CrossMark
← click for updates

RESEARCH ARTICLE

REVISED An *a posteriori* measure of network modularity [version 3; referees: 1 approved, 2 approved with reservations]

Timothée Poisot[1,2]

[1]Département de Biologie, Chimie et Géographie, Université du Québec à Rimouski, Rimouski, G5L 3A1, Canada
[2]Québec Centre for Biodiversity Science, Montréal, H3A 1B1, Canada

Abstract

Measuring the modularity of networks, and how it deviates from random expectations, important to understand their structure and emerging properties. Several measures exist to assess modularity, which when applied to the same network, can return both different modularity values (i.e. different estimates of how modular the network is) and different module compositions (i.e. different groups of species forming said modules). More importantly, as each optimization method uses a different optimization criterion, there is a need to have an a posteriori measure serving as an equivalent of a goodness-of-fit. In this article, I propose such a measure of modularity, which is simply defined as the ratio of interactions established between members of the same modules vs. members of different modules. I apply this measure to a large dataset of 290 ecological networks representing host–parasite (bipartite) and predator–prey (unipartite) interactions, to show how the results are easy to interpret and present especially to a broad audience not familiar with modularity analyses, but still can reveal new features about modularity and the ways to measure it.

Open Peer Review

Referee Status: ?✓?

	Invited Referees		
	1	**2**	**3**
REVISED version 3 published 27 Dec 2013		✓ report	? report
UPDATED version 2 published 30 Aug 2013		✓	? report
version 1 published 23 May 2013	? report	? report	

1   **Carsten Dormann**, University Freiburg Germany

2   **Daniel Carstensen**, Universidade Federal de São Paulo Brazil

3   **Jochen Fründ**, Georg-August-University of Göttingen Germany

Discuss this article

Comments (0)

Corresponding author: Timothée Poisot (timothee_poisot@uqar.ca)

How to cite this article: Poisot T. **An *a posteriori* measure of network modularity [version 3; referees: 1 approved, 2 approved with reservations]** *F1000Research* 2013, **2**:130 (doi: 10.12688/f1000research.2-130.v3)

Competing interests: No competing interests were disclosed.

First published: 23 May 2013, **2**:130 (doi: 10.12688/f1000research.2-130.v1)

Introduction

Modularity, the fact that groups of nodes within a network interact more frequently with themselves than with other nodes, is an important property of several systems, including genetic[1,2], informatics[3], ecological[4], and socio-economic[5] interactions, as well as biogeographic patterns[6,7] and disease spread management[8]. Because of the relevance of modularity for network properties, it is important to assess it correctly. Several methods exist to measure network modularity, some of which rely on the optimization of a given criterion[9,10], label propagation[11], or combination of these approaches[12,13]. These methods return two elements. The first is a value of modularity for the networks, most often within the 0–1 interval. Each method often has a threshold value, above which a network is considered to be modular. Increasing values reflect an increasingly modular structure. The second element is a "community partition", *i.e.* the attribution of each node to a module.

Recently, Thébault[7] showed that different measures of modularity tailored to presence/absence matrices (*i.e.* networks in which links have no weight), gave roughly equal estimates of the significance of modularity, but differed in the community partition they returned (*i.e.* the identity of nodes composing each module varied). In such situations, one might look for a way to choose which community partition should be used. The challenge in this situation is that the criteria used by each optimisation method cannot be meaningfully compared, and so there is a need for *a posteriori* measurement of how strong the modular structure is, regardless of the method used to obtain the community partition. More importantly, this criterion should be different than the one used to track the progress of any optimisation algorithm.

An important feature of modular networks is the occurrence of interactions between nodes of different modules. They contribute to the propagation of disturbances[4], flow of information[14,15], and cross-regulation of biological processes[16], *inter alia*[17]. In addition to measuring how modular the network is, determining to what extent modules are connected, and to identify nodes and edges responsible for connecting modules, is thus valuable information. In this article, I propose an *a posteriori* measure of the proportion of interactions established between modules, *i.e.* edges connecting different communities. I apply this measure to the community partition identified by the Louvain method on 290 ecological networks, and show that it behaves in a similar way to other modularity measures.

The measure

In this contribution I define the *realized modularity*, termed Q_R. Q_R measures the extent to which edges, within a network, are established between nodes belonging to the same module. For E edges

in a network, if W of them are established between members of the same module, then

$$Q_R = \frac{W}{E}. \qquad (1)$$

When there are no between-module links, then $W = E$ and Q_R takes the maximal value of 1. When between-module interactions are as numerous as within-module interactions, then $W = E/2$, and Q_R takes the minimal value of 1/2. To express the *realized modularity* as a value between 0 and 1, it is expressed as:

