

Worked example of the habitat-species network concept

Habitat-species networks

We use data reported in Hill and Bartomeus (2016) to show the possibilities of the habitat-species network framework to analyze species habitat use at the landscape level.

In this study we surveyed all the dominant habitats in 10 study sites (2 km²). The habitat types surveyed in this study were:

- a) Corridors
- b) Forest
- c) Forest/grassland boundary
- d) Non-flowering crop boundary
- e) Semi-natural grassland
- f) Maintained roadside
- g) Maintained drain.

Within each habitat type a 50m x 3m transect was selected with the greatest number of flowering plant species. Bumblebee richness and abundance was recorded for 15 minutes. Each study plot was surveyed twice between 9th July 2014 and 25th August 2014.

First, let's load and format the data:

```
#load libraries
library(reshape)
library(vegan)
library(bipartite)

#next we read and review the data
d <- read.csv("data/powerlines3.csv", h = TRUE)
head(d)
str(d)

#And check for any inconsistencies
levels(d$Gen_sp)
#We remove unidentified species.
d <- subset(d, Gen_sp != "Bombus_spp")
levels(d$Site)
#Due to issues with swedish letters, they are renamed
levels(d$Site)[1] <- "Angeby"
levels(d$Site)[4] <- "Gavsta"
levels(d$Site)[7] <- "Kottgrind"
levels(d$Site)[8] <- "Laby_Norra"
levels(d$Site)[9] <- "Laby_Sodra"

#Each of the two sampling rounds are pooled for this example
str(d)
d2 <- cast(d, formula = Site + Plot + Habitat + Gen_sp ~ . , fun.aggregate = length)
head(d2)
#rename using same nomenclature as in the paper
colnames(d2) <- c("Site", "Patch", "Habitat", "Gen_sp", "Abundance" )
```

Unfortunately we don't have coordinates for each habitat patch, but we can still visualizing one of the networks, even if it's not spatially explicit.

Visualization

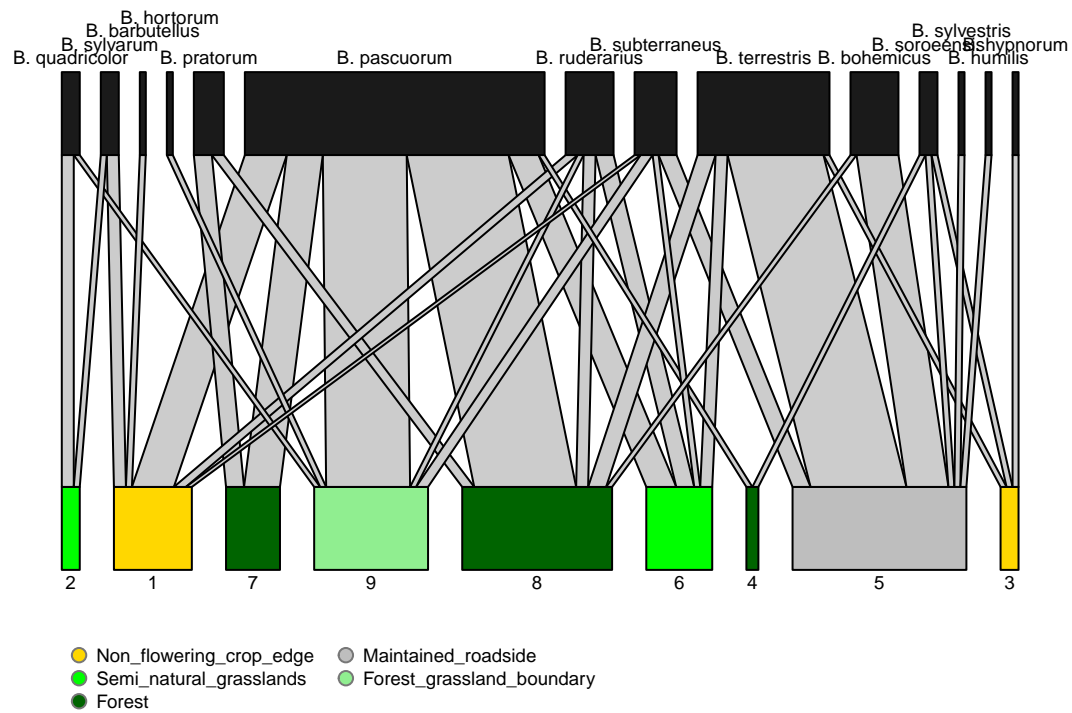
We select one network to work with:

```
#subset one site
site1 <- subset(d2, Site == "Angeby")
site1
#create the network in matrix format
ntw1 <- cast(site1, Patch ~ Gen_sp, fun.aggregate = "sum", value = "Abundance")
#we can remove the first column with rownames
ntw1_ <- ntw1[,-1]
ntw1 <- as.matrix(ntw1_)
colnames(ntw1) <- gsub(pattern = "Bombus_", replacement = "B. ",
                      x = colnames(ntw1_), ignore.case = T)
#create a patch dataframe to store its properties.
patch <- unique(site1[, c("Patch", "Habitat")])
patch$color <- c("gold", "green", "gold", "darkgreen",
                 "grey", "green", "darkgreen",
                 "darkgreen", "lightgreen")

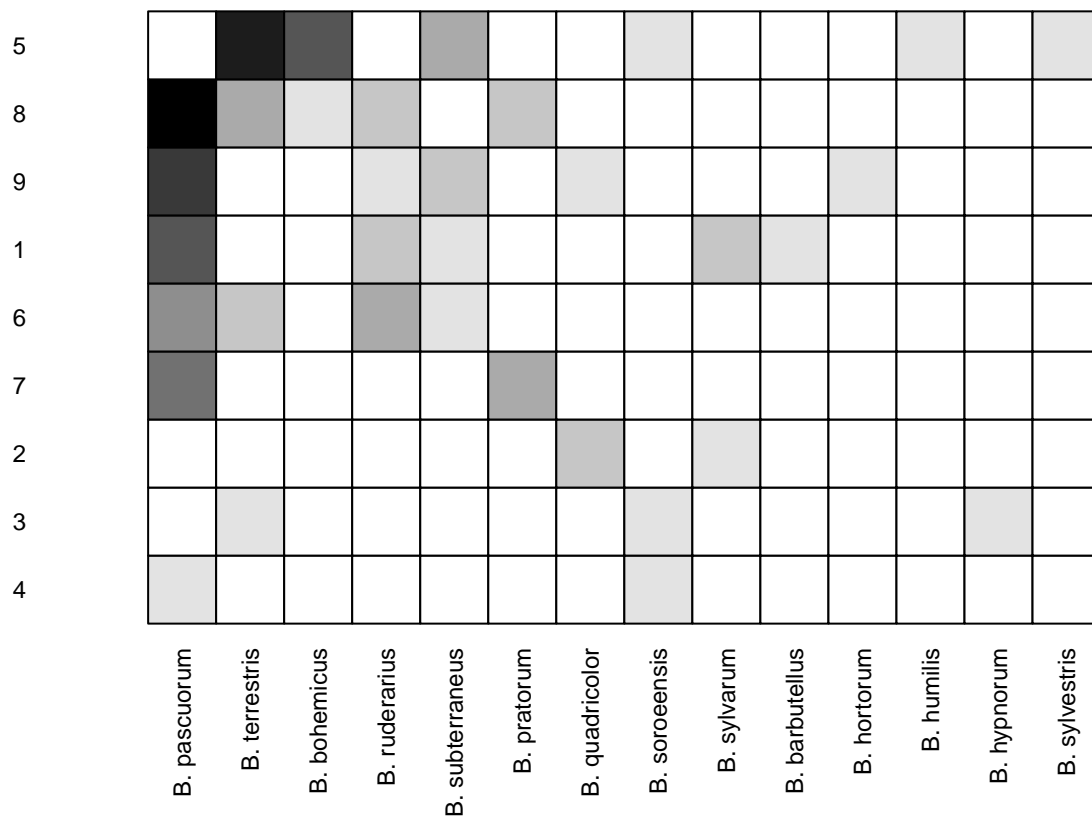
rownames(ntw1) <- patch$Patch
#create a vector of bees
bees <- cast(site1, Gen_sp ~ ., fun.aggregate = sum, value = "Abundance")
colnames(bees)[2] <- "abundance"
bees$labs <- colnames(ntw1)
```

