Response to the reviewers

The authors wish to thank the referees for their remarks. Thank to their valuable comments the paper has been improved. In the following, the point-by-point response to the reviews.

Referee #1

1. I start with a technical comment. Did the authors check how well the Gumbel distribution represents the skewness of the annual maxima in the three areas? This can be a second big assumption that needs to be adequately supported. The L-moments could represent a good metric for it. Should the skewness not be compatible with \xi=0, there are options to use a-priori values different from 0.

The authors addressed this point performing discordancy measure test, homogeneity test and goodness-of-fit based on L-moments as suggested by Hosking and Wallis (1997) for all datasets and for Italy and Estonia. The results allowed to identify suspicious raingauges, confirmed the acceptable regional homogeneity and, finally, confirmed the Gumbel distribution as one of the possible theoretical statistical distributions for the areas.
All these results have been described and discussed in the additional paragraph “2.1.1 Data quality, homogeneity, and goodness-of-fit”.

2. My main concern remains the issue with homogeneity. The authors responded to this major comment by: (a) for the case of Italy, testing the homogeneity of the region with two methods (AD rank test and L-moments homogeneity test); (b) for the case of Estonia referring to a previous study (Olsson & al 2022) in which the 10-year return levels in Northern Europe are examined. While I have some specific concerns (see the points i and ii below), I can see how similar arguments could hold for justifying such a homogeneity approximation in some situations.
(i) Olsson & al actually show variations in the 10-year return level within the study area in Estonia (see Fig. 2 and 3 in Olsson & al 2022).
(ii) Viglione & al 2007 is used to justify the use of AD and L-moments test for homogeneity. For the case of slightly-skewed distributions (such as the Gumbel distribution eventually used here) these authors recommend the L-moments test. This raises 2 questions: why not using only this method? and Why not using any test over Estonia?

L-moments method (Hosking and Wallis, 1997) has been applied confirming acceptable homogeneity both in Italy and in Estonia. Regarding the comment on Olsson et al., 2022, it is worth noting that the study area considered in this paper is smaller than the one analyzed by Olsson et al.: eastern raingauges are outside from the studied area.

One should note that the homogeneity discussed by Viglione & al is within regional frequency analyses, i.e. assuming that some of the parameters of a distribution are homogeneous within the region, while some others are allowed to change. As I mentioned in my first review, assuming all the parameters are in common is quite (too?) strong in the current hydrological practice. This is particularly true in a study based on weather radar. What is the point of using weather radar when eventually one only estimates one distribution? The big advantages of using weather radars for such applications is to (a) get information for ungauged locations and to (b) capture spatial gradients in quantities of interest. But if we assume identical distribution we kill both these advantages. Can’t we just pool all the rain gauge data together for this?To my view, this remains a critical point in this study.

The major advantages of using weather radars for such applications are that information for unmeasured locations can be obtained and spatial gradients of the variables of interest captured. Due to the limited polarimetric weather radar data availability in time (a few years), the present study is limited to climatological homogeneous areas, limiting or losing these advantages.
Nevertheless, previous studies (Overeem et al., 2008, 2009a, b, 2010; Marra and Morin , 2015; Panziera et al., 2018; Marra et al., 2022) analysed weather radar QPEs based on horizontal reflectivity data adjusted with some ground raingauges measurements. Here, the major innovative

aspect is that the QPEs, based on blended algorithm R(Zh, Kdp), are obtained independently from co-located raingauges data availability. This study demonstrates that, by having polarimetric rainfall estimates, it is possible to estimate the rainfall annual maxima even in un-gauged regions. Moreover, as stated by Marra and Morin (2015), dealing with QPEs based on horizontal reflectivity data, the upper threshold used to limit the effect of hail is an issue in rainfall maxima estimation in warm regions, limiting the instantaneous rainfall estimation typically to about 100 mm/h. Involving Kdp for QPEs, the hail contamination issued is overcome, making QPEs independent from neither climatic region nor weather radar attenuation. Future studies will benefit from longer time series allowing investigations in wider non-homogeneous areas.

3. My comment 3 was not answered nor considered in the revision. Section 2.1.1 remains affected by erroneous and inaccurate statements about the extreme value theorem. It is beyond my role to correct point by point this section. The authors should refer to a textbook, to previous papers, or to wikipedia (https://en.wikipedia.org/wiki/Fisher–Tippett–Gnedenko_theorem). Now that the authors moved to using a Gumbel distribution, there is probably less need for details about the limiting types of GEV.

The Section has been completely revised.

Referee #2

Line 65. The R(Zh,Kdp) algorithm used in the study is a combination of a R(Zh) and a R(Kdp) algorithm. Such algorithms have been called "optimal", "composite", or "blended". I suggest to rephrase ".... that R(Zh,Kdp) blended algorithms provides good quality QPEs."

Corrected

Line 116-117: suggested "the annual hourly precipitation maxima concerns years from 1988 to 2020. "

Corrected

Line 118: replace "studly" with "study"

Corrected

Line 135: the text refers to a "shape" parameter that i explained later. Suggest to add something like "see section …"

Corrected

Line 155-156: "When using the default window length of 35, ..." Please specifiy also the size of bis of the 35-bin windows.

Corrected

Line 163-166: There is a discussion about Z-Kdp consistency in hail that in general does not exist. Although the discussion makes sense for Kdp calculation, it is not clear from the text wheter hail is considered in the computation of precipitation statistics or not.

We tested with various window lengths and found length 8 to be the optimal compromise between spatial resolution and smoothness. After the window length change, we obtained realistic looking precipitation fields but the overestimation compared to gauge values increased. This is because Φ_{dp} gradients became steeper as a result of the smaller window length. To mitigate this issue we first decreased the high_z (high limit for reflectivity to remove hail contamination) value from 60 dBZ used in Voormansik et al. (2021a) to 50 dBZ which is the lowest recommended value by Giangrande et al. (2013). Because overestimation was still evident we also reduced the Zh -Kdp self-consistency coefficient. As stated by Kumjian et al. (2019) the R(Zh -Kdp) consistency relationships probably do not exist in hail and it is therefore recommended to reduce the weight of the self-consistency constraint in the case of hail (Reimel and Kumjian, 2021). We tested with various values and found a coefficient value of 0.9 to produce optimal results.

Line 169: Please indicate where the coefficients or rain algorithms are taken from

The references have been included.

Line 173: I would still recommend "data are" instead of "data is"

Corrected

Figure 3: Please fix the title "Italty"

The title has been fixed.

Line 349: "the following table" please replace with "Table 2". In the table is not clear the meaning "0.5ex> 0.5ex"

Fixed