

mangal – making ecological network analysis simpler

Timothée Poisot et al. (see complete author list below)

Author list: Timothée Poisot (1,2), Benjamin Baiser (3), Jennifer A. Dunne (4,5), Sonia Kéfi (6), François Massol (7,8), Nicolas Mouquet (6), Tamara N. Romanuk (9), Daniel B. Stouffer (1), Spencer A. Wood (10,11), Dominique Gravel (12,2)

1. University of Canterbury, School of Biological Sciences, Christchurch, New Zealand
2. Québec Centre for Biodiversity Sciences, Montréal (QC), Canada
3. Department of Wildlife Ecology and Conservation, University of Florida, Gainesville
4. Sante Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501
5. Pacific Ecoinformatics and Computational Ecology Lab, 1604 McGee Ave., Berkeley, CA 94703
6. Institut des Sciences de l'Évolution, UMR CRNS 5554, Université Montpellier 2, 3405 Montpellier, France
7. Laboratoire Génétique et Evolution des Populations Végétales, CNRS UMR 8198, Université Lille 1, Bâtiment SN2, F-59655 Villeneuve d'Ascq cedex, France
8. UMR 5175 CEFE – Centre d'Ecologie Fonctionnelle et Evolutive (CNRS), 1919 Route de Mende, F-34293 Montpellier cedex 05, France
9. Department of Biology, Dalhousie University
10. Natural Capital Project, School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195, USA
11. Department of Biological Sciences, Idaho State University, Pocatello, ID 83209, USA
12. Université du Québec à Rimouski, Département de Biologie, 300 Allées des Ursulines, Rimouski (QC) G5L 3A1, Canada

Author for correspondence: t.poisot@gmail.com.

1 **Short title:** Automated retrieval of ecological networks data

2 **Keywords:** R, API, database, open data, ecological networks, species interactions

3 The study of ecological networks is severely limited by (i) the difficulty to access data, (ii) the lack of a
4 standardized way to link meta-data with interactions, and (iii) the disparity of formats in which ecologi-
5 cal networks themselves are represented. To overcome these limitations, we conceived a data specification
6 for ecological networks. We implemented a database respecting this standard, and released a R package
7 (`rmangal`) allowing users to programmatically access, curate, and deposit data on ecological interactions. In
8 this article, we show how these tools, in conjunctions with other frameworks for the programmatic manipula-
9 tion of open ecological data, streamlines the analysis process, and improves replicability and reproducibility
10 of ecological networks studies.

1 Introduction

2 Ecological networks are efficient representations of the complexity of natural communities, and help discovering mech-
3 anisms contributing to their persistence, stability, resilience, and functioning. Most of the “early” studies of ecological
4 networks were focused on understanding how the structure of interactions within one location affected the ecological
5 properties of this local community. They revealed the contribution of ‘average’ network properties, such as the buffering
6 impact of modularity on species loss (Pimm *et al.* 1991, Yodzis (1981)), the increase in robustness to extinctions along
7 with increases in connectance (Dunne *et al.* 2002), and the fact that organization of interactions maximizes biodiversity
8 (Bastolla *et al.* 2009). New studies introduced the idea that networks can vary from one locality to another. They can
9 be meaningfully compared, either to understand the importance of environmental gradients on the presence of ecological
10 interactions (Tylianakis *et al.* 2007), or to understand the mechanisms behind variation itself (Poisot *et al.* 2012, 2014).
11 Yet, meta-analyses of a large number of ecological networks are still extremely rare, and most of the studies comparing
12 several networks do so within the limit of particular systems (Schleuning *et al.* 2011; Dalsgaard *et al.* 2013; Poisot *et al.*
13 2013b; Chamberlain *et al.* 2014; Olito & Fox 2014). The severe shortage of publicly shared data in the field also restricts
14 the scope of large-scale analyses.

15 It is possible to predict the structure of ecological networks, either using latent variables (Rohr *et al.* 2010; Eklöf *et al.*
16 2013) or actual trait values (Gravel *et al.* 2013). The calibration of these approaches require accessible data, not only
17 about the interactions, but about the traits of the species involved. Comparing the efficiency of different methods is also
18 facilitated if there is an homogeneous way of representing ecological interactions, and the associated metadata. In this
19 paper, we (i) establish the need of a data specification serving as a common language among network ecologists, (ii)
20 describe this data specification, and (iii) describe `rmangal`, a R package and companion database relying on this data
21 specification. The `rmangal` package allows to easily deposit and retrieve data about ecological interactions and networks
22 in a publicly accessible database. We provide use cases showing how this new approach makes complex analyzes simpler,
23 and allows for the integration of new tools to manipulate biodiversity resources.

