual. This ensured that the selected effect of species mingling would continue even beyond $k = 7$.

In a large number of iterations, the construction algorithm optimised the dispersal of species marks based on simulated annealing (Pommerening et al., 2019) so that the final pattern had a species segregation function that resembled one of the two options of Fig. 2 as closely as possible. In this optimisation, the stem-diameter marks were jointly reassigned with the species marks, i.e. the two marks – species and size – were treated as a "bundle", however, there was no optimisation for stem diameters, but only for species, and tree locations remained unchanged. For better understanding our simulations we give two example simulation results (Fig. 3), where panel (A) shows a pattern representing mingling situation (a) and the red curve of $\widehat{\Psi}(k)$ in Fig. 2, whilst panel (B) depicts a pattern with mingling situation (b) based on the black curve of $\widehat{\Psi}(k)$ in Fig. 2. The different ranges of stem diameters are also clearly discernible in Fig. 3.

2.2. Study area and data

Daqingshan forest region (abbreviated as D) forms a part of the Daqingshan Forest Farm of the Experimental Center of Tropical Forestry, Chinese Academy of Forestry (Fig. 4). The research area is situated in Pingxiang city, Guangxi province. The average annual

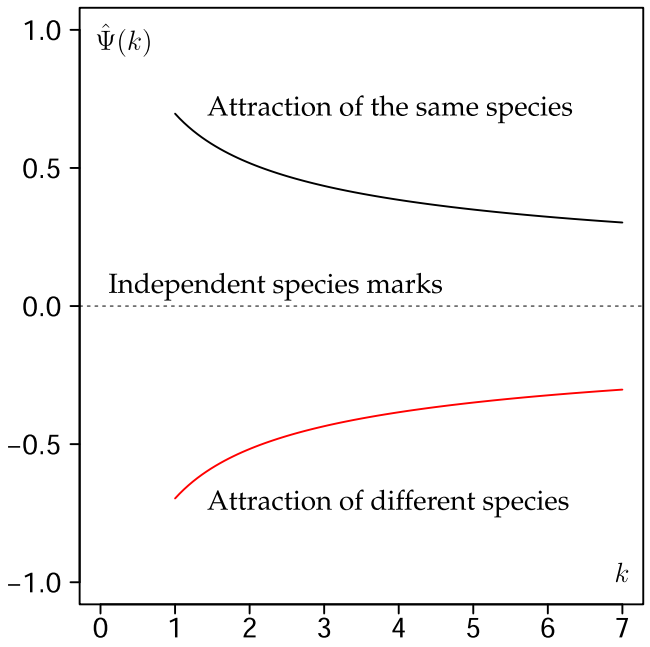

Fig. 2. The two species segregation functions $\widehat{\Psi}(k)$ used in the simulation of the two bivariate mingling patterns studied in this work. For explanations see Section 2.1. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

rainfall varies between 1261 and 1695 mm and each year mainly occurs from May to September. Relative humidity is between 80 and 84% and the average annual temperature is 20.5–21.7 °C. The soil is mainly classified as laterite and red soil. The native vegetation type is characterised by southern subtropical monsoon forests and evergreen broadleaved forests involving diverse species. Plot a in Daqingshan, abbreviated as Da (22°17′ N, 106°42′ E), is dominated by planted *Cunninghamia lanceolata* (Lamb.) Hook. mixed with diverse, naturally seeded broadleaved species.

Jiulongshan Forest (abbreviated as JS) is located in the western suburbs of Beijing (39°57′ N and 116°05′ E) in the northern branch of Taihang Mountain (Fig. 4). The climate in this region is temperate continental and largely influenced by monsoon climate that has a distinctive rain season between June and September. Mean annual rainfall is 623 mm and mean annual temperature is 11.8 °C. The site has a thin brown rocky mountain soil with high stone content. In this analysis, we used forest stands a and b in Jiulongshan, abbreviated as JSa and JSb. Stand JSa is dominated by planted *Platycladus orientalis* (L.) Franco and is mixed with some naturally regenerated species such as *Quercus variabilis* Blume, *Broussonetia papyrifera* (L.) Vent., *Ailanthus altissima* (Mill.) Swingle, *Prunus davidiana* Carr. and *Gleditsia sinensis* Lam. Stand JSb represents a secondary, mixed-species broadleaved deciduous forest, where the main species *Pinus tabuliformis* Carr. and *Larix principis-rupprechtii* Mayr were planted.

Tazigou Experimental Forest Farm (129°56′–131°04′ E, 43°05′–43°40′ N) is located in Jilin Province, China (Fig. 4). This area of secondary forest, abbreviated as TF, is situated on Laoyeling Mountain of the Changbai Mountain range. The elevation ranges from 300 m to 1200 m asl with annual rainfall ranging from 600 mm to 700 mm. The

Fig. 4. Locations of the study areas Daqingshan, Jiulongshan and Tazigou included in this study.

