Species associations in an old-growth temperate forest in north-eastern China

Xugao Wang¹, Thorsten Wiegand², Zhanqing Hao¹*, Buhang Li¹, Ji Ye¹ and Fei Lin¹

¹Institute of Applied Ecology, Chinese Academy of Sciences, PO Box 417, Shenyang 110016, China; and ²Department of Ecological Modelling, UFZ Helmholtz Centre for Environmental Research-UFZ, PF 500136, D-04301 Leipzig, Germany

Summary

1. Studying the spatial pattern of plants may provide significant insights into processes and mechanisms that maintain species richness. We used data from a fully mapped 25-ha temperate forest plot at Changbaishan (CBS), north-eastern China, to conduct a community-wide assessment of the type and frequency of intra- and interspecific spatial association patterns. We analysed complex scale effects in the patterning of large trees of 15 common species. First, we tested for overall spatial patterning at 6, 30 and 50 m neighbourhoods and classified the types of bivariate association patterns at these spatial scales (analysis 1). We then explored small-scale (0–20 m) association patterns conditioning on the larger-scale pattern (analysis 2) and tested for positive large-scale (50–250 m) association patterns (analysis 3).

2. Analysis 1 provided ample evidence for non-random spatial patterning, and the type and frequency of spatial association patterns changed with scale. Trees of most species pairs co-occurred less than expected by chance and positive associations were rare in local neighbourhoods. Analysis 2 revealed a separation of scales in which significant small-scale interactions faded away at distances of 10–15 m. One third of all species pairs showed significant and mostly negative bivariate small-scale association, which occurred more often than expected by chance between species sharing attributes such as family, fruit type and habitat association. This suggests the occurrence of competitive interactions. Analysis 3 showed that only 8% of all species pairs co-occurred at large scales.

3. Comparison of our results with an analogous study conducted in the species-rich tropical forest at Sinharaja, Sri Lanka, revealed several structural similarities including the dominance of segregation and partial overlap in the overall patterning (analysis 1) and the separation of scales (analysis 2). However, species pairs at CBS showed considerably more significant negative small-scale associations (31% vs. 6% at Sinharaja).

4. Synthesis. The techniques presented here allow for a detailed analysis of the complex spatial associations in species-rich forests and have the potential to reveal indicative patterns that may allow researchers to discriminate among competing hypotheses of community assemblage and dynamics. However, this will require comparative studies involving a large number of plots.

Key-words: Changbaishan (CBS), coexistence, pair correlation function, point pattern analysis, spatial segregation, species association, temperate forest

Introduction

A central aim in ecology is to understand processes and mechanisms that control the distribution and abundance of species (Ricklefs 1990). The importance of spatial patterns in this respect has been increasingly recognized (e.g. Watt 1947; Pacala 1997; Tilman & Kareiva 1997; Amarasekare 2003), and investigation of the effect of spatial interactions on population...
patterns can be explained by substantially different theories (e.g. Chave 2004; Gilbert & Lechowicz 2004; Bell 2005). For example, the observed patterning of plants may be explained by environmental niches and trade-offs among species in dispersal and competitive ability (e.g. Tilman 2004). However, models based on neutral theories, where ecological drift is the only process occurring besides stochastic but limited dispersal and speciation (Bell 2001; Hubbell 2001; Chave 2004), can also form patterns of distribution and abundance similar to those found in nature (Mouquet & Loreau 2003; Chave 2004; Purves & Pacala 2005).

One possible reason for the current inability to decide between competing theories is that the patterns used, basically the species abundance distribution and the species–area relationship, have low discriminatory power (e.g. Chave 2004; Purves & Pacala 2005). McGill et al. (2006) pointed to a need to evaluate additional and independent predictions, such as spatial patterns of the location of plants within a community (Hubbell et al. 2001; Wiegand, Gunatilleke & Gunatilleke 2007a; Wiegand et al. 2007b). These point patterns are the results of all mechanisms and processes that have affected the fate of the plants during their life and constitute an ‘ecological archive’ which may contain encoded information on the underlying processes (Hubbell et al. 2001; Wiegand et al. 2003; Grimm et al. 2005; McIntire & Fajardo 2009; Wiegand, Huth & Martinez 2009). The observed spatial patterns may be used for model selection in dynamic and spatially explicit simulation models to infer processes (e.g. Wiegand et al. 2003; Grimm et al. 2005). However, a prerequisite of this research program to decide between competing theories (Levin 1992; Grimm et al. 2005) is identification of indicative patterns in the spatial data. (Note that we use ‘indicative pattern’ to denote patterns that have the potential to reveal underlying biological mechanisms. This differs from the spatial point pattern literature in which ‘pattern’ denotes the spatial pattern of the coordinates of ecological objects within an observation window.) Promising candidates for indicative patterns that could reveal information on community structure include the interspecific spatial associations between pairs of species. However, the scale-dependent nature of spatial processes and mechanisms involved in spatial pattern generation makes examination of these relationships a difficult undertaking. It cannot be expected that all spatial processes and mechanisms act on the same characteristic spatial scales (Wiegand, Huth & Martinez 2009), therefore proper analysis of spatial association patterns must take into account the multi-scale nature of the patterns. For example, two species may be segregated at a larger scale (i.e. they occupy largely disjoint patches), but if plants of the two species are direct neighbours, they may show a positive association. This positive, small-scale association may be overlooked if the analysis is not conducted conditionally on the otherwise overpowering large-scale distribution pattern. Only a few attempts have been made to characterize the spatial association between species within species-rich plant communities (but see Kubota, Kubo & Shimatani 2007; Lieberman & Lieberman 2007; Wiegand, Gunatilleke & Gunatilleke 2007a; Wiegand et al. 2007b; Perry et al. 2009; Illian, Moller & Waagepetersen 2009). However, such studies are of prime importance for advancing our understanding of spatial processes and mechanisms on coexistence in species-rich plant communities.