24 Networks need a data specification

25 Ecological networks are (often) stored as an *adjacency matrix* (or as the quantitative link matrix), that is a series of 0
26 and 1 indicating, respectively, the absence and presence of an interaction. This format is extremely convenient for *use*
27 (as most network analysis packages, *e.g.* `bipartite`, `betalink`, `foodweb`, require data to be presented this way), but
28 is extremely inefficient at *storing* meta-data (this can be done by adding attributes to the matrix objects in the language
29 that support it). In most cases, an adjacency matrix informs on the identity of species (in cases where rows and columns
30 are named), and the presence or absence of interactions. If other data about the environment (*e.g.* where the network
31 was sampled) or the species (*e.g.* the population size, trait distribution, or other observations) are available, they are most
32 either given in other files, or as accompanying text. In both cases, making a programmatic link between interaction data
33 and relevant meta-data is difficult, time-consuming, and error-prone. Because of the lack of a common structure for these
34 data, the process cannot be automated when different datasets needs to be processed.

35 By contrast, a data specification (*i.e.* a set of precise instructions detailing how each object should be represented)
36 provides a common language for network ecologists to interact, and ensure that, regardless of their source, data can be
37 used in a shared workflow. Most importantly, a data specification describes how data are *exchanged*. Each group retains
38 the ability to store the data in the format that is most convenient for in-house use, and only needs to provide export options
39 (*e.g.* through an API, a programmatic interface running on a webserver, returning data in response to queries in a pre-

1 determined language) respecting the data specification. This approach ensures that *all* data can be used in meta-analyses,
2 thus increasing their long-term impact (Piwowar & Vision 2013). Data archival also offers additional advantages for
3 ecology. The aggregation of local observation can reveal large-scale phenomenon (Reichman *et al.* 2011), which would
4 be unattainable in the absence of a collaborative effort. Data archival in databases also prevents data rot and data loss
5 (Vines *et al.* 2014), thus ensuring that data on interaction networks – which are typically hard and costly to produce –
6 continue to be available and usable.

7 Elements of the data specification

8 The data specification (*Fig. 1*) is built around the idea that (ecological) networks are collections of relationships be-
9 tween ecological objects, each element having particular meta-data associated. In this section, we detail the way net-
10 works are represented in the mangal specification. An interactive webpage with the complete data specification can be
11 found online at <http://mangal.uqar.ca/doc/spec/>. The data specification is available either at the API root (*e.g.*
12 <http://mangal.uqar.ca/api/v1/?format=json>), or can be viewed using the `whatIs` function from the R package
13 (see *Supp. Mat. 1*). Rather than giving an exhaustive list of the data specification, this section serves as an overview of
14 each element, and how they interact.

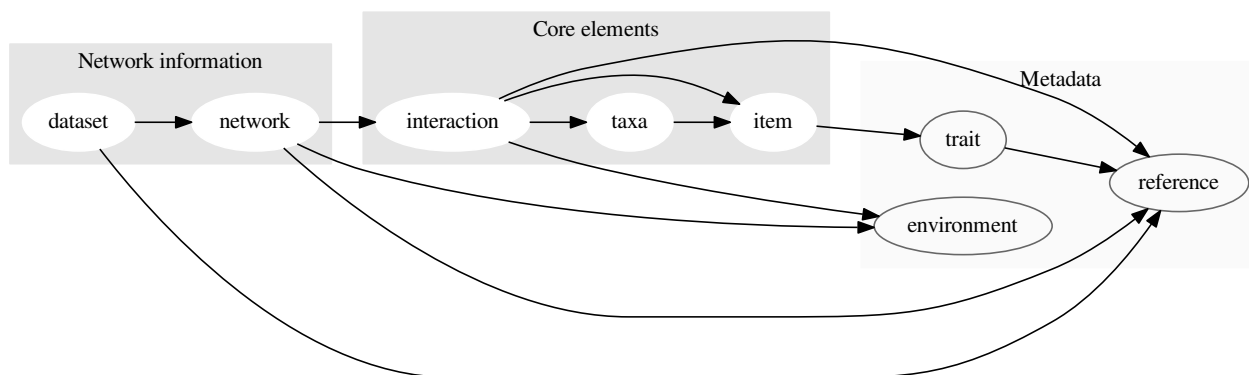


Fig. 1: An overview of the data specification, and the hierarchy between objects. Each box correspond to a level of the data specification, and arrows represent a relationship between objects. For example, `taxa` have associated `items`, and `interactions` have associated `taxa`, `items`, `references`. Not all relationships are mandatory, as we detail below.

15 We propose JSON as an efficient way to uniformise data representation for two main reasons. First, it has emerged as
16 a *de facto* standard for web platforms serving data, and accepting data from users. Second, it allows *validation* of the
17 data: a JSON file can be matched against a scheme, and one can verify that it is correctly formatted. Finally, JSON objects
18 representing individual items are easily and cheaply (memory-wise) parsed in the most common programming languages,
19 notably R (equivalent to `list`) and python (equivalent to `dict`). For most users, the format in which data are transmitted
20 is unimportant, as the interaction happens within R – as such, knowing how JSON objects are organized is only useful for
21 those who want to interact with the API directly. The `rmangal` package takes care of converting the data into the correct
22 JSON format to upload them in the database.

