

Appendix

Measuring local biodiversity with Diversity Measures at Distance

Eric Marcon

Florence Puech

November 21, 2025

Abstract

Code to reproduce the figures of the main text.

Contents

1	Descriptive statistics on the distribution of trees in the <i>Parc des Buttes Chaumont</i>	2
1.1	Preparation of the data	2
1.2	Spatial distribution of the trees	3
2	Measuring diversity in the <i>Parc des Buttes Chaumont</i>	4
2.1	At the park level	4
2.1.1	Evaluation of species' diversity of trees	4
2.1.2	Evaluation of diversity profile	5
2.2	At 5 meters and 25 meters	6

The analysis is made on trees located in the *Parc des Buttes Chaumont* in Paris (France). A local Google map of the park is available at the following address: <https://maps.app.goo.gl/Z2n1CWLB7YjHUdYY9>

All empirical results are based on data freely available on Paris open data (<https://opendata.paris.fr>). In the article, information on trees is extracted from the database “Arbres, Direction des Espaces Verts et de l’Environnement - Ville de Paris” (under license ODbL) in February 2025.

1 Descriptive statistics on the distribution of trees in the *Parc des Buttes Chaumont*

1.1 Preparation of the data

In this study, we select the following characteristics on trees:

- the exact location (coordinates are then converted in Lambert93 projection),
- the species (“*espece*” in French),
- the genus (“*genre*” in French),
- the circumference of the trunk, used to calculate the basal area of trees.

```
library("tidyverse")
library("sf")
# Get the data
if (file.exists("data/BC_Lambert.RData")) {
  # Dataset
  load("data/BC_Lambert.RData")
} else {
  # Build the dataset from scratch
  if (!file.exists("data/les-arbres.geojson")) {
    # Download the original data file if it does not exist (100 MB)
    download.file(
      "https://opendata.paris.fr/explore/dataset/les-arbres/download/?format=geojson",
      destfile = "data/les-arbres.geojson"
    )
  }
  # Project the dataset into Lambert CRS
  st_read("data/les-arbres.geojson") %>%
  # Keep Buttes-Chaumont trees only
  filter(str_detect(adresse, "PARC DES BUTTES CHAUMONT")) %>%
  # Lambert (standard in France) CRS
  st_transform(crs = 2154) -> BC_Lambert
}
# Clean-up the data
BC_Lambert %>%
  # Remove useless columns
  select(
    -typeemplacement, -domanialite, -arrondissement,
    -complementadresse, -adresse,
    -idemplacement, -varieteoucultivar,
    -statedeveloppement, -remarquable, -geo_point_2d
  ) %>%
  # Remove NAs
  filter(
    !is.na(genre)
  ) %>%
  mutate(
    # Binomial botanic species name
    genre_espece = paste(
      genre,
      # Undetermined species are denoted sp. (correct) or n. sp. or NA
      ifelse(is.na(espece) | espece == "n. sp.", "sp.", espece)
    )
  )
```

```

),
# radius
diameter = circumferencecm / pi,
# Basal area
G = circumferencecm^2 / 4 / pi
) %>%
# Factorize character columns
mutate(across(where(is_character), as_factor)) ->
BC_trees

```

1.2 Spatial distribution of the trees

The spatial distribution of trees in the *Parc des Buttes Chaumont* is given hereinafter.

To improve readability:

- only the most abundant genera are presented (in decreasing order),
- the visualization of individual diameters are magnified.

```

library("dbmss")
BC_trees %>%
# Point type and point weight
select(genre, diameter) %>%
rename(PointType = genre, PointWeight = diameter) %>%
# Get the coordinates
bind_cols(st_coordinates(.)) %>%
# dataframe needed by as.wmppp()
as_tibble() %>%
# Make a point pattern
as.wmppp(
  unitname = c("meter", "meters")
) -> BC_wmppp_genera
# Limits of the park
library("divent")
Window(BC_wmppp_genera) <- alphahull(BC_wmppp_genera, alpha = 50)
# Map
library("ggspatial")
autoplot(BC_wmppp_genera) +
# Show the points at scale: max radius, in meters,
# twice bigger for readability
scale_radius(range = c(0, max(BC_trees$diameter / 2 / 100 * 2))) +
labs(color = "Genus", size = "Diameter") +
theme(
  panel.grid.major = element_line(color = "gray50", linetype = "dotted"),
) +
# Scale position: bottom right
annotation_scale(location = "br") +
# North arrow
annotation_north_arrow(
  pad_x = unit(9, "cm"),
  pad_y = unit(1, "cm"),
  style = north_arrow_fancy_orienteering()
)

