Analysing tree species and size diversity patterns in multi-species uneven-aged forests of Northern China

G. Hui a, A. Pommerening a,b,c

a Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Box 1958, 100091 Beijing, China
b Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland
c Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Kreutzwaldi 5, Tartu 51014, Estonia

A R T I C L E   I N F O

Article history:
Available online 20 August 2013

Keywords:
Relative abundance
Species saturation
Non-spatial diversity measure
Spatial diversity measures
Mark mingling
Mark differentiation

A B S T R A C T

The diversity of complex multi-species uneven-aged forests is not easy to describe efficiently and meaningful. In the past, most importance has been assigned to species diversity and there is a wide range of approaches to quantifying this aspect of diversity. In woodlands, size diversity is another important dimension of biodiversity, as trees can be small or large and depending on their size play different ecological roles. Traditionally size diversity has often been neglected in diversity surveys of woody vegetation or considered as difficult to handle. We investigated the size and species diversity of two multi-species uneven-aged forests from monsoon- and subtropical/warm-temperate climate zones in northern China using non-spatial and spatially explicit forestry and diversity characteristics. We found that the analysis of both species and size diversity provided a much better understanding of diversity patterns in complex multi-species forests. Interestingly the relationship between size- and individual-based relative abundances follows a saturation curve. The analysis using non-spatial forestry and biodiversity characteristics led to approximately similar results in all four plots despite the very different woodland communities involved. However, with spatial measures we gained a deeper insight concerning the deviation from random distributions of species and size as well as interaction ranges of trees. We also found an aggregation of similar species and sizes at small inter-tree distances in the monsoon influenced climate and an aggregation of different sizes in the subtropical/warm-temperate climate zone. The use of second-order characteristics was instrumental in elaborating the subtle differences between the diversity of the two study areas and thus made an important contribution to this study.

1. Introduction

Species diversity, a combination of species richness and relative abundance (Newton, 2007), is not the only aspect of biodiversity, but a rather important and the most commonly considered one (Kimmins, 2004, p. 429). Much research has already been dedicated to this topic and many different quantitative species diversity measures have been proposed. They provide important information on the distribution of species in plant communities. This information is crucial for understanding the structure and ecological processes involved in plant communities and for monitoring biodiversity. Species richness, an important aspect of species diversity, is often interpreted as a surrogate measure for other kinds of biodiversity: More species usually lead to greater genetic variation, i.e. there is a greater diversity of genes in the population.

This implies greater ecological variation and a better exploitation of niches and habitats (Magurran, 2004; Gaston and Spicer, 2004, p. 13; Krebs, 1999, p. 17).

A common assumption in diversity surveys and analyses is that all individuals of a population are equal (Magurran, 2004, p. 11f.). In contrast to many other vascular plants tree sizes can range from those of small seedlings to those of mature giants of up to a total height of 40 m or in some ecosystems even up to 100 m. Size diversity hugely contributes to structural diversity of woodlands and it is widely acknowledged that structural diversity largely affects the number and quality of habitats and niches a forest can offer (Kimmins, 2004). Size obviously does matter in diversity surveys of woody vegetation, as the individuals involved occupy different niches and fulfil different ecological roles (Newton, 2007). Also there is a clear relationship between size and abundance – species involving large body sizes are rarer than many species with smaller body sizes (Gaston and Spicer, 2004, p. 10).

Only comparatively recently the interplay between species and size has been addressed in functional diversity (Botta-Dukát, 2005; Mason et al., 2005; Lavorel et al., 2008). Functional diversity is
generally defined as the distribution of traits in a community (Diaz and Cabido, 2001; Tilman, 2001). Species abundance can be combined with these traits to result in a community weighted mean (CWM), where traits can for example be different measures of tree size although they are not limited to quantitative information. Also a number of different traits can be simultaneously included in multivariate analyses (Laliberté and Legendre, 2010; Clark et al., 2012).

At another level, spatial measures of species and size of trees provide even more information on biodiversity patterns in forests than non-spatial diversity indices (Pommerening, 2002). They explain how diversity changes with different scales (Dale, 1999) and extend information derived from non-spatial measures. In addition we learn for example about minimum and maximum interaction ranges between trees. Also we can study whether species and size variables form specific spatial patterns that deviate from random assignment. If so we can continue and find out whether there is an aggregation of similar or different species and sizes at close proximity (Diggles, 2003; Illian et al., 2008). Spatial measures involving species and size have, however, been rarely used so far in biodiversity analyses (Krebs, 1999; Magurran, 2004). These few examples clearly highlight the usefulness of spatially explicit diversity analyses.

In this context, mixed species forests are particularly interesting for diversity research as much variety of tree species and sizes can be expected in this forest type. Pretzsch et al. (2010, 2013) for example were able to demonstrate a number of benefits of mixed-species forests in Central Europe, including increased biomass production, niche complementarity on poor sites and greater resilience to impacts.

Analysing species and size diversity of woodland communities that comprise many different species is challenging (Bagchi et al., 2011). Large numbers of species can often be found in tropical and subtropical forests (Wright, 2002; Wiegand et al., 2007), but occasionally also in temperate climates. These challenges are also addressed in this paper.

