

Spatial patterns of tree species richness in two temperate forests

Xugao Wang¹, Thorsten Wiegand², Amy Wolf³, Robert Howe³, Stuart J. Davies⁴ and Zhanqing Hao^{1*}

¹State Key Laboratory of Forest and Soil Ecology, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110164, China; ²Department of Ecological Modelling, UFZ Helmholtz Centre for Environmental Research-UFZ, D-04301 Leipzig, Germany; ³Department of Natural and Applied Sciences, University of Wisconsin-Green Bay, Green Bay, WI 54311, USA; and ⁴Center for Tropical Forest Science, Arnold Arboretum, Harvard University and Smithsonian Tropical Research Institute, Boston, MA 02131, USA

Summary

1. The relative contribution of external vs. internal clustering mechanisms for determining community structure and its manifestations has been the subject of a continuous debate, but few attempts have been made to examine their single and joint effects in a compound process model.

2. In this study, we tested four *a priori* hypotheses on the relative importance of habitat heterogeneity (topography and soil) and internal clustering mechanisms such as dispersal limitation on the species–area relationship (SAR) in two fully mapped 25-ha plots of temperate forests in the Changbaishan (CBS) Nature Reserve, China, and the Chequamegon-Nicolet National Forest in Wisconsin, USA.

3. We used the distance decay curve to test the generality of the results obtained for the SAR. To find out if the relative importance of internal and external clustering mechanisms changed with life stage, we conducted separate analyses for small, large and all trees.

4. Model selection favoured the most complex hypothesis that assumed an influence of both habitat heterogeneity and internal clustering on SAR and the distance decay curve. For the CBS plot, which shows weak topographical structuring, models were consistent with data only if soil factors were included into assessment of habitat heterogeneity. At the Wabikon plot, we could not test soil variables, but inclusion of topographical variables substantially improved the fit of the distance decay curve.

5. In general, the results of the SAR agreed with those of the distance decay curve, but the latter was sensitive to positive habitat-mediated species associations. The SAR, but not distance decay, distinguished among competing hypotheses for the community of large trees at CBS, where species exhibited only weak clustering.

6. Contrary to our expectations, we did not find substantial differences in the relative importance of internal and external clustering mechanisms with life stage.

7. *Synthesis.* Our analysis of spatial community structure for two relatively diverse temperate forests revealed that the factors governing spatial community structure may not substantially differ from those in highly diverse tropical forests. This result adds to our understanding of the ecological processes underlying the spatial diversity structure in natural forest communities.

Key-words: aggregation, Changbaishan, determinants of plant community diversity and structure, habitat heterogeneity, Poisson processes, temperate forest, Thomas processes, Wabikon

Introduction

The increase in number of species (species richness) with increasing sampling area is one of the most important attri-

butes of biological communities (Holt *et al.* 1999; He & Legendre 2002). This pattern, called the species–area relationship (SAR), quantifies basic aspects of biodiversity in a simple way, allowing comparisons among different study areas and ecological systems. Herein, we focus on tree communities that are completely mapped within similar-sized (25-ha) study areas.

*Correspondence author. E-mail: hzq@iae.ac.cn

In this case, one can calculate ‘local’ species–area relationships in which the area of sampling plots (A) within the 25 ha is consecutively increased (Connor & McCoy 1979). Although spatial patterns of species richness vary widely among natural communities, they show basic similarities that suggest general underlying mechanisms (He & Legendre 2002). Indeed, the local SAR has played an important role in recent debates about whether ecological communities are dispersal assembled or niche assembled, (Hubbell 1997, 2001) because it can be used to test neutral models (McGill, Maurer & Weiser 2006).

One major factor that influences the shape of the local SAR is spatial aggregation. If a species is more aggregated at a given spatial scale, the probability of presence in a randomly selected area (corresponding to this scale) becomes smaller. Consequently, the SAR values for a given area of the sampling plots will decrease if more species are aggregated at this scale (Plotkin *et al.* 2000; He & Legendre 2002; Tjørve *et al.* 2008). However, aggregated distribution patterns in species may be broadly attributed to two major, yet contrasting, factors: (i) external effects of the environment, such as habitat heterogeneity and (ii) internal processes of population and community dynamics. One of the most prominent examples of an internal clustering mechanism is dispersal limitation (Hubbell 2001) that builds one cornerstone of neutral theory (Hubbell 2001), but many other factors such as non-random seed deposition (Howe 1989), facilitation (Kikvidze *et al.* 2005), succession (Felinks & Wiegand 2007) and gap dynamics (Nagel, Svoboda & Diaci 2006) can contribute to clustered patterns in homogeneous environments. However, the relative contributions of external vs. internal clustering mechanisms for generating patterns of species richness in real communities are difficult to quantify (Wiegand, Gunatilleke & Gunatilleke 2007a; Wang *et al.* 2010a). This issue has been the subject of a continuous debate around the question of whether ecological communities are dispersal assembled or niche assembled (Hubbell 2001; McGill, Maurer & Weiser 2006), but researchers are now generally convinced that these assembly mechanisms are mutually complementary rather than mutually exclusive (He & Legendre 2002; Shen *et al.* 2009). However, only a few investigations have attempted to examine the single and joint effects of habitat heterogeneity and dispersal limitation in a compound process model (Shen *et al.* 2009). New insight can be expected when comparing results of such analyses among plant communities with contrasting characteristics (e.g. tropical vs. temperate forest) (Wang *et al.* 2010a, b).

The theory of spatial point processes (Møller & Waagepetersen 2003; Illian *et al.* 2008; Waagepetersen & Guan 2009) offers an opportunity to test if an observed local SAR is predominantly shaped by habitat heterogeneity and/or by internal clustering. Habitat models are fitted to the distribution data of individual species to quantify the influence of environmental factors on the species distribution, and fitting cluster point process models allows quantifying spatial clustering without and with consideration of the underlying habitat model (Shen *et al.* 2009; Waagepetersen & Guan 2009).

Combination of these elements allows testing of four *a priori* hypotheses on single and joint effects of habitat heterogeneity and internal clustering: (i) random placement (no habitat association, no clustering), (ii) habitat heterogeneity (no internal clustering), (iii) internal clustering (no habitat association) and (iv) joint effects of habitat heterogeneity and internal clustering (Shen *et al.* 2009). Comparison of the observed SAR with that of simulated communities corresponding to the different hypotheses helps identify the hypothesis that is most consistent with the data.

An important uncertainty of this approach (and tests of neutral theories in general; McGill, Maurer & Weiser 2006) is whether the SAR is sensitive enough to separate competing hypotheses of the underlying mechanisms or not. It has been suggested that the SAR may have low discriminatory power in distinguishing between niche assembly and dispersal assembly (e.g. Chave 2004; Purves & Pacala 2005), and McGill, Maurer & Weiser (2006) argued that additional predictions should be evaluated. One promising alternative summary statistic of community structure is the decay of similarity with distance curve (Chave & Leigh 2002; Condit *et al.* 2002; Morlon *et al.* 2008; (McGill 2010). As the distance decay curve evaluates other aspects of spatial community structure than the SAR (Morlon *et al.* 2008), it is not clear *a priori* if both will favour the same hypothesis. For example, Morlon *et al.* (2008) showed that one hypothesis for explaining the structure of three tropical forest communities yielded accurate predictions for the SAR, but not for all distance decay curves.

Previous studies have found that species often show different ecological habitat associations (Webb & Peart 2000; Comita, Condit & Hubbell 2007) and different degrees of spatial clustering with life stage (Wiegand *et al.* 2007b; Wang *et al.* 2010b). Do such species-specific variations even out on the community level or do they result in community-wide shifts in the relative importance of internal and external clustering with life stage? For example, strong influence of habitat association for small trees, but a loss of strong habitat associations for large (adult) trees could be interpreted as support for neutral theories in structuring canopy tree communities (e.g. Hubbell 2001).

