We thank the reviewers for their additional suggestions for minor revision and we answered (in blue) to each single comment (in black).

Answers to RC1:
Many thanks for the detailed revisions of the paper. The paper is ready to be accepted after very minor revisions.
We are happy to hear that the present version of the manuscript has been appreciated.

l. 27: Remove "then"
Done.
l. 43: tipically
Done.

You have stopped using "high resolution" and "very high resolution" but continued to use "HR" and "VHR". This is confusing and should be changed.
This is to maintain the reference to the model version name. This choice is in accord with a comment received in the previous round of review.

l. 116: It seems a bit strange to start a new section between the previous and the following paragraph. Maybe change the position of the section heading?
The present organization of subsections puts together model data (2.1) observational data sets (2.2) and the methodological details used for our analysis (2.3). Moving the first paragraph of 2.3 to 2.2 would result in a mix between data and method description. We would prefer to maintain the current structure.

l.181: "compared ot"
Done.

l.206: "for the models"
Done.

ll.235: Does this maybe indicate that you get the daily cycle wrong? -> Again, changes in the parametrisation may help?
Yes, it might also be related to the fact that the daily cycle is not well capture over certain regions, but we can't claim it in the current version: additional analysis would be necessary.

l.242: "worse"
Done.

l.246-247: Should be reworded.

 Rewritten as: " *In addition it is important to note that moving from the standard to the high resolution of CMCC-CM2, the model behaves consistently with the models participating to the HighResMIP project: a tendency to an increased fraction of land precipitation in the highest resolution, and the same tendency for the fraction of land precipitation caused by moisture convergence (Venniere et al. 2019)."*

This is the second time that I am reviewing this manuscript. The manuscript is improved with respect to the first version and most of my previous comments were successfully addressed (two were not entirely, see items 1 and 2 below). I also added a few new (most relatively minor) comments derived from the new material that was added (items 3-6). Finally, I would recommend the manuscript be looked over by a native English speaker, as I found many English mistakes in the document. I have noted a few at the end, but this is not an exhaustive list, and the last section was a bit difficult to understand (see item 6).

We thank the reviewer for this second round of comments. We answered to each specific comment (see below), also improving the language.

1. Figures: The authors are using a diverging palette for non diverging data (upper right in Figures 1-10. While I don't think it is such a problem for temperature in this case, for precipitation it gives the illusions of boundaries where there is none. I would suggest using a sequential scheme, as described on p.10-11 of the IPCC visual style guide.

Figures related to precipitation have been modified following this suggestion.

Also, there are a lot of figures for such a short manuscript. Since the emphasis is on the biases, may I suggest the following format to reduce their number?
Merge Fig. 1-4 such that:
ERA5 DJF tas 24h HR DJF bias 24h VHR DJF bias 24h
ERA5 DJF tas 6h HR DJF bias 6h VHR DJF bias 6h
ERA5 JJA tas 24h HR JJA bias 24h VHR JJA bias 24h
ERA5 JJA tas 6h HR JJA bias 6h VHR JJA bias 6h
and send the absolute values for the simulations (i.e. the second rows of the current figures) to
the supplementary section. Figures 5-8 could be merged in a similar way. And Figures 9-10 could probably be sent to the supplementary section given that they are so similar to Figures 7-8 and only 1 sentence is spent discussing them. If the authors would like to include a 3rd picture, they could include what is currently Figure S17.

The number of figures was lower in the first version of the manuscript but we had to modify it (leading to the current structure) to match reviewer 1 requests.

Finally, the characters for the latitudes and longitudes on certain figures seem to be overlapping
(see for example Figure 5).

Corrected.

2. Something that was pointed out by myself and the other reviewer was the use of high and very high resolution when referring to the two configurations. In their response to the reviewers, the authors mention "In the new version of the manuscript we use the terms standard and high instead of high and very-high". However, I found many instances where the authors refer to the 1deg resolution simulation as high resolution (HR) and the 0.25deg resolution as very high-resolution (VHR). e.g. lines 34-35, 41, 84, 90, etc.

We adopted "standard" and "high" as suggested, but we preferred to maintain the original acronyms (HR and VHR) for consistency with the model name in the CMIP6 repository, since

there is also a SR version provided by CMCC, but not used in the present study. This choice is in accord with a comment by reviewer 1 received in the previous round of reviews.