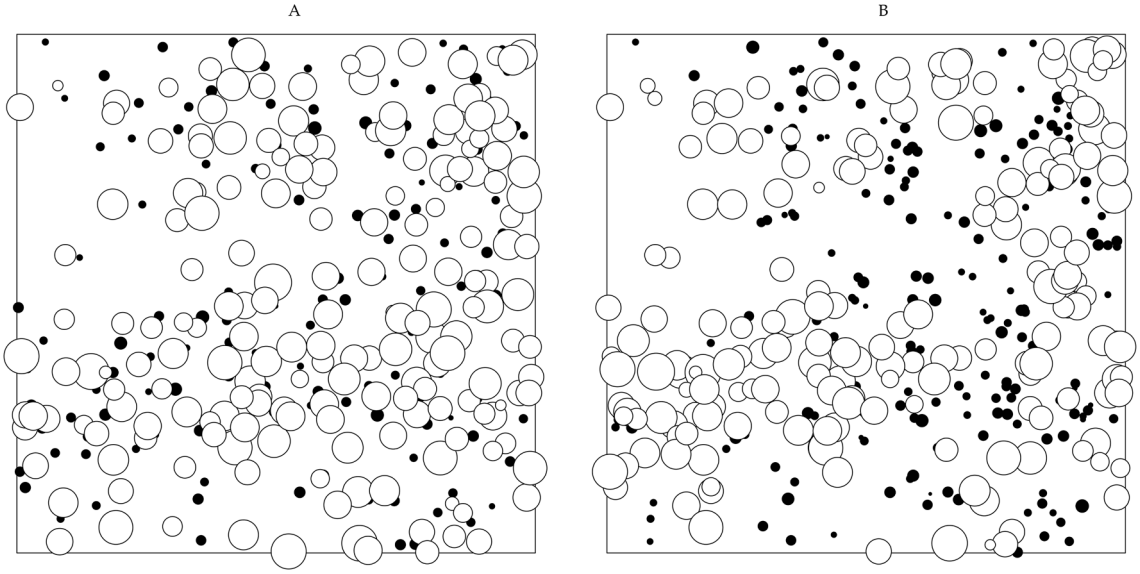

Fig. 3. Sample simulation results showing a bivariate species pattern involving very different size ranges. A: Pattern reflecting attraction of different species (red $\widehat{\Psi}(k)$ curve in Fig. 2). B: Pattern with an attraction of the same species (black $\widehat{\Psi}(k)$ curve in Fig. 2). Both patterns are based on the same tree location coordinates simulated from a Matérn cluster process model (Matérn, 1960).

average annual temperature is 4 °C. The area has predominantly dark brown soil (humic cambisols) with high natural fertility. The main tree species are Mongolian oak (*Quercus mongolica* Fisch. ex Ledeb.), Siberian white birch (*Betula platyphylla* Sukaczev), Korean pine (*Pinus koraiensis* Siebold & Zucc.), Ussuri popular (*Populus ussuriensis* Komarov), Amur lime (*Tilia amurensis* Rupr.), painted maple (*Acer pictum* subsp. *mono* (Maxim.) H. Ohashi) and Korean birch (*Betula costata* Trautv.). The stands included in this research are in plots c and f in Tazigou, abbreviated as TFc and TFf.

2.3. Second-order characteristics for monitoring spatial diversity

Pommerening et al. (2011) and Hui and Pommerening (2014) introduced the mark mingling function $\nu(r)$. This characteristic is ideal for monitoring spatial species diversity in species-rich plant communities and shares similarities with the mark correlation function and the intertype mark connection function. The fundamental idea of $\nu(r)$ is to use the mingling test function $\left(m(\xi_i) \neq m(\xi_j)\right)$, i.e. assessing only, whether the species marks m of a pair of trees under consideration are different (Pommerening and Grabarnik, 2019, p. 162). A suitable estimator of the mark mingling function is given in Eq. (1).

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

Here ξ_i and ξ_j are arbitrary tree locations of a spatial pattern in the observation window W. k_h is a kernel function and we used the Epanechnikov kernel function in this study, see Pommerening and Grabarnik (2019, p. 151f.). $A\left(W_{\xi_i} \cap W_{\xi_j}\right)$ is the area of intersection of W_{ξ_i} and W_{ξ_j} (Illian et al., 2008, p. 481f. and p. 188), relating to the translation edge correction (Ohser and Stoyan, 1981). Expected mingling, $\mathbf{E}M$, serves as a normalising term (Pommerening and Grabarnik, 2019, p. 132). For $\nu(r) > 1$ we can conclude that there is heterospecific attraction, whilst with $\nu(r) < 1$ we have conspecific attraction, i.e. different species are organised in separate neighbourhood clusters. Spatially uncorrelated (=independent) species marks are indicated by $\nu(r) \approx 1$. In analogy to the mark variogram and the mark correlation function, heterospecific attraction can also be termed negative autocorrelation, whilst conspecific attraction qualifies for positive autocorrelation (Suzuki et al., 2008).

