nature portfolio

Peer Review File

Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high-resolution greenhouse warming simulations

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

13th Feb 23

Dear Mr Nellikkattil,

Your manuscript titled "Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high-resolution greenhouse warming simulations" has now been seen by 3 reviewers, whose comments are appended below. You will see that they find your work of some potential interest. However, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version that fully addresses these serious concerns.

We hope you will find the reviewers' comments useful as you decide how to proceed. Should additional work allow you to

• address these comments (that is, either to incorporate the suggestions or provide a compelling argument why the points made by the reviewers are not valid, or relevant to the editorial threshold as outlined below)
AND
• meet our editorial thresholds as outlined below,

Then we would be happy to look at a substantially revised manuscript. In the following, we list our main editorial concerns.

****
****Editorial threshold 1:
Provide an in-depth and reasonable interpretation of the analysis of dynamical versus thermodynamic contributions to future AR probabilities.
****Editorial threshold 2:
Provide an in-depth discussion on the novelty of ARDT method compared to previous studies.
****Editorial threshold 3:
Provide a much more transparent methodology and the code/data, to a point where an independent researcher can reproduce their results.
****

If you choose to take up this option, please either highlight all changes in the manuscript text file, or provide a list of the changes to the manuscript with your responses to the reviewers.

Please bear in mind that we will be reluctant to approach the reviewers again in the absence of substantial revisions.

If the revision process takes significantly longer than three months, we will 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.

We understand that due to the current global situation, the time required for revision may be longer than usual. We would appreciate it if you could keep us informed about an estimated timescale for resubmission, to facilitate our planning. Of course, if you are unable to estimate, we are happy to accommodate necessary extensions nevertheless.

We are committed to providing a fair and constructive peer-review process. Please do not 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 reviewers’ comments with a list of your changes to the manuscript text (which should be in a separate document to any cover letter) and any 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 **

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Min-Hui Lo, PhD
External Editor
Communications Earth & Environment

Heike Langenberg, PhD
Chief Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements (Download the link to your computer as a PDF.)

For your information, you can find some guidance regarding format requirements summarized on the following checklist:(https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf) and formatting guide (https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

Reviewer’s summary of the Manuscript

The paper entitled “Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high-resolution greenhouse warming simulations” by Nellikkattil et al. explores the potential change in atmospheric rivers and associated precipitation. Using a suite of ultra-high-resolution warming simulations and a novel detection algorithm (ARDT), this study demonstrates the spatially inhomogeneous increase in AR frequency, IVT, and precipitation, are mainly determined by the balance/imbalance between dynamics and thermodynamics response to the radiative forcing, which is further reflected on the change in underlying synoptic circulations.

Summary of Review
While this study is generally interesting, there are some major issues, which stop me from suggesting further publication of this work. First, the authors claimed the novel ARDT, SCAFET, is more objective and the geometry-based detection approach makes it less influenced by the warming-induced increase of specific humidity. This argument is correct but it does not address the main source of uncertainty uncovered by the ARTMIP effort: whether the basic climate state for detection is time-dependent or not. The ARDTs can be roughly categorized into two groups: (1) fixed threshold, and (2) relative threshold. While most of the ARDTs fall within the first category as indicated by this study and O'Brien et al. (2022), there are two ARDTs, Tempest (Ullrich et al., 2021) and TECA-BARD(O'Brien et al., 2020) belonging to the second category. When the used thresholds in the former category can lead to a quantitative difference (amplitude) in estimated AR frequency, whether the threshold is time-dependent can lead to a difference in sign (qualitative). SCAFET, obviously, falls into the second category. I don't think this paper provides additional insight into ARDTs other than those that have been discussed in O'Brien et al. (2022). In plain language, it's the difference between categories 1 and 2 that causes significant uncertainty rather than one category being superior to the other. For example, using category 1 of ARDT is more likely to highlight the importance of thermodynamics while using category 2 is more likely to attribute the climate change signals to the dynamics component. Introducing another new AR detection algorithm doesn't help reconcile these differences.

Second, the use of ultra-high-resolution simulations should be better justified. I was expecting more effort in the discussion of how the model resolution influences the estimated AR frequency in warming simulations since it distinguishes this study from other CMIP-based studies, where the model grid spacing is usually around 0.5-1 degree. For example, does the ultra-high-resolution CESM captures more reasonable AR climatology than its course-resolution counterpart? Or how sensitive the result is over different climate states? I think this is the money shot of the entire study but it is barely addressed.

Third, some analyses are flawed, which cast concerns about the result interpretation. For example, the decomposition of thermodynamics and dynamics components suggests a decline in IVT due to the increase of temperature over the tropics (Figs 2e and 2f), which is counter-intuitive and not consistent with previous studies. The authors need to carefully survey the analyses.

Data availability

In addition to the sciencewise issues, the data availability is not properly provided. According to Nature policy, the data availability statement must make the conditions of access to the "minimum dataset" that is necessary to interpret, verify and extend the research in the article, transparent to readers. However, the public FTP server does not provide access to either IVT or the fields necessary to calculate IVT. It is therefore not possible to replicate the analysis. There are currently a few resources for preserving files of very large size. (see https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards for details). Given that it is very unlikely to provide the entire dataset, I suggest authors choose a small chunk of data (~5-10 years) from each simulation for release.

Minor comments:
Line 105: True, but as mentioned in the major comments. The biggest uncertainty arises from

whether the ARDT threshold is time-dependent or not.

Line 152: Tempest uses a similar geometry-based approach (Laplacian in their paper) to identify the regions of AR. Do authors compare SCAFET with Tempest or other geometry-based algorithms to test the robustness of result?

Fig. 1 and lines 169-170: The caption is not clear. Is the AR frequency defined by a year or a single season? 20% of AR days (equivalent to 73 days) for an entire year is an unrealistic high value. This almost extends over the entire winter season (~90 days).

Line 212: I will say most of these features are very noisy. For example, the change in East Asia monsoon shows a dipole feature in meridional direction rather than experiencing an increase as suggested by the authors.

Equation 8: typo for the last term. Should be a "-" sign rather than "+"

References:
O'Brien, T. A., Risser, M. D., Loring, B., Elbashandy, A. A., Krishnan, H., Johnson, J., et al. (2020). Detection of atmospheric rivers with inline uncertainty quantification: TECA-BARD v1. 0.1. Geoscientific Model Development, 13(12), 6131–6148.

Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., & Reed, K. A. (2021). TempestExtremes v2.1: a community framework for feature detection, tracking, and analysis in large datasets. Geoscientific Model Development, 14(8), 5023–5048.

Reviewer #2 (Remarks to the Author):

Summary

Nellikkattil et al. in "Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high resolution greenhouse warming simulations" globally evaluate AR environments and characteristics across 20-year present day, 2XC02, and 4XC02 simulations produced by a high-resolution CESM1.2.2 (0.25 deg atmosphere, 0.10 deg ocean simulation) run and contrast the historical model results against the ERA5 reanalysis dataset. The authors apply a novel tracking algorithm (SCAFET) that better accounts for background IVT shifts from present day that minimizes "stationarity" built into the tracking algorithm in a future climate (i.e., does not rely on a minimum IVT threshold that determines if an AR does/does not exist). The authors provide a thorough, comprehensive, and well-written analysis with several interesting findings such as ARs becoming stronger by 15% and 34% and increased mean annual precipitation by 6% and 15% in 2XC02 and 4XC02 worlds, respectively. Overall, I think the paper definitely fits within the scope of Nature Communications Earth and Environment and could be, given a bit more work, a valuable contribution to the scientific community. The findings have both scientific and societal impact as future changes in ARs have important scientific uncertainties that are still being debated in the community (e.g., the role of ARDT in shaping future AR change responses) and future AR changes will play a large role in driving the hydrologic cycles (and water resource management) in many coastal midlatitude regions. I think there are still several major(ish) revisions that need to happen prior to this paper being accepted.

Suggested Revisions

Line 53 – can you rephrase what you mean by “core axis”?

Line 54 – change “ARs cause vigorous” to “Once ARs make landfall they undergo vigorous”

Line 55 – change “excessive” to “copious”

Line 57 and 58 – be consistent with “mid-latitude” or “midlatitude” throughout the manuscript

Line 64 – The authors might consider citing Payne et al., 2020 - https://www.nature.com/articles/s43017-020-0030-5

Line 71-79 – If the authors are also interested in highlighting paleoclimate AR interactions (using CESM) within glaciated regions, please see this study - https://www.sciencedirect.com/science/article/pii/S0277379121004625

Line 83-84 – The authors might consider citing Gershunov et al., 2019 - https://www.nature.com/articles/s41598-019-46169-w , Patricola et al., 2021 - https://www.sciencedirect.com/science/article/pii/S2212094722000275 - and Rhoades et al., 2020 and 2021 - https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020GL089096 and https://www.sciencedirect.com/science/article/pii/S2212094721000244

Line 105 – the authors might consider citing O’Brien et al., 2020 - https://gmd.copernicus.org/articles/13/6131/2020/gmd-13-6131-2020-discussion.html

Line 106 – “absolute or relative thresholds…” some also use size constraints (e.g., binary mask length/width criteria) which might be worth mentioning too

Line 108-121 – the authors might consider reading/citing Inda-Diaz et al., 2021 - https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033746 - and Zhou et al., 2021 - https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033711 - and adding some more details about the strengths/weaknesses of different ARDTs and how they shape reanalysis dataset estimated AR frequency/intensity, particularly due to ARDT structural decisions (e.g., length/width criteria) and sensitivities to large-scale modes of climate variability (e.g., ENSO, MJO, etc.)

Line 125 – change “Compared to understanding regarding AR characteristics…” to “Compared to physical intuition regarding AR characteristics…”

Line 129-137 – Eloquently said. These are all very important points that need to be more discussed in the literature!

Line 138-141 – Could the authors provide a bit more specificity on how machine learning/image processing techniques could get around issues of embedding historical stationarity into ARDT estimates of AR frequency/intensity/spatial extents (e.g., fixed threshold of 250 kg/m/s)?

Line 151 – is “ultra-high-resolution” a common term? I’m a global/regional climate modeler and I haven’t heard of this term before. It made me initially think the model would be run at convective/storm resolving scales such as the DYAMOND experiments - https://progearthplanetsci.springeropen.com/articles/10.1186/s40645-019-0304-z - yet the atmosphere is at ~0.25 deg. I recognize the ocean model being run at ~0.1 deg is no small feat, but the authors might consider revising this term throughout or provide better justification for why this term is used given the community is now calling the DYAMOND experiments “digital twins” (which is analogous to “ultra-high-resolution”).

Line 185-186 – could you guide the reader to the figure and subpanel label to back this claim up (i.e., Figure 1c)?

Line 186-188 – could it also be the jet stream strength too? What about other drivers such as the MJO? See Zhou et al., 2021 - https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033711 - for more information. Also, do the authors posit that ERA5 might have biases too (e.g., too dissipative in estimating wind speed/specific humidity strength/variability) and could they provide a sentence or two discussing how reanalysis dataset choice might also shape model historical bias estimation?

Fig. S2 – please add in the caption of the figure which line corresponds to ERA5 vs UHR CESM

Line 192-194 – could the authors provide a citation or two that discuss these various “components” of AR conditions (e.g., WCBs, ECs, AFs, etc.) that non-AR experts or new AR researchers could read to spin up on these terms? For example, I’ve found this to be a great resource - http://cw3e.ucsd.edu/ARconf2016/Presentations/Wednesday/S3_TheoryStructureProcesses1/Cordeira_IARC2016.pdf – but I’m not sure if Cordeira has published a study highlighting these “components” of AR conditions (see Part 2).

