hess-2017-150 – At-site and regional frequency analysis of extreme precipitation from radar-based estimates – by Goudenhoofdt et al

The manuscript presents the use of different methods to perform rainfall frequency analysis from weather radar data. The topic is of increasing interest for the community given the growth of radar and remote sensing archives worldwide. Studies proposing methods, testing approaches and evaluating the accuracy of such products are greatly welcome and definitely of interest for the readers of HESS.

This study focuses on a region in Belgium and derives at-site and 'regional' frequency analyses for 1h and 24h durations and provides new contributions (not clearly evident from the text) to the field, such as the use of new (i) methods (i.e. a peak over threshold approach, QQ plot regression) and (ii) regionalization approaches for rainfall frequency analyses from remote sensing data sets.

This study will provide new, interesting information to the field and deserves publication. However, a number of issues are currently preventing it from being published in its present form: literature review is missing key papers that need to be mentioned and, in some cases, discussed; methods are not sufficiently described, motivated and supported by literature; some of the results need to be re-considered and discussed, also in light of the new literature review; presentation and language need some improvement. Below a list of major and minor comments.

Major comments

1. **Literature review** is missing some key papers of the field.
   - Panziera et al., 2016 developed regional rainfall frequency analysis and implemented (and tested) them in an early-warning system for Switzerland – this is probably the first study deriving rainfall frequency analysis from remote sensing data and providing an actual operational, quantitative product;
   - Peleg et al., 2017 analyzed the impact of small scale rainfall variability on frequency analysis from point (rain gauges) and areal (radar) estimates – they found that, due to the relatively short record length, point and areal estimates, are expected to differ (even if both measurements are 'true'), and observed large differences between frequency analyses from rain gauges located within a 1 km^2 pixel. This means that no exact match between point and areal results should be expected (not only because of the areal reduction factors issue). This contribution needs to be mentioned by the literature review, in light of the contribution it provides to the interpretation of the results of this study (also the conclusions [page 11, lines 18-19] could be updated accordingly);
   - Wright et al., 2013 proposed the use of stochastic storm transposition for radar rainfall frequency analysis in order to overcome the limitations due to the short radar records.

2. **Methods** are sometimes insufficiently described and apparently subjective choices are made without providing the reader with rationale, supporting references, and discussion of the implications.
   Frequency analysis*:*
   - the use of PoT approach for the frequency analysis of short records is highly desirable, however the choice of the exponential distribution (a special case in which the shape parameter – driving the long return period tail of the distribution – is assumed uniform in space and equal to 0) is very strong and goes against some literature on the topic. This choice needs to be motivated and supported;
   - [page 6, lines 13-18] how is the return period used in the QQ-plots derived? Is it is done following Willems et al, 2007? This is a key aspect in the methods and in the shaping of the results and needs to be explained and discussed; [7, 12] how is figure 2 created? from Willems et al, 2007, I imagine that the figure shows on the x-axis the return period derived from the exponential distribution that maximizes the linearity of the relation – how is the maximization done?;
   - [6, 14-16] this is not clear for the reader unfamiliar with Willems et at., 2007, please provide more information;

- [7, 20] what do the authors mean with “heavy tailed” (the exponential distribution has no shape parameter)?
- [9, 7-8] this sentence should be somehow supported/motivated;
- [10, 14] why is 60 chosen?

Radar QPE*:*

- [9, 22-23], what do the authors mean with “standard Z-R”? In what cases is a non-standard Z-R used?
- Previous studies found important instability of the MFB factor for short periods (1 h), especially in convective conditions. The use of hourly mean field bias adjustment needs to be supported by sensitivity analyses or references. Please discuss this and provide information on the stability of the factors;
- Is there any motivation for the choice of the hail threshold (80 mm/h looks low for some climates)? Are there cases in which the rain gauges did measure heavier rain intensities? [7, 18] is it possible to check how often the hail filter is activated? Is it possible to check what reported in [7, 23-24]?
- [8, 5] is it possible to check if bright band was contributing to this observation? This would be interesting since VPR impact was rarely discussed in previous studies on radar-based frequency analyses;

3. **Interpretation** of some results.
   - [Section 4.1] The authors select the maximum within 20 km range windows around each analyzed pixel in order to better capture the maximal intensities (see [8, 31]). The motivation for doing this is clear to me and it is a good direction to take to exploit the distributed information provided by the radar (and other gridded datasets). However, I am wondering whether the interpretation one should give to the obtained results still holds: will we still be dealing with the estimation of the frequency of occurrence of a given rain intensity-duration combination at a given location? Especially since the conclusions open with: “… to study extreme precipitation at a given location…”. I’m not sure this is the case. I recommend the authors carefully examine and discuss this issue. At this regard, the stochastic storm transposition approach adopted by Wright et al., 2013, even though much more complicated, provides similar advantages while preserving the interpretation;
   - [Figure 6, 7] Can the circular patterns be caused by the regionalization method (in case problematic pixels are still there, one will choose them when selecting the max value within the circular area – see also [10, 20-21])? Can this represent a weakness of this method?

