Revision of wcd-2021-40: "A dynamical adjustment perspective on extreme event attribution" by Laurent Terray

Response to reviewers

I would like to thank the two reviewers for their detailed reading, constructive comments and suggestions as well as their appreciation of the manuscript. In response to the main issues raised by the reviewers, I have made the following main changes:

- As suggested by reviewer 2, an uncertainty range of the dynamical component, based on bootstrapping, has been added to Figs. 1 and 3. This has led to a more robust assessment of the internal residual significance. The homogeneous distribution of analogues over time has also been verified (new Appendix B, with one figure).
- As asked by both reviewers, a detailed discussion of some of the causal factors behind the results shown in Figs. 5 and 6 has been added to section 4. The influence of the AMV is now being discussed and illustrated with Fig. 7 (see also below).
- The robustness of raw TX and TN trends (with a specific focus on the TNx trend displayed in Fig. 5) has been assessed with another dataset (HadEX3, Dunn et al. 2020). An additional figure (Fig. 7) has been added to support the analysis and discussion as well as a new Appendix, Appendix C, with one figure.

I have also taken into account all minor comments and typos. Note that I have also homogenized as much as possible the size and aspect of the label bar in figures 1, 3, 5, 6, 7.

In the following, the reviewer comments appear in black with the author responses in blue. All the references mentioned in the text are given at the end. The indicated lines refer to the tracked change manuscript.

Reply to reviewer 1 (K. Wehrli):

General Comments

The manuscript is well-written, follows a clear narrative and the conclusions are supported by the analysis and literature cited in the paper. I especially appreciate the comprehensive literature review and that the observation-based results are frequently compared to results from modelling studies and vice versa. The study complements existing work in the field and provides added value to the understanding of past extreme events and long-term changes in extreme indicators. It shows that the dynamical adjustment method is a practical and versatile approach that can even be considered for rapid attribution of extreme events.

I thank the referee for her detailed lecture, helpful comments and suggestions, and appreciation of the manuscript.

I was wondering whether there is similar year-to-year variability in the contribution of the dynamic component to changes in TXx and TNx as there is on a daily basis for the specific events. Below I also added a comment about this point. Maybe the author has already done some analysis in this direction that he can share.

This is an interesting question. The interannual variability of the TXx and TNx changes induced by the dynamical component is substantial and comparable with the daily variability of dynamically-induced changes during extreme events. For instance, Figure R1 shown below suggests that the interannual variability of the dynamic component contribution to TXx changes over western Russia has a magnitude on the order of 2-3 °C, with larger values over the northern and central part of the domain. It is noteworthy that the interannual variability of TXx residual (thermodynamic) changes is slightly larger than that of the TXx dynamically-induced changes. The results are similar for TNx, albeit with smaller amplitudes (not shown). Figure R2 (right panel) shows the daily variability of the dynamically-induced TX changes during the 2010 Russian heatwave (from July 15 to August 14). The daily variability pattern shows very similar amplitudes to the interannual variability one shown in Figure R1c.

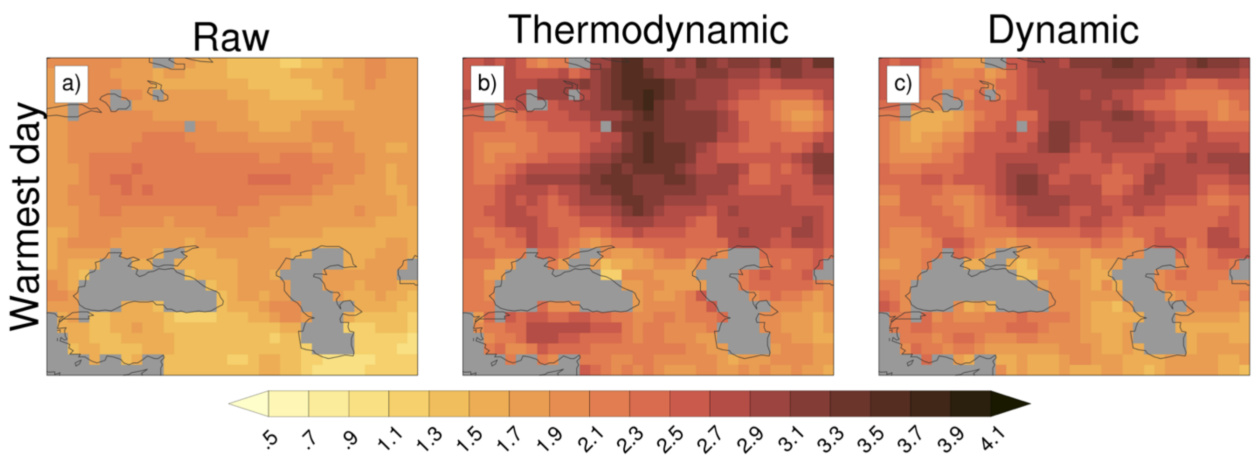

Figure R1: Summer interannual standard deviation over the 1979–2018 period: (a) Raw TXx (b) Thermodynamic contribution to TXx changes (c) Dynamic contribution to TXx changes

