Hydrol. Earth Syst. Sci. Discuss., 6, C1355–C1374, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C1355/2009/

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.

B. Guse

bguse@gfz-potsdam.de

Received and published: 29 June 2009

We would like to thank the anonymous Referee#1 for the very detailed review and the useful comments. These comments should enable us to improve our manuscript significantly. We listed below our point-by-point replies to the comments of the Referee#1.

Main comments

1) 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.

We are thankful for this very helpful comment. In the revised version of our manuscript, we enlarged the description of the method of probabilistic regional envelope curves (PREC). The description of all steps of the PREC concept was more structured (see our replies to Referee's comments 14-23). In this way, we emphasised the main ideas and assumptions of the PREC concept. In our study, we did not modify the equations of the PREC method. Therefore we referred to the original studies (Castellarin et al., 2005; Castellarin, 2007) for a more detailed description of the backgrounds of the PREC concept.

2) A more structured introduction to regional flood frequency estimation and the influence of intersite correlation of flood data would be useful.

We reorganised the introduction to improve the description of regional flood frequency analysis (RFFA) and the influence of intersite correlation on RFFA. The revised introduction briefly listed and recalled the different methods of RFFA (index flood, linear regression models ...). We started the introduction to the influence of intersite correlation on RFFA by the studies of Matalas and Langbein (1962) and Stedinger (1983) and Stedinger and Tasker (1985). We referred to further studies, which considered intersite correlation. According to Referee#2, additional references were incorporated in the revised manuscript. In this way, the novelty of our study was clarified with respect to published studies on the impact of intersite correlation on RFFA and previous PREC studies.

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

See our replies to the examples below.

4) 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).

We modified the notation and wrote short, more elegant notations. For example we used neff consistently instead of EFFOBSNESTED or RN instead of RNESTED.

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

We are improving the language of the manuscript, also with the help of an expert.

6) 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.

This comment is critical, because it clearly indicates that the main focus of our study was not presented clearly enough in the original version of the manuscript. We modified the manuscript accordingly, in particular we better underlined especially in the introduction and in the presentation of the results that our goal was an investigation of the impact of different parameter sets for nested and unnested catchment structures on the recurrence interval derived by the method of PREC. We compared a global approach using one parameter set for the cross-correlation function (Eq.1) for all sites with a nested approach using separate parameter sets for nested and unnested catchment relationships. Since the cross-correlation function represents a very influential part of the PREC concept (see Castellarin, 2007), it is important to quantify the impact of different parameter sets. We agree with the Referee that our analysis was based on the equations in section 2 (Eqs. 1, 2 and 4 of the original manuscript). These equations were used to calculate the recurrence interval of PREC for each pooling group. A validation of the PREC results or a discussion of bias or variance of flood quantiles was not in the purpose of our study. The PREC method has been validated in comparison to the index flood method in the original PREC studies (Castellarin et al., 2005, Castellarin, 2007).

Specific comments:

Abstract:

7) Consider removing the first sentence.

Removed

8) Line 7: Consider “located in Saxony” rather than “belonging to Saxony”.

Changed

9) Line 19: Consider “degree of homogeneity” rather than “homogeneity degree”.

Changed

1. Introduction:

10) 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).

We clarified the introduction to regional flood frequency methods by a clear separation in the index flood method and linear regression models including relevant references such as suggested by the Referee. Since these methods were not in the focus of our study, we think that a short recall is adequate for our study.

11) 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.

Following the recommendation of the Referee, we restructured the literature review in order to improve the understanding and the significance of our results. Also according to the comments of the other Referees, we clarified the significance and novelty of our results with respect to published studies on this topic. The importance of intersite correlation for RFFA was emphasised by referring to studies, which assessed a larger uncertainty due to intersite correlation or highlighted a better performance by a consideration of intersite correlation (e.g., Stedinger and Tasker, 1985; Madsen and Rosbjerg, 1997; Kjeldsen and Jones, 2006). We focused the revised manuscript of our study more on PREC. In this way, the significance of our results were presented with respect to the original PREC studies (Castellarin et al., 2005; Castellarin, 2007) and the published studies on intersite correlation and nested catchment structures. As the Referee has mentioned earlier (see comment 1), the PREC-concept is a relatively new method. Therefore an application, which considers the different aspect of the concept, is helpful to check the PREC method.