1 Node information

2 Taxa

3 Taxa are a taxonomic entity of any level, identified by their name, vernacular name, and their identifiers in a variety of
4 taxonomic services. Associating the identifiers of each taxa allows using the new generation of open data tools, such
5 as `taxize` (Chamberlain & Szöcs 2013), in addition to protecting the database against taxonomic revisions. The data
6 specification currently has fields for `ncbi` (National Center for Biotechnology Information), `gbif` (Global Biodiversity
7 Information Facility), `tsn` (Taxonomic Serial Number, used by the Integrated Taxonomic Information System), `eol`
8 (Encyclopedia of Life) and `bold` (Barcode of Life) identifiers. We also provide the taxonomic status, *i.e.* whether a taxa
9 is a true taxonomic entity, a “trophic species”, or a morphospecies. Taxonomic identifiers can either be added by the
10 contributors, or will be automatically retrieved during the automated curation routine.

11 Item

12 An `item` is any measured instance of a taxon. Items have a `level` argument, which can be either `individual` or
13 `population`; this allows to represent both individual-level networks (*i.e.* there are as many `items` of a given taxa as there
14 were individuals of this sampled), and population-level networks. When `item` represents a population, it is possible to
15 give a measure of the size of this population. The notion of `item` is particularly useful for time-replicated designs: each
16 observation of a population at a time-point is an `item` with associated `trait` values, and possibly population size.

17 Network information

18 All objects described in this sub-section can have a spatial position, information on the date of sampling, and references
19 to both papers and datasets.

20 Interaction

21 An `interaction` links two `taxa` objects (but can also link pairs of `items`). The most important attributes of `interactions`
22 are the type of interaction (of which we provide a list of possible values, see *Supp. Mat. 1*), and its `ob_type`, *i.e.* how
23 it was observed. This field help differentiate direct observations, text mining, and inference. Note that the `obs_type`
24 field can also take `confirmed absence` as a value; this is useful for, *e.g.*, “cafeteria” experiments in which there is high
25 confidence that the interaction did not happen.

26 Network

27 A `network` is a series of `interaction` object, along with (i) informations on its spatial position (provided at the latitude
28 and longitude), (ii) the date of sampling, and (iii) references to measures of environmental conditions.

29 Dataset

30 A `dataset` is a collection of one or several `network(s)`. Datasets also have a field for `data` and `papers`, both of which
31 are references to bibliographic or web resources describing, respectively, the source of the data, and the papers in which

1 these data have been significantly used. Datasets or networks are the preferred entry point into the resources, although in
2 some cases it can be meaningful to get a list of interactions only.

3 **Meta-data**

4 **Trait value**

5 Objects of type `item` can have associated `trait` values. These consist in the description of the trait being measured,
6 the value, and the units in which the measure was taken. As traits may have been measured at a different time and/or
7 location that the interaction was, they have fields for time, latitude and longitude, and references to original publication
8 and original datasets.

9 **Environmental condition**

10 Environmental conditions are associated to datasets, networks, and interactions objects, to allow for both macro and micro
11 environmental conditions. These are defined by the environmental property measured, its value, and the units. Similarly
12 to traits, they have fields for time, latitude and longitude, and references to original publication and original datasets.

13 **References**

14 References are associated to datasets. They accommodate the DOI, JSON or PubMed identifiers, or a URL. When
15 possible, the DOI should be preferred as it offers more potential to interact with other on-line tools, such as the *CrossRef*
16 API.

17 **Use cases**

18 In this section, we present use cases using the `rmangal` package for R, to interact with a database implementing this data
19 specification, and serving data through an API (<http://mangal.uqar.ca/api/v1/>). It is possible for users to deposit
20 data into this database, through the R package. Data are made available under a *CC-0 Waiver* (Poisot *et al.* 2013a).
21 Detailed informations about how to upload data are given in the vignettes and manual of the `rmangal` package.

22 The data we use for this example come from Ricciardi et al. (2010). These were previously available on the *Interaction-*
23 *Web DataBase* as a single `xls` file. We uploaded them in the `mangal` database at <http://mangal.uqar.ca/api/v1/dataset/1>.
24 Before running the examples, users need to install the relevant packages and connect to the database:

25 **Link-species relationships**

26 In the first example, we visualize the relationship between the number of species and the number of interactions, which
27 Martinez (1992) proposed to be linear (in food webs).

```
# Pull the dataset of interest
dataset <- getDataset(api, DSET_ID)
```

```

# Get each network in the dataset as a graph object
graphs <- alply(dataset$networks, 1, function(x) toIgraph(api, x))

# Make a data.frame with the number of links and species
ls <- ldply(graphs, function(x) c(S = length(V(x)), L = length(E(x))))
ls$X1 <- aaply(as.numeric(as.vector(ls$X1)), 1,
              function(x) getNetwork(api, x)$name)
colnames(ls)[1] <- 'Network'

# Now plot this dataset
plot(jitter(L)~jitter(S), ls, log='xy', pch=22, bg='lightgrey',
     lwd=1.5, cex=1.5, xlab='Species richness', ylab='Number of interactions')
# Constant connectance
X <- c(1:max(ls$S))
Y <- X^2 * mean(ls$L/ls$S^2)
lines(X, Y, lty=2)
# Best fit
bfit <- lm(L~S, ls)
Yf <- X * bfit$coefficients[2] + bfit$coefficients[1]
lines(X, Yf)
legend('bottomright', pch=c(22, NA, NA), lty=c(NA, 2, 1),
      pt.cex=c(1.5, 1, 1), lwd=c(1.5, 1, 1), pt.bg=c('lightgrey', NA, NA),
      legend=c('Data', 'Constant connectance', 'Best fit (linear model)'), bty='n')