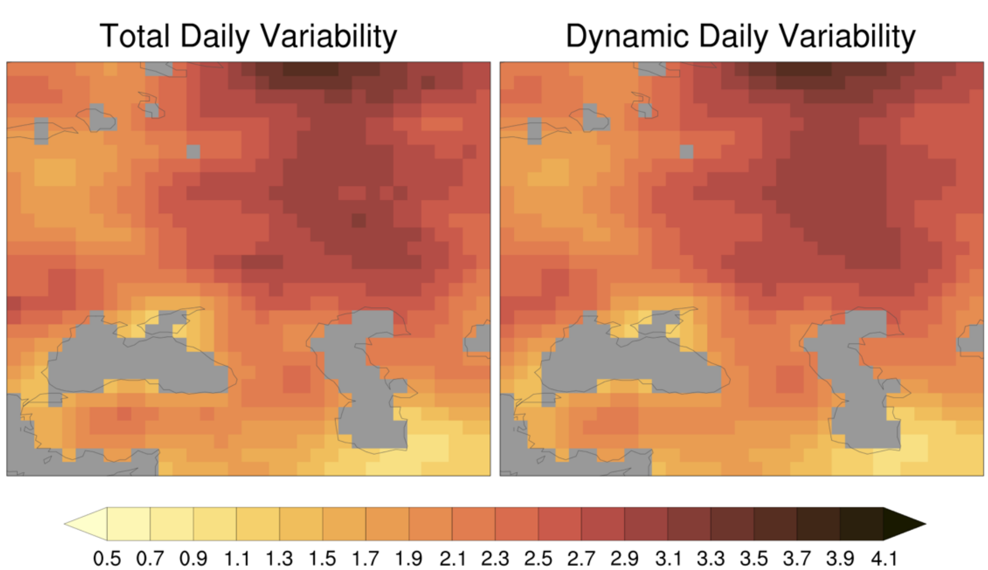

Figure R2: Daily standard deviation during the 2010 summer Russian heatwave: (a) Raw TX (b) Dynamic contribution to daily TX changes

Minor comments

l.37: The second approach of extreme event attribution is introduced as the “process-based or storyline approach”. To me, the term «storyline» might need a little more explanation in this context as I think I would not name every process-based study a storyline. Both are not probabilistic and aim to understand the driving factors. However, following e.g. Shepherd et al. 2018, storylines also explore different plausible climates, which is not done in Wehrli et al. 2019 or the present study. I am aware that it is probably not possible to make a clear distinction between a process-based or storyline study in every case. It might help if the author could share his definition of a storyline.

Thank you for the comment. I do agree that the two expressions (process-based and storyline) are not strictly equivalent. Based on Shepherd et al. (2018), the storyline approach can be defined as a physically self-consistent unfolding of past events, or of plausible future events. For example, they write: “However, we also include past events, because historical events are not simply single data points but involve detailed stories which can be unpacked”.
I would suggest that both the Terray and the Wehrli et al. papers are exactly doing that, the “unpacking” of past events. Perhaps, one can propose that the process-based approach is a sub-category of the broader storyline approach whose main focus is on past events. In a sense, one could say that the process-based approach also explores plausible **past** climates. For instance, from the Wehrli et al. results, it is possible to make some plausible inference on the Russian heatwave characteristics, had the soil moisture been at its climatological state (instead of being in a dry state).

Manuscript change:
I propose to leave only “process-based approach” on line 37 and to add the following sentence at the end of the paragraph: “Note that the process-based approach can also be viewed as a sub-category of the storyline approach that focuses on the key drivers and physically plausible unpacking of past events (Shepherd et al., 2018).”

l.60: I would leave away the word “tropical” as the heatwaves examined in Wehrli et al. 2019 were not really tropical events even if parts of the regions examined for Australia and South Africa can be classified as tropical/sub-tropical climate.

Yes, agreed.
Manuscript change:
“tropical” has been replaced by “subtropical”

l.98: It would be nice to briefly mention what method was used in Horton et al. 2015 as the study will be referenced also later in the manuscript.

