Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2015-536-RC2, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Ordinary kriging as a tool to estimate historical daily streamflow records" *by* W. H. Farmer

J. O. Skøien (Referee)

jon.skoien@gmail.com

Received and published: 2 March 2016

This manuscript analyses ordinary kriging for estimation of historical daily streamflow. The paper is well written, and includes interesting analyses. Some revisions are still necessary before it can be published. Below are some suggestions for improvement.

There is some overlap between results and discussion, where some results are discussed in the results section, and then this is partly repeated in the discussion, which is then more like a summary. I think this part of the manuscript could be better organized.

Some work is based on the author's PhD (Farmer, 2015). I am not sure how easily accessible this PhD is? I think it is ok to include results from the PhD in this manuscript as previously unpublished, as long as they have not been presented in other peer re-

Printer-friendly version

Discussion paper

viewed journals/conference proceedings. The paragraph on P6 L12-17 should anyway be rewritten. I don't see why it is not intuitive to model each day independently? What is the most extreme? What is meant by stationarity of variogram parameters here, that they are temporally constant?

The brief summary of kriging is not brief; it pretty much includes everything in Skoien et al. (2006), which the author refers to, just in a different way. What is missing is actually a variogram model and the sample variogram, which is of more interest for the analyses than the equations for finding the weights. I think it is necessary to discuss in more detail the advantages/disadvantages of OK/TK and pooled/daily variograms beyond the cross-validation results. First of all, OK can be seen as the most extreme case of TK with only one regularization point, and comparing with the results of Skoien et al (2015), it is not surprising that also OK can perform reasonably well for most catchments. However, this is likely to depend on the configuration of observation and prediction locations. Table 1 indicates that the NSE of TK is considerably higher for the 10th percentile. Then the author barely mentions the prediction uncertainty, where I would expect daily variograms to perform better than pooled variograms, and TK should perform better than OK.

Regarding biases of upper/lower extremes, particularly P8 L3-10 and P12 11-14 – I think it is not surprising that kriging tends to over/underestimate the extremes. Kriging is generally unbiased around the 50th percentile, and a range around this percentile. For the conclusions, kriging is a stochastic approach, not deterministic, and using simulations would only create a range of values which is still centred around the under/overestimated prediction. Higher nugget effects will increase the tendency to smooth the extremes in the interpolated field. The average underestimation of 40% for the largest streamflows is still somewhat large, it would be useful to further analyze and discuss the causes for such large deviations, and see if there are particular cases where they are larger. On the other hand, the prediction difference might be smaller, as this refers to exceedance probability, something which could be better explained. It

Printer-friendly version

Discussion paper

is not easy to understand from the text what is meant by exceedance probability in P8. The sentence “For low streamflows, below . . .” is not clear either.

Although not frequently in use as far as I know, there are some methods for unbiased kriging of extreme values. Two of them are IWQSEL (Craigmile et al., 2006) and Modified Ordinary Kriging (Skoien et al, 2008), both methods implemented in the R-package intamap. I don’t think this would be feasible to include for the analyses in this manuscript, but could be a possibility for future work. The results from the analyses of autocorrelation on P8 should probably not focus so much at what I would refer to as relative differences in percent (as is currently done, although referred to as absolute), rather the absolute errors. As mentioned, relative differences when autocorrelation is below 0.1 can be large, but still negligible. A figure could present the autocorrelation as lines.

The discussion about temporal variations of variogram parameters at the end of 3.2 should also include some more thoughts about the reason for the temporal changes, and check the description of the current relationships. I would assume that shorter ranges in Summer is an indication of more heterogeneity (more convective precipitation) and that long ranges is an indication of homogeneity. I would also assume that sill is decreasing in Winter because runoff is decreasing, so the ratio sill/mean runoff could be of interest, the same with the nugget/sill ratio. Long ranges in Winter/spring could be related to snow/snow melt. “beyond the range” is confusing.

Minor comments

P2 L12 “It is postulated” – Who postulated what?

P2 L27-28 There are also kriging methods where the predictions are based on external variables in addition to geolocation.

P3 L9 “Deterministic rainfall-runoff models” is more commonly used than mechanistic.

P3 L9-10 I would say that the comparison is missing from their work, they did not

emphasize the need.

P3 L25 logarithms of UNIT runoff?

P4 L1 The sentence is a bit clumsy (Because . . . because), consider rewriting, such as ". . . reference-quality in the designation (. . .) or in previous flood-frequency studies (. . .)." Remove brackets from next sentence.

P4 L10 One out of 33 days on average or some periods of 1-33 days? These were filled by the author or has previously been filled in by USGS?

P4 L16-17 This sentence seems unnecessary complicated, and I am not sure if it is completely correct.

P5 L17 Which previous hydrologic geostatistics is the OK in this manuscript an extension to? And depending on the answer, is OK really an extension or does the manuscript include analyses which are useful as an addition to other methods?

P5 L30-31 I think "the temporal considerations" can be deleted.

P6 L6 I don't think it is the covariance, but it could be the variance, due to short scale variability or measurement errors.

P6 L8 "the dependent variable" can be deleted, together with "of" (structure of which)?

P6 L21 What is stability of the spatial covariance structure?

P6 L23 Move first sentence to L25 (after daily variogram sentence).

P6 L31-32 Here it is a bit unclear what is meant by "average model". In L15-16 it is referred to averaging of model parameters, which is definitely different than a variogram model fitted to a temporally pooled empirical variogram. If the difference between average and pooled refers to the difference between treating each daily empirical variogram as equal, or giving them weights according to the number of pairs in each bin, then this should be described more explicit.

P7 L11 The usage of top-kriging in this manuscript is not exactly the same as the one described in Skoien and Bloschl (2007). The previous paper uses spatio-temporal regularization, whereas the implementation in rtop only uses spatial regularization. However, it can be assumed that the difference between these is small, and not likely to affect the quality of predictions. A formal comparison has not been done, but the current version of rtop uses a similar method as the one in the manuscript for predicting time series of runoff.

P7 L27 Is this poor performance in the recession period typical for all catchments, or is this example worse than many others? I would expect a bias in the extreme, as presented later in the paper, but also that some catchments will even be underestimated during May and June in the figure.

P7 L31 - P8 L2 I find this sentence somewhat unclear, in addition, is the difference significant? For me there is barely any difference between the curves. Regarding the figure, I usually don't like grids in figures, but this could be an exception where it might add some information.

P8 L17 and -> a

P11 L15-17 I did not understand this sentence. What is meant by hours and days per site?

P11 L18 This was correct in the past, but pooled variogram estimation is now included in the package, together with time series interpolation.

Craigmile, P. F., N. Cressie, T. J. Santner, and Y. Rao. 2006. A loss function approach to identifying environmental exceedances. Extremes, 8, 143-159.

Skoien, J. O., G. B. M. Heuvelink, and E. J. Pebesma. 2008. Unbiased block predictions and exceedance probabilities for environmental thresholds. In: J. Ortiz C. and X. Emery (eds). Proceedings of the eight international geostatistics congress. Gecamin, Santiago, Chile, pp. 831-840.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-536, 2016.

Printer-friendly version

Discussion paper