This paper presents the first application of extreme value statistics to precipitation estimates based on dual polarisation (dual-pol) weather radar. The topic is relevant to the community and to the readers of this journal, and the contribution is timely because dual-pol technology have been shown to improve radar precipitation estimates but extreme value statistics from such data were so far limited by the lack of sufficiently long records of dual-pol observations.

However, I highlight some aspects that need revisions before this study is considered for publication. These revisions may alter the overall results and conclusions of the study.

Major comments:

- The study makes use of strong assumptions on the homogeneity of the extreme precipitation statistics field: all the three parameters of the GEV distribution used to represent the annual maxima are assumed to be uniform within the examined areas. This clearly helps with the problem of parameter estimation. However, while the authors carefully address the issue of independence in the used data, they completely neglect the homogeneity assumption. This homogeneity assumption is quite strong for today’s research and practice standards and is not supported by specific tests/analyses. Spatial variations of precipitation statistics at scales smaller than the examined domains are known and reported in a number of previous studies based on both rain gauge and weather radar data (e.g., https://doi.org/10.1088/1748-9326/ab98b4; https://doi.org/10.5194/hess-26-1659-2022; https://doi.org/10.5194/hess-26-1439-2022). Neglecting the presence of spatial variations in the extreme precipitation statistics violates the identical distribution assumption of the extreme value theorem which underlies the adopted approach. Ultimately, this may lead to erroneous estimates of the parameters (samples from different distributions are mixed) and misleading evaluations of their uncertainty (the
A proper benchmarking of the results of this study with respect to results based on single-pol measurements is missing. Many studies about the analysis of extreme precipitation statistics from single-pol weather radars are not discussed in the introduction. As a consequence the key scientific questions underpinning this study are not well presented and leave the reader with questions like: why do we need to use dual-pol technology for these applications? what are the expected advantages over single-pol? what kind of uncertainties/errors were found in previous results based on single-pol? is dual-pol able to reduce such uncertainties/errors?

Although some concepts about extreme value theory are presented in a reasonable way in lines 219-225, the manuscript misrepresents many key concepts. Section 2.1.1 in particular should be rewritten.

The manuscript is not optimally organised. The results section contains large parts with methodological aspects and the discussion section contains main results rather than discussion.

The main objectives and results should be better stated in the abstract and text. The last sentence of the abstract lists two main results reached by the study as (a) “weather radar observations can provide a reliable QPEs compared to rain gauges” and (b) “even relatively short time series can provide reliable estimation of the rainfall return periods in climatological homogeneous areas”. However, point (a) is a result already addressed by previous literature (also listed in the introduction) and is not actually directly presented as a main result of this paper. Additionally, the study comes short at addressing point (b), because of an improper application of methods for extreme value analyses (see comment 1) and because it lacks a benchmarking of the results against single-polarisation data, neither from the radars used in this study nor from previous studies based on single polarisation technology (see comment 2). The natural question raised by the title and abstract of the paper “can dual-pol estimates improve our quantification of extreme precipitation probability with respect to single-pol estimates?” is not answered by the study.

Minor comments:

- line 14-15: “Single C-band polarimetric…” this sentence is not clear
- line 19-27: it is weird to start the manuscript with text about climate change when the study assumes stationarity and does not provide any further result or discussion about changes
- lines 32 and line 41: the reference to Peleg et al 2016 is a bit misplaced as these authors only focused on one radar pixel. Other more relevant studies could help detailing the state of the art (see above some suggestions among many). Also note that the correct citation is Peleg et al. 2018
- lines 37-48: this part is a bit confusing, please rephrase it
- Figures 1 and 2: I think these two figures could be merged. The quality of the figure should be improved, for instance using similar graphics and symbols
- The division of the Italian domain into two areas tries to address the homogeneity issue described above. It definitely help is some way (although it is in my opinion
insufficient), but should be better motivated and explained. Currently it is mentioned in lines 93-95, but no details are presented
- line 197-225: this part would better fit the methods section
- Figure 4: it is not fully clear what is the difference between the second and third panel of the first row and between the second and the third panels of the second row. The term model-simulated data only appears here and in line 229 of the text.
- Figure 5: some quantitative information on the results presented in this figure should be added, also in relation to the biases of the radar-estimated values with respect to rain gauges and with respect to non-polarimetric radar estimates.
- lines 285-287: this conclusion is underpinned by crucial assumptions (see above). These assumptions should be clearly mentioned