An appropriate test involves the null hypothesis of *a priori marking* or *random superposition*, also referred to as population independence (Illian et al., 2008; Pommerening and Grabarnik, 2019). This test requires simulating $n = 2499$ spatial patterns with independent marks for estimating global envelopes (Myllymäki and Mrkvička, 2019). Since species is the mark of interest, spatial mark independence is simulated by random shifts of species populations (Illian et al. 2008, p. 460f.; Pommerening and Grabarnik, 2019, p. 182f.). For this purpose we randomly shifted all individuals of species 1 by adding the same random values z_x and z_y to the x and y coordinates of these individuals in our simulations. A variant of periodic boundary conditions ensured that all points were inside the observation window. For the data examples from China we selected as many species as were required to shift approximately half of all points.

The mark variogram $\gamma_m(r)$ is a characteristic derived from geostatistical variograms and is designed for quantitative marks such as plant size variables. Its test function $\frac{1}{2}\left(m(\xi_i) - m(\xi_j)\right)^2$ quantifies the difference between two size marks m by subtracting them from one another and squaring the difference. The estimator of the mark variogram we used in this study is given in Eq. (2).

$$\widehat{\gamma}_m(r) = \frac{1}{\sigma_m^2} \sum_{\xi_i,\xi_j \in W}^{\neq} \frac{\frac{1}{2}\left(m(\xi_i) - m(\xi_j)\right)^2 k_h\left(\|\xi_i - \xi_j\| - r\right)}{2\pi r A\left(W_{\xi_i} \cap W_{\xi_j}\right)} \tag{2}$$

Apart from a different test function the normalising term before the sum is also different from the mark mingling function: It is the reciprocal of the mark variance σ_m^2 and in our case stem diameters at 1.3 m above ground level, d, were the size marks used. Large differences between marks are indicated by $\gamma_m(r) > 1$ (also referred to as negative autocorrelation), whilst with $\gamma_m(r) < 1$ (also referred to as positive autocorrelation) both marks are similar in size regardless whether the two marks in question are both large or both small. Spatially uncorrelated (=independent) size marks are indicated by $\gamma_m(r) \approx 1$ (Suzuki et al., 2008; Pommerening and Särkkä, 2013). We also checked alternatives to the mark variogram such as the mark correlation function $k_{mm}(r)$ (Illian et al., 2008; Pommerening and Grabarnik, 2019), but the mark variogram always turned out to be the better indicator for the purposes of this study.

As with any quantitative marks, the null hypothesis relates to *a posteriori marking* or *random labelling*. Simulations under this mark independence hypothesis are typically based on fixed point locations and permutated marks. According to the *traditional* random-labelling method, all size marks are freely permutated without restriction. However, when multivariate patterns involving several species are studied, it is common to *restrict* random labelling in such a way that plant sizes are only permutated *within each species population* (Wiegand and Moloney, 2014, p. 227f.; Wang et al., 2020). As a result the non-spatial empirical size distribution of each species are preserved. Also here we applied $n = 2499$ simulations for estimating global envelopes (Myllymäki and Mrkvička, 2019). We deliberately applied both variants of the random labelling test in order to uncover correlations between spatial species and size diversity. Our hypothesis is that the results of the two random-labelling tests hold vital clues about spatial species-size correlations.

For all simulations and calculations we used our own R (version 3.5.1, R Development Core Team 2020) and C++ code and additionally applied the spatstat (Baddeley et al., 2016) and GET packages (Myllymäki and Mrkvička, 2019).

3. Results

Here we present the results obtained from our spatial simulations and from the analysis of the species-rich forest ecosystems in China.

3.1. Simulation analysis

The point-process simulations offered valuable insights on how the mark mingling function $\widehat{\nu}(r)$, the mark variogram $\widehat{\gamma}_m(r)$ and the associated tests respond to different strategies of tree species mingling. Since the processes were stochastic, we simulated each spatial pattern 100 times and calculated the means of the function graphs as well as of the envelopes (Fig. 5).

The different shapes of the mark mingling function clearly indicated the two intended mingling situations, i.e. (a) species 1 and 2 intimately mingled at short distances (leading to an attraction of different species and negative autocorrelation) and where (b) species 1 and 2 did not mingle much at short distances and as a consequence occurred in segregated clusters (leading to an attraction of the same species and positive autocorrelation). In the top two rows a and b of Fig. 5, we can see that spatial size diversity "follows" spatial species diversity, i.e. there is also a negative autocorrelation of size marks (large difference of tree sizes) at short distances for $\widehat{\nu}(r) > 1$ and a positive autocorrelation of marks (small difference of tree sizes) for $\widehat{\nu}(r) < 1$. This means that size diversity clearly is a result of mingling, provided the species involved are represented by very different size ranges and have unimodal size distributions. From panels B and C in the two top rows a and b of Fig. 5 we also understand that there is a marked difference between the envelopes obtained from traditional random labelling (panel B) and from restricted within-species random labelling (panel C). In panel B, the envelopes were always centred towards the horizontal line through 1.0, whilst in the two top rows a and b and panel C they followed the mark-variogram

