

Final Review Response

HESS-egusphere-2025-3073

Replies to Reviewer #1: **pp. 1-15**

Replies to Reviewer #2: **pp. 16-27**

Replies to Reviewer #3: **pp. 28-38**

Reviewer #1

Climatology and trends of observed daily and hourly extreme precipitation in the French Alps

Authors: Sebastian Berghald, Juliette Blanchet, Antoine Blanc, and David Penot

Recommendation: acceptable after major revision.

Theme

This paper analyses time series of monthly, seasonal, and yearly maxima of daily and hourly precipitation at observation stations in the French Alps using methods based on extreme value theory. Their purpose is twofold: firstly to provide a climatology of extremes in the region, where extremes are defined as 20-year return levels, and secondly to examine the change in these return levels between 1985 and 2022.

For the climatology of extremes, the time series are assumed to be stationary over the years, and a Generalized Extreme Value (GEV) distribution is fitted to each time series. Seasonal non-stationarity is addressed either by analysing the series month by month, or season by season. An annual analysis is also provided for comparison. They determine the circulation patterns involved in producing the largest daily precipitation events. For the trend analysis, the location or the scale parameters of the GEV, or both, are assumed to depend linearly on the year. The paper does not say which of the three models is retained at which station. A known inhomogeneity in the dataset is addressed by introducing a break in the parameters in 1985.

The results reveal a different seasonal cycle for daily and hourly extremes, and, for daily extremes, a different seasonal cycle of the relative contributions of the main circulation patterns in the north and in the south of the French Alps. Thanks to the relatively large number of stations with significant trends for daily extremes, the authors can plausibly speculate on thermodynamic or dynamic changes that might be at the origin of the trends. For hourly extremes, the arguments are more tenuous, as they sometimes depend on a large number of stations for which the trends are not significant. The reasons for the seasonal and spatial variability in significance are not discussed.

The paper is well structured and well written, although it is a little tedious to flip or scroll through so many pages to compare the text with the figures. The information presented is valuable, and the interpretation of the results can invite interesting further investigation. If the comments below can be addressed, I consider the present study suitable for

publication in HESS.

Thank you for your detailed comments! We greatly appreciate the time and effort you put into reviewing our paper and the constructive feedback which helped improve the manuscript. Please find our answers below in blue.

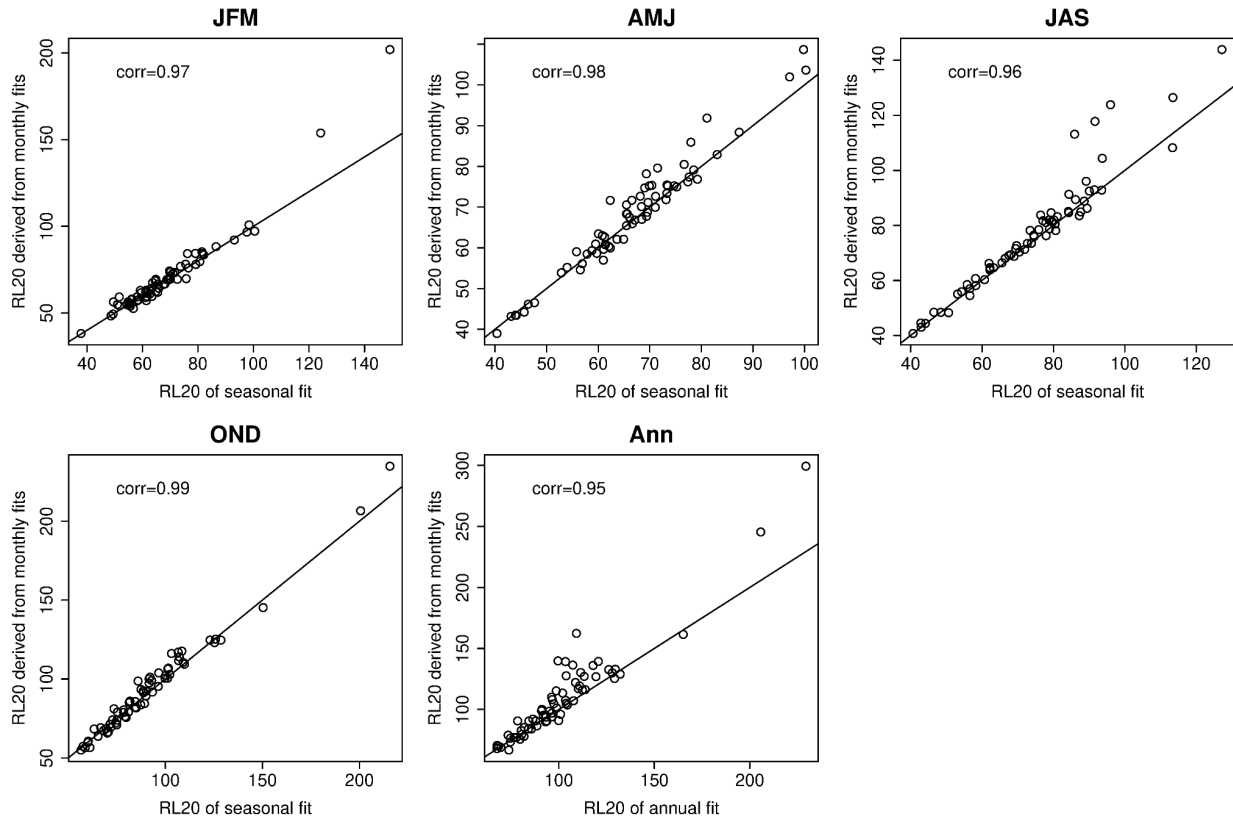
Major Comments

Major Comment 1

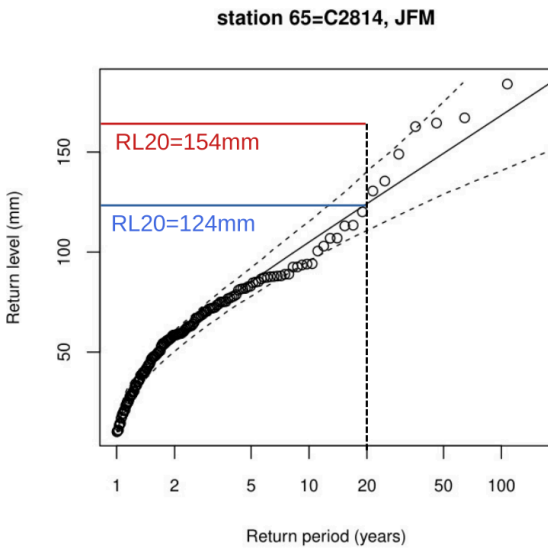
The seasonal non-stationarity of the time series is evident in Fig. 6, especially for hourly precipitation, but also for daily precipitation south of the climatological divide. In fact, the seasonal non-stationarity for hourly extremes by far exceeds the trend over the years. For daily precipitation, the assumption of seasonal stationarity is plausible in the north, but is violated in the south.

The authors painstakingly analyze the extremes month by month. I recommend that for the seasonal and yearly analyses, the return levels be numerically derived based on the monthly GEV distributions. While this represents quite an effort, I believe it would avoid introducing a bias in the results, and their subsequent interpretation would be more trustworthy.

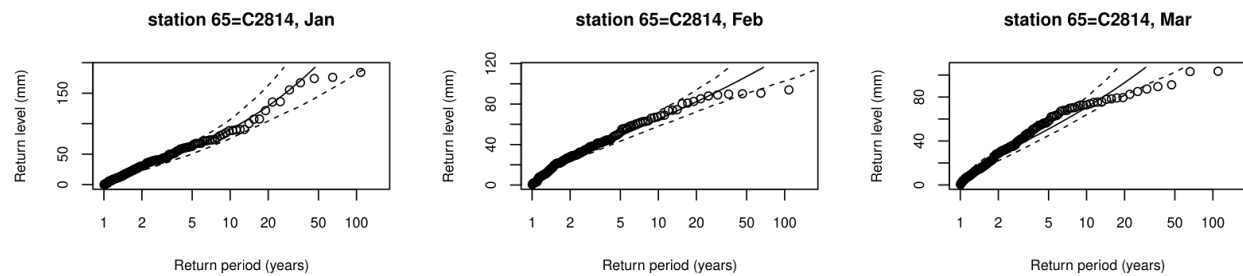
We thank you for suggesting to derive return levels for seasonal and yearly analysis from the monthly GEV distributions. The CDF of, for example, annual maxima can indeed theoretically be derived from the product of the CDF of monthly maxima. Although we understand this could help avoid potential inconsistencies, we see a major drawback in doing this due to potential error propagation. A monthly fit that would, for any reason, be biased at a given station would result in a biased fit at seasonal and annual scales by error propagation. This is illustrated in the figure below where we compare the seasonal (stationary) 20-year return level of the 67 stations with long time series (x-axis) to that derived from the monthly fits (y-axis). In case of a perfect agreement, all points would lie on the diagonal line. The agreement is in general good but we see for example in JFM two stations giving discrepancies.



For example, at station C2651 (EDF Sainte-Croix) one finds, by combination of the monthly fits, a 20-year return level in JFM of 154 mm instead of 124 mm as estimated on the JFM maxima. The figure below shows the return level plot of JFM maxima at this station. We see that the seasonal GEV fit is good. The value 124 mm seems to be a good estimate of the 20 year return level, whereas 154 mm is clearly not supported by the data.



The reason for this is that the GEV fit for the month of March shows discrepancies at this station, although the fits are reasonable for the months of January and February, as shown in the figure below.



A visual inspection of the return level plots of all stations showed that such discrepancies in the GEV fits are only found for some isolated months, whereas all seasonal and annual fits seem good, as illustrated above for station C2651. This shows the drawback of deriving seasonal/annual CDF from the monthly GEVs due to error propagation. Although it was illustrated here on stationary return levels, the same applies of course to return levels trends. For this reason we preferred to keep in the article the more classical method of deriving monthly, seasonal, and annual GEVs separately based on monthly, seasonal and annual maxima respectively.

Major Comment 2

For the trend analysis, I recommend that the authors do not include the stations for which the trend is not significant (neither in the maps, nor in the boxplots), as it suggests more information than there really is, and orients the discussion towards interpreting a change in the 20-year return levels at all stations. In point of fact, for 7 months of the year for daily extremes, and at least for 11 months of the year for hourly extremes, the majority of stations have no trend at all, if one believes the significance test. In my view, a balanced discussion should address the reasons for a non-change, as much as those for a change.