```

- 1 Getting the data to produce Fig. 2 requires less than 10 lines of code. The only information needed is the identifier of the
- 2 network or dataset, which we suggest should be reported in publications as: “These data were deposited in the mangal
- 3 format at <URL>/api/v1/dataset/<ID>” (where <URL> and <ID> are replaced by the corresponding values), preferably
- 4 in the methods, possibly in the acknowledgements. So as to encourage data sharing and its recognition, we encourage
- 5 users of the database to cite the original dataset or publication.

6 Network beta-diversity

- 7 In the second example, we use the framework of network β -diversity (Poisot *et al.* 2012) to measure the extent to which
- 8 networks that are far apart in space have different interactions. Each network in the dataset has a latitude and longitude,
- 9 meaning that it is possible to measure the geographic distance between two networks.
- 10 For each pair of network, we measure the geographic distance (in km.), the species dissimilarity (β_S), the network dissim-
- 11 ilarity when all species are present (β_{WN}), and finally, the network dissimilarity when only shared species are considered
- 12 (β_{OS}).

```

# We first retrieve all informations about the networks
Networks <- alply(dataset$networks, 1, function(x) getNetwork(api, x))

# Extract the lat/lon data

```

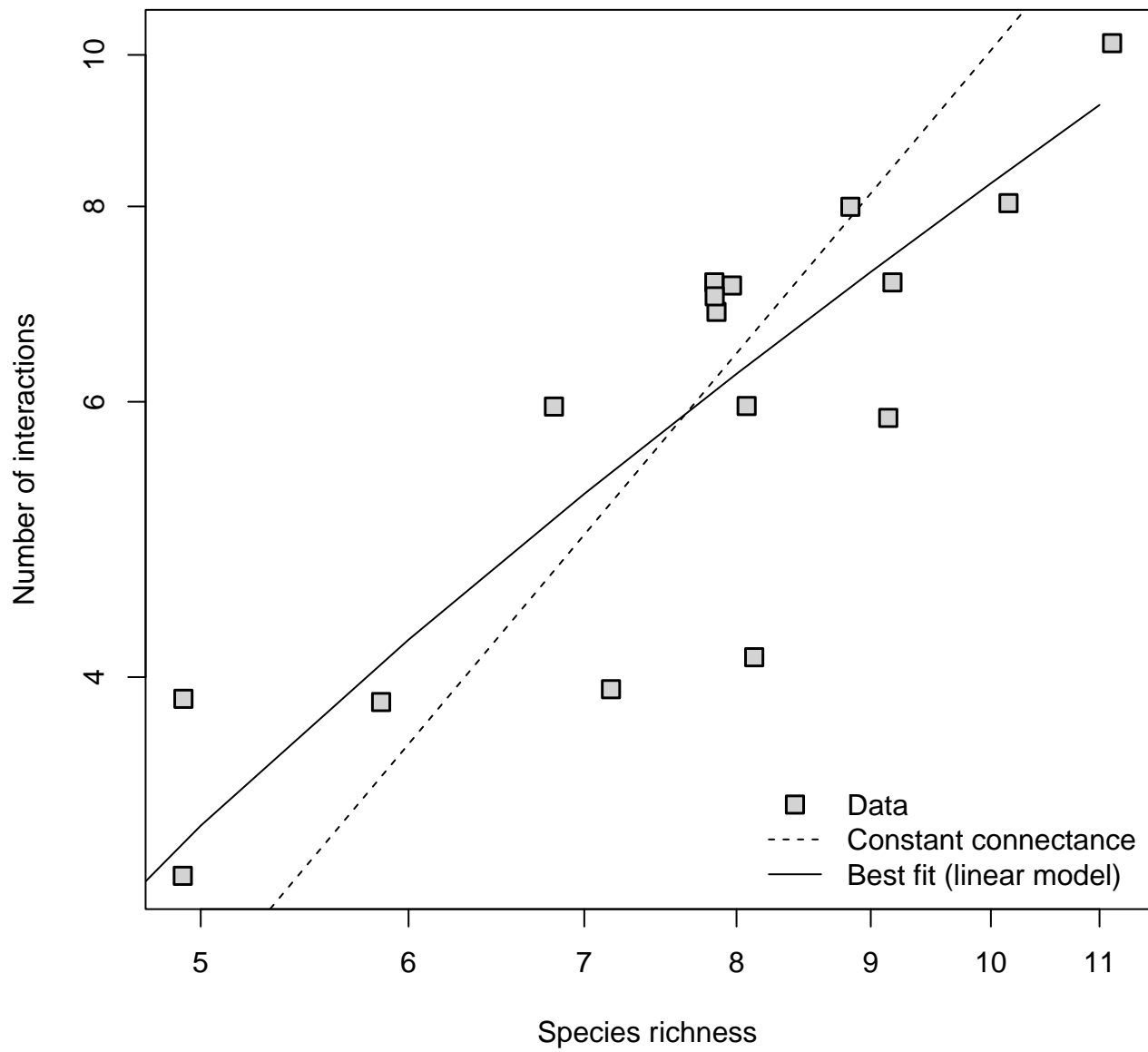


Fig. 2: Relationship between the number of species and number of interactions in the anemonefish-fish dataset. Constant connectance refers to the hypothesis that there exist a quadratic relationship between these two quantities.


```

LatLon <- ldply(Networks, function(x) c(name = x$name, lat = x$latitude, lon = x$longitude))
rownames(LatLon) <- LatLon$name
LatLon$lat <- as.numeric(LatLon$lat)
LatLon$lon <- as.numeric(LatLon$lon)
LatLon <- LatLon[,c('lat', 'lon')]