In this study, we test four *a priori* hypotheses on the relative importance of habitat heterogeneity and internal clustering mechanisms on spatial community structure of two 25-ha fully mapped plots of temperate forests in the Changbaishan (CBS) Nature Reserve, north-eastern China and the Wabikon plot in the Chequamegon-Nicolet National Forest of north-eastern Wisconsin, USA. This analysis allows us to address three specific objectives. Firstly, we explore the relative importance of habitat heterogeneity and internal clustering in explaining the observed SAR. Secondly, we use the distance decay curve as an additional summary statistic to test the generality of the results obtained for the SAR. Lastly, we analyse assemblages of small trees [< 10 cm at breast height (d.b.h.) and large trees (d.b.h. ≥ 10 cm] separately to explore evidence for shifts in the relative importance of habitat heterogeneity and internal clustering with life stage.

Materials and methods

STUDY AREAS AND FIELD METHODS

Two temperate forest tree communities were chosen for this study. The first is represented by a 25-ha (500 × 500 m), fully censused temperate forest plot (42°23'N, 128°05'E) in the Changbaishan (CBS) Nature Reserve, north-eastern China. The Reserve, located along the border of China and North Korea, is one of the largest biosphere reserves in China and has been spared from logging and other severe human disturbances. Mean elevation in the CBS temperate forest plot is 801.5 m a.s.l., and elevation ranges from 791.8 to 809.5 m. All free-standing trees at least one centimetre in d.b.h. were mapped and identified to species, and their geographical coordinates were recorded following a standard field protocol (Condit 1998) by scientists from the Institute of Applied Ecology of the Chinese Academy of Science. The first census in 2004 yielded 38,902 living individuals (d.b.h. ≥ 1 cm) belonging to 52 species, 32 genera and 18 families. The main tree species included *Pinus koraiensis*, *Tilia amurensis*, *Quercus mongolica*, *Fraxinus mandshurica*, *Ulmus japonica* and *Acer mono*. Unlike tropical rain forests without obvious dominant species, eight species were recorded with more than 1000 individuals, together accounting for 83.4% of the total individuals in the plot. Mean stand density was 1556 living trees per hectare. Mean basal area was 43.2 m² ha⁻¹ (Hao *et al.* 2008; Wang *et al.* 2009).

The second data set derives from the 25.2-ha (300 × 840 m) Wabikon Forest Dynamics plot (45°33'N, 88°48'W) established in 2008 by scientists at the University of Wisconsin-Green Bay using the same methods described by Condit (1998). Living trees of at least 1 cm d.b.h. numbered 48 858, belonging to 36 species, 28 genera and 17 families. The Wabikon plot is located within the Chequamegon-Nicolet National Forest in north-eastern Wisconsin, USA, c. 10 km east of Crandon. The glacially formed topography consists of hummocky outwash features, including an esker running through part of the site. Elevations range from 488.3 to 514.2 m, with a mean of 498.1 m. Mesic northern hardwoods occupy most of the plot, dominated by sugar maple (*A. saccharum*), basswood (*T. americana*), white ash (*F. americana*) and ironwood/eastern hop hornbeam (*Ostrya virginiana*). Like the CBS plot, a relatively small number of species comprised the majority of individuals; 10 species were represented by 1000 or more individuals, together comprising 95.1% of all live individuals. The eight most abundant species at the Wabikon plot represented 90.1% of all individuals. Mean stand density was 1939 living trees per hectare. Mean basal area was 32.0 m² ha⁻¹.

To examine the effect of habitat heterogeneity on species–area relationships of the CBS and Wabikon plots, we evaluated three topographical variables (elevation, slope and aspect) and, for the CBS plot, eight soil properties (pH, organic matter, total N, total P, total K and available N, available P and available K). The plots were divided into grid systems using a 5 × 5 m quadrat size, and the mean values for these environmental variables (topographical and soil variables) were then calculated at the 5-m scale using geostatistical methods. Overall tree density in each quadrat (5 × 5 m) was also calculated and used as a comprehensive bioenvironmental index for this analysis.

POINT PATTERN ANALYSIS

We used recent advances in the theory of spatial point processes (Møller & Waagepetersen 2003; Illian *et al.* 2008; Waagepetersen & Guan 2009) to fit corresponding species-specific point process models for each of the four alternative hypotheses. For each species, we sub-

sequently generated 100 realizations of a fitted point process model (i.e. simulated distribution patterns) and independent superposition of the simulated distribution patterns for each hypothesis results in 100 simulated communities. For each simulated community, we calculated SAR and the distance decay curve and compared them with the observed patterns. The algorithms of the four point processes have been described in detail by Møller & Waagepetersen (2003), Illian *et al.* (2008) and others. Herein, we only summarize the basic framework of the four processes.

Homogeneous/inhomogeneous Poisson process

The random placement hypotheses can be represented by a homogeneous Poisson process in which the points are: (i) independently scattered and (ii) the intensity λ of the process (i.e. the mean point density in a unit area) is constant (Stoyan & Stoyan 1994). The habitat heterogeneity hypothesis can be represented by an inhomogeneous Poisson process in which condition: (i) holds, but where the intensity of the process depends on location x (i.e. the probability $\lambda(x)dx$ of a point occurring in an infinitesimally small disc of centre x and area dx depends on location x). The intensity $\lambda(x)$ may be influenced by environmental factors. In general, statistical habitat models or species distribution models (Elith & Leathwick 2009) may be used for parametric estimation of the intensity function. The most obvious parametric model to fit the intensity function for a heterogeneous Poisson process is the loglinear model (Waagepetersen 2007).

Homogeneous/inhomogeneous Thomas process

Considering the unrealistic independence assumption (i) of the two Poisson processes for real data, two kinds of cluster processes are used to model clustered spatial distribution patterns. To represent the internal clustering hypothesis we used the homogeneous Thomas cluster processes (Thomas 1949). It generates a number of randomly and independently distributed clusters, where the cluster centres follow a homogeneous Poisson process with intensity ρ , and the number of points per cluster follows a Poisson distribution with mean $\mu = \lambda/\rho$. The location of the points in a given cluster, relative to the cluster centre, has a bivariate Gaussian distribution with variance σ^2 (Stoyan & Stoyan 1994). The K -function of the Thomas process can be calculated analytically (Stoyan & Stoyan 1994; Wiegand *et al.* 2007b), which allows fitting the parameters of this process to the data for each species. Realizations of the fitted process can be easily simulated (Stoyan & Stoyan 1994). Note that fitting a Thomas process to an inhomogeneous pattern may formally produce a good fit. Thus, some of the effects of environmental heterogeneity may already be accounted for by a homogeneous Thomas process.

The inhomogeneous Thomas process represents the most complex hypothesis, where habitat heterogeneity and internal clustering occur simultaneously. The inhomogeneous Thomas process results from thinning a homogeneous Thomas process with intensity function $\lambda(x)$ (Waagepetersen 2007). If $\lambda(x)$ is known, the parameters of the corresponding homogeneous Thomas process can be fitted using the inhomogeneous K -function (Baddeley, Møller & Waagepetersen 2000; Waagepetersen 2007). This process provides a simple phenomenological description of clustering that explicitly includes the effect of environmental heterogeneity.

Model fitting

We used a two-step approach proposed by Waagepetersen & Guan (2009) to estimate the parameters of our four point processes. There

are basically two sets of parameters: regression parameters for the different covariates to estimate intensity functions $\lambda(x)$ (inhomogeneous Poisson and inhomogeneous Thomas process) and the clustering parameters ρ and σ of the homogeneous and inhomogeneous Thomas process. We used minimum contrast estimation to estimate the clustering parameters (Stoyan & Stoyan 1994). A Poisson likelihood function corresponding to the proposed intensity function was maximized to estimate the regression parameters both in case of the inhomogeneous Poisson process and the inhomogeneous Thomas process. Although this is not a maximum likelihood estimation, one can show that the estimated regression parameters are still consistent and asymptotically normal (Waagepetersen 2007; Waagepetersen & Guan 2009).

