

Peer Review File

Exacerbated summer European warming not captured by climate models neglecting long-term aerosol changes



Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality

16th Oct 23

Dear Dr Schumacher,

Your manuscript titled "Why climate models underestimate the exacerbated warming in Western Europe" has now been seen by 3 reviewers, and we include their comments at the end of this message. They find your work of interest, but some important points are raised. We are interested in the possibility of publishing your study in *Communications Earth & Environment*, but would like to consider your responses to these concerns and assess a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, along with a point-by-point response that takes into account the points raised. In particular, please:

- * make a compelling case that the omission of time-evolving aerosol concentrations in regional climate model simulations explains part of the underestimation of amplified warming in Western and Central Europe.
 - * provide a thorough clarification on the methodology (e.g. choice of the domain and CMPI6 models)
- Please highlight all changes in the manuscript text file.

We are committed to providing a fair and constructive peer-review process. Please don't hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and the completed checklist:

[link redacted]

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to receive your revised paper within six weeks; please let us know if you aren't able to submit it within this time so that we can discuss how best to proceed. If we don't hear from you, and the revision process takes significantly longer, we may close your file. In this event, we will still be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at *Communications Earth & Environment* or published elsewhere in the meantime.

Please do not hesitate to contact us if you have any questions or would like to discuss these revisions further. We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Best regards,

Mengze Li, PhD
Editorial Board Member
Communications Earth & Environment

Heike Langenberg, PhD

Chief Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMATTING

We ask that you ensure your manuscript complies with our editorial policies. Please ensure that the following formatting requirements are met, and any checklist relevant to your research is completed and uploaded as a Related Manuscript file type with the revised article.

Editorial Policy: [Policy requirements](#) (Download the link to your computer as a PDF.)

Furthermore, please align your manuscript with our format requirements, which are summarized on the following checklist:

[Communications Earth & Environment formatting checklist](#)

and also in our style and formatting guide [Communications Earth & Environment formatting guide](#) .

*** DATA: Communications Earth & Environment endorses the principles of the Enabling FAIR data project (<http://www.copdess.org/enabling-fair-data-project/>). We ask authors to make the data that support their conclusions available in permanent, publically accessible data repositories. (Please contact the editor if you are unable to make your data available).

All Communications Earth & Environment manuscripts must include a section titled "Data Availability" at the end of the Methods section or main text (if no Methods). More information on this policy, is available at <http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf>.

In particular, the Data availability statement should include:

- Unique identifiers (such as DOIs and hyperlinks for datasets in public repositories)
- Accession codes where appropriate
- If applicable, a statement regarding data available with restrictions
- If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement.

DATA SOURCES: All new data associated with the paper should be placed in a persistent repository where they can be freely and enduringly accessed. We recommend submitting the data to discipline-specific, community-recognized repositories, where possible and a list of recommended repositories is provided at <http://www.nature.com/sdata/policies/repositories>.

If a community resource is unavailable, data can be submitted to generalist repositories such as [figshare](#) or [Dryad Digital Repository](#). Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publically available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.

Please refer to our data policies at <http://www.nature.com/authors/policies/availability.html>.

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Review of Communications Earth & Environment manuscript COMMSENV-23-1350 “Why climate models underestimate the exacerbated warming in Western Europe” by Schumacher et al.

This paper examines the rising trends in summer temperatures in Western Europe and why climate models, particularly regional CORDEX models, are failing to capture the rapid recent warming. Europe has suffered from severe heatwaves in recent years, and the region is in fact a global 'hotspot' where heatwaves have intensified rapidly in recent years. Attribution studies have shown that climate models struggle to simulate recent changes in summer high temperatures. It is therefore of great scientific interest to investigate the causes behind the differences between observations and models, justifying the topic of this paper.

I liked the story of the paper. It proceeded logically from the observed and simulated summer warming trends, to first separating the dynamic and thermodynamic components of the trends, and then quantifying the model biases due to aerosol changes. The main outcome of the paper was that because the aerosol forcing in certain regional models is constant, the models do not capture the increased radiation due to the reduction in aerosols and thus simulate too weak warming.

The language was understandable and the subject was well backgrounded in the Introduction and also throughout the text. The figures were clear. Overall, I think the topic is well suited to Communications Earth & Environment and should be published. I only have a few rather minor comments which I hope the authors can address before publication.

Minor comments

1. Why did you focus only on western Europe and not the full domain of IPCC WCE region? It seems that the warming and model-observations discrepancy is even greater in the eastern half of Central Europe. So wouldn't the research and the results be even more appropriate for Eastern Europe?
2. I had difficulty understanding how you derive i) regional thermodynamic warming, ii) regional thermodynamic bias and iii) Rnet-based TD bias. Thus, the beginning of Section “Quantifying the aerosol forcing-induced warming mismatch” was somewhat hard to understand. Could you elaborate the calculation of these three terms, perhaps even devote a subsection to them in the Methods section?
3. You only use 7 CMIP6 models, and after limiting the background warming, you end up with two CMIP6 models with 15 realizations. Why didn't you use a wider range of CMIP6 models? Doesn't the model uncertainty remain quite high especially when you only focus on two different models for example in Fig. 3 (the CMIP6 subset). I understand that the focus is on CORDEX models so this might not be a large issue.

Line-to-line comments

L99: Focusing

L231-232: "Given that the dynamic contribution to summer warming in WWCE is a manifestation of unforced internal variability...". Is this 100 % certain statement? If yes, citation would be nice. Otherwise I'd probably add "mostly" or "mainly".

L270-271. What does it mean to "remove the background warming from the total thermodynamic warming"? I understood that Fig. 3d results all have the same background warming. But how can one remove background (global) warming contribution from the total thermodynamic warming? This could be elaborated in the manuscript as it seems to be a key step in the study.

L314-315. "since this is what each model simulation — and, contrary to the dynamic warming, especially the ensemble mean — should capture." This may be a self-evident sentence for climate scientists, but I had to read the statement a couple of times, and I'm still not 100 % sure why each model simulation should capture the thermodynamic contribution. Maybe some elaboration would help?

L318-321: "We estimate ..." What's regional thermodynamic bias? It's not listed in Table 1. This is related to Minor comment 2.

Figures: Figure subtitles (a,b,c ...) are either badly placed or partly missing

Fig. 2b: Does "hottest model" refer to the hottest model realization? When you only have two models in the CMIP6 subset, it might be more reasonable to call it realization than model.

Supplementary information: Reference list seems to be missing.

Reviewer #2 (Remarks to the Author):

Based on ERA5 reanalysis data, RCMs and GCMs, the study analyzed and quantified the impact of aerosol changes on warming in Western Europe. The results show some new phenomena. However, the following must be commented on or addressed before publishing.

1. The temperature simulation has significant uncertainty at the regional scale, although the ERA5 reanalysis data performs well in reproducing the climate globally, it is still a model reanalysis and can not take the place of the benchmark dataset. It is recommended to add in situ data for comparative analysis.

2. Another recent article (Li et al, 2022) pointed out that the model's underestimation/overestimation of regional temperature trends is mainly due to its underestimation of its response to external temperature forcing and its persistence. In my opinion, it also seems very reasonable. What are the substantive differences between this article and that one? Is there a certain correlation between the two?

(Ref: Li, Q., Sheng, B., Huang, J. et al. Different climate response persistence causes warming trend unevenness at continental scales. *Nat. Clim. Chang.* 12, 343–349 (2022).

<https://doi.org/10.1038/s41558-022-01313-9>)

3. How is the boundary of the study area (western West-Central Europe) defined? This study mainly focuses on western West-Central Europe, but the title is Western Europe. It is recommended to change the title to western West-Central Europe as well.
4. The east side outside the study area has a more pronounced difference between CORDEX RCMs evol. and const. of the SW and temperature. It is recommended to include this in the study area.
5. The influence of clouds and aerosols on downward shortwave radiation in CORDEX are expected to be given in Fig.S3.
6. As shown in Fig.S5, the RCMs only underestimate warming significantly in the summer months of June, but also in other months (such as February and April). The authors should analyze these months with a large warming gap.
7. While considering the effects of aerosols and clouds, it is also necessary to consider the effects of greenhouse gases.
8. There seem to be some issues with the annotation of subfigures(a,b,c,d,....) that need improvement.

Reviewer #3 (Remarks to the Author):

The manuscript addresses the issue of underestimation in the historical warming trend within Western Europe, particularly during the summer season, as projected by CMIP6 GCMs and EURO-CORDEX RCMs. The authors commence by characterizing the warming disparity within both ensembles and subsequently endeavor to discern the primary underlying causes. They conclude their analysis by evaluating how these identified factors may also introduce biases in future warming projections within the EURO-CORDEX ensemble, encompassing both mean climate and temperature extremes.

The subject matter of this study holds profound scientific relevance for a broad readership. Western Europe has witnessed a significantly accelerated rate of warming compared to other global regions. Furthermore, climate change adaptation measures are presently being formulated and implemented based on EURO-CORDEX datasets, disseminated through various climate service data portals at European, national, and local levels. Thus, comprehending and reporting on EURO-CORDEX's historical and future warming underestimations in Western Europe assumes paramount importance.

Additionally, I wish to acknowledge the authors for their clear and well-organized presentation in the manuscript. The introduction provides a solid foundation by succinctly outlining the research objectives. The methodology is robust and effectively executed, and the results provide new insights of relevance to a significant scientific community.

However, I do have some general concerns that, in my opinion, fall within the scope of second-order issues that can be addressed during the revision process. Specifically, I recommend the following improvements:

1. Utilize existing references more effectively, not only in the introduction but also throughout the manuscript.
2. Evaluate the performance of the evaluation runs generated within EURO-CORDEX.
3. Enhance the readability of Table 1.
4. Improve the clarity and informativeness of the abstract.
5. Consider providing supplementary annexes with simulation-specific information to aid the

broader user community.

I firmly believe that these general concerns, while essential for enhancing the content and overall relevance of the manuscript, can be adequately addressed by the authors. Therefore, I recommend the publication of this manuscript upon incorporating these recommendations. I am willing to reassess the revised version of the manuscript after the modifications have been made.

1. Below are the detailed descriptions of my main concerns:

1.1. Make better use of existing references

1.1.1. I have noticed instances in the introduction where the use of references may be further optimized to substantiate certain statements. Given my background in the RCM community, I would suggest considering alternative references, particularly in lines 36, 37, and 38. Notably, comprehensive multi-model studies with CPRCMs are now available and could enhance the support for your arguments.

1.1.2. It appears that there are at least three references missing in the introduction: Ribes et al. 2022, around line 52; Nabat et al. 2014, mentioned at lines 75-76 (previously cited as #40); and Lorenz and Jacob 2010 (see the reference list at the end of this review)

1.1.3. I also recommend that the authors review the recent work by Taranu et al. (2023) regarding the causes of GCM-RCM inconsistency in EURO-CORDEX future projections in Europe during the summer. This article, being very recent, may not have been on the authors' radar, but it provides a valuable alternative approach to characterizing and disentangling the sources of discrepancy.

1.1.4. Additionally, I suggest revisiting the use of reference #23 in the introduction in light of the comment on PubPeer, which may warrant a modification of the use of this reference. (PubPeer link: <https://pubpeer.com/publications/2F852866CC945DB4615D4FC40105FC#1>)

1.2. Verify the behavior of the evaluation runs produced in EURO-CORDEX

While the authors have thoroughly examined the past trends in the historical runs of EURO-CORDEX, where RCMs are driven by GCMs with random natural variability, I would like to draw attention to another category within EURO-CORDEX known as "evaluation runs". In these runs, RCMs are driven by reanalyses, such as ERA-Interim in the earlier CORDEX phase or ERA5 in the current phase. Investigating the warming trend in these evaluation runs could provide complementary insights. It is plausible that the evaluation runs may also exhibit underestimation of past trends, which would further corroborate the role of aerosols in most RCMs. Should this be the case, a brief annex or an entry in Table 1 may suffice to convey this information.

1.3. Improve the readability of Table 1

Table 1 contains crucial data summarizing the study, which holds significant value for future research endeavors. However, its current presentation is not as user-friendly as desired. I propose a clearer organization that distinctly conveys four primary pieces of information: Total warming, Background warming, Dynamic, and Local Thermodynamic. It remains unclear why Background warming is not

provided for ERA5.

1.4. Enhance the abstract

Overall, the paper is well-structured, comprehensible, and communicates its implications effectively, particularly in the conclusion section. Paradoxically, I find that the abstract falls slightly short of maintaining the same level of clarity and emphasis. I recommend revising the abstract to enhance its readability and emphasize the key messages.

