

RESEARCH

Open Access



Global woodland structure from local interactions: new nearest-neighbour functions for understanding the ontogenesis of global forest structure

Arne Pommerening¹, Hongxiang Wang^{2*} and Zhonghua Zhao³

Abstract

Background: A number of hypotheses and theories, such as the Janzen-Connell hypothesis, have been proposed to explain the natural maintenance of biodiversity in tropical and temperate forest ecosystems. However, to date the details of the processes behind this natural maintenance are still unclear. Recently two new nearest-neighbour characteristics were proposed and in this paper we demonstrate how they can contribute to a better understanding of the ontogenesis of global forest structure from localised neighbourhoods.

Methods: We applied the new species and size segregation functions together with appropriate test procedures to four example woodland data sets from China at Daqingshan, Jiaohe, Jiulongshan and Xiaolongshan forest regions. In addition we quantified the morphology of the new characteristics and modelled a neighbourhood allometric coefficient linking the two functions.

Results: The results revealed quite different species segregation patterns with both conspecific and heterospecific attraction. We found these to be generally matched by equivalent size segregation patterns of attraction of similar and different sizes. It was straightforward to model the size segregation function from the knowledge of the species segregation function by estimating a neighbourhood allometric coefficient.

Conclusions: The new characteristics have helped to quantify the extent and rate of decline of neighbourhood interactions in terms of spatial species and size diversity. Through the allometric neighbourhood coefficient the analysis highlighted once more how closely related species and size segregation are, thus supporting the mingling-size hypothesis. Using both a traditional and a restricted random-labelling test has provided a valuable tool for understanding the exact nature of species-mingling and size-inequality relationships.

Keywords: Species segregation function, Size segregation function, Mingling-size hypothesis, Neighbourhood allometric coefficient, random labelling

* Correspondence: wanghongxiang168@126.com

²College of Forestry, Guangxi University, Nanning 530004, China
Full list of author information is available at the end of the article

Background

Many theories and hypotheses have been proposed to explain the maintenance of plant and tree diversity in forest ecosystems (Pommerening and Uria-Diez 2017).

The Janzen–Connell hypothesis (Janzen 1970; Connell 1971), for example, proposes that elevated numbers of specialist natural enemies, such as herbivores and pathogens, maintain diversity in plant communities. They reduce the survival rates of conspecific seeds and seedlings located close to reproductive adults or in areas of high conspecific density (Comita et al. 2014) leading to elevated conspecific self-thinning (Yao et al. 2016), i.e. a progressive decline in density in a population of growing individuals of the same species (Begon et al. 2006, p. 157). An important effect of the Janzen–Connell hypothesis is the *negative density/distance dependence* that occurs when nearby conspecific plants negatively affect performance through mechanisms such as intraspecific competition and pest facilitation (Wills et al. 1997; Wright 2002; Piao et al. 2013; Yao et al. 2016).

Another important ecological hypothesis in this context, the *herd immunity* hypothesis, focuses on the variation in heterospecific neighbour densities and predicts that species diversity confers protection from natural enemies by making it more difficult for specialist natural enemies to locate host plants (Wills et al. 1997). Thus the spatial spread of an infection can be slowed down or even stopped by mixtures of susceptible and resistant species (Begon et al. 2006). Therefore individual plant fitness should be enhanced in stands involving many species but reduced in stands with low species diversity (Wills et al. 1997; Murphy et al. 2016).

In addition, Ford (1975) and Weiner and Solbrig (1984) demonstrated how the Janzen–Connell and the herd immunity hypotheses can affect natural populations. One observation in many natural plant populations is that self-thinning leads to size inequality (also referred to as local size hierarchies), where heterospecific stands include dominant plants emerging from a first colonisation cohort, which are often surrounded by patches of smaller sized plants of the same cohort. Small plants of these early colonisers are initially often of the same species as the dominant plants and according to the mechanisms of both the Janzen–Connell and herd immunity hypotheses later decrease in numbers due to self-thinning processes. Eventually the small early colonisers are partially or completely replaced by even smaller individuals of other species from subsequent colonisation cohorts. This combined effect of species and size replacement enforces both local size inequality and the mingling of different plant species in a given area or patch and prevents the development of monocultures (Pommerening and Grabarnik 2019).

Synthesising the key elements of the aforementioned ecological hypotheses, Pommerening and Uria-Diez (2017) and Wang et al. (2018) have concluded that – as a consequence of them – there is often a tendency for plants with high species mingling, i.e. those with heterospecific nearest neighbours, to be larger sized plants and referred to this as the *mingling-size hypothesis*.

Some aspects of these theories could be confirmed in experiments but many processes are still largely unknown. However, the effects described by the Janzen–Connell, herd-immunity, size inequality and mingling-size hypotheses are strictly local effects, i.e. their spatial range hardly exceeds 8–10 m in temperate and boreal forests. It is therefore interesting and important to study how the local woodland structures emerging from these localised effects cooperate in space and time and influence the global structure of a forest stand or woodland. Of particular interest is the question of how localised clusters of tree vegetation evolve in the vicinity of large parent trees, how these localised clusters interact and eventually merge to form overall global stand structure. This question has to be pursued using spatial approaches, since different spatial scales are involved. As typical local clusters of 1–2 mature trees surrounded by smaller offspring essentially constitute neighbourhoods, nearest-neighbour approaches seem most appropriate. These methods are part of point process statistics, one of the main fields of spatial statistics, and good overviews are offered in the textbooks by Illian et al. (2008), Wiegand and Moloney (2014) and Pommerening and Grabarnik (2019).

The objective of this paper is to show how recently proposed nearest-neighbour characteristics can contribute to tracing the ontogenesis of global forest stand structure from localised neighbourhoods and particularly how measures of species and size diversity are related.

Methods

Nearest-neighbour approaches of species and size diversity

Gadow (1993) and Aguirre et al. (2003) extended the species segregation index by Pielou (1977) to general multivariate species patterns involving k neighbours. The mingling index M_i (Eq. 1) is defined as the mean heterospecific fraction of plants among the k nearest neighbours of a given plant i . In the analysis, every plant within a given research plot acts once as plant i (Pommerening et al. 2019).

$$M_i = \frac{1}{k} \sum_{j=1}^k \mathbf{1}(\text{species}_i \neq \text{species}_j) \quad (1)$$

The mingling approach is particularly useful in woodlands with many species, because the only distinction made is between co- and heterospecific pairs of trees.