We can use package *bipartite* to plot it in two different ways:

```
#prepare legend
legend <- unique(patch[,c("Habitat", "color")])
par(xpd = T) #allow plotting outside the plot
plotweb(ntw1, col.low = as.character(patch$color))
legend(x=0, y=0.25, as.character(legend$Habitat), pch=21,
       col="#777777", pt.bg=as.character(legend$color),
       pt.cex=1, cex=.6, bty="n", ncol=2)
```



```
visweb(ntw1, prednames = T, preynames = T, labsiz = 0.6)
```



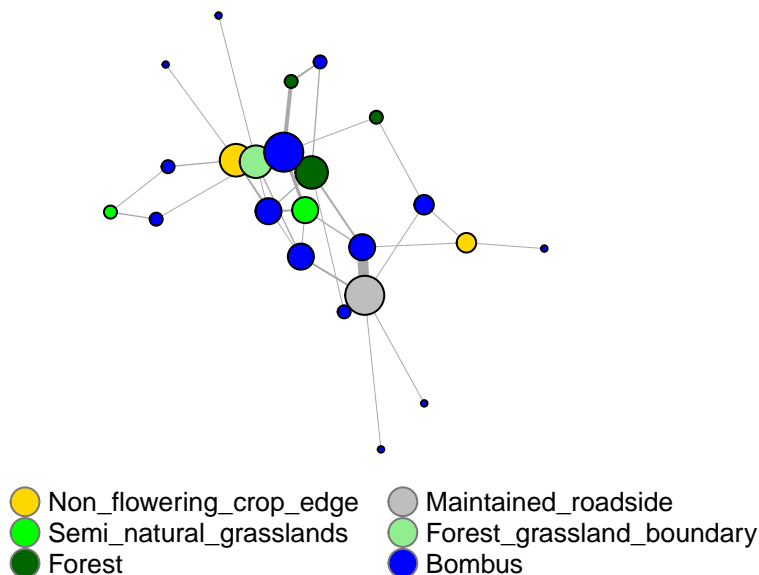
By inspecting the network we can see *Bombus pascuorum* is abundant (large box) and highly connected (6 links), especially to semi-natural habitats (in green). Let's explore other visualization options. Package *igraph* is pretty cool too:

```

#or with prettier tools:
library(igraph)
#prepare the data for igraph
links <- site1[,c("Gen_sp", "Patch", "Abundance")]
colnames(links)[3] <- "weight"
node1 <- unique(site1[,c("Patch", "Habitat")])
colnames(node1) <- c("node", "attribute")
node1$type <- "habitat"
node2 <- data.frame(node = unique(site1[,c("Gen_sp")]),
                    attribute = NA,
                    type = "species")
nodes <- rbind(node1, node2)
#create igraph object
net <- graph_from_data_frame(d=links,
                            vertices=nodes, directed=F)

# Generate colors based habitat:
clrs <- data.frame(nod = V(net)$attribute,
                  cols = c(patch$color, rep("blue", 14)))
V(net)$color <- as.character(clrs$cols)
# Compute node degrees (#links) and use that to set node size:
deg <- degree(net, mode="all")
V(net)$size <- deg*3
# Setting them to NA will render no labels:
V(net)$label <- as.character(nodes$id)
# Set edge width based on weight:
E(net)$width <- E(net)$weight/3
#change arrow size and edge color:
E(net)$arrow.size <- .2 #but note no arrows in Unidirected graphs like this
E(net)$edge.color <- "gray80"
#prepare colors
cl <- unique(clrs)
cl$nod <- as.character(cl$nod)
cl$nod[which(is.na(cl$nod))] <- "Bombus"
plot(net, vertex.label = NA) #force vertex label NA to make visualization clearer.
legend(x=-1.5, y=-1.1, cl$nod, pch=21,
      col="#777777", pt.bg=as.character(cl$cols),
      pt.cex=2, cex=.8, bty="n", ncol=2)

```



Here, we can see that the four larger habitat nodes that are highly connected. Interestingly, these belong to four different habitat types.