Yes, agreed.
Manuscript change:
I have added this at the beginning of the sentence: “Based on a trend analysis of atmospheric circulation patterns derived from a self-organizing map clustering approach, Horton et al. ….

l.278: Do you mean “more persistent” instead of “intense”? The first spell in mid-December looks more intense to me than the second.

Yes, thank you.
Manuscript change:
“intense” has been replaced by “persistent”.

l.397: Could this trend in TX due to dynamics be related to one extreme event in the recent years such as the Russian heatwave? Would the maps look different if you left out 2010? And is there a lot of year-to-year variability in the contribution by dynamics?

The TXx and TNx trend patterns with year 2010 left out are shown in Figure R3. The patterns are very similar to the ones shown in the paper (Figure 5). This suggests that the trends do not significantly depend on the presence of 2010. As shown in Figure R1 and discussed above, the year-to-year variability in the dynamical contribution of TXx changes is quite substantial.

Manuscript change:
A short discussion regarding these two points has been added to section 4.

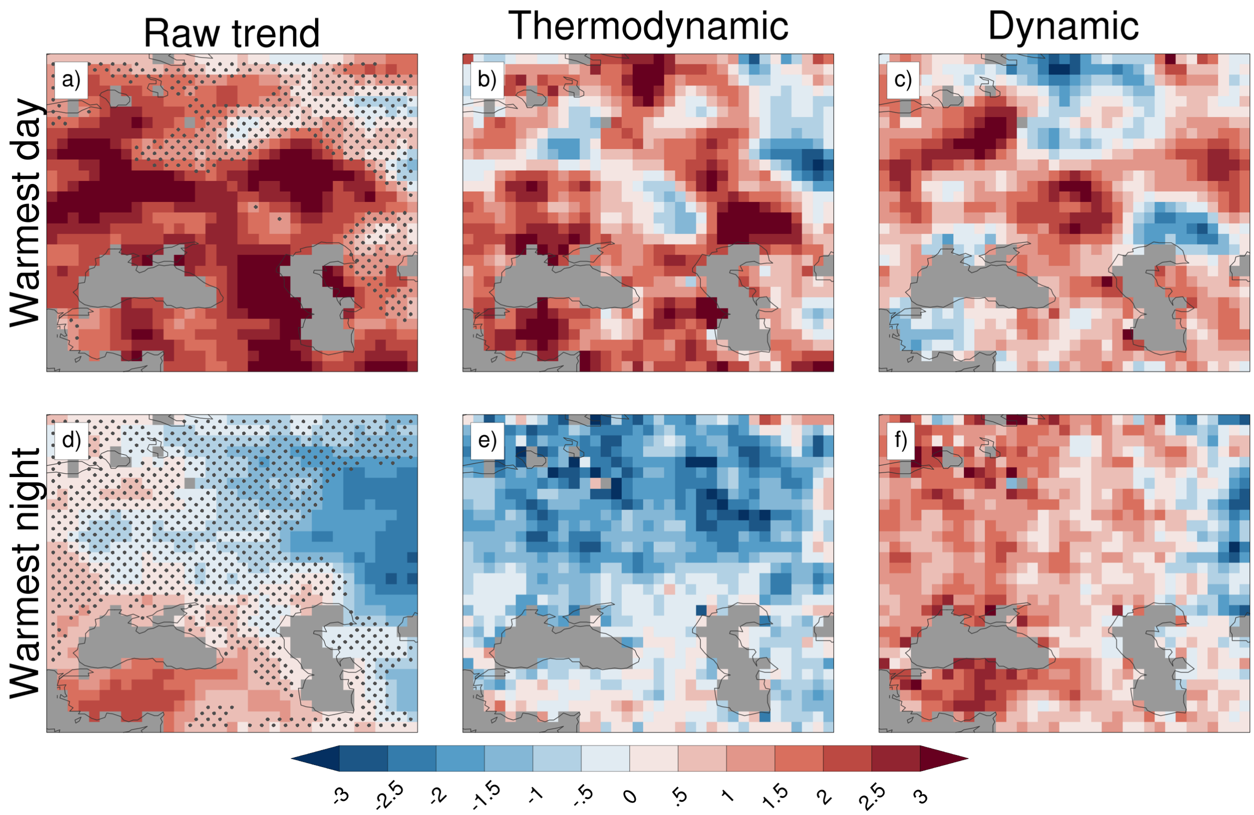

Figure R3: Similar as Figure 5 of the paper but without year 2010

l.401: Do you have a hypothesis why the thermodynamic component trend is overall cooling TNx? Would you expect this to be due to the local or the remote contribution associated with internal variability?