To study tree diversity at different levels we have analysed spatially explicit tree data from semi-natural woodlands in the monsoon- and subtropical/warm-temperate climate zones of northern China.

The objective of this paper is to identify similarities and dissimilarities between patterns of species and size diversity in multi-species uneven-aged forests. For this purpose we (1) analyse the data using non-spatial diversity characteristics. (2) In addition we employ two new second-order characteristics, the mark minter- and mark differentiation functions (Pommerening et al., 2011), as spatial measures of tree diversity for a more detailed analysis. (3) Finally we interpret all results for the benefit of a holistic description of the tree diversity of the analysed forest types and we compare and discuss the performance of and information produced by the different tree diversity metrics.

2. Materials and methods

2.1. Data

For this study we have selected four large replicated monitoring plots from semi-natural forests, which have been taken out of forest management 50 years ago and can now be considered as biodiversity hotspots in northern China. The observed tree patterns are now increasingly the consequence of natural processes without human disturbances and therefore also represent typical zonal plant communities in this part of China.

The first set of two monitoring plots (B and F) is located in the Dongdapo Natural Reserve (43°51′–44°05′N and 127°35′–127°51′E), which has a size of 16,660 ha and is part of the Zhangguangcai mountain range extending from north of the Songhua river to south of the Changbaishan mountains (see Fig. 1, bottom left and Fig. 2, top).

This area is approximately 45 km from Jiaohe in Jilin Province, China, and has a monsoon climate with dry, windy springs and warm, wet summers. The tree vegetation primarily consists of 15 tree species, among which Juglans mandshurica Maxim., Fraxinus mandshurica Rupr. and Phellodendron amurense Rupr. are used to define forest stand types and are known as the “three great hardwoods” in north-eastern China because of their economic importance and increasing value in forest conservation. The two monitoring sites in this area are both 1 ha in size.

The second set of monitoring plots (H and I) is located in mixed pine-oak forest on the north-facing slopes of the Qining Mountain Range (see Fig. 1, right and Fig. 2, bottom). This area in the transition zone between the subtropical and the warm-temperate region also forms a natural reserve and contains 432,914.0 ha of forest land. The plots were located in the 57th compartment of the Baihua forest farm (33°30′–34°49′N, 104°22′–105°43′E) on Xiaolong Mountain, Gansu Province. The forest contains more than 30 species including Quercus liaotungensis Koehne, Carya cathayensis Sarg., and Pinus armandii Franch. The dominant species is Quercus aliena var. acuteserrata Maxim., which accounts for about 20% of the total number of trees in the plots. The forest is highly natural and its spatial structure is complex. The monitoring plots have a size of 0.49 ha each.

Compared to temperate forests in other parts of the world the study woodlands show a high species diversity and are therefore a challenge to any characterisation of tree diversity. This makes them an ideal object for comparative studies using different measures of tree diversity.

2.2. Non-spatial diversity measures

2.2.1. Species diversity measures

For quantifying tree species diversity we used the evenness measure of the Shannon index (Shannon and Weaver, 1949, Eq. (1)), an information theory index. The Shannon index was originally proposed to quantify the entropy, i.e. the uncertainty of information, in strings of text (Krebs, 1999, p. 444). By contrast, the Simpson index (Simpson, 1949, Eq. (2)) is a dominance or concentration measure.

\[
E_p = H / \ln s
\]

with \( H = - \sum p_i \ln(p_i) \)

\[
E_D = D / s
\]

with \( D = 1 - \sum p_i^2 \) (1)
(2)

\( p_i \) is the proportion of individuals found in the ith species and \( s \) is the number of species. The Simpson index has been calculated in various ways and we have followed Magurran’s (2004, p. 116) recommendation to adopt the reciprocal as opposed to the complement form for calculating the evenness measure. Whilst the Shannon evenness measure emphasizes the species richness component of diversity the Simpson evenness index is weighted by abundances of the commonest species. Both Shannon and Simpson measures are among the most meaningful and robust diversity measures available (Krebs, 1999, p. 444ff.; Magurran, 2004, p. 114ff.).

The evenness forms (Eqs. (1) and (2)) are often used as standardisation to allow for comparisons between different monitoring sites (Pretzsch, 2009, p. 280). Both species diversity measures can be calculated based on tree number proportions and basal area proportions. Basal area, \( g \) (measured in \( m^2 \)), relates to stem diameter, \( dbh \) (measured in cm), through \( g = \pi (dbh/200)^2 \). In the case of
basal area proportions the conspecific tree sizes are taken into account.

We have also depicted relative tree species abundance in terms of number of individuals, \( i_{\text{rel}} \), of the ten commonest species over relative species abundance in terms of basal area, \( B_{\text{rel}} \), in order to understand the relationship between species and size diversity. Unexpectedly we found that the trend of the relationship between these two relative abundance measures can be well described using the Michaelis–Menten saturation curve (Michaelis and Menten, 1913; Bolker, 2008, p. 77ff.):

\[
B_{\text{rel}} = \frac{a}{b + r_{\text{rel}}},
\]

where \( a \) and \( b \) are model parameters to be determined through non-linear regression. Model parameter \( a \) can be interpreted as the asymptote of the saturation curve, i.e. the total number/basal area of species in the population, whilst \( b \) represents the half-maximum. Magurran (2004, p. 81f.) extensively discusses the use of the Michaelis–Menten model in species richness estimations.