```

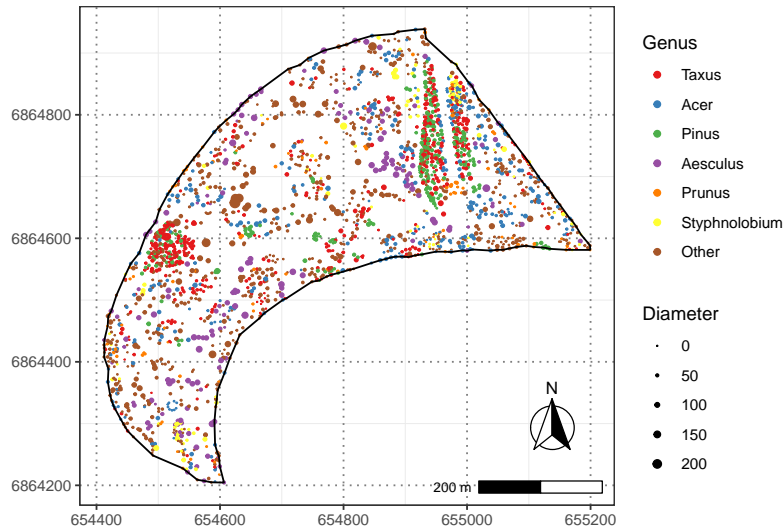


Figure 1: Spatial distribution of trees

2 Measuring diversity in the *Parc des Buttes Chaumont*

We now focus on the diversity of species of trees.

We use the R package *divent* (Marcon and Puech, 2025) for estimating the diversity at the park-level or at the local scale with the DMD.

2.1 At the park level

2.1.1 Evaluation of species' diversity of trees

```
# Build the species distribution object
BC_trees %>%
  pull(genre_espece) %>%
  as_species_distribution() ->
  BC_sad
# Compute richness at the observed level
BC_richness <- div_richness(
  BC_sad,
  estimator = "naive",
  as_numeric = TRUE
)
# Compute Shannon's entropy at the observed level
BC_Shannon <- ent_shannon(
  BC_sad,
  estimator = "naive",
  as_numeric = TRUE
)
# Diversity of order 1
BC_D1 <- exp(BC_Shannon)
# Compute Simpson's entropy at the observed level
BC_Simpson <- ent_simpson(
  BC_sad,
  estimator = "naive",

```

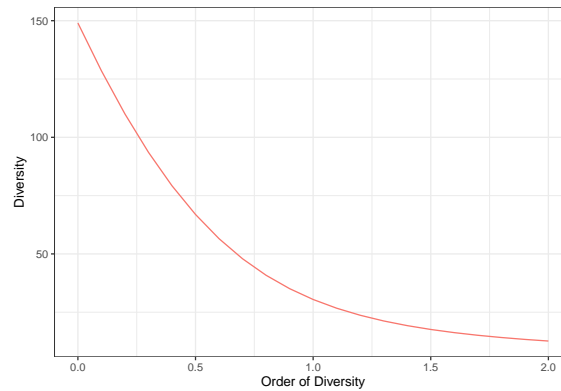


Figure 2: Profile of the Hill numbers

```
as_numeric = TRUE
)
# Diversity of order 2
BC_D2 <- 1 / (1 - BC_Simpson)
```

We find:

- Richness is equal to 149.
- The Shannon index is equal to 3.42,
- The Simpson index is equal to 0.92.

2.1.2 Evaluation of diversity profile

We compute the profile of the Hill numbers (Hill, 1973) on the entire *Parc des Buttes Chaumont*.

As we note in the article, remarkable values of q are:

- $q=0$ for richness,
- $q=1$ for Shannon's index,
- $q=2$ for Simpson's index.

```
BC_sad %>%
# Build the diversity profile
profile_hill(estimator = "naive") %>%
# Plot it
autoplot() +
# No legend
theme(legend.position = "none")
```

2.2 At 5 meters and 25 meters

The local tree diversity (Hill numbers) in the *Parc des Buttes Chaumont* is given hereinafter.

We only choose two distances to evaluate the local biodiversity of trees species:

- 5 meters: local maps are given in the first column,
- 25 meters: local maps are given in the second column.

For both distances, we provide three orders of diversity:

- $q=0$ for richness on the first row,
- $q=1$ for the Shannon index on the second row,
- $q=1$ for the Simpson index on the third row.

