Hydrol. Earth Syst. Sci. Discuss., 7, C2692–C2701, 2010
www.hydrol-earth-syst-sci-discuss.net/7/C2692/2010/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models" *by* B. A. Botero and F. Francés

B. A. Botero and F. Francés

ffrances@hma.upv.es

Received and published: 5 October 2010

Nice to see reactions from the audience! Thanks John for your comments. I will try to answer your very good questions.

GENERAL COMMENTS

"Upper tail function looks like"

See my comment about figure 2.21 below.

"Physically basis of upper bound parameter"

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

In the three cases, the upper bound parameter is the PMF. The physical basis for the PMF is the existence of an upper bound for precipitation (PMP) and the limited size of any catchment. I took in a Pepe Salas' lecture (long time ago, when I was his student), the next Horton's words: "A small stream cannot produce a major Mississippi flood for much the same reason a barnyard fowl cannot lay an egg a yard in diameter" (I couldn't find the exact reference). For me, it is clear.

"Physically basis for LN4"

The paper of Naghetini et al. (1996) is a very important one for several reasons. From my point of view mainly because it gives statistical and physical meaning to the interesting French methodology called GRADEX. Simplifying may be too much, GRADEX (and its modifications AGREGEE and more recently SCHADEX method) is a way of transferring information from extreme precipitation to flood frequency analysis. However, there is not a direct application of Naghetini et al. (1996) work in our paper (first problem: we are using a parametric approach with upper bounded functions) and I cannot see how to justify physically the LN4 from this transference. Which it can be more feasible is apply the Naghetini et al. (1996) complete methodology, assuming upper bounds in the precipitation and floods. It is a good idea, but we will work on it in the future.

"Combining results" in a Bayesian framework (O'Connell et al., 2002) is a good way to avoid the inherent error in the parametric function selection, but to apply the O'Connell et al. (2002) methodology was not the objective of our paper. We were aware of this problem and, in fact, this was the aim of our Robustness Analysis section.

SPECIFIC COMMENTS

"Fernandes et al. (2010) paper"

Because is a very recent paper, we did not have the opportunity to read it before our HESSD paper. From a first "light" reading, I think the work of Fernandes et al. (2010)

HESSD

7, C2692–C2701, 2010

Interactive Comment

is in some sense a continuation of our paper, introducing the PMF uncertainty through a Bayesian framework. A very good job. We will study it and, if appropriate, we will incorporate their results/conclusions in our final version.

"Page 5415, last line onto next page"

Yes, you are right. We want to mean "it is easy to extrapolate a parametric distribution". We need to clarify it in the final version.

"page 5416, near line 10. Here a citation to Enzel et al. (1993) is appropriate"

Thanks! It is difficult to find good scientific references for the PMF concept. We will add this reference for sure in the final version. Concerning the work of Koutsoyiannis (1999), we studied it, but as you said, he assumes no upper limit and tries to assign a return period to the PMP, which is a conceptual contradiction. At the end, to assign a return period to the PMP or PMF is a way of estimating their uncertainties. In my opinion, to address this objective properly you must assume the existence of the upper limit.

"Paper's Section 2 versus Chapter 2 of Botero's dissertation: Figure 2.21"

It is difficult to decide from an excellent work with 245 pages which concepts, ideas and results should be summarized in a scientific paper with 26 equivalent pages. Chapter 2 in Botero's dissertation is dedicated to a deep analysis of statistical properties of the 3 upper bounded distributions EV4, LN4 and TDF. Section 2 in our paper is a brief summary of that chapter. In particular, figure 2.21 of the Botero's dissertation (attached as figure 1) try to show their different behavior when approaching to the same upper bound. It is a very interesting property to be in mind during the model selection. In the present paper is described in words in line 20 page 5425 and in figure 2. However, only in figure 2b (ML-PG estimation method) the upper bound is the same for the three distribution functions, but the upper limit is out of the graph. It can be a good idea to incorporate it in the final version of our paper, instead of figure 2.21, the complete

HESSD

7, C2692–C2701, 2010

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

figure 2b showing the common upper bound.