The spatial segregation hypothesis is an important example of a hypothesis related to species coexistence that involves spatial association between pairs of species (Pacala 1997; Pacala & Levin 1997). Intraspecific aggregation leads to interspecific segregation whereby an average plant in the community will compete mostly locally with con-specific plants. As a result, competitively superior species become suppressed, which prevents (or retards) the elimination of competitively inferior species (Stoll & Prati 2001), thereby promoting species diversity (Kareiva 1990; Tilman 1994; Rees, Grubb & Kelly 1996; Pacala 1997; Stoll & Weiner 2000). Motivated by the segregation hypothesis, we will study for many species pairs how often plants of the second species are located within a specified neighbourhood around plants of the first species. Comparison of such spatial structure of several plant communities with contrasting characteristics (e.g. tropical vs. temperate forest) may finally allow us to identify indicative patterns in the data to distinguish among competing theories of species coexistence (Wiegand et al. 2003; Grimm et al. 2005; McIntire & Fajardo 2009; Wiegand, Huth & Martinez 2009). In this study, we take a step in this direction.

We focused here on detailed point pattern analysis based on maps of the locations of all larger trees in a given forest plot (Kenkel 1988; Duncan 1991; Hoshino, Nishimura & Yamamoto 2001; Kubota, Kubo & Shimatani 2007; Wiegand, Gunatilleke & Gunatilleke 2007a; Wiegand et al. 2007b). We analysed point pattern data from the completely mapped 25-ha Changbaishan (CBS) plot located in the Changbai Nature Reserve, China. The large size of the CBS plot provided the unique opportunity to investigate community-wide interspecific associations for a temperate forest (we included 15 species). To compare the results of the CBS plot with those of the 25-ha plot of tropical forest at Sinharaja, Sri Lanka (Wiegand, Gunatilleke & Gunatilleke 2007a), we used the methodology proposed by Wiegand, Gunatilleke & Gunatilleke (2007a), but conducted additional analyses to relate our results to the species characteristics.

Our overall goal was an assessment of the type and frequency of intra- and interspecific spatial association patterns in the CBS forest. Because of the complex scale effects, we divided this goal into three analyses. First, we tested for overall non-random spatial patterning in uni- and bivariate patterns and classified the types of bivariate association patterns at various spatial scales (analysis 1). Under the segregation hypothesis, we expected mostly aggregated univariate patterns and segregation or partial overlap between species pairs in local neighbourhoods. In analysis 2, we tested selectively for small-scale association patterns (0–30 m) conditional on the observed large-scale (i.e. > 30 m) variations in the intensity (that may be caused by dispersal and/or association to smoothly varying environmental factors). We expected a high proportion of negative associations (caused by competitive interactions) among species that share ecological attributes.
Finally, we used a null model that tests selectively for positive large-scale association patterns (50–250 m; analysis 3). Species belonging to the same successional stage should occupy the same larger-scale patches.

Materials and methods

STUDY AREA

The study area is located in the Changbai Nature Reserve, which extends along the border of China and North Korea from 127°42′ to 128°17′ E and 41°43′ to 42°26′ N (Fig. 1). It is one of the largest biosphere reserves in China and has been spared from logging and other severe human disturbances since it was established in 1960. The Changbai Nature Reserve joined the World Biosphere Reserve Network under the UNESCO Man and the Biosphere Programme in 1980. The CBS plot is located in the core zone of the Changbai Nature Reserve and is representative of broad-leaved Korean pine mixed forest, the most common vegetation type in species composition and ecosystem structure for the region.

The 500 × 500 m CBS plot was established in summer 2004 and is the northernmost plot of the China Forest Biodiversity Monitoring Network (http://www.cfbiodiv.org) initiated by the Chinese Academy of Sciences (CAS) in collaboration with the Center for Tropical Forest Science (CTFS). Mean elevation in the plot is 801.5 m; the elevation ranges from 791.8 to 809.5 m. The soil is classified as dark brown forest soil. Mean annual precipitation is approximately 700 mm; most of this occurs from June to September (480–500 mm). Mean annual temperature is 2.8°C, with a January mean of −13.7°C, and a July mean of 19.6°C (Yang & Li 1985). All trees at least 1 cm in diameter at breast height (d.b.h.) were mapped and identified to species; their geographic coordinates were recorded following a standard field protocol (Condit 1998). The total number of living individuals in the first census of 2004 was 38 902, consisting of 52 species, 32 genera and 18 families (Hao et al. 2008).

