# Quantifying effects of habitat heterogeneity and other clustering processes on spatial distributions of tree species

7 AUTHORS, INCLUDING:

Guochun Shen
East China Normal University
4 PUBLICATIONS   44 CITATIONS

SEE PROFILE

I Fang Sun
National Dong Hwa University
53 PUBLICATIONS   1,016 CITATIONS

SEE PROFILE

Zueng-Sang Chen
National Taiwan University
99 PUBLICATIONS   988 CITATIONS

SEE PROFILE

Mingjian Yu
Zhejiang University
32 PUBLICATIONS   357 CITATIONS

SEE PROFILE

# Quantifying effects of habitat heterogeneity and other clustering processes on spatial distributions of tree species

Guochun Shen,[1,2,3] Fangliang He,[3,4] Rasmus Waagepetersen,[5] I-Fang Sun,[6] Zhanqing Hao,[7] Zueng-Sang Chen,[8] and Mingjian Yu[2,9]

[1]*Tiantong National Field Observation Station for Forest Ecosystem, East China Normal University, Shanghai 200241 China*
[2]*Key Laboratory of Conservation Biology for Endangered Wildlife, Ministry of Education, College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang 310058 China*
[3]*SYSU–Alberta Joint Lab for Biodiversity Conservation, State Key Lab of Biological Control and School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510275 China*
[4]*Department of Renewable Resources, University of Alberta, Edmonton, Alberta T6G 2H1 Canada*
[5]*Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg, Denmark*
[6]*Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualein 974 Taiwan*
[7]*State Key Laboratory of Forest and Soil Ecology, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, Liaoning 110016 China*
[8]*Department of Agricultural Chemistry, Taiwan University, Taipei 10617 Taiwan*

*Abstract.* Spatially explicit consideration of species distribution can significantly add to our understanding of species coexistence. In this paper, we evaluated the relative importance of habitat heterogeneity and other clustering processes (e.g., dispersal limitation, collectively called the non-habitat clustering process) in explaining the spatial distribution patterns of 341 tree species in three stem-mapped 25–50 ha plots of tropical, subtropical, and temperate forests. Their relative importance was estimated by a method that can take one mechanism into account when estimating the effects of the other mechanism and vice versa. Our results demonstrated that habitat heterogeneity was less important in explaining the observed species patterns than other clustering processes in plots with flat topography but was more important in one of the three plots that had a complex topography. Meanwhile, both types of clustering mechanisms (habitat or non-habitat) were pervasive among species at the 50-ha scale across the studied plots. Our analyses also revealed considerable variation among species in the relative importance of the two types of mechanism within each plot and showed that this species-level variation can be partially explained by differences in dispersal mode and growth form of species in a highly heterogeneous environment. Our findings provide new perspectives on the formation of species clustering. One important finding is that a significant species–habitat association does not necessarily mean that the habitat heterogeneity has a decisive influence on species distribution. The second insight is that the large species-level variation in the relative importance of the two types of clustering mechanisms should not be ignored. Non-habitat clustering processes can play an important role on species distribution.

*Key words: clustering processes; dispersal limitation; environmental control hypothesis; habitat heterogeneity; patchy distribution; spatial point pattern modeling; species–habitat association.*

## INTRODUCTION

Plants in nature are often patchily distributed (Greig-Smith 1983, He et al. 1997, Condit et al. 2000). Such aggregated distribution of plant species has great implications for species coexistence, because it can directly affect species interaction and resource use (Atkinson and Shorrocks 1981, Ives and May 1985). Many processes have been hypothesized to explain spatial distribution of species (Harms et al. 2001, Phillips et al. 2003, Seidler and Plotkin 2006, John et al. 2007, Pinto and MacDougall 2010), yet the ecological

mechanisms invoking spatial aggregation remain poorly understood.

Historically, environmental variation in space, or habitat heterogeneity, is considered the major determinative factor in spatially aggregated distribution of species (i.e., the classical environmental control model [Whittaker 1956, Bray and Curtis 1957]). Because the environmental variables are always spatially clustered, their effects will be reflected in species distribution through the species–habitat association (Harms et al. 2001). However, some other spatially clustering processes, e.g., dispersal limitation, can generate similar spatial aggregated distribution of species (Levine and Murrel 2003). Meanwhile, theoretical and empirical results suggest that habitat heterogeneity and other clustering

processes may simultaneously determine spatial distribution of species (Mouquet and Loreau 2003).

Evaluating and separating the relative importance of habitat heterogeneity and other clustering processes is of vital importance to understand the mechanisms of species distribution and coexistence. Several attempts have been made to this end (Harms et al. 2001, Tuomisto et al. 2003, Gilbert and Lechowicz 2004, John et al. 2007, Wiegand et al. 2007, Legendre et al. 2009), but significant gaps regarding the relative importance of the two types of clustering mechanisms on spatial distribution of species still remain. Most previous studies have only tested significance of either habitat heterogeneity or other non-habitat clustering processes on the spatial distribution of species, without properly controlling for the effects of the other factors (Harms et al. 2001, Plotkin et al. 2002, John et al. 2007, Wiegand et al. 2007). Since the effects of the two mechanisms are highly entangled, these results are likely confounded by this oversight (Legendre and Legendre 1998, Seidler and Plotkin 2006). Additionally, such tests provide little information about the relative importance of the two types of mechanisms. A significant species–habitat association, for example, does not necessarily mean that habitat has a big influence on spatial distribution of species.