Thank you for the question. First, I was wondering whether the sign (and amplitude) of the TNx raw trend observed in the eastern part of the domain was realistic. Therefore, I have used the HadEX3 dataset (gridded values of both TXx and TNx, Dunn et al. 2020) to check the BERK results. Figure R4 below shows the TXx and TNx 1979-2018 trends based on HadEX3 (***note that I have slightly extended the domain eastward in Figure R4 compared with Figure 5 in order to make the point below***). Figure R4 does show a reasonably similar pattern to that of BERK for the TXx trend pattern. However, the HadEX3 pattern is different from the BERK one for TNx, with weak warming in the center and eastern part of the domain shown in Figure 5 (or Figure R3). A slight cooling is observed east of the region shown in Figure 5 (or R3). In addition, the HadEX3 TXx and TNx trend patterns both show cooling east of 60°E in the western Siberia plains (with a larger cooling for TXx than TNx). This is in good agreement with recent findings about observed mean temperature changes in this region that have resulted from increased winter and spring precipitation in recent decades, delayed snowmelt and increased soil moisture in summer as well as land-atmosphere interaction (see Sato and Nakamura 2019, Bulygina et al. 2009, Bulygina et al. 2011, Guo et al. 2019).

Therefore, I speculate that the westward-shifted cooling seen in the BERK TNx trend pattern might result from the scarcity of stations in this region for the BERK TN dataset and the use of a large influence ratio in the infilling procedure. See also response to comment N°6 by reviewer 2.

Manuscript change:
I have split Section 4 into 2 sub-sections: 4.1 that discusses the trend results based on BERK as well as observational uncertainty based on HadEX3 results, and 4.2 that details the causal factors of the extreme temperature changes over the 1979–2018 period (focusing on the AMV).

Specifically, I have added a detailed discussion (lines 467–476) in section 4.1 about TXx and TNx trend patterns and observational uncertainty based on a new Figure, Fig. 7, that shows the TX and TN trend patterns based on HadEX3 for both WE and WA regions (an extended version of Fig. R6). Note that Fig. 7 also includes a panel showing a composite-based estimate of the AMV influence on both TXx and TNx changes that is now used in Section 4.2 for the discussion of causal factors responsible for extreme temperature changes. A new Appendix, Appendix C, with one figure (Fig. C1 in the manuscript) has also been added to support the discussion about the TNx cooling and observational uncertainty in section 4.1.

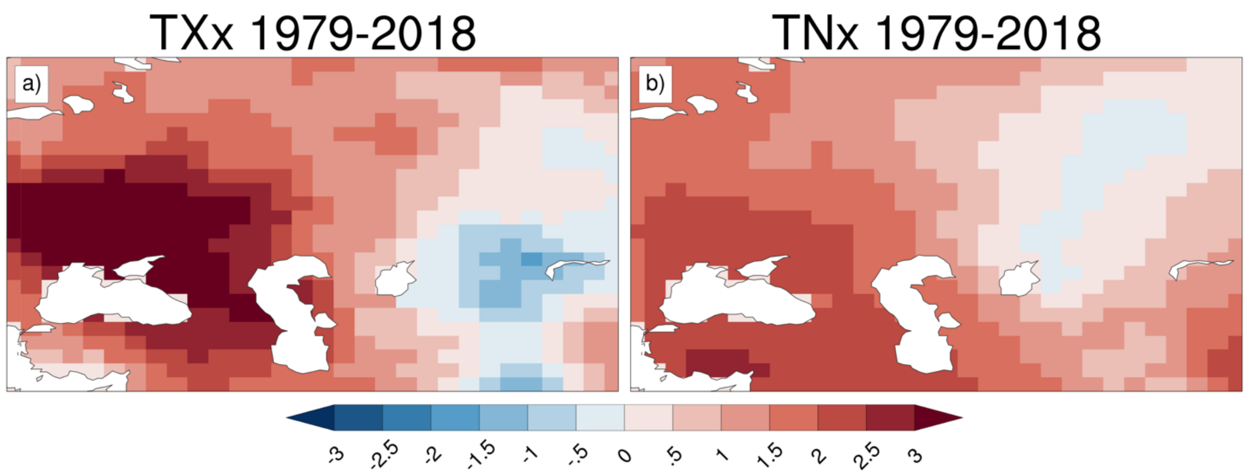

Figure R4: TXx and TNx 1979–2018 trend patterns derived from the HadEX3 dataset

Details, typing errors, etc.
To ensure reproducibility of the study the location and extent of the SLP regions that were used for the two events in 3.1 and 3.2 still need to be specified. I only found the numbers for the TX regions in the tables.
Yes, agreed. Note that the numbers given in table legends correspond to the boundaries of the black boxes in Figure 1c and 3c.
Manuscript change:
Geographical boundaries for SLP and TX have been added in the method section (lines 167–171).
Western Europe cold spell: 25°N–90°N; 60°W–100°E
Russian heat wave: 25°N–90°N; 20°W–80°E