$$Q'_R = 2 \times Q_R - 1. \qquad (2)$$

Note that Q'_R will yield values in the [0; 1] interval only if there are more edges established within than between modules. Although, if modules are determined at random, Q'_R values are expected to be centered on 0, it is expected that they will increase when modules are properly optimized (only as far as the network is modular). The main advantage of Q_R is that it is agnostic with regard to the measure used to optimize modularity (and even to the method by which the nodes were assigned to modules, which can be arbitrary), as it acts *a posteriori*, *i.e.* after nodes have been attributed to modules. Nonetheless, it assumes a simple yet functional definition of modularity: the fact that nodes interact more within than between modules. Given that measuring to which extent this is true, it can therefore be used to select the community detection method maximizing modularity. This measure works on most types of networks, as it makes no difference if links are directional, or if the networks are bipartite/unipartite. An illustration of this measure is given in Figure 1. This measure is purposefully simple, (i) so that it makes only minimal assumptions about what modularity is (except for the fact that in a modular network, nodes interact more within than between modules), or how it should be optimized, and (ii) because it is not meant to be used to optimize modularity, but to either compare the outcome of different methods, or present the value of modularity in a way that is straightforward to interpret.

A python implementation of this measure, using the networkx package, is proposed at https://gist.github.com/tpoisot/4947006. It reads data in the edge list format, and offers additional functions to generate null networks, as detailed in the following section.

Example application: realized modularity in ecological networks

In this section, I analyze the modular structure of a large dataset of 290 ecological networks (187 food webs and 113 host-parasite networks) published in previous meta-analyses[18,19]. Modularity is an important feature of ecological interaction networks, which is linked to their resilience[20,21], stability[7], biogeographic structure[22], functioning[23], and to the evolutionary mechanisms involved in their assembly[24]. Notably, the occurrence of interactions between and within modules plays a central role in the structure of pollination networks[4], and help buffer the effect of species extinctions[21].

The existence of interactions in ecological systems involves a large family of processes, ranging from abudance related[25,26] (abundant
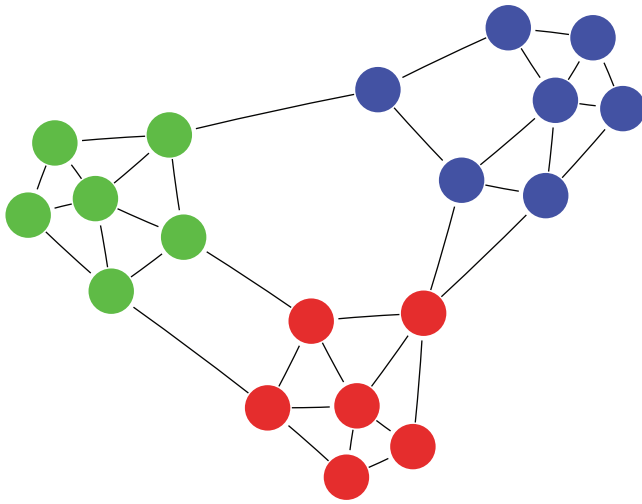
Figure 1. A cartoon depiction of a modular network with links between modules. Nodes of the same modules are identified by different colors. This network has a modularity (Louvain method) of $Q = 0.527$. Out of the 36 interactions, 31 are established within modules, and 5 between modules. This gives a Q_R value of 0.86, and $Q'_R = 0.72$.

species are more likely to interact together) to trait related[27] (pollination depends on the flower and insect having compatible morphologies, predators are constrained by the body-size of their preys). The interaction within these different families of mechanisms will drive heterogeneity in interaction strength[28]. Yet, the analysis of binary matrices (is there an interaction between a pair of species, or not), still has relevance to identify properties that are conserved across systems[29], especially given that one could argue that quantitative information on interaction strength is an additional level of information. The systems analyzed in this section are represented by their adjacency matrix, describing the presence or absence of an interaction. Bipartite networks have further been transformed into unipartite networks before analysis.

Data and analysis

I used the Louvain method[30] to detect modules, due to its rapidity and efficiency on large networks. The Louvain method works in two steps: first it optimizes modularity *locally*, through clustering of neighboring nodes. These clusters are, in the second steps, aggregated together, until modularity ceases to increase. This method is known to give values of modularity comparable to what is found using *e.g.* simulated annealing, and has been observed to give modules that have a functional relevance[30]. Once the partition is returned by the Louvain method, I recorded its realized modularity Q'_R, and its modularity Q (using the Newman and Girvan[31] measure).

For each network, I compared the values of Q and Q'_R on the empirical networks to their random estimate using a network null model. Because random networks will by chance (here meaning, as expected by networks having a given connectance and thus degree distribution, Poisot and Gravel[32]) display a modular (among other) structure, it is important to confront the empirical measures of both

Q and Q'_R to their random expectations. The null model is defined as follows. For each node n of the network, I measured its degree d_n, its number of successors (the number of node it links to, or generality in ecological terms, as *per*[33]) g_n, and its number of predecessors (the number of nodes that link to it, or vulnerability) v_n. In each random network, for each pair of nodes (i, j), the probability that i interacts with j is given by

$$P(i \rightarrow j) = \frac{1}{2}\left(\frac{g_i}{d_i} + \frac{v_j}{d_j} \right), \qquad (3)$$

and conversely for $P(j \rightarrow i)$. This null model allowed the generation of pseudo-random networks through a Bernoulli process (in each replicate, the occurrence of a link is randomly determined), with the same expected connectance, and the same expected distribution of degrees, generality, and vulnerability, as the original one (these properties are also conserved at the *node* level). For each of the 290 networks, 1000 pseudo-random replicates are generated. For each of them, the average value of Q_R and Q'_R are estimated along with their 90% confidence interval. When the empirical value lies outside the confidence interval, it can be assumed that the modular structure of the network is different than expected by chance.