The number and location of significant stations varies considerably from month to month, especially for hourly precipitation. A short discussion would be welcome.

We understand your caution concerning possible overinterpretation of non-significant trends. Unfortunately, in extreme-value statistics, all results are by construction based on a very limited number of values. It is thus inherently difficult to obtain significance in trends. Additionally, the significance test applied here is quite restrictive, as we only consider a trend in 20-year return levels significant if the 5-95% range of the bootstrapped uncertainty interval does not contain zero. Discarding non-significant trends altogether could therefore mean discarding valuable information.

Taking statistical significance as the only criterion also raises another problem: due to the neighbourhood approach we use here, some stations in the peripheral zone of the study region have much less neighbours, and thus data, than stations in the centre. As a consequence, the trend analysis for these stations is just based on very few values (one per year, so 30-60

overall). Here, the presence or absence of single events can largely affect both trend magnitude and statistical significance.

We therefore base our interpretation on both statistical significance and spatial coherence of trends: If a larger region shows a coherent pattern this is likely not due to pure chance, even if not all the trends in that region are statistically significant.

We do, however, acknowledge that it is visually difficult to tell significant and insignificant trends apart. Therefore, we decreased the point size of non-significant trends in Figs. 8 and 10 in the revised manuscript. We also added the information on the relative share of stations with significant trends to Fig. 9. In the Results and Discussion sections, we emphasised more which results and interpretations are statistically significant. We hope that this alleviates some of your concerns.

Minor Comments

General

- The presentation of the datasets is both very detailed and somehow unclear and incomplete. Despite the authors' efforts, it is difficult to understand which data set is used for what, and which period they cover. A table would be most helpful. In particular, the reader would be interested in knowing which of the datasets used in the trend analysis contain known inhomogeneities. It is not clear.
Thank you for this observation. Also in response to Reviewer #2, we completely restructured this section to make it clearer for the reader. Notably we improved the description of the EDF datasets to avoid confusion and added a subsection 2.2.4 "Datasets for stationary and trend analysis" where summarise which subsets of the data is used for stationary and nonstationary analysis, respectively.

- Please consider merging Figs 5 and 8: the dominant weather patterns could be represented by northward and eastward looking arrowheads for the Mediterranean and Atlantic weather patterns. The reader wouldn't need to flip/scroll all these pages to get an idea which weather pattern led to the highest return levels, for instance.

Thank you for this suggestion, we implemented it (Figs. 4 and 7 in the revised manuscript). Unfortunately, the spatial distribution of the dominant weather patterns are much less visible in these merged figures. However, in the trade-off, we consider the information on the spatial distribution or return levels more relevant, so after discussion among the coauthors we agreed to replace the figures following your suggestion.

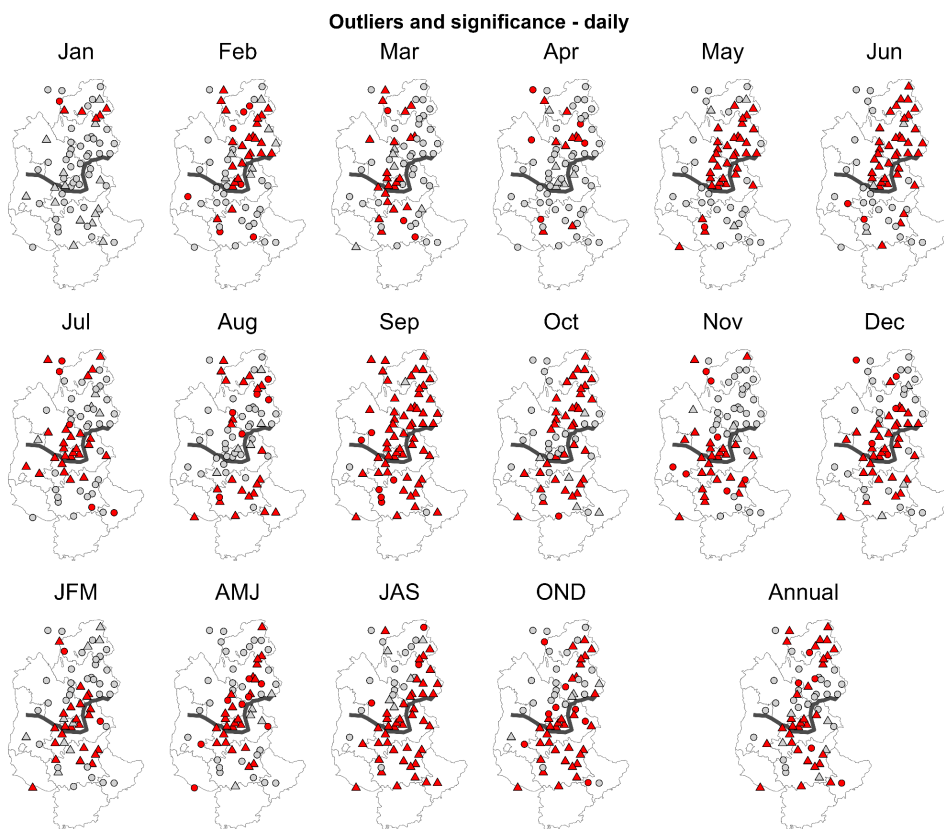
- Fig.4 could be omitted or moved to the appendix, especially since its appearance is not commented in the text.

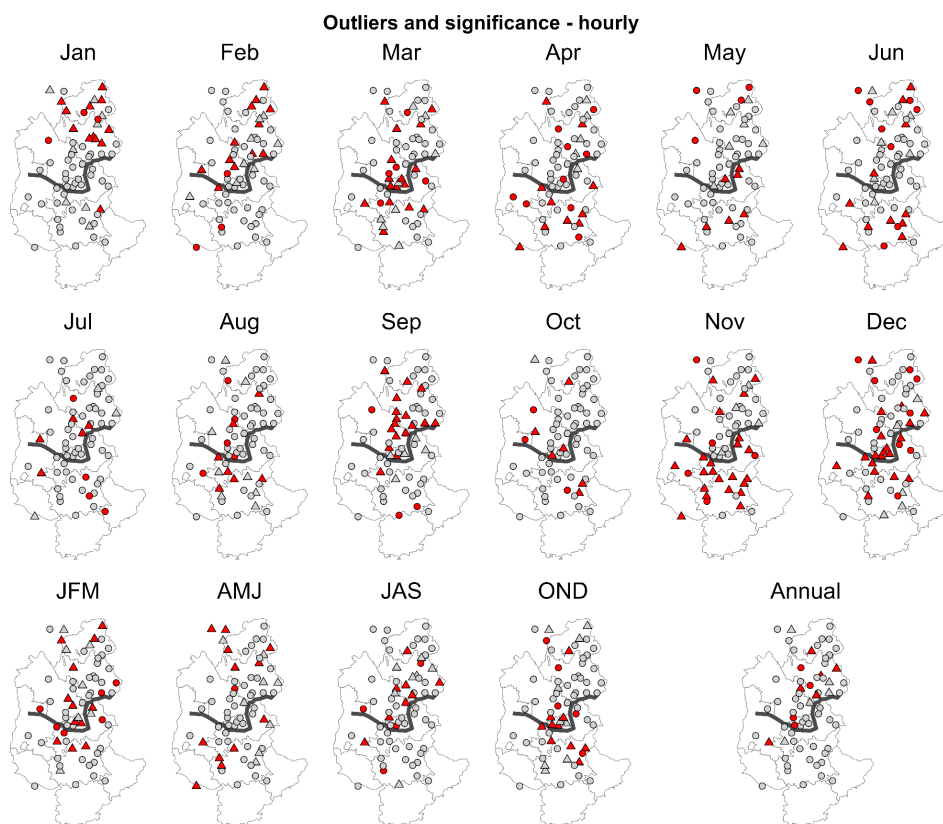
In our view it is important to point out that the robustness of trend estimates depends heavily on the number of neighbouring stations of which the observations are included. We highlighted this in the text (lines 322-324) and moved the figure to the appendix (Fig. A1) to reduce the overall number of figures in the text.

- The uncertainties are neither shown nor discussed. For the trend analysis in particular, it would be important to know if, for instance, a trend of 10mm/d/10y between 1990 and

2020, i.e. an increase 30mm/d, is within the uncertainties of the 20 year return level in 1990.

Our apologies if the following answer is not entirely exact, as we are not sure if we fully understand your comment. In the non-stationary case, uncertainty intervals are included as we use them to assess significance: a trend is considered significant if the 5-95% uncertainty interval does not contain zero. In the stationary analysis, we indeed chose to not show uncertainty intervals. As a test, we plotted the following figures, where significant trends (following our significance test) are shown with triangles, and trends that are outside the 95% confidence interval of the stationary estimate are shown in red. It can be seen that most triangles are also red (and the inverse). We therefore argue that this analysis does not yield much additional information, and chose to not include it in the revised manuscript for the sake of conciseness.





- Avoid “However” at the beginning of sentences. “Blablabla, however, blablabla”.
Thanks for the remark, we modified some of the sentences in question.
- “timeseries” or “time series” : please choose one or the other.
We made sure to use “time series” throughout the paper.

Line-by-line

In the following we only answer where necessary. We adopted stylistic and grammatical corrections as suggested or clarified where needed. We really appreciate your detailed suggestions, which helped improve the manuscript.

Line Number	Comment	Answer
12	What is meant by “partially large differences”?	Also in response to Reviewer #2, we completely reformulated this part of the Introduction and hope this solves the question.

