Comment on egusphere-2022-215
Clemens Schwingshackl (Referee)

This paper analyses the historical occurrence and future projections of heat waves in Europe using simulations from four RCMs driven a) by ERA-Interim and b) by three different GCMs. Heat waves are identified by the Heat Wave Magnitude Index daily (HWMId). The paper finds that RCMs generally reproduce well observed heat wave patterns when driven by ERA-Interim. When using historical simulations driven by GCMs, the (statistical) agreement with observations gets a bit weaker. Future projections reveal increasing heat wave magnitudes throughout Europe but with differing patterns for GCMs and RCMs. This implies that RCM patterns are not only determined by the driving GCM model, but the physical parameterization of RCMs plays an important role as well. Uncertainties of future projections are due to both GCMs and RCMs.

The paper is well structured and generally it is easy to follow the argumentation. Some parts would need a more concise language to be better understandable. The paper mostly focuses on analysing the contributions of RCMs and GCMs to the HWMId patterns, but also includes some statements about impacts. For the latter, I would suggest using a different approach (e.g. focusing on warming levels instead of time periods) due to the usage of the high-emissions scenario RCP8.5, as this scenario does not include any climate mitigation efforts and thus likely overestimated impacts at the end of the century.

General remarks:
Some parts of the result section would benefit from a revision of the language to make the paper easier understandable. Arguments are sometimes rather hard to follow, especially in the parts focusing on the analysis to which extent RCMs and GCMs are responsible for a certain HWMId pattern (e.g. lines 170-194). I would recommend formulating the text in a more concise way and try to keep the same terms throughout the paper to facilitate reading.

In the results section, the coverage of certain topics is sometimes split (e.g. a table relating to a certain figure is described at a different place than the figure). I would suggest restructuring the results such that each topic is only described/discussed once (i.e. table and figure at the same time). This would make it easier for the reader to follow the argumentation. In the specific remarks I mention some examples for text that could be merged.

The paper mentions several times that HWMId rises in an “exponential-like” rate. However, the paper provides no figure or analysis that shows such an exponential increase. I would thus recommend to include a time series plot showing this behaviour. Further, I am not sure if “exponential-like” is the right term: Either the increase is indeed exponentially, or it follows a different functional form.

The study currently uses a matrix of 4 RCMs x 3 GCMs. To my knowledge, even a larger matrix of GCMs and RCMs with full RCM/GCM coverage would be available. Is there any reason why only 4x3 models were used? Have the models been specifically selected?

Specific remarks:

There are some typos and small mistakes in the manuscript that I do not list here. I would recommend to carefully read the manuscript again and correct them.

- Line 5: I would specify the observations (i.e. E-OBS)
- Line 6: Higher resolution compared to what?
- Line 8: What does “satisfactory way” mean? How is this determined?
- Abstract: I think it would be nice to finish the abstract with one or two sentences describing the implications of this study.
- Line 22: Weren’t also other Scandinavian countries in addition to Sweden affected by the 2018 heatwave?
- Lines 23-24: Heatwaves can also affect agricultural productions or cause forest dieback (due to lack of water or insects). Might be worth to mention this here as well.
- Lines 54-59: I think some more argumentation would be good to explain why HWMId is used (instead of other heat wave indicators)?
- Lines 76-77: From the original publication of HWMId (Russo et al., 2015) I understand that Tmax,ref,25p and Tmax,ref,75p are calculated based on the annual maximum temperature distributions at a certain gridpoint (i.e., 30 values per grid point). It seems that in your study, the total summer distribution was used. This should be checked to make sure the calculation agrees with Russo et al.
- Line 80: Remove “is calculated” at end of sentence.
- Line 88: Can you briefly mention why focusing only on RCP8.5?
- Line 101-103: This sentence regarding the different reference periods sounds a bit complicated. As I understand, when RCMs are driven by ERA-Interim the reference period is 1989-2008, and when driven by GCMs it is 1990-2010. Maybe it could be explained like this?
Line 106: At which step were the data remapped? Before or after calculating HWMId? And were the data remapped to the ERA-Interim or E-OBS grid? And what about remapping of ERA-Interim or E-OBS?