There is one more package *visNetwork* that makes impressive interactive plots that may be better to visualize and explore these kind of graphs.

```
library('visNetwork')
colnames(nodes)[1] <- "id"
nodes$shape <- "dot"
nodes$shadow <- TRUE # Nodes will drop shadow
nodes$attribute <- as.character(nodes$attribute)
nodes$attribute[10:23] <- as.character(nodes$id[10:23])
nodes$title <- nodes$attribute # Text on click
nodes$label <- nodes$type # Node label
nodes$size <- deg*3 # Node size
nodes$borderWidth <- 2 # Node border width

nodes$color.background <- clr$cols
nodes$color.border <- "black"
nodes$color.highlight.background <- "orange"
nodes$color.highlight.border <- "darkred"
links$width <- links$weight # line width
links$color <- "gray" # line color
#links$arrows <- "middle" # arrows: 'from', 'to', or 'middle'
links$smooth <- TRUE # should the edges be curved?
links$shadow <- FALSE # edge shadow
colnames(links)[1:2] <- c("from", "to")
visNetwork(nodes, links)
```

Next let's calculate a few network parameters.

Network structure

How is this network structured? To calculate it's *nestedness*, we use weighted NODF, one of the more popular and robust nestedness metrics (Almeida-Neto et al. 2008, Almeida-Neto et al. 2010), but other options are also available.

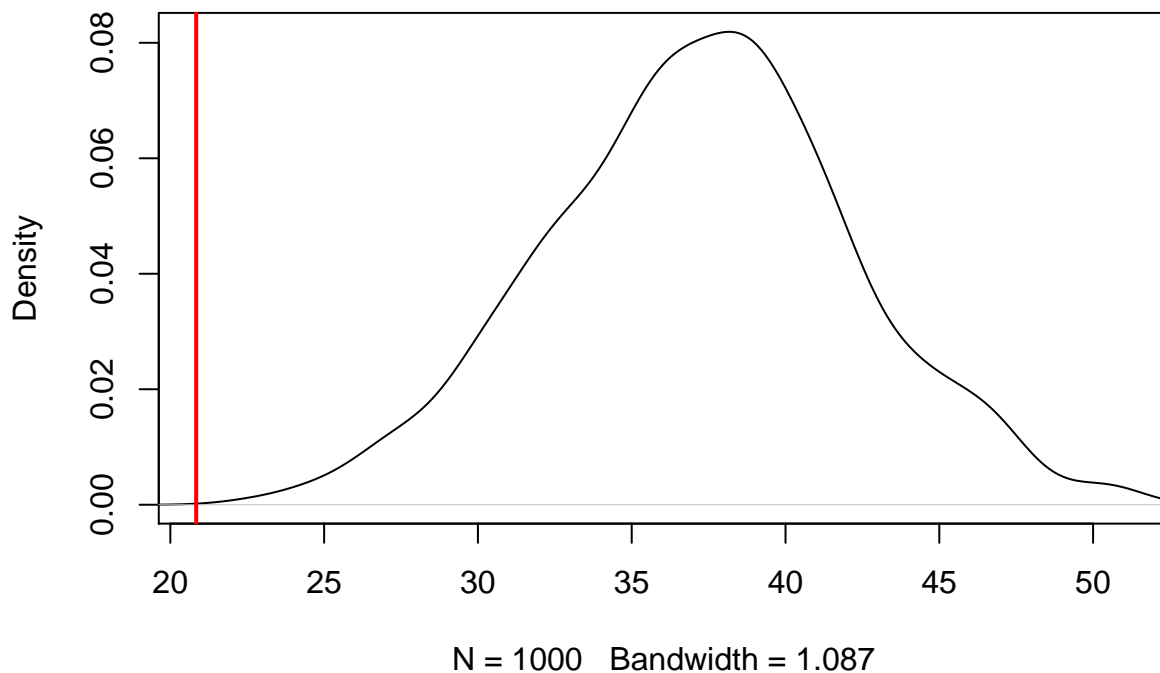
```
(obs <- networklevel(web = ntw1, index = "weighted NODF"))
```

```
## weighted NODF
##      20.8399
```

To know if 20.839895 is more nested than expected by chance for this network, we need to compare it with a null model:

```
nm <- nullmodel(web = ntw1, N=1000, method="r2d")
null <- unlist(sapply(nm, networklevel, index="weighted NODF"))
plot(density(null), xlim=c(min(obs, min(null)), max(obs, max(null))),
     main="comparison of observed with null model Patefield")
abline(v=obs, col="red", lwd=2)
```