12) 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).

We agree with the remarks of the Referee. The sentence (page 2847, line 20) has been dropped in the revised manuscript when restructuring the introduction. The Referee's

remarks are very helpful and we considered them also while restructuring the introduction (see our reply to the Referee's comment 11). Since our study was more focused on a separation in nested and unnested catchment relationships than on a discussion of the relevance of intersite correlation for regional flood frequency analysis, we clarified our core ideas in comparison to these other studies in the revised manuscript.

13) 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).

The validity of index flood hypothesis in the pooling group is a fundamental theoretical assumption of the PREC concept (Castellarin et al., 2005; Castellarin, 2007). In this study, the index flood method was not used explicitly. The index flood method represents a well-known method (see e.g. Stedinger and Lu, 1995; GREHYS, 1996; Robson and Reed, 1999). For these reasons, we think that we should not explain this method in all details. Anyway, we clarified the role of the index flood method as a fundamental assumption of the PREC concept in the revised manuscript.

2. Methods:

2.1 Regional information content and number of effective observations

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

This is an important remark. We explained the calculation of the number of effective observations on Page 2851 (original manuscript). However, apparently an improvement of the description is necessary, which was implemented in the revised version. Especially the explanation of the different terms in the Eq. 2 was clarified. The different steps of the method were clearly marked. This should enable us to give more structure in the description and help the reader to understand our proceeding. We referred to Castellarin et al. (2005) and Castellarin (2007) for a detailed explanation of the backgrounds.

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

We clarified this aspect. The length of the overlapping time series was used in the optimisation process to fit the parameters of the cross-correlation functions.

16) 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.

We are aware of a simple exponential function, which can also adequately describe the correlation-distance relationship. Castellarin (2007) compared the formula by Tasker and Stedinger (1989), which was used in our study, with a simple exponential formula. The study showed that "the possibility to model peak flow correlations allowing them to decrease very slowly as a function of distance offered by formula (6) (T&S) [two parameter correlation formula proposed by Tasker and Stedinger (1989)] is beneficial to the reliability of the PREC flood quantiles.." (p9, passage 48). In this study, we adopted the method of PREC (Castellarin et al., 2005; Castellarin, 2007). This study was focused on a consideration of the impact of different parameters for the same cross-correlation function on PREC results. Therefore we did not change any other aspect of the original PREC concept.

17) 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?

Hopefully, this aspect becomes even clearer after restructuring this section (see our reply to the Referee's comment 14). The number of effective observations was calculated year-wise (see Castellarin, 2007). This means that in a first approach the number of effective observations was calculated for each year separately by using all sites with

discharge data for this year. When only one site had a discharge measurement available for a given year Y, the number of effective observations for this year was one and therefore identical with the number of total observations. The fact that this observed value can be correlated to other unobserved values at different sites in the region does not alter its information content for the regionalisation approach (PREC, but also others): since there is only this discharge observation available for year Y, the regional information content of the observation is equal to 1.

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

It is true. We changed it.

19) 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.

We clarified this description of this equation in the revised form. LS signified the number of sites, which had discharge values in the year under consideration. As mentioned above (replies to the Referee's comment 14 and 17), we restructured this section by also explaining this aspect in detail.

2.2 Probabilistic regional envelope curves (PREC)

20) 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.

We agree with the Referee and we improved the description of PREC method. For example, the relevance of the index flood method for PREC was emphasised. As mentioned earlier (see replies to Referee's comment 1 and 14), we referred to Castellarin et al. (2005) and Castellarin (2007) for the backgrounds of the PREC method. We thought about a detailed description in an appendix, however we prefer referring to the original PREC studies.

21) 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, . . .'.

We adopted the suggested recommendation.

22) 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)?

We agree with the Referee's comment that the second principle leads to Eq. (3). This relationship was clarified in the revised manuscript. Furthermore we made clear that a homogeneous pooling group is a prerequisite for the application of the PREC method. According to the PREC concept, the Eq.(3) is only valid in homogeneous regions.

23) 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.

We changed the sentence in: "... assignment of an exceedance probability to that particular data pair of a specific unit flood of record and its drainage area,...".

24) 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.

