

treespap:an R package for spatial tree diversity analysis

F. Chianucci, N. Puletti, L. Cesaretti

2023-05-22



Abstract

Forest structure is a key element in understanding functionality and resilience of forest ecosystems. Analyses of forest structure and diversity have traditionally employed non-spatial measures, due to the simplicity of such approach. However, spatial structure can provide a deeper understanding of tree patterns, but widespread use of a spatially-explicit approach has been often limited by the complexity of theoretical equations and formulas, and the lack of freely distributed tools to calculate such spatial indices. To fill this gap we created the R package **treespap**, which allows to calculate some of the most diffuse stand-level spatial diversity indices. The document describes theory, basic definition and example application for using the **treespap** package.

Introduction

Forest structure is a key element in understanding functionality and resilience of forest ecosystems (Seidl et al., 2014) and a basic element of species diversity (Staudhammer and LeMay, 2001). An increasing complexity of stand structure is often linked to a higher number of species and, consequently, to greater ecological stability.

Analyses of forest structure and diversity have traditionally employed non-spatial measures, due to the simplicity of such approach. However, spatial structure can provide a deeper understanding of tree patterns (Grotti et al. 2019) and determines the properties of the ecosystem, including total biomass production, biodiversity, ecosystem functioning (Gadow et al., 2012; Castaño-Santamaría et al. 2021). However, use of a spatial approach in studying tree diversity is often hampered by the difficulty of implementing theoretical equations and formulas, and the lack of freely distributed tools to calculate such spatial indices.

According to Pommerening (2002), the majority of indices quantify spatial forest structure can be divided into three groups:

1. **stand-level attributes** based on neighborhood relations;
2. **distance-dependent measures** describing stand structure;
3. **continuous functions**.

While some existing R package already implemented some distance-dependent measures (2) and continuous functions (3) as for example in the **spatstat** family packages (Baddeley et al. 2015) and **spatdiv** package (see <https://ericmarcon.github.io/SpatDiv>), no available R packages allow to calculate stand-level parameters (1). Nonetheless, stand-level spatial indices are important to explore small-scale differences in diversity and also to link stand structure to ecosystem management, functioning and diversity (see example in Grotti et al. 2019).

The R package **treespac** allows to calculate some of the most diffuse stand-level spatial diversity indices, according to work by Pommerening (2002). The next section will describe the basic theory of most diffuse stand-level indices implemented in the package.

Spatial diversity indices

The package implements three of the most diffuse stand-level spatial indices (Pommerening 2002 , Pommerening 2020): *size differentiation*, *size dominance*, *species mingling*.

Their implementation requires mapped tree data, namely horizontal tree (x, y) position and an additional mark (*m*) attribute, e.g. species, tree height, tree diameter measured for all individual in a forest-inventoried plot.

Two additional indices were also calculated by the package:

- *Mean Directional Index* is an angle-based index (Corral-Rivas, 2006) representing the spatial arrangement of trees.
- *Stand Complexity Index* is a stand level measure of structural 3D complexity.

Formula of the indices is reported below.

Size differentiation

An early measure of spatial size inequality in plants was introduced by Gadow (1993) as the mean of the ratio of smaller and larger plant sizes *m* of the *k* nearest neighbors:

$$T_i = 1 - \frac{1}{k} \sum_{j=1}^k \frac{\min(m_i, m_j)}{\max(m_i, m_j)}$$

where *m* can be any quantifiable plant size measure, e.g. height or tree diameter among others. The value of *T_i* increases with increasing average size difference between neighboring trees, with *T_i* = 0 implying that neighboring trees have equal size. This index is based on the size-ratio construction principle.

2. Size dominance

(Aguirre et al. 2003) proposed an index using the size-comparison construction principle. This method turns the continuous size variables into a binary problem. The dominance index is defined as the mean fraction of trees among the *k* nearest neighbors of a given individual *j* that are smaller than *i*.

$$D_i = \frac{1}{k} \sum_{j=1}^k 1(m_i > m_j)$$

where m can be any quantifiable plant size measure, e.g. height or tree diameter among others. The indicator function returns 1, if the size mark of plant i exceeds that of neighboring plant j , otherwise 0. It can have k possible discrete outcomes.

