Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-401-RC2, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "An Extended Kriging method to interpolate soil moisture data measured by wireless sensor network" *by* Jialin Zhang et al.

Anonymous Referee #2

Received and published: 4 November 2016

This paper introduces an interesting alternative model for prediction of a spatial variable by kriging. However, it does not reflect adequately on the motivation of the model, or explore its implications. There is therefore no real justification for the selection of the model, which strikes me as very implausible. Furthermore, the model is not implemented correctly. I expand on these statements below.

First, there is no reason not to extend the geostatistical model from 1, 2 or 3 spatial dimensions to a space of higher dimensions. This is done, afterall, in space-time geostatistics. However, if one is to treat some covariate as defining an additional dimension then, if one is assuming intrinsic stationarity in the new space, as one must for this extension of ordinary kriging, then this implies that there is no systematic relationship between the target variable and the covariate. In terms of the definition of the intrinsic hypothesis of stationarity E[Z(s)-Z(s+h]=0 where Z(s) defines our random variable ex-

Printer-friendly version

Discussion paper

pressed as a random function of the covariate. In words, one is envisaging a situation in which a plot of the observations z against the corresponding values of the covariate s would not show any systematic trend but rather a random fluctuation, exhibiting some degree of correlation. This seems a rather implausible model to me. However it might be worth considering. Despite what the authors say, however, it is not a model one might use in a situation where cokriging or kriging with external drift were also candidate approaches, since in the case of kriging we assume a linear relationship between z and s, and in KED we assume a linear relationship or a relationship linear in the parameters of some simple fixed effects structure such as a polynomial or (non unduly complex) spline basis. This proposed approach and cokriging/KED would be fundamentally incompatible. Given this, one would expect the authors so start by showing that cokriging or KED are not suitable for these data by plots and exploratory statistics that show (for cokriging) that a linear model of coregionalization is not plausible or (for KED) that no reasonable fixed effects structure looks reasonable.

Let us assume that the authors do show a sound motivation for applying their model. They must then estimate it appropriately. I think there is a difficulty here. The authors should read the literature on space-time geostatistics to get a better understanding of this. One problem is that, unlike the space-time case, one cannot compute estimates of two marginal variograms (the spatial variogram with the time (or here s) lag =0, and vice versa). In this paper the authors cannot compute a marginal variogram in s space, because it is not possible to find observations with lag 0 in space but some non-zero lag in s. The variogram shown presumably uses lag bins, but the lack of the marginal variogram is a problem because there is likely to be nugget variation in both dimensions and you cannot resolve their contribution to the joint space-s variogram. There is a second problem. The authors propose a simple model for the space-s variogram which is the sum of a spatial and an s-dimension variogram (Equation 15), but such a model is not in general valid (i.e. it does not define a non-negative definite covariance structure in the overall space). The authors should look at papers such as the one by De Cesare, L., Myers, D.E., Posa, D., 2001.[Estimating and modelling

space–time correlation structures. Stat. Probab. Lett. 51, 9–14.] for an account of this and of some valid models.

I have a further problem. The authors are using the variogram from a remote sensor for kriging from in-situ sensors. Even if they could be confident that the two variables are measurements of the same underlying quantity the variogram of the remote sensor data is a regularization of the variogram of the networked sensor data onto a very different spatial support. The question of how the remote sensor data might be used to help with this problem is interesting, given the fact that you have some data on the desired support it might be possible to do this, but it requires an explicit change of support step, not just ignoring the difference.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-401, 2016.

Printer-friendly version

Discussion paper