4. **Presentation** is sometimes difficult to follow.
   - new contributions brought by this study are not clearly stated in abstract, introduction and conclusions. Reading the manuscript, the main results appear to be: raw radar QPE provides unreliable analyses and bias adjustment is needed; differences are observed between at-site analyses from radar and gauges, but radar analyses lie within the gauge confidence intervals; regionalization approaches provide improved analyses. These results were already reported in literature (see for example Overeem et al, 2009; Marra and Morin, 2015; Panziera et al, 2016; Peleg et al 2017; Marra et al., 2017). In my opinion the study brings a lot of new to the field, in particular the use of (i) new methods (i.e. a peak over threshold approach, QQ plot regression) and (ii) new regionalization approaches for rainfall frequency analyses from remote sensing data sets. Abstract, introduction and conclusions need to be reorganized in these terms, even though the results reported by the authors definitely deserve to be mentioned;
   - the presentation of the gauge networks in [2, 20-25] and in section 2.1 is difficult to follow, I recommend reorganizing and rephrasing these parts (how many networks are used?, why are they considered separately?, what are the differences? What the advantages of including each of them? Why not using them together?);
   - [4, 8-9] please provide information about these methods and move the reference to Goudenhoofdt and Delobbe, 2016 earlier in the text;
   - organization of the radar datasets (QPE, CAP,…) need to be made clearer (sections 2.2, 2.3);
   - [section 3.1] is difficult to follow; in particular [6, 4-9] and [6, 19-21] are not clear to me;
   - [section 3.2] what does “problematic events” mean? How are they identified?; why did the authors focus on 10 extremes?
   - [8, 21-22] why mentioning the index flood approach? Here the shape is actually assumed uniform (by the use of

the exponential distribution), but I guess this is not what the authors mean with 'regionalization';
- how does the 20 km regionalization of the parameters relate to the 10 km and 50 km used in the following parts of the study? how did the authors check/motivate that 20 km "provides a sufficiently large data sample"?
- it is not clear whether the method by Reed et al., 1999 is the one the authors used in this paper;
- [9, 28-29] did the authors check for non-stationarity in the data (e.g. changes in the instrumentation, or other)?
- [11, 14-15] is this expected?
- [11, 24-25] this problem should be solved by the adopted regionalization (20 km);
- [1, 11-13], [4, 8], [4, 20-24], [6, 19-21], [7, 10-11], [9, 1-5], [9, 11-16], [9, 31-32] please rephrase;

Minor comments/edits

- [abstract] the text of the abstract needs to be better organized;
- [page 1, line 3] "independent sliding 1h and 24h rainfall": this is not clear;
- [1, 9] natural rainfall variability in combination with short record lengths is also to be mentioned as a cause of the mismatch between point and areal frequency analyses (see Peleg et al., 2017 and major comment above);
- [1, 11] "assuming that the extremes are correlated": this is not clear to me, I guess it is related to the regionalization, but needs to be better written;
- [1, 18] please remove "very" and "very"; please add a comma after "activities"; please provide a reference for this sentence;
- [1, 19] sewer systems are an example, but I'd insert an example from other applications, such as dams design/management; [1, 21] sewer systems are usually designed for relatively short return periods, applications requiring long return periods are dams, bridges, etc.;
- [1, 20] no need to specify "a branch of statistics";
- [2, 3] an example of what the authors mean with "high-resolution" would be helpful for the reader;
- [2, 5-6] "the best potential is provided by radar QPE". Satellite products are fruitfully being used too and are, often, characterized by longer records. This sentence should be motivated and supported by references;
- [2, 11-12] the reference to Saito and Matsuyama, 2015 looks unrelated to the rest of the text, can you provide some information on its relevant parts;
- [Figure 1] what do the areas in the figure represent? Are they catchments? Are they used in the manuscript?
- [4, 28] how is gauge data validated?
- [5, 5-6] "The hourly bias obtained… convective storms" can be removed;
- [5, 6-10] is this done with a moving window? Or on 24 h blocks?
- [5, 10-11] Marra and Morin, 2015 quantified this uncertainty;
- [Table 1] what is the meaning of the "Avail. All" column? Does it mean that "Both" were available?
- [6, 22 and 7, 25] I'd suggest to change these titles to something focusing on the tested product rather than on the rain gauges against which it is compared;
- [8, 21-22] please change "consider" to "considers";
- [10, 25] please, change "estimate" to "estimation";
- [11, 18-19] the authors may want to check Avanzi et al., 2015 for additional inputs;
- [12, 6] since the work by Frederick et al., 1977, a number of papers are available on the derivation of ARFs from radar data (e.g. Durrans et al., 2002; Overeem et al., 2010; Wright et al., 2014, among others).

