

Tinkering with measures of spatial species diversity

John L. Godlee

24 May, 2020

Preamble

```
# Function for robust line wrapping of output
##' Normally this chunk would be `echoFALSE`,
##' but for demonstration I've left it in.
library(knitr)
hook_output = knitr_hooks$get('output')
knitr_hooks$set(output = function(x, options) {
  # this hook is used only when the linewidth option is not NULL
  if (!is.null(n <- options$linewidth)) {
    x = knitr:::split_lines(x)
    # any lines wider than n should be wrapped
    if (any(nchar(x) > n)) x = strwrap(x, width = n)
    x = paste(x, collapse = '\n')
  }
  hook_output(x, options)
})

# Knitr global options
knitr::opts_chunk$set(echo=TRUE, fig.align="center", linewidth=60)
knitr::opts_chunk$set(tidy.opts=list(width.cutoff=60),tidy=TRUE)
knitr::opts_chunk$set(fig.pos = "H", out.extra = "")

# Packages
library(ggplot2)
library(gridExtra)
library(sf)
library(mngeo)
library(dplyr)
library(rgdal)
library(tidyr)
library(vegan)
library(ggplot2)
library(contoureR)
library(LearnGeom)

# Set seed so random values are maintained between sessions
set.seed("57")
```

Introduction

The theory of niche complementarity in the context of biodiversity-ecosystem function research presumes that ecosystems with a greater species richness will exhibit higher levels of ecosystem function (e.g. greater biomass), because different species inhabit different niches in the ecosystem, allowing a greater portion of the ecosystem-level niche space to be filled and more complete utilisation of resources (e.g. sunlight) (Janzen 1970; Tilman, Isbell, and Cowles 2014). A higher number of species decreases the likelihood of conspecifics competing with one another at the local scale.

To extend this theory, I think, you can also assume that if two species occupy different niche space

to allow niche complementarity, two individuals from those species will compete with each other less than two individuals from the same species, which inhabit the same niche space. It could therefore be assumed that within an ecosystem with a greater species diversity, all else being equal, there will be less interspecific competition than an ecosystem with few species.

Competition between individuals however, occurs over small spatial scales. An adult tree, for example, will compete for light with the individual adjacent to it, but will not compete for light with a tree 50 m away, as their crowns are unlikely to overlap. Imagine a model of two plots in a forestry plantation, each containing five tree species with 20 individuals each (Figure @ref(fig:plots)). In Plot A, individuals of the five species are distributed randomly within the plot, while in Plot B the five species are distributed in monospecific bands.

```
# Create fake data

## Plot A
trees_a <- data.frame(sp = rep(c("A", "B", "C", "D", "E"), times = 20),
  x = round(runif(100, min = 0, max = 100), 1), y = round(runif(100,
    min = 0, max = 100), 1))

## Reference tree Plot A
ref_a <- trees_a[45, ]

## Plot B
trees_b <- data.frame(sp = rep(c("A", "B", "C", "D", "E"), each = 20),
  x = rep(seq(from = 0, to = 99, by = 10), times = 10) + 5,
  y = rep(seq(from = 0, to = 99, by = 10), each = 10) + 5)

## Reference tree Plot B
ref_b <- trees_b[45, ]

# Create plot diagrams
plot_a <- ggplot() + geom_point(data = ref_a, aes(x = x, y = y),
  colour = "black", size = 5) + coord_equal() + geom_point(data = trees_a,
  aes(x = x, y = y, fill = sp), colour = "black", shape = 21) +
  ggtitle("Plot A") + lims(x = c(0, 100), y = c(0, 100)) +
  theme(legend.position = "none")

plot_b <- ggplot() + geom_point(data = ref_b, aes(x = x, y = y),
  colour = "black", size = 5) + geom_point(data = trees_b,
  aes(x = x, y = y, fill = sp), colour = "black", shape = 21) +
  coord_equal() + ggtitle("Plot B") + lims(x = c(0, 100), y = c(0,
  100)) + theme(legend.position = "none")

grid.arrange(plot_a, plot_b, ncol = 2)
```

