Response to review 2

We extend our sincere gratitude for the dedicated time and effort you invested in reviewing our manuscript, are we are grateful for the constructive comments on and valuable improvements to our paper. We have incorporated changes to reflect the suggestions provided. Please see below, in blue, for a point-by-point response to the comments and concerns.

The paper explores precipitation patterns and rainfall statistics in space and time based on weather radar data and the application of the method SMEV. This method is described in detail in Marra et al (2022) for radar pixel values however here the method is extrapolated to also include spatial rainfall and thus dependence of area.

The paper is generally well-written and understandable. Several things are indeed interesting – especially the pixel-area relations depending on rainfall duration as well as the significant climatological differences in the study area.

Thank you very much.

1) The novelty of the paper should be emphasized more in the introduction. I guess compared to former efforts, the novelty here is the application of SMEV also on the areal component rather than single pixels?

Author response: Thank you for the suggestion. Yes, this is the main novelty. The follow has been added to the paper to emphasise this:

> *"In this study we apply the SMEV framework to examine extreme precipitation at various spatial scales for the first time, in order to investigate the impact of area size on local extremes. As stated, while the SMEV framework has demonstrated efficacy in successfully estimating extreme rainfall, its prior applications have been confined to either the point (in the case of rain gauge data) or pixel (when utilising radar rainfall data) scale analyses at different temporal scales. We here extend the application of the SMEV to estimate extreme return levels up to 100 years across multiple spatial and temporal scales."*

2) In my view, it is not clear what the differences between MEV and SMEV are. The authors refer to past publications, but could be relevant to describe the SMEV in a bit more detail here in order to understand how it differs from other methods of extreme value statistics. For example line 157-158 could be detailed further.

Author response: Thank you for the suggestion. The text has been revised and now reads:

> *"Extreme precipitation return levels are estimated across the study area using the novel non-asymptotic SMEV framework proposed by Marra et al. (2019a) and (2020), a simplified version of the original MEV framework proposed by Marani*

and Ignaccolo (2015). The MEV and SMEV approaches are based on the concept of 'ordinary events', which are all the independent realisations of the process of interest. Unlike classic extreme value theory, which only exploit a small subset of the data, i.e. the annual maxima or the peaks exceeding a high threshold, they make use of a greater proportion of observations to fit the distribution parameters, thus decreasing the parameter estimation uncertainty. The SMEV is a modified version of the MEV; it neglects the interannual variability of the distribution of ordinary events and in their number of yearly occurrences (Marra et al., 2019a). The SMEV formulation significantly reduces the number of parameters and allows for a direct interpretation of their meaning. This results in a simpler formulation for the non-exceedance probabilities of extreme rainfall, and more robust parameter estimation. Several studies have applied the SMEV to precipitation frequency analysis over different regions (Marra et al., 2020, 2019a; Miniussi and Marra, 2021; Araujo et al., 2023), including over the study area (Marra et al., 2022), and have demonstrated the robustness of the method's assumptions and its ability to reproduce extreme frequencies from relatively short records. The SMEV is used here to estimate precipitation events of varying sizes and durations, so that spatial and temporal effects on extreme precipitation can be analysed"

3) The equation in line 128: Shouldn't it be Z=316R^1.5?

Author response: Yes, thank you, this has been corrected

4) Do you think the discrepancy between gauge and radar is due to differences in climatology, bias adjustment, or radar artifacts? Consulting figure 4 it seems that there are some radar issues close to the radar – maybe something related to scanning and CAPPI generation?

Author response: We cannot be sure of the exact reason for the discrepancy, however we hypothesise that it is due mainly to issues with the radar. There are no substantial differences in climatology in this area, and there are no indications in the correction factors that they are causing the increased values here. There is, however, no CAPPI generation via interpolation or similar in this radar archive, so CAPPI generation can be excluded from the causes. As stated in the text:

"Despite best efforts, the exact issue with the data here could not be fully identified by the authors. It is suggested that the poor results may be due to the close proximity of the site to the radar station, where radar precipitation intensity estimates are known to be poor and unreliable, or to other types of radar systematic errors (e.g., residual ground echoes)."

The following was also added to the text:

"There are no substantial changes in climatology in this area that could cause these discrepancies. Additionally, there are no indications that the parameter correction factors are related to the increased values here."

The area of low values around the radar station is indeed caused by the close proximity to the radar station and this is stated in the text:

> *"The area of very low values, attributed to data quality issues around the radar location, is clearly visible."*

5) Figure 5. I think it would be very interesting also to include the rain gauges statistics in this figure – and in principle also in figures 6-7. Typically, you would see an underestimation of the radar estimates at short durations. See e.g. Schleiss et al (2020) or Andersen et al 2021. The point scale is indeed interesting to study in addition to the study of the difference between pixel scale and areas of 10, 100 and 500 km2. In regards to point 1 in the conclusion the actual comparison in the manuscript is not made on point scale but on a pixel scale. Even going from point to pixel scale will result in scaling – which will be more dominant for shorter durations than larger ones.

