We are grateful to the two reviewers for the time taken to review our manuscript. We are pleased that both reviewers consider this work to be of high quality and importance to the community. We have considered their suggestions and respond to them below, in red.

General Comments:

This study presents a comprehensive analysis of extreme precipitation events in West Antarctica using a mini-ensemble of three state-of-the-art kilometer-scale regional climate models (RCMs). The authors effectively demonstrate that high-resolution RCMs are crucial for capturing the complex interactions between atmospheric rivers (ARs) and the region's topography, which significantly influences precipitation patterns and phase transitions. The study provides valuable insights into orographic uplift, latent heat release, and the formation of supercooled liquid precipitation, contributing to a deeper understanding of precipitation dynamics in this climatically sensitive area.

I find this study to be well-executed and interesting, addressing an important gap in climate modeling for polar regions. The work is valuable for its use of high-resolution modeling to explore the drivers and impacts of ARs in West Antarctica. However, some clarifications and details are needed to strengthen the findings. Overall, I would consider these changes to be minor, as the study is already well-structured and offers valuable insights into the dynamics of extreme precipitation events in West Antarctica.

Specific Comments:

Abstract:

Line 18-19: Better to quantify the "severely underestimated" and provide more detail on the difference. For example, "In contrast, ERA5 reanalysis data underestimate snowfall by 40-60% compared to in situ measurements, highlighting the limitations of coarse resolution reanalysis in capturing localized precipitation events."

Unfortunately given the word limit constraints we cannot add all the words you suggest here, but we have tried to compromise by increasing a quantitative estimate of the underestimation in brackets, as follows:

"By contrast, snowfall estimates from ERA5 reanalysis for both events are severely underestimated **(by 3-4 times)** compared to the measurements."

We have also added more quantitative statements regarding the underestimate in the main text, e.g. at L297:

"Meanwhile, the median cumulative ERA5-derived estimate of total precipitation over TG ice shelf in the summer case is even lower than the mean RCM estimate, at 28 mm w.e**. (3-4 times lower than the SNOWPACK-derived estimates),** although still within the range of the RCM mini-ensemble (Figure 3a). "

Line 24-25: I would clarify what "resolution insensitive" means and elaborate on the significance of this finding. For example, "Our results demonstrate that RCM-simulated snowfall shows limited sensitivity to spatial resolution, while simulated rainfall increases by up to 50% in 1 km resolution models compared to 12 km models, suggesting that fine-scale modeling is essential to capture rainfall variability accurately."

Thanks for this suggestion. As above, we have limited space available in the abstract, but have revised to: "We also show that while the amount of MetUM-simulated snowfall was comparatively resolution insensitive, the amount of rainfall simulated was not, with rainfall amounts **over Thwaites ice shelf 4-16 times** higher in 1 km simulations compared to 12 km simulations."

And included the following in the main text at L623:

"There are few differences in accumulated snowfall amounts between the various resolution domains of the MetUM, **indicating that simulated snowfall totals are comparatively insensitive to resolution**."

And the following at L627:

"Furthermore, while total accumulated snowfall in the 12 km simulation is comparable to 1 km totals, **the accumulated rainfall totals over TG ice shelf are 16 times lower in the 12 km simulation compared to the 1 km simulations during the summer case and 4 times lower during the winter case** (Figure 8). **Across the entire domain**, rain **is** simulated only over the steepest terrain, such as Thurston Island and King Peninsula."

Summer

1.13 mm w.e. 12 km Thwaites ice shelf

17.84 mm w.e. 1 km Thwaites ice shelf

== 16 x higher

Winter

1.44 mm w.e. 12 km Thwaites ice shelf

6.22 mm w.e. 1 km Thwaites ice shelf

= 4.3 x higher

Introduction:

Line 70-72: Better to highlight the mechanisms leading to rainfall.

