

Curved atmospheric rivers and their moisture remnants: a new detection tool for Antarctica
Victoire Buffet, Vincent Favier, Benjamin Pohl, Jonathan D. Wille

Response to the Reviewers' Comments

We thank the two reviewers for their constructive and insightful comments, which have helped us significantly improve the manuscript. We have addressed all comments in detail below and revised the manuscript accordingly. All changes are described in the responses and summarized at the end of this document.

The changes made in the manuscript are described below and listed in the Summary of Modifications at the end of this document. The major revision consists in adding Guan and Waliser (2024; hereafter G&W v4) as a fourth comparison catalogue throughout all results sections and figures, following the request of both reviewers.

Reviewer #1

This manuscript presents an updated Antarctic atmospheric river (AR) detection tool (ARDT), detecting ARs using their Kurvature (DARK) that builds upon the Wille Antarctic ARDTs (Wille-vIVT and Wille-IWV; Wille et al., 2019). The new framework uses IVT as the primary moisture detection field, applies a 98th percentile climatological threshold, and extends the spatial domain to the South Pole. These modifications allow the detection of ARs with complex geometries (e.g., curved, overturning, or crossing the South Pole) in comparison with Wille ARDTs. The method also introduces "AR-children," interpreted as moisture remnants of ARs that do not meet full detection thresholds. Using this framework, the authors show that DARK ARs capture extreme events more effectively than the Wille-based ARDTs. Overall, the manuscript is well written and represents a valuable contribution to Antarctic AR research. I have several major and minor comments that may help improve clarity, rigor, and the strength of the manuscript.

Major Comments:

1. Interpretation of Guan and Waliser (2015, 2019): The manuscript states that the Guan and Waliser (2015, 2019) AR detection algorithm excludes ARs that curve or reverse (e.g., Lines 519–520). However, this characterization may not be entirely accurate. Guan and Waliser impose a directional coherence constraint, but this does not explicitly exclude curved ARs. Their criterion states that if more than half of the grid cells within an object have IVT directions deviating by more than 45° from the object's mean IVT direction, the object is discarded. This condition does not prohibit curvature per se. In fact, Figures 2 and 3 of Guan and Waliser (2015) show examples of ARs with noticeable curvature that satisfy the coherence criterion. Therefore, it may be more precise to clarify that the Guan and Waliser framework restricts directional incoherence rather than curvature or reversal outright.

We thank the reviewer for this important and precise clarification. We agree that our original wording was too strong and have revised these accordingly. The sentence now reads (Lines 110–113):

“Global AR detection schemes such as G&W v4 also apply directional-coherence constraints that may limit the detection of ARs with strongly curved or reversing geometries, though their primary limitation over Antarctica relates to the effective IVT threshold rather than directionality, as further discussed in Sect. 3.”

We have made the same correction in the Introduction (Lines 56–68), where the original description of G&W has also been substantially revised (see response to Major Comment 2 below and to Reviewer 2’s General Comment).

2. Comparison with Guan and Waliser ARDT: Given that the Guan and Waliser ARDT is one of the most widely used global AR detection algorithms, a direct comparison with it would substantially strengthen the manuscript. Currently, the evaluation focuses primarily on comparisons with Wille-vIVT and Wille-IWV. Since Shields et al. (2022) have already compared Wille-based ARDT with several global ARDTs, including Guan and Waliser ARDT, it may be most useful here, at minimum, to apply the Guan and Waliser ARDT to the specific case studies shown in Figures 8–10. This would provide a complementary evaluation of DARK’s added value. My understanding is that in polar regions, Guan and Waliser apply a minimum IVT threshold of $100 \text{ kg m}^{-1} \text{ s}^{-1}$ to avoid detecting extremely weak moisture plumes. Over inland Antarctica, where IVT values are often below $50 \text{ kg m}^{-1} \text{ s}^{-1}$ (see Figure 1 in Guan & Waliser 2015), this absolute threshold may limit detection. Therefore, over Antarctica, it may not be the directional coherence constraint that limits detection for Guan and Waliser ARDT, but rather the absolute IVT threshold. Including Guan and Waliser ARDT results in Figures 8–10 would clarify: if Guan and Waliser detects these cases, even when curved, this would suggest that curvature is not inherently excluded by its coherence criterion. If it does not detect them, it would be important to determine whether this is due to the $100 \text{ kg m}^{-1} \text{ s}^{-1}$ threshold. Such a comparison would significantly strengthen the manuscript’s central claim.