23	"Also southeastern France was not spared from such events recently". → "Even southeastern France was not spared". Or "has not been".	ok
25	"lead" → "led" (preterite)	ok
31	"are more exposed" → "are particularly exposed"? (exposed more than what?)	ok
36	"is not fixed between seasons and over time". Are you talking about the location of the divide? How about "changes location over the seasons and over the years"?	We are indeed talking about the location, and modified the text accordingly.
38	"the Mediterranean influence on precipitation extremes is retreating in winter and spring (in favour of the Atlantic influence), but increasing in autumn" → "the Mediterranean influence on precipitation extremes retreats in winter and spring (in favour of the Atlantic influence), but increases in autumn". Meaning: "retreats" suggests a change in location (towards the South), while "increases" only suggests a change in the strength of the Mediterranean influence. Please clarify.	Increase/decrease and advance/retreat actually go together in this case. We changed this to "the Mediterranean influence on precipitation extremes is retreating in winter and spring (in favour of the Atlantic influence), but advancing in autumn. Associated with this change in location is also a change in strength of the relative Mediterranean and Atlantic influences."
39	Both thermodynamic and dynamic what? A word is missing.	Processes. Added.
40	"thus complicating the topic of assessing possible trends." → why "complicating"? The trends themselves are not "complicated" to assess. Do you mean "complicating the interpretation of possible trends"?	For clarification, we changed this to "complicating the big picture and the understanding of physical causes of trends."
43	"that involved the analysis of a single aggregated time series of daily data from several stations in southern France" → "that involved the analysis of a single time series of daily data, aggregated from several stations in southern France"	ok

45	Shouldn't Clausius-Clapeyron be mentioned earlier where the authors talk about thermodynamic processes (line 38). Here it comes out of the blue. Why not simply: they found an increase of precipitation with temperature 1 to 3 times larger than the increase in maximum water content of an air parcel with temperature (equation of Clausius-Clapeyron)".	ok
48	"onset of clear trends" → What does "clear trends" mean?	Changed to "onset of statistically significant trends"
48	"a consensus to significant trends" → "a consensus in the significance of trends". I would even prefer "showed that, only in the 2000s did a consensus appear in the significance of trends in observations in southern France."	ok
62	"globally this assumption seems to hold true" -> what does globally mean here? A spatial average? Or does it mean "in most cases"?	Changed to "in most cases studied so far this assumption seems to hold true."
66	Super CC trends -> scaling rates or actual trends?	Changed to scaling rates.
83	"generally agree on the big picture" seems like a strange statement, when the following sentence offers no "big picture". Perhaps simply say that "Studies in the European Alps indicate that trends in daily heavy precipitation"?	Changed to "show how the complex interplay of synoptic patterns and local topography influences the spatial differences in trends in daily heavy precipitation."
83-85	"Trends in daily heavy precipitation are different depending on the region and the period studied" → isn't it normal that the trends should depend on the period studied, regardless of changes in processes?	Changed with previous comment.
90	"much closer correlated" -> "much more closely correlated"	ok
111	"28 stations with more than 45 years, longest 70 years" → "the longest with 70 years"	ok
122-155	Very nice description!	Thank you!

180	“Only weekly maximum precipitation totals for 1h, 2h, 3h, 6h, 12h and 24h were archived and could be retrieved for our study” → are these gliding sums (for instance for 1h, every 6 minutes the sum over the last 10 * 6 minutes, or for 12 h every hour the sum over the last 12 hours, or for 24 h every hour the sum over the last 24)?	Answer to this and following comments: We completely restructured the descriptions of the datasets and also removed some errors regarding the number of stations in the different datasets. We took all your comments into account in this revised version.
183	“The dataset consists of quality-controlled daily and uncontrolled hourly observations for 68 stations” → The dataset consists of daily and hourly data at 68 stations, of which only the daily data is quality-controlled. “uncontrolled” sounds strange.	See above.
184	“A quality control of the hourly data was thus necessary.” → Indicate here that you can use the quality-controlled daily data to calibrate the hourly measurements.	See above.
184	“For this purpose, we scale the hourly values, i.e. we multiply all hourly values of a given day by a uniform factor such that their 24-hour sum from 6am to 6am on the next day equals the corresponding daily value.” → a little unclear. Perhaps it would help to break it up. “The daily precipitation of a given day is the amount of precipitation that has accumulated in the 24 hours between 6am and 6am of the following day. Should the sum of the 24 hourly measurements within this time span differ from the daily sum, we scale the hourly values, day by day, i.e. multiply each hourly measurement by a uniform factor, such that their 24-hour sum from 6am to 6am on the next day equals the corresponding daily value”	See above.
192	“were overlapping” → “overlapped”	See above.
201	“Intensities” → are these mm/h? If not, please use “values” or “amounts” or “observations” or “accumulations”.	See above.
208	“attributing weeks between two months” → “attributing weeks straddling two months”	See above.

212	"contains 177 time series" → 89 MeteoFrance + 67 EDF?	See above.
213	For the trend analysis, we only use the 67 time series including more than 30 years of non-missing data. → of EDF? This is very confusing.	See above.
222	Evolution of relative occurrence of each WP → never used again? Why bother? Especially if no explanation or comment is made?	Thank you, we removed the right side of the figure in the revised manuscript.
224	"we first match" → is the "first" necessary?	Removed.
225	"of of"	ok
255	Which dataset are you using here? The 27 EDF stations with a combination of old and new?	We added "Here it should be recalled that for the non-stationary analysis we use only EDF time series." Together with the reformulated Data section this should be clearer now.
261	"for cases when all the density of maxima shifts in time" -> "when only the density of maxima shifts in time"?	The whole density function/curve, we changed this. See also Fig. 3 of Blanchet et al (2021): https://doi.org/10.1016/j.wace.2021.100356
265	"we adopt a bootstrap approach by resampling the series of maxima 100 times with replacement" → isn't 100 a rather small sample?	Thank you for the remark. We increased the number of resamples to 500.
277	"so these are simply discarded" → the stations?	We did this in a previous stage of the project, but noticed that it is actually not necessary. This sentence was there by mistake and we removed it.
297	"After some initial testing we decided against doing this, following Evin et al. (2016) (Sect. 7.2.3 therein)" → please provide a summarized reason.	We added: "As Evin et al. (2016) (Sect. 7.2.3 therein) have shown, this type of declustering techniques reduces the sample size and tends to alter the statistical properties of the distribution."

319	“In the north Atlantic influence is dominant during the whole year” → “In the north, the Atlantic influence dominates during the whole year”	ok
320	“In the south there are clear differences between the seasons, with Mediterranean dominance extending far north in spring (April and May) and autumn (September and October), while in the other months Atlantic and Mediterranean circulations have roughly equal influence on maxima in this region” • “in this region” is confusing. It gives the impression we are maybe talking about something else than just “the south”. • “far north” is also confusing. “all the way to the climatological divide”? • “while in the other months Atlantic and Mediterranean circulations have roughly equal influence on maxima” → does not seem quite accurate, except if one looks at the seasonal summary. I would go about it differently, and say that “the influence of the Atlantic and Mediterranean circulation oscillate in opposition of phase, the Mediterranean circulation having the upper hand in spring (April and May) and autumn (September and October). When analyzed seasonally, the Atlantic and Mediterranean circulations lead roughly to the same percentage of maxima in winter and summer”.	Following your suggestion, we changed this to: “In the south, the influence of the Atlantic and Mediterranean flow oscillate in opposition of phase, the Mediterranean flow having the upper hand between April and June and September to November. When analysed seasonally, the Mediterranean flow remains dominant throughout the year.” In the revised manuscript, we only show the new seasons (JFM, AMJ etc.) for both stationary and non-stationary results to increase clarity. As a result, Mediterranean influence dominates throughout the year when analysed seasonally (but not monthly).
327	“The stationary 20-year return levels of hourly precipitation are shown in Figs. 6 and 9.” I would suggest immediately bringing the reader’s attention to the seasonal cycle: “The stationary 20-year return levels of hourly precipitation reveal a seasonal cycle remarkably	ok

	different from daily precipitation and across the north-south climatological divide (Figs. 6 and 9)."	
328	Evidently, hourly precipitation extremes are not stationary over the spring and autumn seasons. How does this affect your seasonal analyses?	This shows the added value of looking at the monthly level, which is not frequently done. Regarding the seasonal analyses, that means that the seasonal maxima are not representative of the whole season. We added this aspect to the Discussion (lines 472-475).
335	"almost no EDF stations is located in the Rhône valley" → "almost no EDF stations are located in the Rhône valley"	ok
341	"Here we use the long time series (EDF stations >30 years)" → are these the 27 stations with the combination of old and new data?	Changed to "Here we use only the long EDF time series which, as noted above, 370 include almost no stations in the Rhône valley or along the coast where stationary return levels are highest." Together with the revised Data section this should be clearer.
341	"We now turn to the trends in extremes (20 year return levels)." → Please add "over the period 1985 – 2022" if that's what it is.	ok
342	"in in"	ok
343	"are not for the exact same regions" → "are not for exactly the same regions"	ok
352	"positive trends in May" → how much of this could be attributed to a seasonal shift (maxes that normally would take place in June taking place in May)? The month of June experiences negative trends.	Thank you for this interesting remark! The shifts in the seasonality of extremes would more generally be an interesting question to study, although out of the scope of this paper. We added this aspect to the Discussion section, line 535-536.
372	"Annually, trends are centred around zero with comparatively large variability in the north and overall slightly positive in the south." → Actually, if one only looks at the stations with a significant	Looking at significant trends only actually gives positive trends in the north, and in the south only two stations show a significant trend, one of which is quite peripheral and has

	trend, there is virtually no variability in the north and it is inconclusive in the south.	little data. For a detailed reasoning on the use of statistical significance, we refer to our answer to your major comment #2.
411	"However, even though..." → it would be more elegant not to start the paragraph with "However". "It is interesting to note that...", perhaps?	ok
412	"This is probably due to the combination of yet relatively short timeseries" → "This is probably due to the combination of relatively short timeseries"	ok
416	"and the additional observations from Météo-France's network starting in the 1990s starting to become long enough to be included" → "and the additional observations from Météo-France's network starting in the 1990s soon to become long enough to be included"	ok
465	Please specify what "a drying in this region" means. Number of wet days?	We changed this to "generally less precipitation in this region".
466– 476	Much of this discussion is based on the assumption that there is a west-east contrast in trend, when in fact, the mentioned negative trend in the western part of the domain is significant at only a few stations.	Thank you for this observation. We refer again to our answer to your major comment #2. We also clarified in the text which part of our discussion is based on which level of confidence in the results.
477	"continuous observations by Météo-France are yet too short" → "continuous observations by Météo-France are still too short"	ok
482	"However, they also cover just a limited period of time and come with different limitations, such as masking in mountainous regions." → and inhomogeneities in time?	That is correct, we added this aspect.
492	"as we show for example in Fig. 11" → "as we show in Fig. 11"	ok
497	"generalised extreme value theory" → "extreme value theory"?	ok
498	"Regarding the spatial and temporal of extremes" → is there a word missing?	evolution
511	"which leads us to state that" → "state" makes it seem as if it were a fact, when it is simply an interpretation. "argue" would be better.	ok
517	"On the other side" sounds a little strange. → "In contrast"?	ok

Figures and Tables

Fig. 8	Change “light red Mediterranean” color to dark blue.	Following one of your previous comments, we merged this figure with what is now Fig. 4.
Fig. 5 and 9	Add a note that within a month/season/year, the size of the dots corresponds to the amplitude of the 20-year return levels.	ok
Fig. 11	Indicate that the boxplots include both stations with significant and non-significant trends.	Ok. We also added the information on the proportion of significant trends.
Fig. 10 and 12	Replace stations where the trend is not significant with empty circles.	We preferred to keep them, but decreased the size of points with non-significant trends. See our reply to major comment #2.
Fig. 10 and 12	Specify that upward triangles correspond to an increase and downward triangles correspond to a decrease in the 20-year return level.	ok

Reviewer #2

Summary

The paper analyses recent daily and hourly precipitation extremes in the French Alps using in-situ station data from Météo-France and Électricité de France. Looking at 20 year return levels, the authors present an updated climatology of precipitation extremes as well as a trend analysis. The connection to dominant weather patterns during extreme events provides hints as to the drivers of the observed extremes. The analysis is valuable not only for stakeholders in the region, but to the broader scientific community in providing insights into changes of precipitation extremes in topographically complex regions and the drivers thereof. This is especially useful since observations of extreme precipitation are generally short and model projections are yet highly uncertain. I recommend a publication of the manuscript after the following comments are addressed.