Suggestion: "…This is primarily due to warm air advection and moisture convergence facilitated by ARs, which elevate temperatures and increase the likelihood of liquid precipitation." Also, please add these two recent publications to highlight the AR-driven rainfall events in Antarctica:

- Bromwich, D. H., and Coauthors, 2024: Winter Targeted Observing Periods during the Year of Polar Prediction in the Southern Hemisphere (YOPP-SH). Bull. Amer. Meteor. Soc., 105, E1662–E1684, https://doi.org/10.1175/BAMS-D-22- 0249.1.
- Bozkurt, D., Carrasco, J., Cordero, R., Fernandoy, F., Gómez, A., Carrillo, B., Guan, B., 2024. Atmospheric river brings warmth and rainfall to the northern Antarctic Peninsula during the mid-austral winter of 2023. Geophysical Research Letters, 51, e2024GL108391, https://doi.org/10.1029/2024GL108391.

Thank you, we have added in your text suggestion and reference to these two new papers. We have also added reference later in the text (L523-524): "**Bromwich et al. (2024) and Bozkurt et al. (2024) also highlight the complexity of processes involved in initiating the phase transition from snow to rain in ARs**."

Line 70-7Line 92: I think it is good to specify the dates of case studies or at least to mention summer/winter events.

Added (L91-92): "This study will use a mini-ensemble of kilometre-scale RCMs to explore how ARs over the ASE interact with the TG and PIG ice shelves in two case studies **(during winter and summer 2020)**…"

Line 91-95: I suggest rewriting the main motivation of the study i) clarifying the research gap, ii) specifying the novelty and relevance, and iii) emphasizing the broader impact in a clearer way. Also, please clarify what is meant by "mini-ensemble.":

o "This study uses a mini-ensemble of three state-of-the-art kilometer-scale RCMs to explore the interactions between ARs and the TG and PIG ice shelves in the ASE during two extreme precipitation events—a summer and a winter case study. The study aims to quantify the precipitation deposited by ARs, including rain, which can significantly affect the SMB and stability of these ice shelves. Additionally, we evaluate the performance of these high-resolution RCMs in capturing the dynamics of extreme precipitation events in this climatically sensitive region, where in situ observations are limited. By understanding the drivers and impacts of AR-induced extreme precipitation, this work seeks to determine the extent to which ice mass loss from these critical ice shelves and glaciers can be mitigated or exacerbated by changes in accumulation patterns."

Thank you for your helpful suggestion. We have revised this paragraph as follows:

"This study uses a mini-ensemble of three state-of-the-art kilometre-scale RCMs to explore the interactions between ARs and the TG and PIG ice shelves during two extreme precipitation events in winter and summer 2020. We aim to quantify the precipitation deposited during the ARs, including rain, which may affect the SMB and stability of these ice shelves. Additionally, we evaluate the performance of these RCMs in this region in capturing the dynamics of extreme precipitation in this climatically sensitive region, where observations are limited. By understanding the drivers and impacts of AR-induced extreme precipitation, this work seeks to determine the extent to which ice mass loss from important ice shelves and glaciers in the region can be modulated by accumulation."

Method:

• Line 150: Can you please add which version of Polar-WRF was implemented?

We used v4.1.1. This detail has been added.

• Line 165-170: Which specific variables of ERA5 were used for downscaling? For instance, were SST and SEAICE data provided by ERA5 or another satellite product?

We used ERA5 for the standard variables for RCM initialisation (T, P, q, winds etc), as well as SST and sea ice. We also use precipitation variables (rain, snow, precip) to compare with the RCMs directly. We have inserted clarification of the variables used at L120-121:

"**We use standard atmospheric variables to initialise the model (temperature, pressure, winds and humidity) as well as sea surface temperature and sea ice**."

• Line 175-180: Can you please add the temporal resolution of the observational data?

Added "…as well as firn temperature, snow height and GPS position **at hourly resolution**." At L181

• Line 184-186: Does the SNOWPACK model use directly in-situ measurements in Cavity Camp and Channel Camp described in 2.3? Or any other data source?

SNOWPACK uses the in-situ observations available, but relies for downwelling long- and shortwave radiation on the reanalysis products. It can also be run using variables from reanalysis, but the simulations we have used here rely on observational measurements.