l.103: Do you mean TN maxima?
Yes, thank you.
Manuscript change:
“TN minima” has been replaced by “TN maxima”.

L.115: Later in the manuscript (e.g Table 1) an underscore is used before the version number in 20CR_V3
Manuscript change:
The acronym 20CR_V3 is now used throughout the manuscript.

l.220: The period chosen for the cold European winter is not exactly two weeks but 17 days
Manuscript change:
“two-week period” has been replaced by “seventeen-day period”.

l.237: Do you mean Fig. 1b?
Yes, thank you.
Manuscript change:
The typo has been corrected.

l.254: insert “the” before amplitude
Manuscript change:
‘the” has been inserted before amplitude.

Fig.1: I find it hard to distinguish the one contour line that is thicker. Line thickness could be increased or the zero SLP anomaly line could be highlighted in a different color e.g. violet.
Manuscript change:
The thickness of the 0-line has been increased.

l.269: typo, should be 2009-2010 early winter
Thank you
Manuscript change:
the typo has been corrected.

l.270: Shouldn’t the numbers in the brackets be -3.07 °C and -2.04 °C?
Yes, thank you
Manuscript change:
The numbers have been corrected.

l.271: 20CR_V3
Manuscript change:
20CR_V3C has been changed to “20CR_V3”.

l.336: I would close the first bracket after “anomaly”

Manuscript change:
A bracket has been added after “anomaly”.

l.381: For the TX, TN and the SLP domain it should say °E instead of °W.
Yes, thank you.
Manuscript change:
The typos have been corrected.

l.387: I think the bracket saying “smallest” should be omitted as you are using the largest anomalies whether it is TX or TN.
Yes, thank you.
Manuscript change:
“(smallest)” has been removed.

l.444: Was it not mitigated by around one third from -3.07°C to -2.04°C (instead of 50%)?
Yes, thank you for pointing this.
Manuscript change:
50 % has been replaced by 33 %.

Reply to reviewer 2 (T. Woollings):

General Comments

This is a nice paper which applies the existing method of dynamical adjustment to investigate some aspects of European climate change and extreme weather. The paper is sound and well written, and I am supportive of acceptance after the points below have been considered. My main comment is that the paper could be improved by the addition of some uncertainty analysis.

I thank the reviewer for his detailed lecture, helpful comments and suggestions, and appreciation of the manuscript.

1. The paper is aimed at the event attribution problem, ie in quantifying the role of climate change in extreme weather events. This is very clear in the abstract and the introduction. However, this specific method instead quantifies the circulation contribution, and any climate change effects are left in the residual. This is still very useful, but it could be made clearer early on that this method alone cannot make statements about the role of any particular external forcing, such as greenhouse gases or aerosols.

Thank you for the comment. I think that the scientific field of extreme event attribution encompasses both approaches, the risk-based and the process-based. As mentioned in the text, the two methods can (and perhaps should always …) be combined to enhance the robustness of the results. The referee rightly points out that the dynamical adjustment method as used in the paper cannot make statements about the respective role of different external forcings. First, I would add that this is also true of many risk-based studies that most of the time do not use single-forcing model experiments. Note also that this inability to make single-forcing attribution statements does not come from the dynamical adjustment but rather from the fact that the paper approach only relies on observations. Indeed, dynamical adjustment could be used on large ensembles of single-forcing simulations such as those presented in Deser et al. 2020 or performed in the framework of DAMIP. The use of large ensembles with all- and single-forcing jointly with dynamical adjustment would allow a **model-based** quantification of the total (thermodynamic and dynamic) model response to different forcings as well as more robust statements about any potential forced dynamical response to combined and individual forcings. To this regard, a multimodel community resource similar to the MMLEA (https://www.cesm.ucar.edu/projects/community-projects/MMLEA/) would be extremely helpful.

Manuscript changes:
Based on the above text, I have added a short discussion (lines 85–91) in the introduction pointing out the limitation of using only observations and therefore not being able to make single-forcing attribution statements.