FOREST CHARACTERISTICS

The main tree species at CBS include Pinus koraiensis, Tilia amurensis, Quercus mongolica, Fraxinus mandshurica, Ulmus japonica and Acer mono. Unlike tropical rain forests without obvious dominant species, there were eight species with more than 1000 individuals, which accounted for 83.4% of the total individuals in the plot. Mean stand density was 1556 living trees per ha. Mean basal area was 43.2 m² ha⁻¹ (Hao et al. 2008; Wang et al. 2009). In this study, the spatial pattern of 10 527 non-juvenile trees with d.b.h. ≥ 10 cm was analysed. To obtain a sufficiently large sample size for the point pattern analyses, we included 15 species each with more than 50 such trees (Table 1), making up 97.7% (10 313 trees) of the total. The ecological characteristics of these species are also shown in Table 1. Most species in the studied forest were clustered at some spatial scale (Hao et al. 2007; Wang et al. 2008b) and about 60% of 35 species studied showed significant habitat association (J. Ye, unpubl. data).

Forest vegetation at the CBS plot is largely the result of natural succession (Zhao 1981), which is now at the late-succession stage. The mean age of the canopy trees is about 300 years. One common coexistence hypothesis is niche differentiation in temperate forests (Dai et al. 2004), resulting from heterogeneous environmental conditions (e.g. differences in topographical and soil factors). However, Wang et al. (2008a) showed that pure topographical factors were not sufficient to explain the species diversity pattern at the CBS plot. Another coexistence hypothesis is that natural disturbance (e.g. wind) creates canopy openings and consequently initiates regeneration, maintaining early successional species by local secondary succession, thus enabling the coexistence of tree species (Zang et al. 1998). In addition, species competition may be important for the maintenance of species diversity at small scales (Wang et al. 2008a).

Statistical analyses

POINT PATTERN ANALYSIS: SUMMARY STATISTICS

To quantify the spatial patterns found at the CBS forest ‘from the plant’s eye view’, we used recent techniques of spatial point pattern analysis (Ripley 1981; Stoyan & Stoyan 1994; Diggle 2003; Illian et al. 2008) and summary statistics such as the pair-correlation function (Stoyan & Stoyan 1994), Ripley’s (1981) K-function and the distribution function of nearest neighbour distances (Diggle 2003). The bivariate pair-correlation function g_{12}(r) can be estimated using the quantity λ_{12}(g_{12}(r)), which is the mean density of trees of species 2 at distance r away from trees of the focal species 1, whereby λ_{2} is the mean density of trees of species 2 in the whole study area. Ripley’s (1981) K-function K(r) is the cumulative version of the pair-correlation function, i.e. the quantity λ_{12}K_{12}(r) is the average number of trees of species 2 within distance r from trees of the focal species 1. The univariate K- and pair-correlation functions follow intuitively (Wiegand & Moloney 2004), but the focal point is not counted. To describe additional characteristics of the spatial patterns, we used the bivariate distribution function D_{12}(y) which gives the fraction of trees of the focal species 1 that have their nearest species 2 neighbour within distance y (Diggle 2003; Illian et al. 2008). Note that D(y) is often referred to as G(v) in the literature (e.g. Diggle 2003), but we have adopted the notation of the recent textbook by Illian et al. (2008).

Fig. 1. The location and contour map of the 25-ha (500 × 500 m) Changbai temperate plot.

The \( g \)-, \( K \)- and \( D \)-statistics are usually interpreted for homogeneous patterns to indicate interactions among pairs of points. In this case, they reflect properties of a ‘typical tree’ of the pattern (Illian \textit{et al.} 2008). However, the patterns at our study site are certainly not all homogeneous, which means that a typical tree of a pattern may not exist. Instead, we interpreted the \( g \)- and \( K \)-functions as averages taken over all trees of the focal pattern and designed our analyses and null models so as to account for potential heterogeneities.

**POINT PATTERN ANALYSIS: TESTING SIGNIFICANCE OF PATTERNS AGAINST A NULL MODEL**

For a given species or species pair, we contrasted the observed summary statistics to that expected under an appropriate null model. We used a Monte-Carlo approach to test for significant departures from the null models. Each of the \( n = 199 \) simulations of a point process underlying the null model generates a summary statistic [e.g. a pair-correlation function \( g_2(r) \); simulation envelopes with \( \alpha = 0.05 \) were calculated from the 5th highest and lowest values of \( g_2(r) \) in the 199 simulations (Stoyan & Stoyan 1994)]. Significant departure from the null model occurred at scale \( r \) if the test statistic was outside the simulation envelopes. This approach allowed us to assess scale effects approximately for illustrative purposes and to determine the type of significant effect.

However, to avoid problems due to simultaneous inference (e.g. Loosmore & Ford 2006), we evaluated the overall ability of a given null model to describe the data by means of a goodness-of-fit test (GoF; Diggle 2003; Illian \textit{et al.} 2008). This test reduces the scale-dependent information contained in the summary statistics into a single test statistic, \( u_i \), which represents the total squared deviation between the observed pattern and the theoretical results across the distances of interest (i.e. a Cramer–von Mises type statistic as e.g. used in Plotkin \textit{et al.} 2000). The \( u_i \) values were calculated for the observed data \( (i = 0) \) and for the data created by the \( (i = 1...199) \) simulations of the null model, and the rank of \( u_0 \) among all \( u_i \) values was determined. If the rank of \( u_0 \) was larger than 190, the data showed a significant departure from the null model (across the distances of interest) with error rate \( \alpha = 0.05 \).