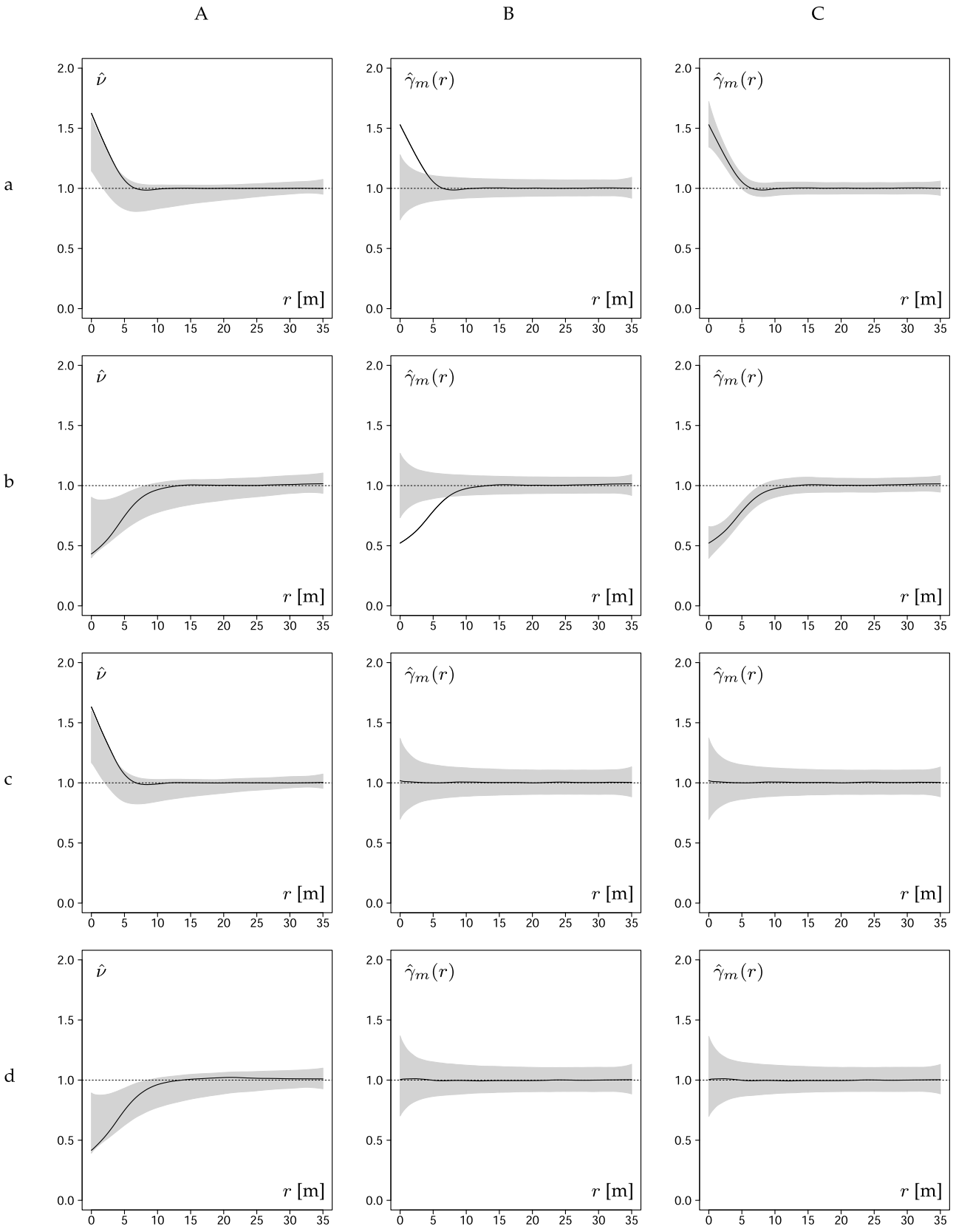

Fig. 5. Results of the point process simulations according to the methods described in Section 2.1. A: Mark mingling function $\widehat{\nu}(r)$ with envelopes resulting from the random superposition test. B: Mark variogram $\widehat{\gamma}_m(r)$ including envelopes from traditional random-labelling testing. C. Mark variogram $\widehat{\gamma}_m(r)$ including envelopes based on restricted within-species random-labelling testing. In the simulations leading to the results in the top two rows a and b, species 1 and 2 had distinctively different size ranges (see Fig. 1). In the bottom two rows, the size range of species 1 and 2 was identical. r is the distance between pairs of trees.

curve. The graphs in panel B also seem to suggest that $\widehat{\gamma}_m(r)$ up to $r = 10$ m is significant, but panel C clearly shows that it is not when the restricted within-species random labelling test is applied. The central argument of our paper is that any difference in the envelopes resulting from the two random-labelling tests clearly indicates spatial correlations between species and size diversity.

The two bottom rows c and d of Fig. 5 show the results of our control experiments: Here all tree sizes were sampled from the Weibull distribution of species 1 regardless of actual species, i.e. as a consequence the two species had no distinctly different size ranges. Although the mark mingling function in these cases clearly indicated the two distinctive species mingling patterns, the mark variograms suggested mark independence, i.e. there was no spatial correlation between the size marks. Also, the envelopes obtained from both random labelling tests were centred towards the horizontal line through 1 and were as good as identical. From these results we can conclude that spatial species-size correlations can only develop, if the size ranges of the species involved are markedly different.