We thank the reviewer for this suggestion, which is also echoed by Reviewer 2. We have fully implemented this comparison. Bin Guan kindly provided the G&W v4 (Guan and Waliser, 2024) AR catalogue for ERA5 1979–2023. G&W v4 is now included as a fourth comparison scheme throughout the manuscript: a new column showing differences relative to G&W v4 has been added to Figures 3, 5, 6, 7, 8, and 9; G&W v4 AR contours (green) are shown in all case study figures (Figs. 10 and 11); and dedicated parts discussing G&W v4 results have been added to each results section. The Discussion has also been expanded accordingly.

The case study analysis confirms the reviewer’s hypothesis: it is primarily the effective IVT threshold of G&W v4, not its directional coherence constraint, that limits inland detection. Despite its polar refinement, the G&W v4 hemispheric threshold ($\sim 46.8 \text{ kg m}^{-1} \text{ s}^{-1}$ in January, $\sim 39.4 \text{ kg m}^{-1} \text{ s}^{-1}$ in July; Guan and Waliser, 2024) prevents detection of ARs as they penetrate inland and lose intensity. Mean AR IVT for DARK and the Wille schemes decreases to $10 \text{ kg m}^{-1} \text{ s}^{-1}$ or below south of the EAID, well below the G&W v4 effective threshold. G&W v4 records no detections at all over some interior regions such as Dome A and the area surrounding the South Pole. We have added a new supplementary figure (Fig. S3) showing the number of AR days for each ARDT and another new supplementary figure (Fig. S4) showing mean AR IVT for all four schemes to document this threshold effect explicitly.

3. Clarification of the Methodological Workflow (Lines 165–173): The methodological description includes several technical terms such as “morphological thinning algorithm”, “skeleton”, “endpoint-to-endpoint paths”, “longest and most directionally consistent path”, “principal component analysis”, “primary centerline”, “B-spline interpolation”. These concepts may not be readily accessible to readers who are not familiar with image-processing techniques. I strongly recommend illustrating the workflow with a concrete example (similar to how Guan and Waliser (2015) demonstrated their detection steps in their Figure 2). A schematic or step-by-step figure showing the progression from “Initial IVT object”, “Skeletonization”, “Centerline extraction”, and “Final smoothed AR axis” would substantially improve accessibility and reproducibility.

A new figure (Fig. 2) has been added showing the step-by-step DARK detection workflow applied to a real ERA5 time step, with six panels: (a) input IVT field, (b) grid cells exceeding the 98th-percentile threshold, (c) objects with area exceeding 200,000 km² retained, (d) skeleton extraction, (e) smoothed centerline derivation, and (f) final DARK AR classification with centerline lengths annotated. The methods text in Sect. 2.2 has been simplified accordingly, removing technical image-processing details that are not essential for reproducibility and replacing them with a reference to the new figure.

Comments on Figures:

1. Add latitude and longitude grids to all maps: None of the figures include latitude and longitude lines. This makes it difficult to locate specific regions, especially for readers not deeply familiar with Antarctic geography. Please add coordinate grids to all maps.

Latitude and longitude gridlines have been added to all maps. For multi-panel figures sharing the same projection and domain, coordinate labels are shown on the top-left panel only to preserve readability.

2. Figure 2a (Line 248): The text refers to the “midlatitude storm-track region” as having the highest occurrence. Please explicitly indicate this region in Figure 2a for clarity.

We have added “~50°S” to the text to identify this region, and the reference Chemke et al., 2022 to support this location (as seen in their Extended Figure 2). The latitude circles now visible on Figure 2a (50°S, 65°S, 75°S) make this region visually identifiable. Note: In the revised manuscript, this figure is now labeled Figure 3 and the text appears around Line 272.

3. Figure 3 (Lines 264–291): Please provide more detail on how Figure 3 was constructed. What specific processing steps or compositing methods were used?