# Then we measure the distances between all pairs of sites
GeoDist <- spDists(as.matrix(LatLon, latlon=TRUE))
colnames(GeoDist) <- rownames(GeoDist) <- rownames(LatLon)
GeoDist <- as.dist(GeoDist)

# Now, we measure the beta-diversity of the networks
names(graphs) <- aapply(names(graphs), 1, function(x) Networks[[x]]$name)
# Finally, we measure the beta-diversity
BetaDiv <- network_betadiversity(graphs)

# We add the geographic distance
BetaDiv$GEO <- GeoDist

# And we do some plots
par(mfrow=c(2,2), pty='s')
with(BetaDiv,{
  plot(GEO, S, pch=22, bg='lightgrey', cex=1.5, lwd=1.5,
       xlab="Geographic distance", ylab="Species composition dissimilarity")
  plot(GEO, WN, pch=22, bg='lightgrey', cex=1.5, lwd=1.5,
       xlab="Geographic distance", ylab="Network dissimilarity (all species)")
  plot(GEO, OS, pch=22, bg='lightgrey', cex=1.5, lwd=1.5,
       xlab="Geographic distance", ylab="Network dissimilarity (shared species)")
})

```

- 1 As shown in *Fig. 3*, while species dissimilarity and overall network dissimilarity increase when two networks are far
- 2 apart, this is not the case for the way common species interact. This suggests that in this system, network dissimilarity
- 3 over space is primarily driven by species turnover. The ease to gather both raw interaction data and associated meta-data
- 4 make producing this analysis extremely straightforward.

5 Spatial visualization of networks

- 6 Bascompte (2009) uses an interesting visualization for spatial networks, in which each species is laid out on a map at the
- 7 center of mass of its distribution; interactions are then drawn between species to show how species distribution determines
- 8 biotic interactions. In this final use case, we propose to reproduce a similar figure (*Fig. 4*).

```

# We fill a community data matrix
sp_by_site <- llply(graphs, function(x) unlist(V(x)$name))