Line 110: I am not sure “effective precipitation” is the right term here, as it would also include runoff. Maybe just use P-E? Alternatively, sometimes “net surface water budget” is used.

Line 132: Which RCM simulations?

Lines 139-140: Would be better to include this in the paragraph of lines 117-127.

Lines 143-144: I guess this is due to the fact that RCMs are driven by ERA-Interim and not E-OBS.

Lines 147-151: I would remove this here and only mention it when the respective figures or tables are discussed (see general remark).

Lines 155-156: The improvement seems to depend strongly on the model. E.g. RCA4 has a relatively weak pattern correlation in Table 4.

Lines 163-166: Combine with lines 152-156.

Lines 172: I would not necessarily call this “error”. Maybe the term “error” could even be removed here.

Lines 174-175: I do not fully understand this. Could this be rephrased to make it better understandable?

Line 193: “are not” instead of “would not be”

Line 199: Where does the manuscript contain the information that there is no difference according to the spatial r?

Lines 201-202: Something is missing in this sentence after “the driving”.

Line 215: “Observed” refers to E-OBS, right? If yes, I think it is best to mention it explicitly.

Line 222: What is meant by “simulations”?

Line 243: Better performance in terms of what?

Lines 249-255: In your study, HWMId is already studied in detail, so I am wondering if you could extend the analysis a bit more to the mentioned events. And what would be the benefit of such a detailed analysis (keeping in mind the other CORDEX evaluations that have been carried out already)?

Line 269: How exactly does this study show added value? What does this refer to exactly?

Line 273: How does this statement relate to line 263?

Line 275: Again, what is the added value referred to here?

Line 286: RCP8.5 is a high-emission scenario and thus, it is unlikely that the future climate will be as projected by the scenario. Thus, statements about the impacts should be made carefully when using RCP8.5. One option would be to use warming levels instead of time periods, if statements about impacts are made.

Line 288: I would remove “with a strong probability”. Also, what is meant by “on the alarm on”?

Lines 292-294 & lines 301-302: I think that missing plant-physiological effects in RCMs might also contribute to the difference between RCMs and GCMs (Schwingshackl et al., 2019; https://doi.org/10.1088/1748-9326/ab4949). I am aware that the suggestion to include more papers is always delicate (in particular, if the paper is written by the reviewer). My future review will not depend on the inclusion of this paper.

Lines 310-328: This paragraph seems rather speculative to me, as it does not include any clear causal links, but remains mostly on comparing patterns. I would suggest to either extend this analysis of potential drivers or to shorten it.

Lines 327-329: This sentence seems rather vague. I think it might be better to highlight which open questions arising from your study would be worth to be analysed by future studies.

Line 335: I am not sure I would expect an exact match with ERA-Interim, given that it only provides the boundary conditions.

Line 341: The exponential-like increase is not shown in the paper. I would suggest to include a figure showing it (see general remark).
- Line 343: What exactly does “relatively more moderate rising trend” mean?
- Lines 346-347: As mentioned above, I am not really convinced by the analysis regarding the impact of drying trends on HWMI\text{d}. Thus, this statement currently seems not convincing to me.
- Line 348: Which figure does this refer to?

Remarks about figures and tables:

- Figure 2: I think a time series plot for the different datasets would be easier to understand than the current Figure 2a. Moreover, I am not sure the distributions in 2b are really needed. I personally find them hard to interpret. Another option would be to replace the violin plots by PDFs.
- Figure 3: The red rectangle is hard to see. A different colour (e.g. blue) might be better.
- Table 2: If I understand correctly, this table refers to Figure 1. I would try to highlight this better because it is not instantly clear to me (same for other tables that are connected to certain figures).
- Figure 4 ff: I find the greyish colormap rather hard to read. Maybe choose a different one? Or adjust the limits?
- Figure 10: I think this figure is too busy to understand. And the skewness is hardly visible to me from the violin plots. As above, I would suggest showing PDFs instead of violin plots as they are probably easier to interpret, and potentially also make the skewness visible more clearly.