Figures 2.7 and 2.8 of the Botero's dissertation (also attached as figures 2 and 3) show the coefficient of variation and skewness coefficient respectively as a function of the EV4 parameters. There are equivalent figures for the LN4 and TDF distributions. The EV4 is the selected distribution from a descriptive point of view, in the paper's case study and in others presented in the Botero's dissertation. In order to do not increase unnecessarily the main paper size, we will think how to include in an annex of the final version some statistical characteristics of the three distributions.

"Other distributions with upper (and lower) bounds"

In this way our work will never be ended! The main objective of our paper is to "put on the table" the possibility of using upper bounded distributions, when you are dealing with very high return period quantiles, because its influence on them. The already discussed figure 2.21 will be important from this point of view. What is not the objective of our paper is to have a complete review and comparison of extreme value distribution functions.

We couldn't access these days to the paper from Nathan and Weinmann (1999), published in Australian Journal of Water Resources, neither the research report from Siriwardena and Weinmann (1998). So, I cannot comment them.

"page 5420, Section 3, data classification"

The censoring threshold can be different each year: the year in equations 6, 7 and 8 is a subindex for the upper (U) and lower (L) censoring. On the other hand, it is true in figure 1 the censoring thershold is fixed. We fixed it for the sack of simplicity. We can add in the figure caption "for a fixed threshold" or change the figure as John England suggests.

Concerning the "credit", of course, the first attempt of historical information classification was done by Stedinger. What was new (at least for me) in the Naulet's dissertation

HESSD

7, C2692–C2701, 2010

(I was in his PhD tribunal) is the consideration of each year as an independent piece of information, which facilitate to have a different censoring threshold each year. We will clarify this point.

“page 5421, last line before section 4.1.”

No, what is wrong is the previous line. It should be “parameters set”.

“page 5422, lines 2-3. PMF estimates are nonstationary”

Unfortunately we know in the long run floods are non stationary due to Global Changes (climatic plus hydrological). In this paper the hypothesis is “stationarity” and we leave non stationarity for future work. But I think what John England is saying is other thing: PMP and PMF estimation has a large uncertainty. See the comment below about the G uncertainty.

“page 5422, equation (9)”

As far as I know, the answer is no. Equation 9 is an estimator of the distribution upper limit and g must be in both sides of the equation.

“page 5423, equn (12) using a pdf for g”

All flood information has errors and I agree it seems G (the a priori deterministic computation of PMF) has more uncertainty than the rest of the information. John England’s comment about including explicitly the G uncertainty is reasonable. In fact, we use this way (in our paper G is a normal random variable with some bias, instead being the lower limit of a LN3 as suggested by John England) not in the model, but in the Monte Carlo uncertainty analysis in our Section 6 for the ML-PG estimation method.

“page 5424, line 21. To avoid a bias estimation in the number....”

We don’t begin the historical period with the first flood date, precisely because the results of the Hirsh and Stedinger (1989) described in their appendix A. Their conclusion was clear for us: ”Clearly, the estimation of n based on the date the first extraordinary flood occurred exacerbates the severe imprecision of any of the plotting position formulas and the severe upward bias that exists in the W formula and less so in the E and B formulas."

And they recommended: "Every effort should be made to establish n accurately on the basis of the quality of historical evidence and not on the basis of the occurrence of the first extraordinary flood."

However, their recommendation cannot be followed in most cases, not in our case study. What we are doing to reduce this problem is to eliminate the first historical flood and consider the beginning of the historical period in the middle year between the first and second historical floods. However, we cannot assure all bias is eliminated with this method. So, in the final paper, we must cite Hirsh and Stedinger's paper (because we used their conclusion) and add a proper discussion. Also, it must be mentioned in this case study the sensitivity of the results to this decision is small (maximum EV4 estimated quantile change is 5%).

page 5425, line 22: Sentence starting with "A very interesting first conclusion..."