### 2.2.2. Size diversity measures

The size variable or mark used in this study is stem diameter at 1.3 m above ground level and the derived variable basal area. A simple and yet effective measure of size diversity is the coefficient of variation, \( cv \), given in

\[
CV = \sqrt{\frac{\sum_{i=1}^{n} (dbh_i - \bar{dbh})^2}{(n-1) \cdot s_{dbh}^2}},
\]

where \( n \) is the number of individual trees and \( \bar{dbh} \) the arithmetic mean \( dbh \) (Porkress, 2004, p. 3; Sterba and Zingg, 2006).

Skewness of the empirical diameter distribution is another important size diversity measure (Sterba and Zingg, 2006) and can be calculated according to

\[
sk = \frac{\sum_{i=1}^{n} (dbh_i - \bar{dbh})^3}{(n-1) \cdot s_{dbh}^3},
\]

For interpretation the following guidance can be used (Porkress, 2004, p. 231):

- \( sk < 0 \) negatively skewed diameter distribution
- \( sk = 0 \) symmetric diameter distribution
- \( sk > 0 \) positively skewed diameter distribution

The last measure of non-spatial size diversity used in this study is the Lorenz curve and the associated Gini coefficient. The Lorenz curve is a graph of relative concentration. After ranking individual trees by their size the cumulative sizes are plotted against their cumulative ranks. We have ranked individual trees according to their basal area, \( g_r \), in ascending order \( g_1^\bullet = \{g_r^\bullet\} \) where \( r = 1, \ldots, n \) and \( g_1^\bullet < g_2^\bullet < \ldots < g_n^\bullet \). The basal areas can be accumulated as

\[
M_r = \sum_{j=1}^{r} p_j^\bullet / (r = 1, \ldots, n),
\]

where \( p_j^\bullet = g_j^\bullet / g \) and \( g = \sum_{j=1}^{n} g_j^\bullet \).

The cumulative proportions of stem density are calculated in relation to the total number of trees, \( n \), and represented on the abscissa:

\[
x_r = \sum_{j=1}^{r} \frac{n_j^\bullet}{n} = \frac{r}{n}
\]

The Lorenz curve therefore depicts the relative increase in basal area, \( M_r \), against the relative increase in the number of trees, \( x_r \), accounted for each additional tree of increasing size (Valbuena et al., 2012). The more the Lorenz curve deviates from the plot diagonal the more diverse is the size structure. One Lorenz curve indicates more diversity than another one if it always runs below the other one throughout its range. The Gini coefficient is a summary characteristic derived from the Lorenz curve. It is defined as

\[
gc = \frac{0.5 - LA}{0.5}
\]

where \( LA \) is the area between the Lorenz curve and the abscissa, i.e. the Lorenz area. The more diverse the diameter structure the larger gc. Gini coefficient and coefficient of variation are correlated. We calculated both the Lorenz curve and the Gini coefficient using the R package lawstat (Hui et al., 2008).

### 2.3. Combined species and size diversity measures

Functional diversity is considered as a key driver of ecosystem processes (Hooper et al., 2005), ecosystem resilience to environmental change (Folke et al., 2004) and ecosystem services (Díaz et al., 2007). This has led to the development of several indices for measuring functional diversity. These combine species abundance and functional traits, which provide additional information (Laliberté and Legendre, 2010) and allow simultaneous calculation.

For illustrating this concept we used a single-trait variant of the recently suggested index of functional dispersion, \( FDis \) (Anderson, 2006; Laliberté and Legendre, 2010). In our case, this single trait is species basal area derived from stem diameter. In a first step, abundance weighted mean basal area is calculated as

\[
g_{w} = \sum_{i=1}^{s} p_i \cdot g_i
\]
Then the absolute distances, $d_i$, between $g_i$ and $g_w$ are calculated using

$$d_i = \left| g_i - g_w \right|$$

For better comparison between different forests these distances are then standardized according to

$$d_i = \frac{d_i - d_{\text{min}}}{d_{\text{max}} - d_{\text{min}}}$$

Finally, $F\text{Dis}$ is calculated as abundance weighted mean distance as in

$$F\text{Dis} = \sum_{i=1}^{s} p_i \cdot d_i$$

2.4. Spatial diversity measures

2.4.1. Mark mingling function

The mark mingling function is a second-order characteristic that is based on the general concept of the mark correlation function (see for example Diggle, 2003; Illian et al., 2008). The core of this characteristic is the test function, $t_1$. Based on Gadow (1993) $t_1$ is defined as

$$t_1(m_1, m_2) = I(m_1 \neq m_2),$$

where $m_1$ and $m_2$ are the marks of pairs of tree locations with inter-tree distance $r$. In the case of this summary characteristic we consider qualitative marks, i.e. the tree species. $I(\cdot)$ is an indicator function that returns the value of 1 if the condition in the brackets
is fulfilled, otherwise it returns the value of 0. $t_1$ only considers the similarity or dissimilarity of species. This is a useful property in forests with many species that are often represented only by few individuals (Pommerening et al., 2011).