If the size ranges of the species involved are quite different, size diversity is a function of species mingling including the sign of spatial autocorrelation. The two variants of the random labelling test are thus able to uncover correlations between spatial species and size diversity.

3.2. Field data analysis

Tree densities ranged from 952 trees per hectare at Tazigou Experimental Forest Farm, stand f (TFf) to 2331 trees per hectare at Jiulongshan, stand a (JSa). Basal area per hectare was lowest at TFf (13.9 m^2) and highest at Daqingshan, stand a (Da) with 33.2 m^2. The latter stand included a maximum of 57 species, whilst the minimum number of eight species occurred at JSa. The stem-diameter coefficient of variation is a non-spatial measure of size diversity and was highest in TFf and lowest at JSa. Incidentally, such size-related coefficients of variation are often regarded as measures of community stability with larger coefficients potentially implying lower stability (Doak et al., 1998; Tilman et al., 1998).

The results of the five field data sites predominantly showed spatial configurations with conspecific aggregation (JSb, TFc, TFf), i.e. trees of the same species occurred at close proximity (Fig. 6A). There was moderate, heterospecific attraction in Da for $r < 5$ m and in JSa for 7 m $< r <$ 30 m. Da was originally planted with *Cunninghamia lanceolata* and later colonised by natural broadleaves, which explains negative autocorrelation of species and size (Wang et al., 2020). In our experience, such heterospecific attraction at close proximity is comparatively rare in nature and often the result of natural or human disturbances. As in Fig. 5, the mark variograms in panels B and C of Fig. 6 followed the general trend of the mark mingling functions (Fig. 6A) thus supporting the same close association between species and size.

With the exception of stand JSa the two random labelling simulations also here led to different test results: Whilst unrestricted random labelling produced envelopes centred towards the horizontal line of mark independence through 1, random labelling restricted to species boundaries gave envelopes that followed the mark variograms more closely. Notably this was not the case for stand JSa. At the same time mark variograms that seemed to be significant when using the traditional, unrestricted random-labelling test, were less or not significant at all (e.g. for JSb) when applying the restricted random-labelling test. The differences in envelope behaviour between stands Da, JSb, TFc and TFf on one hand and JSa on the other can again be explained by different patterns of conspecific size distributions (Fig. 7).

We selected the size distributions of the seven most abundant species, since differences between the size distributions of the most abundant species are most likely to influence species-size correlations. At Daqingshan stand a (Da), for example, the size distributions of the two most abundant species markedly differed but these two distributions together also differed much from the five less abundant species in the stand. We found the most heterogeneous size distributions between species at Tazigou Experimental Forest Farm, stand f (TFf) followed by those of stand c (TFc) of the same forest ecosystem. Here, incidentally, the stem-diameter coefficients of variation were highest (0.79 and 0.71; see Table 1). The lower the heterogeneity of conspecific size distributions the less pronounced is the difference between the envelopes resulting in the two variants of the random-labelling test. The extreme end of this continuum of size distributions was illustrated by stand Jiulongshan a (JSa), where the size distributions of the seven most abundant species were fairly homogeneous and the stem-diameter coefficient of variation was lowest (0.29; see Table 1). This corresponds with simulated random-labelling envelopes in panels B and C of Fig. 6 that are largely identical.

An interesting observation from the analysis of these species-rich forest ecosystems is that the species-size effects uncovered by the two random-labelling tests were weaker than those achieved in our simulations (compare Figs. 5 and 6) where only two species with approximately equal abundances occurred.

4. Discussion

Our simulations have demonstrated that distinctively different conspecific size ratios are an important pre-requisite for spatial correlations of size and species diversity, which often occur in natural and managed forest ecosystems. Then spatial size diversity "follows" spatial species diversity or in other words spatial size inequality is a function of spatial species mingling that even includes the sign of autocorrelation. This implies negative size autocorrelation in cases where spatial species mingling suggests an attraction of different sizes and positive size autocorrelation where there is an attraction of the same species at short intertree distances (Fig. 5). An example of negative size autocorrelation at short distances is Daqingshan (stand a, Da, see Fig. 6B) and a particularly strong example of positive size correlation can be seen at Tazigou Experimental Forest Farm (stand f, TFf, Fig. 6B).

The existence of spatial species-size correlations can be diagnosed by comparing the results of two random-labelling tests applied to the mark variogram, i.e. a traditional, unrestricted variant of the test that is performed regardless of species and a restricted variant where size marks are permutated only within species-population boundaries (Wiegand and Moloney, 2014, p. 227f.; Wang et al., 2020). Spatial species-size correlations exist, if the resulting envelopes differ, specifically, if at least for some r the envelopes simulated with the restricted variant of the test are not centred towards the horizontal line through 1 denoting the case of uncorrelated size marks. The greater the difference between the centres of the two test envelopes the more spatial correlation between species and size diversity exists. The envelope behaviour even offers the opportunity to potentially reduce the number of statistics and associated tests to be computed to just one, i.e. to the mark variogram along with the random-labelling test where permutations are restricted to species populations. If the envelopes of this test are not centred towards the horizontal line through 1 but rather follow the curve of the mark variogram at least for some r, spatial species-size correlations exist at that spatial scale. Owing to these spatial species-size correlations, the shape of the mark mingling function can then be inferred from that of the mark variogram provided the size distributions are unimodal, as they were throughout this study.