Some studies have tried to estimate the relative importance of habitat heterogeneity and non-habitat clustering processes by decomposing the variance of a particular community-level summary statistic (species compositional dissimilarities between sites, i.e., beta diversity) into fractions explained by environmental and geographical distances between sites (Tuomisto et al. 2003, Gilbert and Lechowicz 2004, Legendre et al. 2009). However, this community-level approach cannot distinguish the potential effects of the two types of clustering mechanisms, especially their separated effects on the spatial distribution of individual species. For instance, species with different seed dispersal ability and life forms might respond differently to habitat heterogeneity. Yet these species-level differences have rarely been explored in previous studies. It is still far from clear to what extent the spatial distributions of different species depend on habitat heterogeneity and non-habitat clustering processes (Pinto and MacDougall 2010).

To overcome these problems, we introduce a method that can quantify the relative importance of habitat heterogeneity and other clustering processes on individual species distribution. The method is based on joint modeling of the two clustering processes. By this new method, species-level differences in the relative importance of the two types of mechanisms can be estimated. We can disentangle and quantify the effect of either type of the processes by controlling for the effect of the other. This leads to the estimation of the percentage of species significantly affected by either one or both types of the clustering mechanisms. Furthermore, it decomposes the variation in species distribution into the effects of habitat heterogeneity and non-habitat processes and thus it separates the relative importance of the two clustering processes on the spatial distribution of individual species. This new method relies on advanced spatial statistics and is based on fitting heterogeneous Cox point process models (Møller and Waagepetersen 2004, Jalilian et al. 2011) to individual species distributions. Given the best fitting model for each species, significance of habitat association and other clustering processes and proportions of variances explained by the two types of clustering mechanisms on species distribution can be reliably estimated.

In this study, we applied the method to disentangle and quantify the effects of habitat heterogeneity and non-habitat clustering processes on the spatial distributions of 341 tree species distributed in three forests in tropical, subtropical and temperate ecosystems. To understand the potential species-level differences in the importance of the two types of clustering processes, we examined the relationships between species life-history traits (dispersal mode and life form) and parameters reflecting the importance of the two types of the processes.

## METHODS

### Data

Three sets of data were used to test the roles of habitat heterogeneity and other clustering processes in structuring spatial distributions of trees. These sites form a latitudinal gradient from tropical rain forest on Barro Colorado Island (BCI), Panama, to subtropical forest at Fushan (FSH), Taiwan and temperate forest at Changbaishan (CBS), Northeast China (Table 1). All stems with diameter at breast height (dbh) $\geq$ 1 cm were located in each site plot and were species identified. Because the reliability of spatial pattern modeling requires some reasonable minimum population size (Baddeley et al. 2005), rare species with less than 11 individuals in each plot were excluded from this study, resulting in 341 species in total (Table 1). Since the mechanisms of spatial patterning of trees may undergo substantial changes with life stages, we divided each species in each plot into a small individuals group (dbh is smaller than the population median $dbh_{med}$) and a large individuals group (dbh $\geq$ $dbh_{med}$). The small and large individuals were then analyzed separately.

Four abiotic environmental variables and one biotic environmental variable were used to test the significance of habitat association for each species in each site. They are mean elevation, mean convexity, mean aspect, mean slope and total tree density in each 4 × 4 m quadrat (Harms et al. 2001, Valencia et al. 2004, Shen et al. 2009). Additional soil variables in each 4 × 4 m quadrat were also included in testing the species–habitat association: 13 soil variables in BCI plots, 10 for FSH, and 8 for CBS plots. Detailed information about the soil variables are given in Appendix A. The inclusion of the soil variables in each site is based on the availability of

TABLE 1. Description of the three forest dynamics plots.

| Plot | Location | Forest type | Size (ha) | Year the plot was censused | Number of species with abundance >10 | Number of stems with dbh ≥1 cm |
|---|---|---|---|---|---|---|
| Barro Colorado Island (BCI), Panama | 9.1543° N; 79.8461° W | tropical rain forest | 50 | 2000 | 228 | 197 829 |
| Fushan (FSH), Taiwan | 24.7614° N; 121.555° E | subtropical evergreen forest | 25 | 2004 | 78 | 128 458 |
| Changbaishan (CBS), China | 42.3833° N; 128.083° E | temperate deciduous forest | 25 | 2004 | 35 | 35 631 |

*Note:* Further information about each site can be found in Hubbell et al. (1999, 2005) and Condit (1998) for the BCI plot, Lin et al. (2010) for the Fushan plot, and Hao et al. (2007) for the Changbaishan plot.

the soil data and previous studies (e.g., John et al. 2007, Shen et al. 2009, Yavitt et al. 2009, Lin et al. 2010). A noticeable difference between the tree plots is that the FSH plot has a much more complex topography than the BCI and CBS plots (see Fig. 1).