We used the soil variables together with three topographical parameters and the bioenvironmental index (total tree density in a 5×5 m quadrat) as environmental variables to determine the intensity function $\lambda(x)$. To reduce the risk of over-fitting, we computed the principal components (PCs) from the eight soil variables and used only the first two components as condensed variables because together they explained 97.2% of total variance in soil variables (John *et al.* 2007). The intensity function is then fitted using maximum likelihood estimation to models of the loglinear form $\lambda(x) = \exp(\beta_0 + \beta_1 v_1(x) + \dots + \beta_n v_n(x))$ with coefficients β_i and the variables $v_i(x)$. At CBS we have $n = 6$ variables and at Wabikon $n = 4$ variables. To account for the problem (common to both inhomogeneous Poisson and inhomogeneous Thomas processes) that some random variation is bound to be picked up by covariates even though these covariates in reality do not influence the spatial pattern of trees, we performed stepwise model reduction using Wald-tests (Waagepetersen & Guan 2009). Otherwise, too much variation would be attributed to the covariates.

Model selection

We generated, for each hypothesis, 100 communities by superposing realizations of the fitted point process models for each species. The species–area relationships were constructed by randomly throwing quadrats with increasing sizes in these simulated communities (Shen *et al.* 2009). To calculate the distance decay curve, we divided the plot into 20×20 m quadrats and calculated the similarity among these quadrats using the Jaccard index of similarity (Legendre & Legendre 1998). The predicted summary statistics (i.e. SAR and distance decay curve) for the four models were computed by averaging the simulated patterns of 100 simulated communities, and 95% simulation envelopes were constructed for each predicted summary statistic.

Finally, the observed summary statistics from the original data obtained in the CBS and Wabikon plots were compared with the summary statistics predicted by the four point process models. A model is considered satisfying, if the observed summary statistic falls within the simulation envelopes of the predicted summary statistic. To select the hypothesis that received most support from the data, we used a type of Akaike's Information Criterion (AIC) in which the log likelihood was approximated by the sum of squared residuals for either the SAR or the distance decay summary statistics (Webster & McBratney 1989; Shen *et al.* 2009). Due to the stepwise model reduction used for estimation of the intensity function, the number of parameters for the inhomogeneous Poisson and Thomas process may differ among species. However, as we calculated the AIC on the community level (superposing realizations of the point process models for individual species), we counted all covariates that were used at a given site and model at least once.

In addition, we illustrated the observed and simulated spatial distribution maps of *U. japonica* in the CBS plot. The distribution function $G(r)$ of the nearest-neighbour distances r was then calculated for each simulated distribution pattern to evaluate the goodness of fit of each model (Ripley 1988; Møller & Waagepetersen 2003). All calculations were carried out in R version 2.10.0 (R Development Core Team 2009), using the 'spatstat' package (Baddeley & Turner 2005).

Results

SMALL AND LARGE TREES TOGETHER

The species–area pattern in the two temperate forest plots (CBS and Wabikon) showed a similarly increasing tendency of species richness with increased sampling area (Figs 1a and 2a) although the final increase at areas larger than 12 ha was somewhat steeper at Wabikon. Note that all four hypotheses provide seemingly accurate results near the two ends of the species–area relationships, i.e. near 0 and 25 ha in the CBS plot or 0 and 25.2 ha in the Wabikon plot. However, this was an artefact because the total species richness and plot area were fixed regardless of what models were applied.

The observed distribution of *U. japonica* at the CBS plot (Fig. 3e) together with realizations of the different point process models (Fig. 3a–d) illustrates a typical result (see Fig. S1 Supporting Information). It is clear from visualization of the patterns that the two point processes without internal clustering (i.e. hypothesis i and ii; Fig. 3a,b) miss important aspects of the observed spatial structure. The homogeneous Thomas process (i.e. hypothesis iii) reproduces the small-scale clustering better (Fig. 3c); however, it produces unrealistically large gaps and does not effectively reproduce the spatial variation in tree density. The realization of the inhomogeneous Thomas process with soil factors (hypothesis iv; Fig. 3d) overcomes the shortcomings of the homogeneous Thomas process, and produced patterns that agreed well with the observed distribution of *U. japonica*.

The above observations for a single species generally hold for the entire communities. In most cases, the SAR was able to distinguish among the four hypotheses. Hypotheses i and ii were clearly rejected both on the basis of the simulation envelopes (Figs 1b and 2b) and on the basis of formal model selection using AIC (Table 1). Both hypotheses tended to considerably overestimate species richness. Interestingly, when soil variables were not considered, hypothesis iii (i.e. the internal clustering hypothesis) received the most support for both plots (Table 1), but yielded not fully satisfying fits (Figs 1b and 2b). The more complex model for the Wabikon forest (hypothesis iv) received almost the same support as model iii ($\Delta\text{AIC} = 3.4$), but internal clustering (hypothesis iii) slightly underestimated species richness at larger areas (> 7 ha) whereas consideration of the joined effect of internal clustering and habitat association (hypothesis iv) slightly overestimated species richness (Fig. 2b). Inclusion of the two soil variables into hypothesis iv for the CBS plot produced an excellent fit of the observed SAR (Fig. 1b) and was clearly the most parsimonious model (Table 1).

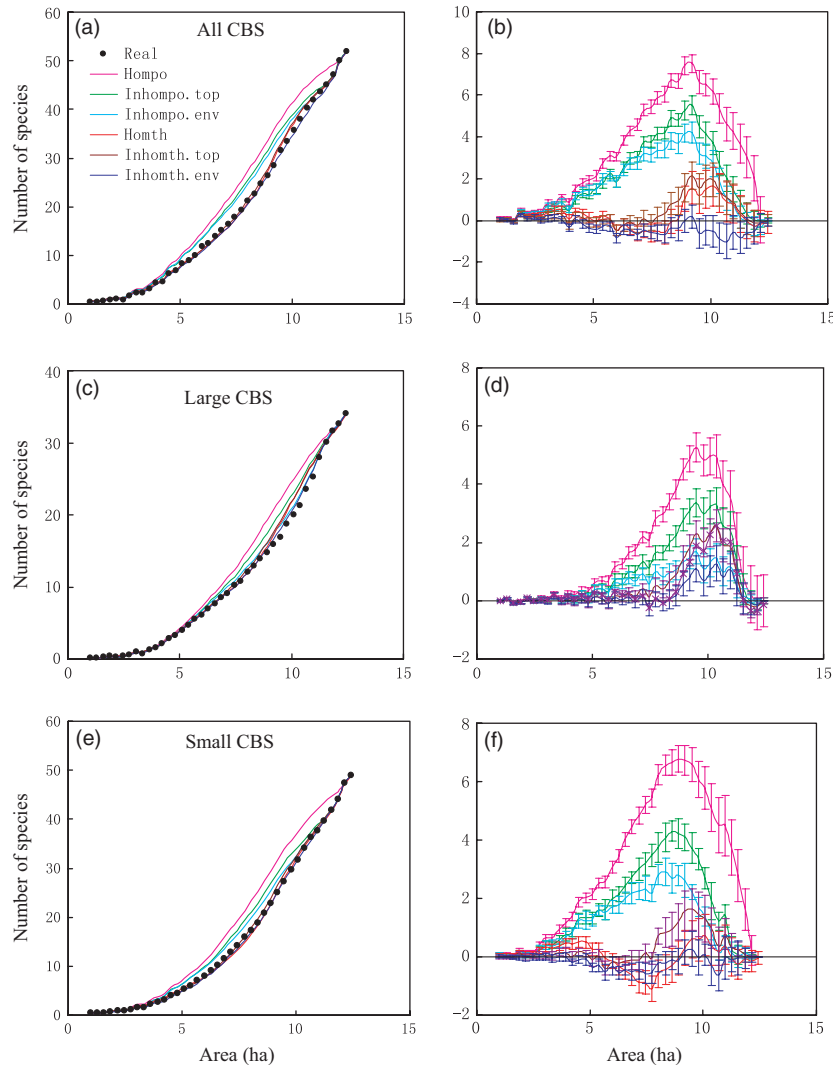


Fig. 1. The observed (dots) and predicted (coloured lines) species–area relationships for the data from the CBS plot. Small and large trees together (a, b), only for large trees (≥ 10 cm d.b.h.; c, d), and only for small trees (< 10 cm d.b.h.; e, f). The left column (a, c, e) shows species absolute richness, whereas the right column shows the relative species richness (i.e. observed – predicted) for the different hypotheses. The vertical bars are the 95% simulation envelopes arising from simulation of the point process models. Hompo, homogeneous Poisson process; Inhompo.top, inhomogeneous Poisson process without soil factors; Inhompo.env, inhomogeneous Poisson process with soil factors; Inhomth.top, inhomogeneous Thomas process without soil factors, Inhomth.env: Inhomogeneous Thomas process with soil factors.