comparison of observed with null model Patefield

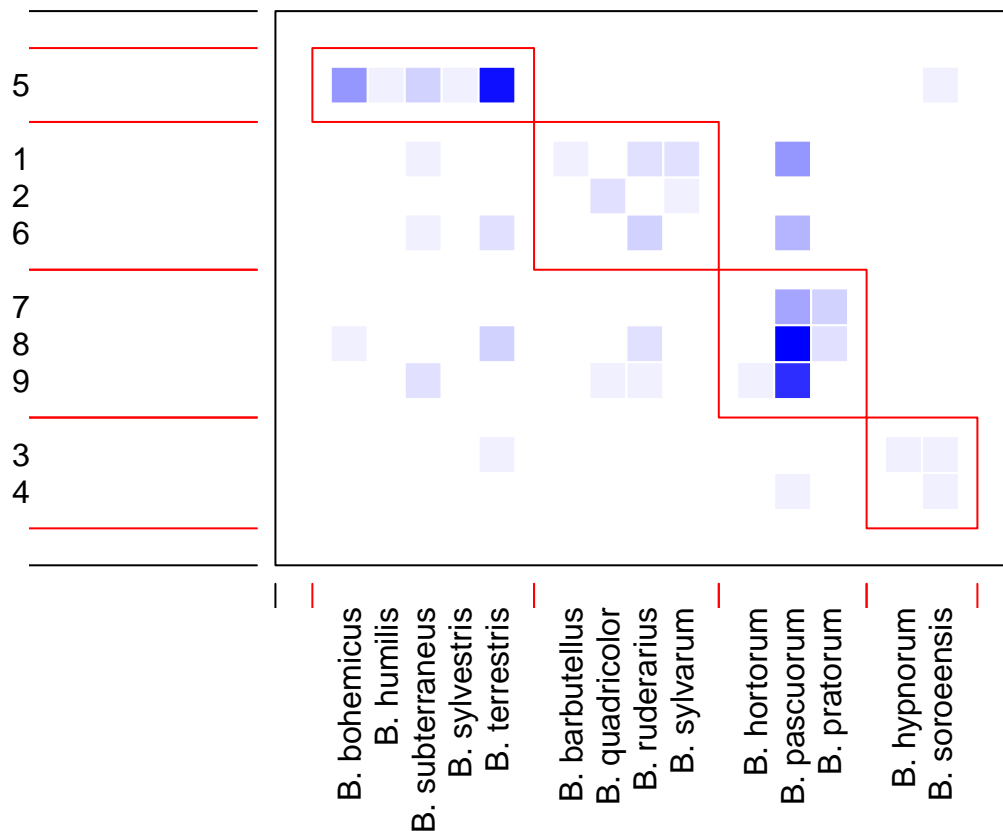


```
praw <- sum(null>obs) / length(null)
```

Here, We can see this network is more nested than expected by chance (p value = 0). This indicates that patches hosting less diversity tend to host also common species, or alternatively, rare species tend to be found in more biodiverse patches.

Now, let's calculate another structural metric. Its modularity. We will use a quantitative version for this end (Dormann et al 2016).

```
res <- computeModules(ntw1)
plotModuleWeb(res, displayAlabels = T)
```

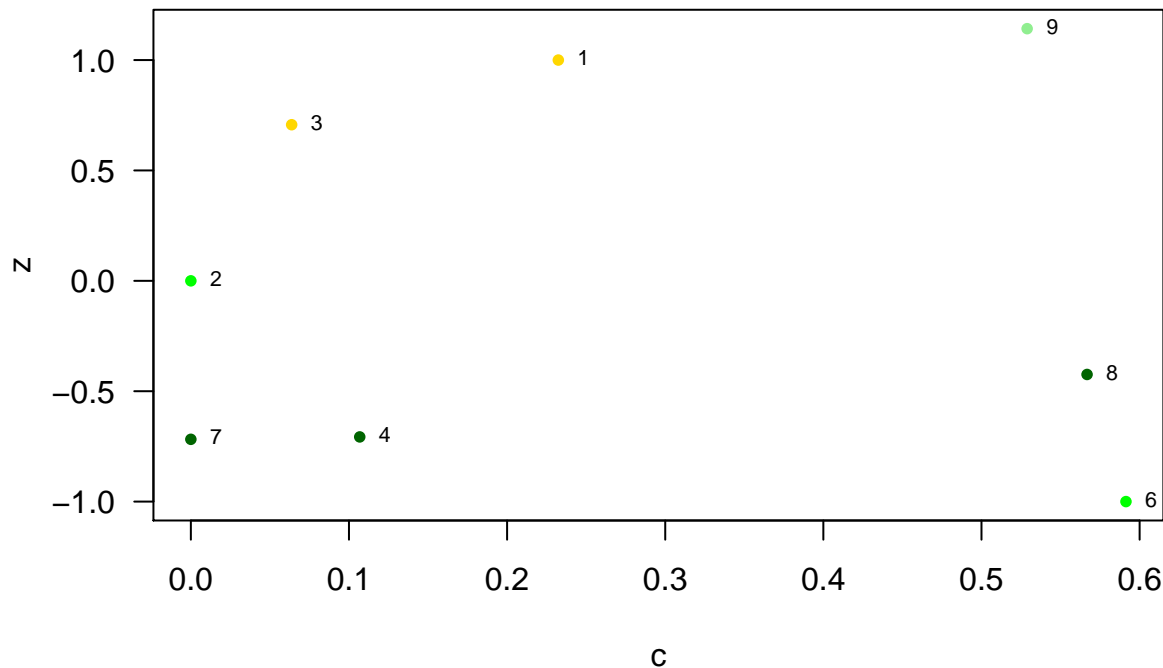


```
#listModuleInformation(res)
#printoutModuleInformation(res)
```

We can identify four modules. To see if this modularity level is larger than expected by chance we can follow a similar approach as above using a null model.

For the purpose of this worked example, we now focus on the description of these modules. We can see patches 7, 8 and 9 form a dense module with common bumblebee species. Those are forested areas. Patches 1,2 and 6 (grasslands and open habitats) host rarer bumblebees. The other modules are very small, but it's interesting that the only roadside patch has it's own module.

```
#we can calculate 2 values for each node
cz <- czvalues(res, weighted = TRUE, level = "lower")
#c : among-module connectivity
#z : within-module connectivity
#Olesen et al. (2006) give critical c and z values of 0.62 and 2.6, respectively. Species exceeding the
plot(cz[[1]], cz[[2]], pch=16, xlab="c", ylab="z",
     cex=0.8, las=1, col = patch$col)
text(cz[[1]], cz[[2]], names(cz[[1]]), pos=4, cex=0.7)
```



```
#we can check congruence between runs with function metaComputeModules
#res2 <- metaComputeModules(as.matrix(ntw1))
```

We can also see the among-module (z) and within-module (c) connectivity of different patches. Interestingly, module 9 and 1 and 3 tend to act as links among modules. This makes sense as patch 9 is the interface between Forest and grassland boundaries. Note c can't be computed for modules composed by a single patch, and hence patch 5 is not plotted.