Results and discussion

There is a strong, positive relationship, between the values of Q'_R and the values of Q (Pearsons's product-moment correlation coefficient, as implemented in R 2.15[34], $\rho = 0.64$, 288 d.f., $p < 10^{-6}$), *i.e.* networks for which a high modularity is detected tend to have relatively few between-module links (Figure 2). It is worth noting

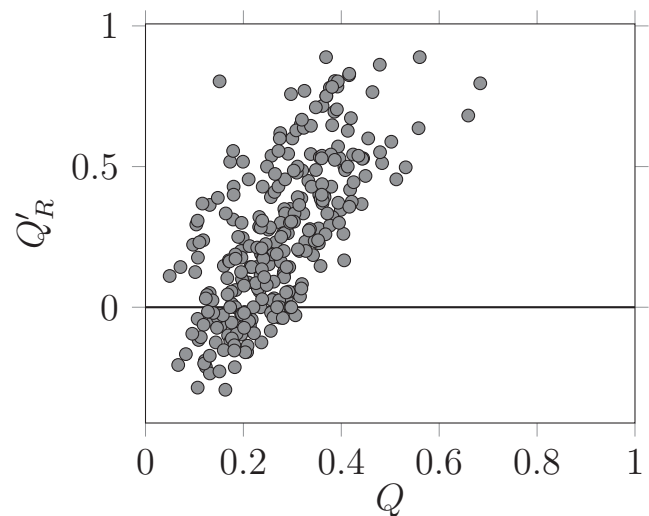**Figure 2. Relationship between the modularity of the best partition using the Louvain method and the *a posteriori* realized modularity.** There exists a strong, positive relationship between the two variables. Worth noting is the fact that, for some networks, the best partition resulted in *negative* versions of Q'_R, *i.e.* there were more interactions between than within modules. Each dot corresponds to a network.

that some Q'_R values were negative: in some cases, the best community division resulted in more interactions between than within modules. This result highlights why using an *a posteriori* measure is useful: other measures of modularity do not reveal the fact that there were more interactions between than within modules. In the dataset examined, most of the networks with a modularity lower than 0.2 had a negative realized modularity. This result suggests that discussing the modularity of such networks makes little sense, as their modules are not more densely connected, within a module, than other random collections of nodes within the graph. Q and Q'_R have different relationships with connectance (Figure 3). Increased connectance values resulted in lower modularity ($\rho = -0.61$, 288 d.f., $p < 10^{-6}$), but had no impact on Q'_R. This is a desirable property, as it allows easy comparison with the Q'_R values of networks with extremely different connectances.

There is a linear relationship between the deviation from random expectation of Q and Q'_R ($\rho = 0.78$, 288 d.f., $p < 10^{-6}$ – Figure 4). The deviations (respectively ΔQ and $\Delta Q'_R$) are calculated as the empirical value, minus the average of the values on the networks generated by the null model. As an example, a ΔQ less than zero indicates that the empirical network is less modular than expected by chance. Confidence intervals for the average of the null models were typically very narrow (not represented in the figure to avoid cluttering – see associated original dataset), probably owing to the fact that the null model is restrictive on the type of networks which are generated. It is worth noting that for some networks, the diagnostic of the null model analysis is conflicted. In a vast majority of the situations, this corresponds to networks having a lower modularity than expected by chance, yet having a higher realized modularity (dots in the upper left corner of Figure 4). In this type of situation, whereas one would usually conclude that the networks are not significantly modular, the identified modules are nonetheless more densely connected (internally) than they are with the rest

of the network. Because the dataset presents these contrasted situations, it allows us to understand how the measure reacts to different network structures. Depending on whether the true modularity, or the realized modularity, is the most relevant metric of the processes studied, the interpretation of the null models for these networks will be different.