### Modeling species distribution by spatial point process

Spatial point processes have been increasingly used to describe stem-mapped tree distribution (Seidler and Plotkin 2006, John et al. 2007, Shen et al. 2009). It commonly uses an intensity function $\lambda(u)$, analog to the probability density function in classical statistics, to describe tree distribution at any location $u$ (e.g., $u = (x,y)$ in two-dimensional space) in a plot (Møller and Waagepetersen 2004, Ripley 2004, Illian et al. 2008). If $du$ is the area of an infinitesimal area centered at $u$, then $\lambda(u)du$ is the probability that there is a tree in this area. In the current study, the distributions of trees are modeled by a random intensity function $\Lambda(u)$. Given a realization of $\Lambda(u)$, i.e., a particular intensity function $\lambda(u)$, the trees occur independently of each other. The resulting point process model is a Cox process (Diggle 1983, Stoyan et al. 1995). In order to model the effects of habitat heterogeneity and other clustering processes on tree distribution, we considered the following log-linear model for the random intensity function (Møller and Waagepetersen 2004):

$$\log \Lambda(u) = \mu + \mathbf{H}(u)\boldsymbol{\beta}^\top + D(u). \quad (1)$$

Here $\mu$ is an intercept, $\boldsymbol{\beta}$ is a vector of regression parameters of length $p$ and $\mathbf{H}(u)$ denotes a vector of habitat variables (e.g., aforesaid environmental variables in the three plots) at the spatial location $u$. The term $D(u)$ is a random, zero-mean, and spatially correlated "residual effect" that serves to model sources of aggregative variation not captured by the habitat regression term. The random fluctuations in the $D(u)$ create random clusters of trees located around peaks in the surface given by the $D(u)$. The spatial correlation of $D(.)$ was modeled by the Matérn covariance function $c(.)$ (Minasny and McBratney 2005), which can model different clustering behavior ranging from tight, well-

defined clusters to more dispersed and diffuse clusters (Jalilian et al. 2011). If $D(.)$ is a Gaussian process, there is a very simple relation between the covariance function $c(.)$ for $D(.)$ and the pair correlation function $g(.)$ of the Cox model. Thus parameters in $D(.)$ can be easily estimated by comparing the pair correlation function of the Cox model with the empirical pair correlation function for the observed data. A detailed introduction of the Cox process model is given in Appendix B.

The salient feature of the Cox model is that it incorporates effects of habitat heterogeneity (the $\mathbf{H}(u)\boldsymbol{\beta}^\top$
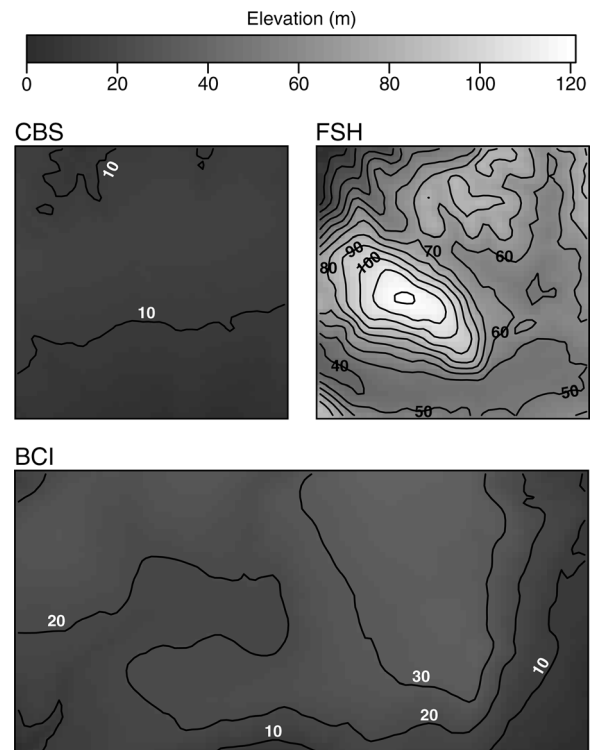


FIG. 1. Topography contours at the BCI, FSH, and CBS forest plots. The different shades represent the different relative elevations as defined in the legend (the same scale is used for all three plots). Darker shades indicate low elevation. See Table 1 for the abbreviations of the plot names.

term in Eq. 1) and other clustering processes (the $D(u)$ term in Eq. 1) simultaneously. This makes it possible to take one term into account when estimating the effect of the other (Waagepetersen and Guan 2009). In the assessment of the significance of habitat effects we adjusted standard errors to take into account spatial correlation due to $D(u)$. Moreover, when estimating the effect of $D(u)$, we filtered out aggregation due to habitat. More details on the estimation method are provided in Appendix B.