3. Mingling Index

One very intuitive extension of taxonomic species diversity (either richness or abundance) is considering spatial mingling, namely how plants of the same (con-specific neighbors) or different (hetero-specific neighbors) species are arranged in space.

The mingling index (Aguirre et al. 2003) calculates the proportion of the k nearest neighbors that do not belong to the same species as the reference tree i :

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

For example, with four neighbors, the mingling attribute M_i can assume five values, ranging from 0 (all trees are of the same species) to 1 (all trees belong to different species).

4. Mean directional index

The mean directional index (Corral-Rivas, 2006) is defined as the sum of the unit vectors from the reference tree i to its k nearest neighbors and represents the spatial arrangement of tree:

$$MDI : \sqrt{\left(\sum_{j=1}^k \cos \alpha_{ji}\right)^2 + \left(\sum_{j=1}^k \sin \alpha_{ji}\right)^2}$$

5. Stand Complexity Index (SCI; Zenner and Hibbs, 2000)

The SCI index has been developed to incorporate the 3D structural complexity into spatial diversity analysis. It can be calculated from mapped tree data, as:

$$SCI : \frac{SCI_p}{AT_p}$$

Where SCI_p is the continuous faceted surface, calculated in the 3D-dimensional space, and AT_p is the projected areas of the triangles forming the faceted surface. For the calculation, the 3D-space consider both the horizontal (x, y) tree position and tree height (z), but other tree attribute can be considered for the third-dimension.

Installation

You can install the development version of **treespap** from [GitLab](#) using **devtools** (Wickham et al. 2021):

```
# install.packages("devtools")
devtools::install_gitlab("fchianucci/treespap")
```

Basic usage

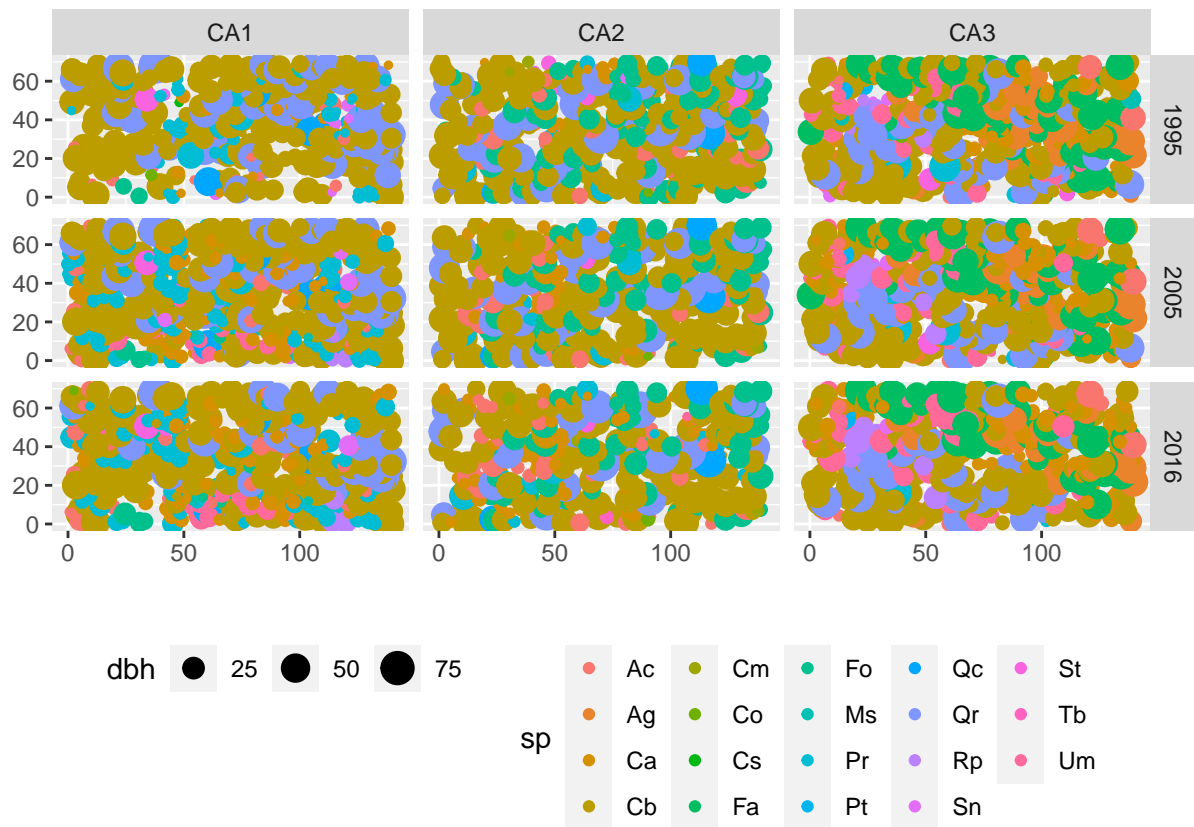
The package implements the function above by calculating the values for each reference i tree and a user-defined number of k neighbors (suggested values is $k \leq 4$).