**ANALYSIS 1: DETECTION OF OVERALL NON-RANDOM PATTERNING**

**Univariate case**

To detect overall departure from randomness, we confronted our data with the null model of complete spatial randomness (CSR; Wiegand & Moloney 2004). To quantify departures from the null model, we used the pair-correlation function. Note that this test is sensitive to effects from heterogeneous environment, dispersal and tree interactions.

**Bivariate case**

Our basic question was conceptually simple: we wanted to know how the trees of a given species 2 were distributed within local neighbourhoods of the trees of a focal species 1. Did they occur on average more (or less) frequently within the neighbourhoods than expected by chance alone, and was this association homogeneous or heterogeneous? In the heterogeneous case this distribution varies substantially among trees of the focal species, e.g. some species 1 trees may have many species 2 neighbours but other species 1 trees have few species 2 neighbours. To distinguish the various types of spatial associations from those that may arise purely by chance, we compared the observed bivariate point patterns with a null model in which the locations of the focal species remained unchanged, but trees of species 2 were distributed randomly and independently of the locations of species 1 (i.e. CSR). Clearly, testing against CSR is often not very informative (Wiegand & Moloney 2004); however, we used this test to quantify and categorize the overall bivariate spatial associations based on a scheme developed by Wiegand, Gunatilleke & Gunatilleke (2007a; see section below). The scheme uses the bivariate \( K_{12}(r) \) and \( D_{12}(y) \) as test statistics and distinguishes four significant types of spatial associations that may occur between two (possibly heterogeneous) patterns (see section below).

A ‘not significant’ type arises if neither \( K_{12}(r) \) nor \( D_{12}(y) \) show significant departures from the CSR null model (as measured by the GoF tests).
Scheme to characterize bivariate associations

The spatial association between two species can be characterized by the distribution function $P_{12}(n, r)$ that gives the probability of finding $n$ trees of species 2 within neighbourhoods of radius $r$ around trees of species 1. If the point configurations between pairs of trees of the two species are the same all over the study plot except for stochastic variation (i.e. homogenous patterns), we do not need the full distribution $P_{12}(n, r)$ to describe the association between the two species. In this case the mean of $P_{12}(n, r)$ with respect to $n$ suffices, which is given by $\lambda_2 K_1(r)$. However, we cannot expect that a typical bivariate point configuration exists at the CBS plot because the known locations of trees randomly within a neighbourhood of radius $R$ (i.e. a heterogeneous Poisson null model; Wiegand, Gunatilleke & Gunatilleke 2007a; Wiegand et al. 2007b). This displacement removes potential patterns for distances $r < R$, but it leaves the larger-scale patterns untouched. Contrasting the observed pattern to realizations of this null model will therefore detect only small-scale effects.

While this analysis can be conducted for any displacement distance $R$, it is desirable to use a distance which is likely to separate biological effects. In general, it is expected that direct interactions among larger trees only occur within a limited spatial separation (i.e. $< 30$ m). For example, Hubbell et al. (2001) found that the neighbourhood effects of conspecific density on survival disappeared within approximately 12–15 m of the focal plant. Several other studies using individual-based analyses of local neighbourhood effects on tree growth and survival confirmed this result (e.g. Uriarte et al. 2004, 2005; Stoll & Newbery 2005) suggesting that direct plant–plant interactions in forests may fade away at larger scales. We therefore used a separation distance of $R = 30$ m (Wiegand, Gunatilleke & Gunatilleke 2007a).

An interesting question is whether separation of scales occurred: this can be tested in a simple way. Because the heterogeneous Poisson process conditions on the spatial structure for scales larger than 30 m, it is only able to indicate significant effects at scales smaller than 30 m. In cases without separation of scales, we expect therefore that the frequency of significant effects, taken over all pairs of species, should fade away smoothly at 30 m. However, if small-scale effects operate only over a short range (i.e. $r \leq 30$ m), the frequency of significant effects should disappear well below the threshold of 30 m.

### Univariate case

The null model was a heterogeneous Poisson process based on a non-parametric intensity estimate using the Epanechnikov kernel with a bandwidth of $R = 30$ m (see e.g. Wiegand, Gunatilleke & Gunatilleke 2007a). This null model is basically equivalent to randomly displacing the tree locations of the species of interest within a radius of 30 m. We used the univariate pair-correlation function $g(r)$ as test statistic and studied species associations with a spatial resolution of 2 m up to 40 m. This is a sufficiently fine resolution to capture detailed variation in the pair-correlation function over the range of scales of interest, but it is coarse enough to yield feasible computation time for the high number of analyses required. We selected 40 m as the maximum scale of analysis in order to check that the heterogeneous Poisson process does indeed only depict significant effects at scales smaller than 30 m. Thus, this analysis has several types of ‘spatial resolutions’: the kernel bandwidth, the grain at which we estimate $g(r)$, and the maximal scale up to which we estimate $g(r)$.

### Bivariate case

The null model was analogous to that of the univariate case, but we left the locations of species 1 untouched and distributed the most trees of the highly clustered species 1. Type IV associations will rarely occur (Wiegand, Gunatilleke & Gunatilleke 2007a).

