Hydrol. Earth Syst. Sci. Discuss.

At-site and regional frequency analysis of extreme precipitation from radar-based estimates

E. Goudenhoofdt, L. Delobbe, P. Willems

In this paper a peak-over-threshold method is used to perform an extreme rainfall analysis and to derive return levels from weather radar and rain gauges in Belgium. The importance of this work is high, as radar archives are nowadays long enough to permit the development of extreme rainfall analyses which are of fundamental importance for many applications, but the common annual maxima approach needs even longer time series. However, some important explanations and discussions, in addition to those already highlighted by the first review by F. Marra, are missing in the manuscript, and need to be provided before the article can be accepted for publication.

Luca Panziera

Major comments

1. **What is new?** It is somehow difficult to understand which new contribution this paper brings with respect to previous studies, and I think that this should be better highlighted in the text. To my understanding, the main novelty of this paper is the use of a POT method for an extreme rainfall analysis for weather radar data.

2. **Temporal Declustering**. As rainfall data need to be declustered in order to remove the temporal correlation in the time series before GPD parameters estimation, the authors choose an interval of 12 hours (for 1-hour rainfall) and 3 days (for daily rainfall) in order for two threshold exceedances to be considered as independent. The choice of these intervals, which should be referred to as run length or run parameters according to the literature, seems reasonable, but it could potentially have a big impact on the derived return levels, as it shapes the exceedances time series whose maxima are used for the parameters estimation. If the data are temporally clustered, such temporal lags could not be long enough to remove dependency, but if the temporal clustering occurs rarely, they could actually lead to a significant bias of the return levels estimates. What do the authors mean as temporal independence? How did the authors choose such temporal lags? Did the authors investigate the effect of changing these values on the parameters estimation and final return levels? The subjective choice of these values should be motivated and discussed in the text.

3. **Exponential distribution**. As the choice of a null shape parameter is fundamental for this work, I think that it should be motivated more in the text. Therefore, I suggest to briefly report and discuss the main results of Willems (2000), in order to better understand the motivation of this choice The text states also that this choice was taken because of the short period. However, with a POT approach the shortness of the period should not be a limiting factor, as many events are considered. It should also be discussed if this is the best choice for both 1-hour and 24-hours accumulations. Did the authors try to estimate also the shape parameters, to see if from the data a value different from 0 could be derived?

4. **Radar and gauge comparison.** The authors present an interesting comparison between the radar and gauges extremes, for 1-hour and 24-hours accumulations. Despite this being very interesting and

instructive, the implications for this study are not very clear. I suggest the authors discuss at least qualitatively the influence of this investigations on the overall results of the study.

5. **Regional frequency analysis**. The regional frequency analysis needs also to be better explained and the choices which were taken need to be motivated and discussed. How did the authors choose the 20-km radius for the analysis? How the resulting return levels at a given pixel should be interpreted, as they stem from exceedances in rainfall values which were observed all around it? Does it still make sense to speak about point measurement? How are the maps of GPD parameters affected by the choice of the 20-km radius circles?

6. **Return levels maps**. I guess the final goal of the study is to derive maps of return levels with relative uncertainty for Belgium. Despite the return levels are shown for given rain gauge locations, it would be desirable to show also maps of return levels for selected return periods. Would it be possible to insert a map or two of the return levels? How would those maps be affected by the 20-km radius selected for the regional frequency analysis? How these maps should be interpreted? Since you are using a constant shape parameter (equal to 0), and the longest return levels are shaped by it, long return periods map will tend to produce maps less variable in space with respect to short return periods. This should be discussed in the text.

Minor comments

1. The title is rater general, and you might want to consider adding the name of the region for which this study was performed (Belgium).
2. In the introduction some relevant papers are missing. I strongly encourage the authors to discuss also the papers referenced by F. Marra in his review.
3. Pag.2, line22."*in this study, we want to demonstrate the potential of this radar-based QPE to derive point rainfall statistics"*. I don't think that the aim of this study is this, as the radar pixel does not represent point rainfall statistics. As the authors know, the radar rainfall estimate comes from the reflectivity measured within the sample volume, representing an area- not a point- measurement. The intrinsic difference among radar and gauges measurements should be at discussed in the paper, since a comparison between rain gauges and radar return levels is performed (see also major comment 1 by F. Marra).
4. Pag.3, line 4: is there a reference for the 5-10% rain gauges underestimation?
5. Pag.3, line 7: improve English. I propose to change "very high" with "10-min" temporal resolution (and delete *"10-min accumulations are available from the database")*
6. Pag. 4, line 25: please clarify the last sentence of section 2.2 which, in its present form, it is not correct. Could change *" In addition, the increasing radar sample volume will give lower extreme values" to "In addition, the increasing radar sample volume will produce an underestimation of local small-scale extremes".*
7. Pag. 5, line 24. First two sentences of section 3.1 need to be reformulated as they are very colloquial.
8. Page 6, line 14. With this method of regression in QQ plots, is there a risk of over fitting? Could you please comment on that?
9. pag.7, line 13-14 and pag.8 lines 9-10. How this percentage would vary by changing the temporal lags considered for independence? (see major comment 2). *"This is what we expect from …."*. According to which theory/observations? Please clarify and give references.
10. Pag. 8, lines 21-28. It would be more appropriate move the literature review to the Introduction, instead of leaving it in this Methodology section.

11. Pag. 8, second paragraph of section 4.1: please clarify the explanation of the regional frequency analysis. Given that your circle has a radius of 20 km, what is the aim of considering all the events within a 50 km radius dependent? Isn't this the same as taking just the max value within the 20-km radius? In case it is, wouldn't be easier just say that you take this maximum within the 20-km radius circle?
12. Pag.10, lines 6-9. Also here I suggest to move the references to other studies in the Introduction.
13. Pag. 10, line 13: *"a few pixels having too much (50) …. removed"* . This sentence is rather unclear, and this seems a rather subjective choice which can hardly be motivated.
14. Figure 2. I suggest to rename "Extreme 1-hour precipitation quantiles" to "1-hour return levels", to be consistent with theory and common nomenclature in the field.
15. Tables 2 and 4. I actually miss how the events in the tables are ordered, if there is a logic.
16. Figures 1, 6 and 7: a scale in km would help the interpretation of the figure, for those who are not familiar with Belgium