Evaluating the Representations of Atmospheric Rivers and Their Associated Precipitation in Reanalyses with Satellite Observations

Weiming Ma[1,2], Gang Chen[1], Bin Guan[3,4], Christine A Shields[5], Baijun Tian[4], and Emilio Yanez[1,1]

[1]Department of Atmospheric and Oceanic Sciences, University of California
[2]

[3]Joint Institute for Regional Earth System Science and Engineering, University of California
[4]Jet Propulsion Laboratory, California Institute of Technology
[5]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research

August 1, 2023

Abstract

Atmospheric rivers (ARs) are filaments of enhanced horizontal moisture transport in the atmosphere. Due to their prominent role in the meridional moisture transport and regional weather extremes, ARs have been studied extensively in recent years. Yet, the representations of ARs and their associated precipitation on a global scale remains largely unknown. In this study, we developed an AR detection algorithm specifically for satellite observations using moisture and the geostrophic winds derived from 3D geopotential height field from the combined retrievals of the Atmospheric Infrared Sounder and the Advanced Microwave Sounding Unit on NASA Aqua satellite. This algorithm enables us to develop the first global AR catalog based solely on satellite observations. The satellite-based AR catalog is then combined with the satellite-based precipitation (Integrated Muti-SatellitE Retrievals for GPM) to evaluate the representations of ARs and AR-induced precipitation in reanalysis products. Our results show that the spreads in AR frequency and AR length distribution are generally small across datasets, while the spread in AR width is relatively larger. In terms of the AR-induced precipitation, both AR-induced mean and extreme precipitation are too weak nearly everywhere in reanalyses. However, all reanalyses tend to precipitate too often under AR conditions, especially over low latitude regions. This finding is consistent with the "drizzling" bias which has plagued generations of climate models. Overall, the findings of this study can help to improve the representations of ARs and associated precipitation in reanalyses and climate models.

Evaluating the Representations of Atmospheric Rivers and Their Associated Precipitation in Reanalyses with Satellite Observations

Weiming Ma[1,5], Gang Chen[1], Bin Guan[2,3], Christine A. Shields[4], Baijun Tian[3] and Emilio Yanez[1]

[1]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA

[2]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA

[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

[4]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

[5]Now at: Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, Washington

Corresponding author: Weiming Ma (mawei@atmos.ucla.edu), Gang Chen (gchenpu@atmos.ucla.edu)

Key Points:

- Developed a novel method to detect atmospheric rivers using moisture and wind data from satellite observations
- Reanalyses and satellite observations show high agreement with each other in atmospheric river frequency distributions
- Under atmospheric river conditions, reanalyses tend to precipitate too often and too lightly

Abstract

Atmospheric rivers (ARs) are filaments of enhanced horizontal moisture transport in the atmosphere. Due to their prominent role in the meridional moisture transport and regional weather extremes, ARs have been studied extensively in recent years. Yet, the representations of ARs and their associated precipitation on a global scale remains largely unknown. In this study, we developed an AR detection algorithm specifically for satellite observations using moisture and the geostrophic winds derived from 3D geopotential height field from the combined retrievals of the Atmospheric Infrared Sounder and the Advanced Microwave Sounding Unit on NASA Aqua satellite. This algorithm enables us to develop the first global AR catalog based solely on satellite observations. The satellite-based AR catalog is then combined with the satellite-based precipitation (Integrated Muti-SatellitE Retrievals for GPM) to evaluate the representations of ARs and AR-induced precipitation in reanalysis products. Our results show that the spreads in AR frequency and AR length distribution are generally small across datasets, while the spread in AR width is relatively larger. In terms of the AR-induced precipitation, both AR-induced mean and extreme precipitation are too weak nearly everywhere in reanalyses. However, all reanalyses tend to precipitate too often under AR conditions, especially over low latitude regions. This finding is consistent with the “drizzling” bias which has plagued generations of climate models. Overall, the findings of this study can help to improve the representations of ARs and associated precipitation in reanalyses and climate models.

Plain language summary

Atmospheric rivers (ARs) are filaments of enhanced horizontal moisture transport in the atmosphere. These weather systems are responsible for most of the poleward atmospheric moisture transport over mid-latitudes and can cause extreme precipitation around the world. For a long time, researchers relied heavily on reanalysis products to study ARs. Albeit incorporating information from observations, reanalyses are produced by numerical models and thus should not be treated as real observations. In this study, for the first time, we developed a near global AR detection algorithm specially for satellite observations. Unlike previous AR detection algorithms designed for satellite observations, which were applicable to mean moisture content and on regional scales, our algorithm utilizes both the moisture field and wind information from satellite observations. Our algorithm can thus better characterize the transport nature of the detected ARs. Using both the developed algorithm and NASA satellite observations, we developed the first satellite-based near global AR catalog. This satellite-based AR catalog together with NASA satellite precipitation product was then used to evaluate how well major reanalyses represent ARs and their associated precipitation. We found that reanalyses generally perform well in representing the AR occurrence frequency and length, but show relatively larger uncertainty in representing the AR width. In terms of AR precipitation, ARs in reanalyses precipitate too lightly and too often. Our findings can help to improve the representation of ARs and associated precipitation in reanalyses and climate models. As the quality of satellite observations continues to improve, the methodology presented here can be applied to other satellite observations to develop higher resolution or higher frequency AR statistics.

1 Introduction

Characterized by filaments of enhanced moisture transport in the atmosphere, atmospheric rivers (ARs) play a critical role in the global hydrological cycle. Despite only covering a very small fraction of the Earth's circumference, an early study revealed that ARs on average contribute to more than 90% of the poleward moisture transport over midlatitudes (Zhu & Newell, 1998). At the regional scale, depending on their strength and duration, ARs can exert either beneficial or detrimental impacts upon landfall (Eiras-Barca et al., 2021; Ralph et al., 2019). It has been shown that ARs are important freshwater suppliers to many coastal regions around the world and can serve as effective drought busters (Dettinger, 2013; Dettinger et al., 2011; Guan et al., 2010; Rutz & Steenburgh, 2012; Viale et al., 2018). For example, it has been estimated that up to half of the annual precipitation over California is delivered by ARs (Dettinger et al., 2011). Meanwhile, intense ARs making landfall usually lead to a wide range of weather hazards, such as wind and precipitation extremes, and flooding (Chen et al., 2018; Henn et al., 2020; Kim et al., 2018; Lamjiri et al., 2017; Lavers & Villarini, 2013; Ma et al., 2020a; Paltan et al., 2017; Ralph et al., 2006; Waliser & Guan, 2017). In recent years, there are an increasing number of studies on ARs' roles in sea ice variability (Hegyi & Taylor, 2018; Wang et al., 2020; Woods & Caballero, 2016; Zhang et al., 2023) and ice shelf stability (Djoumna & Holland, 2021; Mattingly et al., 2018; Wille et al., 2019, 2022), extending the understanding of AR impacts beyond mid-latitude areas.

Most AR studies have treated reanalysis products as observations (DeFlorio et al., 2019; Guan & Waliser, 2017; Massoud et al., 2019). Yet, reanalyses are not obtained by direct observations but produced by models which are constrained by observations through data assimilation. Since reanalyses are model-based "observations", it is expected that each reanalysis would have its own biases intrinsic to the model used to produce it, especially over the regions where the observation networks are sparse (e.g., Guan et al., 2020). However, most AR studies, which use reanalyses as observations, usually assume that ARs in reanalyses are representative of the true observation. Given the inherent errors of the models used to produce these reanalyses, such an assumption needs to be justified. So far, studies on the intercomparison between reanalyses and observations of the AR representations are very limited, which reduces our confidence in the ability of reanalyses in representing ARs. By evaluating six AR events in reanalyses against aircraft observations, Ralph et al. (2012) concluded that Climate Forecast System Reanalysis (CFSR) (Saha et al., 2010) , Modern-Era Retrospective analysis for Research and Applications (MERRA) (Rienecker et al., 2011) and European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERAI) (Dee et al., 2011) exhibit comparable skills in representing the characteristics of these six ARs. These three reanalyses also have better performance compared to National Centers for Environmental Prediction (NCEP) -National Center for Atmospheric Research (NCAR) Reanalysis I (NCEP R1) (Kalnay et al., 1996), Tropospheric Chemistry Reanalysis (TCR) (Miyazaki et al., 2012), and North American Regional Reanalysis (NARR) (Mesinger et al., 2006). Expanding the sample size to 21 AR events, Guan et al. (2018) found that, compared to dropsonde observations, ERAI and MERRA, Version 2 (MERRA-2) (Gelaro et al., 2017) have a mean error of -2% and -8% in AR width, respectively, and +3% and -1% in total integrated water vapor transport (IVT), respectively. Using MERRA-2, ECMWF Reanalysis Version 5 (ERA5) (Hersbach et al., 2020) and Japanese 55-year Reanalysis (JRA-55) (Kobayashi et al., 2015), a recent study from the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) (Shields et al., 2018) found that ARs tend to get detected more frequently in MERRA-2 due to its higher climatological IVT and also noted that ARs in ERA5 tend to be narrower due to its finer spatial resolution (Collow et al., 2022). It is immediately apparent that the above studies either evaluate AR representations in reanalyses using a small sample of observations over a limited area, or evaluate uncertainty in AR representations based on intercomparing between a few reanalyses themselves. In addition, these studies have only examined the representation of basic AR characteristics, such as AR frequency and intensity, in reanalyses, leaving AR-induced precipitation in reanalyses largely unexplored.

Besides reanalyses, satellite observations have also been frequently used to characterize ARs and their associated precipitation (Arabzadeh et al., 2020; Behrangi et al., 2016; Cannon et al., 2017, 2020; Guan et al., 2010; Matrosov, 2013; Neiman et al., 2008; Neiman et al., 2008; Ralph et al., 2004; Wick et al., 2013). Integrated water vapor (IWV) from the Special Sensor Microwave Imager (SSM/I) (Hollinger et al., 1990) has been instrumental since early studies of ARs. For example, by compositing the IWV from SSM/I of 312 AR events over the eastern North Pacific, Ralph et al. (2004) established the IWV AR threshold and found that, on average, the IWV magnitude and width of a typical AR is about 2.81 cm and 388 km, respectively. Focusing on landfalling ARs along western North America from 1997 to 2005, Neiman et al. (2008) used the IWV from SSM/I to investigate AR seasonality and landfall orientation. They found that warm season ARs tend to occur in the North while cool season ARs tend to occur in the South. Winter landfalling ARs tend to extend northeastward from the tropical eastern Pacific while summer landfalling ARs tend to be more zonally oriented. Building on these studies, an automated AR detection method using satellite IWV was introduced in Wick et al. (2013). Besides AR characteristics, AR-induced precipitation has also been studied using satellite observations. For example, using radar reflectivity profiles from the Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM-DPR), Cannon et al. (2020) showed that both stratiform and convective precipitation is abundant in ARs, and that AR-induced precipitation is usually triggered by forced ascent in the vicinity of a cold front in frontogenetic environments. While these satellite-based studies have improved our understanding of ARs in observations, most of these studies focused on ARs occurring over the eastern North Pacific, and a global satellite-based AR study is lacking. Secondly, since 3D satellite observed wind field is currently not available, these studies usually detect ARs by adopting a simple IWV threshold of 2 cm and requiring the detected object to be longer than 2000 km and narrower than 1000 km (Neiman et al., 2008; Ralph et al., 2004; Wick et al., 2013). However, ARs are defined as enhanced moisture transport in the atmosphere (AMS Glossary of Meteorology, 2017). Detecting ARs using only the moisture field would inevitably run the risk of detecting filamentary features which resemble ARs, but are associated with weak moisture transport due to stagnant weather conditions. Furthermore, variability in ARs at different time scales can be controlled by the variabilities in both circulation and moisture (Gao et al., 2015; Ma et al., 2021; Ma et al., 2020; Ma & Chen, 2022; Payne et al., 2020; Zhang et al., 2021). For example, at the interannual time scale, it has been shown that AR variability is predominantly controlled by the circulation variability (Ma & Chen, 2022). But at the decadal time scale, either the circulation variability (Ma et al., 2020) or the moisture variability (Zhang et al., 2021) can dominate the AR variability. Using only IWV in the AR detection can generate AR variability which only reflects the variability in the moisture field and thus fails to capture the variability in the circulation field. Therefore, further improvements are needed to incorporate the wind component into the AR detection methods for satellite observations.

Given the limitations in the previous AR studies discussed above, the goals of this study are threefold: i) improve previous AR detection methods for satellite observations by incorporating satellite-based wind information, ii) perform a comprehensive intercomparison of AR representations between seven reanalyses, which are commonly used in AR community, and satellite observation, iii) evaluate AR precipitation in reanalyses against precipitation from satellite observation. The structure of this paper is organized as follows. Section 2 describes the reanalyses and satellite data used, as well as the approach used to detect ARs in satellite data. Main results will be presented in section 3. A brief conclusion and discussion are provided in section 4.

2 Data and Methods

2.1 Satellite Data and Reanalyses

IWV from SSM/I (Hollinger et al., 1990) has been widely used in AR studies. However, the spatial coverage of SSM/I is confined to oceans while observations over land, which are most relevant for AR impacts, are not available. To circumvent this issue, the version 6 (V6) Level 3 (L3) total integrated column water vapor (TotH2OVap or IWV) product from the combined retrievals of the Atmospheric Infrared Sounder (AIRS) and the Advanced Microwave Sounding Unit (AMSU) instruments on NASA's Aqua satellite is used in this study (hereafter AIRS/AMSU) (AIRS Science Team/Joao Teixeira, 2013). The AIRS/AMSU dataset has a global coverage (land + ocean) with spatial resolution of $1^{\circ} \times 1^{\circ}$ and daily temporal resolution. Observations made each day consist of an ascending orbit and a descending orbit (Tian et al., 2013). We take the average of these two orbits to obtain a much smoother field which is representative of the daily mean. Due to the limited swath width of the satellite observations, gaps with no observation exist between swaths. While averaging the ascending orbit and descending orbit to create the daily mean can effectively remove most of these swath gaps, small gaps remain over subtropical regions after this procedure. We fill in these small gaps using the "poisson_grid_fill" function from NCL (The NCAR Command Language., 2019). Note that the results presented in this study would not be affected by whether these small gaps are being filled or not. In order to get the lower tropospheric wind information, geopotential heights at 925, 850, 700, and 600 mb levels are also obtained from the AIRS/AMSU and processed in the same way as the IWV field (see section 2.2 for the use of this variable in our study). Observed precipitation is based on the Integrated Muti-SatellitE Retrievals for GPM (IMERG) Version 6 Final Run (Huffman etal., 2019). This satellite-based precipitation dataset has been widely used in previous studies for mid-latitude weather systems, such as cyclones and ARs (Arabzadeh et al., 2020; Naud et al., 2020). Due to the temporal coverage of the AIRS/AMSU combined retrievals, the study period of this work is from 08/31/2002 to 09/24/2016.

Seven reanalyses, most of which have been widely used in AR studies, are employed here: they are MERRA-2 (Gelaro et al., 2017), ERA5 (Hersbach et al., 2020), ERAI (Dee et al., 2011), JRA-55 (Kobayashi et al., 2015), CFSR (Saha et al., 2010), NCEP/Department of Energy (DOE) Reanalysis II (NCEP R2) (Kanamitsu et al., 2002) and NCEP R1 (Kalnay et al., 1996). The inclusion of ERAI, NCEP R1 and NCEP R2 here allows us to see if there are any improvements in the AR representation from older generations of reanalyses to newer generations of reanalyses. These datasets have varied temporal resolutions. Data at 00 and 12 UTC are averaged to obtain the daily mean. We also tried calculating the daily mean using data at 00, 06, 12, and 18 UTC. The results presented in this study are not sensitive to how the daily mean is calculated (not shown). Both IWV and geopotential height are bilinearly interpolated to a common resolution of $1^{\circ} \times 1^{\circ}$ before analysis. The precipitation fields from reanalyses also have varied temporal resolutions. Some datasets provide the field at the analysis time step while others provide it at forecast time step, although all reanalysis precipitation data used are essentially short-term forecasts (i.e., no assimilation of observed precipitation) and purely generated by the models without any corrections based on observations. Precipitation data are thus processed accordingly to obtain the daily mean. More specifically, daily precipitation from MERRA-2 and ERA5 is calculated by aggregating hourly total precipitation. Daily precipitation in ERAI is calculated by summing the 12-hour forecasted accumulated precipitation initialized at 00 and 12 UTC. Daily precipitation in JRA-55 is calculated from the forecasted precipitation rate initialized at 00, 06, 12, and 18 UTC with steps of 3 and 6 hours. Daily precipitation in CFSR is calculated by aggregating the 6-hour forecasted accumulated precipitation initialized at 00, 06, 12, and 18 UTC. For NCEP R1 and NCEP R2, mean daily precipitation rate is available for direct download. All reanalysis-based precipitation data and IMERG are regridded to a common $1^{\circ} \times 1^{\circ}$ resolution using an areal conservative method ("area_conserve_remap" from NCL) before analysis.