```

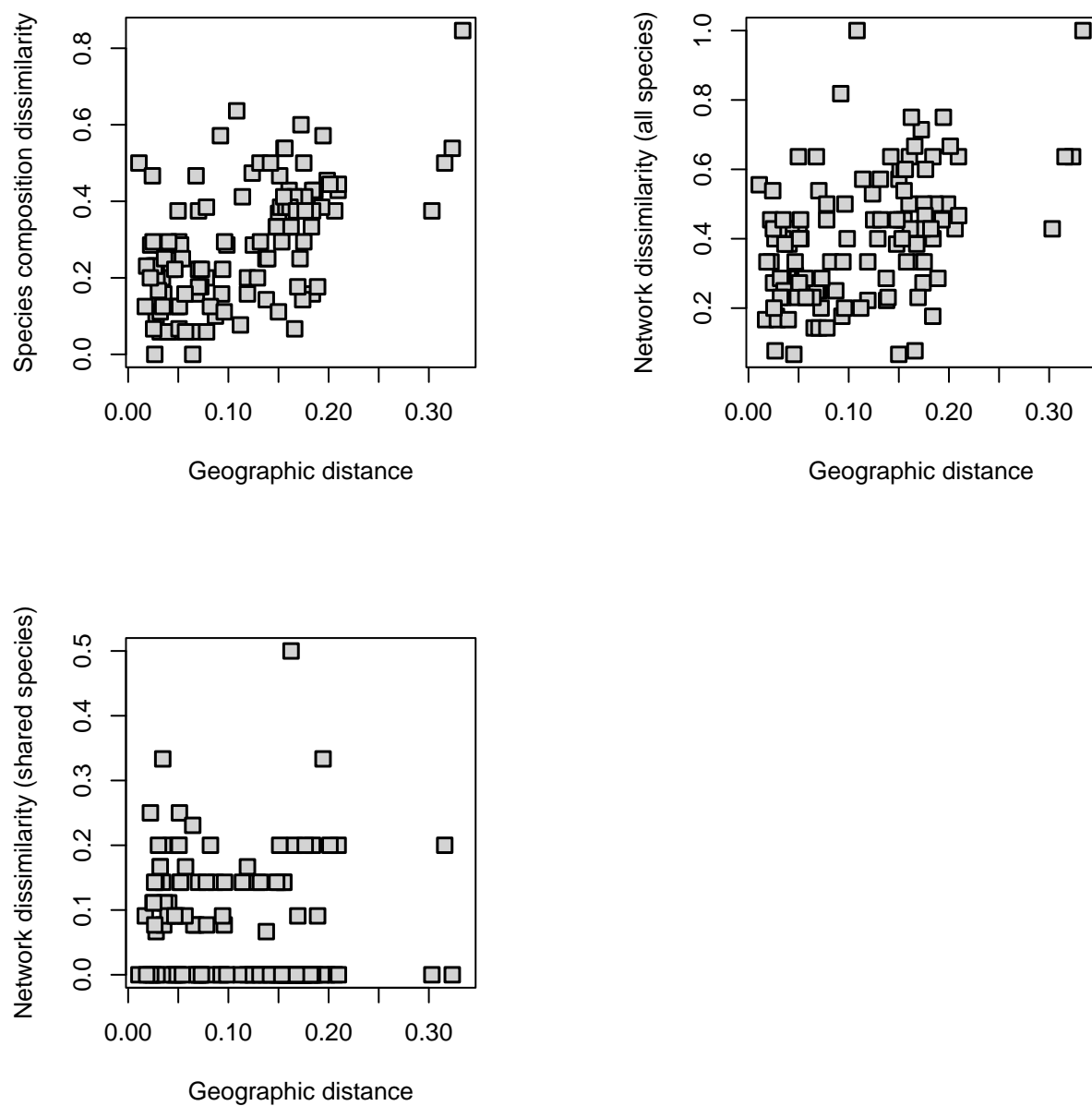


Fig. 3: Relationships between the geographic distance between two sites, and the species dissimilarity, network dissimilarity with all, and only shared, species.

```

sp_list <- unique(unlist(sp_by_site))
M <- matrix(0, ncol = length(sp_list), nrow = length(sp_by_site))
colnames(M) <- sp_list
rownames(M) <- names(sp_by_site)
for (site in c(1:length(sp_by_site))) M[names(sp_by_site)[site], sp_by_site[[site]]] = 1

# Next, we get the center position for each species
# (i.e. the mean position of the sites it occurs at)
sp_center <- apply(M, 2, function(x) colMeans(LatLon[names(x)[x > 0], ]))
rownames(sp_center) <- sp_center[, 1]
sp_center <- sp_center[, -1]

# We now create a regional network using betalink::metaweb
Mw <- metaweb(graphs)

# Plot a map
center_point <- colMeans(sp_center)
vcolors <- c(brewer.pal(9, "Set1"), brewer.pal(8, "Set2"))

Layout <- matrix(c(1,2,2,2), 2, 2)
colSize <- c(1.3, 2.0)
rowSize <- c(1.2, 2)
layout(Layout, colSize, rowSize)

# Inset map is number 1
par(mar=c(4.5, 1, 1, 4))
map("worldHires", xlim=c(90,136), ylim=c(-15,15), col="gray90",
    fill=TRUE, resolution=0)
points(center_point[2], center_point[1], pch=1, cex=2, lwd=2)
box()

par(mar=c(4.1, 4.1, 4.1, 4.1))
map("worldHires", xlim=c(124.0,125.1), ylim=c(1.2,1.9), col="gray90",
    fill=TRUE, resolution = 0)
plot(Mw, layout = jitter(as.matrix(LatLon[,c('lon','lat')])),
    rescale = FALSE, add = TRUE, vertex.color = vcolors, vertex.size = 1,
    vertex.label = NA, edge.arrow.size = 0.25, edge.color = 1)
axis(1)
axis(4)
legend("bottomleft", fill = vcolors, legend = V(Mw)$name, inset = 0.02,
    cex = 0.7, bty = "n", ncol=2)