Thank you for your detailed comments! We greatly appreciate the time and effort you put into reviewing our paper and the constructive feedback which helped improve our manuscript. Please find our answers below in blue.

General comments

Abstract: You focus largely on the motivation and the methods in your abstract. You only describe the results on a very high level and do not mention any interpretation, implications or needs for further investigation you identified. Please extend on this.

Thank you for this comment, we completely reformulated the abstract.

Study area and data: Your presentation of the station data is very confusing:

1. (1) I count 32 EDF old data stations in Fig. 1, but the text in Section 2.2.2 states there should be 27 stations for the long term data, circled in Fig. 1. It is also unclear to me how these 27 are exactly derived from the 65 rain gauges, only 48 of which are from EDF (are only these 48 considered or all 65?), that have long term data, and the 68 provided by EDF with more recent data. You state that these are derived by “combining old and recent data” but it is not clear to me what that means. Do some or all of the 68 recent stations overlap with the old 65, or 48? There seem to be 68 EDF stations (old and recent together) in Fig. 1, so it seems to me that the old datasets always also have recent data because they always have colored circles. You state that if “the two datasets were overlapping, the more recent dataset was used”. Does that mean that there are two separate stations in the same location?
2. (2) I only find 83 Météo-France stations in Fig. 1 as opposed to 89 mentioned in 2.2.1.
3. (3) The introduction and Section 2.2.5 state that there are overall 177 time series in the final “recent” dataset. Summing up the (potentially) 89 time-series from Météo-France with the 68 recent EDF stations only yields a total of 157 time-series. Please elaborate where the other 20 come from. I only count 151 stations in total in Fig. 1, and about 160

the “Annual” panel of Fig. 8. Please check all of your figures and numbers in the text for consistency. If any station is for some reason excluded from any analysis, either exclude it from the whole dataset or clearly describe the procedure which stations are included were.

4. (4) In section 2.2.5 you furthermore state that you use 67 of the 177 for trend analysis. Please mark these either in Fig. 1 or reference Fig. 4 here. Fig. 4 only contains 66 stations according to my (admittedly limited) counting capabilities, not 67 as stated in 2.2.5, please double check.

Overall, you should considerably improve sections 2.1, and 2.2, as well as Fig. 1 to be as concise as possible about which stations you actually analyse and where they come from. Please integrate section 2.2.5 into a new overarching “station observations” section together with 2.2.1 and 2.2.2 or at least move it further up to be directly after the first two subsections on the observation data to avoid further confusion. Please also consider dropping unnecessary information in these sections as for example naming all individual départements in the text, the detailed description of the study area geography including individual valley and mountain names, as well as the history of the station networks in sections 2.2.1 and 2.2.2 (and the introduction). Lastly, please elaborate on the time period you are looking at, do the Meteo France observations also stop in 2022 as do the EDF data? From when to when does your 30 year period reach for the trend analysis?

Thank you for the detailed observations. We see that the presentation of the data was both unclear and inconsistent. We therefore reformulated section 2.2 and hope this helped make it clearer to the reader. We also removed errors in the numbers.

Concerning the details on the study area, we think they contain valuable information to many readers and thus prefer to keep them. Also, we received positive comments on this part from reviewer #1. Keeping it in a separate paragraph should make it easy for readers uninterested in this particular part to skip it.

Neighborhood Approach and trend estimation:

1. (1) In section 3.2.1 you mention that you discard stations where the parameter estimation of ξ is unrealistic, please state how many these are and how many are then left in the dataset. Is this different from month to month / season to season?

We did this in a previous stage of the project, but noticed that the problem could be resolved by constraining $-\infty < \xi < 1$. This sentence was left by mistake and we removed it in the revised version.

Please elaborate if the fact that your study region is in a highly complex terrain affects the validity of the neighborhood approach. Do you employ the neighborhood approach across the climatological divide? How robust is your assumption that stations close in space share a similar climate in this terrain, considering the dominant weather patterns causing extremes vary, even between stations that are close together?

We do indeed apply the neighbourhood approach with a fixed radius regardless of the topography. As we mention in l. 408 we tested for different radii and found that the spatial patterns remain overall similar. In our experience, the strong influence of orography on mean precipitation is much smaller or even non-significant for extremes. In previous studies (Evin et al, 2016, <https://doi.org/10.1016/j.jhydrol.2016.08.024>) we have used a 'crossing distance', taking orography into account, but during preliminary analysis for the study of Blanchet et al (2021, <https://doi.org/10.1016/j.wace.2021.100356>) we did not find that this was significantly better than the simpler Euclidean distance. We added this information in lines 319-322.

As we show in Figs. 4 and 7 and as also Blanchet et al (2021) show, the location of the climatological divide, with regards to extremes, is not fixed between months and also over time, so taking it into account explicitly does not seem relevant. It should also be noted that the climatological divide of Auer et al (2007) has been determined on the basis of several climatological variables and on the much larger scale of the European Alps. In our study, we thus take it more as a guide than a hard border. We also added this aspect in the Introduction (l. 36-39).

2. (2) What exactly is it that you call "trend"? Is it μ_1 ? If so, is the trend for stations that choose $M\sigma$ as their best model, automatically 0? Or do you fit a trend slope to the 20-year return level values inferred from the non-stationary GEVs? How is this trend slope calculated? Are the trends that you show then by definition from 1985 to 2022? A trend is the slope of the 20-year return level derived from the non-stationary GEVs (slope of the solid red line in Fig. 3A). These trends are indeed calculated over the period 1985-2022, and scaled to mm/d/10y for better interpretability. We added "Finally, we derive from the selected model the evolution in 20-year return levels, using Eq. 3 (red solid line in Fig. 3). The choice to focus in this study on the 20-year return level is arbitrary, it aims at documenting changes in very rare events that are encountered only a few times in a human's lifetime. Patterns associated with 10-year and 50-year return levels would show similar patterns." (l. 284-287).

Map figures:

1. Please plot the in-situ maps using a discrete colorbar to enhance the interpretability of these figures.
Thank you for the suggestion, we updated Figs. 4,7,8, and 10 accordingly.
2. I find that the monthly maps do not add significant value to your interpretation and recommend moving them to the appendix. Most arguments you make can be made based on the boxplots or with a few references to the appendix.
Thank you for this suggestion. We understand your point, however, in our opinion a map is more visual and thus easier to interpret with regards to spatial patterns, especially spatial coherence of trends. This is even more important as the location of the climatological divide varies between months and seasons, and thus splitting the region into north and south along a fixed line is to some extent arbitrary (see our reply to one of your previous comments).

3. The motivation for shifting the seasons for the trend analysis is well made and the maps in support of it can also be moved to the appendix.

Thank you for this suggestion. In the revised manuscript, we only used the 'new' seasons (JFM etc.) for both stationary and non-stationary analysis, and explained this decision in the relevant paragraphs of the Methods and Discussion sections. This reduced the number of subfigures in Figs. 10 and 12 and improved the overall clarity of the manuscript, while maintaining a coherence between the stationary and non-stationary analyses.

In this way, you can limit the panels in the map figures to always four seasons and the annual map, greatly reducing the overall number of panels in the main text and consequently improving on overall readability, keeping the readers focus on the significant findings.

For the reasons named above we kept the monthly panels in the figures, but removed the traditional seasons in Figs. 8 and 10 (and changed the seasons from traditional to new seasons in Figs. 4 and 7). Following the suggestions by Reviewer #1 we additionally removed the right panel of Fig. 2, and moved Fig. A1 (previously Fig 4) to the appendix. We also merged the figures showing stationary 20-year return levels and dominant weather patterns (now Figs. 4 and 7), coding the WPs into symbols. Although this removes the visual clarity of the distribution of dominant weather patterns, we considered this easier for direct comparison between return levels and dominant weather patterns. This way, we reduced the overall amount of figures and improved readability.

Please discuss the significance of the trends you find in more depth. Are there differences in significance between seasons/months or the daily and hourly aggregation levels? Do you have hypotheses as to why? How many stations actually show significant trends? You state in the abstract that trends in hourly extremes are rarely significant but do not mention this in the results.

Thank you for this important remark. We added the proportion of significant trends to the boxplots in Fig. 9 and put a higher emphasis on which results and interpretations are statistically significant throughout the manuscript.

However, we would also like to emphasise that in extreme value statistics it is inherently difficult to obtain significant trends due to the limited amount of datapoints (one value per year). We thus base our interpretations not only on the statistical significance of trends for individual stations, but also on the spatial coherence or non-coherence of trends regardless of their level of significance. We further elaborate on our reasoning in our answer to Reviewer #1, major comment #2

Specific comments

Line 8: "In the first part, we use a stationary framework....", you do not mention a second part in the abstract. Please do so and describe where you use the non-stationary framework and why. We completely reformulated the abstract following your first comment.

Line 63: "Furthermore, it is expected that the trends in intensity should be higher for subhourly

durations" please quickly explain why. Later in this paragraph you give explanations for super-CC scaling, please clarify if you therein focus on hourly or sub-hourly extremes. We restructured the paragraph to make this clearer. Specifically, we replaced "expected" by "generally observed". Regarding super-CC scaling, this is observed for both hourly and subhourly extremes, we modified the sentence accordingly.