1.5. Assist the user community by adding annexes with simulation-specific information

The potential impact of this study extends to a broad community, including climate modelers, climate service providers, and climate data users, such as the impact community and the end-users. To facilitate these communities' effective utilization of your findings, I propose the addition of an annex that characterizes individual CMIP6 and CORDEX (and possibly CMIP5) simulations. This could resemble a simplified version of Table 1, providing details for each model simulation. Such an annex would enable climate data users to select EURO-CORDEX simulations post hoc and help the regional climate modeling community select adequate driving GCMs for the CORDEX-CMIP6 initiative. In particular, in CORDEX-CMIP5, six out of 49 simulations are notable for capturing the appropriate warming background due to their driving GCMs and the evolving aerosol dynamics. Listing these simulations would be valuable.

2. Minor comments:

2.1. Title: The word "Summer" is absent from the title, which does not fully reflect the content of the paper.

2.2. Line 41: "RCMS" should be corrected to "RCMs."

2.3. Line 42: It should be noted that coordination of RCM simulations existed before CORDEX since the 1990s. What distinguishes CORDEX is its "worldwide" coordination.

2.4. Lines 59-60: The impact of changes in land-sea contrast and regional lapse rate, particularly during the summer in Europe, is worth mentioning.

2.5. Lines 99-100: Is it feasible to use GCM large ensembles to distinguish human-induced changes from those related to unforced natural variability? This approach could offer valuable insights for interpretation (Line 312) and might be worth including in Table 1 as supplementary information.

2.6. Figure 1b: Noticeable differences along the coastline in Fig. 1b may be linked to an issue with the land-sea mask in handling the ERA5 dataset, where mixed land-sea points cover a larger area than in 12km CORDEX runs. It would be helpful to clarify how mixed land-sea points in ERA5 were addressed.

2.7. Figure 1 caption: It should be clarified that the curves and numbers in Fig. 1 are computed only over land points.

2.8. Line 134: Please specify at least once which dataset is used in this paper, considering that various CORDEX datasets exist over the European domain (e.g., 50 km resolution, Med-CORDEX, evaluation runs).

2.9. Lines 146-149: Over recent years, CMIP5 (hence CORDEX) and CMIP6 differ in terms of GHG concentration and aerosol forcings. The historical period is shorter in CMIP5, which introduces a source of error, potentially small. This difference should be acknowledged.

2.10. Figure 2: Please provide a clear means for the reader to identify which simulations are considered "good," meaning those with favorable DeltaTglo and DeltaT-WWCE.

2.11. Please specify the aerosol dataset used in ERA5 and its reliability. Has this dataset been evaluated in a published paper? Is ERA5 considered reliable for simulating climate-aerosol interactions (direct, semi-direct, indirect effects)? This information is critical for interpreting the results.

2.12. Figure 3a: It might be prudent to reconsider the utility of Fig. 3a, as it appears challenging to create a comprehensive and easily readable representation. For example, notable factors such as aerosol effects and land-sea contrast are omitted. It may be worth considering its removal.

2.13. Lines 236-237: The statement suggests that GCM selection in CORDEX phase 1 was not ideal. It would be beneficial to provide recommendations for improving the situation in CORDEX-CMIP6 in the conclusion.

2.14. Figure 4: The caption indicates 18 CORDEX simulations, while 15 is mentioned in the text (Line 239), and Table 3 lists only 14. Please clarify this discrepancy.

2.15. Table 3: The table lists 43 simulations, whereas 49 are mentioned in Lines 133 and 332. Please ensure consistency.

2.16. Line 376, Figure 5 caption: "Linear trends" is mentioned, while the unit in Fig. 5b is °C. Ensure consistency.

2.17. Lines 388-389: It is stated that "authors assessed surface rather than air temperature changes." Please verify this statement, as the "tas" variable (near-surface air temperature) is used in Boé et al. 2020, similar to your study.

2.18. Line 410-411: The study of the correlation between aerosol conditions and cloud conditions over Europe in Nabat et al. (2020) may provide valuable insights for your interpretation.

2.19. Line 443: You may consider replacing this text with the equivalent statement in IPCC-AR6 WGI, where RCMs are deeply assessed. For example, Chapter 10, Executive Summary states, « The performance of global and regional climate models and their fitness for future projections depend on their representation of relevant processes, forcings and drivers and on the specific context. [...] Including the relevant forcings (e.g., aerosols,[...] in global and regional models is a prerequisite for reproducing historical regional trends and ensuring fitness for future projections (high confidence).»

2.20. Conclusion: It would be valuable to add perspectives and guidance for the RCM community

regarding the upcoming CORDEX-CMIP6 exercise, especially concerning CMIP6 GCM selection and aerosol representation.

4. References of this review not already cited by the authors:

Ribes et al. 2022, <https://doi.org/10.5194/esd-13-1397-2022>

Taranu et al. 2023, <https://doi.org/10.1007/s00382-022-06540-6>

Lorenz and Jacob 2010, <https://doi.org/10.3354/cr00973>

Nabat et al. 2020, <https://doi.org/10.5194/acp-20-8315-2020>

Reviewer #1 (Remarks to the Author):

Review of Communications Earth & Environment manuscript COMMSENV-23-1350 “Why climate models underestimate the exacerbated warming in Western Europe” by Schumacher et al.

This paper examines the rising trends in summer temperatures in Western Europe and why climate models, particularly regional CORDEX models, are failing to capture the rapid recent warming. Europe has suffered from severe heatwaves in recent years, and the region is in fact a global 'hotspot' where heatwaves have intensified rapidly in recent years. Attribution studies have shown that climate models struggle to simulate recent changes in summer high temperatures. It is therefore of great scientific interest to investigate the causes behind the differences between observations and models, justifying the topic of this paper.

I liked the story of the paper. It proceeded logically from the observed and simulated summer warming trends, to first separating the dynamic and thermodynamic components of the trends, and then quantifying the model biases due to aerosol changes. The main outcome of the paper was that because the aerosol forcing in certain regional models is constant, the models do not capture the increased radiation due to the reduction in aerosols and thus simulate too weak warming.

The language was understandable and the subject was well backgrounded in the Introduction and also throughout the text. The figures were clear. Overall, I think the topic is well suited to Communications Earth & Environment and should be published. I only have a few rather minor comments which I hope the authors can address before publication.

We thank the reviewer for their positive and thorough assessment of our work, and reply individually to all specific questions below using blue font.

Minor comments

1. Why did you focus only on western Europe and not the full domain of IPCC WCE region? It seems that the warming and model-observations discrepancy is even greater in the eastern half of Central Europe. So wouldn't the research and the results be even more appropriate for Eastern Europe?

We thank the reviewer for raising this question; the focus on western Europe in the study is mandated by several considerations. One aspect is that Vautard et al. (2023) mainly found circulation-related amplification of heat extremes in western rather than eastern parts of Europe, but it remains unclear to which extent mean temperatures are affected. More importantly, the eastern edge of the entire IPCC WCE region is close to the boundary of the

CORDEX EUR-11 domain, such that the dynamical adjustment approach — in which regional temperatures are predicted by the surrounding large-scale circulation pattern — cannot be applied. This is now also mentioned in the methods:

“Our main analysis is restricted to boreal summer (June–August) in 1980–2022 (or 1980-2099), and we focus on the western half (-10°W to 15°W) of West-Central Europe (WCE), an IPCC climate reference region⁴⁶, referred to WWCE hereafter. This choice is motivated by the fact that the eastern edge of the entire WCE region at 40°W is too close to the regional model domain boundaries to apply the dynamical adjustment approach described further below. All climate variables are area-weighted when calculating regional means.”

It is thus not possible to extend the entire analysis to the full WCE region, or conduct it analogously for ‘EWCE’, the eastern half of WCE: without dynamical adjustment, we cannot disentangle the thermodynamic warming, based upon which the warming mismatch of RCMs with constant aerosols emerges clearly for WWCE.

Moreover, we believe that slightly changing our analysis region to some entirely arbitrary region, e.g. to include some but not all of the area in eastern Europe where we estimate aerosol-related temperature trend differences on the order of 1°C for 1980–2022 during summertime, would result in a selection bias.

Nonetheless, as ERA5 provides the required geopotential height fields (Z500) for the dynamical adjustment globally, we also disentangled the warming contributions for EWCE and complemented our analysis with a comparison of the warming in both WWCE and EWCE in the CORDEX simulations (shown further below in this rebuttal as a separate section).

2. I had difficulty understanding how you derive i) regional thermodynamic warming, ii) regional thermodynamic bias and iii) Rnet-based TD bias. Thus, the beginning of Section “Quantifying the aerosol forcing-induced warming mismatch” was somewhat hard to understand. Could you elaborate the calculation of these three terms, perhaps even devote a subsection to them in the Methods section?

In response to this as well as other comments, we have adjusted and simplified **Table 1**, and no longer derive the regional thermodynamic warming. In the initial submission, we conceptualized regional thermodynamic warming as the difference between total thermodynamic and global background warming (obtained from the respective driving GCM). The “true” background warming in WWCE, however, is arguably much more dependent on neighboring European regions and in particular the Northern Atlantic rather than the entire globe, hence we no longer employ this conceptualization. In **Fig. 4a**, we now show the total thermodynamic warming (previously shown in the Supplementary), whose slope remains effectively unchanged, yet the intercept now indicates the amount of thermodynamic warming

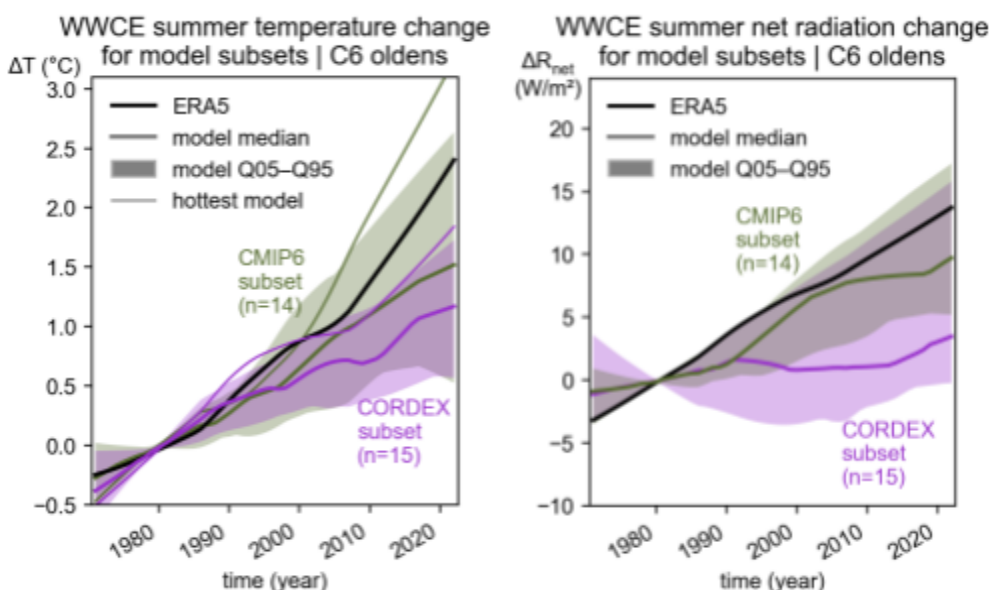
unrelated to regional radiation changes (i.e., background warming and other regional drivers such as land–use changes).

We previously employed this slope — the relationship between thermodynamic warming and net radiation changes — to approximate the overall aerosol-related warming underestimation based on a CORDEX ensemble mean net radiation increase underestimation of about 10 W/m^2 , resulting in $\sim 0.5^\circ\text{C}$. However, additional analyses indicate that there is a certain RCM dependence concerning the temperature sensitivity to net radiation changes; therefore, we now provide one estimate for the model subset, and an additional one for the entire ensemble, since they consist of different GCM–RCM model chains. As such, the discussion revolving around Table 1 has been edited substantially, hence we do not paste all of the text here, and instead refer to the updated manuscript.

3. You only use 7 CMIP6 models, and after limiting the background warming, you end up with two CMIP6 models with 15 realizations. Why didn't you use a wider range of CMIP6 models? Doesn't the model uncertainty remain quite high especially when you only focus on two different models for example in Fig. 3 (the CMIP6 subset). I understand that the focus is on CORDEX models so this might not be a large issue.

