## Anonymous Referee #1

Received and published: 25 April 2016

### General comments

This paper presents the use of weather radar data for regional extreme rainfall analysis together with a relevant application for automatic heavy precipitation alerts. The authors not only clearly explain methodology and describe the data employed, but they also show how they addressed issues that otherwise might have introduced biases in their calculations, as for example the presence of residual ground clutter. The presentation is clear and the text completed with an interesting case study which illustrates how the proposed system is complementary to thunderstorm nowcasting in case of low convective, but heavy precipitation events. Finally, the study not only touches at topics that are particularly relevant for operational weather services, as the production of warnings for severe weather, but it also introduces practical methods for exploiting the potential of increasingly growing weather radar archives.

### Specific comments

p. 8, line 13: The use of monthly block maxima for the regional extreme analysis is certainly appropriate and has the advantage of introducing a robust temporal independence of the extremes. However, given the seasonal variation of the precipitation, just taking one value per month might lead to the exclusion of some of the extremes: the second peak in August may be larger than the highest value in May. The issue is of some importance considering the short duration of the dataset (10 years). Often a peaks-over-threshold approach is proposed to overcome the limitation. Have you considered this approach for your application?

<u>Reply</u> *Applying a peak-over-threshold methodology in order to select the regional maxima would have certainly been a valid alternative to the chosen monthly block maxima approach. However, in such approach the choice of the threshold is a difficult and questionable task; it would imply also some grade of subjectivity since there are no common or objective rules. A preventive study about the stability of the Generalized Pareto Distribution parameters derived with different thresholds would give useful hints, but this would require an extra load on the preparation of partial duration series. Hence, we opted for the monthly maxima because this method is less sensitive to implementation and parameter settings and makes the results more comparable, convinced that for the operational scopes of this study it is a valid alternative to peak-over-threshold. One could also redo the same study using the peak-over-threshold methodology, but this would not change the main message of the paper.*

p. 9, lines 4-20: You are not only assuming that the GEV parameters are spatially constant within each warning region, but also that they are stationary throughout the warm/cold season. The assumption is justified by the need to compensate for the limited number of years in the record, but in my opinion it might be appropriate to mention it more explicitly in the text. There are also examples of studies trying to explicitly account for seasonality (e.g. Rust et al., 2009).

<u>Reply</u> *Thank-you for highlighting this important assumption we made which, as you noticed, is motivated by the need to extend the empirical observations of maxima.*

*Explicitly accounting for seasonality seems to be a too refined methodology for the application of this study, i.e. the issuance of heavy precipitation alerts for the general public in case of threshold exceedance. A comprehensive study about the seasonality of extreme precipitation over Switzerland in order to explicitly model the annual cycle of precipitation, like the one you cited, is beyond the scopes of this work. However, in order to make the reader conscious about this fundamental assumption, we inserted a sentence at the end of section 2.3:*

*Moreover, the analysis was performed separately for the warm (May to October) and cold (November to April) seasons in order to reduce the effect of seasonality on the choice of maxima, assuming that the GEV parameters are constant within each season. Thus, a total of 6 (months) x 10 (years) = 60 monthly maxima are considered for each season.*

Technical corrections

Figure 8: If alert thresholds are defined per region and correspond to given return periods, why is the colorbar on Figure 8 showing both alert levels and rainfall thresholds? As I understand it, two regions at the same alert level (and therefore experiencing precipitation events with similar return periods) may get very different 1-hour rainfall totals (as shown in Figure 6).

<u>Reply</u> *Thank-you for noticing it. Figure 8 represents the visualization of NowPAL alerts as they are currently obtained at MeteoSwiss, where for practical reasons the system is still working with the same rainfall thresholds for all the regions for 1-hour rainfall accumulation. The diversification of rainfall threshold will be inserted in the system in the next few months. In order to not confuse the reader, we deleted the rainfall thresholds from figure 8.*

References

Rust, H. W., Douglas Maraun, and T. J. Osborn. "Modelling seasonality in extreme precipitation." The European Physical Journal Special Topics 174.1 (2009): 99-111.

## Anonymous Referee #2

Received and published: 28 April 2016

This is an interesting and a well written paper. It describes an operational warning system that is based on radar-rainfall combined with nowcasted rainfall to derive several regional statistics. These statistics are compared with thresholds obtained from extreme event analysis applied to block-maxima series. The study presents a one more step forward in operational utilization of radar rainfall in real-time alert applications. The study is a continuations of the substantial scientific efforts taken by MeteoSwiss in developing radar-based QPE algorithms that are accurate enough to allow their use in operational systems.