The estimated Hill numbers are given on the map, whatever the distance and the order of diversity. Contour lines facilitate the lecture because they identify similar estimation values of Hill numbers.

```
# Make a wmpmp with species, all trees have weight 1
BC_trees %>%
  # Point type
  select(genre_espece) %>%
  rename(PointType = genre_espece) %>%
  # Point weight
  mutate(PointWeight = 1) %>%
  # Get the coordinates
  bind_cols(st_coordinates(.)) %>%
  # dataframe needed by as.wmpmp()
  as_tibble() %>%
  # Make a point pattern
  as.wmpmp(
    unitname = c("meter", "meters")
  ) -> BC_wmpmp_species
# Limits of the park already computed
Window(BC_wmpmp_species) <- Window(BC_wmpmp_genera)
# Spatial accumulation of diversity
accum_r <- accum_sp_hill(
  BC_wmpmp_species,
  # Richness, Shannon and Simpson
  orders = 0:2,
  # 5 and 25 meters. 0 is mandatory.
  r = c(0, 5, 25),
  # Save individual neighborhoods
  individual = TRUE
)
# 3 rows, 2 columns for the figure
op <- par(
  mfrow = c(3, 2),
  # Outer right margin must not be 0 to show the legends
  oma = c(0, 0, 0, 1)
)
# Plots
for (q in 0:2) {
  for (r in c(5, 25)) {
    plot_map(
      accum_r,
      q = q,
      neighborhood = r,
      dim_x = 512,
      dim_y = 512
    )
    # Show order and distance on the plot
    text(
      x = 654520,
      y = 6864900,
      paste0("q=", q, ", r=", r)
    )
  }
}
```

```
)  
}  
}
```

```
# Restore parameters  
par(op)
```

References

- Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology* 54(2), 427–432.
- Marcon, E. and F. Puech (2025). *divent: Entropy Partitioning to Measure Diversity*. R package version 0.5-3, <https://github.com/EricMarcon/divent>.

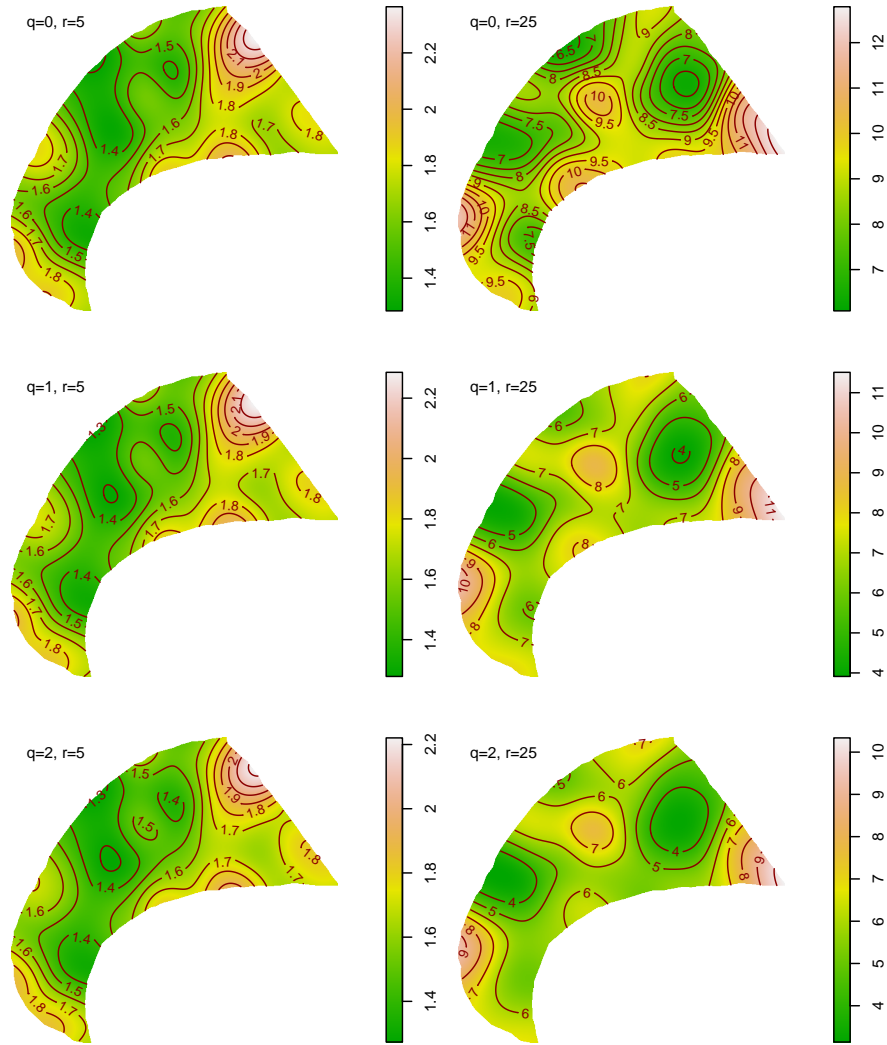


Figure 3: Local tree diversity in the Parc des Buttes Chaumont, estimated by the Hill numbers