Hydrol. Earth Syst. Sci. Discuss., 8, C3300–C3306, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C3300/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "The within-day behaviour of 6 minute rainfall intensity in Australia" *by* A. W. Western et al.

A. W. Western et al.

a.western@unimelb.edu.au

Received and published: 6 August 2011

We thank P. Willems for his thoughtful comments on our paper and respond to each in turn as follows.

***Willems**: The discussion on the choice between a two-parameter versus a three-parameter distribution could be complemented with a discussion on the balance between higher uncertainty in the parameter estimates (when more parameters are used) versus higher potential bias (when less parameters are used). Methods exist to quantify the parameter uncertainties, and the uncertainties on the quantile estimates obtained from the distributions. Because the final aim of the research by the authors is to obtain accurate distribution functions (thus quantile estimates) for applications (e.g. erosion* *modeling), low uncertainty in the quantile estimates might be a better criterion than the goodness-of-fit to observations (as considered by the authors). The authors might add a limited discussion on that, and add references to existing uncertainty quantification methods.*

We have added some discussion on the criterion for goodness-of-fit as an introduction to that section in the methodology. It is as follows: "There are two possible approaches to assessment of the performance of different distributions; either examining how closely the distribution functions fit the data by some sort of analysis of residuals from the distribution function, or examining the uncertainty in the quantile estimates resulting from the fitted distribution. To estimate the uncertainties in the quantile estimates requires either independent samples or a rigourous treatment of any temporal strucutre in the data. Rainfall over a day is both intermittent and exhibits (potentially intensity dependent) serial correlation. This structure would need to be incorporated into the uncertainty estimation for the parameters of each of the distributions and for each of the fitting methods. Because of this complexity we opted to examine the fits based on a residual analysis rather than uncertainty in the quantile estimates."

Responding to this point and reviewer 2, we have also added the following comment to the discussion of selection of distributions. Page 3207 line 18 to 24 is replaced with: "Based on their performance as measured by the goodness-of-fit statistics mCOE and RMSE90, we suggest that the two best performing TDFs were GPT3 and GPT2, where GPT3 has a slightly better fit but GPT2 has the advantage of only two parameters. In selecting between two and three parameter distributions there is likely a trade-off between higher bias in the two parameter distribution (due to less flexibility) and higher uncertainty in parameter estimation in the three parameter distribution. The main advantage of GPT2 over GAMA and EXP is that it outperforms GAMA and EXP at the higher intensities."

Willems*: * Also related to the discussion on the choice between a two-parameter and a three parameter distribution:*

HESSD

8, C3300–C3306, 2011

Interactive Comment

- The EXP can be seen as a special case of the GPT3, where the tail of the distribution is "normal" instead of "heavy" or "light" (shape parameter κ zero, negative or positive; where the EXP equals the GPT3 for κ=0). Methods exist to discriminate between these three classes, based on an analysis of the tail behavior of the distribution (e.g. upper tail in your Figures 2 and 3 asymptotically linear, bending up or bending down).

- Light tail GPT3 distributions would mean that rainfall intensities have an upper limit. This sounds unrealistic (from a meteorological point of view).

- The difference between GPT2 and GPT3 is the application of a location parameter, which typically is taken as a threshold value, above which the distribution is valid. Different methods exist to identify the optimal threshold value. This might be different from the threshold value prior and arbitrarily selected by the authors (probably much higher).

- Another related comment is that the EXP and GPT3 are also commonly applied extreme value distributions (while they are classified as non-EVD by the authors). In fact they are more applicable than the GEV and Gumbel in this case, because GTP3 (and EXP as special case) are limiting extreme value distributions of peak-over-threshold (POT) values.

In response to the fourth point in the methodology on page 3195 line 20 we included the sentence "Of these distributions it is worth noting that the generalised pareato distribution and its special case, the exponential distribution, can be interpreted as peak-over-threshold distributions (Madsen and Rosberg, 1997; Claps and Liao, 2003), which provides some theoretical justification for their suitability here."

In response to these points we added the following paragraph to our discussion of the selected distributions on page 3208.