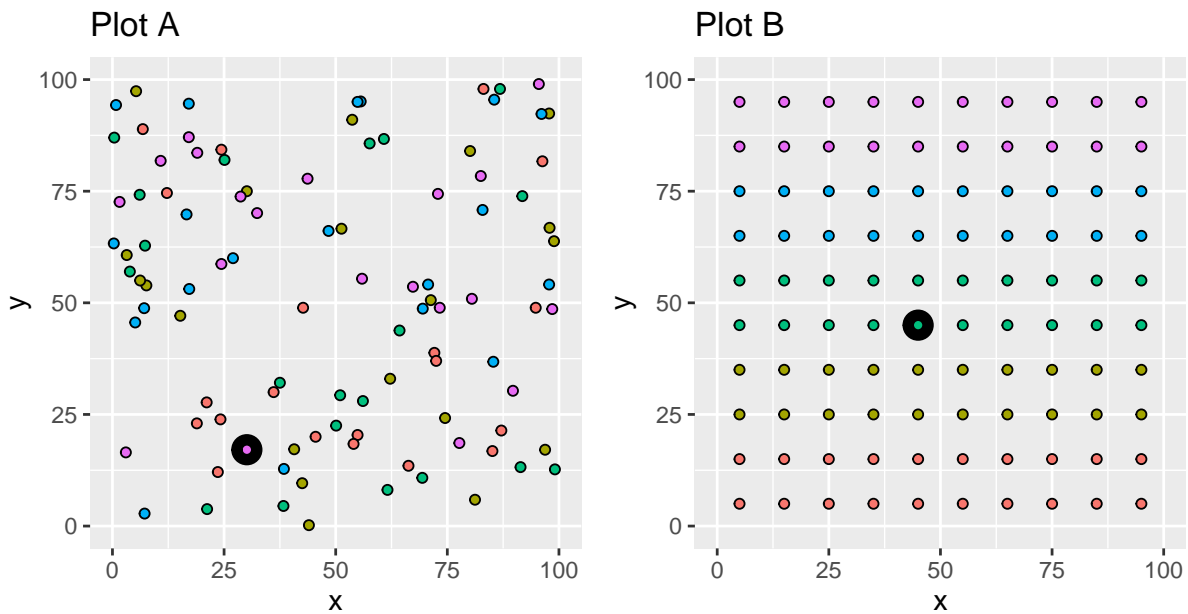


Figure 1: Schematic diagrams of trees in plots.

Consider the reference trees marked in each plot as large black circles. It is likely that in Plot A, the reference tree will have fewer neighbours of the same species than Plot B, but let's check:

```
# Convert to sf objects
trees_a_sf <- st_as_sf(trees_a, coords = c("x", "y"))
trees_b_sf <- st_as_sf(trees_b, coords = c("x", "y"))
ref_a_sf <- st_as_sf(ref_a, coords = c("x", "y"))
ref_b_sf <- st_as_sf(ref_b, coords = c("x", "y"))

# Calculate 10 nearest neighbours to reference
ref_a_nn <- unlist(st_nn(ref_a_sf, trees_a_sf, k = 11, progress = FALSE))
ref_b_nn <- unlist(st_nn(ref_b_sf, trees_b_sf, k = 11, progress = FALSE))

# How many neighbours are conspecific with reference tree?
conspec_a <- sum(trees_a[ref_a_nn[-1], ]$sp == ref_a$sp)
conspec_b <- sum(trees_b[ref_b_nn[-1], ]$sp == ref_b$sp)

# Create plot diagrams
plot_a_nn <- ggplot() + geom_point(data = ref_a, aes(x = x, y = y),
  colour = "black", size = 5) + geom_point(data = trees_a,
  aes(x = x, y = y, fill = sp), shape = 21, colour = "black") +
  geom_point(data = trees_a[ref_a_nn[-1], ], aes(x = x, y = y,
  fill = sp), colour = "black", shape = 22, size = 3) +
  coord_equal() + ggtitle("Plot A") + lims(x = c(0, 100), y = c(0,
  100)) + theme(legend.position = "none")

plot_b_nn <- ggplot() + geom_point(data = ref_b, aes(x = x, y = y),
  colour = "black", size = 5) + geom_point(data = trees_b,
  aes(x = x, y = y, fill = sp), shape = 21, colour = "black") +
  geom_point(data = trees_b[ref_b_nn[-1], ], aes(x = x, y = y,
  fill = sp), colour = "black", shape = 22, size = 3) +
  coord_equal() + ggtitle("Plot B") + lims(x = c(0, 100), y = c(0,
  100)) + theme(legend.position = "none")
```

```
grid.arrange(plot_a_nn, plot_b_nn, ncol = 2)
```