Following Illian et al. (2008, p. 354f.) the estimator of the mark mingling function can be defined in analogy to the estimator of the mark correlation function:

$$\hat{v}(r) = \frac{1}{n \cdot \min(x_1, x_2)} \sum_{i=1}^{n} \frac{1}{2 \pi r A(W_{x_1} \cap W_{x_2})} \sum_{i \neq j} \mathbf{1}(m(x_i) \neq m(x_j)) k_0(|x_i - x_j| - r),$$

(14)

$x_1$ and $x_2$ are arbitrary points of the point pattern in the observation window $W$. $k_0$ is the Epanechnikov kernel function, $A(W_{x_1} \cap W_{x_2})$ is the area of intersection of $W_{x_1}$ and $W_{x_2}$, see Illian et al. (2008, p. 481f. and p. 188), relating to the translation edge-correction (Ohser and Stoyan, 1981).

$EM$, expected mingling, is used as a normalisation term (see Illian et al., 2008, p. 346) and can be calculated according to Lewandowski and Pommerening (1997) as

$$EM = \sum_{i=1}^{s} \frac{n_i(n - n_i)}{n(n - 1)}$$

(15)

with $s$, the number of species, $n_i$ the number of trees in the observation window and $n_i$, the number of trees of species $i$. Information on the theoretical foundation of this new function can be found in Pommerening et al. (2011). For a graphical interpretation of the mark mingling function see Fig. 3 (left).

2.4.2 Mark differentiation function

The mark differentiation function is also a second-order characteristic based on the general concept of the mark correlation function. However, this function requires quantitative marks. The test function of the mark differentiation function, $t_2$, also follows an idea by Gadow (1993) and is defined as

$$t_2(m_1, m_2) = 1 - \frac{\min(m_1, m_2)}{\max(m_1, m_2)},$$

(16)

where $m_1$ and $m_2$ are again the marks of pairs of tree locations with inter-tree distance $r$. As we are interested in size diversity, we have used stem diameters at breast height as quantitative marks. The estimator of the mark differentiation function can again be defined in analogy to the estimator of the mark correlation function:

$$\hat{t}(r) = \frac{1}{n \cdot \min(x_1, x_2)} \sum_{i=1}^{n} \frac{1}{2 \pi r A(W_{x_1} \cap W_{x_2})} \sum_{i \neq j} \mathbf{1}(m(x_i) \neq m(x_j)) k_0(|x_i - x_j| - r),$$

(17)

For calculating expected mark differentiation as a normalisation term, the marks (here tree diameters, $dbh$) need to be sorted in ascending order, i.e. $i < j \Rightarrow dbh_i \leq dbh_j$. As a result we obtain the index set $J$ of a given forest. Then the auxiliary measure $R$ is defined as

$$R_i = \int_{dbh_i}^{dbh_j} \frac{1}{k(k-1)} \sum_{j=1}^{k} R_j$$

(18)

Expected mark differentiation, $ET$, can now be calculated as

$$ET = 1 - \frac{2}{k(k-1)} \sum_{j=1}^{k} R_j$$

(19)

Additional information on the performance of this new function can again be found in Pommerening et al. (2011). For a graphical interpretation of the mark differentiation function see Fig. 3 (right).

Note that “aggregation of similar marks” can also be termed “positive autocorrelation”. In analogy, “aggregation of different marks” corresponds to “negative autocorrelation”. We have decided to use the term “interaction range” for correlation range to stress the biodiversity and ecological focus in this paper. Incidentally, the mark differentiation function can also be understood as a variant of the mark variogram and we therefore also include this summary characteristic as a reference in the analysis. The mark variogram is explained in detail in Pommerening and Särkkä (2013).

We have implemented both functions in R (R Development Core Team, 2013) and C++ using the Rcpp package (Eddelbuettel and François, 2011). The corresponding implementation files are freely available on http://www.cran.r.org.

2.5. Testing

In ecology, the testing of spatial independence of marks plays a fundamental role. In this study, we considered tree species and stem diameters as marks. If marks are independent, there are no interactions between the associated trees as far as the marks under study are concerned and a more detailed spatial analysis cannot be justified (Grabarnik et al., 2011). Independent marks imply a random assignment of species/diameter marks that precludes a definite dispersion pattern in which ecologists are interested in (Kimmins, 2004, p. 367).