Figure 1: This figure shows the study area together with rivers, départements borders, selected cities, orography and the three different station types. Please omit numbers and names of départements that are not actually used in the study. Moreover, please plot the stations above river names or département numbers, some of the stations are almost completely (and some maybe completely?) covered by them, see for example the "Drac" river label.

Thank you for these observations, we agree that labels should not cover stations, and modified the figure accordingly. Please note however that some few stations are very close to each other (less than a few hundred meters apart), and are thus difficult to distinguish on the map (for example in Villard-de-Lans, southwest of Grenoble).

Line 131 and 136: Please explain what "structured" means in this context. Orographically structured. We clarified this.

Lines 139 to 155: Please provide some references for this climatological assessment of the study region.

Thank you for pointing this out, we added relevant references (Blanchard, 1921; Auer et al, 2007; Garavaglia et al, 2010; Metzger, 2023)

Line 157: You already mention EDF as Électricité de France in the Introduction, please give the abbreviation at the first mention in the typical format "Électricité de France (EDF)".

Ok.

Line 175: "a change from mechanical to digital registration of observations took place in the 1980s, which seems plausible given the obtained dataset": please explain how this "seems plausible".

The 'old dataset', of which we only have access to weekly maxima, ends sometime in the 1980s (different end date between stations). We assume that this coincides with the change to digital registration. We changed this sentence to "...which seems plausible given that the datasets we obtained are structured very differently before and after this period".

Line 184-189: The description of the quality control procedure should go into its own subsection, for the sake of not distracting the reader here. You could for example rephrase the part here to "We employ a quality control procedure described in section 2.2.x. One station was removed during this procedure, yielding now a total of ??? stations in the dataset." Moreover, since we do not know the names or numbers of any of the other stations, and neither the quality-controlled nor the raw data is publicly available, it is not very meaningful to provide this for the one station that is removed, in the text.

Thank you for this suggestion. Following your earlier comments we completely restructured the data section, and took this into account. We notably added a subsection on the quality control (2.2.3).

Line 192: Please explain what you mean with the higher resolution of the recent dataset. Do you mean higher temporal resolution?

We do indeed mean higher temporal resolution. We reformulated the dataset descriptions to make it clearer, the information previously in line 192 of the original manuscript is now included in lines 192-197.

Caption Fig. 2: Please provide a reference for the LOESS smoother.

Following reviewer #1, we removed the right side of the figure altogether. Nevertheless thank you for pointing this out.

Line 226: "the maximum of year and season or month m" please rephrase for improved understanding, e.g. "the maximum of season or month m in every respective year t"

Ok.

Line 159: Please give examples of the mentioned literature where ξ is kept constant.

In the revised version we cited some references where ξ is kept constant, including one of the standard references for nonstationary extreme values analysis (Katz, 2013).

Line 264: 100 resamples don't seem very many?

We increased the number of resamples to 500 and updated the figures.

Line 273: Please briefly elaborate on the theoretical considerations that render the bounds $-\infty < \xi \leq 1$ reasonable or give a suitable reference.

We added "Global analyses (e.g. Koutsoyiannis, 2004; Ragulina and Reitan, 2017) found that for extreme daily precipitation $\xi \approx 0.15$. While allowing for some flexibility, we consider cases where $\xi > 1$ unrealistic..."

Fig. 4: Please also plot the climatological divide here and mention in the caption that these are the stations considered for the trend estimation.

Following Reviewer #1, we move this figure to the appendix, but we added the climatological divide.

Line 337: "As the weather pattern classification by Garavaglia et al. (2010) used here is done at daily scale, it is less relevant at hourly scale so we choose to not show the detailed maps here." Please clarify that while the reference work only uses the technique at daily scale, you still apply it to the hourly scale and why you conclude that it is less relevant. Just because the reference did not do so does not mean it is not relevant. In my opinion, the fact that Anticyclonic conditions are more dominant drivers for hourly extremes in summer, is a hint for a larger role of convection for these extremes and is, in fact, very relevant. Please provide the maps for this analysis at least in the Appendix.

Thank you for this comment. Applying these weather patterns for hourly precipitation is less robust as a consequence of the methodology used by Garavaglia et al. (2010) to classify the different weather patterns. However, we agree with you that it still bears some relevance. We also replaced the word 'relevant' by 'robust'.

Following Reviewer #1 we included the spatial distribution of dominant weather patterns in the figures showing 20 year return levels, for both daily and hourly analysis.

Caption Fig. 8: "generating monthly, seasonal or annual maxima..." Please add. Also please make sure all figure captions are harmonized (e.g. "First and second row...." description for all or none of them).

Thank you, we harmonised the captions.

Line 346 and later: It is sufficient to include references for the climatological divide the first time you describe it and can be omitted thereafter.

Ok.

Line 351: "They show significant variability temporally (between months and seasons)" → they show significant variability between individual months/seasons - to avoid confusion. Similarly, please change in lines 363-264.

Ok.

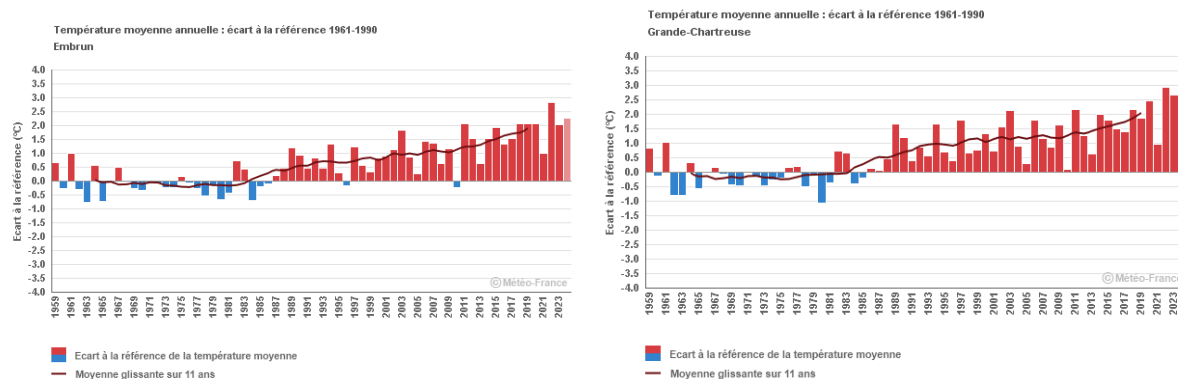
Figures 6 and 11: If you cut off the outliers, one would be better able to see the patterns in the boxplots, especially for the trends in hourly return levels, where the boxplots become very small. In any case, please state at which quantile the whiskers are cut off / what is considered an outlier.

Thank you, we modified the figure following your suggestion.

Line 394 ff.: When you contemplate on the suitability of time as the sole predictor of your non-stationary return level estimation it would be useful to provide evidence that temperature in the study region increases somewhat linearly with time, motivating that a climate change signal also in extreme precipitation should reasonably correlate with time.

Thank you for this remark. Météo-France have done this for selected homogenised time series of temperature measurements, the results of which are available at <https://meteofrance.com/climathd>. For example, the stations of Grande-Chartreuse and Embrun are located in our study region and show approximately linear temperature trends since the mid-1980s (Figs. below, see also Gibelin et al., 2014, DOI:10.4267/2042/54336). We referenced these in the revised version of the manuscript as well (l. 430). Actually, these figures provide another reason as to why choosing a breakpoint with linear trends starting in 1985 as we do in our study seems reasonable.

It should also be noted that the best choice of covariates is an ongoing discussion in the scientific literature (e.g. Roderick et al, 2020, DOI:10.1029/2019WR02692, or Schröer & Kirchengast, 2018, DOI:10.1007/s00382-017-3857-9), and entering this discussion would be out of the scope of our paper.



Mean annual temperature anomalies for Embrun (Hautes-Alpes) and Grand-Chartreuse (Isère). Figures taken from <https://meteofrance.com/climat/d> with data by Gibelin et al. (2014, DOI:10.4267/2042/54336)

Line 406: "However, in our experience this is difficult to implement and not robust, which is why we decided against using this method in our study." Please elaborate on how the mentioned approach is not robust.

In our experience homogeneity tests such as the Hosking and Wallis test tend to be too permissive, i.e. they tend to falsely classify nonhomogeneous region as homogeneous thus resulting in too large neighborhoods (see Evin et al., 2016 for a simulation study). We clarified this in the manuscript (l. 438-441).

Line 415: "we here arrive at the limits of studying precipitation extremes based on surface observations." Please add "to date" at the end or similar, to express that you expect future studies to perform better, not that a trend estimation in extreme precipitation is never and will never be possible with surface observations.

Thank you for pointing it out, we added "current".

Line 444 f.: "Second, the spatial distribution of hourly extremes, showing higher return levels along the southern and western slopes and pre-Alps compared to the inner Alps, a pattern consistent with the findings of Fukutome et al. (2015); Panziera et al. (2018) for Switzerland." Please share with the reader hypotheses as to why these patterns are observed if available. When unstably stratified air masses reach the first mountain ranges, convection can be triggered by the orography. Outside the mountains as well as in the inner Alps, convection occurs mostly due to spontaneous thermodynamic triggering. Fukutome et al. (2015) also note that there is a lack of sufficient moisture convergence in the inner-Alpine zones. We added the sentence: "Panziera et al. (2018) attribute this to the role of the first Alpine slopes triggering convection in unstable airmasses arriving from the west or south, and Fukutome et al. (2015) note that there is a lack of sufficient moisture convergence in the inner-Alpine zones."

Line 448 ff.: You state that you find "clear differences between the north and the south of the region" for both daily and hourly trends in extreme precipitation. However, in section 4.2.2 you

state that "the clear north-south pattern in trends found for daily extremes in autumn is completely absent in the hourly data" and do not mention a clear north south divide for the other seasons; rather you talk about a potential east-west gradient. Please streamline the statements made to be better aligned, presenting a potentially more differentiated view on the data in both sections.

Thank you for highlighting this. We clarified that here we are talking about daily trends, and improved the whole section 5.3 following comments of all reviewers.

Line 468: "but also not totally in line with..." please rephrase, colloquial. It is not clear how the hourly trends being different from the daily ones would support consistency with CC scaling and how both being similar does not. Moreover, it would be useful to give the reader information about the warming in this region over the study period. Without this, it is hard to infer if the results are in line with CC or not.

