Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-5-RC2, 2020

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "HER: an information theoretic alternative for geostatistics" *by* Stephanie Thiesen et al.

Anonymous Referee #2

Received and published: 28 April 2020

This paper presents a method called Histogram via Entropy Reduction (HER) for the interpolation of spatial geophysical data. This method is based on information theory measures of entropy and relative entropy, and has advantages over benchmarks of kriging and nearest neighbor methods in terms of its generality and lack of assumptions. The authors present the methodology which determines spatial dependency structure based on observed data points, and estimates optimal weighting parameters used to predict a given variable at a location. An application to several synthetic datasets shows high effectiveness of the method relative to three existing interpolation techniques.

Overall this paper was interesting and clearly written, and the figures are very informative and help to illustrate complex concepts. Although I do not have a background

Printer-friendly version

Discussion paper

in geostatistics or interpolation of sparse datasets, this paper seems to introduce a promising avenue of how IT measures can be advantageous in this field. Some comments and suggestions listed below, which consist of minor revisions/technical corrections. They mainly highlight places that could use additional explanation or clarification.

Main Comments: Line 111: on the description of creating the "infogram cloud": at first it was not clear to me whether an "infogram" was an existing technique that I was unfamiliar with, or designed by the authors. I think it is the latter (based on line 134) – either way, this aspect could be made more clear earlier in the subsection, that you have developed this graphical technique called an infogram that shows spatial correlation structure.

Line 145: Could you add a bit more information on what effect/advantage this has? It seems like attributing a small probability to every category would make a larger difference to some types of distributions than to others.

Section 2.3.1: For readers less familiar with aggregating probabilities towards a spatial context, I this description could benefit from some sort of illustration or simple example that shows the difference between the different pooling operators in Eqs 4-7. For example, show two measurement points D1 and D2 with a target location A somewhere between them, and show how the measures differ. This actually become more clear to me with the later discussion in Line 445 onward, so maybe some of these aspects could be brought forward earlier.

Section 3.1: Are there implications to model performance of the Gaussian process used to make the test cases? I wonder how this would compare to a realistic landscape (or whether this is considered very close to a "real world" case). Something is mentioned later about this in the discussion regarding the OK method, but more information would be beneficial here.

Line 360: I think this is explaining why the infogram illustration in Figure 2a is very different from that in Figure 4a for the farther distances (which is because with the

data, there are fewer pairs that exist at those farthest distances). If this is correctly interpreted, it could be brought up in the description of Figure 2a and the infogram cloud-shape in general.

Technical/writing style comments: Abstract: There are several sentences here with parenthesis for additional context, I would recommend re-writing these without as many parentheses to potentially simplify and help the flow.

Line 38: applying

Line 90: H(X) is upper bounded by infinity in a continuous case, but as you mention in the next sentence and the equation that this case is discrete – the upper bound should be log2(N) where N is the number of bins or categories. Figure 6: It is hard to see the targets in a few of the maps, I think because the markers are in the background behind the observation markers. It would help to make outlines of the marker shapes bolder or colored to show which target is which.

Line 379: “differentiate between”, instead of differ? Line 429: I found this sentence to be unecessary

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-5, 2020.

Printer-friendly version

Discussion paper