Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-276-RC2, 2018

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Geostatistical interpolation by Quantile Kriging" *by* Henning Lebrenz and Andras Bárdossy

Anonymous Referee #2

Received and published: 27 June 2018

The authors propose a new kriging technique developed to regionalize non-normally distributed spatio-temporal variables. The approach, as I understand it, involves (i) fitting time series observations to a predetermined distribution type at each gauge independantly; (ii) using the fitted distribution to assign a probability of non-exceedance to each observation at each gauge; (iii) fit the subset of these probabilities corresponding to each observation time to (different) beta distributions; (iv) map the fitted probabilities to normal quantiles; (v) use the (now normally distributed and *assumed* second-order stationary) outcomes as a basis to apply OK or EDK, as described (rather cryptically) in section 2.1.2.

While the approach is intriguing, its description in the paper lacks statistical rigor and minimal proofs and intuitions. Its application in cross validation suggests that it does

Printer-friendly version

Discussion paper

a decent job at predicting some measure of rainfall, but so would most properly fitted interpolation methods. This does not mean that the estimator is BLUE (best – or even efficient – unbiased linear estimator) of the considered stochastic process. Actually, the apparent correlation between Z*(x) and sigma2(x) (Figure 4) suggests that the process has some degree of heteroskedasticity. Under these conditions, linear models (like most kriging estimators) are not necessarily efficient (see, for instance, the Gauss Markov Theorem for ordinary least squares). Again, the approach is promising, but its exposition needs a major overhaul to be convincing.

In particular:

Non-Gaussianity: In its canonical form, Ordinary Kriging is based on the method of moment (i.e. variance minimization subject to unbiasedness) and so is not technically restricted to normally distributed processes. Gaussiannity is, however, required for maximum likelihood (ML) type estimations, which some studies have shown can be more efficient for kriging-based regionalization (e.g., Lark 2000). ML approaches become necessary to have unbiased prediction of both the mean and the variance when it comes to EDK or Universal Kriging, particularly when subject to more intricate error correlation structures (e.g., Muller 2015). Since non-gaussianity appears to be a key rationale for the proposed approach, it is important to be specific on that.

Beta-distribution: The use of the beta distribution is intriguiging, but more intuition is needed on the assumed underlaying stochastic process. Please describe clearly the properties of the (space-time) stochastic process that you assume gives rise to the observed sample and use that to demonstrate Eqn 6 and 7 in a rigorous mathematical proof. I find the use of the beta distribution promising because a common interpretation is that it describes the distribution of the probabilities associated to a binomial process observed over a finite sample. Let's say that the binomial process in question is the exceedance of a given threshold (as eluded to in the manuscript). Then, if the underlaying point process is identically distributed in space and if an identically sized sample is taken at each observation point, the proportion of observations lower than a given threshold across all gauges will be beta-distributed. Perhaps that's a start?

Stationarity: More fundamentally, a main issue that I have is that your approach involves fitting distributions independantly at different points in space (gamma or weibull) and time (beta), which implies that the underlaying random point process follows a different distribution at each point in space. Granted, EDK and Universal kriging allow the first moment of the underlaying distribution to vary through space, as allowed by the scaling properties of the expectated value estimator. However, I am not aware of any existing geostatistical approach that allows for higher order moments to vary through space. This may be fine, but please demonstrate that your approach does not violate the second order stationarity assumption (i.e that the variogram is constant through space), which is critical (and arguably more important than gaussianity) for kriging.

References:

Lark, R.M. (2000) "Estimating variograms of soil properties by the method of moments and maximum likelihood", European Jornal of Soil Science

Muller, M.F and Thompson, S.E. (2015), "TopREML: a topological restricted maximum likelihood approach to regionalize trended runoff signatures in stream networks", HESS

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-276, 2018.