2.2 AR Detection Method

As has been discussed above, AR detections for satellite data usually rely only on the IWV because satellite-based 3D wind fields are not available. To introduce wind information into the AR detection, we derive geostrophic winds from the geopotential height at 925, 850, 700, and 600 mb levels using the "z2geouv" function from NCL. Outside of the deep tropics in ERA5, geostrophic winds at these levels are found to temporally correlate with the actual winds with correlation coefficients close to one (not shown). Geostrophic winds at these four levels are then vertically averaged to obtain the mass-weighted vertical average geostrophic zonal $< U_g >$ and meridional $< V_g >$ winds using:

$$< U_g > = \frac{U_{925} \times 37.5 + U_{850} \times 112.5 + U_{700} \times 125 + U_{600} \times 50}{(37.5 + 112.5 + 125 + 50)}$$

$$< V_g > = \frac{V_{925} \times 37.5 + V_{850} \times 112.5 + V_{700} \times 125 + V_{600} \times 50}{(37.5 + 112.5 + 125 + 50)}$$

where the subscripts in U and V denote pressure levels. The IVT based on the weighted vertical average of geostrophic winds, which we will call the geostrophic IVT (GIVT), is then calculated as follows:

$$GIVT = \sqrt{(IWV \times \langle U_g \rangle)^2 + (IWV \times \langle V_g \rangle)^2}$$

We find that the GIVT can serve as a good proxy for the actual IVT in terms of magnitude. As demonstrated in Figures 1a and 1b, the snapshot of the GIVT shown for ERA5 is nearly identical to the snapshot of actual IVT. Filaments of enhanced IVT in the actual IVT field can also be found in the GIVT field. Minor differences between these two fields only exist in the magnitude: GIVT tends to slightly overestimate the magnitude of the actual IVT, especially over regions with enhanced IVT. Indeed, as shown in Figure 2a, which plots the joint probability distribution function (PDF) of the actual IVT versus GIVT in ERA5 for the year 2003, most of the points fall along the one-to-one line, indicating the good correspondence between the GIVT and the actual IVT. As IVT increases, slightly more points are located above the one-to-one line than those located below it. This corroborates the results in Figures 1a and 1b that GIVT tends to be slightly stronger than the actual IVT over enhanced IVT regions. Such sub-geostrophic wind is expected near a low pressure center due to gradient wind balance. As shown later, such a slight overestimate of the IVT magnitude by the GIVT has negligible effect on the ARs detected due to the percentile-based threshold adopted by the AR detection tool (ARDT) used in this study.

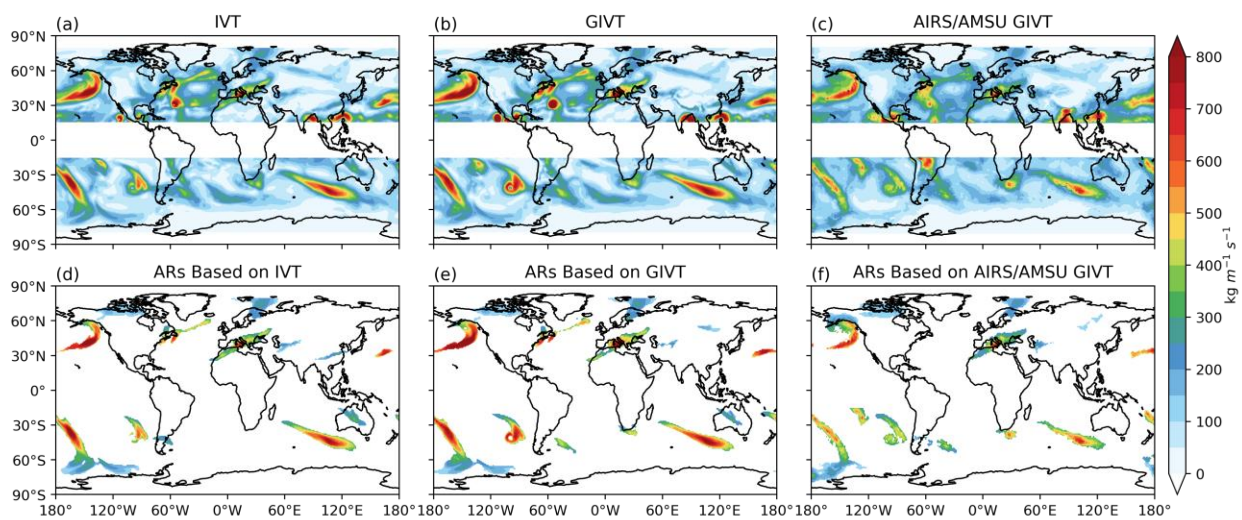


Figure 1. A snapshot of the IVT (a) and GIVT (b) in ERA5 on 10/05/2003. (c) is showing the same snapshot of GIVT, but from AIRS/AMSU. Corresponding ARs detected by the modified algorithm are shown in (d), (e), and (f), respectively.

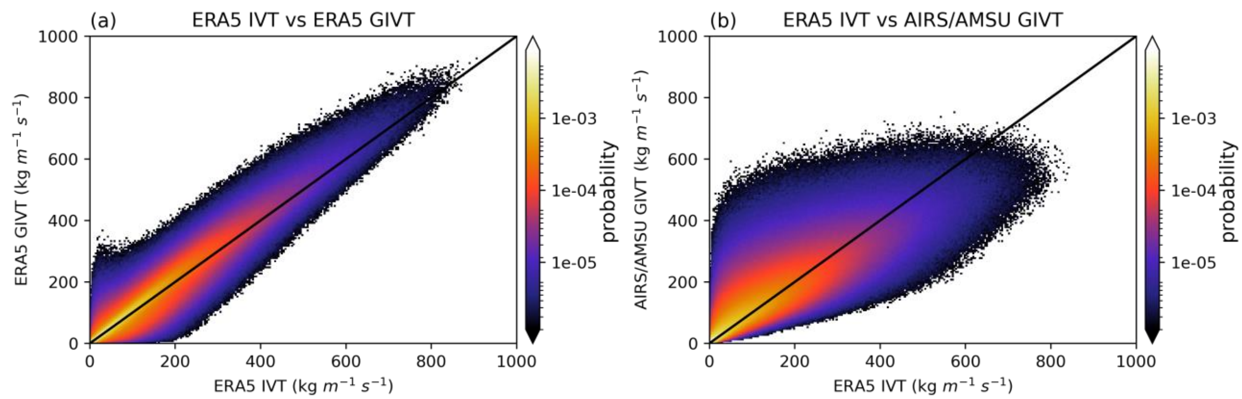


Figure 2. Joint probability distribution function of IVT versus GIVT in ERA5 (a) and IVT from ERA5 versus smoothed GIVT from AIRS/AMSU (b). Note that the color bars are in logarithmic scale and only data from the year 2003 are used.

The ARDT used in this study is based on the IVT-based Guan & Waliser (2015) algorithm, with modifications so that GIVT can be directly used as input. This ARDT is a global algorithm which has been widely used in AR studies (e.g., Arabzadeh et al., 2020; Espinoza et al., 2018; Ionita et al., 2020; Kim et al., 2021; Nash et al., 2018). Notable criteria adopted by this ARDT are listed here. Readers are referred to Guan & Waliser (2015) for a detailed description of the algorithm. In the first step of the detection, a seasonally and regionally dependent 85th percentile of the IVT magnitude, or 100 kg m^{-1} s^{-1}, whichever is greater, is used as the intensity threshold. The detected “objects” are then further filtered by three IVT direction criteria. The detected “object” will be filtered out if 1) more than half of the grid cells have IVT direction deviating more than 45° from the object’s mean IVT direction (the coherence criterion), 2) mean poleward meridional IVT is less than 50 kg m^{-1} s^{-1} (the meridional IVT criterion) and 3) direction of object-mean IVT deviates from the overall orientation of the object’s shape by more than 45° (the consistency criterion). Objects which pass these three IVT direction criteria are subjected to further geometrical screening, such that the final detected ARs are longer than 2000 km in length and have a length/width ratio greater than two.

To test the sensitivity of the algorithm to the input variables, in the first step, we used both the IVT and GIVT from ERA5 as input to the original Guan & Waliser (2015) ARDT. The GIVT-based AR frequency, which is defined as the fraction of time a grid cell experiences AR conditions, is very similar to that based on the IVT (Figure S1). Enhanced AR frequency is found over the mid-latitude storm track regions (Figure S1a). However, compared to the AR frequency based on the IVT, results based on the GIVT underestimate AR frequency over mid-latitude regions and overestimate it over the subtropics (Figure S1b). We then removed the three IVT direction criteria one at a time, and tested the sensitivity of the modified algorithm to the input variables. We found that removing either the coherence criterion or the consistency criterion has very little effects on the AR frequency (Figures S1c and S1e), consistent with the degree of filtering associated with the two criteria reported in Guan and Waliser (2015). The differences between the IVT-based AR frequency and the GIVT-based AR frequency persist (Figures S1d and S1f). However, once we removed the meridional IVT criterion, the differences between the IVT-based AR frequency and the GIVT-based AR frequency mostly vanish (Figure S1h). Meanwhile, the magnitude of the AR frequency also increases nearly everywhere (Figure S1g). These results imply that the differences in the magnitude of the meridional IVT between the IVT and GIVT are likely nonnegligible. This is likely due to boundary layer friction that causes the direction of actual winds to deviate from geostrophic winds. Given the results found in the sensitivity experiments, we removed the three IVT direction criteria of the original algorithm in our modified algorithm. After these three IVT direction criteria are removed, AR frequency increases nearly everywhere and its distribution becomes more uniform (Figure S1i).

Besides removing these three criteria, two additional minor modifications are also made to the algorithm. We found that the modified algorithm tends to detect too many ARs over the Northern Hemisphere continents during boreal summer. To partially alleviate this problem, instead of calculating the IVT threshold for a particular month using all the time steps from the five months centered on that month over the study period, we only use the time steps from that month in the modified algorithm. Furthermore, previous studies suggested that the Guan & Waliser (2015) algorithm may occasionally pick up tropical disturbances as ARs (Guan et al., 2018; Lora et al., 2020). To remedy this problem, we impose that, if the detected object has all its area located within 30°N/S, it will be filtered out. This criterion mostly affects ARs within 30°N/S and it reduces the magnitude of AR frequency over these regions. We want to emphasize that the conclusions presented in this study are not sensitive to whether these two additional modifications are adopted or not (not shown). In summary, our modified algorithm is based on the Guan & Waliser (2015) ARDT and detects ARs with an enhanced IVT/GIVT relative to its background state, a length greater than 2000 km, and a length/width ratio greater than two (see table S1 for the summary of the differences between the original Guan & Waliser (2015) algorithm and the modified algorithm used in this study).

Due partly to the low sampling frequency in satellite observations, the geostrophic winds derived from the geopotential height tend to be noisier compared to the smoother fields in reanalyses (Fetzer et al., 2006; Hearty et al., 2014; North et al., 1993; Tian et al., 2013; Tian & Thomas, 2020). Since geostrophic winds are derived based on the gradient of the geopotential height, these noises in the geopotential height result in the derived geostrophic winds being too strong compared to the geostrophic winds in reanalyses. We applied a simple bias correction to the satellite weighted vertical average geostrophic wind speed so that the satellite mean wind speed over midlatitudes is equal to that in ERA5 (see text S1 in the supplementary for details on how the bias correction is carried out). We found that these noises in the geostrophic wind field in satellite data can also increase the "false negative" rate for AR detection. In other words, features which are detected as ARs in reanalyses are occasionally not picked up as ARs in satellite data. After manually examining those ARs which are detected in reanalyses, but not in satellite data, we found that most of those ARs are either broken into more than one smaller object or with stronger GIVT magnitude concentrated within a smaller single object in satellite data. Both situations can result in the AR failing to meet the geometric criteria of AR. To resolve this issue, Gaussian smoothing was applied to the bias-corrected weighted vertical average geostrophic wind speed of the satellite data. (Note that

Gaussian smoothing is not applied to the reanalyses because the data from reanalyses is already quite smooth.) We adjusted the size of the smoothing kernel by varying the sigma parameter ranging from 1 to 6 in an increment of 0.5. We found that, when the sigma is too small, the resulting field is not smooth enough. This results in the "false negative" rate remaining high. However, when we set the sigma too large, the field is smoothed out too much. This effect can cause some filaments of enhanced IVT to be smoothed out and thus potentially get filtered out during the detection process. We set the sigma to be three in this study as a balance between smoothing the IVT enough yet not inadvertently erasing any filamentary structure. This smoothing on the AIRS/AMSU geostrophic wind field tends to slightly enhance the weak GIVT values while weakening the strong GIVT values (Figures 1c and 2b). Since the ARDT used in this study uses a percentile-based threshold, the smoothing is expected to have very minor effects on the ARs detected (comparing Figure 1f with 1d and 1e). As a result, the AR frequency difference between ERA5 and AIRS/AMSU is small when sigma is set to three. These two additional operations on the satellite data are based on the assumption that the ERA5 winds are better representative of the true observation, as the satellite-based geostrophic winds are not directly measured or dynamically constrained. Whether such an assumption is valid or not is not a concern in this study, since both satellite data and reanalyses have their own biases in representing AR winds, and the purpose of this study is not to treat satellite data as the true observation and evaluate the biases in reanalyses relative to satellite observations. Instead, our goal is to demonstrate the feasibility of including wind information for AR detection based on satellite data and also comprehensively investigate the spread among reanalyses. Therefore, the assumption we made is justifiable for the purpose of this study.

3 Results

3.1 AR Frequency and Characteristics

Figure 3a shows the annual AR frequency distribution in AIRS/AMSU. Enhanced AR frequency is observed over the mid-latitude oceans. Unlike the AR frequency distribution obtained from other global ARDTs that participated in the ARTMIP (Figure S2), the AR frequency distribution in Figure 3a is spatially more uniform, and substantially more ARs are detected over land. It has been shown by previous studies that these features in the AR frequency distribution are unique to algorithms adopting relative AR thresholds (Rutz et al., 2019; Shields et al., 2018). Since our algorithm is modified from the Guan & Waliser (2015) ARDT, it's thus expected that the AR frequency distribution based on our algorithm shares many similarities to that based on the Guan & Waliser (2015) algorithm (Figure S2). The spatial patterns of the AR frequency difference between reanalyses and AIRS/AMSU are very similar across reanalyses: the differences over the storm track regions, where ARs are most active, are generally small. Reanalyses tend to have fewer ARs over the subtropical regions while having more ARs over the higher latitudes poleward of 60°N/S, especially near the coast of Antarctica. As shown in Figure S3, this pattern in the AR frequency difference does not depend on the magnitude of the sigma used in the smoothing function for AIRS/AMSU. Reanalyses also tend to have slightly more ARs over the high latitude land regions over the Northern Hemisphere. Compared to AIRS/AMSU, all reanalyses show more ARs around the date line. This is caused by the temporal discontinuity at the date line in the AIRS/AMSU daily data files which can occasionally prevent the detection of ARs over this region. In each AIRS/AMSU daily L3 product file, observations start at the date line and progress westward. This results in the data immediately west of the date line being farthest apart in time (~24 hours) from those immediately east of the dateline, leading to the temporal discontinuity at the date line in AIRS/AMSU daily data files. For the zonal mean figures shown below, regions around the date line ($\pm 10^{\circ}$) are excluded from the calculations. In addition, pressure levels at 925 and 850 mb are used in the derivation of the geostrophic winds. This can result in unrealistic AR statistics over topographies. Therefore, regions with climatological surface pressure less than 850 mb (based on ERA5) are also excluded in the calculations of zonal mean. After these bias corrections, we see that the spreads among the reanalyses are small, with NCEP R1 and NCEP R2 having only slightly higher AR frequency, suggesting that all reanalyses have similar performance in representing the AR frequency distribution (Figures 3c, 3d, and 3e).

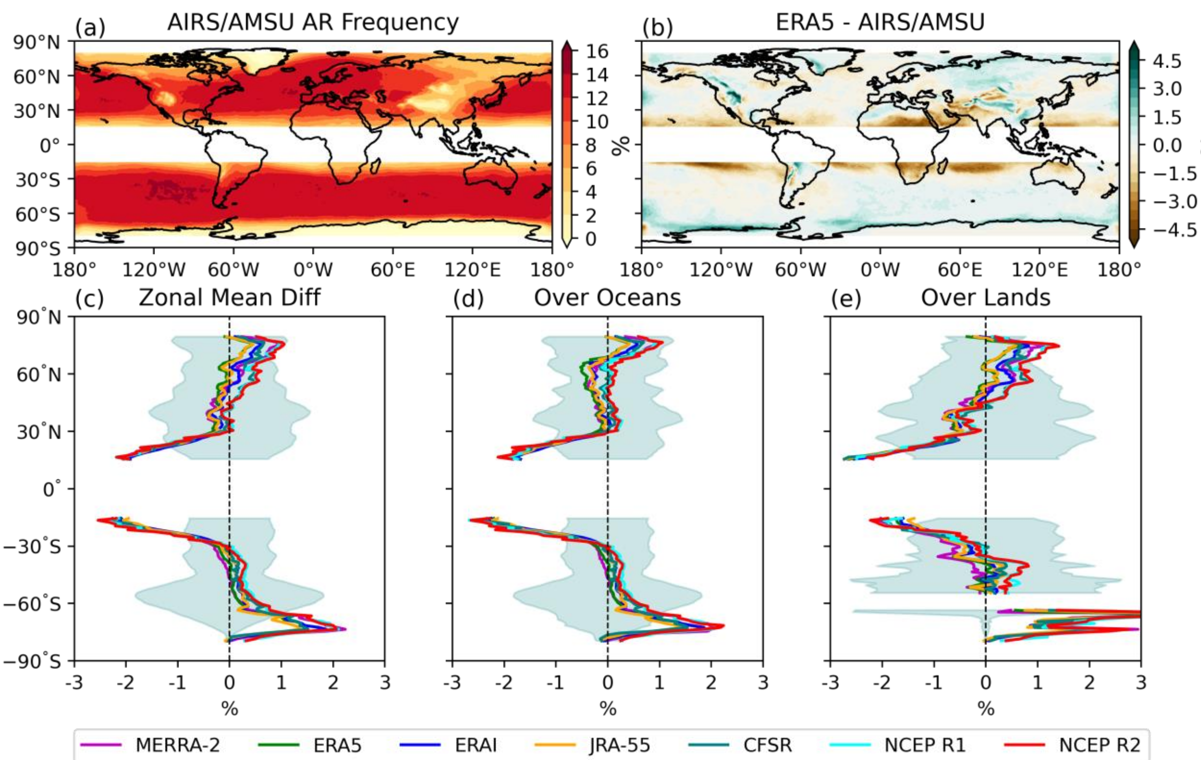


Figure 3. AR frequency detected by the modified algorithm based on AIRS/AMSU (a). AR frequency difference between ERA5 and AIRS/AMSU (b). The differences between the zonal mean AR frequency in reanalysis products and AIRS/AMSU. (d) and (e) are the same as (c), but

for the zonal mean differences over oceans and land, respectively. The shading in (c), (d) and (e) represents one standard deviation of the annual zonal mean AR frequency in AIRS/AMSU.