• Also, it would be useful if you can label the location of in-situ measurements in Fig. 1 or Fig. 2.

Thank you, we have added these into figure 1.

Results:

• Line 201: How do you define extreme precipitation in this study? I couldn't find relevant information in Section 2.

Ordinarily one could define extreme precipitation as being above a threshold (e.g. the 90th percentile) of the climatological average. For instance, Turner et al. (2019) use the "heaviest 10% of daily precipitation amounts" to indicate extreme precip. However, because this observed dataset is so new, it is not long enough to really compute average statistics. We selected the two cases based on the anomalies in the time series – these two were the largest accumulation periods during the dataset. However, they also fall into the top 10th percentile of precipitation for the available data. The summer case was also identified by Maclennan et al (2023) as an extreme precipitation case.

• Line 203-205: Interesting to see above 0°C temperatures during the winter case. Do you have any available atmospheric measurements (radiosonde profile)?

Unfortunately not! This is an extremely remote site and wintertime radiosonde ascents are impossible. The only data that are available are (uncrewed) automatic weather station data and snow height sensor data.

• Line 268-273: For the readers, it would be really useful to have a plot showing AR origin, origin and characteristics (e.g., axis, landfall) to better understand these lines as well as discussions in the next paragraph.

Unfortunately we do not have the relevant variables available to produce a plot such as this (IVT for instance) and the smaller domains are too small to show the origin of the ARs. However, the summer case is evaluated in Maclennan et al. (2023) and those authors present a more in-depth analysis of the event(s).

• Line 279-283: Following my previous point, can you quantify the impact of the spatial variability in AR landfall locations on the snowfall distribution across the TG ice shelf, particularly between the Cavity and Channel Camp stations? Specifically, how does the timing and trajectory of each AR event correlate with the differences in observed accumulation at these two locations?

While the landfall location may impact the wider spatial variability of precipitation across Thwaites, we are not sure this is a satisfactory explanation for the observed differences between the Cavity and Channel stations. Because the stations are so close to one another (just 4 km apart), it is unlikely that the landfall locations are the dominant case of differences between the two stations. Rather, as discussed in the text, it is more probable that highly localised characteristics, for instance the presence of depressions and metre-scale surface texture like sastrugi or ridges, as well as blowing snow etc. are responsible for the observed spatial variability.

• Line 285-301: Given the observed discrepancies in snowfall estimates between SNOWPACK, RCMs, and ERA5, especially in the summer case, can you elaborate on how the representation of AR dynamics (e.g., moisture transport, phase transitions) differs among the models? How might these differences contribute to the underestimation of snowfall?

Thank you for this insightful comment. There are indeed varying parameterisation schemes, model philosophy etc of the models / ERA5 which produce these differences. It is beyond the scope of this study to drill into the exact reason for these differences, but it is certainly useful discussion to include. As such, we have included the following in the paragraph:

"**Differences in simulated precipitation totals and the timing of precipitation events between the RCMs could be related to the way the models parameterise cloud microphysics or differences in their dynamics that determine how, when and where the simulated ARs make landfall.** Meanwhile, the median cumulative ERA5-derived estimate of total precipitation over TG ice shelf in the summer case is even lower than the mean RCM estimate, at 28 mm w.e. **(3-4 times lower than the SNOWPACK-derived estimates)**, although still within the range of the RCM mini-ensemble (Figure 3a). **This discrepancy is likely to be related to the coarser resolution of ERA5 and its orography, which would affect the passage and trajectory of airflow in the simulations, and hence the interactions between the AR and terrain. Convection is also parameterised in ERA5, which could produce different cloud dynamics and lower precipitation totals than the RCMs, which explicitly resolve convection at 1 km resolution.**"

• Line 302-310: Is it possible to provide an estimation of the potential uncertainty in the SNOWPACK accumulation totals due to unmeasured processes (e.g., melt, sublimation, wind erosion)? A brief discussion on this would be useful.