Relationships between raw and realized modularity

1 Data File

http://dx.doi.org/10.6084/m9.figshare.156237

Finally, for the unipartite network dataset, I compare the results of three alternative methods of community detection (the walktrap, spinglass, and edge-betweenness methods, as implemented in the *igraph* library). For each of the unipartite networks, I computed the value of Barber's Q, and Q'_R, on the best partition found. The strong correlation between Q and Q'_R were observed for the spinglass method ($\rho = 0.61$, 165 d.f., $t = 10.02$), and the weakest for the edge-betweenness method ($\rho = 0.04$, non-significant at $\alpha = 0.05$). The walktrap algorithm gave results in between ($\rho = 0.489$, 165 d.f., $t = 7.20$). For both the walktrap and edge-betweenness methods, several networks had negative values of Q'_R, which indicates that the "best" community partition had more links between than within modules. The spinglass method had, by contrast, less than 8% of all networks with values of Q'_R lower than 0, meaning that this algorithm should be prefered when one wants to group nodes in densely connected clusters. This result reinforces the statement made by Thébault[7], i.e. that several modularity optimisation methods will return best modular structures that widely differ in their properties; thus, there is a need for *a posteriori* comparison of these outputs.
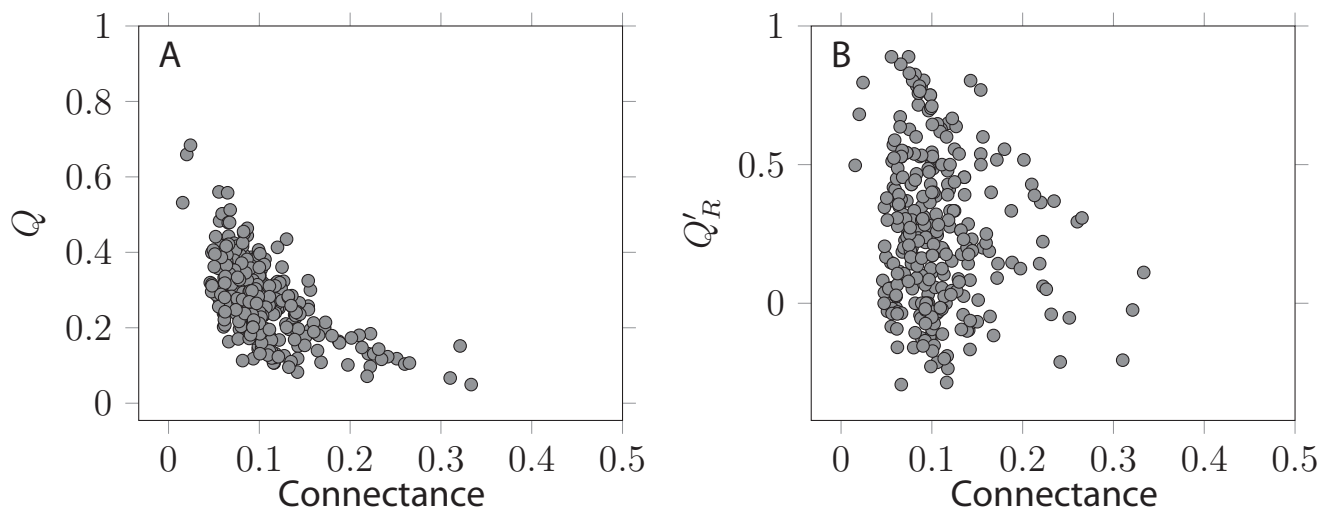

Figure 3. Relationship between the two measures of modularity and network connectance. A. *Q* is negatively affected by connectance, *i.e.* densely connected networks are more likely not to be modular. **B.** *Q'$_R$* is not affected by connectance, allowing to use it to compare different networks. Each dot corresponds to a network.
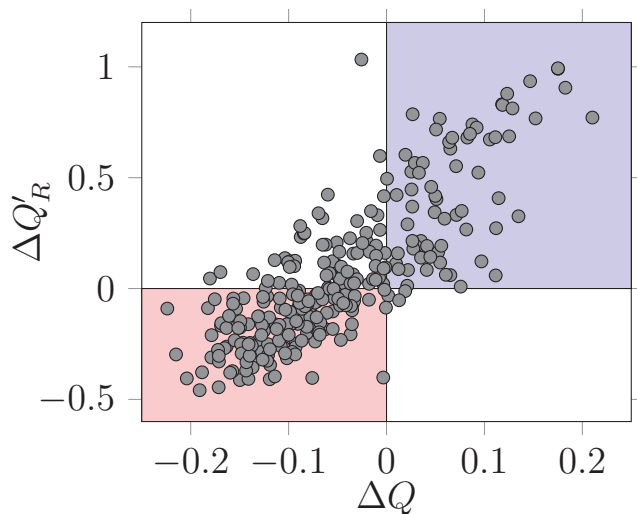
Figure 4. Linear relationship between the deviation from random expectation in Q and Q'_R. Networks in the red area are detected as being less modular than expected both by Q'_R and Q, while networks in the blue area are detected as being more modular. Although the agreement between the two measures is good (see main text for statistics), some networks are detected as having a higher than expected realized modularity Q'_R, despite a lower than expected modularity Q. Each dot correspond to a network.

Conclusions

The Q'_R measure presented here allows the estimation of the proportion of interactions established between different modules in a network. This measure can be analyzed much in the same way as other measures of modularity, but is applied *a posteriori*. As such, it can help choose the "best" community partition according to the property of the network that one wants to maximize. For example, choosing the partition giving the lowest Q'_R can help identify which species are more likely to act as connectors between different modules. Ultimately, this information may have some practical relevance as a decision tool. Saavedra *et al.*[5] showed that different nodes contribute differently to overall network properties. In a context in which networks are increasingly being used as management tools to adress *e.g.* conservation or pest management[8], knowing the realized modularity, and developing methods to estimate which species have the highest impact on it, can allow the design of efficient policies to maximize, or decrease, the ability of network modules to interact.

