Hydrol. Earth Syst. Sci. Discuss., 6, C444–C451, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C444/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Effects of intersite dependence of nested catchment structures on probabilistic regional envelope curves" *by* B. Guse et al.

Anonymous Referee #1

Received and published: 28 April 2009

This manuscript addresses an interesting and relevant issue: the impact of intersite correlation in flood data on regional frequency analysis, exemplified through the use of probabilistic regional envelope curves (PREC).

In my opinion, some additional editorial work is required to improve the presentation of the research, in particular:

- The authors are not being very helpful to readers not intimately familiar with the concept of PREC, which is not yet standard method. A better description of the method and the underlying assumptions would be useful to increase readability.

- A more structured introduction to regional flood frequency estimation and the influ-

ence of intersite correlation of flood data would be useful.

- In several places the description of technical aspects of the methodology is inadequate (see examples below).

- In places the scientific notation is more like FORTRAN variable names (e.g. eqs. 5, 6) which does not look elegant (ok, a minor comment).

- Being a non-English speaker myself, I still think that, in general, the English could be improved.

Also, I am concerned that the results presented in section 4 are merely reflecting the relationships between neff, T and intersite correlation presented in Section 2 rather than providing any form of independent validation of the PREC or the effects of intersite correlation and heterogeneity on bias or variance of the flood quantiles. I think the authors need to clarify this issue before the manuscript can be considered for publication.

Specific comments:

Abstract: Consider removing the first sentence. Line 7: Consider "located in Saxony" rather than "belonging to Saxony". Line 19: Consider "degree of homogeneity" rather than "homogeneity degree".

1. Introduction:

Page 2847: The index flood method is only one of several possible regional flood frequency methods (Madsen and Rosbjerg, 1995). Linking at-site estimates of the design flood directly to catchment descriptors through a linear regression model is another popular method, which also explicitly needs to consider the correlation between flood data, see for example Reis et al., (2005). I think a more structured introduction to regional frequency estimation and the role of intersite correlation would help readers to better understand why this is an important topic, and the significance of the results presented here.

Page 2847, Line 20: It is not necessarily a requirement of a 'pooled analysis' that the data are independent. Assuming that the sites are independent, and if the L-moment ratios do not vary from site to site (and the variance of the L-moment ratios depend on record-length only) will result in the simple record-length weighted pooled average (e.g. Hosking & Wallis 1997, Eq. 1.5). If these assumptions are not fulfilled then the resulting estimates will be either biased or more uncertain (for example, Stedinger (1993) consider the variance to increase as a result of heterogeneity, whereas Hosking and Wallis (1997) argue that heterogeneity will increase bias). There are examples in the literature of index flood models considering the effect of intersite correlation on the variance of the design floods (e.g. Stedinger 1983, Kjeldsen and Rosbjerg, 2002; Bayazit and Önöz, 2003, Kjeldsen and Jones 2006), but I am not aware of any study considering analytically the effect of intersite correlation on the weights (of the L-moment ratios) in an index flood method other than stating that the effect is minor (e.g. Hosking and Wallis, 1997). Other examples might include the effect of using imperfect knowledge of the intersite correlation on model parameters in a regional regression model (Kroll and Stedinger, 1998). I suggest that you make clear that you use the index flood method as presented by Hosking and Wallis (1997) and list all the assumptions (Hosking and Wallis, 1997, page 8).

2. Methods:

2.1 Regional information content and number of effective observations

Page 2851, Line 1: I cannot follow the description of how the number of effective observations is derived.

Page 2851, Line 4-5: The length of the overlapping series does not feature in Eq. (1).

Eq. (1): Why did you choose this particular form of correlation function? Other researchers, notably Robson and Reed (1999), have used a simple exponential function to describe the correlation-distance relationship.

Page 2851, Line 16-17: Perhaps it is the formulation, but I would think that a single observation can be influenced by intersite-correlation, but you might not be able to estimate it?

Page 2851, Line 17: I do not understand the notation Nsub(N-n1). Is there a '=' missing?

Eq. (2): I do not understand the notation used to derive this equation. For example, what does the subscript LS on the correlation coefficient signify? This is an example where the manuscript could be a bit more helpful to the reader.

2.2 Probabilistic regional envelope curves (PREC)

As mentioned above, I think this section needs to provide a better description of the method to readers not familiar with the method. Perhaps more details could be provided in an appendix.

Page 2852, Line 5-7: Consider something like ' First, the index flood method requires that all selected flood series constitute a homogeneous region, i.e. that they are identically distributed except for the scale parameter; the index flood. And Secondly, . . .'.

Eq. 3 + sentence above: Is it necessary for the region to be homogeneous for the relationship in Eq. (3) to apply? Is it not only the second of the two principles that lead to Eq. (3)?

Page 2852, Line 17: Consider replacing 'data pair' with 'data point' to make it clear that it is a single unit runoff measurement with the associated drainage area and not to different measurement of unit runoff.

Eq. (4): Can the authors discuss for example the maximum and minimum possible values of neff (or some typical values for simple regions) and how they effect the return period. This might help the reader better to conceptualise the effect of intersite correlation on the PREC. Also, it will help this section not to end with an equation. This might render Section 4.3.1 redundant.

2.3 Pooling scheme

Page 2853, Line 4: I don't think that section 2.2 made it clear why it is 'essential' to provide homogeneous regions.

Page 2853, Line 6: What is a 'behavioural subset'?