AR frequency is controlled by both the size and the number of detected ARs. In AIRS/AMSU, both the size and the number of detected ARs are sensitive to the value of sigma chosen for the gaussian smoothing. As sigma increases, both the AR length and width increase while the number of detected ARs drops (not shown; consistent with Figure 5k,l of Guan & Waliser (2015)). We set the sigma value to three, which minimizes the AR frequency difference between ERA5 and AIRS/AMSU over midlatitudes (Figure S3). Figure 4a shows that the PDFs of the AR length are consistent across different datasets. Compared to the AR length in reanalyses, ARs in AIRS/AMSU are slightly longer. The spread measured by the standard deviation in the AR length across reanalyses is relatively small and only about 0.5% of the climatology in AIRS/AMSU. However, there is a larger spread in the AR width distribution with the spread across reanalyses reaching about 4% of the climatology (Figure 4b). In particular, the ARs in MERRA-2 are the narrowest. Consistent with the narrowest ARs in MERRA-2, ARs in MERRA-2 also have the smallest area (Figure 4c) and largest length/width ratio (Figure 4d). The total number of ARs in MERRA-2 during the study period is more than other datasets (Figure 4a), consistent with the ARTMIP analysis for MERRA-2 (Collow et al., 2022). This suggests that MERRA-2 tends to simulate more ARs, but with smaller AR size. These two effects in MERRA-2 compensate for each other and result in the AR frequency being comparable to those in other reanalyses. NCEP R1, NCEP R2, and to a lesser extent, CFSR have larger AR size due to the larger AR width in these datasets, but fewer ARs were detected in them. Consequently, the length/width ratio in these three datasets are smallest and AR frequencies in NCEP R1 and NCEP R2 are slightly higher compared to other reanalyses (Figures 3c, 3d and 3e). It is also worth pointing out that NCEP R1, NCEP R2 and CFSR start out with a coarser resolution of 2.5°×2.5°. Even though we have regridded them to a common resolution of 1°×1°, such a coarser native resolution can still be expected to have some impacts on the geometry of the ARs (e.g., Guan & Waliser. (2017)). The wider ARs found in these three datasets thus could be partially caused by their coarser native resolution.

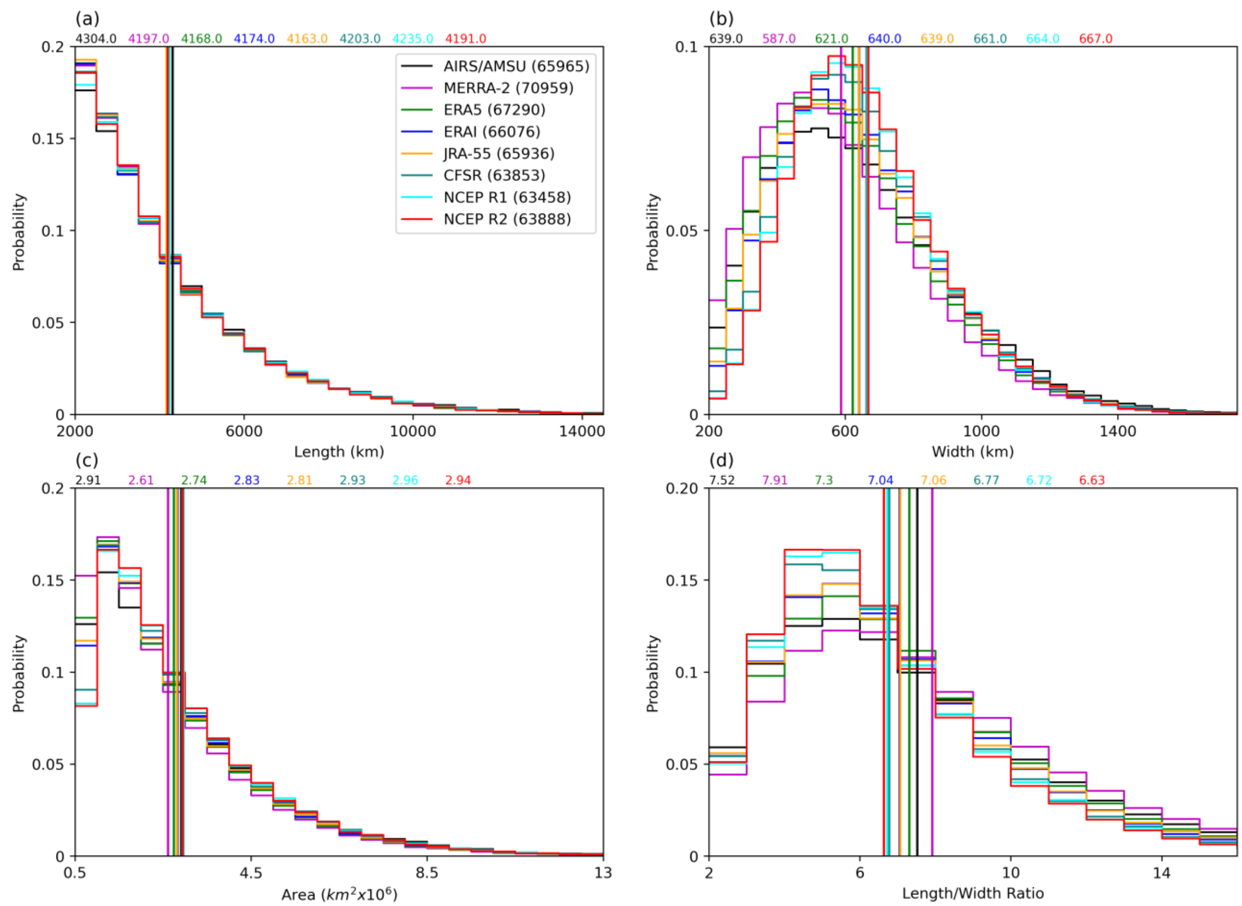


Figure 4. Probability distribution functions of AR length (a), width (b), area (c), and length/width ratio (d) for all reanalyses and AIRS/AMSU. The numbers inside the parentheses in (a) indicate the total number of ARs detected during the study period.

3.2 AR Strength

AR GIVT is defined as the GIVT under AR conditions. As shown in Figure 5a, enhanced AR GIVT occurs over the storm track regions. Unlike AR frequency, the spatial distribution of AR GIVT is less uniform and exhibits more spatial variation, with enhanced AR GIVT concentrated in smaller regions of the storm tracks. Compared to the magnitude of the AR GIVT over oceans, AR GIVT over land is substantially weaker. The inter-product differences in AR GIVT between reanalyses and AIRS/AMSU have considerably similar spatial characteristics (Figures 5b, 5c, 5d, and 5e). Reanalyses show stronger AR GIVT over the midlatitudes, especially the storm track regions. Weaker GIVT can be found over subtropical regions. Since the geostrophic wind in AIRS/AMSU is bias-corrected based on ERA5 (see the Method section or text S1 in the supplementary), the differences in AR GIVT between reanalyses and AIRS/AMSU should not be treated as biases from reanalysis winds. Over the Northern Hemisphere, MERRA-2, NCEP R1, and NCEP R2 tend to have stronger AR GIVT. Over the Southern

Ocean, AR GIVT in NCEP R1 and NCEP R2 are substantially stronger than the AR GIVT in AIRS/AMSU (up to about 27% stronger) and those in other reanalyses. The spread among these five reanalyses is relatively small, though the AR GIVT in MERRA-2 is slightly stronger, consistent with the stronger IVT magnitude in MERRA-2 found in Collow et al. (2022). It is not surprising that we also find that the results of the AR GIVT are consistent with those of the climatological mean GIVT (Figure S3), suggesting that differences in AR GIVT among datasets are mostly due to the differences in the climatological mean GIVT.

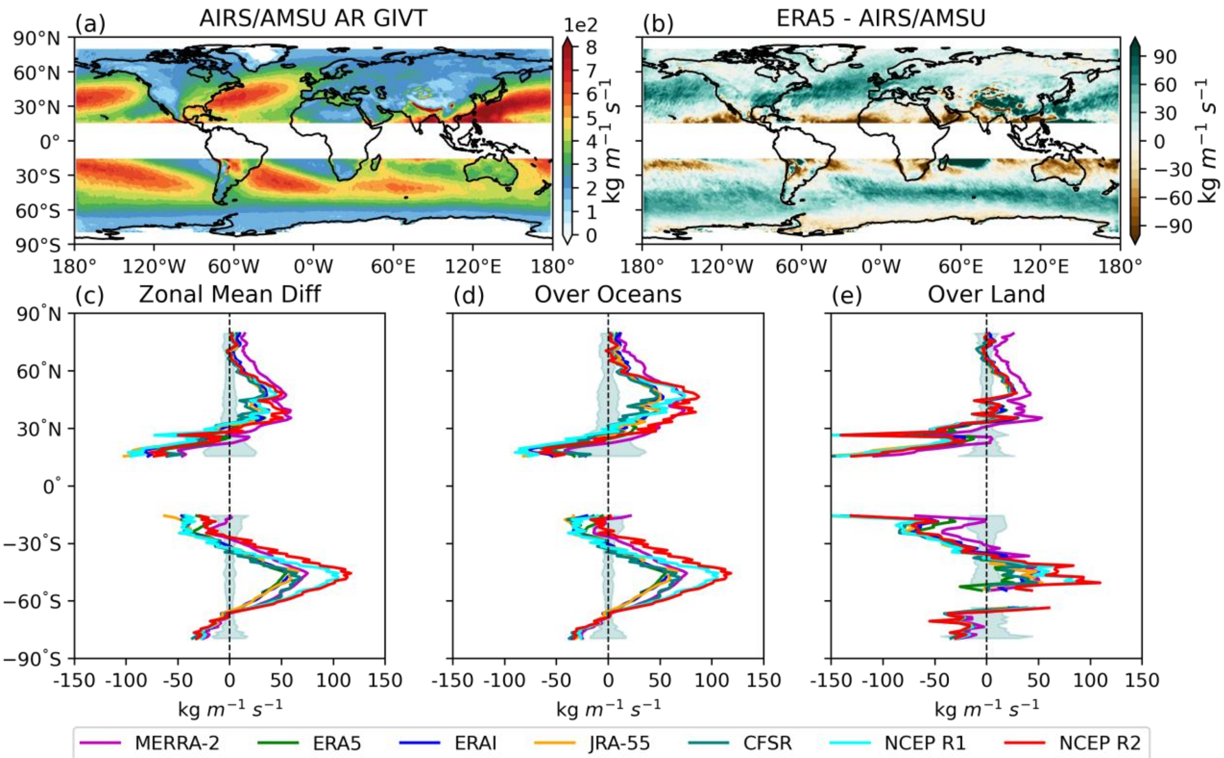


Figure 5. AR GIVT in AIRS/AMSU (a). AR GIVT difference between ERA5 and AIRS/AMSU (b). The differences between zonal mean AR GIVT in reanalyses and AIRS/AMSU (c). (d) and (e) are the same as (c), but for the zonal mean AR IVT over oceans and land, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean AR GIVT in AIRS/AMSU.

The inter-product differences in the AR GIVT can be caused by the differences in AR IWV and AR geostrophic wind magnitude. Figure 6a shows the AR IWV in AIRS/AMSU. Enhanced AR IWV mostly occurs over the subtropical regions equatorward of 30°N/S. Poleward of 30°N/S, AR IWV decreases rapidly. Compared to the AR IWV over oceans, AR IWV over land is substantially weaker. Compared to the

AR IWV in AIRS/AMSU, reanalyses simulate stronger AR IWV over midlatitudes poleward of 30˚N/S and weaker AR IWV can be found equatorward of 30˚N/S. This spatial pattern in the difference is shared by all reanalyses (Figures 6c, 6d and 6e). Over the Northern Hemisphere, the spread among all reanalyses is relatively small, with the AR IWV in NCEP R1 and NCEP R2 being slightly stronger than those in other reanalyses. Over the Southern Ocean, the AR IWV in NCEP R1 and NCEP R2 is substantially stronger than those in the other five reanalyses, which is consistent with the results in AR GIVT. Similar to GIVT, the results of the AR IWV are also reflective of the results in the climatological mean IWV (Figure S4). Since the AIRS/AMSU IWV data are subject to sampling biases and there is no bias correction applied to the AIRS/AMSU IWV field (Hearty et al., 2014; Tian et al., 2013; Tian & Thomas, 2020), the differences between AIRS/AMSU and other reanalyses should not be viewed as the biases in reanalyses. Instead, the spread among datasets should be simply viewed as observed uncertainties.

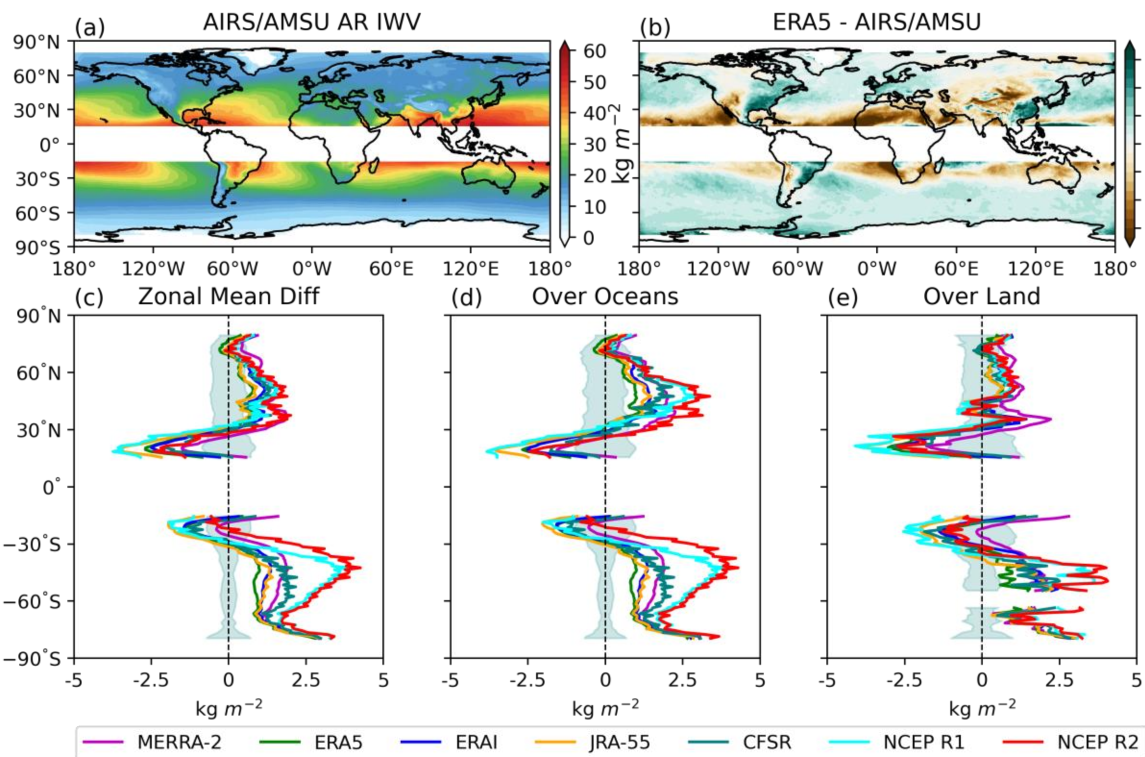


Figure 6. AR IWV in AIRS/AMSU (a). AR IWV difference between ERA5 and AIRS/AMSU (b). The differences between zonal mean AR IWV in reanalyses and AIRS/AMSU (c). (d) and (e) are the same as (c), but for the zonal mean AR IWV over oceans and lands, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean AR IWV in AIRS/AMSU.

Figure 7a shows the AR geostrophic wind magnitude. Unlike the pattern in the AR IWV, enhanced AR wind is found over the regions poleward of 30°N/S. This suggests that the enhanced AR GIVT shown in Figure 5a is dominated by IWV over the subtropics, but by wind over midlatitudes. Consistent with the climatological wind speed (Fig. S5), AR wind over the Southern Hemisphere is stronger compared to that over the Northern Hemisphere. Enhanced wind can also be found along the coastal regions of Antarctica. The enhanced wind over these regions is likely unrealistic, which may be caused by the presence of topography. Compared to the AR wind in AIRS/AMSU, reanalyses overestimate the wind magnitude over the midlatitude Southern Ocean while substantially underestimating it along the coastal regions of Antarctica where sea ice is present. Such a large difference between reanalyses and AIRS/AMSU over these regions usually covered by sea ice likely indicates that the wind over these regions in AIRS/AMSU may be biased high (Yue & Lambrigtsen, 2017, 2020). Weaker wind in reanalyses can be found over subtropical regions and regions at around 60°N. Unlike the AR IWV field, the spread in the AR wind among reanalyses is small (Figures 7c, 7d and 7e), indicating higher skills for reanalyses in simulating the wind field. Note that the AR wind in the AIRS/AMSU shown here has been bias-corrected by the wind filed in ERA5 based on the climatological wind speed over midlatitudes. The differences between reanalyses and AIRS/AMSU thus should not be treated as biases in reanalyses.