Additional insight in the ability of the four hypotheses to explain spatial community structure is provided by the distance decay curve (Fig. 4; Table 1). In agreement with results of the SAR analyses, all hypotheses that did not include internal clustering (i.e. i, ii) failed notably and produced far too high similarity values (Fig. 4a,b). However, in contrast to the SAR results, the support for hypothesis iv over hypothesis iii was unambiguous for the distance decay curve, both for the simulation envelopes (Fig. 4a,b) and for AIC (Table 1). At small distances (i.e. < 300 m), the predictions for the CBS plot were very close to the observations, but for larger distances the corresponding point process models slightly overestimated similarity (Fig. 4a). Interestingly, hypothesis iv with and without soil factors received the same support ($\Delta\text{AIC} < 2$; Table 1). The predictions of hypothesis iii (only internal clustering) overestimate similarity at the CBS

plot at 20 m and underestimate similarity at distances between 60 and 200 m (Fig. 4a). The failure of hypothesis iii to fit the distance decay curve is evident at the Wabikon plot, where it severely underestimated similarity at all distances except 20 m (Fig. 4b). However, inclusion of environmental heterogeneity (hypothesis iv) provided a much better, albeit not fully satisfying, fit. Thus, the distance decay curve shows that environmental heterogeneity must be considered together with internal clustering.

SMALL VS. LARGE TREES

As most trees were small, the total number of species of small trees (49 at CBS and 33 at Wabikon) was similar to the total number of species for all trees (52 at CBS and 36 at Wabikon), but the number of species of large trees at both

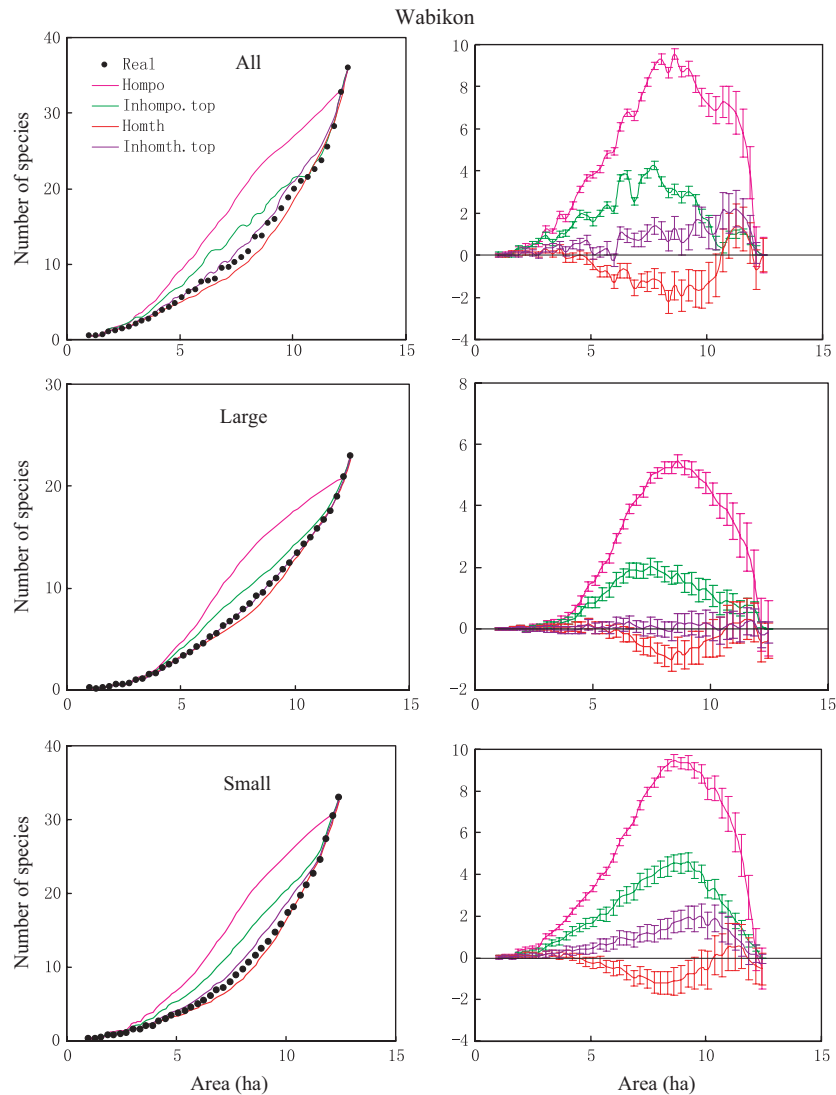


Fig. 2. The observed (dots) and predicted (coloured lines) species–area relationships for the data from the Wabikon plot. Conventions as in Fig. 1.

plots was substantially lower (34 at CBS and 23 at Wabikon). Unsurprisingly, the SAR results obtained for separate analysis of small trees did not differ substantially from those obtained for all trees together (Figs 1 and 2).

Separate analyses of the distance decay function for small and large trees generally supported the results of the analysis for all trees; hypothesis iv was selected in all cases. Again, hypothesis iv with and without soil factors received the same support for separate analysis of small and large trees at the CBS plot ($\Delta\text{AIC} < 2$; Table 1). However, some differences in detail are evident for large trees. At the CBS plot hypotheses ii, iii and iv received similar support for the distance decay (Fig. 4c; Table 1), which means that large trees were relatively less clustered at this plot. This is also evident from the predicted distance decay curves which showed little differences among the four hypotheses (Fig. 4c), whereas the predicted distance decay curves of all trees and small trees differed widely among the four hypotheses (Fig. 4a, e). However, at the Wabikon plot, the distance decay of large trees at smaller distances

(< 150 m) was poorly simulated by all four hypotheses (Fig. 4d), although hypothesis iv yielded accurate SAR predictions for large trees (Fig. 2d).

ENVIRONMENTAL VARIABLES

When analysing small and large individuals together, we found that 61% and 64% of all species distribution patterns at the CBS and Wabikon plots, respectively, showed a significant relationship with environmental variables (i.e. they show a habitat association; see Table S1 Supporting Information). For large trees, these figures dropped to 33% and 42% respectively. This is probably a consequence of smaller statistical power due to the smaller number of larger trees.