Next, we can ask which habitat patches would we prioritize for its conservation value. For this we can calculate its strength (Bascopmte et al 2006).

```
#we transpose the matrix to calculate it for the lower level.
patch$strength <- bipartite::strength(t(ntw1), type="Bascompte")
patch
```

##	Patch	Habitat	color	strength
## 1	1	Non_flowering_crop_edge	gold	2.1995238
## 6	2	Semi_natural_grasslands	green	1.0000000
## 8	3	Non_flowering_crop_edge	gold	1.3787879
## 11	4	Forest	darkgreen	0.3533333
## 13	5	Maintained_roadside	grey	4.3641775
## 19	6	Semi_natural_grasslands	green	0.7087662
## 23	7	Forest	darkgreen	0.7200000
## 25	8	Forest	darkgreen	1.2513636
## 30	9	Forest_grassland_boundary	lightgreen	2.0240476

Interesting, roadsides has the highest strength! This is because they sustain both common and rare species. However, we may want to correct for the fact that this habitat is only represented once in the dataset. Not surprisingly forests tend to rank lower, as they host a moderate number of common species.

Let's see now how each patch influences each other (Müller et al 1999):

```
(inf <- PAC(ntw1))
```

##		1	2	3	4	5	6
## 1		0.30432234	0.05128205	0.000000000	0.010769231	0.03296703	0.12252747


```
## 2 0.22222222 0.55555556 0.000000000 0.000000000 0.00000000 0.00000000
## 3 0.00000000 0.00000000 0.459595960 0.111111111 0.35353535 0.03030303
## 4 0.07000000 0.00000000 0.166666667 0.176666667 0.16666667 0.05000000
## 5 0.01477833 0.00000000 0.036572623 0.011494253 0.73725556 0.06493506
## 6 0.14480519 0.00000000 0.008264463 0.009090909 0.17119244 0.17724321
## 7 0.09333333 0.00000000 0.000000000 0.013333333 0.00000000 0.06666667
## 8 0.11520000 0.00000000 0.005454545 0.013600000 0.12227273 0.10890909
## 9 0.13135338 0.03508772 0.000000000 0.014736842 0.04511278 0.10845865
##          7          8          9
## 1 0.06461538 0.22153846 0.19197802
## 2 0.00000000 0.00000000 0.22222222
## 3 0.00000000 0.04545455 0.00000000
## 4 0.06000000 0.17000000 0.14000000
## 5 0.00000000 0.10540752 0.02955665
## 6 0.05454545 0.24752066 0.18733766
## 7 0.28000000 0.36000000 0.18666667
## 8 0.12960000 0.30456364 0.20040000
## 9 0.08842105 0.26368421 0.31314536
```

We can read this matrix as the influence mediated by shared pollinators between each pair of habitat patches. It looks like influence is low overall (mean = 0.1111111) but some patches influence each others via shared pollinators (e.g. 7->8 0.36, but note this is not reciprocal, as 8->7 influence is moderate: 0.1296)

Finally let's see which are the more selective bumblebees (Blüthgen et al 2007)? For this we regroup at habitat level.

```
d3 <- cast(d, formula = Site + Habitat + Gen_sp ~ . , fun.aggregate = length)
colnames(d3) <- c("Site", "Habitat", "Gen_sp", "Abundance" )
site1b <- subset(d3, Site == "Angeby")
#create the network in matrix format
ntw1b <- cast(site1b, Habitat ~ Gen_sp, fun.aggregate = "sum", value = "Abundance")
#we can remove first column with rownames
ntw1b <- ntw1b[,-1]
#let's visualize it with bipartite
#plotweb(ntw1b)
#and calculate d'
#here low.abun can be used if we know patch attributes like area.
bees$d <- specieslevel(web = ntw1b, index = "d", level = "higher")
bees
```

##	Gen_sp	abundance	labs	d
## 1	Bombus_barbutellus	1	B. barbutellus	0.22629439
## 2	Bombus_bohemicus	8	B. bohemicus	0.36040676
## 3	Bombus_hortorum	1	B. hortorum	0.17833866
## 4	Bombus_humilis	1	B. humilis	0.06033821
## 5	Bombus_hypnorum	1	B. hypnorum	0.22629439
## 6	Bombus_pascuorum	50	B. pascuorum	0.39456717
## 7	Bombus_pratorum	5	B. pratorum	0.35516628
## 8	Bombus_quadricolor	3	B. quadricolor	0.30913755
## 9	Bombus_ruderarius	8	B. ruderarius	0.17267753
## 10	Bombus_soroeensis	3	B. soroeensis	0.01736694
## 11	Bombus_subterraneus	7	B. subterraneus	0.14189519
## 12	Bombus_sylvarum	3	B. sylvarum	0.31301000
## 13	Bombus_sylvestris	1	B. sylvestris	0.06033821
## 14	Bombus_terrestris	22	B. terrestris	0.34582083

We can see *B. terrestris* or *B. pascuorum* have high values of d' (very unselective), while *B. humilis* or *B. soroeensis* are highly selective.

