The authors run two WRF models, a 5-km centered on Italy and a 15-km centered on Europe. The nested models are forced with ERA5 boundary conditions for an historical run and with 3 SSP scenarios using the MIP-ESM model.

REPLY:
Thank you for your valuable feedback on our manuscript, which gives us the opportunity to significantly improve our work and clarify a few critical points.
In fact, we are under the impression that some comments, and specifically those questioning the relevance of our paper to the journal's scope and focus and the overall consistency of our approach, might stem from an over-compressed description of the simulation protocol, which can lead to misunderstanding the hierarchical steps of the adopted double-nesting procedure.

Similarly, we are concerned that the main point of the paper as stated in the abstract and the introduction, i.e. the analysis of a climate simulation run at a grid-step that falls within the so called gray zone for the convection representation (4-10 km), as well as its bearings on the optimal design of numerical simulations, was not fully appreciated. Nevertheless, we acknowledge that our line of reasoning could have been better highlighted along the article, and will proceed to amend the text accordingly.

They describe this a CMIP6-driven, but really it is only a single model. For that title I would expect multiple CMIP6 models to be used. Please remove CMIP6 from the title.

1.REPLY:
We acknowledge that the current title might suggest expectations greater than those we are dealing with. However, the results shown in this paper are among the very first obtained when downscaling a CMIP6 global model and this is a distinctive feature of our work. We propose to change the title with "Impact of spatial resolution on multi-scenario WRF-ARW simulations driven by the CMIP6 MPI-ESM1-2-HR global model: a focus on precipitation distribution over Italy".

Basically, this is a paper not about model development, but about the simple application of an off-the-shelf WRF model run at two resolutions.

2. REPLY:
We trust that once clarified what double-nesting is, some of the subsequent criticisms of the reviewer will be answered as well, and the main results of the paper might be fully understood.
In particular, we will expand section 2 in order to include a more thorough description of the numerical protocol.
In order to ease the current discussion, we anticipate here a brief illustration of how this protocol generally works:

1. *a parent simulation at a relatively coarse resolution (i.e. 15km) is run as a first regional downscaling on a domain much larger than the final domain of interest, whose boundary and initial conditions (BC and IC) are provided by a global model (reanalysis or climate scenario simulations),*
2. *a second child simulation (i.e. 5km) is nested into the first, taking its boundary conditions from the parent simulation (i.e. 15km) and covering the inner domain of interest at higher resolution.*

This approach allows to perform a progressive downscaling on a given region, from the original resolution of a global model (order of 100 km) to the finer resolutions recommended for regional applications and impact studies. The use of increasingly high resolutions allows the system to develop its own dynamics at the intermediate scales that will feed the high-resolution model , as well as to avoid numerical instabilities and unreliable results.

The authors spend much time in the abstract (where it is inappropriate) and in the introduction describing convection permitting (CP) models that need scale of <4 km. But then, they use models that are not at CP scales, so why waste this space which should be better used describing what they did here.

3. REPLY:
*As one of the main results of the model configuration investigated is a model behavior very close to a convection-permitting configuration, we dedicated in depth discussion to the description of models at such resolutions, as background to show the advantages of our configuration at its highest resolution. As also reported in **REPLY#14** we have demonstrated that the model at 5km in the current configuration behaves as a CP one.*

There is also a lot of talk about ECVs, but the comparisons are only to T2m and Precip., so why open discussion on ECVs here?

4. REPLY:
We agree with the reviewer about mentioning ECVs, having dealt with only two of them. We'll remove ECVs from the text.

The two model simulations WRF-15km and WRF-5km are compared against observations (with the ERA5 forcing) but there is less comparison and emphasis on the overlap region, the

only region of interest here since the case being made is what is the difference between 5km and 15km. In fact the experimental design is poorly thought out, Both 5km and 15km should have been run over the same domain, because making the 15 km domain so much larger changes the way the ALPS domain sees the boundaries. Better to run both for the same Italian region

5. REPLY:
*See **REPLY#2** for a more detailed description of double nesting technique that should definitely clarify that it is not possible to run "5km and 15km over the same domain" as the 15 km run should provide BC to the inner 5 km domain and its continental size is thought in favour of satisfying reproduction of large scale condition over the region. The experiment is not merely an exercise of comparison of two different resolutions, but, as specified, a nesting strategy that then also allows to evaluate the benefits of high resolution that is desirable in a morphologically complex region such as the Mediterranean, where the benefits of moving toward the km-scale are largely demonstrated in literature.*

L34 – "demands for high-res" is confusing use of English

6. REPLY:
Thanks for the comment: we will remove "for"

L35 – Really, so many regions over the globe are "critically prone to the impacts of local-scale and severe weather" Certainly the mountains are difficult to model, but not unique in tough weather.

7. REPLY:
As stated in the title of the article we are not studying any other region in the world but the Mediterranean, and how its climate has a negative impact on local *communities. (See references)*

L39 – No one is seriously running 200 km global climate models these days!

8. REPLY:
We'll better highlight in the revised text that several CMIP6 experiments run at this resolution (See Klaver et al., 2020).