2. Given that the main aim of the paper is to quantify the role of atmospheric circulation, it seems the method could be improved by adding an estimate of the uncertainty in this quantification. Even very close analogues in surface pressure will likely have differences in temperature due to large-scale effects not captured by the surface pressure. Hence, to complement the mean effect of circulation as used here, an uncertainty range could be given. More rigorous statements could then be made about the residual terms, in particular as to whether they are within the uncertainty range of the dynamical contribution or not.

Thank you for the comment. I agree with the reviewer that adding an uncertainty analysis for the two extreme events (specifically for Figures 1 and 3) is a very good idea (note that the daily uncertainty analysis is already performed in Figures 2a and 4a. In the case of the Russian heat wave, it suggests that the total residual – as well as the RES_INT term (not shown) – is most of the time much greater than the dynamical uncertainty range).

Manuscript changes:
I have implemented a bootstrapping method inspired by the one used in O'Reilly et al. 2017 that is now described in Sect. 2.2 (see lines 183–188). I have then modified Figures 1 and 3 by adding stippling to panels 1d and 3d when the RES_INT term values are within the dynamical uncertainty range.

3. One term is labelled RES_ADV and frequently discussed as representing advection, but this is never tested. Changes in advection will indeed contribute to this term but it is not clear that other processes do not contribute. This seems especially likely in summer, when advection is relatively less important for temperature variability and other factors such as radiation or adiabatic heating anomalies may play an important role (eg Pfahl and Wernli, Quinting and Reeder). Could this term be re-named, or the role of advection tested (eg fig 1f does look consistent with the changing nature of advection from the warming Arctic...).

This term is estimated by running the dynamical adjustment twice, once with TX detrended and once with the raw TX. By construction, it necessarily includes thermal advection changes related to externally forced changes in zonal and meridional TX gradients (as pointed out by the reviewer in Figure 1f). However, I agree with the reviewer that it does also include other contributions such as those mentioned in the above comment.

Manuscript changes:
I have changed the naming of the two terms RES_FRC and RES_ADV. RES_FRC has been changed to RES_TRD (indeed it is the fraction of the anomaly which is due to the long-term trend). RES_ADV has been changed to RES_FRC. I have also modified in the text (lines 213–214) the definition of the new RES_FRC to make clear that it is not only related to advection as just discussed. As a consequence, I have also updated the meaning of DYN_TOT that can now be defined as an upper bound of the total dynamic contribution.

4. Around Figure 1 there is speculation that the RES_INT anomalies reflect the role of regional SST features. This discussion could be informed by the use of uncertainty analysis as in point 2 above. In particular, are these anomalies outside the range of uncertainty in the dynamic term?

I agree, please also see the response to comment 2.

Manuscript changes:
I have added a few sentences about the significance (or lack of) of the RES_INT term for both events.

5. Has the author tested for trends affecting the estimated dynamical contribution? It seems possible that for a given case the selected analogues may not span the period evenly. Eg, if the analogues happened to sample the most recent decades only, they would then not give a representative sample of the temperatures over the whole period. Hopefully this is unlikely, but could perhaps happen when there is a lack of SLP variance in the early, data-poor, period of the reanalysis?

Thank you for pointing this out. I have checked whether the analogue selection is biased to a specific period or spans evenly the full period.