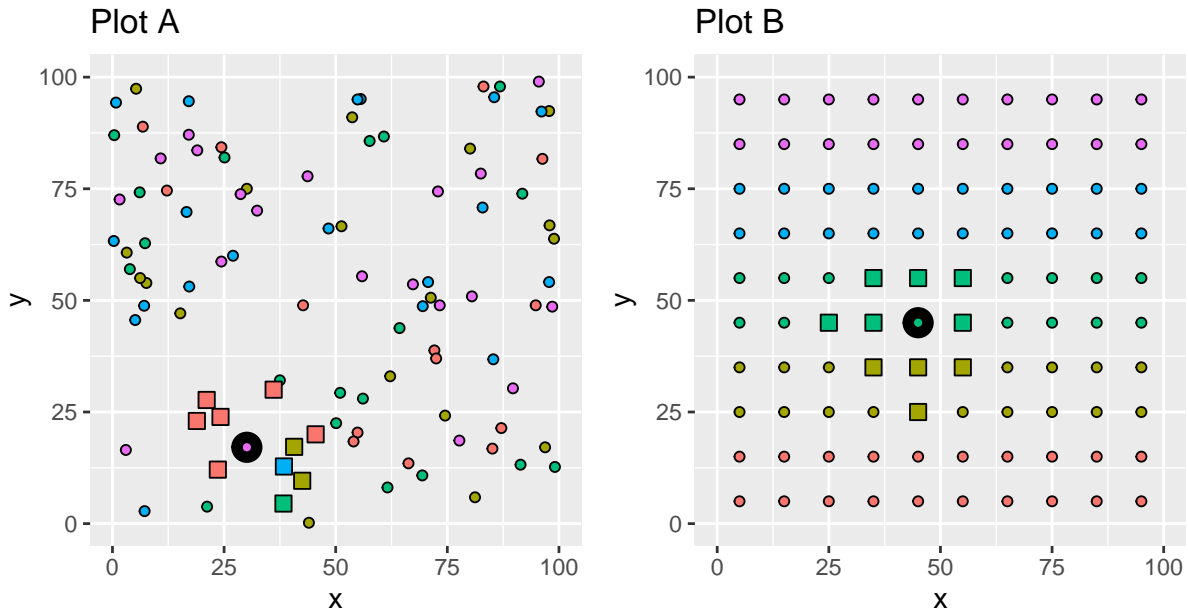


Figure 2: Schematic diagrams showing the 10 nearest neighbours of the reference tree as squares shaded according to species

In Plot A, of the ten nearest neighbours to the reference tree, 0 are the same species as the reference tree. In Plot B however, 6 are conspecifics (Figure @ref(fig:neighbours)). Following the idea that competition is higher between conspecifics due to them occupying the same niche space, ecosystem function may be lower in Plot B than Plot A, despite the two plots having the same species richness.

This conjecture then raises the question of whether species richness alone is enough to predict ecosystem function. Would a spatially explicit estimate of species diversity better represent the true diversity experienced by individuals within a particular ecosystem?

Spatial species distribution

Wiegand et al. (2007) suggest measuring the number of species in the “average neighbourhood” of a chosen focal species as a measure of the “local dominance” of that species. They report $1 - M$, where M is the local dominance of the species, to give a diversity index rather than a dominance index. This measurement is known as the “spatial mingling index”. The mean of spatial mingling index values for each species in a plot can be used to compare the plot-level spatial mingling of plots.

Hui et al. (2011) present a straightforward equation for the spatial mingling index per individual M_i , which they say was first described by Gadow (1993):

$$M_i = \frac{1}{k} \sum_{j=1}^k v_{ij}$$

Where $v_{ij} = 1$ if the j th neighbour is not the same species as the i th reference tree, k is the number of nearest neighbours considered, most commonly $k = 4$.