References

1. Espinosa-Soto C, Wagner A: **Specialization can drive the evolution of modularity.** *PLoS Comput Biol.* 2010; **6**(3): e1000719.
   PubMed Abstract | Publisher Full Text | Free Full Text

2. Bauer-Mehren A, Bundschus M, Rautschka M, *et al.*: **Gene-Disease Network Analysis Reveals Functional Modules in Mendelian, Complex and Environmental Diseases.** *PLoS One.* 2011; **6**(6): e20284.
   Publisher Full Text

3. Fortuna MA, Bonachela JA, Levin SA: **Evolution of a modular software network.** *Proc Natl Acad Sci U S A.* 2011; **108**(50): 19985–19989.
   PubMed Abstract | Publisher Full Text | Free Full Text

4. Olesen JM, Bascompte J, Dupont YL, *et al.*: **The modularity of pollination networks.** *Proc Natl Acad Sci U S A.* 2007; **104**(50): 19891–19896.
   PubMed Abstract | Publisher Full Text | Free Full Text

5. Saavedra S, Stouffer DB, Uzzi B, *et al.*: **Strong contributors to network persistence are the most vulnerable to extinction.** *Nature.* 2011; **478**(7368): 233–235.
   PubMed Abstract | Publisher Full Text

6. Carstensen DW, Dalsgaard B, Svenning JC, *et al.*: **Biogeographical modules and island roles: a comparison of Wallacea and the West Indies.** *J Biogeogr.* 2012; **39**(4): 739–749.
   Publisher Full Text

7. Thébault E: **Identifying compartments in presence-absence matrices and bipartite networks: insights into modularity measures.** *J Biogeogr.* 2013; **40**(4): 759–768. Ed. by Joseph Veech, n/a-n/a.
   Publisher Full Text

8. Chadès I, Martin TG, Nicol S, *et al.*: **General rules for managing and surveying networks of pests, diseases, and endangered species.** *Proc Natl Acad Sci U S A.* 2011; **108**(20): 8323–8328.
   PubMed Abstract | Publisher Full Text | Free Full Text

9. Newman ME: **Modularity and community structure in networks.** *Proc Natl Acad Sci U S A.* 2006; **103**(23): 8577–82.
   PubMed Abstract | Publisher Full Text | Free Full Text

10. Zhang XS, Wang RS: **Optimization analysis of modularity measures for network community detection.** *The Second International Symposium on Optimization and Systems Biology.* Lijiang, China, 2008; 13–20.
    Reference Source

11. Barber MJ: **Modularity and community detection in bipartite networks.** *Phys Rev E Stat Nonlin Soft Matter Phys.* 2007; **76**(6 Pt 2): 066102.
    PubMed Abstract | Publisher Full Text

12. Liu X, Murata T: **Community detection in large-scale bipartite networks.** *Trans Jpn Soc Artif Intell.* 2010; **5**, 184–192.
    Reference Source

13. Marquitti D, Maria F, Roberto GP Jr, *et al.*: **MODULAR: Software for the Autonomous Computation of Modularity in Large Network Sets.** 2013; **1304**(2917): arXiv e-print.
    Reference Source

14. Wiederhold G: **Mediators in the architecture of future information systems.** *The IEEE Computer Magazine.* 1992; 1–36.
    Reference Source

15. Leskovec J, Lang KJ, Dasgupta A, *et al.*: **Statistical properties of community structure in large social and information networks.** *Proceeding of the 17th international conference on World Wide Web - WWW ' 08.* New York, New York, USA: ACM Press, 2008; 695.
    Publisher Full Text

16. Hartwel LH, Hopfield JJ, Leibler S, *et al.*: **From molecular to modular cell biology.**

Nature. 1999; **402**(6761 Suppl): C47–52.
PubMed Abstract | **Publisher Full Text**

17. Rosvall M, Bergstrom CT: **Maps of random walks on complex networks reveal community structure.** *Proc Natl Acad Sci U S A.* 2008; **105**(4): 1118–23.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Gravel D, Massol F, Canard E, *et al.*: **Trophic theory of island biogeography.** *Ecol Lett.* 2011; **14**(10): 1010–1016.
PubMed Abstract | **Publisher Full Text**

19. Poisot T, Canard E, Mouillot D, *et al.*: **The dissimilarity of species interaction networks.** *Ecol Lett.* 2012; **15**(12): 1353–1361.
PubMed Abstract | **Publisher Full Text**

20. Fortuna MA, Stouffer DB, Olesen JM, *et al.*: **Nestedness versus modularity in ecological networks: two sides of the same coin?** *J Anim Ecol.* 2010; **79**(4): 811–817.
PubMed Abstract | **Publisher Full Text**