For each index, plot-level estimates were then calculated by calculating their arithmetic mean.

In addition a NN1 edge correction method is also applied for all the indices (with the exception of SCI), following Pommerening and Stoyan (2006).

We will illustrate the calculation of all indices in **treespac** using an example dataset (**bf**) of a semi-natural oak-hornbeam forest sampled in Northern Italy. it provides replicated, repeated measurements at single-tree level conducted in three ~ 1-ha permanent plots (*Core Areas*, abbreviated as CA); all standing trees in these plots were mapped and inventoried in 1995, 2005, and 2016.

The dataset is described in Fardusi et al. 2018. The original dataset included some coppice stools, which have identical tree coordinates, and dead trees. For usage in **treespac**, data were elaborated to consider only living trees, while for coppice stools we calculated their cumulative basal area, and then we kept only an individual, whose diameter was calculated from inverting the cumulative basal area, and whose height was that of the tallest stool. For details see Puletti et al. (2023).



We can inspect the dataset (**bf**):

```
#> # A tibble: 6 x 7
#>   CA      yr sp      x      y    TH  dbh
#>   <chr> <int> <chr> <dbl> <dbl> <dbl> <dbl>
#> 1 CA1   1995 Cb     5.2   5.3   23   37
```

```
#> 2 CA1      1995 Cb      8.3   5      22      19
#> 3 CA1      1995 Ac      7.2  10.5     5       5
#> 4 CA1      1995 Cb       1   11.9    25      44
#> 5 CA1      1995 Ac      9.1  19.2     5       7
#> 6 CA1      1995 Fo      5.6  24     10      11
```

- *CA* is the plot identifier;
- *yr* is the sampling year;
- *sp* is the code representing the tree species;
- *x*, *y* are the horizontal coordinates from the bottom-left plot origin (0,0), expressed in m;
- *TH* is the tree top height, in m;
- *dbh* is the diameter at the breast height, in cm.

All the spatial functions implemented requires an input dataframe with:

- the tree positions (*.x*, *.y*);
- a quantitative *.mark* (e.g. diameter or height) or *.species* (for mingling index);
- information about plot size (*xmax*, *ymax* or *radius*) and *shape* (to apply edge corrections);
- the *k* number of neighboring trees to consider (*max.k*);
- a grouping column(s) *.groups*.

Below we provide examples with *k*=3 neighboring trees.