In the basic spatial mingling index however, the number of different tree species in the structural unit is not taken into account. One can therefore multiply M_i by $\frac{S_i}{n_{max}}$ where S_i is the number of tree species in the neighbourhood of the reference tree and n_{max} is the maximum number of species in the structural unit (plot). This revised index is often reported as MS_i . Rare species will produce higher values of MS_i .

The plot-level value of \overline{MS} can be calculated as:

$$\overline{MS} = \frac{1}{(k+1)N} \sum_{i=1}^k N(M_i S_i)$$

Where N is the total number of reference trees in the community and k is the number of neighbours considered per reference tree.

The spatial status of the average species is given by the same formula, except N is replaced by N_{sp} , the number of tree species in the plot.

Calculate spatial mingling in R

Import some data to play with:

```
trees <- read.csv("data/trees.csv", stringsAsFactors = FALSE)

str(trees)
```

```
## 'data.frame': 22041 obs. of 5 variables:
## $ plotcode: chr "TKW-001" "TKW-001" "TKW-001" "TKW-001"
...
## $ gen_sp : chr "Margaritaria indet" "Diplorhynchus
condylocarpon" "Combretum apiculatum" "Combretum
apiculatum" ...
## $ dbh_cm : num 28.6 18.1 45 21.8 86.2 ...
## $ x : num 1 7.1 44.8 47.9 46.9 ...
## $ y : num 0.8 30.6 12.1 13.2 29.9 ...
```

This dataset consists of 52 plots one hectare plots, each with the location of all living trees >5 cm DBH noted on a cartesian x,y grid system. Each tree also has the main stem DBH and species.

First find the four nearest neighbours of each tree. Do this by converting each plot to a `SpatialPointsDataFrame` object, then to an `sf` object, then using `ngeo::st_nn()` to calculate nearest neighbours:

This creates a nested list, structured in the following way, with plot lists (e.g. ABG-001) which contain two lists each, the first containing the nearest neighbour row indexes of each tree (`nn`) and the second containing the distances of each nearest neighbour per tree (`dist`):

```
nn_list
  ABG-001 - list
  | nn - list
  | | [[1]] - integer [1:5]
  | | [[2]] - integer [1:5]
  | | [[3]] - integer [1:5]
  | | [[4]] - integer [1:5]
  | | [[5]] - integer [1:5]
  | | ...
  | | [[541]] - integer [1:5]
  | dist - list
  | [[1]] - numeric [1:5]
  | [[2]] - numeric [1:5]
  | [[3]] - numeric [1:5]
  | [[4]] - numeric [1:5]
  | [[5]] - numeric [1:5]
  | ...
  | [[541]] - numeric [1:5]
  ABG-002 - list
  | nn - list
```

```
| | ...
| dist - list
| ...
TKW-011 - list
...
```

Then calculate the spatial mingling index. This code is incredibly messy, basically just a bunch of `sapply()` statements to drill down into the nested list and check whether the species of the reference tree is the same as the species of each neighbour tree:

```
mi_vec <- sapply(seq_along(nn_list), function(x) {
  # For each plot (1:52) For each tree For each neighbour
  1/4 * mean(sapply(seq_along(nn_list[[x]][[1]]), function(y) {
    sum(sapply(seq_along(nn_list[[x]][[1]][[y]][-1]), function(z) {
      sf_list[[x]]$gen_sp[nn_list[[x]][[1]][[y]][[1]]] !=
      sf_list[[x]]$gen_sp[nn_list[[x]][[1]][[y]][[z]]]
      # which neighbours are not the same species
    })
  })
})
```

Finally, this vector of plot-level spatial mingling index values can be put into a dataframe with their respective plotnames, ready for further analysis:

```
div_df <- data.frame(plotcode = names(nn_list), mi = mi_vec)
```