196-197 – Patricola et al., 2021 - https://www.sciencedirect.com/science/article/pii/S2212094722000275 - could provide some further guidance on proportionality of EC/AR relationship and how the presence of an EC shapes AR characteristics. This might be worth a sentence or two of discussion.

Line 218-237 and 275-277 – I’m not sure what the global temperature response in UHR CESM is at 2XC02 or 4XC02, but I think this would be important to include throughout the manuscript and maybe in figure headers, especially when comparing with RCP8.5 scenario literature. With that said, the global mean change in AR frequency seems very consistent with Clausius-Clapeyron (6-7%/K) in the 2XCO2 and 4XCO2 experiments. Do the authors want to comment on how the dynamical response of 2XCO2 and 4XCO2, at least in Fig. 2, seems to play less of a role than the thermodynamic response? Figure 3 in Payne et al., 2020 - https://www.nature.com/articles/s43017-020-0030-5 – makes it seem like the dynamic component under climate change (e.g., storm track shift poleward, slow down of jet/winds due to equator-to-pole temperature gradient reduction, etc.) would be more equal and opposite to the thermodynamic component under climate change. Might be worth mentioning that the strength/direction of dynamic vs thermodynamic component changes in IVT with warming could be model dependent too?

Line 218-237 – Rhoades et al., 2020 - https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020GL089096 – found a decrease in mean climatological AR frequency over the North Pacific under RCP8.5 with CESM, but an uptick in the number of landfalling ARs and, importantly, the strength/lifetime of the ARs. Mean AR frequency is an important metric, but landfalling AR frequency (particularly extreme ARs) are the most important metric from an impacts perspective, particularly under 2XC02/4XC02/RCPs/SSPs. Could the authors highlight how the strength/intensity/duration of landfalling AR events shift under different levels of CO2/warming? Another neat metric that the authors could assess (thanks to the impressive decomposition framework they have generated!) would be to assess the orthogonality of landfalling ARs which is particularly important when assessing the orographic uplift potential (which has been found to better correlate with AR total precipitation - https://journals.ametsoc.org/view/journals/hydr/23/9/JHM-D-21-0119.1.xml) at AR landfall in regions where large coastal mountains are (e.g., western North America and South America)?

Fig. S3c,4c,5c,6c – change "Themal" to "Thermodynamic"

Fig. S3e,4e,5e,6e – change "Thermal" to "Thermodynamic"

Line 266-273 – "…AR frequency is largely modulated by DIVT…", could the authors provide some quantitative evidence for this statement (maybe using spatial pattern correlations or ? between IVTy and the thermo/dynamic components) At the moment, the reader is asked to do an "eyeball" test between subpanel a) with b) and c) and in some regions this statement about DIVT doesn't seem to hold true (but maybe at the global scale).

Line 295-329 – very thorough and striking analysis, nice work! Did the authors assess rain-snow partitioning of ARs too? Would it be easy enough to recreate Fig. 4 but compare and contrast precipitation extremes by precipitation phase (i.e., extreme snowfall events vs rainfall events)? This would be a critical analysis from an AR impacts perspective (e.g., flood and landslide hazards and water resource management of downstream reservoirs due to rain-on-snow/spring snowmelt reduction), particularly if done for all the major AR terminus regions (as shown in Fig. 3 of Payne et al., 2020 - https://www.nature.com/articles/s43017-020-0030-5).

Line 337-338 – could you provide your % change values from your findings here and the literature range of % change values so that the reader can determine "significant" more easily (especially if a reader was to quickly read abstract/conclusions sections)?

Line 355-356 – what is meant by "AR circulations systems"?

Line 364 – change "though this is admittedly difficult to globally" to "though this is admittedly difficult to systematically apply globally"

Line 439-440 – minor detail, but do you also use the surface pressure variable for the integration from the Earth's surface to the top of atmosphere (or 300 hPa for ERA5) or simply assume 1000 hPa (as is shown in Section 5.1 equations 1-3)?

Line 456 – I think it would be helpful if the authors highlight the step-by-step AR extraction using the workflow outlined for SCAFET in the Methods (SFig. 8) earlier on in the study (maybe on line 164-165). Also, how does the SCAFET workflow differ from TempestExtremes and some of the algorithms outlined in Inda-Diaz et al., 2021 -
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033746 ?

Line 460 – could the authors provide a citation (or two) explaining the Gaussian filter used?

Line 479 – wouldn't this minimum mean precipitation criteria (1 mm/day) bias the AR detection (particularly in a warmer world where these precipitation rates might systematically increase)? How sensitive are estimates when not using or altering this criterion, especially when applying this to sub-daily fields (since its fixed at the daily scale, mm/day)?

Line 491-494 – could you define ps and pt (some AR researchers don't integrate through the entire atmosphere, while others do)? I assume you followed the same workflow as in ERA5 in Section 5.1, but just thought it would be important to be clear to readers.

Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high-resolution greenhouse warming simulations

[Reply to Reviewers]

Arjun Babu Nellikkattil[1,2], June-Yi Lee[1,2,3], Bin Guan[4,5], Axel Timmermann[1,6], Sun-Seon Lee[1,6], and Jung-Eun Chu[7]

[1]Center for Climate Physics, Institute for Basic Science
[2]Department of Climate System, Pusan National University
[3]Research Center for Climate Sciences, Pusan National University
[4]Joint Institute for Regional Earth System Science and Engineering
[5]Jet Propulsion Laboratory, California Institute of Technology
[6]Pusan National University
[7]Low-Carbon and Climate Impact Research Centre, School of Energy and Environment, City University of Hong Kong

Contents

IBS Center for Climate Physics

Rooom No. 1105, IBS Center for Climate Physics
Busandaehak-ro 63beon-gil 2 (Jangjeon-dong),
Geumjeong-gu, Busan 46241, Republic of Korea

T +82-51-510-7636
F +82-51-510-7694

Email: arjunbabun@pusan.ac.kr
Web: ibsclimate.org

1 Response to Reviewer #1

Re: Re-submission of manuscript COMMSENV-22-1342-T

Dear Reviewer,

Thank you for the comments and suggestions regarding the manuscript. We recognize that the concerns raised are important and have tried our best to address them faithfully. The main results and interpretations in the revised manuscript have become more robust and reliable, thanks to your suggestions.

A point-by-point response to all the remarks are attached. A PDF document, highlighting the changes between the two versions of the manuscript is also attached for your convenience.

The revisions were done in consultation with all authors. We hope that the current version of the manuscript is competent enough to be published in *Communications Earth and Environment*. However, we are open to further revisions if necessary.

1.1 Reviewer Comments

Recommendation: While this study is generally interesting, there are some major issues, which stop me from suggesting further publication of this work.

1.2 Revision Highlights

- Novelty and intercomparison of the ARDT.
  - AR event detection and mean spatial distribution obtained from 8 other ARDTs are compared with SCAFET.
  - ARs detected from CESM UHR using SCAFET is compared with Guan and Waliser (2015) and Ullrich and Zarzycki (2017).
  - Included new reference describing the detection framework. "Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction

from large climate datasets" (`https://egusphere.copernicus.org/preprints/2023/egusphere-2023-592/`).

- Better attribution of thermodynamic and dynamic components of IVT changes to mean changes in AR frequency. Additional analysis has been conducted and Fig. 2 has been modified in the manuscript.

- The AR frequencies and characteristics obtained from low resolution and high resolution simulations are compared.

- Code and data availability

  - Latest version of SCAFET can be found in GitHub (`https://github.com/nbarjun/SCAFET/`)

  - The codes and sample datasets used for the manuscript can be downloaded from `10.5281/zenodo.6979522`

1.3 Point-by-Point Response

1.3.1 Major Comments

First, the authors claimed the novel ARDT, SCAFET, is more objective and the geometry-based detection approach makes it less influenced by the warming-induced increase of specific humidity. This argument is correct but it does not address the main source of uncertainty uncovered by the ARTMIP effort: whether the basic climate state for detection is time-dependent or not. The ARDTs can be roughly categorized into two groups: (1) fixed threshold, and (2) relative threshold. While most of the ARDTs fall within the first category as indicated by this study and O'Brien et al. (2022), there are two ARDTs, Tempest (Ullrich et al., 2021) and TECA-BARD(O'Brien et al., 2020) belonging to the second category. When the used thresholds in the former category can lead to a quantitative difference (amplitude) in estimated AR frequency, whether the threshold is time-dependent can lead to a difference in sign (qualitative). SCAFET, obviously, falls into the second category. I don't think this paper provides additional insight into ARDTs other than those that have been discussed in O'Brien et al. (2022). In plain language, it's the difference between categories 1 and 2 that causes significant uncertainty rather than one category being superior to the other. For example, using category 1 of ARDT is more likely to highlight the importance of thermodynamics while using category 2 is more likely to attribute the climate change signals to the dynamics component. Introducing another new AR detection algorithm doesn't help reconcile these differences.

The reviewer is right in pointing out that SCAFET does not remove one of the main uncertainties in AR detection, which is **whether ARs should be defined based on time-independant (fixed) or time-dependant (relative) thresholds, and we would like to clarify that it is not the objective of the current study to remove such uncertainties**. We would like to point out that, answering this question will involve theoretically defining ARs. Similar ambiguities exist in studying other phenomena like cyclones, atmospheric fronts etc. Thus, it is quite normal to set different threshold values depending on the impact of the phenomena and the focus of the study. For instance, in the current manuscript we have set a minimum threshold on precipitation (although this does not affect our conclusions) as our focus is the changes in mean and extreme precipitation due to ARs in response to greenhouse gas forcing.

Lacking clarity in defining ARs, we believe that SCAFET has some unique features compared to other ARDTs. The classification of ARDTs into three categories, 1) absolute 2) fixed relative and 3) relative as in O'Brien et al. (2020) is still a simplified view of the various characteristics of ARDTs. SCAFET has characteristics of all three types of ARDTs, depending on what you are looking at. Like an ARDT with absolute thresholds, SCAFET applies one single value (shape index > 0.375) to extract AR structures independent of mean state, resolution etc. The application of one absolute value is also done for TempestExtremes. However, unlike TempestExtremes and absolute threshold ARDTs, SCAFET shows hemispherically homogeneous AR patterns. The hemispherical inhomogeneity for absolute thresholds is due to the lower water vapor content in the Southern Hemisphere compared to the Northern Hemisphere. In identifying present day ARs the results from SCAFET are similar to fixed relative ARDTs which shows hemispherically homogeneous AR distribution. As pointed out by the reviewer, climate change studies using SCAFET yield results similar to that of ARDTs based on relative thresholds. Also, as discussed in supplementary Fig. S9, by changing the value of minimum precipitation threshold for ARs, one can make the results from climate change studies using SCAFET look like any of the three types of ARDTs. None of this discussion is to claim SCAFET performs better that other ARDTs, but to emphasize the uniqueness of the methodology in combining various aspects across different types of the previously published ARDTs.

The issue of choosing a threshold is not limited to the detection of atmospheric rivers. Other features like cyclones, jet streams, and fronts also struggle to identify the right kind of threshold for feature identification. The framework developed in SCAFET can be extended to identify any such phenomenon. Different types of features can be extracted based on their spatial scale and geometric shape. Please refer to the preprint "Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction from large climate datasets" (`https://egusphere.copernicus.org/preprints/2023/eg` usphere-2023-592/) for more details on how SCAFET can be applied to detect three and two-dimensional phenomenons. The manuscript also uses some simple examples to explain the uniqueness of the approach compared to TempestExtremes and other traditional feature extraction methods.