- **Diameter differentiation** (*DIFF*):

```
DIFF(bf,
     .x = x,
     .y = y,
     .mark = dbh,
     xmax = 70,
     ymax = 140,
     shape = 'square',
     max.k = 3,
     .groups=c('CA','yr'))
#> # A tibble: 9 x 3
#> # Groups:   CA, yr [9]
#>   CA      yr  DIFF
#>   <chr> <int> <dbl>
#> 1 CA1    1995 0.422
#> 2 CA2    1995 0.430
#> 3 CA3    1995 0.484
#> 4 CA1    2005 0.409
#> 5 CA2    2005 0.460
#> 6 CA3    2005 0.478
#> 7 CA1    2016 0.451
#> 8 CA2    2016 0.464
#> 9 CA3    2016 0.493
```

- Size dominance (*DDOM*):

```
DDOM(bf,
      .x = x,
      .y = y,
      .mark = dbh,
      xmax = 70,
      ymax = 140,
      shape = 'square',
      max.k = 3,
      .groups=c('CA','yr'))
#> # A tibble: 9 x 3
#> # Groups:   CA, yr [9]
#>   CA      yr DBH.DOM
#>   <chr> <int>   <dbl>
#> 1 CA1    1995   0.460
#> 2 CA2    1995   0.486
#> 3 CA3    1995   0.474
#> 4 CA1    2005   0.486
#> 5 CA2    2005   0.486
#> 6 CA3    2005   0.485
#> 7 CA1    2016   0.504
#> 8 CA2    2016   0.488
#> 9 CA3    2016   0.486
```

In the above example we used tree diameter, but any other quantitative `.mark` (e.g. tree height) can be used.

- Mingling index (*MING*):

```
MING(bf,
      .x = x,
      .y = y,
      .species = sp,
      xmax = 70,
      ymax = 140,
      shape = 'square',
      max.k = 3,
      .groups=c('CA','yr'))
#> # A tibble: 9 x 3
#> # Groups:   CA, yr [9]
#>   CA      yr MINGLING
#>   <chr> <int>   <dbl>
#> 1 CA1    1995   0.557
#> 2 CA2    1995   0.627
#> 3 CA3    1995   0.693
#> 4 CA1    2005   0.595
#> 5 CA2    2005   0.609
#> 6 CA3    2005   0.683
#> 7 CA1    2016   0.619
#> 8 CA2    2016   0.652
#> 9 CA3    2016   0.690
```

- Mean directional index (*MDIR*):

```
MDIR(bf,
      .x = x,
      .y = y,
      xmax = 70,
      ymax = 140,
      shape = 'square',
      max.k = 3,
      .groups=c('CA', 'yr'))
#> # A tibble: 9 x 3
#> # Groups:   CA, yr [9]
#>   CA      yr MeanDirI
#>   <chr> <int>    <dbl>
#> 1 CA1    1995    1.51
#> 2 CA2    1995    1.63
#> 3 CA3    1995    1.55
#> 4 CA1    2005    1.53
#> 5 CA2    2005    1.47
#> 6 CA3    2005    1.56
#> 7 CA1    2016    1.49
#> 8 CA2    2016    1.55
#> 9 CA3    2016    1.50
```

- Stand complexity index (*SCI*):

```
SCI(bf,
     .x = x,
     .y = y,
     .mark = TH,
     .groups=c('CA', 'yr'))
#> # A tibble: 9 x 5
#> # Groups:   CA, yr [9]
#>   CA      yr mean_SCI sd_SCI N_tri
#>   <chr> <int>    <dbl> <dbl> <dbl>
#> 1 CA1    1995    4.77    9.42   713
#> 2 CA2    1995    4.73    8.26   866
#> 3 CA3    1995    6.77   10.3   1183
#> 4 CA1    2005    4.79    9.36  1100
#> 5 CA2    2005    4.89    8.17   746
#> 6 CA3    2005    6.41    9.48   977
#> 7 CA1    2016    4.90   10.1   878
#> 8 CA2    2016    6.46   26.3   692
#> 9 CA3    2016    7.12   13.1   811
```

Note: we did not implemented edge correction in this function as it requires different approach compared to NN1 (see Pöldveer et al. 2021).

Funding The work was supported by the Agritech National Research Center (project CN00000022) and National Biodiversity Future Center (project CN00000033), which received funding from National Recovery and Resilience Plan (NRRP) – MISSION 4 COMPONENT 2, INVESTMENT 1.4.