Compare the spatial mingling index with a non-spatially explicit Shannon diversity index to see how the two are related:

```
# Define a function to create abundance matrices
tree_ab <- function(data, id, sp) {
  id <- enquo(id)
  sp <- enquo(sp)
  tree_ab_mat <- data %>% filter(!is.na(!sp)) %>% group_by(!id,
    !!sp, .drop = FALSE) %>% tally() %>% spread(as_label(sp),
    n, fill = 0) %>% ungroup() %>% rename(plotcode = !!names(.[1]))
  return(tree_ab_mat)
}

# Create abundance matrices
ab_mat <- tree_ab(trees, id = plotcode, sp = gen_sp)

# Use vegan::diversity() to calculate Shannon index
div_df$shannon <- diversity(ab_mat[, -1])

# Comparison of shannon with mi
ggplot(div_df, aes(x = shannon, y = mi)) + geom_point() + geom_smooth(method = "lm",
  formula = "y ~ x")
```

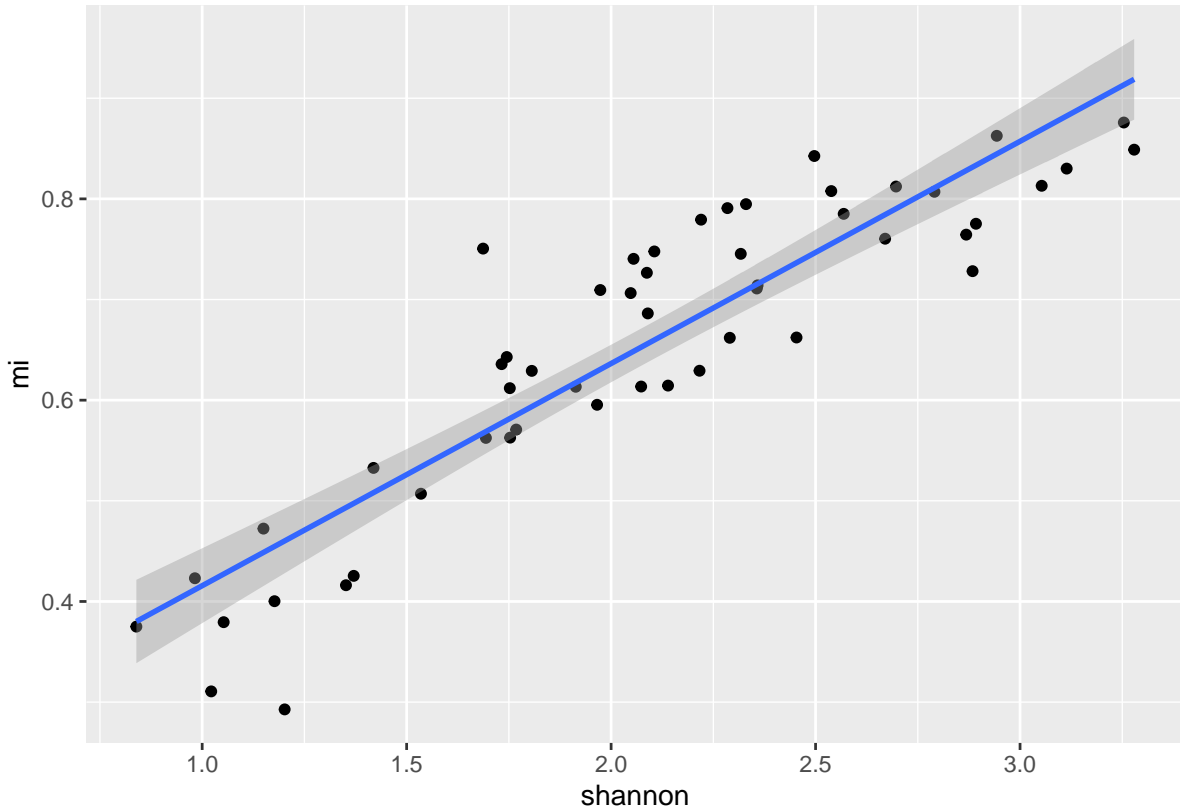


Figure 3: Comparison of Spatial Mingling Index and Shannon index

As the Shannon index increases so does spatial mingling, in a linear fashion, because with more species it's more likely that a neighbour will be a different species, as expected.

Spatial regularity of trees

von Gadow and Hui (2001) introduces a new index which relates to the spatial regularity of trees in a forest ecosystem. If individuals are distributed evenly throughout the plot, there will be less competition among them than if individuals are clustered together. The name of this spatial regularity index is the “Winkelmass”.