Author response: Thank you for this comment. Actually, radar return levels are adjusted to exactly match the rain gauge ones at the pixel scale (see Marra et al., 2022). This means that the effects of areal mismatch between rain gauges and radar cannot be seen in the estimated return levels. The referee is kindly referred to Marra et al. 2022 for an analysis and discussion of the mismatch between weather radar at the pixel scale and rain gauges (see Fig. 2 in the mentioned paper and related discussion). As this was already examined in the previous study, we keep it out of the objectives of this study.

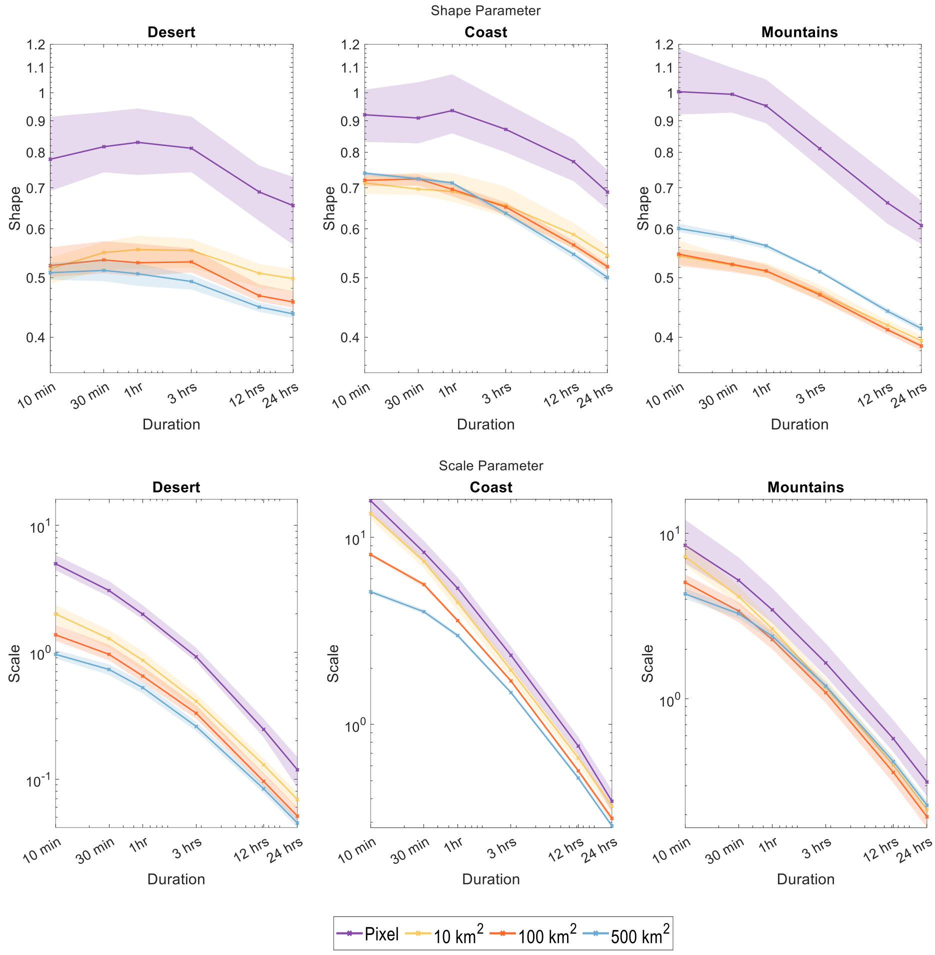
Thank you for raising the terminology mismatch, the “point” has been corrected to “pixel” throughout the text. Although we note that, as mentioned above, after adjustment the pixel scale is representative of the rain gauge (i.e., point) scale.

6) It would be relevant to present fitted parameter values of shape, scale, kappa, lambda in order to study how the vary over different durations and area. Maybe present selected values in a table. Potentially an empirical relation between model parameter estimates and duration and area could be sought.

Author response: Thank you for the suggestion, which is in line also with Reviewer 1 comment. We have therefore added a figure (rather than table) showing the shape and scale parameters (see figure below) along with the following text:

> *"The calculated shape and scale parameters, after correction factors have been applied, are presented in Fig. S4. The effect of both duration and area is clearly visible: the scale parameter decreases with increasing duration and increasing area, with the values converging at long durations – mirroring the behaviour of the return levels presented in Fig. 5. Unlike the scale parameter, the values of the shape parameter do not become more similar for long durations. The parameter displays non-monotonic behaviour, with generally minimal change for durations between 10 min and 1 h, and decreasing for durations between 1 and 6 h (implying an increasing tail heaviness). Very low parameters, between 0.4 and 0.75 (indicating heavy tails), are observed for*

area sizes greater than the pixel scale, especially over the desert and mountains, while exponential tales (i.e. values close to 1) are observed for the pixel scale."



Shape and scale parameters (after correction factors have been applied) as a function of area and duration estimated for the desert, coast, and mountains. Shaded areas represent the 90 % confidence interval from 100 bootstrap samples.