We finally note that commonly used homogeneous Poisson cluster point processes such as the Thomas process (e.g., Seidler and Plotkin 2006, John et al. 2007) are in fact also special cases of Cox processes (e.g., Møller and Waagepetersen 2004). However, in addition to random clustering our model has the benefit of also including aggregation due to habitat. The so-called inhomogeneous Thomas process does have a random intensity function that can be expressed in the form of Eq. 1, but our use of the flexible Matérn covariance function implies that a greater range of clustering behavior can be modeled.

### Parameter estimation, model selection, goodness-of-fit test, and spatial variance decomposition

Given an observed spatial distribution of a species and observed habitat maps, parameters of the above Cox process model were estimated by the two-step approach of Waagepetersen and Guan (2009). The model selections of the Cox process (e.g., which habitat variable should be included in $H(u)$ and whether there was a significant clustering $D(u)$ in the random intensity function Eq. 1) were conducted by step-wise model reduction (Waagepetersen and Guan 2009) and Loosmore's goodness-of-fit test (Loosmore and Ford 2006), respectively. The goodness of fit for the best-fitting Cox model for each species was assessed using the pair-correlation function (e.g., Illian et al. 2008). Details of parameter estimation, model selection, and goodness-of-fit tests are given in Appendix B.

From the formal definition of the Cox process model, the variance of a given random intensity function $\Lambda$ is totally determined by variance in $H(u)$ and variance in $D(u)$ (Jalilian et al. 2011). From the log linear random intensity function perspective, variance of $\log(\Lambda)$ can be simply expressed as $\mathrm{Var}[\log \Lambda(u)] = \mathrm{Var}\ H(u) + \mathrm{Var}\ D(u) = \mathrm{VH} + \mathrm{VD}$. Hence, we can use the following two ratios:

$$\frac{\mathrm{VH}}{\mathrm{VH} + \mathrm{VD}}, \ \mathrm{PVH} \qquad (2)$$

$$\frac{\mathrm{VD}}{\mathrm{VH} + \mathrm{VD}}, \ \mathrm{PVD} \qquad (3)$$

to measure relative importance of habitat and other clustering processes to each species. We refer to Eqs. 2 and 3, respectively, as the proportions of variance explained by habitat heterogeneity and by other clustering processes with abbreviations PVH and PVD.

Note that PVD = 1 − PVH. Estimation of VD and VH are based on the best-fitting Cox point processes model for each species. Specifically, VD is a parameter of $D(.)$ in Eq. 1 that can be obtained from the fitted Cox point process model. VH is estimated by a simple empirical estimate

$$\mathrm{VH} = \frac{1}{n_I - 1} \sum_{u=(x,y)\in I} (\hat{\hat{H}}(u) - \bar{H})^2$$

where $\hat{\hat{H}}(u) = H(u)\hat{\beta}^\top$, $\bar{H} = (1/n_I)\sum_{u=(x,y)\in I} \hat{\hat{H}}(u)$, $\hat{\beta}$ is the estimate of $\beta$, and $I$ denotes the set of the $n_I$ locations $u$ where $H(u)$ is observed.

Simulation studies in Appendix C showed that the estimation of the proportions PVH and PVD are nearly unbiased under various situations. The Supplement contains the R functions for estimating parameters of the Cox process, model selection, goodness-of-fit test, and calculation of PVH and PVD given by Eqs. 2 and 3 for each species and the simulation studies. All analyses were conducted in R 2.15.2 (R Development Core Team 2012). Performance of the model and fitted parameters for each species were given in Appendix D. A more exhaustive and general mathematical exploration of the spatial variance decomposition method can be found in Jalilian et al. (2011).

### Comparative analyses between the spatial distribution of species and their life-history traits

To understand the effects of life-history traits of species (dispersal mode and life form) on the spatial distribution of species, we examined the relationships between the variance explained by habitat VH, variance explained by other clustering processes VD, proportion of variance explained by habitat heterogeneity (PVH), and dispersal mode/species life form in each plot by the phylogenetic comparative method in which VH, VD, PVH, and life-history traits were transformed into phylogenetic independent contrasts before correlation analysis (Felsenstein 1985). Following Seidler and Plotkin (2006), dispersal mode was divided into three ordered states: from explosive, mechanical, or gravity mode with the smallest expected seed dispersal distance, to wind mode with intermediate expected seed dispersal distance, to animals (including mammals and birds) mode with the largest expected seed dispersal distance (Janson 1983). Life form was also divided into three states: tallest canopy trees, subcanopy trees, and understory shrubs. We then labeled each species in the three plots with a particular state of dispersal mode and state of life form (Wright et al. 2007).

