Nat. Hazards Earth Syst. Sci. Discuss.,
doi:10.5194/nhess-2016-183-AC1, 2016

Natural Hazards and Earth System Sciences Discussions
Open Access

Interactive comment on "Effects of sample size on estimation of rainfall extremes at high temperatures" *by* Berry Boessenkool et al.

Berry Boessenkool et al.

berryboessenkool@hotmail.com

Received and published: 4 October 2016

1. The main issue raised by referee #1 is that a weighted average is computed from different distributions which might not be comparable per se.

We use a weighted average to better visualize the results from different distributions and to reduce the effect of distribution choice. We want to address the problem of choosing distributions / methods to single out unreasonable distribution choices from being included in the average. This will likely deal with the issues also pointed out by referee #1, like the suspiciously high results in figure 8 or the peak in figure 3, which stems from the Rice distribution. It might also explain why parametric quantiles in figure 5 do not converge to the 'true' value, as criticized by referee #2. As a strategy, we will inspect the sample size dependency of each distribution function and for several GPD

Printer-friendly version

Discussion paper

methods. We will repeatedly take random samples with different sample sizes from the full set of precipitation records. The 5 distributions showing the lowest dependency on sample size (the earliest convergence to the value for the full sample as per figure 6) will be selected for the weighted average. To make this selection more robust, the procedure will be repeated with different datasets.

2. Referee #2 points out that we combine many distributions, parameter estimation procedures, truncations and software packages and requests a more complete description of these methods. Referee #1 suspects that our main conclusions heavily rely on the choice of distributions, estimations, methods and R packages.

Part of our objective is to show exactly that different distributions and fitting methods yield different results, where some are more biased with sample size than others. We will try to write this in a more comprehensible way, as suggested by referee #2. Considering different methods and software packages (apart from using different distributions) has a pragmatic value. Readers should become aware that using the same distribution with the same estimation method (GPD MLE) might yield different results if different software implementations are used. This is worrying and unsurprising at the same time. We think it is beyond the scope of this paper to scrutinize all these implementations in detail, as suggested by referee #2. The main purpose is to show that sample size dependency may partly explain the drop in precipitation scaling. Our second aim is to point out that quantile estimation methods have an influence on the estimated P-T relationship. However, we will try to provide more background on the methods while remaining concise. We will try to better communicate that it matters to use distribution fitting methods and computational implementations that have the smallest possible sample size bias.

3. Referee #2 demands that we consider dew point temperatures in order to analyze whether the quantile drop is caused by moisture limitation.

In our opinion, this requirement is beyond the scope of this paper. At least for all our analyses with synthetic data, we "know" that there is no moisture effect. The aim of this paper is to point out the potential significance of sample size artefacts. We would like to leave it to other studies to actually discriminate between sampling and moisture effects, to avoid that this article becomes overloaded.

4. Referee #2 notes that L-moment estimates are usually less biased than MLE, and claims to find the opposite relation in Figure 6.

We show that 6 out of 7 MLE implementations demonstrate exactly the expected behavior, so we cannot quite understand this concern. We will add a note on the GMLE GPD estimate, the one exception, which seems to be connected with the singular behavior for GMLE in figure 4.

5. Minor comments from referee #1 that we will almost all implement: - revise figure 3: show only one percentile, deal with the peak at a truncation value of 0.05 (comes from the rice distribution, see point 1). "Other distributions" will not be removed from figure but referenced to in the text. - replace the word "quantile" with "percentile" - define super-CC scaling - add the note that precipitation events at higher temperature are generally convective in nature and very localized in space (and thus often missed by the observing network), resulting in smaller sample sizes compared to large-scale precipitation at lower temperatures

6. A few things we would want to add to help clarify point 2: We will describe in the figure captions that gpa in figure 3 is the same method as lmomco_LM_gpa in figure 4. Figure 4 shows that GPD MLE has no increase with truncation as shown with GPD PWM/LM. We believe this to be due to the fact that a higher truncation produces smaller samples, thus GPD MLE underestimates the true value increasingly (fig 6), thus shows no increase with truncation.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2016-183, 2016.