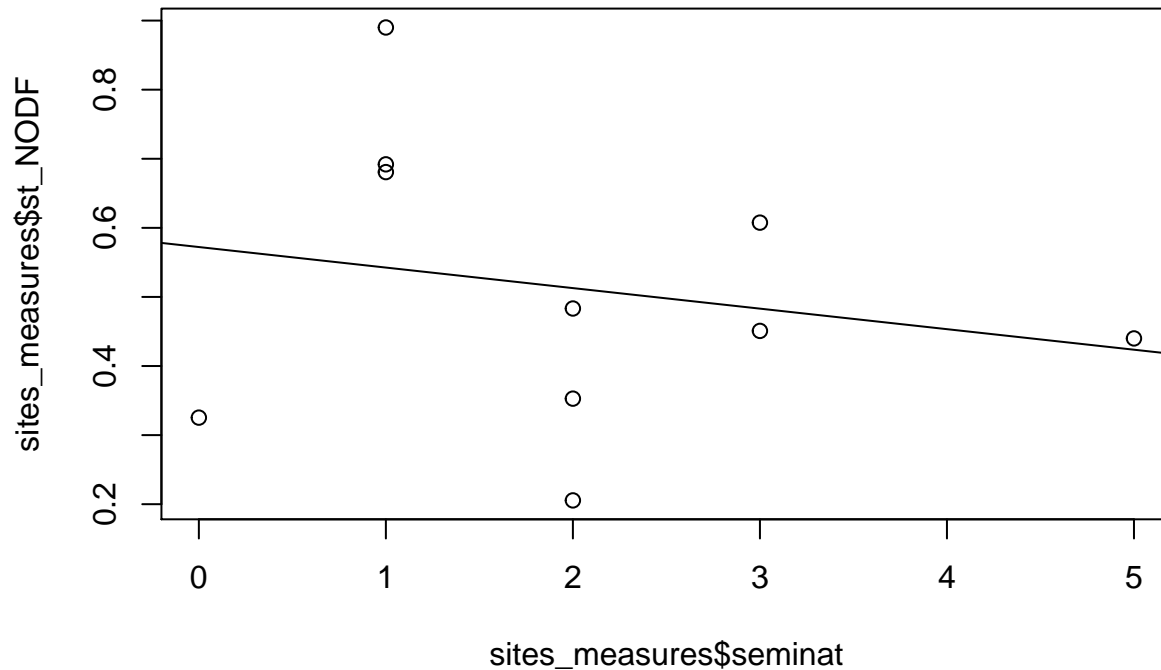
At the site levels we can also determine which site is more nested and if the nestedness pattern correlates with amount of semi-natural habitats or species richness levels. Further, we can determine if more nested networks are also more robust (Memmot et al. 2004)?

Let's calculate nestedness for all networks as well as other network parameters like species richness, robustness and number of semi-natural patches. Remember that to compare nestedness values, we need to standardize them. We suggest to follow Song et al 2017 method, as is the more robust:

```
source("toolbox.R") #load code developed by Song et al. and available in his paper.
#let's loop through all sites
#first we create empty objects to store the data
sites <- unique(d2$Site)
ntwks <- list()
nested <- c()
NODF <- c()
st_NODF <- c()
rob_rand <- c()
rob_real <- c()
rich <- c()
seminat <- c()
for(i in 1:length(sites)){
  sitex <- subset(d2, Site == sites[i])
  #create the network in matrix format
  ntwx <- cast(sitex, Patch ~ Gen_sp, fun.aggregate = "sum",
              value = "Abundance")
  #we can remove first column with rownames
  ntwx <- ntwx[,-1]
  #let's visualize it with bipartite
  #plotweb(ntwx)
  ntwks[[i]] <- ntwx
  #calculate nestedness
  nested[i] <- networklevel(web = ntwx, index = "weighted NODF")
  rob_rand[i] <- robustness(second.extinct(web = ntwx, nrep = 50, participant = "lower", method = "random"))
  NODF[i] <- nestedness_NODF(ntwx)
  st_NODF[i] <- comb_nest(web = ntwx, NODF = NODF, max_NODF = max_nest(ntwx))
  #reate a realistic extinction sequence
  ext_seq <- unique(sitex[,c("Patch", "Habitat")])
  #quick and dirty way to order habitats
  levels(ext_seq$Habitat) <- c("gCorridor", "bForest",
                              "cForest_grassland_boundary",
                              "fMaintained_drain",
                              "eMaintained_roadside",
                              "dNon_flowering_crop_edge",
                              "aSemi_natural_grasslands")
  ext_seq$Patch <- order(as.character(ext_seq$Habitat))
  rob_real[i] <- robustness(second.extinct(web = ntwx, participant = "lower", method = "external", ext_seq = ext_seq))
  rich[i] <- ncol(ntwx)
  seminat[i] <- length(subset(ext_seq, Habitat %in%
                              c("aSemi_natural_grasslands",
                                "bForest"))$Patch)
}
sites_measures <- data.frame(sites, nested, NODF, st_NODF, rob_rand, rob_real, rich, seminat)
```

Are more nested networks correlated with the amount of semi-natural habitats?