### RESULTS

#### Effects of habitat heterogeneity and non-habitat clustering processes on species distribution

Results in Table 2 show that most of species (94.3% on average) in the three plots were either significantly affected by heterogeneous habitat or/and non-habitat

TABLE 2. Percentages of species whose spatial distributions are significantly ($P \leq 0.05$) affected by habitat association only, non-habitat clustering processes only, jointly affected by habitat association and non-habitat clustering processes, and affected by neither habitat association nor non-habitat clustering processes in small and large individual groups in each plot.

| Group and plot | Habitat association only (%) | Non-habitat clustering processes only (%) | Both types of clustering processes (%) | Neither habitat nor non-habitat processes (%) |
|---|---|---|---|---|
| Small individuals | | | | |
| BCI | 23.68 | 29.82 | 35.53 | 10.96 |
| FSH | 16.67 | 12.82 | 70.51 | 0.00 |
| CBS | 11.43 | 48.57 | 37.14 | 2.86 |
| Large individuals | | | | |
| BCI | 26.32 | 21.49 | 35.53 | 16.67 |
| FSH | 19.23 | 20.51 | 56.41 | 3.85 |
| CBS | 25.71 | 34.29 | 40.00 | 0.00 |

*Note:* See Table 1 for the abbreviation of the plot names.

clustering processes. Percentages of species significantly affected only by habitat heterogeneity or both types of clustering mechanisms increased from the small individual groups to the large individual groups at the BCI and CBS plots, but decreased at the FSH plot. For all plots and both small and large individuals, the majority of species were affected by both of the two types of clustering mechanisms.

### Relative contributions of habitat heterogeneity and non-habitat clustering processes to species distribution

The relative contributions of the two types clustering mechanisms to the spatial distribution of species are measured by the proportions of variances explained by habitat heterogeneity (PVH) and other clustering processes (PVD), given by the complementary Eqs. 2 and 3. Scatter plots of PVH against the rank of PVH for all species in each site are shown in Fig. 2. In the BCI and CBS plots with flat topography (Fig. 1), the average PVH was around 36.5% and most species (62.3–82.9% for the two sites) had PVH less than 0.5 (panels A, C, D, and F in Fig. 2). Although each separate estimate of PVH may have a fairly large uncertainty, the overall trend is that non-habitat clustering processes dominate the spatial variance for most species. In the FSH plot with complex topography (Fig. 1), the average PVH was slightly larger than 0.5 (53.0%) and about one-half (51.3%) of species had PVH less than 0.5 (Panels B and C in Fig. 2). For the BCI and CBS plots, the average PVH is smaller for the small individuals than for the large individuals.

Fig. 2 also highlights the considerable variation in PVH among species within each plot as well as the fact that there is a large proportion of species whose spatial patterns are dominated by habitat heterogeneity (e.g., PVH > 0.5 in Fig. 2).

### Correlations between the spatial patterns of species and their life-history traits

In general, no consistent significant correlation between the spatial characteristics of species distribution (PVH, VH, and VD) and species' life-history traits (dispersal mode and life form) was found in the BCI and CBS plots that have flat topographies, while consistent significant correlations were found in the FSH plot that has a complex topography (Table 3).

Specifically, species life form of the large individual group in FSH plot was positively correlated with PVH and was negatively correlated with VD. These correlations suggest that the relative importance of habitat association was highest in large canopy tree species and lowest in the large individual group of understory shrub species in FSH plot. Table 3 also showed that dispersal mode of species in the FSH plot was negatively correlated with PVH. Thus the relative importance of habitat association was highest in species with the smallest expected dispersal distance (e.g., explosive, mechanic or gravity dispersal mode), and was lowest in species with the longest expected dispersal distance (animals dispersal mode).