Page 2853, Line 19: It is not clear what combinations of catchment descriptors are being used. However, if different combinations are used, will that not create different scales for the similarity distance measure? If that is the case, is it reasonable to use fixed thresholds as they might not compare across sites? There is probably something here I don't understand, but perhaps the authors need to provide a better descriptions of the subsets and how similarity measure is calculated.

2.3.1 Homogeneity test

Page 2854, Line 3: Can the authors clarify where this bias comes from?

2.4 Application and interpretation. . .

Page 2854, Line 12: Perhaps replace 'approaches' with 'cross-correlation functions'?

Eq. 5,6: Be aware of FORTRAN-like notation. Is EFFOBS=neff as defined in Eq. (2)?

It is not necessary with a separate section 2.4.3, and also consider merging Eqs. 7 and 8.

3. Study area:

First sentence: Consider something like "The study area is the ???km2 federal state. . ."

Page 2857: Why not include data from all catchments that have all or parts of their drainage area in Saxon? Why is an administrative boundary considered an appropriate delineation of the study area?

Page 2857, Line 13-14: Consider ". . .at least 30 years and include data at least up to the year 2002."

Page 2857, Line 23: SRTM DEM?

3.1 Selection of catchment descriptors

It is not specified in the beginning of this section what the selected catchment descriptors are used for. I assume that it is for the definition of pooling groups, in which case it might be more relevant to consider catchment descriptors that are correlated with the higher-order moments that define the growth curve as opposed to the index flood.

4 Results

Section 4.3: Could perhaps be deleted with reference to my comment on Eq. (4) above.

Section 4.4: I have some problems with the results presented in this section, as I am not sure that what exactly the authors are trying to convey. The plots in Figure 6 show the effect of the total number of events and the correlation structure on neff. However, this relationship, if I understand it correctly, is determined solely through Eq. (2). Thus, data points on these graphs are not a validation of the method or in any way surprising but merely reflect Eq. (2). Also, rather than create 'random' pooling-groups as described in Section 4.4, why not just take each site, start with a very small pooling group and incrementally add a site until H>2 or N=89? There is nothing stopping you from defining all sites as a region. Granted, the homogeneity measure might be larger than 2, but that seems not to be of any relevance to the evaluation of Eq. (2)? Alternatively, consider removing section 4.5.

Section 4.6: I think the logic in last sentence is the wrong way around. I would argue that adding more sites to a pooling group typically results in a larger degree of heterogeneity. Is the larger return period a result of the larger heterogeneity or simply because more sites means a higher value of neff, which then result in a higher return period (Eq 4)?

Section 4.7: As in Section 4.4, the differences in neff reported here derive from the evaluation of Eq. (2) with different correlation functions, and are not surprising. It is not

clear to me what is being communicated in the last sentence 'The differentiation. . .'?

Section 4.8: The first sentence is not very informative, and the particular effects should have been made clear in the introduction. Second sentence: perhaps a more formal introduction of the H1 test would make it easier to follow these arguments? Mention that increased variance (between sites) of L-CV will reduce the H1 test-statistic.

Section 4.9: The results in this section are similar to the results presented in section 4.7, with the difference that more pooling-groups are being created, and I can't follow that the authors investigate 'the effect of regional heterogeneity on intersite correlation' in any direct manner. As mentioned before, why not just increase the size of the pooling groups in steps rather than introduce these thresholds for H?

5: Discussion

Page 2866: Consider removing line 15-22.

Page 2866, line 15 onwards: It seems here that the authors are discussing some investigations that they have carried out but not reported. For example, the relationship between rainfall-event types and the degree of correlation. Perhaps it would be interesting to report these data-based findings rather than some of the model-dependent results presented in Section 4? A simple test would be to investigate the correlation of a subset of the data where flood events across the section occurred on the same day (e.g. $\pm$2days) or correlation as a function of flooding season etc.

6. Conclusions

Page 2868, line 11-13: This sentence implies that the results presented I the case-study brought about this insight. However, as I have argued above, this really was implicit in Eq. (2).

References

Hosking, J. R. M. and Wallis, J. R. (1997) Regional frequency analysis: an approach

based on L-moments. Cambridge University Press, New York.

Kjeldsen, T. R. and Rosbjerg, D. (2002) Comparison of regional index flood estimation procedures based on the extreme value type I distribution. Stochastic Environmental Research and Risk Assessment, 16(5), 358-373

Kjeldsen T. R. and Jones D.A. (2006) Prediction uncertainty in a median based index flood method using L-moments. Water Resources Research, 42, W07414, doi:10.1029/2005WR004069.

Kroll, C. N. and Stedinger, J. R. (1998) Regional hydrologic analysis: Ordinary and generalized least squares revisited. Water Resources Research, 34(1), 121-128 Reis D. S., Jr., Stedinger, J. R. and Martins, E. S. (2005), Bayesian generalized least squares regression with application to log Pearson type 3 regional skew estimation, Water Resour. Res., 41, W10419, doi:10.1029/2004WR003445.

Robson and Reed (1999) Statistical procedures for flood frequency estimation. Flood Estimation Handbook, Vol 3, Institute of Hydrology, Wallingford, UK.

Rosbjerg, D. and Madsen, H. (1995) Uncertainty measures in of regional flood frequency estimators. Journal of Hydrology, 167 (1-4), 209-224.

Stedinger, J. R. (1983) Estimating a regional flood frequency distribution. Water Resources Research, 19(2), 503-510.

Stedinger, J. R., Vogel, R. M. and Foufoula-Georgiou, E. (1993) Frequency analysis of extreme event, in Handbook of Hydrology, edited by D. Maidment, chapter 18, McGraw-Hill, New York.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 2845, 2009.