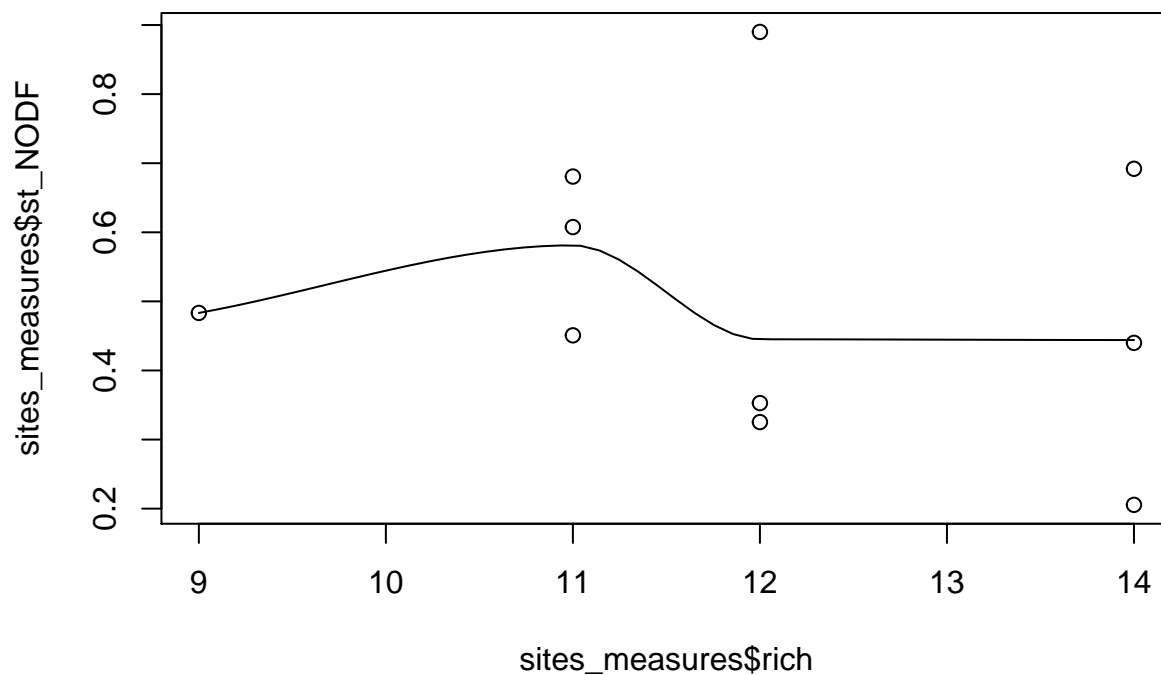
```
plot(sites_measures$st_NODF ~ sites_measures$seminat)
abline(lm(sites_measures$st_NODF ~ sites_measures$seminat))
```



A quick look tells us that there is a weak trend for sites with more seminatural habitat patches (forests and grasslands) to be less nested. We would need better information on the proportion of semi-natural habitats in the landscape to run more accurate statistical tests.

Are more nested networks also more biodiverse?

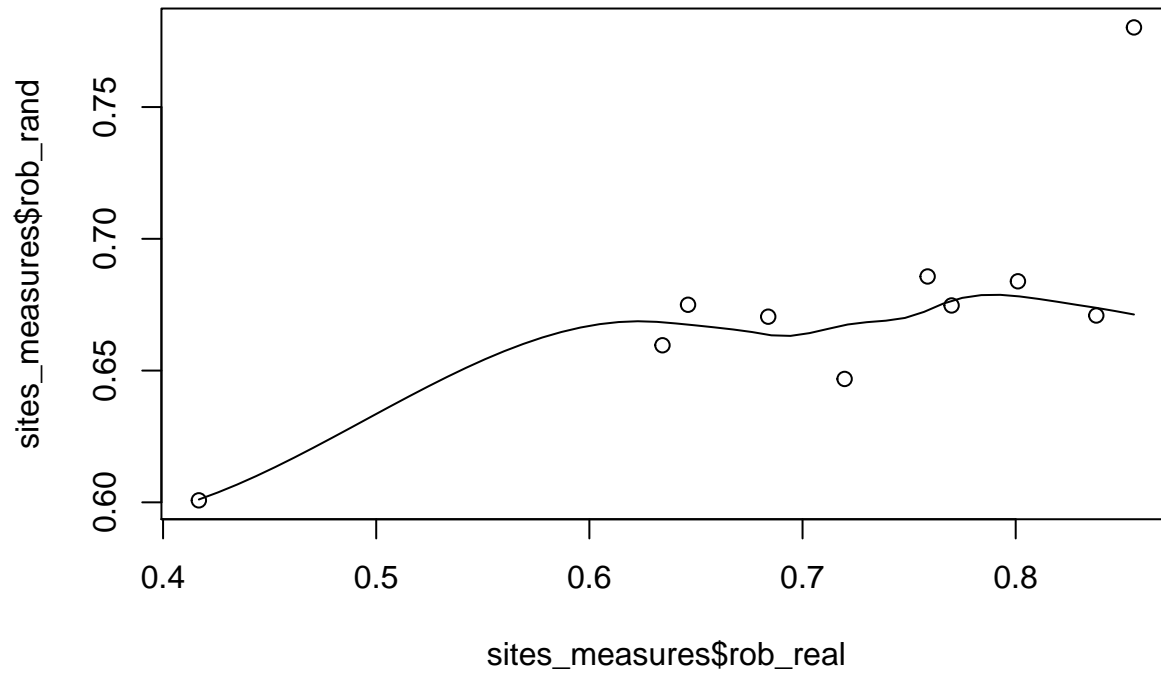
```
scatter.smooth(sites_measures$st_NODF ~ sites_measures$rich)
```



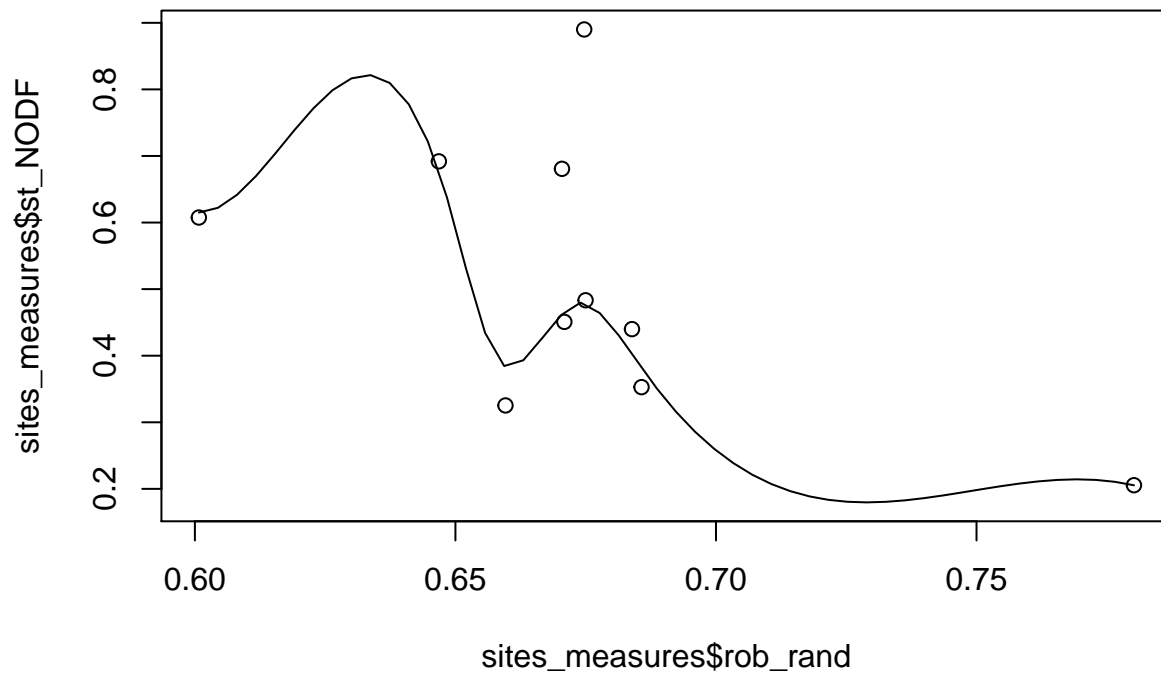
In this case, the plot suggest that nestedness is not related to bumblebee richness levels.