We have added a dedicated methods section (Sect. 2.5: “Characterisation of AR moisture transport directionality”) describing the computation of the three metrics shown in Figure 4 (previously labeled as Figure 3). The figure caption now references this section. Briefly, at each grid cell and 6-hourly time step flagged as AR, we compute: (1) the fraction of AR time steps where $|v_{IVT}| > |u_{IVT}|$ (meridional-dominant flow) versus $|u_{IVT}| > |v_{IVT}|$

(zonal-dominant); (2) the fraction with eastward-dominant zonal flow ($u_{IVT} > 0$) versus westward-dominant zonal flow ($u_{IVT} < 0$); and (3) the fraction with northward-dominant meridional flow ($v_{IVT} > 0$) versus southward-dominant meridional flow ($v_{IVT} < 0$).

4. Figure 9: Please mask the McMurdo Dry Valleys in the figure.

We were not entirely sure whether the reviewer meant to mask the region or simply to indicate its location on the figure. We have therefore added a small grey rectangle indicating the McMurdo Dry Valleys domain used in the case study analysis, which we believe addresses this comment. We chose not to overlay a mask on the filled contour fields, since the domain is small enough to avoid obscuring the key precipitation and temperature discussed. Note: In the revised manuscript, this figure is now labeled Fig. 10.

5. In Line 423, it is unclear where “another farther north directed a zonal moisture flux from the South Pacific into the western Ross Sea sector” is located in Figure 9. Please annotate or clarify this.

The synoptic description has been revised. The sentence previously referred ambiguously to two troughs; we have clarified that on 1 December, it is the western flank of the ridge centered over Marie Byrd Land that drives the zonal moisture flux into the Ross Sea sector, while the trough northwest of the Ross Sea becomes the dominant channeling feature only on 2 December. Note: In the revised manuscript, Figure 9 is now labeled Figure 10 and the text appears around Line 505.

Additional Comments:

Line 46–48: Please provide references supporting the statement that ARs “play a central role in the Antarctic surface mass balance.”

We have added Wille et al. (2025) in Line 46, a recent Nature Reviews Earth & Environment review paper on Antarctic ARs, as a general reference at this location. The specific SMB impacts (snowfall, surface melt, ice-shelf instability) are individually referenced in the sentences immediately following.

Lines 59–61: The description of Guan and Waliser as identifying “extensive zonal moisture bands” is not clear. Their method also includes minimum IVT thresholds and directional coherence constraints. Please revise for precision.

We have substantially revised this paragraph (Lines 56–68). It now accurately describes the G&W v4 method, including its seasonally and spatially varying IVT threshold (85th percentile), geometric and directional filters, and the polar refinement introduced in Guan and Waliser (2024). G&W v4 is also explicitly introduced as a fourth comparison catalogue used throughout the study.

Line 416: There is an extra comma.

Corrected, thank you.

Line 492: The term “non-circular structure” is unclear. The figure appears to show a clear low–high pressure couplet channeling moisture poleward. Please clarify how “non-circular structure” implies weak wind speeds and why this explains the failure of detection by vIVT or DARK.

We have revised the sentence to read (Lines 591–593):

“The disorganized structure of the 500-hPa geopotential height anomaly contours, with wavy rather than well-defined closed centers of action, reflects relatively weak pressure gradients and wind speeds, which likely explains why this event is not detected by either vIVT or DARK.”

Lines 501–504: By Day 4, although DARK AR exists, its footprint does not appear to align precisely with the region of strongest 2 m air temperature anomalies. Please clarify whether this indicates spatial offset, lagged response, or limitations of the detection.

We have clarified this in the revised text. The offset reflects the persistence of the warm air mass advected during peak AR activity, which remains anchored over the region covered by the AR core on the previous day, rather than following the contracted AR on Day 4. The revised sentence reads (Lines 602–606):

“By the fourth day, the AR had contracted into a thinner filament along the southern flank of the EAID and continued advancing westward, while the strongest warm anomalies remained anchored over the region previously covered by the AR core, reflecting the persistence of the advected warm air mass rather than active transport at this stage.”

Section 2.5: Please provide appropriate references for the definition and application of the risk ratio used in this section.

We have added Katz et al. (1978) and Scholz and Lora (2024) as citations at the first mention of risk ratio in the methods (now Sect. 2.6 following the addition of Sect. 2.5).