Our finding re-confirms the results of recent research suggesting that spatial species and size diversity are often correlated (Pommerening and Uria-Diez, 2017; Pommerening et al., 2019; Wang et al., 2020). The importance of this finding cannot be overestimated, since it implies that conservation management in mixed-species forests only has to focus on one of these two spatial tree diversity aspects whilst getting the other "for free" as a side product. At the same time this correlation between spatial species and size diversity contributes an explanation to the long-standing debate on how nature maintains size diversity: It now seems very likely that size hierarchies and size diversity in forest ecosystems

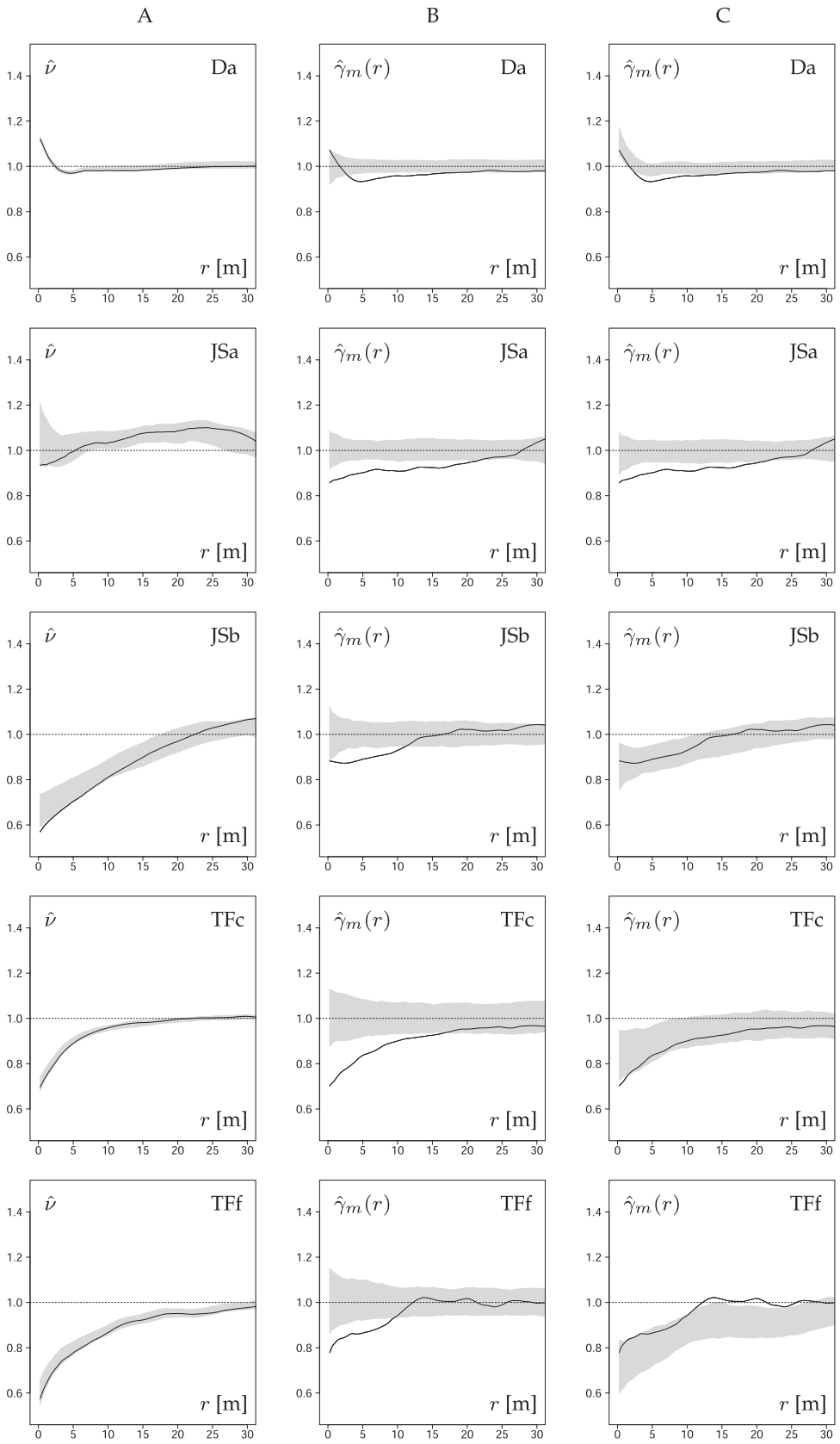

Fig. 6. Results of analysing the five observed species-size patterns Da, JSa, JSb, TFc and TFf described in Section 2.2. A: Mark mingling function $\hat{\nu}(r)$ with envelopes produced by the random superposition test. B: Mark variogram $\hat{\gamma}_m(r)$ including envelopes from traditional random labelling testing. C. Mark variogram $\hat{\gamma}_m(r)$ including envelopes simulated by restricted within-species random labelling testing. r is the distance between pairs of trees.