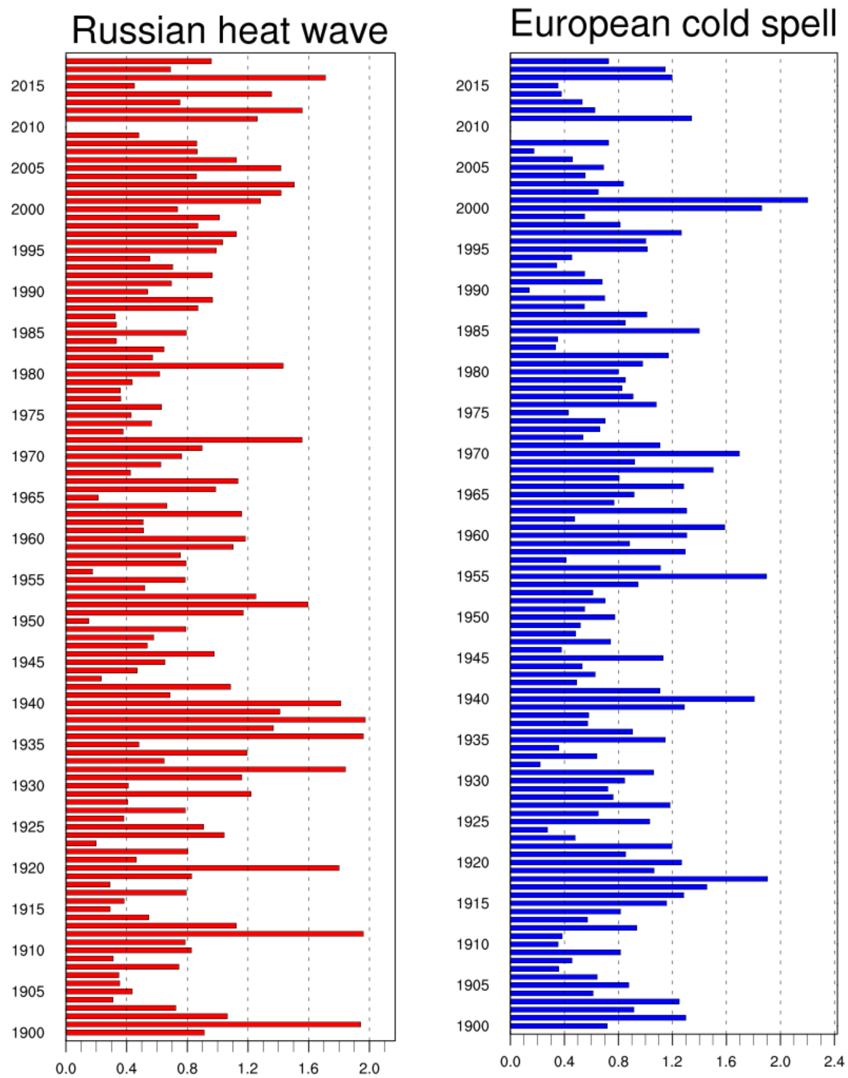

Figure R5: distribution of analogues (X-axis, unit in percent of the total number of used analogues given above) versus their year of occurrence (Y-axis) for the two extreme events

Figure R5 shows the distribution of analogues with respect to the years for the two extreme events (for the entire event, the total number of analogues used is equal to N_r x N_s x N_d, with N_r and N_s defined as in the paper and N_d the number of days of the event). It clearly shows that the selected sample of analogues does not exhibit any particular trend and that specific years with a large number of analogues can be found throughout the entire period.

Manuscript changes:
A new Appendix, Appendix B (lines 615–626), has been added with the above figure (Figure B1 in the manuscript) and a short discussion. One sentence (lines 191–192) has also been added to the method section to refer to the new Appendix B.

6. Some of the results in section 4 are interesting and some further discussion might be helpful. eg why is there a 'thermodynamic' cooling trend in fig 5e, and what mechanisms underlie the

strong dynamical trends shown here. Could these be consistent with internal variability associated with AMV, as described for example by Sutton and Dong (2012)?

Thank you for the question. Following a similar comment from reviewer 1, I have first investigated whether the cooling trend seen in Figure 5e is robust to observational uncertainty. Therefore, I have used the HadEX3 dataset (gridded values of both TXx and TNx, Dunn et al. 2020) to check the BERK results. Figure R6 below shows the TXx and TNx 1979-2018 trends based on HadEX3 (***note that I have slightly extended the domain eastward in Figure R6 compared with Figure 5 in order to make the point below***). Figure R6 does show a reasonably similar pattern to that of BERK for the TXx trend pattern. However, the HadEX3 pattern is different from the BERK one for TNx, with weak warming in the center and eastern part of the domain shown in Figure 5. A slight cooling is observed, but to the east of the region shown in Figure 5. In addition, the HadEX3 TXx and TNx trend patterns both show this cooling east of 60°E in the western Siberia plains (with a larger amplitude for TXx than TNx). This is in good agreement with recent findings about observed mean temperature changes in this region that have resulted from increased winter and spring precipitation in recent decades, delayed snowmelt and increased soil moisture in summer as well as land-atmosphere interaction (see Sato and Nakamura 2019, Bulygina et al. 2009, Bulygina et al. 2011, Guo et al. 2019). Therefore, I speculate that the westward-shifted cooling seen in the BERK TNx trend pattern might result from the scarcity of stations in this region for the BERK TN dataset and the use of a large influence ratio in the infilling procedure.