Expected mingling, EM , (implying independent species marks) as explained in Pommerening et al. (2019) and in analogy to Pielou’s segregation index, can now be used to normalise species mingling in such a way that the resulting measure Ψ is independent of the number of species and refers to the null hypothesis of independent species marks according to

$$\Psi = 1 - \frac{\bar{M}}{EM}. \tag{2}$$

Pommerening and Uria-Diez (2017) referred to the index defined in Eq. (2) as *species segregation index*. Here \bar{M} is mean mingling of a tree population. Consequently, $\Psi = 0$, if the species are independently or randomly dispersed. If the nearest neighbours and a given tree always share the same species, $\Psi = 1$ (attraction of the same species leading to species segregation). $\Psi = -1$ is approached, when many neighbours are increasingly of a different species than the tree under consideration (attraction of different species).

Since the number k of nearest neighbours is related to different scales of local neighbourhood that can vary considerably within and between forest stands, Pommerening et al. (2019) considered a function $\Psi(k)$ that depends on variable k (Eq. 3).

$$\Psi(k) = 1 - \frac{\bar{M}(k)}{EM}, k = 1, 2, 3, \dots \tag{3}$$

With increasing k , $\Psi(k) \rightarrow 0$. Our hypothesis is that the exact way in which $\Psi(k)$ approaches 0 is a better quantitative description of the mingling pattern in a given population than \bar{M} or Ψ alone. For plant patterns with completely random species dispersal (independent species marks) $\Psi(k) = 0$ for all k (Pommerening et al. 2019).

To better account for different spatial scales of local neighbourhood, we replaced k by the population mean distance $\bar{r}_k = \frac{1}{N} \sum_i^N r_{ik}$ between any individual i and its k^{th} nearest neighbour. This resulted in the new scale-dependent *species segregation function* $\Psi'(r)$ (Eq. 4).

$$\Psi'(r) = \begin{cases} \Psi(k), & \text{for } r = \bar{r}_k, k = 1, 2, 3, \dots \\ \Psi(k) + \frac{\Psi(k+1) - \Psi(k)}{\bar{r}_{k+1} - \bar{r}_k} (r - \bar{r}_k), & \text{for } \bar{r}_k < r < \bar{r}_{k+1}, k = 1, 2, 3, \dots \end{cases} \tag{4}$$

The interpretation of the function is straightforward and basically follows the interpretation of Ψ (Eq. 2).

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 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. 460 f.; Pommerening and Grabarnik 2019, p. 182 f.). For this purpose we selected as many species as required to shift approximately half of all points in the observation window. In the simulations, all individuals of the selected species are shifted by adding the same random values z_x and z_y to the x and y coordinates of these individuals. A variant of periodic boundary conditions ensures that all points are inside the observation window.

Size differentiation is defined as the mean of the ratio of smaller and larger plant sizes u of the k nearest neighbours subtracted from one. In the analysis, every plant within a given research plot acts once as plant i (Gadow 1993).

$$T_i = 1 - \frac{1}{k} \sum_{j=1}^k \frac{\min(u_i, u_j)}{\max(u_i, u_j)} \tag{5}$$

Here u can be any quantifiable plant size measure. The value of T_i increases with increasing average size difference between neighbouring trees. $T_i = 0$ implies that neighbouring trees have equal size, whilst larger size differences approach an upper value of 1.

Expected size differentiation across species, ET , is independent of the number of neighbours, k , and can be combined with the arithmetic mean \bar{T} in an index Υ expressing the relationship between observed size differentiation and spatially uncorrelated size marks according to

$$\Upsilon = 1 - \frac{\bar{T}}{ET}. \tag{6}$$

Pommerening and Uria-Diez (2017) referred to the index defined in Eq. (6) as *size segregation index*. Since the index takes the expectation into account, it is normalised for the range of sizes involved in each study

population, which the original size differentiation index is not. Consequently, $\Upsilon = 0$, if plant sizes are independently dispersed without any spatial correlation of marks. If the sizes of the nearest neighbours and a given plant are more frequently of similar size than expected by chance, $\Upsilon \approx 1$ (attraction or aggregation of similar sizes leading to a segregation of sizes). Negative values tending towards -1 are achieved, if more neighbours have sizes quite different from that of a tree under study (attraction of different sizes; aggregation in the classical sense of Pielou's original approach) than expected in a pattern of independently dispersed sizes (Wang et al. 2020).

In analogy to the species segregation function $\Psi'(r)$, we can consider a function $\Upsilon(k)$ and eventually the *size segregation function* $\Upsilon'(r)$ (Wang et al. 2020, Eq. 7).

$$\Upsilon'(r) = \begin{cases} \Upsilon(k), & \text{for } r = \bar{r}_k, k = 1, 2, 3, \dots \\ \Upsilon(k) + \frac{\Upsilon(k+1) - \Upsilon(k)}{\bar{r}_{k+1} - \bar{r}_k} (r - \bar{r}_k), & \text{for } \bar{r}_k < r < \bar{r}_{k+1}, k = 1, 2, 3, \dots \end{cases} \quad (7)$$

With increasing r , $\Upsilon'(r) \rightarrow 0$, i.e. spatial independence of size marks is approached with distance used in the index calculation. The interpretation of $\Upsilon'(r)$ follows that of Υ . For plant patterns with spatially independent plant sizes, $\Upsilon'(r) = 0$ for all r (Wang et al. 2020).

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 permuted marks. According to the *traditional* random-labelling method, all size marks are freely permuted 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 permuted *within each species population* (Wiegand and Moloney 2014, p. 227 f.; 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).

Our hypothesis is that the exact way in which $\Psi'(r)$ and $\Upsilon'(r)$ approach 0, i.e. the morphology of $\Psi'(r)$ and $\Upsilon'(r)$, and the two aforementioned random-labelling tests give vital clues about local pattern formation in forests.

Characterising the morphology of species and size segregation functions

For a better quantitative description and interpretation of the function morphology we performed a linear regression on the values of $\Psi'(r)$ and $\Upsilon'(r)$ at \bar{r}_k for $k = 1,$

$2, 3, \dots, 30$. We considered the intercept and the slope of the linear regression as shape parameters, where the intercept is an expression of maximum species or size segregation (Pommerening et al. 2019). The slope parameter expresses the average slope of the (in most cases nonlinear) function, i.e. how rapidly on average the functions tend towards a value of 0 denoting spatial species or size independence. This characteristic is therefore a measure of decay of the species or size effect with distance. In addition we calculated the range $\bar{r}_k - \bar{r}_1$, i.e. the difference between the average distance to the first and the k^{th} nearest neighbour, which we set to $k = 30$. Short ranges occur in tree patterns with high densities and in patterns that are highly clustered whilst long ranges are common in those with low densities. Finally we also fitted polynomials of degree two to the

values of the two functions at \bar{r}_k and calculated the sum of squares of deviation (SSDL, Liu et al. 2009) from linearity (Eq. 8).

$$\text{SSDL} = \sum_{j=1}^k \left(f_j^P(r) - f_j^L(r) \right)^2 \quad (8)$$

The values of $f_j^P(r)$ and $f_j^L(r)$ (Eq. 8) were calculated for the j^{th} nearest neighbour from the polynomial (P) and linear (L) models, respectively, where $f(r)$ can refer to both species or size segregation function. Deviation from linearity as expressed by SSDL is a measure of the decreasing effect of species segregation or size inequality with distance. Nonlinear curves of $f(r)$ often indicate a fast decline of neighbourhood effect with distance (Pommerening et al. 2019; Wang et al. 2020).