### DISCUSSION

The classical environmental control model emphasizes the role of environmental factors in forming spatial patterns of species (Whittaker 1956, Bray and Curtis 1957). It is expected that most species will show significant habitat association. Our results in Table 2 confirm this expectation (e.g., around 75% of species significantly affected by habitat heterogeneity in Table 2) and were not consistent with previous studies in which only a small percentage of species had significant habitat associations (Harms et al. 2001, John et al. 2007). The inconsistency arises because other clustering processes are not appropriately considered in those studies. Note that although the torus shift and homogeneous Poisson cluster process were used by Harms et al. (2001) and John et al. (2007) to control for the effect of other non-habitat clustering processes, variance of species distribution captured by these models could also be raised by habitat heterogeneity. The more appropriate models are the inhomogeneous Cox point process (Møller and Waagepetersen 2004, Shen et al. 2009, Jalilian et al. 2011). As shown in Table 2, habitat and other clustering processes are highly confounded. The improper control of non-habitat clustering processes will depress the effect
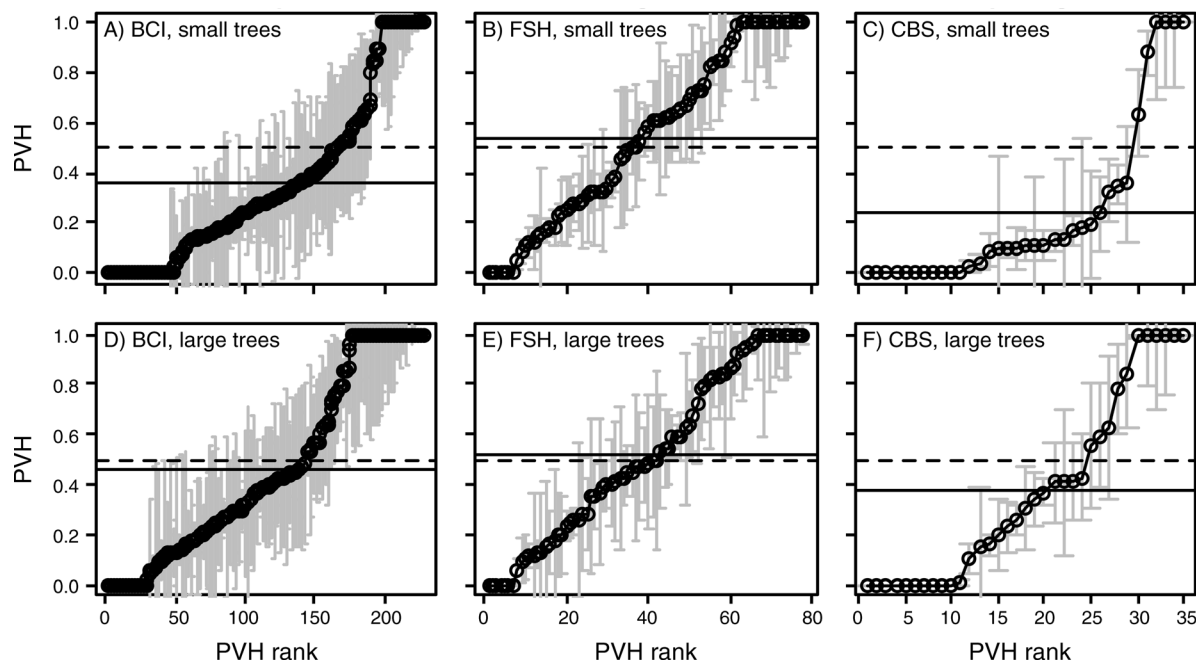
Fig. 2.   Proportion of variance explained by habitat association (PVH) in (A–C) small and (D–F) large individuals for each species plotted against its PVH rank for each plot. The vertical gray lines show one standard deviation generated by the bootstrap method. The dashed horizontal lines are the reference lines at which PVH equals 0.5. The solid horizontal lines represent the mean PVHs of all species in each plot. See Table 1 for the abbreviations of the plot names.

of habitat and lead to inaccurate conclusions concerning the importance of habitat conditions.

Our results confirmed that a significant species-habitat association does not necessarily mean that habitat determines the spatial distribution of the species. In the BCI and CBS plots, although habitat heterogeneity significantly affected the majority of species, it played a smaller role than non-habitat clustering processes in forming spatial patterns of species (Fig. 2). These results suggest that there is still a significant gap in our understanding about the importance of habitat hetero-geneity on spatial distribution of species in different forest communities, because most of previous studies only conducted significant tests of species-habitat association. The results of this study highlight the importance of clustering processes other than habitat conditions (e.g., dispersal limitation) for understanding aggregated distribution of species. Hubbell et al. (1999) showed that there was a strong seed dispersal limitation in the BCI plot.

The relative importance of habitat heterogeneity and non-habitat clustering processes seems to change with complexity of topography. The FSH plot has a much more complex topography than the BCI and CBS plots (Fig. 1). Correspondingly, the proportion of species with significant habitat association was largest (81%) in FSH plot (Table 2). The species mean PVH in FSH was also largest among the three plots and was even above 0.5 (Fig. 2). The percentage of species for which habitat association was more important than non-habitat

Table 3.   Pearson correlations between spatial distribution characteristics (the proportion of variance explained by habitat heterogeneity [PVH] and variances VH and VD associated with habitat heterogeneity or other clustering processes, respectively) and dispersal mode/life form of species.

| Group, trait, and plot | PVH | VH | VD |
|---|---|---|---|
| Small individuals | | | |
| Life form | | | |
| BCI | −0.046 | 0.05 | 0.043 |
| FSH | 0.078 | −0.048 | −0.076 |
| CBS | 0.228 | 0.259 | 0.384† |
| Dispersal model | | | |
| BCI | 0.008 | −0.071 | 0.014 |
| FSH | −0.347** | −0.161 | 0.159 |
| CBS | −0.036 | −0.255 | −0.02 |
| Large individuals | | | |
| Life form | | | |
| BCI | 0.126† | −0.005 | −0.092 |
| FSH | 0.327** | −0.209 | −0.294* |
| CBS | 0.237 | 0.167 | −0.017 |
| Dispersal model | | | |
| BCI | 0.035 | −0.05 | 0.038 |
| FSH | −0.229† | 0.007 | 0.152 |
| CBS | −0.253 | −0.211 | −0.007 |