To emphasize the uniqueness of the ARDT, we have modified paragraphs in section 2.1 (lines 199-215) and section 2.3 (lines 265-291). Further, reference to the preprint describing the feature extraction framework is also added.

Second, the use of ultra-high-resolution simulations should be better justified. I was expecting more effort in the discussion of how the model resolution influences the estimated AR frequency in warming simulations since it distinguishes this study from other CMIP-based studies, where the model grid spacing is usually around 0.5-1 degree. For example, does the ultra-high-resolution CESM captures more reasonable AR climatology than its course-resolution counterpart? Or how sensitive the result is over different climate states? I think this is the money shot of the entire study, but it is barely addressed.

Addressing the effect of resolution in the simulation of ARs and its projected changes is a very important question to answer. However, to answer the question better than any of the previous studies, we would have to apply multiple ARDTs on datasets. This is to account for the large uncertainties introduced by the ARDTs. Also, to keep the uncertainties due to datasets low, we will have to run the same model configuration but with varying resolution. Performing these analyses are out of the scope of this paper.

To address the comments on a limited scale, we have applied SCAFET AR detection on daily mean IVT data from CMIP5 and CMIP6 outputs. To keep the consistency in the datasets considered, we used CCSM4 and CESM2 historical simulations (see supplementary Fig. S4 and Fig. S5). The overlapping time period was chosen to be from 1986 to 2005. The major inferences from the comparison between AR characteristics in low resolution (LR) and high resolution (HR) datasets are as follows.

- As ARs are large-scale phenomenons, there is no significant difference in the spatial distribution of AR frequency. In a way this analysis also shows that the AR detection using SCAFET is resolution independent, unlike some other ARDTs that belong to the relative category and supporting the uniqueness of SCAFET.
- ARs in HR simulations tend to give narrower ARs. This is similar to some previously published works, which also found HR datasets to have smaller widths (Collow et al., 2022).

- The maximum convection strength and maximum precipitation intensity in LR ARs are much smaller than that in HR ARs. Similar inferences were made in previous studies too (Collow et al., 2022; Chang et al., 2020; Shields et al., 2023).

Apart from the addition of new supplementary figures, we have also added discussion about the resolution sensitivity in AR characteristics and the resolution insensitivity in the algorithm in the manuscript text (see lines 185-198).

Third, some analyses are flawed, which cast concerns about the result interpretation. For example, the decomposition of thermodynamics and dynamics components suggests a decline in IVT due to the increase of temperature over the tropics (Figs 2e and 2f), which is counter-intuitive and not consistent with previous studies. The authors need to carefully survey the analyses.

While we agree that the IVT decomposition is confusing, it is mathematically and physically sound. The previous Figs 2e and 2f shows the zonal component of dynamic IVT changes. Since IVT is a vector, negative sign does not indicate a reduction in magnitude of IVT but an increase in westward transport. The full decomposition is shown in supplementary Fig. S11-S14. The residuals from the decomposition was also found to be much smaller than the thermodynamic or dynamic terms.

To better demonstrate the relationship between AR frequency and IVT components, we have modified the analysis. In the revised version, instead of showing the zonal and meridional components of IVT changes, we now show vector projection of DIVT and TIVT onto the mean IVT in present climate. This step will allow us to see the changes in each component with respect to the present day mean state allowing us to focus on the magnitude of change in relation to the existing mean transport. In the revised Fig 2e-h, positive values indicate that the component will reinforce the prevailing water vapor transport while negative values indicate that the component acts to weaken it. From the revised Fig. 2, we see that regions which show reduction in mean AR frequency show DIVT acting to reduce the mean IVT. The strongest increases are observed in regions where both TIVT and DIVT act to reinforce the mean IVT.

The materials and methods section have been modified to reflect these new changes. We have also described the new analysis within the manuscript in lines 312-337.

In addition to the sciencewise issues, the data availability is not properly provided. According to Nature policy, the data availability statement must make the conditions of access to the "minimum dataset" that is necessary to interpret, verify and extend the research in the article, transparent to readers. However, the public FTP server does not provide access to either IVT or the fields necessary to calculate IVT. It is therefore not possible to replicate the analysis. There are currently a few resources for preserving files of very large size. (see https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards for details). Given that it is very unlikely to provide the entire dataset, I suggest authors choose a small chunk of data (5-10 years) from each simulation for release.

1. The latest version of the detection algorithm is available in GitHub (https://github.com/nbarjun/SCAFET).
2. All the codes and data used for detecting ARs, and generating the figures in the paper is hosted in Zenodo (https://zenodo.org/deposit/6979522).
3. A sample of the high resolution data is also hosted in Zenodo (https://zenodo.org/deposit/6979522).

1.3.2 Minor Comments

Line 105: True, but as mentioned in the major comments. The biggest uncertainty arises from whether the ARDT threshold is time-dependent or not.

Yes. It is true that the biggest uncertainty does lie in the definition of AR thresholds. However, the issue of defining thresholds exits in the detection of all types of phenomenons in climate. Thus, depending on the focus of the study, various type of thresholds can be used. As discussed in the previous section, using SCAFET for identifying ARs provides a unique blend of different types of ARDTs. The framework is robust enough to be applied to a variety of weather phenomenons as shown in "Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction from large climate datasets" (https://egusphere.copernicus.org/preprints/2023/egusphere-2023-592/).

Line 152: Tempest uses a similar geometry-based approach (Laplacian in their paper) to identify the regions of AR. Do authors compare SCAFET with Tempest or other geometry-based algorithms to test the robustness of result?

Yes. Thank you for the suggestion. We have applied TempestExtremes (Ullrich et al., 2021) and Guan Waliser (Guan and Waliser, 2015) for estimating the projected changes.

The results are included in the supplementary section (Fig. S3, Fig. S8). The results are discussed within the manuscript. The results are as expected considering the type of detection algorithms and previous studies.

SCAFET was developed independently of Tempest. Tempest uses the Laplacian of the Integrated Vapor Transport (IVT) to identify AR like structures. In a more technical way SCAFET uses the two eigenvalues k_1, and k_2 to construct Shape Index (SI) and apply threshold on it. Tempest applies threshold on k_2 to identify ARs. The fundamental difference between the two approaches is that while SI is a resolution-independent, bounded variable, k_2 is not. A more detailed comparison of detection methods on simple atmospheric circulation can be found in "Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction from large climate datasets" (`https://egusphere.copernicus.org/preprints/2023/egusphere-2023-592/`).

Fig. 1 and lines 169-170: The caption is not clear. Is the AR frequency defined by a year or a single season? 20% of AR days (equivalent to 73 days) for an entire year is an unrealistic high value. This almost extends over the entire winter season (90 days).

The AR frequency shown in the figure is for the annual mean. Thus, 20% of AR days would be equivalent to 73 AR days. While it is a high number of ARs, it is not an outlier. Other detection methods like Lora et.al. (see lora in Fig. S2) show similarly high numbers.

The large number of ARs is because we do not apply any kind of thresholding in IVT (although it can be implemented). As shown in Fig. S9, application of minimum thresholds on mean precipitation of the AR structure can reduce the number of ARs. However, then the debate would be on how to choose this threshold. In the current setup we have kept the minimum AR precipitation to be higher than 1mm/day. The results from this threshold look very similar to not having a threshold at all (see Fig. S9).

Line 212: I will say most of these features are very noisy. For example, the change in East Asia monsoon shows a dipole feature in meridional direction rather than experiencing an increase as suggested by the authors.

Yes. It is true that the patterns appear nosier than other ARDTs (see also Fig. S8). This again is due to the fact that no lower threshold is set on IVT. The inference would then be that these AR like structures are present in current climate as well as the future climate. But as the climate warms, these structures carry more water vapor and leads to higher precipitation (see Fig. 2 and Fig. 3). The changes get less and less noisy as we apply stricter thresholds (Fig. S9). Further, the noisiness is not very surprising as the AR detection in SCAFET is sensitive to the changes in mean circulation patterns (see Fig. 2). As discussed in many previous studies, the circulation changes in response to greenhouse warming is very noisy. If we apply a fixed or fixed relative threshold, the changes in AR frequency becomes more coherent, while the changes in mean AR properties like IVT or precipitation become less so. The discussion then comes back to on how to do you define ARs. We believe that this would be a complementary approach of looking at the response of climate features to increased greenhouse gas.

Also, the two-tailed t-test is used to identify regions that show significant changes in frequency distribution.

Thank you for pointing out the error in the description of AR changes. We have modified the sentences to reflect you comments.

Equation 8: typo for the last term. Should be a "-" sign rather than "+"

Thank you. The correction has been made.

IBS Center for Climate Physics
Rooom No. 1105, IBS Center for Climate Physics
Busandaehak-ro 63beon-gil 2 (Jangjeon-dong),
Geumjeong-gu, Busan 46241, Republic of Korea
T +82-51-510-7636
F +82-51-510-7694
Email: arjunbabun@pusan.ac.kr
Web: ibsclimate.org

2 Response to Reviewer #2

Re: Re-submission of manuscript COMMSENV-22-1342-T

Dear Reviewer,

First and foremost, we would like to thank you for the time and effort you took to meticulously go through our manuscript. We have tried to incorporate nearly all of your suggestions into the new version of the manuscript, and it shows considerable improvement in quality, thanks to your suggestions. Your attention to details and overall enthusiasm about the paper is much appreciated.

A point-by-point response to all the remarks are attached. A PDF document, highlighting the changes between the two versions of the manuscript is also attached for your convenience.

All authors have gone through the revised version of the manuscript. Please feel free to communicate any further revisions if necessary. We hope that the current version is suitable for publishing with Communications Earth and Environment.

2.1 Reviewer Comments

Recommendation: Overall, I think the paper definitely fits within the scope of Nature Communications Earth and Environment and could be, given a bit more work, a valuable contribution to the scientific community. The findings have both scientific and societal impact as future changes in ARs have important scientific uncertainties that are still being debated in the community (e.g., the role of ARDT in shaping future AR change responses) and future AR changes will play a large role in driving the hydrologic cycles (and water resource management) in many coastal midlatitude regions. I think there are still several major(ish) revisions that need to happen prior to this paper being accepted.

2.2 Revision Highlights

- Novelty and intercomparison of the ARDT.

- AR event detection and mean spatial distribution obtained from 8 other ARDTs are compared with SCAFET.
- ARs detected from CESM UHR using SCAFET is compared with Guan and Waliser (2015) and Ullrich and Zarzycki (2017).
- Included new reference describing the detection framework. “Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction from large climate datasets” (`https://egusphere.copernicus.org/preprints/2023/egusphere-2023-592/`).

• Better attribution of thermodynamic and dynamic components of IVT changes to mean changes in AR frequency. Additional analysis has been conducted and Fig. 2 has been modified in the manuscript.

• The AR frequencies and characteristics obtained from low resolution and high resolution simulations are compared.

• Code and data availability

- Latest version of SCAFET can be found in GitHub (`https://github.com/nbarjun/SCAFET/`)
- The codes and sample datasets used for the manuscript can be downloaded from `10.5281/zenodo.6979522`

2.3 Point-by-Point Response

2.3.1 Suggested Revisions

Line 53 – can you rephrase what you mean by “core axis”?

Modified to remove the word "core axis"

Line 54 – change “ARs cause vigorous” to “Once ARs make landfall they undergo vigorous”

changed "ARs cause vigorous upward motion" to
ARs undergo vigorous upward motion along topographic barriers like mountains or coastlines where they make landfall, leading to torrential downpours.

Line 55 – change “excessive” to “copious”

Changed to

ARs undergo vigorous upward motion along topographic barriers like mountains or coastlines where they make landfall, leading to torrential downpours.