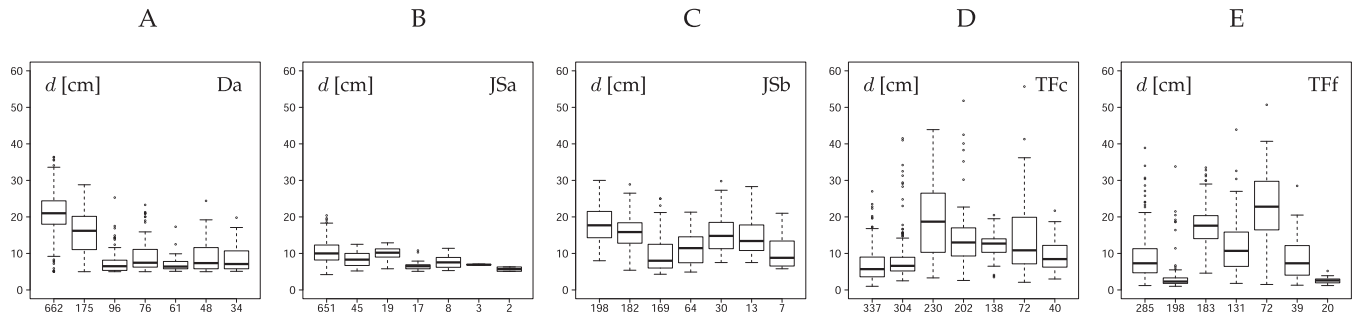

Fig. 7. Empirical stem-diameter distributions of the seven most abundant species in the five observed species-size patterns Da, JSa, JSb, TFc and TFf described in Section 2.2. The numbers on the abscissa indicate species abundances. *d* is stem diameter at breast height (1.3 m above ground level).

Table 1
Summary data of five Chinese tree diversity monitoring plots. *d* is stem diameter at breast height (1.3 m above ground level).

Plot	Slope (°)	Mean altitude (m)	Plot size (m × m)	Density (trees ha^{-1})	Number of species	Mean *d* (cm)	*d* coeff. of variation	Basal area (m^{2}ha^{-1})
Da	23	725	90 × 110	1445	57	15.4	0.486	33.16
JSa	17	145	40 × 80	2331	8	10.1	0.289	20.24
JSb	15	990	100 × 50	1346	12	14.4	0.392	25.44
TFc	8	675	100 × 100	1344	13	11.3	0.704	20.28
TFf	7	645	100 × 100	952	12	10.7	0.793	13.88

increase with species diversity. Frequently occurring small-to-medium-scaled natural and human-induced disturbances provide opportunities for new cohorts of tree species to colonise gaps of moderate size and this process causes distinctive patterns of spatial species-size correlations that sometimes are negative and sometimes positive depending on scale. Thus local diversity is maintained by immigration from surrounding patches, which provides the source of variation on which selection can act to favour adapted types (Loreau et al., 2003). The mingling-size hypothesis predicting that larger trees have a tendency towards high species mingling (Pommerening and Uria-Diez, 2017; Wang et al., 2018) is but a special case of positive spatial correlations between species and size diversity. It may also hold in the spatial context that the more negative autocorrelation there is in spatial species and size diversity, the stronger the stabilising effect is (Valone and Barber, 2008). These insights on how natural maintenance of diversity works are indispensable to conservation practitioners for mitigating diversity losses due to climate change.

Finally it was an intriguing outcome of this study that the species-size correlation effects as indicated by the random-labelling tests were generally weaker in the multi-species Chinese forest ecosystems (Fig. 6) than they were in the theoretical simulations (Fig. 5) involving only two species. This can be explained by a "dilution effect" resulting from the great number of species populations in these ecosystems. As Fig. 7 illustrates, even among the seven most abundant species some size distributions of different species are quite similar whilst only a few differ more markedly. Therefore in very species-diverse forests such as many forest ecosystems in China and elsewhere, spatial species-size correlations can be weaker than in ecosystems with less species, although in general non-spatial species diversity is high. In multi-species temperate and subtropical forest ecosystems therefore care must be taken in the analysis, since spatial species-size correlations can be masked, because some of the many species populations may have similar size distributions.

5. Conclusions

Spatial species dispersal and conspecific size distributions are key drivers of spatial species-size correlations. Such correlations are an important feature of the natural maintenance of tree diversity. Simple simulation tests based on random size-labelling techniques can efficiently diagnose species-size correlations. The existence of these correlations is crucial to conservation practice, as they imply that conservation efforts can be rationalised: Promoting tree species diversity in most cases also promotes size diversity at the same time. All that is necessary to do in mixed-species woodlands is to encourage either species diversity or size diversity, e.g. through canopy openings or by thinning monospecific patches of juvenile trees. Such small-scaled induced disturbances can lead to a diversification of tree species and size and only one of these two diversity aspects needs to be monitored.

Author contributions

A.P. conceived the ideas and designed the methodology. G.Z. and X. Z. collected the data and all authors analysed the data, carried out the modelling/analyses and contributed to the text. All authors gave final approval for publication.

