The linkage methods we used to connect cluster two generic clusters A and B were the maximum linkage method (at least one realised connection from the nodes of A to the nodes of B to consider a connection from A to B), intermediate linkage methods (at least 25%, 50%, 75% of the possible connections from the nodes of A to B were realised to consider a connection from A to B) and minimum linkage method (all connections realised from the nodes of A to the nodes of B to consider a connection from A to B). The interaction strength methods that we considered were the new interaction strength from A to B was the minimum/maximum/average interaction strength from the nodes of A to B and the sum of the interaction strengths going from the nodes of A to the nodes of B.

This is the first time prey based and predator based were used for aggregation.

However, we did not have enough time to see what is the best one in that sense. This is why we decided to use structure as a proxy for dynamics. We used centrality indices as a proxy for network structure. Centrality indices tell us about many things of nodes. This is why we used them, as a proxy for structure. So, we can actually say that this is a study where we check whether large food webs can be aggregated into smaller ones by maintaining the properties of the nodes.

<

Large food webs often need to be reduced in size, especially if we want to simulate them dynamically. Future work might be focused on the effects of aggregation methods on dynamical indices.

Different centrality indices have different predictive power to find keystone species.

The most reliable indices (e.g. keystoneness index, see Libralato, Christensen, & Pauly, (2006)) are the ones based on a dynamical model of the network.

However, building a good dynamical model is really time consuming and might not be feasible for multiple networks.

This is why often researchers use topological indices, such as the ones studied in this work.

To check the reliability of topological indices, Gouveia, Móréh, & Jordán (2020) tested how much they agreed with the dynamical index keystoneness (KS).

They found that the most reliable topological index was the weighted degree (wDC).

It could predict the most important species 70.06% of the times.

It was followed by the 5-step weighted topological importance (WI5).

A combination of wDC and WI5 increased this percentage to 78.42\%.

It seems like a good aggregation should then maintain these two indices.

Another index that a good aggregation method should maintain is the trophic level.

This can be calculated, for example, as the trophic position in this study.

Trophic level is associated with many biological properties of species.

For example, the direction by which a perturbation propagates to (Curtsdotter et al., 2011).

Or the risk of extinction of a species (Binzer et al., 2011).

Or the success of a species to colonise a new environment (Holt, 2010).

(Add: (Gouveia et al., 2020) didn't take STO and wSTO into consideration, as well as the trophic level.)

By looking at the results in Figure 5, we see clear differences between algorithms.

Density modularity is the worst aggregation algorithm.

This is probably because it produces communities based on other factors (such as phylogeny, body mass, and habitat structure, see Rezende, Albert, Fortuna, & Bascompte (2009)) rather than trophic links between modules.

Predator modularity, prey modularity and the group model also performed poorly, considering they performed worse than REGE and Jaccard across all centrality indices (except for predator modularity that performed the best with regards to trophic position).

Jaccard and REGE seem to be the best aggregation methods.

They both maintained well the rankings of all centrality indices.

Jaccard maintained the greatest number of centrality indices slightly better than REGE.

However, REGE maintained better the pattern of the two most important indices according to Gouveia et al. (2020), wDC and WI5.

In light of these findings, we suggest using either the Jaccard index or the REGE index for aggregation in keystone species research.

The clusters produced by these two similarity indices can be useful to answer different questions.

The Jaccard index maintains all the information about a certain network, but it doesn't allow you to compare different food webs (Luczkovich et al., 2003).

The REGE index produces a controlled loss of information but it allows to compare between different food webs (Luczkovich et al., 2003).

Future work should be focused on developing new algorithms for the aggregation of species in food webs.

It should follow five steps:

1. Developing a singular index that tells us the goodness of the aggregation.
2. Deciding what is the best method to assign the interaction strength of the link between the clusters.
3. Deciding what is the best percentage of realised links between clusters to decide that there is a link between them.
4. Deciding what are the indices we should use. Are more indices better or worse? Should we combine all indices? Or should we develop different aggregation methods according to the type of pattern we want to preserve (e.g., redundancy, capacity of spreading an effect)?

This research, however, is complicated by our lack of understanding of keystone species.

It is not clear what the best centrality indices are.

Also, it seems like trophic links might not be the only interactions important in determining keystone species.

Donohue et al. (2017) found that secondary extinctions were better modelled when they took into consideration competition.

This might mean that a multilayer approach (see Pilosof, Porter, Pascual, & Kéfi (2017)) could be a better way of finding keystone species.

So, we would argue that developing algorithms of aggregation should go hand in hand with this type of research.

Another possible direction is checking how different data aggregations influence dynamical indices instead of topological indices.

For example, the keystoneness index (KS) (Libralato et al., 2006) or eigenvector centrality (Allesina & Pascual, 2009b).

An interesting way of using aggregation, however, could be to find keystone species.

It has been suggested by Bond (1994) that keystone species might be the species that cannot be aggregated.

This makes us wonder whether it would be possible to use aggregation to find keystone species in topological networks.

And whether they would work better than centrality indices.

We always need to remind ourselves that aggregated food webs bias our analyses.

Ultimately, the question should be also whether we should aggregate our data.

Aggregating entities in ecology has been found to always bias models and produce worse results.

For example, Woodward et al., (2010) showed how using the Allometric Diet Breadth Model (ADBM) (Owen L. Petchey, Beckerman, Riede, & Warren, 2008) to predict the links between nodes predicted 52\% of the links when building a species-based food web.

This number increased to 83\% when considering size-based classes.

Another example is from Petchey & Gaston (2002).

When species were put into functional groups, the community showed high functional redundancy.

When single species were considered singularly, the community showed low functional redundancy.

We might argue that the best way of working with food webs is not aggregating.

However, the problem is that it is never possible to do this.

This is because we have a limited resolution power when we sample and the data that have been collected in the past and are available for our analyses have low resolution, at least in some part of the network.

This could be solved by the way we sample.

Ideally, we should develop technologies and methods that allow us to sample a food web in its highest resolution possible.

However, this might be really expensive, and ecology might not have enough money to finance this type of research.

This should be especially directed toward the lowest trophic levels of the food web, for example plankton in aquatic food webs.

Aggregation research should also inform sampling methods and how to deal with missing data ( see Patonai & Jordán (2017).

It is also important to remember that it seems like species who are unique trophospecies seem to be really important for secondary extinction.

In particular, Petchey, Eklöf, Borrvall, & Ebenman (2008) found that they are particularly vulnerable to secondary extinctions when trying to model their dynamic food web.

The fact that the concept of trophospecies not only is important to understand which ones are the most vulnerable species in the system, but it seems also important in something related to how different species are related to each other.

If, as someone said but I don't remember who, keystone species are the ones that are unique in their trophospecies, it means that secondary extinctions happen only if you hit the network close or on keystone species.

This is because keystone species, if my interpretation is correct, are not only the most important, but also the most vulnerable.

This might be due to the fact that they have such low abundance as well.

Keystone species might be also the ones that we cannot aggregate in a food web (Bond, 1994).

So, at this point, maybe data aggregation would reveal keystone species.

The fact that the aggregation of according to Jaccard not only can reveal the keystone species, but also would maintain the relative importance of the nodes, seems like a great way of aggregating data.

Wait: I need to check whether the species that are unique in their cluster also have high centrality indices.

The problem with aggregation algorithms is that they might give you a solution where certain clusters are not connected to each other.