The initial analysis for this study was based on a far larger CMIP6 ensemble, and we present Figs. 2b and c below for this original ensemble, or more accurately a 14-member subset with near-ERA5 background warming, consisting of the following models (**RFig. 1**):

BCC-CSM2-MR, CMCC-ESM2, CNRM-CM6-1, FGOALS-f3-L, GFDL-ESM4, IITM-ESM, INM-CM4-8, INM-CM5-0, MIROC6, MIROC-ES2L, MPI-ESM1-2-HR, MPI-ESM1-2-LR, NorESM2-LM, NorESM2-MM.



RFig. 1: Like Fig. 2b & c, but using an alternative CMIP6 model subset.

Based on this, we noticed that CMIP6 models with global warming since 1980 roughly in line with ERA5 typically underestimate the summer warming in WWCE, yet not as much as a CORDEX GCM–RCM simulation subset with (by design) comparable background warming. The CMIP6 simulations also tend to better capture the net radiation increases evident for ERA5, raising the question why the simulated warming is generally weaker than observed.

We then extended our analysis by applying a dynamical adjustment approach to a different CMIP6 ensemble that showed the same general underestimation of summer warming in WWCE, and found that weaker simulated dynamic contributions than in ERA5, largely explain the warming discrepancies. We believe that the work of Vautard et al. (2023), which is based on a large CMIP6 ensemble consisting of 30+ models, already conclusively demonstrates that the below-observed simulated rise in heat extremes has predominantly dynamical origins, and that realizations from a single model can portray a large spread in dynamic contributions (see, e.g., MPI-ESM1-LR in their Suppl. Fig. 6). Our results are based on a CMIP6 ensemble consisting of fewer models, but instead many initial condition ensemble members, and we are confident that repeating our entire analysis for a different model ensemble would not change any of our conclusions.

As the reviewer points out, our focus is indeed on the CORDEX RCM simulations, and we mainly employ CMIP6 data to provide an additional line of evidence when contrasting simulated temperature and net radiation changes with reanalysis data.

Line-to-line comments

L99: Focusing

Thank you for spotting this typo, it has been corrected.

L231-232: “Given that the dynamic contribution to summer warming in WWCE is a manifestation of unforced internal variability...”. Is this 100 % certain statement? If yes, citation would be nice. Otherwise I’d probably add “mostly” or “mainly”.

This is not a 100% certain statement, hence we have adjusted the sentence as follows:

“Given that the dynamic contribution to summer warming in WWCE is mostly a manifestation of unforced internal variability, [...]”

Following the “classic” interpretation (nicely elaborated in, e.g., [Trenberth et al., 2015](#)), circulation changes often reflect internal variability rather than the response to external forcing (e.g., [Lehner et al., 2023](#)). As Vautard et al. (2023) elaborate in detail, whether this holds in the

case of Western Europe has profound implications, and it remains unclear whether the widespread underestimation of the dynamic contribution to the warming of TXx (annual maximum daily maximum temperature) in CMIP6 simulations “is due to uncaptured internal variability or missing (dynamical) forcing/processes”. We note here that for summer mean daily maximum temperatures, which are far more comparable to the summer mean temperatures we employ than TXx, Vautard et al. (2023) find that the dynamic contribution according to ERA5 is at the upper end yet within the range of CMIP6 simulations, corresponding well to our results (see, e.g., **Fig. 3c** or **SFig. 1**).

L270-271. What does it mean to "remove the background warming from the total thermodynamic warming"? I understood that Fig. 3d results all have the same background warming. But how can one remove background (global) warming contribution from the total thermodynamic warming? This could be elaborated in the manuscript as it seems to be a key step in the study.

As detailed in our response to question 2, we have adapted our analysis to no longer directly rely on the ‘regional’ thermodynamic warming, which was conceptualized as the difference of total thermodynamic and global background warming. Instead of removing the background warming, we simply proceed with the total thermodynamic warming obtained from the dynamical adjustment, and group our RCM simulations by driving GCMs with similar background warming (as is the case for the subset) or by individual GCMs (new **SFig. 5**).

L314-315. “since this is what each model simulation — and, contrary to the dynamic warming, especially the ensemble mean — should capture.” This may be a self-evident sentence for climate scientists, but I had to read the statement a couple of times, and I'm still not 100 % sure why each model simulation should capture the thermodynamic contribution. Maybe some elaboration would help?

We appreciate that this sentence has been flagged, it is indeed hard to read. It no longer exists in the revised manuscript. Based on other research now cited in the manuscript (and shown below), the response to external — or in this case simply anthropogenic — forcing mainly manifests in the thermodynamic rather than the dynamic warming, with the latter being more indicative of internal variability (which can dampen or amplify forced thermodynamic responses). We have adjusted the manuscript as follows:

“[...] Most CMIP6 simulations feature positive, but generally far weaker dynamic contributions [than ERA5], leading to an underestimation of the temperature increase. These findings are consistent with a recent study that suggests dynamics as the main “culprit” for discrepancies between observed and simulated trends in heat extremes³³. Given that dynamic warming in GCM simulations is primarily an expression of (unforced) internal variability rather than externally forced (e.g., Deser et al., 2016; Lehner et al., 2023)~~Given that the dynamic contribution to~~

~~summer warming in WWCE is a manifestation of unforced internal variability~~, dynamically inflicted differences between ensemble mean and observed temperature trends do not imply systematic model biases.”

L318-321: “We estimate ...” What’s regional thermodynamic bias? It’s not listed in Table 1. This is related to Minor comment 2.

We believe this is no longer applicable and kindly refer to the response to Minor comment 2.

Figures: Figure subtitles (a,b,c ...) are either badly placed or partly missing

We apologize for this mishap, which was caused by improper file conversions.

Fig. 2b: Does “hottest model” refer to the hottest model realization? When you only have two models in the CMIP6 subset, it might be more reasonable to call it realization than model.

Thank you for pointing this out; it is actually 3 unique models, MIROC-ES2L MPI-ESM1-2-LR and **GFDL-ESM4**, with the latter only contributing one initial condition ensemble member and (falsely) not shown in bold font in Table 4 of the originally submitted manuscript. We agree with the suggestion and have changed ‘hottest model’ to ‘hottest realization’ in Fig. 2b as well as Fig. 1c,d.

Supplementary information: Reference list seems to be missing.

Thanks for noticing, we have added the missing references.

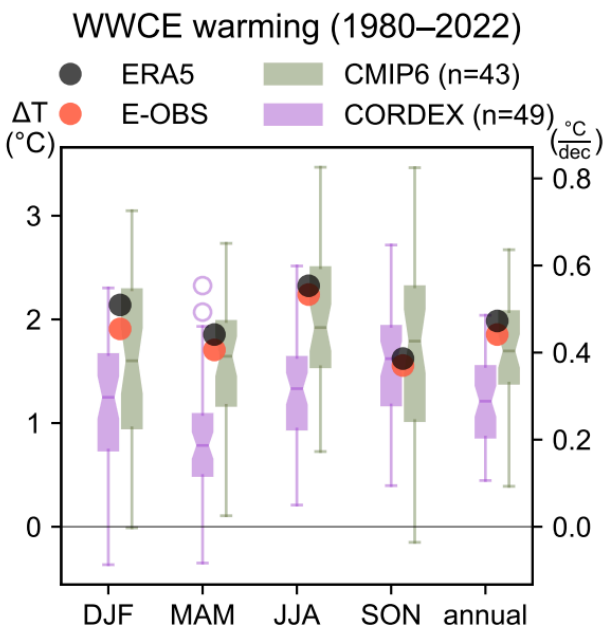
Reviewer #2 (Remarks to the Author):

Based on ERA5 reanalysis data, RCMs and GCMs, the study analyzed and quantified the impact of aerosol changes on warming in Western Europe. The results show some new phenomena. However, the following must be commented on or addressed before publishing.

We appreciate the reviewer’s questions, to which we reply individually below.

1. The temperature simulation has significant uncertainty at the regional scale, although the ERA5 reanalysis data performs well in reproducing the climate globally, it is still a model reanalysis and can not take the place of the benchmark dataset. It is recommended to add in situ data for comparative analysis.

E-OBS is already included in our analysis, but we have changed the marker to make it more visible in **Fig. 1e**:



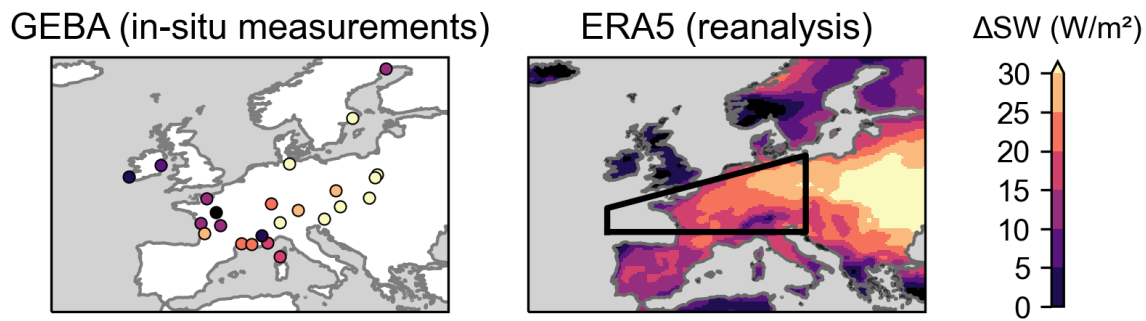
RFig. 2: Fig. 1e from the manuscript with an updated E-OBS marker.

While E-OBS temperatures could also be shown in Figs. 1c,d and Fig. 2b, we believe each of these subplots is already busy enough, and that it would not add much value since both ERA5 and E-OBS indicate WWCE summer warming well above 2 °C since 1980, thereby residing at the upper end of the simulated range by a large CORDEX ensemble.

Our analysis requires not just temperature, but also geopotential height at 500 hPa, and radiation data. For this reason, we only employ E-OBS to show that, although there are indeed slight differences between “observed” (though interpolated to the grid-scale, and not homogenized!) temperatures and those from reanalysis, especially in winter (Fig. 1e), we are abundantly certain that CORDEX simulations tend to strongly underestimate the summer warming in WWCE.

Concerning the ERA5-based finding that shortwave radiation trends in parts of Europe are very strong, i.e. in excess of 30 W/m² since 1980, we have extended our analysis and now include measurements from the Global Energy Balance Archive (GEBA) and show linear trends in incoming shortwave (or global) radiation from 1980 to 2021 in **RFig. 3**:

Summer global radiation changes across 1980–2021



RFig. 3: Changes between 1980 and 2021 in summer incoming shortwave radiation based on in-situ GEBA data and ERA5.

This figure is now referred to in the main text:

“Consequently, at the eastern edge of our analysis region, and particularly even further to the east, downward shortwave trends for 1980–2022 are underestimated more severely in simulations with constant aerosol forcing than in any other part of Europe (**Fig. 4c**). This finding is corroborated by ERA5 downward shortwave radiation trends, which also depict stronger increases towards the east, as well as in-situ measurements that reveal the same east–west gradient (SFig. 4).”

2. Another recent article (Li et al, 2022) pointed out that the model's underestimation/overestimation of regional temperature trends is mainly due to its underestimation of its response to external temperature forcing and its persistence. In my opinion, it also seems very reasonable. What are the substantive differences between this article and that one? Is there a certain correlation between the two?
(Ref: Li, Q., Sheng, B., Huang, J. et al. Different climate response persistence causes warming trend unevenness at continental scales. Nat. Clim. Chang. 12, 343–349 (2022). <https://doi.org/10.1038/s41558-022-01313-9>)

The reviewer points to an interesting study, which shows that the warming since 1900 has unfolded at markedly different rates in the United States (US) and China, with stronger warming in the latter. The authors show that, based on several large CMIP6 ensembles, most (but not all) models do not capture the extent to which the warming in the US is dampened compared to China. The authors then demonstrate that the climate in China responds in a more persistent manner to both external forcing and internal variability compared to the US, possibly due to differing roles of atmosphere–ocean interactions. This suggests that the warming in the US, primarily a response to external forcing, is more susceptible to be modulated by internal variability.

This is quite consistent with the recent work of Singh et al. (2023), who showed that for heat extremes on a near-weekly timescale, large-scale circulation changes have acted to dampen (enhance) the warming over the midwestern US (Western Europe). We emphasize that circulation changes may impact summer mean temperatures differently than heat extremes, although in the case of Western Europe, Vautard et al. (2023) have reported strong circulation contributions to both annual maximum temperature trends as well as summer mean maximum temperatures, similar to our analysis for summer mean temperatures. These circulation changes are typically interpreted as an expression of internal variability, but we cannot preclude that these changes (also) represent a response to external forcing.