```

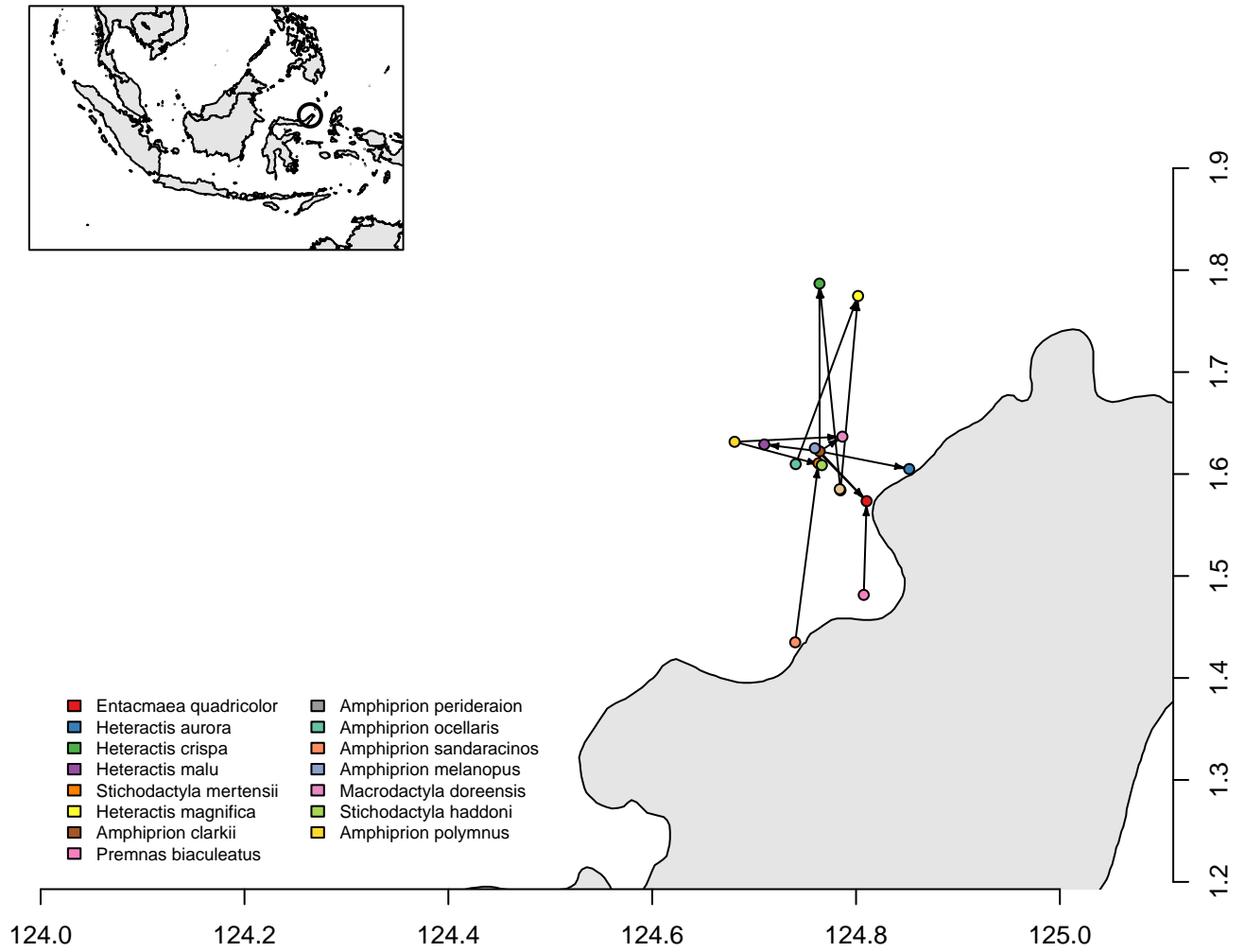


Fig. 4: Spatial plot of a network, using the maps and rmangal packages. The circle in the inset map show the location of the sites. Each dot in the main map represents a species, with (symbiotic mutualism) drawn between them. The land is in grey.

1 Conclusions

2 The mangal data format will allow researchers to put together dataset with species interactions and rich meta-data, that
3 are needed to adress emerging questions about the structure of ecological networks. We deployed an online database with
4 an associated API, relying on this data specification. Finally, we introduced rmangal, a R package designed to interact
5 with APIs using the mangal format. We expect that the data specification will evolve based on the needs and feedback
6 of the community. At the moment, users are welcome to propose such changes on the project issue page: <https://github.com/mangal-wg/mangal-schemes/issues>. A python wrapper for the API is also available at <http://github.com/mangal-wg/pymangal/>. Additionally, there are plans to integrate this database with GLOBI, so that data
9 can be accessed from multiple sources (Poelen *et al.* 2014).

10 **Acknowledgements** This paper was developed during a workshop hosted at the *Santa Fe Institute*. TP, DBS, and DG
11 acknowledge funding from the Canadian Institute of Ecology and Evolution. We thanks James Caveen for setting up the
12 database and web server on which the data are hosted. We thank Scott Chamberlain and one anonymous reviewer for
13 comments on the manuscript.

