Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-586-RC1, 2016

Hydrology and Earth System Sciences Discussions

Open Access

Interactive comment on "Simple Scaling of extreme precipitation in North America" *by* Silvia Innocenti et al.

Anonymous Referee #1

Received and published: 22 December 2016

Overall comments:

The article is overall well written, and the scientific subject is very topic and well addressed. The literature review and scientific context given in the introduction and Section 2 are very well written! The data description is also very clearly outlined. I had difficulties understanding (some technical details of) Section 4, and consequently (some of the results of) Sections 5 and 6. Therefore most of my comments aim to improve (my understanding) of this section. (I had also some perplexities about the chosen regions, which I address in the major comments too). I hope my comments lead to an improvement of the manuscript.

Major comments:

1) It is very difficult for me to comprehend Section 4 (mainly page 7 and Figures 1 and

2):

1a). the definition of scaling interval needs to be clearer: from page 7, lines 7-8, I understand that for each data-set (SD, ID and LD) you define several scaling intervals. These intervals have fixed durations equal to 6, 12,18 and 24 times the reference duration unit d* (which is 15', 1h and 6h, respectively, for each data-set SD, ID and LD). I assume the 6, 12, 18 and 24 durations are associated to a dilatation factor of λ = 6, 12, 18, 24 (the same as in equations 2,3,4). The several scaling intervals you consider have all the same length (duration) but are distinct for their initial time. If this is correct, the text at line 11 (and thereafter) need to be revised: I suggest replacing "its first duration" with "its initial time", because using the word "duration" for both (initiation and duration) indices is very confusing.

1b). I am trying to understand how you estimate H: I understand that you use Eq. 4, with a fixed q and a fixed λ (say, 6h). You have to regress however on more than one estimated moment, so I assume you consider the aforementioned scaling intervals with the different initial times (e.g. for the ID data-set and 6h-duration you consider the 19 6h-long scaling intervals you show in Figure 1b, left panel). For each (of the 19) initial times, you select the annual max, and then an AMS, from which you compute the q-moment (so you have 19 q-moments for the LHS of equation 4). What do you use for the first term of the RHS of Equation 4? will you have 6 X 19 q-moments corresponding to the AMS of 1h accumulated precipitation within each 6h long scaling interval?

1c). Page 7, line 16 "and the corresponding durations" are different durations, or different initial times for a fixed duration λ?

1d). I am not sure I understand what are you testing with the slope test. My hypothesis is that you have evaluated the regression coefficient Kq = -Hq log(λ) (the equality is from equation 4). Since the final goal is to estimate H, you want to regress the Kq versus q (for all the q=0.2,0.4, .. ,3), with a fixed λ, to finally find H. Please explain this better.

1e). GOF test: Page 7, line 23, “for each duration d”: this time seems to refer to a real duration. Whereas four lines later “for all durations dj in the scaling interval” suggests dj are initial times. I do not understand what are you testing with the GOF test: what is the “pooled sample of the rescaled AMS x’_dj for all durations dj in the scaling interval”? I got completely lost in Eq. 8 and 9 . . . (I came back to this page few times, in separate days, to make sure my lack of understanding was not due to a particular bad moment. This is why I am describing in details what I do not understand, I hope this helps to point out what needs to be rephrased).

1f). Page 8, lines 10-11: here duration seems again to refer to a time duration. So, in my understanding, you consider 1h, 2h, 4h, 5h and 6h; from these you evaluate H; then you evaluate the 3h AMS with the SS. Then you evaluate the RMSE for the high quantiles of this synthetic 3h AMS versus the high quantiles of the empirical 3h AMS. I am not sure this is correct, but this is my guess . . ..

I think it would be very useful if you could show a figure with one (or maybe two) concrete example of your regression, for a fixed q and λ, so that the reader can better understand what youu regress, and what are the statistical tests you perform to be confident in your results.

2) When you perform your analysis, in my understanding, you consider scaling intervals of a duration of 6h (as an example) spanning the whole diurnal cycle as intervals with equivalent statistical properties. However, physically, precipitation occurring in the afternoon (which in the summer is triggered by convection) can present very different temporal structure (and physical properties) than precipitation occurring early in the morning: can you pull together these data, or (given that your focus is on extremes, which in summer often is related to convective events), should you consider a stratification based on the diurnal cycle? (The inhomogeneity of your phenomena along the dyurnal cycle might be also the cause of test rejections for longer scaling intervals, as you state at page 8, line 28-29)

3) Figure 1: your lowest confidence (largest proportion of rejected stations) is associated at the "durations" in the beginning of the scaling interval: why? This seems a sampling problem (is this what you "expect" in your statement at page 8, lines 26-27)? All the possible causes you list from page 8, line 29 to page 9 line 6 are plausible, but should hold also for 15m durations within the scaling interval (not just in the beginning). Possibly, my lack of understanding is linked to my lack of understanding on how the calculation is performed (page 7).