It's difficult to estimate these potential uncertainties. The largest uncertainty in the accumulated mass would be from the density of the accumulated snow and the interpretation of accumulation in terms of snowfall. The latter assumes that there was no net convergence or divergence of drifting snow in the area. Generally speaking, melt and sublimation during these cases are too small to result in substantial height changes, compared to the height changes caused by the snowfall. We have included some discussion of uncertainties in the manuscript at the end of this paragraph as follows:

"**As discussed in Section 2.4, the biggest sources of error in the SNOWPACK analysis probably come from the estimation of new snow density by the model, and the uncertainty related to blowing snow convergence or divergence. Based on our analysis, we expect that other, negative, mass balance components were small during these events.**"

• Line 327-328: Please simplify the sentence for clarity. "The lower precipitation estimates from the satellite data, compared to the multi-model mean of the RCMs, likely result from the satellite product's coarse resolution (0.1° × 0.1°, approximately 12 km) and the challenges in detecting precipitation over cold, snow-covered surfaces."

Thank you, we have revised as follows:

"The lower **precipitation estimates derived from satellite data, as compared** to the multi-model mean of the RCM estimates, **is likely related to** the lower resolution of the satellite product (0.1° × 0.1° or approximately 12 km) and may also reflect the difficulty of satellite sensors in measuring precipitation over cold snow and ice surfaces."

• Line 337-340: Better to break this into two sentences for clarity. "The RCM 5th-95th percentile range suggests that these models capture some of the spatial heterogeneity indicated by differences between Cavity and Channel. However, the time series at grid points corresponding to each station's location (not shown) are very similar, indicating unresolved spatial variations below the 1 km grid scale or inadequately represented processes."

Thank you, we have revised as follows:

"The RCM 5th-95th percentile range demonstrates that the RCMs can capture the magnitude of spatial heterogeneity implied by the differences between Cavity and Channel. **However,** the time series at grid points corresponding to each station's location (not shown) are extremely similar, suggesting that there are still unresolved spatial variations below the 1 km grid scale of the RCMs, or processes that are incorrectly represented."

• Line 346-356: Have you considered evaluating the impact of using different temperature thresholds to diagnose rain in SNOWPACK and how this might influence the discrepancies between the snowfall totals estimated by SNOWPACK and the RCMs? A brief discussion on the potential effects of this diagnostic choice could help clarify its implications for your surface mass balance estimates.