Line 57 and 58 – be consistent with “mid-latitude” or “midlatitude” throughout the manuscript

Accepted. The changes were made throughout. Thank you for pointing it out.

Line 64 – The authors might consider citing Payne et al., 2020 - `https://www.nature.com/articles/s43017-020-0030-5`

Added the citation

Line 71-79 – If the authors are also interested in highlighting paleoclimate AR interactions (using CESM) within glaciated regions, please see this study - `https://www.sciencedirect.com/science/article/pii/S0277379121004625`

Added a sentence and relevant reference referring to paleo-climate changes. The new sentence reads the following.

The mean hydroclimate over these areas are shaped by ARs in the present as well as past climates [25–27].

Line 83-84 – The authors might consider citing Gershunov et al., 2019 - `https://www.nature.com/articles/s41598-019-46169-w` , Patricola et al., 2021 - `https://www.sciencedirect.com/science/article/pii/S2212094722000275` - and Rhoades et al., 2020 and 2021 - `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020GL089096` and `https://www.sciencedirect.com/science/article/pii/S2212094721000244`

Added new references. Thank you for pointing them out.

Line 105 – the authors might consider citing O’Brien et al., 2020 - `https://gmd.copernicus.org/articles/13/6131/2020/gmd-13-6131-2020-discussion.html`

Added.

Line 106 – "absolute or relative thresholds…" some also use size constraints (e.g., binary mask length/width criteria) which might be worth mentioning too

The sentence has been modified to the following.
The majority of pre-existing ARDTs use absolute or relative thresholds based on climate variables, such as integrated water vapor content (IWV) or integrated water vapor transport (IVT) (hereafter thresholds), along with few shape, size and location constraints to identify ARs in the present climate [50, 51].

Line 108-121 – the authors might consider reading/citing Inda-Diaz et al., 2021 - `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033746` - and Zhou et al., 2021 - `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033711` - and adding some more details about the strengths/weaknesses of different ARDTs and how they shape reanalysis dataset estimated AR frequency/intensity, particularly due to ARDT structural decisions (e.g., length/width criteria) and sensitivities to large-scale modes of climate variability (e.g., ENSO, MJO, etc.)

We have not explored the relationship between AR frequencies and major modes of variability because 1) due to the design of the algorithm, mean AR frequencies are not sensitive to the biases in the datasets (Fig. 1, Fig. S4, and Fig. S7), and 2) we are focussing on the responses to forced changes and a detailed analysis into the responses to internal variability is outside the scope of this article.

However, we have added more discussion on the strengths and weaknesses of different ARDTs. We have further compared the mean AR frequencies obtained from 8 other ARDTs in present climate. Two of the ARDTs are applied to study the responses of ARs to greenhouse gas warming (see Fig. S2, S3, S8).

Line 125 – change "Compared to understanding regarding AR characteristics…" to "Compared to physical intuition regarding AR characteristics…"

Intuition is defined as *the ability to understand something immediately, without the need for conscious reasoning*. So I have changed the word *understanding* to *interpreting*. The final sentence is as follows.
Compared to interpreting AR characteristics in the present climate, analyzing projected AR changes in future climate becomes even more complicated as methods can be based on historical thresholds (i.e., absolute or fixed relative schemes) or based on mean state thresholds (i.e., relative schemes) [34].

Line 129-137 – Eloquently said. These are all very important points that need to be more discussed in the literature!

Thank you.

Line 138-141 – Could the authors provide a bit more specificity on how machine learning/image processing techniques could get around issues of embedding historical stationarity into ARDT estimates of AR frequency/intensity/spatial extents (e.g., fixed threshold of 250 kg/m/s)?

As mentioned in the sentence these are alternate methods, where one doesn't have to use a threshold on climate variables. Thus, they get rid of the discussion on how to select a threshold.

Whether ARs should be detected based on historical thresholds or mean state dependent thresholds is more fundamental and cannot be addressed fully by using these approaches. Such ambiguities exists in the detection of not just ARs but any other climate feature. However, image-processing, machine learning or topological methods provides us an alternate way to identify ARs. For example Ullrich and Zarzycki (2017) applied the threshold not on a climate variable but on the Laplacian of IVT. This makes the results less sensitive to the mean state changes while being consistent with other ARDTs in the present climate. Muszynski et al. (2019) did not apply any thresholds but used pattern recognition to label ARs. The image-processing study mentioned also did not set a hard threshold for climate variable (Xu et al., 2020). Instead, they use top-hat reconstruction technique to identify ARs. All these methods are examples of detection of ARs with less sensitivity to climate thresholds. They cannot give a theoretical definition for ARs, but presents a different way of identifying them and understanding the changes.

Line 151 – is "ultra-high-resolution" a common term? I'm a global/regional climate modeler and I haven't heard of this term before. It made me initially think the model would be run at convective/storm resolving scales such as the DYAMOND experiments -`https://progearthplanetsci.springeropen.com/articles/10.1186/s40645-019-0304-z`- yet the atmosphere is at 0.25 deg. I recognize the ocean model being run at 0.1 deg is no small feat, but the authors might consider revising this term throughout or provide better justification for why this term is used given the community is now calling the DYAMOND experiments "digital twins" (which is analogous to "ultra-high-resolution").

The DYAMOND experiements/simulations are termed as cloud resolving simulations.

Generally high resolution refers to climate model simulations up to 0.25^o. The term ultra-high resolution is used to convey that we use a higher resolution. The term is used to represent the experients conducted in our center (see Ultra-high-resolution climate simulation project – IBS Center for Climate Physics. Other studies using the same datasets also refer it by the same name. It would be great if we can keep the consistent nomenclature.

Line 185-186 – could you guide the reader to the figure and subpanel label to back this claim up (i.e., Figure 1c)?

Thank you for pointing it out. I have added the reference to the figure. The sentence now reads.

The most notable difference between ERA5 and the CESM UHR PD simulation is the location of the maximum AR frequency and associated precipitation (Fig. 1c, f)

Line 186-188 – could it also be the jet stream strength too? What about other drivers such as the MJO? See Zhou et al., 2021 - `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033711` - for more information. Also, do the authors posit that ERA5 might have biases too (e.g., too dissipative in estimating wind speed/specific humidity strength/variability) and could they provide a sentence or two discussing how reanalysis dataset choice might also shape model historical bias estimation?

Since the ARDT is not sensitive to the magnitude of IVT, we do not think the jet stream strength is an important factor in the discrepancy. As shown in Fig. S6 the zonal mean of mean AR frequency shows similar latitudinal shifts as the maximum values of wind speed and column integrated eddy kinetic energy.

CESM 1.2 does not simulate MJO very realistically (Small et al., 2014; Chang et al., 2020). Interpreting the relationship and the changes would then require to go beyond the scope of the manuscript. Hence, we have not included the analysis.

It is expected that ERA5 will have some biases. However, the differences in mean AR frequency between ERA5 and MERRA2 shows very little differences (Fig. S7). This leads us to think that the choice of the reanalysis dataset will not affect the historical biases for mean AR characteristics. A deep dive into the reasons for these differences between reanalysis datasets are outside the scope of the paper. Also, a recent study already discusses the issue (Collow et al., 2022).

A new sentence and reference have been added to convey that choice reanalysis datasets will not affect the historical biases.

Fig. S2 – please add in the caption of the figure which line corresponds to ERA5 vs UHR CESM

Thank you for pointing out the error. The figure and caption has been modified.

Line 192-194 – could the authors provide a citation or two that discuss these various "components" of AR conditions (e.g., WCBs, ECs, AFs, etc.) that non-AR experts or new AR researchers could read to spin up on these terms? For example, I've found this to be a great resource - `http://cw3e.ucsd.edu/ARconf2016/Presentations/Wednesday/S3_TheoryStructureProcesses1/Cordeira_IARC2016.pdf` – but I'm not sure if Cordeira has published a study highlighting these "components" of AR conditions (see Part 2).

The following new references have been added. Dacre et al. (2015, 2019) explains the various components of ARs. The new sentence reads.
From a cyclone-centered view, ARs can be viewed as vestiges of a poleward moving extratropical storm [71, 72], with robust co-occurrences and moderate correlations between AR and storm strength [73, 74].

196-197 – Patricola et al., 2021 - `https://www.sciencedirect.com/science/article/pii/S2212094722000275` - could provide some further guidance on proportionality of EC/AR relationship and how the presence of an EC shapes AR characteristics. This might be worth a sentence or two of discussion.

The following new references are added. Zhang et al. (2019); Guo et al. (2020) discusses the statistical relationship between extratropical cyclones and ARs. The new sentence reads.
From a cyclone-centered view, ARs can be viewed as vestiges of a poleward moving extratropical storm [71, 72], with robust co-occurrences and moderate correlations between AR and storm strength [73, 74].

Line 218-237 and 275-277 –I'm not sure what the global temperature response in UHR CESM is at 2XC02 or 4XC02, but I think this would be important to include throughout the manuscript and maybe in figure headers, especially when comparing with RCP8.5 scenario literature. With that said, the global mean change in AR frequency seems very consistent with Clausius-Clapeyron (6-7%/K) in the 2XCO2 and 4XCO2 experiments. Do the authors want to comment on how the dynamical response of 2XCO2 and 4XCO2, at least in

Fig. 2, seems to play less of a role than the thermodynamic response? Figure 3 in Payne et al., 2020 https://www.nature.com/articles/s43017-020-0030-5– makes it seem like the dynamic component under climate change (e.g., storm track shift poleward, slow down of jet/winds due to equator-to-pole temperature gradient reduction, etc.) would be more equal and opposite to the thermodynamic component under climate change. Might be worth mentioning that the strength/direction of dynamic vs thermodynamic component changes in IVT with warming could be model dependent too?

Scaling changes with mean surface temperature changes is a good idea. It has been implemented in the revised manuscript. See lines 342-350.

The analysis for the revised Fig. 2 has been modified to better show the relationship between each component and mean AR changes. More discussion on the IVT decomposition method and its caveats and applicability to different ARDTs are included in the text (see lines 303-325).

The changes in thermodynamic and especially the dynamic component will be model dependent. We have mentioned that in the corresponding section (see lines 314-340).

Line 218-237 –Rhoades et al., 2020 -(https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020GL089096) – found a decrease in mean climatological AR frequency over the North Pacific under RCP8.5 with CESM, but an uptick in the number of landfalling ARs and, importantly, the strength/lifetime of the ARs. Mean AR frequency is an important metric, but landfalling AR frequency (particularly extreme ARs) are the most important metric from an impacts' perspective, particularly under 2XC02/4XC02/RCPs/SSPs. Could the authors highlight how the strength/intensity/duration of landfalling AR events shift under different levels of CO2/warming? Another neat metric that the authors could assess (thanks to the impressive decomposition framework they have generated!) would be to assess the orthogonality of landfalling ARs which is particularly important when assessing the orographic uplift potential (which has been found to better correlate with AR total precipitation - https://journals.ametsoc.org/view/journals/hydr/23/9/JHM-D-21-0119.1.xml at AR landfall in regions where large coastal mountains are (e.g., western North America and South America)?

While detailed analysis on to the landfalling ARs and their extreme precipitation would be very interesting, we have not included it in our manuscript as any useful discussion on to their impacts on coastal area would go beyond the word limit of the journal. This could be a future analysis just by focussing on the most extreme of the ARs.

To partially address the comment, we have presented a global map illustrating the changes in vertical velocity at 500 hPa, as depicted in Fig. S16.

Fig. S3c,4c,5c,6c – change “Themal” to “Thermodynamic”

The change has been accepted and required modifications has been made.

Fig. S3e,4e,5e,6e – change “Thermal” to “Thermodynamic”

The change has been accepted and required modifications has been made.