### ANALYSIS 2: ISOLATION OF SMALL-SCALE EFFECTS

Acknowledging the multi-scale nature of the spatial association patterns, we selectively studied the small-scale association pattern by using a null model which randomizes the data conditionally on the observed large-scale pattern. In practice, this can be done by displacing the known locations of trees randomly within a neighbourhood with radius $R$ (i.e. a heterogeneous Poisson null model; Wiegand, Gunatilleke & Gunatilleke 2007a; Wiegand et al. 2007b). This displacement removes potential patterns for distances $r < R$, but it leaves the larger-scale patterns untouched. Contrasting the observed pattern to realizations of this null model will therefore detect only small-scale effects.

The two-axis scheme allows for four fundamental types of bivariate association. In the case of ‘segregation’ (type I), both the average number of neighbours within distance $r$ and the proportion of nearest neighbours within distance $r$ are smaller than expected [i.e. $\bar{M}(r) < 0$ and $\bar{P}(r) < 0$]. In the case of ‘mixing’ (type III), both are larger than expected [i.e. $\bar{M}(r) > 0$ and $\bar{P}(r) > 0$]. In the case of ‘partial overlap’ (type II), the mean number of trees of species 2 within neighbourhoods of radius $r$ around trees of species 1 is larger than would be expected according to the null model [i.e. $\bar{M}(r) > 0$], and the probability that a tree of species 1 has no neighbour of species 2 is smaller than expected [i.e. $\bar{D}_{12}(r) < 0$]. This configuration is only possible for heterogeneous patterns if some trees of species 1 are surrounded at the given neighbourhood $r$ by many trees of species 2 but others are surrounded by few (or no) trees of species 2. Finally, in the case opposite to partial overlap (type IV), trees of species 1 are highly clustered and trees of species 2 overlap the cluster of species 1. As a result, the mean number of species 2 neighbours is smaller than expected [i.e. $\bar{M}(r) < 0$], but the probability to have the nearest neighbour of species 2 within distance $r$ is larger than expected [i.e. $\bar{D}_{21}(r) > 0$]. This is because a few trees of species 2 are in fact the nearest neighbour of
trees of species 2 using the heterogeneous Poisson process described above. We used the bivariate pair-correlation function $g_{12}(r)$ as test statistic. Again we studied species associations with a spatial resolution of 2 m up to 40 m. Because the association may be asymmetric, we tested all pairs, i.e. both species 1 versus species 2 and species 2 versus species 1.

**LARGE-SCALE ASSOCIATIONS (ANALYSIS 3)**

Approximately 60% of species at the CBS plot showed significant habitat association (J. Ye, unpubl. data); it is therefore interesting to find out whether and how often two species share roughly the same areas within the 25-ha study plot. An appropriate null model for this question leaves the trees of species 1 unchanged and has the trees of species 2 distributed in accordance with the large-scale pattern of species 1. We implemented this null model as a heterogeneous Poisson process whereby the locations of the trees of species 2 were randomized in accordance with the intensity function of species 1. We used an Epanechnikov kernel with a bandwidth of 50 m to estimate the intensity function of species 1 (Stoyan & Stoyan 1994; see Wiegand & Moloney 2007a). Note that the bandwidth of 50 m was selected to capture effects of larger scale association.

**SETTINGS FOR ANALYSES AND GOF TEST**

For analyses 1 and 2, we used a spatial resolution $dw = 2$ m; we used 5 m for analysis 3. Note that the scale $r = 0$ m for the pair correlation function refers to effects ranging from 0 m to half of the spatial resolution, and scale $r$ refers to effects ranging within scales $r - dw/2$ and $r + dw/2$. For analysis 2, we conducted the GoF test as in Wiegand et al. (2007b) over the distance interval of 0–20 m because we expected significant effects to occur primarily over this range of scales. For analysis 3, we conducted the GoF test over the full range of scales (i.e. 50–250 m) at which we expected the null model to be met. We retained data sets for further analysis only if the observed $P$ value of the GoF test was smaller than 0.05 (analyses 2) or larger than 0.05 (analysis 3). Note that analysis 3 assumes a specific association, thus data sets with $P < 0.05$ are cases where the null model was met, that is, indicating significant large-scale association. All analyses were performed with the Programata software (Wiegand & Moloney 2004). Details on edge correction and estimators of the $g_{12}(r)$, $g(r)$, $K(r)$ and $K_{12}(r)$ can be found in Wiegand & Moloney (2004). Following Diggle (2003), we calculated the $D_{12}$ function without edge correction.

**Results**

**ANALYSIS 1: DETECTION OF OVERALL NON-RANDOM PATTERNING**

**Univariate case**

The 15 species studied at the CBS plot showed stem maps with diverse spatial patterns (Fig. 2). Use of the CSR null model revealed that all species showed significant clustering at some or all scales, thus confirming our expectation.

**Bivariate case**

The various types of non-random effects in interspecific association were not equally distributed among the 210 species pairs analysed, and their relative frequency depended on scale (Fig. 3; Fig. S1 in Supporting Information). The most notable result is that roughly half of the species pairs showed segregation at all three spatial scales analysed (113, 115 and 110 cases at scales 6, 30 and 50 m, respectively). Thus, trees of different species co-occur less often than expected by chance. Non-significant associations (i.e. no departure from the null model detected by the two test statistics $K_{12}$ and $D_{12}$) were relatively frequent in small neighbourhoods (60 cases at 6 m), but stabilized for larger neighbourhoods at a frequency of approximately 20% (38 and 34 cases at scales of 30 and 50 m, respectively). Interestingly, partial overlap was rare in the 6-m neighbourhood (nine cases; Fig. 3a), but more frequent in the 30 and 50-m neighbourhoods (42 and 53 cases, respectively; Fig. 3b,c). Mixing occurred in 28 cases at the 6-m scale but only 13 and 10 times at the 30 and 50-m scale respectively. The dominance of segregation and partial overlap confirms our expectation based on the segregation hypothesis. Figure S1 indicates that most changes in the relative frequency of the different bivariate association types occur at scales smaller than 20 m.