L43-– what does "(120 divided by 20 km) mean? 6 km?

9.REPLY:
It means in the range between 120 km and 20 km, we will exchange the symbol for the periphrasis

L66- CP is not defined except in the abstract (does not count as text)

10.REPLY:
We will move the definition of the acronym CP from the abstract to the text

L73- "model uncertainties" or model errors?

11.REPLY:
In this context we reported what stated in Fosser et al 2024
In the revised text we'll better remark the proposed reference

These are some examples of why this paper is difficult to read and interpret.

12.REPLY:
We'll improve the text readability accordingly to the previous replies

Fig.5 – the description of this figure in the paper is contrary to what I see in it. The summertime dip in the 5km model is described as producing a much better seasonal fit than the 15km one. I see that the summer (months 6-7-8) in 5km is better, but for the rest of the year, it is still just like the 15km and overestimates the rain. "more realistic" is a marginal call. " better in summer" is OK

13. REPLY:
The description of Figure 5 states that the 5 km simulation reduces the summer bias from April to November. We will better explain Figure 5 and avoid to say "more realistic"

Figure 6 is even more of a shock. The 5km run has NO cumulus rain, only annually uniform stratiform rain. The 15km run develops cumulus in summer, but the total (Fig 5) is far too large vs. observed. This indicates the your scale-aware convection is totally failing: as we go form 15km to 5km your cumulus convection parameterization shuts down.

14.REPLY:
Figure 6 is one of the main points of the paper. In the revised text we'll remark that it demonstrates that the scale-aware parameterization is correctly working, at least for the experimental design here proposed. The 5 km model is run in the gray zone, this means that we did not switch off the convection parameterization (as it happens in CP models). Nevertheless, the G-F parameterization, due to the scale-awareness, dynamically manages the cumulus activity based on vertical mass-flux within a grid cell. Actually, the model behaves similarly to a CP one and benefits of an improved representation of the precipitation, at the same time saving computational time and data storage. We will make an effort to state more clearly the point in the text, as it seems prone to misunderstandings. The use of the term "stratiform" is actually misleading, as it's referred to the precipitation comes from the microphysics, due to both large and convective scale processes. The term "cumulus" here is referred to the parametrized contribution to the precipitation (i.e. it might be zero, or very small, if G-F is "switched off" and the convection is explicitly resolved).

Figure 8 – The identical power spectrum for 5km and 15km tells me that both are doing the same thing because both have the same parameterization??

15.REPLY:
In Fig.8 is reported the PDF of daily precipitation. However, it is unusual to have similar behavior for 5 km and 15 km. We already notice this point and it deserves further investigation. We will state more clearly this limit in the text, additional explanation and open questions will be added.

Figures 12&13: If you want us to learn anything from this experiment, both should be plotted on the same grid so we can compare. As is, there is no useful information here.

16. REPLY:
These figures actually refer to different variables for the same experiment. We suppose that the reviewer meant fig 12 and 14. In any case the set of figures from 12 to 16 display the climate signal for T and P for the parent (15 KM) and the child experiment (5 KM), respectively. Given the above explanation about double nesting, we think that it is clear why we do not plot them on the same grid. The useful information is that we are undoubtedly facing a warming and alterations of the precipitation patterns for the end of the century. In particular, the signal in precipitation for 5 km experiment is generally more statistically robust.

L381 – Conclusions:

I would say this is a downscaling example, but certainly not a strategy.

We have seen no evidence the your G-F cumulus scheme mimics a CP scheme.

In fact it seems to shut down.

The 5km runs improved a wet bias ONLY in summer, and they did that oddly (above)

How does this work show that your 5km or 15km runs are "adequate to provide the boundary conditions for" CP scale and further downscaling. This is simply not shown.

17. REPLY:
We think that the above replies should have shed a light also on the conclusions and clarified the comments made by the reviewer.

I am confused as to why this paper is a GMD paper, there is really no model development.

18. REPLY:
We thank the reviewer for the comment which helps us clarify the scope of the paper. This paper deals with the description of a regional climate downscaling experiment. The development is not to be intended in model physics but in the setup. In particular, we recall here that one of the most important novelties of the paper is the use of CMIP6 forcing.

An odd question: since you are forcing at the somewhat distant boundaries, do you not have synoptic climate variability in the interior, especially with regard to convection? Do you force internally? Therefore, would different ensembles produce different statistics?

19. REPLY:
We think that the above replies should have shed a light also on this aspect.

The code and data availability is very weak. Is the exact WRF version used here the "current version" with the github reference? Were any mods made? I agree it Is not the responsibility of these GMD authors to archive the ERA5 data. But where is the reference to the MPI-ESM data? (also on another archive). What is missing here is the actual values from the figures posted here. That is a useful minimum requirement. What code was used to process the data sets?

20. REPLY:
We'll modify these aspects in the revised text accordingly to the GMD policies . We anticipate that the MPI-ESM, similarly to the other CMIP6 simulations, are freely available on the Earth System Grid Federation (ESGF).