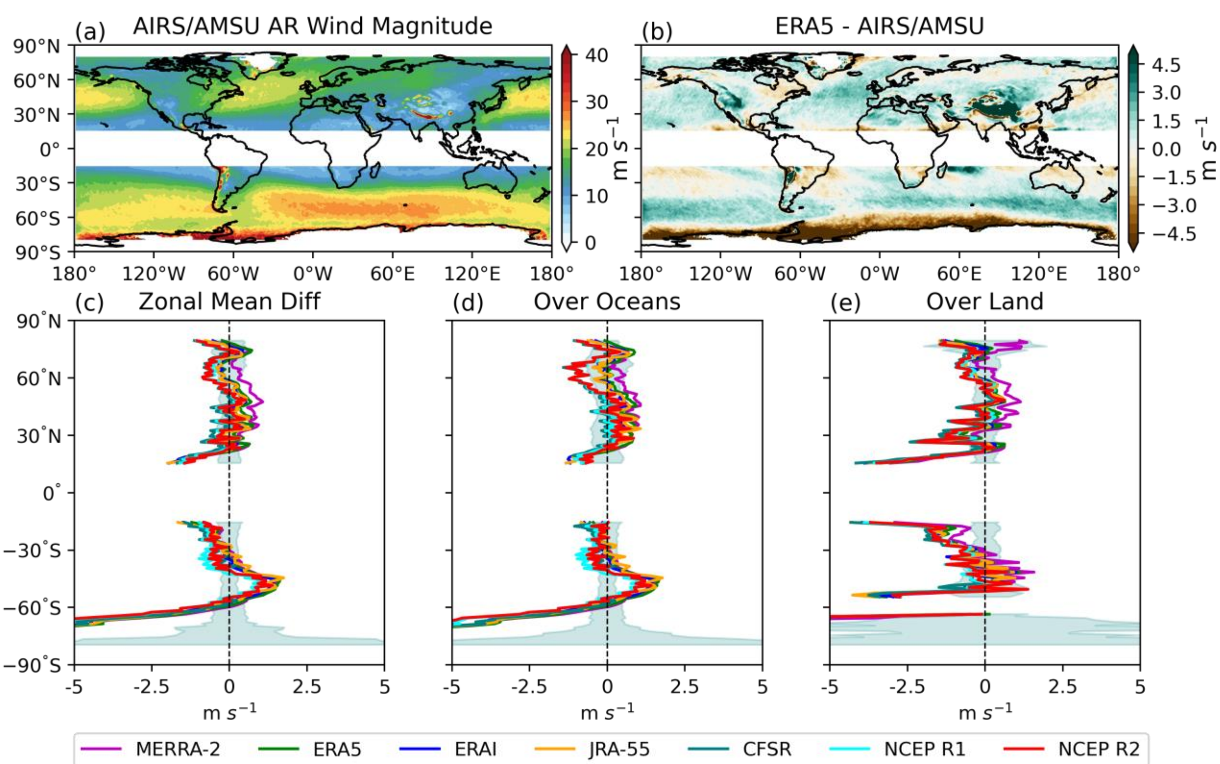


Figure 7. AR geostrophic wind magnitude in AIRS/AMSU (a). AR geostrophic wind magnitude difference between ERA5 and AIRS/AMSU (b). The differences between zonal mean AR geostrophic wind magnitude in reanalyses and AIRS/AMSU (c). (d) and (e) are the same as (c), but for the zonal mean AR geostrophic wind magnitude over oceans and land, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean AR wind magnitude in AIRS/AMSU. The geostrophic winds shown here are the mass-weighted vertical average geostrophic winds.

3.3 AR Precipitation

It has been well documented that ARs are associated with enhanced precipitation (e.g., Arabzadeh et al., 2020; Gao et al., 2016; Lavers & Villarini, 2013). It's also quite common that reanalysis-based precipitation is directly used in AR studies (Collow et al., 2020; Gao et al., 2016; Kim et al., 2022; Maclennan et al., 2022; Pasquier et al., 2019; Zhou et al., 2022). Yet, reanalysis-based precipitation is not directly constrained by observations. Their performance against observed precipitation thus requires further evaluation. In this section, we will focus on AR precipitation which is defined as the precipitation that falls within the AR boundaries. Figure 8a shows the mean AR precipitation intensity. Since observations of precipitation in IMERG are scarce poleward of 60°N/S, the analyses of AR-related precipitation will be restricted to regions within 60°N/S. Over the Northern Hemisphere, enhanced AR precipitation occurs over the poleward flank of the AR active regions over both the North Pacific and North Atlantic. Intense AR precipitation can be observed extending from the southwest of the ocean basin into the northeast of the ocean basin. We note a discontinuity between the western North Pacific and the eastern North Pacific. The intensity west of the date line is much weaker than that over east of the dateline. The exact cause of this discontinuity is unclear. However, after manual examinations of the identified ARs and the precipitation field, there seems to be a time lag between the AR footprints detected from AIRS/AMSU and the precipitation systems in IMERG over the western North Pacific: the precipitation systems tend to locate east/northeast (ahead) of the AR footprints. This spatial mismatch between the AR footprints and the precipitation systems likely contributes to the abnormally weak AR precipitation over this region. Therefore, we will exclude this region from the following analyses and discussion, as well as in the zonal mean calculations followed.

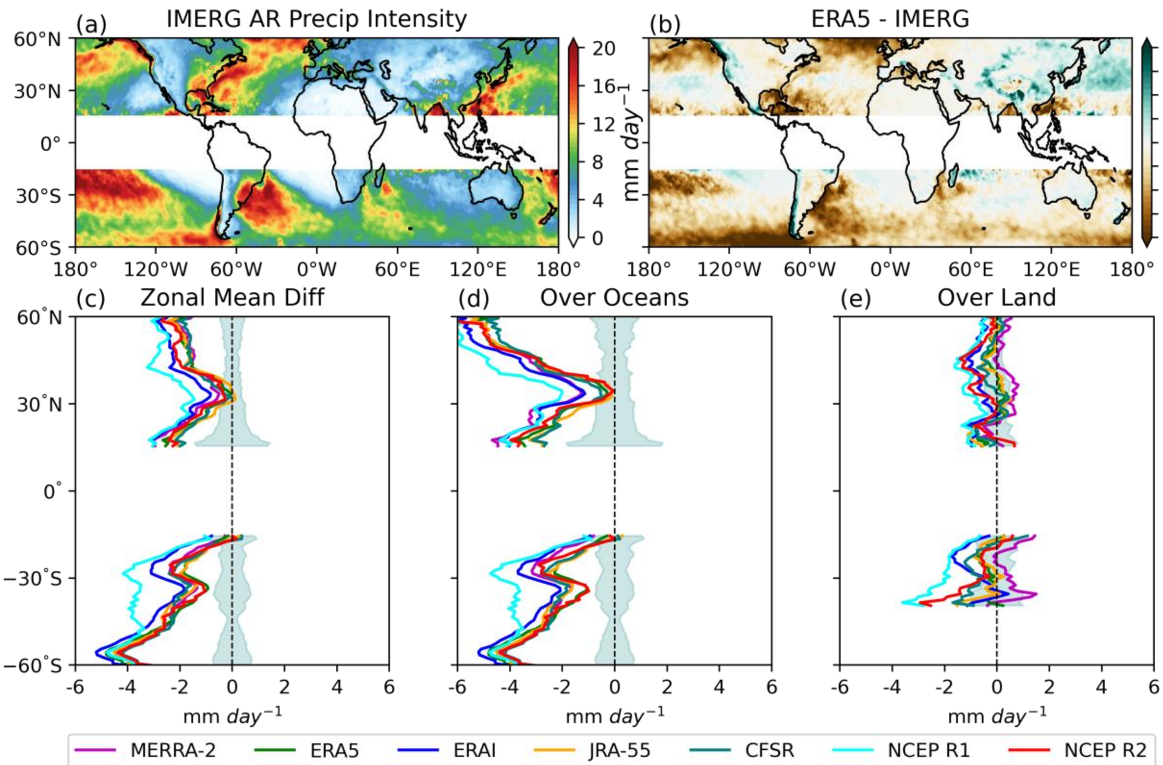


Figure 8. AR precipitation intensity in IMERG (a). AR precipitation intensity difference between ERA5 and IMERG (b). The differences between zonal mean AR precipitation intensity in reanalyses and IMERG (c). (d) and (e) are the same as (c), but for the zonal mean AR precipitation intensity over oceans and land, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean AR precipitation intensity in IMERG.

Over the Southern Hemisphere, consistent with Collow et al. (2022), strong AR precipitation occurs over the South Pacific and South Atlantic while AR precipitation over the South Indian Ocean is relatively weak. Compared to the AR precipitation over oceans, AR precipitation over land is much weaker, except over some coastal regions (Figures 8a). Reanalyses underestimate the AR precipitation over regions with strong AR precipitation intensity, such as the northeastern North Pacific, the North Atlantic poleward of $50^{\circ}N$, western subtropics over the South Pacific and the Southern Ocean Poleward of around 40˚S. The differences between reanalyses and satellite observation are relatively small over regions between 30˚ and 40˚ latitude in both hemispheres (Figure 8c). The biases over land are smaller compared to those over oceans, likely due to the smaller AR precipitation intensity over land (Figure 8e). Due to their small areas, land regions poleward of 40˚S are excluded from the zonal mean calculation over land. Compared to other reanalyses, AR precipitation intensity is weakest in NCEP R1 while the spread among the other six reanalyses is generally small.

As shown in Figure 9a, ARs contribute substantially to the total annual precipitation. Over many of the oceanic and coastal regions, AR precipitation can account for up to half of the total annual precipitation. The contribution of ARs to total precipitation over land is spatially more heterogeneous. For example, ARs can contribute up to half of the annual precipitation over Australia and North Africa. At the same time, East Asia only receives 10-30% of its annual precipitation from ARs. Compared to observation, reanalyses generally underestimate ARs' contribution to the total precipitation over oceans, especially over regions equatorward of 30˚N/S and poleward of 50˚N/S. Smaller differences between reanalyses and satellite observation can be found over regions between 30˚N/S and 40˚N/S. Over land, ARs in reanalyses can contribute more to the total precipitation, such as over Australia and East Asia. At the same time, they can also underestimate ARs' contribution to the total precipitation over regions such as northern North America and South Africa. The spread across reanalyses is generally small. However, over the Northern Hemisphere and equatorward of about 30˚S over the Southern Hemisphere, NCEP R1 simulates the lowest contribution to total precipitation by ARs while CFSR produces the lowest contribution to total precipitation by ARs over oceans poleward of 30˚S over the Southern Hemisphere.

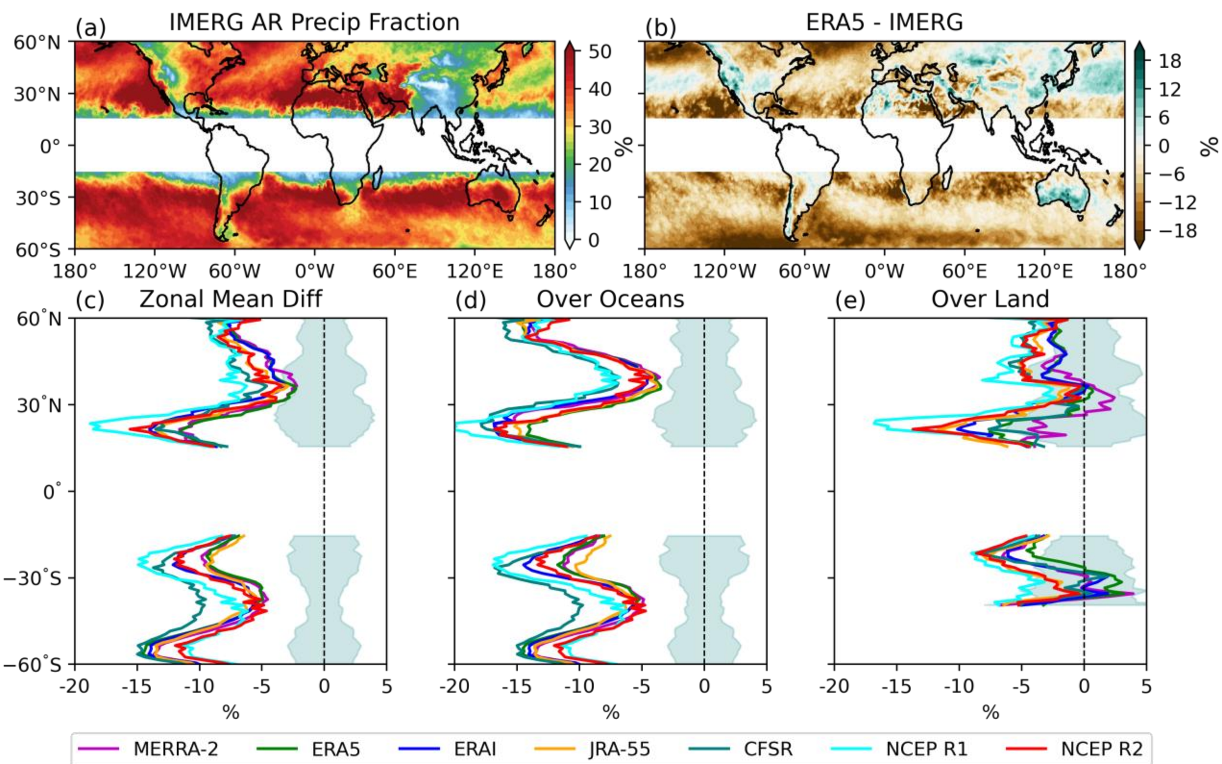


Figure 9. AR precipitation fraction in IMERG (a). AR precipitation fraction difference between ERA5 and IMERG (b). The differences between zonal mean AR precipitation fraction in reanalyses and IMERG (c). (d) and (e) are the same as (c), but for the zonal mean AR precipitation fraction over oceans and land, respectively. The shading in (c), (d),

and (e) represents one standard deviation of the annual zonal mean AR precipitation fraction in IMERG.

It has been shown by previous studies that ARs play an even more important role in extreme precipitation compared to mean precipitation (Arabzadeh et al., 2020; Gao et al., 2016; Guan et al., 2023; Waliser & Guan, 2017). Defined as the 95th percentile intensity of all the AR precipitation (including non-precipitating days), AR extreme precipitation intensifies substantially (Figure 10a) compared to AR mean precipitation (Figure 8a). Over the western South Pacific subtropics, northwest of the South Atlantic and southwest of the North Atlantic, the intensity of AR extreme precipitation can exceed 60 mm/day. Over the Southern Ocean, the spatial pattern of the AR extreme precipitation intensity is very similar to the AR mean precipitation intensity, with very strong intensity observed over the South Pacific and South Atlantic. The intensity over the South Indian Ocean is substantially weaker. Over the Northern Hemisphere oceans, enhanced AR extreme precipitation can be observed over the entire ocean basins of the North Pacific and North Atlantic, except over the southeast ocean basins. Enhanced AR extreme precipitation can also be found over the west coasts of North America and Chile due to orographic lifting, as well as in eastern North America and eastern South America. The spatial patterns of the difference between reanalyses and satellite observation (Figure 10b) are similar to those for the differences in AR mean precipitation between reanalyses and satellite observation (Figure 8b). The AR extreme precipitation in reanalyses is substantially weaker than the satellite observation nearly everywhere (Figure 10b). The largest underestimate occurs over the subtropical oceans and regions poleward of 50°N/S. The differences over land are smaller compared to the differences over oceans due likely to the weaker AR extreme precipitation over land. Over land, the magnitude in MERRA-2 is most comparable to the observation (Figure 10e). The magnitude of the AR extreme precipitation is weakest in NCEP R1, followed by ERAI (Figure 10c). The spread of the other five reanalyses is generally small (Figure 10c). There is an improvement in simulating AR precipitation intensity from NCEP R1 to NCEP R2 and CFSR. Similar improvement can also be found from ERAI to ERA5.

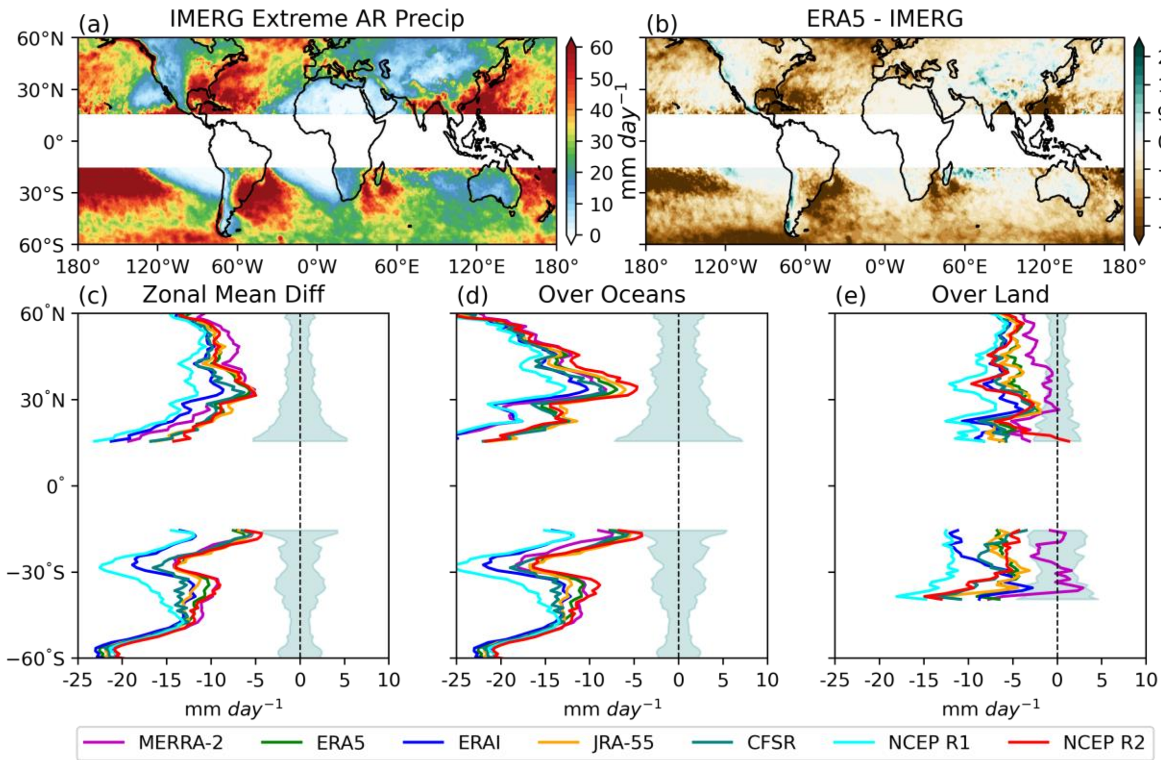


Figure 10. Extreme AR precipitation intensity in IMERG (a). Extreme AR precipitation intensity difference between ERA5 and IMERG (b). The differences between zonal mean extreme AR precipitation intensity in reanalyses and IMERG (c). (d) and (e) are the same as (c), but for the zonal mean extreme AR precipitation intensity over oceans and land, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean extreme AR precipitation intensity in IMERG.