Allometric relationship between species segregation and size inequality

Spatial species segregation and spatial size inequality act on the same neighbourhood. As work on confirming the size-mingling hypothesis by Pommerening and Uria-Diez (2017) and Wang et al. (2018) has shown, size inequality and species mingling are often closely related. This was an interesting finding highlighting that in mixed-species forests large size inequality can be found where there is large species mingling and vice versa. This association of the two aspects of tree diversity suggests that they can be interpreted as allometric neighbourhood relationships, i.e. it should be possible to estimate one of the two functions from the knowledge of

the other one using an allometric coefficient. To estimate the allometric coefficient, first the equivalents of relative growth rates (RGR) are calculated from $\Psi'(r)$ and $Y'(r)$ at successive \bar{r}_k thus exchanging time for space and quantifying rates of decline of the respective tree diversity effect with distance:

$$\bar{p}_{\bar{r}_k}^{(\Psi)} = \frac{\log \Psi'(\bar{r}_k) - \log \Psi'(\bar{r}_{k+1})}{\bar{r}_{k+1} - \bar{r}_k} \quad (9)$$

RGRs for $Y'(r)$, $\bar{p}_{\bar{r}_k}^{(Y)}$, can be calculated in analogy. Finally the allometric coefficient is calculated as

$$m_{\bar{r}_k} = \frac{\bar{p}_{\bar{r}_k}^{(\Psi)}}{\bar{p}_{\bar{r}_k}^{(Y)}} = \frac{\log \Psi'(\bar{r}_k) - \log \Psi'(\bar{r}_{k+1})}{\log Y'(\bar{r}_k) - \log Y'(\bar{r}_{k+1})}. \quad (10)$$

Considering mean RGRs or rate of declines and the mean allometric coefficient, $\bar{p}_{\bar{r}_k}^{(Y)}$ can now be estimated from $\bar{p}_{\bar{r}_k}^{(\Psi)}$ as $\bar{p}_{\bar{r}_k}^{(Y)} = \bar{p}_{\bar{r}_k}^{(\Psi)} / \tilde{m}$, where \tilde{m} is the geometric mean of $m_{\bar{r}_k}$. Finally using the growth-multiplier concept, values of $Y'(r)$ for $r > r_1$ can be estimated from the first observation, $Y'(r_1)$, and the growth multiplier $M_{\bar{r}_k}^{(Y)} = \exp(\bar{p}_{\bar{r}_k}^{(Y)}(\bar{r}_{k+1} - \bar{r}_k))$, see Pommerening and Grabarnik (2019, Chap. 5). In computational practice, however, it is more robust to estimate $m_{\bar{r}_k}$ from the two growth multipliers, i.e.

$$m_{\bar{r}_k} = \frac{M_{\bar{r}_k}^{(\Psi)}}{M_{\bar{r}_k}^{(Y)}} \quad (11)$$

yielding the same results as Eq. (10). For all calculations we have written our own code in R (version 3.5.1), C++ and additionally used the GET package (R Development Core Team 2019; Myllymäki and Mrkvička 2019).

Example data

To illustrate the new nearest-neighbour characteristics, the associated statistical tests and other curve characteristics, we selected four example forest sites from China.

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. The research area is situated in Pingxiang City, Guangxi Province, which is close to the border between China and Vietnam. The native vegetation type is characterised by southern subtropical monsoon forests and evergreen broad-leaved forests involving diverse broad-leaved species. Native species mainly include *Quercus griffithii* HOOK. f. & THOMSON ex MIQ., *Erythrophleum fordii* OLIVER, *Castanopsis hystrix* MIQ., *Mytilaria laosensis* LECOMTE, *Betula alnoides* BUCH.-HAM. ex D. DON

and *Dalbergia lanceolata* ZIPP. ex SPAN. Plot Da (22°17' N, 106°42' E) was established in 2018 and the prevailing main-canopy species are *Cunninghamia lanceolata* and *Mytilaria laosensis* LEC. and subdominant broadleaved species include *Macaranga denticulata* (BL.) MUELL. ARG., *Schefflera octophylla* (LOUR.) HARMS., *Machilus chinensis* (CHAMP. ex BENTH.) HEMSL., *Castanopsis hystrix* MIQ. and *Liquidambar formosana* HANCE.

Jiaohe forest region (abbreviated as J) is an experimental forest situated in the Dongdapo Nature Reserve (43°51'–44°05' N and 127°35'–127°51' E), Jilin Province, north-eastern China. The climate is largely influenced by monsoons from the Pacific Ocean and cold air from the interior Asian continent. The stand, denoted as Jb, is a mixed *Fraxinus mandshurica*-*Juglans mandshurica* forest. It mainly includes *Fraxinus mandshurica* RUPR., *Juglans mandshurica* MAXIM., *Acer mandshurica* MAXIM., *Carpinus cordata* BL., *Acer mono* MAXIM. and *Pinus koraiensis* SIEB. ET ZUCC.

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. The climate in this region is temperate continental and largely influenced by monsoon conditions. The selected 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.

Xiaolongshan Forest (abbreviated as XS) is part of the Xiaolongshan Nature Reserve, Gansu Province, north-west China. The forest is situated on the north-facing slopes of the West Qinling Mountain range (33°30'–34°49' N and 104°22'–106°43' E) and constitutes a natural mixed pine-oak forest. The research plot labelled as XSc represents a natural deciduous broad-leaved mixed forest and *Quercus aliena* var. *acuteserrata* MAXIM., *Dendrobenthamia japonica* (DC.) FANG. var. *chinensis* and *Acer davidii* FRANCH. are the most abundant species. This forest stand was restored since the 1970s after commercial harvesting (Pommerening et al. 2019; Wang et al. 2020).