At the CBS plot, approximately half of all habitat models for all trees yielded only one significant environmental variable (Table S2), but this was true for only 9% of the habitat models at the Wabikon plot. In general, the abiotic variables were more frequently selected at Wabikon for all trees (elevation: 19

Table 1. Comparison of Akaike's Information Criterion (AIC) among the four process models for the different analyses and forest plots

		Poisson process		Thomas process	
		Homogeneous random placement	Inhomogeneous habitat heterogeneity	Homogeneous internal clustering	Inhomogeneous habitat heterogeneity and internal clustering
SAR					
All	CBS	260.9	229.2♂, 212.1♀	117.8	136.2♂, 94.7
	Wabikon	288.1	210.9	152.3	155.7
Large	CBS	221.1	183.4♂, 144.9♀	139.4	145.8♂, 82.5♀
	Wabikon	242.2	160.7	79.9	18.9
Small	CBS	257.2	211.7♂, 182.1♀	81.6	109.6♂, 45.4♀
	Wabikon	288.4	228.2	110.6	159.2
Distance decay curve					
All	CBS	153.4	112.8♂, 107.8♀	44.0	41.1♂, 39.1
	Wabikon	171.9	140.7	95.7	40.0
Large	CBS	34.1	11.6♂, 9.1♀	9.4	11.2♂, 11.1♀
	Wabikon	51.5	39.4	26.8	18.2
Small	CBS	95.5	62.9♂, 58.3♀	17.1	15.2♂, 14.5♀
	Wabikon	115.3	89.2	22.6	17.7

♂, without soil data, ♀, with soil data.

of 23 species, slope: 7/23, aspect: 9/23) compared with the CBS plot (elevation: 7/32, slope: 5/32, aspect: 3/32, PCA1: 9/32, PCA2: 12/32). As expected, the coefficients for analyses with all individuals were in close accordance with those for small individuals, but coefficients for large trees were often quite different (Tables S2 and S3).

Discussion

Our results have shown that the combined effects of habitat heterogeneity (including soil factors) and internal clustering lead to good approximations of the observed species–area relationships and distance decay curves in two representative temperate forests, one in north-eastern China and the other in north-central United States. The species–area relationship proved to be a sensitive summary statistic that can detect subtle differences among competing point process models and yielded results that, in general, agreed with those of the distance decay curve. Contrary to our expectations, the most complex model including habitat association and internal clustering received, in most cases, the most support. Thus, we did not find substantial differences in the relative importance of habitat association and internal clustering with life stage (this would be given if different models would be selected for different life stages). Exceptions were the SAR of small trees at Wabikon, where the homogeneous Thomas process was clearly favoured and two cases (SAR of all trees at Wabikon and distance decay for all trees at CBS) where several models including the most complex one received similar support (Table 1). Comparison of our results from these two relatively diverse temperate forests with spatial patterns of trees in highly diverse tropical and subtropical forests suggests that community structure may be governed by the same underlying mechanisms.

RELATIVE IMPORTANCE OF INTERNAL AND EXTERNAL AGGREGATION MECHANISMS

In our analyses, the random placement hypothesis yielded the poorest model for predicting spatial community structure. In particular, this model overestimated species richness for most of the sampling range and did not account for the observed decline in similarity with distance. This result is consistent with previous studies that have identified spatial clustering as the dominant pattern of species from temperate forests to tropical forests (He, Legendre & LaFrankie 1997; Condit *et al.* 2000; Morlon *et al.* 2008; Wang *et al.* 2010b). However, the random placement hypothesis performed relatively well at the CBS plot and for the analysis of all trees yielded a maximal overestimation of species richness of *c.* 28% at the 0.9-ha area: (36 species were predicted and 28 observed). At Wabikon, the maximal overestimation yielded 64% at the 0.6-ha area (23 species were predicted and 14 observed). This result and the steeper increase of the SAR for larger areas at Wabikon suggests that the CBS plot shows a better mixing of species within the plot compared with the Wabikon plot. Indeed, some species at the Wabikon plot show very patchy distribution patterns that result in hotspots of species richness (and number of individuals), related in part to the elevation pattern (Fig. S2). These trends are much weaker at the CBS forest (Fig. S3). Consequently, the consideration of habitat heterogeneity (*i.e.* hypothesis ii) improved the SAR prediction at Wabikon considerably (the maximum error dropped from nine to four species), but only moderately at the CBS plot. The Wabikon plot experienced localized logging during the 1900s, as evidenced by historical records and air photos dating to 1938. These disturbances, in addition to pronounced variation in elevation, have contributed to spatial heterogeneity of trees in parts of the plot. Nevertheless, we found for analyses of all trees that the distribution pattern of approximately two-thirds of all

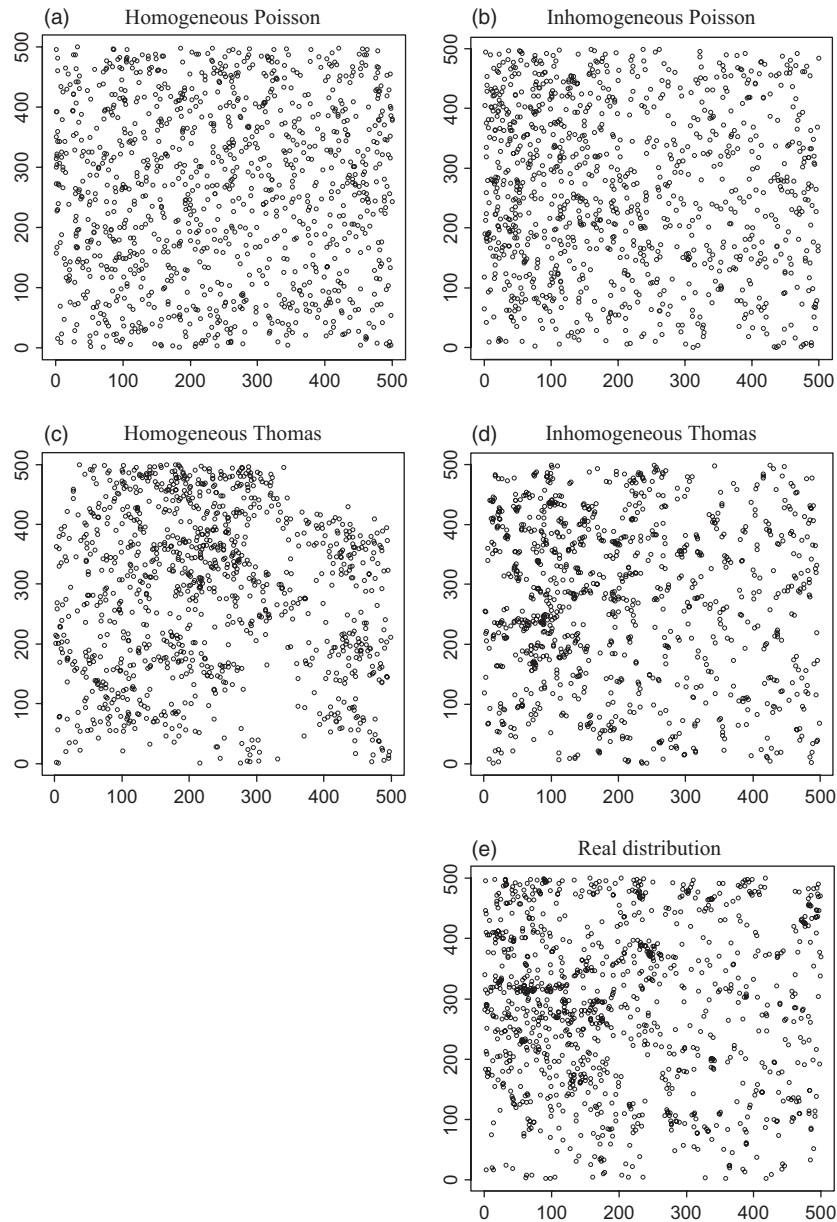


Fig. 3. The observed distribution of *Ulmus japonica* in the CBS plot and its distribution generated from the homogeneous Poisson process, inhomogeneous Poisson process with soil factors, homogeneous Thomas process and inhomogeneous Thomas process with soil factors.

species showed a significant relationship with environmental variables (Table S1). Thus, habitat heterogeneity cannot be neglected at our two study plots, but it is not sufficient to produce the observed patterns in SAR and distance decay. This result suggests a need to consider additional clustering mechanisms.