"It is worth briefly discussing the results from a more theoretical perspective. First the GPT3 (and by inference its special cases) are peak-over-threshold distributions, which matches with the analysis undertaken here, albeit with a low threshold. Also the GPT2

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

and EXP are both special cases of GPT3, with GPT2 being equivalent to GPT3 with the location parameter set to zero and EXP being equivalent to GPT3 with κ = zero (Claps and Liao, 2003). Some inferences can be made from the fitted parameters for GPT3. First κ<0, κ=0 and κ>0 implies light, normal and heavy tailed distributions respectively. Light tails are not expected as they imply an upper bound, which is unlikely for rainfall intensity. We examined the results from both Melbourne and Darwin and found κ varied from slightly negative (-0.23 and -0.24 respectively) to strongly positive (a few values >1 and >2 respectively) with the average being 0.11 and 0.15 respectively. This indicates a slight tendency towards heavy tailed distributions. Second the location parameter, ξ, for GPT3 can be interpreted as a threshold above which the distribution holds. We thresholded the data a 1mm/h before fitting the distributions. For Melbourne and Darwin respectively, we found 35

Willems*: The application of the techniques to calibrate the distribution function parameters to data requires independent data. The authors did not discuss this independence requirement. They neither tested this assumption. Different 6-min rainfall intensities within the same day are probably strongly dependent. Same comment reg. the applicability of the Chi-square test.*

We have pointed out that this is an issue at the end of the discussion of distributions and fitting methods (page 3196, line 21). The actual paragraph added is: “It should be noted that there is temporal structure to within-day rainfall that involves both intermittency and serial correlation during rainfall periods. This structure impacts on fitting techniques and in particular uncertainty estimation for fitted parameters. In this study we have not attempted to estimate the uncertainty in the fit of parameters for each type of distribution because of this issue.”

We have also noted this issue where we describe the Chi-square testing by adding to page 3198 line 8. “It should be noted that any intermittency and serial correlation should be accounted for in implementing these tests. We did not do this and this means the power of the Chi-squared test is over-estimated (i.e. more days are found to be statistically different to the hypothesised distribution that is really the case given that there will be some temporal structure to the data). This is a limitation of the testing that was attempted."

***Willems**: * Some additional references that the authors might consider (and which address some of the issues mentioned above) are:*

- Claps, P., Laio, F., 2002 [should be 2003]. Can continuous streamflow data support flood frequency analysis? An alternative to the partial duration series approach. Water Resour. Res. 39 (8), 1216. doi: 10.1029/2002WR001868

- Madsen, H., Rasmussen, P.F., Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. At-site modelling. Water Resour. Res. 33 (4), 747–757

- Smith, R.L., 1987. Estimating tails of probability distributions. Ann. Statist. 15, 1174–1207.

- Willems, P., Guillou, A., Beirlant, J. (2007). Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function, Journal of Hydrology, 338, 221-236

We have included references to Claps and Laio (2003), Madsen et al. (2997) and Willems et al. (2007).

***Willems**: On page 3191 lines 11-12, the authors explain the application of their distribution function of 6-minutes rainfall intensities. The application of a threshold to the rainfall intensities will, however, deliver incomplete information on that distribution. The authors may comment on that. The conditional distribution (above the threshold) should at least be complemented with information of the frequency of low 6-min intensities (below the threshold).*

This information is already provided on page 3194 lines 23-25 and the insensitivity of the results to the choice of threshold is also mention on lines 25-27.

***Willems**: * Some minor comments:*

- Page 3194 line 17: how was a “valid day” defined?

This is defined on Page 3194, line 1 (it is simply a day with a complete pluviograph record); however, our terminology is clearly confusing since two referees asked the same question. We have therefore amended the text in places where “valid day” was used to indicate that we are referring to days with complete record, all of which were analysed.

- Page 3194 line 26: “the results”: which ones?

We have amended “results” to “the fitted parameter values and quality of fits”

- Page 3196 lines 4-5: Whether the rainfall intensity distribution is heavy-tailed depends on the data or location; also for EVDs (see also my previous comment on the discrimination between normal-heavy-light tails).

We have inserted “in some cases”

- Page 3196 lines 6-14 Page 3199 lines 1-11: Why do you spread the description of the goodness-of-fit statistics over two different sections of the manuscript? It would be better to combine these in one of the sections.

We have reviewed this and decided to keep the paper as it is. This is because subsection 2.51. discusses how results were summarised across all rain days at a station while the earlier part of section 2.1. discusses the statistics used for individual days.

- Page 3197: Should the choice of the rainfall variable under study (e.g. five highest intensity periods) not depend on the application (erosion modeling, urban flood modeling, ...; e.g. depending on the concentration time of the catchment)? Why did you choose the “five highest”?

Yes the choice should depend on the application. We chose this due to the application of interest as pointed out. We think the existing text is sufficiently explicit about this on

P3197. We also point out tha the results are insensitive to exactly which measure is chosen.

- Page 3197 lines 9-10 (and several other places in the manuscript): There is no need to repeat so often to the reader that the focus of your paper is on the high intensities

We have edited the text to remove this repetition.

- Table 3b: These are well known distributions. There is no need to give an overview of the mathematical formulae.

We have deleted table 3b and references to it in the text.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 3189, 2011.