In the study by Li et al. (2022), aerosols are considered to have similar estimated total effects on the warming in the US and China throughout the whole period since 1900, and therefore do not explain the differences in warming between these two regions. In our study, we do not compare warming across regions, but rather point out that since aerosol-related shortwave radiation changes are by design neglected in a majority of CMIP5-CORDEX simulations, even the thermodynamic warming in much of Europe since about 1980 is underestimated.

3. How is the boundary of the study area (western West-Central Europe) defined? This study mainly focuses on western West-Central Europe, but the title is Western Europe. It is recommended to change the title to western West-Central Europe as well.

We define western West-Central Europe as the western half of the official IPCC 'WCE' region, i.e. we longitudinally split this region, spanning from -10 to 40 °E, in half (such that it only covers -10 to 15 °E). We now also provide the resulting corner coordinates in the methods:

“The western corner coordinates of WWCE are identical to WCE (-10 °W, 45 or 48 °N, respectively), whereas the eastern corners are at 15 °W and 45 or 54.65 °N.”

Concerning the suggested title change, we believe that our analysis region is sufficiently consistent with typical definitions of “Western Europe”, and thus prefer to keep the title as simple as possible. Additionally, and this is now hopefully more clear in the revised paper, the warming underestimation due to neglected aerosol changes is really not restricted to a specific region, which is why we think that putting more emphasis on our analysis region would not be ideal.

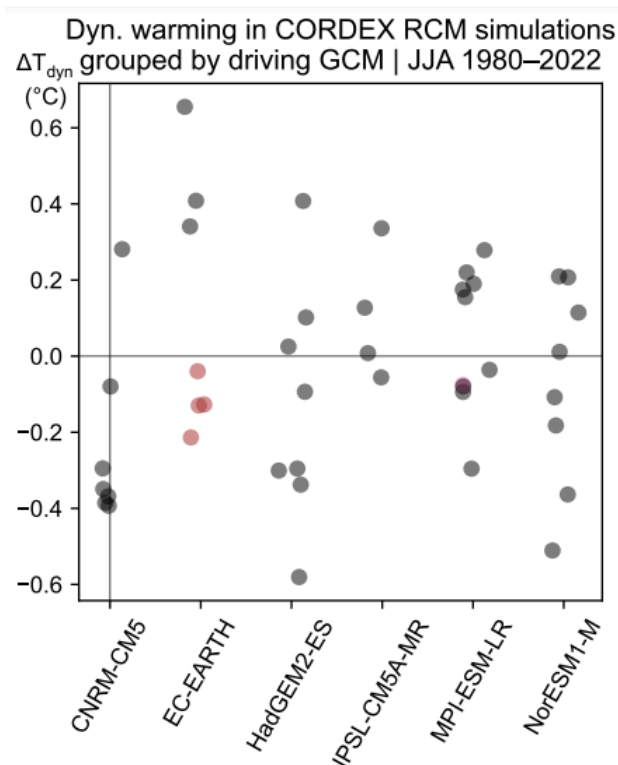
4. The east side outside the study area has a more pronounced difference between CORDEX RCMs evlo. and const. of the SW and temperature. It is recommended to include this in the study area.

We thank the reviewer for this question, which was also raised by reviewer 1, and hence refer to our response to their first question; in short, we extended our analysis for EWCE, the eastern half of WCE, but cannot apply the dynamical adjustment due to domain boundary limitations. The additional analysis is now mentioned in the main text:

“We emphasise that our analysis region does not include the areas with the strongest aerosol-related warming, as the estimated effect peaks east of WWCE. We hence repeated the disentangling of summer warming into dynamic and thermodynamic contributions based on ERA5 for eastern West-Central Europe (EWCE; displayed in SFig. 6), and found that dynamic contributions are nearly identical to WWCE, whereas the thermodynamic warming is clearly higher (2.0°C as opposed to 1.6 °C; SFig. 7). We note here that even CORDEX simulations with constant aerosols point to slightly higher warming in EWCE than WWCE, which might reflect a weaker oceanic influence in the eastern half of WCE, and implies that the thermodynamic warming difference between east and west evident for ERA5 should not be entirely attributed to the difference in aerosol-related warming. Still, according to the CORDEX simulations with evolving aerosols, and whether based on the subset (SFig. 7) or the full ensemble (SFig. 8), aerosol effects are about 0.15 °C stronger in EWCE than WWCE. Although these estimates are based on total temperature changes rather than thermodynamic warming, we conclude that even faster warming due to aerosol changes in eastern parts of West-Central Europe might cause (and at least contribute to) the east-west gradient in model–reanalysis warming discrepancies evident in Fig. 1b.”

This is also described in more detail at the bottom of this rebuttal in a separate section.

We additionally considered applying the dynamical adjustment to the driving CMIP5 GCMs rather than the RCMs themselves, as they essentially provide the large-scale circulation pattern through the domain boundaries. It has, however, been noted that the large-scale circulation in RCMs can substantially deviate from the driving GCMs, especially in summer (due to the weaker dynamical forcing from the GCMs compared to winter; [Vautard et al., 2020](#)). Grouping the dynamical contributions to WWCE summer warming by driving GCM, we find that some simulations cluster (**RFig. 4**): e.g., most of those driven by CNRM-CM5 indicate a similarly negative contribution, and the simulations for EC-Earth based on two initial condition ensemble members are either clearly positive (r1i1p1-driven) or slightly negative for (r12i1p1-driven). This is precisely the behavior we would expect if the circulation in RCM simulations is close to the respective driving GCM simulation, and shows that our dynamical adjustment approach — at least for these simulations — yields a plausible outcome. However, for other GCMs such as HadGEM2-ES and NorESM1-M, no such clustering is apparent, and the dynamic contributions have a comparatively wide range.



RFig. 4: Dynamical contributions to WWCE summer warming (1980–2022) in RCM simulations, grouped by several driving CMIP5 GCMs. Black = r1i1p1, red = r12i1p1, pink = r3i1p1, MPI-ESM-LR; single)

While this spread could, in theory, also originate from the dynamic adjustment itself, which provides merely estimates of these contributions, we consider it more likely that RCM simulations do not necessarily mirror the atmosphere’s internal variability (and hence multi-decadal circulation changes) from the driving GCMs. This conclusion is consistent with Vautard et al. (2020); not only does the topography vary drastically between RCM and driving GCM, but of course also the horizontal resolution, such that the RCMs may produce storms that are not found in GCMs themselves. Therefore, we refrain from quantifying the dynamic contributions from the CMIP5 driving GCMs as a proxy for the circulation effects in GCM–RCM simulations.

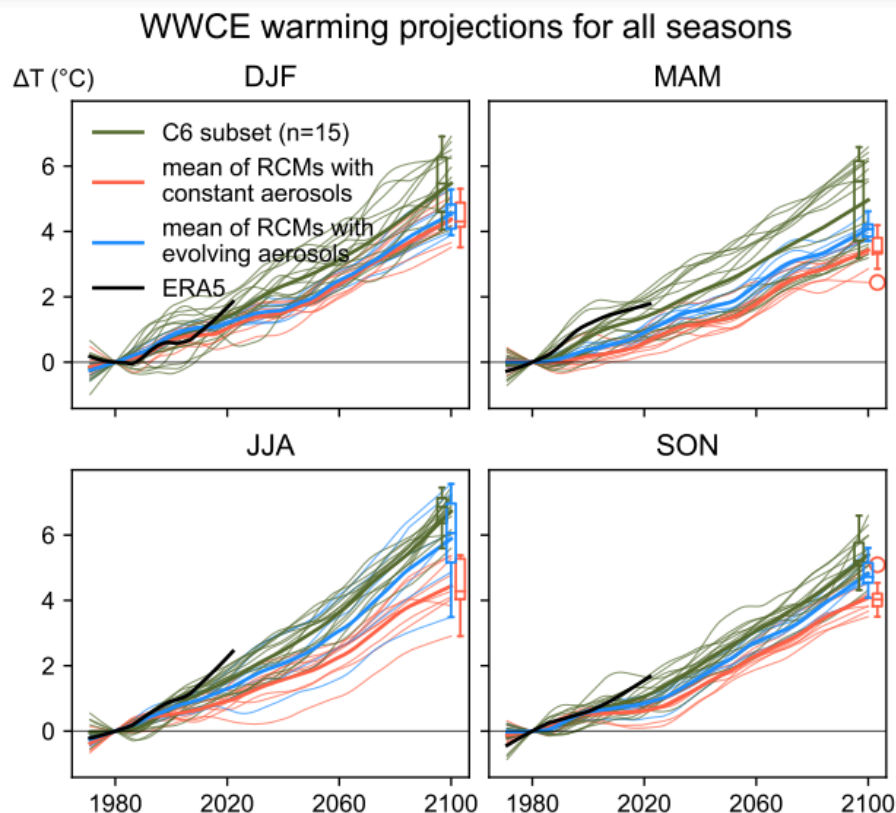
5. The influence of clouds and aerosols on downward shortwave radiation in CORDEX are expected to be given in Fig.S3.

Unfortunately, and unlike for CMIP6 data, the required variable for this (“rsdscs”, clear-sky shortwave downward radiation at the surface) is not available for the CORDEX simulations. While there are certain methods to estimate this variable using other model output, we do not believe that this is worth the effort as our analysis already demonstrates that only CORDEX

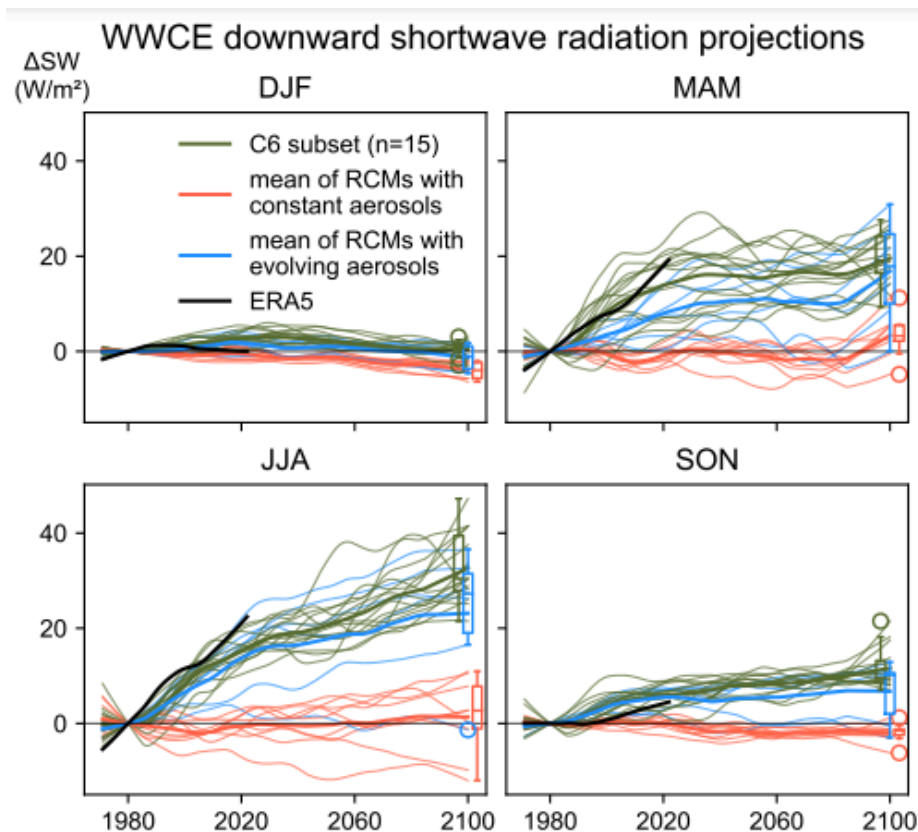
RCMs with evolving aerosols capture the strong increases in downward shortwave radiation evident in both ERA5 and CMIP6 data (**Fig. 4b**).

6. As shown in Fig.S5, the RCMs only underestimate warming significantly in the summer months of June, but also in other months (such as February and April). The authors should analyze these months with a large warming gap.