4) Figure 2 (and related text, at page 9, lines 15-21): your largest relative errors are always associated at the durations in the beginning and at the end of the scaling interval: similarly to Figure 1, is it possible that this is a sampling problem? (Or maybe, for longer time scaling, the inhomogeneity of the weather phenomena along the diurnal cycle plays a role ..)

5) Figure 3: I find it very difficult to understand the results shown in Figure 3:

5a) It is not well set what is the scope of this section is (in fact, at my first reading, I had the feeling it was a aimless technical analysis ... which instead is not). Later, I came to this hypothesis: in my understanding, H should be a scale-invariant parameter. Therefore large changes in H (aka large ΔH) are "bad", whereas if ΔH is near zero the SS model is a good approximation. Is this correct? Can you please state this clearly in the beginning of this section. Then, the reader will be able to search for the wanted results while analyzing Figure 3.

5b) Technical question: (the distribution of) H is computed for each duration (e.g. 1h, 2h, 3h, . . . in Figure 3 i-b). From Equations 2,3,4 I understood that H is scale invariant; then there should be one H which enables to describe all time scales. Why this is not the case? (Similar for the following sections, e.g Figure 8). Again, it is possible that my lack of understanding is linked to my lack of understanding on how the calculation of H is performed (page 7).

5c) I suggest as new title for Section 4.2: "Variability of the estimated scaling exponents” (to mirror the results shown in Figure 3, which shows ΔH).

6) Section 5: from the maps you show in Figures 4 and 5, the only two clearly distinct homogeneous regions (at a first eye analysis) seem to be SW_pacif and NW_pacif. I can see (the signal is more mild) also the regions C and E. It seems to me that the South-East of the United Stated could also be split in two regions (e.g. from Fig 4 ID and LD, and figure 5 ID). The Boreal region (as you conlude yourself at page 12, line 16) is very etherogeneous, and I do not see any reason for clustering these stations together (sole common factor is the network sparseness . . ..). Have you attempted a cluster analysis to define your own regions, rather than considering the Bukovsky regions?

7) Figure 7 and S4 are needed -in my view- solely for supporting the description of Fig.8d (page 12, lines 26-31). (The results related to the other panels are less interesting, inmy view). I suggest to move also Figure 7 in the supporting material, along with the text at page 11, lines 11-22.

Minor comments:

1) There is a typo at line 7 of the abstract: should be 15’, and not 15h.

2) page 4, equation 2: I suggest to eliminate “D” and explicitely write “λd” instead (the less symbols you introduce, the more readable is the article).

3) page 5, line 1: eliminate “and the frequency . . . F(x)”.

4) page 5, line 9: I suggest writing “Approaches aimed at increasing the sample size may be used . . . ”

5) page 5, Equation 7: notation is too complex, and should be simplified.

6) Page 6, lines 6-9: the cause-effect is not entirely clear to me: why two different periods where chosen (JJAS for the north and MJJASO for the south), rather the same period (either JJAS or MJJASO) for all stations?

7) Page 6, line 13-14 (and thereafter): rather than saying “coarser resolution” I suggest writing “more discretized recording procedure” (or something similar). The effect of recording discretizations are a well known problem in statistics, which can be bypassed simply by adding a uniformly distributed random noise (ranging between 0 and 2.54) to your data.

8) Page 6, line 30: this condition is not clear to me, please explain it more explicitely.

9) Figures 1 and 2: I suggest using a notation as 2h30’, rather than 2.5h. Similarly for the days, 2d12h (or 50h) rather than 2.5d.

10) page 7, line 10-12: “This procedure . . . evaluate the variability of the SS estimates . . . ” could it be the sensitivity instead? This text is not clear.

11) page 7, line 20: student t-test (the t is associated to the test, not to the student).

12) page 8, lines 14-20: r is usually used ijn statistics for correlation: it is possible to use a different notation? Please specify that the normalized RMSE is zero for a good fit, and it gets larger and larger for a worse fit.

13) Page 9, line 23: the acronym for inter-quartile range is, traditionally, IQR.

14) Figures 7,8 would be more easy to read if you put a title on each panel with the name of the region.

15) Page 11 lines 6-9: join this paragraph to the previous (they both pertain to the physical explanation of the different panels of Figure 8).

16) Page 12, lines 2-3: this apply to the SW_pac region as well.

17) Page 13, lines 10-12: I agree that scaling regimes are weaker for short d1 than for longer d1, however you need to rephrase the sentence at line 12. In fact, despite “smaller”, the scaling regimes for short duration exceed 0.5 (most of them 0.6). Therefore effect of the scaling factor (λˆ{-H}) is not negligible on the AMS distribution moments.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-586, 2016.