Next, we examine the AR fractional contribution to total extreme precipitation amount. For a grid point, we define an extreme precipitation threshold as the 95th percentile of all precipitation with intensity greater than 0.01 mm/day (including both AR days and non-AR days). This threshold is calculated separately for each dataset. The total extreme precipitation for a grid point is then calculated by summing all the daily precipitation with intensity greater than or equal to this extreme precipitation threshold. As shown in Figure 11a, the spatial pattern of the fractional contribution to extreme precipitation by ARs is very similar to the spatial pattern in the fractional contribution to mean precipitation by ARs (Figure 9a). However, ARs contribute more to the total extreme precipitation. Compared to the 30-50% in the fractional contribution to mean precipitation, ARs account for 50-70% of the total extreme precipitation over most of the oceanic regions. Unlike the fractional contribution to mean precipitation which shows the largest fraction at around 30°N/S, the spatial pattern in the fractional contribution to extreme precipitation is spatially more uniform over oceans. Over land, large fractional contribution can be found over the west coast of North America, Chile, South Africa, eastern North America, eastern South America, Australia, and interestingly North Africa.

In contrast to the differences in the fractional contribution to mean precipitation which shows strong underestimates by reanalyses nearly everywhere, the differences in fractional contribution to extreme precipitation are smaller. Underestimates by reanalyses can be found over regions poleward of 50°N/S and some regions over the subtropics. Over midlatitudes and many land regions, ARs in reanalyses tend to contribute slightly more to the total extreme precipitation. Compared to other reanalyses, NCEP R1 and R2 tend to simulate a slightly smaller fraction of extreme precipitation contributed by ARs, especially over oceanic regions equatorward of about 40°N/S.

Despite using different definition of AR days and/or different precipitation dataset, the satellite-based spatial patterns of the AR precipitation intensity (Figure 8a), fraction (Figure 9a), extreme AR precipitation intensity (Figure 10a) and fraction (Figure 11a) show marked similarity to those based on the original Guan & Waliser (2015) algorithm using reanalysis data (Arabzadeh et al., 2020; Zhao, 2022). These results indicate that our algorithm can correctly capture those precipitating systems with filaments of enhanced moisture transport. Compared to the original Guan & Waliser (2015) algorithm, the modified algorithm used here produces slightly higher AR frequency and with a more uniform spatial pattern (compared Figure 3a to Figure S2d). The consistency in the spatial patterns of the AR precipitation characteristics between our algorithm and the original Guan & Waliser (2015) algorithm suggests that both algorithms likely have high agreement in detecting those relatively strong ARs with intense precipitation. The disagreement in the AR frequency distribution is thus likely caused by the modified algorithm used here being able to detect more relatively weak ARs due to its less stringent criteria.

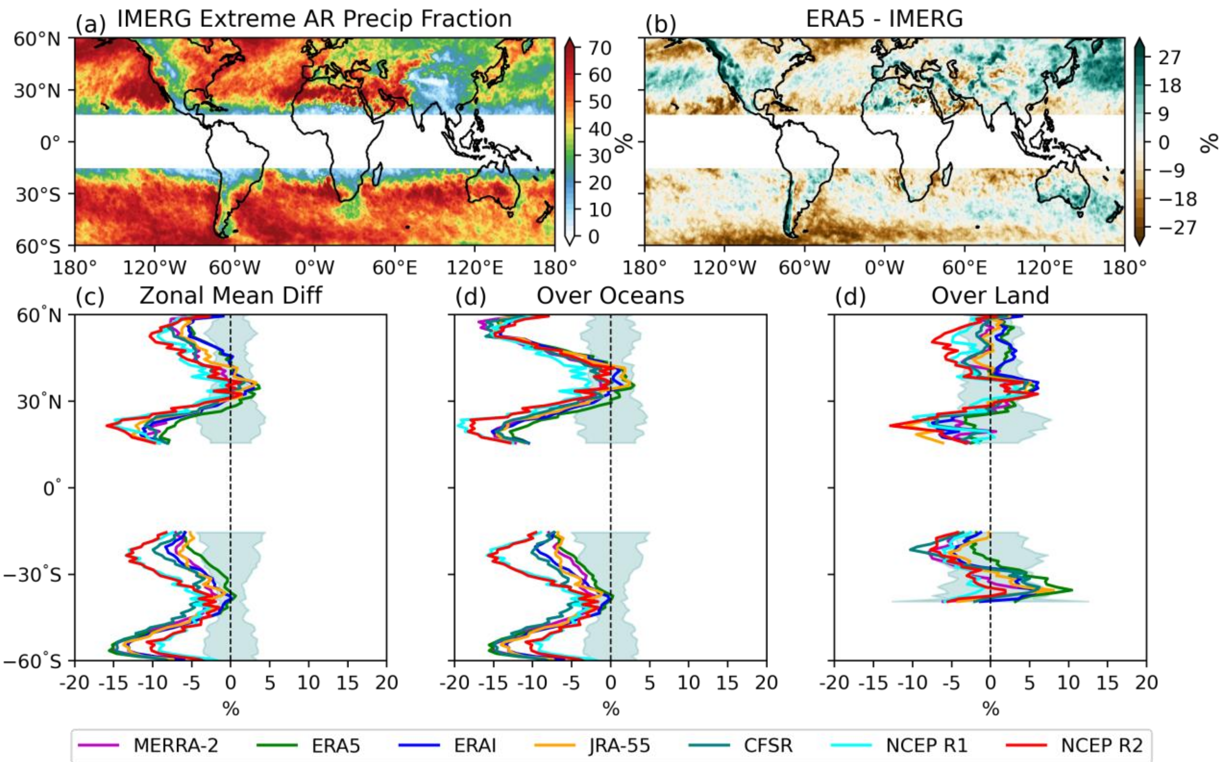


Figure 11. Extreme AR precipitation fraction in IMERG (a). Extreme AR precipitation fraction difference between ERA5 and IMERG (b). The differences between zonal mean extreme AR precipitation fraction in reanalyses and IMERG (c). (d) and (e) are the same as (c), but for the zonal mean extreme AR precipitation fraction over oceans and land, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean extreme AR precipitation fraction in IMERG.

It has been well documented that climate models tend to suffer from the so-called "drizzling" bias in which models tend to rain too frequently and too lightly (Chen et al., 2021; Dai, 2006). To investigate whether this problem is also present in AR precipitation, we define a metric called AR precipitation frequency. There is no guarantee that precipitation must occur when a grid point is experiencing AR conditions. We thus define AR precipitation frequency as the fraction of AR days a grid point experiences noticeable precipitation. Here, noticeable precipitation is defined to be greater than or equal to 0.5 mm/day. We tried varied thresholds ranging from 0.1 mm/day to 2 mm/day. The conclusions presented here are not sensitive to the threshold used (not shown). As shown in Figure 12a, over most of the mid-latitude regions, more than 80% of the AR days are associated with noticeable precipitation. This suggests that our algorithm is able to effectively identify precipitating systems with enhanced moisture transport. Smaller AR precipitation frequency is generally found over land and subtropical regions, suggesting that ARs are usually less efficient in generating precipitation over these regions. Compared to observation, ARs in reanalyses tend to precipitate too often, especially over subtropics equatorward of 30˚N/S, which is consistent with previous studies showing that "drizzling" bias is most severe over lower latitude regions (Chen et al., 2021; Dai, 2006). The "drizzling" bias of ARs in reanalyses discovered here seems to be consistent with previous studies in which they found that reanalyses tend to simulate too many wet days compared to observations (Herold et al., 2016; Naud et al., 2020; Zhou & Wang, 2017). Other regions which also suffer from this problem include regions poleward of 50˚S, Eurasia, and the west coasts of North America and South America. Interestingly, this “drizzling” bias is greatly alleviated poleward of about 30˚N over North Pacific and North Atlantic, as well as over the Southern Ocean between about 30˚S to 40˚S. Surprisingly, the biases in the AR precipitation frequency in NCEP R1 and R2 are relatively smaller compared to other newer generation of reanalyses (Figures 12c, d and e). This seems to be in line with the finding by Zhou & Wang (2017) who showed that the frequency of the drizzle days over China is lower in NCEP R1 and R2 compared to other reanalysis products. In these two reanalyses, ARs precipitate even less often compared to observation over oceanic regions poleward of about 30˚N over the North Atlantic. As shown in Figure S6, the biases in the AR precipitation frequency in reanalyses are also consistent with the biases in the precipitation frequency, which is simply defined as the fraction of days with daily precipitation greater than or equal to 0.5 mm/day. The reasons for the smaller AR-related “drizzling” bias in these two older generation reanalyses are unknown and warrant further studies.

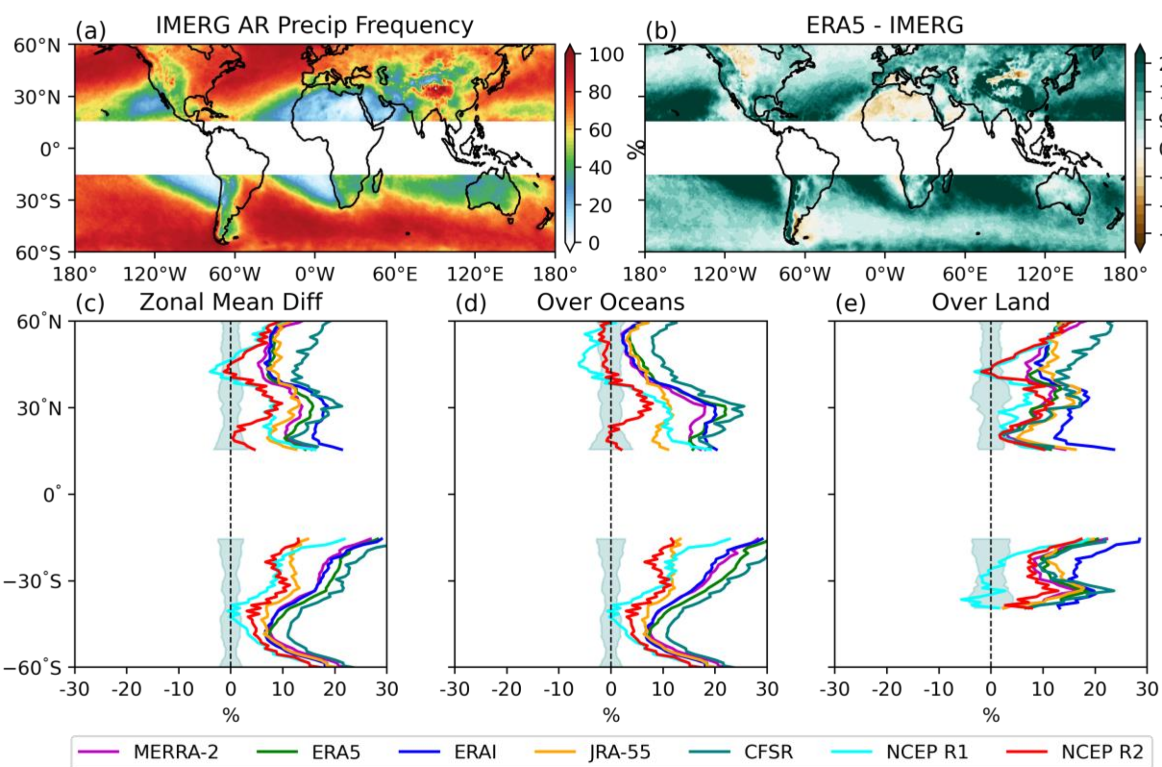


Figure 12. AR precipitation frequency in IMERG (a). AR precipitation frequency difference between ERA5 and IMERG (b). The differences between zonal mean AR precipitation frequency in reanalyses and IMERG (c). (d) and (e) are the same as (c), but for the zonal mean AR precipitation frequency over oceans and land, respectively. The shading in (c), (d), and (e) represents one standard deviation of the annual zonal mean AR precipitation frequency in IMERG.

5 Conclusions and Discussion

Satellite observations and reanalyses have been indispensable in characterizing ARs, which are associated with changes in both moisture and wind fields. However, previous AR studies using satellite data usually detect ARs based only on the IWV while studies using reanalyses usually make the assumption that reanalyses are representative of the true observation (Matrosov, 2013; Neiman et al., 2008; Ralph et al., 2004; Wick et al., 2013). In this study, we improve previous satellite-based AR studies by incorporating wind into the AR detection. Low-level geostrophic winds derived from satellite-based geopotential height are combined with the satellite-based IWV to obtain the GIVT. We demonstrate that GIVT can serve as a good proxy for IVT in terms of magnitude. By removing the three IVT direction criteria in the Guan & Waliser (2015) algorithm, namely the coherence criterion, the meridional IVT criterion and the consistency criterion, we show that the AR frequency based on the GIVT and the one based on the IVT are nearly identical. The modified ARDT is then applied to the GIVT from the satellite observation by AIRS/AMSU and seven commonly used reanalyses: MERRA-2, ERA5, ERAI, JRA-55, CFSR, NCEP R1, and NCEP R2. We find that all datasets show high agreement on AR frequency. Given that the IWV and wind field in reanalyses are strongly constrained by satellite observations, this result should be expected: a weather system which is present in satellite observations should also be present in reanalyses. While the spread in AR length across datasets is relatively small, larger spread in the AR width can be found across datasets. Consequently, the spread in width leads to the spreads in the AR area and AR length/width ratio. Compared to other datasets, MERRA-2 tends to simulate narrower and more ARs while CFSR, NCEP R1, and NCEP R2 tend to simulate broader and fewer ARs. Compared to ARs in satellite observation, ARs in reanalyses have stronger GIVT over midlatitudes. The spread in the AR GIVT among reanalyses is mostly caused by the spread in AR IWV while the spread in AR wind magnitude is small.

Unlike IWV and winds in reanalyses which are heavily constrained by observations, precipitation in reanalyses is produced by the models without any direct observational constraints. Larger biases are thus expected in reanalysis-based precipitation. We evaluate the AR-induced precipitation in reanalyses against that based on IMERG. We reveal systematic biases in the reanalysis-based AR precipitation characteristics. Specifically, we find that reanalyses systematically underestimate both the mean and extreme AR precipitation intensity over oceans, with the strongest underestimates found in NCEP R1. Consequently, the fractional contributions to both the mean and extreme precipitation by ARs are all underestimated by reanalyses. It has long been known that climate models suffer from the so-called “drizzling” bias problem (Chen et al., 2021; Dai, 2006). Namely, models tend to rain too often and too lightly. Defining AR precipitation frequency as the fraction of AR days when a grid point experiences noticeable precipitation, we discover that ARs in reanalyses tend to rain too often, especially over the lower latitude regions. Combined with the weak biases in the AR precipitation intensity, we demonstrate that the "drizzling" bias also exists for AR precipitation in reanalyses. These findings cast doubts on the direct uses of reanalysis-based precipitation in AR studies.

Studies have shown that the statistics of ARs and AR precipitation are sensitive to the ARDT used (Collow et al., 2022; Rutz et al., 2019; Shields et al., 2018). In this study, we employed the modified ARDT based on Guan & Waliser (2015) to demonstrate the feasibility of using GIVT for detecting ARs. In this regard, this study thus serves as a proof of concept. We have demonstrated that GIVT can be a good proxy for IVT given that the direction of GIVT/IVT is not considered. Therefore, as long as the IVT-based ARDT doesn't have any IVT direction criteria, GIVT can be used readily as input to the algorithm and produce AR statistics comparable to those based on IVT. As has been shown (Figure S2), the AR statistics based on Guan & Waliser (2015) are quite different from those based on other ARDTs. The results presented in this study are likely algorithm-dependent. For example, we show in this study that there is a high agreement among datasets on the AR frequency. However, this is mostly due to the percentile-based threshold used in our ARDT. It can be expected that reanalyses with larger climatological GIVT would have larger AR frequency if an absolute threshold is used. Nevertheless, since an identical ARDT is consistently applied to all datasets, the results regarding the spreads among datasets in this study should be more robust.

It is well known that the satellite products have sampling biases (Fetzer et al., 2006; Hearty et al., 2014; Lin et al., 2002; North et al., 1993; Tian et al., 2013; Tian & Thomas, 2020). As a product obtained from both infrared (AIRS) and microwave (AMSU) sensors, the retrieval quality of AIRS/AMSU under cloudy conditions degrades rapidly, and the sampling frequency under cloudy regions, such as the ITCZ and mid-latitude regions, is lower than that over the clear regions, such as the subtropics and some land regions. Furthermore, the Aqua satellite on which the AIRS and AMSU sensors board is in a Sun-synchronous polar orbit, both AIRS and AMSU can only sample the atmosphere twice daily at low latitude regions and thus cannot adequately resolve the diurnal cycle. These sampling issues can result in the sampling biases in the AIRS/AMSU observations (Hearty et al., 2014; Tian et al., 2013; Tian & Thomas, 2020). In addition, the AIRS/AMSU may also have measurement errors due to the AIRS/AMSU retrieval algorithm (Hearty et al., 2014). Given these factors, the results based on AIRS/AMSU are subject to sampling biases and measurement errors and the differences between AIRS/AMSU and other reanalyses also should not be viewed as biases in reanalyses. Instead, the spread among datasets should be simply viewed as observed uncertainties. Smaller spread of a quantity thus gives us more confidence in the ability of reanalyses and satellite observations in representing the true observation of that quantity and vice versa. In this sense, we have more confidence in the reanalyses and satellite observation's ability in representing AR frequency, AR length and AR wind magnitude while our confidence in the AR width and AR IWV representations is lower.

While potential biases exist in the IWV and geostrophic wind field from AIRS/AMSU, it is also possible that the satellite-based precipitation product IMERG also contains biases. Evaluations on the performance of IMERG against other ground observations and remotely sensed products have shown varied results (Pradhan et al., 2022). Whether IMERG has better performance than other observed precipitation products remains inconclusive. However, IMERG is produced by merging the precipitation estimates with highest quality passive microwave sensors and infrared sensors. It thus can be regarded as the state-of-the-art observational precipitation product (Li et al., 2023; Ma et al., 2020; Watters et al., 2021; Xin et al., 2022). Compared to IMERG, it has been well known that models have trouble simulating the precipitation realistically (Chen et al., 2021; Christopoulos & Schneider, 2021; Frei et al., 2003; Kim et al., 2021). It thus can be expected that the IMERG precipitation is closer to the true observation compared to reanalysis-based precipitation. The differences between IMERG-based AR precipitation and reanalysis-based AR precipitation can also be viewed as biases in reanalyses. Furthermore, a previous study (Naud et al., 2020), which uses IMERG to evaluate the modeled precipitation in oceanic extratropical cyclones, found that IMERG-based cyclone precipitation is stronger over the warm sector of the cyclone compared to other model- and reanalysis-based cyclone precipitation. Given that ARs are usually located over the warm sector of the cyclone, this result adds further confidence to the findings presented in this study.