It is true that there are also discrepancies in other seasons or months, but we believe that separate, detailed investigations are required to fully understand why these differences between simulated and measured temperatures occur. For example, it was found that atmospheric internal variability is responsible for both the sudden winter warming in central Europe in the 1980s and the ensuing moderate cooling, (temporarily) masking the long-term thermodynamic warming ([Sippel et al., 2020](#)). In other words, the role of circulation changes strongly differs from summer, where the dynamic contributions to the warming in recent decades are clearly positive and arguably not as complex. Given our focus on the warming mismatch inflicted by neglecting changes in aerosol emissions, we present two new Supplementary Figures that provide information on both mean temperature and downward shortwave radiation for all four seasons (**RFigs. 5 & 6**). The effects of employing a time-invariant aerosol representation are most evident in summer when the insulation is strongest, but also emerge for fall and spring.



RFig. 5: like Fig. 5a, but for each season and also showing CMIP6 data.



RFig 6: Like RFig. 3, but for downward shortwave radiation instead.

We now mention this in the text, but refrain from shifting our focus to other seasons and repeating our full analysis for each of them (or even each month), since that might dilute our analysis:

“We repeat our analysis for all seasons, and find the strongest aerosol representation effects on shortwave radiation changes and hence also temperatures in summer, followed by spring and autumn (~ 0.7 °C discrepancy in 2100), whereas no clear differences emerge for winter. This is precisely the result we would expect considering the strong seasonal cycle of downward shortwave radiation in a region such as WWCE, with comparatively weak insolation in winter and short daytimes.”

7. While considering the effects of aerosols and clouds, it is also necessary to consider the effects of greenhouse gases.

All simulations we employ consider greenhouse gases: the 1980–2022 period is either based on historical (up to 2005) and RCP85-based simulations in the case of RCM simulations, or historical up to 2015 for the CMIP6 GCMs followed by SSP585. For both the CMIP5

GCM-driven RCM simulations and the CMIP6 GCM ones, we thus rely on historical emissions combined with a high-emission scenario to enable better comparability. But since we don't attribute any of the warming mismatch with respect to observations to issues with the greenhouse gas forcing or the global response to said forcing, but rather a lack of aerosol forcing within the RCM domain (due to assuming constant aerosols), this is what we focus on and discuss at length in the manuscript.

8. There seem to be some issues with the annotation of subfigures(a,b,c,d,...) that need improvement.

Thank you for phrasing it in such a diplomatic manner; the annotation was messed up when converting to .docx, we will avoid this for future submissions.

Reviewer #3 (Remarks to the Author):

The manuscript addresses the issue of underestimation in the historical warming trend within Western Europe, particularly during the summer season, as projected by CMIP6 GCMs and EURO-CORDEX RCMs. The authors commence by characterizing the warming disparity within both ensembles and subsequently endeavor to discern the primary underlying causes. They conclude their analysis by evaluating how these identified factors may also introduce biases in future warming projections within the EURO-CORDEX ensemble, encompassing both mean climate and temperature extremes.

The subject matter of this study holds profound scientific relevance for a broad readership. Western Europe has witnessed a significantly accelerated rate of warming compared to other global regions. Furthermore, climate change adaptation measures are presently being formulated and implemented based on EURO-CORDEX datasets, disseminated through various climate service data portals at European, national, and local levels. Thus, comprehending and reporting on EURO-CORDEX's historical and future warming underestimations in Western Europe assumes paramount importance.

Additionally, I wish to acknowledge the authors for their clear and well-organized presentation in the manuscript. The introduction provides a solid foundation by succinctly outlining the research objectives. The methodology is robust and effectively executed, and the results provide new insights of relevance to a significant scientific community.

However, I do have some general concerns that, in my opinion, fall within the scope of second-order issues that can be addressed during the revision process. Specifically, I recommend the following improvements:

1. Utilize existing references more effectively, not only in the introduction but also throughout the manuscript.
2. Evaluate the performance of the evaluation runs generated within EURO-CORDEX.
3. Enhance the readability of Table 1.
4. Improve the clarity and informativeness of the abstract.
5. Consider providing supplementary annexes with simulation-specific information to aid the broader user community.

I firmly believe that these general concerns, while essential for enhancing the content and overall relevance of the manuscript, can be adequately addressed by the authors. Therefore, I recommend the publication of this manuscript upon incorporating these recommendations. I am willing to reassess the revised version of the manuscript after the modifications have been made.

We would like to express our appreciation not only for the reviewer's positive assessment, but also for their detailed, clearly structured and exceptionally helpful review. Based on these and the other reviewers' comments, we have partly re-written our manuscript and performed several additional analyses, and once again respond individually to each suggestion or comment below.

1. Below are the detailed descriptions of my main concerns:

1.1. Make better use of existing references

1.1.1. I have noticed instances in the introduction where the use of references may be further optimized to substantiate certain statements. Given my background in the RCM community, I would suggest considering alternative references, particularly in lines 36, 37, and 38. Notably, comprehensive multi-model studies with CPRCMs are now available and could enhance the support for your arguments.

We thank the reviewer for the suggestion to refer to multi-model CPRCM studies, and have adjusted the text as follows:

“Dynamical downscaling, that is, running a regional climate model (RCM) for a limited domain with coarser GCM data at the boundaries, has been established as an important modeling branch as it entails more realistic regional climate representations⁵, although the benefits depend on the variable and domain of interest⁶ [Ban et al., 2021, Soares et al., 2022, Sangelatoni et al., 2023](#). ~~Even though RCMs can be run at~~ [The first multi-model RCM experiments at convection-permitting km-scale resolution are already available \(Ban et al., 2021; Soares et al., 2022\)](#), ~~yet~~ this is typically only performed for short periods (e.g., ref. [7, Ban et al., 2021; Soares et al., 2022](#)) due to the immense computational demands.”

1.1.2. It appears that there are at least three references missing in the introduction: Ribes et al. 2022, around line 52; Nabat et al. 2014, mentioned at lines 75-76 (previously cited as #40); and Lorenz and Jacob 2010 (see the reference list at the end of this review)

We have added these three references, including the following edit to already mention Nabat et al. (2014) in the introduction:

“Therefore, [as first suggested by an analysis based on a single RCM for the Euro-Mediterranean domain \(Nabat et al., 2014\)](#), different aerosol representations in CORDEX simulations might also contribute to the 1980–2022 warming discrepancy with respect to observations, which we aim to analyze [further](#) in this study.”

Thank you for providing these specific reference suggestions.

1.1.3. I also recommend that the authors review the recent work by Taranu et al. (2023) regarding the causes of GCM-RCM inconsistency in EURO-CORDEX future projections in Europe during the summer. This article, being very recent, may not have been on the authors' radar, but it provides a valuable alternative approach to characterizing and disentangling the sources of discrepancy.

Thank you for bringing this article to our attention, it is now mentioned twice in the manuscript:

“Besides these studies, model experiments with a single RCM suggest that nearly a quarter of the simulated 1980–2012 annual mean warming in Europe is achieved by aerosol changes⁴⁰.

Recently, Taranu et al. (2023) presented a set of experiments based on a specific GCM–RCM pair that can share the same grid and utilize identical model physics and forcings, designed to pinpoint the causes behind discrepancies in RCM and GCM simulations. They found that different aerosol representations in RCMs strongly affect shortwave radiation and temperature projections, with the same east–west gradient across Europe emerging as in our analysis (cf. their Fig. 10) and similar results despite obvious methodological differences.”

[...]

“We emphasise that our estimates are based on subgroups consisting of RCMs with different model physics; for example, most RCMs do not consider plant physiological responses to rising CO₂ (i.e., closing stomata and hence reducing transpiration due to increased water use efficiency), which could lead to a general underestimation of the rise in future heat extremes (Schwingshackl et al., 2019). ALADIN63 with time-evolving aerosols does represent physiological responses, however, which could distort our analysis, but on the other hand and for this specific RCM, physiological effects do not exert strong control on projected summer temperature changes in WWCE (Taranu et al., 2023).”

1.1.4. Additionally, I suggest revisiting the use of reference #23 in the introduction in light of the comment on PubPeer, which may warrant a modification of the use of this reference. (PubPeer link: <https://pubpeer.com/publications/2F852866CC945DB4615D4FC40105FC#1>)

Reference #23 in the initial submission, i.e. Bartok et al. (2017), compare projected shortwave radiation and temperature changes over Europe from CORDEX RCM simulations to their driving GCMs. One of the main messages from the paper is that RCMs should include time dependent aerosols to depict more realistic surface solar radiation changes and projections, since only the GCMs (with evolving aerosols) manage to reproduce the observed historical brightening. The authors were not aware that the ALADIN RCM, even for an earlier version than the one employed here, did in fact already include time dependent aerosols, which would have only

strengthened their conclusion. For this reason, we still cite reference 23 in the introduction, but solely rely on the work of Gutiérrez et al. (2020) or originally reference 24 when we state that most RCMs employ time invariant aerosols:

“The CORDEX multimodel mean has, however, been found to underestimate the observed increase in downward shortwave radiation²³, since most RCMs do not consider time-evolving aerosols^{23,24}.”

1.2. Verify the behavior of the evaluation runs produced in EURO-CORDEX

While the authors have thoroughly examined the past trends in the historical runs of EURO-CORDEX, where RCMs are driven by GCMs with random natural variability, I would like to draw attention to another category within EURO-CORDEX known as "evaluation runs". In these runs, RCMs are driven by reanalyses, such as ERA-Interim in the earlier CORDEX phase or ERA5 in the current phase. Investigating the warming trend in these evaluation runs could provide complementary insights. It is plausible that the evaluation runs may also exhibit underestimation of past trends, which would further corroborate the role of aerosols in most RCMs. Should this be the case, a brief annex or an entry in Table 1 may suffice to convey this information.

We highly welcome the suggestion to examine past trends in reanalysis-driven RCM simulations. Unfortunately, we could not find any ERA5-driven CORDEX EUR-11 simulations on ESGF nodes. Instead, we worked with ERA-Interim driven simulations, which typically only cover about 30 years or less.

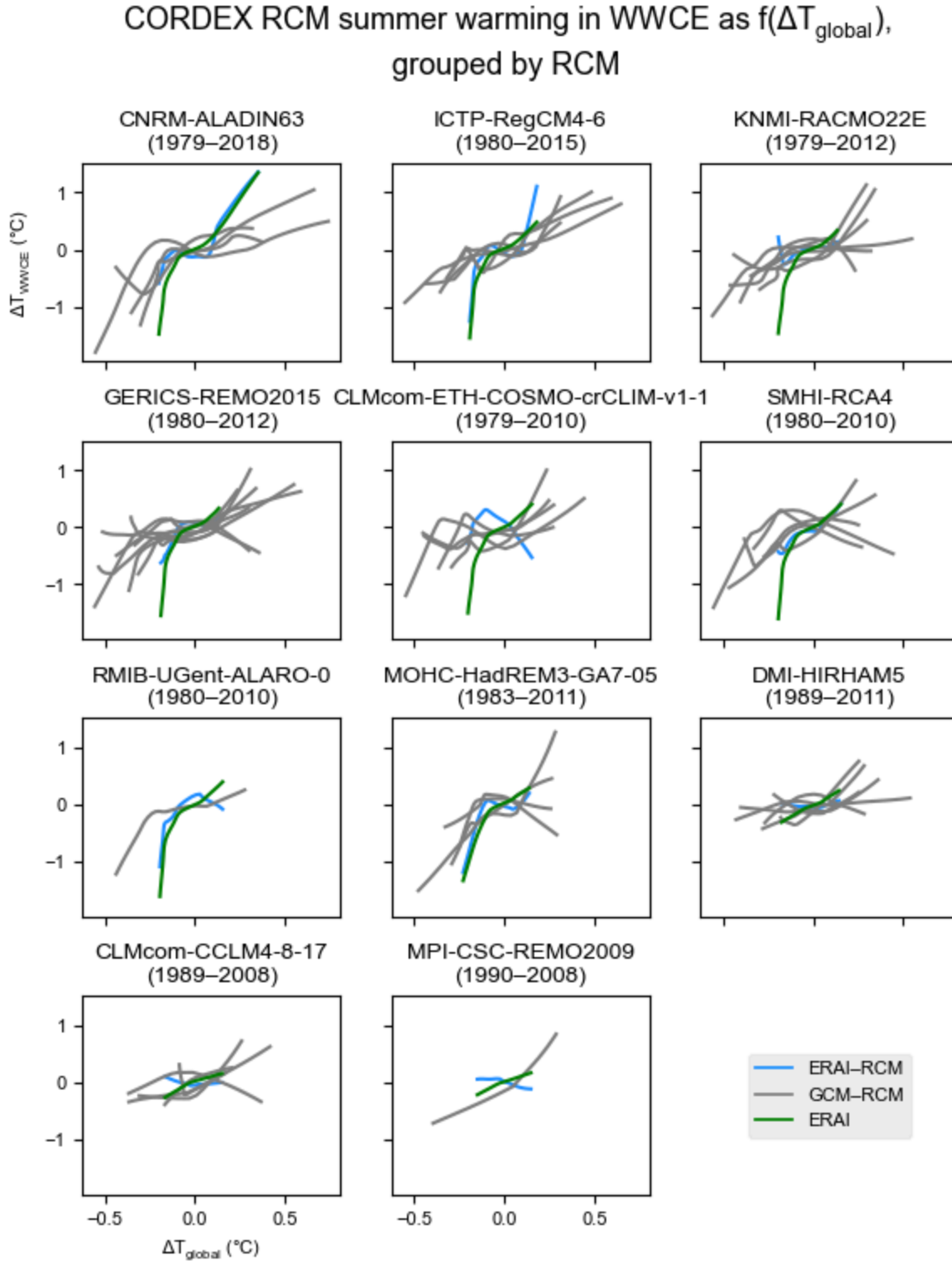
The different simulation lengths complicate our analysis, hence we always use as much data as available in the respective ERA-Interim driven simulation for each RCM when we construct **RFig 7**. This figure shows that ERA-Interim generally features stronger WWCE summer warming than the ERA-Interim driven RCMs, as we would expect since most of those RCMs do not represent time-varying aerosols. Curiously, the only ERA-Interim driven simulation that warms *more* than ERA-Interim itself is based on RegCM4-6 with constant aerosols. HadREM3-GA7-05 with evolving aerosols comes very close to the warming of ERA-Interim itself, and so does ALADIN63 except for the very first years in the 1980s with extremely rapid warming in ERA-Interim. RACMO22-E is the third RCM with time-varying aerosols, but its warming trajectory, similarly to the ETH-COSMO model, does not closely follow ERA-Interim. For the ETH-COSMO model, there is in fact even a *net cooling* across 1979 to 2010 when driven by ERA-Interim.