Recently, both theoretical and empirical studies have emphasized the importance of dispersal limitation in controlling species distributions (Condit *et al.* 2002; Seidler & Plotkin 2006; Wiegand, Martinez & Huth 2009). However, the effect of dispersal limitation on spatial patterns is rarely straightforward and many other biological interactions and processes (other than habitat association) may contribute to the spatial pattern of a species. Seidler & Plotkin (2006)

found that the extent and scale of conspecific spatial clustering in the 50-ha Pasoh Forest Plot was correlated with the mode of seed dispersal. Thus, dispersal limitation is likely to leave a signal on the species clustering process and thus on species–area patterns and distance decay curves. Our results indicate that homogeneous Thomas processes mimicking the spatial clustering of species (i.e. hypothesis iii) reproduce the SAR fairly well (see also Plotkin *et al.* 2000; Morlon *et al.* 2008), but failed in reproducing observed distance decay curves. This is because the homogeneous Thomas process can produce good fits to aggregation caused by environmental heterogeneity, but it cannot describe positive dependency in the distribution pattern among species mediated by habitat association. As a consequence, it produces communities

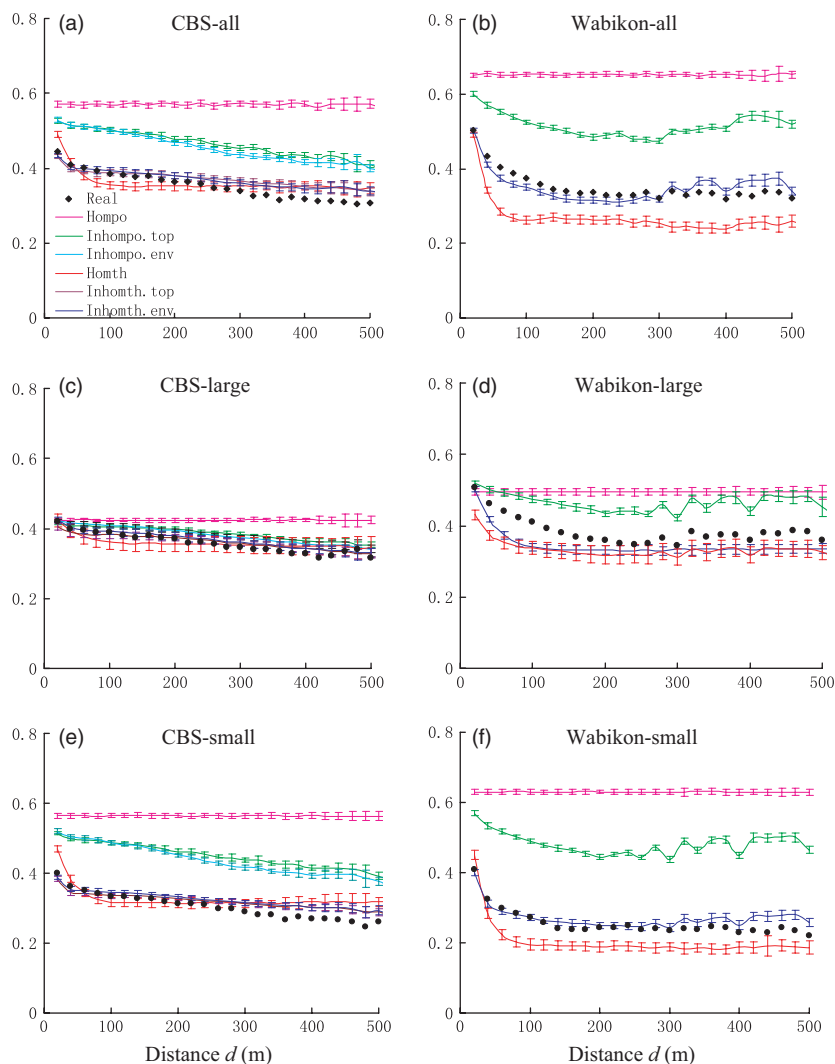


Fig. 4. The observed (dots) and predicted (coloured lines) decay of similarity with distance curve for the CBS and the Wabikon plot. We used the Jaccard index to quantify the similarity of 20×20 m quadrats that were distance d away. Colour conventions as in Fig. 1.

that are too well mixed (i.e. similarity decays too quickly; Fig. 4a,b). However, the inhomogeneous Thomas process (hypothesis iv), which places clusters only at suitable areas, can generate the positive association among species necessary to produce the observed decay in similarity with distance. In particular, the inhomogeneous Thomas process is able to account for the patchy diversity hotspots at Wabikon forest. In other words, this model is better able to simulate spatial patterns for communities that show strong spatial structuring (Figs S2 and S3). Thus, at the local spatial scale studied here, both habitat association and internal clustering are required to explain the observed patterns in spatial community structure.

Our results on the relative importance of internal vs. external aggregation mechanisms agree with the findings of Shen *et al.* (2009) and Morlon *et al.* (2008) for highly diverse subtropical and tropical forests. This is a further indication that the effects of the environment and internal clustering mechanisms contribute in a complementary way to the assembly of

species rich communities (He & Legendre 2002; Shen *et al.* 2009).

However, it is important to emphasize that species–area patterns are dramatically dependent on spatial scales (Palmer & White 1994; Whittaker, Willis & Field 2001). Our study was confined to the local community scale, where habitat types usually do not change dramatically, but rather more gradually (as depicted by the habitat models for the intensity function). Beyond the local scale, however, the relative effects of different processes may change. For example, dispersal limitation may have a dominant effect on the species–area relationships at the local community scale, whereas habitat heterogeneity may become more pronounced at the regional scale (Kallimanis *et al.* 2008).

SAR VS. DISTANCE DECAY CURVE

Our results support the analysis of more than one summary statistic for understanding spatial patterns of

community structure (McGill, Maurer & Weiser 2006). In general, the distance decay curve supported the results found for the SAR, but there was a somewhat curious reversal of the AIC model ranking between the CBS and Wabikon plots (Table 1). At the Wabikon plot, the distance decay curve favoured for all trees and small trees the more complex model iv that included the influence of both habitat association and internal clustering, whereas the SAR favoured the simpler model iii that included internal clustering only (although the AIC difference for all trees was small). At the CBS plot, the distance decay curve could not discriminate between models with or without habitat associations, whereas the SAR clearly favoured the more complex model iv with soil variables included. The distance decay curve of hypothesis iii for all trees at Wabikon (Fig. 4b) clearly underestimated similarity, but this discrepancy was only slight at CBS (Fig. 4a). As discussed above, this may be due to the inability of hypothesis iii to generate positive associations among species caused by shared habitats. Consequently, hypothesis iv, which included habitat association, improved the fit of the distance decay curve at Wabikon considerably. This effect is also visible at the CBS plot, but habitat associations appear to be weaker and influential over a smaller range of distances (100–200 m). Thus, the distance decay curve proved to be sensitive to violation of the independence assumption of hypothesis iii due to habitat association, but the SAR was not as sensitive. This important result shows that the independence assumption underlying most theories of stochastic geometry of biodiversity (McGill 2010) may indeed not be valid. If local diversity is patchy with marked diversity hotspots as shown by the Wabikon plot, positive association among species (mediated by habitat heterogeneity) is required to account for the observed distance decay curve. The CBS plot, however, was more strongly characterized by interspecific segregation patterns (Wang *et al.* 2010a) that resulted in a more even distribution of local diversity. Consequently, the error of not considering habitat heterogeneity was less severe for the distance decay curve.

The distance decay curve was not always able to distinguish among competing hypotheses. For example, the distance decay curve for large trees at the CBS plot was not able to distinguish among hypotheses ii, iii and iv with and without soil factors (Table 1), unlike the SAR. The reason for this is probably that the community of large trees at CBS was only weakly clustered, an attribute for which the distance decay curve is especially sensitive. Morlon *et al.* (2008) have demonstrated that distance decay is closely related to the pair correlation function, which describes details of clustering of individual species (Wiegand, Martinez & Huth 2009), whereas the SAR is related to the spherical contact distribution $H_S(r)$, the probability that there is no tree of species *i* within distance *r* from the centre of a circular sampling plot with radius *r* (Illian *et al.* 2008). Note that the spherical contact distribution basically describes the empty space between the clusters (Illian *et al.* 2008), and is therefore less sensitive to details of smaller-scale clustering

as long as a simplified point process model represents the gaps.