Results

The four example research plots have tree densities that range from 748 tree per hectare (at Jiaohe, plot b) to 2331 tree per hectare (at Jiulongshan, plot a) and from 20.3 m².ha⁻¹ (at Jiulongshan, plot a) to 33.2 m².ha⁻¹ (at Daqingshan, plot a) (Table 1).

The number of species varies from 9 (at Jiulongshan, plot a) to 57 (at Daqingshan, plot a). The stem-diameter coefficient of variation is highest at Jiaohe (0.7, plot b) and lowest at Jiulongshan (0.3, plot a). For the remainder of this text we will mostly drop the plot affiliation for

Table 1 Basic characteristics of the four example research plots. D – Daqingshan forest region, J – Jiaohe forest region, JS – Jiulongshan Forest, XS – Xiaolongshan Forest. \hat{M} , $\hat{\psi}$, $\hat{\tau}$ and $\hat{\gamma}$ were calculated for $k = 4$ neighbours

Plot	Plot size (m × m)	Density (trees·ha ⁻¹)	Number of species	Mean dbh (cm)	dbh coeff. of variation	Basal area (m ² ·ha ⁻¹)	\hat{M}	$\hat{\psi}$	$\hat{\tau}$	$\hat{\gamma}$
Da	90 × 110	1445	57	15.4	0.486	33.16	0.782	-0.030	0.425	-0.030
Jb	100 × 100	748	21	21.7	0.724	27.95	0.785	0.116	0.453	0.042
JSa	40 × 80	2331	9	10.5	0.289	20.26	0.198	0.153	0.252	0.048
XSc	50 × 65	1486	30	15.6	0.524	28.58	0.713	0.177	0.399	0.007

ease of reading. The hat symbol, e.g. in $\hat{\gamma}$, used throughout the following test denotes estimators.

Mean mingling, \hat{M} , is highest at Daqingshan (0.78) and Jiaohe (0.79) and lowest at Jiulongshan (0.20). In addition the species segregation index, $\hat{\psi}$, highlights that there is weak heterospecific attraction at Daqingshan and mainly conspecific attraction in the other three research plots, particularly at Xiaolongshan. The hints given by the species segregation index are much more

clearly elaborated on by the species segregation function $\hat{\psi}'(r)$ (Fig. 1): There is indeed a mild attraction of heterospecific individuals at Daqingshan Forest up to a distance of $r = 5$ m, while in all other forest areas there is conspecific attraction only. The different ranges of $\hat{\psi}'(r)$ reflect the different tree densities at the respective sites in terms of number of trees per hectare.

Figure 1 also clearly shows different shapes of $\hat{\psi}'(r)$, even where the general attraction pattern is the same. Particularly the decline of $\hat{\psi}'(r)$ with distance is of im-

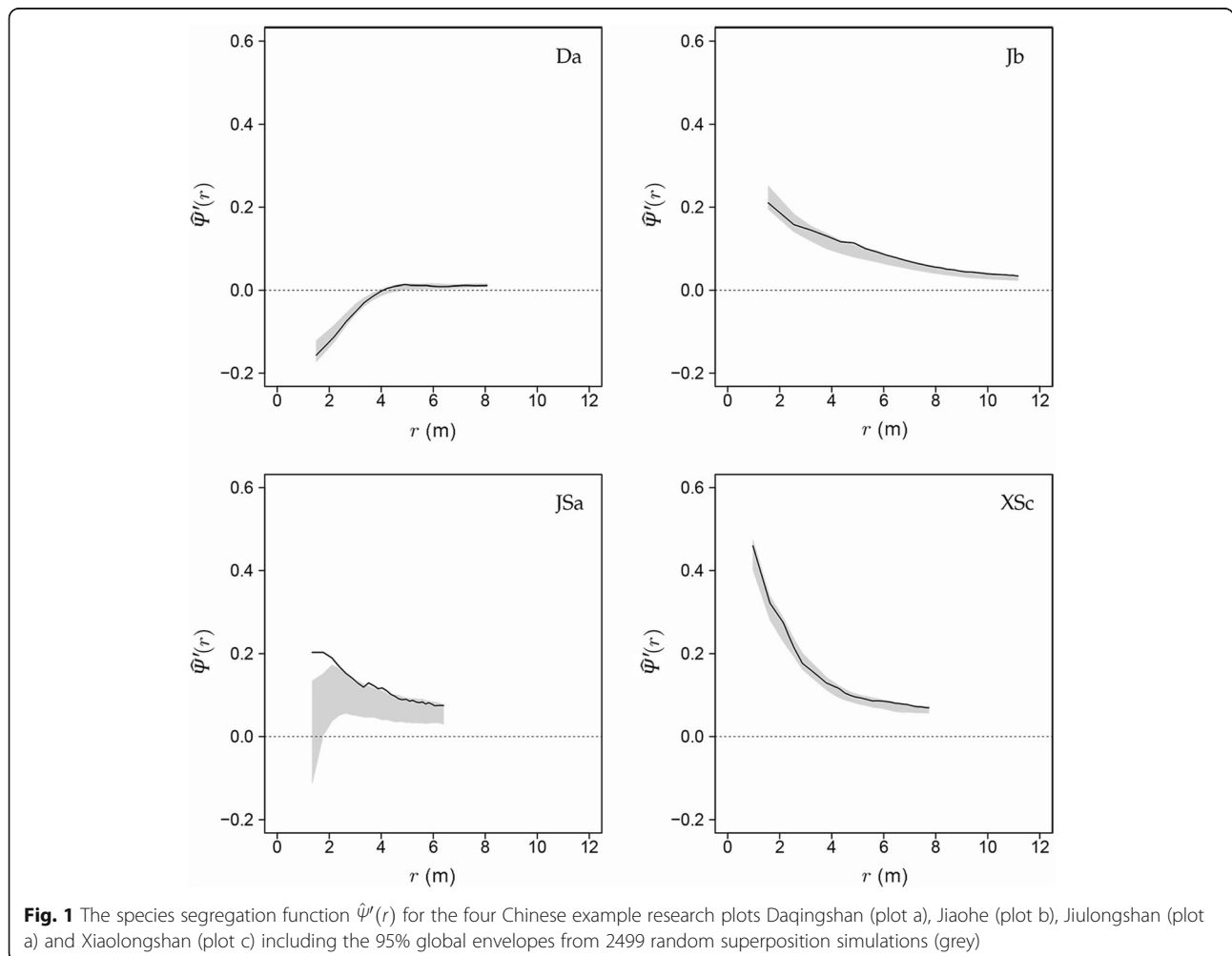


Fig. 1 The species segregation function $\hat{\psi}'(r)$ for the four Chinese example research plots Daqingshan (plot a), Jiaohe (plot b), Jiulongshan (plot a) and Xiaolongshan (plot c) including the 95% global envelopes from 2499 random superposition simulations (grey)

portance. At Jiaohe, there is a rather gradual decline of the mingling effect with distance, while the decline is rapid at Xiaolongshan. The random superposition simulations highlight that observed $\hat{\Psi}'(r)$ is significant at Daqingshan, Jiaohe, Jiulongshan and Xiaolongshan at least for some distances r . The curves in Fig. 1 also emphasise that even with $k = 30$ nearest neighbours the correlation range (beyond which species marks are independent) has not yet been reached for all forest patterns with the exception of Daqingshan, since the corresponding curves have not yet closely and asymptotically approached the horizontal line through 0, which would indicate spatial independence.