Yes. See my comment for figure 2.21, which will be the "behaviour generalization" in the final version.

page 5426, line 1: delete "d" in "appropriated"

Ok

"page 5427, about line 10. The PMF uncertainty estimates chosen by the authors are too small"

John England finds more frequently PMF underestimation of about 44%. We are assuming a 10% positive bias (overestimation) plus a coefficient of variation of 30% with a normal error distribution. It means a theoretical error in terms our eq. (13) of 32%, which is in the same magnitude (i.e., not too small), but with slight more frequent overestimation than underestimation (i.e., not in accordance with John England expe-

HESSD

7, C2692–C2701, 2010

Interactive Comment

rience). Of course, John has much more experience than us dealing with PMF estimation, but what I can say from our results is the values for bias and coefficient of variation don't change the main conclusions (at least for the EV4):

i) The uncertainty for high return period quantiles using ML-PG method is controlled by the uncertainty in G (the a priori deterministic estimation of the PMF), as it is shown in figure 3 in our paper.

ii) And, unfortunately, introducing a prior estimation of G with relative small errors (as we are using in our paper) is worst than don't use it (figure 3 of our paper).

We will better state these conclusions in the final paper.

"page 5430, line 14. The results are dependent on your EV4 distribution assumption"

The caution is already there: we refer to EV4/ML-PG model. On the opposite: I am almost sure (by intuition of 20 years of flood frequency analyses) this result is general. But we will maintain the caution.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 5413, 2010.

HESSD

7, C2692–C2701, 2010

Interactive Comment

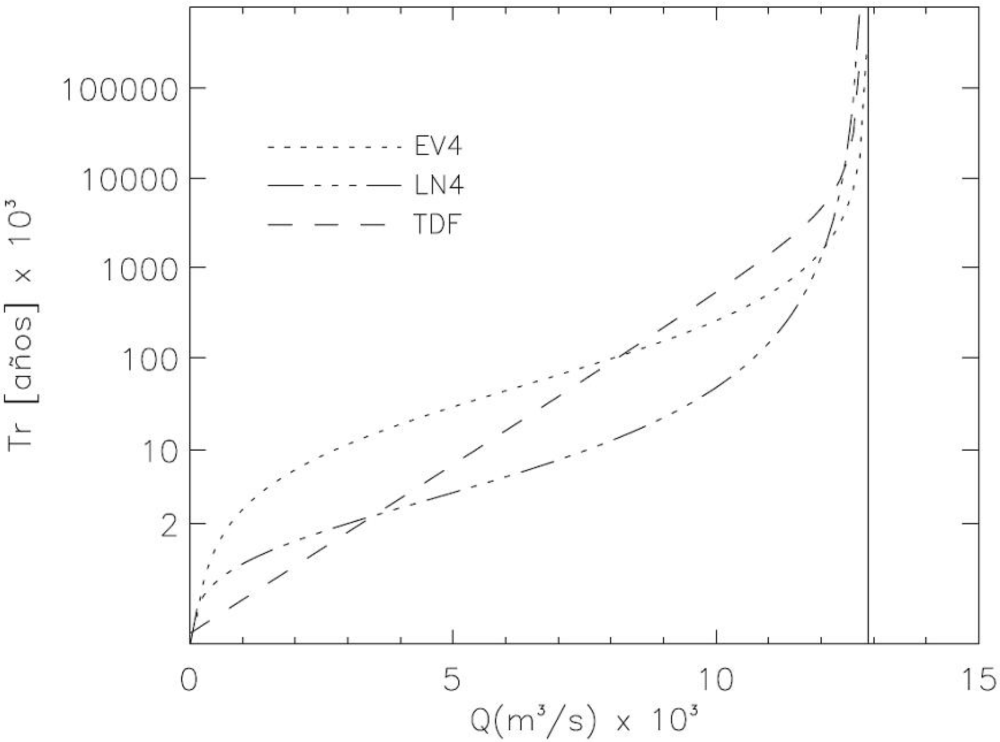


Fig. 1. Figure 2.21 in Botero's dissertation. How the three upper bounded distributions approach to the same upper limit "g".