References

Avanzi, F. et al. 2015. "Orographic Signature on Extreme Precipitation of Short Durations." Journal of Hydrometeorology 16(1): 278–94. http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-14-0063.1.

Durrans, S. Rocky, Lesley T. Julian, and Michael Yekta. 2002. "Estimation of Depth-Area Relationships Using Radar-Rainfall Data." Journal of Hydrologic Engineering 7(5): 356–67.

Frederick, Ralph H., Vance A. Myers, and Eugene P. Auciello. 1977. "Storm Depth-area Relations from Digitized Radar Returns." Water Resources Research 13(3): 675–79.

Goudenhoofdt, E. and Delobbe, L.: Generation and Verification of Rainfall Estimates from 10-Yr Volumetric Weather Radar Measurements, Journal of Hydrometeorology, 17, 1223–1242, doi:10.1175/JHM-D-15-0166.1, http://dx.doi.org/10.1175/JHM-D-15-0166.1, 2016

Marra, F. and E. Morin 2015: Use of radar QPE for the derivation of Intensity–Duration–Frequency curves in a range of climatic regimes. Journal of Hydrology 531 (2015) 427–440. http://dx.doi.org/10.1016/j.jhydrol.2015.08.064

Marra, F., Morin, E., Peleg, N., Mei, Y., and Anagnostou, E. N.: Intensity–duration–frequency curves from remote sensing rainfall estimates: comparing satellite and weather radar over the eastern Mediterranean, Hydrol. Earth Syst. Sci., 21, 2389-2404, doi:10.5194/hess-21-2389-2017, 2017

Overeem, A., T. A. Buishand, and I. Holleman. 2009. “Extreme Rainfall Analysis and Estimation of Depth-Duration-Frequency Curves Using Weather Radar.” Water Resources Research 45(10): 1–15.

Overeem, A., T. A. Buishand, I. Holleman, and R. Uijlenhoet. 2010. “Extreme Value Modeling of Areal Rainfall from Weather Radar.” Water Resources Research 46(9): 1–10.

Panziera, L., Gabella, M., Zanini, S., Hering, A., Germann, U., and Berne, A.: A radar-based regional extreme rainfall analysis to derive the thresholds for a novel automatic alert system in Switzerland, Hydrol. Earth Syst. Sci., 20, 2317–2332, doi:10.5194/hess-20-2317-2016, 2016.

Peleg, N., Marra, F., Fatichi, S., Paschalis, A., Molnar, P., and Burlando, P.: Spatial variability of rainfall at radar subpixel scale, J. Hydrol., in press, doi:10.1016/j.jhydrol.2016.05.033, 2017.

Reed, D. W., Faulkner, D. S., and Stewart, E. J.: The FORGEX method of rainfall growth estimation II: Description, Hydrology and Earth System Sciences, 3, 197–203, doi:10.5194/hess-3-197-1999, http://www.hydrol-earth-syst-sci.net/3/197/1999/, 1999.

Willems, P., Guillou, A., and Beirlant, J.: Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function, Journal of Hydrology, 338, 221, doi:10.1016/j.jhydrol.2007.02.035, 2007.

Wright, Daniel B., James A. Smith, Gabriele Villarini, and Mary Lynn Baeck. 2013. “Estimating the Frequency of Extreme Rainfall Using Weather Radar and Stochastic Storm Transposition.” Journal of Hydrology 488: 150–65. http://dx.doi.org/10.1016/j.jhydrol.2013.03.003.

Wright, Daniel B., James A. Smith, and Mary Lynn Baeck. 2014. “Critical Examination of Area Reduction Factors.” Journal of Hydrologic Engineering 19(4): 769–76.