Lines 191–192: “Temporally Associated with AR Occurrence”, does it mean for each grid point, please clarify this.

We have revised the sentence to remove the ambiguous term and clarify that the association is computed at each grid cell, with the attribution window extended to include adjacent days depending on the variable. The revised text reads (Lines 206–209):

“To attribute surface impacts to ARs (or AR-children), we identify precipitation and temperature events associated with AR occurrence at each grid cell. Because AR-related precipitation and temperature anomalies can persist beyond the period when a given grid cell is directly affected by an AR, the attribution window is extended to include adjacent days, depending on the variable.”

We also identified and corrected a typographical error in the Methods (Sect. 2.2) and Introduction: the minimum area threshold for DARK AR objects was incorrectly stated as 20,000 km² in the original manuscript, when the correct value applied in the algorithm is 200,000 km². This threshold is physically motivated by the typical spatial scale of ARs, which are generally hundreds of kilometers wide and thousands of kilometers long (here at least 2,000 km). The 20,000 km² threshold cited in the original text corresponds exclusively to the

AR-children module (Sect. 2.3), where a lower area criterion is appropriate to capture the smaller post-landfall remnant structures. This distinction has been clarified in both sections.

Reviewer #2

Buffet et al. introduce a novel algorithm to detect ARs in Antarctica. The novelty of the algorithm lies in the ability to detect a wider range of geometric features, in addition to features that penetrate farther inland on the Antarctic Ice Sheet. They compare output from their algorithm to output from a common polar ARDT (Wille) and note that their algorithm identifies a higher number of impactful features over Antarctica. Overall, I think this is a very strong paper: well-written, with a clear story and an effective use of analysis to back up their main points. The algorithm's development is well-justified, and the algorithm itself is well-presented and explained. I think this algorithm will be a good fit into the existing community of AR detection methods. I have one key concern I highlight below, but after considering this concern, and other comments, I'm happy to recommend publication!

General Comment:

My only concern with the manuscript is that, in placing their algorithm into context within the broader AR detection community, the authors only use one algorithm (Wille, albeit with different versions) for comparison. Given the robust selection of ARDTs available through the ARTMIP project (<https://ncar.github.io/ARTMIP/intro.html>), I'd be curious to see how this algorithm performs compared to other ARDTs. Of course, there may be limitations in the availability of algorithm output for ERA5 (especially polar algorithms). However, Guan and Waliser (2024), for example, recently updated their algorithm to be more regionally-refined. They tested this algorithm on ERA5 data and reported the potential for improved performance in polar regions. Since the authors specifically mention limitations related to older versions of Guan and Waliser, I encourage the authors to either a) add Guan and Waliser (2024) as an additional point of comparison (perhaps in the supplement) or b) provide some discussion about how global algorithms, such as Guan and Waliser, are similarly being adapted to perform better in polar regions.

We thank the reviewer for this suggestion, which is consistent with Major Comment 2 of Reviewer 1. We have adopted option (a) and added G&W v4 as a full fourth comparison throughout the manuscript, not only in a supplement but integrated into all results figures and discussed in detail in the text and Discussion. Bin Guan kindly provided the G&W v4 catalogue for ERA5 1979–2023 (acknowledged in the Acknowledgements section). See the response to Reviewer 1 Major Comment 2 above for the complete list of changes.

Lines 30–44: I'm concerned that the Plain Language Summary seems a bit too technical in places. In its current form, it reads very closely to the abstract, so I suggest the authors find places where they can simplify or adjust the text to make it less technical. One possible avenue is to reduce the number of statistics discussed in the summary.

We have revised the Plain Language Summary to reduce technical terminology and the number of statistics, making it more accessible to a non-specialist audience. The revised version focuses on the key physical concepts (ARs as moisture corridors, their role in Antarctic weather, the novelty of DARK and AR-children) while retaining only the most impactful statistics.

Lines 59–61: How is “skill” determined in this context? I gather that the authors are highlighting how the Guan and Waliser algorithm depicts fewer ARs. Guan and Waliser (2024) highlights an updated version of their algorithm, which is designed for better performance in polar regions.

We have removed the ambiguous term “skill” and have substantially revised this paragraph to accurately describe G&W v4, including its hemispheric threshold approach and polar refinement. G&W v4 is now explicitly introduced as a fourth comparison catalogue in the Introduction, with a clear statement that it is compared throughout the study.

