Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-148-RC4, 2020

Hydrology and Earth System Sciences
Discussions
Open Access

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

Anonymous Referee #4

Received and published: 19 June 2020

Dear Authors,

I've read your manuscript with enthusiasm, as this topic is of my great interest. I do agree with the authors that there is no consensus in the design of monitoring networks, but I do not agree with the conclusions found in this paper. In general, I find that the scoping of the problem validate the obtained results, and therefore, beats its purpose. This is further detailed in the comments below. In addition, I believe that critical methods and literature are overlooked, especially in the use of metaheuristics for the design of sensor networks. Also, I believe that the document could be better structured, as sometimes results and methodology sections overlap.

Printer-friendly version

Discussion paper

Objective functions for information-theoretical monitoring network design: what is optimal?

General comments

What is the definition of optimallity?

In the introduction you point out that monitoring objectives are posed as part of a wider decision problem (l18-19). Later, you state that the objective in the design of a sensor network is to maximise its joint entropy (l26-27) as sensor networks may support many decisions. At this point, this becomes a normative approach to the design problem, where you define what optimallity is, and that other objectives should be secondary (l347-348). As a consequence, it is clear that the problem is not a multi-objective optimisation problem anymore, and any trade-off with other objectives (such as minimising redundancy) will directly reflect in a performance loss. This is actually seen in Table 2, where you clearly point it out.

Defining the design problem in these terms makes it sufficiently narrow to justify the use of single-objective optimisation, but the point is that not every problem is.

This also connects with the three main arguments presented in the motivation (s 1.2). l44. "Firstly, we argue that objective functions for optimizing monitoring networks can, in principle, not be justified by case studies" - I do not agree with this postulate. The objective of monitoring is to provide information about the state of a system, to support a given action. Measuring for the sake of measuring do not serve any purpose.

l48. "However, from case studies, we cannot draw any normative conclusions as to what objective function should be preferred." - Preferences are not normative, but relative to the decision problem, objectives and context. These are particular to each case study.

l50. "Secondly, we argue that the joint entropy of all signals together is in principle sufficient to characterize information content and can therefore serve as single optimization objective" - I do agree with this postulate, as long as the objective is to maximise joint entropy. This goes back to the first comment.

l53. "Thirdly, multi-objective approaches that use some quantification of dependency or redundancy as a secondary objective, next to joint entropy, could only be justified if redundancy is interpreted as beneficial for creating a robust network" - I do partially agree with this postulate, but then again is linked to the definition of optimality. Given the decision problem, a decision-maker may opt to trade some improvement in joint entropy for redundancy (as an example), and that is out of the scope of what this paper presents.

In general, once you assume that joint entropy corresponds to the definition of optimality, the problem is self-validated. Leading to the conclusions that you are presenting such as: "Information theory provides a valid framework for monitoring network design" (l344), and that single-objective optimisation is sufficient to approach the sensor network design problem.

What are the alternatives for solving the optimisation problem?

From the methodological point of view, I see that you propose the greedy drop algorithm, as an alternative to greedy selection, especially attractive in the cases where exhaustive search is not feasible (l216-223). However, you omitted mentioning meta-heuristics to approach this issue. These are often used in the design of monitoring networks when the combinatorial problems are too large. This being the argument for greedy approaches.

What are the cases of sensor network design that are considered?

When designing a monitoring network, you may have one of 3 possible scenarios: augmentation, reduction or relocation. Augmentation makes for the case when additional sensors are to be placed in the network. Reduction, accounts for the case when sensors would like to be removed from the network. Relocation, deals with the issue of changing the placement of the sensors. The case you illustrate in this paper corresponds to reduction, as the objective is to select a sub-set of sensors from a larger pool. However, you use justifications from the augmentation case (l221-222) to justify the use of a greedy algorithm. I think it is necessary to better define the optimisation problem in this respect.

Specific comments

The section on presenting the basics of information theory can be better summarised. There is plenty of "well-known" material on it.

l172-173 requires a reference.

l187 requires explanation about what is objective GR3

l200-201 These are not MOO methods. These are objective functions.

l202- 203 requires Preferences

l218-220 It is not true that the only way of selecting stations is using greedy algorithms.

l221 is not combinatorial "explosion". Instead we can argue that the problem is exponentially complex (Oˆn)

l222 I think here you are mixing two design problems. One of the problems is of design (where to measure at several locations), and other of augmentation (Where to put additional stations). Of course these are clearly different processes.

l232 I think it should be necessary to include methods using metaheuristics for comparison.

l233 It is not clear what logistical reasons are. Should not these be included in the optimisation constraints?

l235-236 This seems speculative at this point, and better be moved to other place in the document (perhaps introduction?)

l233 This "golden standard" expression seem somewhat loose talk. Can just point out is the only way to prove optimallity?

Figure 4 can be improved. labels are hard to read, and would be more informative just to keep the ID's?. Also if its for monthly data, have you consider using flow duration curves instead? alternatively, please consider using a log y-scale, as discharge distributions are positively skewed. In addition, y-scale label is missing.

in Table 2 are presented the results of different optimisation criteria (defined as MOO), but there is no indication of the selecting strategies, thus the information of the whole Pareto set is unavailable. In top, that is the whole point of MOO, that there exists trade-offs between objectives that cannot be assessed by the modeller.

S4.2 includes parts that should have been presented in the methodology and not in the results section.

l276. H(X,X) = H(X), therefore two completely "dependent" are exactly as informative as one.

Visualisation using Venn diagrams are an excellent way of presenting concepts, but are really hard to follow to describe precise quantities. Will it be possible to re-think Figure 5 in a simpler manner?

l287-288 Those are precisely the trade-offs that decision-makers do in MOO, and the reason of its relevance. If you claim that maximum joint entropy is somewhat equivalent to minimum total correlation, then these two objective functions are not conflicting, and therefore, by maximising one, you are maximising the other. Precisely as shown in Figure 7.

l297-298 The whole point in obtaining the Pareto front between maximum joint entropy and minimum total correlation explores the trade-offs between a network that is able to capture most information, vs a network that has little "information overlap", which is not the same as the individual entropies are different. Therefore, these are different metrics, and is the reasoning behind finding the Pareto set in the dashed line of Fig 7.

S4.3 A lot of the text in this section can be part of the methodology.

Table 4 is quite hard to read. Also, this table is precisely showing you that optimallity is not found using greedy algorithms.

Table 5 Be consistent in the amount of significant digits through your document

l344 No information system is justified if there is not objective to tackle. What problem am I addressing if I do not know what the problem is.

l361 Large optimisation problems are tackled using metaheuristics, and has been a widely used approach. This has not been mentioned here at all.

l364 the differences between the greedy and exhaustive search approaches have not been presented quantitatively. In this problem they may seem "little" (not being explicit about what little or big is in the context of the problem), but this cannot be ensured for larger problems that the ones presented here.

l367-368 Language has to be precise (how to numerically calculate this objective function, or other objective functions used in other approaches)

l370 Any information metric is hard to calculate with limited data.

l378-379 I completely agree with this line "before thinking about how to optimize, we should be clear on what to optimize".

l381 I visited the GitHub repository, but I was unable to find the code to reproduce these results. Only a reply to a WRR paper of 2018, and a fork of pysheds.

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