14 References

- 15 Bascompte, J. (2009). Disentangling the web of life. *Science (New York, N.Y.)*, **325**, 416–9.
- 16 Bastolla, U., Fortuna, M.A., Pascual-García, A., Ferrera, A., Luque, B. & Bascompte, J. (2009). The architecture of
17 mutualistic networks minimizes competition and increases biodiversity. *Nature*, **458**, 1018–1020.
- 18 Chamberlain, S.A. & Szöcs, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Research*. Retrieved January
19 8, 2014,
- 20 Chamberlain, S.A., Cartar, R.V., Worley, A.C., Semmler, S.J., Gielens, G., Elwell, S., Evans, M.E., Vamosi, J.C. & Elle,
21 E. (2014). Traits and phylogenetic history contribute to network structure across Canadian plant–pollinator communities.
22 *Oecologia*, 1–12. Retrieved September 11, 2014,
- 23 Dalsgaard, B., Trøjelsgaard, K., González, A.M.M., Nogués-Bravo, D., Ollerton, J., Petanidou, T., Sandel, B., Schleuning,
24 M., Wang, Z., Rahbek, C., Sutherland, W.J., Svenning, J.-C. & Olesen, J.M. (2013). Historical climate-change influences
25 modularity and nestedness of pollination networks. *Ecography*, no–no. Retrieved May 14, 2013,
- 26 Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002). Network structure and biodiversity loss in food webs: robustness
27 increases with connectance. *Ecology Letters*, **5**, 558–567.
- 28 Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P., Dalsgaard, B., Sassi, C. de, Galetti, M.,
29 Guimarães, P.R., Lomáscolo, S.B., Martín González, A.M., Pizo, M.A., Rader, R., Rodrigo, A., Tylianakis, J.M.,
30 Vázquez, D.P. & Allesina, S. (2013). The dimensionality of ecological networks (J. Dunne, Ed.). *Ecology Letters*,
31 n/a–n/a. Retrieved February 27, 2013,
- 32 Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013). Inferring food web structure from predator-prey body
33 size relationships. *Methods in Ecology and Evolution*.
- 34 Martinez, N.D. (1992). Constant connectance in community food webs. *The American Naturalist*, **139**, 1208–1218.
- 35 Olito, C. & Fox, J.W. (2014). Species traits and abundances predict metrics of plant–pollinator network structure, but not
36 pairwise interactions. *Oikos*, n/a–n/a. Retrieved September 10, 2014,

- 1 Pimm, S.L., Lawton, J.H. & Cohen, J.E. (1991). Food web patterns and their consequences. *Nature*, **350**, 669–674.
- 2 Piwowar, H.A. & Vision, T.J. (2013). Data reuse and the open data citation advantage. *PeerJ*, **1**, e175. Retrieved October
3 5, 2013,
- 4 Poelen, J.H., Simons, J.D. & Mungall, C.J. (2014). Global Biotic Interactions: An open infrastructure to share and analyze
5 species-interaction datasets. *Ecological Informatics*. Retrieved September 15, 2014,
- 6 Poisot, T., Canard, E., Mouillot, D., Mouquet, N. & Gravel, D. (2012). The dissimilarity of species interaction networks.
7 *Ecology Letters*, **15**, 1353–1361.
- 8 Poisot, T., Mounce, R. & Gravel, D. (2013a). Moving toward a sustainable ecological science: don't let data go to waste!
- 9 Poisot, T., Stanko, M., Miklisová, D. & Morand, S. (2013b). Facultative and obligate parasite communities exhibit
10 different network properties. *Parasitology*, **FirstView**, 1–6.
- 11 Poisot, T., Stouffer, D.B. & Gravel, D. (2014). Beyond species: why ecological interaction networks vary through space
12 and time. *Oikos*.
- 13 Reichman, O.J., Jones, M.B. & Schildhauer, M.P. (2011). Challenges and opportunities of open data in ecology. *Science*,
14 **331**, 703–5.
- 15 Ricciardi, F., Boyer, M. & Ollerton, J. (2010). Assemblage and interaction structure of the anemonefish-anemone mutu-
16 alism across the manado region of sulawesi, indonesia. *Environmental Biology of Fishes*, **87**, 333–347. Retrieved January
17 10, 2014,
- 18 Rohr, R.P., Scherer, H., Kehrli, P., Mazza, C. & Bersier, L.-F. (2010). Modeling food webs: exploring unexplained
19 structure using latent traits. *The American naturalist*, **176**, 170–7.
- 20 Schleuning, M., Blüthgen, N., Flörchinger, M., Braun, J., Schaefer, H.M. & Böhning-Gaese, K. (2011). Specialization
21 and interaction strength in a tropical plant-frugivore network differ among forest strata. *Ecology*, **92**, 26–36.
- 22 Tylianakis, J.M., Tscharnkte, T. & Lewis, O.T. (2007). Habitat modification alters the structure of tropical host–parasitoid
23 food webs. *Nature*, **445**, 202–205. Retrieved January 28, 2013,
- 24 Vines, T.H., Albert, A.Y.K., Andrew, R.L., Débarre, F., Bock, D.G., Franklin, M.T., Gilbert, K.J., Moore, J.-S., Renaut,
25 S. & Rennison, D.J. (2014). The Availability of Research Data Declines Rapidly with Article Age. *Current Biology*, **24**,
26 94–97. Retrieved September 24, 2014,
- 27 Yodzis, P. (1981). The stability of real ecosystems. *Nature*, **289**, 674–676.