Fig. 3. Schematic diagram of the mark mingling, $v(r)$, (left) and the mark differentiation function, $t(r)$, (right). Dotted lines: Aggregation of different tree characteristics; Dashed lines: Aggregation of similar tree characteristics (after Pommerening et al., 2011). $r_{int}$ is the interaction range indicating the maximum extent of spatial correlation between marks (tree species and stem diameters in this study).
When analysing tree species usually the random superposition (or population independence) test is used (Illian et al., 2008; Grabarnik et al., 2011; Pommerening et al., 2011). This test assumes that the species involved form independent point patterns in the same area. To test this hypothesis, in the bivariate case usually the locations of the trees of one species are fixed whilst those of the other species is shifted by a random vector. Periodic boundary conditions apply to ensure that the trees are not shifted over the boundaries of the observation window (Illian et al., 2008). Since we have on average 27 different species in each monitoring plot we decided to always shift a group of species consisting of those with the largest numbers of individuals so that their sum roughly equals half of the total number of trees. We fixed the locations of the other tree species representing the other half of the total tree numbers.

For quantitative marks such as tree diameters the random labeling or random reallocation test is recommended (Illian et al., 2008; Grabarnik et al., 2011). This test assumes that the tree diameters have been randomly assigned to the tree locations.

As a first step we applied the traditional envelope test. To this end 1000 Monte Carlo simulations were performed and we calculated from the resulting functions the 95% envelopes marking the acceptance region of the independence hypothesis. The lower and upper envelopes give the 2.5% and 97.5% quantiles of the sorted values, respectively.

Since Grabarnik et al. (2011) have demonstrated that deviation tests are superior to envelope tests, we have also included deviation tests in our study.

In analogy to the mark-weighted $L_{d}$ function ($L_{c}$) (corresponding to the mark variogram) we have developed $L_{m}(r)$ and $L_{s}(r)$ corresponding to the mark mingling and the mark differentiation functions. We use these characteristics to carry out both envelope and deviation tests as detailed in Grabarnik et al. (2011).

Finally, following Illian et al. (2008) the bandwidth of all summary characteristics using non-parametric density functions was selected to produce a clear and smooth trend line whilst avoiding the masking of important details.

### Table 1

Forestry and diversity characteristics of the two study sites and the four plots. $d_{g}$ – quadratic mean diameter, $d_{100}$ – quadratic mean diameter of the 100 largest trees per hectare, SPH – relative species frequencies based on stem numbers, BA – relative species frequencies based on basal area.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Plot B</th>
<th>Plot F</th>
<th>Plot H</th>
<th>Plot I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees (ha)</td>
<td>797.0</td>
<td>936.0</td>
<td>932.7</td>
<td>842.9</td>
</tr>
<tr>
<td>Basal area (m²/ha)</td>
<td>31.7</td>
<td>28.7</td>
<td>27.9</td>
<td>25.3</td>
</tr>
<tr>
<td>$d_{g}$ (cm)</td>
<td>22.5</td>
<td>19.8</td>
<td>19.5</td>
<td>19.6</td>
</tr>
<tr>
<td>$d_{100}$ (cm)</td>
<td>44.8</td>
<td>40.6</td>
<td>41.5</td>
<td>37.5</td>
</tr>
<tr>
<td>Minimum dbh (cm)</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.4</td>
</tr>
<tr>
<td>Maximum dbh (cm)</td>
<td>77.7</td>
<td>79.2</td>
<td>61.0</td>
<td>70.5</td>
</tr>
<tr>
<td>Shannon evenness (SPH)</td>
<td>0.83</td>
<td>0.81</td>
<td>0.74</td>
<td>0.77</td>
</tr>
<tr>
<td>Shannon evenness (BA)</td>
<td>0.76</td>
<td>0.72</td>
<td>0.55</td>
<td>0.63</td>
</tr>
<tr>
<td>Simpson evenness (SPH)</td>
<td>0.49</td>
<td>0.48</td>
<td>0.27</td>
<td>0.29</td>
</tr>
<tr>
<td>Simpson evenness (BA)</td>
<td>0.37</td>
<td>0.33</td>
<td>0.11</td>
<td>0.17</td>
</tr>
<tr>
<td>dbh skewness</td>
<td>1.25</td>
<td>1.37</td>
<td>1.35</td>
<td>1.07</td>
</tr>
<tr>
<td>dbh coefficient of variation</td>
<td>0.71</td>
<td>0.67</td>
<td>0.72</td>
<td>0.64</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.63</td>
<td>0.62</td>
<td>0.64</td>
<td>0.58</td>
</tr>
<tr>
<td>Lorenz area</td>
<td>0.185</td>
<td>0.190</td>
<td>0.180</td>
<td>0.210</td>
</tr>
<tr>
<td>Functional diversity</td>
<td>0.33</td>
<td>0.55</td>
<td>0.45</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Fig. 4. Relative abundances in terms of number of trees and relative basal area of the ten most frequent tree species in the four plots. According to Gaston (1994) and Magurran (2004, p. 666) the rarest and most common among the ten most frequent tree species were identified by the first and third quartile (dashed lines) in terms of species proportions. The numbers represent the original species codes and the parameters $a$ and $b$ are those of the Michaelis-Menten saturation function (Eq. (3)) used as trend line (grey).
3. Results

3.1. Non-spatial forestry and diversity characteristics

The highest tree density in terms of number of individuals per hectare can be found in plot F, followed by plot H and plot I. The lowest density is in plot B (see Table 1 and Fig. 2). In terms of basal area per hectare the density ranks follow the order of plots in Table 1. Plots B and F obviously have larger diameter trees than H and I. This is also confirmed by the two mean diameters, \(d_g\) and \(d_{100}\), in Table 1 although the differences are comparatively small.