Like with mean mingling, \widehat{M} , mean diameter differentiation, \widehat{T} , is also largest at Jiaohe (0.45) and Daqingshan (0.43) and smallest at Jiulongshan (0.25), thus supporting the mingling-size hypothesis. At Daqingshan Forest, there is a moderate attraction of different tree sizes. Here $\hat{\Psi}$ and $\hat{\Upsilon}$ are even of the same magnitude. At Jiaohe and Jiulongshan the size segregation index suggests an attraction of similar sizes, which is strongest at Jiulongshan (0.05). In plot c at Xiaolongshan, the observed size marks appear to be largely independent. We can refine the impressions obtained from $\hat{\Upsilon}$ by studying the size segregation function $\hat{\Upsilon}'(r)$ (Fig. 2).

Comparing the curves of Fig. 1 and Fig. 2 we can confirm that size segregation generally operates at a scale smaller than species segregation. In Daqingshan Forest, the attraction of different sizes appears to be much stronger than index $\hat{\Upsilon}$ suggested, which supports the use of a function providing more details as opposed to a simple index. It is actually the strongest effect of all four example sites. At Jiaohe and Xiaolongshan there is an attraction of similar sizes with different patterns of decay with increasing distance r : At Jiaohe the differentiation effect decreases much more gradual than at Xiaolongshan. By contrast, at Jiulongshan Forest there is even hardly any change with distance.

Not only for testing the significance of the estimated characteristics but also for understanding the spatial relationship between species and size it is instructive to carry out the aforementioned two different types of random labelling tests: The first test is the “traditional” one (results in the left column of Fig. 2), where the different species occurring in the research plots are ignored and all sizes are permuted regardless of species. An important consequence of this test is that the species-specific mark distributions change from simulation to simulation and can largely deviate from those observed in the original plot. The second variant of the test (results in the right column of Fig. 2) is preferred by many quantitative ecologists and statisticians (Wiegand and Moloney 2014, p. 227 f.). Here the permutations are restricted by species boundaries, i.e. tree sizes are permuted within species

populations only. As a consequence species-specific mark distributions are preserved in the simulations, i.e. they are exactly the same in each simulation and correspond with those originally observed. However, for the less abundant species this can lead to situations where within the tiny species populations hardly any permutation of sizes is possible.

When comparing the envelopes in the left with those in right column of Fig. 2 we clearly see that they are centred towards zero for the traditional test but not for the restricted random labelling test. In some cases the envelopes of the restricted test follow the observed curves more closely than those of the traditional test, e.g. for Daqingshan and Xiaolongshan Forests. In other cases, such as for Jiaohe and Jiulongshan, the differences in the envelopes of the two tests appear to be marginal. As a consequence, the strongly significant test outcome indicated by the traditional method when applied to Daqingshan turns out to be merely an artefact and the observed size segregation function is only marginally significant when the correct test procedure, i.e. the restricted random-labelling method, is applied. A similar situation exists at Xiaolongshan. However, the question why the test envelopes differ so markedly between the two test methods for some patterns but not for others remains and will be discussed in the next section, as it is rather important.

The modelling of $\hat{\Upsilon}'(r)$ using the mean allometric coefficient \tilde{m} is straightforward and leads to good results (Fig. 3). This emphasises that species and size segregation functions are indeed related. *RMSE* is a fit characteristic and generally shows that the modelling based on only one model parameter, i.e. the mean allometric coefficient, was satisfactory. The curves show the general trend that the values of $\hat{\Upsilon}'(r)$ are usually much smaller than those of $\hat{\Psi}'(r)$ with the notable exception of Daqingshan Forest where both are of similar magnitude. The mean allometric coefficient is in all four cases close to 1 indicating an isometric decline (increase in Daqingshan Forest) of the size differentiation and mingling effects in tree neighbourhoods. The isometric relationship is particularly strong in Xiaolongshan ($\tilde{m} = 1.00$) and Daqingshan ($\tilde{m} = 1.01$). Deviations from isometry of 3% occur in Jiaohe and Jiulongshan.

The morphology of the species and size segregation functions is of particular interest for making inference about processes of natural maintenance of tree diversity. Table 2 summarises four characteristics of morphology, intercept and slope from applying a linear regression model, the spatial range of the two segregation functions and SSDL, a measure of deviation from linearity.

The differences in intercept between the two functions reflect the aforementioned differences in the function values. The signs in the slopes are consistent