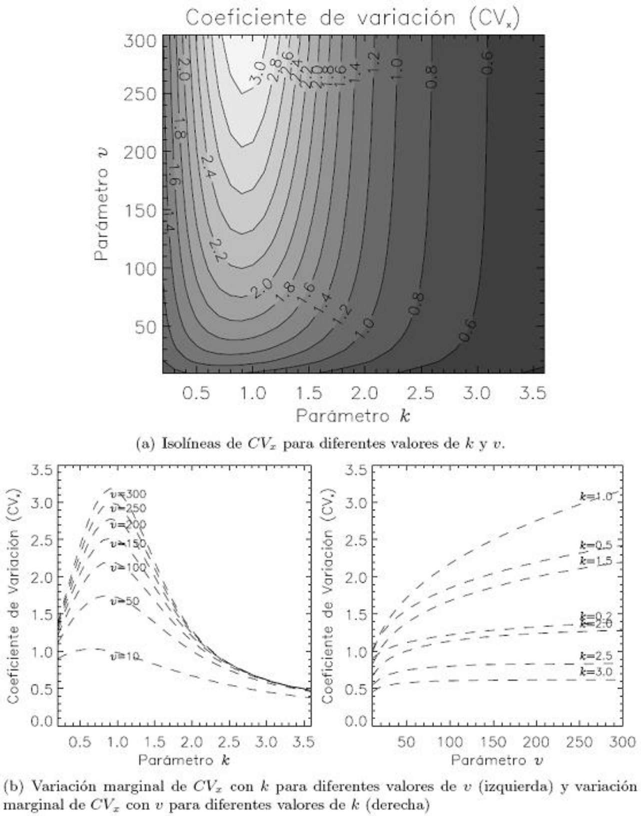


Fig. 2. Figure 2.7 in Botero's dissertation. Coefficient of variation for the EV4 distribution as a function of parameters k and v, for g = 10 000 and a =1.

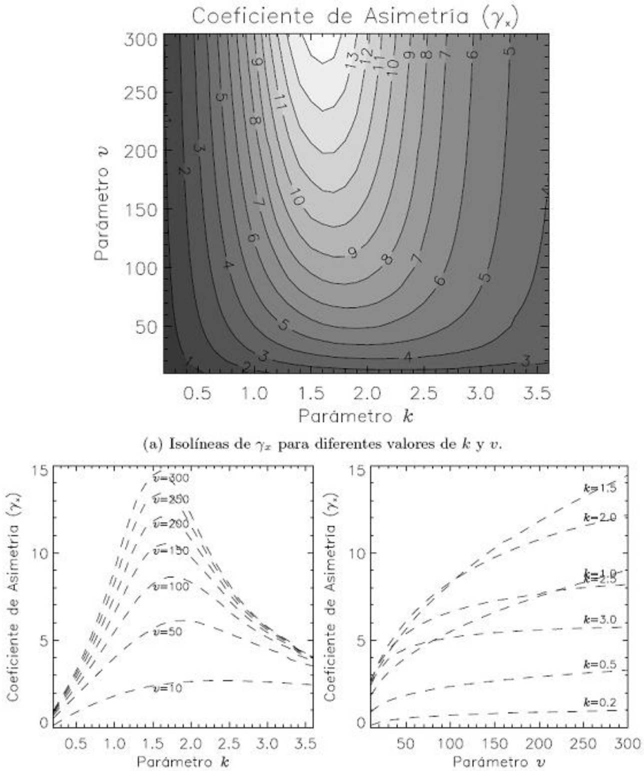
HESSD
7, C2692–C2701, 2010

Interactive Comment

HESSD

7, C2692–C2701, 2010

Interactive Comment



(a) Isolíneas de γ_x para diferentes valores de k y v.

(b) Variación marginal de γ_x con k para diferentes valores de v (izquierda) y variación marginal de γ_x con v para diferentes valores de k (derecha)

Fig. 3. Figure 2.8 in Botero's dissertation. Skewness coefficient for the EV4 distribution as a function of parameters k and v, for g = 10 000 and a =1.