Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-148-RC2, 2020

Hydrology and Earth System Sciences Discussions
Open Access

HESSD

Interactive comment

Interactive comment on "Objective functions for information-theoretical monitoring network design: what is optimal?" *by* Hossein Foroozand and Steven V. Weijs

Anonymous Referee #2

Received and published: 17 June 2020

This paper discussed the objective functions in information theory-based hydrometric network design problems and suggested a backward greedy approach instead of other optimization methods. While I was interested in reading this paper because taking more correct, reasonable, and meaningful objective functions and proper optimization techniques is very important in the network design using information theory, unfortunately, I couldn't get the answer of the question in the title, "what is optimal?", throughout this paper.

1. The authors argue that a higher amount of redundant information should be preferred because it reinforces the robustness of the network. However, this is not a

Printer-friendly version

Discussion paper

general statement but can be applied only for a specific condition. World Meteorological Organization has recommended the minimum network densities, which do not represent the optimal number of stations, rather they are suggested to avoid any serious deficiency in water resources management. That is, just meeting the guideline couldn't be sufficient while there is no clear definition of sufficient density. However, many regions in the world, such as developing countries, do not meet even the WMO guideline. Even in the study area in this paper, Brazos River Watershed, the drainage area is about 116,000 sq.km. and there are 12 USGS stations considered as the existing stations. The current network density becomes 9,667 sq.km per station while the minimum network density of the WMO guideline for the Interior Plains is 1,875 sq.km., which is more than five times sparser network. Considering the WMO's minimum network is the baseline not to lose critical information, it should be noted that the network is seriously under-gauged, such that we need to consider network expansion and network efficiency rather than network robustness. In this case, minimizing total correlation is more meaningful to optimize network efficiency. Besides, installing and maintaining monitoring stations often highly depend on financial budgets which cannot be satisfactory in practice. On the other hand, if the network density largely exceeds the minimum network density, and the water resources managers consider shrinking network by closing stations, the authors' argument may become agreeable.

2. The authors also argue that maximizing joint entropy already connotes minimizing total correlation; however, this is not an absolute case, even in the case study in this paper. In Figure 7, the red bar on the top right represents Pareto-front given by maximizing both joint entropy and total correlation. If maximizing joint entropy is equivalent to minimizing total correlation, there should be only one optimal solution rather than Pareto-fronts, and its total correlation should be minimum. It seems like there are three (authors') optimal solutions in Figure 7, and it represents the solution which has the maximum joint entropy does not have maximum total correlation, this conflicts with the authors' argument.

3. The original objective function of the MIMR method by Li et al. (2012) is not the one in Function (9) in the paper. To convert the multiobjective problem into a single-objective problem, Li et al. applied the information-redundancy tradeoff weights and maximized the single objective function. In this case, proper weights should be predefined because the optimal solutions can differ due to the weights. If the weight, $\lambda 1$, is equal to one, the problem will become the same with the maxJE what the authors are proposing. To apply the objective functions in Function (9), multiobjective optimization technique should be employed and it will of course yield multiple optimal solutions on Pareto-fronts. In this case, which optimal solution was selected and discussed, such as in Tables 2 and 3, and why the optimal solution was selected should be addressed.

4. Calculating streamflow information for network design from the monthly time series is quite skeptical. Is a hydrometric network which was numerically designed by monthly time series also good for short-term analysis, such as flood forecasting?

5. The authors finally suggested the greedy optimization algorithm. However, the greedy algorithm is not guaranteed to find the global optimum solution and is easy to fall into a local optimum, even though global optimum can be found in the case study of this paper. Also, in the reviewer's opinion, taking 20 years of monthly time series at 12 stations could be not enough to make a general conclusion. Why do we need an optimization technique if we can calculate the objective values of all populations?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-148, 2020.