This is a good point, that the choice of threshold may play a role. We therefore tested varying the threshold between –2°C and +2°C for the implications of producing rainfall in the SNOWPACK simulation for the Cavity and Channel sites. The figure below (Fig R1) illustrates the results we found. For ERA5 during the summer case, a small change in threshold results in a strong increase in prognosed rainfall. However, as discussed in the manuscript, very strong rainfall rates at the AWS sites would be associated with a decrease in the snow depth because rainfall wets snow, increasing its density and darkening the surface. However, an increase in snow depth is generally found during this period. For MERRA-2 and for both models in the winter case, the threshold would need to be lowered below the freezing point to obtain substantial rainfall.
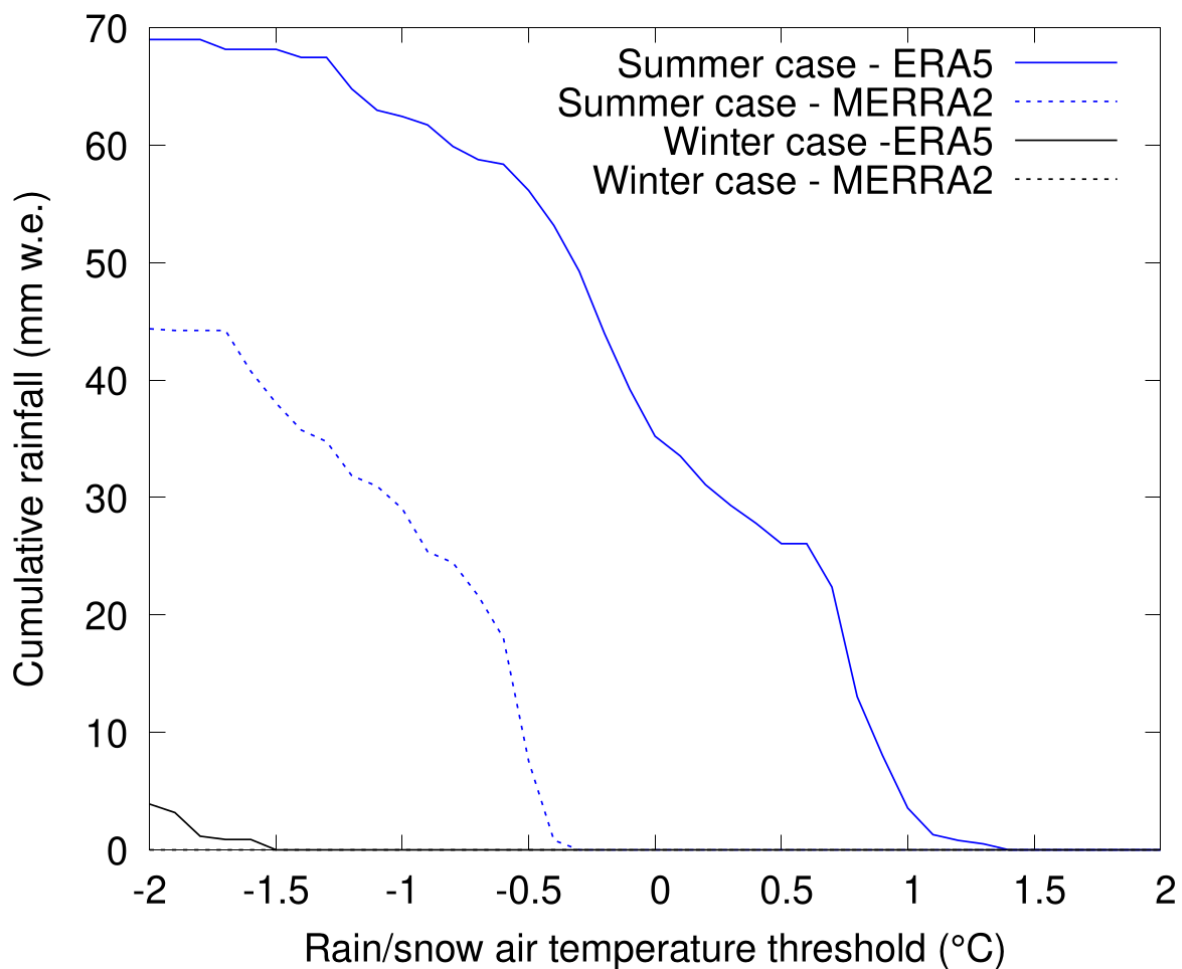
We have briefly summarised the above results in the revised manuscript as follows:

L367: "**Sensitivity experiments (not shown) where the diagnostic threshold was varied between - 2°C and + 2°C showed that in the summer case, lower thresholds resulted in considerably higher SNOWPACK-derived rainfall totals. However, during winter and the threshold needed to be lowered to considerably below zero to prognose rainfall.** Note that…"

• Line 374: Change to "Figures 3a and 3b" for consistency

Revised accordingly.

• Line 380-382: Please strengthen this by providing a brief explanation of the specific conditions (e.g., exact temperature ranges) that suggest mixed-phase precipitation. This would clarify why mixed precipitation is "likely."

We have included the exact temperature range observed at the AWS stations as follows:

"Given that the near-surface atmospheric conditions during this case were near the melting point (see Figures 6 and S2), **with warm air overlying the snowpack and observed 2 m temperatures fluctuating between -2°C and +2°C,** it is likely that precipitation was mixed phase."

• Line 420 and Fig. 5: Considering that Polar-WRF (Figure 5g) shows more extensive rainfall over the PIG ice shelf than MetUM (Figure 5f), could these differences be linked to how each model simulates the intensity and landfall of ARs over the region?

This is certainly a possibility. Polar-WRF is a somewhat wetter, warmer model in these cases, so it could also be related to the availability of heat and moisture.