Line 266-273 – “. . . AR frequency is largely modulated by DIVT. . . ”, could the authors provide some quantitative evidence for this statement (maybe using spatial pattern correlations or ? between IVTy and the thermo/dynamic components) At the moment, the reader is asked to do an “eyeball” test between subpanel a) with b) and c) and in some regions this statement about DIVT doesn’t seem to hold true (but maybe at the global scale).

All other reviewers had similar concerns. Considering the comments, we have modified our analysis which now shows the vector projection of dynamic and thermodynamic components of IVT changes onto the mean IVT from PD. Thus, in the revised Fig 2e-f positive values indicate that the components act to reinforce the mean IVT from present climate while negative values act to weaken the prevailing IVT patterns in PD. It is now clearer from the figure that that regions that show a reduction in mean AR frequencies are characterized by the DIVT acting to reduce the mean transport. Regions showing large increases are characterized by both DIVT and TIVT reinforcing the prevailing transport. The conclusions are true for ARDTs like SCAFET and TempestExtremes that are less sensitive to global mean increases in humidity. However as discussed in section 2.2, the conclusions may vary depending on ARDTs. Further discussion is also added in the text from lines 314-340.

Line 295-329 – very thorough and striking analysis, nice work! Did the authors assess rain-snow partitioning of ARs too? Would it be easy enough to recreate Fig. 4 but compare and contrast precipitation extremes by precipitation phase (i.e., extreme snowfall events vs rainfall events)? This would be a critical analysis from an AR impacts perspective (e.g., flood and landslide hazards and water resource management of downstream reservoirs due to rain-on-snow/spring snow melt reduction), particularly if done for all the major AR terminus regions (as shown in Fig. 3 of Payne et al., 2020 - `https://www.nature.com/articles/s43017-020-0030-5`.

The additional analysis is shown in Figure 1. As the snow-rain partitioning in CESM UHR PD is not validated against observed or reanalysis dataset, further discussion on it is not added in the main manuscript.

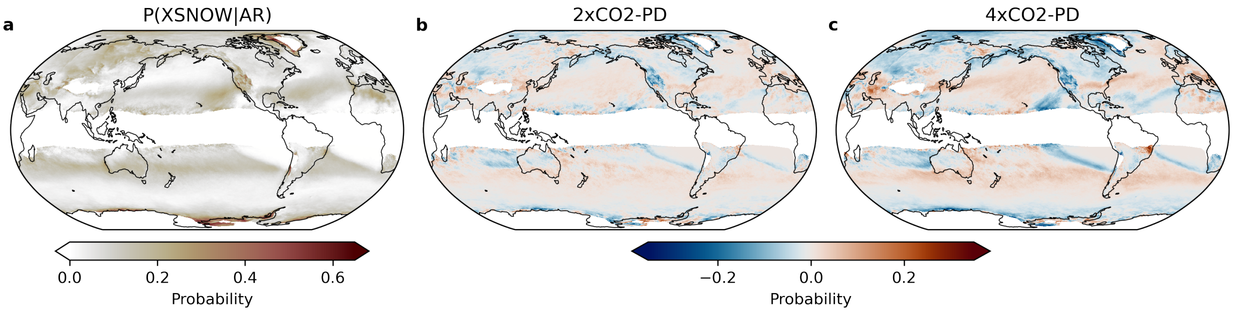

Figure 1: **Changes in the probability of extreme snowfall events given the presence of AR**. (a) The probability of extreme snowfall (95th percentile of daily mean snowfall) given the presence of an AR, P(XSNOW|AR) for (a) CESM-UHR PD. Changes in the probability of AR-induced extreme snowfall for 2CO2 (c) and 4CO2 (d) simulations.

Line 337-338 –could you provide your % change values from your findings here and the literature range of % change values so that the reader can determine "significant" more easily (especially if a reader was to quickly read abstract/conclusions sections)?

The sentence has been modified to include quantitative results for the changes in AR frequency and its contribution to mean annual precipitation. The sentence now reads. The results are consistent with previous studies indicating an increase in the mean frequency (by 6% and 12%) and annual mean precipitation associated with ARs (by 6% and 15%) in response to increased atmospheric CO_2 concentrations (for 2CO2 and 4CO2 simulations).

Line 355-356 – what is meant by "AR circulations systems"?

By AR circulation system, we mean the components of an AR like the WCBs, AFs and the cyclones as described in Dacre et al. (2015). The reference is already added in the section 2.1

Line 364 – change "though this is admittedly difficult to globally" to "though this is admittedly difficult to systematically apply globally"

Accepted.

Line 456 – I think it would be helpful if the authors highlight the step-by-step AR extraction using the workflow outlined for SCAFET in the Methods (Fig. S8) earlier on in the study (maybe on line 164-165). Also, how does the SCAFET workflow differ from TempestExtremes and some of the algorithms outlined in Inda-Diaz et al., 2021 - `https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2020JD033746`?

Comparison to other ARDTs, especially to TempestExtremes was suggested by other reviews too. Hence I have added a new supplementary section to compare ARDTs in present climate (Fig. S2) high resolution datasets (Fig. S3) and in enhanced GHG conditions (Fig. S8, S10). Futher discussion have been added to the manuscript in lines 200-216.

For more theoretical comparison between the methods, please refer to "Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction from large climate datasets" (`https://egusphere.copernicus.org/preprints/2023/egusphere-2023-592/`). The reference to the preprint is also added to the manuscript.

Line 460 – could the authors provide a citation (or two) explaining the Gaussian filter used?

The Gaussian filter is a standard method used in many other ARDTs including Ullrich and Zarzycki (2017); Xu et al. (2020). Here we have used a slightly modified version of the Gaussian smoothing from MetPy (`https://unidata.github.io/MetPy/latest/api/generated/metpy.calc.smooth_gaussian.html`).

Line 479 – wouldn't this minimum mean precipitation criteria (1 mm/day) bias the AR detection (particularly in a warmer world where these precipitation rates might systematically increase)? How sensitive are estimates when not using or altering this criterion, especially when applying this to sub-daily fields (since its fixed at the daily scale, mm/day)?

This is an important point. Thank you for pointing it out. Yes. The estimates will be sensitive to the precipitation criteria. However, in this study the value for minimum precipitation is kept at $1mm/day$. This criterion is just to make sure that the AR structures are precipitating. Even if we remove this criterion (i.e, no thresholds on IVT and precipitation), we essentially get the same result as putting a $1mm/day$ threshold on precipitation (see Fig. S9). Interestingly, increasing the threshold will result in the AR frequency changes becoming similar to ARDTs that use a fixed relative or absolute thresholds. We have discussed it with in the manuscript in lines 289-293

Line 491-494 – could you define ps and pt (some AR researchers don't integrate through the entire atmosphere, while others do)? I assume you followed the same workflow as in ERA5 in Section 5.1, but just thought it would be important to be clear to readers.

Clarified the values of p_s and p_t. The sentence now reads.

We have used the same pressure levels as in the calculation of IVT. That is, $p_s = 1000$ hPa and $p_t = 300$ hPa.

IBS Center for Climate Physics
Rooom No. 1105, IBS Center for Climate Physics
Busandaehak-ro 63beon-gil 2 (Jangjeon-dong),
Geumjeong-gu, Busan 46241, Republic of Korea
T +82-51-510-7636
F +82-51-510-7694
Email: arjunbabun@pusan.ac.kr
Web: ibsclimate.org

3 Response to Reviewer #3

Re: Re-submission of manuscript COMMSENV-22-1342-T

Dear Reviewer,

We appreciate the effort and time you put into going through the manuscript. Your suggestions have proven to be invaluable in improving the quality of the manuscript. Thank you.

A point-by-point response to all the remarks is attached. A PDF document, highlighting the changes between the two versions of the manuscript is also attached for your convenience.

The suggestions have been discussed and deliberated with all authors. We hope that the revisions can convince you to recommend the publication of the article in *Communications Earth and Environment*. We would be happy to provide further corrections and clarifications.

3.1 Reviewer Comments

Recommendation: Overall, I found the study interesting and the ARDT method has the potential to answer important research questions regarding future ARs and implications for precipitation. However, the analysis of dynamical versus thermodynamic contributions to future AR probabilities was not presented as clearly as needed. It wasn't clear to me how some of the figures supported the interpretations and conclusions.

3.2 Revision Highlights

- Novelty and intercomparison of the ARDT.
  - AR event detection and mean spatial distribution obtained from 8 other ARDTs are compared with SCAFET.
  - ARs detected from CESM UHR using SCAFET is compared with Guan and Waliser (2015) and Ullrich and Zarzycki (2017).

- Included new reference describing the detection framework. "Scalable Feature Extraction and Tracking (SCAFET): A general framework for feature extraction from large climate datasets" (`https://egusphere.copernicus.org/preprints/2023/egusphere-2023-592/`).

- Better attribution of thermodynamic and dynamic components of IVT changes to mean changes in AR frequency. Additional analysis has been conducted and Fig. 2 has been modified in the manuscript.
- The AR frequencies and characteristics obtained from low resolution and high resolution simulations are compared.
- Code and data availability
  - Latest version of SCAFET can be found in GitHub (`https://github.com/nbarjun/SCAFET/`)
  - The codes and sample datasets used for the manuscript can be downloaded from `10.5281/zenodo.6979522`

3.3 Point-by-Point Response

Abstract: The abstract states that the "mean changes in AR frequency in a warmer climate is mainly driven by the underlying synoptic scale circulation systems". I'm not sure this is generally considered to be true, and other studies have concluded that thermodynamic effects were more important. In this study you tried to separate thermodynamic from dynamical changes, but it wasn't clear that dynamical changes were the primary cause of the changes in AR frequency. Perhaps I don't fully understand the interpretation of Figures S3-S6.

The reviewer is right in pointing out that the conclusion that "mean changes in AR frequency in a warmer climate is driven by the underlying synoptic scale circulation systems" is not universally accepted. The conclusion will depend on the Atmospheric River Detectors (ARDTs) employed to identify ARs as described in Section 2. ARDTs that use an absolute or fixed relative threshold will conclude that the changes in AR frequency is driven by thermodynamic effect (see Fig. S8a,b). While ARDTs which account for the mean state changes like TempestExtremes, TECA-BARD, and SCAFET identify dynamics effects to be driving the AR frequency changes (see Fig. 2, Fig. S8c,d).

As you pointed out there is not enough evidence that the changes in AR frequency are due to *synoptic* circulation systems. We have modified the abstract and the manuscript to say that the changes in AR frequencies are driven by the in *changes in global atmospheric circulation systems* as discussed in section 2.2. The new sentence now reads,

Mean changes in AR frequency under global warming are tightly linked to changes in global atmospheric circulation systems, while AR properties like precipitation and water vapor transport follow the Clausius-Clapeyron scaling.

The section 2.2 and Fig. 2 has been modified to better show the influence of thermodynamic and dynamic component of IVT on the mean IVT changes. We have also tried to explain the caveats of this analysis by discussing the inferences obtained from two other ARDTs too (see lines 314-340).

Abstract Lines 43-44: Is the increased AR amplitude also associated with changes in synoptic scale circulation? I don't think you showed this.

The reviewer is right that the evidence was not enough to make that statement. With the new analyses in section 2.2 and the modified Fig. 2 we can show that the changes in mean AR frequency is primarily modulated by how the dynamic component of IVT changes influence the mean IVT in present day. Which leads us to conclude that the changes in AR frequency observed by applying SCAFET is due to the changes in mean circulation changes. We have also pointed out that these conclusions may not hold if one uses an ARDT that uses a historical thresholds (Fig S8a, b).