21. Stouffer DB, Bascompte J: **Compartmentalization increases food web persistence.** *Proc Natl Acad Sci U S A.* 2011; **108**(9): 3648–3652.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Flores CO, Valverde S, Weitz JS: **Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages.** *ISME J.* 2013; **7**(3): 520–532.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Thébault E, Loreau M: **Food-web constraints on biodiversity-ecosystem functioning relationships.** *Proc Natl Acad Sci U S A.* 2003; **100**(25): 14949–14954.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

24. Flores CO, Meyerb JR, Valverdec S, *et al.*: **Statistical structure of host-phage interactions.** *Proc Natl Acad Sci U S A.* 2011; **108**(28): E288–297.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

25. Bluthgen N, Menzel F, Bluthgen N: **Measuring specialization in species interaction networks.** *BMC Ecol.* 2006; **6**: 9.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Canard E, Mouquet N, Marescot L, *et al.*: **Emergence of Structural Patterns in Neutral Trophic Networks.** *PLoS One.* 2012; **7**(8): e38295.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

27. Bartomeus I: **Understanding Linkage Rules in Plant-Pollinator Networks by Using Hierarchical Models That Incorporate Pollinator Detectability and Plant Traits.** *PLoS One.* 2013; **8**(7): e69200.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Berlow EL, Navarrete SA, Briggs CJ, *et al.*: **Quantifying variation in the strengths of species interactions.** *Ecology.* 1999; **80**(7): 2206–2224.
Publisher Full Text

29. Dunne JA: **The Network Structure of Food Webs.** In: *Ecological networks: Linking structure and dynamics.* Ed. by Jennifer A Dunne and Mercedes Pascual. Oxford University Press, 2006; 27–86.
Reference Source

30. Blondel VD, Guillaume JL, Lambiotte R, *et al.*: **Fast unfolding of communities in large networks.** *Journal of Statistical Mechanics: Theory and Experiment.* 2008; **2008**(10): P10008.
Publisher Full Text

31. Newman MEJ, Girvan M: **Finding and evaluating community structure in networks.** *Phys Rev E Stat Nonlin Soft Matter Phys.* 2004; **69**(2 Pt 2): 26113.
PubMed Abstract | **Publisher Full Text**

32. Timothée P, Dominique G: **When is a network complex? Connectance drives degree distribution and emerging network properties.** *PeerJ Preprints.* 2013; **1**: e50v3.
Publisher Full Text

33. Schoener TW: **Food webs from the small to the large.** *Ecology.* 1989; **70**(6): 1559–1589.
Publisher Full Text

34. **R. Development Core Team.** *R: A Language and Environment for Statistical Computing.* Vienna, Austria, 2008.
Reference Source

Open Peer Review

Current Referee Status: ☐? ☑ ☐?

Version 3

Referee Report 08 April 2014

☑ **Daniel Carstensen**
Department of Bioscience, Universidade Federal de São Paulo, São Paulo, Brazil

I have no further issues with this submission.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 06 January 2014

☐? **Jochen Fründ**
Agroecology, Georg-August-University of Göttingen, Göttingen, Germany

I appreciate the improvements to the article done during the revision. Both the abstract and the main text now describe better where the article is going and explain the purpose of the new metric. Several minor concerns have been overcome and questionable points have been clarified.

However, my main reservations largely persist. This includes:

1. I question whether Q'_R is generally useful for choosing which community partition to use. For choosing the best partition for a particular purpose, I think it would be most efficient to use the metric of choice during the optimisation and not *a posteriori*. In my view, a simple *a posteriori* metric is mostly useful for presenting and describing modularity. For this purpose, I would prefer Q_R over Q'_R because it makes even less assumptions.

2. Fig. 4 still suggests to me that most of the empirical networks used here are not more modular than expected by chance (more negative than positive deltaQ values) and I am thus still uncertain how suitable the dataset is for exploring modularity.

3. Less fundamentally, in the comparison of the different algorithms from the *igraph* package I am still missing the percentage of Q'_R that were below 0 for the walktrap and edge-betweenness method, in order to compare it to the 8% for the spinglass method (currently it only says "several").

In conclusion, I think the metric Q_R presented here can be helpful, but should be applied with caution.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Author Response 06 Jan 2014

Timothée Poisot, Université du Québec à Rimouski, Canada

I believe these points were all addressed in the revision and associated replies. I would be willing to provide more arguments, but the referee is merely stating their feeling or ideas, not making any factual criticism of the paper. In these circumstances, it is very hard for me to decide what to revise.

Competing Interests: No competing interests were disclosed.

Version 2

Referee Report 05 December 2013

doi:10.5256/f1000research.2230.r2678

Jochen Fründ
Agroecology, Georg-August-University of Göttingen, Göttingen, Germany

This article suggests a simple intuitive measure of network modularity. The suggested measure, Q'R, is related to established measures but calculated slightly differently. It is proposed as an *a posteriori* measure, which means it is not suggested to be used for assigning nodes (e.g. species) to modules, but only to evaluate partitions based on other methods that calculate modularity, identify modules and assign nodes to them.