ENVIRONMENTAL VARIABLES

Our analysis revealed some interesting differences between temperate forests and species rich tropical and subtropical forests (Shen *et al.* 2009). For example, we found that, without including soil factors, consideration of environmental heterogeneity did not improve the predicted SAR based on internal clustering, unlike the results of Shen *et al.* (2009) for the tropical forest plot on Barro Colorado Island (BCI). Interestingly, this result applies to a separate analysis of small and large tree communities. Thus, soil factors may have a stronger effect on species–area patterns in the CBS forests compared with the BCI plot. The flat topography of the CBS plot (elevation varies only 18 m; Wang *et al.* 2008) may explain the weak impact of topographical variables compared with that at BCI, where elevation within the plot varies 40 m.

The stronger effect of topography at BCI could also be related to the specific hydrological conditions, where slopes are wetter than plateaus and experienced a shorter drought during the dry season (Daws *et al.* 2002; Leigh *et al.* 2004). In addition, it should be emphasized that we interpreted the effect of soil factors on these spatial patterns as directional responses of species to variations in soil properties in the study. However, tree species in forests may both affect and respond to soil properties through litterfall inputs and effects on microbial communities and decomposition rates, etc. (Boerner & Koslowsky 1989; Finzi, Canham & Van Breemen 1998). In other words, soil variables are not necessarily abiotic factors generated by processes extrinsic to population and community dynamics. Although some studies have argued that biotic feedback effects are less likely to influence spatial variation in soil nutrient availability in species rich tropical forests (Powers, Kalicin & Newman 2004), their influence on soils in temperate forests is still poorly known.

At both of our temperate forest plots, the intensity function for most species was positively related with the bioenvironmental index (overall tree density in each 5 × 5 m quadrat; Tables S2 and S3). This suggests that some environmental constraint, common for most species, was not captured by the environmental variables used in our analysis.

SIMPLIFICATIONS OF THE APPROACH

The point process models that combined habitat heterogeneity and internal clustering provided relatively good, but not always perfect, fits to the SAR. This is to a large extent due to the power of the Thomas processes to represent clustering (see also Plotkin *et al.* 2000; Morlon *et al.* 2008). However, the predictions for the distance decay curve may be improved using point process models that are able to capture more complex distribution patterns and underlying processes. For example, Wiegand *et al.* (2007b); Wiegand, Martinez and Huth (2009) showed that species may often cluster at several critical scales. Consideration of only one scale of clustering did not severely reduce

the power of our point process models to describe the observed spatial community structure. However, the analysis of Morlon *et al.* (2008) for the tropical forest at the Korup National Park (Cameroon) showed that hypothesis iii overestimated similarity, which is the opposite of our results. They attributed this to the inability of the Thomas process to describe the complex multi-clustered spatial patterns at this plot (Morlon *et al.* 2008). The fit with the homogeneous Thomas process probably captured only the larger scale of clustering and produced patterns that lacked the small-scale clustering necessary for producing lower similarities among subplots.

An interesting result of our study is that the assumption of independence among the patterns of individual species led to a good fit of models with the observed species–area patterns (Plotkin *et al.* 2000). Herein, we found that inclusion of positive, habitat-mediated species associations were important to fit the distance decay curve (but not SAR) at the Wabikon plot, where diversity hotspots are present. The segregation patterns that dominated the intraspecific species relationships of larger trees (> 10 cm d.b.h.) at the CBS plot (Wang *et al.* 2010a) are probably also produced by the impact of environmental heterogeneity, because the point process model that assumed homogeneous clustering (without habitat heterogeneity) failed to reproduce both the SAR and the distance decay curve. However, intraspecific interactions of large trees that were detected for 1/3 of all pairs of large species at the CBS plot (after removing large-scale habitat effects; Wang *et al.* 2010a) cancelled out and did not compromise the fit of the SAR and the species decay curve. Thus, smaller-scale species interactions may not be major factors structuring spatial community structure.

Finally, the point process models used here are static, and do not incorporate the effects of temporal processes or site history (Ripley 1988; Plotkin *et al.* 2000; Møller & Waagepetersen 2003). In all, more sophisticated dynamic models are required to further explain the underlying mechanisms controlling the relationship between species richness and sampling area at these spatial scales.

Acknowledgements

We thank the CBS and Wabikon plot census and data management teams, the latter including many student assistants from the University of Wisconsin-Green Bay. In particular, we thank Buhang Li, Ji Ye, Fei Lin, Zuoqiang Yuan, Xuejiao Bai, Liwei Wang and Dingliang Xing for their field work and data collection in the CBS plot. At the Wabikon plot, Gary Fewless, Kathryn Corio and Juniper Sundance were instrumental in collecting and managing the data. This study was sponsored by the Knowledge Innovation Program of the Chinese Academy of Sciences (KZCX2-YW-QN402, KSCX2-EW-Z-5 and KZCX2-EW-401), the National Natural Science Foundation of China (40971286) and the National Key Technologies R&D Program of China (2008BAC39B02). The ERC advanced grant 233066 supported travelling of X.W. The Center for Tropical Forest Science of the Smithsonian Tropical Research Institute and the HSBC Climate Partnership provided support for X.W. and for establishment of the Wabikon Plot. Primary funding for the Wabikon Plot was provided by the Cofrin Center for Biodiversity with additional support from the Department of Natural and Applied Sciences at the University of Wisconsin-Green Bay. We are grateful for permission and important logistic contributions by the U.S. Forest Service, particularly Chequamegon-Nicolet National Forest ecologists Linda Parker and Steven Janke. Other significant contributors to the project include Richard Condit, Steve Dhein and Kimberlee McKeefry.