Lines: 220-221: The wording suggests that this finding comes by surprise when the result is what you would expect given the design of the two different algorithms. Maybe remove the word "interestingly".

Agreed. The wording has been changed. The new sentence reads the following

Estimated global mean changes in AR frequency are considerably less than those cited in previous studies

Lines 235-237: You state that "threshold-dependent ARDTs may overestimate AR activity as they do not account for a future mean state that has a higher specific humidity". I see where you're coming from, but whether this is true or not depends on the definition of an AR, which historically does have an IVT magnitude requirement. It's true that you trigger this threshold more often in a warmer climate even if the dynamics do not change but I don't know if I would call it an "overestimate" without clarifying this point. I agree that SCAFET is useful for understanding thermodynamic versus dynamical changes in relation to ARs.

We understand how "overestimate" may sound like an error in other ARDTs. We had removed the wording. The whole section has also been revamped to include discussion on ARDT dependence of the inferences discussed (see lines 314-340).

Figure S2: There is no legend to identify the color scale.

Sorry about the oversight. A new legend has been added. We have also modified the caption to make the message clearer. Thank you for bringing it to our attention.

Lines 266-269: I don't really see how the conclusion that "AR frequency changes are largely modulated by DIVT" is evident from Figures S4 and S6. The zonal change in IVT seems to be mostly from thermodynamic effects (Figure S3) so I'm confused about how AR frequency is attributed to dynamical changes.

We understand the IVT decomposition is a little confusing. Other reviewers have also raised similar concerns. We have modified our analysis to make the relationship between AR frequency changes and the dynamic and thermodynamic components of IVT clearer. In the updated methodology instead of relying on the reader's knowledge on prevailing circulations to figure out the relationship from Figs 2 a,b,e-h, we show the vector projection of the thermodynamic (TIVT) and dynamic component (DIVT) of IVT onto the mean IVT. It is much clearer from Fig. 2 that regions that show a reduction in mean AR frequencies are characterized by the DIVT acting to reduce the mean transport. Regions showing large increases are characterized by both DIVT and TIVT reinforcing the prevailing transport. The conclusions are true for ARDTs like SCAFET and TempestExtremes that are less sensitive to global mean increases in humidity. However, as discussed in section 2.2, the conclusions may vary depending on ARDTs.

Figure 3: Some of the changes discussed do not look significant. For example, the 1.6% and 2.3% increase in duration. Is this change significant? Except for 3a, these changes are very small, and I wonder if they would fall within the range of uncertainty in a multi-ensemble or multi-model experiment.

It is true that some changes discussed does not look significant. However, due to the large number of samples, simple statistical tests shows significant changes. Due to the design of the algorithm without any thresholds on IVT, we tend to identify AR structures in present as well as warming simulations. Consequently, it can be inferred that these AR structures exist both in the current climate and projected future climate. However, as the climate continues to warm, these structures carry a greater amount of water vapor, resulting in increased precipitation (refer to Fig. 2 and Fig. 3). The observed changes become less noisy when more stringent thresholds are applied (see Fig. S9). Moreover, the presence of noise in the results is expected due to the inherent sensitivity of AR detection in SCAFET to changes in the mean circulation patterns (refer to Fig. 2). As numerous prior studies have discussed, the changes in circulation patterns as a response to greenhouse warming exhibit significant uncertainty, which consequently affects the detected changes in ARs. If we apply a fixed or fixed relative threshold, the changes in AR frequency becomes more coherent, while the changes in mean AR properties like IVT or precipitation become less so.

Similar to other ARDTs that are less sensitive to mean state changes, geometric characteristics like length, width, and area does not show large changes. But the properties within ARs like upward velocity, precipitation, IVT etc. show large changes.

While we agree that looking at a multiple ensemble or models would be helpful to clear this issue, running high resolution multi-ensemble simulations are currently infeasible with the computing resources we have. To justifiably answer the question, one as to implement multiple ARDTs on large ensemble data. This is because similar studies have already been done using large-ensembles and multimodel and multi-ARDTs in low resolution datasets (Hagos et al., 2016; O'Brien et al., 2020). We wanted to limit the scope of the study to changes in AR characteristics due to increased greenhouse gas concentration on a high resolution simulation. To partially answer this question, we have used CESM2 GCM output to identify changes in AR frequency in response to anthropogenic warming.

Lines 353-358: Has the SCAFET methodology been tested on coarser resolution GCM output? I agree with the point about the importance of high resolution for precipitation simulations, but I'm curious about the benefit for representing the large-scale circulation associated with the AR-like features you're detecting with the algorithm. Having multiple model output would be beneficial for climate change studies.

Yes. SCAFET has been tested in different datasets with varying resolution to detect ARs. For instance, SCAFET has now contributed to the tier 1 experiments of ARTMIP, and Paleo ARTMIP. Tier 1 experiments uses MERRA2 reanalysis data with a spatial resolution of $\approx 0.5^o$ while Paleo ARTMIP uses $\approx 1^o$. To be more self-contained, we have shown the results obtained from CMIP5 and CMIP6 datasets from NCAR (CCSM4 for CMIP5 and CESM2 for CMIP6) with that obtained from ERA5 and CESM-UHR for present climate. The results indicate that SCAFET can identify ARs insensitive to resolution. Other inferences that can be drawn from comparing Low Resolution (LR) and high resolution (HR) datasets are.

- As ARs are large-scale phenomenons, there is no significant difference in the spatial distribution of AR frequency. In a way this analysis also shows that the AR detection using SCAFET is resolution independent.
- ARs in HR simulations tend to give narrower ARs. This is similar to some previously published works, which also found HR datasets to have smaller widths (Collow et al., 2022).
- The maximum convection strength and maximum precipitation intensity in LR ARs are much smaller than that in HR ARs. Similar inferences were made in previous studies too (Collow et al., 2022; Chang et al., 2020; Shields et al., 2023).

Studies using CMIP5 and CMIP6 simulations have already been conducted and might not add new insights now that we have shown SCAFET is insensitive to resolution and performs similar to existing ARDTs that use a relative threhold. However, we have also shown the changes in mean AR frequency in response to anthropogenic warming captured using SCAFET in CESM2. The results show similar changes (see Fig. S10). The changes in AR distribution are dependent on the mean circulation changes in CESM2 in response to warming. Thus, a different model simulation might give us a different output.

Figure S7. It would helpful so show where these changes are statistically significant.

We have applied student t-test. The figures have been modified.

Figure 3. There is no description in the caption for panel e.

The panel e in Figure 3 is described in the caption.

Figure 2 a,b and lines 214-218: It seems worth mentioning that in many of the regions listed, these changes are not significant under 2CO2. For the southwestern US, they are also insignificant under 4CO2.

Agreed. The lack of significance in comparison with other ARDTs are discussed in the next paragraph (lines 257-258).

References

Ping Chang, Shaoqing Zhang, Gokhan Danabasoglu, Stephen G. Yeager, Haohuan Fu, Hong Wang, Frederic S. Castruccio, Yuhu Chen, James Edwards, Dan Fu, Yinglai Jia, Lucas C. Laurindo, Xue Liu, Nan Rosenbloom, R. Justin Small, Gaopeng Xu, Yunhui Zeng, Qiuying Zhang, Julio Bacmeister, David A. Bailey, Xiaohui Duan, Alice K. DuVivier, Dapeng Li, Yuxuan Li, Richard Neale, Achim Stössel, Li Wang, Yuan Zhuang, Allison Baker, Susan Bates, John Dennis, Xiliang Diao, Bolan Gan, Abishek Gopal, Dongning Jia, Zhao Jing, Xiaohui Ma, R. Saravanan, Warren G. Strand, Jian Tao, Haiyuan Yang, Xiaoqi Wang, Zhiqiang Wei, and Lixin Wu. An unprecedented set of high-resolution earth system simulations for understanding multiscale interactions in climate variability and change. *Journal of Advances in Modeling Earth Systems*, 12(12), dec 2020. doi: 10.1029/2020ms002298. URL `https://doi.org/10.1029%2F2020ms002298`.

A. B. Marquardt Collow, C. A. Shields, B. Guan, S. Kim, J. M. Lora, E. E. McClenny, K. Nardi, A. Payne, K. Reid, E. J. Shearer, R. Tomé, J. D. Wille, A. M. Ramos, I. V. Gorodetskaya, L. R. Leung, T. A. O'Brien, F. M. Ralph, J. Rutz, P. A. Ullrich, and M. Wehner. An overview of ARTMIP's tier 2 reanalysis intercomparison: Uncertainty in the detection of atmospheric rivers and their associated precipitation. *Journal of Geophysical Research: Atmospheres*, 127(8), apr 2022. doi: 10.1029/2021jd036155. URL `https://doi.org/10.1029%2F2021jd036155`.

H. F. Dacre, P. A. Clark, O. Martinez-Alvarado, M. A. Stringer, and D. A. Lavers. How do atmospheric rivers form? *Bulletin of the American Meteorological Society*, 96(8):1243–1255, aug 2015. doi: 10.1175/bams-d-14-00031.1. URL `https://doi.org/10.1175%2Fbams-d-14-00031.1`.

H. F. Dacre, O. Martínez-Alvarado, and C. O. Mbengue. Linking atmospheric rivers and warm conveyor belt airflows. *Journal of Hydrometeorology*, 20(6):1183–1196, jun 2019. doi: 10.1175/jhm-d-18-0175.1. URL `https://doi.org/10.1175%2Fjhm-d-18-0175.1`.

Bin Guan and Duane E. Waliser. Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research: Atmospheres*, 120(24):12514–12535, dec 2015. doi: 10.1002/2015jd024257. URL `https://doi.org/10.1002%2F2015jd024257`.

Yanjuan Guo, Toshiaki Shinoda, Bin Guan, Duane E. Waliser, and Edmund K. M. Chang. Statistical relationship between atmospheric rivers and extratropical cyclones and anticyclones. *Journal of Climate*, 33(18):7817–7834, sep 2020. doi: 10.1175/jcli-d-19-0126.1. URL `https://doi.org/10.1175%2Fjcli-d-19-0126.1`.

Samson M. Hagos, L. Ruby Leung, Jin-Ho Yoon, Jian Lu, and Yang Gao. A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western north america from the large ensemble CESM simulations. *Geophysical Research Letters*, 43(3):1357–1363, feb 2016. doi: 10.1002/2015gl067392. URL `https://doi.org/10.1002%2F2015gl067392`.

Grzegorz Muszynski, Karthik Kashinath, Vitaliy Kurlin, and Michael Wehner and. Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets. *Geoscientific Model Development*, 12(2):613–628, feb 2019. doi: 10.5194/gmd-12-613-2019. URL `https://doi.org/10.5194%2Fgmd-12-613-2019`.

Travis A. O'Brien, Mark D. Risser, Burlen Loring, Abdelrahman A. Elbashandy, Harinarayan Krishnan, Jeffrey Johnson, Christina M. Patricola, John P. O'Brien, Ankur Mahesh, Sarahí Arriaga Ramirez, Alan M. Rhoades, Alexander Charn, Héctor Inda Díaz, and William D. Collins and. Detection of atmospheric rivers with inline uncertainty quantification: TECA-BARD v1.0.1. *Geoscientific Model Development*, 13(12):6131–6148, dec 2020. doi: 10.5194/gmd-13-6131-2020. URL `https://doi.org/10.5194%2Fgmd-13-6131-2020`.