This is rather surprising given that, as shown in **RFig. 8**, the driving GCM typically manages to dictate the shape of the warming trajectory in the global-regional warming space. The work of Tanaru et al. (2023) suggests that dynamical inconsistencies between specific GCM-RCM pairs

are at least in part caused by structural differences, e.g. different atmospheric physics, so perhaps driving the ETH-COSMO model with ERA-Interim results in a particularly strong incoherence between the driving model and the RCM's large-scale circulation, and similarly for RACMO22-E.

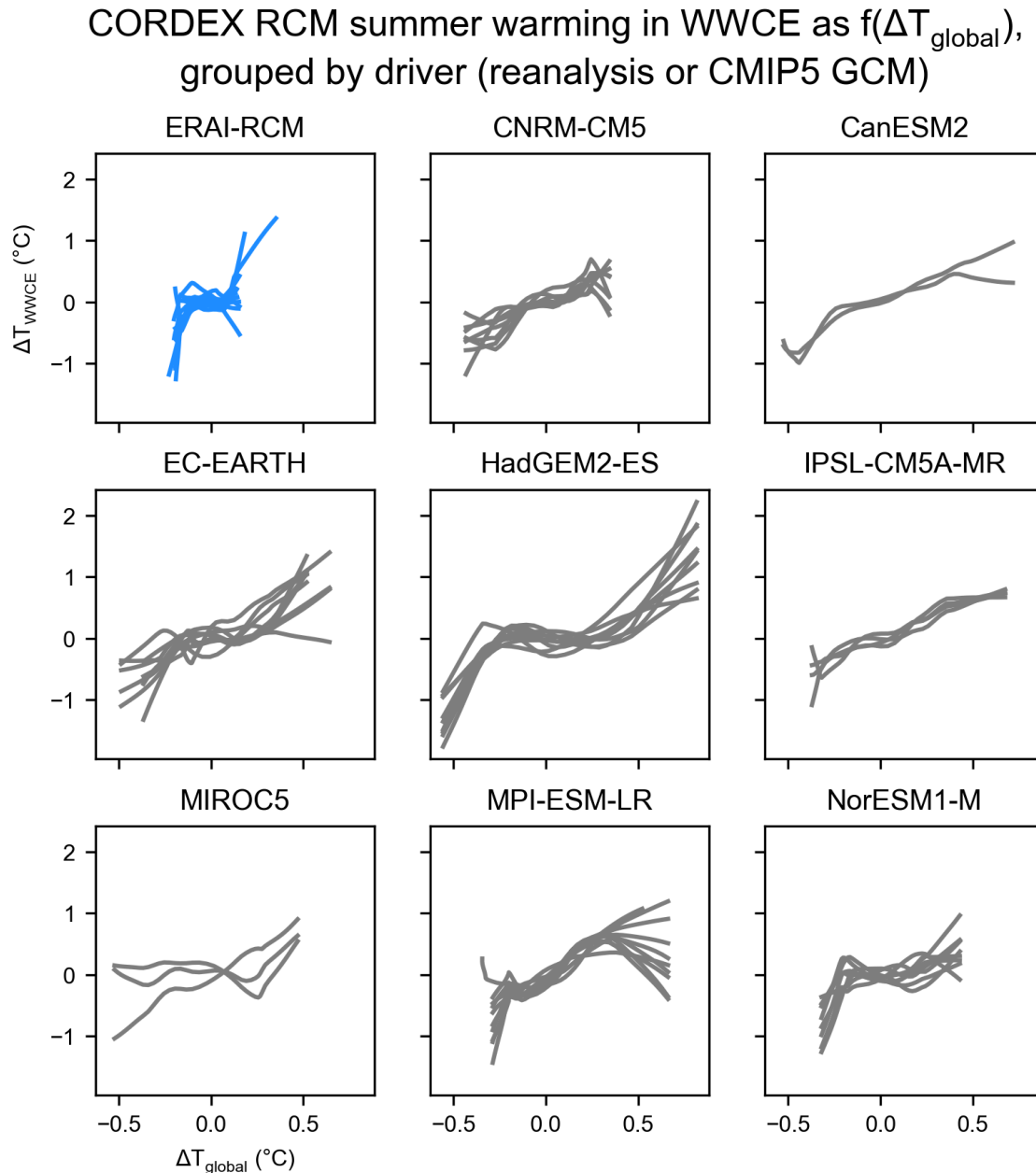
Therefore, even though we cannot explain some of the behavior evident in **RFig. 7** and it is difficult to quantify the mismatch due to the limited length of the timeseries, the overall pattern of weaker simulated temperature trends in RCMs forced by ERA-Interim compared to ERA-Interim itself is consistent with our main analysis. We have amended the manuscript as follows:

“This further substantiates that neglecting time-evolving aerosols causes a lack of regional thermodynamic forcing (and resulting warming response). As an additional line of evidence, we also examined CORDEX RCM simulations driven by ERA-Interim, the predecessor of ERA5, and found that all RCMs with constant aerosols except for RegCM4-6 underestimate the WWCE summer warming compared to ERA-Interim itself (SFig. 3).”



RFig. 7: CORDEX RCM summer warming in WWCE for ERA-Interim- and GCM-driven simulations as a function of global warming, with the latter taken from ERA-Interim or the driving GCM, respectively. Temperatures obtained directly from ERA-Interim are also shown as green lines. All temperatures are processed with a 19-year LOWESS filter to remove

high-frequency variability, and expressed as anomalies with respect to the mean calculated over 1990–2010 to facilitate the comparison. Note that here, more RCMs than used for analysis are displayed, since this figure is solely based on temperature and does not require any further variables (which are missing for some RCMs). RCMs are sorted by the respective duration of the ERA-Interim forced simulation.



RFig 8: As RFig. 7, but grouped by the driving model instead. Showing as much available data as possible for ERA-Interim as a driver (at most 1979–2018), and 1980–2020 for all GCM–RCM simulations.

1.3. Improve the readability of Table 1

Table 1 contains crucial data summarizing the study, which holds significant value for future research endeavors. However, its current presentation is not as user-friendly as desired. I propose a clearer organization that distinctly conveys four primary pieces of information: Total warming, Background warming, Dynamic, and Local Thermodynamic. It remains unclear why Background warming is not provided for ERA5.

Table 1 has been simplified as suggested; since we no longer subtract the global background warming from the thermodynamic warming to arrive at the regional thermodynamic warming (see, e.g., our response to the second question of reviewer #1), we show the total warming, disentangled into dynamic and thermodynamic warming, and in addition the background warming as well as estimated aerosol effects.

As for the original Table 1, it did not contain any background warming for ERA5 because that column was actually defined as ‘biases’ in background warming compared to ERA5, and hence was empty by definition. We acknowledge that our first attempt at Table 1 was unnecessarily confusing, and hope the updated version is more clear.

1.4. Enhance the abstract

Overall, the paper is well-structured, comprehensible, and communicates its implications effectively, particularly in the conclusion section. Paradoxically, I find that the abstract falls slightly short of maintaining the same level of clarity and emphasis. I recommend revising the abstract to enhance its readability and emphasize the key messages.

We appreciate the suggestion and hope that the revised abstract is more clear than our initial attempt:

“In much of western and central Europe, summer temperatures have surged three times faster than the global mean warming since 1980. Particularly regional climate models (RCMs), and to a lesser extent global climate models (GCMs) underestimate this strong observed warming of ~2.3 °C. Our analysis reveals weaker-than-observed circulation-related contributions in most model simulations, but this only partly accounts for the warming discrepancy. Crucially, a majority of RCMs from the Coordinated Regional Downscaling Experiment (CORDEX) additionally underestimates the thermodynamic warming. We find that this is mainly caused by constant aerosol representations in most CORDEX RCMs, hence missing the regional brightening and associated warming due to aerosol reductions in Europe. We estimate an effect of ~0.5 °C over western and central Europe, although this depends on the GCM–RCM ensemble subset, and find even stronger aerosol-related warming further east. Heatwave intensity changes since 1980 are already underestimated by RCMs with constant aerosols by ~1 °C, and the warming

discrepancies relative to the GCMs become even larger in projections. To better reap the benefits of high-resolution RCM simulations, it is thus imperative to represent the relevant external forcings and associated responses in the entire GCM–RCM modelling chain.”

1.5. Assist the user community by adding annexes with simulation-specific information

The potential impact of this study extends to a broad community, including climate modelers, climate service providers, and climate data users, such as the impact community and the end-users. To facilitate these communities' effective utilization of your findings, I propose the addition of an annex that characterizes individual CMIP6 and CORDEX (and possibly CMIP5) simulations. This could resemble a simplified version of Table 1, providing details for each model simulation. Such an annex would enable climate data users to select EURO-CORDEX simulations post hoc and help the regional climate modeling community select adequate driving GCMs for the CORDEX-CMIP6 initiative. In particular, in CORDEX-CMIP5, six out of 49 simulations are notable for capturing the appropriate warming background due to their driving GCMs and the evolving aerosol dynamics. Listing these simulations would be valuable.

We agree that it could be useful to provide details for the entire CORDEX ensemble used in the study, and hence we have added an extended version of Table 1 that lists WWCE temperature trends as well as disentangled dynamic and thermodynamic contributions for each GCM–RCM model chain (**STable 1**). While we no longer explicitly list biases with respect to ERA5 both in Table 1 and the new STable 1, all the required data is provided such that users can extract the information of interest.

2. Minor comments:

2.1. Title: The **word "Summer" is absent from the title**, which does not fully reflect the content of the paper.

Although the other seasons now receive slightly more attention compared to the initial submission, we agree that mentioning ‘summer’ in the title makes sense. Thank you for pointing this out.

2.2. Line 41: "RCMS" should be corrected to "RCMs."

Thanks for noticing this typo, it has been fixed.

2.3. Line 42: It should be noted that coordination of RCM simulations existed before CORDEX since the 1990s. What distinguishes CORDEX is its "worldwide" coordination.

We hope that our slightly revised sentence makes this more clear:

“While RCMs were not widely used for the IPCC AR5 (ref. 13) as worldwide coordinated intercomparison studies were still emerging, in the IPCC AR6, RCMs from the Coordinated Regional Downscaling Experiment (CORDEX; refs. 14–16) provide the basis for sub-continental climate information (e.g., Chapters 11, 12 and Atlas).”

2.4. Lines 59-60: The impact of changes in land-sea contrast and regional lapse rate, particularly during the summer in Europe, is worth mentioning.

