Hydrol. Earth Syst. Sci. Discuss., 8, C2724–C2726, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C2724/2011/

Hydrology and Earth System Sciences Discussions

HESSD

8, C2724–C2726, 2011

Interactive Comment

Interactive comment on "Interpolation of groundwater quality parameters with some values below the detection limit" *by* A. Bárdossy

F. Serinaldi (Referee)

francesco.serinaldi@ncl.ac.uk

Received and published: 12 July 2011

General comments

In this work, A. Bárdossy proposes an original application of the copula-based geostatistical interpolation framework introduced by Bárdossy and Li (2008). The main novelty, which is surely of interest for the HESS audience, is the handling of the values of environmental variables below the detection limit of the measuring devices. In more detail, the Author introduces suitable marginal distributions to model the observations below the detection limit and the corresponding univariate and bivariate likelihood functions, which are fundamental for the appropriate estimation of the model parameters. I share the opinion of the other reviewers and suggest the publication of this work once the issues raised in the review process have been addressed to the Editor's satisfaction. In particular, as already mentioned by P.J. Smith, I further recommend to follow the detailed and valuable comments by G. Pegram and to improve the overall presentation. Some further general remarks are reported in the following.

HESSD

8, C2724–C2726, 2011

Interactive Comment

Specific comments

The marginal distribution G implies a parametric component for the values below the detection limit; however, I'm not able to recognize the parametric family used in the case study. Which family is used? Does the Author use the same $F(Z|\theta)$ for the three data sets? Fig. 1 is not very clear. Do the cross symbols denote the ECDF of the observations below the detection limit? Do the curves denote $F(z|\theta)$? What about the full G? In general, I strongly recommend to improve figures and captions. As this work contains valuable materials, the graphical output must exhibit the same quality.

The distribution of the process $Z(x)$ should be G. However, in Section 2.2, the distribution of $Z(x)$ is F. Therefore, it can be worth clarifying if $Z(x)$ denotes the whole process or the process below z_{lim}.

In my opinion, the comparison between the copula method and the ordinary kriging is not very fair. Probably, OK can yield better results by preprocessing the data through the normal quantile transform. This way the OK linear interpolator can likely provide results closer to those of the Gaussian copula approach.

My last concern is about the performance assessment. While figures 8 and 9 are rather effective, Tables 3, 4, and 5 are not very much. The improvement given by the copula methods is not so evident based on these indices. For instance, in Table 3, v-copula and IK show the same value of LEPS score, and the difference in terms of rank correlation could be no significant. In general, a measure of performance only allows for a classification of the models based on their score, but does not give information about the significance of the differences, unless the significance is verified through some formal tests. Moreover, the measures in Tables 3, 4 and 5 do not take into account the different complexity of the tested interpolators. It is also worth noting that the rank correlation cannot be properly considered as a measure of performance. It tells us if the data follow a monotonic curve, but does not measure the signed distance between observed and interpolated values, which in turn can be very large. For instance, if the interpolated values are log-transformed or shifted (i.e. a monotonic transformation is applied) the rank correlation do not vary at all, but the interpolations can be strongly biased towards meaningless values. Finally, I did not find any discussion about the last row of Tables 3, 4 and 5. Please, consider to improve the table captions.

Technical corrections

In Eq. 7, X_j is used to denote the v-transformed variable; however, x_j also denotes the spatial coordinate of the random process. Please, check the notation throughout the paper.

Page 5279 (5-10) "Results for the two censored variables and for an artificially censored case chloride are displayed in Tables 3, 4 and 5"

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 5263, 2011.