Following the reviewer suggestion, the manuscript now includes an estimate of the AMV influence on TXx and TNx changes, based on a HadEX3-based simple composite analysis of the difference between periods with positive and negative AMV phases (as in Figure 6 from O'Reilly et al. 2017 for mean summer temperature). Given the recent time evolution of the AMO/AMV index during the 1979–2018 (going from a cold phase to a warm phase), this composite analysis provides a simple fingerprint that can be used to estimate the AMV induced contribution to TXx and TNx 1979–2018 changes.

Manuscript changes:
I have split Section 4 into 2 sub-sections: 4.1 that discusses the trend results based on BERK as well as observational uncertainty based on HadEX3 results, and 4.2 that details the causal factors of the extreme temperature changes over the 1979–2018 period (focusing on the AMV).

Specifically, I have added a detailed discussion (lines 467–476) in section 4.1 about TXx and TNx trend patterns and observational uncertainty based on a new Figure, Fig. 7, that shows the TX and TN trend patterns based on HadEX3 for both WE and WA regions (an extended version of Fig. R6). Note that Fig. 7 also includes a panel showing a composite-based estimate of the AMV influence on both TXx and TNx changes that is now used in Section 4.2 (lines 493–517) for the discussion of causal factors responsible for extreme temperature changes. A new Appendix, Appendix C, with one figure (Fig. C1 in the manuscript) has also been added to support the discussion about the TNx cooling and observational uncertainty in section 4.1.

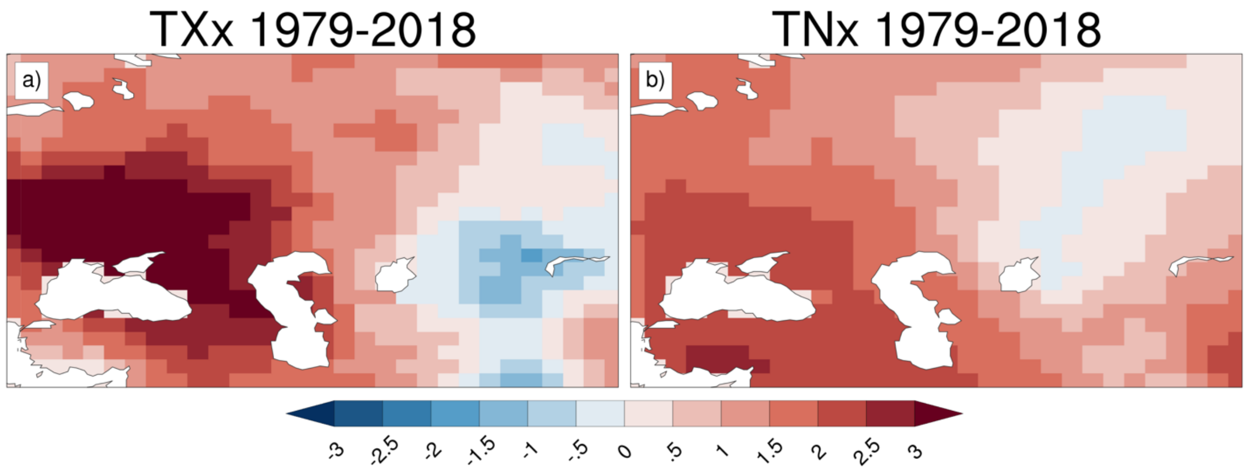

Figure R6 (identical to R4): TXx and TNx 1979–2018 trend patterns derived from the HadEX3 dataset

7. On the Russian heatwave: the text suggests this has been 'mainly linked' to La Nina. I agree this is very likely a factor (see also Drouard and Woollings 2018, GRL), but even so this statement feels a bit strong.

Yes, Agreed

Manuscript changes:

I have removed “mainly” and added the suggested reference.

References

Bulygina, O.N. et al., 2011 *Environ. Res. Lett.* 6 045204

Bulygina, O.N. et al., 2009 *Environ. Res. Lett.* **4** 045026

Dunn, R. J. H., Alexander, L. V., Donat, M. G., Zhang, X., Bador, M., Herold, N., et al. (2020). Development of an updated global land in situ-based data set of temperature and precipitation extremes: HadEX3. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD032263. https://doi.org/10.1029/2019JD032263

Guo, R., Deser, C., Terray, L., & Lehner, F. (2019). Human influence on winter precipitation trends (1921–2015) over North America and Eurasia revealed by dynamical adjustment. Geophysical Research Letters, 46. https://doi.org/10.1029/2018GL081316

Sato, T., Nakamura, T. Intensification of hot Eurasian summers by climate change and land–atmosphere interactions. *Sci Rep* 9, 10866 (2019). https://doi.org/10.1038/s41598-019-47291-5