We now mention the link to regional lapse rate changes at the suggested location in the manuscript:

“Regional warming is the consequence of (i) the “background” global warming, (ii) large-scale circulation changes, often conceptualized as “dynamics”, and (iii), regional changes in the surface energy budget and partitioning, including snow/ice–albedo–temperature and soil moisture/vegetation–temperature feedbacks (“thermodynamics”). We note here that the (projected) near-surface summer temperature rise is constrained by the vertical warming structure, which enables enhanced near-surface warming over most extratropical land areas due to moisture limitations (Brogli et al., 2021).”

Since variations in lapse rate changes also explain land–sea warming contrasts, we opted to focus on the former in the additional sentence. We hope that our distinction of physical warming drivers and constraints (such as these lapse rate changes) is helpful for the readers despite not going into details, as we do not want to distract from the main story.

2.5. Lines 99-100: Is it feasible to use GCM large ensembles to distinguish human-induced changes from those related to unforced natural variability? This approach could offer valuable insights for interpretation (Line 312) and might be worth including in Table 1 as supplementary information.

We begin our response with a quote from Vautard *et al.*, 2023, who investigated thermodynamic and dynamic contributions to (changes in) heat extremes over a comparable domain:

“In the 170 available simulations from 32 different models that we analyzed, including 3 large model ensembles, none have a circulation-induced heat trend as large as observed. **This can be due to underestimated circulation response to external forcing, or to a systematic underestimation of low-frequency variability, or both.** The former implies that future projections are too conservative, the latter that we are left with deep uncertainty regarding the pace of future summer heat in Europe. This calls for caution when interpreting climate projections of heat extremes over Western Europe, in view of adaptation to heat waves.”

That is, the finding that GCM simulations tend to underestimate the dynamic contributions (to heat extremes in Western Europe) does not inform us on the nature of this mismatch. It is, however, interesting to note that overall (i.e., in terms of ensemble means), both the CESM2 large ensemble and our CMIP6 multi-model ensemble point to weaker than “observed”, yet still positive circulation contributions to the summer mean warming, which we would not expect if circulation changes in the GCMs were solely governed by internal variability (in this case, averages calculated over a sufficiently high number of initial condition ensemble members should be near zero). But what is a sufficiently high number, and do the GCMs really have an issue with the dynamic contributions? In our analysis of mean temperature changes, and contrary to Vautard *et al.* (2023), we find that the “observed” warming attributable to circulation changes is, though at the upper end, well within the range of GCM simulations.

Due to these difficulties, we focus on thermodynamic biases that can clearly be attributed to simplified aerosol representations in our study, and believe that further research well beyond our scope is required to adequately address the reviewer’s question.

2.6. Figure 1b: Noticeable differences along the coastline in Fig. 1b may be linked to an issue with the land-sea mask in handling the ERA5 dataset, where mixed land-sea points cover a larger area than in 12km CORDEX runs. It would be helpful to clarify how mixed land-sea points in ERA5 were addressed.

Thank you for pointing out that this information was missing; we have extended the description of ERA5 as follows:

“[...]. We use ERA5 as a reference for climate models and to estimate the dynamic and thermodynamic warming contributions, including 2-meter temperature, radiation and 500-hPa geopotential height data from 1950 to 2022 at a horizontal resolution of 0.25°. Only grid cells that consist of at least 50% land are considered. This choice does not noticeably impact our analysis of WWCE averages, but can introduce slight inconsistencies in the vicinity of coastlines when comparing against re-gridded CORDEX simulations (Fig. 1b).”

2.7. Figure 1 caption: It should be clarified that the curves and numbers in Fig. 1 are computed only over land points.

We have edited the caption to hopefully make this more clear:

“In both a.) and b.), the analysis region — the western half of West–Central Europe (WWCE) — is delineated by a black contour. For Calculated using all land areas within this domain, c and d depict smoothed regional average summer temperatures ~~over land~~, comparing ERA5 (black line) to CORDEX and CMIP6 simulations (colored range and lines), respectively.”

2.8. Line 134: Please specify at least once which dataset is used in this paper, considering that various CORDEX datasets exist over the European domain (e.g., 50 km resolution, Med-CORDEX, evaluation runs).

Thank you for the suggestion, we have edited the sentence as follows:

“Based on a state-of-the-art 49-member RCM simulation ensemble from the CORDEX EUR-11 experiment at 0.11 degrees or about 12.5 km horizontal resolution, however, [...]”

2.9. Lines 146-149: Over recent years, CMIP5 (hence CORDEX) and CMIP6 differ in terms of GHG concentration and aerosol forcings. The historical period is shorter in CMIP5, which introduces a source of error, potentially small. This difference should be acknowledged.

We agree that mentioning this potential source of discrepancies makes sense, and have edited the text:

“The CMIP6 multi-model ensemble (green shading in **Fig. 1d**) also features weaker than observed trends in most simulations, yet the upper warming range is considerably higher compared to CORDEX, and the overall distribution actually includes the ERA5 data, thus not necessarily implying a discrepancy. We remark here that we do not expect a perfect correspondence between WWCE warming in CMIP5 GCM-driven RCM simulations and CMIP6 GCMs, for example because the latter employ a longer historical period in CMIP6 compared to CMIP5 (up to 2015 rather than 2005) and different aerosol and greenhouse gas emission scenarios. Nevertheless, this does not explain the widespread lack of WWCE warming in The CORDEX simulations, which also feature less warming in winter and spring than observation-derived datasets, whereas linear trends in simulated and observed fall temperature are fairly consistent (**Fig. 1e**).”

2.10. Figure 2: Please provide a clear means for the reader to identify which simulations are considered "good," meaning those with favorable DeltaT_{glo} and DeltaT-WWCE.

We agree that a bit more guidance is helpful for the interpretation of Fig. 2 (particularly 2a), and have amended the text:

“The background warming is indicated by linear trends in global annual mean 2m-temperatures from the corresponding driving GCM of each CORDEX simulation, and visualized together with WWCE summer warming since 1980 (purple dots in **Fig. 2a**). Ideally, we would find most model simulations near ERA5 (black marker) with respect to both dimensions, i.e. the global background warming as well as the WWCE summer temperature rise. Within the CORDEX ensemble used here is, forced by 8 individual GCMs, that is, there are only 8 independent realizations of background warming. Most of these GCMs, most models clearly exceed ERA5’s 1980–2022 global mean temperature increase of 0.8 °C.”

2.11. Please specify the aerosol dataset used in ERA5 and its reliability. Has this dataset been evaluated in a published paper? Is ERA5 considered reliable for simulating climate-aerosol interactions (direct, semi-direct, indirect effects)? This information is critical for interpreting the results.

To address this question, we delved deep into the available publications and technical model descriptions. While the Hersbach et al. (2020) paper describing ERA5 is vague about how exactly the aerosol forcing from CMIP5 is employed, the procedure seems to be identical to ERA 20-CM, for which only long-term changes in anthropogenic sulphate emissions are considered (in addition to natural volcanic forcings). All other aerosol species are based on respective climatologies.

We have added a paragraph to the description of ERA5 in the Methods to make it clear that ERA5 only aims to capture long-term changes in the “dominant” anthropogenic aerosol species, and therefore should not be seen as a perfect reference.

“ERA5 uses sulfate aerosol forcings from CMIP5 (Hersbach et al., 2020), in turn based on a reconstructed aerosol emission inventory (Lamarque et al., 2010) up to 2009, and thereafter using emission scenarios (specifically RCP2.6; Herbach et al., 2015). In other words, long-term changes in anthropogenic aerosols are only considered for the dominant species (Hersbach et al., 2015), and to the author’s knowledge, only direct aerosol effects are represented (i.e., the scattering and absorption of radiation). On the one hand, this implies that ERA5 only partly captures the impacts of anthropogenic aerosol emissions, but on the other hand, this simplifies our comparison to the CMIP5 GCM-driven RCM simulations, since none of the RCMs with evolving aerosols represents all indirect effects either (i.e., cloud albedo and interactions between aerosols and cloud microphysics).”

Our comparison of ERA5’s shortwave radiation trends to in-situ data indicates that strong increases since 1980 on the order of 20–30 W/m² and beyond in certain parts of Europe are plausible (see also our response to the first question of Reviewer #2). Even though the simplifications employed to produce ERA5 might introduce biases, the product thus still captures the well-established regional brightening in Europe — unlike CORDEX simulations with constant aerosol representations.

2.12. Figure 3a: It might be prudent to reconsider the utility of Fig. 3a, as it appears challenging to create a comprehensive and easily readable representation. For example, notable factors such as aerosol effects and land-sea contrast are omitted. It may be worth considering its removal.

We believe that we did not make it sufficiently clear what Fig. 3a attempts to achieve. To this end, we have edited the figure caption as follows:

“a Schematic representation of our analysis region, western West-Central Europe, and the drivers of regional warming, depicting: the (human-induced) global mean temperature increase that provides a ‘background’ warming to the region, the presence of a regional thermodynamic forcing (primarily due to increasing net radiation) and the resulting warming response, and dynamic changes. Note that this schematic focuses on the direct physical drivers of regional warming, which can in turn be affected by other processes and interactions; e.g., land cover changes may affect the surface energy partitioning and/or the surface albedo and hence ΔSW_{net} and aerosols primarily modulate the attenuation of shortwave radiation (and hence also ΔSW_{net}).”

We agree with the reviewer that it would be difficult to add *all* the relevant factors, but since this schematic is mainly a graphic representation of the conceptualization of physical warming drivers in the introduction, we think that it could still be helpful for the broad target audience of Nature Communications Earth & Environment.

2.13. Lines 236-237: The statement suggests that GCM selection in CORDEX phase 1 was not ideal. It would be beneficial to provide recommendations for improving the situation in CORDEX-CMIP6 in the conclusion.

In these and the following lines, we refer to the fact that our analysis points to a reduced range in dynamic contributions to WWCE summer warming since 1980 compared to CMIP6 GCM simulations, and that this is to be expected given that, at least in theory, the driving GCM largely prescribes the synoptic weather patterns within the domain (although reality is more complex, as nicely described by Tararu et al., 2023). Given the uncertainty revolving around whether these dynamic contributions are in part or even entirely due to either a circulation response to external forcing or internal variability, we are not sure if it would make sense to try and select the driving GCMs based on dynamic warming for a specific region.

We are convinced, however, that it would make sense to ensure realistic historic thermodynamic warming patterns in the driving GCMs; based on Di Virgilio *et al.* (2022), the main focus with respect to simulated climate change signals is to obtain a wide range of projections. We now mention this in our conclusion.

2.14. Figure 4: The caption indicates 18 CORDEX simulations, while 15 is mentioned in the text (Line 239), and Table 3 lists only 14. Please clarify this discrepancy.

Once again we thank the reviewer for spotting a mistake in the figure caption describing the subset with near-ERA5 background warming: 18 is a relic from earlier analysis without dynamical adjustment, for which we do not always have the required geopotential height data and finally only use 15 simulations. This mistake has been corrected.

As for table 3, we cannot identify any errors. There are indeed 14 simulations with time-evolving aerosols, but only 6 of those are part of the model subset (the remaining 9 simulations use constant aerosols). To find out which simulations are actually part of the subset, it is helpful to additionally consider Table 2 that highlights the two driving GCMs of the subset (IDs 2 and 8). Counting the occurrences of IDs 2 and 8 in Table 4 then yields a total of 15 simulations used for the subset, consistent with the text and updated figure caption.

2.15. Table 3: The table lists 43 simulations, whereas 49 are mentioned in Lines 133 and 332. Please ensure consistency.

We obtain a total of 49 simulations when summing up the numbers of individual simulations per RCM in Table 4, which has not changed since the initial submission. We do use 43 CMIP6 simulations, however, which are listed in Table 4; perhaps this explains the suspected inconsistency?

2.16. Line 376, Figure 5 caption: "Linear trends" is mentioned, while the unit in Fig. 5b is °C. Ensure consistency.

We have edited the caption accordingly, thank you for noticing:

“Long term changes ~~linear trends~~ of the hottest n-day periods in each year between 1980–2022 and 1980–2099 in the upper and lower panel for ERA5 and the RCM simulations, determined by fitting a linear trend and calculating the total temperature change across the respective time period.”

2.17. Lines 388-389: It is stated that "authors assessed surface rather than air temperature changes." Please verify this statement, as the "tas" variable (near-surface air temperature) is used in Boé et al. 2020, similar to your study.

Thanks for catching this mistake, we have corrected the sentence.