*Notes:* PVH, VH, VD, and life-history traits were trans-formed into phylogenetically independent contrasts to remove bias from evolutionary relationships of trees. There are three dispersal modes, denoted as: 1, explosive/mechanic/gravity; 2, wind; and 3, animal. Life forms are: 1, shrub; 2, subcanopy tree; and 3, canopy tree. Numbers in the table are correlation coefficients. See Table 1 for the abbreviation of the plot names.
  * $P < 0.05$; ** $P < 0.01$; † $P < 0.1$.

clustering processes (e.g., PVH larger than 0.5) was again highest in the FSH plot and exceeded 50% for the small individual group. The results from the FSH plot seem to support the niche prediction that habitat has a stronger influence on species distribution when the habitat (represented here by topography of the plot) is more complex.

There was also some weak evidence in our results regarding the relationship between species life stage and the relative importance of the habitat association. From the small individual group to the large individual group, the percentage of species with significant habitat association (Table 2 and species mean PVH in Fig. 2) both increased in the BCI and CBS plots. These trends suggest that the effect of habitat heterogeneity on species distribution could accumulate over species life stages. Compared to the spatial patterns that can be quickly generated by non-habitat clustering processes (e.g., dispersal limitation), habitat heterogeneity might take time to leave observable effects on species distribution, especially at sites with less complex topography like the BCI and CBS plots. In sites of complex topography like the FSH, species life stages and the relative importance of habitat association (Table 2 and Fig. 2) were observed. This result implies that different complexities of topography can have different impacts on species distribution as predicted by the niche theory (Chase and Leibold 2003).

The different effects of topography on species distribution were also reflected in the relationship between species life-history traits and species-level importance of habitat heterogeneity. In the BCI and CBS plots (Fig. 1), no significant correlation between species life-history traits and importance of habitat was observed (Table 3), while in the FSH plot with complex topography, we observed that canopy tree species were more sensitive to the habitat condition than the understory shrubs, and species with short dispersal distance were more influenced by habitat than long distance dispersal species. These correlations also imply that species-level variation on the relative importance of habitat heterogeneity may matter to coexistence of tree species with different life-history traits under a highly heterogeneous environment, and is worth exploring in future studies.

LITERATURE CITED

Atkinson, W. D., and B. Shorrocks. 1981. Competition on a divided and ephemeral resource: a simulation model. Journal of Animal Ecology 50:461–471.

Baddeley, A., R. Turner, J. Møller, and M. Hazelton. 2005. Residual analysis for spatial point processes. Journal of the Royal Statistical Society: Series B 67:617–666.

Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325–349.

Chase, J. M., and M. A. Leibold. 2003. Ecological niches. University of Chicago Press, Chicago, Illinois, USA.

Condit, R. 1998. Tropical forest census plots. Springer-Verlag and R. G. Landes Company, Berlin, Germany, and Georgetown, Texas, USA.

Condit, R., et al. 2000. Spatial patterns in the distribution of tropical tree species. Science 288:1414–1418.

Diggle, P. J. 1983. Statistical analysis of spatial point patterns. Academic Press, London, UK.

Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Gilbert, B., and M. J. Lechowicz. 2004. Neutrality, niches, and dispersal in a temperate forest understory. Proceedings of the National Academy of Sciences USA 101:7651–7656.

Greig-Smith, P. 1983. Quantitative plant ecology. Third edition. University of California Press, Berkeley, California, USA.

Hao, Z., J. Zhang, B. Song, J. Ye, and B. Li. 2007. Vertical structure and spatial associations of dominant tree species in an old-growth temperate forest. Forest Ecology and Management 252:1–11.

Harms, K. E., R. Condit, S. P. Hubbell, and R. B. Foster. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. Journal of Ecology 89:947–959.

He, F., P. Legendre, and J. V. LaFrankie. 1997. Distribution patterns of tree species in a Malaysian tropical rain forest. Journal of Vegetation Science 8:105–114.

Hubbell, S. P., R. Condit, and R. B. Foster. 2005. Barro Colorado forest census plot data. https://ctfs.arnarb.harvard.edu/webatlas/datasets/bci

Hubbell, S. P., R. B. Foster, S. T. O'Brien, K. E. Harms, R. Condit, B. Wechsler, J. Wright, and S. Loo de Lao. 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science 283:554–557.

Illian, J., A. Penttinen, H. Stoyan, and D. Stoyan. 2008. Statistical analysis and modeling of spatial point pattern. John Wiley, Chichester, UK.

Ives, A. R., and R. M. May. 1985. Competition within and between species in a patchy environment: relations between microscopic and macroscopic models. Journal of Theoretical Biology 115:65–92.

Jalilian, A., Y. Guan, and R. P. Waagepetersen. 2011. Decomposition of variance for spatial Cox processes. Scandinavian Journal of Statistics. http://dx.doi.org/10.1111/j.1467-9469.2012.00795.x

Janson, C. H. 1983. Adaptation of fruit morphology to dispersal agents in Neotropical forest. Science 219:187–189.