The tree-level Winkelmass varies between zero and one. Where zero indicates that trees in the vicinity of the reference tree are positioned in an entirely regular manner. The Winkelmass is calculated as the fraction of the angles α_j which are smaller than the “standard angle” α_0 . α_j angles are the angles between the immediate neighbours of the n neighbour trees of the reference tree. An immediate neighbour is the next tree following a clockwise direction around the reference tree. The angle measurement is always measured as the smallest angle, so if the angle between two immediate neighbours is $>180^\circ$, the angle is taken in the opposite direction, therefore α_j is always $<180^\circ$.

$$W_i = \frac{1}{k} \sum_{j=1}^4 v_j$$

Where $v_j = 1$ if α_j is greater than α_0 . In the case of a 4-tree sample ($k=4$) the standard angle α_0 would be $360/4=90^\circ$.

von Gadow (1998) describes the “Uniform Angle Index”, which as far as I can tell is measured and calculated identically to the Winkelmass.

As with the Spatial Mingling Index, the Winkelmass can be calculated per plot, and per species within a plot, depending on the application.

Calculate the Winkelmass in R

Next we can estimate the Winkelmass, using a similar approach, with a nested list describing neighbour relationships between trees.

```

wm_list <- lapply(seq_along(nn_list), function(x) {
  # For each plot (1:52) For each tree create a matrix
  lapply(seq_along(nn_list[[x]][[1]]), function(y) {
    mat <- cbind(sf_list[[x]][nn_list[[x]][[1]][[y]], ]$x,
                 sf_list[[x]][nn_list[[x]][[1]][[y]], ]$y)
    centre <- mat[1, ] # Centre location
    neighb <- mat[-1, ] # Neighbours
    neighb_order <- neighb[orderPoints(neighb[, 1], neighb[,
      2], xm = centre[1], xy = centre[2], clockwise = TRUE),
      ] # Neighbours ordered clockwise
    sapply(seq_along(neighb_order[, 1]), function(z) {
      # For each neighbour, find angle to next neighbour
      unname(Angle(neighb_order[z, ], centre, neighb_order[ifelse(z +
        1 == 5, 1, z + 1), ]))
    })
  })
})

div_df$wm <- sum_list <- sapply(wm_list, function(x) {
  mean(sapply(x, function(y) {
    sum(ifelse(y < 90, TRUE, FALSE), na.rm = TRUE)
  })), na.rm = TRUE)
})

```

Then compare the Winkelmass with both the Shannon Index and the Spatial Mingling Index:

```

div_df_g <- gather(div_df, key = "index", value = "value", -wm,
  -plotcode)

ggplot(div_df_g, aes(x = wm, y = value)) + geom_point() + geom_smooth(method = "lm",
  formula = "y ~ x") + facet_wrap(~index, scales = "free")