The contour lines in Fig. 2 convey the visual impression of larger localised density differences, particularly in plots B and F. Still the assumption can be accepted that the four patterns belong to an underlying point process, which is homogeneous (or stationary) and isotropic, i.e. their probability distributions are invariant to translations and rotations (Diggle, 2003; Illian et al., 2008, p. 42 and p. 35ff.). The point densities do not vary systematically.

In general, the forestry characteristics (upper half of Table 1) of the two plots belonging to the same study areas are largely similar. Compared to temperate forests in Europe the Shannon and Simpson measures of evenness are very high in all four plots. The basal-area based Shannon and Simpson measures of evenness always have smaller values than the tree-number based indices. This is often the case and can be explained by an uneven-distribution of tree basal area over the different species. The differences between basal-area and tree-number based Shannon and Simpson measures are largest in plot H.

All skewness indices indicate positively skewed diameter distributions, which are confirmed by Fig. 5. Both \(dbh\) coefficient of variation and Gini coefficient show the same trends in the four plots and identify plot H as the most diverse. The values of these two indices are of similar magnitude in the four plots.

The values of the functional diversity index, \(F_D\), are largest in plot F followed by plot H, lowest in plot B. These results greatly differ from Shannon and Simpson evenness and also from the size diversity measures.

Fig. 4 shows relative abundance in terms of tree numbers over relative abundance in terms of basal area of the ten commonest tree species in each of the four monitoring plots. As plots B and F and plots H and I are replicates it is not surprising that the most common species are similar. There is always at least one species among the ten main species, which can be classified as rare and at least three species form a group of the most frequent species (indicated by the dashed quartile lines). However, the rarest and commonest species do not coincide in the replicates. Only in plot H the most common species, Quercus aliena, reaches a basal-area proportion of 0.5. In all other plots the species specific basal-area proportions are considerably lower. The tree-number proportions are usually not larger than 0.2, again with the notable exception of Quercus aliena in plot H. There is always one lead species with the exception of plot F (where there are two), i.e. Juglans mandshurica in plot B, Quercus aliena in plot H and plot I and the basal-area proportions of these species are well in excess of 0.2. The many species occurring in these multi-species forests are only represented by a comparatively small number of individuals. Interestingly the relationship between tree-count based and basal-area based relative abundances follows a saturation curve, which can be modelled well by the Michaelis–Menten equation (Eq. (3)). From rarest to commonest species the abundances increase more in basal area then in number of individuals or in other words the most frequent species have individuals of larger size than the rarer species.

![Fig. 5](image-url)
Fig. 5 shows the diameter distributions of the four plots both as histograms and as trend lines. Particularly the trend lines give the impression of negative exponential diameter distributions in plots B, F and H whilst it could be argued that there is a bimodal distribution in plot I. Plot I also has the lowest skewness, lowest diameter coefficient of variance values and smallest Gini coefficient. In all four plots there is obviously a large proportion of small trees representing the understorey-reinitiation phase according to Oliver and Larson (1996) and only a small proportion of very large trees belonging to what could be described as old-growth phase. The negative exponential shape is indicative of a highly diverse size structure and can be found in many natural and near-natural forests (Kimmins, 2004).

In addition to the values of the Gini index (and the associated Lorenz areas) in Table 1 we also plotted the Lorenz curves (Fig. 6). The large area between the plot diagonal and the Lorenz curve emphasizes again the diverse size structure in all four monitoring plots and confirms the impression made by the results of the size-based diversity indices in Table 1 and by the empirical diameter distributions of Fig. 5. The shape of the Lorenz curves is very similar in all four plots perhaps with the slight exception of plot I where the Lorenz area is slightly smaller than in the three other plots. This again is also confirmed by the Gini coefficient, the Lorenza area and the diameter coefficient of variance in Table 1. H is the most diverse plot. This trend is only partly confirmed by the index of functional diversity that identified plot F as the most and plot B the less diverse of all monitoring plots.

3.2. Second-order characteristics as spatial measures of tree diversity

The mark mingling and the mark differentiation functions have been included in this study to refine our understanding of the tree diversity in the two geographic areas.

In plot B (Fig. 7, top left) there is an aggregation of similar species or positive autocorrelation up to a distance of \( r = 9\) m, the interaction range. This implies that conspecific trees are preferably arranged in clusters with a diameter of up to 9 m. The mark differentiation function (Fig. 7, bottom left) is indicative of an aggregation of similar-sized trees in plot B, however, this aggregation effect is weaker than for species mingling. Also the interaction range is larger than for species mingling. The mark variogram in Fig. 11 (top) confirms this finding and offers the additional information that the diameter variance in plot B is the largest of all plots. The results of the envelope tests in Figs. 7 and 11, respectively, indicate that the spatial species and size patterns deviate from the hypothesis of complete spatial randomness or mark independence, i.e. the described spatial trends in this plot are significant. The deviation test, however, suggests that the mingling patterns reflect independent marks (Table 2).