In this study, we have provided a proof of concept for the feasibility of detecting ARs in satellite observations using both moisture and wind information. As the quality of the satellite observations continues to improve, the methodology presented here can be applied to other satellite observations such as geostationary satellites to develop higher resolution or higher frequency AR statistics. We have also conducted a comprehensive intercomparison between reanalyses and satellite observations and among reanalyses on their skills in representing AR characteristics and AR-related precipitation. Our findings provide better guidance on the direct uses of reanalyses and reanalysis-based precipitation in future AR studies. The satellite-based AR statistics and AR precipitation developed in this study can also be used to evaluate the climate models' skills in representing ARs and AR precipitation. Such evaluation will be presented in a future study.

Acknowledgments

We thank Hui Su for valuable discussion on the differences between satellite products and reanalyses. Gang Chen and Weiming Ma were supported by NSF grant AGS-1832842 and NASA grant 80NSSC21K1522. Bin Guan was supported by NASA grants 80NSSC20K1344, 80NSSC21K1007, and 80NSSC22K0926, and the California Department of Water Resources. Part of Baijun Tian's research was performed at Jet Propulsion Laboratory, California Institute of Technology under a contract with National Aeronautics and Space Administration. Weiming Ma also acknowledges support by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research, Regional and Global Model Analysis program area.

The Pacific Northwest National Laboratory (PNNL) is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RLO1830.

Open Research

AIRS/AMSU can be found at: https://disc.gsfc.nasa.gov/datasets/AIRX3STD_006/summary

IMERG can be found at: https://gpm.nasa.gov/data/directory

MERRA-2 (https://doi.org/10.1175/JCLI-D-16-0758.1) can be found at: https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/data_access/

ERA5 (https://doi.org/10.1002/qj.3803) can be found at: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5

ERAI (https://doi.org/10.1002/qj.828) can be found at: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim

JRA-55 (https://doi.org/10.2151/jmsj.2015-001) can be found at: https://rda.ucar.edu/datasets/ds628.0/

CFSR (https://doi.org/10.1175/2010BAMS3001.1) can be found at: https://rda.ucar.edu/datasets/ds093.0/ and https://rda.ucar.edu/datasets/ds094.0/

NCEP R1 (https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2) can be found at: https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html

NCEP R2 (https://doi.org/10.1175/BAMS-83-11-1631) can be found at: https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.html

References

AMS Glossary of Meteorology. 2017. Atmospheric River. Accessed February 17, 2023. http://glossary.ametsoc.org/wiki/Atmospheric_river.

Arabzadeh, A., Ehsani, M. R., Guan, B., Heflin, S., & Behrangi, A. (2020a). Global intercomparison of atmospheric rivers precipitation in remote sensing and reanalysis products. *Journal of Geophysical Research: Atmospheres*, *125*(21), e2020JD033021.

Arabzadeh, A., Ehsani, M. R., Guan, B., Heflin, S., & Behrangi, A. (2020b). Global Intercomparison of Atmospheric Rivers Precipitation in Remote Sensing and Reanalysis Products. *Journal of Geophysical Research: Atmospheres*, *125*(21), 1–15. https://doi.org/10.1029/2020JD033021

Behrangi, A., Guan, B., J. Neiman, P., Schreier, M., & Lambrigtsen, B. (2016). On the Quantification of Atmospheric Rivers Precipitation from Space : Composite Assessments and Case Studies over the Eastern North Pacific Ocean and the Western United States. *Journal of Hydrometeorology*, *17*(1), 369–382. https://doi.org/10.1175/JHM-D-15-0061.1

Cannon, F., Ralph, F. M., Wilson, A. M., & Lettenmaier, D. P. (2017). GPM satellite radar measurements of precipitation and freezing level in atmospheric rivers: Comparison with ground-based radars and reanalyses. *Journal of Geophysical Research: Atmospheres*, *122*(23), 12,747-12,764. https://doi.org/10.1002/2017JD027355

Cannon, F., Cordeira, J. M., Hecht, C. W., Norris, J. R., Michaelis, A., Demirdjian, R., & Martin Ralph, F. (2020). GPM satellite radar observations of precipitation mechanisms in atmospheric rivers. *Monthly Weather Review*, *148*(4), 1449–1463. https://doi.org/10.1175/MWR-D-19-0278.1

Chen, D., Dai, A., & Hall, A. (2021). The Convective-To-Total Precipitation Ratio and the “Drizzling” Bias in Climate Models. *Journal of Geophysical Research: Atmospheres*, *126*(16), 1–17. https://doi.org/10.1029/2020JD034198

Chen, X., Leung, L. R., Gao, Y., Liu, Y., Wigmosta, M., & Richmond, M. (2018). Predictability of extreme precipitation in western US watersheds based on atmospheric river occurrence, intensity, and duration. *Geophysical Research Letters*, *45*(21), 11–693.

Christopoulos, C., & Schneider, T. (2021). Assessing biases and climate implications of the diurnal precipitation cycle in climate models. *Geophysical Research Letters*, *48*(13), e2021GL093017.

Collow, A. B.Marquardt, Shields, C. A., Guan, B., Kim, S., Lora, J. M., McClenny, E. E., et al. (2022). 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), 1–20. https://doi.org/10.1029/2021JD036155

Collow, Allison B.Marquardt, Mersiovsky, H., & Bosilovich, M. G. (2020). Large-scale influences on atmospheric river–induced extreme precipitation events along the coast of Washington State. *Journal of Hydrometeorology*, *21*(9), 2139–2156. https://doi.org/10.1175/JHM-D-19-0272.1

Dai, A. (2006). Precipitation characteristics in eighteen coupled climate models. *Journal of Climate*, *19*(18), 4605–4630. https://doi.org/10.1175/JCLI3884.1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

DeFlorio, M. J., Waliser, D. E., Guan, B., Ralph, F. M., & Vitart, F. (2019). Global evaluation of atmospheric river subseasonal prediction skill. *Climate Dynamics*, *52*(5–6), 3039–3060. https://doi.org/10.1007/s00382-018-4309-x

Dettinger, M. D. (2013). Atmospheric rivers as drought busters on the US West Coast. *Journal of Hydrometeorology*, *14*(6), 1721–1732.

Dettinger, M. D., Ralph, F. M., Das, T., Neiman, P. J., & Cayan, D. R. (2011). Atmospheric rivers, floods and the water resources of California. *Water (Switzerland)*, *3*(2), 445–478. https://doi.org/10.3390/w3020445

Djoumna, G., & Holland, D. M. (2021). Atmospheric Rivers, Warm Air Intrusions, and Surface Radiation Balance in the Amundsen Sea Embayment. *Journal of Geophysical Research: Atmospheres*, *126*(13). https://doi.org/10.1029/2020jd034119

Eiras-Barca, J., Ramos, A. M., Algarra, I., Vázquez, M., Dominguez, F., Miguez-Macho, G., et al. (2021). European West Coast atmospheric rivers: A scale to characterize strength and impacts. *Weather and Climate Extremes*, *31*. https://doi.org/10.1016/j.wace.2021.100305

Espinoza, V., Waliser, D. E., Guan, B., Lavers, D. A., & Ralph, F. M. (2018). Global Analysis of Climate Change Projection Effects on Atmospheric Rivers. *Geophysical Research Letters*, *45*(9), 4299–4308. https://doi.org/10.1029/2017GL076968

Fetzer, E. J., Lambrigtsen, B. H., Eldering, A., Aumann, H. H., & Chahine, M. T. (2006). Biases in total precipitable water vapor climatologies from Atmospheric Infrared Sounder and Advanced Microwave Scanning Radiometer, *111*, 1–14. https://doi.org/10.1029/2005JD006598

Frei, C., Christensen, J. H., Déqué, M., Jacob, D., Jones, R. G., & Vidale, P. L. (2003). Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *Journal of Geophysical Research: Atmospheres*, *108*(D3).

Gao, Y., Lu, J., Leung, L. R., Yang, Q., Hagos, S., & Qian, Y. (2015). Dynamical and thermodynamical modulations on future changes of landfalling atmospheric rivers over western North America. *Geophysical Research Letters*, *42*(17), 7179–7186. https://doi.org/10.1002/2015GL065435

Gao, Y., Lu, J., & Leung, L. R. (2016a). Uncertainties in projecting future changes in atmospheric rivers and their impacts on heavy precipitation over Europe. *Journal of Climate*, *29*(18), 6711–6726. https://doi.org/10.1175/JCLI-D-16-0088.1

Gao, Y., Lu, J., & Leung, L. R. (2016b). Uncertainties in projecting future changes in atmospheric rivers and their impacts on heavy precipitation over Europe. *Journal of Climate*, *29*(18), 6711–6726.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Guan, B., & Waliser, D. E. (2015). Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. *Journal of Geophysical Research*, *120*(24), 12,514-12,535. https://doi.org/10.1002/2015JD024257

Guan, B., & Waliser, D. E. (2017). Atmospheric rivers in 20 year weather and climate simulations: A multimodel, global evaluation. *Journal of Geophysical Research*, *122*(11), 5556–5581. https://doi.org/10.1002/2016JD026174

Guan, B., Molotch, N. P., Waliser, D. E., Fetzer, E. J., & Neiman, P. J. (2010). Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements. *Geophysical Research Letters*, *37*(20), n/a-n/a. https://doi.org/10.1029/2010GL044696

Guan, B., Waliser, D. E., & Martin Ralph, F. (2018). An intercomparison between reanalysis and dropsonde observations of the total water vapor transport in individual atmospheric rivers. *Journal of Hydrometeorology*, *19*(2), 321–337. https://doi.org/10.1175/JHM-D-17-0114.1

Guan, B., Waliser, D. E., & Ralph, F. M. (2020). A multimodel evaluation of the water vapor budget in atmospheric rivers. *Annals of the New York Academy of Sciences*, *1472*(1), 139–154. https://doi.org/10.1111/nyas.14368

Guan, B., Waliser, D. E., & Ralph, F. M. (2023). Global Application of the Atmospheric River Scale. *Journal of Geophysical Research: Atmospheres*, *128*(3). https://doi.org/10.1029/2022JD037180

Hearty, T. J., Savtchenko, A., Tian, B., Fetzer, E., Yung, Y. L., Theobald, M., et al. (2014). Estimating sampling biases and measurement uncertainties of AIRS/AMSU-A temperature and water vapor observations using MERRA reanalysis. *Journal of Geophysical Research: Atmospheres*, *119*(6), 2725–2741.

Hegyi, B. M., & Taylor, P. C. (2018). The Unprecedented 2016–2017 Arctic Sea Ice Growth Season: The Crucial Role of Atmospheric Rivers and Longwave Fluxes. *Geophysical Research Letters*, *45*(10), 5204–5212. https://doi.org/10.1029/2017GL076717

Henn, B., Musselman, K. N., Lestak, L., Ralph, F. M., & Molotch, N. P. (2020). Extreme Runoff Generation From Atmospheric River Driven Snowmelt During the 2017 Oroville Dam Spillways Incident. *Geophysical Research Letters*, *47*(14), 1–11. https://doi.org/10.1029/2020GL088189

Herold, N., Alexander, L. V., Donat, M. G., Contractor, S., & Becker, A. (2016). How much does it rain over land? *Geophysical Research Letters*, *43*(1), 341–348. https://doi.org/10.1002/2015GL066615

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hollinger, J. P., Peirce, J. L., & Poe, G. A. (1990). SSM / I Instrument Evaluation, *28*(5).

Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, J. T. (2019). GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V06. Retrieved May 23, 2020, from https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06/summary?keywords=%22IMERG final%22

Ionita, M., Nagavciuc, V., & Guan, B. (2020). Rivers in the sky, flooding on the ground: The role of atmospheric rivers in inland flooding in central Europe. *Hydrology and Earth System Sciences*, *24*(11), 5125–5147. https://doi.org/10.5194/hess-24-5125-2020

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–472. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., & Potter, G. L. (2002). Ncep--doe amip-ii reanalysis (r-2). *Bulletin of the American Meteorological Society*, *83*(11), 1631–1644.

Kim, H., Kang, S. M., Takahashi, K., Donohoe, A., & Pendergrass, A. G. (2021). Mechanisms of tropical precipitation biases in climate models. *Climate Dynamics*, *56*(1–2), 17–27.

Kim, J, Guan, B., Waliser, D. E., Ferraro, R. D., Case, J. L., Iguchi, T., et al. (2018). Winter precipitation characteristics in western US related to atmospheric river landfalls: observations and model evaluations. *Climate Dynamics*, *50*, 231–248.

Kim, Jinwon, Moon, H., Guan, B., Waliser, D. E., Choi, J., Gu, T. Y., & Byun, Y. H. (2021). Precipitation characteristics related to atmospheric rivers in East Asia. *International Journal of Climatology*, *41*(S1), E2244–E2257. https://doi.org/10.1002/joc.6843

Kim, S., Leung, L. R., Guan, B., & Chiang, J. C. H. (2022). Atmospheric river representation in the Energy Exascale Earth System Model (E3SM) version 1.0. *Geoscientific Model Development*, *15*(14), 5461–5480. https://doi.org/10.5194/gmd-15-5461-2022

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*(1), 5–48. https://doi.org/10.2151/jmsj.2015-001

Lamjiri, M. A., Dettinger, M. D., Ralph, F. M., & Guan, B. (2017). Hourly storm characteristics along the U.S. West Coast: Role of atmospheric rivers in extreme precipitation. *Geophysical Research Letters*, *44*(13), 7020–7028. https://doi.org/10.1002/2017GL074193

Lavers, D. A., & Villarini, G. (2013). The nexus between atmospheric rivers and extreme precipitation across Europe. *Geophysical Research Letters*, *40*(12), 3259–3264. https://doi.org/10.1002/grl.50636

Li, R., Guilloteau, C., Kirstetter, P.-E., & Foufoula-Georgiou, E. (2023). How well does the IMERG satellite precipitation product capture the timing of precipitation events? *Journal of Hydrology*, *620*, 129563.

Lin, X., Fowler, L. D., & Randall, D. A. (2002). Flying the TRMM satellite in a general circulation model. *Journal of Geophysical Research: Atmospheres*, *107*(D16), ACH--4.

Lora, J. M., Shields, C. A., & Rutz, J. J. (2020). Consensus and Disagreement in Atmospheric River Detection: ARTMIP Global Catalogues. *Geophysical Research Letters*, *47*(20), 1–10. https://doi.org/10.1029/2020GL089302

Ma, W., & Chen, G. (2022). What Controls the Interannual Variability of the Boreal Winter Atmospheric River Activities over the Northern Hemisphere? *Journal of Climate*, 1–39. https://doi.org/10.1175/jcli-d-22-0089.1

Ma, W., Chen, G., & Guan, B. (2020). Poleward Shift of Atmospheric Rivers in the Southern Hemisphere in Recent Decades. *Geophysical Research Letters*, *47*(21), 1–11. https://doi.org/10.1029/2020GL089934

Ma, W., Norris, J., & Chen, G. (2020). Projected Changes to Extreme Precipitation Along North American West Coast From the CESM Large Ensemble. *Geophysical Research Letters*, *47*(1), 1–10. https://doi.org/10.1029/2019GL086038

Ma, W., Chen, G., Peings, Y., & Alviz, N. (2021). Atmospheric River Response to Arctic Sea Ice Loss in the Polar Amplification Model Intercomparison Project. *Geophysical Research Letters*, 1–12. https://doi.org/10.1029/2021gl094883

Ma, Z., Xu, J., Zhu, S., Yang, J., Tang, G., Yang, Y., et al. (2020). AIMERG: a new Asian precipitation dataset (0.1°/half-hourly, 2000--2015) by calibrating the GPM-era IMERG at a daily scale using APHRODITE. *Earth System Science Data*, *12*(3), 1525–1544.

Maclennan, M. L., Lenaerts, J. T. M., Shields, C., & Wille, J. D. (2022). Contribution of atmospheric rivers to antarctic precipitation. *Geophysical Research Letters*, *49*(18), e2022GL100585.

Massoud, E. C., Espinoza, V., Guan, B., & Waliser, D. E. (2019). Global Climate Model Ensemble Approaches for Future Projections of Atmospheric Rivers. *Earth's Future*, *7*(10), 1136–1151. https://doi.org/10.1029/2019EF001249

Matrosov, S. Y. (2013). Characteristics of landfalling atmospheric rivers inferred from satellite observations over the Eastern North Pacific Ocean. *Monthly Weather Review*, *141*(11), 3757–3768. https://doi.org/10.1175/MWR-D-12-00324.1

Mattingly, K. S., Mote, T. L., & Fettweis, X. (2018). Atmospheric River Impacts on Greenland Ice Sheet Surface Mass Balance. *Journal of Geophysical Research: Atmospheres*, *123*(16), 8538–8560. https://doi.org/10.1029/2018JD028714

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360. https://doi.org/10.1175/BAMS-87-3-343

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., Van Weele, M., & Boersma, K. F. (2012). Simultaneous assimilation of satellite NO 2, O 3, CO, and HNO 3 data for the analysis of tropospheric chemical composition and emissions. *Atmospheric Chemistry and Physics*, *12*(20), 9545–9579. https://doi.org/10.5194/acp-12-9545-2012

Nash, D., Waliser, D., Guan, B., Ye, H., & Ralph, F. M. (2018). The Role of Atmospheric Rivers in Extratropical and Polar Hydroclimate. *Journal of Geophysical Research: Atmospheres*, *123*(13), 6804–6821. https://doi.org/10.1029/2017JD028130

Naud, C. M., Jeyaratnam, J., Booth, J. F., Zhao, M., & Gettelman, A. (2020). Evaluation of modeled precipitation in oceanic extratropical cyclones using IMERG. *Journal of Climate*, *33*(1), 95–113. https://doi.org/10.1175/JCLI-D-19-0369.1

Neiman, P. J., Ralph, F. M., Wick, G. A., Kuo, Y. H., Wee, T. K., Ma, Z., et al. (2008). Diagnosis of an intense atmospheric river impacting the pacific northwest: Storm summary and offshore vertical structure observed with COSMIC satellite retrievals. *Monthly Weather Review*, *136*(11), 4398–4420. https://doi.org/10.1175/2008MWR2550.1

Neiman, P. J., Ralph, F. M., Wick, G. A., Lundquist, J. D., & Dettinger, M. D. (2008). Meteorological characteristics and overland precipitation impacts of atmospheric rivers affecting the West coast of North America based on eight years of SSM/I satellite observations. *Journal of Hydrometeorology*, *9*(1), 22–47. https://doi.org/10.1175/2007JHM855.1

Neiman, P. J., Ralph, F. M., Wick, G. A., Lundquist, J. D., Dettinger, M. D., Neiman, P. J., et al. (2008). Meteorological Characteristics and Overland Precipitation Impacts of Atmospheric Rivers Affecting the West Coast of North America Based on Eight Years of SSM/I Satellite Observations. *Journal of Hydrometeorology*, *9*(1), 22–47. https://doi.org/10.1175/2007JHM855.1

North, G. R., Shen, S. S. P., & Upson, R. (1993). Sampling Errors in Rainfall Estimates by Multiple Satellites. *Journal of Applied Meteorology and Climatology*, *32*, 339–410.

