Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-658-AC1, 2017
© Author(s) 2017. CC-BY 3.0 License.

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Estimating sediment thickness by using horizontal distance to outcrop as secondary information" *by* Nils-Otto Kitterød

NOK Kitterød

nils-otto.kitterod@nmbu.no

Received and published: 29 March 2017

Thanks a lot for constructive questions and comments! I have worked through all of them, and I have revised the manuscript according to the reply below.

On C1. The purpose of the paper was to explore possible cross-correlations between the thickness of sediments, D(u), and horizontal distance to exposed bedrock, L(u). Cross-correlations might be capitalized later to improve estimates of D(u). This question was tested by using ordinary kriging and co-kriging. Thus, the paper was not intended to describe "the value of kriging and co-kriging", as indicated by Dr. Farmer. Nevertheless, I appreciate this general comment very much.

The typographical error (P2, L5) is corrected, and the term "significant" is substituted with other synonyms as for example "large".

I agree that scientific writing should be as precise as possible. This is a motivations for using mathematics! At the same time there is a lot of useful (common) words that also has a distinct mathematical or statistical meaning. "Expectation, variance, and correlation" are all examples of such words, they have a specific mathematical definition, but they might also be used in another context to express something "significant".

On C2. P3, L5: My intention was not to be amusing, therefore "Zoo" is substituted with "number" in the manuscript. However, a competition is going on out there and the fittest method will survive ("jungle" might therefore be a more adequate metaphor than "Zoo"!). My point is simply that the current study is not a part of the on going competition, I just want to investigate whether the secondary variable (L) might be used to improve the estimates or not.

P5, L14: I do not agree. There are two populations of wells, those located on exposed bedrock, L(u)=0, and those not located on exposed bedrock, L(u)>0. The 750 wells belongs to the second population and should therefore not be described in the previous paragraph. Even though L(u)>0, there are 750 where D(u)=0. These wells are removed in the current study to keep the analysis as simple as possible. It is interesting to notice that number of wells (750) is only a minor fraction (3.7%) of the wells with L(u)>0 (20432). One reason is that the soil cover is very patchy some places, but there is no serious inconsistency between the mapping (and definitions) of D and L.

P5, L15: I think descriptive statistics belongs to exploratory data analysis, and I recognize that others do the same (e.g. Goovaerts et al. 2005, Water Resour.Res., Vol. 41, W07013, doi:10.1029/2004WR003705, 2005). Thus, at this point I suggest no revision of the text.

P5, L16: I've made a (sloppy) normality plot, which indicate that log10(D) and log10(L) did not belong to a Gaussian pdf (c.f. enclosed figure). The main reason for not using a lognormal variable, however, is the (large) estimation error associated to the log-transform. The estimation error of a lognormal variable becomes usually very large

in some (extreme) locations because (by definition) the error of a lognormal variable includes the expected value. This is the main reason why the lognormal transformation has fallen into disuse in science.

P5, L22: No, it is not the same searching window as mentioned on P7, L10. To avoid any confusion I have rewritten the paragraph to: "Mean and standard deviation of D and L as a function of separation distance h, is given in Fig.3 for $\Delta h = 20$ m and $\Delta h = 150$ m."

P6, L24: I would prefer "origo" (location of point zero), which I find more specific than "origin", but both terms may work equally well.

P7, L20: I tried both a computer approximation and a analytical expression of the Gaussian pdf, and according to my experience, the analytical expression was more precise especially for the extreme parts of the distribution.

P7, L27: The paragraph ("The back transformation, however, does not reproduce the censored part of the pdf.") is deleted in the revised manuscript. The normal-score transform is done on declustered data, which removed (smoothed) the censored character of the raw data. This is not related to the problem of over-estimation of low D and under-estimation of large D.

P9, L12: The first order exponential model (\alpha = 1 in eq. 14 and 16) is usually referred as the "the exponential semivariogram model". Even though \alpha is close to one, this is not the "chosen" model. The alpha parameter in eq. (14) and (16) may vary between 1 (exponential model) and 2 ("gaussian" model). Cross-validation is also done with C_0 ∼= C_1, which is similar to a "constant" model (c.f. Tab. 3, case A).

P13, L28: Thanks a lot for your humble reference to Farmer and Vogel, 2016! Quantification of uncertainties is a part of the estimation problem, I totally agree! Over-estimation of (extreme) small values and under-estimation of (extreme) large values is a well known problem for Gaussian least square methods. The purpose of the paper is

not to improve ordinary kriging or co-kriging, but to demonstrate the effect of including a secondary variable (L). Of that reason, I do not intend to elaborate this subject any more in the current manuscript. There are however, made significant progress on the problem of Gaussian and non-Gaussian estimation, and I therefore include references to: Omre and Haldorsen, 1989; Rue et al., 2009; Lindgren and Linström, 2011; Leblois and Creutin, 2013; Ingebrigtsen et al., 2014; 2015.

P14, L1: The paragraph is rewritten to: "In cases with minor difference in absolute error, the estimation results might be ranked according to criteria for estimation accuracy (28) and precision (31)."

P16, L32: Yes, I agree, and the paragraph is rewritten! However, I think the value of public databases on hydrology and environmental databases in general, should be more honored by the scientific society. of that reason I suggest to include the paragraph: "Hence, in this context, the present study is a call to explore public data to obtain important estimates for science and society."

Figure 10 and 11: I decided to include a simple map of Scandinavia, which also indicate the three subsections: Southern Norway; Northern Norway; and the Oslo region. If possible, I would suggest to merge Fig. 10 and Fig. 11.

Table 3 and 4: From my point of view, the presentation of Tab. 3 and 4 does not belong to the method section. Both tables are results of the applied methods. I think the results should be presented as clear as possible without any comments, and interpretation and explanations therefore belongs to the discussion section. Thus, I suggest to keep the text on this point as it is.

With respect to grammar, style and typographical errors, I appreciate all help and corrections! Thanks a lot!

Nils-Otto Kitterød <nils-otto.kitterod@nmbu.no>

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-658, 2017.

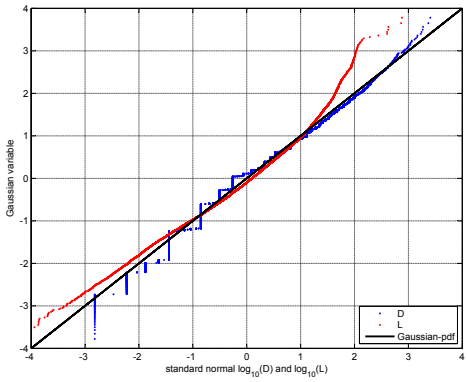

Fig. 1. Standard normalized log10(D) and log10(L) plotted against a Gaussian variable, N(0,1).