Section 1: My only suggestion is that it would be cool to also see an example of how DARK adequately detects a representative equatorward-curving feature.

Figure 1b shows three representative landfall events illustrating the range of AR geometries captured by DARK, including curved and overturning structures that the Wille schemes miss due to their directional and geometric constraints. In particular, the 2020-11-16 event shown schematically in Fig. 1b is an equatorward-turning AR, exhibiting both southward and northward vIVT components, which is precisely the type of feature that DARK captures and the Wille vIVT scheme cannot detect. With the addition of G&W v4 contours in the revised Figure 1b, all four schemes are now shown simultaneously on schematic outlines derived from real landfall events, making the detection differences directly apparent.

Section 2.2: I was curious if the authors have a schematic that visually depicts what’s being described. I realize that this methodology may not be easily distilled into a graphic, but a visual aid, if possible, could enhance this section.

A workflow figure has been added (Fig. 2), as described in the response to Reviewer 1 Major Comment 3.

Section 2.3: It’s cool how this functionality is included as an option with the algorithm. I have a few questions related to this add-on: 1) How much additional computational cost is required to enable this parent-child add-on feature? (And in general, what are estimates for computational cost for this method as a whole?) 2) Are there instances when this functionality wouldn’t be desired? I’d be curious to hear about example use cases for each “version” of the algorithm. This could provide useful guidance for future users.

We have added a sentence at the end of Sect. 2.3 about the computational cost (Lines 201–202): “The AR-children module requires approximately 9h of CPU time and ~17 GB of memory for the full 1979–2023 climatology, compared to 1h and ~8 GB for the core DARK algorithm (see Code and Data Availability section).” We replicated that in the Code and Data availability section.

We have added a sentence in the Discussion about the use of the AR-children module (Lines 693–696): “From a practical standpoint, users whose focus is limited to coastal AR landfalls and their immediate impacts may find the core DARK algorithm sufficient, given its substantially lower computational cost, while the AR-children module is recommended for studies focusing on inland moisture transport and surface impacts.”

Lines 213–216: Is there a citation (or a group of citations) that introduces the risk ratio?

We have added Katz et al. (1978) and Scholz and Lora (2024) as citations. See also the response to Reviewer 1's comment on Section 2.5 above.

Line 241: This may be a simple answer, but perhaps the authors could provide a sentence explaining why they adopted the log-normal approximation method.

The following sentence has been added (Lines 263–265): “The statistical significance of the RRs was assessed following the log-normal approximation method of Katz et al. (1978), because the distribution of the risk ratio is asymmetric while its logarithm is approximately normally distributed, allowing reliable confidence intervals to be derived using standard normal theory.”

Line 297: Very minor, but it looks like the section number (3.2) is repeated here.

Corrected, thank you.

Figure 5: I like the practice here, and in prior figures, where the authors add summary statistics to the figure labels. I suggest doing something similar for the overall frequency plot (Figure 2).

Land-area weighted mean and maximum values are now indicated below each panel of Figure 3 (previously labeled as Figure 2), consistent with the practice in subsequent figures.

Line 416: Very minor, but there are double commas here.

Corrected, thank you.

Section 3.4: I find the use of case studies to be very helpful in a study such as this one. This section provides some visual intuition, and it also provides some physical understanding of the differences between algorithms.

We thank the reviewer for this positive comment. G&W v4 AR contours have been added to all case study figures (Figs. 10 and 11), and the corresponding text now discusses the G&W v4 detections for each event in detail.

Summary of Modifications

Major Additions:

- G&W v4 (Guan and Waliser, 2024) added as fourth comparison throughout manuscript
 - New columns in Figs. 3, 5, 6, 7, 8, 9
 - Green contours in Figs. 1b, 10, 11
 - Expanded Discussion section on G&W v4 limitations and RR interpretation
 - Dedicated discussions in each Results section
- New workflow figure (Fig. 2) showing DARK detection steps
- New supplementary figures: S3 (AR day counts), S4 (mean AR IVT)
- New Section 2.5: "Characterisation of AR moisture transport directionality"

Key Text Revisions:

- Plain Language Summary (Lines 30–44): Reduced technical terminology and statistics
- Sect. 2.2 and Introduction: Corrected typographical error in minimum AR object area threshold (200,000 km², not 20,000 km²); clarified that the 20,000 km² threshold applies only to AR-children (Sect. 2.3).
- Lines 46: Added Wille et al. (2025) reference for SMB role
- Lines 56–68: Substantially revised G&W v4 description
- Lines 110–113: Corrected characterization of G&W directional coherence constraint
- Lines 161–162 : Added G&W v4 ARs to the data
- Lines 170–178 : Simplified explanation of the algorithm, with added Fig. 2
- Lines 201–202 : Added computational cost for AR children versus DARK ARs only
- Line 207: Clarified AR attribution window at “each grid cell”
- Lines 228-234 : Added the section 2.5 “Characterisation of AR moisture transport directionality”
- Line 237: Added reference (Scholz and Lora, 2024)
- Lines 263–269: Added explanation for log-normal approximation
- Line 272: Added reference (Chemke et al, 2022)
- Lines 282–297 : Added a part on G&W v4 (occurrence)
- Lines 331–333 : Added a part on G&W v4 (direction)
- Lines 350–358 : Added a part on G&W v4 (contribution to precipitation)
- Lines 386–395 : Added a part on G&W v4 (precipitation intensity)
- Lines 410–415 : Added a part on G&W v4 (contribution to extreme precipitation)
- Lines 423–426 : Added a part on G&W v4 (contribution to extreme temperature)
- Lines 437–438 : Added a part on G&W v4 (contribution to compound extremes)
- Lines 471–480 : Added a part on G&W v4 (risk ratio)
- Line 501 : Rephrased
- Lines 508–516 : Added G&W v4 in the text (case study McMurdo)
- Lines 531–534 : Added a part on G&W v4 (case study McMurdo)
- Lines 569–570 : Added G&W v4 in the text (case study South Pole)
- Lines 575–579 : Added G&W v4 in the text (case study South Pole)
- Line 583 : Added G&W v4 in the text (case study South Pole)
- Lines 591–592 : Clarified the synoptic structure
- Lines 593–601 : Added G&W v4 in the text (case study South Pole)
- Lines 603–606 : Addressed the footprint of the being not aligned with the strongest 2 m air temperature anomalies
- Line 622: Refined the description of G&W v4 directional constraints
- Lines 629–637 : Added a part on G&W v4 (specified the fundamental detection differences with DARK and Wille schemes)
- Lines 667–680 : Added a part on G&W v4 (Discussion of the contribution to precipitation and the risk ratio)

- Lines 693–696 : Added a recommendation for the use of AR children
- Lines 737–742 : Added the computational cost of DARK ARs and AR children
- Lines 756–757 : Added acknowledgements of the reviewers and Bin Guan

Figure Changes:

- All maps: Added latitude/longitude gridlines, with labels
- Fig. 1b: Added G&W v4 contours
- Fig. 2: NEW workflow schematic showing detection steps
- Fig. 3: Added land-area weighted statistics; added G&W v4 column
- Fig. 4: Updated caption to reference new Sect. 2.5; added G&W v4 column
- Figs. 5–9: Added G&W v4 columns (differences)
- Figs. 10–11: Added G&W v4 AR contours (green)
- Fig. 10: Added a grey box outlining the McMurdo–Dry Valleys region
- Fig. S3: NEW – Number of AR days for all four schemes
- Fig. S4: NEW – Mean AR IVT for all four schemes

Computational Cost Additions:

- Sect. 2.3, Lines 201–202, 737–742: Added CPU/memory requirements for AR-children module
- Code and Data Availability: Duplicated for user guidance

Methodology Additions:

- New Sect. 2.5 (Lines 228–234): Methods for directional analysis
- Sect. 2.6 (formerly 2.5): Risk ratio with new citations (Katz et al. 1978; Scholz & Lora 2024)

References Added:

- Chemke et al. (2022) – Storm track dynamics (for ~50°S region)
- Katz et al. (1978) – Risk ratio definition
- Scholz and Lora (2024) – Risk ratio applications
- Guan and Waliser (2024) – Used throughout

Acknowledgements:

- Added thanks to Bin Guan for providing G&W v4 catalogue
- Added thanks to two anonymous reviewers