We rephrased the sentence on question to "While this is clearly different to the daily trends, negative trends or trends close to zero are also not in line with the increase which could be expected by the Clausius-Clapeyron relationship, given that we observe a temperature increase in the region (Gibelin et al., 2014)."

Line 469: "Thus, these trends might be the result of..." → It is not clear to me how this statement derives from the finding before, consider omitting "thus" or improve on the explanation.

We removed "Thus".

Line 474 ff.: Please elaborate on the mentioned albedo feedback loop and how it would affect rainfall, this is very interesting.

We added "This could be attributed to albedo feedback loops, where decreasing snow cover leads to additional surface warming, an effect that has been shown for intermediate altitudes (1500-2000m a.s.l.) in our study region (Beaumet et al., 2021). This in turn increases thermodynamic instabilities and the potential for convective precipitation."

Line 490 f.: Here you give trend estimates in %. It would be very useful to have changes in percent or percent per decade also in the Appendix. Otherwise, please put (not shown) here.

In order to limit the total amount of figures also in the appendix, we preferred to add "not shown".

Line 505 f.: "Regarding recent trends in extremes, we find a mixed picture of both positive and negative trends in different areas and seasons." → "Regarding recent trends in extremes, we find a mixed picture of both positive and negative trends in different areas and seasons with considerable differences in patterns between daily and hourly data." Moreover, inserting a line break before this sentence would improve the readability of your conclusion section.

Thank you, we completely reformulated the conclusion following your last comment.

Line 507: "The strongest positive trends in daily 20-year return levels tend to occur in the shoulder seasons in the south". Please rephrase: "We find significant positive trends in daily

20-year return levels in...". This clarifies that the trends are actually high, and does not allow for speculation if the strongest trends of overall low trends are found there. Moreover, please explain what shoulder seasons are, or, since you do not use this term anywhere else, just name the seasons you mean explicitly. Lastly, from here on until the end of this section, it is not entirely clear what you mean with "summer" or any other season, if it is the traditional seasons or the ones you newly defined. Removing the traditional seasons from the figures would help clear this up as well.

Thank you for your suggestions. As discussed earlier, we focused only on the new seasons in the revised version of the manuscript, thus removing any ambiguities in the text. Furthermore we reformulated the conclusion, taking this comment into account.

Line 509 ff.: You do not explicitly mention the June hourly trends anywhere else in the manuscript, it comes as a bit of a surprise here in the conclusion. Consider mentioning this already in the results and discussion section.

We included this in the results (l. 382, 394) and discussion sections (l. 501, 535).

Line 509 (continued): You say: "Trends of hourly extremes are overall small, with the most notable exception being the month of June where we find locally high positive trends. In summer, hourly trends are almost systematically higher than the daily trends,... " Please clarify if trends in hourly extremes are now overall small or systemically higher than daily trends. You could also replace "summer" with "June", to avoid this conflict. Moreover, please explain how you identified that hourly trends are "higher" than daily trends, because a comparison of absolute values, as you show in your figures, does not make sense, and you do not show percentage changes in the manuscript.

We acknowledge that our formulation was unclear, and we rephrased this in the revised version (l. 497-503)

Line 511: "...which leads us to state that a decoupling is taking place between processes generating extremes on the different timescales. While a general drying of the region takes place in summer, higher potential energy as a result of a temperature increase may lead to an increase in occurrence and/or intensity of short-duration convective precipitation events." Were the processes causing hourly and daily extremes ever coupled? What you describe is also not a decoupling but simply different drivers developing differently. Suggestion to simplify this: "An explanation of why summer daily extremes seem to become weaker while hourly ones become stronger, could be that a general drying of the region takes place in summer, limiting the overall potential for intensive precipitation over a longer time scale. At the same time, higher potential energy as a result of a temperature increase may lead to an increase in occurrence and/or intensity of short-duration convective precipitation events."

Thank you for your suggestion. Following your last comment, we completely reformulated the conclusions, taking this comment into account.

Line 513 f.: "This can especially be seen in early summer (June), when there is still enough

moisture available for local recycling in convective processes." This can not be seen in your data as you do not show available moisture. Please provide a reference for this or clarify / better support this argument.

Following your last comment, we completely reformulated the conclusions.

Line 516 ff.: "Overall, the increases we find are strongest for daily extremes in autumn (October to December) and in June and November for hourly extremes. On the other side, September is the month experiencing strongest negative trends of extremes (although less for hourly extremes), which we attribute to an increase of aridity in summer and a later onset of the rainy autumn season as compared to the past." Please consider inserting another line break before this part. Moreover, similar to my comment to line 507, consider rephrasing this to be less ambiguous. Lastly, I recommend giving a more high-level summary here instead of reiterating on these specifics. What are the main take away messages of this work? For example, that considering a shift between seasons is essential to identify trends in extreme precipitation? Or that you do not find consistent trends over the whole region because underlying drivers of extreme precipitation differ between time scales, seasons and the North and South of the domain? You could also highlight the "tug of war" between increased potential for convection and moisture limitation during summer.

Thank you for these suggestions, we completely reformulate the conclusion, elaborating more on the implications of our study.

Technical corrections

Line 6: in situ → in-situ

Line 13: driven by → caused by

Line 25: In October 2024?

Line 55: Given that in smaller catchments,

Line 152: "As in the wider region of the north-western Mediterranean" → "Similar to the wider region...."

Line 173: "registered by mechanic (past) or digital (tipping gauge, today) means" → registered by mechanic (past) or digital (today) tipping gauge.

Line 207: "for a given season" → for a given season/month

Line 230: "to characterize the general patterns" → please be more concise

Line 295: "sometimes" → too colloquial

Line 296: "after some initial testing" → too colloquial

Line 296: "Also, ..." → too colloquial

Line 319 and 320: In the north, and In the south, to avoid confusion (could read "the North Atlantic...")

Line 335: "stations is" → stations are

Line 431: "the the"

Line 435: "the higher amount of humidity"

Line 437: "(6)" forgot "Fig." ?

Line 459-461: Please split up this long sentence for better readability.

Line 471: "We-could-then further hypothesize that"

Line 498: "Regarding the spatial and temporal ? of extremes...." forgot a word?

Line 517: “On the other side,” → on the other **hand**,

Thank you for these corrections, we included them in our manuscript.

Reviewer #3

Climatology and trends of observed daily and hourly extreme precipitation in the French Alps

Authors: Sebastian Berghald, Juliette Blanchet, Antoine Blanc, and David Penot

Recommendation: major revision

The study of Berghald et al. describes the development of extreme precipitation in the French Alps since 1970 based on an analysis of daily and hourly station data using extreme value statistics. A better understanding of changes in extreme precipitation worldwide is of general interest, yet current knowledge is insufficient. Therefore, this study is well suited for HESS. The study follows a well-described methodological approach and describes changes for the study area. However, I see some necessary improvements to be done before the paper can be accepted.

Thank you for your detailed comments! We greatly appreciate the time and effort you put into reviewing our paper and the constructive feedback which helped improve the manuscript. Please find our answers below in blue.

Major comments:

One main problem is that the analyses are not very in-depth, which gives the impression of a technical report. This is evident from the figures presented, the analyses performed and the conclusions drawn. The paper frequently points out that the results of this study are confirmed by other studies. While it is important to demonstrate agreement/disagreement, this cannot be the ultimate goal; rather, the current work must achieve new results. This should be made much clearer in the text which results go beyond those of other studies. While the authors correctly state that their study is the first on trends in extreme hourly precipitation for Alpine region, at the same time they miss to include some relevant recent papers on extreme hourly precipitation (in particular Zeder and Fischer, 2020 and Zeder and Fischer, 2024, see the list at the end of the paper) which are relevant for the study region as well.

Thank you for making us aware of these references, we included them in the manuscript. We also made sure to highlight the novelties of our study more throughout the manuscript.

A useful extension to the work presented could be to visualize the station time series used in the analyses (e.g. the 67 series used in the trend analysis). I fully understand the authors' motivation not to interpret individual time series, due to the risk of misinterpretation. However, visualizing the time series would give readers a better understanding of changes over time, as well as providing an impression of the length of the time series and the data gaps.

Thank you for this suggestion. Unfortunately, as the dataset provided by EDF is not publicly available, we cannot show individual time series.

Another question concerns the homogeneity of the time series used. Extreme values are particularly prone to inhomogeneity. While there are currently no usable methods for

homogenising hourly values, but homogenisation methods are already available for daily values (see e.g. Mestre et al., 2013). These methods are not without their weakness, but at least they indicate breaks that are/are not present in the series, an important information that should be possible to obtain using existing break detection methods. Even if no homogenisation has been carried out, this should be mentioned and with an explanation as to why this was not done.

Regarding homogeneity, as you say this is difficult for hourly time series, especially in the case of annual (or monthly, seasonal) maxima because of the small number of values. During preliminary analysis we did a simple test, looking at the season in which the annual maximum falls for each year, see figures below. Only one station (C2430) failed this test due to a dubiously high amount of hourly maxima falling in the winter season before 1985. We mentioned this in the revised version of the manuscript.