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.

Acknowledgements

The authors thank Hongxiang Wang (Guangxi University, China) for letting us use his Daqingshan data in this study. A.P. gratefully acknowledges a guest professor grant from the University of Natural Resources and Life Sciences BOKU that allowed him to spend three weeks of this study at Vienna in October 2019 and to collaborate with X.Z. He thanks Manfred Lexer and the colleagues at the Institute of Silviculture for their hospitality. G.Z.'s research visit to Umeå was funded by the Fundamental Research Funds for the Central Non-profit Research Institution of the Chinese Academy of Forestry (project number CAFYBB2019GC001-2). X.Z. received funding from the Fundamental Research Funds for the Central Non-profit Research Institution of the Chinese Academy of Forestry (project number CAFYBB2019GC001-6) for a research stay at Vienna (Austria). We thank Mari Myllymäki (LUKE, Finland) for interesting discussions about random-labelling and global-envelope testing.

Data accessibility statement

The data of this study are available at https://zenodo.org/record/4056565 or using DOI 10.5281/zenodo.4056565.

References

Baddeley, A., Rubak, E., Turner, R., 2016. Spatial Point Patterns. Methodology and Applications with R. CRC Press, Boca Raton, FL.

Doak, D.F., Bigger, D., Harding, E.K., Marvier, M.A., O'Malley, E.E., Thomson, D., 1998. The statistical inevitability of stability-diversity relationships in community ecology. Am. Nat. 151, 264–276.

Gaston, K.J., Spicer, J.I., 2004. Biodiversity. An Introduction. Blackwell Publishing, Oxford.

Ford, E.D., 1975. Competition and stand structure in some even-aged plant monocultures. J. Ecol. 63, 311–333.

Hui, G., Pommerening, A., 2014. Analysing tree species and size diversity patterns in multi-species uneven-aged forests of Northern China. For. Ecol. Manage. 316, 125–138.

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

Krebs, C.J., 1999. Ecological Methodology, 2nd edition. Addison Wesley Longman, Inc., Menlo Park, CA.

Leary, D.J., Petchey, O.L., 2009. Testing a biological mechanism of the insurance hypothesis in experimental aquatic communities. J. Anim. Ecol. 78, 1143–1151.

Loreau, M., Mouquet, N., Gonzalez, A., 2003. Biodiversity as spatial insurance in heterogeneous landscapes. Proc. Natl. Acad. Sci. USA 100, 12765–12770.

Magurran, A.E., 2004. Measuring Biological Diversity. Blackwell Publishing, Oxford.

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

Matérn, B., 1960. Spatial variation. Meddelanden fran Statens Skogsforskningsinstitut 49, 1–144.

Myllymäki, M., Mrkvička, T., 2019. GET: Global envelopes in R. arXiv:1911.06583 [stat. ME].

Nagel, J., Biging, G.S., 1995. Schätzung der Parameter der Weibullfunktion zur Generierung von Durchmesserverteilungen. [Estimation of the parameters of the Weibull function for generating diameter distributions.]. Allgemeine Forst- und Jagd-Zeitung [Ger. J. For. Res.] 166, 185–189.

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

Pommerening, A., Gonçalves, A.C., Rodríguez-Soalleiro, R., 2011. Species mingling and diameter differentiation as second-order characteristics. Allgemeine Forst- und Jagd-Zeitung [Ger. J. For. Res.] 182, 115–129.

Pommerening, A., Särkkä, A., 2013. What mark variograms tell about spatial plant interaction. Ecol. Model. 251, 64–72.

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

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. Ecol. Ind. 105, 116–125.

Pommerening, A., Grabarnik, P., 2019. Individual-Based Methods of Forest Ecology and Management. Springer, Cham.

R Development Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria http://www.r-project.org.

Shanafelt, D.W., Dieckmann, U., Jonas, M., Franklin, O., Loreau, M., Perrings, C., 2015. Biodiversity, productivity, and the spatial insurance hypothesis revisited. J. Theor. Biol. 380, 426–435.

Suzuki, S.N., Kachi, N., Suzuki, J.-I., 2008. Development of a local size-hierarchy causes regular spacing of trees in an even-aged *Abies* forest: analyses using spatial autocorrelation and the mark correlation function. Ann. Bot. 102, 435–441.

Tilman, D., Lehman, C.L., Bristow, C.E., 1998. Diversity-stability relationships: Statistical inevitability or ecological consequence? Am. Nat. 151, 277–282.

Valone, T.J., Barber, N.A., 2008. An empirical evaluation of the insurance hypothesis in diversity-stability models. Ecology 89, 522–531.

Wang, H., Peng, H., Hui, G., Hu, Y., Zhao, Z., 2018. Large trees are surrounded by more heterospecific neighboring trees in Korean pine broad-leaved natural forests. Sci. Rep. 8, 9149.

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. Ecol. Inf. 56, 101054.

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

Wiegand, T., Moloney, K.A., 2014. Handbook of Spatial Point Pattern Analysis in Ecology. CRC Press.

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