Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-113-SC1, 2018

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Optimal Design of Hydrometric Station Networks Based on Complex Network Analysis" *by* Ankit Agarwal et al.

F. Cioffi

francesco.cioffi@uniroma1.it

Received and published: 31 May 2018

The paper is novelty and innovative and in my opinion it should be published. I believe that the 'complex network paradigm' offers a very useful reference framework that can stimulate the development of new approaches and methods to solve relevant issues in the field of hydrology. Therefore, the exploration of new methods within such framework, as those proposed in the present manuscript, should be encouraged. Despite this, I think that the quality of the paper should be improved in some points, because in the present form could be misunderstood, especially from people who are more familiar with the traditional hydrological approaches.

Mathematical critical points:

Printer-friendly version

Discussion paper

A) I think that there is an error in the formula (3) since I was not able to calculate the values of WDB, for the simple networks of figure 2, reported in table 2. In each case an effort to provide an hydrological meaning of such metric could help the readers to better evaluate the worth of the proposal. B) Betweenness in figure 2, coherently with the definition given in paragraph 2.2, it should be 4(12) – 5(12) – 6(10) instead of 4(24) 5(24) 6(20). C) In paragraph 4.2, in order to select the most synchronized nodes of the network, authors assume the 95th percentile of the values of adjacency matrix elements. May authors provide a justification of such assumption? Just a suggestion: what about if you use all that edges that are significant? An edge is significant if the synchronization value exceeds the 95th percentile of the synchronization obtained by two synthetic variables that have the same number of events positioned randomly in the time series. Authors can run synchronization on 100 random time series (couples) and you get the 95th percentile of synchronization for these synthetic values. D) In the appendix authors use an event synchronization formula which is a modified version of Quiroga et al 2002. However this formulation is still wrong (Q can exceed 1, as stated in Conticello et al 2018. This is due to double count of the events.). If authors use a dynamical tau, you should use the formulation of Kreuz et al 2015, 2016 (Spiky), if you use a defined tau, you should use Conticello et al 2018. (In this case it seems it is dynamical, so Kreuz will be fine).

Physical Interpretation critical point:

D) I have a doubt: what about single nodes which are not synchronized with others, but due to a particular location or specific hydrological conditions have a particular informative content in representing a spatial or temporal variability of the region considered? Maybe, in the case of a very dense network this is not the case, but this issue should be useful to clarify. Let's make a simple example: look at the figure below (attached). You have 10 nodes. The maximum number of undirected edges you can have is $\frac{1}{2}$*(10x10)-10 = 40 (1/2 because it's symmetric, 10x10 is n_nodes x n_nodes, -10 because you don't want to get self synch). In this case, if you want to retain those edges that exceed

Printer-friendly version

Discussion paper

95th percentile you get only 2 edges. With this configuration you will drop a lot of stations. Maybe I misunderstood and you want to retain the 95th percentile of all the 100 edges (you should explain clearly in the paper in order to make it reproducible). In any case, following your optimization approach, your network improves when you drop 1-2-3-8 nodes. Your efficiency grows because the area that you are describing it is getting smaller. With this approach you lose all the unique information coming from parts of the areas that need more station. So, from my point of view, maybe the stations that you should drop are 4-6-9-10 because they are more synchronized with 5 and 7, and you need only these stations to represent these subnetworks. I think the betweennes (and WDB) should be used to synthesize data and reduce the uncertainty due to the curse of dimensionality.

Description and Style critical points:

E) in the introduction there is just a list of the classical approaches used for the design of hydrometric networks. These approaches have the goal to identify the optimum number and locations of measurement points able to allow the most reliable possible representation of the spatial and temporal variability of the hydrological variable observed. Since a completely new method is proposed, and not a simple improving of existing ones, some justifications could be provided: a) why is it necessary, or opportune, to explore new approaches offered by the application of the complex network paradigm ; b) what are the potentiality of such more general approach to hydrology. In the specific case faced in the manuscript, some issues about the limitations or drawbacks of the classical approaches could be better underlined, as for instance: they are too much complex, empirical, based on not fully consistent hypothesis, etc, etc. Furthermore, is it possible to emphasize that the proposed approach, based on the complex network paradigm, is potentially more promising than existing ones? F) the metrics of the complex network reported in table 1 - degree, betweeness, etc - are not clearly explained to allow people who are not familiar with complex network, to understand well their meaning. G) In paragraph 3.3.1 the Decline Rate of Network Efficiency is introduced. I had some difficulties to understand the meaning of such variable. N is not defined. I try to read the cited paper (Liu 2016) but it use exactly the same words, and this paper in turn cites a paper in Chinese language. How do you measure dij ? i.e. is it the shortest path between nodes ni and nj equal to the sum of the links between ni and nj or it is the number of the shortest paths? Can any hydrological meaning be attributed to such measurement of network efficiency. H) A figure about the network of the German hydrometric stations, obtained by the application of the method, could help to understand how the proposed method works. I figure out that likely more independent networks should arise from the application of such approach because the 95th percentile threshold of the values of adjacency matrix elements .

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-113, 2018.