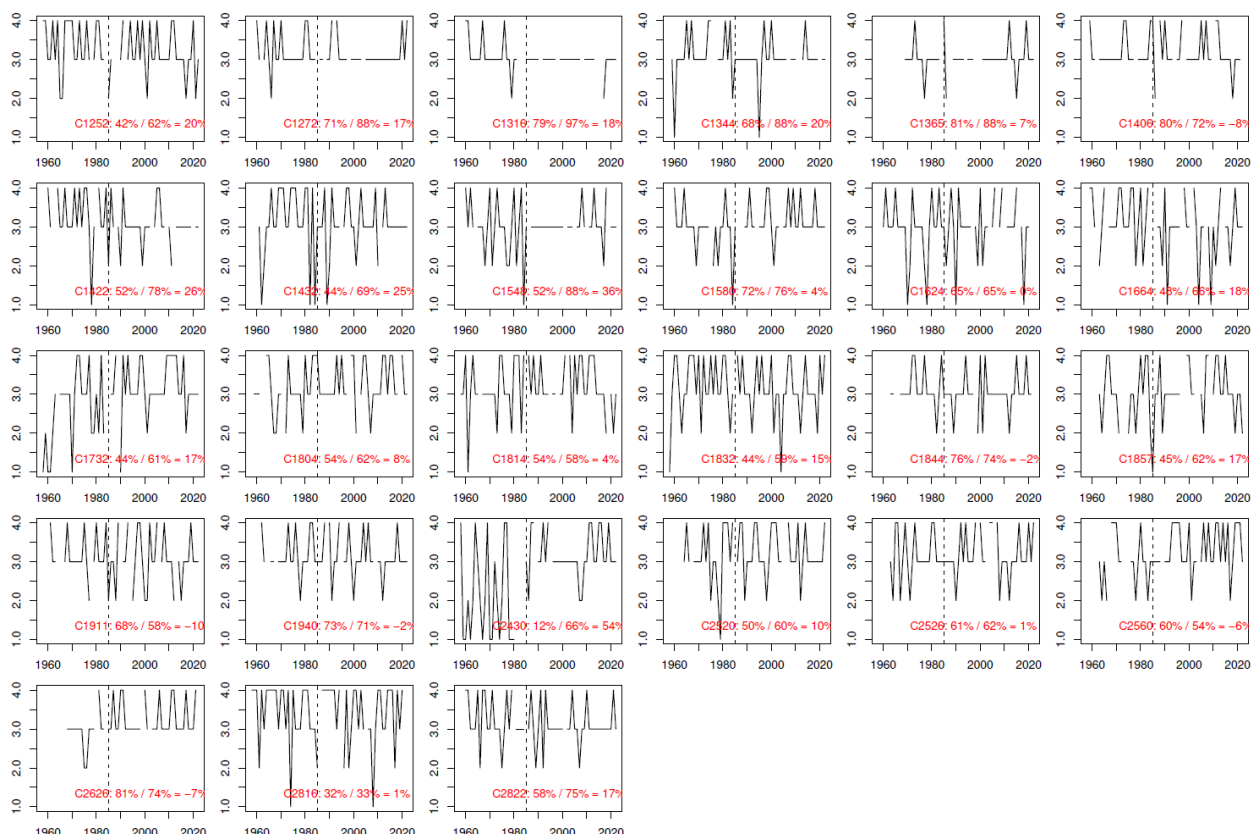
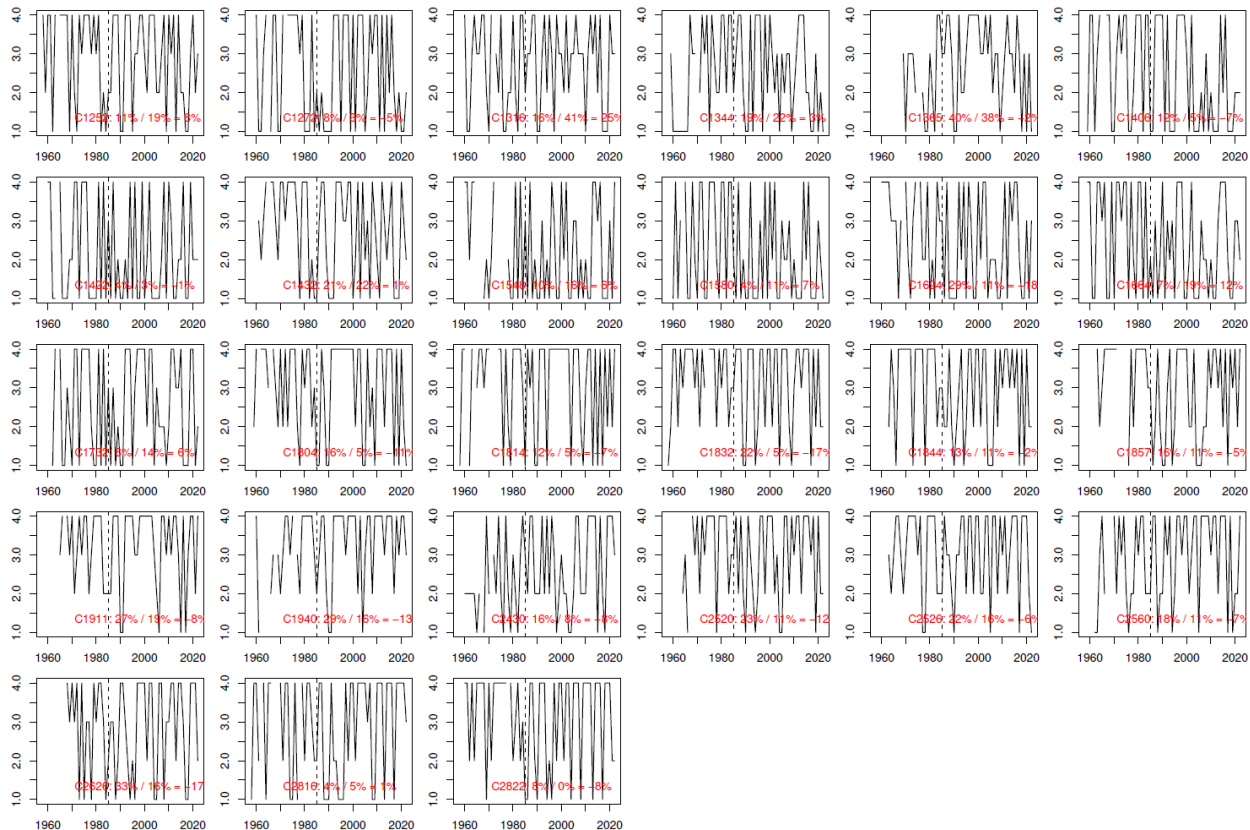


Figure showing for each year the season of the annual maximum of hourly precipitation (1=DJF, 2=MAM, 3=JJA, 4=SON). The red number gives the relative amount of maxima falling in the summer season before and after 1985 (dashed vertical line). A strong difference is found for station C2430, which led us to discard this time series.



Same as above, but for maxima of daily precipitation.

I also see potential risk of a bias being introduced by the neighbourhood approach used, as stations at the margins of the study area have fewer or no neighbours within the specified search radius. Consequently, stations in the center of the study area are therefore used much more frequently for adjusting the frequency distribution parameters.

We share the concern that as a consequence of the neighbourhood approach the results are less robust for stations at the margin of the study area. We highlighted this aspect in the manuscript (l. 322-323) and included Fig. A1 to avoid possible misinterpretations.

For regionalization of subregions (North and South of Alpine main divide) the border defined in the HISTALP data set is used (which is on a much larger region). It should be discussed in the paper if such divide is really the best choice in order to show +- robust subregions with similar extreme precipitation climatology. The main divide of the Alps must not be the relevant border for precipitation extremes.

Regarding the climatological boundary, we actually use the more detailed boundary of Blanchard (1921), which coincides roughly with the HISTALP boundary of Auer et al (2007). In our study, we use this divide purely as a guide for interpretation and mention this now in l. 36-38). In our opinion, the relevant border for precipitation extremes can best be seen by the influence of the different weather patterns (Figs. 4, 7, symbols). While these influences change spatially over the year, the boundary between Mediterranean and Atlantic influence for annual maxima actually coincides quite well with the line of Blanchard (1921) and Auer (2007).

As indicated above, another shortcoming of the work is that some of the figures presented (figures 5 to 12, which are the key figures of the study) are of low quality, both in terms of content and layout (differentiation of the values shown). Probably, some of the trend/climatology figures (maps) for individual month could be moved to the appendix and only the relevant one for interpretation could be kept in the main paper and increased in size and cartographic quality (figures 5, 8,, 9, 10, 12). Figure 11 serve as the basis for relevant interpretation and discussion in the paper but miss the information on significance of shown trends. This information could be added by bars showing the number of significant/non-significant series per box-plot. Generally, all figures have rather sparse description which should be extended to full information available in order to make understanding easy for the reader.

In response also to the comments by Reviewer #1 and #2, we made some changes to our figures:

- We removed the differentiation between traditional (DJF, MAM...) and new (JFM, AMJ...) for the whole manuscript and now show only the new seasons for both stationary and non-stationary analysis.
- In Figs. 8 and 10 (previously 10 and 12), we increased the size of the triangles (significant trends) and decrease the size of the round points (non-significant trends) to better highlight visually where the trends are significant.
- In Fig. 11, we added the information on the relative amount of stations with significant trends for the respective months.
- We extended the figure captions to provide more information to the reader.

Minor comments:

Language: could be improved (e.g. repeating “However” etc. should be avoided), though the paper is generally very well to understand.

Title: I suggest to use a more informative title (which also would avoid the impression of technical report)

Abstract: Add some information to make the presented results of statistics data comprehensible. Otherwise the pure statistical results give impression of randomness. There is the result/indication of thermodynamic vs. weather pattern driven changes in the precipitation extremes, which could clearly go into the abstract as well. Distill your results carefully.

Thank you for your suggestions, we reformulated the abstract taking them into account. After discussion with all co-authors, we preferred to keep the title as it is.

Introduction: (tried not to repeat comments from other reviewers)

Thank you for your detailed corrections, we included them in our manuscript.

75: **Unfortunately**, there is no consensual definition of what characterises an 'extreme'.

ok.

79: More precisely, we here define extremes as 20-year return levels. Can the authors give a motivation for exactly using this definition of extreme precipitation?

As explained in the manuscript, we define an 'extreme event' as an event that occurs only rarely in a human's lifetime (as opposed to, for example, the 99th percentile which by definition occurs about three times a year). Of course, the choice of the exact return level to analyse is somewhat arbitrary but, at least in our opinion, also secondary.

85/90: large-scale circulation better to write: "weather patterns" or "synoptic patterns"

ok.

90-92: large-scale circulation patterns while short-duration high-intensity rainfall seems to be much closer correlated to air temperature through increased moisture holding capacity and increased energy available for convection (see also Giorgi et al., 2016). clarify by writing ... much closer correlated to air temperature through dependency of moisture holding capacity and energy available for convection on air temperature.

ok.

93-95: ... Barbero et al. (2017) found overall lower significance and temperature sensitivity of annual/seasonal maximum precipitation of hourly compared to daily precipitation over the contiguous United States,.... Improve (significance of what?, what is mean with "contiguous US")

Thank you, improved this sentence.

99-100: Only now are observational time series becoming long enough to attempt an analysis of extreme precipitation trends on an hourly timescale ... this is bit misleading (if thought as general statement and not only for France) as the main challenge hasn't been series length but the data to be digitized. There are many data on precipitation at time scale of e.g. 10 minutes in data archives of weather services which are still not digitized going back to 1950 and earlier.

Thank you for this important remark, we included this in the manuscript (l. 105).

104 and throughout the ms: do not mix up rainfall and precipitation

We replaced 'rainfall' by 'precipitation' throughout the manuscript.