The maximum and minimum possible values of neff depend on the number of sites in the pooling groups and their cross-correlation relationships. Furthermore the PREC method assumes that the pooling group of sites is homogeneous (see our reply to

Referee's comment 22). Castellarin et al. (2005) illustrated in a theoretical approach how the recurrence interval was affected by different numbers of sites with concurrent series of length under consideration of different average cross-correlations (see Figure 9 of Castellarin et al., 2005). We agree with the Referee#1 that the readability and the understanding of the manuscript should be improved by dropping section 4.3.1 and adding the relevant information to the methodical part of the manuscript.

2.3 Pooling scheme

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

PREC is based on the validity of the index flood hypothesis. This implies that the regions have to be homogeneous in terms of the index flood hypothesis. Castellarin et al. (2005) remarked that "...we expect the practical utility of RECs derived for highly heterogeneous regions to be limited, because a few of the discordant sites, i.e., sites for which the distributions of floods have very thick tails, are likely to dominate the REC." [page 9, passage 41] We emphasised the comment of the Referee in the revised version. By reorganising the chapter 2.3, the importance of homogeneous regions for the PREC concept was clarified.

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

We selected the term "behavioural subset" to reveal that several subsets of catchment descriptors have a similar performance. In our study, all subsets with a performance larger than the selected criteria (correlation coefficient to the empirical index flood values >0.6) were used. The term "behavioural subset" was introduced by Beven and Freer (2001) and considers the equifinality of parameter sets. We see that the term "behavioural subset" might lead to a misunderstanding. Therefore we replace it by "candidate set of catchment descriptors" in the revised version of our manuscript.

27) 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.

This is another very useful comment, which helped us to improve the clarity the manuscript. In our study, we used all combinations of catchment descriptors, which were illustrated in Tab.2 in the original manuscript. Therefore, all thirteen catchment descriptors were standardised (mean=0, std=1) before summing up them to all possible combinations of 1, 2 or 3 catchment descriptors. All combinations with a correlation coefficient to the empirical index flood values larger than 0.6 were selected as candidate set of catchment descriptors (see our reply to Referee's comment 26). Therefore there were no different scales and the similarity measure could be applied as it is. We restructured the section 2.3 and described the derivation of the pooling groups more in detail. In this way, the understanding and relevance of this section (derivation of pooling groups) should be more precise.

2.3.1 Homogeneity test

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

The impact of intersite correlation on the heterogeneity measure (Hosking and Wallis, 1993) was discussed by Castellarin et al. (2008). They demonstrated that intersite correlation leads to a decrease of the H-value. Then it is possible that a heterogeneous cross-correlated region is seen as homogeneous. The higher the intersite correlation, the higher is the risk of a categorisation error. Castellarin et al. (2008) stated that "Since the synthetic sequences are uncorrelated by definition, the sample variability of L-moment ratios for the synthetic group of sites is expected to be higher than the sample variability for the original sequences when cross-correlation is present, even when real and synthetic regions are characterised by the same degree of heterogeneity"

(p.69, line 8). Consequently a high intersite correlation results in a linear displacement of the H-value resulting in a lower H-value (see Fig. 6 on page 72 in Castellarin et al., 2008). We clarified this aspect in the revised manuscript.

2.4 Application and interpretation. . .

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

Changed

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

We homogenised the notation and always use neff instead of effobs. Following Referee's remarks, we avoided FORTRAN-like notation. Examples of these modifications were given above in the replies to the main comments (see Referee's comment 4).

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

We merged the Eqs.7 and 8. According the comments of Referee#3, we dropped the separate sub-subsections. In the revised manuscript, the section 2.4 has no separate sub-subsections.

3.Study area:

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

We adopted the recommendation.

33) 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?

The removal of catchments with large portions of catchment area out of Saxony was due to data availability. This study was part of a project in Saxony. Therefore, the selection of an administrative boundary resulted from the project. The catchment descriptors were applied as average values for all catchments. Therefore it was not practicable to use catchment descriptors covering only a small part of the catchment.

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

Modified, also according to a comment of Referee#3.

35) Page 2857, Line 23: SRTM DEM?

In the revised manuscript, we wrote: " Digital Elevation Model from the Shuttle Radar Topographic Mission (SRTM) "and we included a reference.

3.1 Selection of catchment descriptors

36) 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.