**ANALYSIS 2: ISOLATION OF SMALL-SCALE EFFECTS**

**Univariate case**

The GoF test revealed for nine of the 15 species (60%) significant departures from the heterogeneous Poisson null model at the scale of 0–20 m. Seven species (47%) showed small-scale aggregation (Fig. 4a,b), the two species *Acer mono* (Fig. 4c) and *Quercus mongolica* (Fig. 4d) showed significant regularity at scale $r = 0–2$ m, and the other six species followed the null model.

In order to roughly estimate the effect of scale on species spatial pattern (i.e. regularity and aggregation), we counted the number of species (using only species where the rank of GoF test was $> 190$) for each scale $r$ where the pair-correlation function was above or below the fifth-highest or fifth-lowest value of the pair-correlation function in the 199 Monte Carlo simulations. The frequency of aggregation peaked at a scale between 0 and 4 m, and aggregation effects disappeared at scales $r > 10$ m (Fig. 5a). Figure 5a reveals repulsion at intermediate scales larger than 10 m. Thus, the clusters have a tendency to be regularly distributed. Most significant effects already disappeared at 10 m, confirming that separation of scales occurred. Note that strong aggregation at small scales suggests repulsion at larger scales when tested with the heterogeneous Poisson null model. The regularity effects at larger...
scales (e.g. > 20 m) depicted in Fig. 5a may be primarily due to this effect.

**Bivariate case**

At a total of 15 × 14 = 210 bivariate point pattern analyses for all pairs of the n = 15 species were performed (Fig. 6; Appendix S1, Table S1). The GoF test revealed significant departures from the null model for 65 species pairs (31% of all cases); in 55 cases, the small-scale association was negative (repulsion), thus confirming our expectations, and in the other 10 cases, the small-scale association was positive (attraction). However, note that approximately five of the 10 cases of significant attraction may arise by chance because this analysis has a

Fig. 2. Classification of large-scale associations within the CBS plot at a 30-m scale. Diagram (a) shows the allocation of the large-scale association of the 210 species pairs based on the classification axes defined in eqn (1). Diagrams b, c, d and e represent examples for type II association with partial overlap, type III association with mixing, type I association with segregation and type IV association, respectively. Solid circle = species 1, open circle = species 2.

Fig. 3. Overall assessment of non-random interspecific patterns at the neighbourhood scale r = 6, 30 and 50 m. The circle symbols indicate the ‘not significant’ associations (i.e. no departure from the null model detected by \( K_{12} \) and \( D_{12} \)). Other conventions are as in Fig. 2a.
The rank of the GoF test correlated strongly and positively with the number of trees of the two species ($r_{SP} = 0.52; P < 0.01$), moderately and positively with the number of trees of the focal species ($r_{SP} = 0.38; P < 0.01$), and negatively with clustering of the focal species ($r_{SP} = -0.39; P < 0.01$; Appendix S1, Table S2). As expected, we found that negative associations occurred more frequently among species with the same attribute family (which was in our case the same as genus), fruit type and habitat associations (Appendix S1, Table S3). Positive associations occurred less frequently among species with the same attribute family and more frequently for those species occupying the same canopy layer (Appendix S1, Table S3).

We also counted, for each scale $r$, the number of species for which the pair-correlation function was above or below the simulation envelopes (using only species where the rank of GoF test was $> 190$). Repulsion occurred more frequently than attraction; pair frequencies for repulsion and attraction peaked at the 0 and 2-m scales with 49 and 9 pairs, respectively (Fig. 5b). Significant effects were rare at scales $r > 10$ m, again demonstrating separation of scales. If no separation of scales was present, significant effects would only gradually disappear at the neighbourhood radius of $R = 30$ m used for the heterogeneous Poisson null model.

**Large-scale associations (Analysis 3)**

The GoF test revealed that 17 species pairs (8.1%) followed the null model at the distance interval 50–250 m (compared to 3.3% at Sinharaja), i.e. they showed significant large-scale association (see Appendix S2, Fig. S2). However, significant effects were only symmetric for one species pair (Ulmus japonica – Maackia amurensis; see Appendix S2). Note that the specific null model tested here is also met if species 2 is only distributed inside a subarea of the area broadly occupied by species 1 (see Appendix S3). Thus, large-scale associations were mostly of the partial overlap type.