113-115: not really correct the study "get an overview over the weather patterns behind high-intensity precipitation in the region." This would need much more meteorological information.

Thank you for this remark, we replaced 'overview' by 'idea'.

139-140: The Northern Alps, under Atlantic influence, are among the rainiest regions in France with a mean precipitation of 1172 mm/year Would be good to describe the rainiest region with highest values within the region (1172mm is not that much compared to other regions in the Alps where values above 3500mm/year have been measured).

Thank you for this remark, we indeed did not know that such high annual amounts are observed in the European Alps (assuming you are referring to some specific places in the Julian Alps in Italy). However, since we here refer to France and a regional level, we consider a regional average more relevant, and in the French context a regional average of 1200 mm/year is very high.

147: can you specify the “East Return” e.g. by reference.

ok.

161: What means “large-scale automatisation”?

We rephrased this to “large-scale replacement of manually operated weather stations in France by automatised ones”.

161-162: “continuous precipitation observation”: they are not continuous but has some sampling interval (e.g. 10 minutes), which would be interesting to know, e.g. how are hourly precipitation values derived from precipitation values of highest temporal resolution? (for both Météo-France and EDF).

As tipping gauges are used for these measurements the observations per se are continuous down to a resolution of typically 0.2mm. The registration intervals are then sums over 6 minutes for MétéoFrance stations and sums over 1 hour for EDF stations. For MétéoFrance, the 1-hour sum is the sum over all 6-min intervals in a given hour. We clarified this.

171: observations are not continuous for snow, runoff and precip (have some time resolution) Numbers of stations in section 2 of the paper are bit confusing. Would be very useful to get information on how precipitation were measured from the begin of the temporally high resolution measurements until today. Are always values of hourly and daily precipitation from the same sensors? (in some countries this could be different)

Thank you for this remark. To the best of our knowledge the hourly and daily values have always been measured with the same sensor for both Météo-France and EDF stations, and these are tipping gauge sensors with a resolution of typically 0.2mm. It should be kept in mind that in the case of EDF stations only the daily values are routinely controlled for spurious values, and also that the individual measurement devices have changed over time.

Data: Section: Please provide a table showing the data used in the study providing/summarizing some meta-information (e.g. length, missing data, relocations if known, ...) and help the reader to distinguish the different data sets used in the analyses part.

Thank you for this suggestion. In answer also to Rev1 and Rev2 we completely restructured this part and removed known inconsistencies in the text and the figures. Following your suggestion, we also updated the table in the appendix of Berghald (2024, <https://dumas.ccsd.cnrs.fr/dumas-04723319v1>) and will provide it as supplementary material to the manuscript.

215: Do not mix up “weather pattern” and “circulation pattern”

ok.

216: Do not use “circulation mode over the Alps”, e.g. use “flow condition/pattern over the Alps” or “weather patterns over the Alps” (dependent on what should be explained)

ok.

223: make your grouping of WPs visible in Fig. 2 from the colors used for the time series (similar colors for same group)

In response to Rev1 we removed the right side of Fig. 2 (the timeseries). On the left side the arrows are coloured sequentially in a counter-clockwise order, and the colours used in Fig. 6 are aligned with the ones in Fig. 2.

242: I guess it is meant that “The quantiles of the distribution can now be obtained by transforming Eq. 1 or 2”

ok.

3.1.2 Non stationary GEV using a breakpoint in 1985: this is bit weak, as series could have other breakpoints as well and this was not tested in any way. So would be good to give more motivation for the selection of 1985.

In lines 267-273, we provide different reasons motivating our choice of this specific breakpoint, including the findings of previous studies (e.g. Blanchet et al., 2016), considerations regarding data availability as well as our best knowledge of changes in the instrumentation.

302: We then turn to the trends in extreme intensities. Be more precise by saying “trends in 20-years return levels”.

ok.

Throughout the paper (and particular chapter 4) authors should be precise and clear describing the synoptic drivers from weather patterns. They shouldn't use the terms Atlantic and Mediterranean circulation, which are misleading (see also my previous comment). Use the terms from the Garavaglia (which are not very well suited as well, in my opinion, but at least published). For the paper air flow against the Alps is the key interest from the weather patterns, so focus should be given to this.

Thank you for your comment. We made sure to talk about flow directions instead of circulations. We would, however, like to note that our study is to some extent a follow-up on previous studies by Blanchet et al (2021a, <https://doi.org/10.1016/j.wace.2021.100356>) and Blanchet et al (2021b, <https://doi.org/10.1088/1748-9326/abb5cd>) where the terms Atlantic and Mediterranean circulation have been employed.

320: “Mediterranean dominance” replace by something like “Mediterranean impact”

Ok.

322: Why is anticyclonic influence associated with convective events (which should be more related to gradient weak weather patterns)? In fact this goes already into the discussion part.

Thank you for this remark. We here mainly mean local convective events linked to the daily temperature cycle (typically highly localised thunderstorms at the end of a summer day) which occur primarily in anticyclonic conditions. We do, however, acknowledge that convective events are per se not limited to anticyclonic conditions, and removed the second part of this sentence ("indicating the role of convective events like thunderstorms during this season.")

334: incomplete sentence

The sentence seems complete to us.

335-336: not clear what is meant

This refers to the comparison of results between the recent dataset (since 1980s, EDF and MétéoFrance) on the one side and the long time series only (the EDF time series we base the trend analysis on) on the other side. The EDF network is more spatially confined and does not cover the areas where the estimated 20-year return values are highest, which is in the Rhône valley and along the Mediterranean coast. We clarified this in the manuscript. We rephrased this to "For comparison, we also did the analysis using only EDF stations, but with all available data (i.e., the dataset used for the trend analysis). This did not result in significantly different spatial patterns (not shown). It should, however, be noted that almost no EDF stations are located in the parts of the study area with highest 20-year return levels (Rhône valley and along the coast)."

343: ... not exact the same regions ... be more specific (it probably means that there are less stations per region included?)

See our comment above.

387: again here, authors should be more precise in the discussion. If there results are confirmed by previous studies the question arises, what is new in the presented paper. In fact they could show confirmation of their results for the daily precipitation extremes with previous studies, but not for the hourly values (as this is new, as stated in the introduction).

For the discussion in 5.2 and 5.3 it would be useful to take into account which month of the year (in which region) are particularly relevant from stakeholder's perspective, as known to have precipitation events resulting in flood events. This could maybe show that those months with significant trends are also those known as risky from the stakeholder's perspective.

Thank you for these suggestions, we included them in the discussion and in the reformulated conclusion.

395-399: speculating on using temperature as estimator for parameter distribution leads to question why not giving it a try or at least showing the temperature evolution for the study region to get some feeling if it could work.

Thank you for this remark, which was also raised by Rev2. As we explain in more detail in the reply to Rev2's comment, there are homogenised time series of annual mean temperatures for some selected reference stations available online (<https://meteofrance.com/climat/hd>). Although providing just anecdotal evidence, these time series show roughly linear trends since the mid-1980s (e.g. Grande-Chartreuse in Rhône-Alpes and Embrun in Provence-Alpes-Côte d'Azur). We included the reference in our manuscript.

440: give more details on findings from Fukutome et al. (2015) here, if this could be useful for your results.

We added the following sentence: "Panziera et al. (2018) attribute this to the role of the first Alpine slopes triggering convection in unstable airmasses arriving from the west or south, and Fukutome et al. (2015) note that there is a lack of sufficient moisture convergence in the inner-Alpine zones."

447-448: Overall, the southern part in the study region experience much stronger trends How does this fit to Figure 11 for hourly data?

As can be seen in Fig. 11, for most of the months and seasons, both positive and negative trends tend to be higher (in absolute terms) in the south compared to the north of the study region.

450: what is meant with "large scale circulations patterns"?

We replaced this by 'weather patterns' (the ones we use in this study).

454: We thus partly attribute this make clear that this is speculation, as you have no prove for really attributing it.

ok.

459: ...is never much lower than zero .. I guess it should read is close to zero

ok.

468: .. with an increase with increasing temperature Rephrase

ok.

474: Make clear that Georgi et al (2016) is result from model simulation and explain in more details, otherwise is hard to understand for the reader.

Thanks for this remark, we included this information.

477-479: again here (and mentioned earlier), would important to know the time resolution of precipitation sensors used by EDF and Meteo France

See our answer to your previous comments, we also rephrased this sentence to “However, stations maintained by EDF unfortunately do not register at such small timesteps, and observations registered by Météo-France at 6-minute intervals are still too short for a robust analysis of trends.”

498: Regarding the spatial and temporal of extremes, ... word is missing here.

‘evolution’, we added it.

499-500: more precision is needed here “hourly extremes” and add the word “precipitation” to make it easier for the reader.

In the revised manuscript, we completely reformulated the conclusion following your last comment (and suggestions by Reviewer #2).

501-502: I do not understand this sentence “While the general spatial and seasonal patterns confirm previous studies, our main contribution here lies in updating previous estimates of hourly extremes based on surface observations”. This is not what is described at the beginning of the paper.

Thank you for pointing this out, we took this into account when reformulating the conclusions.

501-505: this is not something for the conclusions → should be delete here

Thank you, we removed this.

509: trend in June needs to be introduced earlier as well

Thank you, we highlighted this at relevant points in the results (l. 382, 394) and discussion sections (l. 501, 535).

511: you need to explain the “decoupling”. How could/should this work in a physical sense?
Thank you for this remark, we took it into account when reformulating the conclusions.

Generally, the conclusions part is still a bit weak and should be improved. There is clear potential for doing so.

Thank you for your remark, we reformulated the conclusions, elaborating more on the implications of our findings.

References:

Jiang, R., X. Cui, J. Lin, and J. Tian, 2023: 40-Year Statistics of Warm-Season Extreme Hourly Precipitation over Southwest China. J. Appl. Meteor. Climatol., 62, 1891–1908, <https://doi.org/10.1175/JAMC-D-23-0018.1>.

Zeder E. and M. Fischer. 2020. Observed extreme precipitation trends and scaling in Central Europe. *Weather and Climate Extremes*, Volume 29, September 2020, 100266, <https://doi.org/10.1016/j.wace.2020.100266>

Zeder E. and M. Fischer. 2024. Decadal to centennial extreme precipitation disaster gaps — Long-term variability and implications for extreme value modelling, *Weather and Climate Extremes*, Volume 29, September 2024, 100636, <https://doi.org/10.1016/j.wace.2023.100636>