In principle I welcome the suggestion of a simple, easy to interpret measure. The metric suggested here can help presenting modularity. I see that the amount of between-module links in relation to within-module links may have direct functional consequences. Established measures of modularity measure basically the same, they only correct for the expectation of within-module links in non-modular random networks in a different way.

However, I have a number of serious concerns making the study potentially misleading. These concerns include reservations about whether the analyses address the study aim, whether the dataset is suitable for testing modularity, how the proposed metric is interpreted and that it is suggested as an *a posteriori* measure.

General issues:

The study aim set out in the abstract and introduction, to compare different methods and approaches

detecting modularity, is not reflected by the analyses. Neither is a functional meaning of the new metric demonstrated to support the case that the metric can be used to evaluate other methods, and decide which method to use. The abstract claims that new insights are gained about the modularity of the food webs in the empirical example dataset, but I struggle to find these new insights.

A paragraph added during the revision does some comparison, but it is not integrated with the rest of the paper and neither does it demonstrate the usefulness or added value of the new measure. For the most part, the paper rather compares values of one standard measure of modularity with values for the new metric in an example dataset of 290 unipartite and bipartite food webs. Using one method of module assignment, the paper shows how the two metrics are related to each other, to randomizations of the webs, and to network connectance. The meaning of these relationships for the study purpose is unclear.

Importantly, the usefulness of the empirical dataset for evaluating modularity methods is questionable. Typically, studies proposing modularity methods test them on networks of known modularity. However, the nullmodel analysis brings to attention that the vast majority of the networks used here are not more modular (based on Q) than expected by chance, and even less might be significantly modular. This means that this study tries to evaluate modularity methods on networks that are mostly not modular. This questions the value of the whole study and calls again for external information for validation. If networks are not modular, then the practical value of measuring modularity becomes negligible: the variation in module assignment in networks not significantly modular is probably much less worrying than failure to detect a known modular structure (which is not given here).

The straightforward interpretation of QR is changed in Q'R, the version the author describes as being scaled between 0 and 1, only to report negative values later on in the paper (Fig. 2). For networks of unknown modularity, Q'R can actually take values between -1 and 1. Furthermore, the notion that negative values of Q'R detect cases of spuriously significant modularity is not generally correct. The threshold of meaningful modularity depends on the purpose and may be above or below Q'R =0.

This brings me to a fundamental problem with the study – what is modularity and why should it be measured? The author states that the new metric "makes no assumption about what modularity is". If this is really the case, then there is no point in defining a measure for it. To be useful, an assumption about what is being measured has to be made. This questions the claims and even policy recommendations made by the author. The difference in concepts and goals is likely a major reason why previous methods differ (e.g. unipartite vs. bipartite modularity suggested by Guimera *et al.*, 2007). Only when a concept of interest is defined can methods be compared in how well they serve the purpose.

I am not sure how useful the whole idea of an *a posteriori* measure is. The author stresses that the measure is not aimed at maximizing modularity in an algorithm, but just to select which algorithm to use. This is not convincing: either the measure reflects the property of interest, then it should be maximized in the first place in the algorithm to find the best partition; or it is not a sensible measure, then it cannot be used for selection at all. The approach proposed here appears very inefficient and almost certainly not to give the best partition. Furthermore, any measure of modularity could be calculated *a posteriori* or during modularity optimization. The description of this index specifically as an *a posteriori* measure gives no real sense, without additional data or simulations showing that it is more meaningful than others. If the functional meaning was demonstrated, there could be some value in using it *a posteriori* for those who don't have access to source algorithms.

Alternative methods (algorithms) paragraph:

As said above, this is not connected to the rest of the paper. Of course it improves the paper to consider alternative methods for module assignment. However, this paragraph has several shortcomings. First, restricting this analysis to the unipartite networks makes it hard to compare to the other results. Second, it remains unclear why this focuses on the correlation between Q and Q'R. The modularity of the partition returned by each method would be compared more directly by comparing the values of Q or Q'R between methods. At the moment, for judging the three methods the reader is just left at guessing that "several" negative values (for methods walktrap and edge-betweenness) are more than "less than 8%" (method spinglass). Third, it looks like an inconsistent ad-hoc addition: citations for the methods and the igraph library (package) are missing, the methods are not mentioned before or described and correlation coefficients are called r here but rho above.

Null model:

The description of the nullmodel leaves unclear whether the connectance and degree are fixed exactly or just determine the expected value probabilistically. Moreover, the nullmodel is discussed as reflecting "chance". Given that many links likely remain unobserved in ecological network datasets, a reasonable simulation of chance should ideally consider detection probability. Binary network data (e.g. the data used here) are often problematic: ecological network data are virtually always just samples of all realized interactions – this likely applies to the examples used here (even expert opinion may be influenced by observation bias). This can lead to strong biases in measures of network structure between the real web and its sample, but these problems are ignored here. As the simulations are called "pseudo-random", they may be acceptable within the constraints of binary data – which then casts questions about the usefulness of the test dataset for the study purpose (see above).