Paltan, H., Waliser, D., Lim, W. H., Guan, B., Yamazaki, D., Pant, R., & Dadson, S. (2017). Global Floods and Water Availability Driven by Atmospheric Rivers. *Geophysical Research Letters*, *44*(20), 10,387-10,395. https://doi.org/10.1002/2017GL074882

Pasquier, J. T., Pfahl, S., & Grams, C. M. (2019). Modulation of atmospheric river occurrence and associated precipitation extremes in the North Atlantic region by European weather regimes. *Geophysical Research Letters*, *46*(2), 1014–1023.

Payne, A. E., Demory, M.-E., Leung, L. R., Ramos, A. M., Shields, C. A., Rutz, J. J., et al. (2020). Responses and impacts of atmospheric rivers to climate change. *Nature Reviews Earth & Environment*, *1*(3), 143–157. https://doi.org/10.1038/s43017-020-0030-5

Pradhan, R. K., Markonis, Y., Vargas Godoy, M. R., Villalba-Pradas, A., Andreadis, K. M., Nikolopoulos, E. I., et al. (2022). Review of GPM IMERG performance: A global perspective. *Remote Sensing of Environment*, *268*(May 2021), 112754. https://doi.org/10.1016/j.rse.2021.112754

Ralph, F. M., Wick, G. A., Neiman, P. J., Moore, B. J., Spackman, J. R., Hughes, M., et al. (2012). Atmospheric rivers in reanalysis products: A six-event comparison with aircraft observations of water vapor transport. *Extended Abstracts, WCRP Reanalysis Conference*, 1.

Ralph, F. Martin, Neiman, P. J., & Wick, G. A. (2004). Satellite and CALJET aircraft observations of atmospheric rivers over the Eastern North Pacific Ocean during the winter of 1997/98. *Monthly Weather Review*, *132*(7), 1721–1745. https://doi.org/10.1175/1520-0493(2004)132<1721:SACAOO>2.0.CO;2

Ralph, F. Martin, Neiman, P. J., Wick, G. A., Gutman, S. I., Dettinger, M. D., Cayan, D. R., & White, A. B. (2006). Flooding on California's Russian River: Role of atmospheric rivers. *Geophysical Research Letters*, *33*(13), L13801. https://doi.org/10.1029/2006GL026689

Ralph, F Martin, Rutz, J. J., Cordeira, J. M., Dettinger, M., Anderson, M., Reynolds, D., et al. (2019). A scale to characterize the strength and impacts of atmospheric rivers. *Bulletin of the American Meteorological Society*, *100*(2), 269–289.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Rutz, J. J., & Steenburgh, W. J. (2012). Quantifying the role of atmospheric rivers in the interior western United States. *Atmospheric Science Letters*, *13*(4), 257–261. https://doi.org/10.1002/asl.392

Rutz, J. J., Shields, C. A., Lora, J. M., Payne, A. E., Guan, B., Ullrich, P., et al. (2019). The Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Quantifying

Uncertainties in Atmospheric River Climatology. *Journal of Geophysical Research: Atmospheres*, *124*(24), 13777–13802. https://doi.org/10.1029/2019JD030936

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015–1057. https://doi.org/10.1175/2010BAMS3001.1

Shields, C. A., Rutz, J. J., Leung, L. Y., Martin Ralph, F., Wehner, M., Kawzenuk, B., et al. (2018). Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project goals and experimental design. *Geoscientific Model Development*, *11*(6), 2455–2474. https://doi.org/10.5194/gmd-11-2455-2018

Teixeira, A. S. T. (n.d.). AIRS/Aqua L3 Daily Standard Physical Retrieval (AIRS+AMSU) 1 degree x 1 degree V006.

The NCAR Command Language. (2019). https://doi.org/http://dx.doi.org/10.5065/D6WD3XH5

Tian, B., & Hearty, T.J. (2020). Estimating and Removing the Sampling Biases of the AIRS Obs4MIPs V2 data. *Earth and Space Science*, *7*. https://doi.org/10.1029/2020EA001438

Tian, B., Fetzer, E. J., Kahn, B. H., Teixeira, J., Manning, E., & Hearty, T. (2013). Evaluating CMIP5 models using AIRS tropospheric air temperature and specific humidity climatology. *Journal of Geophysical Research Atmospheres*, *118*(1), 114–134. https://doi.org/10.1029/2012JD018607

Tian, B., Manning, E., Fetzer, E., Olsen, E., Wong, S., Susskind, J., & Iredell, L. (2017). AIRS / AMSU / HSB Version 6 Level 3 Product User Guide Edited by :

Viale, M., Valenzuela, R., Garreaud, R. D., & Ralph, F. M. (2018). Impacts of atmospheric rivers on precipitation in southern South America. *Journal of Hydrometeorology*, *19*(10), 1671–1687.

Waliser, D., & Guan, B. (2017a). Extreme winds and precipitation during landfall of atmospheric rivers. *Nature Geoscience*, *10*(3), 179–183. https://doi.org/10.1038/ngeo2894

Waliser, D., & Guan, B. (2017b). Extreme winds and precipitation during landfall of atmospheric rivers. *Nature Geoscience*, *10*(3), 179–183. https://doi.org/10.1038/ngeo2894

Wang, Z., Walsh, J., Szymborski, S., & Peng, M. (2020). Rapid arctic sea ice loss on the synoptic time scale and related atmospheric circulation anomalies. *Journal of Climate*, *33*(5), 1597–1617. https://doi.org/10.1175/JCLI-D-19-0528.1

Watters, D., Battaglia, A., & Allan, R. P. (2021). The diurnal cycle of precipitation according to multiple decades of global satellite observations, three CMIP6 models, and the ECMWF reanalysis. *Journal of Climate*, *34*(12), 5063–5080.

Wick, G. A., Neiman, P. J., & Ralph, F. M. (2013). Description and validation of an automated objective technique for identification and characterization of the integrated water vapor signature of atmospheric rivers. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(4), 2166–2176. https://doi.org/10.1109/TGRS.2012.2211024

Wille, J. D., Favier, V., Dufour, A., Gorodetskaya, I. V., Turner, J., Agosta, C., & Codron, F. (2019). West Antarctic surface melt triggered by atmospheric rivers. *Nature Geoscience*, *12*(11), 911–916. https://doi.org/10.1038/s41561-019-0460-1

Wille, J. D., Favier, V., Jourdain, N. C., Kittel, C., Turton, J. V., Agosta, C., et al. (2022). Intense atmospheric rivers can weaken ice shelf stability at the Antarctic Peninsula. *Communications Earth and Environment*, *3*(1). https://doi.org/10.1038/s43247-022-00422-9

Woods, C., & Caballero, R. (2016). The role of moist intrusions in winter arctic warming and sea ice decline. *Journal of Climate*, *29*(12), 4473–4485. https://doi.org/10.1175/JCLI-D-15-0773.1

Xin, Y., Yang, Y., Chen, X., Yue, X., Liu, Y., & Yin, C. (2022). Evaluation of IMERG and ERA5 precipitation products over the Mongolian Plateau. *Scientific Reports*, *12*(1), 21776.

Yue, Q., Lambrigtsen, B., Behrangi, A., Chen, L., Dang, V., Fetzer, E.J., Fishbein, E., Irion, F.W., Kahn, B., Kalmus, P., Manning, E., Olsen, E., Tian, B., Wilson, C., & Wong, S. (2017). AIRS V6 Test Report Supplement: Performance of AIRS+AMSU vs. AIRS-only retrievals. *Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, USA*.

Yue, Q., Lambrigtsen, B., & Coauthors (2020). AIRS version 7 Level 2 performance test and validation report. *Jet Propulsion Laboratory, California Institute of Technology: Pasadena, CA, USA*.

Zhang, P., Chen, G., Ma, W., Ming, Y., & Wu, Z. (2021). Robust atmospheric river response to global warming in idealized and comprehensive climate models. *Journal of Climate*, *34*(18), 7717–7734. https://doi.org/10.1175/JCLI-D-20-1005.1

Zhang, P., Chen, G., Ting, M., Ruby Leung, L., Guan, B., & Li, L. (2023). More frequent atmospheric rivers slow the seasonal recovery of Arctic sea ice. *Nature Climate Change*, *13*(March). https://doi.org/10.1038/s41558-023-01599-3

Zhao, M. (2022). A study of AR-, TS-, and MCS-associated precipitation and extreme precipitation in present and warmer climates. *Journal of Climate*, *35*(2), 479–497.

Zhou, C., & Wang, K. (2017). Contrasting Daytime and nighttime precipitation variability between observations and eight reanalysis products from 1979 to 2014 in China. *Journal of Climate*, *30*(16), 6443–6464. https://doi.org/10.1175/JCLI-D-16-0702.1

Zhou, Y., O'Brien, T. A., Collins, W. D., Shields, C. A., Loring, B., & Elbashandy, A. A. (2022). Characteristics and variability of winter northern Pacific atmospheric river flavors. *Journal of Geophysical Research: Atmospheres*, *127*(23), e2022JD037105.

Zhu, Y., & Newell, R. E. (1998). A proposed algorithm for moisture fluxes from atmospheric rivers. *Monthly Weather Review*, *126*(3), 725–735.

Journal of Geophysical Research: Atmospheres

Supporting Information for

Evaluating the Representations of Atmospheric Rivers and Their Associated Precipitation in Reanalyses with Satellite Observations

Weiming Ma[1], Gang Chen[1], Bin Guan[2,3], and Christine A. Shields[4], Baijun Tian[3] and Emilio Yanez[1]

[1]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA

[2]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA

[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

[4]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

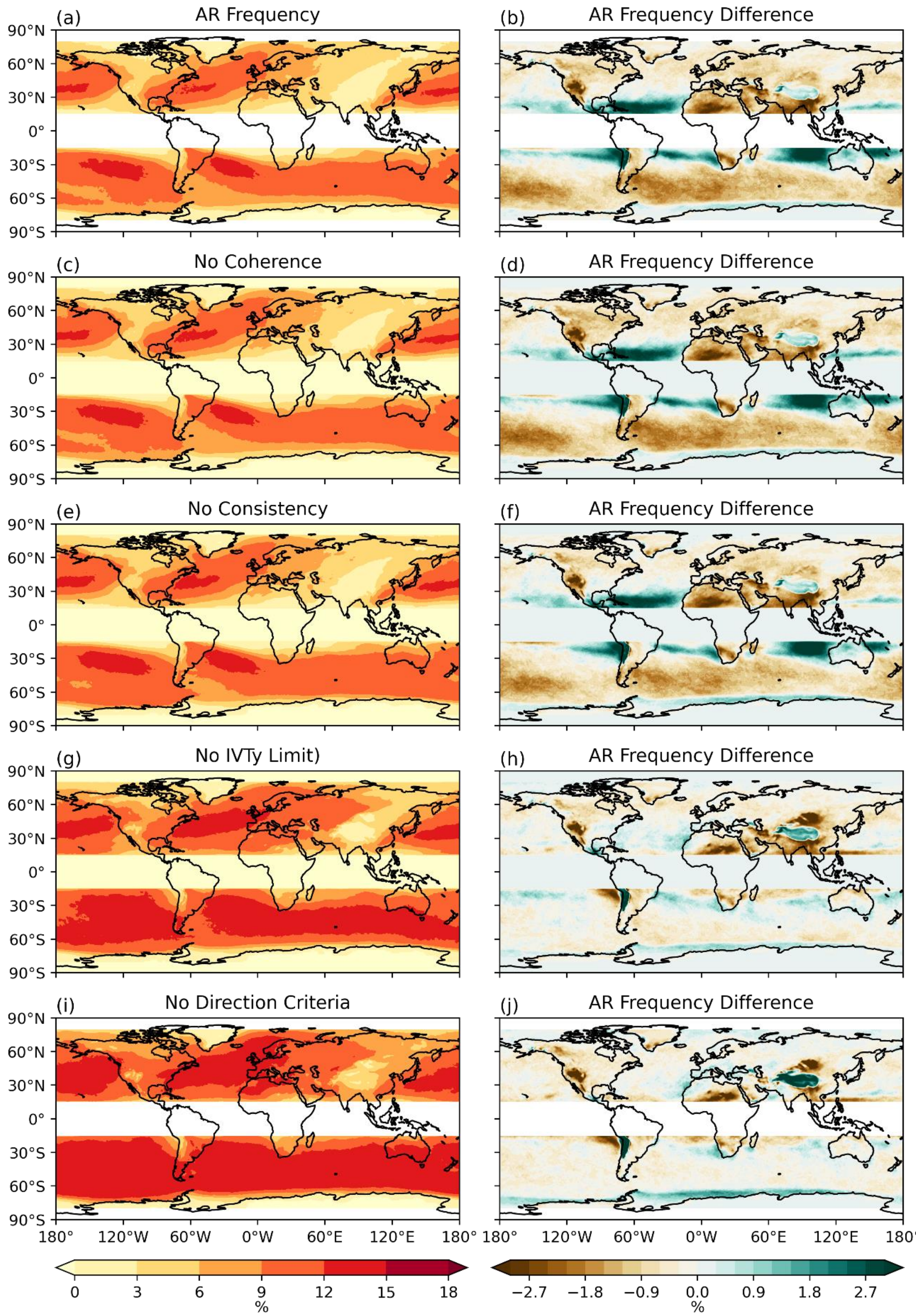
Contents of this file

Text S1.

Due to the presence of the noises in the geopotential height field in AIRS/AMSU, the derived geostrophic winds are too strong compared to those derived from reanalyses. To alleviate this problem, a bias correction procedure was applied to the geostrophic wind magnitude in AIRS/AMSU. More specifically, the weighted vertical average geostrophic winds were first calculated for each grid point using data at 925, 850, 700 and 600 mb. The climatological mean geostrophic wind magnitude for each grid point over mid-latitude oceans from 30° N/S to 60° N/S in ERA5 was then calculated and regressed onto the unsmoothed climatological mean geostrophic wind magnitude of the corresponding grid point over mid-latitude oceans from 30° N/S to 60° N/S in AIRS/AMSU (spatial regression). The obtained regression coefficient and intercept were then applied to the geostrophic wind magnitude at every grid point and time step in AIRS/AMSU. This bias correction procedure would correct the mean in AIRS/AMSU so that the climatological mean geostrophic wind magnitude spatially averaged over mid-latitude oceans from 30° N/S to 60° N/S in AIRS/AMSU would equal to that in ERA5. Since the AR detection algorithm used in this study adopts a percentile-based threshold, such bias correction on the geostrophic wind magnitude would not affect the ARs detected.

Criteria	Original Guan & Waliser (2015)	Modified Guan & Waliser (2015)
Threshold	85^{th} percentile IVT for a particular month is based on all the time steps from the five months centered on that month over the study period and a lower limit of 100 kg/m/s	85^{th} percentile IVT for a particular month is based on all the time steps from that month over the study period and a lower limit of 100 kg/m/s
Coherence criterion	Yes	No
Meridional IVT criterion	Yes	No
Consistency Criterion	Yes	No
Length >= 2000 km	Yes	yes
Length/Width Ratio >= 2	Yes	Yes
Detected Object filtered out if all its area located within 30° N/S	No	Yes

Table S1. Summary of the differences between the original Guan & Waliser (2015) algorithm and the modified algorithm used in this study. The definition of coherence criterion, meridional IVT criterion, and consistency criterion can be found in section 2.2 of the main text.



(a)
AR Frequency
(b)
AR Frequency Difference
(c)
No Coherence
(d)
AR Frequency Difference
(e)
No Consistency
(f)
AR Frequency Difference
(g)
No IVTy Limit)
(h)
AR Frequency Difference
(i)
No Direction Criteria
(j)
AR Frequency Difference
90°N
60°N
30°N
0°
30°S
60°S
90°S
180°
120°W
60°W
0°W
60°E
120°E
180°
0
3
6
9
12
15
18
%
−2.7
−1.8
−0.9
0.0
0.9
1.8
2.7
%

1260

Figure S1. AR frequency distribution based on GIVT using the original Guan & Waliser (2015) algorithm (a), the original algorithm but with the coherence criterion removed (c), the original algorithm but with the consistency criterion removed (e), the original algorithm but with the meridional IVT criterion removed (g), the original algorithm but with all three IVT direction criteria removed (i). (b), (d), (f), (h), and (j) are showing the AR frequency differences between those based on GIVT and those based on IVT (GIVT – IVT) using the corresponding modified algorithms. See the main text for more information on the IVT direction criteria used in the original algorithm. All results shown here are based on ERA5.