Christine A Shields, Ashley E Payne, Eric Jay Shearer, Michael F Wehner, Travis Allen O'Brien, Jonathan J Rutz, L Ruby Leung, F Martin Ralph, Allison B Marquardt Collow, Paul A Ullrich, Qizhen Dong, Alexander Gershunov, Helen Griffith, Bin Guan, Juan Manuel Lora, Mengqian Lu, Elizabeth McClenny, Kyle M Nardi, Mengxin Pan, Yun Qian, Alexandre M Ramos, Tamara Shulgina, Maximiliano Viale, Chandan Sarangi, Ricardo Tomé, and Colin Zarzycki. Future atmospheric rivers and impacts on precipitation: Overview of the ARTMIP tier 2 high-resolution global warming experiment. *Geophys. Res. Lett.*, 50(6), March 2023.

R. Justin Small, Julio Bacmeister, David Bailey, Allison Baker, Stuart Bishop, Frank Bryan, Julie Caron, John Dennis, Peter Gent, Hsiao ming Hsu, Markus Jochum, David Lawrence, Ernesto Muñoz, Pedro diNezio, Tim Scheitlin, Robert Tomas, Joseph Tribbia, Yu heng Tseng, and Mariana Vertenstein. A new synoptic scale resolving global climate simulation using the community earth system model. *Journal of Advances in Modeling Earth Systems*, 6(4):1065–1094, dec 2014. doi: 10.1002/2014ms000363. URL `https://doi.org/10.1002%2F2014ms000363`.

Paul A. Ullrich and Colin M. Zarzycki. TempestExtremes: a framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, 10(3):1069–1090, mar 2017. doi: 10.5194/gmd-10-1069-2017. URL `https://doi.org/10.5194%2Fgmd-10-1069-2017`.

Paul A. Ullrich, Colin M. Zarzycki, Elizabeth E. McClenny, Marielle C. Pinheiro, Alyssa M. Stansfield, and Kevin A. Reed. TempestExtremes v2.1: A community framework for feature detection, tracking and analysis in large datasets. jan 2021. doi: 10.5194/gmd-2020-303. URL https://doi.org/10.5194%2Fgmd-2020-303.

Guangzhi Xu, Xiaohui Ma, Ping Chang, and Lin Wang. Image-processing-based atmospheric river tracking method version 1 (IPART-1). *Geoscientific Model Development*, 13 (10):4639–4662, sep 2020. doi: 10.5194/gmd-13-4639-2020. URL https://doi.org/10.5194%2Fgmd-13-4639-2020.

Zhenhai Zhang, F. Martin Ralph, and Minghua Zheng. The relationship between extratropical cyclone strength and atmospheric river intensity and position. *Geophysical Research Letters*, 46(3):1814–1823, feb 2019. doi: 10.1029/2018gl079071. URL https://doi.org/10.1029%2F2018gl079071.

30th Jun 23

Dear Mr Nellikkattil,

Your manuscript titled "Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high-resolution greenhouse warming simulations" has now been seen by our reviewers, whose comments appear below. In light of their 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 address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript 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,

Heike Langenberg, PhD
Chief Editor
Communications Earth & Environment

Min-Hui Lo, PhD
External Editor
Communications Earth & Environment

On Twitter: @CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

I appreciate the time and effort invested by the authors in revising the manuscript, particularly the comparison of results using different ARDTs. It is crucial for readers to understand the main source of uncertainty. The manuscript has seen significant improvement, and most of my previous comments have been adequately addressed. I only have a few minor comments:

Minor comments:

Abstract: Considering the sensitivity of the results to ARDTs, I am wondering if the statement "Mean changes in AR frequency under global warming... follow the Clausius-Clapeyron scaling" is still appropriate?

Line 266: Are you suggesting that the amplitude doubles from the 2XCO2 simulation to the 4XCO2 simulation?

Eq. 5: It would be helpful if the author could provide further elaboration on the SI index and explain the additional information it offers to readers compared to traditional ARDTs.

Eq. 6: I find the decomposition in this equation quite confusing. The use of \bar, which represents the monthly mean, may not necessarily reflect the change in AR. I encounter similar confusion with equations 11 and 12, as the changes in dynamics and thermodynamics are defined based on monthly timescales rather than AR timescales.

Eq. 8: There seems to be a typographical error in the last term on the right-hand side. It should be \bar{q_1'u_1'}. The "prime" symbol is missing.

Eq. 8: I am wondering if the change in the eddy term across different climate states is negligible, as numerous studies have suggested that a decrease in the equator-to-pole temperature gradient can effectively reduce eddy activity, leading to weaker baroclinic waves.

Reviewer #2 (Remarks to the Author):

I sincerely appreciate the thoroughness of the authors in responding to the many reviewer comments. The point-by-point responses were extremely organized and well-reasoned. I have no further comments.

Increased amplitude of atmospheric rivers and associated extreme precipitation in ultra-high-resolution greenhouse warming simulations

[Reply to Reviewers]

Arjun Babu Nellikkattil[1,2], June-Yi Lee[1,2,3], Bin Guan[4,5], Axel Timmermann[1,6], Sun-Seon Lee[1,6], Jung-Eun Chu[7], and Danielle Lemmon[1,6,8]

[1]Center for Climate Physics, Institute for Basic Science
[2]Department of Climate System, Pusan National University
[3]Research Center for Climate Sciences, Pusan National University
[4]Joint Institute for Regional Earth System Science and Engineering
[5]Jet Propulsion Laboratory, California Institute of Technology
[6]Pusan National University
[7]Low-Carbon and Climate Impact Research Centre, School of Energy and Environment, City University of Hong Kong
[8]Office of Clean Energy Demonstrations, Department of Energy, Washington D.C., United States

Contents

IBS Center for Climate Physics

Rooom No. 1105, IBS Center for Climate Physics
Busandaehak-ro 63beon-gil 2 (Jangjeon-dong),
Geumjeong-gu, Busan 46241, Republic of Korea

T +82-51-510-7636
F +82-51-510-7694

Email: arjunbabun@pusan.ac.kr
Web: ibsclimate.org

1 Response to Reviewer #1

Re: Re-submission of manuscript COMMSENV-22-1342A

Dear Reviewer,

Thank you for your valuable comments and suggestions on our manuscript. We appreciate the time and effort you invested in reviewing our work and providing thoughtful feedback. Your insights have been instrumental in improving the quality and reliability of our research. We have carefully addressed each of the concerns you raised and made revisions to strengthen the main results and interpretations.

In response to your queries, we have included a point-by-point response addressing all the remarks raised during the review process. Additionally, we have attached a PDF document that highlights the changes made between the initial submission and the revised version for your convenience.

We believe that the current version of the manuscript meets the standards for publication in *Communications Earth and Environment*. Nevertheless, we remain open to further revisions if needed to ensure the utmost accuracy and clarity of our work.

1.1 Revision Highlights

- The abstract has been modified to include suggestion from reviewer #1 and to limit the word count to 150.
- A new subfigure showing the change in distribution of maximum intensity of precipitation within ARs is added to Figure 3. The inferences are discussed in the abstract as well as the main manuscript.
- Methods section is rearranged to comply with the journal recommendation of "only one layer of subheadings".
- The mathematical notation of variables and reference to supplementary figures has been modified to comply with journal requirements.

1.2 Reviewer Comments

Recommendation: The manuscript has seen significant improvement, and most of my previous comments have been adequately addressed. I only have a few minor comments.

1.3 Point-by-Point Response

1.3.1 Minor Comments

Abstract: Considering the sensitivity of the results to ARDTs, I am wondering if the statement "Mean changes in AR frequency under global warming ... follow the Clausius-Clapeyron scaling" is still appropriate?

The reviewer is right in pointing out the lack of clarity. The abstract has been now modified to only include the AR characteristic changes which are largely insensitive to the detection method. To justify the changes, an additional subfigure has been added to the Figure 3. The addition is not just in response to your comments but also fits well with some recent publications. The following additions are made.

- Lines 350-363 in Section 2.2
- Lines 415-418 in Section 3

Line 266: Are you suggesting that the amplitude doubles from the 2XCO2 simulation to the 4XCO2 simulation?

Not exactly. Compared to 2xCO2 conditions, the global mean amplitude in 4xCO2 shows an increase by a factor of 1.8. By linear response, we wanted to bring the reader's attention to the fact that regions that show increased/decreased AR frequencies in response to 2xCO2 forcing also show similar changes with higher amplitude in response to 4xCO2 forcing. The response is indeed linear, however the strength of the relationship depends on the region considered.

Eq. 5: It would be helpful if the author could provide further elaboration on the SI index and explain the additional information it offers to readers compared to traditional ARDTs.

The following additional sentences have been included to convey the added value of shape based AR detection compared to traditional ARDTs. This point has been discussed in detail in the related publication Nellikkattil et al. (2023).

The SI is a unitless, bounded quantity between -1 to 1, independent of the absolute magnitude or resolution of the scalar field. This invariance is crucial as it ensures that the SI focuses solely on the nature of curvature, independent of its magnitude or size. A traditional ARDT may use the absolute magnitude (Guan and Waliser, 2015) of the field or its curvature (Ullrich and Zarzycki, 2017) to identify AR like structures while SI would characterize all ridge structures on the $|\mathbf{IVT}|$ as ARs. The basic philosophy of SI can be summarized as "all circles are circles regardless of its radius".The SI provides a single-value representation that captures different shapes: values close to -1 represent concave surfaces, values near +1 indicate convex shapes, and the intermediate range corresponds to transitional shapes like flat or saddle-shaped regions. This intuitive interpretation makes the SI an effective tool for characterizing and analyzing local shapes across a wide range of applications. To obtain AR-like structures, the SI threshold of 0.375 is chosen to extract local shapes such as ridges, caps, and domes. Additionally, points where the $\mathbf{IVT}$ deviates by more than 45° from the local ridge are also removed to ensure coherence in the transport direction. The local ridge is aligned in the direction of the eigenvector corresponding to k_2.

Eq. 6: I find the decomposition in this equation quite confusing. The use of $\bar{u}$, which represents the monthly mean, may not necessarily reflect the change in AR. I encounter similar confusion with equations 11 and 12, as the changes in dynamics and thermodynamics are defined based on monthly timescales rather than AR timescales.

As discussed in Newman et al. (2012) and other publications, AR structures encompass moisture from both low and high-frequency variations in midlatitude moisture transport. While it is confusing that the mean changes in AR frequency, which is inherently a synoptic-scale phenomenon is explained by low frequency changes in atmospheric circulation, that is what our analysis showed. Supplementary Fig. 11-14 demonstrate that the eddy terms are considerably smaller, about one order of magnitude, than the dynamic or thermodynamic terms. Moreover, the patterns of changes in eddy terms do not align with those of AR frequency changes. Notably, the IVT decomposition remains consistent even when comparing AR frequency changes obtained using other ARDTs (Supplementary Fig. 8), provided that we consider their sensitivity to humidity changes.

Currently, we are working on identifying the specific circulation systems responsible for the detected AR changes, but this remains an ongoing endeavor. Our belief is that the changes in ARs result from the response of upper-level jets to warming. Applying a similar detection framework to upper-level wind speed, we find a close match between the changes in jet frequency and AR frequency changes (see the Figure 1). Given the tight link between ARs and upper level jets in present climate, this should not be surprise (Payne and Magnusdottir, 2014; Salimi et al., 2020).