Unipartite vs. bipartite webs:

It should be better explained how the different data types were handled with the same methods. Bipartite networks have additional (conceptual) ambiguities in how modularity should be calculated, which may be a core reason for discrepancy between modularity methods (Guimera *et al*., 2007, Thébault, 2013). To be able to interpret the data better, it is warranted to present or identify the unipartite and bipartite webs separately in the graphs and results.

Minor points:
- More information on the datasets should be provided; the bipartite dataset is not even found in the reference provided for it, but must be traced back several steps to the original reference.

- Why is an algorithm chosen that is recommended for large networks (many thousands to millions of nodes, Blondel *et al.*, 2008) when the webs analyzed here have less than 200 nodes?

- Without defining the purpose or demonstrating the functional meaning of Q'R, it is difficult to know whether no correlation with connectance is desirable or not.

- "Results" should actually be entitled "Results and Discussion".

- The terminological differentiation between true modularity and realized modularity is confusing.

Overall, the study is inconsistent and doesn't live up to its promises. A study evaluating modularity measures should look at additional information to validate it (especially if it is not a formal comparison of multiple metrics). As shown by previous papers on the topic, this additional information could be the

correspondence to biological traits in empirical networks (e.g. Martín González *et al.*, 2012), the detection of build-in module structure in simulated networks (e.g. Thébault, 2013) or the demonstration of functional consequences (e.g. by a model). To be useful, the study should be put on a more solid foundation.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 02 September 2013

doi:10.5256/f1000research.2230.r1638

Daniel Carstensen
Department of Bioscience, Universidade Federal de São Paulo, São Paulo, Brazil

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Version 1

Referee Report 11 July 2013

doi:10.5256/f1000research.1000.r1055

Daniel Carstensen
Department of Bioscience, Universidade Federal de São Paulo, São Paulo, Brazil

The aim of the author is interesting and relevant. I am intrigued by the development of a method to quickly evaluate different modularity measures, and an *a posteriori* method might well be a good solution.

Overall the manuscript is generally well written. However, it is not clear to me how much is gained with this approach. Poisot only uses one method to detect the module configuration (Louvain 2008) and one method to calculate the modularity (Newman & Girvan 2004). It would be interesting to explore if the Q_R differs markedly when applied to the results of different methods. It would also be good to see what existing modularity measures do when optimizing modularity; do they minimize between module links? This is why Q_R measures and the strong correlation in Figure 2 is not surprising. What is more interesting about Figure 2 is that it shows that below a certain value of Q (~ 0,2 ?) it is not sensible to talk about modularity even if the empirical data is more modular than a random system. In such cases, the presented method seems useful to evaluate results.

Other minor revisions

- An earlier reference could be used for the use of modularity in biogeographic networks instead of Cummings *et al.* 2010 (reference 6). Cummings *et al.* does not handle modularity. THe author should onsider citing Carstensen & Olesen 2009.
- In the 'Data and analysis' section the statement starting on line six in this paragraph needs a reference.
- Null model: What is meant by generality/successors and vulnerability/predecessors?

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 26 June 2013

Carsten Dormann
Biometry and Environmental System Analysis, University Freiburg, Freiburg, Germany

The proposed index of modularity is of striking simplicity - and thus likely to be prone to artifacts. In the opening paragraph, Poisot forgot to mention that random networks are also modular. Thus, a $Q\_R > 0$ means, in itself, nothing, as Poisot rightly assumes when employing a null model.

The typically log-normal abundance of species in nature will introduce apparent structure into networks, even if the links simply reflect probabilistic interactions (i.e. any species interacts more with a common than a rare species). Thus, without a null model correcting for number of species, for their abundance and for the possibility of random networks also being modular, any index may report only spurious, artefactual results. Poisot uses a null model, but because his example data are binary networks (containing no information about the strength of a link), the best he (or anyone) can do is to use a null model based on degrees, which is only a very poor reflection of the actual abundance. Given that often more than a third of the species in a network are singletons, I believe that their contributions to modularity are overemphasized by any binary measure.

Suggested Revisions:

1. Simulate networks (ideally weighted ones) and compare their Q_R values to quantitative null models. How much information does Q_R (and Q) actually contain?
2. Comparison of Q_R not only with Q and connectance but also with other network metrics, such as linkage density or dependence asymmetry (and particularly those with a more or less clear ecological interpretation, such as H2'). The question, again, is: what does Q_R provide in addition to current metrics?
3. Gain in ecological knowledge (which follows from 1. and 2.): If there is additional information, what does it mean? Which ecological features (specialization, number of functionally similar species, number of trophic level, number of habitats sampled etc.) contribute to Q_R? (For example along the lines of Pocock *et al.* 2012, who work on different types of sub networks put together into one large, or Clauset *et al.* 2008. Are these different sub networks identifiable as modules? If so, what does Q_R stand for?)

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.