Actually, the catchment descriptors were used to derive the pooling groups. We added a sentence in the revised manuscript, in which we explained why we used the catchment descriptors. The correlation to the index flood was selected because PREC is based on the index flood method. The empirical index flood values were used to derive the slope of the regional envelope curve in the PREC concept. The description of the pooling groups (section 2.3) was enhanced in the revised version of our manuscript (see also our reply to the Referee's comment 27).

4 Results

Section 4.3:

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

As we mentioned above (see our reply to the Referee's comment 24), section 4.3.1 was dropped in the revised form. However the section 4.3 with the distance-cross-correlation plot (Fig. 5) was a relevant aspect of our study and should be maintained. The large correlation coefficients for nested catchment relationships became clear in this figure. Figure 5 demonstrated why it is relevant to separate nested and unnested catchment relationships. The results presented in the next sections (4.4 to 4.9) are based on the differences illustrated in Fig. 5. We are convinced that this section was a major point of our study and that it was indispensable to understand our manuscript.

Section 4.4:

38) 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).

In Figure 6a, the effective observations were related to the total observations for each pooling group, which was used to derive a probabilistic regional envelope curve. The ratio of the effective observations to the total observations (information content) was presented in Fig. 6b. The number of effective observations resulted from the Eq. (2) under consideration of the correlation function (Eq. 1) for the different catchment structures (global, nested and unnested). We had not the intention to illustrate a validation with these plots. And we agree with the Referee that it was expected that the information content decreased with increasing total number of observations. However, this study was focused on the impact of different parameter sets of the correlation function on the recurrence interval of PREC. Therefore we were especially interested in the differences for both approaches (nested-unnested vs. global). According to the Referee#3, we dropped the Figure 6b in the revised version.

39) 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.

We are aware that there are some variants of the Region of Influence (RoI) approach. The size of a RoI might be determined by the threshold in the Euclidean space (as it is done in our study), by a stepwise increase of the RoI until a defined heterogeneity measure (suggestion of the Referee) or by a target recurrence interval (see e.g. Pfaundler, 2001). We compared our RoI approach with the RoI variant suggested by the Referee. However we did not find significant differences concerning our main conclusions. Therefore we reported in this study only the results referring to the first RoI variant (threshold in Euclidean space). A region with all 89 sites is largely heterogeneous according to the heterogeneity measure.

40) Granted, the homogeneity measure might be larger than 2, but that seems not to be of any relevance to the evaluation of Eq. (2)?

A PREC could only be derived using the equations 1, 2 and 4 for homogeneous regions according to the fundamental assumption of the PREC concept. This includes a validity of the homogeneity criteria in terms of the index flood hypothesis. In this study, we used a threshold of 2 for the heterogeneity measure to drop all heterogeneous regions. The validity of the Eqs. (1), (2) and (4) is only given when the homogeneity criteria is fulfilled (see also our replies to Referee's comment 13).

41) Alternatively, consider removing section 4.5.

We are convinced that the section 4.5 is an important point of our study, because it presented the differences between the effective and total number of observations of

the different applications of the cross-correlation function. In addition, it showed the impact of cross-correlated sites and the reduction of the information content.

Section 4.6:

42) 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)?

We agree with the Referee that the last sentence has been written the wrong way around. Anyway, this sentence was removed from the revised manuscript, because the section ‘Results’ was more focused on the core idea of the study and less on the effect of heterogeneity according to the remarks of the Referee#2 and #3. A larger return period resulted from larger number of sites and from the correlation structure. In the case of the nested-unnested approach of the cross-correlation function, the return period was also depending on the degree of nesting within the pooling group.

Section 4.7:

43) 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. . .’?

We agree with the Referee#1 that the differences in neff resulted from the comparison of cross-correlation functions with different parameter sets. Our study aimed to assess the impact of the different parameters on neff and the recurrence interval. Therefore, the results illustrated in section 4.7 quantified the effect of the degree of nesting on the PREC results. This impact was relevant for a degree of nesting larger than 0.2. In the revised manuscript we removed the last sentence in order to get a more focused presentation of the results.

Section 4.8:

44) 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.

We removed the first sentence. And we used this comment to improve this section. In the revised version of our manuscript, we mentioned that an increasing variance of L-CV reduces the H1 value.

Section 4.9:

45) 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?