For example, this is clear in Figures R2a to R2d, which show mean synoptic conditions (including mean near-surface temperature) during the summer case in all 3 RCMs and ERA5.

We have added the following at the end of the section:

"The differences between the RCMs in terms of the spatial distribution of rainfall and its intensity may be related to many factors, including the exact trajectory of the simulated ARs, temperature, moisture availability, model orographic interactions and cloud microphysics."
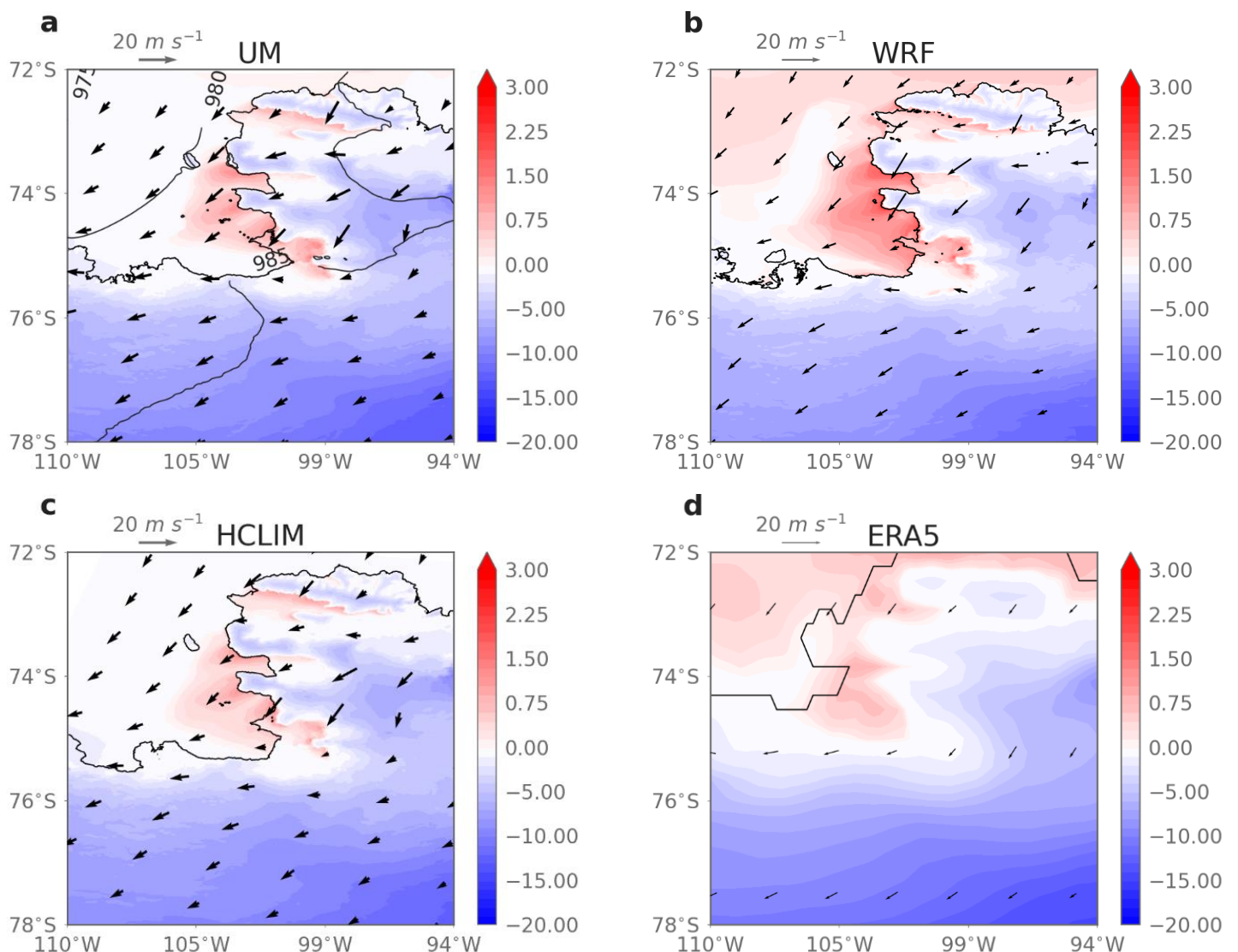

Fig R2. Mean synoptic conditions, averaged during the summer case, for a) MetUM, b) WRF, c) HCLIM and d) ERA5 over the RCM 1 km model domain. Colours show mean near-surface temperature, while arrows show mean 10 m wind direction. A scale vector is given at the top of each panel for reference.