John, R., et al. 2007. Soil nutrients influence spatial distributions of tropical tree species. Proceedings of the National Academy of Sciences USA 104:864–869.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Elsevier, Amsterdam, The Netherlands.

Legendre, P., X. Mi, H. Ren, K. Ma, M. Yu, I. F. Sun, and F. He. 2009. Partitioning beta diversity in a subtropical broad-leaved forest of China. Ecology 90:663–674.

Levine, J. M., and D. J. Murrel. 2003. The community-level consequences of seed dispersal patterns. Annual Review of Ecology, Evolution, and Systematics 34:549–574.

Lin, Y., L. Chang, K. Yang, H. Wang, and I. F. Sun. 2010. Point patterns of tree distribution determined by habitat heterogeneity and dispersal limitation. Oecologia 165:175–184.

Loosmore, N. B., and E. D. Ford. 2006. Statistical inference using the G or K point pattern spatial statistics. Ecology 87:1925–1931.

Minasny, B., and A. B. McBratney. 2005. The Matérn function as a general model for soil variograms. Geoderma 128:192–207.

Møller, J., and R. P. Waagepetersen. 2004. Statistical inference and simulation for spatial point processes. Chapman and Hall/CRC Press, New York, New York, USA.

Mouquet, N., and M. Loreau. 2003. Community patterns in source–sink metacommunities. American Naturalist 162:54–57.

Phillips, O. L., P. N. Vargas, A. L. Monteagudo, A. P. Cruz, M. E. C. Zans, W. G. Sánchez, M. Yli-Halla, and S. Rose. 2003. Habitat association among Amazonian tree species: a landscape-scale approach. Journal of Ecology 91:757–775.

Pinto, S. M., and A. S. MacDougall. 2010. Dispersal limitation and environmental structure interact to restrict the occupation of optimal habitat. American Naturalist 175:675–686.

Plotkin, J., J. Chave, and P. Ashton. 2002. Cluster analysis of spatial patterns in Malaysian tree species. American Naturalist 160:629–644.

R Development Core Team. 2012. R 2.15.2. R Project for Statistical Computing, Vienna, Austria. www.r-project.org

Ripley, B. D. 2004. Spatial statistics. John Wiley, Hoboken, New Jersey, USA.

Seidler, T. G., and J. B. Plotkin. 2006. Seed dispersal and spatial pattern in tropical trees. PLoS Biology 4:e344.

Shen, G., M. Yu, X.-S. Hu, X. Mi, H. Ren, I. F. Sun, and K. Ma. 2009. Species–area relationships explained by the joint effects of dispersal limitation and habitat heterogeneity. Ecology 90:3033–3041.

Stoyan, D., W. S. Kendall, and J. Mecke. 1995. Stochastic geometry and its applications. Second edition. Wiley, Chichester, UK.

Tuomisto, H., K. Ruokolainen, and M. Yli-Halla. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. Science 299:241–244.

Valencia, R., R. B. Foster, G. Villa, R. Condit, J. C. Svenning, C. Hernandez, K. Romoleroux, E. Losos, E. Magård, and H. Balslev. 2004. Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. Ecology 92:214–229.

Waagepetersen, R. P., and Y. Guan. 2009. Two-step estimation for inhomogeneous spatial point processes. Journal of the Royal Statistical Society: Series B 71:252–258.

Whittaker, R. H. 1956. Vegetation of the great smoky mountains. Ecological Monographs 26:1–80.

Wiegand, T., S. Gunatilleke, and N. Gunatilleke. 2007. Species associations in a heterogeneous Sri Lankan dipterocarp forest. American Naturalist 170:E77–E95.

Wright, S. J., A. Hernandéz, and R. Condit. 2007. The bushmeat harvest alters seedling banks by favoring lianas, large seeds, and seeds dispersed by bats, birds, and wind. Biotropica 39:363–371.

Yavitt, J. B., K. E. Harms, M. N. Garcia, S. J. Wright, F. He, and S. J. Mirabello. 2009. Spatial heterogeneity of soil chemical properties in a lowland tropical moist forest, Panama. Australian Journal of Soil Research 47:674–687.

## SUPPLEMENTAL MATERIAL

### Appendix A

Details of soil variables included in our analysis (*Ecological Archives* E094-223-A1).

### Appendix B

A detailed introduction to the Cox process model (*Ecological Archives* E094-223-A2).

### Appendix C

Simulation studies of the performance of the Cox process model (*Ecological Archives* E094-223-A3).

### Appendix D

Performance and estimated parameters of the best-fit model for each species in the three plots (*Ecological Archives* E094-223-A4).

### Supplement

R functions for estimating parameters of the Cox process, model selections, and calculating percentage of variances explained by habitat heterogeneity and other clustering processes (*Ecological Archives* E094-223-S1).