2.18. Line 410-411: The study of the correlation between aerosol conditions and cloud conditions over Europe in Nabat et al. (2020) may provide valuable insights for your interpretation.

We greatly appreciate this suggested reference, as it provides a source for the suggested dependence of the aerosol effect on cloudiness, and further emphasizes the role of variable aerosol loads. We now mention this in the text:

“We argue that aerosol representation-inflicted biases should emerge most clearly in cloud-free conditions, which are closely related to the large-scale circulation patterns — typically

atmospheric blockings — that enable heatwaves in WWCE (e.g., ref. 42). In fact, enhanced direct aerosol effects during blocking conditions in Europe have been linked both to reduced cloud cover and increased aerosol load (Nabat et al., 2020).»

2.19. Line 443: You may consider replacing this text with the equivalent statement in IPCC-AR6 WGI, where RCMs are deeply assessed. For example, Chapter 10, Executive Summary states, « The performance of global and regional climate models and their fitness for future projections depend on their representation of relevant processes, forcings and drivers and on the specific context. [...] Including the relevant forcings (e.g., aerosols,[...] in global and regional models is a prerequisite for reproducing historical regional trends and ensuring fitness for future projections (high confidence).»

Thank you for suggesting to cite the AR6 instead; the sentence now reads:

“The IPCC has emphasized that our confidence in RCM projections depends on whether the ~~key~~ relevant processes, forcings (including aerosols) and drivers of climate change are adequately represented ~~“well-simulated and well-projected”~~ by the models.”

2.20. Conclusion: It would be valuable to add perspectives and guidance for the RCM community regarding the upcoming CORDEX-CMIP6 exercise, especially concerning CMIP6 GCM selection and aerosol representation.

We once again agree and now provide two recommendations in this regard at the end of our conclusion:

“In light of the upcoming next generation of CORDEX RCM experiments driven by CMIP6 GCMs, we provide two specific recommendations: First, the driving GCMs are selected based on several criteria including model performance, model independence and projected climate change diversity (di Virgilio et al., 2020), but it might make sense to additionally consider the accuracy of historic thermodynamic warming patterns. Second, we recommend increased efforts to improve the aerosol representations in the RCMs, such that they at the very least include anthropogenic long-term changes in the direct aerosol effect. This would improve both the consistency of CORDEX simulations with respect to the driving CMIP6 GCMs and their ability to reliably depict historical and future climate change.”

4. References of this review not already cited by the authors:

Ribes et al. 2022, <https://doi.org/10.5194/esd-13-1397-2022>

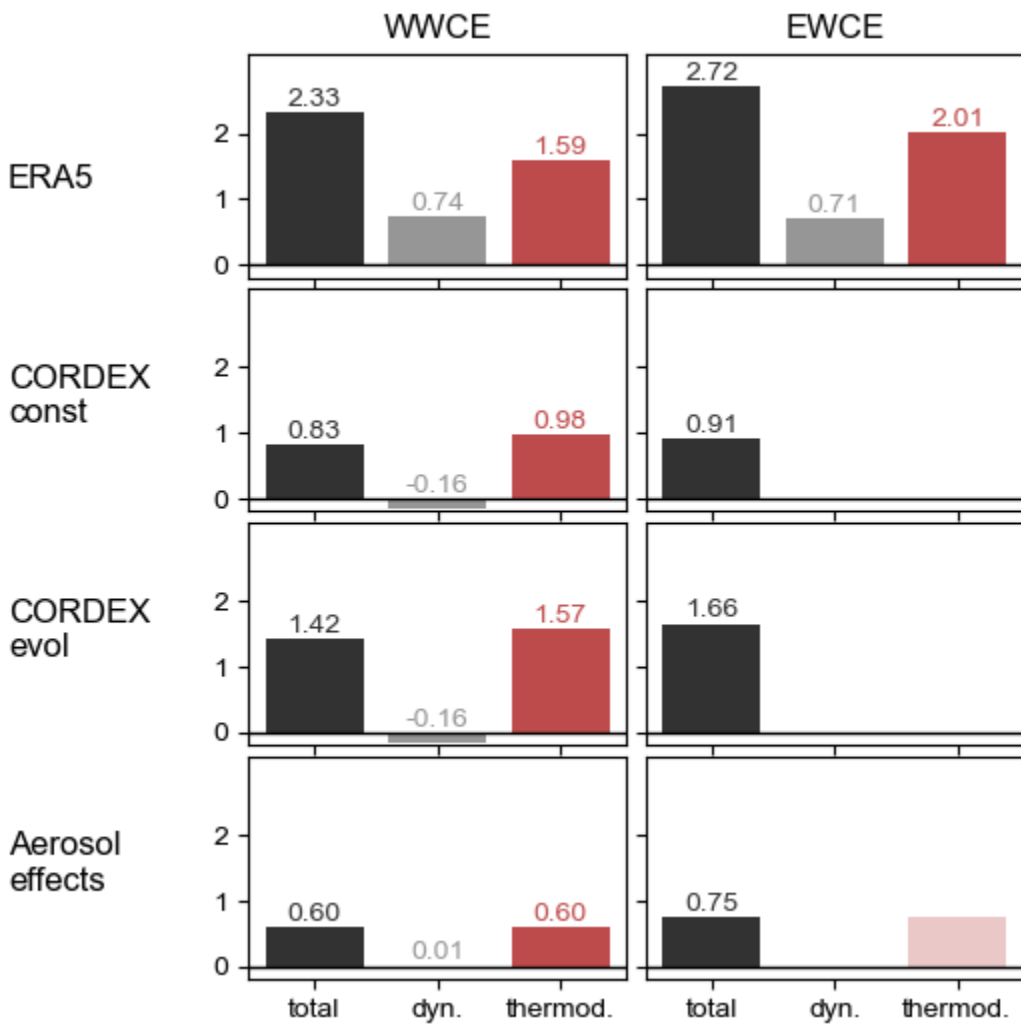
Taranu et al. 2023, <https://doi.org/10.1007/s00382-022-06540-6>

Lorenz and Jacob 2010, <https://doi.org/10.3354/cr00973>

Nabat et al. 2020, <https://doi.org/10.5194/acp-20-8315-2020>

Additional analysis for the eastern half of West-Central Europe (EWCE)

We have repeated the dynamical adjustment analysis for EWCE using ERA5 data. The eastern half of WCE shows even stronger thermodynamic warming since 1980 than the western half. In the following, we outline how our analysis based on CORDEX data, although incomplete due to the lack of warming disentangling, consistently reveals even stronger aerosol-related warming in the east, which likely also contributes to the warming difference between WWCE and EWCE indicated by ERA5.



RFig. 9: 1980–2022 summer warming in WWCE and EWCE in ERA5 and CORDEX RCM simulations (black bars), further disentangled into dynamic (grey) and thermodynamic (red) contributions wherever possible. In the bottom row, the aerosol effect is estimated based on the differences of RCM simulations with evolving and constant aerosols. This is based on the same model subset as employed in the paper; using the entire ensemble, aerosol effects are

weaker overall, but we also find even stronger aerosol effects in the eastern than the western half of the WCE region.

We begin with the contributions from circulation changes, which are estimated to be nearly identical at about 0.7 °C (**RFig. 9, top row**). This is an interesting result given that Vautard *et al.* (2023), as mentioned before, found dynamic contributions predominantly in western Europe (see their Fig. 1 for summer mean maximum temperatures). We expect some differences primarily due to the fact that we investigate the drivers of summer mean temperatures rather than heat extremes, but the inconsistent time periods (1980–2022 here vs. 1950–2022 in Vautard *et al.*, 2023) likely also contribute to the contrasting results for eastern Europe.

Next, we refer again to **RFig. 9**, which reveals stronger thermodynamic warming in the eastern (2.0 °C) than the western half of WCE (1.6 °C) for ERA5. It might be tempting to view this as the difference in aerosol-related warming, but if this were the case, the CORDEX simulations with constant aerosols should show identical warming across WWCE and EWCE. We assume here that the dynamic contributions are similar on average for the western and eastern half of WCE, which we deem reasonable based on the typically near-zero effects found for WWCE, although the domain boundary limitations prevent us from checking whether this holds true.

Interestingly, both for the CORDEX RCM subset (**RFig. 9 or SFig. 7**) and the entire ensemble (**SFig. 8**), EWCE warms slightly more — by about 0.08 °C and 0.17 °C, respectively — than WWCE. We interpret this as the consequence of a stronger oceanic influence for WWCE than EWCE: given the prevalence of westerlies (see, e.g., Fig. S11 in [Keune *et al.*, 2019](#)), in the last days prior to arrival, the winds that reach WWCE, tend to, if at all, interact with the Northern Atlantic, whereas for EWCE, it is a more rapidly warming European land surface (including WWCE itself; the land–ocean warming contrast can be seen in, e.g., [GISTEMP](#) data). Furthermore, as the entire ensemble is driven by GCM simulations with a higher average GMST increase since 1980 than the subset, the CORDEX simulations indicate a more pronounced east–west gradient under stronger background warming. Nevertheless, this gradient seems insufficient to single-handedly explain the east–west thermodynamic warming difference in ERA5.

For simulations with evolving aerosols, we thus expect even stronger differences between the two regions, which is indeed the case: 0.24 °C for the subset (0.16 °C more than for constant aerosols; **RFig. 9**), and 0.33 °C for the full ensemble (again 0.16 °C more; **SFig. 8**). Whether we rely on the subset or the full ensemble, the aerosol effect is always stronger in the east, by about 0.15 °C in absolute terms (**SFig. 7 & 8**, bottom rows). The difference in thermodynamic warming suggested by ERA5 thus likely arises both due to stronger aerosol effects in the east as well as the naturally occurring warming gradient. It is also possible that other factors such as stronger soil moisture–temperature feedbacks play a role, which are known to amplify (the rise in) summer heat extremes unevenly across Europe (e.g., [García-García *et al.*, 2023](#)).

In conclusion, the gradient in model–observation discrepancy from west to east mentioned by reviewer #1 (and shown in Fig. 1b) is at the very least strengthened — if not caused — by stronger aerosol-related changes, which the majority of CORDEX RCM simulations fails to capture.

References

Please refer to the manuscript for all references not listed below.

García-García, A., Cuesta-Valero, F.J., Miralles, D.G. *et al.* Soil heat extremes can outpace air temperature extremes. *Nat. Clim. Chang.* **13**, 1237–1241 (2023).
doi=[10.1038/s41558-023-01812-3](https://doi.org/10.1038/s41558-023-01812-3)

Keune, J., & Miralles, D. G. (2019). A precipitation recycling network to assess freshwater vulnerability: Challenging the watershed convention. *Water Resources Research*, **55**, 9947–9961 (2019). doi=[10.1029/2019WR025310](https://doi.org/10.1029/2019WR025310)

Lehner, F. and Deser, C. *Environ. Res.: Climate* **2**, 023001 (2023).
doi=[10.1088/2752-5295/accf30](https://doi.org/10.1088/2752-5295/accf30)

Singh, J., Sippel, S. & Fischer, E.M. Circulation dampened heat extremes intensification over the Midwest USA and amplified over Western Europe. *Nat. Commun. Earth. Environ.* **4**, 432 (2023). doi=[10.1038/s43247-023-01096-7](https://doi.org/10.1038/s43247-023-01096-7)

Vautard, R., Kadyrov, N., Iles, C. *et al.* Evaluation of the large EURO-CORDEX regional climate model ensemble, *J. Geophys. Res.* **126**, e2019JD032344 (2021).
doi=[10.1029/2019JD032344](https://doi.org/10.1029/2019JD032344)

Vautard, R., Cattiaux, J., Hap  , T. *et al.* Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends. *Nat. Commun.* **14**, 6803 (2023).
doi=[10.1038/s41467-023-42143-3](https://doi.org/10.1038/s41467-023-42143-3)

2nd Feb 24

Dear Dr Schumacher,

We have now assessed your revised manuscript titled "Why climate models underestimate the exacerbated summer warming in Western Europe". In light of your responses to our reviewers' advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to revise your paper one last time to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at <https://www.nature.com/documents/commsj-file-checklist.pdf>.

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a [CC BY license](#) (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit <https://www.nature.com/commsenv/article-processing-charges>

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing

information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:

[link redacted]

**** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first ****

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Mengze Li, PhD
Editorial Board Member
Communications Earth & Environment

Heike Langenberg, PhD
Chief Editor
Communications Earth & Environment

On Twitter: @CommsEarth

REVIEWERS' COMMENTS: