Supplementary Material: Atmospheric river trajectories organize along a global transport network

Tobias Braun[*1,2], Sara M. Vallejo-Bernal[1,2], Norbert Marwan[2],
Juergen Kurths[2], Johannes Quaas[3], Albert Díaz-Guilera[4,5],
Luis Gimeno[6], and Miguel D. Mahecha[1,7,8]

[1]Institute for Earth System Science and Remote Sensing,
Leipzig University, Leipzig, 04103, Germany
[2]Potsdam Institute for Climate Impact Research, Potsdam,
14473, Germany
[3]Institute of Meteorology, Leipzig University, Leipzig, 04103,
Germany
[4]Departament de Física de la Matèria Condensada, Universitat
de Barcelona, Barcelona, 08028, Spain
[5]Universitat de Barcelona Institute of Complex Systems
(UBICS), Universitat de Barcelona, Barcelona, 08028, Spain
[6]Centro de Investigación Mariña, Universidade de Vigo,
Environmental Physics Laboratory (EPhysLab), Campus da
Auga, Ourense, 32004, Spain
[7]German Centre for Integrative Biodiversity Research (iDiv)
Halle-Jena-Leipzig, Leipzig, 04103, Germany
[8]Helmholtz Centre for Environmental Research, Leipzig, 04318,
Germany

1

AR network

A schematic to better illustrate the construction of the atmospheric river transport network (ARTN) is shown in Fig. S1. The sensitivity to the threshold ϵ is showcased in Fig. S2. As an alternative to AR centroids, we studied ARTNs based on AR core (Fig. S3) and AR head coordinates (Fig. S4). We found that the latter track AR landfall and inland penetration more appropriately. However, AR head trajectories can be geometrically more complex and non-smooth (see Fig. S5). The fact that some of the observed displacements are caused by ARs deforming rather than actually advancing is an important caveat of the introduced method. If we locate ARs by their head coordinate instead of their centroids, all bands of enhanced edge weight shift closer towards coasts and into inland areas (Fig. S4d-f). The tARget-4 network (Fig. S4e) suggests considerably higher AR frequencies along edges in polar regions and across some regions of elevated topography, e.g., North America, in line with previous findings that the tARget differs from other ARDTs here [1].

Figure S6 demonstrates how rectangular and hexagonal gridding can yield substantially different results at high latitudes, with PageRank being inflated for some rectangular grid cells over the Antarctic interior due to enhanced cell proximity.

Edge directions are largely consistent between both catalogs (Fig. S7): mid-latitudinal AR transport orients towards the West with north/southward edges closer to the poles, respectively. Some of the ARs responsible for these north/southward edges make landfall along Arctic/Antarctic coastlines, with potentially severe impacts [2, 3].

The overall structure of the global ARTN stays consistent between the full and the post-1979 ERA5 period, as well as for MERRA2-based network.

The four different random network types that were designed as null models for this study appear visually random, apart from the properties conserved from the real ARTN (Fig. S9). The Fully Random Walker (FRW) model produces spatially isotropic edge weights (Fig. S9a). The Rewired Graph (RWG) model recovers large-scale connectivity but randomizes edges at sub-synoptic scales (Fig. Sb). The Genesis-Constrained Random Walker and the Termination-Constrained Random Walker models reproduce the prominent genesis and termination regions and edge weights fade out from there, re-

---

*tobias.braun@uni-leipzig.de

spectively (Fig. S9c/d). Four distinct realizations for each model are shown in (Fig. S10).
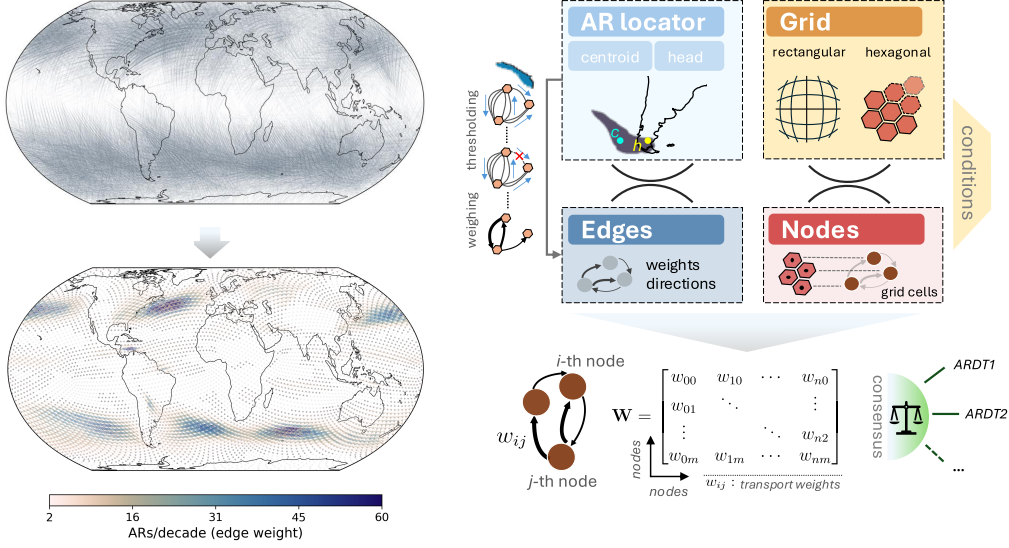

Figure S1: **Schematic illustration of how an AR network is generated from an ensemble of AR trajectories, i.e., Lagrangian data from one (or multiple) AR catalogs.** ARs are represented by a single 2-dimensional locator, e.g., their centroids or heads. AR trajectories are sequences of grid cells traversed by this locator. The grid can be defined as a conventional rectangular grid or a hexagonal grid to reduce spherical distortions. Edges are directed and weighted, denoting how frequently and into which directions AR move from a given node. This defines the (asymmetric) transport matrix $\mathbf{W}$ in which each element stores the transport frequency from node n to node m. The transport matrix can be thresholded and defines the AR network as its adjacency matrix. Given a set of different AR catalogs, an adjacency matrix can be derived from each catalog individually and a consensus derivation yields the consensus network.
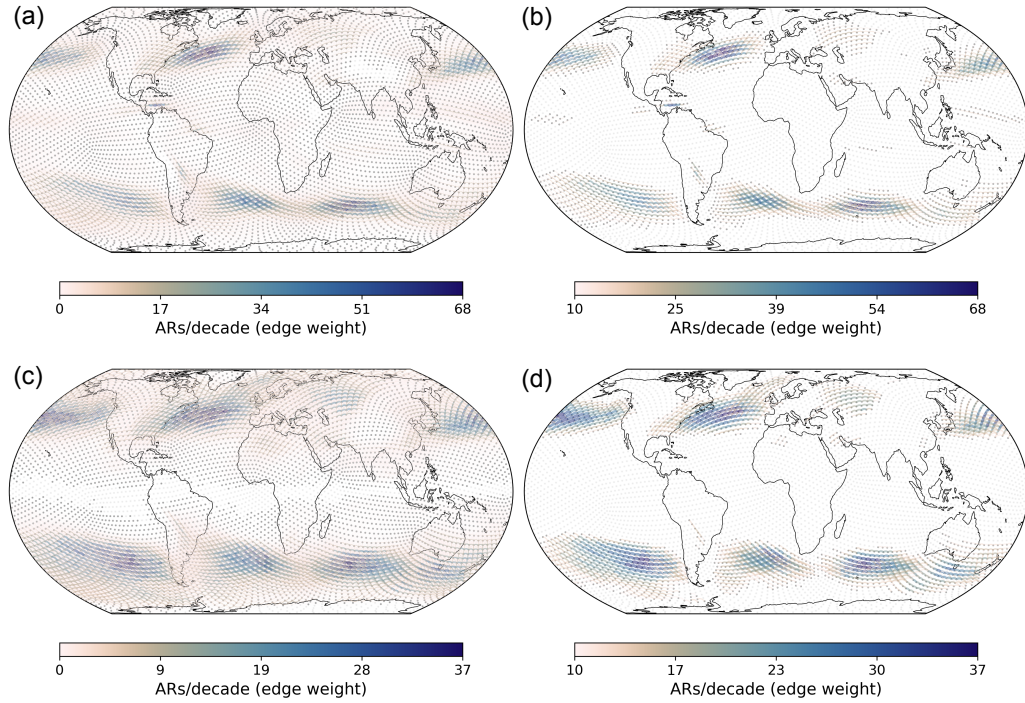
Figure S2: **AR networks for different thresholds.** No thresholding is applied in panel (a) and (c) for the AR networks derived from the PIKART and the tARget-4 catalog, respectively. This gives rise to a high number of edges with low AR frequency. Conversely, thresholding with a relatively high minimum AR frequency of at least ten ARs per decade is applied in panel (b) and (d) for the AR networks derived from the PIKART and the tARget-4 catalog, respectively. The resulting networks are more disconnected and concentrated around the main transport bands.
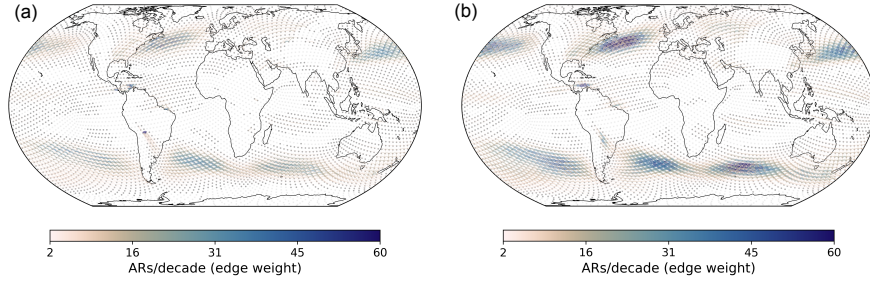
Figure S3: **AR network from the PIKART-1 catalog with ARs located by their cores (a) as compared to their centroids (b) with $\epsilon = 2\textbf{ARs/decade}$.** The AR core is the location with highest integrated vapor transport over an AR's contours. It marks a location that is potentially conducive of precipitation. While core trajectories appear less spatially centralized, the overall structures of the core-based and centroid-based AR networks align well.
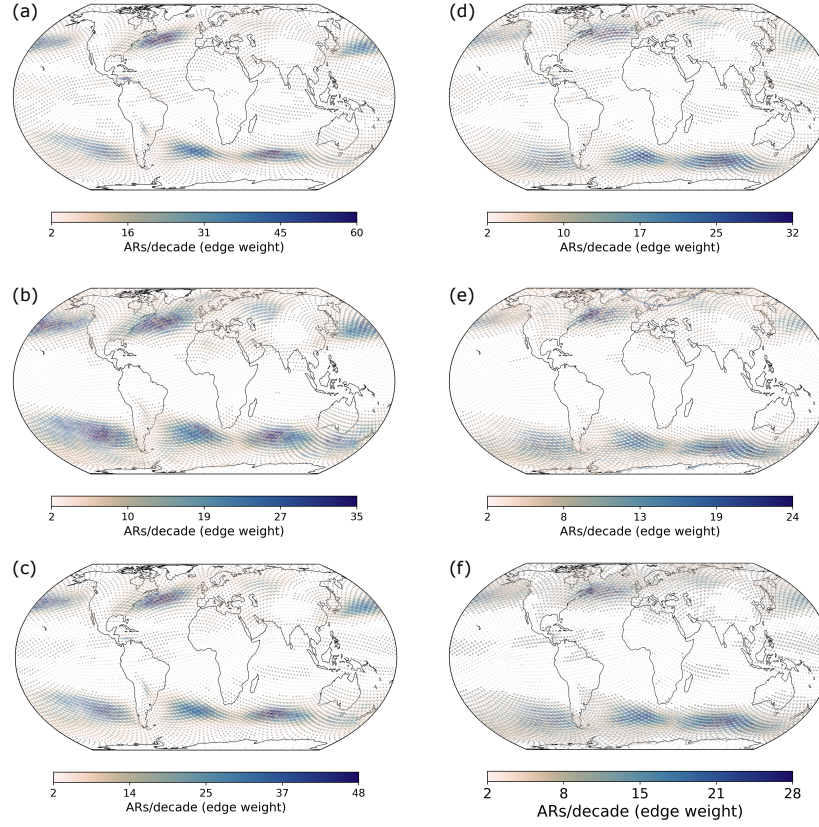
Figure S4: **AR network from each catalog ((a)/(b) and the consensus network (c), computed with ARs located by their heads.** Networks are derived from, (a) the PIKART-1 catalog, (b) the tARget-4 catalog, (c) their synthesis by means of their consensus network with $\epsilon = 2\text{ARs/decade}$. The AR head marks a location that is more informative about AR inland penetration and potential AR-induced continental precipitation. All displayed head-based networks are shifted towards the interior of continents and the poles than their centroid-based counterparts.
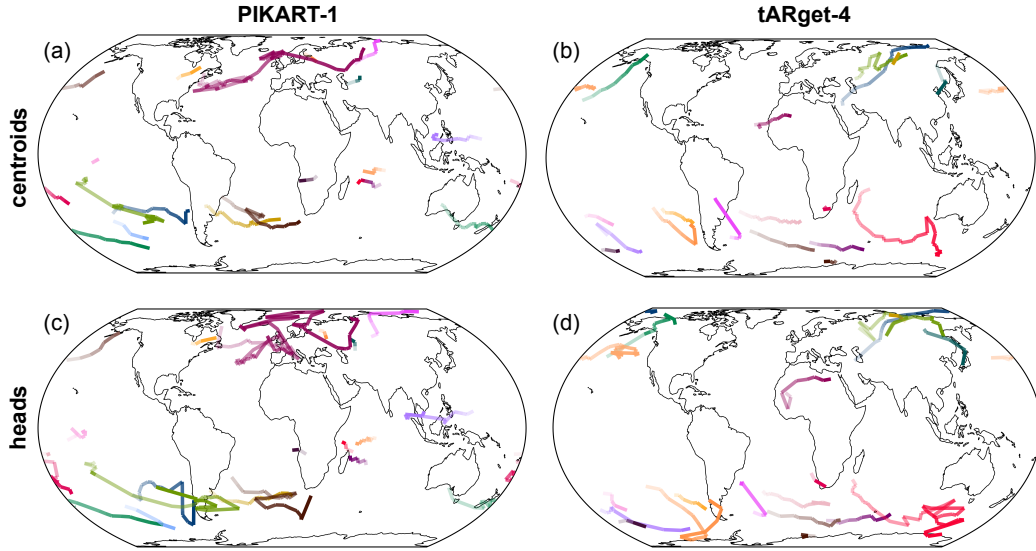
Figure S5: **Examples of AR trajectories for AR centroids (a)/(b) and AR heads (c)/(d).** The 20 displayed AR trajectories have been randomly selected from, (a)/(c) the PIKART-1 catalog, (b)/(d) the tARget-4 catalog. Individual AR trajectories, e.g., the darkred PIKART trajectory over the United Kingdom ((a)/(b)), can linger within a confined spatial region for several steps. During such a period, ARs may reside over the same location, e.g., due to orographic barriers, but due to deformations, the AR centroid/head still move. AR heads are spatially displaced from AR centroids, usually showing a more northwesterly/southwesterly tendency in the NH/SH extratropics, respectively. Their motion yields more erratic trajectories, indicating that AR head locations are more susceptible to AR deformations than AR centroids.
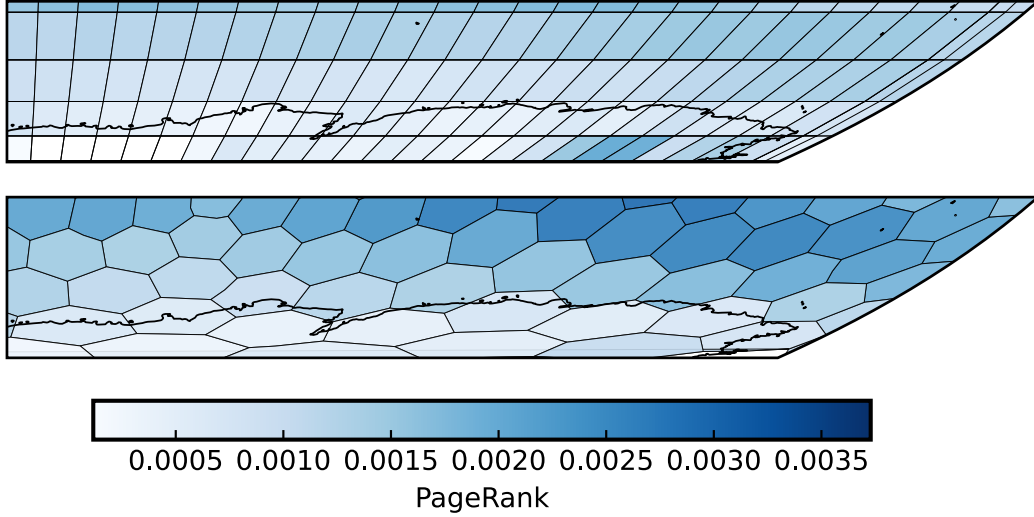
Figure S6: **PageRank scores over east Antarctica, derived from a network on a rectangular (top) and a hexagonal (bottom) grid using an equal area projection.** Grid cell sizes are selected to cover similar total areas. Towards the poles, rectangular grid cells get distorted more significantly and individual cells cover smaller and smaller areas. On the other hand, the number of hexagonal grid cells significantly shrinks towards the poles, approximately recovering equal areas over each grid cell. Rectangular grid overestimate node importance for nodes that are placed at high latitude grid cells (here given by PageRank scores): for the rectangular grid, two nodes over interior east Antarctica overprint the spatial variability in PageRank scores while for the hexagonal grid, less nodes over interior east Antarctica yield lower PageRank scores and reveal more continuous variability of node importance over the displayed region.
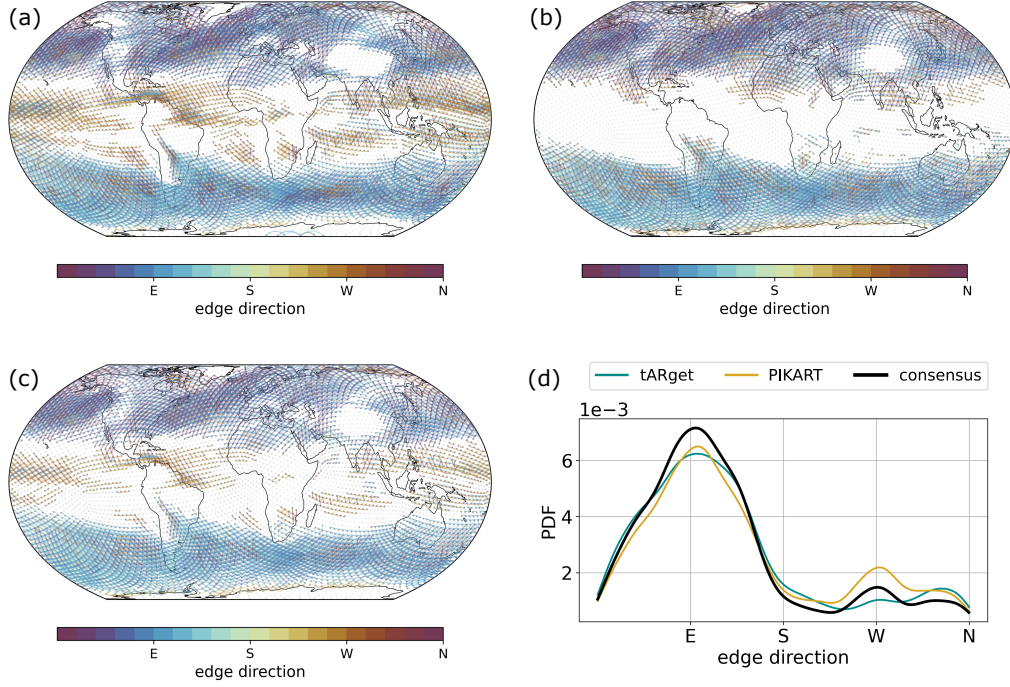
Figure S7: **AR network from each catalog ((a)/(b)) and the consensus network (c), colored by edge direction.** Networks are derived from, (a) the PIKART catalog, (b) the tARget V4 catalog, (c) their synthesis by means of their consensus network with $\epsilon = 2$ARs/decade. Edges are predominately aligned with the westerlies in both extratropical hemispheres and deviate more poleward closer to the poles. In the (sub)tropics, edges tend towards the west. Edge directions are overall consistent between both catalogs and the consensus network (d). The largest deviation results from PIKART-1 tracking ARs in the (sub)tropics more permissively than tARget-4, giving rise to a larger population of westward directed edges.
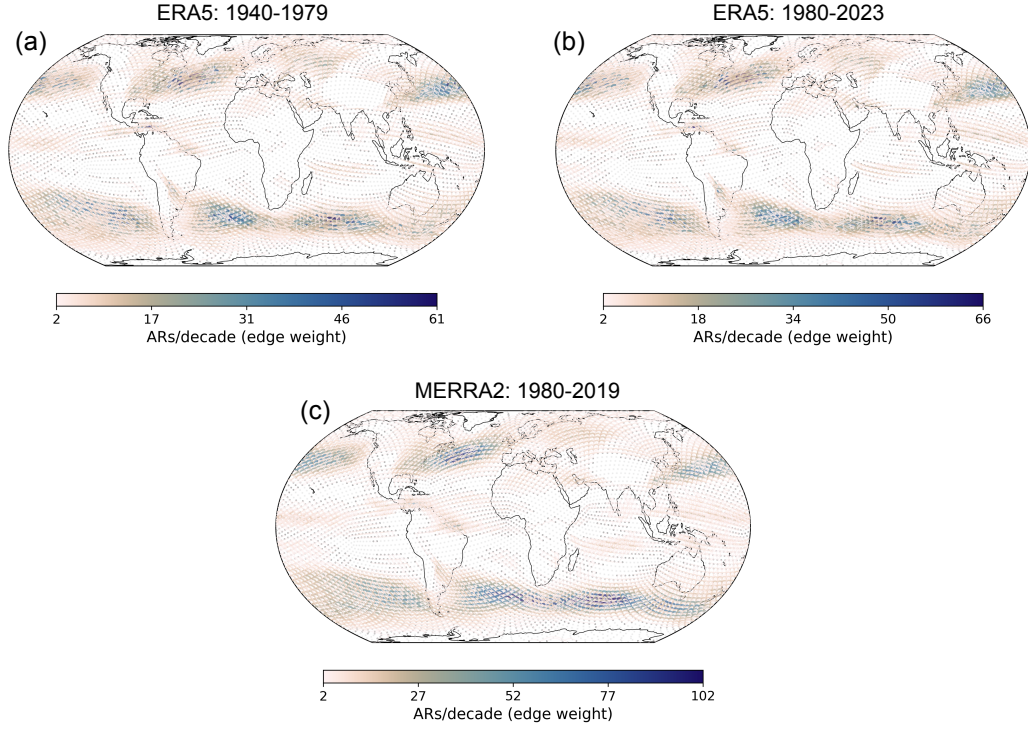
9

Figure S8: **AR consensus network computed from (a) pre-1980 ERA5, (b) post-1980 ERA5, and (c) MERRA2.** Edge weights are scaled logarithmically with $\epsilon = 2\mathrm{ARs/decade}$. Overall, the main large-scale transport patterns are recovered across all three datasets. The MERRA2-based AR network yields transport frequencies twice as high as those for ERA5 as AR trajectories have 3h- instead of 6h-resolution. While pre-1979 ERA5 is known to suffer from observational biases, the general global structure of the AR network remains robust.
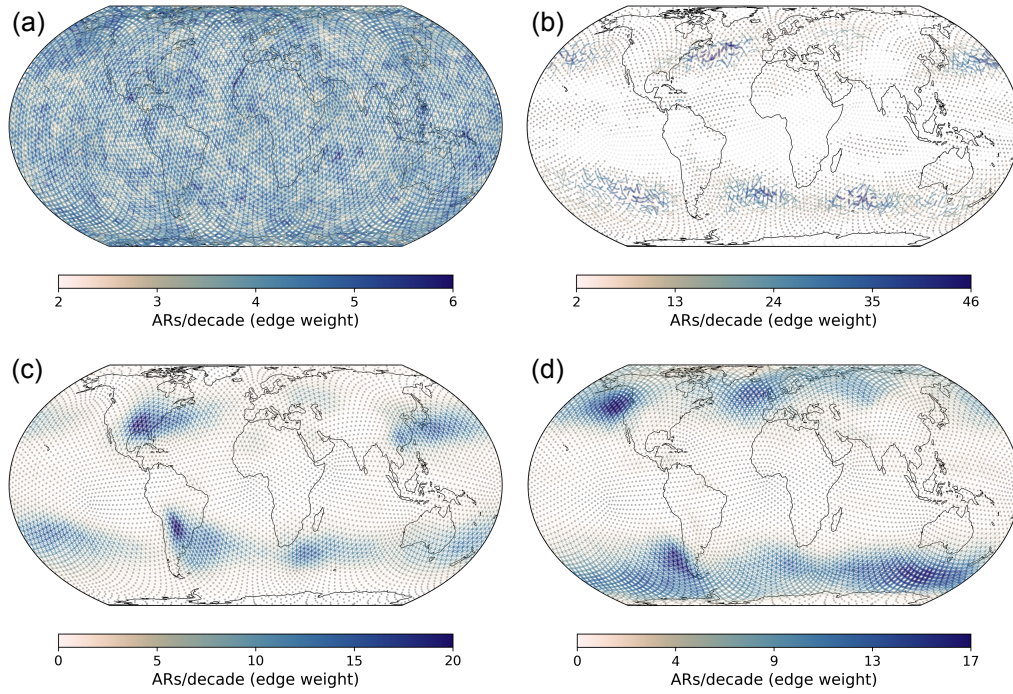
Figure S9: **The four conceptually distinct types of statistical null models considered in this study.** A single realization of the (a) Fully Random Walker (FRW), (b) Rewired Graph (RWG), (c) Genesis-Constrained Random Walker (GCW) and, (d) Termination-Constrained Random Walker (TCW) network. The resulting networks (a) exhibit fully random connectivity, (b) conserve large-scale but randomise short-range transport patterns, (c) preserve AR genesis, and (d) preserve termination regions, respectively. Colors indicate the frequency of AR transport, represented as edge weights and expressed as ARs/decade (individually for each panel).
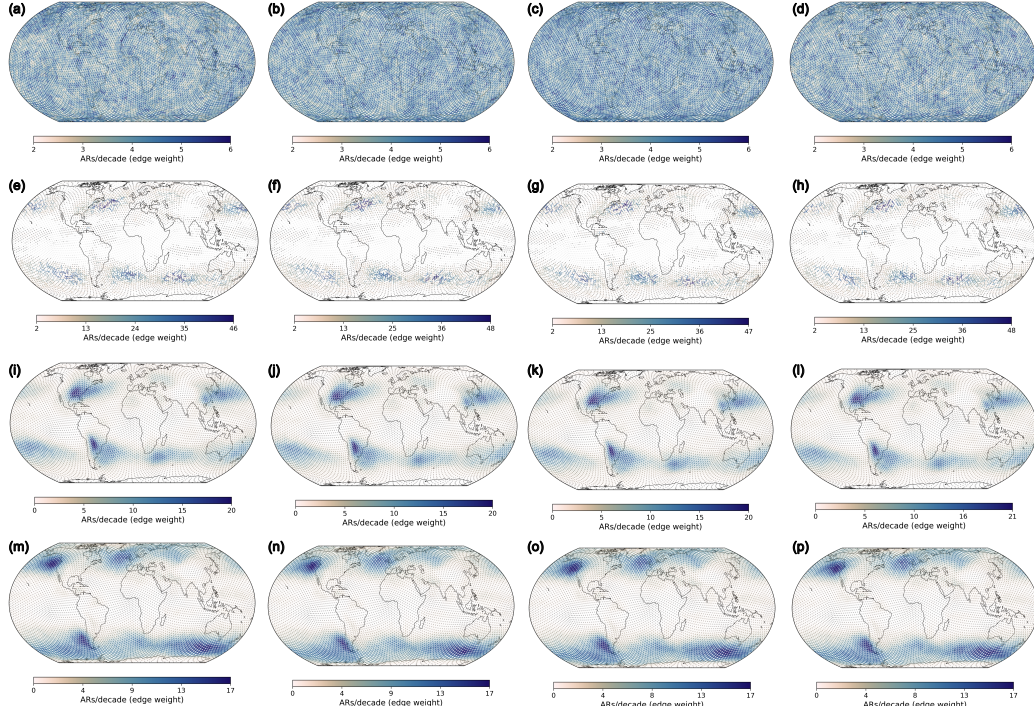
Figure S10: **Four example realizations of each null model.** (a)-(d) FRW model, (e)-(h) RWG model, (i)-(l) GCW model, (m)-(p) TCW model. All models visibly recover the constrained features, i.e., local connectivity in the RWG model ((e)-(h)), AR genesis regions in the GCW model ((i)-(l)), and AR termination regions in the TCW model ((m)-(p)). The FRW model ((a)-(d)) only constrains the number of AR trajectories and their lifetimes but clearly generates networks with random spatial structure. Differences between the realization solely arise from the stochastic, unconstrained components of the models and can be seen upon closer visual inspection of each network's small-scale connectivity.

AR hubs

Clustering coefficients characterize triadic closure, i.e., if a node's neighbour is also its neighbour's neighbour. In the ARTN, this pinpoints regions of cyclical AR transport, e.g. due to orographic barriers (Fig. S11).

In Fig. S12, we show that AR hubs shift more towards continental interiors when locating ARs by their heads, consistent with the network shift observed in Fig. S4. Separating both AR catalogs lends further credibility to the main AR hubs identified in Fig. 1 but also shows that the relative importance of individual hub regions may differ between ARDTs (Fig. S13).

Figure S14 shows the overall relationship between the discretized integrated vapor transport (IVT) changes and IVT at landfall displayed spatially in Fig. 2b (separated by AR catalog). The indicated correlation values are used to normalise the defined predictability measure: the maximum predictability is set to the correlation over all spatial locations, averaged over both catalogs. PIKART-1 gives rise to a more well-expressed monotic relationship while the high fraction of low-IVT ARs in tARget-4 yields a more complex relationship. The latter is skewed towards low values which do not appear to have strong predictive power for IVT at landfall.
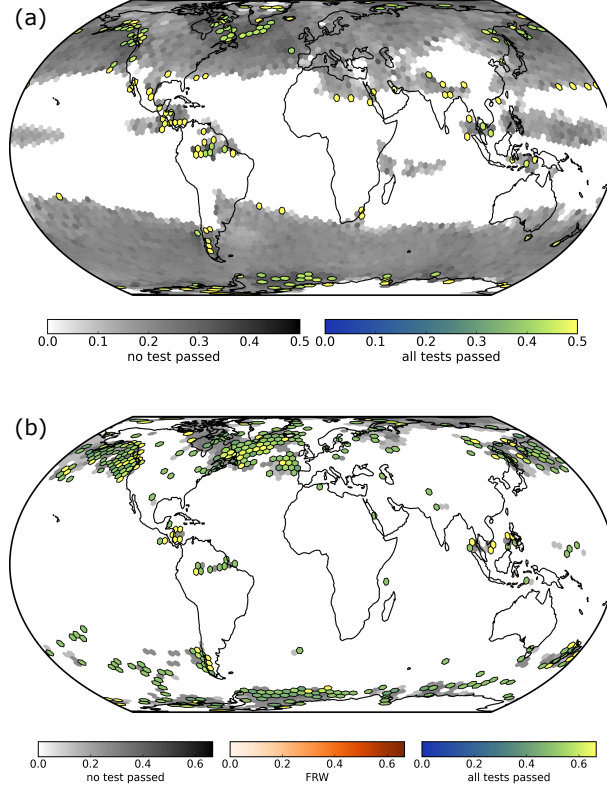
Figure S11: **Clustering coefficients for the AR consensus network.**
(a) Regular clustering coefficients, highlighting regions where ARs tend to
take various routes into all directions, resulting in closed triangles in the
AR network. (b) Cycle clustering coefficient, i.e., fraction of closed triangles
that form cycles in the AR network. Cycle clustering coefficients highlight
regions where ARs may exhibit circular movement, e.g., in regions where
they reside for prolonged time periods due to enhanced topography or where
AR transport is dense along main transport corridors. The distinct colour
maps in each panel indicate which set of tests was passed while still showing
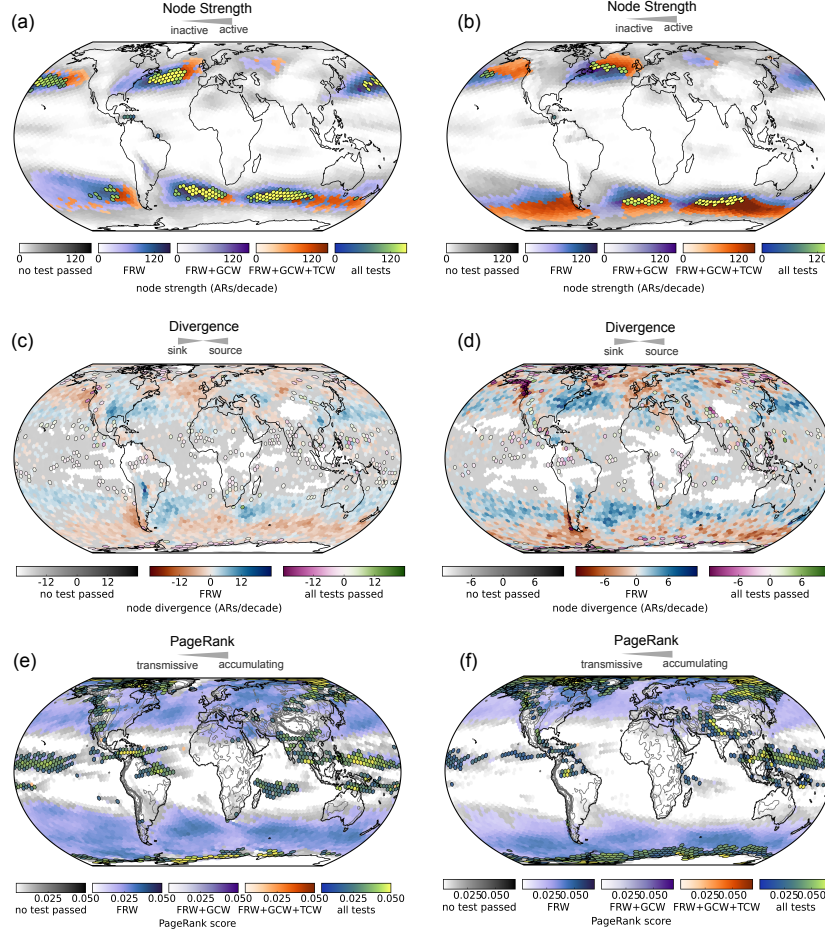continuous variations of the respective network metric.

14

Figure S12: **AR hubs, captured by node strength (a)/(b), node divergence (c)/(d) and PageRank (e)/(f) for ARs located by their centroids (a)/(c)/(e) and heads (b)/(d)/(f).** (a)/(b) Node strength of each node in the AR centroid/head network, tested against all null models in an aggregated fashion. (c)/(d) Divergence score of each node in the AR centroid/head network, tested against the FRW and all null models. (e)/(f) PageRank score (normalized to % of total accumulated 'AR walker mass') of each node in the AR centroid/head network. The distinct colour maps in each panel indicate which set of tests was passed while still showing continuous variations of the respective network metric.
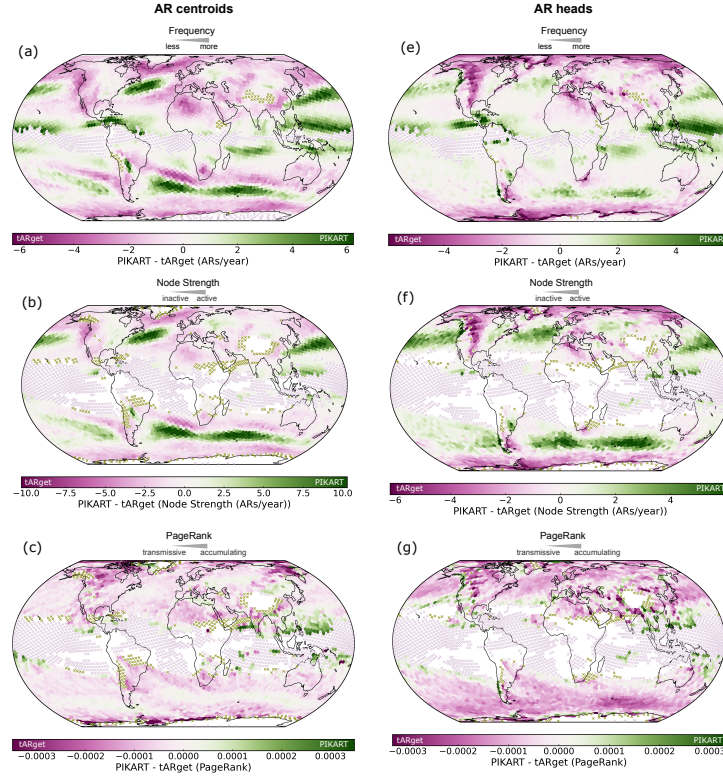
Figure S13: **Comparison between node importance in PIKART-1 and tARget-4.** All panels show differences between a property derived from the PIKART-1 catalog to those derived from the tARget-4 catalog at each location. The left column (a)-(c) locates ARs by their centroids, the right column (d)-(f) by their heads. Green (pink) colors denote higher importance in PIKART-1 (tARget-4). Green (pink) crosses mark nodes where there are not sufficient ARs in the PIKART-1 (tARget-4) catalog. Independent of any AR networks, (a)/(d) show AR frequency differences across both catalogs. PIKART-1 tracks more ARs in the tropics, along the main transport basins and over parts of interior South America. tARget-4 exhibits enhanced AR frequencies at the poles, over interior North America and other continental regions, e.g., North Africa. Differences in node importance based on node strength (b)/(e) largely align with these patterns. Nodes with crosses in (b)/(e) and no crosses in (a)/(d) arise due to thresholding of the network. Differences in node importance based on PageRank scores (c)/(f) show that tARget-4 assigns higher PageRanks to most nodes, except for tropical regions (for AR centroids and heads) and some coastal regions such as western North American, Chile, Greenland and Antarctica for AR heads.
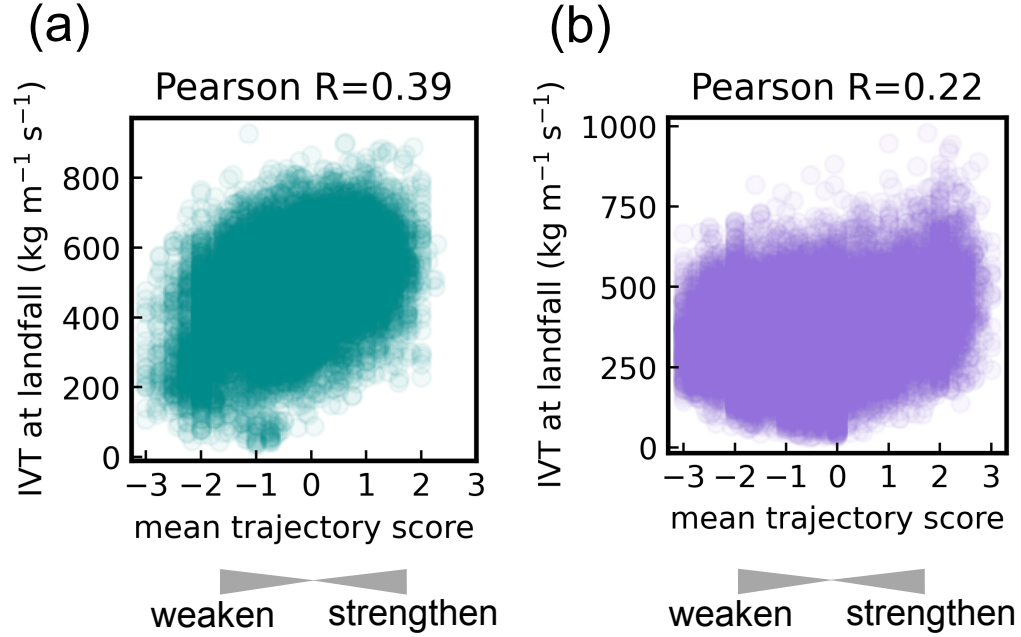
16

Figure S14: **Scatter plots of mean trajectory score (based on IVT differences) versus IVT at landfall for both catalogs.** First, discretized IVT difference values (between -3 and 3) have been assigned to each node in the ARTN (Fig. S16). For each AR that lives longer than 18hrs, the trajectory score is computed as the average of all IVT difference scores it encounters during its lifecycle. These are predictive of IVT at landfall as ARs tend to intensify/weaken their IVT along their paths, resulting in the indicated statistically significant correlation values. The dependencies are shown separately for the (a) PIKART-1 and (b) tARget-4 catalog. The predictability scores shown in Fig. 2b of the main manuscript are derived from the residuals of a robust Huber regression through the displayed points, scaled by average of the indicated correlation values.

17

AR highways

All edge betweenness centralities (EBC) from the real consensus network are tested against the ones expected from the four null models (Fig S15). Interestingly, the FRW model yields the largest significance threshold (upper 90%-quantile, Fig S15a) whereas the more systematic models entail lower EBC. We hypothesize that the reason for this counter-intuitive result could be that in a more random network, edge weights are homogeneously distributed so that many edges can bridge different parts of the network. If thresholded by the highest displayed EBC-threshold from the FRW model (Fig S15a, orange line), the main AR highways discussed in the manuscript are still deemed significant.

The discretized IVT classifiers used for Fig. 3 (and also Fig. 2a) are defined through the quantiles displayed in Fig S16. The IVT difference log-distributions of both AR catalogs are smoothed using a kernel density estimate for better comparability. While the intervals increase in size towards the distributions' tails, the covered probability mass shrinks.

While Fig. 3 directly shows the AR highways for the seasonally and ENSO-conditioned ARTNs, the underlying networks are shown in Fig. S17. Differences are not as striking as they are through the AR highways but visible. The classification of El Niño/La Niña phases is based on the time series discretizazion shown in Fig. S18. To better illustrate where seasonal/ENSO-phase induced IVT changes are most pronounced, Fig. 19 plots these differences explicitly. We clearly observe the strongest contrasts in the (sub)tropics whereas in the mid-latitudes, regions of enhanced IVT strengthening/weakening do not shift that strongly. The most marked exceptions are likely the enhanced IVT intensification in boreal summer over the eastern North Atlantic and the eastern North Pacific, as well enhanced IVT depletion towards Australia in austral winter.

The notion of 'AR highways' prompts the question of how closely individual ARs actually follow them and whether this could offer enhanced predictability. We classify ARs into three groups based on the mean EBC of all edges they traverse: *Conformists* (top 33%-EBC), *Straddlers* (middle 33%-EBC), and *Strays* (bottom 33%-EBC). These categories reflect decreasing alignment with AR highways (Fig. S20a-c). Trajectories were smoothed using B-spline interpolation. While *Conformists* tend to follow AR highways most rigidly, *Straddlers* visibly drift off AR highways more often and behave more intermittently. *Stray* trajectories exhibit marked anomalies, e.g., cir-

cular movement, or venture far off the main branches. *Conformists* differ significantly in mean IVT from both *Straddlers* and especially from *Strays* (moderate standardized Wasserstein distance d_w of $0 \leq 0.27 \leq 1$, $p < 0.01$ for two-sided Kolmogorov-Smirnov and Mann-Whitney U-test, Fig. S20d), while the latter two are similar ($d_w = 0.09$). Combined with visible differences in lifetimes (Fig. S20e), this suggests that ARs that stray off highways tend to "live" longer and be more intense.
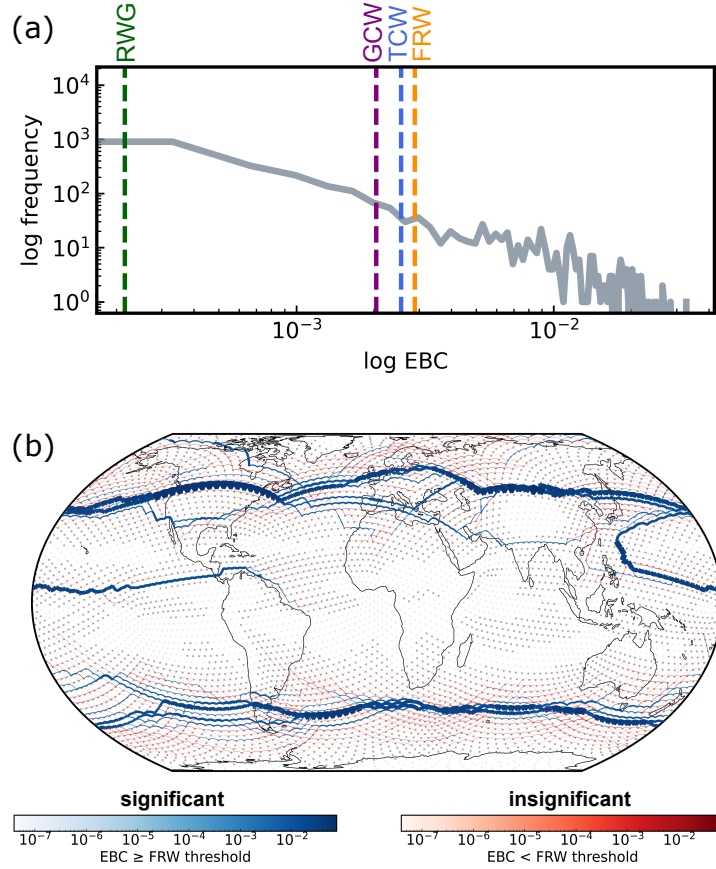

Figure S15: **Significance testing of the identified global AR highways.** (a) Histogram of edge betweenness centralities from (unconditioned) real consensus network with $\epsilon = 2\text{ARs/decade}$ (log-log-scale). The upper 90%-quantiles for each null model ensemble (based on 200 model realizations) are shown as colored dashed lines. (b) If thresholded by the highest of these quantiles (FRW model), the main AR highways remain significant.
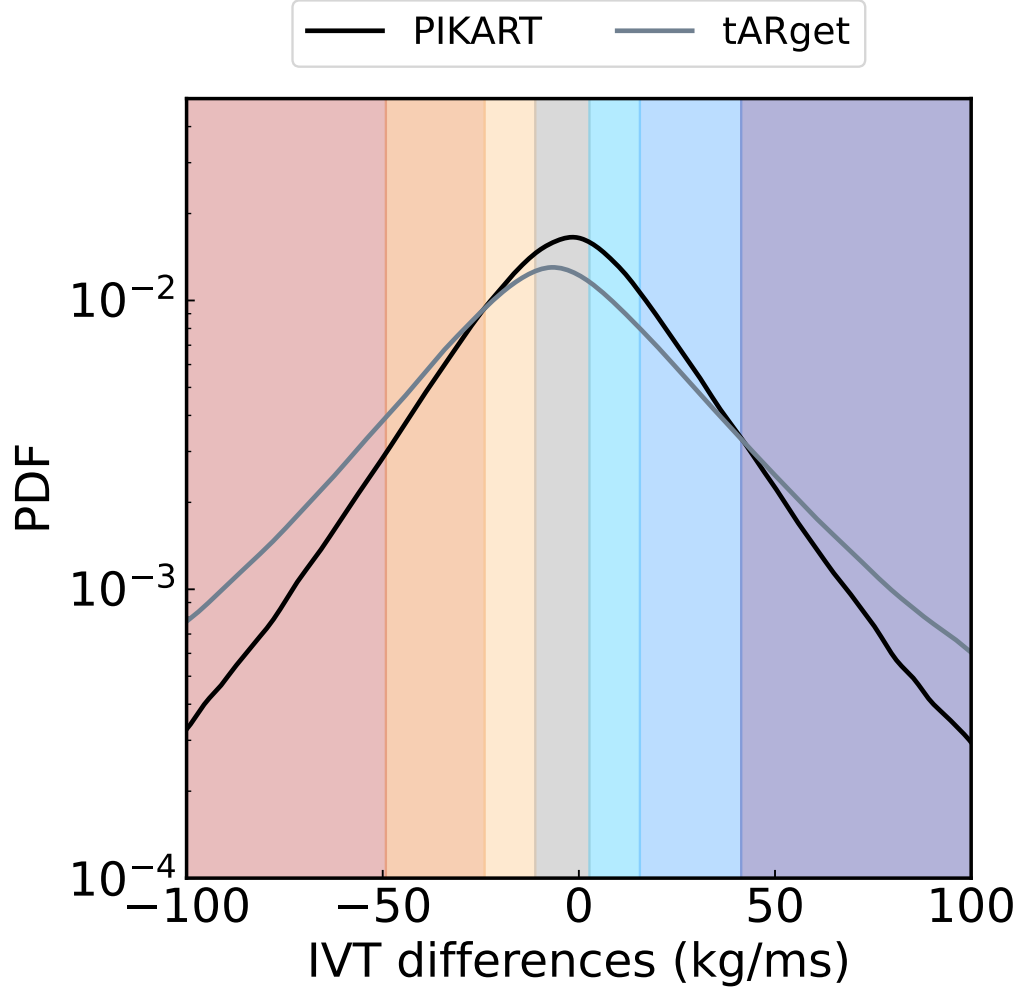
19

Figure S16: **Kernel density estimate of log-probability density functions of differences in integrated vapor transport**. Differences are computed between each pair of consecutive AR steps for each of the two catalogs. Colored shading highlights the corresponding thresholds used for discretization of IVT differences (10%, 25%, 40%, 60%, 75%, 90% quantiles derived from IVT differences in both catalogs together, from left to right). Red denotes large IVT losses and blue denotes large IVT gains. IVT losses and gains can be both caused by changes in wind speed and relative humidity.
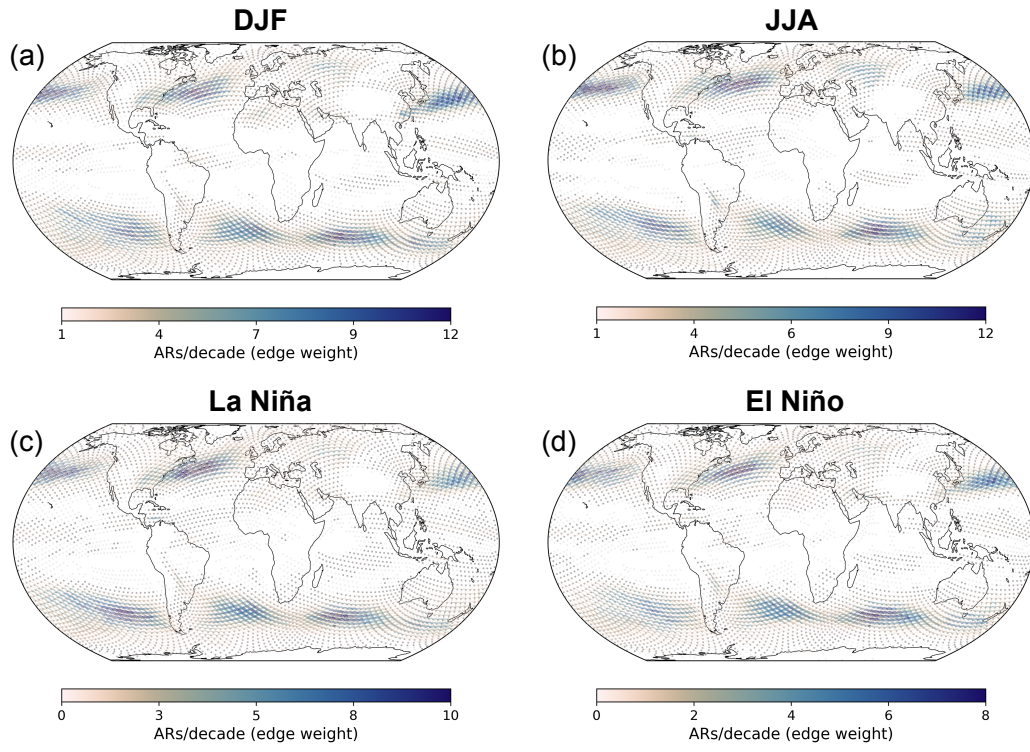
Figure S17: **Conditional consensus networks.** (a)/(b) AR networks conditioned on (a) boreal winter (DJF) and (b) summer (JJA). (c)/(d) AR networks conditioned on (a) La Niña and (b) EL Niño conditions.
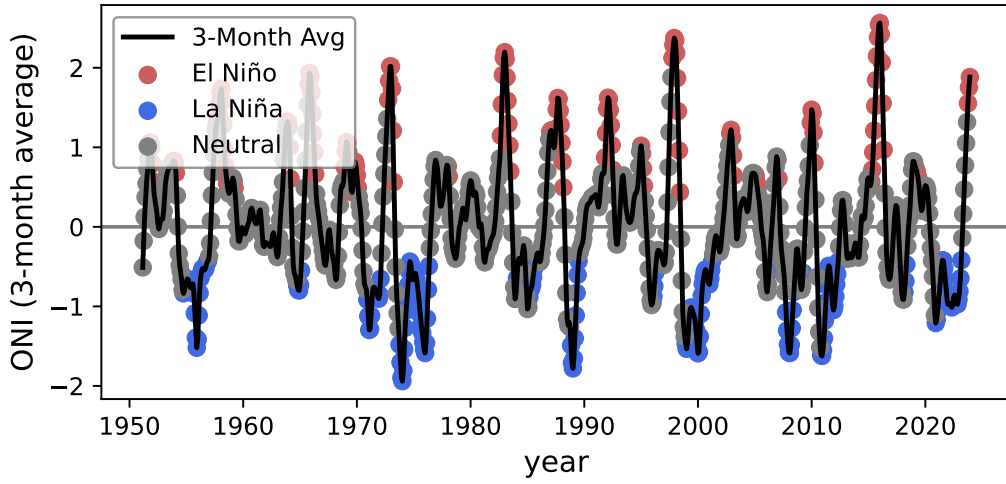
Figure S18: **Moving average of Oceanic Niño Index (ONI) z-scores between 1950-2023.** Grey dots denote neutral conditions. Red (blue) dots denote El Niño (La Niña) conditions, identified as excursions of at least five consecutive 3-months averages above 0.5 (below -0.5). The identified conditions are used to condition the AR network.
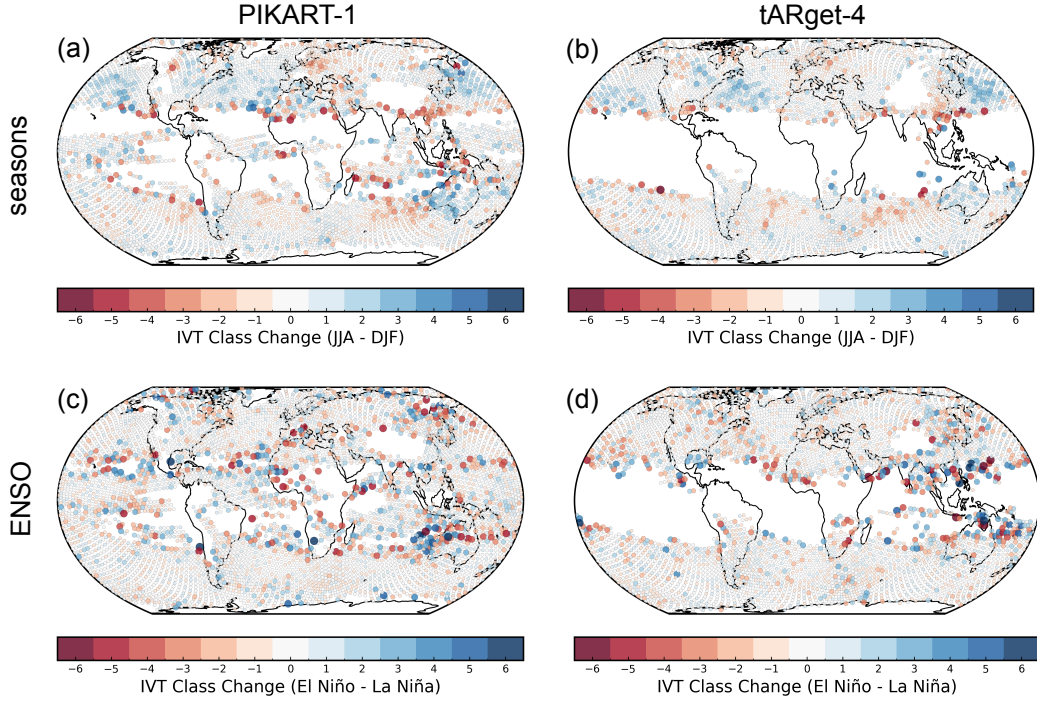
Figure S19: **Differences between net IVT changes across nodes for seasonally end ENSO-conditioned AR networks.** (a)/(b) Difference of discretized net IVT changes between boreal summer (JJA) and winter (DJF) for each node in the (a) PIKART-1 and (b) tARget-4 network. (c)/(d) Difference of discretized net IVT changes between El Niño and La Niña for each node in the (c) PIKART-1 and (d) tARget-4 network. As net IVT changes were discretized into seven classes (ranging from large losses (-3) to large gains (+3)), nodes can change their role in 13 ways (ranging from a change from large gains (+3) to large losses (-3), yielding -6, to a change from large losses (-3) to large gains (+3), yielding +6). Red bubbles denote nodes that become more of a sink (larger losses) during JJA or El Niño conditions, respectively. Blue bubbles denote nodes that become more of a source (larger gains) during JJA or El Niño conditions, respectively.
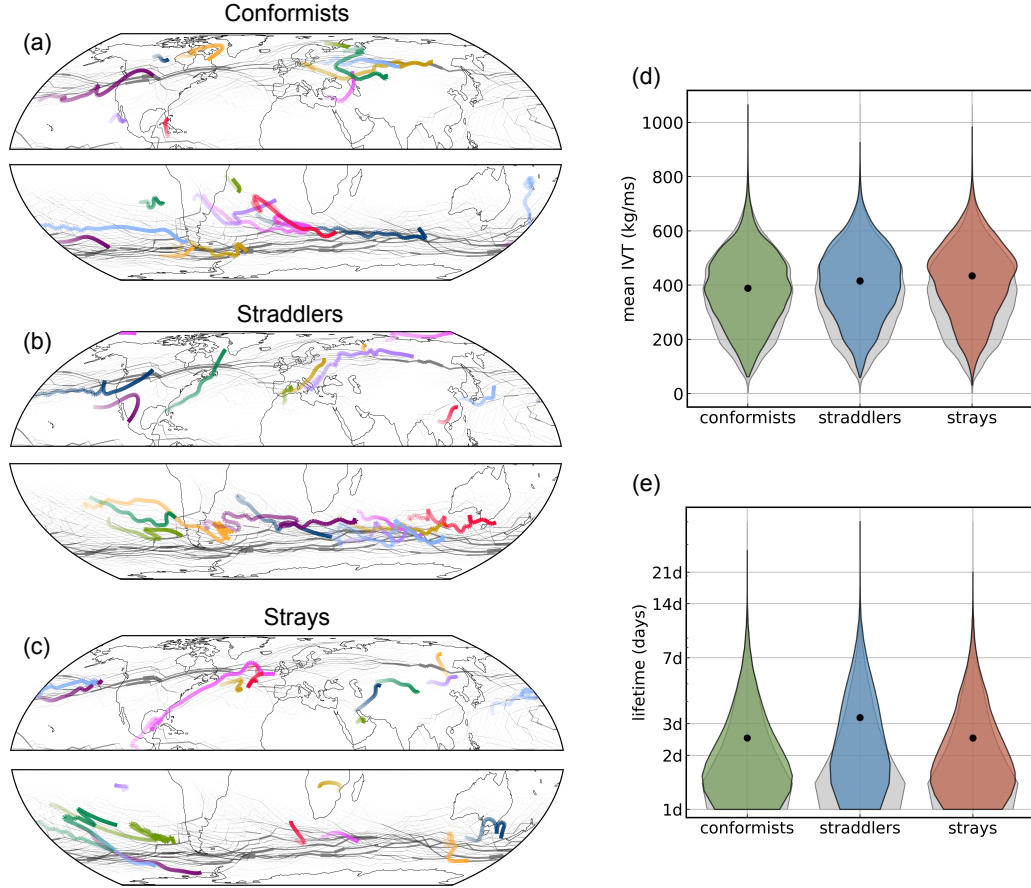
Figure S20: **Trajectories of three AR populations that exhibit different tendencies in following the identified boreal winter (DJF) AR highways.** (a) Trajectories of ten randomly selected ARs that were classified as *Conformists* (top-33% highest EBC-trajectories), (b) same but for ARs that were classified as *Straddlers* (middle-33% highest EBC-trajectories), (c) same but for ARs that were classified as *Strays* (lowest-33% EBC-trajectories). Network edges (grey) are weighed by their EBC. Opacity indicates temporal progression. (d) IVT distribution of distinct AR populations. (e) Logarithmic lifetime distribution of distinct AR populations. In (d)/(e), the overall distribution is plotted in grey in the background for reference.

AR basins

Evaluation metrics of the non-hierarchical AR basins are shown in Fig. 20. Overall, all continental AR basins are less well-separated and are detected less consistently across distinct runs of Infomap (Fig. S20b/d). Hence, their exact contours cannot be delineated as robustly. Some isolated basins (e.g., over Alaska) exhibit high flow ratios simply due to their low degree of connectivity to other communities (Fig. S20b, purple). Stability and quality of one null model realization per null model are shown in Fig. S21.

In comparison to the four null models (Fig. S22), most of the detected AR communities are significantly more stable and well-separated than those obtained from the FRW network model and, on average, also slightly more stable than those detected in the GCW/TCW network models (which naturally have a strong community structure in the genesis/termination regions, respectively). Only the RWG network ensemble yields significantly more robust and separated communities, likely since the implemented distance constraint restricts rewiring to local connections, resulting in confined, well-structured communities. Finally,
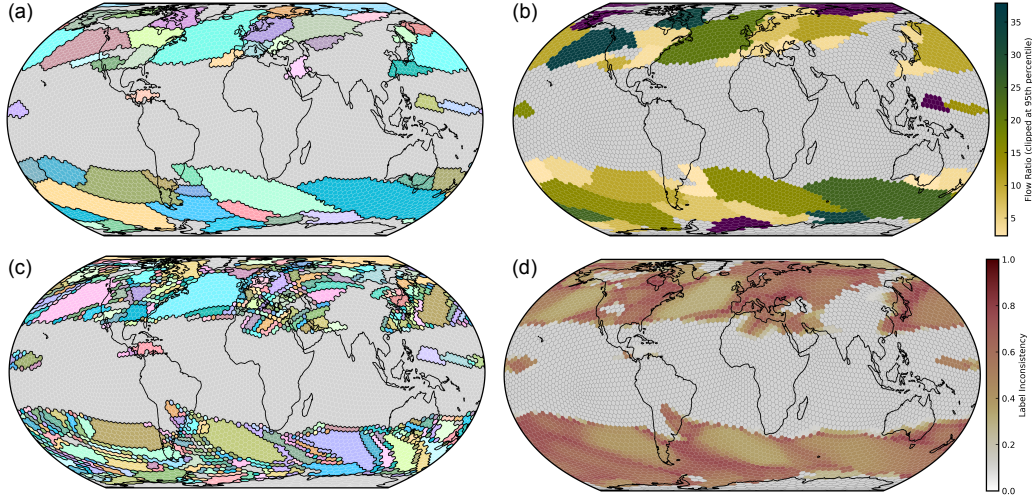
Figure S21: **Stability and quality of AR basins identified through Infomap community detection on AR consensus network ($\epsilon = 1$ARs/decade)**. (a) Most central partition, defined as the set of communities that minimizes overall label incosistency across all nodes. (b) Flow ratios, indicating partition quality by means of intra- and inter-community flow. (c) Consensus partition, i.e., the most central community a node belongs to considering all of its distinct community labels across runs. (d) Label inconsistency, i.e., the heterogeneity of a node labels from all of its distinct community labels across runs. Both of the latter are computed from the Jaccard similarity between the different communities a node is assigned to across different runs. All measures are derived from 100 runs of the (non-hierarchical) Infomap algorithm, each seeded with a distinct random seed.
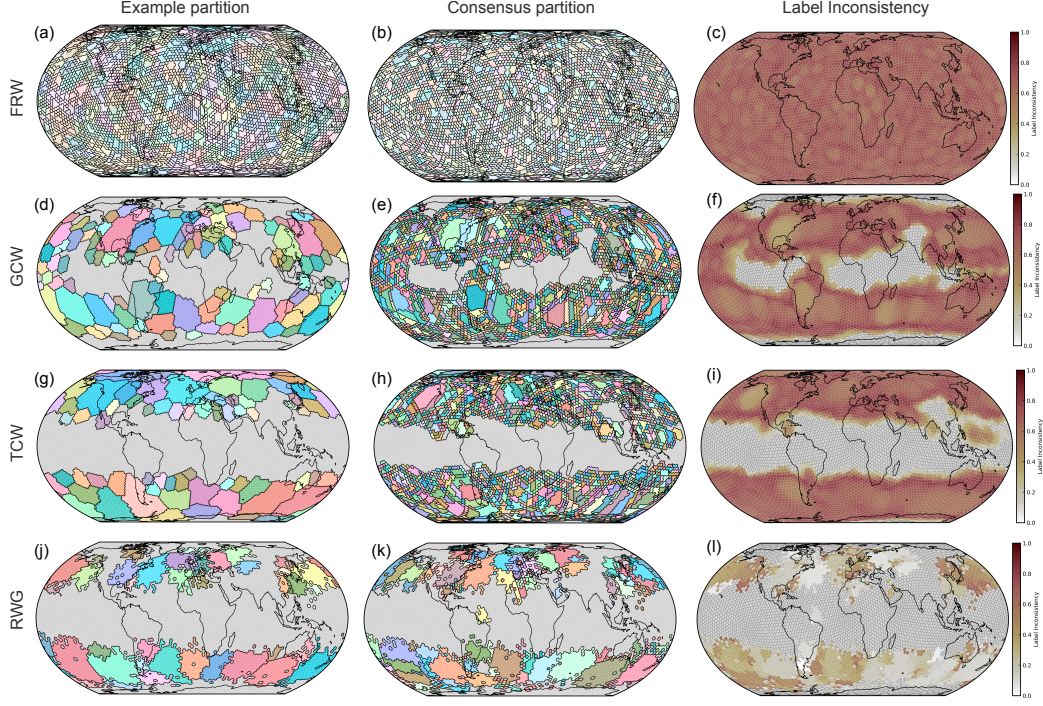
Figure S22: **Stability and quality of AR basins identified through Infomap community detection on a single example realization of each of the null model networks ($\epsilon = 1\mathbf{ARs/decade}$)**. Rows show output for distinct null models. (a)/(d)/(g)/(j) Selected example partition. (b)/(e)/(h)/(k) Consensus partition, i.e., the most central community a node belongs to considering all of its distinct community labels across runs. (c)/(f)/(i)/(l) Label inconsistency, i.e., the heterogeneity of a node labels from all of its distinct community labels across runs. Both of the latter are computed from the Jaccard similarity between the different communities a node is assigned to across different runs. All measures are derived from 100 runs of the (non-hierarchical) Infomap algorithm, each seeded with a distinct random seed.
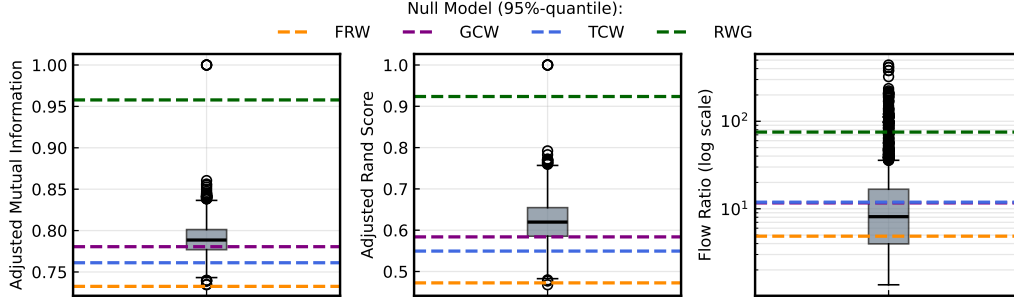
Figure S23: **Significance testing of stability and quality of AR basins identified through Infomap community detection.** Stability is measured through (a) Adjusted Mutual information and (b) Adjusted Rand Index. (c) Separation of communities is measured through flow ratios. The distribution of each metric is shown as a boxplot, providing the range of values observed for the communities identified across different runs of Infomap on the real AR consensus network. Horizontal dashed lines indicate the 95%-quantiles of all four null models. The real AR basins are consistently more stable and well-separated than expected from the FRW model (yellow line). The GCW and TCW model yield comparable stability and more well-separated communities as AR transport concentrates around genesis/termination hubs (blue/purple line). The RWG model yields very stable and well-separated communities, likely due to the hard constraint on local rewiring of nodes (green line).

References

[1] Juan M Lora, CA Shields, and JJ Rutz. Consensus and disagreement in atmospheric river detection: ARTMIP global catalogues. *Geophysical Research Letters*, 47(20):e2020GL089302, 2020.

[2] Michelle L Maclennan, Andrew C Winters, Christine A Shields, Rudradutt Thaker, Léonard Barthelemy, Francis Codron, and Jonathan D Wille. Rising atmospheric moisture escalates the future impact of atmospheric rivers in the Antarctic climate system. *Communications Earth & Environment*, 6(1):1–10, 2025.

[3] Zhibiao Wang, Qinghua Ding, Renguang Wu, Thomas J Ballinger, Bin Guan, Deniz Bozkurt, Deanna Nash, Ian Baxter, Dániel Topál, Zhe Li, et al. Role of atmospheric rivers in shaping long term Arctic moisture variability. *Nature Communications*, 15(1):5505, 2024.