Twelve of the 15 species studied at the CBS plot showed significant large-scale association with at least one other species (compared with 42 of 46 species at Sinharaja). In the role of species 1, the shade-tolerant canopy species *Tilia amurensis* showed significant association with five species followed by the shade-tolerant midstorey species *Acer mono* with four associations. The shade-tolerant midstorey species *Acer tegmentosum* showed, in the role of species 2, four associations to other species. The species without significant large-scale association to any other species were *Pinus koraiensis*, *Tilia mandshurica* and *Phellodendron amurense*. Relating significant large-scale association to species properties, however, did not reveal such clear rules as for analysis 2. The reason for this is probably that large-scale association patterns. The figures show the number of species (univariate) and species pairs (bivariate) where the observed pair-correlation function for a given scale is outside the Monte Carlo simulation envelopes, being the fifth-lowest and fifth-highest $g(r)$ values of the 199 simulations of the null model.
ger-scale associations may depend in a complex way on factors such as habitat association, dispersal characteristics, historic events and successional dynamics.

Discussion

In this study, we conducted a comprehensive spatial pattern analysis to assess species associations among large trees of 15 common species at the fully mapped 25 ha CBS temperate forest plot in north-eastern China. These species comprised 97.7% of all trees at the study plot with d.b.h. larger than 10 cm. Our study revealed marked differences to findings of a similar study of a hyper-diverse tropical forest (Wiegand, Gunatilleke & Gunatilleke 2007a). Most importantly, we found that one third of all species pairs at the CBS plot showed significant and mostly negative small-scale associations. Negative associations occurred more frequently than expected by chance between species which shared one or more attributes such as family, fruit type or habitat association. This suggests that the spatial pattern of larger trees showed signatures of competitive effects between species sharing ecological attributes. Analogous analyses of data from tropical forests at Sinharaja (Sri Lanka) and Barro Colorado Island (Panama) revealed significant small-scale associations (Wiegand, Gunatilleke & Gunatilleke 2007a; Table 2). Thus, individuals of different species show a clear tendency to ‘avoid’ each other, both at small and larger scales.

**THE SEGREGATION HYPOTHESIS**

Spatial segregation is a mechanism known from theoretical models to promote coexistence (Pacala 1997; Pacala & Levin 1997; Chesson 2000). The underlying mechanism is that spatial segregation among species decreases the probability of interspecific encounters, with the effect that the importance of interspecific competition decreases relative to intraspecific competition and competitive exclusion is retarded. We found that all 15 species were significantly aggregated and that spatial segregation and partial overlap were the most common bivariate association types at the CBS plot. Thus, trees of different species have a tendency to avoid each other. This observation is in line with the segregation hypothesis. Suppression of competitively superior species by this spatial configuration prevents...
Table 2. Comparison of results between the temperate forest at Changbaishan (CBS) and the tropical Sinharaja forest. Results for Sinharaja are taken from Wiegand, Gunatilleke & Gunatilleke (2007a) or were obtained after reanalysis of the raw data. Given are the percentages of cases in a given category. At the CBS plot, we analysed 15 species and 210 species pairs; at Sinharaja 46 species and 2070 species pairs

(a) Bivariate association type (analysis 1)*

<table>
<thead>
<tr>
<th>Study site</th>
<th>Segregation (type I)</th>
<th>Partial overlap (type II)</th>
<th>Mixing (type III)</th>
<th>Type IV</th>
<th>No sig. effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBS</td>
<td>54.8%</td>
<td>20%</td>
<td>6.2%</td>
<td>0.9%</td>
<td>18.1%</td>
</tr>
<tr>
<td>Sinharaja</td>
<td>50.2%</td>
<td>34%</td>
<td>5.4%</td>
<td>0.1%</td>
<td>10.3%</td>
</tr>
</tbody>
</table>

(b) Univariate analysis 2

<table>
<thead>
<tr>
<th></th>
<th>Small-scale aggregation</th>
<th>Small-scale regularity</th>
<th>No sig. effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBS</td>
<td>13%</td>
<td>40%</td>
<td></td>
</tr>
<tr>
<td>Sinharaja</td>
<td>6%</td>
<td>48%</td>
<td></td>
</tr>
</tbody>
</table>

(c) Bivariate analysis 2

<table>
<thead>
<tr>
<th></th>
<th>Small-scale attraction</th>
<th>Small-scale repulsion</th>
<th>No sig. effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBS</td>
<td>26%</td>
<td>69%</td>
<td></td>
</tr>
<tr>
<td>Sinharaja</td>
<td>3%</td>
<td>94%</td>
<td></td>
</tr>
</tbody>
</table>

*The ‘no significant effect’ type showed for the two axes $M(r)$ and $P(r)$ no significant departure from the null model, and the four types I to IV are defined by positive or negative departures from the null model in the $M(r)$ and/or $P(r)$ axis (see Fig. 2a).

(or retards) the elimination of competitively inferior species (Stoll & Prati 2001). This effect can tip the balance from competitive exclusion to coexistence and promote species diversity (Kareiva 1990; Tilman 1994; Rees, Grubb & Kelly 1996; Pacala 1997; Stoll & Weiner 2000).

The observed segregation may also explain why only 31% of all species pairs showed significant small-scale associations (analysis 2). This hypothesis is supported by our finding that significant small-scale effects were more likely if the two species were more abundant. This result is somewhat to be expected because the power of the GoF test (type-II error rate) is a function of abundance (e.g. Plotkin et al. 2000). This in turn is because fewer direct encounters between a species pair at a given spatial scale $r$ will weaken the power of the pair-correlation function to detect significant effects. However, this is not only a statistical issue but may also have additional biological relevance. Recent studies on species abundance and interaction strength in ecological networks suggest a strong dependency of interaction strength on abundance, i.e. interactions tend to be stronger if individuals of two species have more direct encounters (e.g. the ‘abundance-asymmetry hypothesis’; Vázquez et al. 2007). Interestingly, the interaction strength (measured as the rank of the GoF test) at the tropical forest in Sinharaja depended only weakly on abundances (Wiegand, Gunatilleke & Gunatilleke 2007a).