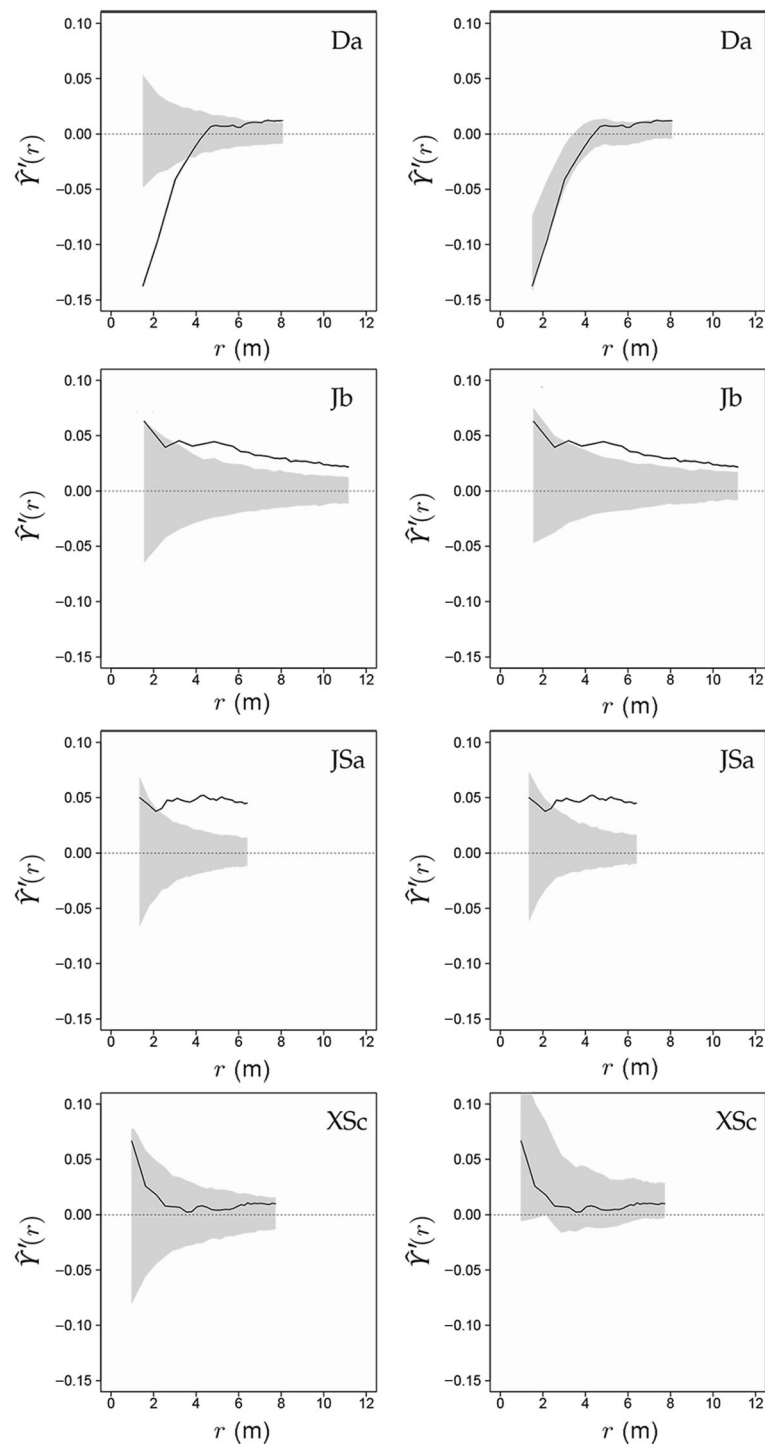
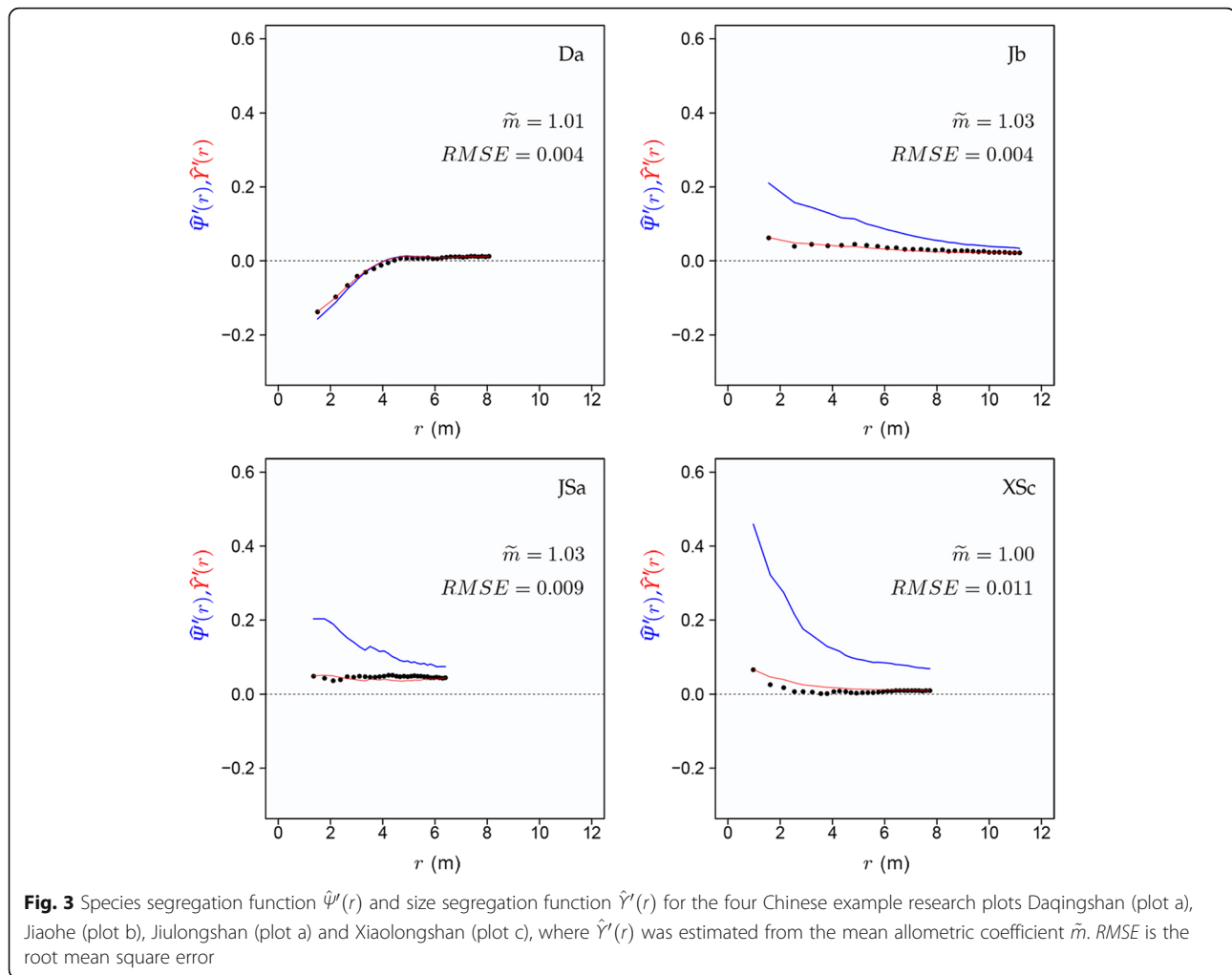


Fig. 2 The size segregation function $\hat{Y}'(r)$ based on stem diameters as size marks for the four Chinese example research plots Daqingshan (plot a), Jiaohe (plot b), Jiulongshan (plot a) and Xiaolongshan (plot c) including the 95% global envelopes from 2499 traditional random labelling simulations (grey, left) and from 2499 random labelling simulations restricted by species (grey, right)

between $\hat{\Psi}'(r)$ and $\hat{Y}'(r)$, only for Jiulongshan the slope of the linear fit is almost zero thus confirming the visual impression we gained from Fig. 2. Interestingly the absolute values of slope for both functions are only close for

Daqingshan but not for the other three woodlands. Perhaps the slope of a linear approximation is not the most effective way to estimate average decline of interactions. The range is naturally the same for both



functions and is a reflection of tree density as shown in Table 1. The greatest deviation from linearity for both $\hat{\Psi}'(r)$ and $\hat{Y}'(r)$ can be confirmed for Xiaolongshan and Daqingshan. This implies that the decline of mingling and size inequality effect with distance is greatest here whilst it is much more gradual for Jiaohe and Jiulongshan.