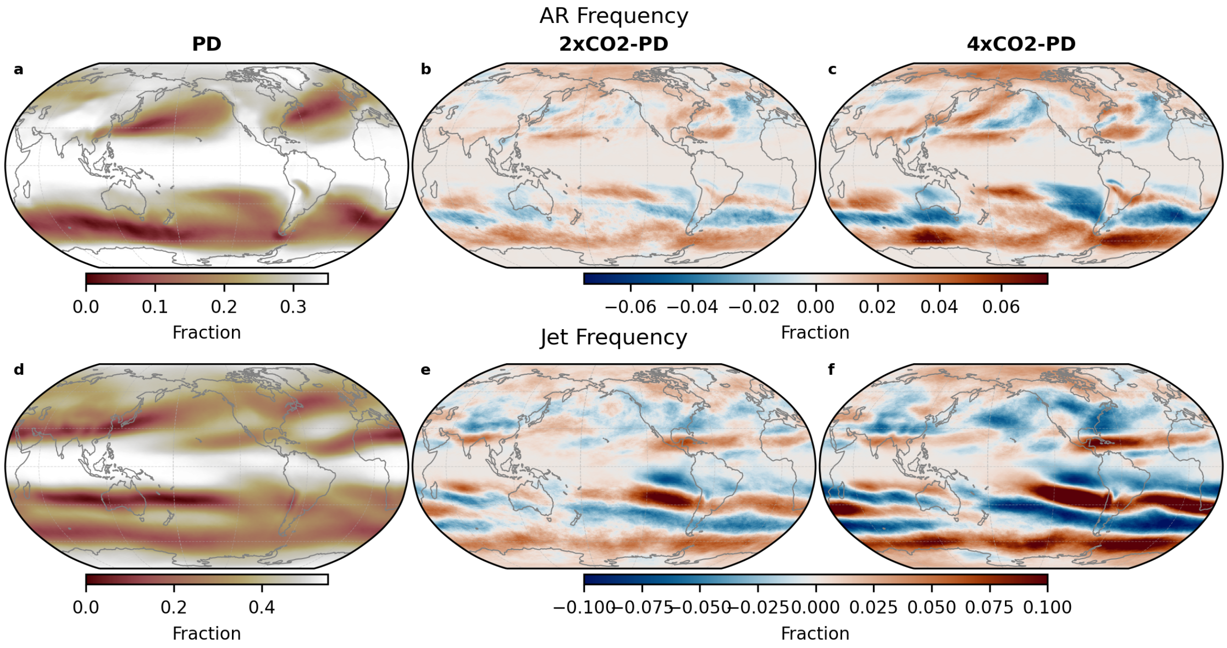

Figure 1: **Comparison changes in AR frequency to upper level jet frequency**. The upper panel shows (a) mean AR frequency in present climate simulation and its changes under (b) 2xCO2 and (c) 4xCO2 forcing. Similarly, (d) mean frequency of upper level jets and its response to (e) 2xCO2 and (f) 4xCO2 forcing are shown in lower panel.

Eq. 8: There seems to be a typographical error in the last term on the right-hand side. It should be $q_1'\bar{u}_1'$. The "prime" symbol is missing.

Absolutely right. The mistake has been rectified. Thank you.

Eq. 8: I am wondering if the change in the eddy term across different climate states is negligible, as numerous studies have suggested that a decrease in the equator-to-pole temperature gradient can effectively reduce eddy activity, leading to weaker baroclinic waves.

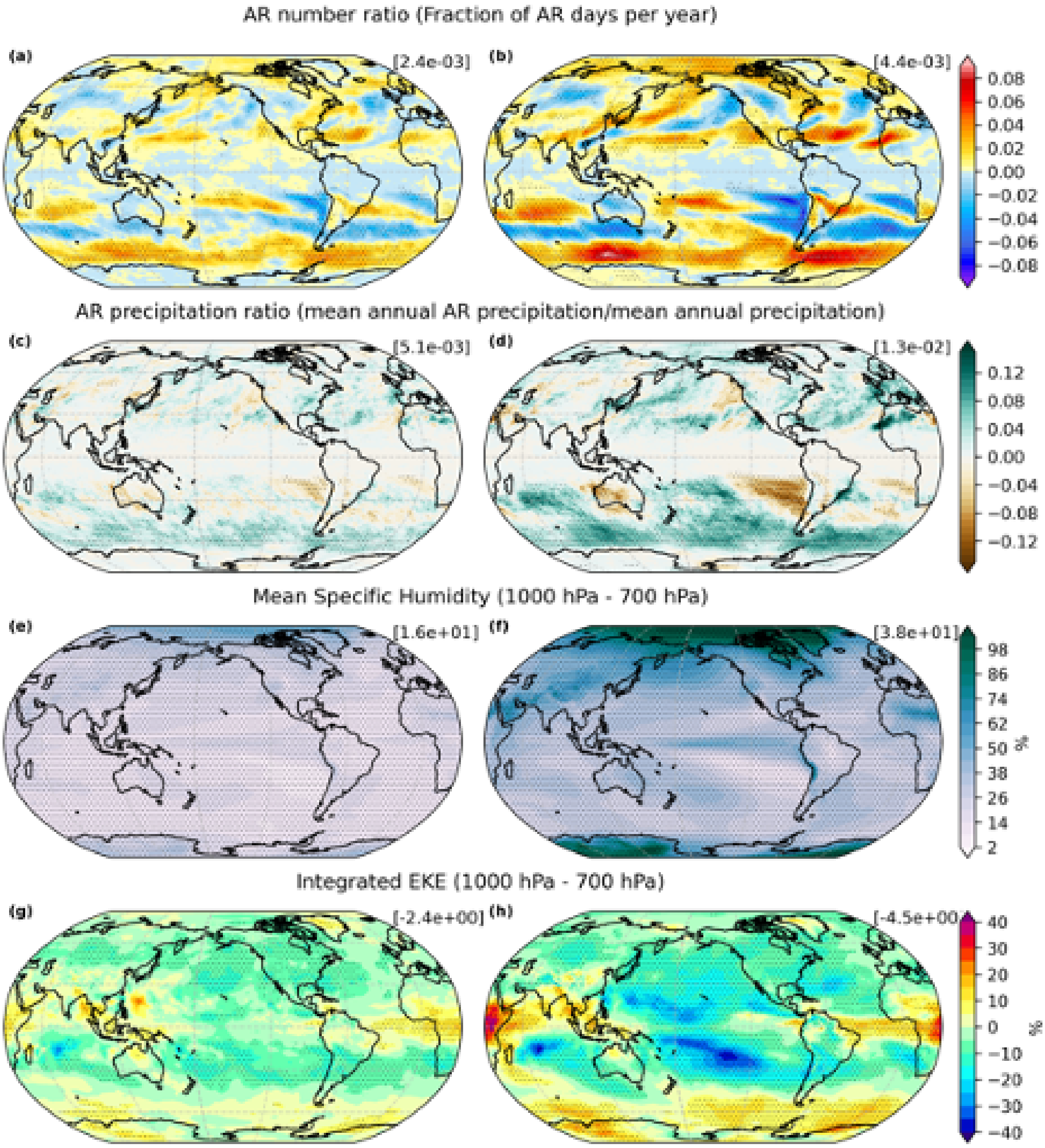

Figure 2: **Projected changes in AR characteristics, specific humidity, and lower tropospheric eddy kinetic energy**. Future changes in the annual fraction of (a, b) AR days, (c, d) AR to annual mean precipitation ratio (e, f) mean specific humidity, and (g, h) mean column integrated lower tropospheric (1000 hPa-700 hPa) eddy kinetic energy from (a, c) 2xCO2 and (b, d) 4xCO2 experiments in comparison to PD. Stippling denotes significant changes at 0.05 level in a two-tailed t-test. The numbers in the top right corner indicate the global mean of the change.

The values of eddy term from the equation is an order less than the other terms as shown in Supplementary Figs. 11-14. We would like to point out that eddy terms in Eq. 8 are eddy component of **IVT** and not atmospheric Eddy Kinetic Energy. We initially tried to explain the detected changes in AR frequency as a function of increased atmospheric humidity and reduced eddy kinetic energy. The figure produced for the first internal draft of the paper showing humidity changes and vertically integrated eddy kinetic energy changes are shown below (see Figure 2). As expected we see a ubiquitous increase in mean specific humidity and a significant reduction in the integrated eddy kinetic energy over the midlatitudes. However, the patterns in the changes could not satisfactorily explain the AR frequency changes. This lead to the development of the new decomposition method.

IBS Center for Climate Physics
Rooom No. 1105, IBS Center for Climate Physics
Busandaehak-ro 63beon-gil 2 (Jangjeon-dong),
Geumjeong-gu, Busan 46241, Republic of Korea
T +82-51-510-7636
F +82-51-510-7694
Email: arjunbabun@pusan.ac.kr
Web: ibsclimate.org

2 Response to Reviewer #2

Re: Re-submission of manuscript COMMSENV-22-1342A

Dear Reviewer,

Thank you for your kind words and valuable feedback on our manuscript. My coauthors and I sincerely appreciate your thoughtful suggestions, which have undoubtedly contributed to a significant improvement in the quality of our work. Your insights and guidance have been instrumental in refining the manuscript, and we are grateful for your time and effort in reviewing our research.

2.1 Reviewer Comments

Recommendation: I sincerely appreciate the thoroughness of the authors in responding to the many reviewer comments. The point-by-point responses were extremely organized and well-reasoned. I have no further comments.

2.2 Revision Highlights

- The abstract has been modified to include suggestion from reviewer #1 and to limit the word count to 150.
- A new subfigure showing the change in distribution of maximum intensity of precipitation within ARs is added to Figure 3. The inferences are discussed in the abstract as well as the main manuscript.
- Methods section is rearranged to comply with the journal recommendation of "only one layer of subheadings".
- The mathematical notation of variables and reference to supplementary figures has been modified to comply with journal requirements.

References

Bin Guan and Duane E. Waliser. Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research: Atmospheres*, 120(24):12514–12535, dec 2015. doi: 10.1002/2015jd024257. URL https://doi.org/10.1002%2F2015jd024257.

Arjun Babu Nellikkattil, Travis Allen O'Brien, Danielle Lemmon, June-Yi Lee, and Jung-Eun Chu. Scalable feature extraction and tracking (SCAFET): A general framework for feature extraction from large climate datasets. may 2023. doi: 10.5194/egusphere-2023-592. URL https://doi.org/10.5194%2Fegusphere-2023-592.

Matthew Newman, George N. Kiladis, Klaus M. Weickmann, F. Martin Ralph, and Prashant D. Sardeshmukh. Relative contributions of synoptic and low-frequency eddies to time-mean atmospheric moisture transport, including the role of atmospheric rivers. *Journal of Climate*, 25(21):7341–7361, nov 2012. doi: 10.1175/jcli-d-11-00665.1. URL https://doi.org/10.1175%2Fjcli-d-11-00665.1.

Ashley E. Payne and Gudrun Magnusdottir. Dynamics of landfalling atmospheric rivers over the north pacific in 30 years of MERRA reanalysis. *Journal of Climate*, 27(18):7133–7150, sep 2014. doi: 10.1175/jcli-d-14-00034.1. URL https://doi.org/10.1175%2Fjcli-d-14-00034.1.

Saadoun Salimi, Jalil Helali, Mohammad Lotfi, Hossein Momenzadeh, Seyed Asaad Hosseini, Ebrahim Asaadi Oskuei, Alireza Izadi, Zahra Yarmoradi, and Ismaeil Bakhshi. Investigating the origin and pathways of atmospheric rivers in the world. *Theoretical and Applied Climatology*, 142(1-2):165–175, jul 2020. doi: 10.1007/s00704-020-03299-w. URL https://doi.org/10.1007%2Fs00704-020-03299-w.

Paul A. Ullrich and Colin M. Zarzycki. TempestExtremes: a framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, 10(3):1069–1090, mar 2017. doi: 10.5194/gmd-10-1069-2017. URL https://doi.org/10.5194%2Fgmd-10-1069-2017.