

Interactive comment on “Network Analysis of the American Geophysical Union’s Fall Meetings” by Tom Narock et al.

Tom Narock et al.

tnarock@ndm.edu

Received and published: 21 December 2018

Mark,

Thank you for the time and effort you put into reviewing our paper. We found your comments valuable and provide additional avenues of study we had not considered. Below, we address each comment individually.

- Most of the analysis relies on absolute numbers of presentations, components, pairwise comparisons, etc., but these really needed to be considered relatively. There is little consideration of how the different sizes of the sections, the longevity of authors, and the overall growth of AGU might affect the analyses. For example, the hydrology section probably has the greatest connection with the informatics section simply be-

C1

cause it is the largest session, not that it is relatively more engaged with informatics than other sessions. This might also make the Natural Hazards growth even more notable. The authors should also consider the limits of self-classification when discussing keyword use, not to mention how keywords may have changed over time.

These are good points. We agree that normalizing the numbers by section size would highlight additional insights. We'd like to make this change in the revised paper as well as present additional data on author longevity. We agree that keyword change could also play a role. However, we do not have data on when keywords changed or when new keywords were added (other than the one example given in figure 8).

- The authors also need to decide if they really want to present this as a scientometrics paper or not. On the one hand, they present a clear argument for measuring and assessing data around collaboration in the Earth and space sciences. On the other hand, they are understandably cautious in their assessment because of the limitations of the data. They seem to recognize the limits of scientometrics when they explicitly state that they “are not advocating for any new sort of new metric” (Line 317), but they don't really explore that. It is worth considering this a “scientometrics” paper, but then the authors must more explicitly grapple with critiques of scientometrics (and infometrics and bibliometrics) and how they are dealing with known shortcomings. The paper as currently written might be considered as an exploration of possible approaches to developing scientometrics. Personally, I would simply avoid the term. I find it fraught and misleading.

We were using the term “scientometrics” in the general sense of the “science of science” and understanding how science operates - and can be improved. We agree with your latter point that this paper is an “exploration of possible approaches to developing scientometrics” within the Earth and space science. To this end, we would like to keep the term, but revise the title and text to reflect this exploration and make it clear that we are not at a point of having detailed metrics. We have additional suggestion for how our work can be operationalized, but we are not currently at such.

C2

- It is excellent that the authors provide ready and open access to the data and software used in the study. That said, the data and software should be explicitly cited and listed in the reference list (even if it is self-citation) . This is the new norm in Earth and space sciences. See <http://www.copdess.org/enabling-fair-data-project/commitmentto-enabling-fair-data-in-the-earth-space-and-environmental-sciences/>

Excellent point. We received this comment from other reviewers as well and we are happy to oblige.

- The title is a bit bland and vague. Maybe borrow from the conclusion and say something like 'Identifying and improving collaborations within AGU using network analysis'.

Thank you. We certainly want to avoid bland and vague :) We will give this more thought and update in the revision.

- The abstract reads like an introduction rather than a summary. Please summarize major findings and insights in the abstract and say less about what AGU is and does.

Will do. Other reviewers have made similar comments. We will update both the title and the abstract.

- Line 74: Good point about ORCID. Perhaps AGU could allow members to access their own data and add orcid and resolve conflicts.

Agreed. This would be ideal; however, we have not had success in advocating for this.

- Line 140 and table 1: Please indicate when the different sections were established and if there have been any mergers or changes in the sections over the years. Ideally, you need to account for different time periods in making comparisons. The size of the membership and number of presentations in each section must also be considered when making comparisons.

Will do.

- Perhaps I misunderstand, but the following two statements seem to contradict each

C3

other. I believe you mean the density is not lower on Line 89 Line 89: "We know that the actual values for network density and connected components are not higher than the values reported here, and they would likely be a bit smaller had we been able to uniquely identify all authors in our dataset."

We did. Thank you for catching this typo. We'll fix this.

- Figure 4 would be more useful if it was shown as ratios or somehow dealt with the huge variance in numbers in each section. Which section is most interconnected relatively?

We received a comment from another reviewer to normalize by section size and show figure 4 as a color matrix where the rows and columns are the sections and color in each cell indicates the number of connections. We think this would be a good way of showing this data and we plan to revise accordingly.