Discussion and conclusions

Our analyses have clearly shown that the species and the size segregation functions are a very useful addition to existing point process characteristics, since they relate nearest-neighbour characteristics to spatial scales. There are only few existing characteristics in point process statistics that combine neighbourhood and spatial scales (Wiegand and Moloney 2014). Our new segregation functions are nearest-neighbour characteristics that depend on distance and thus give valuable insights on how mingling and size differentiation effects decrease with distance. This information is crucial for understanding how local forest structure evolves and eventually

contributes to global forest structure. As part of this, morphology of the segregation functions correlation range and nonlinearity are of particular interest. Applying the new methods to the four example woodlands has demonstrated the extent of neighbourhood effects and that the decline of these effects largely differs between them. Of particular interest is also the correlation range of the two new functions, since it indicates the spatial extent of influence of local neighbourhood. From applying the two functions we have understood that local neighbourhoods in temperate woodlands in China, such as those that we studied, extend to 8–12 m.

Even more information on local neighbourhoods and their structure can be derived from applying species and size segregation functions simultaneously. From this we have learnt that the size segregation effect is often weaker than the species segregation effect. Also, the morphology of both function graphs is usually very similar to such an extent that we can consider this similarity as an allometric neighbourhood relationship. The mean allometric neighbourhood coefficient \tilde{m} clearly con-

Table 2 Shape characteristics of the estimated species segregation function $\hat{\Psi}(r)$ and the size segregation function $\hat{\Upsilon}(r)$. Intercept – from linear regression, expression of maximum size segregation; Slope – from linear regression, expression of how rapidly $\hat{\Psi}(r)$ and $\hat{\Upsilon}(r)$ tend towards a value of 0; Range – difference between average distance to first and 30th nearest neighbour, SSDL – sum of squares of deviation from linearity (Eq. 8)

Plot	Intercept		Slope		Range (m)	SSDL	
	$\hat{\Psi}(r)$	$\hat{\Upsilon}(r)$	$\hat{\Psi}(r)$	$\hat{\Upsilon}(r)$		$\hat{\Psi}(r)$	$\hat{\Upsilon}(r)$
Da	-0.101	-0.094	0.017	0.016	6.6	0.0185	0.0102
Jb	0.190	0.057	-0.016	-0.003	9.6	0.0036	0.0001
JSa	0.224	0.045	-0.026	0.000	5.0	0.0022	0.0001
XSc	0.328	0.024	-0.039	-0.003	6.8	0.0548	0.0021

firming the close relationship between the decline rates of both characteristics. This finding strongly supports the mingling-size hypothesis suggesting that trees with high spatial species mingling often tend to be larger sized trees with considerable size differentiation in their vicinity.

The two random labelling test methods applied in Fig. 2 revealed another important aspect of size-species relationships: While the restricted random labelling method,

which takes the existence of distinct species populations into account, is undoubtedly the better option for testing spatial mark independence, there were two sites, Jiaohe (plot b) and Jiulongshan (plot a) where the choice of test procedure mattered little. This begs the important question why some forest patterns are sensitive to the random labelling test method and others are not. In situations, where the envelopes greatly differ between tests, the non-spatial size-mark distributions greatly