• Line 465: For clarity and precision, modify to: "temperatures remain below freezing along the TG transect."

Thank you for this suggestion, edited accordingly.

• Line 491-497: The observed supercooled liquid precipitation reaching up to 2500 m (Figure 6) suggests unusual atmospheric stability or cloud dynamics. I think this is very interesting, and it would be good to see a detailed vertical profile of temperature, humidity, and wind within these ARs provide further insights.

• Line 496: It is not clear to me what exactly means "This same effect."

The phrase refers to the process described in the preceding sentence. We have added the following clarification: ". This same effect **(snow transitioning to rain as it falls through a warm layer)** has been observed during combined AR / foehn conditions at Davis station in East Antarctica by Gehring et al. (2022)."

• Line 520-540 and Fig. 7: Figure 7 indicates a high rain-to-snow ratio over certain regions, particularly over steep slopes. Is there a possibility that the model is overestimating liquid precipitation due to its handling of mixed-phase clouds or inadequate representation of ice nucleation processes? Could this observation suggest a potential bias in how the model simulation handles phase transitions during AR events?

It's certainly a possibility – although I would note that in most cases, RCMs tend to overestimate **ice** in mixed phase clouds rather than liquid.

• Line 585: Rewrite to improve clarity: "The absolute maximum temperatures are also lower, with median air temperatures over the PIG and TG ice shelves approximately 10°C lower in ERA5 than in the RCMs (Table 2)."

Revised accordingly, thank you.

Conclusions:

• Line 620-664: Considering the computational trade-offs mentioned, what resolution do you propose as the optimal balance between accuracy and efficiency for simulating extreme precipitation events in regions with complex topography like West Antarctica?

Could a dynamic resolution approach, where finer grids are applied only in regions of steep topography or during periods of intense precipitation, provide a more efficient modeling strategy?

Using dynamic resolution, or indeed the prospect of integrating machine learning and/or other statistical techniques with physically based models could offer considerable advantages at a fraction of the computational cost – both are very exciting prospects for important regions like this.

• Line 620-664: I suggest highlighting the necessity of conducting sensitivity experiments by altering key microphysical parameters, such as cloud droplet number concentrations and ice nucleation rates, to evaluate their impact on precipitation phase transitions. This would help identify which microphysical processes most significantly influence the partitioning between rain and snow in the region.

We have added (L661): "**Additional sensitivity experiments, for example exploring the impact of specific microphysical parameters and processes like droplet or ice crystal number, are beyond the**

scope of this study but would offer insight into the processes governing precipitation phase partitioning during extreme events."

Technical Corrections:

• In several sections, the text refers to complex atmospheric processes (e.g., supercooled liquid precipitation, seeder-feeder mechanisms) without fully explaining their implications or how they are represented in the models. Adding brief explanations or providing context would improve clarity, particularly for readers less familiar with these concepts.

We have included a more detailed definition of the seeder-feeder process in the introduction (at L81):

"Seeder-feeder processes involve a higher-level cloud (the 'seeder') precipitating ice crystals into a lower cloud containing liquid droplets (the 'feeder') where they grow via riming or vapour deposition, thus enhancing liquid precipitation (He et al., 2022)."

• Line 374: Change "Figure 3b, d" to "Figures 3a and 3b" for consistency.

Revised accordingly.

• Line 386: Use consistent hyphenation: "ground-truthing" or "ground truth."

Thank you, we have revised for consistency.

• Line 550: Correct "Ars" to "ARs."

Done, thanks.

• Line 606: Remove the extra punctuation after "Abel et al., 2017."

Done, thanks.