- Section 3.4. How has the keyword hierarchy changed over time? Scenario 3 highlights how this could be significant.

Unfortunately, we do not have data on this. We are only aware of scenario 3 because the first author was a member of the section's leadership at the time.

- Figure 5: When did NH start? The lack of connection between 2013 and 2016 is as interesting as the growth from 2016 to 2017. The one year jump suggest an event driven cause.

We do not currently have data on when each section started. However, we will inquire with AGU and track down these numbers for the revision.

- Figures 5-7: Why do they have different time scales? Why not the full 17 years?

That is an error on our part. We let the analysis software auto-scale the axes. This will be corrected in the revision.

- Paragraph starting at Line 246: The authors suggest an uptick in use of the education

C4

keyword by PA and U. This does not appear significant in the figure, especially since it's such a short period. Please explain.

Our intention was to show that U use of education keywords increased from less than 10 to 100 in 2 years while PA use of those same keywords took 8 years to increase from less than 10 to 100. We consider the U scenario significant and the PA scenario an example of gradual growth. However, we see your point and will revise this with better statistics to quantify "significant" as well as edits to the text.

- Scenario 3 is interesting but speculative. Some discussion of the limits of self-categorization is warranted.

Point taken and without keyword history data we can not defend this further. Scenario 3 should be removed.

- Line 288: Why should single authors be considered worrisome? Bucking a trend is not necessarily a bad thing.

There is nothing wrong with single author presentations. Yet, we do feel that science conferences, such as AGU, should not be all about dissemination of science. We feel they should also be places to identify new collaborations. AGU seems to be a supporter of this as well with their various networking and collaboration events. Given this, we would have expected the rate of new edges in the graph to keep pace with the rate of new nodes in the graph. This does not seem to be the case. We feel that quantifying these rates - as best we can given the errors in the data - can lead to discussions regarding the effectiveness of current collaboration and cross-disciplinary efforts. Are scientists resistant to these efforts, not aware of them, or are the efforts not effective? We can not answer these questions with our current dataset; yet, we think looking at single author presentations and the rates at which nodes and edges are added are the first steps.

- Line 295: "We could make this process more proactive by providing section leadership

C5

with connected component data and encouraging connections between specific AGU members. This could range from informal networking events to suggesting session coconveners." As chair of the Program Committee for Informatics section next year I am interested in how we could do this. It would be nice if you could expand on the idea a bit more. One concern is that we might end up simply reinforcing existing connections rather than expanding the network.

Certainly. We have been thinking about this since the paper was first submitted. We have some additional thoughts we can add to the revision.

- Section 4.1 is interesting. It may be worth noting that AGU has already started clustering sections into "neighborhoods" of related interest. The component analysis could help here, if it could be done in advance of each meeting (Don't want to just follow last years pattern). This is another area where the size of the section would also need to be considered.

Agreed and we can add text to that effect. Although, AGU has been reluctant to release meeting data prior to the meeting. Making this point in the paper may help change their mind.

- Line 311: you need to consider how long the author has been participating in AGU in conjunction with the use of the keyword. Old timers are more likely to have used the keyword repeatedly.

That was our intention here. We were showing raw counts so that new attendees could see historically which authors presented the most on a given topic. After reading your comment, we think it would be better to provide an "old timer" recommendation as well as highlight early career researchers as well. Give new attendees both perspectives in the recommendation.

- Thank you for the appendix. There appear to be many more interesting scenarios in there. Improved graphical design would make them easier to detect. The coloring of

C6

the lines for the different sections in the graphs should be consistent across figures if possible. It is better to label the actual lines instead of providing a legend or at least list the legend in the same order as the lines appear (top to bottom). Be mindful of the colorblind.

Excellent point. Thank you. We will gladly fix this.

- *Technical corrections:* Line 44: delete “only” Line 13: delete “self-“ Line 383: change ‘. Yet if’ to ‘, yet it’

Thank you for catching these. We will fix all of them.

Interactive comment on Geosci. Commun. Discuss., <https://doi.org/10.5194/gc-2018-15>, 2018.