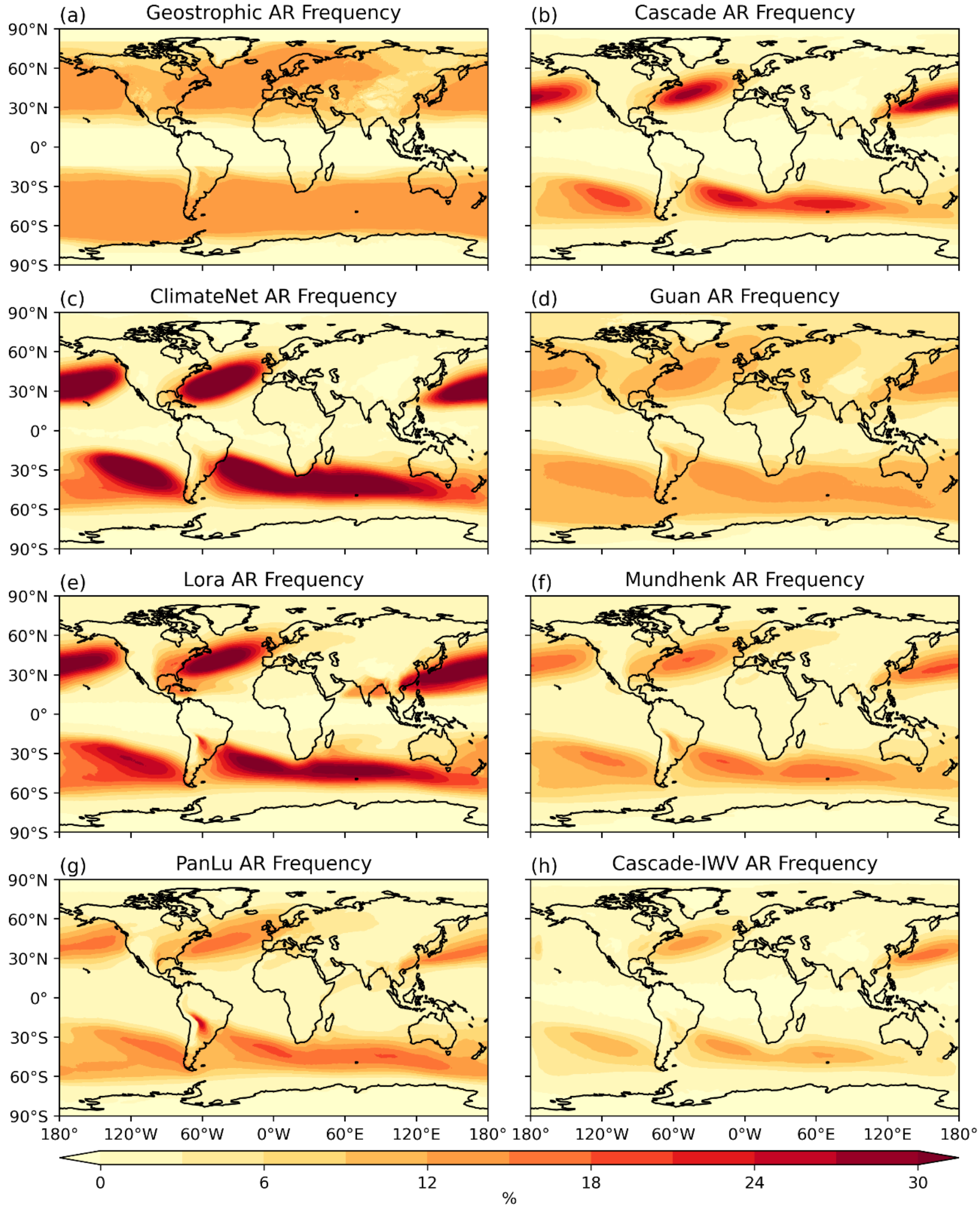


Figure S2. Annual AR frequency based on the AR detection algorithm for satellite data and seven global AR detection algorithms participated in the ARTMIP. All panels are based the 6-hourly data from MERRA-2.

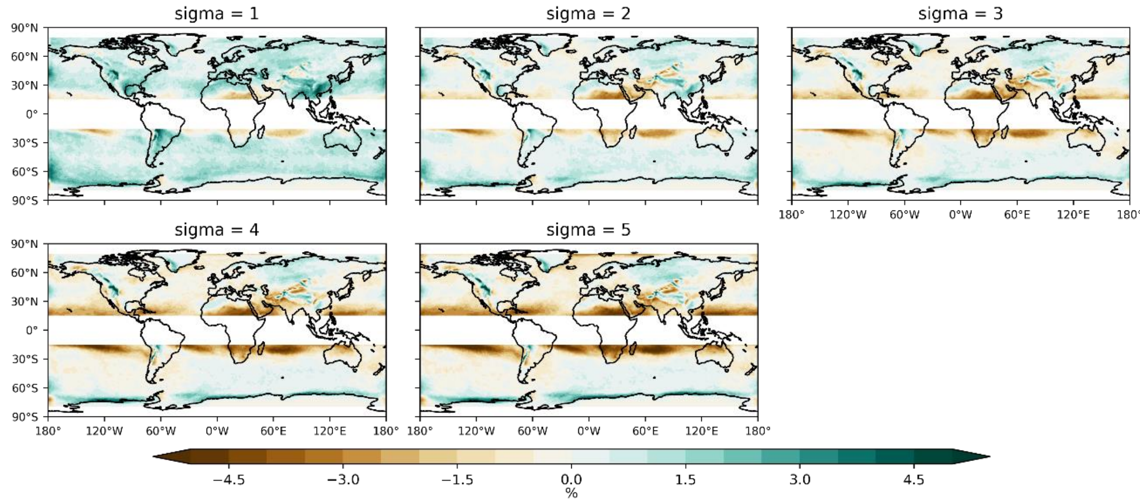


Figure S3. AR frequency difference between ERA5 and AIRS/AMSU as a function of sigma.

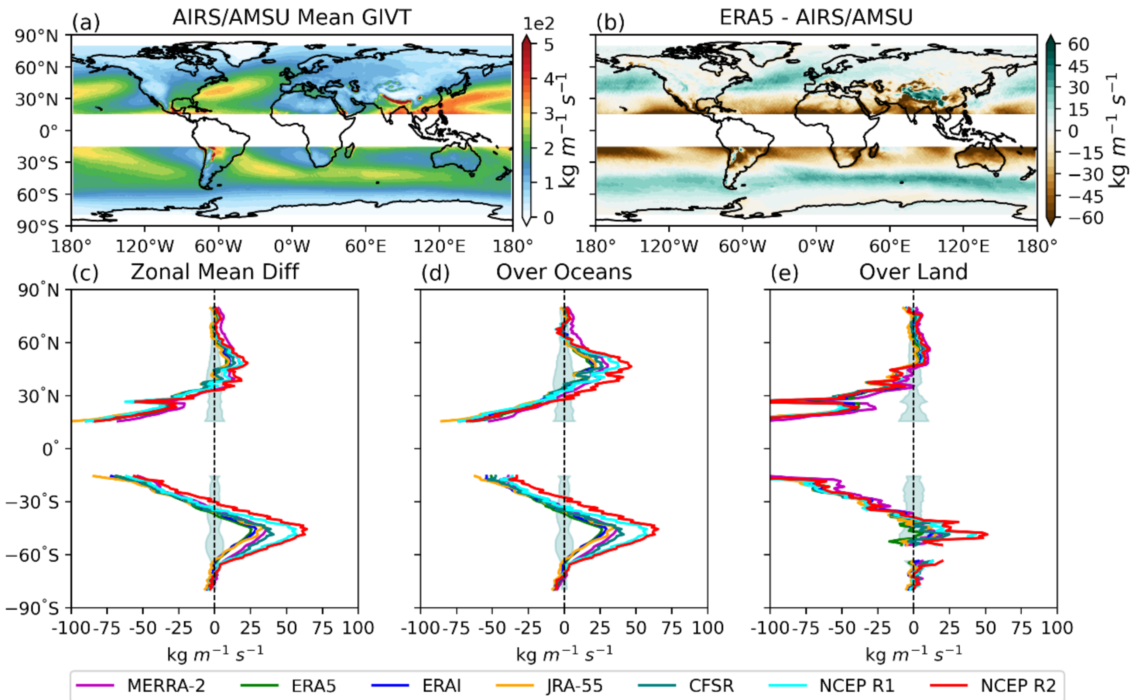


Figure S4. Climatological mean GIVT in AIRS/AMSU (a). Climatological mean GIVT difference between ERA5 and AIRS/AMSU (b). The differences between climatological zonal mean GIVT in reanalyses and AIRS/AMSU (c). (d) and (e) are the same as (c), but for the climatological zonal mean GIVT over oceans and land, respectively. The shading in (c), (d) and (e) represents one standard deviation of the annual zonal mean GIVT in AIRS/AMSU

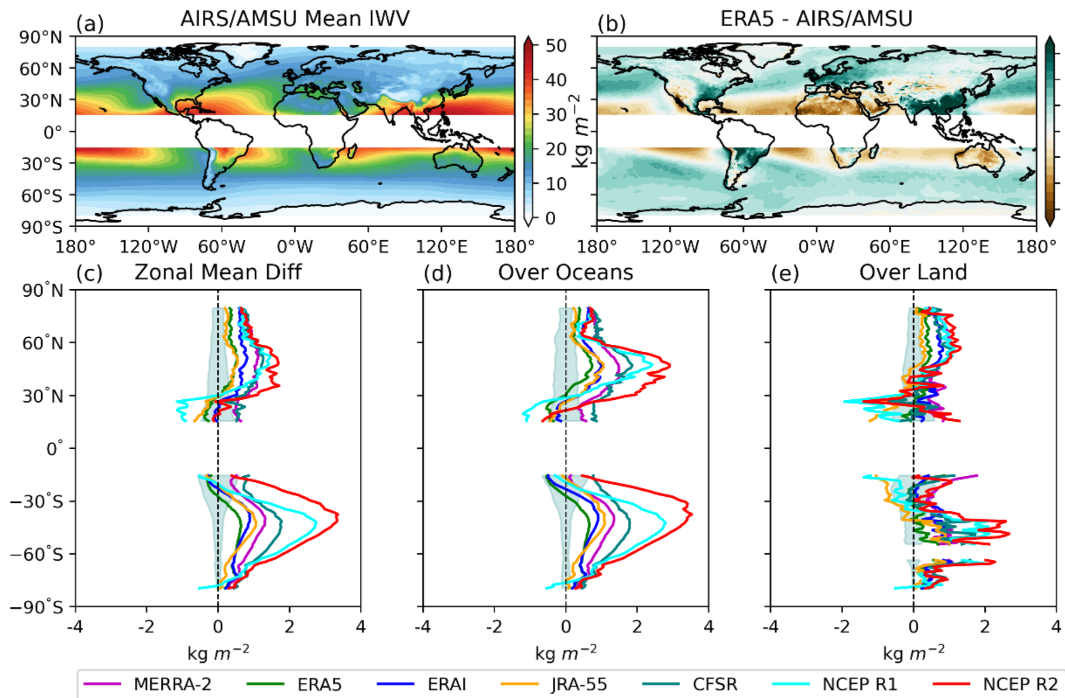


Figure S5. Climatological mean IWV in AIRS/AMSU (a). Climatological mean IWV difference between ERA5 and AIRS/AMSU (b). The differences between climatological zonal mean IWV in reanalyses and AIRS/AMSU (c). (d) and (e) are the same as (c), but for the climatological zonal mean IWV over oceans and land, respectively. The shading in (c), (d) and (e) represents one standard deviation of the annual zonal mean IWV in AIRS/AMSU

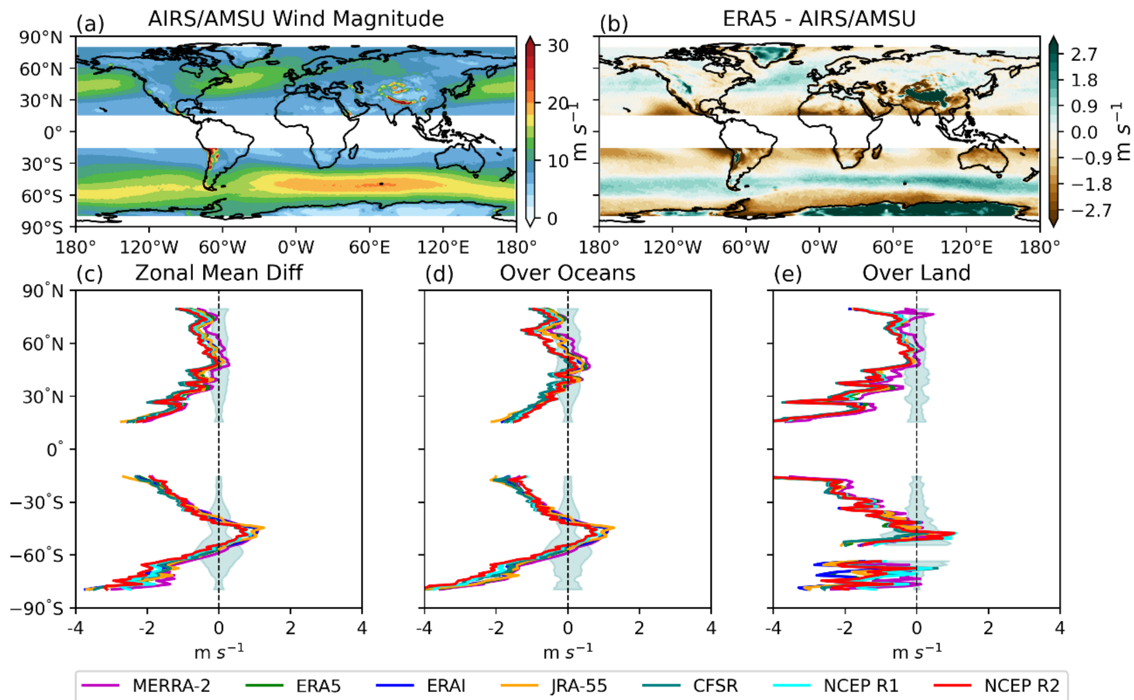


Figure S6. Climatological mean geostrophic wind speed in AIRS/AMSU (a). Climatological mean geostrophic wind speed difference between ERA5 and AIRS/AMSU (b). The differences between climatological zonal mean geostrophic wind speed in reanalyses and AIRS/AMSU (c). (d) and (e) are the same as (c), but for the climatological zonal mean geostrophic wind speed over oceans and land, respectively. The shading in (c), (d) and (e) represents one standard deviation of the annual zonal mean wind magnitude in AIRS/AMSU

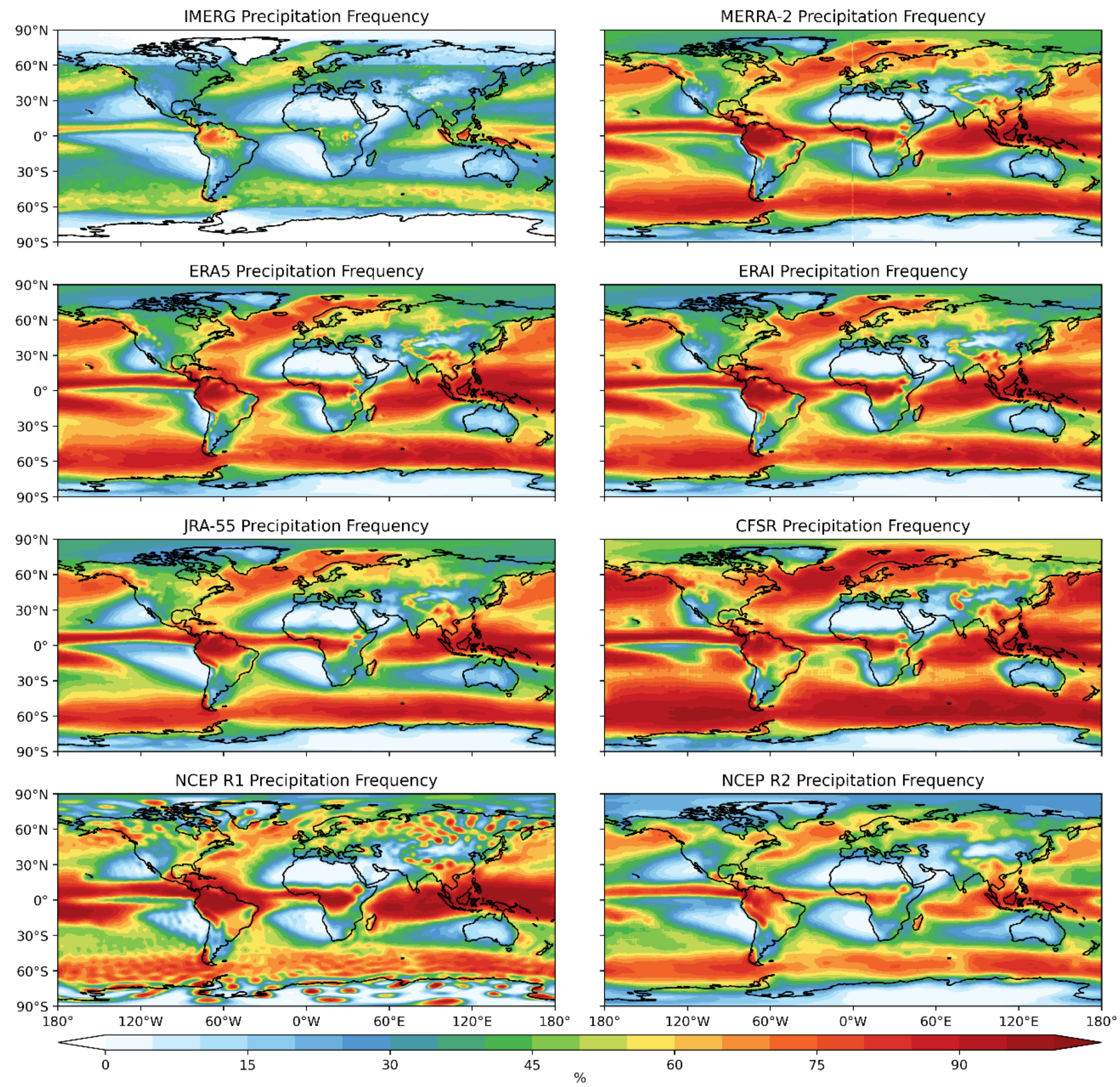


Figure S7. Precipitation frequency in satellite observation and reanalysis products.