References

- Baddeley, A.J., Møller, J. & Waagepetersen, R. (2000) Non- and semiparametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329–350.
- Baddeley, A. & Turner, R. (2005) SPATSTAT: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, **12**, 1–42.
- Boerner, R.E.J. & Koslowsky, S.D. (1989) Microsite variations in soil chemistry and nitrogen mineralization in a beech–maple forest. *Soil Biology and Biochemistry*, **21**, 795–801.
- Chave, J. (2004) Neutral theory and community ecology. *Ecology Letters*, **7**, 241–253.
- Chave, J. & Leigh, E.G. (2002) A spatially explicit neutral model of beta-diversity in tropical forests. *Theoretical Population Biology*, **62**, 153–168.
- Comita, L.S., Condit, R. & Hubbell, S.P. (2007) Developmental changes in habitat associations of tropical trees. *Journal of Ecology*, **95**, 482–492.
- Condit, R. (1998) *Tropical Forest Census Plots*. Springer, Berlin.
- Condit, R., Ashton, P., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N. *et al.* (2000) Spatial patterns in the distribution of common and rare tropical tree species: a test from large plots in six different forests. *Science*, **288**, 1414–1418.
- Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B. *et al.* (2002) Beta-diversity in tropical forest trees. *Science*, **295**, 666–669.
- Connor, E.F. & McCoy, E.D. (1979) The statistics and biology of the species–area relationship. *American Naturalist*, **113**, 791–833.
- Daws, M.I., Mullins, C.E., Burslem, D., Paton, S.R. & Dalling, J.W. (2002) Topographic position affects the water regime in a semideciduous tropical forest in Panama. *Plant and Soil*, **238**, 79–90.
- Elith, J. & Leathwick, J. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, **40**, 677–697.
- Felinks, B. & Wiegand, T. (2007) Exploring spatiotemporal patterns in early stages of primary succession on former lignite mining sites. *Journal of Vegetation Science*, **19**, 267–276.
- Finzi, A.C., Canham, C.D. & Van Breemen, N. (1998) Canopy tree–soil interactions within temperate forests: species effects on pH and cations. *Ecological Applications*, **8**, 447–454.
- Hao, Z., Li, B., Zhang, J., Wang, X., Ye, J. & Yao, X. (2008) Broad-leaved Korean pine mixed forest plot in Changbaishan (CBS) of China: community composition and structure. *Acta Phytocologica Sinica*, **32**, 238–250. (in Chinese).
- He, F. & Legendre, P. (2002) Species diversity patterns derived from species–area models. *Ecology*, **83**, 1185–1198.
- He, F., Legendre, P. & LaFrankie, J.V. (1997) Distribution patterns of tree species in a Malaysian tropical rain forest. *Journal of Vegetation Science*, **8**, 105–114.
- Holt, R.D., Lawton, J.H., Polis, G.A. & Martinez, N.D. (1999) Tropical rank and the species–area relationship. *Ecology*, **80**, 1495–1504.
- Howe, H.F. (1989) Scatter and clump dispersal and seedling demography: hypothesis and implications. *Oecologia*, **79**, 417–426.
- Hubbell, S.P. (1997) A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs*, **16**, S9–S21.
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.
- Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008) *Statistical Analysis and Modeling of Spatial Point Patterns*. Wiley, Chichester.
- John, R., Dalling, J.W., Harms, K.E., Yavitt, J.B., Stallard, R.F., Mirabello, M., Hubbell, S.P., Valencia, R., Navarrete, H., Vallejo, M. & Foster, R.B. (2007) Soil nutrients influence spatial distributions of tropical tree species. *Proceedings of the National Academy of Sciences*, **104**, 864–869.
- Kallimanis, A.S., Mazaris, A.D., Tzanopoulos, J., Halley, J.M., Pantis, J.D. & Sgardelis, S.P. (2008) How does habitat diversity affect the species–area relationship? *Global Ecology and Biogeography*, **17**, 532–538.
- Kikvidze, Z., Pugnaire, F., Brooker, R., Choler, P., Lortie, C., Michalet, R. & Callaway, R. (2005) Linking patterns and processes in alpine plant communities: a global study. *Ecology*, **86**, 1395–1400.
- Legendre, P. & Legendre, L. (1998) *Numerical Ecology*, 2nd English edn. Elsevier Science BV, Amsterdam.
- Leigh, E.G.J., Lao, S., Condit, R., Hubbell, S.P., Foster, R.B. & Perez, R. (2004) Barro Colorado island forest dynamics plot, Panama. *Tropical Forest Diversity and Dynamism: Findings from a Large-Scale Plot Network* (eds E. Losos & E.G.J. Leigh), pp. 451–463. University of Chicago Press, Chicago.
- McGill, B.J. (2010) Towards a unification of unified theories of biodiversity. *Ecology Letters*, **13**, 627–642.

- McGill, B.J., Maurer, B.A. & Weiser, M.D. (2006) Empirical evaluation of neutral theory. *Ecology*, **87**, 1411–1423.
- Møller, J. & Waagepetersen, R. (2003) *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall CRC Press, New York.
- Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R. & Green, J.L. (2008) A general framework for the distance-decay of similarity in ecological communities, theory. *Ecology Letters*, **11**, 904–91775:1.
- Nagel, T., Svoboda, M. & Diaci, J. (2006) Regeneration patterns after intermediate wind disturbance in an old-growth Fagus-Abies forest in southeastern Slovenia. *Forest Ecology and Management*, **226**, 268–278.
- Palmer, M.W. & White, P.S. (1994) Scale dependence and the species-area relationship. *American Naturalist*, **144**, 717–740.
- Plotkin, J.B., Potts, M.D., Leslie, N., Manokaran, N., Lafrankie, J. & Ashton, P.S. (2000) Species–area curves, spatial aggregation, and habitat specialization in tropical forests. *Journal of Theoretical Biology*, **20**, 81–99.
- Powers, J.S., Kalicin, M.H. & Newman, M.E. (2004) Tree species do not influence local soil chemistry in a species rich Costa Rica rain forest. *Journal of Tropical Ecology*, **20**, 587–590.
- Purves, D.W. & Pacala, S.W. (2005) Ecological drift in niche-structured communities: neutral pattern does not imply neutral process. *Biotic Interactions in the Tropics* (eds D. Burslem, M. Pinard & S. Hartley). pp. 108–138, Cambridge University Press, New York.
- R Development Core Team. (2009) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL: <http://www.R-project.org>.
- Ripley, B.D. (1988) *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.
- Seidler, T.G. & Plotkin, J.B. (2006) Seed dispersal and spatial pattern in tropical trees. *PLoS Biology*, **4**, 2132–2137.
- Shen, G., Yu, M., Hu, X., Mi, X., Ren, H., Sun, I. & Ma, K. (2009) Species–area relationships explained by the joint effects of dispersal limitation and habitat heterogeneity. *Ecology*, **90**, 3033–3041.
- Stoyan, D. & Stoyan, H. (1994) *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. Wiley, Chichester, UK.
- Thomas, M. (1949) A generalization of Poisson's binomial limit for use in ecology. *Biometrika*, **36**, 18–25.
- Tjørve, E., Kunin, W.E., Polce, C. & Tjørve, K.M.C. (2008) Species–area relationship: separating the effects of species abundance and spatial distribution. *Journal of Ecology*, **96**, 1141–1151.
- Waagepetersen, R. (2007) An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics*, **63**, 252–258.
- Waagepetersen, R. & Guan, Y. (2009) Two-step estimation for inhomogeneous spatial point processes. *Journal of the Royal Statistical Society, Series B*, **71**, 685–702.
- Wang, X., Hao, Z., Ye, J., Zhang, J., Li, B. & Yao, X. (2008) Spatial pattern of diversity in an old-growth temperate forest in Northeastern China. *Acta Oecologica*, **33**, 345–354.
- Wang, X., Hao, Z., Zhang, J., Lian, J., Li, B., Ye, J. & Yao, X. (2009) Tree size distributions in an old-growth temperate forest. *Oikos*, **118**, 25–36.
- Wang, X., Wiegand, T., Hao, Z., Li, B., Ye, J. & Lin, F. (2010a) Species associations in an old-growth temperate forest in north-eastern China. *Journal of Ecology*, **98**, 674–686.
- Wang, X., Ye, J., Li, B., Zhang, J., Lin, F. & Hao, Z. (2010b) Spatial distributions of species in an old-growth temperate forest, northeastern China. *Canadian Journal of Forest Research*, **40**, 1011–1019.
- Webb, C.O. & Peart, D.R. (2000) Habitat associations of trees and seedlings in a Bornean rain forest. *Journal of Ecology*, **88**, 464–478.
- Webster, R. & McBratney, A.B. (1989) On the Akaike information criterion for choosing models for variograms of soil properties. *Journal of Soil Science*, **40**, 493–496.
- Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: towards a general hierarchical theory of species diversity. *Journal of Biogeography*, **28**, 453–470.
- Wiegand, T., Gunatilleke, C.V.S. & Gunatilleke, I.A.U.N. (2007a) Species associations in a heterogeneous Sri Lankan dipterocarp forest. *American Naturalist*, **170**, E77–E95.
- Wiegand, T., Martinez, I. & Huth, A. (2009) Recruitment in tropical tree species: revealing complex spatial patterns. *American Naturalist*, **174**, E106–E140.
- Wiegand, T., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N. & Okuda, T. (2007b) Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology*, **88**, 3088–3102.

Received 16 February 2011; accepted 24 May 2011

Handling Editor: Kyle Harms

Supporting Information

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

Table S1. Number of cases for which the intensity function $\lambda(x)$ could not be estimated (number of individuals < 10), for which the stepwise procedure did not yield a significant model, and the number of species for which a significant model was found.

Table S2. Detailed results of the habitat models for the Wabikon plot.

Table S3. Detailed results of the habitat models for the CBS plot.

Figure S1. The nearest neighbour distance function $G(r)$ for the *U. japonica* in the CBS plot and the theoretical $G(r)$ for the fitted homogeneous Poisson process, the inhomogeneous Poisson process, the homogeneous Thomas process and the inhomogeneous Thomas process.

Figure S2. Elevation, species abundance and richness at the Wabikon plot.

Figure S3. Elevation, species abundance and richness at the CBS plot.

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.