Below are few **comments** to be considered:

1. It was not clear (at least to me) how the thresholds are actually set for each region? Is it selected by the value corresponds a specific recurrence interval? if yes,

what recurrence interval? for which duration? How the different levels (Alert level 1 – Alert level 4) of alerts are set?

<u>Reply</u> *Even though the NowPAL system can ingest any threshold according to subjective user needs, the paper proposes to set them to the return levels corresponding to specific return periods. The study shows how such threshold values can be found by means of an extreme rainfall analysis, but, as you correctly noticed, it does not tell which return periods correspond to the different alert levels in the NowPAL version currently running at MeteoSwiss. The reason is that at MeteoSwiss this choice has not yet been definitively taken, as significant discussion is still taking place about the severity of the rainfall events for which alerts should be issued to the general public and authorities. In our current provisional configuration framework, the system is working with alert levels which correspond to 6 months, 1, 2 and 4 years return period. Note that the number of alert levels is also a configuration of the system, and more than 4 levels might also be set.*

*In order to highlight that the choice of the return periods corresponding to the alert levels is part of the configurations which have to be set for the customers, the first sentence of the third paragraph of section 6 was slightly modified and now it reads:*

*Return levels are used as rainfall thresholds and associated to different alert levels by NowPAL. In fact, our assumption is that an alert of a given level should be expected the same number of times during a season in every warning region. Under this assumption, the extreme rainfall analysis objectively provides the threshold values for each warning region and for a given regional statistic, once the return periods corresponding to the alert levels have been fixed.*

2. A main problem related to the question above is that, typically, if we set the thresholds relatively low, we will have high false alarm rate but also high hits rate, while setting the thresholds high, and vice versa for a too high threshold. The threshold selection is not a straightforward task and it affects the general performance of the alert system. Therefore, guidance for how to select the thresholds or the appropriate duration and recurrence interval is important.

<u>Reply</u> *The system was designed such that it can be easily configured with different sets of configurations, including thresholds, for different customers. The heavy rainfall alerts which are issued for the general public have certainly a different scope with respect to alerts issued for very specific customers which have to monitor very small and/or remote catchments. So, it is rather difficult to provide a general guidance for the choice of thresholds. The basic assumption of the paper is that the frequency of heavy rainfall determines the return levels which are used as thresholds by the system. In other words, the frequency of the hazard determines its severity and this is the only criterion which should drive the choice of the thresholds according to the presented work. Of course this might not be the case for specific applications, and this is the reason for which the system can work with any rainfall threshold.*

3. The authors state that "a verification of the alerts issued by NowPAL is beyond the scope of this work" (P. 4, L. 25), which is understood. However, to my opinion it would benefit the reader to get some statistics on the rate of alerts issued in each region, beyond the single case shown in this paper.

*Reply The statistics about the rate of alerts issued in each region is strongly connected to the presented extreme value analysis. By construction, in fact, setting in NowPAL a threshold corresponding to a return period of x-months, it means that in the last 10 years that particular alert level would have been issued on average every x-months, if the system would have been available in real-time. Since the system is new and it runs operationally since just a few months, we cannot really obtain some statistical study about the rate of alerts issued. In any case, such an analysis would be dominated by the skill of both the forecasting system and the product used to estimate past precipitation. In other words, it is not possible to evaluate the skill of the NowPAL real-time system itself, without evaluating the skill of past and forecast precipitation systems, which is beyond the scopes of this article.*

4. "The local maxima used for the statistical analysis have not necessarily been measured at the same place, but they might have occurred at different locations within the region of interest". This means that the return periods are dependent (among other things) on the relation between the window size and the region size, i.e., for a relatively small window applied once to a small region and once to a large region, the return periods of the former will be generally lower than the latter, without a real change in the precipitation regime. Does this affect in any way the threshold selection?