Not surprisingly Fig. 8 suggests spatial diversity patterns in plot F, which are similar to those of its replicate, plot B. However, the interaction range is larger in terms of mingling and shorter in terms of diameter differentiation than in plot B. Also in plot F there is less aggregation of similar tree diameters at close proximity than in plot B. The mark variogram of this plot reveals the same trend...
(Fig. 11, 2nd row from the top), although the diameter variance is much lower. Again the $L$ functions and the corresponding envelope tests suggest significant mingling and differentiation patterns. The deviation tests again reveal a mingling pattern corresponding to independent marking. Interestingly the deviation test based on mark differentiation leads to different results than the one based on the mark variogram (Table 2).

In the completely different geographic area of plots H and I the mingling and differentiation patterns are similar to those in plots B and F. However, the aggregation of similar species or positive autocorrelation at close proximity is, particularly in plot I, are more pronounced then in plots B and F. Whilst the interaction range for species mingling is approximately 20 m in plot H, it is only ca. 7 m in plot I.

The diameter differentiation patterns in plots H and I differ somewhat from plots B and F. Around $r = 2.5$ m there is evidence of a local maximum indicating an aggregation of different tree sizes or negative autocorrelation. This effect, however, is not very strong and the patterns otherwise seem to be close to mark independence. Interestingly the envelope tests (Figs. 9 and 10, Table 2)

---

**Table 2**

Results of the envelope and the deviation tests. The test statistics $\hat{p}_5$ and $\hat{p}_6$ are those related to Eqs. (5) and (6) in Grabarnik et al. (2011). A rejection of the independence hypothesis if marked by a negative sign and the positive signs indicate no rejection.

<table>
<thead>
<tr>
<th>Plots</th>
<th>$L$ function used</th>
<th>$\hat{p}_5$</th>
<th>$\hat{p}_6$</th>
<th>Envelope test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plot B</td>
<td>$\hat{l}_r$</td>
<td>0.223*</td>
<td>0.402*</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_t$</td>
<td>0.001*</td>
<td>0.001*</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_c$</td>
<td>0.020</td>
<td>0.009</td>
<td>–</td>
</tr>
<tr>
<td>Plot F</td>
<td>$\hat{l}_r$</td>
<td>0.395*</td>
<td>0.201*</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_t$</td>
<td>0.006*</td>
<td>0.013</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_c$</td>
<td>0.379*</td>
<td>0.444</td>
<td>–</td>
</tr>
<tr>
<td>Plot H</td>
<td>$\hat{l}_r$</td>
<td>0.082*</td>
<td>0.049</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_t$</td>
<td>0.001*</td>
<td>0.250*</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_c$</td>
<td>0.180*</td>
<td>0.837*</td>
<td>–</td>
</tr>
<tr>
<td>Plot I</td>
<td>$\hat{l}_r$</td>
<td>0.100*</td>
<td>0.071</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_t$</td>
<td>0.100</td>
<td>0.096</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$\hat{l}_c$</td>
<td>0.160*</td>
<td>0.086</td>
<td>–</td>
</tr>
</tbody>
</table>
reject the hypothesis of mark independence, i.e. the weak diameter differentiation observed in plots H and I is significant. However, the $p_6$ measure of the deviation test suggests independent marking of tree diameters. The mark variogram confirms negative autocorrelation in plot H but not in plot I (Fig. 11). Diameter variance in plots H and I is similar to plot F.

4. Discussion

The main objective of this study has been to carry out a combined analysis of species and size diversity of tree vegetation. This analysis followed two strategies, a parallel study of species and size diversity both non-spatially and spatially explicit, which we extended by a non-spatial combined analysis using the functional diversity approach.

The saturation trend curves of Fig. 4 clearly show that there is a marked difference between size and species abundances in the forests studied. The shape of the curves suggests that species abundances alone obviously do not suffice to describe the pattern of the ten most abundant species and the inclusion of size gives a much better idea of the differences between species and size abundance patterns. In the same way there are marked differences between species and size diversity indices, i.e. the order of plot ranking differs.

The relationship between relative abundance measured by a size variable and relative abundance measured in number of individuals has reminded us of allometric relationships (see Pretzsch, 2009, p. 387ff.), for example those of tree stem diameters and total heights. Allometry simply means that the relative increment of one characteristic is proportional to the relative increment of another characteristic (Huxley, 1932; Teissier, 1934). In the case of relative species abundance we could describe the relationships in Fig. 4 as negative allometry, i.e. from the rarest to the commonest species the relative tree-count abundance increases less than the relative basal-area abundance. To some degree this has something to do with the well-known fact that larger trees are rarer (Gaston and Spicer, 2004, p. 10), i.e. with the relationship between size and species abundance. The abundance of more frequent species increases relative to rarer species more in size than in number of individuals, i.e. more frequent species have larger individuals. Using only species abundance indices would lead to an overestimation of the abundance of rare species and an underestimation of the abundance of common species.