References

- Aguirre, O., Hui, G., von Gadow, K. and Jiménez, J., 2003. An analysis of spatial forest structure using neighbourhood-based variables. *Forest ecology and management*, 183(1-3), pp.137-145. doi: [https://doi.org/10.1016/S0378-1127\(03\)00102-6](https://doi.org/10.1016/S0378-1127(03)00102-6)
- Baddeley, A., Rubak, E. and Turner, R., 2015. *Spatial point patterns: methodology and applications with R*. CRC press.
- Castano-Santamaria, J., López-Sánchez, C.A., Obeso, J.R. and Barrio-Anta, M., 2021. Structure, environmental patterns and impact of expected climate change in natural beech-dominated forests in the Cantabrian Range (NW Spain). *Forest Ecology and Management*, 497, p.119512. doi: <https://doi.org/10.1016/j.foreco.2021.119512>
- Corral-Rivas, J.J., 2006. Models of tree growth and spatial structure for multi-species, uneven-aged forests in Durango (Mexico). PhD thesis, University of Göttingen.
- Fardusi, M.J., Castaldi, C., Chianucci, F., Corona, P., Mason, F., Minari, E. and Puletti, N., 2018. A spatio-temporal dataset of forest mensuration for the analysis of tree species structure and diversity in semi-natural mixed floodplain forests. *Annals of Forest Science*, 75(1), pp.1-5. doi: <https://doi.org/10.1007/s13595-018-0688-8>
- Gadow, K.V., Zhang, C.Y., Wehenkel, C., Pommerening, A., Corral-Rivas, J., Korol, M., Myklush, S., Hui, G.Y., Kiviste, A. and Zhao, X.H., 2012. Forest structure and diversity. *Continuous cover forestry*, pp.29-83.
- Grotti, M., Chianucci, F., Puletti, N., Fardusi, M.J., Castaldi, C. and Corona, P., 2019. Spatio-temporal variability in structure and diversity in a semi-natural mixed oak-hornbeam floodplain forest. *Ecological Indicators*, 104, pp.576-587. doi: <https://doi.org/10.1016/j.ecolind.2019.04.014>
- Poldveer, E., Potapov, A., Korjus, H., Kiviste, A., Stanturf, J.A., Arumäe, T., Kangur, A. and Laarmann, D., 2021. The structural complexity index SCI is useful for quantifying structural diversity of Estonian hemiboreal forests. *Forest Ecology and Management*, 490, p.119093. doi: <https://doi.org/10.1016/j.foreco.2021.119093>
- Pommerening, A., 2002. Approaches to quantifying forest structures. *Forestry: An International Journal of Forest Research*, 75(3), pp.305-324. doi:<https://doi.org/10.1093/forestry/75.3.305>
- Pommerening, A., Szmyt, J. and Zhang, G., 2020. A new nearest-neighbour index for monitoring spatial size diversity: The hyperbolic tangent index. *Ecological modelling*, 435, p.109232. doi: <https://doi.org/10.1016/j.ecolmodel.2020.109232>
- Pommerening, A. and Stoyan, D., 2006. Edge-correction needs in estimating indices of spatial forest structure. *Canadian Journal of Forest Research*, 36(7), pp.1723-1739. doi: <https://doi.org/10.1139/x06-060>
- Puletti, N., Grotti, M., Cesaretti, L., Chianucci F., 2023. How to consider tree positions in multi-stemmed trees. Rpubs, https://rpubs.com/npuletti/tree_positions_in_multi_stemmed_trees
- Staudhammer, C.L. and LeMay, V.M., 2001. Introduction and evaluation of possible indices of stand structural diversity. *Canadian journal of forest research*, 31(7), pp.1105-1115. doi: <https://doi.org/10.1139/x01-033>
- Zenner, E.K. and Hibbs, D.E., 2000. A new method for modeling the heterogeneity of forest structure. *Forest ecology and management*, 129(1-3), pp.75-87. doi: [https://doi.org/10.1016/S0378-1127\(99\)00140-1](https://doi.org/10.1016/S0378-1127(99)00140-1)