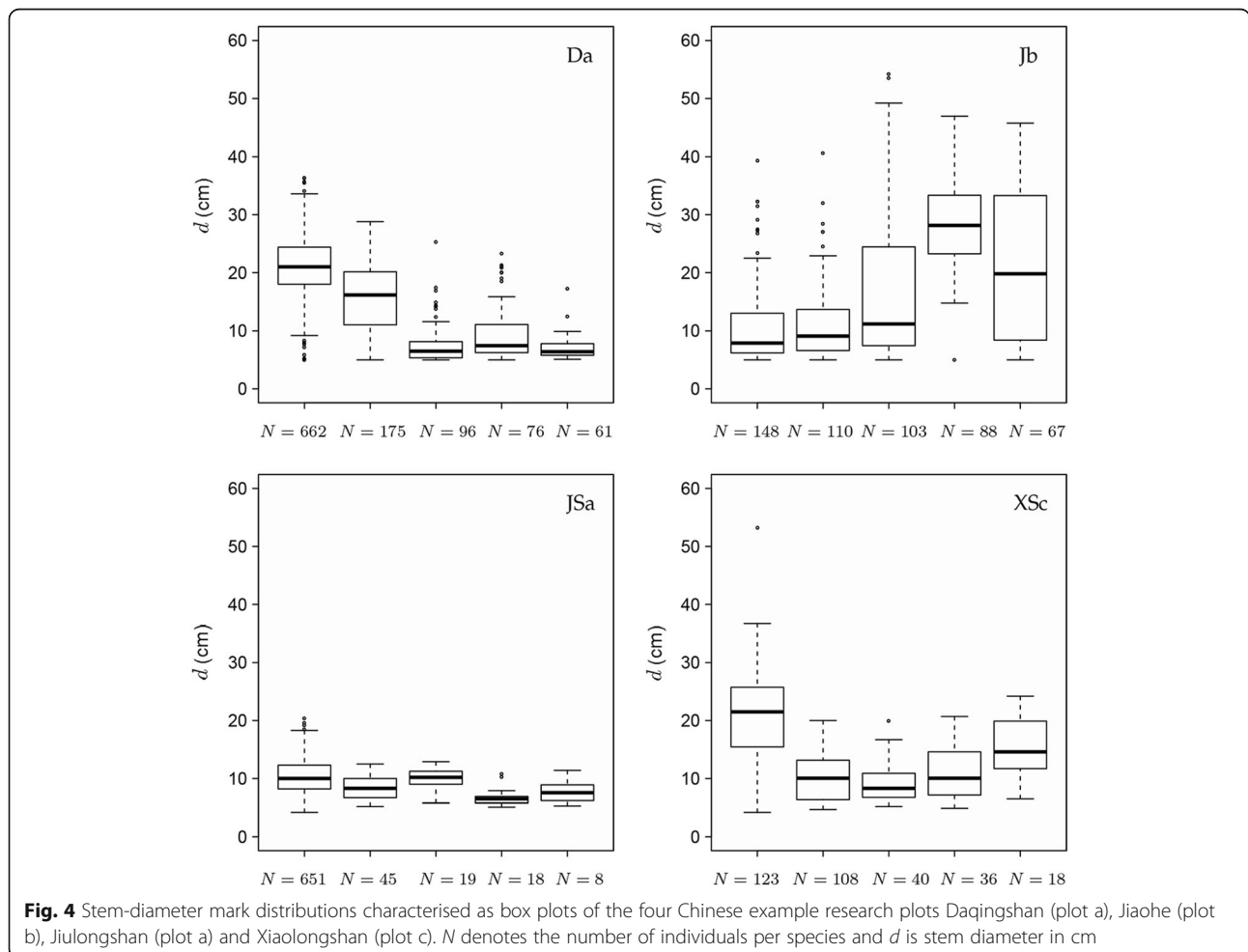


Fig. 4 Stem-diameter mark distributions characterised as box plots of the four Chinese example research plots Daqingshan (plot a), Jiaohe (plot b), Jiulongshan (plot a) and Xiaolongshan (plot c). N denotes the number of individuals per species and d is stem diameter in cm

differ between the most abundant species (Fig. 4). This in turn affects spatial size differentiation where local neighbourhoods include several different species, i.e. where attraction of different species is high.

Apparently this is the situation in Daqingshan (plot a). In this plot, unrestricted size permutations across species boundaries produce spatial size differentiation among nearest neighbours that does not match observed size differentiation, hence the envelopes of traditional random labelling are markedly different from the observed curves. Thus the strong effect of quite different sizes at close proximity at Daqingshan is mainly a consequence of species with different size-mark distributions mingling at this range. By contrast, the three other forest sites show a con-specific attraction, i.e. here the within-species size variability is more important than the between-species differences in size range, particularly at Xiaolongshan (plot c). However, at Jiaohe (plot b) and Jiulongshan (plot a) also the size distributions of the most abundant species are quite similar (see Fig. 4) which contributes to an outcome where spatial size differentiation in unrestricted random labelling is quite similar to that produced by restricted random labelling. Figure 4 confirms this explanation, except that in the case of our four example plots the size-mark distributions of the two most abundant species appear to be the most influential ones.

Thus the comparison between the two random-labelling test methods offers vital insights on how species and size segregation are related in mixed-species forests. It tells a particularly interesting story about the interaction between local and global forest structure that otherwise would not have emerged, if only the seemingly “correct” test were applied.

Acknowledgements

H.W. gratefully acknowledges a grant from the IUFRO-EFI Young Scientists Initiative that provided him with the opportunity to work three months at the Swedish University of Agricultural Sciences at Umeå (Sweden) in 2019. We thank Mari Myllymäki (LUKE, Finland) for interesting discussions about the random-labelling test results.

Authors' contributions

All authors equally contributed to all aspects of this work.

Funding

H.W. was partly supported by the Guangxi Innovation Driven Development Project (No. AA17204087–8). Z.Z. was funded by the National Natural Science Foundation of China (project No. 31670640).

Availability of data and materials

The datasets and programming code used for current study are available from the authors on request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Faculty of Forest Sciences, Department of Forest Ecology and Management, Swedish University of Agricultural Sciences SLU, Skogsmarksgränd 17, SE-901 83 Umeå, Sweden. ²College of Forestry, Guangxi University, Nanning 530004, China. ³Key Laboratory of Tree Breeding and Cultivation of National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Box 1958, Beijing 100091, China.

Received: 20 November 2019 Accepted: 26 February 2020

Published online: 19 April 2020

References

- Aguirre O, Hui GY, Kv G, Jiménez J (2003) An analysis of spatial forest structure using neighbourhood-based variables. *Forest Ecol Manag* 183:137–145
- Begon M, Townsend CR, Harper JL (2006) *Ecology. From individuals to ecosystems*. 4th edition. Blackwell publishing, Malden
- Comita LS, Queenborough SA, Murphy SJ, Eck JL, Xu K, Krishnadas M, Beckman N, Zhu Y (2014) Testing predictions of the Janzen-Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *J Ecol* 102:845–856
- Connell JH (1971) On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forests. den Boer PJ, Gradwell GR (eds) dynamics of populations. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands, pp 298–312
- Development Core Team R (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <http://www.r-project.org>.
- Ford ED (1975) Competition and stand structure in some even-aged plant monocultures. *J Ecol* 63:311–333
- Gadow KV (1993) Zur Bestandesbeschreibung in der Forsteinrichtung. [New variables for describing stands of trees.] *Forst und Holz* 48:602–606
- Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns. John Wiley & Sons, Chichester
- Janzen DH (1970) Herbivores and the number of tree species in tropical forests. *Am Nat* 104:501–528
- Liu JP, Chow SC, Hsieh TC (2009) Deviations from linearity in statistical evaluation of linearity in assay validation. *J Chem* 23:487–494
- Murphy SJ, Xu K, Comita LS (2016) Tree seedling richness, but not neighbourhood composition, influences insect herbivory in a temperate deciduous forest community. *Ecol Evol* 6:6310–6319
- Myllymäki M, Mrkvička T (2019) GET: global envelopes in R. [arXiv:1911.06583](https://arxiv.org/abs/1911.06583) [stat. ME]. <https://arxiv.org/abs/1911.06583>.
- Piao T, Comita LS, Jin G, Kim JH (2013) Density dependence across multiple life stages in a temperate old-growth forest of Northeast China. *Oecologia* 172: 207–217
- Pielou EC (1977) *Mathematical ecology*. John Wiley & Sons, New York
- Pommerening A, Grabarnik P (2019) *Individual-based methods of forest ecology and management*. Springer, Cham
- 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 Indic* 105:116–125
- Pommerening A, Uriá-Diez J (2017) Do large forest trees tend towards high species mingling? *Ecol Inform* 42:139–147
- 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 Inform* 56:101054
- Weiner J, Solbrig OT (1984) The meaning and measurement of size hierarchies in plant populations. *Oecologia* 61:334–336
- Wiegand T, Moloney KA (2014) *Handbook of spatial point pattern analysis in ecology*. CRC Press, Boca Raton
- Wills C, Condit R, Foster RB, Hubbell SP (1997) Strong density- and diversity-related effects help to maintain tree species diversity in a neotropical forest. *Proc Natl Acad Sci U S A* 94:1252–1257
- Wright SJ (2002) Plant diversity in tropical forests: a review of mechanics of species coexistence. *Oecologia* 130:1–14
- Yao J, Zhang X, Zhang C, Zhao X, Kv G (2016) Effects of density dependence in a temperate forest in northeastern China. *Sci Rep* 6:32844