```

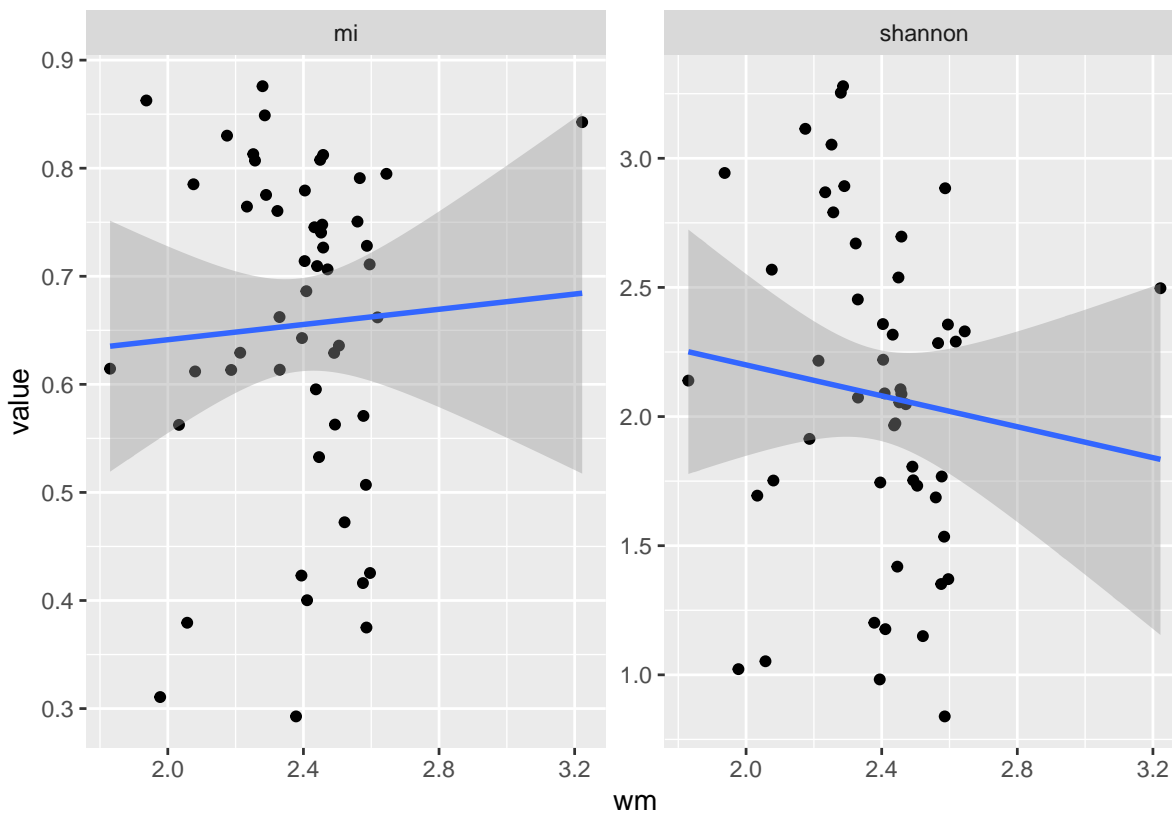



Figure 4: Comparison of the Winkelmass with Shannon index and Spatial Mingling Index

The Winkelmass appears to be very uncorrelated with both the Shannon index and Spatial Mingling Index. Possibly Winkelmass would be a good addition to a model predicting productivity in a woodland plot.

Things to think about

All the analysis above used 1 ha square plots. How do estimates change with plots of different shapes? Square plots have a higher perimeter:area ratio than circular plots. How do these edge effects bias the results? Is it possible to account for edge effects by excluding trees found at the edge of the plot as reference trees? How would one set the width of a plot edge exclusion buffer?

References

- Gadow, Klaus. 1993. "Zur Bestandesbeschreibung in Der Forsteinrichtung." *Forst Und Holz* 48 (January): 602–6.
- Hui, G., X. Zhao, Z. Zhonghua, and K. von Gadow. 2011. "Evaluating Tree Species Spatial Diversity Based on Neighbourhood Relationships." *Forest Science* 57 (4): 292–300. <https://doi.org/http://dx.doi.org/10.1093/forestscience/57.4.292>.
- Janzen, Daniel H. 1970. "Herbivores and the Number of Tree Species in Tropical Forests." *The American Naturalist* 104 (940): 501–28. <https://doi.org/10.1086/282687>.
- Tilman, D., F. Isbell, and J. M. Cowles. 2014. "Biodiversity and Ecosystem Functioning." *Annual Review of Ecology, Evolution, and Systematics* 45: 471–93. <https://doi.org/http://dx.doi.org/10.1146/annurev-ecolsys-120213-091917>.
- von Gadow, K. 1998. "The Neighbourhood Pattern—a New Parameter for Describing Forest Structures." *Centralbl Ges Forstwesen* 115: 1–10. <https://ci.nii.ac.jp/naid/10029315806/en/>.

- von Gadow, K., and G. Y. Hui. 2001. *Characterizing Forest Spatial Structure and Diversity*. Lund, Sweden: Sustainable Forestry in Temperate Regions.
- Wiegand, Thorsten, Savitri Gunatilleke, Nimal Gunatilleke, and Toshinori Okuda. 2007. "Analyzing the Spatial Structure of a Sri Lankan Tree Species with Multiple Scales of Clustering." *Ecology* 88 (12): 3088–3102. <https://doi.org/10.1890/06-1350.1>.