According to the comments of all Referees, we decided to reduce the section 4.9 to one example. Then, the focus of our study should become clearer. We clarified that we investigated different thresholds of the heterogeneity measure on the formation of pooling groups and on the recurrence interval of PREC. Therefore we dropped the formulations "the effect of regional heterogeneity on intersite correlation". In accordance to all three reviews, we dropped the results reporting to the threshold H1<4 and reduced the section 4.9 to the comparison of H1<1 and H1<2. The comparison of two different threshold of the heterogeneity measure, which represented different degrees of homogeneity, coincided with the original PREC concept and homogeneity criteria of the index flood hypothesis.

5: Discussion

46) Page 2866: Consider removing line 15-22.

In the discussion, we related our results to general hydrologic statements. Therefore

we discussed the relevance of a separation in nested and unnested catchment relationships in relationship to the spatial extent of flood events (local floods vs. regional wide-spread floods). The lines 15-22 were part of this discussion. Therefore we think that it was a valuable part, which should be maintained.

47) 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.

This part did not comprise any supplementary investigations. Therefore, we were not able to perform the test proposed by Referee#1. In this section, we discussed the effects of our results in different hydrologic situations (local floods vs. wide-spread floods).

6. Conclusions

48) 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).

We agree with the Referee and modified this sentence.

References:

Bayazit, M. and Önöz, B.: Sampling variances of regional flood quantiles affected by intersite correlation, J. Hydrol., 291(1-2), 42-51, 2004.

Beven, K.J. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249(1-2), 11-29, 2001.

Castellarin, A.: Probabilistic envelope curves for design flood estimation at ungauged sites, Water Resour. Res., 43(4), W04406, doi:10.1029/2005WR004384, 2007.

Castellarin, A., Burn, D.H. and Brath, A.: Homogeneity testing: How homogeneous do heterogeneous cross-correlated regions seem?, J. Hydrol., 360(1-4), 67-76, 2008.

Castellarin, A., Vogel, R.M. and Matalas, N.C.: Probabilistic behaviour of a regional envelope curve, Water Resour. Res., 41, W06018, doi: 10.1029/2004WR003042, 2005.

GREHYS: Presentation and review of some methods for regional flood frequency analysis, J. Hydrol., 186(1-4), 63-84, 1996.

Hosking, J.R.M. and Wallis, J.R.: Some statistics useful in regional frequency analysis, J. Hydrol., 29(2), 271-281, 1993.

Hosking, J.R.M. and Wallis, J.R.: Regional frequency analysis: an approach based on L-moments. Cambridge University Press, New York, 1997.

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

Kjeldsen, T. R. and Rosbjerg, D.: Comparison of regional index flood estimation procedures based on the extreme value type I distribution, Stoch. Env. Res. Risk A., 16(5), 358-373, 2002.

Kroll, C. N. and Stedinger, J. R.: Regional hydrologic analysis: Ordinary and generalized least squares revisited, Water Resour. Res., 34(1), 121-128, 1998.

Madsen, H. and Rosbjerg, D.: Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modelling, Water Resour. Res., 33(4), 771-781, 1997.

Matalas, N.C. and Langbein, W.B.: Information content of the mean, J. Geophys. Res., 67(9), 3441-3448, 1962.

Pfaundler, M.U.: Adapting, analysing and evaluating a flexible Index Flood regionalisation approach for heterogeneous geographical environments, Dissertation, ETH Zürich, 2001.

Reis D.S., Jr., Stedinger, J. R. and Martins, E. S.: Bayesian generalized least squares regression with application to log Pearson type 3 regional skew estimation, Water Resour. Res., 41, W10419, doi:10.1029/2004WR003445, 2005.

Robson, A. and Reed, D.: Flood Estimation Handbook 3: Statistical procedures of flood frequency estimation, Institute of Hydrology, Wallingford, UK, 1999.

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

Stedinger, J.R.: Estimating a Regional Flood Frequency Distribution, Water Resour. Res., 19(2), 503-510, 1983.

Stedinger, J. R. and Lu, L.: Appraisal of regional and index flood quantile estimators, Stoch. Hydrol. Hydraul., 9(1), 49-75, 1995.

Stedinger, J. R., and Tasker, G. D.: Regional hydrologic analysis. 1. Ordinary, weighted, and generalised least squares compared, Water Resour. Res., 21, 1421-1432, 1985.

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

Tasker, G.D. and Stedinger, J.R.: An operational GLS model for hydrologic regression, J. Hydrol., 111(1-4), 361-375, 1989.

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