Small-scale attraction Small-scale repulsion No sig. effect

SEPARATION OF SCALES

Our analyses revealed, in line with the study at Sinharaja and a study of individual species area relationships (Wiegand et al. 2007b), separation of scales in the spatial association structure of the forest community. The observed separation of scales is in need of explanation and is a promising candidate for an indicative pattern. Note that this finding did not require any biological assumptions, because conditioning of the null model to preserve the large-scale spatial patterns selectively is a purely technical issue. Our analysis revealed that there is a separation of scales but does not provide direct evidence on the processes and mechanisms involved. We found a pattern in the data, but subsequent research needs to derive specific hypotheses on processes and mechanisms that could be tested in the field. One hypothesis is that the patterns at scales $> 30$ m would be driven by limited dispersal and habitat conditions, which typically vary along environmental gradients, which, in turn, are often related to topographical features (Harms et al. 2001; Valencia et al. 2004; Gunatilleke et al. 2006; John et al. 2007), and that the patterns at scales smaller than 10–15 m are driven by direct intraspecific interactions. Variation in the local light environment is another potential factor that may explain these
fine-scale patterns. However, while local light environment is certainly important for recruits, it may be less important for the larger trees analysed here. Clearly, in small neighbourhoods one would expect repulsion because there are physical limits to how close individual stems can be based on the sizes of their crowns (Wiegand et al. 2007b). However, our finding that most of the significant negative small-scale associations occurred within guilds suggests involvement of interactions other than those due to physical tree size.

**POINT PATTERN ANALYSIS AND POINT PROCESS MODELLING**

To answer our ecological questions, we used conceptually simple nonparametric techniques in the spirit of an exploratory analysis (Law et al. 2009). This is an important initial step to reveal basic indicative patterns of complex multivariate data sets such as the one used in our study. A promising next step to elucidate species interactions in more detail is the use of recent developments in point-process modelling (e.g. Grabarnik & Särkkä 2009; Illian, Møller & Waagepetersen 2009). Such models assume parametric interaction structures whereby the fitted parameters contain information on the strength and direction of the interactions. Point-process models also allow adjustments for the effects of the remaining species on the association of species pairs (Illian, Møller & Waagepetersen 2009). This is not possible with our approach, but may not be a severe limitation given the strong segregation between species found in our study forest. However, point-process models that consider all pairwise interactions quickly become intractable for species-rich communities. Therefore, Grabarnik & Särkkä (2009) and Illian, Møller & Waagepetersen (2009) used hierarchical models to simplify the interactions structure based on biological arguments (e.g. small trees do not influence the pattern of large trees). Future development of point-process models will allow more detailed investigation of the spatial organization of species-rich communities.

**Conclusions**

Our study is unique in that it presents comprehensive analyses on the spatial association structure of most of the tree community in a megaplot of a temperate forest. Clearly, there is a long tradition of spatial pattern analyses in temperate forests (e.g. Kenkel 1988; Duncan 1991), but most of these studies have focused on a few dominant or canopy species in these forests (Hao et al. 2007; but see Kubota, Kubo & Shimatani 2007). Because of the analogy in sampling and analyses with megaplots of tropical forests, our study is a first step towards more detailed investigations of similarities and differences between tropical and temperate forests in their complex spatial association structures. We found for large trees both indications for competitive interactions due to niche overlap and avoidance effects due to spatial segregation patterns. Analogous analyses in the species-rich tropical forests at Sinharaja showed almost no evidence for competitive interactions, but strong segregation patterns (Wiegand, Gunatilleke & Gunatilleke 2007a). We hypothesize that the balance between these two effects may vary with species richness and that species-poorer forests should show stronger (or more frequent) significant interspecific association patterns than species-rich forests. We took a first step by revealing patterns in the complex spatial data, but comparative studies among a larger number of plots will be necessary in order to prove our hypotheses and to determine whether these observations are indeed indicative patterns. We are confident that analysis of spatial patterns may eventually provide the additional information required in order to decide between competing hypotheses of community assemblage and dynamics, and to stimulate new theoretical developments.

**Acknowledgements**

This study was sponsored by the Knowledge Innovation Program of the Chinese Academy of Science (KZCX2-YW-QN402), the National Natural Science Foundation of China (40971286 and 30870400) and the National Key Technologies R&D Program of China (2008BAC39B02). X.W. and T.W. revised the manuscript during a workshop funded by the Center for Tropical Forest Science (CTFS). T.W. was supported by the ERC advanced grant 233066. We thank Egbert Leigh, Micheal Papiak and Justin Calabrese for checking the English and providing useful comments. We also thank anonymous referees for constructive criticism on earlier drafts of the manuscript.

**References**


Received 9 September 2009; accepted 1 February 2010

Handling Editor: Kyle Harms
Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Additional information on analysis 2.

Appendix S2. Association matrix for large-scale association.

Appendix S3. Why is the null model 3 met if species 2 is only distributed inside a subarea of the area occupied by species 1?

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.