*Reply The size of the pre-defined geographical regions to be monitored plays a role in the process of monthly maxima extraction which drives the statistical analysis presented in this study. In fact, as you correctly noticed, the same precipitation regime, or spatio-temporal pattern, might lead to slightly different return periods for regions of significantly different area. In particular, the return levels (return periods) for a fixed return period (return levels) will be larger (lower) for a large region with respect to a smaller region located in the same climatological regime, since in a larger region it is likely that more rainfall events of the same magnitude have been observed. However, this effect is of secondary importance for the results of the regional extreme rainfall analysis presented in the paper, as the maps presented in figures 5 and 6 generally agree with previous rainfall climatology studies; thus, it appears that the different return levels are associated with different climatological regimes more than with different sizes of regions. The strong precipitation gradients observed between the Po Valley, the Southerly Alpine slopes, the inner Alpine crest, the Northerly Alpine Slopes, the Swiss plateau and the Jura Mountains somehow hide the change in return levels due to the different size of nearby regions located within the same precipitation regime.*
*As a practical remark, the user which would like to be warned every time the rainfall reaches a given magnitude independent on the extension of the region where he is located, should work with the same thresholds even for regions of different areas using* $max(RS(tot_i)_j)$ *as regional statistic.*

More **specific comments** or required clarifications:

1. Did the goodness of fit of GEV was examined?

*Reply* *In the present study we did not examine in detail the goodness of the GEV fit, but we only derived the 95% confidence intervals which also gives an indication of the goodness of the estimation of model parameters. Confidence intervals were computed for all the regions, and, in general, they increase with decreasing the area of the moving window used for the regional statistic computation. The smaller confidence intervals are provided by the mean of regional rainfall. In the paper we only show confidence interval in figure 4, for summer monthly maxima of rainfall in the Schaffhausen region. We think that for the operational application of this study, i.e. for finding rainfall threshold for the operational system, a more detailed assessment of the goodness of the fit is not required. On the other hand, for climatological applications this would be a very important task. In fact, we started to extend the radar-based extreme rainfall analysis working at the pixel-scale for climatological purposes, and in this case we are going to assess also the goodness of the fit by means of several statistical indices.*

2. What about the effect of snow? How well it is estimated by the radar?

*Reply* *Winter precipitation estimates from radar suffer from systematic underestimation, which affect both the winter return levels estimated from the 10-years data set and the winter rainfall measurements used by the NowPAL system in real-time operation. Thus, even though the winter thresholds might result not realistic since underestimated, this would be compensated by the real-time underestimation, making the two quantities comparable.*
*However, we should also mention that the estimation of solid precipitation is poor for both radar and rain-gauges, and that, although we obtain return levels also for the cold season, the NowPAL system is designed to be used mainly in summer, when liquid precipitation makes rainfall estimates more reliable and precipitation intensities are indeed more critical. In order to highlight this concept in the paper, at the end of section 2.3 we added the following sentence:*

*Since the estimation of solid precipitation from radar suffers from underestimation, the cold season return levels might result underestimated as well. However, we should mention that the NowPAL system was designed to issue alerts mainly in summer, when liquid precipitation makes rainfall estimates more reliable and precipitation intensities are indeed more critical.*

3. How does the window apply close to the region border?

*Reply* *The process of averaging within the moving window of configurable size is independent on region borders, since the smoothing of the precipitation field is performed prior to regional statistics computation. When smoothing, each pixel of the radar domain is given the value of the mean of all the pixels of the surrounding square window. In the subsequent step of computing the regional statistics, however, only the pixels belonging to the region of interest are considered. As a result, also pixels out of the region but close to it have an influence on the regional statistic computation. The larger the size of the moving window, the larger the influence of pixels not belonging to the region. However, the weight of pixels external to the region will be generally lower than pixels within it, excluding rare very barbed regions, since they are smaller in number. In this way the system will react to very intense rainfall located very close to the region of interest, which might constitute a danger since it might easily be advected over it.*

4. P. 8, L. 29 - correct to "exceeded"

Reply *Corrected, thank-you.*

5. Fig. 10 - are the 5-min data in mm or in mm/h? if the latter, it should be noted on the axis title.

Reply *Only the 5-min rain rate is in mm/h, as indicated in the legend. The other quantities which are plotted are all in mm. The y-axis is thus given in mm, as 3 out of 4 quantities are in mm.*

6. Fig. 10 - what is the explanation for the large difference between the "last 30-min" and the "next 30-min". Is it because one is observed and the other is nowcast?

Reply *Yes, the “last 30-min” is the rainfall measured by radar in the previous 30 minutes, whereas the “next 30-min” is the rainfall predicted by MAPLE for the following 30 minutes. Since this is a moving storm not producing stationary rainfall, the two curves are different. For the sake of clarity, we added “radar” and “MAPLE” in the legend of the figure.*