3. Lines 52-56:" Regarding the extreme precipitation representation, based on simulations from single GCM, some improvement in skill at higher resolution for some measures of extreme precipitation over certain regions of the globe have been found in the past (Wehner et al. 2014, Kopparla et al. 2013) and only recently, multi-model assessment on this topic have been done, confirming that increasing the horizontal resolution to ¼ of degree (the highest adopted by the model object of this study), the magnitude of simulated daily (Bador et al. 2020) and sub-daily precipitation (Wehner et al. 2021) extremes is increased."
I feel this ignores all the work done on this topic in the RCM community, and it somewhat contradicts what is written on line 40: ("high resolution models, when implemented with a resolution similar to VHR, achieve skills comparable to state-of-the-art Regional Climate Models in reproducing precipitation distributions"). I would suggest reformulating.
We now specify in the second sentence that this statement (line 52-56) is related to GCMs: "Regarding the extreme precipitation representation in GCMs"

4. Line 151: "On the other hand, the negative JJA bias of about -8C over north-eastern Canada shown by the HR model is even worse in the VHR version, where a larger portion of the domain is subject to a bias of about 10C."
This negative bias sticks out like a sore thumb. Do the authors have an explanation for this very large negative bias? Is it linked to an excess of sea ice in the summer?
We don't have a definitive explanation for this bias, but we confirm that the two CMCC-CM2 model versions object of this study tend to overestimate the sea ice over the Northern Hemisphere during JJA. We added a comment on this in the new version:
"This negative bias is also consistent with the tendency of the two versions of the CMCC-CM2 model to overestimate the sea ice cover during summer over the Northern Hemisphere (not shown)"

5. I recommend computing the mean bias for the various variables and the different configurations. That value could be inserted directly in the corner of the relevant figure. That would help supporting statements such as:
"In terms of average precipitation, the VHR model shows less pronounced biases with respect to the HR model…(line 171) "on average, the highest resolution CMCC model is better than the lower resolution model in representing…" (line 213)
Done: In the new version of the manuscript, within the figure panels referring to the biases, there is also the indication of the average value.

6. I have to confess that I had a difficult time understanding the conclusion section, from line 225 onward. Part of it might be due to the English I think (e.g. sentence on lines 238-242 is too long; the following sentence is missing a verb), but some of the sentences in that section that were added as answers to the reviewers comments are not really well integrated in the text, which makes the message a little confusing. I am afraid I don't have a good suggestion here, other than spending a bit of time to make sure that the text flows a little better and making sure the main conclusions/ideas are clearly put forth.
Long sentences have been divided and the section has been then improved as follow (see also the provided version of the manuscript with "track-change on"):

"Regarding the precipitation distribution, the VHR model performs better in representing averages and intense events, but more pronounced biases appear in VHR compared to HR when focusing on extreme events, with a more evident degradation in the daily statistics compared to the 6-hourly. This latter result reduces the confidence we usually attribute to the highest horizontal resolution in modelling extreme precipitation, and is consistent with single model analysis based on CAM5.1 atmospheric model (Wehner et al. 2014) suggesting a positive bias over most of the globe in the representation of extreme events at ¼ degree horizontal resolution. This is also in agreement with recent findings (Bador et al. 2020) suggesting that highest resolution models tend to produce more pronounced extremes than lower resolution ones. In addition many of them show lower skill in representing observed patterns, both in terms of intensity and spatial distribution, at the higher resolution, compared to their corresponding lower resolution version.

This emphasizes the need to focus not only on the horizontal resolution to improve the model ability in representing the climate system, but also on physics and tuning. It is important to note that in the model object of this analysis the tuning parameters were kept constant, moving from the HR to the VHR version, in order to be compliant with the HigResMIP protocol. The different biases, obtained based on daily and 6-hourly time frequencies, also suggest that for the setup of model physics and tuning we need to consider the event distributions at different time frequencies, to take into account the representation of the different processes responsible of the extreme conditions emerging at the different frequencies (Scoccimarro et al. 2015).