To our knowledge this interesting relationship between size- and individual-based relative abundances has not been reported so far from multi-species forests. It is also worth noting that the Michaelis–Menten model as yet has only be applied to the relationship between sample size or area and cumulative number of species to produce species accumulation curves (see Magurran, 2004, p. 78ff.; Newton, 2007, p. 125f.; Plotkin et al., 2000; Wiegand et al., 2007). This study has clearly shown that the same saturation model is very suitable to depict the relationship between size- and individual-based abundances.

The concept of functional dispersion with its simultaneous use of species abundance and size information as abundance weighted size values is intriguing (Lavorel et al., 2008; Laliberté and Legendre, 2010). It even allows the use of multiple size characteristics at the same time. In our study, the results of the functional dispersion index FDIs had some similarities with those of the species-abundance only indices, however, there were also similarities with the results of...
Fig. 9. Mark mingling (top left) and mark differentiation (bottom left) functions using the Epanechnikov kernel and a bandwidth of $h = 2 \text{ m}$ along with the corresponding mark-weighted $L$ functions of plot H. The envelopes (grey areas) were constructed from the 2.5% and 97.5% quantiles of 1000 Monte Carlo simulations. The dashed lines denote mark independence as a reference.

Fig. 10. Mark mingling (top left) and mark differentiation (bottom left) functions using the Epanechnikov kernel and a bandwidth of $h = 2 \text{ m}$ along with the corresponding mark-weighted $L$ functions of plot I. The envelopes (grey areas) were constructed from the 2.5% and 97.5% quantiles of 1000 Monte Carlo simulations. The dashed lines denote mark independence as a reference.
the size-based indices and differences to both. The simultaneous computation of both diversity aspects clearly has a theoretical advantage over separate computations and future methodological research needs to elaborate more clearly the advantages and disadvantages of this concept.

The results of the non-spatial diversity measures stress the similarities in species and size diversity between these two very different sites. Spatially explicit diversity functions additionally produce evidence that there is an aggregation of conspecific trees and similar trees sizes at close proximity, although the mingling patterns were not significant in the Dongdapo Natural Reserve. This finding may be a reflection of the typical seed dispersal and development processes in the study sites although details of these processes are yet unknown for the studied woodland communities.

Fig. 11. Mark variograms (left) using the Epanechnikov kernel and a bandwidth of \( h = 2 \) m and the mark-weighted \( L_1 \) functions of the four plots. The envelopes (grey areas) were constructed from the 2.5% and 97.5% quantiles of 1000 Monte Carlo simulations and the dashed lines indicate the diameter variance, i.e. mark independence as a reference.
The tendency of species segregation that exists in all four research sites, is potentially also the reason for the high level of species diversity (Wright, 2002), as a pattern of intimately mixed species would inevitably lead to a reduction of species diversity due to competitive exclusion.

In the Qining Mountain Range there is some evidence of negative autocorrelation, i.e. trees with very different diameters occur at short inter-tree distances. This is a special characteristic of plots at Tartu funded by the DoRa programme of the European Union.

In addition we have learned that the mingling pattern in the Dongdapo Natural Reserve agrees with the hypothesis of independent marking, i.e. the species are arranged in a random fashion, which is not uncommon in forests with high species diversity (Jillian et al., 2008; Pomerening, 2002). This is contrasted by the findings from the Qining Mountain range where species mingling forms a definite pattern that deviates from randomness.

5. Conclusions

The study has demonstrated that biodiversity analyses should also include the size of individuals if larger differences in size are typical of these organisms. Variations in size are crucial for a full understanding of the diversity patterns. Parallel and even simultaneous investigations of species and size diversity therefore give much more complete information, particularly in forest ecosystems with complex forest structures.

If spatially explicit data are available this additional information should definitely be used for a refined analysis of diversity patterns. Forests with the same non-spatial measures can turn out to be quite different when spatial characteristics are applied. The methods available in spatial statistics including tests for different hypotheses allow a much more efficient use of ecological field data. Also for a holistic understanding it is clearly recommendable to use a number of different summary characteristics and concepts rather than relying on just one measure. The use of spatial diversity characteristics can make an important contribution to diversity analyses and we therefore propose to include them in similar studies.

Acknowledgements

The authors gratefully acknowledge the financial support of the national Sci-Tech Support Plan of China (2012BAD22B03). They wish to thank Pavel Grabarnik (Russian Academy of Sciences, Pushchino, Russia) and Mari Myllymäki (Aalto University, Finland) for their invaluable advice regarding the envelope and the deviation tests. Klaus von Gadow has kindly facilitated the cooperation and communication between the two authors. The second author also wishes to thank Andres Kiviste and Henn Korjus (both Estonian University of Life Sciences, Tartu, Estonia) for their hospitality, as the revision of this paper has been carried out during a sabbatical at Tartu funded by the DoRa programme of the European Union. Helpful comments of two anonymous reviewers and two guest editors are gratefully acknowledged.

References