Are more nested sites also more robust to in silico patch removal?

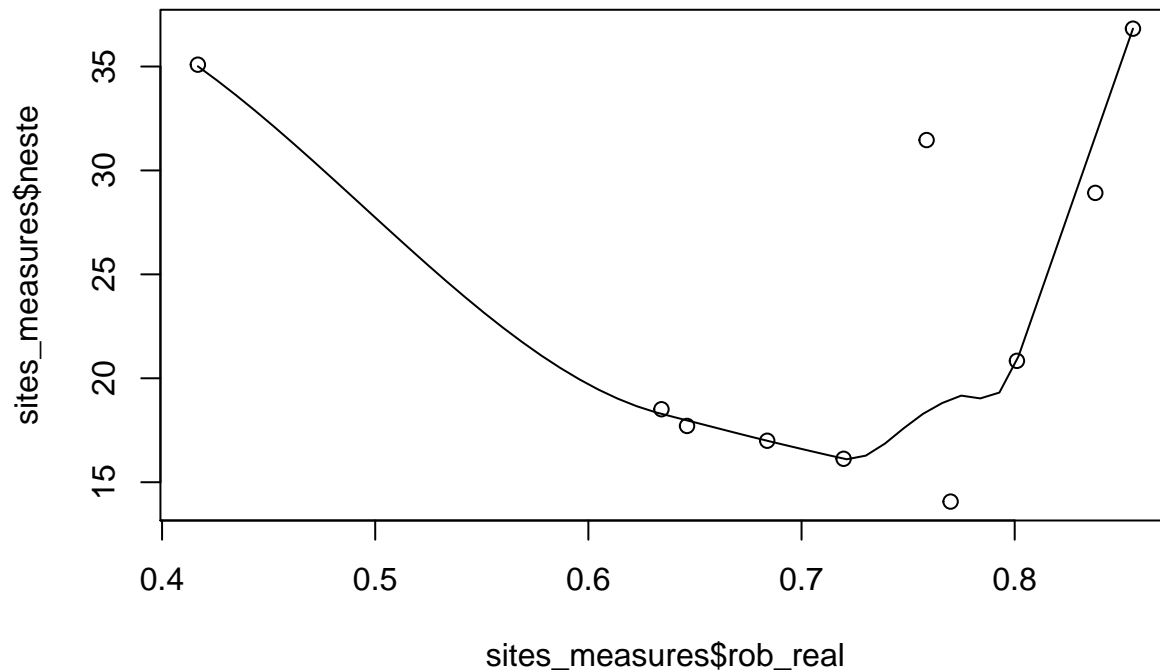
```
scatter.smooth(sites_measures$rob_rand ~ sites_measures$rob_real)
```



```
scatter.smooth(sites_measures$st_NODF ~ sites_measures$rob_rand)
```



```
scatter.smooth(sites_measures$neste ~ sites_measures$rob_real)
```



In this case, both the random removal of patches, and the “semi-natural habitats first” removal are correlated. Contrary to expectations, in this situation robustness is negatively correlated with nestedness level when removals are random, but this relationship is more complex when semi-natural patches are removed first.

What we have learned?

We have seen here how to apply network tools to analyze habitat-species networks, and highlight that the type of analysis will depend on your question and the type of data available.

References

- Almeida-Neto, M., Loyola, R.D., Ulrich, W., Guimaraes, P., Guimaraes, Jr., P.R. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* 117, 1227–1239
- Almeida-Neto, M. & Ulrich, W. (2011) A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling & Software* 26, 173–178
- Bascompte, J., Jordano, P. and Olesen, J. M. 2006 Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* 312, 431–433
- Blüthgen, N., Menzel, F., Hovestadt, T., Fiala, B. and Blüthgen N. 2007 Specialization, constraints and conflicting interests in mutualistic networks. *Current Biology* 17, 1–6
- Burgos, E., H. Ceva, R.P.J. Perazzo, M. Devoto, D. Medan, M. Zimmermann, and A. Maria Delbue (2007) Why nestedness in mutualistic networks? *Journal of Theoretical Biology* 249, 307–313
- Dormann, C. F., and R. Strauß. 2014. Detecting modules in quantitative bipartite networks: the QuanBiMo algorithm. *Methods in Ecology & Evolution* 5 90–98 (or arXiv [q-bio.QM] 1304.3218.)
- Memmott, J., Waser, N. M. and Price M. V. 2004 Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society B* 271, 2605–2611

Müller, C. B., Adriaanse, I. C. T., Belshaw, R. and Godfray, H. C. J. 1999 The structure of an aphid-parasitoid community. *Journal of Animal Ecology* 68, 346–370