The poor performance of climate models in representing extreme precipitation is not improved in the last CMIP6 generation models, compared to the previous CMIP5 generation (Scoccimarro et al. 2020). In the present work we have shown that this lack is even more evident moving to the highest resolution version of the CMCC-CM2 model adopted for HighResMIP, consistently with multi-model analysis performed at the same horizontal resolution (Bador et al. 2020): GCMs whose parameterizations were not retuned at higher resolution lead to worse results. The high-resolution version of the model tends to overestimate extreme precipitation in the wet and warm regions, consistently with findings based on experiments carried out with the CAM5 atmospheric model at the same resolutions (Wehner et al, 2014), highlighting once again the importance of an extensive model tuning at the high resolution. In addition it is important to note that moving from the standard to the high resolution of CMCC-CM2, the model behaves consistently with the models participating to the HighResMIP project: a tendency to an increased fraction of land precipitation in the highest resolution, and the same tendency for the fraction of land precipitation caused by moisture convergence (Venniere et al. 2019). Also, in CMCC-CM2 model, the orographic precipitation captures most of the change of precipitation due to resolution, consistently with most of HighResMIP models (Venniere et al. 2019)."

Minor points
Line 40: "Demory et al. (2020) have shown that high-resolution models"
Done.
Line 43: typically, not tipically
Done.
Line 45-47: "Regarding the extreme temperature representation, based on data at the daily

frequency, it has been shown that GCMs tend to have warm bias over most land areas (Li et al.,2021) and the horizontal resolution plays a minor role with respect to the one played in the

extreme precipitation representation"

1) I would suggest rephrasing this sentence, as the wording is a bit awkward.

We rephrase the sentence as follow:

"Regarding the <u>extremely high</u> temperature representation, based on data at the daily frequency, it has been shown that GCMs tend to have warm bias over most land areas (Li et al., 2021) and the horizontal resolution plays a minor role <u>in affecting the bias</u>, with respect to the one played in the extreme precipitation representation (Kharin et al. 2013, Wei et al. 2019)."

2) Do you mean to say that models overestimate both warm and cold extremes?

I meant warm extremes, now specified.

3) Play a minor role in what?

The horizontal resolution plays a minor role <u>in affecting the bias</u>, as now specified.

Lines 47-51: "Typically, the warm extremes are computed based on maximum daily temperature, but in this work we want to verify the potential improvements induced by the increased resolution in the representation of extreme temperature events defined at two different time frequency (daily and 6-hourly). For this reason we investigate the distribution of daily and 6-houry average temperature, instead of maximum daily temperature."

This should be moved to the methodology section.

Done.

Line 53: "...based on simulations from a single GCM..."

Done.

Line 82: "The two models object of this study differ only..."

Sorry, I don't see the typo. The sentence is as you suggest in the manuscript too.

Line 95: "The model performance in representing the temperature distribution is evaluated by comparing results to..."

Done.

Line 109: "Since we aim to characterize different types of extreme events..."

Done.

Line 119:" This time period is sufficiently long to capture..."

Done.

Line 122: "The grid differences are minor and therefore the interpolation introduces very little differences in the fields."

I don't know if I agree with this statement. It is true that the difference is small between the VHR resolution and ERA 5, but CHIRPS resolution is 0.05deg while the atmospheric models used here have a 0.25 deg and 1 deg resolution.

We agree. This sentence was acceptable in the first version of the manuscript where we didn't use MSWEP 0.1degree data. Anyway, the interpolation is done only to compute biases on the

final fields, and not before the percentile analysis as stated: comparable results are obtained using bicubic interpolation. This is now added to the text as:
"The kind of interpolation introduces very little differences in the fields (not shown)."

Line 155: "The positive bias over the north western part of South America…"
Done.

Line 184: "suggesting that the worst VHR extreme precipitation representing during DJF is mainly due to a too much pronounced stretching of the right part of the precipitation distribution only"
Please rephrase.
Rewritten as:
"… suggesting that the bad representation of DJF extreme precipitation in VHR (Figure 5) is mainly due to a too much pronounced stretching of the right part of the precipitation distribution."

Line 213: "Anyway, on average…"
Done.

Line 227: "...and also with multi-model recent findings, suggesting that higher resolution models…"
I don't quite understand what the authors mean with "and also with multi-model recent findings".
It was just to highlight the multi-model nature of the cited work. Anyway, we use "recent findings" instead of "multi-model findings" in the new version of the manuscript.